Moment-SOS hierarchy for large scale set approximation. Application to power systems transient stability analysis

Matteo Tacchi

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :
l'Institut National des Sciences Appliquées de Toulouse (INSA de Toulouse)

Présentée et soutenue le *29 Juin 2021* **par :**

Matteo TACCHI

Moment-SOS hierarchy for large scale set approximation. Application to power systems transient stability analysis.

JURY

Leo LIBERTI	DR LIX-CNRS	Président, Rapporteur
Sorin OLARU	Pr. CentraleSupélec	Rapporteur
Carmen CARDOZO	Ph.D. Ingénieur RTE	Examinatrice
Colin N. JONES	Assoc. Pr. EPFL	Examinateur
Monique LAURENT	Pr. CWI Amsterdam	Examinatrice
Line ROALD	Asst. Pr. UW-Madison	Examinatrice
Didier HENRION	DR LAAS-CNRS	Directeur de Thèse

École doctorale et spécialité :
EDSYS : Automatique 4200046
Unité de Recherche :
Equipe MAC, LAAS-CNRS (UPR 8001)
Directeur de Thèse :
Didier HENRION
Rapporteurs :
Leo LIBERTI et *Sorin OLARU*

À Isaac, petit miracle dans notre quotidien.
À Giuseppe *Pino* (1933 – 2021), mémoire éternelle.
À Tatiana, coéquipière de choc en toutes circonstances.

"La vie est une cerise
La mort est un noyau
L'amour un cerisier."
Jacques Prévert, *Chanson du Mois de Mai.*

Remerciements

"La gratitude est le secret de la vie. L'essentiel est de remercier pour tout. Celui qui a appris cela sait ce que vivre signifie. Il a pénétré le profond mystère de la vie."
Albert Schweitzer - médecin, philosophe et théologien alsacien.

C'est paradoxalement au début de mon manuscrit de thèse que je fais le bilan, sinon scientifique, au moins personnel, de cette longue période qui touche à sa fin. Regarder ainsi en arrière donne quelque peu le vertige... Il est sans aucun doute de bon aloi de tenter d'exprimer la profonde gratitude que j'éprouve à l'égard des personnes qui ont contribué, de près ou de loin, à l'aboutissement de ce long travail.

En premier lieu bien sûr, je pense à l'équipe qui a encadré mes recherches : Didier Henrion, mon directeur de thèse, Jean Bernard Lasserre, co-encadrant régulier de mes travaux, Carmen Cardozo, ma référente entreprise, et Patrick Panciatici, conseiller scientifique RTE.

Didier, nous n'aurons finalement jamais eu cette grave discussion à la bibliothèque... Et pour cause ! Tu m'as apporté un encadrement bienveillant et respectueux de mon autonomie, tout en me donnant de précieux conseils, tant sur le fond et la forme de nos travaux que sur leur direction. Ta pédagogie et tes suggestions de lecture m'ont permis de m'approprier rapidement les concepts difficiles qui sous-tendent la thèse. Enfin, tu as toujours été accessible en dehors du laboratoire, fût-ce pour m'aider à gérer des galères ou des conflits, ou simplement pour boire un verre. Je garde un souvenir mémorable de nos missions ensemble, à Prague et à Zürich, expériences fantastiques que j'ai été heureux de partager avec toi. J'espère qu'on pourra continuer ce partage, tant scientifique qu'amical, longtemps après la fin de cette thèse.

Jean, tu as été un pilier de nos innovations scientifiques, toujours là pour suggérer une nouvelle piste de réflexion ou proposer une idée que personne n'avait eue, sans pour autant brimer ma propre créativité, et toujours avec une admirable facilité à faire abstraction du superflu pour te plonger dans l'essentiel scientifique. Avec le recul, j'ai le sentiment qu'une vraie synergie s'est mise en place, qui a permis des développements dont je suis fier. Mais aussi, tu as partagé avec moi des trésors insoupçonnés, comme la fameuse recette de l'omelette sans œufs, la localisation du dernier type qui t'a mal parlé ou encore l'énigme de Gibraltar. Nos sorties au Bikini, au Filochard et à la Mécanique des Fluides vont me manquer...

Carmen, c'est à toi que je dois d'avoir pu envisager cette thèse. Je n'oublierai jamais ton rôle déterminant dans l'élaboration de ce projet d'encadrement au LAAS, ni comment tu as su transformer le stagiaire timide que j'étais en un doctorant épanoui. Ton expertise technique des systèmes électriques a été d'une aide précieuse, et c'est toi qui m'as permis de réaliser ce rêve que j'avais d'utiliser des connaissances théoriques pour m'attaquer à des problèmes appliqués. Ta patience, ta pédagogie et ta rigueur ont été la clef du réalisme de mes contributions les plus techniques. Je n'oublie pas non plus mes séjours au siège de RTE où tu m'as accueilli avec beaucoup de bienveillance. Pour tout cela, merci à toi.

Pour son suivi assidu de mes travaux et l'initiative de ce projet collaboratif entre RTE et le LAAS, je veux rendre un hommage tout particulier à Patrick Panciatici, dont la vision scientifique extrêmement large, couvrant à la fois les problématiques industrielles liées à la gestion des réseaux de transport d'électricité, et les méthodes théoriques ayant le potentiel pour contribuer à l'innovation dans ce domaine, a permis la rencontre entre les hiérarchies moment-SOS et l'analyse de stabilité des systèmes électriques. Ta curiosité et ton ouverture d'esprit m'ont fait forte impression, et je sais que c'est aussi grâce à elles que cette thèse a pu avoir lieu. Je me réjouis que notre collaboration se poursuive au-delà de la thèse.

I would also like to deeply thank the members of my defense committee. I am particularly grateful to Leo Liberti and Sorin Olaru for their very careful review of the present manuscript. Your comments and suggestions helped a lot improving the overall quality of this work, including the PhD defense. Also, thank you Leo for accepting the additional responsibility of presiding over the committee, that was really appreciated. Many thanks also to the whole committee: Leo Liberti, Sorin Olaru, Carmen Cardozo, Colin Jones, Monique Laurent, Line Roald and Didier Henrion, for their benevolent compliments on my work, as well as their very interesting and relevant questions after my talk. Thanks to you all, my defense was a priceless reward for all the work of the past years.

Then, I am also grateful to all my collaborators. Our stimulating interactions allowed me to quickly enlarge my scientific horizon to a point that I would have taken much more than 3 years to reach without you.

Immédiatement après celles et ceux qui ont contribué au lancement, à l'accompagnement et à la conclusion de ce marathon scientifique, je tiens à mentionner un certain nombre d'amis chers à mon coeur : ce dernier volet de mes études a apporté, outre la bénéfique expérience dans le métier de la recherche, son lot d'épreuves et de rencontres, celles-ci ayant été déterminantes pour surmonter celles-là.

Swann, je te le dis tout de go : tu me dois un nugget. Initialement je voulais entamer ce paragraphe avec une série de private jokes à base de langoustin et de "c'est où Zouane ?", mais en me relisant j'ai trouvé ça bancal alors je change mon fusil d'épaule. On a été co-bureaux pendant 2 ans au bas mot, ce qui fait que tu partages la plupart des souvenirs que j'évoquerai dans les prochains paragraphes. Il faut donc que je sois original ici. Il y a un truc unique et je dois dire assez sidérant avec toi, c'est que cette complicité entre nous, qui est si vite devenue centrale dans mon horizon relationnel, continue à se développer même à distance. Je me rappelle les lectures de mails au bureau, les soirées "gaufres" en ville, le Kraken (sacrée histoire) ; mon EVG organisé par tes soins, les voyages ensemble en France et à l'étranger, les Utopiales ; notre goût commun pour l'absurde et la science fiction... Mais en fait je crois que ce qui compte le plus pour moi, c'est ta disponibilité pour parler à coeur ouvert, de visu ou à distance. Merci pour cette amitié profonde, extrêmement précieuse. Et merci pour l'exemple et l'espoir que tu me donnes, par ton admirable développement scientifique et professionnel.

Matteo (*a.k.a.* MDR, *a.k.a* Jean-Michel Pizza le slovène), merci de porter un si beau prénom (il te va bien). Merci aussi pour l'excellence de ton humour poétique,

pour les sous-problèmes du capitalisme, les descentes de police, la polémique, la retraite à l'italienne... Merci enfin pour tes paroles et tes gestes d'affection toujours très touchants. Ti ringrazio per tutto questo e spero di vederti presto!!!

Tillmann, j'ai été très heureux de partager ton bureau pendant presque un an, merci de m'avoir enseigné les ficelles pour être un bon doctorant, merci pour toutes nos conversations, scientifiques ou non, au bureau, au Filochard, à la Méca, au Biergarten, au Père Peinard, au Concorde... merci aussi pour ta présence au mariage. Et merci pour le magnifique concert avant ton départ ! Guten tag!

Victor, sacré Magron, merci à toi d'avoir si souvent volé à mon secours (ton aide logistique pour ma soutenance a été déterminante), de m'avoir rassuré quand ma confiance en moi déclinait, de m'avoir appris à chercher sur internet. Merci aussi pour les sorties à Toulouse, Autrans, Bilbao, Nice, Nantes, Guidel... J'espère qu'on aura encore beaucoup de bons **moments** à partager ! Mes amitiés à toute la famille.

Flavien, très cher collègue daron au sang chaud, merci pour toute la complicité qu'on a pu avoir, pour ta solidité aux 50 ans du LAAS, pour ta générosité à l'EVG (déguisements, rafraîchissements et cascades inclus), pour ta camaraderie au week-end ski, pour ton ton expertise en randonnées, pour tes biscottos si bien taillés, pour ta solidarité jusque dans la paternité (félicitations à vous trois)... On se reverra !

Alex, merci pour ton accessibilité amicale, j'ai un souvenir impérissable de nos sorties au Melting Pot, au Senso et à la Dynamo ! Tu as fait tomber le mur imaginaire qui me séparait des chercheurs permanents. Merci aussi pour ton soutien psychologique dans la dernière ligne droite avant la soutenance, sans lequel j'aurais bien pu perdre les pédales ! À très bientôt j'espère.

Lucie, qui as partagé avec moi la couronne de la galette deux années d'affilée, qui t'es mariée la même année que moi (et que Jean), merci pour ton infinie gentillesse, ta précieuse amitié, tes conseils avisés, ton soutien moral... Merci de nous avoir prêté ton époux le temps d'une soirée jeu, et d'avoir accueilli mon pot de thèse dans ton magnifique jardin. Pense à nous à ton prochain passage en Suisse !

Comme l'ambiance de travail est primordiale pour la créativité scientifique, je veux aussi remercier ceux de mes collègues de bureau que je n'ai pas encore eu l'occasion de citer : Clément, pour les gazzinades et pour le cours de sabre japonais ; Yohei, pour nous avoir suivis très très loin, Swann et moi (belly check) ; Mathias, pour les sympathiques discussions entre deux lignes de code (faut surtout pas s'énerver) ; Marianne, pour la profonde gentillesse que j'ai trouvée derrière ta discrétion réservée ; Corbi, pour ta bonne humeur permanente et contagieuse, et pour ta belle moustache.

Évidemment, je n'oublie pas non plus les camarades doctorants, post-docs **et** stagiaires dont j'ai pu croiser le chemin, notamment Matthieu, toujours avenant, pour ton sourire et pour les nombreux moments de convivialité partagés (bises à Harmony et Léo) ; Antoine, pour la forte émulation scientifique que j'ai ressentie à ton contact lors de ton stage chez RTE ; Florent, pour ton précieux cocktail de compétence, de modestie et de sympathie ; Hoang, pour tes retours sur certains de mes articles ; Roxana, pour ta soutenance tenue le jour de mon arrivée et pour ton plaid dont j'ai hérité ; Alessandra, pour les sorties avec les copains et pour la salsa ; Constantinos, pour la culture orthodoxe, de Grèce, de Russie ou d'ailleurs ; Quentin,

pour la cordialitay et les soirées jeux (bisous à Blanche et KP) ; Larbi, on s'est à peine croisés mais j'ai beaucoup apprécié nos échanges ; et enfin Yoni et Mathieu, pour votre engagement dans la "vie estudiantine" des doctorants MAC, à travers le séminaire doc, les ateliers oenologie, les gâteaux, et les surprises que vous organisez à l'occasion des soutenances (vous êtes géniaux)... Merci à vous tous, que votre nom soit écrit ci-dessus ou non (confinement aidant, je ne vous connais malheureusement pas tous personnellement...), chacun de vous contribue à sa manière à faire vivre le groupe des "éphémères" MAC !

Et enfin, puisqu'on parle des "éphémères", merci aussi à ceux qui ne le sont pas, et qui voient les stagiaires, doctorants et post-doc arriver dans l'équipe, grandir et s'en aller... Après avoir assisté au départ de tous mes prédécesseurs, je crois que je comprends ce que ressentent les permanents lorsqu'un doctorant qu'ils ont côtoyé pendant trois ans soutient finalement sa thèse et part vers de nouveaux horizons. Merci Milan pour t'être joint à nous dans nos soirées festives ; Dmitry, pour m'avoir partagé ton expérience dans tous les domaines imaginables ; Sophie, pour ces plaisanteries qui vont me manquer (j'adore ce que tu fais !) ; Isabelle, pour les discussions dans le couloir ou en salle café ; Aneel, pour tes questions toujours stimulantes et pour ton sens de l'humour ; Luca, pour tes précieux conseils de présentation ; Fred, pour avoir partagé avec moi tes astuces d'enseignement ; Mateusz, pour t'être amicalement joint à nous en pleine crise sanitaire.

Maintenant que j'ai cité les membres du LAAS que j'ai fréquentés régulièrement, je veux aussi exprimer ma gratitude envers des personnes que je n'ai que très rarement (voire jamais) croisées, mais sans qui mes recherches (ainsi que la recherche en général au LAAS) n'auraient pas pu se faire dans d'aussi bonnes conditions. Malheureusement trop souvent oubliés lorsque vient le temps des remerciements, ces collègues fournissent un travail colossal pour nous permettre à nous, enseignants et chercheurs, de nous focaliser sur notre travail scientifique et nos diverses responsabilités organisationnelles. Merci, donc, au personnel du restaurant, de permettre chaque jour à des hordes d'employés affamés de venir recharger leurs batteries à la cantine ; à l'équipe de Sysadmin, pour m'avoir installé (et réparé quand c'était nécessaire) l'ordinateur sur lequel cette thèse a été rédigée, et pour m'avoir aidé à faire fonctionner la visioconférence et la diffusion vidéo de ma soutenance sur les canaux dédiés ; à l'équipe administrative, pour avoir piloté la mise en place de mon contrat d'encadrement avec RTE, pris en charge la logistique de mes diverses missions, et sans aucun doute effectué bien d'autres tâches dont j'ignore jusqu'à l'existence même (mention spéciale à Amandine qui a dû gérer ma mobilité à Zürich) ; aux membres du service documentation pour avoir suivi et enregistré toutes mes publications, y compris ce manuscrit de thèse ; au personnel de la reprographie, pour m'avoir régulièrement imprimé des documents reliés que je garde précieusement, et pour avoir édité le présent ouvrage ; à toute l'équipe d'entretien et de nettoyage, grâce à qui mon bureau et toutes les installations que j'ai pu utiliser étaient toujours nickel ; aux employées du standard, pour leurs salutations avenantes le matin (et pour avoir retrouvé ma gourde après ma soutenance) ; enfin, à Éric et Corine du magasin, sans qui j'aurais bien vite manqué de papier, crayons et marqueurs.

Et comme il n'y a pas eu que le LAAS dans ma thèse, il me faut étendre encore mes remerciements à des zones géographiques plus exotiques. Dans un autre bureau, loin, très loin à La Défense (et avant cela à Versailles), j'ai eu le plaisir de faire la connaissance d'Emeline, Thibault, Guillaume, Philippe, Gilles, Manuel et Jean, des équipes INT et optimisation chez RTE. Merci pour votre accueil chaleureux, pour votre curiosité envers mes travaux, et pour avoir fait de mes séjours à Paris des expériences enrichissantes non seulement scientifiquement mais aussi humainement ; grâce à vous je n'ai jamais eu le mal du pays.

Merci également à mes trois camarades de perdition à l'IRCCyN, Kossi, Jérémy et Yankaï, on sait tous les quatre quelles épreuves on a eues à surmonter, et sans vous peut-être n'aurais-je pas pris les bonnes décisions aux bons moments. Merci pour le temps passé ensemble à Nantes et pour votre solidarité en béton armé !

And I cannot forget those who welcomed me during my mobility in Zürich: Florian, for being my local advisor, and as such patiently and benevolently teaching me so many things (I have not forgotten our ongoing project and I have great hopes in it); Sabrina, for organizing everything for my arrival at IfA; Daniele and Liviu, for sharing your office with me; Irina, for explaining so many things about power converter control; Ali, for sharing your knowledge and your codes with me; Michael, for your enthusiastic curiosity about my work; Verena, for joining our power systems stabiltiy group; Lukas, for welcoming me in the team and adding me to the Riot network; Marcello, for your valuable insights on optimization for power system control.

Mes pensées vont ensuite à toute ma famille, dont le soutien moral et les encouragements m'ont indubitablement aidé à surmonter les épreuves qui se sont présentées à moi.

Claude, Jacky, vous rendre fiers de votre petit-fils a toujours été une importante source de motivation pour moi, et je me réjouis du soutien affectueux que vous m'apportez sans cesse.

Frédérique, Pascal, votre attention, vos encouragements et votre capacité à toujours me consoler dans les moments de cafard n'ont jamais cessé de me donner la force dont j'avais besoin.

Nicolas, ta fierté, ton enseignement martial et mental, ainsi que nos moments de détente père-fils, nourrissent continuellement mon développement personnel.

Gaetano, avoir un petit frère aussi brillant que toi (même si tu refuses de l'admettre) a suscité en moi cette fierté qui donne envie de se donner à fond pour être au niveau. Courage pour ta suite d'études, tu feras un médecin admirable. Et n'oublie jamais : Banach, Bourbaki, Alaoglu.

Néo, Angelo, mes chers jeunes petits frères, je suis d'ores et déjà fier de vous deux. Les bons moments partagés dans notre petite tribu m'apportent ce qu'il faut de bonne humeur pour me plonger sereinement dans mon travail.

Edith, Giuseppe, je n'oublie pas vos conseils, vos félicitations et votre attention à mon égard, notamment au sujet de cette thèse dont nous avons longuement discuté.

Valérie, Jean-Luc, Lisa, Rémi, merci de nous avoir tant reçus (le plus souvent à l'improviste) pendant ces trois années où nous fûmes voisins et où nous refîmes le monde ensemble, et merci également pour vos visites chez nous. Ces contacts

réguliers étaient un vrai rayon de soleil.

Marianne, Christophe, Florence, je vous remercie tous les trois d'avoir partagé avec moi votre expérience de docteurs ès sciences ; vos conseils, vos témoignages et vos argumentaires pendants nos débats ont nourri ma réflexion sur ma thèse et le monde de l'enseignement-recherche en général.

Louise, Alice, Juliette, merci pour les conversations au téléphone, les soirées jeux à Noël et en été, et tous ces bons moments que nous partageons et qui ont rechargé mes batteries à bloc quand je devais affronter mes systèmes électriques.

Enfin, Danièle, Alain, François, Sabine, merci pour vos appels et messages d'encouragements, aux moments cruciaux, qui renforçaient en moi le sentiment de ne jamais être seul face à une difficulté.

Avant de terminer ce chapitre, il y a quelques amis de longue date qui, je pense, méritent aussi que je leur signifie ma reconnaissance, pour avoir été là depuis un bon moment déjà, et avoir donc contribué à me nourrir scientifiquement et émotionnellement.

Thieum, merci à toi, pour tellement de choses ! Pour m'avoir filé de gros coups de pouce sur certains problèmes de maths que nos profs s'amusaient à nous poser à Lyon, pour ta fabuleuse musique, pour ton accueil à Nantes quand je cherchais un toit... Pour toutes nos discussions politiques, philosophiques et littéraires, qui ont profondément changé ma vision de beaucoup de choses et qui m'ont énormément aidé à grandir, notamment l'impressionnant préambule de ta thèse [28], qui m'a donné une véritable leçon d'autocritique. Merci aussi pour l'ossau iraty, le paprika, les youhouuu et les wololooo, sans doute plus bas de plafond mais non moins importants pour l'épanouissement émotionnel.

Idriss, quand il est question de discussions, scientifiques, littéraires ou politiques entre doctorants, ton nom me vient naturellement à l'esprit. Merci pour ton accueil à Jussieu, et pour ton passage à Toulouse, on s'est bien amusés ! Merci aussi de m'avoir rafraîchi les idées sur certains concepts mathématiques qui m'échappaient encore, mon chapitre 4 devrait te rappeler des choses... Et puis, merci pour ton humour piquant : "Adieu rasta blanc, bonjour mocassins à glands !".

Juliane, tu as beau détester les maths, il fallait bien que je mette ton nom quelque part dans ce chapitre : même à distance, on a partagé pas mal de choses ces dernières années ! Merci pour l'EVG à Toulouse, pour ton témoignage à Burey, pour ta présence à Daru, pour les quelques fois où on a réussi à se voir à Nancy, pour tes séjours à Toulouse, pour m'avoir reproché mes annonces par messages, pour avoir pensé à mon anniversaire, pour le temps passé au téléphone... On continue sur cette lancée !

Un petit mot amical aux camarades Troyens, spécialistes en organisation de vacances de rêve, des montagnes du Vercors aux plages vendéennes, en passant par les hospices bourguignons. Ma dernière soirée enquête avec vous commence à dater, il va falloir remettre ça prochainement ! Merci Lisa, Laeti, Thibault, Zoé, Patricia, Simon, Olivier, Julia, Mathilde, Marie, Cécile, Max, Merveille et Péleg, vous êtes un peu comme une deuxième famille : vos petits messages maintiennent le moral et aident à bien bosser. Courage aux doctorant·e·s (Olivier, ces points médians sont pour toi), force aux profs (mes héros) et plein de bonnes choses à tous les autres !

Merci enfin à Solène, pour t'acquitter à merveille de tes rôles de marraine et de disciple en Hapkido, et pour la relation unique qui en résulte. Ton année en poste à Toulouse, alors que j'y étais moi-même établi, a été une excellente surprise (même si c'était beaucoup trop court...) ! Les entraînements ensemble au Shaolin étaient la parfaite excuse pour se voir toutes les semaines. Merci aussi pour les dîners chez toi, et enfin merci pour l'invitation à ton mariage, c'était un sacré événement ! Bisous à Alex et Béatrice...

Je ne pourrais conclure ce chapitre de remerciements autrement qu'en exprimant mon infinie gratitude envers **la** personne qui a sans doute joué le rôle le plus important dans l'accomplissement quotidien de mon travail, en rendant possible mon installation à Toulouse malgré les difficultés pratiques qui en ont résulté, et surtout en m'apportant un soutien moral continu et sans faille, ainsi qu'un épanouissement personnel total qui a décuplé ma motivation : Tatiana, ma chère et tendre épouse. J'ai longtemps cherché un moyen de te remercier pour tout ce que tu fais et tout ce que tu es pour moi, mais les mots me manquent pour exprimer réellement ma reconnaissance envers toi. Je me refuse à te remercier de me supporter, car il y a une facilité toute patriarcale à se complaire dans un rôle de boulet et à se féliciter d'avoir une épouse aimante et patiente pour nous "supporter" quand on rentre fourbu du travail. Pour autant, je veux souligner à quel point le soutien actif que tu m'as apporté a été précieux. J'espère être moi aussi un solide appui pour ton propre accomplissement. Étant toi aussi en thèse, tu m'as également offert une émulation intellectuelle particulièrement stimulante, et nos discussions sur tes propres recherches ont été autant de motivations pour faire avancer moi aussi l'état des connaissances dans mon domaine. Et par dessus tout, au plan strictement personnel, tu m'as comblé de bonheur, d'abord en m'accordant ta main (au premier mois de ma thèse), puis en donnant naissance à notre fils Isaac (au dernier mois de ma thèse, amusante coïncidence...). Merci à toi d'être qui tu es. Я тебя люблю.

Abstract

This thesis deals with approximating sets using Lasserre's moment-SOS hierarchy. The motivation is the increasing need for efficient methods to approximate sets of secure operating conditions for electrical power systems. Indeed, recent and ongoing changes in the European power network, such as the increase in renewable energy sources interfaced by power electronic devices, are bringing up new challenges in terms of power grid security assessment. The aim of the present thesis is to investigate the suitability of the moment-SOS hierarchy as a tool for large scale stability assessment.

In this regard, the very scheme of moment-SOS hierarchies is analysed in-depth, and general results regarding the convergence and accuracy of the framework are stated, along with specific computational methods inspired from differential geometry and partial differential equations theory, in order to improve the convergence of the numerical scheme.

From the computational viewpoint, the core of this thesis is the exploitation of problem structure to alleviate the computational burden of high dimensional, large scale industrial problems. The structure of power grids leads us to consider general sparsity patterns and design methods which distribute our computations accordingly, drastically reducing computational costs in implementation.

In addition to stability analysis, a special interest is put on the theoretical problem of volume computation, whose applications rather concern the field of integral calculus and probability evaluation, as understanding this problem turns out to be a prerequisite for approximating stability regions of differential systems, such as regions of attraction or positively invariant sets, with the moment-SOS hierarchy. Indeed, the moment-SOS approach to volume computation is the core of moment-SOS stability analysis.

Keywords: Power systems transient stability – infinite dimensional optimization – moments – polynomial sums of squares – set approximation – direct methods for stability analysis – sparsity – semidefinite relaxations.

Résumé

Cette thèse a pour objet l'approximation d'ensembles au moyen de la hiérarchie moment-sommes-de-carrés (abrégée moment-SOS) de Lasserre. Elle est motivée par le besoin croissant de méthodes efficaces pour approcher des ensembles de points de fonctionnement stables dans le domaine des réseaux électriques. En effet, les récents développements et les changements en cours au sein du système électrique européen, comme l'augmentation de la part des énergies renouvelables dans la génération d'électricité, et leur raccordement au réseau par des interfaces d'électronique de puissance, soulèvent de nouveaux défis en termes d'évaluation de la sécurité des réseaux électriques. L'objectif de cette thèse est d'étudier la pertinence de la hiérarchie moment-SOS dans les études de stabilité à grande échelle.

Dans cette optique, le schéma numérique que constituent les hiérarchies moment-SOS est étudié en détails, et des résultats généraux sur la convergence et la précision de cet outil sont formulés, et accompagnés de méthodes de calcul spécifiques, inspirées de notions de géométrie différentielle et de théorie des équations aux dérivées partielles, visant à améliorer la convergence du schéma numérique.

Du point de vue purement calculatoire, l'élément central de cette thèse est l'exploitation de la structure des problèmes, en vue d'alléger le coût des calculs liés aux problèmes industriels à grande échelle, modellisés en très grande dimension. La structure en réseau des systèmes électriques nous conduit à nous intéresser aux configurations dites parcimonieuses, et à concevoir des méthodes distribuant les calculs suivant ces configurations, permettant ainsi de réduire drastiquement le coût en calcul de nos implémentations.

Enfin, en plus de l'analyse de stabilité, un intérêt particulier est accordé au problème théorique du calcul de volumes, dont les applications se situent plutôt dans le domaine du calcul intégral et de l'évaluation probabiliste, la compréhension de ce problème étant un prérequis pour l'approximation de régions de stabilité pour les systèmes différentiels, comme par exemple les régions d'attractions ou les ensembles positivement invariants, au moyen des hiérarchies moment-SOS. En effet, l'approche du calcul de volumes par les hiérarchies moment-SOS est à l'origine de l'analyse de stabilité par ces mêmes hiérarchies.

Titre en français : Hiérarchie moments-SOS pour approximation ensembliste à grande échelle. Application à l'analyse de stabilité transitoire des systèmes électriques.

Mots clés : Analyse de la stabilité transitoire des systèmes électriques – optimisation en dimension infinie – moments – polynômes sommes de carrés – approximations ensemblistes – méthodes directes pour l'analyse de stabilité – parcimonie – relaxations semidéfinie.

Contents

List of Figures

List of Tables

Notation

This section provides the notations used all along the thesis.

Usual sets

- $\mathbb{N}$: set of natural integers,
- $\mathbb{N}^\star := \mathbb{N} \setminus \{0\}$: set of positive integers,
- $\mathbb{N}^\star_N := \{1, \dots, N\}$: set of N first consecutive positive integers,
- $\mathbb{R}$: set of real numbers,
- $\mathbb{R}_+ := \{x \in \mathbb{R} : x \geq 0\}$: set of nonnegative real numbers,
- $\mathbb{R}_{++} := \{x \in \mathbb{R} : x > 0\}$: set of positive real numbers,
- $\mathbf{I} := [x^-, x^+],\ x^- < x^+ \in \mathbb{R}$: a real interval.

Linear algebra

- $\mathbb{R}^{n \times m}$: space of matrices with n rows and m columns with coefficients in $\mathbb{R}$,
- $m_{i,j}$: for $\mathrm{M} \in \mathbb{R}^{n \times m}$, refers to the coefficient on the i^{th} row and j^{th} column,
- $\mathrm{I}_n \in \mathbb{R}^{n \times n}$: identity matrix, $\mathrm{I} := \mathrm{I}_2$, $\mathrm{J} := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathbb{R}^{2 \times 2}$,
- $\mathrm{Tr}\ \mathrm{M} := \sum_{1 \leq i \leq n} m_{i,i}$: trace of a square matrix M,
- $\mathrm{M}^\top$: transposition of a matrix M,
- $\mathbb{S}^n$: space of real symmetric matrices with n rows ($\mathrm{M}^\top = \mathrm{M}$),
- $\mathbb{S}^n_+$: cone of symmetric positive semi-definite matrices, $\mathrm{M} \in \mathbb{S}^n_+ \Leftrightarrow \mathrm{M} \succeq 0$,
- $\mathbb{S}^n_{++}$: open cone of symmetric positive definite matrices, $\mathrm{M} \in \mathbb{S}^n_{++} \Leftrightarrow \mathrm{M} \succ 0$,
- $\mathrm{M} \preceq 0 \Leftrightarrow -\mathrm{M} \succeq 0$, $\mathrm{M} \prec 0 \Leftrightarrow -\mathrm{M} \succ 0$ and $\mathrm{M} \preceq \mathrm{N} \Leftrightarrow \mathrm{N} - \mathrm{M} \succeq 0$.

Euclidean geometry

- $\mathbf{x} := (x_1, \dots, x_n)^\top \in \mathbb{R}^n$: a real vector with n rows,
- $\mathbf{0}$: zero finite dimensional vector, $\mathbf{x} \geq \mathbf{0} \Leftrightarrow x_i \geq 0, i \in \mathbb{N}^\star_n$,
- $\mathbf{x} \cdot \mathbf{y} := \mathbf{x}^\top \mathbf{y}$: inner product of two finite dimensional real vectors,
- $|\mathbf{x}| := \|\mathbf{x}\|_2 = \sqrt{\mathbf{x} \cdot \mathbf{x}}$: euclidean norm of a real vector $\mathbf{x} \in \mathbb{R}^n$,
- $\mathbf{B}^n := \{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x}| \leq 1\}$: the unit ball of $\mathbb{R}^n$,

- $\mathbf{S}^{n-1} := \{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x}| = 1\}$: the unit sphere of $\mathbb{R}^n$,
- $\mathbf{B}_R := R\,\mathbf{B}^n$ and $\mathbf{S}_R := R\,\mathbf{S}^{n-1}$: the ball and sphere of radius $R > 0$,
- $\pi_{\mathbb{X}} : \mathbb{R}^n \to \mathbb{X}$: orthogonal projection on the vector subspace $\mathbb{X} \subset \mathbb{R}^n$,
- $\mathbb{X}^\perp := \ker \pi_{\mathbb{X}} = \{\mathbf{y} \in \mathbb{R}^n : \forall \mathbf{x} \in \mathbb{X}, \mathbf{x} \cdot \mathbf{y} = 0\}$: vector space orthogonal to $\mathbb{X}$.

Differential analysis

Let $\mathbf{\Omega} \subset \mathbb{R}^n$ be an open or compact set, $k \in \mathbb{N}^\star$.

- $\dot{\mathbf{x}} := \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}$: derivative of the vector function $t \mapsto \mathbf{x}(t)$,
- ∂_x: partial differentiation operator with respect to the variable x,
- $\partial^k_{x_{i_1},\dots,x_{i_k}} := \partial_{x_{i_1}} \cdots \partial_{x_{i_k}}$: k^{th} partial differentiation operator w.r.t. $x_{i_1}, \dots, x_{i_k}$,
- $\partial\,\mathbf{f} := (\partial_{x_j} f_i)_{(i,j) \in \mathbb{N}^\star_m \times \mathbb{N}^\star_n}$: jacobian matrix of function $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$,
- $\mathbf{grad}\ f := (\partial\, f)^\top = (\partial_{x_1} f, \dots, \partial_{x_n} f)^\top$: gradient of function $f : \mathbb{R}^n \to \mathbb{R}$,
- $\operatorname{div} \mathbf{f} := \operatorname{Tr}(\partial\,\mathbf{f})$: divergence of function $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$,
- $\partial^2 f := \partial\,(\mathbf{grad}\ f) = (\partial^2_{x_i,x_j} f)_{i,j \in \mathbb{N}^\star_n}$: hessian matrix of function $f : \mathbb{R}^n \to \mathbb{R}$,
- $\Delta f := \operatorname{Tr}(\partial^2 f) = \operatorname{div}(\mathbf{grad}\ f)$: laplacian of function $f : \mathbb{R}^n \to \mathbb{R}$,
- $\mathcal{C}^0(\mathbf{\Omega}) = \mathcal{C}(\mathbf{\Omega})$: space of continuous functions on $\mathbf{\Omega}$,
- $\mathcal{C}^k(\mathbf{\Omega}) := \left\{ f \in \mathcal{C}(\mathbf{\Omega}) : \mathbf{grad}\ f \in \mathcal{C}^{k-1}(\mathbf{\Omega})^n \right\}$, $\mathcal{C}^\infty(\mathbf{\Omega}) := \bigcap_{l \in \mathbb{N}} \mathcal{C}^l(\mathbf{\Omega})$.

Integration

Let $\mathbf{A} \subset \mathbf{\Omega}$ be a Borel set (countable intersection & union of closed & open sets).

- $\int_{\mathbf{A}} f(\mathbf{x})\ \mathrm{d}\mathbf{x}$: Riemann integral of $f \in \mathcal{C}(\mathbf{\Omega})$ on $\mathbf{A}$.
- $\mathcal{C}_c(\mathbf{\Omega})$: space of continuous functions on $\mathbf{\Omega}$ vanishing outside a compact.
- $\mathcal{M}(\mathbf{\Omega})$: space of signed measures, *i.e.* continuous linear forms on $\mathcal{C}_c(\mathbf{\Omega})$.
- $\int f\ d\mu := \langle f, \mu \rangle$: duality or Lebesgue integral of $f \in \mathcal{C}_c(\mathbf{\Omega})$ w.r.t. $\mu \in \mathcal{M}(\mathbf{\Omega})$.
- $\int_{\mathbf{A}} f\ d\mu := \int \mathbb{1}_{\mathbf{A}}\, f\ d\mu$: integral of $f \in \mathcal{C}_c(\mathbf{\Omega})$ w.r.t. $\mu \in \mathcal{M}(\mathbf{\Omega})$ on $\mathbf{A}$.
- λ: Lebesgue measure s.t. $\forall f \in \mathcal{C}_c(\mathbf{\Omega})$, $\int f\ d\lambda = \int_{\mathbf{\Omega}} f(\mathbf{x})\ \mathrm{d}\mathbf{x}$.
- $\lambda_{\mathbf{A}} := \mathbb{1}_{\mathbf{A}}\, \lambda$: restriction of λ to $\mathbf{A}$ s.t. $\forall f \in \mathcal{C}_c(\mathbf{\Omega})$, $\int f\ d\lambda_{\mathbf{A}} = \int_{\mathbf{A}} f\ d\lambda$.

Algebraic geometry

- $\mathbf{k} := (k_1, \ldots, k_n) \in \mathbb{N}^n$: a multi-index made of n integers,
- $\mathbf{1}_i := (0, \ldots, 0, 1, 0, \ldots, 0)$: multi-index with i^{th} coordinate equal to 1, all others being 0,
- $|\mathbf{k}| := \|\mathbf{k}\|_1 = k_1 + \ldots + k_n$: range of $\mathbf{k} \in \mathbb{N}^n$,
- $\mathbb{N}^n_d := \{\mathbf{k} \in \mathbb{N}^n : |\mathbf{k}| \leq d\}$: index set with bounded range,
- $\mathbf{x}^\mathbf{k} := x_1^{k_1} \cdots x_n^{k_n}$: $\mathbf{k}^{\text{th}}$ power of a vector $\mathbf{x} = (x_1, \ldots, x_n)^\top \in \mathbb{R}^n$,
- $\mathbf{f}^\mathbf{k} := \mathbf{x} \mapsto \mathbf{f}(\mathbf{x})^\mathbf{k}$: $\mathbf{k}^{\text{th}}$ power of a vector function $\mathbf{f} = (f_1, \ldots, f_n)^\top : \mathbb{R}^m \to \mathbb{R}^n$,
- $\mathbb{R}[\mathbf{x}] := \{p(\mathbf{x}) = \sum_{|\mathbf{k}| \leq d} a_\mathbf{k}\, \mathbf{x}^\mathbf{k} : d \in \mathbb{N} \wedge a_\mathbf{k} \in \mathbb{R}\}$: space of polynomials in $\mathbf{x}$,
- $\mathrm{d}^\circ p := \max\{|\mathbf{k}| : a_\mathbf{k} \neq 0\}$: degree of $p \in \mathbb{R}[\mathbf{x}]$,
- $\Sigma[\mathbf{x}] := \{s = p_1^2 + \cdots + p_k^2 : k \in \mathbb{N}^\star \wedge p_1, \ldots, p_k \in \mathbb{R}[\mathbf{x}]\}$: cone of sums of squares of polynomials,
- $\mathbb{R}_d[\mathbf{x}] := \{p \in \mathbb{R}[\mathbf{x}] : \mathrm{d}^\circ p \leq d\}$: space of polynomials of degree at most d,
- $\Sigma_d[\mathbf{x}] := \Sigma[\mathbf{x}] \cap \mathbb{R}_{2d}[\mathbf{x}]$: cone of SOS polynomials of degree at most $2d$.

1 Introduction

In this introductory chapter, we provide an overview of the concepts and problems that we will study in this thesis. Section 1.1 presents the general context of the thesis, namely the need for new tools to assess the stability of industrial power systems. Section 1.2 briefly reviews some of the different existing approaches to power systems stability analysis. In Section 1.3 we discuss in more details the moment-SOS hierarchical approach, which is the method of interest in this thesis. The chapter ends with Section 1.4, listing the contributions of the thesis and outlines the structure of this manuscript.

Contents

1.1 Context and motivation of the thesis

In the wake of the energy transition, large scale electrical power systems are evolving faster and faster, with increasing complexity and stochastic behaviors, mostly due to the massive introduction of partially uncontrollable renewable energy sources as well as corresponding new technologies, especially in power electronics. Among such devices, one can cite High Voltage Direct Current (HVDC) lines [18] as well as power converters [38]. In order to guarantee the functioning of these sophisticated

systems, it is necessary to consider new methods for analysing them, see e.g. [49]. In particular, the stability analysis of nonlinear systems subject to large perturbations has always been an extremely difficult problem, and it is going to complexify even more with the upcoming evolutions. Estimating the largest perturbation that the system can endure without impacting consumers and industrial loads remains a key strategy for large scale electrical power systems management. To that end, the "bruteforce" method would consist of running a large number of simulations of the system's behavior, corresponding to a sample of all possible perturbations, and determining which ones would present a serious threat for system security, and which ones would automatically be subsided by the system controls. Of course, such an approach is not compatible with large scale systems, for which the number of variables and possible perturbations is way too large to be tractable on a computer. Then, new approaches are needed, which should satisfy a certain number of requirements listed below:

Be compatible with nonlinearities: Electrical power systems include nonlinearities brought upon by the presence of alternative current modelled with trigonometric functions, as well as power controls involving bilinearities, and eventually more sophisticated technologies, such as saturations. Since the aim is to assess the system's behavior subject to *large* perturbations, linearization around equilibrium, which is the classical method for *small signal stability*, is not an option here, hence the need for *nonlinear* computational methods.

Avoid false negatives: Given a scenario, we want to decide whether it will lead to an instability or, on the contrary, if it will pose no threat to the grid security. In that case, supposing that only approximate solutions can be given, so that an error is possible, it is crucial to control such error: misclassifying a scenario as "secure" while it actually endangers the power network (*i.e.* a false negative) is forbidden here, as it could have catastrophic consequences. On the contrary, any scenario classified as "unstable" would be subject to further analysis, which would reveal the eventual false positives. In other words, some *certifications* should be given along with the stability analysis.

Provide accuracy guarantees: Although false positives can be allowed, it is important to control their occurrences. Indeed, a false positive would require additional work to be detected, potentially increasing too much the computational burden of the analysis. For this reason, any guarantees that the probability of false positive ultimately vanishes, would be highly appreciated. Most often such requirement translates into *convergence* of the analysing algorithm.

Be (at least potentially) scalable: For low-dimensional systems such as local grids, some methods already exist that will be presented in more details in Section 1.2. The central challenge of this thesis is to pave the way for large scale stability analysis, which supposes that we find a way to tackle hundreds (ideally tens of thousands) of variables in a reasonable amout of time. As we will expose in Section

1.3, the network structure has the potential to allow for *distributed* computations, which would drastically reduce the computational burden of our task.

Be perturbation-independent: Another factor that would increase the computational burden of a stability assessment method is the dependence to the analysed perturbation. Indeed, as the number of possible perturbations increases, the number of required computations would also grow very quickly, eventually leading to an intractability. Consequently, a method free from such a dependence, such as *geometric* stability characterizations, would be much more efficient and suitable.

In this thesis, we will focus on set approximation schemes, mostly based on Lasserre's moments-sums-of-squares (moment-SOS) hierarchy as well as semidefinite programming (SDP). However, several other approaches to power systems stability analysis have been proposed in the past, that we are now going to review.

1.2 Power system stability analysis

1.2.1 Definition on a classical example

For the purpose of illustration let us first focus on the simplest representation of a synchronous machine electromechanical dynamic: the rotating mass or swing equation, see e.g. [90, 21, 6].

Consider a power system composed of N synchronous generators with respective voltages $\mathbf{v}_1, \dots, \mathbf{v}_N$. We assume, as it is common in the literature, that the voltage magnitudes $|\mathbf{v}_1|, \dots, |\mathbf{v}_N|$ are fixed after the fault is cleared, while the phase angles vector $\boldsymbol{\theta} := (\theta_1, \dots, \theta_N)$ is variable (expressed in a rotating frame) with respective angular speed $\boldsymbol{\omega} := (\omega_1, \dots, \omega_N)$. In addition, the loads in the network are considered to be constant and passive impedances. In normal operation conditions, the phases will satisfy the following set of differential equations (with physical variables expressed in SI units here):

$$\dot{\boldsymbol{\theta}} = \boldsymbol{\omega}, \tag{1.1a}$$

$$M_k\, \dot{\omega}_k = -D_k \omega_k + \left(P_k^{\mathrm{mec}} - P_k^{\mathrm{elec}}(\boldsymbol{\theta})\right), \qquad k \in \mathbb{N}_N^\star \tag{1.1b}$$

where P_k^{mec} is the (fixed) mechanical power input at bus k and $P_k^{\mathrm{elec}}(\boldsymbol{\theta})$ is the electrical power output of each generator k with value given by

$$P_k^{\mathrm{elec}}(\boldsymbol{\theta}) = G_{kk}|\mathbf{v}_k|^2 + \sum_{l \neq k} |\mathbf{v}_k|\,|\mathbf{v}_l| \left\{B_{kl} \sin(\theta_k - \theta_l) + G_{kl} \cos(\theta_k - \theta_l)\right\}. \tag{1.2}$$

The quantities B_{kl} and G_{kl} denote the line susceptances and conductances, and M_k refers to the generator inertia constant H_k ($M_k = 2H_k$). The constant D_k denotes the damping coefficient of each generator. Equation (1.1) is called the *swing equation* and models the electromechanical conversion characteristic of the synchronous machine.

We assume that there exists an equilibrium $\overline{\boldsymbol{\theta}} := (\overline{\theta}_1, \dots, \overline{\theta}_N)$ to these equations, that satisfies

$$P_k^{\text{mec}} = P_k^{\text{elec}}(\overline{\boldsymbol{\theta}}), \quad k \in \mathbb{N}_N^\star. \tag{1.3}$$

In other words, $\overline{\boldsymbol{\theta}}$ corresponds to a steady-state operating point of an AC transmission system. As phases are defined up to a reference value, we choose one bus, denoted by subscript "ref", to serve as the reference bus, with $\theta_{\text{ref}} = \overline{\theta}_{\text{ref}} = 0$ (often referred to as slack bus). Indeed, the equations are invariant up to a phase shift.

We first introduce working notations that we will use all along the thesis. Consider a vector field $\mathbf{f} \in \mathcal{C}^1(\mathbb{R}^n)^n$ with equilibrium point $\overline{\mathbf{x}} \in \mathbb{R}^n$ such that $\mathbf{f}(\overline{\mathbf{x}}) = \mathbf{0}$, along with the differential system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}). \tag{1.4}$$

According to the Cauchy-Lipschitz theorem [105, 77], (1.4) admits a continuously differentiable solution map

$$\begin{cases} \mathbb{R} \times \mathbb{R}^n & \longrightarrow & \mathbb{R}^n \\ (t, \mathbf{x}_0) & \longmapsto & \mathbf{x}(t|\mathbf{x}_0) \end{cases}$$

such that $\mathbf{x}(0|\mathbf{x}_0) = \mathbf{x}_0$ (initial condition) and $\partial_t \mathbf{x}(t|\mathbf{x}_0) = \mathbf{f}(\mathbf{x}(t|\mathbf{x}_0))$ (dynamics). Moreover, if $\mathbf{x}(t|\mathbf{x}_0) = \overline{\mathbf{x}}$ holds for one value of t, then it holds for all $t \in \mathbb{R}$ (in particular, $\mathbf{x}_0 = \overline{\mathbf{x}}$).

Then, transients are defined as follows:

Definition 1.1: Transients

Consider a power system (P) described by differential equation (1.4). Suppose that at $t = t_{\text{p}}$ a perturbation occurs that drastically modifies the differential equation, into $\dot{\mathbf{x}} = \widehat{\mathbf{f}}(\mathbf{x})$, $\widehat{\mathbf{f}} \in \mathcal{C}^1(\mathbb{R}^n)^n$. Its trajectory is $\widehat{\mathbf{x}}(t - t_{\text{p}}|\mathbf{x}_0)$, $t \geq t_{\text{p}}$. At $t = t_{\text{cl}} > t_{\text{p}}$, the perturbation is cleared and (P) goes back to nominal equation (1.4). We define the clearing state $\mathbf{x}_{\text{cl}} := \widehat{\mathbf{x}}(t_{\text{cl}} - t_{\text{p}}|\overline{\mathbf{x}})$ and post-perturbation trajectory $\mathbf{x}(t - t_{\text{cl}}|\mathbf{x}_{\text{cl}})$. The behaviour of (P) for $t \geq t_{\text{p}}$ is called *transient*.

In particular, we want to assess the system's transient stability, which is defined as follows:

Definition 1.2: Transient stability

In this thesis, we call *transient stability* of a power system (P) its ability to go back to an operating equilibrium point $\overline{\mathbf{x}}$ from a post-disturbance state $\mathbf{x}_{\text{cl}}$ far from $\overline{\mathbf{x}}$, with the state variables staying in a secure zone of the state space.

This general definition rules out any linearization-based local stability analysis: we are bound to carry out a sharp, nonlinear *transient stability analysis* (TSA).

Remark 1.1 (Power system transient stability)

In terms of power systems, transient stability more specifically denotes a feature of the so-called rotor angle stability, based on the functioning of synchronous

machines described by (1.1) *(see figure 1.1). However, from the mathematical methodological viewpoint, the specificity of transient stability boils down to the notion of large disturbances and the impossibility of linearization. Thus, the works presented in this thesis apply to all stability analyses related to large disturbances, for example in voltage stability. As a result, we included all such stability properties in our definition of transient stability.*

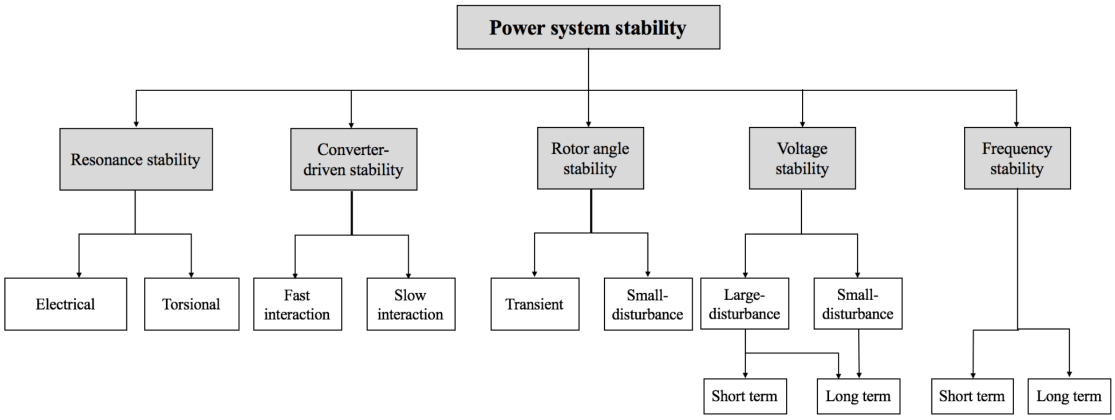


Figure 1.1 – Power systems stability classification.

Image source – [41]; this classification is an upgrade of the one found in [65].

When a perturbation occurs, the dynamics of the system are modified, so that the state leaves its former equilibrium point $\overline{\mathbf{x}}$ and goes into large excursion. For a given perturbation, the longer the perturbation, the further from $\overline{\mathbf{x}}$ the clearing state $\mathbf{x}_{\text{cl}}$. Then, after a certain time, the system state will be too far from equilibrium to be able to go back to equilibrium, even if the perturbation is cleared, and the rotor angles will quickly go to infinity. We call such time the *critical clearing time* t_{cc}.

Definition 1.3: Critical Clearing Time (CCT)

$$t_{\text{cc}} := \max_{\Delta t} \Delta t$$
$$\text{s.t. } \mathbf{x}_{\text{cl}} := \widehat{\mathbf{x}}(\Delta t|\overline{\mathbf{x}}) \text{ is stable in the sense of Definition 1.2.}$$

Critical clearing time is a very useful metric to assess transient stability of a power system under a given perturbation. The CCT depends on the studied perturbation and initial state $\mathbf{x}_0$, so that what we are interested in is the whole function (perturbation, $\mathbf{x}_0$) $\mapsto$ CCT.

Definition 1.4: Critical Clearing State (CCS)

CCT is accompanied with the notion of *critical clearing state (CCS)*, which will also be of interest:

$$\mathbf{x}_{\text{cc}} := \widehat{\mathbf{x}}(t_{\text{cc}}|\overline{\mathbf{x}}).$$

In the context of time-invariant systems (1.4), CCS has the advantage to be independent from the perturbation and the on-fault behavior. It only defines a limit which, if it is passed, compromises the transient stability of the system.

1.2.2 Overview of some existing TSA approaches

In this section we present some of the considered methods for transient stability analysis, based on [90, 117].

Currently used methods

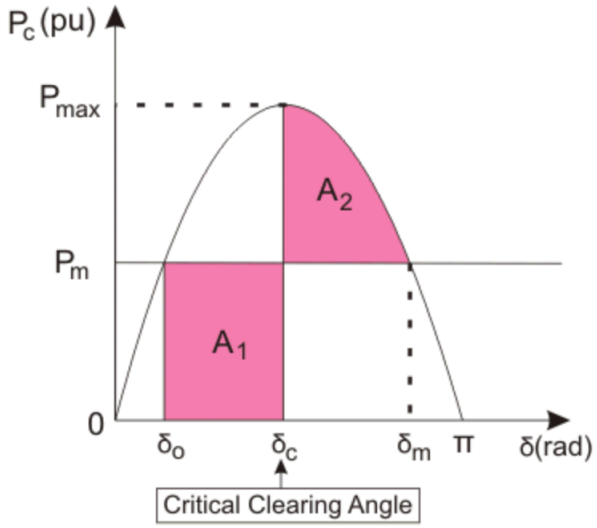


Figure 1.2 – Power-angle characteristic of a synchronous machine.

Corresponding to system (1.1) ($N = 1$) subject to a short-circuit. The plotted curve is the electrical power that the machine should inject into the grid in normal functioning conditions.

Image source – `https://www.electrical4u.com/equal-area-criterion/`

A first basic method for TSA, called the Equal Area Criterion, consists of plotting the graph of $P^{\text{elec}}(\theta)$, denoted $\mathrm{P_c}(\delta)$ in Figure 1.2. Then, one can identify the *critical clearing angle* δ_{c}, such that areas A_1 and A_2 are equal, and t_{cc} is the time at which δ_{c} is attained. Indeed, on the figure δ_0 represents normal functioning conditions (*i.e.* equilibrium), in which the machine injects exactly as much electrical power in the grid as it receives mechanical power: $\mathrm{P_c}(\delta_0) = \mathrm{P_m}$ so that the system (1.1) is at steady-state. However, during the fault (here a short-circuit), the electrical power delivered by the machine to the grid vanishes, so that all mechanical power input $\mathrm{P_m}$ is stored in the rotor as kinetic energy (represented by area A_1 on the figure), making δ sharply increase. Then, it can be returned only after fault clearing, under the form of electrical energy, resulting in at most area A_2. Thus, if $\mathrm{A}_1 \leq \mathrm{A}_2$ all stored mechanical power can be returned to the grid before synchronism is lost, *i.e.* before δ exceeds π, after which δ decreases back to δ_0. On the contrary, if $\mathrm{A}_1 > \mathrm{A}_2$,

then the machine will lose synchronism before all stored energy is returned, resulting in transient instability.

In addition to the intuition that this criterion gives of transient behaviours, it has been extended and studied deeply by power engineers in the past decades, under the name of *Extended Equal Area Criterion* (EEAC, see [140]), although the present thesis does not focus on this TSA method.

As stated before, the general method used to assess transient stability of general power systems consists of numerical time-domain simulations: if one knows *a priori* what the perturbation will be and how long it will last, then one only has to simulate the trajectory of the system for $t \geq t_{\mathrm{p}}$ and see whether or not it goes back to equilibrium after t_{cl}. To that end, equation (1.1) or more generally (1.4) is discretized over time, after which the discretized system is numerically integrated to compute the trajectory of the system. Then, one can directly observe transient stability or instability by checking the result of the time-domain simulation.

However, transmission system operators (TSO) are actually interested not only in deciding the stability of a system but also in assessing stability margins. Indeed, the actual clearing time t_{cl} is fixed, determined by the speed of the involved protection devices. Thus, studying the transient stability of a system boils down to computing t_{cc} and comparing it to the fixed t_{cl}. Then, $t_{\mathrm{cc}} < t_{\mathrm{cl}}$ means by definition that the system will be unstable. However, uncertainties on the model and behaviour of the actual system leads TSOs to look for some kind of robustness propeties for stable situations, under the form of stability margins, *i.e.* lower bounds on the value of $t_{\mathrm{cc}} - t_{\mathrm{cl}}$. For this reason, TSOs are interested in computing the actual CCT for a given perturbation, and not only assessing stability through a single simulation of the whole transients. However, such computation is very costly, as the most straightforward method consists of a bisection that would involve a time-domain simulation at each step.

In order to lighten this computational burden, one can look for randomized solutions, such as the classical Monte-Carlo methods, which consist of sampling the possible scenarios according to their probability, and only computing CCTs associated with the chosen sample. Then, instead of obtaining CCTs for all perturbations, one only has access to an estimation of the probability that the CCT is within a given interval (see [5, 66]). While such a method would surely decrease the computational cost of the TSA, it poses the problem of forecasting rare transient events with severe consequences. Indeed, especially in power grids that have been functioning 24/7, for decades, even very unlikely events have already occurred, and the possibility for randomized approaches to ignore a part of the potential perturbations poses an issue for conservative transient stability analyses. However, such methods have been studied for a long time and paved the way to data-analysis related methods [102].

AI methods

The current times have seen artificial intelligence progressively gain momentum, up to a point where it competes with most of the existing state-of-the art methods. Here we briefly mention different AI-related approaches that have been proposed to

carry out transient stability analyses.

- **Clustering:** A possibility is to train an algorithm to identify classes of perturbations which will lead to similar behaviors, after which one can study the details of a small number of representing events and deduce the various outcomes of a large number of perturbations. Compared to classical randomization, this allows for taking outliers into account as specific clusters. Several contributions were made in this domain, including [37, 19, 98, 55]. Depending on authors, different classification methods were used, such as k-nearest neighbours, bayesian classification or hierarchical clustering (see [40, 146] for details on these classification methods).

- **Artificial Neural Networks (ANN):** Bio-inspired by the functioning of human brain, ANNs consist of automatically splitting a complex task into simpler ones that are ditributed along a pre-determined network of "neurones" (basically boolean classifiers), whose historical ancestor was the Perceptron algorithm. The definition of the subtask is called "training" the ANN and consists of tuning the neurones in a way that minimizes the error rate of the whole, when analysing a "training set" of known perturbations. Such methods were applied in [31, 59, 144, 97].

- **Support Vector Machine (SVM):** Generalizing linear classification, SVM belongs to the category of *lift-and-project* methods; the idea consists of embedding the space in which the TSA problem is considered into a vector space in which discriminating between stable and unstable scenarios is reduced to checking the sign of a linear form one has to determine. Such a method can be combined with classical clustering as was done in [94, 142] or with various optimization methods such as [133, 141].

While AI-based method bear a great potential for power systems TSA, in this thesis we decided to study methods capable of both competing with machine learning approaches and giving a deeper understanding of the physics involved in the transient phenomena. Such methods, called *direct methods* as they do not require any time-domain simulation at all, have been around for a long time and are based on the physical notion of energy.

1.2.3 Our approach: Direct methods for stability regions

Direct methods can be seen as dual to time-domain simulations, in the sense that instead of computing the successive states of the system, one looks for an observable that takes the state as input, and outputs a real value on which the stability analysis is based. This duality will be explained more in details in Chapters 2 and 3. Such direct methods are based on general differential system stability theory [81, 67] and were specifically applied to power systems stability analysis in [6, 21, 22, 53, 99].

Transient stability regions

A promising alternative to the previously reviewed methods consists in determining *a priori* an approximation of a transient stability region for system (1.4), *i.e.* a region $\mathbf{A}$ of the state space for which one can guarantee that if the clearing state $\mathbf{x}_{\text{cl}} \in \mathbf{A}$ then the system will go back to equilbrium within a given time horizon. We now define a natural candidate for a transient stability region. Indeed, we are looking for a geometric criterion for convergence of the post-perturbation system, which leads us to define the region of attraction of a set.

Definition 1.5: Region of Attraction (ROA)

Let $\mathbf{M} \subset \mathbf{X} \subset \mathbb{R}^n$, $T \in (0, +\infty]$. We define the *time T region of attraction* (ROA) of $\mathbf{M}$ (with state constraint $\mathbf{X}$) as:

$$\mathbf{A}_T^{\mathbf{X}}(\mathbf{M}) := \left\{ \mathbf{x}_0 \in \mathbb{R}^n : \begin{array}{l} \forall t \in [0, T), \mathbf{x}(t|\mathbf{x}_0) \in \mathbf{X} \\ \text{dist}(\mathbf{x}(t|\mathbf{x}_0), \mathbf{M}) \underset{t \to T}{\longrightarrow} 0 \end{array} \right\}. \tag{1.5}$$

where $\text{dist}(\mathbf{x}, \mathbf{Y}) := \inf_{\mathbf{y} \in \mathbf{Y}} |\mathbf{x} - \mathbf{y}|$ is the euclidean distance between a point $\mathbf{x} \in \mathbb{R}^n$ and a set $\mathbf{Y} \subset \mathbb{R}^n$. If $\mathbf{X} = \mathbb{R}^n$ then we only write $\mathbf{A}_T(\mathbf{M}) := \mathbf{A}_T^{\mathbb{R}^n}(\mathbf{M})$.

This definition allows us to state a transient stability analysis problem:

Problem 1: Direct transient stability analysis

Compute an approximating subset $\widehat{\mathbf{A}}$ of some well-chosen ROA $\mathbf{A}_T^{\mathbf{X}}(\mathbf{M})$, so that $\mathbf{x}_{\text{cl}} \in \widehat{\mathbf{A}} \Longrightarrow \mathbf{x}_{\text{cl}}$ can be attained without compromising stability.

Remark 1.2 (False negatives and conservativeness)

*In Problem 1, we already allow ourselves to compute an **approximation** of some ROA. Indeed, in most of the general cases, the exact $\mathbf{A}_T^{\mathbf{X}}(\mathbf{M})$ is out of reach for standard computational methods, hence the resort to approximation.*

*However, we stated at the beginning of this Chapter that a good TSA should avoid false negatives. In terms of ROA approximation, this means that we want to compute **inner** ROA approximations, so that we can miss stable scenarios ($\mathbf{A}_T^{\mathbf{X}}(\mathbf{M}) \setminus \widehat{\mathbf{A}} \neq \varnothing$ is allowed), but we cannot miss instability issues ($\widehat{\mathbf{A}} \setminus \mathbf{A}_T^{\mathbf{X}}(\mathbf{M}) \neq \varnothing$ is forbidden).*

Remark 1.3 (Access to CCT)

*Given a solution to Problem 1, one only has access to potential critical clearing **states**. Relating such result to critical clearing **times** requires some kind of time-domain simulation of the faulted system. In other words, direct methods mostly allow their user to get rid of the **post-fault** trajectory simulations. For this reason, ideally, the boundary of the computed transient stability region should be close to a set of critical clearing states. Other ways to access CCT are not reviewed in the present thesis.*

Stability oracle functions

What we call stability oracle functions is a most general class of functions whose value at state $\mathbf{x}_{\text{cl}}$ gives insight on the stability of the trajectory $\mathbf{x}(t|\mathbf{x}_{\text{cl}})$. Such information can in turn be used to assess the transient stability of the studied system. Here we only give the most general definition, as several more specific stability oracles will be studied in the rest of the thesis.

Definition 1.6: Stability Oracle Functions (SOFs)

Let $\mathbf{D} \subset \mathbb{R}^n$. A *stability oracle function (SOF)* is a $v : \mathbb{R}^n \to \mathbb{R}$ non-increasing along trajectories $\mathbf{x}(t|\mathbf{x}_0)$ in $\mathbf{D}$, which is characterized by

$$\forall \mathbf{x} \in \mathbf{D}, \mathbf{f}(\mathbf{x}) \cdot \mathbf{grad}\, v(\mathbf{x}) \leq 0. \tag{1.6}$$

Remark 1.4 (Positively invariant sublevel sets)
A direct consequence of equation (1.6) *is that any sublevel set*

$$\mathbf{\Omega} := \{\mathbf{x} \in \mathbb{R}^n : v(\mathbf{x}) \leq l\}$$

with $l \in \mathbb{R}$ *such that* $\mathbf{\Omega} \subset \mathbf{D}$*, is positively invariant for system* (1.4).

Example 1.5

- *Lyapunov functions, whose definition will be recalled in Section 3.2, are an instance of SOF, related to* ***infinite*** *time ROA.*
- *Energy functions [22] are an instance of SOF, related to* ***infinite*** *time ROA.*
- *Dual decision variables of [42] are an instance of SOF that we will study in Section 3.1, related to* ***finite*** *time ROA.*
- *Control barrier functions [4] are an instance of SOF, related to* ***positively invariant sets****.*

Lyapunov arguments are based on the fact that a stable equilibrium point of a physical system (such as a power network) is characterized as a local minimum of its energy. As transient stability requires that the system converges to an equilibrium after the perturbation is cleared, TSA then reduces to deciding whether the energy of a system in post-fault state $\mathbf{x}_{\text{cl}}$ will decrease to a minimum or not. As it is well known, set $\mathbf{\Omega}$ from Remark 1.4, when considering a specific kind of Lyapunov function, is a subset of the Lyapunov ROA $\mathbf{A}_\infty(\{\overline{\mathbf{x}}\})$ of Definition 1.5. Of course, depending on the considered SOF, set $\mathbf{\Omega}$ will have different properties. Finding relevant SOF for power systems TSA is part of this work, more specifically addressed in Chapters 3 and 6.

Then, the crucial question remains: how does one compute such functions? Due to the physical inspiration of direct methods, the first studied oracles were energy functions, followed by Lyapunov functions, and most of the time they were determined out of physical reasonings or after some analytical computations [22, 131]. However, starting in the early 2000s, some algorithmic methods were developped to:

- Systematize the computation of SOFs,
- Optimize over SOFs to obtain the "best" transient stability region estimates.

We now introduce some mathematical concepts that will be instrumental in our computations of transient stability region estimates.

1.3 The moment-SOS approach

1.3.1 Introduction to measures

This section is addressed to a reader unfamiliar with the mathematical notion of measure. It is an intent to give an intuition on the concept of measure, without introducing the sophisticated set theory elements that are necessary for a rigorous definition. In contrast, here we build only on the knowledge of normed vector spaces, linear forms and continuity. Given a real vector space $\mathcal{X}$ equipped with a norm $\|\cdot\|$, we recall that:

- A linear form on $\mathcal{X}$ is a function $\phi : \mathcal{X} \to \mathbb{R}$ s.t. $\forall \psi, \psi' \in \mathcal{X}$, $s \in \mathbb{R}$, $\phi(\psi + s\,\psi') = \phi(\psi) + s\,\phi(\psi')$; we denote $\langle \psi, \phi \rangle := \phi(\psi)$.
- A continuous function on $\mathcal{X}$ is a function $f : \mathcal{X} \to \mathbb{R}$ s.t.

$$\|\psi - \psi_k\| \underset{k\to\infty}{\longrightarrow} 0 \Longrightarrow |f(\psi) - f(\psi_k)| \underset{k\to\infty}{\longrightarrow} 0.$$

- $\mathcal{X}'$ is the set of continuous linear forms on $\mathcal{X}$, also called *dual space*[1].

Consider the set $\mathcal{C}_c(\mathbb{R}^n)$ of continuous functions on $\mathbb{R}^n$ that vanish outside a ball. This is an infinite dimensional real vector space, equipped with the uniform norm

$$\|f\|_\infty := \sup_{\mathbf{x}\in\mathbb{R}^n} |f(\mathbf{x})|.$$

We define the dual space $\mathcal{M}(\mathbb{R}^n) := \mathcal{C}_c(\mathbb{R}^n)'$, and call it the space of *signed measures*. For $\mu \in \mathcal{M}(\mathbb{R}^n)$ and $f \in \mathcal{C}_c(\mathbb{R}^n)$ we define the *integral* of f w.r.t. μ as

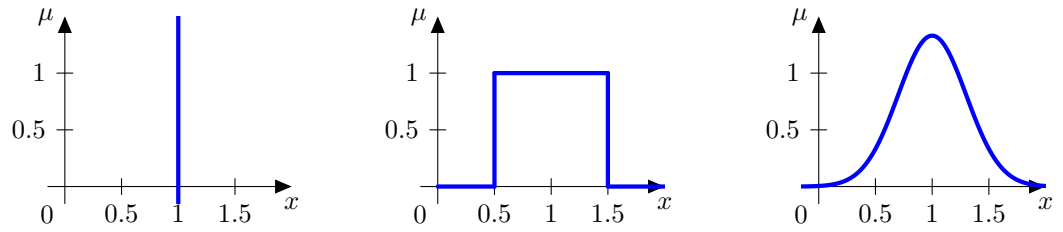
$$\int f \, d\mu := \langle f, \mu \rangle.$$

Then, we define the set of nonnegative functions

$$\mathcal{C}_c(\mathbb{R}^n)_+ := \{f \in \mathcal{C}_c(\mathbb{R}^n) : \forall \mathbf{x} \in \mathbb{R}^n, f(\mathbf{x}) \geq 0\}$$

as well as its dual $\mathcal{M}(\mathbb{R}^n)_+ := \{\mu \in \mathcal{M}(\mathbb{R}^n) : \forall f \in \mathcal{C}_c(\mathbb{R}^n)_+, \int f \, d\mu \geq 0\}$. Those are cones in the sense that they are invariant through multiplication by a positive

[1] If $\mathcal{X}$ has finite dimension, then all linear forms are continuous, and the Riesz representation theorem states that $\mathcal{X}'$ can be identified with $\mathcal{X}$; however, it is not the case for infinite dimensional vector spaces.



(a) Dirac measure in $x = 1$.

(b) Lebesgue measure on $[0.5, 1.5]$.

(c) Gaussian law with mean 1 and standard deviation 0.3.

Figure 1.3 – Representation of classical measures.

number. $\mathcal{M}(\mathbb{R}^n)_+$ is the cone of *measures* on $\mathbb{R}^n$. Eventually, for $\mathbf{X} \subset \mathbb{R}^n$, we define the sets

$$\mathcal{M}(\mathbf{X}) := \left\{ \mu \in \mathcal{M}(\mathbb{R}^n) : \text{if } f \in \mathcal{C}(\mathbb{R}^n) \text{ vanishes on } \mathbf{X} \text{ then } \int f \, d\mu = 0 \right\},$$

$$\mathcal{M}(\mathbf{X})_+ := \left\{ \mu \in \mathcal{M}(\mathbb{R}^n)_+ : \text{if } f \in \mathcal{C}(\mathbb{R}^n) \text{ vanishes on } \mathbf{X} \text{ then } \int f \, d\mu = 0 \right\},$$

which correspond to measures supported on $\mathbf{X}$.

From a physical viewpoint, measures represent mass distributions, or superpositions of points. As a result, they can encode many things, from probability laws (taking a random variable X of law μ, $\int f \, d\mu = \mathbb{E}[f(\mathrm{X})]$) to superpositions of trajectories of a dynamical system, including distributions of solutions to a given optimization problem.

Example 1.6

- *The Dirac distribution $\delta_{\mathbf{x}}$ represents a unitary mass concentrated in $\mathbf{x} \in \mathbb{R}^n$, so that $\int f \, d\delta_{\mathbf{x}} = f(\mathbf{x})$ (Fig. 1.3a).*
- *The Lebesgue measure $\lambda_{[a,b]}$ represents a uniform mass distribution on the segment $[a, b] \subset \mathbb{R}$, so that $\int f \, d\lambda_{[a,b]} = \int_a^b f(x) \, \mathrm{d}x$ (Fig. 1.3b).*
- *The Gaussian measure with mean m and standard deviation s represents a normal probability distribution μ such that (see Fig. 1.3c)*

$$\int f \, d\mu = \frac{1}{s\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{e}^{-\frac{(x-m)^2}{2s^2}} f(x) \, \mathrm{d}x.$$

Some measures, such as the Gaussian measure, are absolutely continuous w.r.t. the Lebesgue measure, meaning that they correspond to actual functions on $\mathbb{R}^n$: they represent densities of mass. Others are not absolutely continuous: for example, the Dirac measure does not represent a density of mass, but rather a singularity with mass 1.

In short, measures have two major characteristics:

- They are the mathematical formalization of point distributions, and hence appear in a large variety of domains, from probability theory to fluid mechanics and statistical physics; in particular, trajectories of a differential system can be encoded under the form of a measure [143, 111, 112, 130, 76];
- They are in duality with functions (such as polynomials as well as stability oracle functions) through Lebesgue's integration theory: indeed, the integral of a function is always computed *with respect to a measure* (in the case of the standard integral, the involved measure is the Lebesgue measure).

Remark 1.7 (Link between measures and stability analysis)
In particular, measures are the mathematical concept that explains the relationship between time-domain simulations and direct methods for stability analysis. Indeed, on the one hand, as a formalization of point distributions, measures can be used to represent trajectories, that are nothing more than time-state point distributions. On the other hand, Lebesgue integration theory allows for integrating SOFs with respect to such trajectory representing measures, introducing a duality between trajectories and SOFs.

As polynomials are continuous functions, $\int p \, d\mu$ represents a particular duality between $p \in \mathbb{R}[\mathbf{x}]$ and $\mu \in \mathcal{M}(\mathbb{R}^n)$. Such a duality is instrumental for implementing the moment-SOS hierarchy. Indeed, historically, moment-SOS hierarchies were made possible by new theorems on polynomials, that exploited their duality with measures.

1.3.2 A brief history

The moment-SOS or Lasserre hierarchies are the fundamental tools that we will be studying and using all along this thesis (except in Section 3.2). They are an elegant framework that brings together real algebraic geometry as well as functional analysis, to define a plug-and-play scheme for solving a very large variety of problems.

The famous 17^{th} Hilbert problem [46] (solved by Emil Artin in [8]) is at the root of the moment-SOS approach:

"Can a nonnegative rational function be written as a sum of squares of rational functions?"

More precisely, Hilbert had proved that the statement is false if one considers polynomials instead of rational functions [45], which Motzkin complemented with a counterexample [93], nonnegative but not sum of squares:

$$p_{\text{M}}(x, y) := x^4y^2 + x^2y^4 - 3x^2y^2 + 1.$$

However, the question of identifying polynomials asked to be positive only on a subset of $\mathbb{R}^n$ still remained to be studied, and motivated numerous works in real algebraic geometry, leading to the so-called Positivstellensätze [60, 119, 115, 107]. A P-satz reduces the question of positivity of a polynomial over a given subset of $\mathbb{R}^n$ to finding P-satz *certificates*, involving sums of squares of polynomials (SOS, see Chapter 2 for details). In parallel, equivalence between SOS characterization

and spectral analysis of symmetric matrices was also established [70], allowing for numerical characterization of positivity over subsets of $\mathbb{R}^n$.

These results were put together, along with fundamental functional analysis arguments, in 2001 to solve the *polynomial optimization problem (POP)* using Putinar's P-satz [68]: Lasserre's moment-SOS hierarchy was born. The moment-SOS hierarchy is presented in depth in Chapter 2, along with some original contributions of this thesis.

Briefly, it consists in modelling a difficult problem using Borel measures such that, if the analysis is carried out properly, the initial difficult problem is rephrased as a *linear* problem on measures. Then, a series of arguments (which we detail in Chapter 2) based on Putinar's P-satz, makes possible to approximate this infinite dimensional linear problem with a sequence of finite dimensional relaxed problems, that can individually be solved numerically by a computer.

Remark 1.8 (Moment-SOS hierarchies and AI)
It is possible to categorize the moment-SOS hierarchy as a lift-and-project method (similar to previously discussed SVM). The lift part consists of the infinite dimensional modelling with Borel measures, and the project part is the formulation and solution of the finite dimensional relaxed problems.

For now, the crucial point is that the high level of abstraction at which the moment-SOS hierarchy is formulated allowed to apply it in a large number of very different domains:

- After contributing to the POP, the moment-SOS hierarchy was used to solve the optimal control problem in [72] in 2008. Indeed, from optimizing over a static point of $\mathbb{R}^n$, a natural extension was optimization over trajectories of a control system with values in $\mathbb{R}^n$. This work was a combination of measure modelling of trajectories with moment-SOS hierarchies, and a duality was proved between the involved measures and solutions to the Hamilton-Jacobi-Bellman equation [95].

- After that, the problem of computing the volume of a semialgebraic set $\mathbf{K}$ was addressed in [44]. This time, the formulation as an optimization problem was only a heuristic to get the volume as optimal solution. An interesting byproduct of this framework was its duality with functions whose unit sublevel set approximated $\mathbf{K}$; we call this feature the set approximation property. Due to this property, the present thesis also focused on this seminal set approximation problem.

- Eventually, combining trajectory-measure modelling from [72] with the set approximation property stated in [44] allowed to make a decisive contribution in the field of direct methods for stability analysis of polynomial control systems, with references [42, 58, 57], in which inner and outer approximations are computed for various instances of the constrained region of attraction.

Comprehensive details on the polynomial optimization problem, as well as the three aforementioned problems, will be given throughout this thesis.

1.3.3 Moment-SOS hierarchies and power systems

In this section we highlight the very general applicability of the moment-SOS hierarchies to power systems, through a fundamental example. The first application of moment-SOS hierarchies to the study of power systems was not related to transient stability considerations, but to the so-called Optimal Power Flow (OPF) problem.

Consider a power network with several power sources as well as loads, modelled as an AC power system. The AC-OPF problem consists of matching power consumption and production while minimizing the energy loss in the process.

The optimal power flow can be written as

$$\begin{aligned} \mathrm{OPF}^\star = \min_{\mathbf{v}_d, \mathbf{v}_q} \; & f(\mathbf{v}_d, \mathbf{v}_q) \\ \text{s.t. } & \mathbf{p}^{\min} \leq \mathbf{p}(\mathbf{v}_d, \mathbf{v}_q) \leq \mathbf{p}^{\max} \\ & \mathbf{q}^{\min} \leq \mathbf{q}(\mathbf{v}_d, \mathbf{v}_q) \leq \mathbf{q}^{\max} \\ & \mathbf{v}_{\mathrm{mag}}^{\min} \leq \mathbf{v}_{\mathrm{mag}}(\mathbf{v}_d, \mathbf{v}_q) \leq \mathbf{v}_{\mathrm{mag}}^{\max} \\ & \mathbf{s}_{\mathrm{mag}}(\mathbf{v}_d, \mathbf{v}_q) \leq \mathbf{s}_{\mathrm{mag}}^{\max} \end{aligned}$$

where f is a convex quadratic cost function, the decision variables $\mathbf{v}_d$ and $\mathbf{v}_q$ represent the direct and quadratic coordinates of the voltage at each node of the grid, $\mathbf{p}$, $\mathbf{q}$ and $\mathbf{v}_{\mathrm{mag}}$ are vectors of nonconvex, degree 2 polynomial functions representing respectively active power, reactive power and voltage magnitude at each node of the power grid, and $\mathbf{s}_{\mathrm{mag}}$ is a vector of nonconvex, degree 2 polynomial functions representing apparent-power line-flows through each line of the power grid (see e.g. [92] and the comprehensive monography [91]).

This is a particular instance of the polynomial optimization problem, with convex polynomial objective f and nonconvex feasible set, making the OPF problem hard to solve in general. Within the literature, one can cite a variety of works that are referenced in surveys [48, 101, 35, 148]. Most recent contributions consist of relaxing the problem to make it more tractable in practice, see e.g. [75, 79, 80]. Among those works, we highlight the use of the moment-SOS hierarchy to formulate a hierarchy of convex semidefinite relaxations of the AC-OPF problem, that are studied in [92, 39]. In these papers, the authors study in depth the potential that the moment-SOS hierarchy bears for solving *exactly* the AC-OPF problem.

Also, building upon [69, 132], the sparse structure of large scale power systems was related to mathematical sparsity patterns such as correlative sparsity, allowing for approximate numerical solution of AC-OPF instances with ~ 10000 variables in [52]. In a similar direction, it is worth mentioning [1, Chapter 5] for the use of graph decomposition for Lyapunov stability analysis.

The aim of this thesis is to contribute to the application of moment-SOS hierarchies to real-life power system transient stability analysis problems. It has been an opportunity for several contributions in various domains, from generic, theoretical questions on the convergence of the moment-SOS schemes, to pratical stability analysis of differential systems, and including considerations on the problem of computing the volume of a semialgebraic set.

1.4 Publications and outline

1.4.1 Thesis organization

The thesis is organized as follows:

- We conclude this introductory chapter with a list of submitted, accepted and published contributions.

- Chapter 2 progressively introduces the concept of moment-SOS hierarchy in more detail. We first focus on the general method of infinite dimensional, measure-based modelling in Section 2.1, with two introductory examples (POP, **K**-moment problem) followed by a general formulation of the problem and some general contributions on the structure of such problems. Then, we introduce the actual moment-SOS hierarchy, which consists of two dual approximating hierarchies (moment relaxations and sum-of-squares strengthenings), and we state all the results that we will build upon in the rest of the thesis. Most of the contributions in this chapter were only formulated in particular cases such as polynomial optimization or volume approximation, but for the first time we give systematic proofs of the classical theorems concerning the general moment-SOS hierarchy. The contributions of this chapter were accepted for publication in [122].

- Chapter 3 then proceeds to the problem of interest here, namely the problem of hierarchical power systems transient stability analysis. While both hierarchical stability analysis and power system TSA have been studied for years, contributions that bring together both fields are scarce. Actually, apart from the contributions in the present thesis, it seems that [6] was the first attempt in such direction, drawing upon SOS programming for Lyapunov stability analysis of power system (1.1). In Section 3.1 (based on publication [53]), we introduce the moment-SOS based method that was developed in [42] and adapt it to the study of the non-polynomial system (1.1) in a similar fashion as in [6], after what Section 3.2 (published in [125]) pushes the method of [6] to its computational limits, by enhancing the considered model with voltage and mechanical power regulations.

- In Chapter 4, we switch to a more theoretical question, namely the volume approximation problem. The link between this problem and hierarchical TSA lies in what we call the "set approximation property", a feature of the hierarchical volume approximation method hilghlighted in [44], and which inspired the works of [42]. More precisely, this chapter is dedicated to enhancing the convergence properties of the moment-SOS hierarchy by improving it with additional constraints deduced from differential geometry arguments. Interestingly, the demonstration of the convergence improvement involves basic results on partial differential equations. The contributions of this chapter were submitted for publication as [124].

- Then, Chapter 5 introduces the notion of sparsity, again in the particular context of volume approximation, which is simpler than stability analysis. A new

scheme is designed to exploit correlative sparsity patterns in the description of the set whose volume we intend to approximate. This scheme allows us to partially parallelize the computations and organize them along a tree, with a transfer of information from the leaves to the root of the tree, at which a solution to the original problem is computed. A first, very simple path decomposition structure is studied with examples in dimension up to 100, after which we formulate a global scheme for generic correlative sparsity patterns, illustrated on some relevant examples. The corresponding paper [126] was recently accepted for publication.

- Eventually, most contributions of the previous chapters are used as an inspiration for Chapter 6. Section 6.1 focuses on the inner approximation of Maximal Positively Invariant sets that was published in [100], and the properties of the corresponding moment-SOS hierarchy are directly deduced from results of Chapter 2. Then, Section 6.2 (published as [123]) consists of a first intent to generalize the results of Chapter 5 on sparsity-exploiting volume computation, to the problem of transient stability region approximation. In this first attempt the sparsity pattern of interest is again a path decomposition structure that paves the way for general correlative sparsity structures.

1.4.2 List of publications

Published in peer-reviewed international journals

[100] A. Oustry, **M. Tacchi** and D. Henrion. Inner approximations of the maximal positively invariant set for polynomial dynamical systems. *IEEE Control System Letters*, 3(3):733–738, 2019.

[126] **M. Tacchi**, T. Weisser, J. B. Lasserre and D. Henrion. Exploiting sparsity in semi-algebraic set volume computation. *Journal of Foundations of Computational Mathematics*. Published online in Feb. 2021. arXiv: 1902.02976.

[122] **M. Tacchi**. Convergence of Lasserre's hierarchy: the general case. *Optimization Letters*. Published online in June 2021. arXiv: 2011.08139

Published in peer-reviewed international conferences

[125] **M. Tacchi**, B. Marinescu, M. Anghel, S. Kundu, S. Benahmed. Power system transient stability analysis using sum-of-squares programming. In *Proceedings of the Power Systems Computation Conference*. IEEE, 2018.

[53] C. Josz, D. K. Molzahn, **M. Tacchi** and S. Sojoudi. Transient stability analysis of power systems via occupation measures. In *Proceedings of the 10th Annual Innovative Smart Grid Technologies Conference*. IEEE, 2019.

[123] **M. Tacchi**, C. Cardozo, D. Henrion and J. B. Lasserre. Approximating regions of attraction of a sparse polynomial differential system. In *Proceedings of the 21st IFAC World Congress*, 2020.

Submitted to peer-reviewed international journals

[124] **M. Tacchi**, J. B. Lasserre and D. Henrion. Stokes, Gibbs and volume computation of semialgebraic sets. Submitted in Sept. 2020. arXiv: 2009.12139

2

Numerical analysis of moment problems

This preliminary chapter is an intent to summarize the mathematical tools that will be used all along this thesis. Briefly, most of the present work is based on the following methodology:

1) Formulating mathematical questions that correspond to the engineering challenge we want to address,

2) Modelling these mathematical questions under the particular form of the generalized moment problem (GMP),

3) Deploying the moment-SOS technology to design schemes that will give approximate solutions to the GMP.

Step 3) in this methodology does not depend on the engineering and mathematical questions that arise in the process (these questions will be the subject of the next chapters). For this reason, we give in the present chapter all the theoretical tools that will be needed to implement this last step in practice: Section 2.1 introduces the generic problem that our numerical schemes will approximate, along with two illustrating examples (polynomial optimization and **K**-moment problem), while Section 2.2 details the hierarchical approximation method that we use. Each section ends with a technical subsection, marked with a $\star$.

The original results contained in this chapter were submitted for publication in [122].

Contents

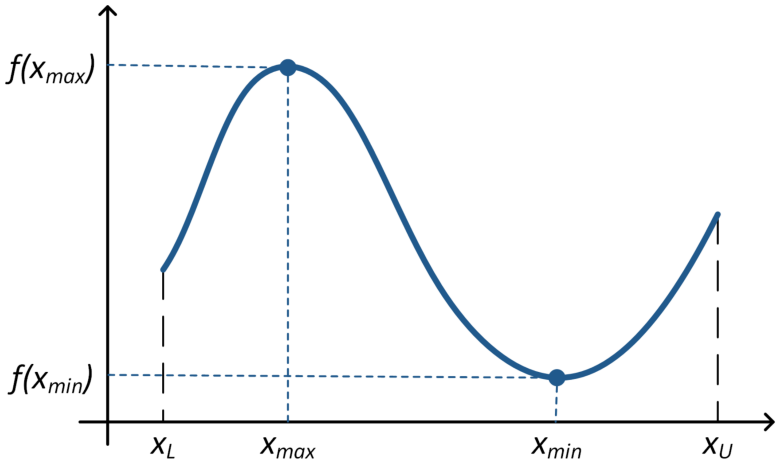


Figure 2.1 – An example of global polynomial optimization problem.

Here $n = 1$ and $\mathbf{K} = [x_L, x_U]$ corresponds to $m = 1$ and $g(x) = (x_U - x)(x - x_L)$. As a continuous function on the compact $\mathbf{K}$, f is bounded and attains its bounds.
Image source – Freiburg university, lecture on real optimization, from the computational economics program.

2.1 The generalized moment problem

This section intends to introduce step by step the concept of *generalized moment problem* (GMP). We first present two particular instances of GMP: first the polynomial optimization problem (POP), which is related to pratical computations involved in engineering and decision making, and then the $\mathbf{K}$-moment problem, being at the root of moment-based numerical analysis. We then proceed to give a most general formulation of the GMP, and we end this section with a synthetic theorem that will be a fundamental tool for all the numerical analyses presented in this thesis. Most of the notions and results presented in this section can be found in a detailed fashion in [70].

2.1.1 Global polynomial optimization

A very common instance of the GMP is the global polynomial optimization problem, which consists in looking for a system configuration (represented by a vector of $\mathbb{R}^n$) that minimizes a given (polynomial) cost function under some given (polynomial) state constraints.

Problem 2: Global polynomial optimization (POP)

Let $f \in \mathbb{R}[\mathbf{x}]$ and $\mathbf{g} \in \mathbb{R}[\mathbf{x}]^m$ such that $\mathbf{K} := \{\mathbf{x} \in \mathbb{R}^n : \mathbf{g}(\mathbf{x}) \geq \mathbf{0}\}$ is compact and nonempty. We consider the problem of minimizing f over $\mathbf{K}$:

$$\begin{aligned} f^\star := \inf_{\mathbf{x}} \ & f(\mathbf{x}) \\ \text{s.t. } & \mathbf{x} \in \mathbf{K}. \end{aligned} \tag{2.1}$$

This problem is known as the *global polynomial optimization problem* (POP).

Problem 2 has many practical applications among which one can cite the Optimal Power Flow (OPF) problem (see [48, 101, 35, 148, 39] and references therein), but is difficult to solve in practice since neither f nor $\mathbf{K}$ are assumed to be convex (for a detailed overview of the approaches to the particular OPF problem, see [91]). This makes it challenging in the field of optimization.

A classical approach to global POP consists in using the definition of the infimum $f^\star$ as the greatest lower bound of f on $\mathbf{K}$ to reformulate problem (2.1) as:

$$\begin{aligned} f^\star = \mathrm{d}_f^\star := \max_{\ell} \ & \ell \\ \text{s.t. } & \forall \mathbf{x} \in \mathbf{K}, f(\mathbf{x}) \geq \ell \\ & \ell \in \mathbb{R}. \end{aligned} \tag{2.2}$$

Note that problem (2.2) is the *definition* of the inf operation, so it is litteraly equivalent to (2.1). We end up faced to a question that is known to be difficult: deciding whether a polynomial is nonnegative on a whole set. However, compared to (2.1), (2.2) has the advantage to be linear in ℓ and always admit a unique minimizer $\ell^\star = f^\star$.

Another approach to problem 2 relies on randomization: given a random variable $\mathrm{X} \in \mathbf{K}$ with probability law $\mathbb{P} \in \mathcal{M}(\mathbf{K})_+$, one intends to minimize the expected value $\mathbb{E}[f(\mathrm{X})]$ of $f(\mathrm{X})$, where the decision variable is the probability distribution $\mathbb{P}$ of X. Such a problem is formulated as follows:

$$\begin{aligned} f^\star = \mathrm{p}_f^\star := \min_{\mathbb{P}} \ & \mathbb{E}[f(\mathrm{X})] \\ \text{s.t. } & \mathrm{X} \sim \mathbb{P} \\ & \mathbb{P} \in \mathcal{P}(\mathbf{K}), \end{aligned} \tag{2.3}$$

where $\mathcal{P}(\mathbf{K}) := \{\mathbb{P} \in \mathcal{M}(\mathbf{K})_+ : \mathbb{P}(\mathbf{K}) = 1\}$ is the set of probability measures on $\mathbf{K}$.

The obvious optimal solution to (2.3) would be a deterministic X constantly equal to a minimizer of f on $\mathbf{K}$: noting $\operatorname{argmin}_{\mathbf{K}} f := \{\mathbf{x} \in \mathbf{K} : f(\mathbf{x}) = f^\star\}$, if $\mathbf{x}^\star \in \operatorname{argmin}_{\mathbf{K}} f$, then the Dirac probability distribution $\mathbb{P} = \delta_{\mathbf{x}^\star}$ yields $\mathrm{X} = \mathbf{x}^\star$ almost surely, so that $\mathbb{E}[f(\mathrm{X})] = f(\mathbf{x}^\star) = f^\star$. More generally, any $\mathbb{P}^\star \in \mathcal{P}(\operatorname{argmin}_{\mathbf{K}} f)$ is a minimizer for problem (2.3).

The striking fact about formulations (2.2) and (2.3) of problem 2, is that they are two sides of the same coin. Indeed, both of them can be rewritten under the form of semi-infinite linear programming problems that are dual to one another:

$$\begin{aligned} \mathrm{p}_f^\star = \min_{\mu} \ & \int f \, d\mu \\ \text{s.t. } & \mu \in \mathcal{M}(\mathbf{K})_+ \\ & \mu(\mathbf{K}) = 1 \end{aligned} \tag{2.4a}$$

$$\begin{aligned} \mathrm{d}_f^\star = \max_{\ell} \ & \ell \\ \text{s.t. } & f - \ell \in \mathcal{C}(\mathbf{K})_+ \\ & \ell \in \mathbb{R} \end{aligned} \tag{2.4b}$$

where we recall that $\mathcal{M}(\mathbf{K})_+$ is the cone of nonnegative measures on $\mathbf{K}$ and $\mathcal{C}(\mathbf{K})_+$ is the cone of nonnegative continuous functions on $\mathbf{K}$.

Here the decision variable ℓ is the Lagrange multiplier corresponding to the (scalar) linear constraint that μ is a probability measure, and conversely the decision variable μ is the Lagrange multiplier corresponding to the (infinite dimensional) conic constraint that $f - \ell$ is a nonnegative continuous function on $\mathbf{K}$.

Remark 2.1 (Retrieving minimizers)
Both of these problem formulations not only give access to the optimal value $f^\star$, but also to the minimizers. Indeed, if $\mu^\star$ and $\ell^\star$ are minimizers for problems (2.4a) *and* (2.4b) *respectively, then* $\operatorname{argmin}_{\mathbf{K}} f$ *is exactly the set of roots of* $f - \ell^\star$ *and contains the support of* $\mu^\star$. *In particular, the uniform probability measure on* $\operatorname{argmin}_{\mathbf{K}} f$ *is an optimal solution for* (2.4a) *that gives access to the whole set* $\operatorname{argmin}_{\mathbf{K}} f$ *as its support.*

In this section we have shown that $\mathrm{p}_f^\star = f^\star = \mathrm{d}_f^\star$, which means that there is no duality gap between problems (2.4a) and (2.4b).

Historically, POP was the motivation for the design of the Lasserre, moment-sum-of-squares hierarchy that we will present in details later on. The duality properties mentioned in the above are a recurrent, fundamental feature of moment-SOS hierarchies, as we will show. For now, we proceed to our second illustrating example of the GMP.

2.1.2 The K-moment problem

We consider a more theoretical problem that is a key for measure-based numerical analysis, namely the $\mathbf{K}$-*moment problem*. It is formulated as follows:

Problem 3: K-moment problem

Given nonempty sets $\mathbf{K} \subset \mathbb{R}^n$, $\mathbf{\Gamma} \subset \mathbb{N}^n$ and a real sequence $\mathbf{z} := (z_\mathbf{k})_{\mathbf{k}\in\mathbf{\Gamma}} \in \mathbb{R}^{\mathbf{\Gamma}}$, is there a measure $\mu \in \mathcal{M}(\mathbf{K})_+$ satisfying

$$\forall \mathbf{k} \in \mathbf{\Gamma}, \qquad \int_{\mathbf{K}} \mathbf{x}^\mathbf{k}\, d\mu(\mathbf{x}) = z_\mathbf{k}, \tag{2.5}$$

where $\mathbf{x}^\mathbf{k} := x_1^{k_1} \cdots x_n^{k_n}$.

This problem has found a variety of solutions during the latest century, including two interesting theorems that we will present here. We first need to introduce the concept of Riesz functional.

Definition 2.1: Riesz functional

Let $\mathbf{z} := (z_\mathbf{k})_{\mathbf{k}\in\mathbb{N}^n} \in \mathbb{R}^{\mathbb{N}^n}$ be a real sequence. For $p(\mathbf{x}) := \sum p_\mathbf{k}\, \mathbf{x}^\mathbf{k} \in \mathbb{R}[\mathbf{x}]$, we define

$$\mathrm{L}_\mathbf{z}(p) := \sum p_\mathbf{k}\, z_\mathbf{k}.$$

The linear form $\mathrm{L}_\mathbf{z} : p \mapsto \mathrm{L}_\mathbf{z}(p)$ is called the *Riesz functional* of $\mathbf{z}$. Note that the map $\mathbf{z} \mapsto \mathrm{L}_\mathbf{z}$ is itself linear.

Example 2.2 *Let us focus on univariate polynomials ($n = 1$):* $\mathbf{z} = (z_k)_{k\in\mathbb{N}}$.
Let $p(x) := x^2$, $q(x) := 1 - x$, $g(x) := 1 - x^2$. *Then,*

- $\mathrm{L}_\mathbf{z}(p) = z_2$; $\mathrm{L}_\mathbf{z}(q) = z_0 - z_1$; $\mathrm{L}_\mathbf{z}(g) = z_0 - z_2$,
- $\mathrm{L}_\mathbf{z}(p\, q) = \mathrm{L}_\mathbf{z}(x^2 - x^3) = z_2 - z_3$,

- $\mathrm{L}_{\mathbf{z}}(g\ p\ q) = \mathrm{L}_{\mathbf{z}}(x^2 - x^3 - x^4 + x^5) = z_2 - z_3 - z_4 + z_5$.

If $n = 2$, $\mathbf{z} = (z_{ij})_{i,j\in\mathbb{N}}$ *and* $p(x, y) = (x - y)^2 = x^2 - 2xy + y^2$ *for instance, then* $\mathrm{L}_{\mathbf{z}}(p) = z_{20} - 2z_{11} + z_{02}$.

Remark 2.3 (Link between $\mathbf{L_z}$ and integrals)
If $\mu \in \mathcal{M}(\mathbf{K})_+$ *is a solution to Problem 3 with* $\mathbf{\Gamma} = \mathbb{N}^n$*, then by definition for all* $p \in \mathbb{R}[\mathbf{x}]$

$$\int p\, d\mu = \mathrm{L}_{\mathbf{z}}(p).$$

Note that Problem 3 only asks for *existence* of a suitable measure μ here. The question of its *uniqueness* is a distinct problem, to which a major contribution was given in the case where $\mathbf{K}$ is compact, $\mathbf{\Gamma} = \mathbb{N}^n$ and μ is asked to have finite total variation and be interior-regular (it is then called a *Radon* measure, see e.g. [110, Section 21.3]): indeed, in this compact $\mathbf{K}$ setting, the Riesz-Markov theorem [110, Section 21.4] ensures that a signed Radon measure μ is *uniquely* represented by the integrals $\int f\ d\mu$ of continuous functions $f \in \mathcal{C}(\mathbf{K})$, and the Stone-Weierstraß theorem [110, Section 12.3] states that any continuous function $f \in \mathcal{C}(\mathbf{K})$ can be approximated uniformly with polynomials $(p_\epsilon)_{\epsilon>0}$, so that

$$\int f\ d\mu = \lim_{\epsilon\to 0} \int p_\epsilon\ d\mu = \lim_{\epsilon\to 0} \mathrm{L}_{\mathbf{z}}(p_\epsilon)$$

is *uniquely* determined by the data $z_{\mathbf{k}} = \int \mathbf{x}^{\mathbf{k}}\ d\mu$ for all $\mathbf{k} \in \mathbb{N}^n$.

The notion of Riesz functional is instrumental in the formulation of a fundamental theorem (see [70, Theorem 3.1]) for solving Problem 3:

Theorem 2.2: Riesz-Haviland

Suppose that the set $\mathbf{K}$ is closed and that $\mathbf{\Gamma} = \mathbb{N}^n$. Then, Problem 3 has a solution *iff*

$$\mathrm{L}_{\mathbf{z}}(p) \geq 0$$

for all $p \in \mathbb{R}[\mathbf{x}]$ that is nonnegative on $\mathbf{K}$.

Theorem 2.2 provides us with a first criterion to decide the feasibility of Problem 3. However, we are again faced with the difficult problem of discriminating polynomials with respect to their sign over a set $\mathbf{K}$. With additional assumptions on $\mathbf{K}$, it is however possible to formulate a decisive feasibility theorem for Problem 3, at the price of the notion of basic semi-algebraic set:

Definition 2.3: Basic semialgebraic sets

$\mathbf{K} \subset \mathbb{R}^n$ is said to be *basic semialgebraic* if there exists $m \in \mathbb{N}^\star$ as well as $\mathbf{g} := (g_1, \ldots, g_m) \in \mathbb{R}[\mathbf{x}]^m$ s.t.

$$\mathbf{K} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{g}(\mathbf{x}) \geq \mathbf{0}\},$$

where $\mathbf{g}(\mathbf{x}) \geq \mathbf{0}$ means that for all $i \in \mathbb{N}^\star_m$, $g_i(\mathbf{x}) \geq 0$. We call $\mathbf{K}$ a *simple semialgebraic set* if $m = 1$, *semialgebraic set* if it is a finite union of basic semialgebraic sets.

Example 2.4

- *The half space* $\mathbf{H}^n := \{\mathbf{x} \in \mathbb{R}^n : x_n \geq 0\}$ *is simple semialgebraic.*
- *The positive orthant* $\mathbb{R}^n_+ := \{\mathbf{x} \in \mathbb{R}^n : \forall i \in \mathbb{N}^\star_n, x_i \geq 0\}$ *is basic semialgebraic.*
- *The unit ball of* $\mathbb{R}^n$, $\mathbf{B}^n := \{\mathbf{x} \in \mathbb{R}^n : 1 - |\mathbf{x}|^2 \geq 0\}$ *is a compact, simple semialgebraic set.*
- *The hypercube* $[-1,1]^n = \{\mathbf{x} \in \mathbb{R}^n : \forall i \in \mathbb{N}^\star_n, 1 - x_i^2 \geq 0\}$ *is a compact, basic semialgebraic set.*

Such notions lead to the formulation of Putinar's P-satz [107, Lemma 3.2], that is at the root of the moment-SOS hierarchy:

Theorem 2.4: Putinar's (primal) Positivstellensatz (P-satz)

Suppose that $\mathbf{\Gamma} = \mathbb{N}^n$ and that there exists $\mathbf{g} \in \mathbb{R}[\mathbf{x}]^m$, $R > 0$ such that

$$\mathbf{K} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{g}(\mathbf{x}) \geq \mathbf{0} \wedge |\mathbf{x}|^2 \leq R^2\}.$$

Then, Problem 3 has a unique solution $\mu \in \mathcal{M}(\mathbf{K})_+$ *iff* for all $p \in \mathbb{R}[\mathbf{x}]$:

- $\mathrm{L}_\mathbf{z}(p^2) \geq 0$,
- $\mathrm{L}_\mathbf{z}((R^2 - |\mathbf{x}|^2)\, p^2) \geq 0$,
- $\forall i \in \mathbb{N}^\star_m, \mathrm{L}_\mathbf{z}(g_i\, p^2) \geq 0$.

Theorem 2.4 makes it possible to replace the difficult positivity constraint on p of Theorem 2.2 with much more easily checked constraints on squares of polynomials and, by linearity, sums of squares of polynomials, which is at the origin of the name *moment-sum-of-squares* or *moment-SOS*: the hierarchy relies on the link drawn by Putinar's P-satz between moments of measures and sums of squares of polynomials.

Remark 2.5 (Ball constraint)

Let $\mathbf{g} \in \mathbb{R}[\mathbf{x}]^m$ *and* $\mathbf{K} := \{\mathbf{x} \in \mathbb{R}^n : \mathbf{g}(\mathbf{x}) \geq \mathbf{0}\}$ *be a generic compact basic semialgebraic set. For* $R > 0$*, define*

$$\mathbf{K}_R := \{\mathbf{x} \in \mathbb{R}^n : \mathbf{g}(\mathbf{x}) \geq \mathbf{0} \wedge |\mathbf{x}|^2 \leq R^2\} = \mathbf{K} \cap \mathbf{B}_R,$$

where $\mathbf{B}_R := \{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x}| \leq R\}$ *is the euclidean ball of radius* R*. In such setting, if* $\mathbf{K}$ *is compact then it is bounded, so that there exists* $R_0 > 0$ *s.t.* $\forall R \geq R_0$, $\mathbf{K} \subset \mathbf{B}_R$*, and thus* $\mathbf{K} = \mathbf{K}_R$*. This shows that if* $\mathbf{K}$ *is compact, it is always possible to add a redundant ball constraint so that Putinar's P-satz 2.4 holds for* $\mathbf{K}_R$*. For this reason, we will omit the technical ball constraint in most of this thesis.*

2.1.3 Generalizations

Problems 2 and 3 are actually both particular instances of a more general problem presented in details in [70, Chapter 1].

Let: • $\mathbf{K} \subset \mathbf{X} \subset \mathbb{R}^n$, • $\mathbb{I}, \mathbb{J}$ be sets of indices,
• $\varphi_\mathbf{i} \in \mathcal{C}(\mathbf{X}), \mathbf{i} \in \mathbb{I}$, • $\psi_\mathbf{j} \in \mathcal{C}(\mathbf{K}), \mathbf{j} \in \mathbb{J}$,
• $r_\mathbf{i} \in \mathbb{R}, \mathbf{i} \in \mathbb{I}$, • $z_\mathbf{j} \in \mathbb{R}, \mathbf{j} \in \mathbb{J}$, • $c \in \mathcal{C}(\mathbf{K})$.

The generalized moment problem is the most general linear optimization problem with measures as decision variables and generalized moment (*i.e.* integral) objective function and constraints.

Problem 4: Original generalized moment problem (GMP)

$$\begin{aligned} \mathrm{p}^\star_{\mathrm{GM}} := \sup_{\mu} \ & \int c \, d\mu & (2.6)\\ \text{s.t.} \quad & \mu \in \mathcal{M}(\mathbf{K})_+ \\ & \int \varphi_\mathbf{i} \, d\mu \leq r_\mathbf{i} \qquad \mathbf{i} \in \mathbb{I} \\ & \int \psi_\mathbf{j} \, d\mu = z_\mathbf{j} \qquad \mathbf{j} \in \mathbb{J}. \end{aligned}$$

Problems 2 and 3 are both instances of Problem 4. Indeed, on the one hand, taking $c = f$, $\mathbb{I} = \varnothing$, $\mathbb{J} = \{0\}$, $\psi_0(\mathbf{x}) = 1$, $z_0 = 1$, and minimizing instead of maximizing, one obtains exactly the formulation (2.3) of Problem 2.

On the other hand, taking $c = 0$, $\mathbb{I} = \varnothing$, $\mathbb{J} = \mathbf{\Gamma} \subset \mathbb{N}^n$, $\psi_\mathbf{j}(\mathbf{x}) = \mathbf{x}^\mathbf{j}$, one ends up with a feasibility problem formulation that exactly matches Problem 3.

In addition to including both Problems 2 and 3 (as well as a variety of other moment problems) as particular instances, Problem 4 allows for more general formulations that include inequality constraints on some moments of the decision variable μ.

Example 2.6 *The volume computation problem is an instance of the GMP* (2.6)*:*

$$\begin{aligned} \mathrm{p}^\star_{\mathbf{K}} := \sup_{\mu} \ & \int 1 \, d\mu \\ \text{s.t.} \quad & \mu \in \mathcal{M}(\mathbf{K})_+ \\ & \int \varphi \, d\mu \leq \int_{\mathbf{X}} \varphi(\mathbf{x}) \, \mathrm{d}\mathbf{x} \qquad \forall \varphi \in \mathcal{C}(\mathbf{X})_+ \end{aligned}$$

where $\mathbf{K} \subset \mathbf{X} \subset \mathbb{R}^n$, $\mathbf{X}$ *being compact. We shall prove later on that* $\mathrm{p}^\star_{\mathbf{K}} = \mathrm{vol}\ \mathbf{K}$.

One can compute the dual of the GMP, simply by defining the Lagrange multipliers $w_\mathbf{i} \in \mathbb{R}_+$, $\mathbf{i} \in \mathbb{I}$ (corresponding to the inequality constraints) and $v_\mathbf{j} \in \mathbb{R}$, $\mathbf{j} \in \mathbb{J}$ (corresponding to the equality constraints) and performing the standard Lagrange dualization process, leading to the following problem:

$$\begin{aligned} \mathrm{d}^\star_{\mathrm{GM}} := \inf_{\substack{(w_{\mathbf{i}})_{\mathbf{i}\in\mathbb{I}} \\ (v_{\mathbf{j}})_{\mathbf{j}\in\mathbb{J}}}} & \sum_{\mathbf{i}\in\mathbb{I}} w_{\mathbf{i}}\ r_{\mathbf{i}} + \sum_{\mathbf{j}\in\mathbb{J}} v_{\mathbf{j}}\ z_{\mathbf{j}} \\ \text{s.t.} & \sum_{\mathbf{i}\in\mathbb{I}} w_{\mathbf{i}}\ \varphi_{\mathbf{i}} + \sum_{\mathbf{j}\in\mathbb{J}} v_{\mathbf{j}}\ \psi_{\mathbf{j}} - c \in \mathcal{C}(\mathbf{K})_+ \\ & w_{\mathbf{i}} \in \mathbb{R}_+ \qquad\qquad \mathbf{i}\in\mathbb{I} \\ & v_{\mathbf{j}} \in \mathbb{R} \qquad\qquad \mathbf{j}\in\mathbb{J}. \end{aligned}$$

However, this dual formulation is somehow more difficult to interpret, with sums over possibly infinite or even uncountable index sets $\mathbb{I}$ and $\mathbb{J}$. Thus, we first work on a rephrazing of Problem 4 allowing for more synthetic dualization: by linearity of the integration operation, we can consider without loss of generality that the elements of $\boldsymbol{\varphi} := (\varphi_{\mathbf{i}})_{\mathbf{i}\in\mathbb{I}}$ and $\boldsymbol{\psi} := (\psi_{\mathbf{j}})_{\mathbf{j}\in\mathbb{J}}$ are linearly independent. Thus, we define:

- $\mathcal{Y}'_{\mathbb{I}} := \operatorname{span} \boldsymbol{\varphi} \subset \mathcal{C}(\mathbf{X})$ (resp. $\mathcal{Y}'_{\mathbb{J}} := \operatorname{span} \boldsymbol{\psi} \subset \mathcal{C}(\mathbf{K})$) the vector space spanned by the basis $\boldsymbol{\varphi}$ (resp. $\boldsymbol{\psi}$),
- $\mathcal{Y}'_{\mathbb{I}+} := \operatorname{span}_+ \boldsymbol{\varphi} := \left\{ \sum_{k=1}^K w_k\ \varphi_{\mathbf{i}_k} : K \in \mathbb{N}^\star \wedge \forall k \in \mathbb{N}^\star_K,\ (w_k \geq 0\ \wedge\ \mathbf{i}_k \in \mathbb{I}) \right\}$ the convex cone spanned by inequality constraints $\boldsymbol{\varphi}$,
- $\mathcal{Y}_{\mathbb{I}} := (\mathcal{Y}'_{\mathbb{I}})'$ & $\mathcal{Y}_{\mathbb{J}} := (\mathcal{Y}'_{\mathbb{J}})'$ where $\mathcal{X}' := \{\phi \in \mathcal{C}(\mathcal{X}) : \phi \text{ is linear}\}$ denotes the topological dual of the vector space $\mathcal{X}$, and $\langle \chi, \phi \rangle_{\mathcal{X}} := \phi(\chi)$ for $\phi \in \mathcal{X}'$, $\chi \in \mathcal{X}$,
- $\mathcal{Y}_{\mathbb{I}+} := (\mathcal{Y}'_{\mathbb{I}+})' = \{\phi \in \mathcal{Y}_{\mathbb{I}} : \forall \varphi \in \mathcal{Y}'_{\mathbb{I}+}, \langle \varphi, \phi \rangle_{\mathcal{Y}_I} \geq 0\}$ the dual cone of $\mathcal{Y}'_{\mathbb{I}+}$,
- $\Phi := \left\{ \begin{array}{ccc} \mathcal{M}(\mathbf{K}) & \longrightarrow & \mathcal{Y}_{\mathbb{I}} \\ \mu & \longmapsto & [\varphi_{\mathbf{i}} \mapsto \int \varphi_{\mathbf{i}}\, d\mu] \end{array} \right.$ the inequality constraint linear map,
- $\Psi := \left\{ \begin{array}{ccc} \mathcal{M}(\mathbf{K}) & \longrightarrow & \mathcal{Y}_{\mathbb{J}} \\ \mu & \longmapsto & [\psi_{\mathbf{j}} \mapsto \int \psi_{\mathbf{j}}\, d\mu] \end{array} \right.$ the equality constraint linear map,
- $\rho := [\varphi_{\mathbf{i}} \mapsto r_{\mathbf{i}}] \in \mathcal{Y}_{\mathbb{I}}$ & $\zeta := [\psi_{\mathbf{j}} \mapsto z_{\mathbf{j}}] \in \mathcal{Y}_{\mathbb{J}}$ synthetic optimization parameters.

This allows us to give a synthetic formulation of (2.6) and its dual.

Problem 5: Synthetic GMP

$$\begin{aligned} \mathrm{p}^\star_{\mathrm{GM}} := \sup_{\mu} & \int c\, d\mu \qquad (2.7a) \\ \text{s.t.} & \ \mu \in \mathcal{M}(\mathbf{K})_+ \\ & \ \rho - \Phi\mu \in \mathcal{Y}_{\mathbb{I}+} \\ & \ \zeta - \Psi\mu = 0_{\mathcal{Y}_{\mathbb{J}}} \end{aligned} \qquad \begin{aligned} \mathrm{d}^\star_{\mathrm{GM}} := \inf_{v,w} & \langle w, \rho \rangle_{\mathcal{Y}_{\mathbb{I}}} + \langle v, \zeta \rangle_{\mathcal{Y}_{\mathbb{J}}} \qquad (2.7b) \\ \text{s.t.} & \ \Phi' w + \Psi' v - c \in \mathcal{C}(\mathbf{K})_+ \\ & \ w \in \mathcal{Y}'_{\mathbb{I}+} \\ & \ v \in \mathcal{Y}'_{\mathbb{J}} \end{aligned}$$

where Φ' (resp. Ψ') is the adjoint of Φ (resp. Ψ), defined by

$$\int(\Phi' w)\, d\mu = \langle w, \Phi\mu\rangle_{\mathcal{Y}_{\mathbb{I}}} \text{ (resp. } \int(\Psi' v)\, d\mu = \langle v, \Psi\mu\rangle_{\mathcal{Y}_{\mathbb{J}}}).$$

Problems (2.6) and (2.7a) are equivalent, and so are their duals.
Indeed, $\sum_{\mathbf{i}} w_{\mathbf{i}}\, \varphi_{\mathbf{i}}, w_{\mathbf{i}} \in \mathbb{R}_+$, and $\sum_{\mathbf{j}} v_{\mathbf{j}}\, \psi_{\mathbf{j}}$, $v_{\mathbf{j}} \in \mathbb{R}$, can be approximated with $w \in \mathcal{Y}'_{\mathbb{I}_+}$ and $v \in \mathcal{Y}'_{\mathbb{J}}$ respectively.

Example 2.7 *The volume problem of Example 2.6 can be reformulated as an instance of* (2.7) *with no equality constraints, which allows for easy dual formulation (see [44]):*

$$\begin{array}{rl} \mathrm{p}^\star_{\mathbf{K}} = \sup\limits_{\mu} & \int 1\, d\mu \\ \text{s.t. } & \mu \in \mathcal{M}(\mathbf{K})_+ \\ & \lambda - \mu \in \mathcal{M}(\mathbf{X})_+ \end{array} \qquad (2.8a) \qquad \begin{array}{rl} \mathrm{d}^\star_{\mathbf{K}} := \inf\limits_{w} & \int w\, d\lambda \\ \text{s.t. } & w - 1 \in \mathcal{C}(\mathbf{K})_+ \\ & w \in \mathcal{C}(\mathbf{X})_+ \end{array} \qquad (2.8b)$$

where λ is the Lebesgue measure on $\mathbf{X}$. Now we show that $\mathrm{p}^\star_{\mathbf{K}} = \mathrm{vol}\ \mathbf{K} = \mathrm{d}^\star_{\mathbf{K}}$:

On the primal side: $\mu^\star := \lambda_{\mathbf{K}} := \mathbb{1}_{\mathbf{K}}\ \lambda$ is clearly feasible (since $\mathbb{1}_{\mathbf{K}} \leq 1$) and for any feasible $\mu \in \mathcal{M}(\mathbf{K})_+$, $\mu^\star - \mu \in \mathcal{M}(\mathbf{K})_+$, thus $\mu^\star$ is optimal and $\mathrm{p}^\star_{\mathbf{K}} = \int 1\, d\mu^\star = \lambda(\mathbf{K}) = \mathrm{vol}\ \mathbf{K}$.

On the dual side: constraints on w can be synthesized as $w \geq \mathbb{1}_{\mathbf{K}}$; then, by density of $\mathcal{C}(\mathbf{X})$ in the space $L^1(\mathbf{X})$ of Lebesgue-integrable functions, any minimizing sequence $(w_\epsilon)_\epsilon$ is such that $\int w_\epsilon\, d\lambda \underset{\epsilon\to\infty}{\longrightarrow} \int \mathbb{1}_{\mathbf{K}}\, d\lambda = \lambda(\mathbf{K}) = \mathrm{vol}\ \mathbf{K}$.

The volume problem will serve as an illustrative example in the rest of this chapter. It will also be at the center of Chapters 4 and 5.

However, some already existing applications of the moment-SOS hierarchy to dynamical and control systems rely on GMP formulations involving multiple measures. For this reason, we eventually propose the following generalization, where instead of considering a single measure $\mu \in \mathcal{M}(\mathbf{K})_+$, one takes:

- $\mu_1 \in \mathcal{M}(\mathbf{X}_1)_+$, …, $\mu_N \in \mathcal{M}(\mathbf{X}_N)_+$, where $\mathbf{X}_k \subset \mathbf{Y}_k \subset \mathbb{R}^{n_k}$,
- corresponding costs and parameters $c_1 \in \mathcal{C}(\mathbf{X}_1)$, …, $c_N \in \mathcal{C}(\mathbf{X}_N)$,
- inequality constraint parameters $\boldsymbol{\varphi}_1 \in \mathcal{C}(\mathbf{Y}_1)^{\mathbb{I}}$, …, $\boldsymbol{\varphi}_N \in \mathcal{C}(\mathbf{Y}_N)^{\mathbb{I}}$
- equality constraint parameters $\boldsymbol{\psi}_1 \in \mathcal{C}(\mathbf{X}_1)^{\mathbb{J}}$, …, $\boldsymbol{\psi}_N \in \mathcal{C}(\mathbf{X}_N)^{\mathbb{J}}$,

and define the multi-measure integral

$$\int \mathbf{c} \cdot d\boldsymbol{\mu} := \sum_{k=1}^{N} \int c_k\, d\mu_k,$$

together with the standard operator notation:

- $\mathcal{X} := \mathcal{M}(\mathbf{X}_1) \times \ldots \times \mathcal{M}(\mathbf{X}_N) \quad , \quad \mathcal{X}_+ := \mathcal{M}(\mathbf{X}_1)_+ \times \ldots \times \mathcal{M}(\mathbf{X}_N)_+$
- $\mathcal{X}' := \mathcal{C}(\mathbf{X}_1) \times \ldots \times \mathcal{C}(\mathbf{X}_N) \ni \mathbf{c} \quad , \quad \mathcal{X}'_+ := \mathcal{C}(\mathbf{X}_1)_+ \times \ldots \times \mathcal{C}(\mathbf{X}_N)_+,$
- $\mathcal{Y} := \mathcal{Y}_{\mathbb{I}} \times \mathcal{Y}_{\mathbb{J}} \quad , \quad \mathcal{Y}_+ := \mathcal{Y}_{\mathbb{I}_+} \times \{0_{\mathcal{Y}_{\mathbb{J}}}\} \quad , \quad \mathbf{b} := (\boldsymbol{\rho}, \boldsymbol{\zeta}) \in \mathcal{Y},$

- $\mathcal{Y}' := \mathcal{Y}'_{\mathbb{I}} \times \mathcal{Y}'_{\mathbb{J}} \quad , \quad \mathcal{Y}'_+ := \mathcal{Y}'_{\mathbb{I}+} \times \mathcal{Y}'_{\mathbb{J}} \quad , \quad \mathcal{A}\,\boldsymbol{\mu} := (\Phi\boldsymbol{\mu}, \Psi\boldsymbol{\mu}) \in \mathcal{Y}$,
- $\forall \boldsymbol{\ell} := (\mathbf{w}, \mathbf{v}) \in \mathcal{Y}', \quad \langle \boldsymbol{\ell}, \mathbf{b} \rangle_{\mathcal{Y}} := \langle \mathbf{w}, \boldsymbol{\rho} \rangle_{\mathcal{Y}_{\mathbb{I}}} + \langle \mathbf{v}, \boldsymbol{\zeta} \rangle_{\mathcal{Y}_{\mathbb{J}}}$
- $\forall \boldsymbol{\ell} := (\mathbf{w}, \mathbf{v}) \in \mathcal{Y}', \quad \mathcal{A}'\boldsymbol{\ell} := \Phi'\mathbf{w} + \Psi'\mathbf{v}$ so that

$$\langle \mathbf{w}, \Phi\boldsymbol{\mu} \rangle_{\mathcal{Y}_{\mathbb{I}}} + \langle \mathbf{v}, \Psi\boldsymbol{\mu} \rangle_{\mathcal{Y}_{\mathbb{J}}} = \langle \boldsymbol{\ell}, \mathcal{A}\,\boldsymbol{\mu} \rangle_{\mathcal{Y}} = \int (\mathcal{A}'\boldsymbol{\ell}) \cdot d\boldsymbol{\mu} = \int (\Phi'\mathbf{w} + \Psi'\mathbf{v}) \cdot d\boldsymbol{\mu},$$

which allows us to write the following abstract rephrazing of the GMP and its dual:

Problem 6: Abstract multivariate GMP

$$\begin{aligned} \mathrm{p}^\star_{\mathrm{GM}} := \sup_{\boldsymbol{\mu}} \ & \int \mathbf{c} \cdot d\boldsymbol{\mu} && (2.9a) \\ \text{s.t. } & \boldsymbol{\mu} \in \mathcal{X}_+ \\ & \mathbf{b} - \mathcal{A}\,\boldsymbol{\mu} \in \mathcal{Y}_+ \end{aligned} \qquad \begin{aligned} \mathrm{d}^\star_{\mathrm{GM}} := \inf_{\boldsymbol{\ell}} \ & \langle \boldsymbol{\ell}, \mathbf{b} \rangle_{\mathcal{Y}} && (2.9b) \\ \text{s.t. } & \mathcal{A}'\boldsymbol{\ell} - \mathbf{c} \in \mathcal{X}'_+ \\ & \boldsymbol{\ell} \in \mathcal{Y}'_+ \end{aligned}$$

By construction, Problems 4 and 5 are particular instances of Problem 6. In fact, all the infinite dimensional optimization problems that we will consider in this thesis are instances of this general Problem 6.

Example 2.8 *The finite time region of attraction (ROA) problem is another particular instance of the GMP* (2.9)*, see [42]:*

$$\begin{aligned} \mathrm{p}^\star_{\mathrm{ROA}} := \sup_{\mu,\nu,\xi} \ & \int 1 \, d\mu \\ \text{s.t. } & \mu \in \mathcal{M}(\mathbf{X})_+ \\ & \nu \in \mathcal{M}(\mathbf{I} \times \mathbf{X})_+ \\ & \xi \in \mathcal{M}(\mathbf{K})_+ \\ & \lambda - \mu \in \mathcal{M}(\mathbf{X})_+ \\ & \partial_t \nu + \mathrm{div}(\nu\, \mathbf{f}) = \delta_0\, \mu - \delta_T\, \xi \end{aligned} \qquad \begin{aligned} \mathrm{d}^\star_{\mathrm{ROA}} := \inf_{v,w} \ & \int w \, d\lambda \\ \text{s.t. } & w - v(0,\cdot) - 1 \in \mathcal{C}(\mathbf{X})_+ \\ & -\partial_t v - \mathbf{f} \cdot \mathbf{grad}\ v \in \mathcal{C}(\mathbf{I} \times \mathbf{X})_+ \\ & v(T,\cdot) \in \mathcal{C}(\mathbf{K})_+ \\ & w \in \mathcal{C}(\mathbf{X})_+ \\ & v \in \mathcal{C}^1(\mathbf{I} \times \mathbf{X}) \end{aligned}$$

where $\mathbf{I} := [0, T]$*,* $T > 0$ *and* $\mathbf{K} \subset \mathbf{X} \subset \mathbb{R}^n$*,* $\mathbf{X}$ *being compact. Here* $\mathbf{c} = (1, 0, 0)$*,* $\Phi(\mu, \nu, \xi) = \mu$*,* $\Psi(\mu, \nu, \xi) = \partial_t \nu + \mathrm{div}(\nu\, \mathbf{f}) + \delta_T\, \xi - \delta_0\, \mu$*,* $\boldsymbol{\rho} = (\lambda, 0, 0)$*,* $\boldsymbol{\zeta} = (0, 0, 0)$*. Then,* $\Phi' w = (w, 0, 0)$ *and* $\Psi' v = (-v(0,\cdot), -\partial_t v - \mathbf{f} \cdot \mathbf{grad}\ v, v(T,\cdot))$*.*

We will have the opportunity to give more detailed explanations on this problem in Chapters 3 and 6.

2.1.4 Infinite dimensional duality ⋆

This section only aims at stating a technical theorem that we will use all along this thesis to prove strong duality in our GMP instances.

Lagrangian duality is a rich notion that makes it possible to consider two viewpoints when faced to an optimization problem. We first introduced it in Section 2.1.1 where we gave a primal probabilistic viewpoint (2.4a) as well as a dual analytic viewpoint (2.4b) to the POP (2.1). Then we introduced generic duality in the GMP with formulations (2.7) and (2.9).

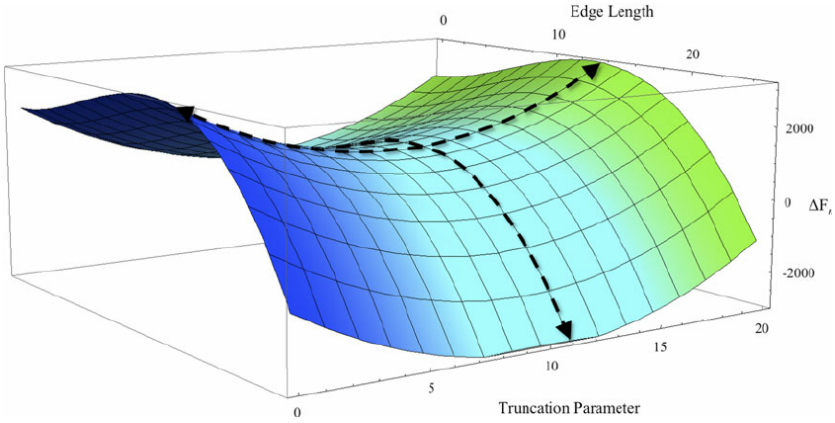


Figure 2.2 – A possible aspect of Lagrangian for univariate optimization.

The primal problem maximizes the Lagrangian while staying on the concave curve, and the dual problem minimizes the Lagrangian while staying on the convex curve.
Image source – "A saddle point", Intro to optimization in deep learning: Gradient Descent, Ayoosh Kathuria.

Let us first go back to the dual formulation. Starting from (2.9a), the dual (2.9b) is obtained by rephrazing the problem in terms of Lagrangian:

$$\mathrm{p}^\star_{\mathrm{GM}} = \sup_{\boldsymbol{\mu}\in\mathcal{X}_+} \inf_{\boldsymbol{\ell}\in\mathcal{Y}'_+} \mathcal{L}(\boldsymbol{\mu},\boldsymbol{\ell}) \quad (2.10a) \quad \longrightarrow \quad \mathrm{d}^\star_{\mathrm{GM}} := \inf_{\boldsymbol{\ell}\in\mathcal{Y}'_+} \sup_{\boldsymbol{\mu}\in\mathcal{X}_+} \mathcal{L}(\boldsymbol{\mu},\boldsymbol{\ell}) \quad (2.10b)$$

where the Lagrangian functional is defined by

$$\mathcal{L}(\boldsymbol{\mu},\boldsymbol{\ell}) := \int \mathbf{c}\cdot d\boldsymbol{\mu} + \langle \boldsymbol{\ell}, \mathbf{b} - \mathcal{A}\,\boldsymbol{\mu}\rangle_{\mathcal{Y}} = \int (\mathbf{c} - \mathcal{A}'\boldsymbol{\ell})\cdot d\boldsymbol{\mu} + \langle \boldsymbol{\ell}, \mathbf{b}\rangle_{\mathcal{Y}},$$

both expressions being equivalent by definition of the adjoint operator $\mathcal{A}'$.

Indeed, by definition of the dual cone $\mathcal{Y}'_+$, for any $\boldsymbol{\mu}\in\mathcal{X}_+$,

$$\inf_{\boldsymbol{\ell}\in\mathcal{Y}'_+} \mathcal{L}(\boldsymbol{\mu},\boldsymbol{\ell}) = \begin{cases} \int \mathbf{c}\cdot d\boldsymbol{\mu} & \text{if } \mathbf{b} - \mathcal{A}\,\boldsymbol{\mu} \in \mathcal{Y}_+ \\ -\infty & \text{else.} \end{cases}$$

Then, since $-\infty < \int \mathbf{c}\cdot d\boldsymbol{\mu}$, (2.10a) is actually equivalent to (2.9a). Eventually, to formulate the dual problem, one only needs to switch the sup and inf operators, and do the reverse reasoning to show that (2.10b) is equivalent to (2.9b).

No matter the subject, considering several viewpoints for a given problem helps understanding it, and each viewpoint has its special features that allow for elegant proofs or convenient practical implementations. However, in order to be able to freely switch between primal and dual formulations of a given problem, one needs an essential property: the optimum should not depend on the viewpoint, meaning that one has to ensure that $\mathrm{p}^\star_{\mathrm{GM}} = \mathrm{d}^\star_{\mathrm{GM}}$. This is the *strong duality* property.

In general, strong duality is not guaranteed, and one only has:

Proposition 2.5: Weak duality

It always holds that $\mathrm{p}^\star_{\mathrm{GM}} \leq \mathrm{d}^\star_{\mathrm{GM}}$.

Proof : Let $\boldsymbol{\mu} \in \mathcal{X}_+, \boldsymbol{\ell} \in \mathcal{Y}'_+$. It is clear that $\inf_{\mathbf{m} \in \mathcal{Y}'_+} \mathcal{L}(\boldsymbol{\mu}, \mathbf{m}) \leq \mathcal{L}(\boldsymbol{\mu}, \boldsymbol{\ell})$. Then, one can take the supremum over all $\boldsymbol{\mu} \in \mathcal{X}_+$ to get that $\mathrm{p}^\star_{\mathrm{GM}} \leq \sup_{\boldsymbol{\mu} \in \mathcal{X}_+} \mathcal{L}(\boldsymbol{\mu}, \boldsymbol{\ell})$. Eventually, taking the infimum over all $\boldsymbol{\ell} \in \mathcal{Y}'_+$ yields $\mathrm{p}^\star_{\mathrm{GM}} \leq \mathrm{d}^\star_{\mathrm{GM}}$. ♢

For the GMP, one can prove an elegant theorem to easily guarantee strong duality in practice.

Theorem 2.6: Strong duality in the GMP

Suppose that there exists $C > 0$ s.t. if $\boldsymbol{\mu} = (\mu_1, \dots, \mu_N) \in \mathcal{X}_+$ is such that $\mathbf{b} - \mathcal{A}\,\boldsymbol{\mu} \in \mathcal{Y}_+$, then for all $k \in \mathbb{N}^\star_N$, $\mu_k(\mathbf{X}_k) \leq C$. Suppose that there exists such $\boldsymbol{\mu}$. In that case,

$$\mathrm{p}^\star_{\mathrm{GM}} = \mathrm{d}^\star_{\mathrm{GM}}.$$

Moreover, (2.9a) has an optimal solution $\boldsymbol{\mu}^\star \in \mathcal{X}_+$ s.t.

$$\mathbf{b} - \mathcal{A}\,\boldsymbol{\mu}^\star \in \mathcal{Y}_+ \qquad \& \qquad \int \mathbf{c} \cdot d\boldsymbol{\mu}^\star = \mathrm{p}^\star_{\mathrm{GM}}.$$

Proof : We rely on [11, Theorem (7.2), Lemma (7.3)]. Consider the cone

$$\mathcal{K} := \left\{ \left(\mathcal{A}\,\boldsymbol{\mu}, \int \mathbf{c} \cdot d\boldsymbol{\mu} \right) : \boldsymbol{\mu} \in \mathcal{X}_+ \right\} \subset \mathcal{Y} \times \mathbb{R}.$$

According to [11, Theorem (7.2)] we only have to prove that $\mathrm{p}^\star_{\mathrm{GM}} < \infty$ and $\mathcal{K}$ is closed. Clearly,

$$\mathrm{p}^\star_{\mathrm{GM}} \leq \|\mathbf{c}\|_\infty \, \|\boldsymbol{\mu}\|_{TV} \leq \|\mathbf{c}\|_\infty \; N \; C < \infty,$$

where $\|\mathbf{c}\|_\infty := \max_k \sup_{\mathbf{x}_k \in \mathbf{X}_k} c_k(\mathbf{x}_k)$ is finite and $\|\boldsymbol{\mu}\|_{TV} := \sum_k \mu_k(\mathbf{X}_k) \leq N \; C$. Besides, [11, Lemma (7.3)] states that for $\mathcal{K}$ to be closed, it is sufficient to prove that $\mathcal{X}_+$ has a weak-$*$ compact, convex base, and that

$$\forall \boldsymbol{\mu} \in \mathcal{X}_+, \qquad \left(\mathcal{A}\,\boldsymbol{\mu}, \int \mathbf{c} \cdot d\boldsymbol{\mu} \right) = (\mathbf{0}_{\mathcal{Y}}, 0) \Longrightarrow \boldsymbol{\mu} = \mathbf{0}_{\mathcal{X}}. \qquad (*)$$

We first exhibit a weak-$*$ compact convex base for $\mathcal{X}_+ = \mathcal{M}(\mathbf{X}_1)_+ \times \ldots \times \mathcal{M}(\mathbf{X}_N)_+$. Let

$$\mathcal{P} := \left\{ \boldsymbol{\phi} \in \mathcal{X}_+ : \int \mathbf{1} \cdot d\boldsymbol{\phi} = 1 \right\},$$

where $\mathbf{1} = (\mathbf{x}_1 \mapsto 1, \dots, \mathbf{x}_N \mapsto 1) \in \mathcal{C}(\mathbf{X}_1) \times \ldots \mathcal{C}(\mathbf{X}_N)$. $\mathcal{P}$ is a base of $\mathcal{X}_+$ in the sense that $\mathcal{X}_+ \setminus \{\mathbf{0}_{\mathcal{X}}\}$ is isomorphic to $\mathbb{R}_{++} \times \mathcal{P}$ through the bijective application $\chi : (t, \boldsymbol{\phi}) \mapsto t \; \boldsymbol{\phi}$. Indeed, any $\boldsymbol{\mu} \in \mathcal{X}_+ \setminus \{\mathbf{0}_{\mathcal{X}}\}$ has a unique antecedent by χ, given by

$$\boldsymbol{\mu} = t \; \boldsymbol{\phi} \text{ with:} \qquad t := \int \mathbf{1} \cdot d\boldsymbol{\mu} = \sum_{k=1}^{N} \mu_k(\mathbf{X}_k) > 0 \qquad \& \qquad \boldsymbol{\phi} := \frac{1}{t} \boldsymbol{\mu}.$$

$\mathcal{P}$ is convex since $\mathcal{X}_+$ is convex and for any $t \in [0,1]$, $\boldsymbol{\phi}_1, \boldsymbol{\phi}_2 \in \mathcal{P}$, $\widetilde{\boldsymbol{\phi}} := t\boldsymbol{\phi}_1 + (1-t)\boldsymbol{\phi}_2$,

$$\int \mathbf{1} \cdot d\widetilde{\boldsymbol{\phi}} = t \int \mathbf{1} \cdot d\boldsymbol{\phi}_1 + (1-t) \int \mathbf{1} \cdot d\boldsymbol{\phi}_2 = t + 1 - t = 1.$$

$\mathcal{P}$ is weak-$*$ closed as the intersection between the level-1 set of the weak-$*$ continuous functional $\boldsymbol{\mu} \mapsto \int \mathbf{1} \cdot d\boldsymbol{\mu}$ and the weak-$*$ closed cone $\mathcal{X}_+$.

Eventually, $\mathcal{P} \subset \mathcal{X}_{\leq 1} := \{\boldsymbol{\nu} = (\nu_1, \ldots, \nu_N) \in \mathcal{X} : \forall k \in \mathbb{N}_N^\star, \|\nu_k\|_{TV} \leq 1\}$, where $\|\nu\|_{TV} := \sup\{\int \varphi \, d\nu : |\varphi| \leq 1\}$ is the total variation norm. The Banach-Alaoglu theorem [15, Theorem 3.16] ensures that $\mathcal{X}_{\leq 1}$ is weak-$*$ compact, yielding that $\mathcal{P}$ is a weak-$*$ closed subset of a weak-$*$ compact set and thus a weak-$*$ compact set itself.

It remains to prove $(*)$. Let $\boldsymbol{\mu}^\perp \in \mathcal{X}_+$ s.t. $\mathcal{A}\,\boldsymbol{\mu}^\perp = \mathbf{0}_\mathcal{Y}$ and $\int \mathbf{c} \cdot d\boldsymbol{\mu}^\perp = 0$. We want to prove that $\boldsymbol{\mu}^\perp = \mathbf{0}_\mathcal{X}$ so that $(*)$ holds.

Let $\boldsymbol{\mu}^{(0)} \in \mathcal{X}_+$ s.t. $\mathbf{b} - \mathcal{A}\,\boldsymbol{\mu}^{(0)} \in \mathcal{Y}_+$. Define for $t \geq 0$ $\boldsymbol{\mu}^{(t)} := \boldsymbol{\mu}^{(0)} + t\ \boldsymbol{\mu}^\perp$. Let $t \geq 0$. Since $\mathcal{X}_+$ is a convex cone, $\boldsymbol{\mu}^{(t)} \in \mathcal{X}_+$; in addition, by construction $\mathbf{b} - \mathcal{A}\,\boldsymbol{\mu}^{(t)} = \mathbf{b} - \mathcal{A}\,\boldsymbol{\mu}^{(0)}$ is in $\mathcal{Y}_+$, so that our assumption ensures that for any $k \in \mathbb{N}_N^\star$, $\mu_k^{(t)}(\mathbf{X}_k) \leq C$. However, $\mu_k^{(t)}(\mathbf{X}_k) = \mu_k^{(0)}(\mathbf{X}_k) + t\ \mu_k^\perp(\mathbf{X}_k) \geq t\ \mu_k^\perp(\mathbf{X}_k)$.

This yields that for all $t \geq 0$, $k \in \mathbb{N}_N^\star$,

$$0 \leq t\ \mu_k^\perp(\mathbf{X}_k) \leq C,$$

which is only possible if $\mu_k^\perp(\mathbf{X}_k) = 0\ \forall k \in \mathbb{N}_N^\star$, *i.e.* if $\boldsymbol{\mu}^\perp = \mathbf{0}_\mathcal{X}$. ♢

Example 2.9 *Consider the global POP:*

$$\begin{aligned} \mathrm{p}_f^\star = \min_\mu \ & \int f\ d\mu && (2.4a) \qquad \mathrm{d}_f^\star = \max_\ell \ \ell && (2.4b) \\ \text{s.t. } & \mu \in \mathcal{M}(\mathbf{K})_+ && \text{s.t. } f - \ell \in \mathcal{C}(\mathbf{K})_+ \\ & \mu(\mathbf{K}) = 1 && \ell \in \mathbb{R} \end{aligned}$$

We already know that optimal values are attained and that $\mathrm{p}_f^\star = \mathrm{d}_f^\star = f^\star$. *However, this fact can also be proved using Theorem 2.6: any* $\mu \in \mathcal{M}(\mathbf{K})_+$ *s.t.* $\mu(\mathbf{K}) = 1$ *satisfies* $\mu(\mathbf{K}) \leq C$ *as long as* $C \geq 1$, *and the set of probability measures over nonempty* $\mathbf{K}$ *is nonempty. Thus,* $\mathrm{p}_f^\star$ *is attained and strong duality holds:* $\mathrm{p}_f^\star = \mathrm{d}_f^\star$.

Example 2.10 *Consider the volume problem:*

$$\begin{aligned} \mathrm{p}_\mathbf{K}^\star = \sup_\mu \ & \int 1\ d\mu && \mathrm{d}_\mathbf{K}^\star := \inf_w \ \int w\ d\lambda \\ \text{s.t. } & \mu \in \mathcal{M}(\mathbf{K})_+ && \text{s.t. } w - 1 \in \mathcal{C}(\mathbf{K})_+ \\ & \lambda - \mu \in \mathcal{M}(\mathbf{X})_+ && w \in \mathcal{C}(\mathbf{X})_+ \end{aligned}$$

We already know that primal optimum is attained and that $\mathrm{p}_\mathbf{K}^\star = \mathrm{d}_\mathbf{K}^\star = \mathrm{vol}\ \mathbf{K}$. *However, this fact can also be proved using Theorem 2.6: any* $\mu \in \mathcal{M}(\mathbf{K})_+$ *s.t.* $\lambda - \mu \in \mathcal{M}(\mathbf{X})_+$ *satisfies* $\mu(\mathbf{K}) \leq \lambda(\mathbf{K}) =: C$, *and the set of such measures is nonempty (for instance, it contains the null measure* 0*). Thus,* $\mathrm{p}_\mathbf{K}^\star$ *is attained and strong duality holds:* $\mathrm{p}_\mathbf{K}^\star = \mathrm{d}_\mathbf{K}^\star$.

Example 2.11 *In Chapter 3 we will use Theorem 2.6 to prove strong duality for the finite time region of attraction problem of Example 2.8.*

Remark 2.12 (Motivation for strong duality results)

As stated at the beginning of this section, strong duality first allows tapping into both versions of the same problem, depending on which property one intends to use, without changing the optimal value of the problem.

However, it is not the only interest for proving strong duality: as we will show later on, strong duality in the GMP is a first sign that the derived finite dimensional approximation problems should be well behaved when numerically addressed through optimization algorithms.

2.2 The moment-SOS hierarchy

After giving an insight on the generalized moment problem and infinite dimensional linear programming in Section 2.1, we now proceed to detail a method to compute approximate solutions of such problems, namely Lasserre's moment-SOS hierarchy. We first introduce the primal, moment hierarchy, that directly exploits previously introduced results to give a finite dimensional counterpart to the moment problem (2.9a). Then, we detail the dual, SOS hierarchy for problem (2.9b), manipulating simpler objects such as functions and polynomials instead of measures and moments. Eventually, as in the previous section, we state a strong duality theorem for the moment-SOS hierarchy, as well as some useful additional results.

2.2.1 Moments

We first focus on the primal (2.9a) on measures, as the moment hierarchy is quickly stated using theorem 2.4. For the sake of clarity and without loss of generality, we work with the particular instance of the volume problem:

$$\begin{aligned} \mathrm{p}^\star_{\mathbf{K}} = \sup_{\mu} \ & \int 1 \, d\mu \\ \text{s.t. } & \mu \in \mathcal{M}(\mathbf{K})_+ \\ & \lambda - \mu \in \mathcal{M}(\mathbf{X})_+, \end{aligned} \tag{2.8a}$$

with $\mathbf{K} \subset \mathbf{X} \subset \mathbb{R}^n$ compact sets. The moment hierarchy consists in replacing the decision variable $\mu \in \mathcal{M}(\mathbf{K})_+$ with the sequence of its moments $z_{\mathbf{k}} = \int \mathbf{x}^{\mathbf{k}} \, d\mu$. To ensure that this does not change the problem, we invoke Putinar's P-satz 2.4 to guarantee that we really work with moment sequences corresponding to actual measures, and not generic multi-index sequences. However, using Theorem 2.4 requires an additional assumption on $\mathbf{K}$ and $\mathbf{X}$.

Assumption 2.7: Compact basic semialgebraic sets

$\exists\, m_{\mathbf{K}}, m_{\mathbf{X}} \in \mathbb{N}$, $R_{\mathbf{K}}, R_{\mathbf{X}} \in \mathbb{R}_{++}$, $\mathbf{g}_{\mathbf{K}} \in \mathbb{R}[\mathbf{x}]^{m_{\mathbf{K}}}$, $\mathbf{g}_{\mathbf{X}} \in \mathbb{R}[\mathbf{x}]^{m_{\mathbf{X}}}$ s.t.

$$g_{\mathbf{K},m_{\mathbf{K}}}(\mathbf{x}) = R_{\mathbf{K}}^2 - |\mathbf{x}|^2, \qquad g_{\mathbf{X},m_{\mathbf{X}}}(\mathbf{x}) = R_{\mathbf{X}}^2 - |\mathbf{x}|^2,$$

$$\mathbf{K} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{g}_{\mathbf{K}}(\mathbf{x}) \geq \mathbf{0}\}, \qquad \mathbf{X} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{g}_{\mathbf{X}}(\mathbf{x}) \geq \mathbf{0}\}.$$

In words, $\mathbf{K}$ and $\mathbf{X}$ are basic semialgebraic sets with a ball constraint in their description. According to Remark 2.5, up to adding a redundant ball constraint, this is equivalent to assuming $\mathbf{K}$ and $\mathbf{X}$ to be *compact basic semialgebraic sets*.

Assumption 2.7 allows us to use Theorem 2.4 and reformulate the volume prob-

lem (2.8a):

$$\begin{aligned}
\mathrm{p}^\star_{\mathbf{K}} = \sup_{\mathbf{z}}\ & z_{\mathbf{0}} \\
\text{s.t. } & \mathbf{z} = (z_{\mathbf{k}})_{\mathbf{k}\in\mathbb{N}^n} \in \mathbb{R}^{\mathbb{N}^n} \\
& \forall p \in \mathbb{R}[\mathbf{x}]\ 0 \leq \mathrm{L}_{\mathbf{z}}(p^2) \leq \int_{\mathbf{X}} p(\mathbf{x})^2\ \mathrm{d}\mathbf{x} \\
& \forall p \in \mathbb{R}[\mathbf{x}]\ \mathrm{L}_{\mathbf{z}}(g_{\mathbf{K},i}\ p^2) \geq 0 && i \in \mathbb{N}^\star_{m_{\mathbf{K}}} \\
& \forall p \in \mathbb{R}[\mathbf{x}]\ \mathrm{L}_{\mathbf{z}}(g_{\mathbf{X},i}\ p^2) \leq \int_{\mathbf{X}} g_{\mathbf{X},i}(\mathbf{x})\ p(\mathbf{x})^2\ \mathrm{d}\mathbf{x} && i \in \mathbb{N}^\star_{m_{\mathbf{X}}}.
\end{aligned}$$

Here, according to Putinar's P-satz 2.4, we ensure that $\mathbf{z}$ is the moment sequence of a measure $\mu \in \mathcal{M}(\mathbf{K})_+$ by enforcing $\mathrm{L}_{\mathbf{z}}(p^2) \geq 0$ and $\mathrm{L}_{\mathbf{z}}(g_{\mathbf{K},i}\, p^2) \geq 0$ for all $p \in \mathbb{R}[\mathbf{x}]$ and $i \in \mathbb{N}^\star_{m_{\mathbf{K}}}$. Similarly, the upper bounding constraints on the Riesz functionals ensure that the measure $\lambda - \mu$, whose moment sequence is $(\int_{\mathbf{X}} \mathbf{x}^{\mathbf{k}}\ \mathrm{d}\mathbf{x} - z_{\mathbf{k}})_{\mathbf{k}\in\mathbb{N}^n}$, is nonnegative on $\mathbf{X}$.

Then, Lasserre's moment hierarchy only consists in relaxing the moment constraints by replacing the infinite dimensional test space $\mathbb{R}[\mathbf{x}]$ with the dimension $D^d_n := \binom{n+d}{n}$ test space $\mathbb{R}_d[\mathbf{x}]$ of degree at most d polynomials, reducing the search to finite *pseudo-moment* sequences:

$$\begin{aligned}
\mathrm{p}^d_{\mathbf{K}} := \sup_{\mathbf{z}}\ & z_{\mathbf{0}} \\
\text{s.t. } & \mathbf{z} = (z_{\mathbf{k}})_{|\mathbf{k}|\leq 2d} \in \mathbb{R}^{\mathbb{N}^n_{2d}} \\
& \forall p \in \mathbb{R}_d[\mathbf{x}]\ 0 \leq \mathrm{L}_{\mathbf{z}}(p^2) \leq \int_{\mathbf{X}} p(\mathbf{x})^2\ \mathrm{d}\mathbf{x} \\
& \forall p \in \mathbb{R}_{d-d_{\mathbf{K},i}}[\mathbf{x}]\ \mathrm{L}_{\mathbf{z}}(g_{\mathbf{K},i}\ p^2) \geq 0 && i \in \mathbb{N}^\star_{m_{\mathbf{K}}} \\
& \forall p \in \mathbb{R}_{d-d_{\mathbf{X},i}}[\mathbf{x}]\ \mathrm{L}_{\mathbf{z}}(g_{\mathbf{X},i}\ p^2) \leq \int_{\mathbf{X}} g_{\mathbf{X},i}(\mathbf{x})\ p(\mathbf{x})^2\ \mathrm{d}\mathbf{x} && i \in \mathbb{N}^\star_{m_{\mathbf{X}}},
\end{aligned} \tag{2.11}$$

where $|\mathbf{k}| = k_1 + \cdots + k_n$, $\mathbb{N}^n_{2d} := \{\mathbf{k} \in \mathbb{N}^n : |\mathbf{k}| \leq 2d\}$, $d_{\mathbf{K},i} := \lceil \mathrm{d}^\circ g_{\mathbf{K},i}/2 \rceil$ and $d_{\mathbf{X},i} := \lceil \mathrm{d}^\circ g_{\mathbf{X},i}/2 \rceil$ so that the moment constraints only involve the $z_{\mathbf{k}}$ for $|\mathbf{k}| \leq 2d$.

The fundamental difference between problems (2.8a) and (2.11) is that the former is an infinite dimensional linear cone programming problem, while the latter is a finite dimensional semidefinite programming problem whose constraints can be reformulated as Linear Matrix Inequalities (LMI), through the following definition.

Definition 2.8: Localizing & moment matrices

Let $d, d_g \in \mathbb{N}$, $g \in \mathbb{R}_{d_g}[\mathbf{x}]$. Let $\mathbf{e}_d(\mathbf{x}) := (e_i(\mathbf{x}))_{i\leq D^d_n}$ be a basis of $\mathbb{R}_d[\mathbf{x}]$.
Let $\mathbf{z} = (z_{\mathbf{k}})_{|\mathbf{k}|\leq 2d+d_g} \in \mathbb{R}^{\mathbb{N}^n_{2d+d_g}}$, where we recall that $|\mathbb{N}^n_d| = D^d_n = \binom{n+d}{n}$.

- The *degree d localizing matrix* $\mathrm{M}_d(g\ \mathbf{z})$ of $\mathbf{z}$ in g is defined as the size D^d_n matrix representation in basis $\mathbf{e}_d(\mathbf{x})$ of the bilinear application
$$(p, q) \in \mathbb{R}_d[\mathbf{x}]^2 \longmapsto \mathrm{L}_{\mathbf{z}}(g\ p\ q).$$
- The *degree d moment matrix* of $\mathbf{z}$ is defined as $\mathrm{M}_d(\mathbf{z}) := \mathrm{M}_d(1\ \mathbf{z})$.

The localizing matrix is defined so that if $p(\mathbf{x}) = \mathbf{p} \cdot \mathbf{e}_d(\mathbf{x})$ and $q(\mathbf{x}) = \mathbf{q} \cdot \mathbf{e}_d(\mathbf{x})$, $\mathbf{p}, \mathbf{q} \in \mathbb{R}^{D_n^d}$, then

$$\mathrm{L}_{\mathbf{z}}(g\ p\ q) = \mathbf{p}^\top \mathrm{M}_d(g\ \mathbf{z})\ \mathbf{q} \qquad \& \qquad \mathrm{L}_{\mathbf{z}}(g\ p^2) = \mathbf{p}^\top \mathrm{M}_d(g\ \mathbf{z})\ \mathbf{p}.$$

Example 2.13 *Let us focus on univariate polynomials ($n = 1$): $\mathbf{z} = (z_k)_{1 \le k \le 2d+d_g}$.*
Let $d = 2$, $g(x) := 1 - x^2$, $\mathbf{e}_2(x) := (1, x, x^2)$. Then, $\mathbf{z} = (z_1, z_2, z_3, z_4, z_5, z_6)$ and

- $\mathrm{M}_2(\mathbf{z}) = \begin{pmatrix} z_0 & z_1 & z_2 \\ z_1 & z_2 & z_3 \\ z_2 & z_3 & z_4 \end{pmatrix}$
- $\mathrm{M}_2(g\ \mathbf{z}) = \begin{pmatrix} z_0 - z_2 & z_1 - z_3 & z_2 - z_4 \\ z_1 - z_3 & z_2 - z_4 & z_3 - z_5 \\ z_2 - z_4 & z_3 - z_5 & z_4 - z_6 \end{pmatrix}$

Now let $n = 2$, $d = 1$, $g(x, y) = 1 + x - y$, $\mathbf{e}_1(x, y) := (1, x, y)$.
Then, $\mathbf{z} = (z_{00}, z_{10}, z_{01}, z_{20}, z_{11}, z_{02}, z_{30}, z_{21}, z_{12}, z_{03})$ and

- $\mathrm{M}_1(\mathbf{z}) = \begin{pmatrix} z_{00} & z_{10} & z_{01} \\ z_{10} & z_{20} & z_{11} \\ z_{01} & z_{11} & z_{02} \end{pmatrix}$
- $\mathrm{M}_1(g\ \mathbf{z}) = \begin{pmatrix} z_{00} + z_{10} - z_{01} & z_{10} + z_{20} - z_{11} & z_{01} + z_{11} - z_{02} \\ z_{10} + z_{20} - z_{11} & z_{20} + z_{30} - z_{21} & z_{11} + z_{21} - z_{12} \\ z_{01} + z_{11} - z_{02} & z_{11} + z_{21} - z_{12} & z_{02} + z_{12} - z_{03} \end{pmatrix}$

Remark 2.14 (Localizing matrix computation)
One can observe the general property that if $\mathrm{M}_d(\mathbf{z}) = (m_{ij})_{ij}$ is represented by its coefficients, and $\mathbf{e}_d(\mathbf{x}) = (\mathbf{x}^{\mathbf{k}_i})_i$ is a basis of monomials, then $m_{ij} = m_{ji} = z_{\mathbf{k}_i + \mathbf{k}_j}$, so that the moment matrix computation is straightforward.
Also, always taking a basis of monomials $(\mathbf{x}^{\mathbf{k}_i})_i$ as $\mathbf{e}_d(\mathbf{x})$, if $g(\mathbf{x}) = \sum_{\mathbf{k}} g_{\mathbf{k}}\ \mathbf{x}^{\mathbf{k}}$ and $\mathrm{S}_{\mathbf{k}_0} := (z_{\mathbf{k}})_{\mathbf{k}} \mapsto (z_{\mathbf{k}+\mathbf{k}_0})_{\mathbf{k}}$ is the shift operator, then $\mathrm{M}_d(g\ \mathbf{z}) = \sum_{\mathbf{k}} g_{\mathbf{k}}\ \mathrm{M}_d(\mathrm{S}_{\mathbf{k}}\ \mathbf{z})$, so that the localizing matrix is quickly deduced from the moment matrix.
For this reason, we define the shift map $\mathbf{z} \mapsto g\ \mathbf{z} := \sum_{\mathbf{k}} g_{\mathbf{k}}\ \mathrm{S}_{\mathbf{k}}\ \mathbf{z}$.
Eventually, if $z_{\mathbf{k}} = \int \mathbf{x}^{\mathbf{k}}\, d\mu$ for some nonnegative measure μ, then one has the following relation between μ and $\mathrm{M}_d(\mathbf{z})$:

$$\mathrm{M}_d(\mathbf{z}) = \int \mathbf{e}_d(\mathbf{x})\ \mathbf{e}_d(\mathbf{x})^\top\, d\mu(\mathbf{x}) \succeq 0,$$

where for a symmetric $\mathrm{M} \in \mathbb{S}^n$, $\mathrm{M} \succeq 0$ means that M is positive semidefinite i.e. $\mathrm{M} \in \mathbb{S}^n_+$.

In the rest of this thesis, we work with a basis of monomials $\mathbf{e}_d(\mathbf{x}) = (\mathbf{x}^{\mathbf{k}_i})_i$. The interest of Definition 2.8 is that it characterizes nonnegativity of $p \mapsto \mathrm{L}_{\mathbf{z}}(g\ p^2)$ as positive semidefiniteness of $\mathrm{M}_d(g\ \mathbf{z})$, allowing us to reformulate problem (2.11) as:

Problem 7: Moment relaxation hierarchy

For $d \in \mathbb{N}^\star$ large enough, compute

$$\begin{aligned} \mathrm{p}^d_{\mathbf{K}} := \sup_{\mathbf{z}}\ & z_{\mathbf{0}} \\ \text{s.t. } & \mathbf{z} = (z_{\mathbf{k}})_{|\mathbf{k}| \le 2d} \in \mathbb{R}^{\mathbb{N}^n_{2d}} \\ & 0 \preceq \mathrm{M}_d(\mathbf{z}) \preceq \mathrm{M}_d(\mathbf{l}) \\ & \mathrm{M}_{d - d_{\mathbf{K},i}}(g_{\mathbf{K},i}\ \mathbf{z}) \succeq 0 \qquad i \in \mathbb{N}^\star_{m_{\mathbf{K}}} \\ & \mathrm{M}_{d - d_{\mathbf{X},i}}(g_{\mathbf{X},i}\ (\mathbf{l} - \mathbf{z})) \succeq 0 \qquad i \in \mathbb{N}^\star_{m_{\mathbf{X}}}, \end{aligned} \tag{2.12}$$

where $\mathbf{l} := (\int_{\mathbf{X}} \mathbf{x}^{\mathbf{k}} \, \mathrm{d}\mathbf{x})_{|\mathbf{k}|\leq 2d}$ is the truncated moment sequence of the Lebesgue measure on $\mathbf{X}$ and $\mathrm{M} \preceq \mathrm{N}$ means that $\mathrm{N} - \mathrm{M} \succeq 0$. Here we are faced with a finite dimensional optimization problem whose constraints are written under the form of size D_n^d LMIs, so that it can be tackled by semidefinite programming solvers, at least for small values of D_n^d.

In order to formulate the moment hierarchy, we *relaxed* the infinite dimensional constraint that a sequence should represent a measure, into *less restrictive* finite dimensional semidefinite positivity constraints on localizing matrices. In other words, the feasible set for (2.12) is *larger* than the feasible set for (2.8a), in the sense that if $\mathbf{z}$ represents a feasible measure μ for (2.8a), then it is feasible for the (2.12). For these reasons, the moment hierarchy is said to be a *hierarchy of relaxations*, and $\mathrm{p}_{\mathbf{K}}^{\star} \leq \mathrm{p}_{\mathbf{K}}^{d}$.

The same reasoning allows us to define $\mathrm{p}_{\mathrm{GM}}^{d} \geq \mathrm{p}_{\mathrm{GM}}^{\star}$ as long as $\mathbf{c}$ and the $\boldsymbol{\varphi}_k$ and $\boldsymbol{\psi}_k$ can be chosen as families of polynomials, which is the case if $\mathcal{Y}'$ is dual to a space of $\mathcal{C}^k$ functions with compact supports, allowing the use of the Stone-Weierstraß theorem to reformulate all constraints into moment constraints.

We end this section with a quite useful result .

Lemma 2.9: Pseudo moment sequences boundedness

Let $d \in \mathbb{N}^{\star}$, $R > 0$, $\mathbf{z} \in \mathbb{R}^{\mathbb{N}^n}$ s.t. $\mathrm{M}_d(\mathbf{z}) \succeq 0$ & $\mathrm{M}_{d-1}((R^2 - |\cdot|^2)\,\mathbf{z}) \succeq 0$. Then,

$$\max_{|\mathbf{k}|\leq 2d} |z_{\mathbf{k}}| \leq z_{\mathbf{0}} \ \max(1, R^{2d}).$$

In words, asking for the moment matrix and the ball-localizing matrix to be positive semidefinite ensures uniform boundedness of the pseudo-moment sequence.

Proof : $\mathrm{M}_d(\mathbf{z}) \succeq 0$ is equivalent to

$$\forall p \in \mathbb{R}_d[\mathbf{x}], \mathrm{L}_{\mathbf{z}}(p^2) \geq 0, \tag{a}$$

while $\mathrm{M}_{d-1}((R^2 - |\cdot|^2)\,\mathbf{z}) \succeq 0$ means that

$$\forall p \in \mathbb{R}_{d-1}[\mathbf{x}], \mathrm{L}_{\mathbf{z}}((R^2 - |\cdot|^2)\,p^2) \geq 0. \tag{b}$$

(a) with $p(\mathbf{x}) = \mathbf{x}^{\mathbf{k}}$, $|\mathbf{k}| \leq d$ yields $z_{2\mathbf{k}} \geq 0$.

(b) with $p(\mathbf{x}) = 1$ yields $R^2 z_{\mathbf{0}} \geq \sum_{|\mathbf{k}|=1} z_{2\mathbf{k}}$, since $|\mathbf{x}|^2 = \sum_{j=1}^n x_j^2 = \sum_{|\mathbf{k}|=1} \mathbf{x}^{2\mathbf{k}}$. Hence, since the $z_{2\mathbf{k}}$ are nonnegative, one has $|\mathbf{k}| = 1 \Rightarrow z_{2\mathbf{k}} \leq R^2 z_{\mathbf{0}}$.

Going forward, if $|\mathbf{k}| = 1$, (b) with $p(\mathbf{x}) = \mathbf{x}^{\mathbf{k}}$ yields $R^2 z_{2\mathbf{k}} \geq \sum_{|\mathbf{k}'|=1} z_{2(\mathbf{k}+\mathbf{k}')}$ with $z_{2(\mathbf{k}+\mathbf{k}')} \geq 0$ by (a), so that $R^4 z_{\mathbf{0}} \geq R^2 z_{\mathbf{k}} \geq z_{2(\mathbf{k}+\mathbf{k}')}$ as long as $|\mathbf{k}| = |\mathbf{k}'| = 1$, and thus, if $|\mathbf{k}| = 2$, $R^4 z_{\mathbf{0}} \geq z_{2\mathbf{k}}$. By induction, one has for $\mathbf{k} \in \mathbb{N}_d^n$ that

$$0 \leq z_{2\mathbf{k}} \leq R^{2|\mathbf{k}|} z_{\mathbf{0}} \leq z_{\mathbf{0}} \ \max(1, R^{2d}). \tag{c}$$

Let $\mathbf{k}, \mathbf{k}' \in \mathbb{N}_d^n$. Then, (a) with $p(\mathbf{x}) = \mathbf{x}^{\mathbf{k}} \pm \mathbf{x}^{\mathbf{k}'}$ yields $0 \leq \mathrm{L}_{\mathbf{z}}(p^2) = z_{2\mathbf{k}} \pm 2 z_{\mathbf{k}+\mathbf{k}'} + z_{2\mathbf{k}'}$ so that

$$|z_{\mathbf{k}+\mathbf{k}'}| \leq \frac{z_{2\mathbf{k}} + z_{2\mathbf{k}'}}{2} \leq \max(z_{2\mathbf{k}}, z_{2\mathbf{k}'}) \overset{(c)}{\leq} z_{\mathbf{0}} \ \max(1, R^{2d}). \tag{d}$$

Eventually, any $\mathbf{k} \in \mathbb{N}^n_{2d}$ can be written $\mathbf{k} = \mathbf{k}' + \mathbf{k}''$ with $\mathbf{k}', \mathbf{k}'' \in \mathbb{N}^n_d$, so that it satisfies (d), and $|z_\mathbf{k}| \leq z_\mathbf{0}\ \max(1, R^{2d})$.

♢

This lemma proves several important things, among which any nonzero feasible vector $\mathbf{z}$ for (2.11) satisfies $z_\mathbf{0} > 0$, and if $R_\mathbf{X} \leq 1$ in Assumption 2.7 and $z_\mathbf{0} \leq C$ is enforced, then for all $d \in \mathbb{N}$, $\mathbf{z}$ feasible for (2.11) satisfies $|z_\mathbf{k}| \leq C\ \forall \mathbf{k}$. These features of pseudo-moment sequences will be useful to state important results in Section 2.2.3.

2.2.2 Sums of squares

We then proceed with the dual (2.9b) on continuous functions, that are more commonly used than measures in the field of automatics. For the sake of clarity and without loss of generality, we keep working with the particular instance of the volume problem:

$$\begin{aligned} \mathrm{d}^\star_\mathbf{K} = \inf_w \ & \int w\, d\lambda \\ \text{s.t. } & w - 1 \in \mathcal{C}(\mathbf{K})_+ \\ & w \in \mathcal{C}(\mathbf{X})_+, \end{aligned} \tag{2.8b}$$

with $\mathbf{K} \subset \mathbf{X} \subset \mathbb{R}^n$ compact sets. First, the Stone-Weierstraß theorem allows approximating the function w with polynomials without changing the value of the problem, yielding:

$$\begin{aligned} \mathrm{d}^\star_\mathbf{K} = \inf_w \ & \int w\, d\lambda \\ \text{s.t. } & w - 1 \in \mathcal{R}(\mathbf{K})_+ \\ & w \in \mathcal{R}(\mathbf{X})_+, \end{aligned}$$

where $\mathcal{R}(\mathbf{X})_+ := \mathbb{R}[\mathbf{x}] \cap \mathcal{C}(\mathbf{X})_+$ is the cone of polynomials that are nonnegative on $\mathbf{X}$. Here the problem of deciding whether a polynomial is nonnegative on a set appears once again. However, we now give a theorem to tackle this difficulty when $\mathbf{X}$ and $\mathbf{K}$ are compact basic semialgebraic sets.

To state our theorem, we introduce the notion of quadratic module:

Definition 2.10: Quadratic module

Let $m \in \mathbb{N}^\star$, $\mathbf{g} \in \mathbb{R}[\mathbf{x}]^m$. We define:

- The cone $\Sigma[\mathbf{x}] := \{p_1^2 + \ldots + p_K^2 : K \in \mathbb{N} \wedge p_1, \ldots, p_K \in \mathbb{R}[\mathbf{x}]\}$ of polynomial *sums of squares* (SOS)
- The *quadratic module* (cone) $\Sigma(\mathbf{g}) := \{s_0 + \mathbf{s} \cdot \mathbf{g} : s_0 \in \Sigma[\mathbf{x}] \wedge \mathbf{s} \in \Sigma[\mathbf{x}]^m\}$.

The notion of quadratic module is instrumental in the statement of a dual version of Theorem 2.4, that is also called Putinar's Positivstellensatz.

Theorem 2.11: Putinar's (dual) Positivstellensatz (P-satz)

Let $\mathbf{g} \in \mathbb{R}[\mathbf{x}]^m$, $R > 0$, $\mathbf{K} := \{\mathbf{x} \in \mathbb{R}^n : \mathbf{g}(\mathbf{x}) \geq \mathbf{0} \wedge |\mathbf{x}|^2 \leq R^2\}$. Then,

$$\mathcal{R}(\mathbf{K})_{++} \subset \Sigma\left(\mathbf{g}, R^2 - |\cdot|^2\right),$$

where $\mathcal{R}(\mathbf{X})_{++} := \{p \in \mathbb{R}[\mathbf{x}] : \forall \mathbf{x} \in \mathbf{X}, p(\mathbf{x}) > 0\}$.

In words, any polynomial $q \in \mathbb{R}[\mathbf{x}]$ that is positive on $\mathbf{K}$ can be written

$$q(\mathbf{x}) = s_0(\mathbf{x}) + \mathbf{s}(\mathbf{x}) \cdot \mathbf{g}(\mathbf{x}) + s_R(\mathbf{x})\ (R^2 - |\mathbf{x}|^2)$$

with $(s_0, \mathbf{s}, s_R) \in \Sigma[\mathbf{x}]^{m+2}$.

Theorem 2.11 is the actual formulation [107, Theorem 1.3] Putinar presented as the Positivstellensatz in his paper, while its primal formulation 2.4 is a lemma he stated to prove it. In [107], Putinar proved that these two different formulations are equivalent. In terms of duality: positivity of q is dual to existence of the measure μ, s_0 is dual to $\mathrm{L}_{\mathbf{z}}(p^2) \geq 0$, $\mathbf{s}$ is dual to the $\mathrm{L}_{\mathbf{z}}(g_i\ p^2) \geq 0$, and s_R is dual to $\mathrm{L}_{\mathbf{z}}((R^2 - |\mathbf{x}|^2)\ p^2) \geq 0$.

Thus, under Assumption 2.7, Theorem 2.11 ensures that any nonnegative polynomial $p \in \mathcal{R}(\mathbf{K})_+$ (resp. $\mathcal{R}(\mathbf{X})_+$) can be approximated with positive polynomials $p + \epsilon \in \Sigma(\mathbf{g_K})$ (resp. $\Sigma(\mathbf{g_X})$), $\epsilon > 0$, so that (2.8b) can be reformulated as

$$\begin{aligned} \mathrm{d}^\star_{\mathbf{K}} = \inf_w & \int_{\mathbf{X}} w(\mathbf{x})\, \mathrm{d}\mathbf{x} \\ \text{s.t. } & w - 1 \in \Sigma(\mathbf{g_K}) \\ & w \in \Sigma(\mathbf{g_X}). \end{aligned}$$

Then, the Lasserre SOS hierarchy simply consists in restricting the feasible set to given finite degree $d \in \mathbb{N}$. To write such a rephrazing in a synthetic way, we introduce a notion of finite dimensional quadratic module.

Definition 2.12: Bounded degree quadratic module

Let $\mathbf{g} \in \mathbb{R}[\mathbf{x}]^m$, $d \in \mathbb{N}$. For $i \in \mathbb{N}^\star_m$, define $d_i := \lceil \mathrm{d}^\circ g_i/2 \rceil$ the half degree of the i-th component of $\mathbf{g}$. We define:

- The cone $\Sigma_d[\mathbf{x}] := \Sigma[\mathbf{x}] \cap \mathbb{R}_{2d}[\mathbf{x}]$ of SOS polynomials of degree $2d$ or less
- The degree $2d$ quadratic module

$$\Sigma_d(\mathbf{g}) := \{s_0 + \mathbf{s} \cdot \mathbf{g} : s_0 \in \Sigma_d[\mathbf{x}] \wedge \forall i \in \mathbb{N}^\star_m, s_i \in \Sigma_{d-d_i}[\mathbf{x}]\} \subset \mathbb{R}_{2d}[\mathbf{x}].$$

This last definition allows us to formulate a new finite degree optimization problem:

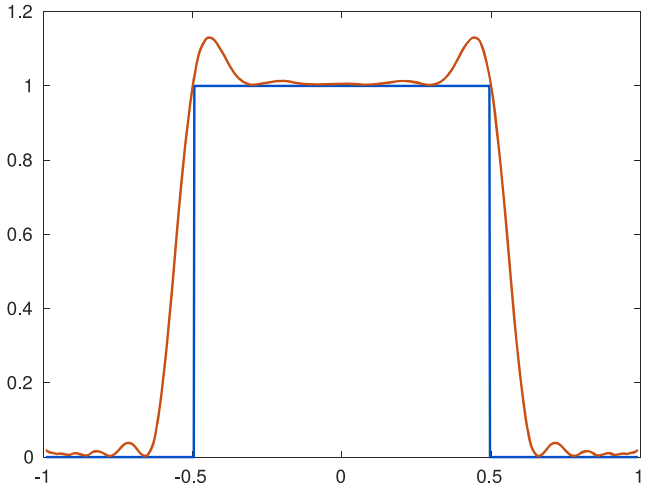


Figure 2.3 – SOS hierarchy for the length of $[-0.5, 0.5]$.

The SOS hierarchy yields a polynomial w_d represented in red, whose integral approximates by above the area below the blue curve.

Problem 8: Sums-of-squares strengthening hierarchy

For $d \in \mathbb{N}^\star$ large enough, compute

$$\begin{aligned} \mathrm{d}_{\mathbf{K}}^d := \inf_{w} \ & \int_{\mathbf{X}} w(\mathbf{x}) \, \mathrm{d}\mathbf{x} \\ \text{s.t. } & w - 1 \in \Sigma_d(\mathbf{g}_{\mathbf{K}}) \\ & w \in \Sigma_d(\mathbf{g}_{\mathbf{X}}). \end{aligned} \tag{2.13}$$

The fundamental difference between problems (2.8b) and (2.13) is that the former is an infinite dimensional linear cone programming problem, while the latter is a finite dimensional semidefinite programming problem. Indeed, the following result holds (see [70, Proposition 2.1.]):

Proposition 2.13: SOS programming

Let $s \in \mathbb{R}_{2d}[\mathbf{x}]$ be a degree $2d$ polynomial, $d \in \mathbb{N}$. Let $\mathbf{e}_d(\mathbf{x})$ be a basis of $\mathbb{R}_d[\mathbf{x}]$.
Then, $s \in \Sigma_d[\mathbf{x}]$ *iff* there exists a positive semidefinite real matrix $\mathrm{S} \in \mathbb{S}_+^{D_n^d}$ s.t.

$$s(\mathbf{x}) = \mathbf{e}_d(\mathbf{x})^\top \, \mathrm{S} \, \mathbf{e}_d(\mathbf{x}).$$

In words, deciding whether s is SOS reduces to a size D_n^d Linear Matrix Inequality (LMI) feasibility problem, tractable on a computer for small values of D_n^d.

In order to formulate the SOS hierarchy, we *strengthened* the infinite dimensional nonnegativity constraints into *more restrictive* finite dimensional SOS constraints. In other words, the feasible set for (2.13) is *smaller* than the feasible set for (2.8b), in the sense that feasibility for the former is sufficient to ensure feasibility for the latter. For these reasons, the SOS hierarchy is said to be a *hierarchy of strengthenings*, and

$\mathrm{d}_{\mathbf{K}}^\star \leq \mathrm{d}_{\mathbf{K}}^d$ (see Figure 2.3).

The same reasoning allows us to define $\mathrm{d}_{\mathrm{GM}}^d \geq \mathrm{d}_{\mathrm{GM}}^\star$ as long as $\mathbf{c}$ and the $\boldsymbol{\varphi}_k$ and $\boldsymbol{\psi}_k$ can be chosen as families of polynomials, which is the case if $\mathcal{Y}'$ is dual to a space of $\mathcal{C}^k$ functions with compact supports, allowing the use of the Stone-Weierstraß theorem to reformulate all constraints into moment constraints.

The approximation power of the SOS hierarchy comes from the following theorem:

Theorem 2.14: Convergence of the SOS hierarchy

Under Assumption 2.7, $(\mathrm{d}_{\mathrm{GM}}^d)_d$ converges monotonically towards $\mathrm{d}_{\mathrm{GM}}^\star$:

$$\mathrm{d}_{\mathrm{GM}}^d \underset{d\to\infty}{\searrow} \mathrm{d}_{\mathrm{GM}}^\star.$$

Proof : Monotonicity of $(\mathrm{d}_{\mathrm{GM}}^d)_d$ is a direct consequence of the fact that we are dealing with a hierarchy of strengthenings: increasing d is equivalent to relaxing constraints, making it possible to go further in the minimization problem, so that $\mathrm{d}_{\mathrm{GM}}^{d+1} \leq \mathrm{d}_{\mathrm{GM}}^d$.

Convergence comes from the way we formulated the hierarchy, so that it is sufficient to prove it for the volume problem (2.8b): let $\epsilon > 0$. Since

$$\begin{aligned} \mathrm{d}_{\mathbf{K}}^\star = \inf_w \ & \int_{\mathbf{X}} w(\mathbf{x})\,\mathrm{d}\mathbf{x} \\ \text{s.t. } & w - 1 \in \Sigma(\mathbf{g}_{\mathbf{K}}) \\ & w \in \Sigma(\mathbf{g}_{\mathbf{X}}), \end{aligned}$$

there exists $w_\epsilon \in \Sigma(\mathbf{g}_{\mathbf{X}})$ such that $w_\epsilon - 1 \in \Sigma(\mathbf{g}_{\mathbf{K}})$ and $\mathrm{d}_{\mathbf{K}}^\star \leq \int w_\epsilon(\mathbf{x})\,\mathrm{d}\mathbf{x} \leq \mathrm{d}_{\mathbf{K}}^\star + \epsilon$. Since w_ϵ is a (finite degree) polynomial, there exists $d_1, d_2 \in \mathbb{N}$ s.t. $w_\epsilon - 1 \in \Sigma_{d_1}(\mathbf{g}_{\mathbf{K}})$ and $w_\epsilon \in \Sigma_{d_2}(\mathbf{g}_{\mathbf{X}})$. Thus, taking $d = d_0 := \max(d_1, d_2)$, w_ϵ is feasible for problem (2.13), which yields that

$$\mathrm{d}_{\mathbf{K}}^\star \leq \mathrm{d}_{\mathbf{K}}^{(d_0)} \leq \int w_\epsilon(\mathbf{x})\,\mathrm{d}\mathbf{x} \leq \mathrm{d}_{\mathbf{K}}^\star + \epsilon.$$

Monotonicity finally ensures that for all $d \geq d_0$, $\mathrm{d}_{\mathbf{K}}^\star \leq \mathrm{d}_{\mathbf{K}}^d \leq \mathrm{d}_{\mathbf{K}}^\star + \epsilon$, which is the definition of convergence of $(\mathrm{d}_{\mathbf{K}}^d)_d$ to $\mathrm{d}_{\mathbf{K}}^\star$. ♢

2.2.3 Duality in the hierarchy ⋆

As in the infinite dimensional case of section 2.1, we study the duality properties of the moment-SOS hierarchy.

Skipping the details, duality between the P-satz formulations 2.4 and 2.11 as well as the Lagrangian function

$$\mathcal{L}(\mathbf{z}, w) := z_{\mathbf{0}} + \int_{\mathbf{X}} w(\mathbf{x})\,\mathrm{d}\mathbf{x} - \mathrm{L}_{\mathbf{z}}(w),$$

where $\mathbf{z} = (z_{\mathbf{k}})_{|\mathbf{k}| \leq 2d} \in \mathbb{R}^{\mathbb{N}^n_{2d}}$ has positive semidefinite moment and $\mathbf{g_K}$ localization matrices and $w \in \Sigma_d(\mathbf{g_X})$, can be used to prove that the semidefinite programming problems (2.12) and (2.13) are dual w.r.t. each other.

A direct consequence of this fact is that the weak duality property of Proposition 2.5 also holds for the moment-SOS hierarchy: $\mathrm{p}^d_{\mathbf{K}} \leq \mathrm{d}^d_{\mathbf{K}}$, and more generally

$$\mathrm{p}^d_{\mathrm{GM}} \leq \mathrm{d}^d_{\mathrm{GM}},$$

as the moment and SOS hierarchies are always dual w.r.t. each other.

The weak duality property allows us to prove convergence of the moment hierarchy.

Theorem 2.15: Convergence of the moment hierarchy

Suppose that strong duality holds for the GMP: $\mathrm{p}^\star_{\mathrm{GM}} = \mathrm{d}^\star_{\mathrm{GM}}$. Then, under Assumption 2.7, $(\mathrm{p}^d_{\mathrm{GM}})_d$ converges monotonically towards $\mathrm{p}^\star_{\mathrm{GM}}$:

$$\mathrm{p}^d_{\mathrm{GM}} \underset{d\to\infty}{\searrow} \mathrm{p}^\star_{\mathrm{GM}}.$$

Proof : Monotonicity of $(\mathrm{p}^d_{\mathrm{GM}})_d$ is a direct consequence of the fact that we are dealing with a hierarchy of relaxations: decreasing d is equivalent to relaxing constraints, making it possible to go further in the maximization problem, so that $\mathrm{p}^{d-1}_{\mathrm{GM}} \geq \mathrm{p}^d_{\mathrm{GM}}$.

Then, one has, for $d \in \mathbb{N}$,

$$\mathrm{d}^\star_{\mathrm{GM}} \overset{\text{strong duality}}{=} \mathrm{p}^\star_{\mathrm{GM}} \overset{\text{relaxation}}{\leq} \mathrm{p}^d_{\mathrm{GM}} \overset{\text{weak duality}}{\leq} \mathrm{d}^d_{\mathrm{GM}},$$

so that one easily concludes using Theorem 2.14 and the sandwich rule. ♢

Remark 2.15 (Application of strong duality in the GMP)

Theorem 2.16 is a good example of the relevance of Theorem 2.6. Indeed, for Theorem 2.16 to hold, i.e. for the moment hierarchy to actually approximate the GMP, one needs strong duality in the GMP. More generally, strong duality in the GMP is a necessary condition for the moment-SOS hierarchy to work at its full potential.

Of course, the strong duality Theorem 2.6 has a hierarchy counterpart that we are now going to state. For the sake of simplicity, in the rest of this section we suppose that $N = 1$ in Problem 6. All the following results still hold with several measures as decision variable, at the price of less readable proofs.

Theorem 2.16: Strong duality in the hierarchy

Suppose that there exists $C > 0$ s.t. if $\mathbf{z} \in \mathbb{R}^{\mathbb{N}^n_{2d}}$ is feasible for the degree d relaxation of (2.9a), then $z_{\mathbf{0}} \leq C$. Suppose that there exists such $\mathbf{z}$. In that case, under Assumption 2.7,

$$\mathrm{p}^d_{\mathrm{GM}} = \mathrm{d}^d_{\mathrm{GM}}.$$

Moreover, the degree d relaxation of (2.9a) has an optimal solution $\mathbf{z}_d \in \mathbb{R}^{\mathbb{N}^n_{2d}}$.

Proof : We illustrate the proof with the moment-SOS hierarchy associated to the volume problem (2.8). The exact same proof can be generalized to the GMP (2.9) at the price of introducing all the notations for the general moment-SOS hierarchy, which does not bring any additional theoretical insight. The proof is similar in spirit to the one of Theorem 2.6. The only change is that the cone $\mathcal{K}$ that we consider is now different, so we only detail the part on closedness of

$$\mathcal{K}_d := \left\{(\mathrm{A}_d\ \mathbf{z}, \mathbf{c}\cdot\mathbf{z}) : \mathbf{z} \in \mathbb{R}^{\mathbb{N}^n_{2d}} \text{ feasible for the degree } d \text{ relaxation}\right\},$$

where in the case of (2.11) $\mathrm{A}_d\ \mathbf{z} = \left(\mathrm{M}_d(\mathbf{z}), (\mathrm{M}_{d-d_{\mathbf{X},i}}(g_{\mathbf{X},i}\ \mathbf{z}))_{i\in\mathbb{N}^\star_{m_\mathbf{X}}}\right)$ is a localizing matrix operator, and $c_\mathbf{k} = 0$ if $\mathbf{k} \neq \mathbf{0}$, $c_\mathbf{0} = 1$. In general, $\mathrm{A}_d : \mathbb{R}^{D^{2d}_n} \to \mathbb{V}$ is a generic linear map onto a finite dimensional real vector space $\mathbb{V}$, and $\mathbf{c} \in \mathbb{R}^{D^{2d}_n}$ a generic vector of coefficients, so we will do the proof with abstract A_d and $\mathbf{c}$. As for Theorem 2.6, we only need to exhibit a compact convex base $\mathcal{P}_d$ of

$$\mathcal{X}_{d+} := \left\{\mathbf{z} \in \mathbb{R}^{D^{2d}_n} : \mathrm{M}_d(\mathbf{z}) \succeq 0 \wedge \forall i \in \mathbb{N}^\star_{m_\mathbf{K}}, \mathrm{M}_{d-d_{\mathbf{K},i}}(g_{\mathbf{K},i}\ \mathbf{z}) \succeq 0\right\}$$

to prove that $\mathcal{K}_d$ is closed. To that end, we define

$$\mathcal{P}_d := \{\mathbf{z} = (z_\mathbf{k})_\mathbf{k} \in \mathcal{X}_{d+} : z_\mathbf{0} = 1\}.$$

$\mathcal{P}_d$ is a base of $\mathcal{X}_{d+}$ in the sense that $\mathcal{X}_{d+}\setminus\{\mathbf{0}\}$ is isomorphic to $\mathbb{R}_{++}\times\mathcal{P}_d$ through the bijective application $\chi_d : (t, \mathbf{p}) \mapsto t\ \mathbf{p}$, with $\chi_d^{-1}(\mathbf{z}) = (z_\mathbf{0}, \mathbf{z}/z_\mathbf{0})$ (using Lemma 2.9, for any $\mathrm{M}_d(\mathbf{z})$ and $\mathrm{M}_{d'}((R^2 - |\mathbf{x}|^2)\ \mathbf{z})$ to be simultaneously positive semi-definite with $\mathbf{z} \neq \mathbf{0}$, it is necessary that $z_\mathbf{0} > 0$).

$\mathcal{P}_d$ is convex. Indeed, let $\mathbf{p}_1, \mathbf{p}_2 \in \mathcal{P}_d$, $t \in [0, 1]$, $\widetilde{\mathbf{p}} := t\ \mathbf{p}_1 + (1-t)\ \mathbf{p}_2$. Then, by linearity of the localizing matrix operator,

$$\forall g \in \mathbb{R}[\mathbf{x}]\ \mathrm{M}_d(g\ \widetilde{\mathbf{p}}) = t\ \mathrm{M}_d(g\ \mathbf{p}_1) + (1-t)\ \mathrm{M}_d(g\ \mathbf{p}_2),$$

so that its semidefinite positiveness is preserved by convex combination, by convexity of $\mathbb{S}^n_+$ for any $n \in \mathbb{N}$. Thus, $\widetilde{\mathbf{p}} \in \mathcal{X}_{d+}$. Besides,

$$\widetilde{p}_\mathbf{0} = t\ p_{1\mathbf{0}} + (1-t)\ p_{2\mathbf{0}} = t + 1 - t = 1$$

so that $\widetilde{\mathbf{p}} \in \mathcal{P}_d$, which proves convexity.

We now move on to showing compactness of $\mathcal{P}_d$. According to Lemma 2.9, the ball constraint in the description of $\mathbf{K}$ and the upper bound $z_\mathbf{0} \leq C$ yield boundedness of $\mathcal{P}_d$.

In addition, $\mathcal{P}_d$ is closed as the intersection between the level-1 set of the continuous function $\mathbf{z} \mapsto z_\mathbf{0}$ and the closed cone $\mathcal{X}_{d+}$. Indeed $\mathcal{X}_{d+}$ is closed as the pre-image of the closed cone $(\mathbb{S}^{D^{2d}_n}_+)^{m_\mathbf{K}+1}$ by the (continuous) linear map

$$\mathbf{z} \mapsto \left(\mathrm{M}_d(\mathbf{z}), (\mathrm{M}_{d-d_{\mathbf{K},i}}(g_{\mathbf{K},i}\ \mathbf{z}))_{i\in\mathbb{N}^\star_{m_\mathbf{K}}}\right).$$

Since finite dimensional closed bounded sets are compact, this proves that $\mathcal{P}_d$ is compact.

Finally, $\mathcal{P}_d$ is indeed a compact convex base of $\mathcal{X}_{d+}$, and the rest of the proof is identical to what we did for Theorem 2.6. ♢

Remark 2.16 (Motivation for strong duality in the hierarchy)

Strong duality in the hierarchy has a special flavor as it has a direct influence on numerical computation of approximate solutions to the GMP. Indeed, most SDP solvers are primal-dual solvers that solve both the primal and dual instances of the studied semidefinite programming problem.

Moreover, most often a duality gap $\mathrm{d}_{\mathrm{GM}}^d - \mathrm{p}_{\mathrm{GM}}^d > 0$ *will lead to numerical issues in the associated LMIs, which makes checking strong duality in the moment-SOS hierarchy an important condition to assess the efficiency of the numerical scheme.*

Eventually, the existence of an optimal pseudo-moment sequence is also a very important guarantee that the semidefinite solvers will converge.

We conclude this chapter with a strong theorem on the pointwise convergence of the Lasserre hierarchy's pseudo-moment sequences.

Theorem 2.17: Convergence of the pseudo-moment sequences

Suppose that (2.9a) has a unique minimizer $\mu^\star$ with support $\mathbf{K}$ included in the unit ball $\mathbf{B}$. Then, under the hypotheses of Theorem 2.16, there exists an optimal sequence $(\mathbf{z}_d)_{d\in\mathbb{N}}$ of optimal feasible pseudo-moment sequences for the hierarchy s.t. $\mathrm{L}_{\mathbf{z}_d}(c) = \mathrm{p}_{\mathrm{GM}}^d$ and for all $\mathbf{k} \in \mathbb{N}^n$,

$$z_{d,\mathbf{k}} \underset{d\to\infty}{\longrightarrow} \int \mathbf{x}^{\mathbf{k}}\, d\mu^\star(\mathbf{x}).$$

Moreover, this automatically yields strong duality $\mathrm{p}_{\mathrm{GM}}^d = \mathrm{d}_{\mathrm{GM}}^d$ & $\mathrm{p}_{\mathrm{GM}}^\star = \mathrm{d}_{\mathrm{GM}}^\star$.

Proof : Existence of $(\mathbf{z}_d)_d$ follows from Theorem 2.16, so we focus on the proof of convergence. Let $d \in \mathbb{N}$. For $\mathbf{k} \in \mathbb{N}^n$, define

$$\hat{z}_{d,\mathbf{k}} := \begin{cases} z_{d,\mathbf{k}} & \text{if } |\mathbf{k}| \le 2d, \\ 0 & \text{else,} \end{cases}$$

so that $\hat{\mathbf{z}}_d \in \mathbb{R}^{\mathbb{N}^n}$ with

$$\|\hat{\mathbf{z}}_d\|_{\ell^\infty(\mathbb{N}^n)} := \max_{\mathbf{k}\in\mathbb{N}^n} |\hat{z}_{d,\mathbf{k}}| = \max_{|\mathbf{k}|\le 2d} |z_{d,\mathbf{k}}| \overset{\text{Lemma 2.9}}{\le} z_{d,\mathbf{0}} \le C.$$

Then, $(\hat{\mathbf{z}}_d)_{d\in\mathbb{N}}$ is a uniformly bounded sequence of

$$\ell^\infty(\mathbb{N}^n) := \left\{ \mathbf{u} \in \mathbb{R}^{\mathbb{N}^n} : \max_{\mathbf{k}\in\mathbb{N}^n} |u_{\mathbf{k}}| < \infty \right\} = \ell^1(\mathbb{N}^n)',$$

where $\ell^1(\mathbb{N}^n) := \left\{ \mathbf{u} \in \mathbb{R}^{\mathbb{N}^n} : \sum_{\mathbf{k}\in\mathbb{N}^n} |u_{\mathbf{k}}| < \infty \right\}$. Thus, the Banach-Alaoglu theorem yields a weak-$*$ converging subsequence $(\hat{\mathbf{z}}_{d_r})_{r\in\mathbb{N}}$: $\exists \mathbf{z}_\infty \in \ell^\infty(\mathbb{N}^n); \forall \mathbf{u} \in \ell^1(\mathbb{N}^n)$,

$$\sum_{\mathbf{k}\in\mathbb{N}^n} u_{\mathbf{k}}\; z_{d_r,\mathbf{k}} \underset{r\to\infty}{\longrightarrow} \sum_{\mathbf{k}\in\mathbb{N}^n} u_{\mathbf{k}}\; z_{\infty,\mathbf{k}}.$$

In particular, if $\mathbf{k} \in \mathbb{N}^n$, $z_{d_r,\mathbf{k}} \underset{r\to\infty}{\longrightarrow} z_{\infty,\mathbf{k}}$. Thus, what we want to show is that for $\mathbf{k} \in \mathbb{N}^n$, $z_{\infty,\mathbf{k}} = z_{\mathbf{k}}^\star := \int \mathbf{x}^{\mathbf{k}}\, d\mu^\star(\mathbf{x})$.

Without loss of generality we focus on the relaxations of problem (2.8a).

Let $\mathbf{l} := (\int_{\mathbf{X}} \mathbf{x}^{\mathbf{k}}\, \mathrm{d}\mathbf{x})_{\mathbf{k}\in\mathbb{N}^n}$ be the moment sequence of the Lebesgue measure on $\mathbf{X}$.

Let $p \in \mathbb{R}[\mathbf{x}]$. Then, for all $r \in \mathbb{N}$ big enough, by feasibility of $\mathbf{z}_{d_r}$ for the relaxation of degree d_r, one has

- $[0, \mathrm{L}_{\mathbf{l}}(p^2)] \ni \mathrm{L}_{\mathbf{z}_{d_r}}(p^2) = \mathrm{L}_{\hat{\mathbf{z}}_{d_r}}(p^2) \underset{r\to\infty}{\longrightarrow} \mathrm{L}_{\mathbf{z}_\infty}(p^2)$
- $0 \leq \mathrm{L}_{\mathbf{z}_{d_r}}(g_{\mathbf{K},i}\, p^2) = \mathrm{L}_{\hat{\mathbf{z}}_{d_r}}(g_{\mathbf{X},i}\, p^2) \underset{r\to\infty}{\longrightarrow} \mathrm{L}_{\mathbf{z}_\infty}(g_{\mathbf{K},i}\, p^2)$
- $0 \leq \mathrm{L}_{\mathbf{l}-\mathbf{z}_{d_r}}(g_{\mathbf{X},i}\, p^2) = \mathrm{L}_{\mathbf{l}-\hat{\mathbf{z}}_{d_r}}(g_{\mathbf{X},i}\, p^2) \underset{r\to\infty}{\longrightarrow} \mathrm{L}_{\mathbf{l}-\mathbf{z}_\infty}(g_{\mathbf{X},i}\, p^2)$

so that according to Putinar's P-satz 2.4, $\mathbf{z}_\infty$ is the actual moment sequence of a measure μ_∞ that is feasible for problem (2.8a). Then, one directly has

$$\mathrm{p}^\star_{\mathbf{K}} \geq \int 1\, d\mu_\infty = z_{\infty,\mathbf{0}} = \lim_{r\to\infty} \hat{z}_{d_r,\mathbf{0}} = \lim_{r\to\infty} \mathrm{p}^{d_r}_{\mathbf{K}} \geq \mathrm{p}^\star_{\mathbf{K}}$$

since for any $d \in \mathbb{N}$ $\mathrm{p}^d_{\mathbf{K}} \geq \mathrm{p}^\star_{\mathbf{K}}$. Hence, $\int 1\, d\mu_\infty = \mathrm{p}^\star_{\mathbf{K}}$, *i.e.* μ_∞ is *optimal* for problem (2.8a). By our uniqueness assumption, this yields $\mu_\infty = \mu^\star$, *i.e.* $\mathbf{z}_\infty = \mathbf{z}^\star$. Thus, $(\hat{\mathbf{z}}_d)_d$ has a unique weak-$*$ accumulation point $\mathbf{z}^\star$, which means that for any $\mathbf{k} \in \mathbb{N}^n$,

$$z_{d,\mathbf{k}} \underset{d\to\infty}{\longrightarrow} z^\star_{\mathbf{k}} = \int \mathbf{x}^{\mathbf{k}}\, d\mu^\star(\mathbf{x}).$$

Eventually, Theorem 2.16 ensures strong duality $\mathrm{p}^d_{\mathrm{GM}} = \mathrm{d}^d_{\mathrm{GM}}$, so that putting together weak GMP duality and the strenghtening property, one has

$$\mathrm{p}^\star_{\mathrm{GM}} \overset{\text{weak duality}}{\leq} \mathrm{d}^\star_{\mathrm{GM}} \overset{\text{strenghtening}}{\leq} \mathrm{d}^d_{\mathrm{GM}} \overset{\text{strong duality}}{=} \mathrm{p}^d_{\mathrm{GM}} \overset{\text{convergence}}{\underset{d\to\infty}{\longrightarrow}} \mathrm{p}^\star_{\mathrm{GM}}$$

and the sandwich rule again yields strong GMP duality. ♢

Remark 2.17 (An "all inclusive" theorem)

Theorem 2.17 shows that the hypotheses of Theorem 2.16 for strong duality in the hierarchy are almost sufficient for the pseudo-moment sequences of the relaxations to converge pointwise to the moments of the optimal solution of the GMP (2.9a) *and for full strong duality to hold.*

In practice we will design instances of the GMP such that Theorem 2.17 holds, so that we automatically obtain strong duality in the GMP and the corresponding hierarchy as well as pointwise convergence of the moment sequence.

In particular, up to rescaling, $\mathbf{K} \subset \mathbf{B}$ *can always be enforced.*

Synthesis of sections 2.1 & 2.2

In this Chapter, the fundamental tools that will be used in this thesis were introduced. Section 2.1 focused on infinite dimensional optimization, introducing the GMP and its strong modelling power, as well as a useful strong duality theorem for the analysis of such problem. Then, Section 2.2 derived the Lasserre, moment-SOS hierarchy that allows numerically approximating solutions of the GMP, along with a variety of duality and convergence properties. Those two aspects (modelling infinite

dimension and approximate finite dimension) of the moment-SOS hierarchy pave the way for the smooth demonstration of approximating properties for the schemes that will be presented in the rest of this thesis. As a result, they will be instrumental in all the remaining chapters. We recall in Figure 2.4 the successive steps of the moment-SOS hierarchy scheme.

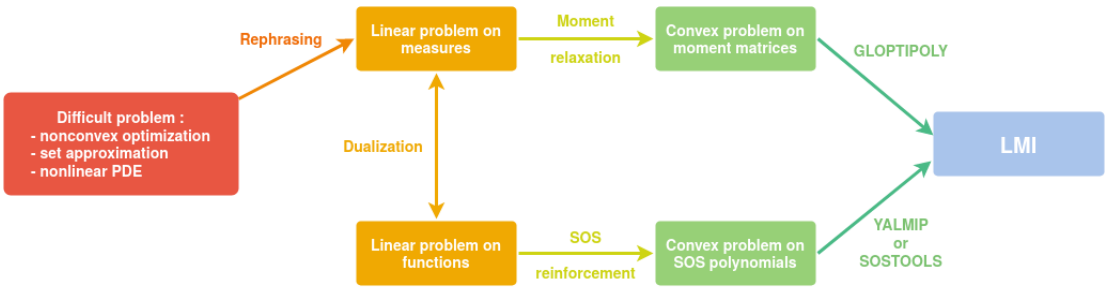


Figure 2.4 – Illustration of the moment-SOS hierarchy.

The GMP is used to model a variety of difficult, nonconvex problems, after what a dual to the GMP is computed, and the moment and SOS hierarchies are formulated in parallel, and solved as SDP problems with LMI constraints.

3

Transient stability of power systems

The aim of this thesis is to contribute to the application of semidefinite programming (SDP) methods in power systems transient stability analysis (TSA). A natural first step is thus to give original applications of existing SDP methods to low-dimensional power systems. Currently, two schools have issued promising results in the field of SDP-based direct stability analysis:

- the French school draws on the developments presented in Chapter 2 and moment-SOS hierarchy-based schemes to approximate various stability regions of differential systems [42, 58, 57] and its first contribution to power systems rather focused on the static optimization AC-OPF problem [52]. Section 3.1 gives a first application of such methods to power systems stability analysis, based on publication [53].

- the US school focuses on SOS programming for Lyapunov methods [51] and it started contibuting to the particular problem of power systems in 2013 [6]. The second section of this chapter is an intent to test the computational limits of such methods on a more sophisticated system model, that was first published in [125].

Contents

3.1 A moment-SOS based approach

In this section, we show that the problem of TSA in power systems can be addressed using the moment-SOS hierarchy that we presented in Chapter 2. TSA considers the behavior of a power system following a major disturbance. Following considerations from Chapter 1, we focus on direct methods, which consist in computing stability oracle functions (SOF) parameterizing approximations of well-chosen regions of attraction (ROA). We recall that according to Remark 1.2, *inner* approximations are preferred to outer ones, as they prevent false negatives; however, we allow ourselves to also consider outer approximations, if it helps the simplicity of exposition. In this section, we focus on uncontrolled dynamics as a first step towards certified estimations of the ROA around a given operating point.

From a theoretical viewpoint, the idea is to use the moment-SOS hierarchy to compute approximated solutions of the finite time ROA problem presented in Example 2.8 (the interest of finite time will be exposed below). This section is organized as follows. Section 3.1.1 introduces the notion of occupation measure, which will be instrumental to deriving the finite time ROA problem in a way that will allow for moment-SOS hierarchy implementation in Section 3.1.2. Section 3.1.3 describes numerical experiments conducted to show the practical relevance of the proposed method and gives future research directions regarding computational tractability.

3.1.1 Occupation measures

It turns out that there exist numerical tools inspired from Chapter 2 on the GMP, which give some methods to compute approximations of constrained regions of attraction. While these numerical tools were initially designed only for polynomial differential systems, the following sections will introduce their conceptual basis, and how they can be applied to some nonpolynomial differential systems, without resorting to polynomial approximations of the vector fields.

As shown in [42, 58], the moment-SOS hierarchy makes it possible to numerically approximate $\mathbf{A}_T^{\mathbf{X}}(\mathbf{M})$ for well-chosen T and $\mathbf{M}$. We recall here how the method works. The key idea of the authors is to elegantly combine the set approximating property of the volume problem (see Example 2.6 and the corresponding reference [44]) with the notion of occupation-measure-based moment-SOS hierarchy introduced in [72]. More precisely, their method consists of grouping all the necessary time-domain simulations into the resolution of a single, linear partial differential equation (PDE). It can be compared to a Monte-Carlo algorithm, but offers the advantage of giving strong inclusion guarantees on the approximation that the computation yields.

Here we introduce the key element to design the ROA problem, following the contributions in [72]. Suppose that instead of having one deterministic initial condition $\mathbf{x}_0$ for the system (1.4), one is given a random variable $\mathbf{x}_0$ with probability law $\mathbb{P}_0 \in \mathcal{P}(\mathbb{R}^n)$. This means that for a given Borel set $\mathbf{Y} \subset \mathbb{R}^n$, $\mathbb{P}(\mathbf{x}_0 \in \mathbf{Y}) = \mathbb{P}_0(\mathbf{Y})$. Instead of considering a sample for the law $\mathbb{P}_0$ and running time-domain simulations for its realizations, as a Monte-Carlo algorithm would do, we directly work on the probability $\mathbb{P}_0$. Indeed, if $\mathbf{x}_0$ is a random variable that follows the probability $\mathbb{P}_0$,

then for any $t \geq 0$, $\mathbf{x}(t|\mathbf{x}_0)$ is also a random variable, so that we call $\mathbb{P}_t \in \mathcal{P}(\mathbb{R}^n)$ its probability law[1]. It is actually pretty easy to deduce $\mathbb{P}_t$ from $\mathbb{P}_0$, through the notion of transfer operator:

Definition 3.1: Ruelle's transfer operator

Let $\mathbf{E}, \mathbf{F}$ be finite dimensional open or compact sets, $\varphi : \mathbf{E} \to \mathbf{F}$ continuous. Ruelle's *transfer operator* [113, 89] $\varphi_\# : \mathcal{M}(\mathbf{E})_+ \to \mathcal{M}(\mathbf{F})_+$ maps a measure μ onto its pushforward through φ, denoted $\varphi_\#\mu$ and defined, for any Borel set $\mathbf{Y} \subset \mathbf{F}$, by

$$\varphi_\#\mu(\mathbf{Y}) := \mu(\varphi^{-1}(\mathbf{Y})) = \mu\{\mathbf{x} \in \mathbf{E} : \varphi(\mathbf{x}) \in \mathbf{Y}\}.$$

Proposition 3.2: Probability transfer

For all $t \geq 0$, $\mathbb{P}_t$ is deduced from $\mathbb{P}_0$ through the transfer operator $\mathbf{x}(t|\cdot)_\#$ as follows: $\mathbb{P}_t = \mathbf{x}(t|\cdot)_\#\mathbb{P}_0$, *i.e.* for any Borel set $\mathbf{Y} \subset \mathbb{R}^n$,

$$\mathbb{P}_t(\mathbf{Y}) = \mathbb{P}_0\left(\mathbf{x}(t|\cdot)^{-1}(\mathbf{Y})\right) = \mathbb{P}_0\{\mathbf{x}_0 \in \mathbb{R}^n : \mathbf{x}(t|\mathbf{x}_0) \in \mathbf{Y}\}.$$

Remark 3.1 (Transfer and composition)

From a duality perspective, the adjoint to the Ruelle transfer operator $\mu \mapsto \varphi_\#\mu$ is the Koopman composition operator $v \mapsto v \circ \varphi$. Indeed, besides the above definition, for $\varphi : \mathbb{R}^m \to \mathbb{R}^n$, $v \in \mathcal{C}_c(\mathbb{R}^n)$ and $\mu \in \mathcal{M}(\mathbb{R}^m)_+$ one has

$$\int v \, d(\varphi_\#\mu) = \int (v \circ \varphi) \, d\mu.$$

The Ruelle operator is also called Perron-Frobenius operator, after the Perron-Frobenius theorem whose infinite dimensional extension (named Krein-Rutman theorem) allows to prove that its eigenvectors are nonnegative measures. However, the Perron-Frobenius theorem has many other applications.

Conversely, knowing $\mathbb{P}_0$ and $\mathbb{P}_t$ gives an information on $\mathbf{x}(t|\cdot)$, without having to perform any simulation. Moreover, the data $\mathbb{P}_t$ for all $t \geq 0$ defines a function

$$\psi : \left\{ \begin{array}{ccc} \mathbb{R}_+ & \longrightarrow & \mathcal{P}(\mathbb{R}^n) \\ t & \longmapsto & \psi(t) := \mathbb{P}_t. \end{array} \right.$$

Then, a theorem due to Joseph Liouville ensures that ψ is subject to the *continuity equation* [129, Theorem 5.34]:

[1]Such time-dependent measure is called a Young measure, after Laurence Chisholm Young, who invented them to account for randomized control laws [143].

Theorem 3.3: Liouville

If $\mathbf{f}$ is Lipschitz, then ψ is the unique solution with initial condition $\psi(0) = \mathbb{P}_0$ of the following infinite dimensional, linear ODE:

$$\dot{\psi} = -\mathrm{div}(\psi\ \mathbf{f}), \tag{3.1}$$

where the derivative is defined w.r.t. the weak-$*$ topology on $\mathcal{P}(\mathbb{R}^n)$.
In particular, if there exists a density function $\rho : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}$ such that $\mathbb{P}_t = \rho(t, \cdot)\ \lambda$, then an integration by part yields the classical PDE form

$$\partial_t \rho + \mathrm{div}(\rho\ \mathbf{f}) = 0. \tag{3.1'}$$

Remark 3.2 (Divergence of a vector measure)
In Theorem 3.3, we introduce the term $-\mathrm{div}(\psi\ \mathbf{f})$*, which is defined as follows:*

$$-\mathrm{div}(\psi\ \mathbf{f})(t) := \left[v \in \mathcal{C}^1_c(\mathbb{R}^n) \mapsto \int \mathbf{f} \cdot \mathbf{grad}\ v\ d\psi(t) \right], \tag{3.2}$$

where $\mathcal{C}^1_c(\mathbf{\Omega})$ *is the subset of* $\mathcal{C}^1(\mathbf{\Omega})$ *of functions that vanish outside a compact set.*

Notice that ψ *is an* ***univariate*** *function (of* t*), with value in a* ***functional space*** *(in* $\mathbf{x}$*), so that equation* (3.1) *is indeed a linear ODE* $\dot{\psi} = \mathrm{A}\ \psi$ *while* (3.1′) *is a linear PDE in the* ***multivariate*** *unknown function* ρ *(of* t *and* $\mathbf{x}$*).*

Thus, Problem 1 can be rewritten in terms of Young measures, if one restricts to a finite time horizon $T \in (0, +\infty)$:

Problem 9: Stable initial measure

Find the maximal support for $\mathbb{P}_0$ such that (3.1) holds and:

- For all $t \in \mathbf{I} := [0, T]$, $\psi(t)$ is supported in $\mathbf{X}$,
- $\psi(T)$ is supported in $\mathbf{M}$

so that $\mathrm{spt}\,\mathbb{P}_0 \overset{\mathrm{def}}{=} \mathbf{A}^{\mathbf{X}}_T(\mathbf{M})$.

Remark 3.3 (Finite VS infinite time horizon)
Historically, stability analysis was first considered in astrophysics, with scientists concerned about the long term stability of the planet orbits in the solar system. For this reason, the most common notion of ROA is the unconstrained, infinite time ROA $\mathbf{A}_\infty(\mathbf{M})$*.*

However, in the context of power system TSA, all events occur within a finite, narrow time horizon, due to the speed of electro-mechanical dynamics. For this reason, it is only natural that we consider finite time ROAs here. Moreover, we will see in this chapter that finite time horizons are more convenient for computations that infinite time horizons.

Nevertheless, infinite time horizons can also be considered, as we will show in Sections 3.2 and 6.1.

Now, instead of considering only probability measures $\mathbb{P}_0 \in \mathcal{P}(\mathbb{R}^n)$ that represent a random initial condition $\mathbf{x}_0$, we propose to consider a *superposition* of initial conditions, which is represented by a measure $\mu \in \mathcal{M}(\mathbb{R}^n)_+$ s.t. $\mu(\mathbb{R}^n) < \infty$. One can easily switch from the probabilistic viewpoint to this statistical physics viewpoint by the following operations: $\mu = C\ \mathbb{P}_0$ for a given $C > 0$, or $\mathbb{P}_0 = \frac{1}{\mu(\mathbb{R}^n)}\ \mu$. By linearity of the continuity equation, Theorem 3.3 still holds for $\psi := t \mapsto \mathbf{x}(t|\cdot)_{\#}\mu$ with initial condition $\psi(0) = \mu$ in equation (3.1). It can even be extended to the notion of occupation measure.

Definition 3.4: Occupation measure

For a given initial measure $\mu \in \mathcal{M}(\mathbb{R}^n)_+$ and time horizon $T > 0$, we define the *occupation measure* $\nu_\mu \in \mathcal{M}(\mathbf{I} \times \mathbb{R}^n)_+$, for $a < b \in \mathbf{I}$ and $\mathbf{Y} \subset \mathbb{R}^n$:

$$\nu_\mu([a,b] \times \mathbf{Y}) := \int_a^b \left(\int_{\mathbb{R}^n} \mathbb{1}_{\mathbf{Y}}(\mathbf{x}(t|\mathbf{x}_0))\ d\mu(\mathbf{x}_0) \right) \mathrm{d}t = \int_a^b \psi(t)(\mathbf{Y})\ \mathrm{d}t.$$

In particular, if $\mu = \mathbb{P}_0 \in \mathcal{P}(\mathbb{R}^n)$, then $\nu_\mu([a,b] \times \mathbf{Y}) = \int_a^b \mathbb{P}_t(\mathbf{Y})\ \mathrm{d}t$ is the average time that the random trajectory $\mathbf{x}(t|\mathbf{x}_0)$ spends in $\mathbf{Y}$ between times a and b.

Remark 3.4 (Alternative definition of occupation measures)

Using the Riesz-Markov theorem, one can define the occupation measure ν through its action on $\mathcal{C}^1$ functions instead of its action on Borel sets: for $v \in \mathcal{C}^1_c(\mathbf{I} \times \mathbb{R}^n)$,

$$\int v\ d\nu_\mu := \int_0^T \left(\int_{\mathbb{R}^n} v(t, \mathbf{x}(t|\mathbf{x}_0))\ d\mu(\mathbf{x}_0) \right) \mathrm{d}t = \int_0^T \langle v(t,\cdot), \psi(t) \rangle\ \mathrm{d}t.$$

Remark 3.5 (Relevance of occupation measures)

While Young measures such as $\psi(t)$ are measure-valued functions of time, occupation measures are actually standard measures, and as such they can be represented on compact sets by their moments, allowing for a GMP rephrazing and thus the implementation of the Moment-SOS hierarchy, as we will show in the next section.

Liouville's theorem also holds for occupation measures, under a slightly different form.

Theorem 3.5: Transport of occupation measures

Let $\mu \in \mathcal{M}(\mathbb{R}^n)_+$, $T > 0$, $\nu_\mu \in \mathcal{M}(\mathbf{I} \times \mathbb{R}^n)$ as in Definition 3.4.
Let $\xi_\mu := \psi(T) = \mathbf{x}(T|\cdot)_{\#}\mu$ (if $\mu = \mathbb{P}_0 \in \mathcal{P}(\mathbb{R}^n)$ then $\xi_\mu = \mathbb{P}_T$).
If $\mathbf{f}$ is $\mathcal{C}^1$, then, (ν_μ, ξ_μ) is the unique solution of the following linear PDE with unknown $(\nu, \xi) \in \mathcal{M}(\mathbf{I} \times \mathbb{R}^n) \times \mathcal{M}(\mathbb{R}^n)$:

$$\partial_t \nu + \mathrm{div}(\nu\ \mathbf{f}) + \delta_T\ \xi = \delta_0\ \mu, \tag{3.3}$$

with parameters $\mathbf{f}, \mu, \delta_0, \delta_T$, where $\delta_t = [\mathbf{Y} \mapsto \mathbb{1}_{\mathbf{Y}}(t)] \in \mathcal{P}(\mathbb{R})$ is the Dirac measure in $t \in \mathbb{R}$.

Proof :

- (ν_μ, ξ_μ) is solution to (3.3):

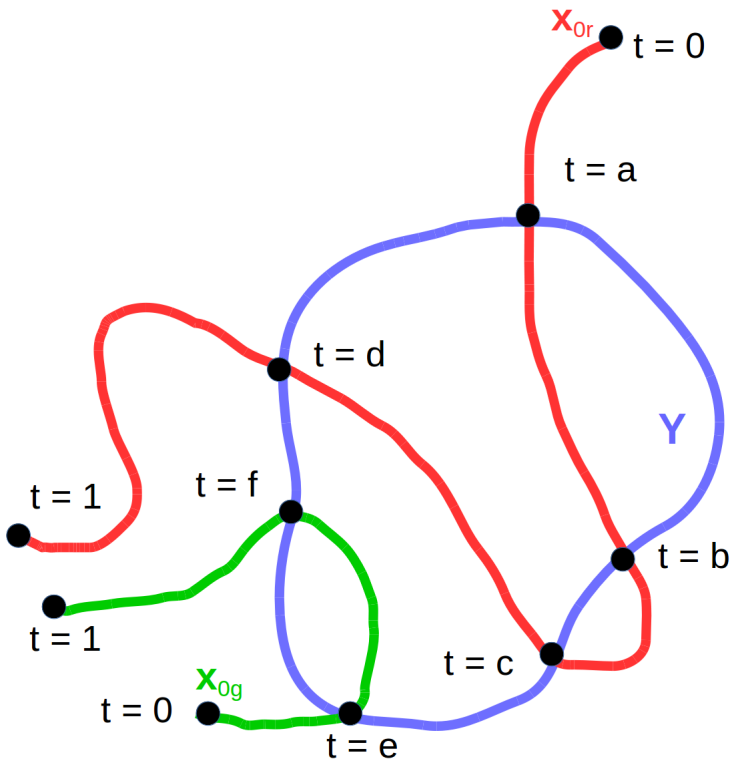


Figure 3.1 – Illustration of occupation measures.

Here the red trajectory is initialized at x_{0r} and the green one at x_{0g}, so that one can assign probabilities to both trajectories:

$$p_r := \mathbb{P}(\text{red traj.}) = \mathbb{P}(\mathbf{x}_0 = \mathsf{x}_{0r}) = \mu(\{\mathsf{x}_{0r}\}),$$
$$p_g := \mathbb{P}(\text{green traj.}) = \mathbb{P}(\mathbf{x}_0 = \mathsf{x}_{0g}) = \mu(\{\mathsf{x}_{0g}\}).$$

Then, the time spent by the red trajectory in Y is $T_r = \mathsf{d} - \mathsf{c} + \mathsf{b} - \mathsf{a}$, and the time spent by the green trajectory in Y is $T_g = \mathsf{f} - \mathsf{e}$, so that

$$\nu_\mu([0,1] \times \mathsf{Y}) \overset{\text{def}}{=} \mu(\{\mathsf{x}_{0r}\})(\mathsf{d} - \mathsf{c} + \mathsf{b} - \mathsf{a}) + \mu(\{\mathsf{x}_{0g}\})(\mathsf{f} - \mathsf{e}) = p_r T_r + p_g T_g$$

is the average of the times spent by the red and green trajectories in Y.

Let $v \in \mathcal{C}^1_c(\mathbf{I} \times \mathbb{R}^n)$, $\phi_\mu := \partial_t \nu_\mu + \operatorname{div}(\nu_\mu\ \mathbf{f}) \in \mathcal{C}^1(\mathbf{I} \times \mathbb{R}^n)'$. Then by definition one has

$$\begin{aligned}
\langle v, \phi_\mu \rangle &= -\int (\partial_t v(t,\mathbf{x}) + \mathbf{f}(\mathbf{x}) \cdot \mathbf{grad}\ v(t,\mathbf{x}))\, d\nu_\mu(t,\mathbf{x}) \\
&= -\int_0^T \int_{\mathbb{R}^n} (\partial_t v(t,\mathbf{x}(t|\mathbf{x}_0)) + \mathbf{f}(\mathbf{x}(t|\mathbf{x}_0)) \cdot \mathbf{grad}\ v(t,\mathbf{x}(t|\mathbf{x}_0)))\, d\mu(\mathbf{x}_0)\ \mathrm{d}t \\
&= -\int_0^T \int_{\mathbb{R}^n} \frac{\mathrm{d}}{\mathrm{d}t}\left[v(t,\mathbf{x}(t|\mathbf{x}_0))\right]\ d\mu(\mathbf{x}_0)\ \mathrm{d}t \\
&\overset{(*)}{=} -\int_{\mathbb{R}^n} \int_0^T \frac{\mathrm{d}}{\mathrm{d}t}\left[v(t,\mathbf{x}(t|\mathbf{x}_0))\right]\ \mathrm{d}t\ d\mu(\mathbf{x}_0)
\end{aligned}$$

$$
\begin{aligned}
&= -\int \left[v(t,\mathbf{x}(t|\mathbf{x}_0))\right]_0^T \, d\mu(\mathbf{x}_0) \\
&= \int v(0,\mathbf{x}_0)\, d\mu(\mathbf{x}_0) - \int v(T,\mathbf{x}(T|\mathbf{x}_0))\, d\mu(\mathbf{x}_0) \\
&= \langle v, \delta_0\ \mu - \delta_T\ \xi_\mu\rangle,
\end{aligned}
$$

where $(*)$ is obtained using Fubini's theorem [36, pages 243–249].

By our definition of measures, this proves that $\phi_\mu = \delta_0\,\mu - \delta_T\,\xi_\mu$, which is exactly equation (3.3).

- Such solution is unique:

Let (ν,ξ) be a solution to (3.3) with $\mu = 0$. By linearity of equation (3.3), showing $(\nu,\xi) = (0,0)$ will prove uniqueness for the generic $\mu \in \mathcal{M}(\mathbb{R}^n)_+$. Let $v \in \mathcal{C}^1_c(\mathbf{I}\times\mathbb{R}^n)$ and define for $(t,\mathbf{x}_0) \in \mathbf{I}\times\mathbb{R}^n$

$$
\varphi(t,\mathbf{x}_0) := -\int_t^T v(s,\mathbf{x}(s-t|\mathbf{x}_0))\,\mathrm{d}s.
$$

Then, since $\mathbf{f}$ is $\mathcal{C}^1$, for all $t \in \mathbf{I}$ one has $\mathbf{x}(t|\cdot) \in \mathcal{C}^1(\mathbb{R}^n)$, and thus $\varphi \in \mathcal{C}^1_c(\mathbf{I}\times\mathbb{R}^n)$ and for any $(t,\mathbf{x}_0) \in \mathbf{I}\times\mathbb{R}^n$ one has

$$
\begin{aligned}
\partial_t\varphi(t,\mathbf{x}(t|\mathbf{x}_0)) + \mathbf{f}(\mathbf{x}(t|\mathbf{x}_0))\cdot\mathbf{grad}\,\varphi(t,\mathbf{x}(t|\mathbf{x}_0)) &= \partial_t\left(\varphi(t,\mathbf{x}(t|\mathbf{x}_0))\right) \\
&= \partial_t\left(-\int_t^T v(s,\mathbf{x}(s-t|\mathbf{x}(t|\mathbf{x}_0)))\,\mathrm{d}s\right) \\
&= \partial_t\left(-\int_t^T v(s,\mathbf{x}(s|\mathbf{x}_0))\,\mathrm{d}s\right) \\
&= v(t,\mathbf{x}(t|\mathbf{x}_0))
\end{aligned}
$$

so that taking $\mathbf{x}_0 = \mathbf{x}(-t|\mathbf{x}_1)$ for all $\mathbf{x}_1 \in \mathbb{R}^n$ one has

$$
\partial_t\varphi(t,\mathbf{x}_1) + \mathbf{f}(\mathbf{x}_1)\cdot\mathbf{grad}\,\varphi(t,\mathbf{x}_1) = v(t,\mathbf{x}_1).
$$

Moreover, it is clear that $\varphi(T,\cdot) \equiv 0$ so that one also has

$$
\begin{aligned}
0 &= \int \varphi(T,\cdot)\, d\xi \\
&\overset{(3.3)}{=} -\langle\varphi, \partial_t\nu + \mathrm{div}(\nu\,\mathbf{f})\rangle \\
&= \int \partial_t\varphi + \mathbf{f}\cdot\mathbf{grad}\,\varphi\, d\nu \\
&= \int v\, d\nu.
\end{aligned}
$$

This last equality holding for all $v \in \mathcal{C}^1_c(\mathbf{I}\times\mathbb{R}^n)$, this proves that $\nu = 0$, after which $\xi = 0$ is deduced from (3.3). ♢

Remark 3.6 (Regularity of f)

For the use of Theorem 3.5, so far we only ask $\mathbf{f}$ *to be* $\mathcal{C}^1$ *so that the flow* $\mathbf{x}(t|\cdot)$ *of* $\mathbf{f}$ *exists at all* $t \in \mathbf{I}$*, is unique, and is a* $\mathcal{C}^1$*- diffeomorphism.*

3.1.2 Outer ROA approximation for polynomial systems

As stated in the introduction, this thesis aims at applying direct methods to assess transient stability of a power system. Such method resorts to determining as accurately as possible a transient stability region, by computing some stability oracle function. In this section, we put the focus on solving Problem 9, which is associated to a **finite** time horizon. In contrast, Section 3.2 deals with **infinite** time horizons.

We now give the numerical scheme to approximate the finite time ROA of a polynomial differential system. Such scheme resorts to the aforementioned notion of occupation measure, combined with the set approximation property.

The set approximation property

We consider again the volume problem from [44], which we described in Chapter 2:

$$\begin{aligned} \mathrm{p}^\star_{\mathbf{K}} = \sup_{\mu} \int 1 \, d\mu \qquad & (2.8a) \\ \text{s.t. } \mu \in \mathcal{M}(\mathbf{K})_+ & \\ \lambda - \mu \in \mathcal{M}(\mathbf{X})_+ & \end{aligned} \qquad \begin{aligned} \mathrm{d}^\star_{\mathbf{K}} := \inf_{w} \int w \, d\lambda \qquad & (2.8b) \\ \text{s.t. } w - 1 \in \mathcal{C}(\mathbf{K})_+ & \\ w \in \mathcal{C}(\mathbf{X})_+, & \end{aligned}$$

with compact sets $\mathbf{K} \subset \mathbf{X} \subset \mathbb{R}^n$. Here one can see that the inequality constraints of the dual problem (2.8b) can be rephrased into

$$w \geq \mathbb{1}_{\mathbf{K}} \qquad \text{on } \mathbf{X}. \tag{3.5}$$

Moreover, the convergence properties proved in Chapter 2 yield existence of a minimizing sequence $(w_d)_{d\in\mathbb{N}}$ such that

$$\int (w_d - \mathbb{1}_{\mathbf{K}}) \, d\lambda \underset{d\to\infty}{\longrightarrow} 0. \tag{3.6}$$

Then, defining $\hat{\mathbf{K}}_d := \{\mathbf{x} \in \mathbf{X} : w_d(\mathbf{x}) \geq 1\}$, (3.5) yields that $\mathbf{K} \subset \hat{\mathbf{K}}_d$, and (3.6) ensures a vanishing volume approximation error:

$$\mathrm{vol}\left(\hat{\mathbf{K}}_d \setminus \mathbf{K}\right) \leq \int (w_d - \mathbb{1}_{\mathbf{K}}) \, d\lambda \underset{d\to\infty}{\longrightarrow} 0.$$

In other words, sequentially solving the SOS hierarchy corresponding to the dual volume problem (2.8b) yields a *converging* outer approximation $(\hat{\mathbf{K}}_d)_d$ of $\mathbf{K}$. In the context of volume approximation, where $\mathbf{K}$ is already known, such a byproduct is useless. However, it paves the way for general set approximation, and one can modify problem (2.8) to approximate unknown sets, including some specific regions of attraction.

From the volume problem to the ROA problem

The ROA problem is then obtained by combining the notion of occupation measure and Liouville equation (3.3) with the set approximation property, through adding the Liouville equation to the volume problem (2.8a):

Let $\mathbf{M} = \mathbf{K} \subset \mathbf{X} \subset \mathbb{R}^n$ be a closed target set, $\mathbf{X}$ being compact, and $T < \infty$ the time horizon, and define $\mathbf{I} := [0, T]$. Then, consider the duality pair of problems

Problem 10: ROA approximating GMP

$$\begin{aligned}
\mathrm{p}^\star_{\mathrm{ROA}} := \sup_{\mu,\nu,\xi} \int 1\, d\mu \qquad & (3.7a) \\
\text{s.t. } & \mu \in \mathcal{M}(\mathbf{X})_+ \\
& \nu \in \mathcal{M}(\mathbf{I}\times\mathbf{X})_+ \\
& \xi \in \mathcal{M}(\mathbf{K})_+ \\
& \lambda - \mu \in \mathcal{M}(\mathbf{X})_+ \\
& \partial_t \nu + \mathrm{div}(\nu\, \mathbf{f}) = \delta_0\, \mu - \delta_T\, \xi,
\end{aligned}
\qquad
\begin{aligned}
\mathrm{d}^\star_{\mathrm{ROA}} := \inf_{v,w} \int w\, d\lambda \qquad & (3.7b) \\
\text{s.t. } & w - v(0,\cdot) - 1 \in \mathcal{C}(\mathbf{X})_+ \\
& -\partial_t v - \mathbf{f}\cdot\mathbf{grad}\, v \in \mathcal{C}(\mathbf{I}\times\mathbf{X})_+ \\
& v(T,\cdot) \in \mathcal{C}(\mathbf{K})_+ \\
& w \in \mathcal{C}(\mathbf{X})_+ \\
& v \in \mathcal{C}^1(\mathbf{I}\times\mathbf{X}).
\end{aligned}$$

Here, taking $\mathbf{f} = \mathbf{0}$ reduces (3.7a) to (2.8a). Then, imposing $v = 0$ in (3.7b) does not change the optimal value, so that it is equivalent to (2.8b). For this reason, (3.7) is a generalization of (2.8). Moreover, for a feasible pair (v_d, w_d) for (3.7b), consider

$$\hat{\mathbf{A}}_d := \{\mathbf{x}_0 \in \mathbf{X} : v_d(0, \mathbf{x}_0) \geq 0\}.$$

Then, the following finite time ROA approximation theorem holds [42]:

Theorem 3.6: Outer finite time constrained ROA approximation

One has

$$\mathbf{A}^{\mathbf{X}}_T(\mathbf{K}) \subset \hat{\mathbf{A}}_d$$

and the SOS hierarchy provides a *converging* sequence $(\hat{\mathbf{A}}_d)_{d\in\mathbb{N}}$ of outer approximations of $\mathbf{A}^{\mathbf{X}}_T(\mathbf{K})$, in the sense that

$$\mathrm{vol}\left(\hat{\mathbf{A}}_d \setminus \mathbf{A}^{\mathbf{X}}_T(\mathbf{K})\right) \underset{d\to\infty}{\longrightarrow} 0.$$

Proof : Let $\mathbf{x}_0 \in \mathbf{A}^{\mathbf{X}}_T(\mathbf{K})$. By definition, for all $t \in \mathbf{I}$, $\mathbf{x}(t|\mathbf{x}_0) \in \mathbf{X}$, and $\mathbf{x}(T|\mathbf{x}_0) \in \mathbf{K}$. Thus, for any feasible v_d for (3.7b),

$$\frac{\mathrm{d}}{\mathrm{d}t} v_d(t, \mathbf{x}(t|\mathbf{x}_0)) = \partial_t v(t, \mathbf{x}(t|\mathbf{x}_0)) + \mathbf{f}(\mathbf{x}(t|\mathbf{x}_0)) \cdot \mathbf{grad}\, v(t, \mathbf{x}(t|\mathbf{x}_0)) \leq 0$$

and $v_d(T, \mathbf{x}(T|\mathbf{x}_0)) \geq 0$, *i.e.* $v_d(\mathbf{x}(\cdot|\mathbf{x}_0))$ decreases along the trajectories and is nonnegative at final time T, so that $v_d(0, \mathbf{x}_0) \geq 0$, yielding $\mathbf{x}_0 \in \hat{\mathbf{A}}_d$. We have just proved that $\hat{\mathbf{A}}_d$ is an outer approximation of $\mathbf{A}^{\mathbf{X}}_T(\mathbf{K})$. It remains to show that the approximation error vanishes when we go through the moment-SOS hierarchy of Chapter 2.

First, the inequality constraint linking w_d and $v_d(0,\cdot)$ ensures that $w_d \geq \mathbb{1}_{\hat{\mathbf{A}}_d}$, so that

$$\mathrm{vol}\left(\hat{\mathbf{A}}_d \setminus \mathbf{A}^{\mathbf{X}}_T(\mathbf{K})\right) = \int \left(\mathbb{1}_{\hat{\mathbf{A}}_d} - \mathbb{1}_{\mathbf{A}^{\mathbf{X}}_T(\mathbf{K})}\right) d\lambda \leq \int w_d\, d\lambda - \mathrm{vol}\, \mathbf{A}^{\mathbf{X}}_T(\mathbf{K}).$$

Second, the choice of compact sets for $\mathbf{K}, \mathbf{X}, \mathbf{I}$ ensures that Assumption 2.7 holds, up to Remark 2.5, allowing for the use of Theorem 2.14: the SOS hierarchy gives access to minimizing sequences for problem (3.7b), *i.e.* $\forall d \in \mathbb{N}$, $\exists v_d \in \mathbb{R}_{2d}[t, \mathbf{x}], w_d \in \mathbb{R}_{2d}[\mathbf{x}]$ feasible for (3.7b) and such that

$$\int w_d \, d\lambda \underset{d\to\infty}{\longrightarrow} \mathrm{d}^\star_{\mathrm{ROA}}.$$

Then, we go back to Example 2.8, to prove strong duality: $\mathrm{d}^\star_{\mathrm{ROA}} = \mathrm{p}^\star_{\mathrm{ROA}}$. Let us show that Theorem 2.6 holds. The existence of feasible (μ, ν, ξ) is trivial: $(\mu, \nu, \xi) = (0, 0, 0)$ is feasible. It only remains to prove that all feasible occupation measures are uniformly bounded in mass:

- As for the volume (2.8), $\lambda - \mu \in \mathcal{M}(\mathbf{X})_+$ yields that $\mu(\mathbf{X}) \leq \mathrm{vol}\ \mathbf{X} =: C_\mu < \infty$.
- Testing (3.3) against $(t, \mathbf{x}) \mapsto 1$ yields $\xi(\mathbf{K}) = \mu(\mathbf{X}) \leq C_\mu =: C_\xi < \infty$.
- Testing (3.3) against $(t, \mathbf{x}) \mapsto t$ yields

  $$\nu(\mathbf{I} \times \mathbf{X}) = \langle t, \partial_t \nu + \mathrm{div}(\nu\ \mathbf{f}) \rangle = \langle t, \delta_0\ \mu - \delta_T\ \xi \rangle = T\xi(\mathbf{K}) \leq T\ C_\xi =: C_\nu < \infty.$$

Then, Theorem 2.6 ensures that strong duality holds, and thus

$$\int w_d \, d\lambda \underset{d\to\infty}{\longrightarrow} \mathrm{p}^\star_{\mathrm{ROA}}.$$

Eventually, we show that $\mathrm{p}^\star_{\mathrm{ROA}} = \mathrm{vol}\ \mathbf{A}^{\mathbf{X}}_T(\mathbf{K})$. Theorem 3.5 together with the definition of the pushforward ensure that $\xi \in \mathcal{M}(\mathbf{K})_+ \Rightarrow \mu \in \mathcal{M}(\mathbf{A}^{\mathbf{X}}_T(\mathbf{K}))_+$, so that $\mu = \mathbb{1}_{\mathbf{A}^{\mathbf{X}}_T(\mathbf{K})}\ \mu$. Together with constraint $\lambda - \mu \in \mathcal{M}(\mathbf{X})_+$, this yields that $\mathbb{1}_{\mathbf{A}^{\mathbf{X}}_T(\mathbf{K})}\ \lambda - \mu \in \mathcal{M}(\mathbf{X})_+$, and then $\mathrm{p}^\star_{\mathrm{ROA}} \leq \int \mathbb{1}_{\mathbf{A}^{\mathbf{X}}_T(\mathbf{K})}\, d\lambda = \mathrm{vol}\ \mathbf{A}^{\mathbf{X}}_T(\mathbf{K})$.

Moreover, $\mu = \mathbb{1}_{\mathbf{A}^{\mathbf{X}}_T(\mathbf{K})}\ \lambda \in \mathcal{M}(\mathbf{X})_+$ is such that $\xi = \mathbf{x}(T|\cdot)_\# \mu \in \mathcal{M}(\mathbf{K})_+$, so that Theorem 3.5 ensures that $\mathbb{1}_{\mathbf{A}^{\mathbf{X}}_T(\mathbf{K})}\ \lambda$ is a feasible value for μ, for which $\int 1\, d\mu = \mathrm{vol}\ \mathbf{A}^{\mathbf{X}}_T(\mathbf{K})$. Then, taking the supremum, one obtains $\mathrm{p}^\star_{\mathrm{ROA}} \geq \mathrm{vol}\ \mathbf{A}^{\mathbf{X}}_T(\mathbf{K})$, and by double inequality, $\mathrm{p}^\star_{\mathrm{ROA}} = \mathrm{vol}\ \mathbf{A}^{\mathbf{X}}_T(\mathbf{K})$, so that

$$\int w_d \, d\lambda \underset{d\to\infty}{\longrightarrow} \mathrm{vol}\ \mathbf{A}^{\mathbf{X}}_T(\mathbf{K}).$$

In conclusion, the approximation error indeed vanishes when d tends to infinity:

$$\mathrm{vol}\left(\hat{\mathbf{A}}_d \setminus \mathbf{A}^{\mathbf{X}}_T(\mathbf{K})\right) \leq \int w_d \, d\lambda - \mathrm{vol}\ \mathbf{A}^{\mathbf{X}}_T(\mathbf{K}) \underset{d\to\infty}{\longrightarrow} 0.$$

♢

Remark 3.7 (Contribution with respect to [42])

Though Theorem 3.6 was already proved in [42], the systematic reduction of the proof to checking boundedness of the feasible measures, using the original results from Chapter 2, is new.

Remark 3.8 (Extension of [42])

It is also possible to prove Theorem 3.6 using Theorem 2.17. In such a case, in addition to the aforementioned results, one also obtains a convergence proof for the pseudo-moment sequences of the moment hierarhy. This, combined with Christoffel-Darboux kernel methods from [87], paves the way for completely new approaches for moment-hierarchy-based set approximations.

Remark 3.9 (Relevance of the constrained finite time ROA)

The proof of Theorem 3.6 heavily relies on the finite time horizon T and state constraint sets $\mathbf{X}$ and $\mathbf{K}$. Indeed, this choice of parameters is required for $\mathbf{I}$, $\mathbf{K}$ and $\mathbf{X}$ to be compact, which is instrumental in the proof. In other words, state constraints and a finite time horizon are the price for the convergence of the ROA approximation.

However, such constraints are quite relevant for TSA: as pointed in Remark 3.3, a finite time horizon is suitable in this context. Moreover, operational TSA does not only expect that the system quickly converges to equilibrium, but also that the trajectory remains secure during all the transients; this involves additional state constraint security specifications (e.g. current, voltage and power boundedness), which can be encoded in $\mathbf{X}$.

Thus, even though the notion of constrained finite time ROA is less usual than the notion of free infinite time ROA, it is particularly relevant in the context of this thesis. As a result, it will be the first TSA approach to which we will apply structure decomposition to facilitate computations (see Section 6.2).

3.1.3 Finite time ROA estimation for a 3 machines model

Apart from rephrazing the proof for the ROA approximation, an important contribution of this section is the extension of the framework to non-polynomial equations of power systems, with numerical application to the three machine model. In terms of numerical application, we are interested in analysing the transient behaviour of system (1.1)–(1.3). More precisely, to numerically illustrate the TSA problem, we use the three-bus numerical example from Chiang et al. [21], which is composed of three synchronous machines connected in a meshed grid. The third bus is arbitrarily chosen as the reference angle (i.e., $\theta_3 = 0$), so that one only needs two phase angle variables, θ_1, θ_2, and two rotor speed variables, ω_1, ω_2, to describe the dynamics. [21] neglected conductances G_{kl} and attributed numerical values to M_k, D_k, P_k^{mec}, $|\mathbf{v}_k|$ and B_{kl}, obtaining:

$$\begin{aligned}
\dot{\theta}_k &= \omega_k, \qquad k = 1, 2,\\
\dot{\omega}_1 &= -\sin(\theta_1) - 0.5\sin(\theta_1 - \theta_2) - 0.4\ \omega_2,\\
\dot{\omega}_2 &= -0.5\sin(\theta_2) - 0.5\sin(\theta_2 - \theta_1) - 0.5\ \omega_2 + 0.05,
\end{aligned}$$

where the ω_k are expressed in rad/s and the θ_k are in rad. A stable equilibrium is given by $\overline{\boldsymbol{\theta}} = (0.02,\ 0.06)$.

This very basic example will serve as a numerical test case for our moment-SOS TSA method.

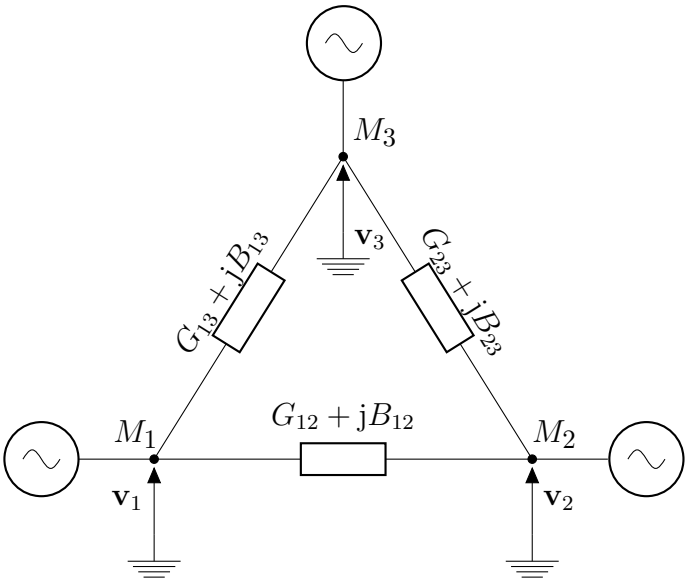


Figure 3.2 – The three machines cycle.

Tackling trigonometric dynamics

Following [6], the coordinates can be shifted so that $(\overline{\theta}_1, \overline{\theta}_2) = (0.00, 0.00)$ is a stable equilibrium. This dynamical system can in turn be formulated as a polynomial differential algebraic system, as suggested in [6]. To that end, we introduce auxiliary variables

$$s_k := \sin(\theta_k) \quad \text{and} \quad c_k := 1 - \cos(\theta_k), \quad k = 1, 2 \tag{3.8}$$

and define $\boldsymbol{\omega} := (\omega_1, \omega_2)$, $\mathbf{s} := (s_1, s_2)$, $\mathbf{c} := (c_1, c_2)$ as well as $\mathbf{x} := (\boldsymbol{\omega}, \mathbf{s}, \mathbf{c}) \in \mathbb{R}^6$. One then obtains the following differential algebraic equation system (DAEs)

$$\begin{aligned}
\dot{\omega}_1 &= 0.4996 s_2 - 0.4\omega_1 - 1.4994 s_1 - 0.02 c_2 + 0.02 s_1 s_2 \\
&\quad + 0.4996 s_1 c_2 - 0.4996 c_1 s_2 + 0.02 c_1 c_2, \\
\dot{\omega}_2 &= 0.4996 s_1 + 0.02 c_1 - 09986 s_2 + 0.05 c_2 - 0.5\omega_2 \\
&\quad - 0.02 s_1 s_2 - 0.4996 s_1 c_2 + 0.4996 c_1 s_2 - 0.02 c_1 c_2, \\
\dot{s}_k &= (1 - c_k)\omega_k \qquad\qquad k = 1, 2, \\
\dot{c}_k &= s_k \omega_k \qquad\qquad k = 1, 2, \\
0 &= s_k^2 + c_k^2 - 2c_k \qquad\qquad k = 1, 2,
\end{aligned}$$

We will show later on that one can actually avoid increasing the number of variables and immediately obtain a polynomial differential system of equations in *complex-valued quantities*.

A particularity of the occupation measure approach is that the state set $\mathbf{X}$ should have an interior point such that the computed volumes are non-zero. Hence, constraints $\mathbf{0} = \mathbf{h}(\mathbf{x}) := (s_k^2 + c_k^2 - 2c_k)_{k=1,2}$ in the dynamics derived from our change

of variable may be troublesome, since the manifold $\mathbf{M} := \{\mathbf{x} \in \mathbb{R}^6 : \mathbf{h}(\mathbf{x}) = 0\}$ is a hyper-surface with no interior point. A simple method to address this issue consists of ignoring the equality constraints when computing the ROA approximation $\hat{\mathbf{A}}_d$, and then consider $\hat{\mathbf{A}}_d \cap \mathbf{M}$ as the desired ROA estimation. Such an exact method does not work with any arbitrary equality constraints. However, in the case of constraints derived from a change of variable, this approach is valid due to the fact that the reformulated vector field $\mathbf{f}$ then satisfies $(\mathbf{grad}\ \mathbf{h})\cdot\mathbf{f} \equiv 0$. Thus, the dynamics are tangent to $\mathbf{M}$, yielding that any trajectory starting in $\mathbf{M}$ will remain in $\mathbf{M}$, which is exactly the constraint $\mathbf{h}(x(t|\mathbf{x}_0)) = 0,\ \forall\ t \in [0, T]$. For equality constraints that are not derived from changes of variables, this does not hold anymore; in such case, one could think of replacing the classical Lebesgue measure λ with a Hausdorff measure supported on $\mathbf{M}$, and computed using Stokes theorem if $\mathbf{M}$ is the boundary of a full-dimensional set [73].

To the best of our knowledge, this is the first time that algebraic equality constraints derived from a change of variable are addressed within the occupation measure approach. This facilitates the novel application of occupation measures theory to non-polynomial systems. The general non-polynomial dynamics framework that can be tackled with this method will be discussed in details in Section 3.2.3.

Case study

For our numerical experiments, we use MATLAB R2015b, YALMIP [78], SeDuMi 1.3 [120], and the "ROA" code of [42] to apply occupation measure theory to the three-bus example from [21] that is described above.

We define the finite time ROA parameters as follows:

- $\mathbf{X} := [-\pi, \pi]^2 \times [-1, 1]^2 \times [0, 2]^2$
  $= \{\mathbf{x} \in \mathbb{R}^6 : (\omega_{1,2}^2 \leq \pi^2) \wedge (s_{1,2}^2 \leq 1) \wedge (0 \leq c_{1,2} \leq 2)\}$
- $\mathbf{K} := \mathbf{B}_\varepsilon = \{\mathbf{x} \in \mathbb{R}^6 : |\mathbf{x}|^2 \leq \varepsilon^2\}$, $\varepsilon = 0.1$, $T = 8$ s.

Following the developments in Chapter 2, problem (3.7b) admits a SOS strenghtening that can be expressed as follows:

$$
\begin{aligned}
\mathrm{d}_{\mathrm{ROA}}^d := \inf_{w}\ & \mathbf{w}\cdot\mathbf{l} \\
\text{s.t. } & w(\mathbf{x}) - v(0,\mathbf{x}) - 1 \in \Sigma_d(\mathbf{g}(\mathbf{x})) \\
& -\partial_t v(t,\mathbf{x}) - \mathbf{f}(\mathbf{x})\cdot\mathbf{grad}\ v(t,\mathbf{x}) \in \Sigma_d(\widetilde{\mathbf{g}}(t,\mathbf{x})) \\
& v(T,\mathbf{x}) \in \Sigma_d(0.01 - |\mathbf{x}|^2) \\
& w(\mathbf{x}) \in \Sigma_d(\mathbf{g}(\mathbf{x})) \\
& v(t,\mathbf{x}) \in \mathbb{R}_{2d}[t,\mathbf{x}],
\end{aligned}
\tag{3.9}
$$

where:

- $\mathbf{l}$ is the vector of moments of the Lebesgue measure λ on $\mathbf{X}$ in a moment basis,
- $\mathbf{w}$ is the vector of coefficients of w in the corresponding monomial basis,
- $\mathbf{g}(\mathbf{x}) = (\pi^2 - \omega_{1,2}^2, 1 - s_{1,2}^2, (2 - c_{1,2})c_{1,2}, 30 - |\mathbf{x}|^2)$ describes the state set $\mathbf{X}$,

- $\mathbf{f}(\mathbf{x}) = (\dot{\boldsymbol{\omega}}, \dot{\mathbf{s}}, \dot{\mathbf{c}})$ describes the polynomial dynamics and
- $\widetilde{\mathbf{g}}(t, \mathbf{x}) = ((T-t)t, \pi^2 - \omega_{1,2}^2, 1 - s_{1,2}^2, (2 - c_{1,2})c_{1,2}, 100 - t^2 - |\mathbf{x}|^2)$.

Again, there is no duality gap between these truncated problems and the corresponding moment relaxations of (3.7a), at every order of the hierarchy, thanks to Theorem 2.16 (the proof is identical to the strong duality proof for Theorem 3.6). Though we already have proved the convergence of the outer approximation, the computational burden increases sharply as the order d increases.

We complement this explanation by briefly discussing the approach for computing inner approximations. The machinery for inner ROA approximations is very similar to the outer approximation approach discussed above. The key distinction is that the inner approximations consider an outer approximation to the complement of the ROA, $\mathbf{A}_T^{\mathbf{X}}(\mathbf{K})^{\mathrm{c}} := \mathbf{X} \setminus \mathbf{A}_T^{\mathbf{X}}(\mathbf{K})$, inducing additional technicalities. See [58] for further details.

We note that practical power system analyses require the ability to address significantly larger problems than the test case considered in this section. However, constructing certified approximations for the ROA leads to difficult computational challenges, see e.g. Section 3.2. Similar to the demonstrations of previous algorithms [6],[50], this section focuses on a small system as an initial step towards practical applications. Future work that exploits network sparsity and other problem structures will be crucial for scalability. Decomposition approaches may also prove to be valuable [61, 62, 63].

With final time $T = 8$ and radius $\varepsilon = 0.1$, we find the following polynomial, $v_5(0, x)$, at fifth-order relaxation ($d = 5$):

$$\begin{aligned} v_5(0,x) = \; & 1.8707 - 4.9538x_1 + 0.0017x_2 \\ & - 4.7856x_3 + 0.0018x_4 - 0.0037x_5 \\ & - 4.8546x_6 - 0.0131x_1^2 + 10.9412x_1x_2 \\ & - 0.0356x_1x_3 + 13.9529x_1x_4 + 0.0208x_1x_5 \\ & + 0.0142x_1x_6 + 16.4121x_2^2 + 0.0609x_2x_3 \\ & - 0.2755x_5^5x_6^5 + 0.0017x_5^4x_6^6 + 0.0170x_5^3x_6^7 \\ & - 0.0002x_5^2x_6^8 - 0.0021x_5x_6^9 - 0.0003x_6^{10}. \end{aligned}$$

The zero super-level set $\hat{\mathbf{A}}_5 := \{\mathbf{x} \in \mathbb{R}^6 : v_5(0, \mathbf{x}) \geq 0\}$ provides an outer approximation to the ROA. We illustrate the polynomial $v_5(0, \cdot)$ in Fig. 3.3 as a function of the original state variables (θ_1, θ_2). We consider $(\omega_1, \omega_2) = (0, 0)$ in order to visualize the ROA, but this is not a necessary restriction. We illustrate the outer approximation to the ROA in Fig. 3.4.

Likewise, with $T = 8$ and $\varepsilon = 0.1$, we find at the third-order relaxation ($d = 3$) the inner approximation to the ROA presented in Fig. 3.5 (again with $\boldsymbol{\omega} = (0, 0)$ used only for representation purposes).

We next show how one could use Hermitian SOS to obtain better numerical results. For optimal power flow problems, applying Hermitian SOS yields computational advantages while preserving convergence guarantees [54]. The idea is to exploit the structure that comes from alternating current physics in order to reduce the computational burden. We consider the transient dynamics of a system after

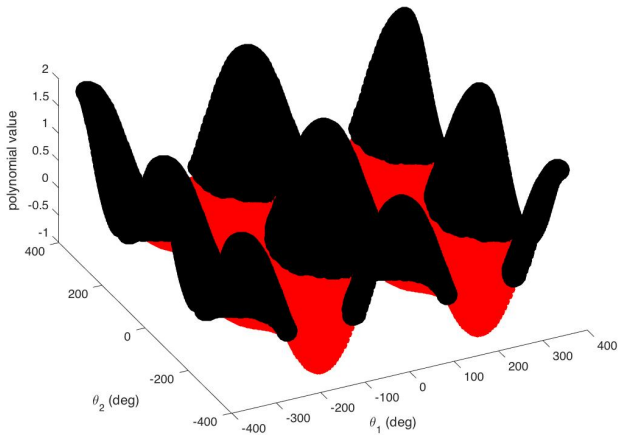


Figure 3.3 – Plot of the graph of $v_d(0, \cdot)$, $d = 5$.

The polynomial for the three-bus system whose zero super-level set, which is indicated by the back region, provides an outer approximation to the ROA. The projection shown is for $\boldsymbol{\omega} = (0, 0)$.

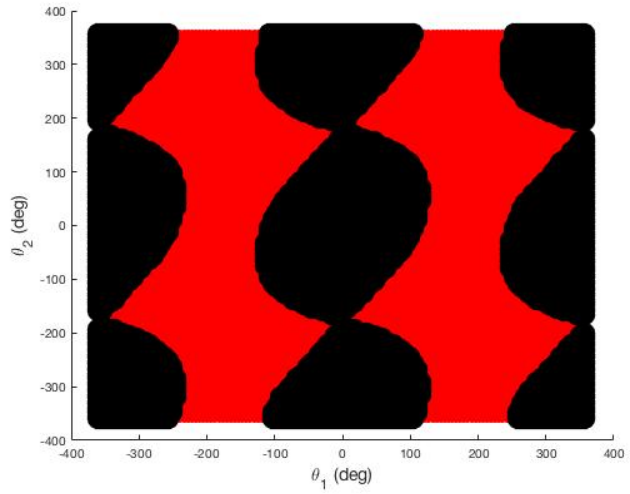


Figure 3.4 – Outer ROA approximation of degree $d = 5$.

An outer approximation of the ROA is indicated by the back region. The projection shown is for $(\omega_1, \omega_2) = (0, 0)$.

the fault has disappeared and we assume that there is no voltage instability. In that case, it is reasonable to assume that the magnitudes $|\mathbf{v}_k|$ of the complex voltages are fixed such that only the phase angles θ_k are variables. This allows us to define $\mathbf{v}_k := \exp(\mathrm{j}\,\theta_k)$ (up to proper rescaling), such that $\dot{\mathbf{v}}_k = \mathrm{j}\,\dot{\theta}_k \exp(\mathrm{j}\,\theta_k)$, where $\mathrm{j}^2 = -1$. The dynamics can thus immediately be written as a polynomial differential system of equations (with physical variables now expressed in perunits, s.t. the voltages have magnitude 1):

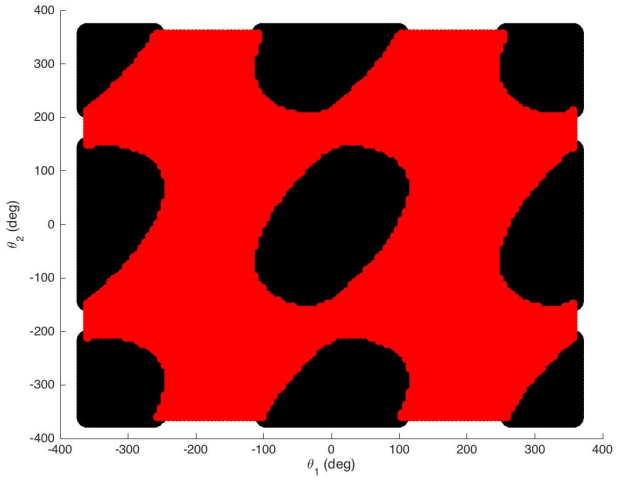


Figure 3.5 – Inner ROA approximation of degree $d = 3$.

An inner approximation of the ROA is indicated by the back region. The projection shown is for $(\omega_1, \omega_2) = (0, 0)$.

$$\begin{aligned} \dot{\mathbf{v}}_k &= \mathrm{j}\ \omega_k \mathbf{v}_k, \\ \dot{\omega}_k &= -\tfrac{D_k}{M_k}\omega_k + \tfrac{1}{M_k}\left(P_k - \tfrac{1}{2}\textstyle\sum_{l\neq k} -G_{kl}|v_k|^2 - \overline{Y}_{kl} v_k \overline{v}_l - Y_{kl} v_l \overline{v}_k\right), \\ 0 &= |\mathbf{v}_k|^2 - 1, \end{aligned} \tag{3.10}$$

where Y_{kl} denotes the mutual admittance of the line connecting buses k and l.

It is straightforward to adapt the theory of occupation measures to complex states by leveraging recent results in complex algebraic geometry [27]. A future work would consist of implementing a complex version of the hierarchy proposed in [42] in order to reduce the computational burden at a given relaxation order.

In conclusion, in the context of the transient stability analysis of power systems, this section demonstrates the potential for using the theory of occupation measures (along with convex optimization techniques) to compute inner and outer approximations to the region of attraction for a stable equilibrium point. To the best of our knowledge, this is the first time that occupation measure theory has been applied to analyze transient stability problems for electric power systems. The resulting approximations have the potential to provide analytically rigorous guarantees that can reduce the need for computationally expensive time-domain simulations. With computational tractability remaining an important challenge, the next chapters will investigate how to exploit sparsity for set approximation, in particular when using occupation measures.

3.2 An SOS, Lyapunov-based approach

In this section, we investigate an alternative SOS method for ROA approximation, which is based on Lyapunov theory rather than occupation measures. We recall here that, as our variable v in Section 3.1, Lyapunov functions are an instance of SOF (historically, they even are the archetype of SOF, since the first direct methods for stability analysis resorted to what we now call Lyapunov functions) as listed in Example 1.5. The Lyapunov approach was used in [6] for transient stability analysis of a power system. However, the generators were modeled only by their swing equations. As the voltage and frequency regulators have an important impact on stability analysis, in the present section this approach is extended by considering a generator with voltage dynamics and both voltage and frequency regulations. This also paves the way to taking into account high penetration of power electronics elements in the grid due to the integration of renewable energies and HVDC transmission lines, thus having an important impact on the transient stability of the system. SOS method gives an analytical solution for the construction of Lypunov functions in order to estimate the infinite time ROA for a locally asymptotically stable equilibrium point of the system.

The section is organized as follows: the problem is formulated and modelled in Section 3.2.1 for a Single Machine-Infinite Bus (SMIB) system. Section 3.2.2 sums up the elements of Lyapunov theory that are at stake. SOS formalism and algorithms are briefly recalled in Section 3.2.3, along with a necessary change of variables (i.e., reformulating the model of the system with trigonometric non linearities into a set of polynomial differential algebraic equations). The reformulating procedure is proved to be a Lie-Bäcklund transformation, which means that the transformed system has equivalent trajectories and stability properties [34]. Next, in Section 3.2.4 we relax Lyapunov's conditions for stability and model constraint equations to suitable SOS conditions using theorems from real algebraic geometry in order to formulate the problem as an optimization one. Hence, a Lyapunov function for the asymptotically stable equilibrium point is constructed using the *expanding interior algorithm* developed in [51, 6]. An estimate of the ROA is given by a level set of the Lyapunov function. We finally test the ROA estimation error by numerically computing the real one in all state directions via full nonlinear simulations. The codes are implemented in MATLAB using SOSTOOLS [106] which is a free, third-party MATLAB toolbox that models SOS problems.

3.2.1 The single machine - infinite bus system

We consider a synchronous machine connected to a power grid through two transmission lines in parallel (see Figure 3.6). The power grid is modelled as an infinite bus. The infinite bus imposes a nominal voltage of amplitude V_g and frequency ω_g at node B. Each transmission line is a series of two impedance matrices $\mathrm{Z}_t := R_t\mathrm{I}+X_t\mathrm{J}$ (or equivalently one synthetic impedance matrix $\mathrm{Z}_{\mathrm{eq}} = 2\mathrm{Z}_t$), where $R_t, X_t \in \mathbb{R}_+$ and

$$\mathrm{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad , \quad \mathrm{J} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

The synchronous machine has two rotating axes d (direct) and q (quadratic), which support the currents i_d and i_q (we use the notation $\mathbf{i}_s := (i_d, i_q)^\top$), generating a voltage $\mathbf{v}_s := (v_d, v_q)^\top$ at the machine's terminal.

The synchronous machine is characterized by an internal resistance R_i and equivalent reactance matrices $\mathrm{X}_i := \mathrm{J}\ \mathrm{diag}(x_d, x_q)$ and $\mathrm{X}'_i := \mathrm{J}\ \mathrm{diag}(x'_d, x_q)$, with $x_d, x_q, x'_d \in \mathbb{R}_+$. It receives a mechanical power P_m and rotates at frequency ω, generating a field voltage with magnitude E_f orthogonal to the rotor axis, and thus providing an e.m.f $\mathbf{e} \in \mathbb{R}^2$. Let θ be the phase between the machine internal voltage and the grid voltage and H the machine's mechanic inertia constant.

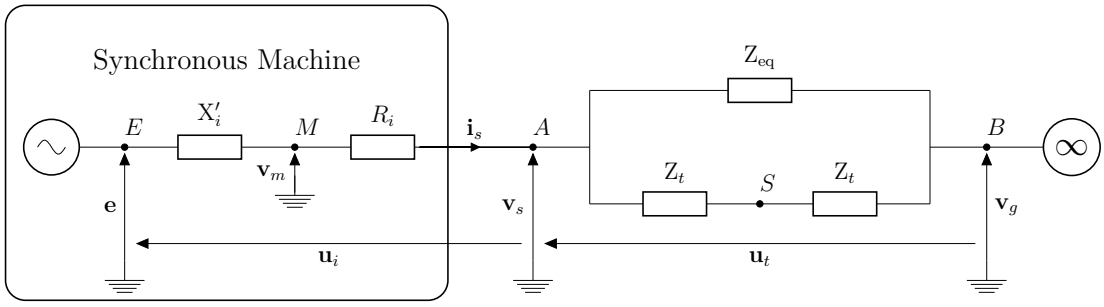


Figure 3.6 – Synchronous machine connected to an infinite bus (SMIB).

The equations describing the dynamics of this system are given in [114, p. 105]. In this section we give some details on how these equations can be derived as a 5^{th} order model. First, we take as a starting point the swing equation that models the synchronous machine, but we do not assume its voltage magnitude $|\mathbf{v}_s|$ to be constant anymore. Thus, we enrich the model with the dynamics of the e.m.f. $\mathbf{e}$, obtaining a 3^{rd} order model. Then, we proceed to modelling the outer control loops on the voltage magnitude $|\mathbf{v}_s|$ and frequency ω, that appear in the dynamics of the field voltage E_f and mechanical power input P_m respectively, leading to a 5^{th} order model.

Synchronous machine dynamics

As stated above, we write the equations of the system in a rotating frame whose direct axis is aligned with the rotor's axis, and whose quadratic axis is aligned with the field voltage. In such system, the synchronous machine's dynamics write:

$$\dot{\theta} = 100\pi(\omega - \omega_g) \tag{3.11a}$$

$$2H\ \dot{\omega} = P_m - (\mathbf{v}_m \cdot \mathbf{i}_s) \tag{3.11b}$$

$$T_0\ \dot{\mathbf{e}} = E_f \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \mathbf{e} - (\mathrm{X}_i - \mathrm{X}'_i)\ \mathbf{i}_s, \tag{3.11c}$$

where θ is expressed in rad, and all other variables are expressed in p.u.

Notice that equation (3.11c), which models how the field voltage E_f generates

the e.m.f $\mathbf{e}$, is given in vector form and can be decomposed as

$$\begin{aligned} T_0\ \dot{e}_d &= -e_d \\ T_0\ \dot{e}_q &= E_f - e_q - (x_d - x'_d)\ i_d \end{aligned} \tag{3.11c$'$}$$

so that the direct component e_d of $\mathbf{e}$ is discarded by our choice of convention (taking initial value $e_d(0) = 0$).

Then, equation (3.11b) can be rephrased using Ohm's law

$$\mathbf{v}_m = \mathbf{e} - \mathrm{X}'_i\ \mathbf{i}_s,$$

so that $\mathbf{v}_m \cdot \mathbf{i}_s = e_q\ i_q - (x'_d - x_q)\ i_d\ i_q = (e_q + (x_q - x'_d)\ i_d)\ i_q$, allowing us to rewrite

$$2H\ \dot{\omega} = P_m + ((x'_d - x_q)\ i_d - e_q)\ i_q. \tag{3.11b$'$}$$

Thus, at this point the synchronous machine's dynamics takes the general form of a polynomial control system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u})$$

with state variable $\mathbf{x} = (\theta, \omega, e_q)^\top$ and "control" input $\mathbf{u} = (E_f, P_m, i_d, i_q)^\top$. The values of i_d and i_q are determined through the transmission lines' equations (which take the general algebraic form $\mathbf{g}(\mathbf{x}, \mathbf{y}) = \mathbf{0}$, $\mathbf{g} \in \mathbb{R}[\mathbf{x}, \mathbf{y}]^m$ where $\mathbf{y} = (i_d, i_q)^\top$), while E_f and P_m are themselves subject to a broader control loop.

Synchronous machine outer control loop

The machine is governed by two regulators. First, an Automatic Voltage Regulator (AVR) is implemented as follows. E_f is set to be proportional to the voltage magnitude error $V_{\text{ref}} - V_t$, where $V_t := |\mathbf{v}_s| = \sqrt{v_d^2 + v_q^2}$ is the voltage magnitude at the machine terminal, and V_{ref} is the voltage reference:

$$E_f^\star := K_a(V_{\text{ref}} - V_t).$$

Then, we model the actuator response as a first order

$$T_a\ \dot{E}_f = E_f^\star - E_f,$$

obtaining the control equation

$$T_a\ \dot{E}_f = -E_f + K_a(V_{\text{ref}} - V_t). \tag{3.12}$$

When the field voltage E_f is subject to such control law, the voltage magnitude at the synchronous machine's terminal will automatically respond to any change in the reference V_{ref}.

Second, a turbine governor regulates the mechanical power injected into the synchronous machine as follows. P_m is set to equate to a reference P_{ref} plus a term that is proportional to the frequency error $\omega_{\text{ref}} - \omega$, where ω_{ref} is the reference frequency value:

$$P_m^\star = P_{\text{ref}} + K_g(\omega_{\text{ref}} - \omega).$$

Again, we model the actuator response as a first order

$$T_g \ \dot{P}_m = P_m^\star - P_m,$$

obtaining the control equation

$$T_g \ \dot{P}_m = P_{\text{ref}} - P_m + K_g(\omega_{\text{ref}} - \omega). \tag{3.13}$$

Now considering the system composed with equations (3.11a), (3.11b′), (3.11c′), (3.12), (3.13), one is seemingly faced to a new polynomial control system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u})$ with state $\mathbf{x} = (\theta, \omega, e_q, E_f, P_m)^\top$ and input $\mathbf{u} = (i_d, i_q, V_t, V_{\text{ref}}, P_{\text{ref}}, \omega_{\text{ref}})^\top$. However, i_d, i_q and V_t are not properly speaking "control inputs", but rather algebraic variables: as stated before, they are determined by the transmission line equations, that we will now derive. Thus, the actual control input is $\mathbf{u} = (V_{\text{ref}}, P_{\text{ref}}, \omega_{\text{ref}})^\top$, and it is supposed to be fixed in the rest of this section.

Transmission lines equations

The circuit represented in Figure 3.6 is equivalent to the much simpler one that we display in Figure 3.7.

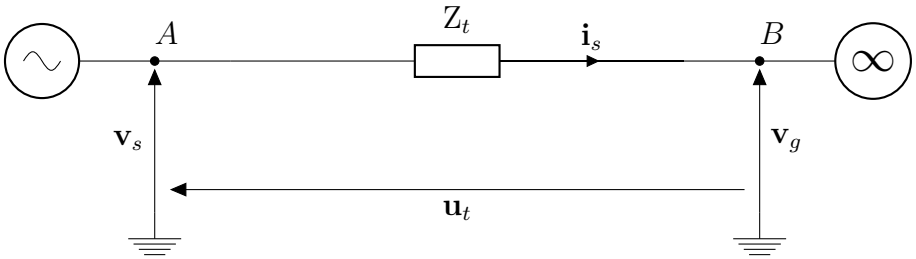


Figure 3.7 – The equivalent simplified SMIB model.

By construction of our state variable θ as the angle from the grid to the machine, by definition of the infinite bus, and since phases are defined up to an additive constant, one has

$$\mathbf{v}_g = V_g \begin{pmatrix} \cos(\theta_0 - \theta) \\ \sin(\theta_0 - \theta) \end{pmatrix}$$

So that choosing the phase reference as $\theta_0 = \frac{\pi}{2}$ yields

$$\mathbf{v}_g = V_g \begin{pmatrix} \sin\theta \\ \cos\theta \end{pmatrix}.$$

On another hand, Ohm's laws yield that

$$\mathbf{v}_s - \mathbf{v}_g = \mathbf{u}_t = \mathrm{Z}_t \ \mathbf{i}_s \text{ (Fig. 3.7)} \qquad \& \qquad \mathbf{e} - \mathbf{v}_s = \mathbf{u}_i = (R_i \ \mathrm{I} + \mathrm{X}_i') \ \mathbf{i}_s \text{ (Fig. 3.6)}$$

so that $\mathbf{v}_s = \mathbf{v}_g + \mathrm{Z}_t \ \mathbf{i}_s$ and $\mathbf{e} - \mathbf{v}_g = (\mathrm{Z}_t + R_i \ \mathrm{I} + \mathrm{X}_i') \ \mathbf{i}_s$, and the SMIB model is now turned into a semi-explicit polynomial differential algebraic system of equations (DAEs)

$$\begin{cases} \dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{y}) \\ \mathbf{g}(\mathbf{x}, \mathbf{y}) = \mathbf{0} \end{cases}$$

$T_0 = 9.67$	$x_d = 2.38$	$x'_d = 0.336$	$x_q = 1.21$
$H = 3$	$R_i = 0.002$	$\omega_g = \omega_{\text{ref}} = 1$	$R_t = 0.01$
$X_t = 1.185$	$V_g = 1$	$T_a = 1$	$K_a = 70$
$V_{\text{ref}} = 1$	$T_g = 0.4$	$K_g = 0.5$	$P_{\text{ref}} = 0.7$

Table 3.1 – Parameter values for the SMIB model (p.u.).

with $\mathbf{f}, \mathbf{g}$ polynomial maps, state variable $\mathbf{x} = (\theta, \omega, e_q, E_f, P_m)^\top$ and algebraic variable $\mathbf{y} = (i_d, i_q, V_t)$. Eventually, this DAEs has index 1: for all $\mathbf{x} \in \mathbb{R}^5$, the map $\mathbf{y} \mapsto \mathbf{g}(\mathbf{x}, \mathbf{y})$ is invertible, with

$$\mathbf{i}_s = (\mathrm{Z}_t + R_i\ \mathrm{I} + \mathrm{X}'_i)^{-1}(\mathbf{e} - \mathbf{v}_g) \qquad \& \qquad V_t = |\mathbf{v}_g + \mathrm{Z}_t\ \mathbf{i}_s|.$$

After some elementary computations and simplifications, one obtains the closed formulae

$$i_q = \frac{(X_t + x'_d)V_g \sin\theta + (R_t + R_i)(e_q - V_g \cos\theta)}{(R_t + R_i)^2 + (X_t + x'_d)(X_t + x_q)} \tag{3.14a}$$

$$i_d = \frac{X_t + x_q}{R_t + R_i} i_q - \frac{1}{R_t + R_i} V_g \sin\theta \tag{3.14b}$$

$$v_d = x_q i_q - R_i i_d \tag{3.14c}$$

$$v_q = R_t i_q + X_t i_d + V_g \cos\theta \tag{3.14d}$$

Introducing transients

We next introduce a temporary short-circuit at node S, according to the following protocol:

- At a time t_{sc} a short-circuit occurs and we switch from the nominal system (Figure 3.6) to a new short-circuited system (Figure 3.8), and thus we are no longer at an equilibrium point.
- During the short-circuit, the system leaves the equilibrium point of the nominal model and follows the short-circuit equations for the duration Δt, until the short-circuit is eliminated.
- At a time $t_{\text{cl}} = t_{\text{sc}} + \Delta t$, we switch back to the nominal topology and equations: the problem is to know whether the system will converge to an equilibrium point or not.

The short-circuit equations are computed as follows using the representation of Figure 3.8: Kirchhoff laws now yield that

$$\mathbf{v}_s = \mathbf{v}_g + \mathbf{u}_t \qquad \& \qquad \mathbf{i}_s = \mathbf{i}_a + \mathbf{i}_b,$$

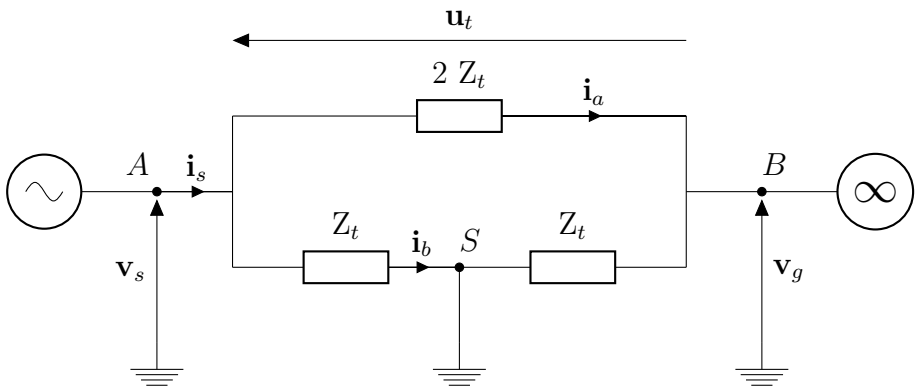


Figure 3.8 – The short-circuited system.

and according to Ohm's laws

$$\mathbf{u}_t = 2\ \mathrm{Z}_t\ \mathbf{i}_a \qquad \& \qquad \mathbf{v}_s = \mathrm{Z}_t\ \mathbf{i}_b,$$

so that $3\ \mathbf{v}_s = \underbrace{\mathbf{v}_g + 2\ \mathrm{Z}_t\ \mathbf{i}_a}_{\mathbf{v}_s} + 2\ \underbrace{\mathrm{Z}_t\ \mathbf{i}_b}_{\mathbf{v}_s} = \mathbf{v}_g + 2\ \mathrm{Z}_t\ \mathbf{i}_s$, and eventually

$$\mathbf{v}_s = \frac{1}{3}\mathbf{v}_g + \frac{2}{3}\mathrm{Z}_t\ \mathbf{i}_s$$

that one can compare with the nominal equation $\mathbf{v}_s = \mathbf{v}_g + \mathrm{Z}_t\ \mathbf{i}_s$. Since the short-circuit only affects the transmission line equation, we conclude that the short-circuited system's equations are the same as (3.11)–(3.14), but with R_t (resp. X_t) replaced by $\frac{2}{3}R_t$ (resp. $\frac{2}{3}X_t$) and V_g replaced by $\frac{1}{3}V_g$. Thus, equations (3.11)–(3.13) are not affected by the short-circuit.

In order to decide if, after the short-circuit is cleared at time t_{cl}, the trajectory of nominal system (3.11)–(3.14) will return to its stable operating point, we use SOS programming tools and Lyapunov function arguments to compute a numerical approximation of its region of attraction.

3.2.2 Lyapunov-based inner ROA approximation

The approach that we use in this section is based on Lyapunov stability theory, and more precisely on a specific definition of stability. We again focus on the differential system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}), \tag{1.4}$$

with equilibrium point $\overline{\mathbf{x}} \in \mathbb{R}^n$, *i.e.* $\mathbf{f}(\overline{\mathbf{x}}) = \mathbf{0}$.

We are interested in the stability properties of (1.4). There exist several notions of stability, so we now detail only the one that interests us in this section.

Definition 3.7: Local Asymptotic Stability

The equilibrium point $\overline{\mathbf{x}}$ of ODE (1.4) is said to be:

- *Lyapunov-stable* (L-S) if

$$\forall\ \varepsilon > 0, \exists\ \eta > 0; |\mathbf{x}_0 - \overline{\mathbf{x}}| < \eta \Longrightarrow \forall\ t \geq 0, |\mathbf{x}(t|\mathbf{x}_0) - \overline{\mathbf{x}}| < \varepsilon$$

  i.e. if trajectories initialized close enough to $\overline{\mathbf{x}}$ do not go too far from it.

- *locally asymptotically stable* (LAS) if it is L-S and

$$\exists\ R > 0; |\mathbf{x}_0 - \overline{\mathbf{x}}| < R \Longrightarrow \mathbf{x}(t|\mathbf{x}_0) \underset{t\to\infty}{\longrightarrow} \overline{\mathbf{x}}$$

  i.e. if $\overline{\mathbf{x}}$ attracts all trajectories initialized close enough to it.

Remark 3.10 (Alternative characterization of LAS points)

Using Definition 1.5, the equilibrium point $\overline{\mathbf{x}}$ is LAS iff $\mathbf{A}_\infty(\overline{\mathbf{x}})$ contains a ball of positive radius $R > 0$.

Several stability estimation theorems are due to Alexandr Mikhailovich Lyapunov and his use of observable functions that are named after him [81]. In particular, the following result holds.

Theorem 3.8: Lyapunov-Persidsky

Suppose that the equilibrium of (1.4) is $\overline{\mathbf{x}} = \mathbf{0}$. Then $\mathbf{0}$ is L-S *iff* there exists an open domain $\mathbf{D} \subset \mathbb{R}^n$ and a $\mathcal{C}^1$ "*Lyapunov function*" $V : \mathbf{D} \to \mathbb{R}$ such that:

(i) $\mathbf{0} \in \mathbf{D}$ and $V(\mathbf{0}) = 0$,

(ii) $\forall \mathbf{x} \in \mathbf{D} \setminus \{\mathbf{0}\}$, $V(\mathbf{x}) > 0$ (*positive definiteness*),

(iii) $\forall \mathbf{x} \in \mathbf{D}$, $\dot{V}(\mathbf{x}) := \mathbf{f}(\mathbf{x}) \cdot \mathbf{grad}\ V(\mathbf{x}) \leq 0$ (*negative semi-definiteness*).

The *if* part is due to Lyapunov in [81], while the *only if* comes from Persidsky [104].

At this point, it is worth recalling that the question we want to answer is the following: after given transients, is our SMIB system (3.11)–(3.14) still able to converge to stable operating point? One can prove that such stable operating point is actually a LAS equilibrium, which means that if the post-fault state is close enough to equilibrium, then it will indeed converge.

Definition 1.5 allows us to mathematically formulate the transient stability problem that we presented in Section 3.2.1 for an infinite time horizon.

Problem 11: Convergence to operating point

Consider the system (3.11)–(3.14), along with a post-fault condition $\mathbf{x}_{\text{cl}}$ and a LAS equilibrium point $\overline{\mathbf{x}}$. Determine whether

$$\mathbf{x}_{\text{cl}} \in \mathbf{A}_\infty(\overline{\mathbf{x}}).$$

When $T = +\infty$, Definition 1.5 is accompanied with the Krasovsky-LaSalle principle, which was proved approximately at the same time by Nikolai Nikolaievich Krasovsky [10] and Joseph Pierre LaSalle [67, Theorem 2].

Theorem 3.9: Krasovsky-LaSalle

Let $V \in \mathcal{C}^1(\mathbb{R}^n)$ and $l \in \mathbb{R}$ such that $\mathbf{\Omega} := \{\mathbf{x} \in \mathbb{R}^n : V(\mathbf{x}) \leq l\}$ is bounded and $\forall\, \mathbf{x} \in \mathbf{\Omega}$, $\dot{V}(\mathbf{x}) \leq 0$. Let

$$\mathbf{M} := \{\mathbf{x}_0 \in \mathbf{\Omega} : \forall\, t \in \mathbb{R}, \dot{V}(\mathbf{x}(t|\mathbf{x}_0)) = 0\}.$$

Then, $\mathbf{\Omega}$ is a positively invariant subset of $\mathbf{A}_\infty(\mathbf{M})$.

From Theorems 3.8 and 3.9, one can elegantly deduce a characterization for local asymptotic stability, due to Lyapunov [81] (*if*) and Massera [88] (*only if*).

Theorem 3.10: Lyapunov-Massera

Suppose that the equilibrium of (1.4) is $\overline{\mathbf{x}} = \mathbf{0}$. Then $\mathbf{0}$ is LAS *iff* there exists an open domain $\mathbf{D} \subset \mathbb{R}^n$ and a $\mathcal{C}^1$ function $V : \mathbf{D} \to \mathbb{R}$ such that:

- Conditions *(i)* and *(ii)* of Theorem 3.8 hold,
- $\forall \mathbf{x} \in \mathbf{D} \setminus \{\mathbf{0}\}$, $\dot{V}(\mathbf{x}) < 0$.

Then, for any $l \in \mathbb{R}$ such that $\mathbf{\Omega} = \{\mathbf{x} \in \mathbb{R}^n : V(\mathbf{x}) \leq l\} \subset \mathbf{D}$, one has that $\mathbf{\Omega}$ is a positively invariant subset of $\mathbf{A}_\infty(\mathbf{0})$.

Proof : For the considered V, $\mathbf{M} = \{\mathbf{0}\}$ in Theorem 3.9. The proof then directly follows from application of Theorem 3.8. Note that the original proof of Lyapunov and Massera is more sophisticated since they did not have access to Theorem 3.9, which is actually posterior to Theorem 3.10 but facilitates its demonstration. ♢

This gives a first answer to Problem 11, up to translation on the state space such that $\overline{\mathbf{x}} = \mathbf{0}$: if there exist $\mathbf{D}$, V and l as in Theorem 3.10 such that $\mathbf{x}_{\text{cl}} \in \mathbf{\Omega}$, then $\mathbf{x}_{\text{cl}} \in \mathbf{A}_\infty(\mathbf{0})$. Thus, what remains to do is find such $\mathbf{D}$, V and l.

However, we want to be able to answer Problem 11 for all possible values of $\mathbf{x}_{\text{cl}}$. In other words, we are looking for the exact ROA, and not only a subset of it. Such problem is a very difficult open question in general, but it is still possible to look for inner approximations of the ROA. For example, the Krasovsky-LaSalle principle allows us to strenghten our problem into a more tractable question:

Problem 12: Inner ROA approximation

Numerically approximate the solution to problem

$$\begin{aligned} \mathrm{q}^{\star}_{\mathrm{ROA}} = \sup_{\mathbf{D},V,l} \quad & \mathrm{size}(\mathbf{\Omega}) \\ \text{s.t.} \quad & \mathbf{D} \subset \mathbb{R}^n, V \in \mathcal{C}^1(\mathbf{D}), l \in \mathbb{R} \\ & \mathbf{0} \in \mathbf{D} \ \& \ V(\mathbf{0}) = 0 \\ & \forall\, \mathbf{x} \in \mathbf{D} \setminus \{\mathbf{0}\}, V(\mathbf{x}) > 0 \ \& \ \dot{V}(\mathbf{x}) < 0 \\ & \mathbf{\Omega} = \{\mathbf{x} \in \mathbb{R}^n : V(\mathbf{x}) \leq l\} \\ & \mathbf{\Omega} \subset \mathbf{D}. \end{aligned} \tag{3.15}$$

for an appropriate "size" function.

Remark 3.11 (Convergence of the inner approximation)

Strenghtening Problem 11 into 12 may introduce a strenghtening gap, i.e. one might always only have $\mathbf{\Omega} \subsetneq \mathbf{A}_\infty(\mathbf{0})$. *Up to this day, to the best of our knowledge, the problem of knowing if, for a given maximizing sequence* $\mathbf{D}_k, V_k, l_k$ *and corresponding* $\mathbf{\Omega}_k$, *one has*

$$\mathbf{\Omega}_k \underset{k\to\infty}{\longrightarrow} \mathbf{A}_\infty(\mathbf{0})$$

for at least one Hausdorff[2] *topology on sets, is still an open question.*

Remark 3.12 (Relevance of the approach)

While the infinite time horizon is not a must, this approach has the advantage to rely on Lyapunov's well-known stability theory, making it easy to understand. Moreover, this approach natuarlly favors ***inner*** *approximations of the ROA, which we identified in Remark 1.2 as a crucial feature for TSA.*

3.2.3 Computing an SOS Lyapunov function

In [103], an SOS programming based method was given for certifying whether a polynomial map is a Lyapunov function or not. In this section, we display a numerical scheme, based on this method, which was designed in [51] to compute approximate solutions to Problem 12 for polynomial dynamics $\mathbf{f} \in \mathbb{R}[\mathbf{x}]^n$, and extended to an instance of nonpolynomial dynamics in [6]. Our contribution here is to systematize such extension to a class of dynamics that we name *algebraic dynamics* (not to be confused with differential algebraic systems).

From algebraic to polynomial dynamics

As highlighted in Chapter 2, reformulating an optimization problem under the form of a SOS programming problem can only be done if all the data of the initial problem are polynomial. However, in our case, equations (3.11)–(3.14) include trigonometric

[2] *i.e.* a topology that separates sets $\mathbf{A}$ and $\mathbf{B}$ when $\mathbf{A} \neq \mathbf{B}$.

functions and non-squared euclidean norms. In addition, Lyapunov functions need the considered equilibrium point to be $\mathbf{0}$.

Those two difficulties lead us to perform a change of variables, as described in [6, 53]. Let $\overline{\mathbf{x}} := (\overline{\theta}, \overline{\omega}, \overline{e}_q, \overline{E}_f, \overline{P}_m)^\top$ be the LAS equilibrium point for (3.11)–(3.14), which we summarize under the form $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ with state $\mathbf{x} = (\theta, \omega, e_q, E_f, P_m)^\top$ and *nonpolynomial* vector field $\mathbf{f} \in \mathcal{C}^1(\mathbb{R}^5)^5$, and let $\overline{V}_t$ be the corresponding value for V_t. We define:

$$\begin{aligned}
\phi_1(\mathbf{x}) &:= \sin(\theta - \overline{\theta}) \\
\phi_2(\mathbf{x}) &:= 1 - \cos(\theta - \overline{\theta}) \\
\phi_3(\mathbf{x}) &:= \omega - \overline{\omega} \\
\phi_4(\mathbf{x}) &:= e_q - \overline{e}_q \\
\phi_5(\mathbf{x}) &:= E_f - \overline{E}_f \\
\phi_6(\mathbf{x}) &:= P_m - \overline{P}_m \\
\phi_7(\mathbf{x}) &:= V_t - \overline{V}_t \\
\phi_8(\mathbf{x}) &:= \frac{1}{V_t} - \frac{1}{\overline{V}_t}.
\end{aligned}$$

along with the change of variable:

$$\mathbf{y} := \Phi(\mathbf{x})$$

and the equality constraints:

$$\begin{cases}
0 = h_1(\mathbf{y}) := y_1^2 + (1 - y_2)^2 - 1 \\
0 = h_2(\mathbf{y}) := (y_7 + \overline{V}_t)^2 - V_t(\mathbf{y})^2 \\
0 = h_3(\mathbf{y}) := (y_7 + \overline{V}_t)\left(y_8 + \frac{1}{\overline{V}_t}\right) - 1,
\end{cases}$$

where equation (3.14) ensures that $V_t(\mathbf{y})^2$ is a degree 2 polynomial in $\mathbf{y}$. We then define $\mathbf{X} := \mathbb{R}^5$ as well as the algebraic set:

$$\mathbf{Y} := \left\{\mathbf{y} \in \mathbb{R}^8 : \mathbf{h}(\mathbf{y}) = \mathbf{0}\right\}.$$

By construction of $\mathbf{h}$, $\Phi : \mathbf{X} \longrightarrow \mathbf{Y}$ is a diffeomorphism with smooth inverse Ψ defined by $\psi_1(\mathbf{y}) = \overline{\theta} + \text{sign}(y_1)\arccos(1 - y_2)$ and $\psi_i(\mathbf{y}) = y_{i+1} + \overline{x}_i$ for $i = 2, 3, 4, 5$.

Remark 3.13 (Validity of the change of variables)
To prove that $\Phi : \mathbf{X} \to \mathbf{Y}$ is well defined, we also need $\forall \mathbf{y} \in \mathbf{Y}\ V_t(\mathbf{y}) \neq 0$. In fact, the physics ensure that $\forall t\ V_t \neq 0$ because we do not consider short-circuits inside the synchronous machine. So we already know that starting from any physically relevant point, the system will not reach the set $\{\mathbf{x} \in \mathbf{X} : V_t(\mathbf{x}) = \mathbf{0}\}$, which is sufficient here.

According to [34], Φ is a particular case of endogenous transformation between systems

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}), \qquad \mathbf{x} \in \mathbf{X} \tag{1.4}$$

$$\dot{\mathbf{y}} = \partial\Phi(\Psi(\mathbf{y}))\ \mathbf{f}(\Psi(\mathbf{y})), \qquad \mathbf{y} \in \mathbf{Y}. \tag{1.4'}$$

As such it preserves the stability properties as well as all trajectory properties of system (3.11)–(3.14) summarized as (1.4), and thus studying the stability of $\mathbf{y}$ is equivalent to studying the stability of $\mathbf{x}$.

Moreover, by construction of Φ it now holds that $\dot{\mathbf{y}} = \partial\Phi(\Psi(\mathbf{y}))\ \mathbf{f}(\Psi(\mathbf{y})) \in \mathbb{R}[\mathbf{y}]$ and $\mathbf{0} = \Phi(\overline{\mathbf{x}})$ is the LAS equilibrium point of interest in the reformulated system (1.4'), which corresponds to the general context in which the following scheme (as well as the hierarchy presented in Section 3.1) can be applied. Before going to the

next step of the method, we give the extended framework to which SOS programming schemes can be applied.

Definition 3.11: Algebraic dynamics

Let $\mathbf{X} := \mathbb{R}^n$. $\mathbf{f} \in \mathcal{C}^\infty(\mathbf{X})^n$ is said to represent *algebraic dynamics iff*:

1. there exists $\mathbf{h} \in \mathbb{R}[\mathbf{y}]^q$ with $\mathbf{y} \in \mathbb{R}^m$, that define the basic semialgebraic set
$$\mathbf{Y} := \{\mathbf{y} \in \mathbb{R}^m : \mathbf{h}(\mathbf{y}) = \mathbf{0}\}$$
2. there exists a smooth diffeomorphism $\Phi : \mathbf{X} \longrightarrow \mathbf{Y}$ with smooth inverse Ψ such that $(\partial\Phi \circ \Psi)\ (\mathbf{f} \circ \Psi) \in \mathbb{R}[\mathbf{y}]^m$.

Remark 3.14 (Application to occupation measures)
Algebraic dynamics always have the property that $(\partial\Phi \circ \Psi)\ (\mathbf{f} \circ \Psi)$ *is tangent to* $\mathbf{Y}$*, allowing for systematic use of the heuristic described in Section 3.1.3.*

The algorithm that we present hereafter applies to any differential system with LAS equilibrium point $\overline{\mathbf{x}} \in \mathbf{X}$ and algebraic dynamics such that $\Phi(\overline{\mathbf{x}}) = \mathbf{0}$.

Another Positivstellensatz

Similarly to Chapter 2, we are now faced with an optimization problem (3.15). However, this problem is much more complex than those we discussed earlier, and cannot be directly tackled using the standard moment-SOS hierarchy. Indeed, Problem 12 not only includes semialgebraic set inequality constraints, but also non-nullity constraints, which require a generalization of the notion of semialgebraic set.

Definition 3.12: General polynomially constrained set

Let $\mathbf{g} \in \mathbb{R}[\mathbf{x}]^p$, $\mathbf{h} \in \mathbb{R}[\mathbf{x}]^q$, $\boldsymbol{\ell} \in \mathbb{R}[\mathbf{x}]^r$, $p, q, r \in \mathbb{N}^\star$. Then, we define

$$\mathbf{U}(\mathbf{g}, \mathbf{h}, \boldsymbol{\ell}) := \left\{ \mathbf{x} \in \mathbb{R}^n \;\middle|\; \begin{array}{l} g_1(\mathbf{x}) \geq 0, \ldots, g_p(\mathbf{x}) \geq 0 \\ h_1(\mathbf{x}) = 0, \ldots, h_q(\mathbf{x}) = 0 \\ \ell_1(\mathbf{x}) \neq 0, \ldots, \ell_r(\mathbf{x}) \neq 0 \end{array} \right\}.$$

Then, real algebraic geometry provides us with results that allow for more generality than Putinar's P-satz 2.11, at the price of new computational technicalities that we are going to explain now.

We first introduce general algebraic structures that will be at the center of another Positivstellensatz. Let $\mathbf{g} := (g_1, \ldots, g_m) \in \mathbb{R}[\mathbf{x}]^m$, $m \in \mathbb{N}^\star$.

Definition 3.13: Multiplicative monoid

We define the *multiplicative monoid* generated by $\mathbf{g}$ as

$$(\mathbf{g})^\star := \left\{ \mathbf{g}^{\mathbf{k}} : \mathbf{k} \in \mathbb{N}^m \right\}.$$

Definition 3.14: SOS cone

We define the *SOS cone* generated by $\mathbf{g}$ as

$$\Sigma[\mathbf{x}][\mathbf{g}] := \left\{ \sum_{i=1}^{N} s_i \ \alpha_i : N \in \mathbb{N} \wedge \forall i \in \mathbb{N}_N^\star, (\alpha_i \in (\mathbf{g})^\star \wedge s_i \in \Sigma[\mathbf{x}]) \right\}.$$

Definition 3.15: Ideal

We define the *ideal* generated by $\mathbf{g}$ as

$$\langle \mathbf{g} \rangle := \{ \mathbf{p} \cdot \mathbf{g} : \mathbf{p} \in \mathbb{R}[\mathbf{x}]^m \}.$$

With these definitions we can now state the following fundamental theorem [70, Theorem 2.11].

Theorem 3.16: Krivine-Stengle weak Positivstellensatz (P-satz)

Let $\mathbf{g} \in \mathbb{R}[\mathbf{x}]^p$, $\mathbf{h} \in \mathbb{R}[\mathbf{x}]^q$, $\boldsymbol{\ell} \in \mathbb{R}[\mathbf{x}]^r$, $p, q, r \in \mathbb{N}^\star$. Then, the following are equivalent:

1. $\mathbf{U}(\mathbf{g}, \mathbf{h}, \boldsymbol{\ell}) = \varnothing$.
2. There exist $g \in \Sigma[\mathbf{x}][\mathbf{g}]$, $h \in \langle \mathbf{h} \rangle$ and $\ell \in (\boldsymbol{\ell})^\star$ such that

$$g + h + \ell^2 = 0, \tag{3.16}$$

The LMI based tests for SOS polynomials provided by Proposition 2.13 can be used to prove that the set emptiness condition from Theorem 3.16 holds, by finding specific g, h and k such that $g + h + \ell^2 = 0$. g, h and ℓ^2 are known as P-satz certificates or P-satz refutations, since they certify emptiness of $\mathbf{U}(\mathbf{g}, \mathbf{h}, \boldsymbol{\ell})$ and refute its nonemptiness. The search for bounded degree P-satz certificates can be done using semidefinite programming (SDP). If the degree bound is chosen large enough the SDP will be feasible and give the refutation certificates. As for the moment-SOS hierarchy, by putting an upper bound on the P-satz certificates degrees and checking whether (3.16) has a solution within these bounds, one can create a series of tests for the emptiness of $\mathbf{U}(\mathbf{g}, \mathbf{h}, \boldsymbol{\ell})$. Each of these tests requires the construction of some SOS and polynomial multipliers, resulting in a SOS program that can be modelled as an SDP using SOSTOOLS.

The expanding interior algorithm

In the case of problem 12 with reformulated *polynomial* dynamics (1.4′) and LAS equilibrium point at $\mathbf{0}$, two algorithms making this research possible are discussed in [51]: the *expanding* $\mathbf{D}$ *algorithm* and the *expanding interior* (EI) algorithm. Since the latter is more efficient than the former in practice, we only implemented the EI algorithm.

The EI algorithm consists of using a positive definite $W \in \Sigma[\mathbf{y}]$ and an expansion

parameter $b > 0$ in order to define a variable sized region,

$$\mathbf{P}_b := \{\mathbf{y} \in \mathbf{Y} : W(\mathbf{y}) \leq b\},$$

included in the sublevel set $\mathbf{\Omega} := \{\mathbf{y} \in \mathbf{Y} : V(\mathbf{y}) \leq 1\}$ of a yet unknown Lyapunov function V, which we choose to be positive on the whole $\mathbf{Y} \setminus \{\mathbf{0}\}$, so that its domain $\mathbf{D}$ satisfies

$$\mathbf{D} = \{\mathbf{0}\} \cup \{\mathbf{y} \in \mathbf{Y} : \dot{V}(\mathbf{y}) < 0\}.$$

The optimization problem is to expand b as long as we can find a V such that $\mathbf{P}_b \subset \mathbf{\Omega}$. The level set $\mathbf{\Omega}$ corresponding to the largest b is our best approximation of the ROA. In order to enforce $\mathbf{\Omega} \subset \mathbf{D}$, we must have

$$\{\mathbf{y} \in \mathbf{Y} : V(\mathbf{y}) \leq 1\} \setminus \{\mathbf{0}\} \subseteq \{\mathbf{y} \in \mathbf{Y} : \dot{V}(\mathbf{y}) < 0\}.$$

Positive definiteness of V on $\mathbf{Y} \setminus \{\mathbf{0}\}$, $\mathbf{P}_b \subset \mathbf{\Omega}$ and $\mathbf{\Omega} \subset \mathbf{D}$ are then respectively rephrased as

$$\{\mathbf{y} \in \mathbf{Y} \setminus \{\mathbf{0}\} : V(\mathbf{x}) \leq 0\} = \varnothing \quad \& \quad \mathbf{P}_b \cap \mathbf{\Omega}^c = \varnothing \quad \& \quad \mathbf{\Omega} \cap \mathbf{D}^c = \varnothing.$$

Consequently, the EI algorithm optimization problem can be formulated using set emptiness constraints as:

Problem 13: Expanding interior optimization problem

$$\begin{aligned} \mathrm{b}_W^\star := \max_{b,V} \quad & b \qquad (3.17)\\ \text{s.t.} \quad & b \in \mathbb{R}, V \in \mathbb{R}[\mathbf{y}], V(\mathbf{0}) = 0 \\ & \varnothing = \{\mathbf{y} \in \mathbb{R}^m : (V(\mathbf{y}) \leq 0) \wedge (\mathbf{h}(\mathbf{y}) = \mathbf{0}) \wedge (\mathbf{y} \neq 0)\} \\ & \varnothing = \{\mathbf{y} \in \mathbb{R}^m : (W(\mathbf{y}) \leq b) \wedge (\mathbf{h}(\mathbf{y}) = \mathbf{0}) \wedge (V(\mathbf{y}) > 1)\} \\ & \varnothing = \{\mathbf{y} \in \mathbb{R}^m : (V(\mathbf{y}) \leq 1) \wedge (\mathbf{h}(\mathbf{y}) = \mathbf{0}) \wedge (\dot{V}(\mathbf{y}) \geq 0) \wedge (\mathbf{y} \neq 0)\}. \end{aligned}$$

It is worth noting that the particular choice of the EI algorithm is a way to strenghten (3.15) into (3.17).

Then, taking positive definite $\ell_1, \ell_2 \in \Sigma[\mathbf{y}]$ (e.g. $\ell_i(\mathbf{y}) = \alpha_i \, |\mathbf{y}|^2$) to reformulate the constraint $\mathbf{y} \neq \mathbf{0}$ as a low dimensional polynomial constraint, we can rephraze (3.17) in terms of emptiness of some well chosen $\mathbf{U}(\mathbf{g}, \mathbf{h}, \boldsymbol{\ell})$ sets:

$$\begin{aligned} \mathrm{b}_W^\star = \max_{b,V} \quad & b \qquad (3.18)\\ \text{s.t.} \quad & b \in \mathbb{R}, V \in \mathbb{R}[\mathbf{y}], V(\mathbf{0}) = 0 \\ & \varnothing = \mathbf{U}(-V, \mathbf{h}, \ell_1) \\ & \varnothing = \mathbf{U}((b - W, V - 1), \mathbf{h}, V - 1) \\ & \varnothing = \mathbf{U}((1 - V, \dot{V}), \mathbf{h}, \ell_2). \end{aligned}$$

By applying the P-satz Theorem 3.16, (3.18) can in turn be formulated as the following SOS programming problem:

$$
\begin{aligned}
\mathrm{b}_W^\star = \max_{\substack{b,V,(g_i)_i \\ (h_i)_i,(k_i)_i}} \quad & b \\
\text{s.t.} \quad & b \in \mathbb{R}, V \in \mathbb{R}[\mathbf{y}], V(\mathbf{0}) = 0 \\
& k_1, k_2, k_3 \in \mathbb{N}, h_1, h_2, h_3 \in \langle \mathbf{h} \rangle \\
& g_1 \in \Sigma[\mathbf{y}][-V]; \qquad\qquad 0 = g_1 + h_1 + \ell_1^{2k_1} \\
& g_2 \in \Sigma[\mathbf{y}][b - W, V - 1]; \quad 0 = g_2 + h_2 + (V-1)^{2k_2} \\
& g_3 \in \Sigma[\mathbf{y}][1 - V, \dot{V}]; \qquad 0 = g_3 + h_3 + \ell_2^{2k_3}.
\end{aligned}
$$

At this stage, our problem could still include an infinite number of SOS constraints, due to the structure of $\Sigma[\mathbf{y}][\mathbf{g}]$ that gives the possibility to multiply any number of powers of $\mathbf{g}$ with SOS coefficients. Thus, a first strenghtening can be implemented by replacing $\Sigma[\mathbf{y}][\mathbf{g}]$ in equation (3.16) with

$$
\widehat{\Sigma}[\mathbf{y}][\mathbf{g}] := \left\{ \sum_{i=1}^{N} s_i \, \mathbf{g}^{\mathbf{k}_i} : N \in \mathbb{N} \wedge \forall i \in \mathbb{N}_N^\star, s_i \in \Sigma[\mathbf{y}] \wedge \forall j \in \mathbb{N}_p^\star, k_{ij} \in \{0,1\} \right\}
$$

which does not remove too much information when one is looking for small degree certificates. Then, one obtains the following SOS programming problem:

$$
\begin{aligned}
\widehat{\mathrm{b}}_W^\star := \max_{\substack{b,V,(k_i)_i \\ (\mathbf{p}_i)_i,(s_j)_j}} \quad & b \\
\text{s.t.} \quad & b \in \mathbb{R}, V \in \mathbb{R}[\mathbf{y}], V(\mathbf{0}) = 0 \\
& k_1, k_2, k_3 \in \mathbb{N}, \mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3 \in \mathbb{R}[\mathbf{x}]^q, s_1, \ldots, s_{10} \in \Sigma[\mathbf{y}] \\
& 0 = s_1 - s_2 V + \mathbf{p}_1 \cdot \mathbf{h} + \ell_1^{2k_1} \\
& 0 = s_3 + s_4(b - W) + s_5(V-1) + s_6(b-W)(V-1) \\
& \qquad + \mathbf{p}_2 \cdot \mathbf{h} + (V-1)^{2k_2} \\
& 0 = s_7 + s_8(1-V) + s_9\dot{V} + s_{10}(1-V)\dot{V} + \mathbf{p}_3 \cdot \mathbf{h} + \ell_2^{2k_3}.
\end{aligned}
$$

In order to limit the size of the SOS problem that we will solve numerically, we restrict ourselves to $k_1 = k_2 = k_3 = 1$ and we simplify the first constraint by enforcing $s_2 = \ell_1$ and factoring ℓ_1 out of s_1 and $\mathbf{p}_1$. Since the second constraint contains quadratic terms in the coefficients of V, we select $s_3 = s_4 = 0$, and factor $V - 1$ out of all the terms. Finally, we select $s_{10} = 0$ in the third constraint in order to eliminate the quadratic terms in V and factor ℓ_2 out. Thus, after renumbering the remaining SOS polynomials and bounding their degree, we reduce the SOS problem to:

Problem 14: Expanding interior strenghtened problem

$$
\begin{aligned}
\mathrm{b}_W^d := \max_{\substack{b,V,(s_i)_i \\ (\mathbf{p}_i)_i}} \quad & b \qquad (3.19)\\
\text{s.t.} \quad & b \in \mathbb{R}, V \in \mathbb{R}[\mathbf{y}], V(\mathbf{0}) = 0\\
& s_1 \in \Sigma_{d-d_W}[\mathbf{y}], s_2 \in \Sigma_{d-d_V}[\mathbf{y}], s_3 \in \Sigma_{d-d'_V}[\mathbf{y}]\\
& \mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3 \in \mathbb{R}_{d-\mathrm{d}^\circ \mathbf{h}}[\mathbf{y}]^q\\
& \Sigma_d[\mathbf{y}] \ni s_4 := V - \mathbf{p}_1 \cdot \mathbf{h} - \ell_1\\
& \Sigma_d[\mathbf{y}] \ni s_5 := -((b-W)s_1 + \mathbf{p}_2 \cdot \mathbf{h} + (V-1))\\
& \Sigma_d[\mathbf{y}] \ni s_6 := -((1-V)s_2 + \dot{V}s_3 + \mathbf{p}_3 \cdot \mathbf{h} + \ell_2),
\end{aligned}
$$

where $d_W := \lceil \mathrm{d}^\circ W/2 \rceil$, $d_V := \lceil \mathrm{d}^\circ V/2 \rceil$, $d'_V := \lceil \mathrm{d}^\circ \dot{V}/2 \rceil$.

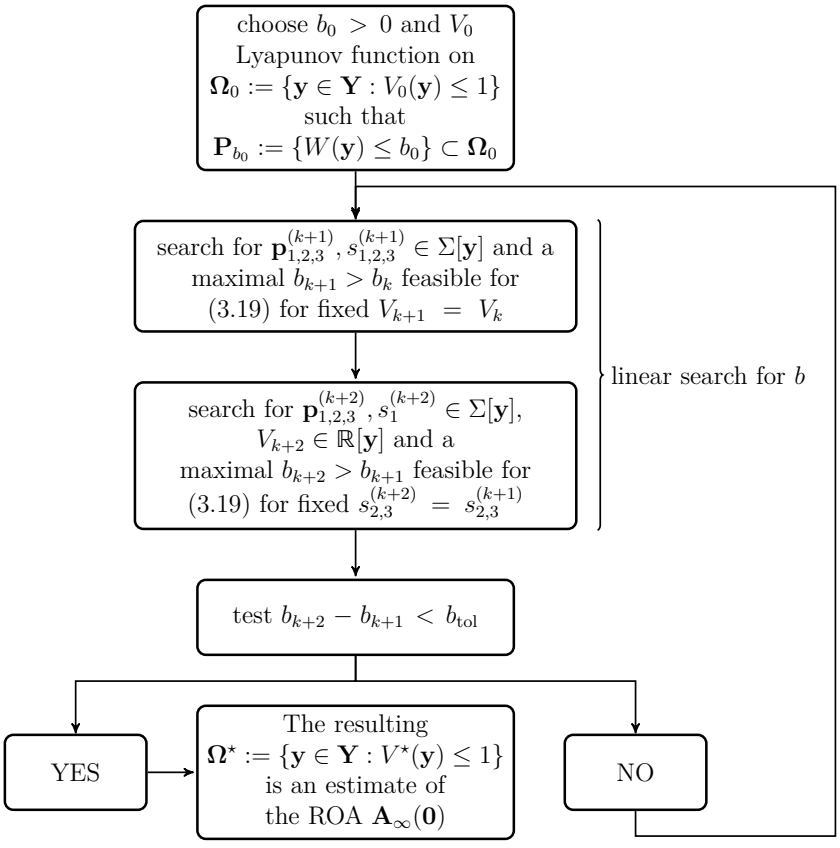


Figure 3.9 – The expanding interior algorithm.

According to Proposition 2.13, problem (3.19) is reduced to an optimization problem with matrix inequalities. However, those inequalities are no longer linear, because of the presence of the terms $b\ s_1$, Vs_2 and $\dot{V}s_3$ that are bilinear in the polynomial decision variables. As a consequence, Proposition 2.13 only allows reformulating problem (3.19) as an optimization problem with bilinear matrix inequalities

(BMI). Such an issue makes problem (3.19) nonconvex and thus more difficult to solve than the convex problems introduced in Chapter 2. Moreover, Lagrange duality theory does not apply anymore so our convergence proofs cannot be used in this context. Nevertheless, problem (3.19) can still be numerically approximated within an iterative algorithm that resorts to a bisection search and alternately optimizes over the variables $b, \mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, s_1, \underline{s_2, s_3}$ on the one hand, and $b, \mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, s_1, \underline{V}$ on the other hand (e.g., [120], [96]), see Figure 3.9.

Remark 3.15 (Dealing with bilinearities)
Fixing some polynomials while optimizing over the others helps dealing with polynomial bilinearities $V s_2$ and $\dot{V} s_3$. On the other hand, the bilinearity $b\ s_1$ cannot be tackled this way, because fixing b would fix the objective function and thus there would be no optimization problem left. Instead, since b is a scalar decision variable, it is possible to perform a bisection algorithm to optimize over both b and s_1 at the same time.

Remark 3.16 (Outer iterations)
In [6], the authors improved the EI algorithm by adding an "outer iteration" loop. The idea is as follows: once the EI algorithm has stopped and given an optimal pair $(b^\star, V^\star)$, one can restart the whole process with $V^\star$ instead of W in the description of $\mathbf{P}_b := \{\mathbf{y} \in \mathbf{Y} : V^\star(\mathbf{y}) \leq b\}$. This makes the algorithm depend less on the arbitrary initial choice of W.

Remark 3.17 (Convergence of the EI algorithm)
Contrary to the moment-SOS hierarchy, the EI algorithm solves a nonconvex optimization problem, which makes any proof of convergence much more difficult. However, simultaneously to the first publication [125] of the present work, [50] proposed an improved version of the EI algorithm with a proof of convergence to a local optimum $(V^\star, b^\star)$.

Some features of SOSTOOLS include the setting of polynomial optimization problems and the search for a polynomial Lyapunov function after expressing the SOS problem as an SDP problem. In [6], SOSTOOLS is used to implement the expanding interior algorithm for a dynamic model without regulation.

3.2.4 ROA estimation of the SMIB model

We finally introduce the second contribution of this section, which is the application of the EI algorithm to the SMIB model of Section 3.2.1. Briefly, the improvement that we bring with respect to [6] is that we tackle a much more complex differential system, which is more computationally challenging than the 2^{nd} order model that was then used as a proof of concept. Then, since the additional control loops help stabilizing the system, one expects to obtain larger inner ROA approximation than in the case of [6].

In fact, the 8 dimensional problem one ends up with after reformulating the SMIB dynamics already exceeds the computational power of a standard laptop. For this reason, we could not directly solve problem (3.19) for system (1.4′). Thus,

we propose to slightly modify the AVR equation (3.12), and transform the voltage droop into a *quadratic droop*, yielding

$$T_a \frac{\mathrm{d}E_f}{\mathrm{d}t} = -E_f + K_a(V_{\mathrm{ref}}^2 - V_t^2). \tag{3.12'}$$

This small change makes the AVR dynamics *polynomial*, by getting rid of the square root operator in the euclidean norm definition. Thus, we no longer need to introduce y_7 and y_8 in the reformulated system which then consists of only the first 6 equations of (1.4′), with $y_1^2 + (1 - y_2)^2 = 1$ as the only remaining polynomial constraint.

From a physical point of view, instead of comparing the magnitude (or *modulus*) of the voltage (phasor) over the synchronous machine to a reference, we are comparing its squared magnitude to the square of the reference.

Then, using an expanding domain $\mathbf{P}_b = \{\mathbf{y} \in \mathbb{R}^6 : W(\mathbf{y}) := |\mathbf{y}|^2 \leq b\}$, the algorithm in Figure 3.9 builds an increasing sequence of inner approximations to the ROA which, when the iterations stop for $b^\star = 0.08544$, returns the following Lyapunov function,

$$\begin{aligned} V^\star(\mathbf{y}) \;=\; & 2.010\, y_1^2 + 0.07823\, y_1y_2 + 3.1961\, y_1y_3 \\ & -2.244\, y_1y_4 - 0.02231\, y_1y_5 + 0.2172\, y_1y_6 \\ & +0.9483\, y_2^2 + 3.422\, y_2y_3 - 2.246\, y_2y_4 \\ & -0.003099\, y_2y_5 + 0.1913\, y_2y_6 + 22.92\, y_3^2 \\ & -0.07196\, y_3y_4 - 0.07616\, y_3y_5 + 2.998\, y_3y_6 \\ & +4.058\, y_4^2 - 0.0003899\, y_4y_5 - 0.1467\, y_4y_6 \\ & +0.004611\, y_5^2 + 0.008518\, y_5y_6 + 0.2425\, y_6^2\,, \end{aligned} \tag{3.20}$$

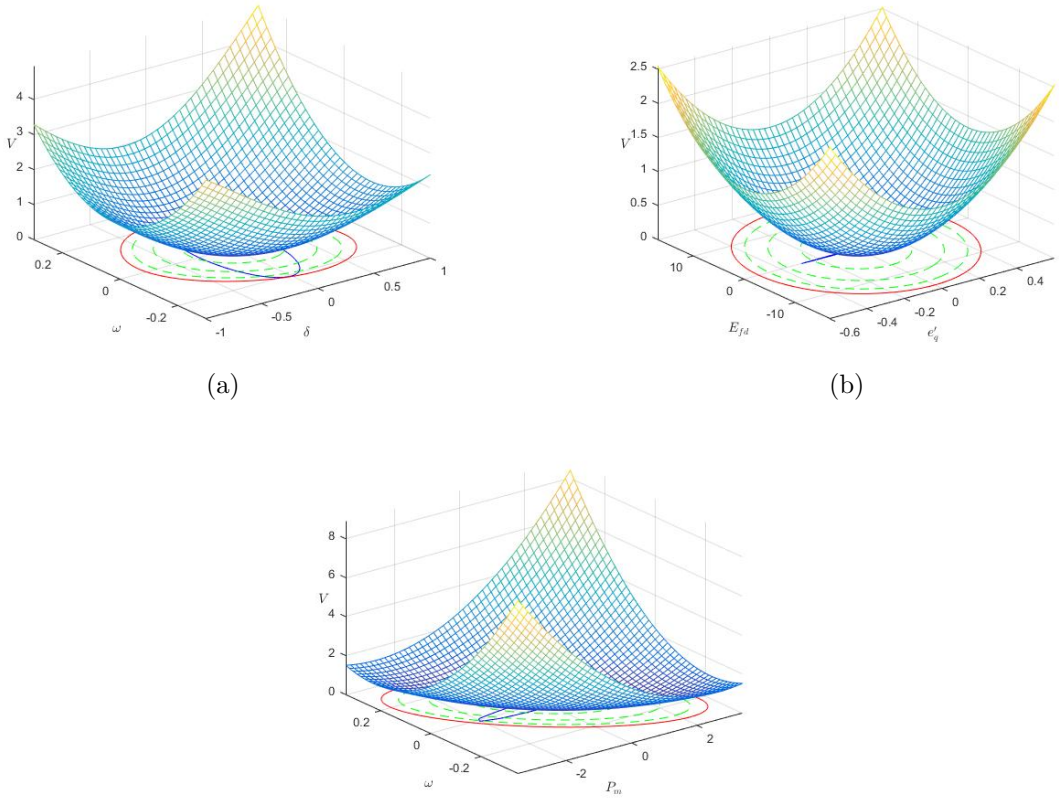
whose 1-sublevel set $\mathbf{\Omega}^\star := \{\mathbf{y} \in \mathbb{R}^6 : V^\star(\mathbf{y}) \leq 1\}$ provides the largest estimation of the ROA.

Two-dimensional projections of the resulting ROA in original coordinates are plotted in Figures 3.10 and 3.11 in red lines. While these estimations are quite large, a comparison to the exact (numerically estimated) ROA, thick blue lines in Figure 3.11, shows that further improvements are possible, for example by increasing the degrees of the polynomials in the SOS program, or by going through more outer iteration loops.

We aim to validate the estimate of the ROA found by the SOS approach, by testing, in simulation, the limits of stability of the system in all the state-space directions. For this, the system is systematically initialized at a starting point far from the considered equilibrium point, and we check by simulation if it goes back to equilibrium or not.

Since the system has 5 state variables which means a huge number of combinations and because θ and ω are the most important state variables for transient stability analysis, we decide to make the test in a projection of the state space on the planes (θ, ω), (e_q, E_f), and (P_m, ω).

Figure 3.11 shows that the estimated ROA (red lines) is inside the real one (thick blue line, computed numerically by simulation) and this validates the previous results. The arrows in the plots show that the real ROA is larger in the direction of the arrows.



(a) (b)

(c)

Figure 3.10 – Projected graph of the Lyapunov function.

On the plots are displayed 2D projections of the estimated ROA (red line) and expanding domain $\mathbf{P}_{b^\star}$ (blue line), and 3D views of the LF in the coordinate pairs: (a) (θ, ω), (b) (e_q, E_f), and (c) (P_m, ω).

The same figure shows that the trajectories of the system initialized in several points inside the numerically computed ROA converge asymptotically to the considered equilibrium point.

In conclusion, SOS approaches and tools have been successfully used to quantify transient stability of a SMIB system for which the generator has been modeled in more detail than in previous studies. Indeed, voltage dynamics and voltage and frequency regulations were taken into account in this formalism. First, this provides more accuracy in estimation of the stability margin in terms of the ROA. Indeed, the estimated ROA is large enough compared to the exact ROA computed by simulation. Next, the Lyapunov approach is well suited for the control synthesis and this quantification can be further exploited to build/tune regulators in order to maximize ROA. As a matter of fact, in the SOS optimization one can next include the regulators' parameters as decision variables. Future work could focus on

- estimation with the full model (without the approximation (3.12′)),
- inclusion of the non linearities of the machine related to saturations of the

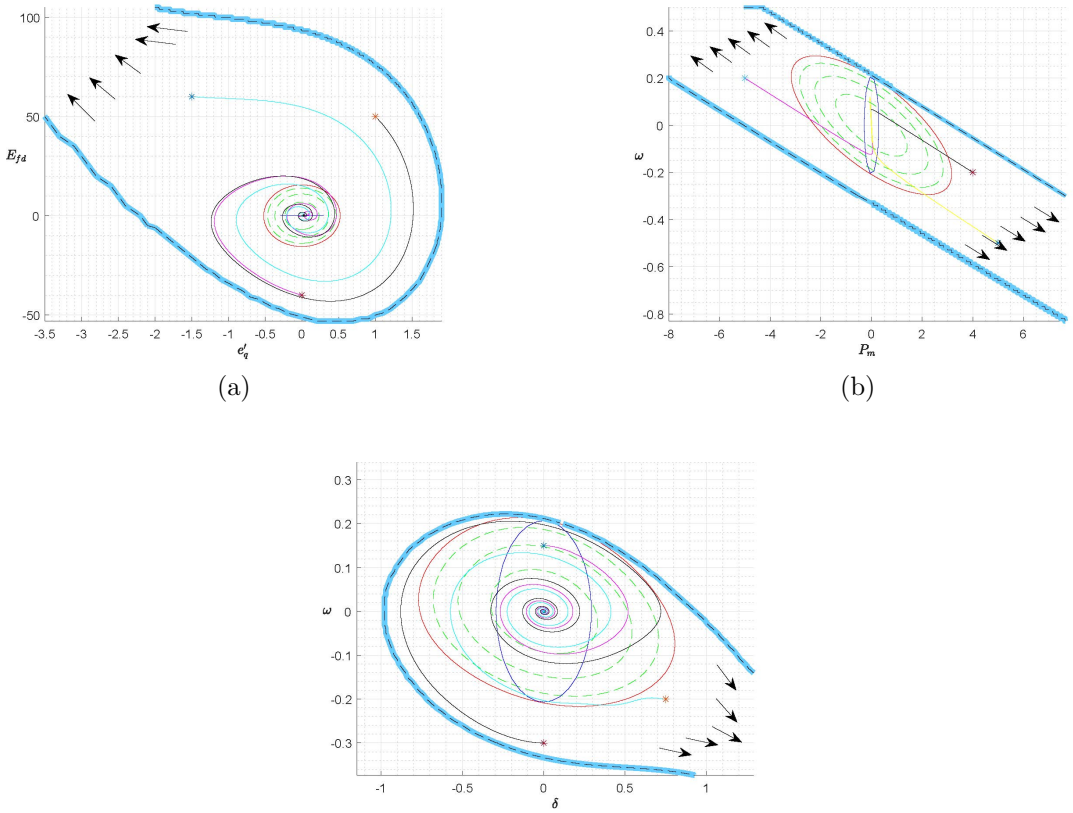


Figure 3.11 – Comparison between ROA estimate and simulation built ROA.

We represented 2D projections of the real ROA (thick blue line), the SOS estimate (thin red line), and expanding domain $P_{\beta_{\max}}$ (thin blue line) on the coordinate subspaces: (a) (e'_q, E_{fd}). (b) (P_m, ω), and (c) (δ, ω). Trajectories of the system initialized in several points (stars) inside the numerically computed ROA converge asymptotically to the considered equilibrium point.

actuating variables,

- application to larger grids using structure-related model reductions,
- extension to the tuning of regulators' parameters.

Synthesis of sections 3.1 & 3.2

This Chapter has been dedicated to proposing new power-systems applications of existing SOS methods for transient stability analysis. The 2$^{\text{nd}}$ order model to which occupation measures were applied in section 3.1 had already been studied in [6] using the Lyapunov framework, in which the generality of the *Krivine-Stengle* P-satz allowed for direct integration of equality constraints; in this matter, our contribution mainly consisted of adapting the more specific *Putinar*-based occupation measure

method to variable-change equality constraints, making possible the study of non-polynomial algebraic dynamics. Section 3.1 was also a first occasion to apply the results of Chapter 2 to an instance of the moment-SOS hierarchy, highlighting the systematic convergence proof allowed by these theoretical results. Section 3.2 then took SOS methods to the next level of complexity, by considering a 5^{th} order model, within the Lyapunov framework. This gives us an opportunity to compare Lyapunov and occupation measure methods in terms of potential applications. We sum up our comparisons in table 3.2.

Both approaches rely on the approximation of the system's ROA with level sets of polynomials. A first important feature of such methods is that a degree d polynomial in N variables is defined by $D_N^d := \binom{N+d}{d} = \frac{(N+d)!}{N!d!}$ coefficients, so that the space of decision variables has dimension D_N^d. One can notice (using the asymptotic equivalent notation $f(x) \underset{x\to a}{\sim} g(x) \leftrightarrow f(x)/g(x) \underset{x\to a}{\longrightarrow} 1$) that $D_N^d \underset{d\to\infty}{\sim} \frac{d^N}{N!}$ grows polynomially in d with exponent N. As a result, one wants N to be as small as possible when implementing those algorithms. For the Lyapunov method, $N = n$ is the number of state variables, while the occupation measure method involves $N = n + 1$ variables, due to the inclusion of time as a variable for function v. Thus, the Lyapunov method is slightly less costly than the occupation measure method, even though both methods hardly scale when n grows as one tackles more realistic models. This highlights the crucial importance of using problem structure to formulate problems with smaller N.

Then, the main difference between the two approaches is that the moment-SOS hierarchy presented in Section 3.1 reduces to semidefinite programming, while the expanding interior algorithm of Section 3.2 resorts to BMI constraints. Consequently, the Lyapunov method loses all the convexity and duality properties that allowed for the strong convergence proof of the occupation measures approach. On another note, a weaker convergence was recently proved for the EI algorithm in [50], in the sense that the EI algorithm terminates, yielding a local optimum for the optimization criterion b. Whether such local optimum corresponds to a tight approximation of the infinite time ROA remains an open question to this day.

However, convergence of the moment-SOS hierarchy comes with a price to pay: the global optimum strongly depends on the parameters $\mathbf{X}, \mathbf{K}$ and T. Dependence on $\mathbf{X}$ is actually good, since it allows taking into account industrial security state constraints in the stability analysis, ruling out stable trajectories that would still damage the power system (such dependence can also be taken into account in the Lyapunov approach, as a constraint in the Krivine-Stengle P-satz). On the contrary, dependence on the target set $\mathbf{K}$ and time horizon should be considered carefully: even though such dependence is also motivated by the TSA specifications of Definition 1.2, the sensitivity to such parameters makes it necessary to precisely determine $\mathbf{K}$ and T in advance, using solid power engineering arguments. An interesting perspective would be to use a Lyapunov-based infinite time ROA inner approximation $\mathbf{\Omega}^\star \subset \mathbf{A}_\infty^{\mathbf{X}}(\overline{\mathbf{x}})$ as a target set for a finite time ROA inner approximation hierarchy. One would then obtain a sequence of sets $\hat{\mathbf{A}}_{d,T} \subset \mathbf{A}_T^{\mathbf{X}}(\mathbf{\Omega}^\star) \subset \mathbf{A}_\infty^{\mathbf{X}}(\overline{\mathbf{x}})$. Then, intuitively, one would have

$$\left[\lim_{d\to\infty} \text{vol}\left(\mathbf{A}_\infty^{\mathbf{X}}(\overline{\mathbf{x}}) \setminus \hat{\mathbf{A}}_{d,T}\right)\right] \underset{T\to\infty}{\longrightarrow} 0.$$

	Lyapunov	Occupation measures
Nb of variables	n (state)	$n+1$ (state & time)
Constraints	BMI	LMI
Convexity	no	yes
Scheme convergence	local	global
Parameters	none	$\mathbf{X}, \mathbf{K}, T$
Time horizon	infinite	finite
Target	equilibrium	eq. neighbourhood

Table 3.2 – Comparison between Lyapunov and occupation measures for TSA.

4

Volume computation and Stokes theorem

As suggested in Section 3.1.2, the hierarchy-based ROA approximation method relies on the set approximation property that appears in the volume computation problem (2.8). Then, one "only" has to include occupation measures constrained by Liouville's PDE (3.3) to go from approximating a given semialgebraic set to estimating an unknown ROA. In other words, the volume computation problem is the simplest set approximation problem one could think of. As a result, it can be useful to look for ways to enhance the volume problem tractability, so that it can be applied to other set approximation problems such as estimating the ROA of a power system. The two following chapters introduce our contributions in this direction, that one can find in [124, 126].

Contents

Consider the problem of computing the Lebesgue volume of a compact basic semi-algebraic set $\mathbf{K} \subset \mathbb{R}^n$. For simplicity of exposition we will restrict to the case where $\mathbf{K}$ is a smooth, simple semialgebraic set, *i.e.* the super-level set $\{\,\mathbf{x} : g(\mathbf{x}) \geq 0\,\} \subset \mathbb{R}^n$ of a single polynomial g.

Computing or even approximating the volume of a *convex body* is already hard theoretically and in practice as well. Even if $\mathbf{K}$ is a convex polytope, exact computation of its volume or integration over $\mathbf{K}$ is a difficult challenge. Computational complexity of these problems is discussed in, e.g. [14, 29, 32]. In particular, any deterministic algorithm with polynomial-time complexity that would compute an upper bound and a lower bound on the volume cannot yield an estimate on the bound ratio better than polynomial in the dimension n. For more detail, the interested reader is referred to the discussion in [44] and to [16] for a comparison. Even *approximating* $\lambda(\mathbf{K})$ by deterministic methods is still a hard problem as explained in [25] and references therein.

If one accepts randomized algorithms that fail with small probability, then the situation is more favorable. Indeed, the probabilistic approximation algorithm of [30] computes the volume to fixed arbitrary relative precision $\varepsilon > 0$ in time polynomial in $1/\varepsilon$. The algorithm uses approximation schemes based on rapidly mixing Markov chains and isoperimetric inequalities; see also hit-and-run algorithms described in e.g. [17, 116, 145]. So far, it seems that the recent works [24, 25] have provided the best algorithm of this type.

In full generality with no specific assumption on $\mathbf{K}$ such as convexity, the only general method available is Monte-Carlo, that is, one samples N points according to Lebesgue measure λ normalized on a simple set $\mathbf{X}$ (e.g. a hypercube or an ellipsoid) that contains $\mathbf{K}$. If P_N is the proportion of points that fall into $\mathbf{K}$ then the random variable $P_N\,\lambda(\mathbf{X})$ provides a good estimator of $\lambda(\mathbf{K})$ with convergence guarantees as N increases. However this estimator is non deterministic and fails to provide lower or upper bounds on $\lambda(\mathbf{K})$.

When $\mathbf{K}$ and $\mathbf{X}$ are compact basic semi-algebraic sets, a deterministic numerical scheme described in [44] provides a sequence $\{\mathrm{d}^d_{\mathbf{K}}\}_{d\in\mathbb{N}} \subset \mathbb{R}$ of upper bounds that converges to $\lambda(\mathbf{K})$ as d increases. Briefly, we recall as stated in Chapter 2

$$\lambda(\mathbf{K}) \;=\; \mathrm{d}^\star_{\mathbf{K}} \;=\; \inf_{w\in\mathbb{R}[\mathbf{x}]} \left\{ \int w\,d\lambda \;:\; w \geq \mathbb{1}_{\mathbf{K}} \text{ on } \mathbf{X} \right\} \tag{4.1}$$

$$\mathrm{d}^d_{\mathbf{K}} \;=\; \inf_{w\in\mathbb{R}_{2d}[\mathbf{x}]} \left\{ \int w\,d\lambda \;:\; w \geq \mathbb{1}_{\mathbf{K}} \text{ on } \mathbf{X} \right\}. \tag{4.2}$$

One can notice that minimizing sequences for (4.1) and (4.2) also minimize the $L^1(\mathbf{X},\lambda)$-norm $\|w - \mathbb{1}_{\mathbf{K}}\|_{L^1(\mathbf{X})} := \int |w - \mathbb{1}_{\mathbf{K}}|\,d\lambda$ (with convergence to 0 in the case (4.1)). As the upper bound $\mathrm{d}^d_{\mathbf{K}} > \lambda(\mathbf{K})$ is obtained by restricting the search in (4.2) to polynomials of degree at most $2d$, the infimum is attained and an optimal solution can be obtained by solving a semidefinite program. Of course, the size of the resulting semidefinite program increases with the degree d; for more details the interested reader is referred to [44].

Then clearly, a Gibbs phenomenon[1] takes place as one tries to approximate, on

[1]The Gibbs phenomenon appears at a jump discontinuity when one numerically approximates a

$\mathbf{X}$ and from above, the discontinuous function $\mathbb{1}_{\mathbf{K}}$ by a polynomial of degree at most d. This makes the convergence of the upper bounds $\mathrm{d}_{\mathbf{K}}^d$ very slow (even for modest dimension problems). A trick was used in [44] to accelerate this convergence but at the price of loosing monotonicity of the resulting sequence.

In fact (4.1) is a dual of the following infinite-dimensional Linear Program (LP) on measures

$$\mathrm{p}_{\mathbf{K}}^\star = \sup_{\mu \in \mathcal{M}(\mathbf{K})_+} \{ \mu(\mathbf{K}) \ : \ \lambda - \mu \in \mathcal{M}(\mathbf{K})_+ \} \tag{4.3}$$

(where $\mathcal{M}(\mathbf{K})_+$ is the space of finite Borel measures on $\mathbf{K}$). Its optimal value is also $\lambda(\mathbf{K})$ and is attained at the unique optimal solution $\mu^\star := \lambda_{\mathbf{K}} = \mathbb{1}_{\mathbf{K}}\, \lambda$ (the restriction of λ to $\mathbf{K}$).

A simple but key observation. As one knows the unique optimal solution $\mu^\star = \lambda_{\mathbf{K}}$ of (4.3), *any* constraint satisfied by $\mu^\star$ (in particular, linear constraints) can be included as a constraint on μ in (4.3) without changing the optimal value and the optimal solution. While these constraints provide additional *restrictions* in (4.3), they translate into additional *degrees of freedom* in the dual (hence a *relaxed* version of (4.1)), and therefore better approximations when passing to the finite-dimensional strengthened version of (4.2). A first set of such linear constraints, experimented in [71] and later in [139, 73], resulted in drastic improvements but with no clear rationale behind such improvements.

Contribution. This chapter is based on [124], whose main message and result is that there is an appropriate set of additional linear constraints on μ in (4.3) such that the resulting dual (a relaxed version of (4.1)) has an explicit *continuous* optimal solution with value $\lambda(\mathbf{K})$. These additional linear contraints (called Stokes constraints) come from an appropriate modelling of Stokes' theorem for integration over $\mathbf{K}$, a refined version of that in [71]. Therefore the optimal continuous solution can be approximated efficiently by polynomials with no Gibbs phenomenon, by the hierarchy of semidefinite relaxations defined in [44] (adapted to these new linear constraints). Interestingly, the technique of proof and the construction of the optimal solution invoke classical results from the field of elliptic partial differential equations (PDE), namely the Lax-Milgram and Poincaré-Wirtinger inequalities as well as regularity theorems for solutions to elliptic PDEs.

Outline. In Section 4.1 we recall the primal-dual linear formulation of the volume problem, explain why the dual value is not attained, resulting in a Gibbs phenomenon, and present the existing Stokes-based heuristics to tackle this issue. In Section 4.2 we revisit the acceleration strategy based on Stokes' theorem, with the aim of introducing a more general acceleration strategy and a new primal-dual linear formulation of the volume problem. Our main result, attainment of the dual value in this new formulation, is stated and proved as Theorem 4.5 in Section 4.3. The drastic improvement in the convergence to $\lambda(\mathbf{K})$ is illustrated on a simple example of the Euclidean ball in Section 4.4.

---

piecewise $\mathcal{C}^1$ function with a polynomial function, e.g. by its Fourier series; see e.g. [128, Chapter 9].

4.1 Existing Stokes-based heuristics

Consider a compact simple semi-algebraic set

$$\mathbf{K} := \{\mathbf{x} \in \mathbb{R}^n : g(\mathbf{x}) \geq 0\}$$

with $g \in \mathbb{R}[\mathbf{x}]$. We suppose that $\mathbf{K} \subset \mathbf{X}$ where $\mathbf{X}$ is a compact basic semi-algebraic set for which we know the moments $\int_{\mathbf{X}} \mathbf{x}^{\mathbf{k}}\, \mathrm{d}\mathbf{x}$ of the Lebesgue measure $\lambda_{\mathbf{X}}$, where $\mathbf{x}^{\mathbf{k}} := x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}$ denotes a multivariate monomial of degree $\mathbf{k} \in \mathbb{N}^n$. We assume that

$$\mathbf{\Omega} := \{\mathbf{x} \in \mathbb{R}^n : g(\mathbf{x}) > 0\}$$

is a nonempty open set with closure

$$\overline{\mathbf{\Omega}} = \mathbf{K},$$

and that its boundary

$$\partial\mathbf{\Omega} = \partial\mathbf{K} = \mathbf{K} \setminus \mathbf{\Omega}$$

is $\mathcal{C}^1$ in the sense that it is locally the graph of a continuously differentiable function. We want to compute the Lebesgue volume of $\mathbf{K}$, i.e., the mass of the Lebesgue measure $\lambda_{\mathbf{K}}$:

$$\lambda(\mathbf{K}) := \int_{\mathbf{K}} \mathrm{d}\mathbf{x} = \int 1\, d\lambda_{\mathbf{K}}(\mathbf{x}).$$

4.1.1 Linear reformulation of the volume problem

If $\mathbf{X} \subset \mathbb{R}^n$ is a compact set, denote by $\mathcal{M}(\mathbf{X})$ the space of signed Borel measures on $\mathbf{X}$, which identifies with the topological dual of $\mathcal{C}(\mathbf{X})$, the space of continuous functions on $\mathbf{X}$. Denote by $\mathcal{M}(\mathbf{X})_+$ the convex cone of non-negative Borel measures on $\mathbf{X}$, and by $\mathcal{C}(\mathbf{X})_+$ the convex cone of non-negative continuous functions on $\mathbf{X}$.

In [44] a sequence of upper bounds converging to $\lambda(\mathbf{K})$ is obtained by applying the moment-SOS hierarchy to approximate, as closely as desired, the (primal) infinite-dimensional LP on measures:

$$\begin{aligned} \mathrm{p}^\star_{\mathbf{K}} = \max_{\mu} \; & \int 1\, d\mu && \text{(2.8a)} \\ \text{s.t. } & \mu \in \mathcal{M}(\mathbf{K})_+ \\ & \lambda - \mu \in \mathcal{M}(\mathbf{X})_+ \end{aligned}$$

whose optimal value is $\lambda(\mathbf{K})$, attained for $\mu^\star := \lambda_{\mathbf{K}}$ (see Section 2.1.3). The LP (2.8a) has an infinite-dimensional LP dual on continuous functions, which reads:

$$\begin{aligned} \mathrm{d}^\star_{\mathbf{K}} = \inf_{w} \; & \int w\, d\lambda && \text{(2.8b)} \\ \text{s.t. } & w - 1 \in \mathcal{C}(\mathbf{K})_+. \\ & w \in \mathcal{C}(\mathbf{X})_+ \end{aligned}$$

Observe that (2.8b) consists in approximating the discontinuous indicator function $\mathbb{1}_{\mathbf{K}}$ (equal to one on $\mathbf{K}$ and zero elsewhere) from above by continuous functions w,

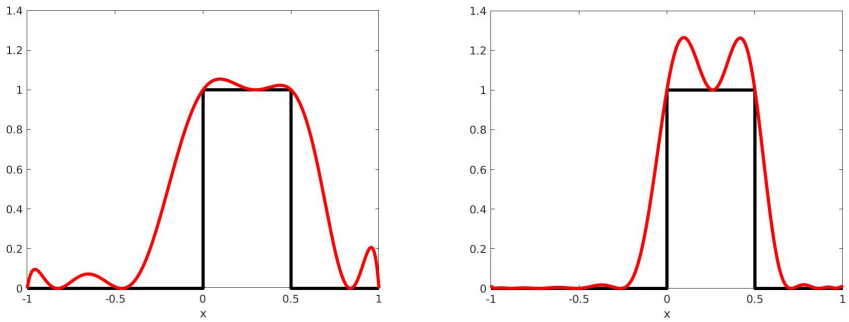


Figure 4.1 – Illustration of the Gibbs phenomenon.

Gibbs effect occurs when approximating from above with a polynomial of degree 10 (left red curve) and 20 (right red curve) the indicator function of an interval (black curve).

by minimizing the $L^1(\mathbf{X})$-norm $\|w - \mathbb{1}_{\mathbf{K}}\|_{L^1(\mathbf{X})}$. Clearly the infimum $\lambda(\mathbf{K})$ is not attained.

Since $\mathbf{K}$ is generated by a polynomial g, one may apply the moment-SOS hierarchy for solving (2.8), as demonstrated in details in Chapter 2. The main drawback of such numerical scheme is its typical slow convergence, observed already for very simple univariate examples, see e.g. [44, Figs. 4.1 and 4.5]. The best available theoretical convergence speed estimates are also pessimistic, with an asymptoptic rate of $\log \log d$ [42]. Slow convergence is mostly due to the so-called Gibbs phenomenon which is well-known in numerical analysis [128, Chapter 9]. Indeed, as already mentioned, solving (2.8b) numerically amounts to approximating the discontinuous function $\mathbb{1}_{\mathbf{K}}$ from above with polynomials of increasing degree, which generates oscillations and overshoots and slows down the convergence, see e.g. [44, Figs. 4.2, 4.4, 4.6, 4.7, 4.10, 4.12].

Example 4.1 *Let* $\mathbf{K} := [0, 1/2] \subset \mathbf{X} := [-1, 1]$. *The degree 10 and degree 20 polynomials* w *obtained by solving the SOS strengthenings of problem* (2.8b) *are displayed in Figure 4.1. We can clearly see bumps, typical of a Gibbs phenomenon at points of discontinuity.*

An idea to bypass this limitation consists of adding certain linear constraints to the finite-dimensional semidefinite relaxations, to make their optimal values smaller and so closer to the optimal value $\lambda(\mathbf{K})$. Such linear constraints must be chosen appropriately:

(i) they must be *redundant* for the infinite-dimensional moment LP on measures (2.8a), and

(ii) become *active* for its finite-dimensional relaxations.

This is the heuristic proposed in [71] to accelerate the Moment-SOS hierarchy for evaluating transcendental integrals on semi-algebraic sets. These additional linear constraints on the moments $\mathbf{z}$ of $\mu^\star$ are obtained from an application of Stokes' theorem for integration on $\mathbf{K}$, a classical result in differential geometry. It has been also observed experimentally that this heuristic accelerates significantly the

convergence of the hierarchy in other applied contexts, e.g. in chance-constrained optimization problems [137].

4.1.2 Stokes' Theorem and its variants

We then explain the heuristic introduced in [44] to accelerate convergence of the Moment-SOS hierarchy by adding linear constraints on the moments of $\mu^\star$. These linear constraints are obtained from a certain application of Stokes' theorem for integration on $\mathbf{K}$.

Theorem 4.1: Stokes

Let $\mathbf{\Omega} \subset \mathbb{R}^n$ be a piecewise $\mathcal{C}^1$ open set. For any $(n-1)$-differential form ω on $\overline{\mathbf{\Omega}}$, one has

$$\int_{\partial\mathbf{\Omega}} \omega = \int_{\mathbf{\Omega}} d\omega.$$

.

Corollary 4.2: Gauss formula

In particular, for $\mathbf{u} \in \mathcal{C}^1(\overline{\mathbf{\Omega}})^n$ and $\omega(\mathbf{x}) = \mathbf{u}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x})\, d\sigma(\mathbf{x})$, where the dot is the inner product, σ is the surface or Hausdorff measure on $\partial\mathbf{\Omega}$ and $\mathbf{n}$ is the outward pointing normal to $\partial\mathbf{\Omega}$, we obtain the Gauss formula

$$\int_{\partial\mathbf{\Omega}} \mathbf{u} \cdot \mathbf{n}\, d\sigma = \int_{\mathbf{\Omega}} \text{div } \mathbf{u}(\mathbf{x})\, \mathrm{d}\mathbf{x}. \tag{4.4}$$

Corollary 4.3: Dual Gauss formula

With the choice $\mathbf{u}(\mathbf{x}) := u(\mathbf{x})\ \mathbf{e}_i$ where $u \in \mathcal{C}^1(\overline{\mathbf{\Omega}})$ and $\mathbf{e}_i$ is the vector of $\mathbb{R}^n$ with one at entry i and zeros elsewhere, for $i = 1, \ldots, n$, we obtain the dual Gauss formula

$$\int_{\partial\mathbf{\Omega}} u\ \mathbf{n}\, d\sigma = \int_{\mathbf{\Omega}} \mathbf{grad}\ u(\mathbf{x})\, \mathrm{d}\mathbf{x}. \tag{4.5}$$

Proof : These are all particular cases of [47, Theorem 6.10.2]. ♢

4.1.3 Original Stokes constraints

Associated to a sequence $\mathbf{z} = (z_\mathbf{k})_{\mathbf{k}\in\mathbb{N}^n} \in \mathbb{R}^{\mathbb{N}^n}$ of moments, introduce the Riesz linear functional $\mathrm{L}_\mathbf{z} : \mathbb{R}[\mathbf{x}] \to \mathbb{R}$ as in Definition 2.1. Thus, if $\mathbf{z}$ is the sequence of moments of $\lambda_\mathbf{K}$, i.e. if for all $\mathbf{k} \in \mathbb{N}^n$, $z_\mathbf{k} := \int_\mathbf{K} \mathbf{x}^\mathbf{k}\, \mathrm{d}\mathbf{x}$, then $\mathrm{L}_\mathbf{z}(p) = \int_\mathbf{K} p(\mathbf{x})\, \mathrm{d}\mathbf{x}$ and by (4.5) with $u(\mathbf{x}) := \mathbf{x}^\mathbf{k} g(\mathbf{x})$:

$$\begin{aligned}\mathrm{L}_\mathbf{z}(\mathbf{grad}(\mathbf{x}^\mathbf{k} g(\mathbf{x}))) &= \int_\mathbf{K} \mathbf{grad}(\mathbf{x}^\mathbf{k} g(\mathbf{x}))\, \mathrm{d}\mathbf{x} \\ &= \int_{\partial\mathbf{K}} \mathbf{x}^\mathbf{k} g(\mathbf{x})\ \mathbf{n}_\mathbf{K}(\mathbf{x})\, d\sigma(\mathbf{x}) \ = \ \mathbf{0},\end{aligned}$$

since by construction g vanishes on $\partial\mathbf{K}$. Thus while in the infinite-dimensional LP (2.8a) one may add the linear constraints

$$\int_{\mathbf{K}} \mathbf{grad}(\mathbf{x}^{\mathbf{k}} g(\mathbf{x}))\, d\mu(\mathbf{x}) \;=\; \mathbf{0} \qquad \forall \mathbf{k} \in \mathbb{N}^n,$$

without changing its optimal value $\mathrm{p}^\star_{\mathbf{K}} = \lambda(\mathbf{K})$, on the other hand inclusion of the linear moment constraints

$$\mathrm{L}_{\mathbf{z}}(\mathbf{grad}(\mathbf{x}^{\mathbf{k}} g(\mathbf{x}))) = \mathbf{0}\,, \quad |\mathbf{k}| \le 2d + 1 - \deg(g) \tag{4.6}$$

in the moment relaxation with pseudo-moments $\mathbf{z}$ of degree at most d, will decrease the optimal value of the initial relaxation.

In practice, it was observed that adding constraints (4.6) dramatically speeds up the convergence of the moment-SOS hierarchy, see e.g. [71, 137]. One main goal of this section is to provide a qualitative mathematical rationale behind this phenomenon.

4.2 Contribution to Stokes constraints heuristics

4.2.1 Infinite-dimensional Stokes constraints

In [126], we formulated Stokes constraints in the infinite-dimensional setting, and a dual formulation was obtained in the context of the volume problem. Using (4.4) with $\mathbf{u} = g\ \mathbf{v}$ (which vanishes on $\partial\mathbf{K}$) and $\mathbf{v} \in \mathcal{C}^1(\mathbf{K})^n$ arbitrary, yields:

$$\int_{\mathbf{K}} (\mathbf{grad}\ g(\mathbf{x}) \cdot \mathbf{v}(\mathbf{x}) + g(\mathbf{x})\ \mathrm{div}\ \mathbf{v}(\mathbf{x}))\ \mathrm{d}\mathbf{x} = \int_{\partial\mathbf{K}} g\ \mathbf{v} \cdot \mathbf{n}\ d\sigma = 0\,,$$

which can be written equivalently (in the sense of distributions) as

$$(\mathbf{grad}\ g)\lambda_{\mathbf{K}} - \mathbf{grad}(g\lambda_{\mathbf{K}}) = \mathbf{0}\,.$$

This allows us to rewrite problem (2.8a) as

$$\begin{aligned} \mathrm{p}^\star_{\mathbf{K}} = \max_{\mu}\ & \int 1\ d\mu \\ \text{s.t. } & \mu \in \mathcal{M}(\mathbf{K})_+ \\ & \lambda - \mu \in \mathcal{M}(\mathbf{X})_+ \\ & (\mathbf{grad}\ g)\mu - \mathbf{grad}(g\mu) = \mathbf{0} \end{aligned} \tag{2.8a$'$}$$

without changing its optimal value $\mathrm{p}^\star_{\mathbf{K}} = \lambda(\mathbf{K})$ attained at $\mu^\star = \lambda_{\mathbf{K}}$.

Using the infinite-dimensional convex duality method described in Section 2.1.4, the dual of LP (2.8a$'$) reads

$$\begin{aligned} \mathrm{d}'_{\mathbf{K}} := \inf_{\mathbf{v},w}\ & \int w\ d\lambda \\ \text{s.t. } & w - \mathrm{div}(g\mathbf{v}) - 1 \in \mathcal{C}(\mathbf{K})_+ \\ & w \in \mathcal{C}(\mathbf{X})_+ \\ & \mathbf{v} \in \mathcal{C}^1(\mathbf{K})^n. \end{aligned} \tag{2.8b$'$}$$

A direct application of Theorem 2.6 yields that $\mathrm{d}'_{\mathbf{K}} = \mathrm{p}^\star_{\mathbf{K}} = \mathrm{d}^\star_{\mathbf{K}}$ (we already verified that the theorem's working assumptions hold in such case, see Example 2.10).

Crucial observation. Notice that w in (2.8b′) is *not* required to approximate $\mathbb{1}_{\mathbf{K}}$ from above anymore. Instead, it should approximate $1 + \mathrm{div}(g\mathbf{v})$ on $\mathbf{K}$ and 0 outside $\mathbf{K}$. Hence, provided that $1 + \mathrm{div}(g\mathbf{v}) = 0$ on $\partial\mathbf{K}$, w might be a continuous function for some well-chosen $\mathbf{v} \in \mathcal{C}^1(\mathbf{K})^n$, and therefore an optimal solution of (2.8b′) (i.e., the infimum is a minimum). As a result, the Gibbs phenomenon would disappear and convergence would be faster.

The issue is then to determine whether the infimum in (2.8b′) is attained or not. And if not, are there other special features of problem (2.8b′) that can be exploited to yield more efficient semidefinite relaxations ?

4.2.2 New Stokes constraints and main result

In the previous paragraph, the Stokes constraint

$$\int (\mathbf{v} \cdot \mathbf{grad}\ g + g\ \mathrm{div}\ \mathbf{v})\ d\mu = 0$$

or equivalently (in the sense of distributions)

$$(\mathbf{grad}\ g)\mu - \mathbf{grad}(g\mu) = \mathbf{0} \tag{4.7}$$

(with $\mu \in \mathcal{M}(\mathbf{K})_+$ being the Lebesgue measure on $\mathbf{K}$) was obtained as a particular case of Stokes' theorem with $\mathbf{u} = g\ \mathbf{v}$ in (4.4). Instead, we can use a more general version with $\mathbf{u}$ not in factored form, and also use the fact that

$$\forall \mathbf{x} \in \partial\mathbf{K}, \mathbf{0} \neq \mathbf{grad}\ g(\mathbf{x}) = -|\mathbf{grad}\ g(\mathbf{x})|\ \mathbf{n}_{\mathbf{K}}(\mathbf{x}),$$

to obtain

$$\int \mathrm{div}\ \mathbf{u}\ d\mu = -\int \mathbf{u} \cdot \mathbf{grad}\ g\ d\nu\,,$$

or equivalently (in the sense of distributions)

$$\mathbf{grad}\ \mu = (\mathbf{grad}\ g)\nu\,, \tag{4.8}$$

with $\mu \in \mathcal{M}(\mathbf{K})_+$ being the Lebesgue measure on $\mathbf{K}$ and $\nu \in \mathcal{M}(\partial\mathbf{K})_+$ being the measure having density $1/|\mathbf{grad}\ g(\mathbf{x})|$ with respect to the $(n-1)$-dimensional Haussdorff measure σ on $\partial\mathbf{K}$. The same linear equation was used in [73] to compute moments of the Hausdorff measure. In fact, equation (4.8) is a generalization of equation (4.7) in the following sense.

Lemma 4.4: Stokes constraints generalization

If $\nu \in \mathcal{M}(\partial\mathbf{K})_+$ is s.t. $\mu \in \mathcal{M}(\mathbf{K})_+$ satisfies (4.8), then μ also satisfies (4.7).

Proof : Equation (4.8) means that

$$\int \mathrm{div}\ \mathbf{u}\ d\mu + \int \mathbf{u} \cdot \mathbf{grad}\ g\ d\nu = 0$$

for all $\mathbf{u} \in \mathcal{C}^1(\mathbf{K})^n$. In particular if $\mathbf{u} = g\ \mathbf{v}$ for some $\mathbf{v} \in \mathcal{C}^1(\mathbf{K})^n$ then (4.8) reads

$$\int (\mathbf{v} \cdot \mathbf{grad}\ g + g\ \mathrm{div}\ \mathbf{v})\ d\mu = 0\,,$$

which is precisely (4.7). ♢

Hence we can incorporate linear constraints (4.8) on μ and ν, to rewrite problem (2.8a) as

$$\begin{aligned} \mathrm{p}^\star_{\mathbf{K}} = \max_{\mu,\nu}\ & \int 1\ d\mu \qquad (2.8a'') \\ \text{s.t. } & \mu \in \mathcal{M}(\mathbf{K})_+ \\ & \nu \in \mathcal{M}(\partial\mathbf{K})_+ \\ & \lambda - \mu \in \mathcal{M}(\mathbf{X})_+ \\ & (\mathbf{grad}\ g)\nu - \mathbf{grad}\ \mu = \mathbf{0} \end{aligned}$$

without changing its optimal value $\mathrm{p}^\star_{\mathbf{K}} = \lambda(\mathbf{K})$ attained at $\mu^\star = \lambda_{\mathbf{K}}$ and $\nu^\star = \sigma/|\mathbf{grad}\ g|$. Notice that LP (2.8a″) involves *two* measures μ and ν whereas LP (2.8a′) involves only one measure μ.

Next, by convex duality, the dual of (2.8a″) reads

Problem 15: Stokes dual GMP

Find a minimizer for

$$\begin{aligned} \mathrm{d}''_{\mathbf{K}} = \inf_{\mathbf{u},w}\ & \int w\, d\lambda \qquad (2.8b'') \\ \text{s.t. } & w - \mathrm{div}\ \mathbf{u} - 1 \in \mathcal{C}(\mathbf{K})_+ \\ & -(\mathbf{u} \cdot \mathbf{grad}\ g) \in \mathcal{C}(\partial\mathbf{K})_+ \\ & w \in \mathcal{C}(\mathbf{X})_+ \\ & \mathbf{u} \in \mathcal{C}^1(\mathbf{K})^n. \end{aligned}$$

Our main result states that the optimal value of the dual (2.8b″) is attained at some continuous function $(w^\star, \mathbf{u}^\star) \in \mathcal{C}(\mathbf{X})_+ \times \mathcal{C}^1(\mathbf{K})^n$. Therefore, in contrast with problem (2.8b), there is no Gibbs phenomenon at an optimal solution of the (finite-dimensional) semidefinite strengthening associated with (2.8b″). On the contrary, existence of a regular optimum yields at least one sequence of polynomials that converge uniformly to a global optimizer, due to the Stone-Weierßtrass theorem.

Let $\mathbf{\Omega}_i$, $i = 1, \ldots, N$ denote the connected components of $\mathbf{\Omega}$, and let

$$m_{\mathbf{\Omega}_i}(g) := \frac{1}{\lambda(\mathbf{\Omega}_i)} \int_{\mathbf{\Omega}_i} g(\mathbf{x})\ \mathrm{d}\mathbf{x}$$

be the *mean* of g on $\mathbf{\Omega}_i$.

Theorem 4.5: Optimum for dual Stokes volume problem

In dual LP (2.8b″) the infimum is a minimum, attained at

$$w^\star(\mathbf{x}) := g(\mathbf{x}) \sum_{i=1}^{N} \frac{\mathbb{1}_{\mathbf{\Omega}_i}(\mathbf{x})}{m_{\mathbf{\Omega}_i}(g)}, \quad \mathbf{x} \in \mathbf{X},$$

and

$$\mathbf{u}^\star(\mathbf{x}) := \mathbf{grad}\ u(\mathbf{x}),$$

where u solves the Poisson PDE

$$\begin{cases} -\Delta u(\mathbf{x}) &= 1 - w^\star(\mathbf{x}), & \mathbf{x} \in \mathbf{\Omega} \\ \partial_{\mathbf{n}} u(\mathbf{x}) &= 0, & \mathbf{x} \in \partial\mathbf{\Omega}. \end{cases}$$

Remark 4.2 (Computing lower bounds for the volume)

The moment-SOS hierarchy associated to LPs (2.8a″) *and* (2.8b″) *yields upper bounds for the volume. Theorem 4.5 is designed for these LPs but it has a straightforward counterpart for lower bound volume computation, obtained by replacing* $\mathbf{K}$ *with* $\mathbf{X} \setminus \mathbf{\Omega}$ *in the previous developments, i.e. computing upper bounds of* $\lambda(\mathbf{X} \setminus \mathbf{\Omega})$*. However, two additional technicalities should then be considered:*

- *This work only deals with semi-algebraic sets defined by a single polynomial; actually, it immediately generalizes to finite intersections of such semi-algebraic sets, as long as their boundaries do not intersect (i.e. here* $\mathbf{K}$ *should be included in the interior of* $\mathbf{X}$*): the constraints on boundaries should just be splitted between the boundaries of the intersected sets.*
- *This work heavily relies on the fact that the boundary of the considered set should be smooth; for this reason, computing lower bounds of the volume implies that one chooses a smooth bounding box* $\mathbf{X}$ *(typically a euclidean ball, ellipsoid or* ℓ^p *ball), which rules out simple sets like the hypercube* $[-1, 1]^n$*.*

Upon taking into account these technicalities, Theorem 4.5 still holds, allowing to deterministically compute upper and lower bounds for the volume, with arbitrary precision. Of course in practice, one is limited by the performance of state-of-art SDP solvers.

4.3 Solving a PDE to attain an optimum

Theorem 4.5 is proved in several steps as follows:

- we show that solutions to a Poisson PDE are optimal for (2.8b″);
- we study the Poisson PDE on a connected domain;
- we study the Poisson PDE on a union of connected domains;
- we construct an explicit optimum for problem (2.8b″).

4.3.1 Equivalence to a Poisson PDE

Lemma 4.6: Optimality condition

Problem (2.8b″) has an optimal solution *iff* there exist $\mathbf{u} \in \mathcal{C}^1(\overline{\mathbf{\Omega}})^n$, $\widetilde{w} \in \mathcal{C}(\overline{\mathbf{\Omega}})_+$ solving

$$
\begin{aligned}
-\operatorname{div} \mathbf{u} &= 1 - \widetilde{w} && \text{in } \mathbf{\Omega}, && (4.9a)\\
\mathbf{u} \cdot \mathbf{n} &= 0 && \text{on } \partial\mathbf{\Omega}, && (4.9b)\\
\widetilde{w} &= 0 && \text{on } \partial\mathbf{\Omega}. && (4.9c)
\end{aligned}
$$

Proof : Let $(\mathbf{u}, \widetilde{w})$ solve (4.9). Using (4.9c), one can define

$$
w(\mathbf{x}) = \begin{cases} \widetilde{w}(\mathbf{x}) & \text{if } \mathbf{x} \in \overline{\mathbf{\Omega}} \\ 0 & \text{if } \mathbf{x} \in \mathbf{X} \setminus \overline{\mathbf{\Omega}}. \end{cases}
$$

Then $(\mathbf{u}, w)$ is feasible for (2.8b″) and one has

$$
\begin{aligned}
\int w \, d\lambda &= \int \widetilde{w} \, d\lambda \\
&\overset{(4.9a)}{=} \int_{\mathbf{\Omega}} (1 + \operatorname{div} \mathbf{u}) \, d\lambda \\
&\overset{(4.4)}{=} \lambda(\mathbf{\Omega}) + \int_{\partial\mathbf{\Omega}} \mathbf{u} \cdot \mathbf{n} \, d\sigma \\
&\overset{(4.9b)}{=} \lambda(\mathbf{\Omega})
\end{aligned}
$$

so that $(\mathbf{u}, w)$ is optimal.

Conversely, let $(\mathbf{u}, w)$ be an optimal solution of problem (2.8b″). We know that $(\mu^\star, \nu^\star) = (\lambda_{\mathbf{K}}, \sigma/|\mathbf{grad}\ g|)$ is optimal for problem (2.8a″). Then, duality theory ensures complementarity:

$$
\int (w - \operatorname{div} \mathbf{u} - 1) \, d\lambda_{\mathbf{K}} = 0, \tag{4.10a}
$$

$$
\int \mathbf{u} \cdot \frac{\mathbf{grad}\ g}{|\mathbf{grad}\ g|} \, d\sigma = 0. \tag{4.10b}
$$

Since $w - \operatorname{div} \mathbf{u} - 1$ is nonnegative on $\mathbf{K}$, (4.10a) yields (4.9a) with $\widetilde{w} := w|_{\mathbf{\Omega}}$. Likewise, since $-(\mathbf{u} \cdot \mathbf{grad}\ g)$ is nonnegative on $\partial\mathbf{\Omega}$, (4.10b) yields (4.9b) and thus, using (4.4), one has $\int_{\mathbf{\Omega}} \operatorname{div} \mathbf{u} \, d\lambda = 0$. Eventually, (4.10a) yields $\int_{\mathbf{\Omega}} w \, d\lambda = \lambda(\mathbf{\Omega}) = \int w \, d\lambda$ by optimality of w, so that $\int_{\mathbf{X}\setminus\mathbf{\Omega}} w \, d\lambda = 0$ and, since w is nonnegative, $w|_{\mathbf{X}\setminus\mathbf{\Omega}} \equiv 0$. Continuity of w finally allows us to conclude that $w = 0$ on $\partial\mathbf{\Omega}$, which is exactly (4.9c). ♢

From Lemma 4.6, existence of an optimum for (2.8b″) is then equivalent to existence of a solution to (4.9), which we rephrase as follows, defining $f := 1 - \widetilde{w}$ and $\mathbf{u} = \mathbf{grad}\ u$ with $u \in \mathcal{C}^2(\overline{\mathbf{\Omega}})$, and where $\Delta u := \operatorname{div} \mathbf{grad}\ u$ is the Laplacian of u, and $\partial_{\mathbf{n}} u := \mathbf{grad}\ u \cdot \mathbf{n}$.

Lemma 4.7: Optimum as solution to Poisson's PDE

If there exist $u \in \mathcal{C}^2(\overline{\mathbf{\Omega}})^n$ and $f \in \mathcal{C}(\overline{\mathbf{\Omega}})$ solving

$$\begin{aligned} -\Delta u &= f && \text{in } \mathbf{\Omega}, && (4.11a)\\ \partial_{\mathbf{n}} u &= 0 && \text{on } \partial\mathbf{\Omega}, && (4.11b)\\ f &\leq 1 && \text{in } \mathbf{\Omega}, && (4.11c)\\ f &= 1 && \text{on } \partial\mathbf{\Omega}, && (4.11d) \end{aligned}$$

then problem (2.8b″) has an optimal solution.

This rephrazing is a Poisson PDE (4.11a) with Neumann boundary condition (4.11b), whose source term f is a parameter subject to constraints (4.11c) and (4.11d).

Remark 4.3 (Loss of generality)
Looking for $\mathbf{u}$ *under the form* $\mathbf{u} = \mathbf{grad}\ u$ *makes us loose the equivalence. Indeed, while* (2.8b″) *and* (4.9) *are equivalent, existence of a solution to* (4.11) *is only a sufficient condition for existence of an optimum for* (2.8b″)*, since* (4.9) *might have only solutions* $\mathbf{u}$ *that are not gradients.*

Remark 4.4 (Invariant set for gradient flow)
From a dynamical systems point of view, constraint (4.11b) *means that we are looking for a velocity field or control* $\mathbf{u}$ *in the form of the gradient of a potential* u *such that* $\overline{\mathbf{\Omega}}$ *is an invariant set for the solutions* $t \in \mathbb{R} \mapsto \mathbf{x}(t) \in \mathbb{R}^n$ *of the Cauchy problem*

$$\dot{\mathbf{x}}(t) = -\mathbf{grad}\ u(\mathbf{x}(t)), \quad \mathbf{x}(0) \in \mathbf{X}$$

after what we just have to define $\widetilde{w} := 1 + \Delta u$ *on* $\mathbf{\Omega}$ *and enforce constraints* (4.11a), (4.11c), (4.11d).

4.3.2 Poisson PDE on a connected domain

It remains to prove existence of solutions to problem (4.11). First, notice that PDE (4.11a) together with its boundary condition (4.11b) enforces an important constraint on the source term f, namely its mean must vanish:

$$\int_{\mathbf{\Omega}} f \, d\lambda = 0. \tag{4.12}$$

Indeed, if (f, u) solves (4.11), then

$$\begin{aligned} \int f \, d\lambda &\overset{(4.11a)}{=} -\int \Delta u \, d\lambda \\ &\overset{(4.4)}{=} -\int \mathbf{grad}\ u \cdot \mathbf{n} \, d\sigma \\ &= -\int \partial_{\mathbf{n}} u \, d\sigma \overset{(4.11b)}{=} 0\,. \end{aligned}$$

Moreover, the following holds.

Lemma 4.8: Existence on a connected domain

Suppose that $\mathbf{\Omega}$ is connected. Let the source term $f \in L^2(\mathbf{\Omega}) \cap \mathcal{C}^\infty(\mathbf{\Omega})$ have zero mean on $\mathbf{\Omega}$, where $L^p(\mathbf{\Omega}) := \{f \in \mathbb{R}^{\mathbf{\Omega}} : \int |f|^p \, d\lambda < \infty\}$. Then there exists $u \in \mathcal{C}^\infty(\mathbf{\Omega})$ satisfying (4.11a) and (4.11b).

Proof : We first recall some basic definitions:

- for $p \in \mathbb{R}_{++}$ and $f \in L^p(\mathbf{\Omega})$ the L^p norm is defined by

$$\|f\|_{L^p(\mathbf{\Omega})} := \left(\int_{\mathbf{\Omega}} |f|^p \, d\lambda\right)^{\frac{1}{p}}.$$

- The L^2 norm is a Hilbert norm with associated inner product

$$\langle u|v\rangle_{L^2(\mathbf{\Omega})} := \int_{\mathbf{\Omega}} u \, v \, d\lambda,$$

  such that $\|u\|_{L^2(\mathbf{\Omega})} = \sqrt{\langle u|u\rangle_{L^2(\mathbf{\Omega})}}$.

- The Sobolev Hilbert spaces are defined by

$$H^1(\mathbf{\Omega}) := \{u \in L^2(\mathbf{\Omega}) : \mathbf{grad}\ u \in L^2(\mathbf{\Omega})^n\}$$

  and by induction on $k > 1$,

$$H^k(\mathbf{\Omega}) := \{u \in L^2(\mathbf{\Omega}) : \mathbf{grad}\ u \in H^{k-1}(\mathbf{\Omega})^n\}.$$

- The Sobolev inner product defines a Hilbert norm:

$$\langle u|v\rangle_{H^1(\mathbf{\Omega})} := \langle u|v\rangle_{L^2(\mathbf{\Omega})} + \langle \mathbf{grad}\ u|\mathbf{grad}\ v\rangle_{L^2(\mathbf{\Omega})^n} \quad \& \quad \|u\|_{H^1(\mathbf{\Omega})} := \sqrt{\langle u|u\rangle_{H^1(\mathbf{\Omega})}}.$$

Now, let us rephrase the Poisson PDE with Neumann boundary condition under a variational form:

Problem 16: Variational form for (4.11a)-(4.11b)

Find $u \in H^1(\mathbf{\Omega})$ such that for any $v \in H^1(\mathbf{\Omega})$ one has

$$\int_{\mathbf{\Omega}} \mathbf{grad}\ u \cdot \mathbf{grad}\ v \, d\lambda = \int_{\mathbf{\Omega}} f \, v \, d\lambda. \tag{4.13}$$

Then, for such u, since $f \in L^2(\mathbf{\Omega})$, the interior H^2-regularity theorem (see [33, Theorem 1 in Section 6.3.1]) ensures that $u \in H^2_{loc}(\mathbf{\Omega}) := \cap_{\mathbf{Y} \Subset \mathbf{\Omega}} H^2(\mathbf{Y})$ ($\Subset$ standing for *compact* inclusion), and Green's "integration by part" theorem writes, for all $v \in H^1(\mathbf{\Omega})$:

$$\int_{\partial\mathbf{\Omega}} v \, \partial_{\mathbf{n}} u \, d\sigma = \int_{\mathbf{\Omega}} v \Delta u \, d\lambda + \int_{\mathbf{\Omega}} \mathbf{grad}\ u \cdot \mathbf{grad}\ v \, d\lambda = \int_{\mathbf{\Omega}} (\Delta u + f) \, v \, d\lambda.$$

Especially, for $v \in \mathcal{C}^\infty_c(\mathbf{\Omega}) \subset H^1(\mathbf{\Omega})$ the left hand side is zero, and by density of $\mathcal{C}^\infty_c(\mathbf{\Omega})$ in $L^2(\mathbf{\Omega})$, we deduce that $-\Delta u = f$ for the $L^2(\mathbf{\Omega})$ Hilbert topology and thus

almost everywhere in $\mathbf{\Omega}$. The left hand side is then zero for any $v \in H^1(\mathbf{\Omega})$ and especially for any $v \in \mathcal{C}^\infty(\overline{\mathbf{\Omega}}) \subset H^1(\mathbf{\Omega})$, so that, again by density of $\mathcal{C}^\infty(\partial\mathbf{\Omega})$ in $L^2(\mathbf{\Omega})$, one has $\partial_{\mathbf{n}} u = 0$ in $L^2(\partial\mathbf{\Omega})$ and then almost everywhere on $\partial\mathbf{\Omega}$.

Eventually, the interior $\mathcal{C}^\infty$-regularity theorem (see [33, Theorem 3 in section 6.3.1]) ensures that since $f \in \mathcal{C}^\infty(\mathbf{\Omega})$, $u \in \mathcal{C}^\infty(\mathbf{\Omega})$ and we obtain the announced result: u is a smooth strong solution of the Poisson PDE.

Next we invoke Lax-Milgram's theorem which provides existence and uniqueness of a solution to a given PDE (see e.g. [33, Section 6.2.1]). In our context the goal is to solve (4.13) for which it is clear that if u is a solution then any $\hat{u} := u + C$, $C \in \mathbb{R}$ is also solution, which makes it impossible to obtain uniqueness of the solution in $H^1(\mathbf{\Omega})$. We thus restrict ourselves to the hyperplane of zero-mean functions

$$\mathcal{H} := \left\{ u \in H^1(\mathbf{\Omega}) : \int_{\mathbf{\Omega}} u \, d\lambda = 0 \right\}$$

which is closed by continuity of the Lebesgue integral, so that $\mathcal{H}$ is a Hilbert space for the scalar product

$$\langle u|v\rangle_{\mathcal{H}} := \langle u|v\rangle_{H^1(\mathbf{\Omega})} = \int_{\mathbf{\Omega}} (uv + \mathbf{grad}\ u \cdot \mathbf{grad}\ v)\, d\lambda.$$

We then define the applications

$$\mathcal{A} : \left\{ \begin{array}{ccc} \mathcal{H} & \longrightarrow & \mathbb{R} \\ v & \longmapsto & \displaystyle\int_{\mathbf{\Omega}} f\ v\ d\lambda. \end{array} \right.$$

and

$$\mathcal{B} : \left\{ \begin{array}{ccc} \mathcal{H} \times \mathcal{H} & \longrightarrow & \mathbb{R} \\ (u, v) & \longmapsto & \displaystyle\int_{\mathbf{\Omega}} \mathbf{grad}\ u \cdot \mathbf{grad}\ v\ d\lambda \end{array} \right.$$

The Lax-Milgram theorem then states that if $\mathcal{A}$ and $\mathcal{B}$ are continuous and if $\mathcal{B}$ is moreover coercive, then there is a unique $u \in \mathcal{H}$ so that $\mathcal{A} = \mathcal{B}(u, \cdot)$, which is the announced equality (4.13). Let us show that these hypotheses are met.

- Continuity of $\mathcal{A}$. Since $\mathcal{A}$ is a linear operator, it is sufficient to show that it is bounded. Let $v \in \mathcal{H}$. Then, Hölder's inequality yields

$$\begin{aligned} |\mathcal{A}(v)| &= \left| \int_{\mathbf{\Omega}} f\ v\ d\lambda \right| \\ &\leq \|f\|_{L^2(\mathbf{\Omega})} \|v\|_{L^2(\mathbf{\Omega})} \\ &\leq \|f\|_{L^2(\mathbf{\Omega})} \|v\|_{\mathcal{H}} \end{aligned}$$

  because $\|v\|_{\mathcal{H}} = \sqrt{\|v\|^2_{L^2(\mathbf{\Omega})} + \|\mathbf{grad}\ v\|^2_{L^2(\mathbf{\Omega})}} \geq \|v\|_{L^2(\mathbf{\Omega})}$. Thus, $\mathcal{A}$ is a bounded operator and $|||\mathcal{A}||| = \|f\|_{L^2(\mathbf{\Omega})}$ (equality is obtained by taking $v = f \in \mathcal{H}$, made possible by (4.12)).

- Continuity of $\mathcal{B}$. Since $\mathcal{B}$ is a bilinear operator, it is sufficient to show that it is bounded. Let $u, v \in \mathcal{H}$. Again, Hölder's inequality yields

$$\begin{aligned} |\mathcal{B}(u,v)| &= \left| \int_{\mathbf{\Omega}} \mathbf{grad}\ u \cdot \mathbf{grad}\ v\ d\lambda \right| \\ &\leq \|\mathbf{grad}\ u\|_{L^2(\mathbf{\Omega})} \|\mathbf{grad}\ v\|_{L^2(\mathbf{\Omega})} \\ &\leq \|\mathbf{grad}\ u\|_{\mathcal{H}} \|\mathbf{grad}\ v\|_{\mathcal{H}} \end{aligned}$$

because $\|v\|_{\mathcal{H}} = \sqrt{\|v\|^2_{L^2(\mathbf{\Omega})} + \|\mathbf{grad}\ v\|^2_{L^2(\mathbf{\Omega})}} \geq \|\mathbf{grad}\ v\|_{L^2(\mathbf{\Omega})}$. Then, $\mathcal{B}$ is a bounded bilinear operator.

- Coercivity of $\mathcal{B}$. First, let us recall the following classical result, proved e.g. in [33, Theorem 1 in Section 5.8.1].

Lemma 4.9: Poincaré-Wirtinger inequality

Let $\mathbf{\Omega} \subset \mathbb{R}^n$ be a bounded, connected, $\mathcal{C}^1$ open set. There is a constant $C_{\mathbf{\Omega}} \geq 0$ such that for any $u \in H^1(\mathbf{\Omega})$:

$$\|u - m_{\mathbf{\Omega}}(u)\|_{L^2(\mathbf{\Omega})} \leq C_{\mathbf{\Omega}} \|\mathbf{grad}\ u\|_{L^2(\mathbf{\Omega})^n}$$

where $m_{\mathbf{\Omega}}(u) := \frac{1}{\lambda(\mathbf{\Omega})} \int_{\mathbf{\Omega}} u\ d\lambda$.

Now let us look for a coercivity constant $C \in \mathbb{R}$ such that $\|u\|_{\mathcal{H}} \leq C\ \mathcal{B}(u,u)$. Let $u \in \mathcal{H}$. Then, since $\partial\mathbf{\Omega}$ is $\mathcal{C}^1$, we can use Lemma 4.9:

$$\|\mathbf{grad}\ u\|_{L^2(\mathbf{\Omega})^n} + \|u - m_{\mathbf{\Omega}}(u)\|_{L^2(\mathbf{\Omega})} \leq (1 + C_{\mathbf{\Omega}}) \|\mathbf{grad}\ u\|_{L^2(\mathbf{\Omega})^n}$$

and thus

$$\begin{aligned} \mathcal{B}(u,u) &= \int_{\mathbf{\Omega}} |\mathbf{grad}\ u|^2\ d\lambda \\ &= \|\mathbf{grad}\ u\|^2_{L^2(\mathbf{\Omega})^n} \\ &\geq \frac{\|\mathbf{grad}\ u\|^2_{L^2(\mathbf{\Omega})^n} + \|u - m_{\mathbf{\Omega}}(u)\|^2_{L^2(\mathbf{\Omega})}}{1 + C^2_{\mathbf{\Omega}}} \\ &= \frac{\|u\|^2_{\mathcal{H}}}{1 + C^2_{\mathbf{\Omega}}} \end{aligned}$$

since $u \in \mathcal{H}$ implies that $m_{\mathbf{\Omega}}(u) := \frac{1}{\lambda(\mathbf{\Omega})} \int_{\mathbf{\Omega}} u\ d\lambda = 0$.

Thus, the conditions of the Lax-Milgram theorem are satisfied, which gives us a unique $u \in \mathcal{H}$ such that for all $v \in \mathcal{H}$, equation (4.13) holds. To conclude, we still need to extend this property to functions v that have nonzero mean. Let $v \in H^1(\mathbf{\Omega})$, not necessary in $\mathcal{H}$. We define $\hat{v} := v - m_{\mathbf{\Omega}}(v)$, so that $\hat{v} \in \mathcal{H}$ and $\mathbf{grad}\ v = \mathbf{grad}\ \hat{v}$.

Then,

$$\begin{aligned}\int_{\mathbf{\Omega}} f\ v\ d\lambda &= \int_{\mathbf{\Omega}} f\ \hat{v}\ d\lambda + m_{\mathbf{\Omega}}(v) \int_{\mathbf{\Omega}} f\ d\lambda \\ &\overset{(4.12)}{=} \int_{\mathbf{\Omega}} f\ \hat{v}\ d\lambda \\ &\overset{(4.13)}{=} \int_{\mathbf{\Omega}} \mathbf{grad}\ u \cdot \mathbf{grad}\ \hat{v}\ d\lambda \\ &= \int_{\mathbf{\Omega}} \mathbf{grad}\ u \cdot \mathbf{grad}\ v\ d\lambda,\end{aligned}$$

which concludes the solution of the variational formulation and the proof of Lemma 4.8. ♢

4.3.3 General Poisson PDE with boundary regularity

In Lemma 4.8, we assumed that $\mathbf{\Omega}$ is connected, so that we could apply the Poincaré-Wirtinger inequality to use the Lax-Milgram theorem, obtaining both existence and uniqueness of a solution in a well-chosen space. However, we are not interested in the uniqueness property, and we would like to tackle non-connected sets. Since $\mathbf{\Omega}$ is a semi-algebraic set, it has a finite number of connected components $\mathbf{\Omega}_1, \dots, \mathbf{\Omega}_N$.

Corollary 4.10: Existence on a disconnected domain

Let the source term $f \in L^2(\mathbf{\Omega}) \cap \mathcal{C}^\infty(\mathbf{\Omega})$ have zero mean on each connected component of $\mathbf{\Omega}$. Then there exists $u \in \mathcal{C}^\infty(\mathbf{\Omega})$ solving (4.11a) and(4.11b).

Proof : Let $i = 1, \dots, N$. Since $\int_{\mathbf{\Omega}_i} f\ d\lambda = 0$, we can apply the result of Lemma 4.8 replacing $\mathbf{\Omega}$ with $\mathbf{\Omega}_i$ to obtain $u_i \in \mathcal{C}^\infty(\mathbf{\Omega}_i)$ such that $-\Delta u_i = f$ in $\mathbf{\Omega}_i$ and $\partial_{\mathbf{n}} u_i = 0$ on $\partial\mathbf{\Omega}_i$.

Then, we notice that since $\partial\mathbf{\Omega}$ is $\mathcal{C}^1$, the $\mathbf{\Omega}_i$ cannot be mutually tangent, so that $\partial\mathbf{\Omega} = \bigsqcup_{i=1}^N \partial\mathbf{\Omega}_i$. Thus, for any $\mathbf{x} \in \mathbf{\Omega}$, the following sum has exactly one non-zero term:

$$u := \sum_{i=1}^{N} \mathbb{1}_{\overline{\mathbf{\Omega}}_i} u_i.$$

By definition of the $\mathbf{\Omega}_i$ as the connected components of $\mathbf{\Omega}$, $u \in \mathcal{C}^\infty(\mathbf{\Omega})$.

Let $\mathbf{x} \in \mathbf{\Omega}$. There is an i such that $\mathbf{x} \in \mathbf{\Omega}_i$, so that $u = u_i$ on a neighbourhood of $\mathbf{x}$. Thus, $-\Delta u(\mathbf{x}) = -\Delta u_i(\mathbf{x}) = f(\mathbf{x})$.

Let $\mathbf{x} \in \partial\mathbf{\Omega}$. There is an i such that $\mathbf{x} \in \partial\mathbf{\Omega}_i$, so that $u = u_i$ on a neighbourhood of $\mathbf{x}$ in $\mathbf{\Omega}$. Thus, $\partial_{\mathbf{n}} u(\mathbf{x}) = \partial_{\mathbf{n}} u_i(\mathbf{x}) = 0$. ♢

In the previous paragraph we have proved that under suitable conditions on the source term f, equations (4.11a) and (4.11b) have a solution $u \in \mathcal{C}^\infty(\mathbf{\Omega})$. However, the existence of an optimum for problem (2.8b″) requires u to be in $\mathcal{C}^1(\mathbf{K})$: we need to establish regularity at the boundary. For this, an additional assumption on $\mathbf{\Omega}$ is needed to state the following corollary to Lemma 4.8.

Lemma 4.11: Regularity on the boundary

Let the source term $f \in \mathcal{C}^\infty(\overline{\mathbf{\Omega}})$ have zero mean on each connected component of $\mathbf{\Omega}$. Suppose that $\partial\mathbf{\Omega}$ is $\mathcal{C}^\infty$. Then, there exists $u \in \mathcal{C}^\infty(\overline{\mathbf{\Omega}})$ solving (4.11a) and (4.11b).

Proof : First, since $\mathcal{C}^\infty(\overline{\mathbf{\Omega}}) \subset L^2(\mathbf{\Omega}) \cap \mathcal{C}^\infty(\mathbf{\Omega})$, we can use Lemma 4.10 to get a suitable $u \in \mathcal{C}^\infty(\mathbf{\Omega})$. The only thing that remains to be proved is the regularity of u on $\partial\mathbf{\Omega}$. For this, we use the boundary $\mathcal{C}^\infty$-regularity theorem [33, Theorem 6 in Section 6.3.2]: since $f \in \mathcal{C}^\infty(\overline{\mathbf{\Omega}})$ and $\partial\mathbf{\Omega}$ is $\mathcal{C}^\infty$, we conclude that $u \in \mathcal{C}^\infty(\overline{\mathbf{\Omega}})$. ♢

Remark 4.5 (Regularity of Ω)

Assuming that $\partial\mathbf{\Omega}$ is $\mathcal{C}^\infty$ instead of $\mathcal{C}^1$ is actually without loss of generality since $\mathbf{\Omega}$ is a semi-algebraic set: as soon as $\partial\mathbf{\Omega}$ is locally the graph of a $\mathcal{C}^1$ function, it is smooth.

4.3.4 Explicit optimum for Stokes-enhanced hierarchy

Our optimization problem does not feature only the Poisson PDE with Neumann condition: it also includes constraints (4.11c) and (4.11d) on the source term. Consequently, a function $f \in \mathcal{C}^\infty(\overline{\mathbf{\Omega}})$ with zero integral over any connected component of $\mathbf{\Omega}$ and satisfying (4.11c) and (4.11d) remains to be constructed. We keep the notations of Lemma 4.10 and suggest as candidate

$$\mathbf{x} \mapsto f(\mathbf{x}) := 1 - g(\mathbf{x}) \sum_{i=1}^{N} \frac{\mathbb{1}_{\mathbf{\Omega}_i}(\mathbf{x})}{m_{\mathbf{\Omega}_i}(g)}. \tag{4.14}$$

By definition, $g = 0$ on $\partial\mathbf{\Omega}$, so that (4.11d) automatically holds. Moreover, both g and $\mathbb{1}_{\mathbf{\Omega}_i}$ are nonnegative on $\mathbf{\Omega}$, so that (4.11c) also holds.

In terms of regularity, f is polynomial on each connected component of $\mathbf{\Omega}$ and since g smoothly vanishes on $\partial\mathbf{\Omega}$, $f \in \mathcal{C}^\infty(\mathbf{K})$.

Eventually, let $i \in 1, \dots, N$ so that $\mathbf{\Omega}_i$ is a connected component of $\mathbf{\Omega}$. Then, by definition, $\partial\mathbf{\Omega}_i \subset \partial\mathbf{\Omega}$, and one has

$$\begin{aligned} \int_{\mathbf{\Omega}_i} f \, d\lambda &= \int_{\mathbf{\Omega}_i} \left(1 - g(\mathbf{x}) \sum_{j=1}^{N} \frac{\mathbb{1}_{\mathbf{\Omega}_j}(\mathbf{x})}{m_{\mathbf{\Omega}_j}(g)}\right) \mathrm{d}\mathbf{x} \\ &= \lambda(\mathbf{\Omega}_i) - \frac{1}{m_{\mathbf{\Omega}_i}(g)} \int_{\mathbf{\Omega}_i} g(\mathbf{x}) \, \mathrm{d}\mathbf{x} \;=\; 0, \end{aligned}$$

since by definition $m_{\mathbf{\Omega}_i}(g) = \frac{1}{\lambda(\mathbf{\Omega}_i)} \int_{\mathbf{\Omega}_i} g(\mathbf{x}) \, \mathrm{d}\mathbf{x}$.

We finally obtain our couple (u, f) solution to problem (4.11) with f defined in (4.14) and u given by Lemma 4.11. Then we retrieve the couple $(\mathbf{u}, \widetilde{w})$ solution to problem (4.9) by defining $\mathbf{u} := \mathbf{grad}\ u$ and for all $\mathbf{x} \in \mathbf{\Omega}$:

$$\widetilde{w}(\mathbf{x}) := 1 - f(\mathbf{x}) = g(\mathbf{x}) \sum_{i=1}^{N} \frac{\mathbb{1}_{\mathbf{\Omega}_i}(\mathbf{x})}{m_{\mathbf{\Omega}_i}(g)}.$$

Eventually, the optimization problem (2.8b$''$) has a (global) minimizer $(\mathbf{u}, w)$ with, for all $\mathbf{x} \in \mathbf{X}$,

$$w(\mathbf{x}) = g(\mathbf{x}) \sum_{i=1}^{N} \frac{\mathbb{1}_{\mathbf{\Omega}_i}(\mathbf{x})}{m_{\mathbf{\Omega}_i}(g)}.$$

Indeed, one can check that

$$\begin{aligned} \int w \, d\lambda \; &= \sum_{i=1}^{N} \frac{1}{m_{\mathbf{\Omega}_i}(g)} \int_{\mathbf{\Omega}_i} g \, d\lambda \\ &= \sum_{i=1}^{N} \lambda(\mathbf{\Omega}_i) \; = \; \lambda(\mathbf{\Omega}) \; = \; \lambda(\mathbf{K}) \,, \end{aligned}$$

which concludes the proof of Theorem 4.5.

4.4 Numerical experiments and general heuristics

To illustrate how efficient can be the introduction of Stokes constraints for volume computation, we consider the simple setting where $\mathbf{K}$ is a Euclidean ball included in $\mathbf{X} = \mathbf{B}$ the unit Euclidean ball. Indeed drastic improvements on the convergence are observed. All numerical examples were processed on a standard laptop computer under the Matlab environment with the SOS parser of YALMIP [78], the moment parser GloptiPoly [43] and the semidefinite programming solver of MOSEK [26].

4.4.1 Practical implementation

Following the Moment-SOS hierarchy methodology for volume computation as described in [44], in the (finite-dimensional) degree d semidefinite strengthening of dual problem (2.8b$''$):

- $w \in \mathbb{R}_d[\mathbf{x}]$ and $\mathbf{u} \in \mathbb{R}_d[\mathbf{x}]^n$ are polynomials of degree at most d;
- the positivity constraint $w \in C^0(\mathbf{B})_+$ is replaced with a Putinar certificate of positivity on $\mathbf{B}$, that is:

  $$w(\mathbf{x}) \; = \; s_{10}(\mathbf{x}) + s_{11}(\mathbf{x})(1 - |x|^2) \,, \quad \forall \mathbf{x} \in \mathbb{R}^n \,,$$

  where s_{10} (resp. s_{11}) is an SOS polynomial of degree at most $2d$ (resp. $2d-2$);
- the positivity constraint $w - \text{div}\mathbf{u} - 1 \in C^0(\mathbf{K})_+$ is replaced with a Putinar certificate of positivity on $\mathbf{K}$, that is:

  $$w(\mathbf{x}) - \text{div } \mathbf{u}(\mathbf{x}) - 1 \; = \; s_{20}(\mathbf{x}) + s_{21}(\mathbf{x}) \, g(\mathbf{x}) \,, \quad \forall \mathbf{x} \in \mathbb{R}^n \,,$$

  where s_{20} (resp. s_{21}) is an SOS polynomial of degree at most $2d$ (resp. $2d - \text{d}^\circ g$);
- the positivity constraint $\mathbf{u} \cdot \mathbf{grad} g \in C^0(\partial\mathbf{K})_+$ is replaced with a Putinar certificate of positivity on $\partial\mathbf{K}$, that is:

  $$-\mathbf{u}(\mathbf{x}) \cdot \mathbf{grad} g(\mathbf{x}) \; = \; s_{30}(\mathbf{x}) + s_{31}(\mathbf{x}) \, g(\mathbf{x}) \,, \quad \forall \mathbf{x} \in \mathbb{R}^n \,,$$

  where s_{30} (resp. s_{31}) is an SOS polynomial of degree at most $2d$ (resp. $2d - \text{d}^\circ g$);

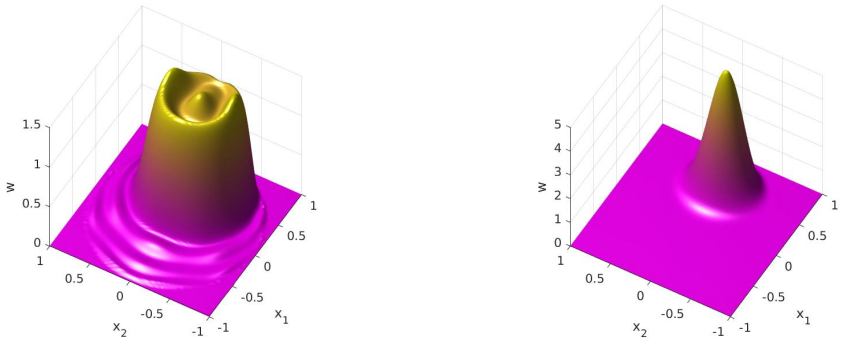


Figure 4.2 – Polynomials obtained with and without Stokes constraints.

Here are represented the degree 16 polynomial approximations obtained without Stokes constraints (left) and with Stokes constraints (right).

- the linear criterion $\int_{\mathbf{B}} w d\lambda$ translates into linear criterion on the vector of coefficients of w, as $\int_{\mathbf{B}} \mathbf{x}^\alpha \, d\lambda$ is available in closed-form.

The above identities define linear constraints on the coefficients of all the uknown polynomials. Next, stating that some of these polynomials must be SOS translate into semidefinite constraints on their respective unknown Gram matrices. The resulting optimization problem is a semidefinite program; for more details the interested reader is referred to e.g. [44].

4.4.2 Bivariate disk

Let us first illustrate Theorem 4.5 for computing the area of the disk $\mathbf{K} := \{\mathbf{x} \in \mathbb{R}^2 : g(\mathbf{x}) = 1/4 - (x_1 - 1/2)^2 - x_2^2 \geq 0\}$ included in the unit disk $\mathbf{B} := \{\mathbf{x} \in \mathbb{R}^2 : 1 - x_1^2 - x_2^2 \geq 0\}$.

The degree $d = 16$ polynomial approximation w obtained by solving the SOS relaxation of linear problem (2.8b) is represented at the left of Figure 4.2. We can see bumps and ripples typical of a Gibbs phenomenon, since the polynomial should approximate from above the discontinuous indicator function $\mathbb{1}_{\mathbf{K}}$ as closely as possible. A rather loose upper bound of 1.1626 is obtained on the volume $\lambda(\mathbf{K}) = \frac{\pi}{4} \approx 0.7854$.

In comparison, the degree $d = 16$ polynomial approximation w obtained by solving the SOS relaxation of linear problem (2.8b″) is represented at the right of Figure 4.2. As expected from the proof of Theorem 4.5, the poynomial should approximate from above the continuous function $g\mathbb{1}_{\mathbf{K}} \, \lambda(\mathbf{K})/(\int g \lambda_{\mathbf{K}})$. The resulting polynomial approximation is smoother and yields a much improved upper bound of 0.7870.

4.4.3 Higher dimensions

In Table 4.1 we report on the dramatic acceleration brought by Stokes constraints in the case of the Euclidean ball $\mathbf{K} := \{\mathbf{x} \in \mathbb{R}^3 : g(\mathbf{x}) = (3/4)^2 - |\mathbf{x}|^2 \geq 0\}$

n	d	without Stokes	with Stokes
3	4	88% (0.03s)	18% (0.04s)
3	8	57% (0.16s)	1.0% (0.44s)
3	12	47% (1.97s)	0.0% (4.63s)
3	16	43% (23.9s)	0.0% (30.1s)
3	20	41% (142s)	0.0% (206s)

Table 4.1 – Stokes constraints performances for increasing relaxation degrees.

Relative errors (%) and computational times (in brackets in seconds) for solving moment relaxations of increasing degrees d approximating the volume of ball of dimension $n = 3$.

n	d	without Stokes	with Stokes	n	d	without Stokes	with Stokes
1	10	17% (0.05s)	0.0% (0.03s)	6	4	190% (0.25s)	45.1% (1.03s)
2	10	35% (0.09s)	0.2% (0.25s)	7	4	203% (0.32s)	60.0% (4.88s)
3	10	56% (0.52s)	0.3% (1.19s)	8	4	221% (0.42s)	78.6% (8.45s)
4	10	72% (9.74s)	0.4% (22.8s)	9	4	245% (1.15s)	102% (45.1s)
5	10	79% (150s)	0.6% (669s)	10	4	278% (3.10s)	131% (176s)

Table 4.2 – Stokes constraints performances for increasing problem dimensions.

Relative errors (%) and computational times (in brackets in seconds) for solving the degree $d = 10$ (left) and $d = 4$ (right) moment relaxation approximating the volume of a ball of increasing dimensions n.

of dimension $n = 3$ included in the unit ball $\mathbf{B}$. We specify the relative errors on the bounds obtained by solving moment relaxations with and without Stokes constraints, together with the computational times (in seconds), for a relaxation degree d ranging from 4 to 20. We observe that tight bounds are obtained already at low degrees with Stokes constraints, sharply contrasting with the loose bounds obtained without Stokes constraints. However, we see also that the inclusion of Stokes constraints has a computational price.

In Table 4.2 we report the relative errors on the bounds obtained with and without Stokes constraints, together with the computational times (in seconds), for a relaxation degree equal to $d = 10$ (left) resp. $d = 4$ (right) and for dimension n ranging from 1 to 5 (left) resp. from 6 to 10 (right). When $d = 10$ and $n = 5$ the semidefinite relaxation features 6006 pseudo-moments without Stokes constraints, and 12194 pseudo-moments with Stokes constraints. We see that introducing Stokes constraints incurs a computational cost, to be compromised with the expected quality of the bounds.

Higher dimensional problems can be addressed only if the problem description has some sparsity structure, as explained in the next chapter. Also, depending on the geometry of the problem, and for larger values of the relaxation degree, alternative polynomial bases may be preferable numerically than the monomial basis which is used by default in Moment and SOS parsers.

4.4.4 General heuristics

As stated in Section 4.2.2, Problem (2.8a$''$) is a generalization of (2.8a$'$) in the sense of Lemma 4.4. However, such generalization is only valid when $\mathbf{K}$ is a smooth simple semialgebraic set $\mathbf{K} = \{\mathbf{x} \in \mathbb{R}^n : g(\mathbf{x}) \geq 0\}$, such that $-\mathbf{grad}\ g/|\mathbf{grad}\ g|$ is well defined on $\partial\mathbf{K}$ and coincides with $\mathbf{n}_\mathbf{K}$. In this section we propose some heuristics for the case of general compact basic semialgebraic sets:

$$\mathbf{K} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{g}(\mathbf{x}) \geq \mathbf{0}\}$$

with $\mathbf{g} \in \mathbb{R}[\mathbf{x}]^m$, $m > 1$. In such case, $\partial\mathbf{K}$ is only piecewise smooth:

$$\partial\mathbf{K} = \mathbf{K} \cap \left(\bigcup_{i=1}^{m} \{\mathbf{x} \in \mathbb{R}^n : g_i(\mathbf{x}) = 0\} \right) = \mathbf{K} \cap \{\mathbf{x} \in \mathbb{R}^n : h(\mathbf{x}) = 0\}$$

where $h(\mathbf{x}) := g_1(\mathbf{x}) \cdots g_m(\mathbf{x})$. In such case, it is straightforward to state a formulation similar to (2.8a$'$)-(2.8b$'$):

$$\begin{array}{lll} \mathrm{p}^\star_\mathbf{K} = \max\limits_{\mu} \int 1\, d\mu & & \mathrm{d}'_\mathbf{K} := \inf\limits_{\mathbf{v},w} \int w\, d\lambda \\ \quad \text{s.t. } \mu \in \mathcal{M}(\mathbf{K})_+ & & \quad \text{s.t. } w - \mathrm{div}(h\mathbf{v}) - 1 \in \mathcal{C}(\mathbf{K})_+ \\ \quad\quad \lambda - \mu \in \mathcal{M}(\mathbf{X})_+ & & \quad\quad w \in \mathcal{C}(\mathbf{X})_+ \\ \quad\quad (\mathbf{grad}\ h)\mu - \mathbf{grad}(h\mu) = 0, & & \quad\quad \mathbf{v} \in \mathcal{C}^1(\mathbf{K})^n, \end{array} \tag{4.15}$$

since $h = 0$ on $\partial\mathbf{K}$, using the same reasoning as in Section 4.2.1. In practice, this general formulation with h is the one that has always been used for compact basic semialgebraic sets. Implementation of this problem drastically improved the convergence of the corresponding moment-SOS hierarchies. However, Theorem 4.5 does not hold for this formulation. At this stage, two possibilities exist for future work:

- Extending Theorem 4.5 to formulation (2.8b$'$),
- Extending formulation 2.8b$''$ to general basic semialgebraic sets and extending Theorem 4.5 to the obtained generalization.

If one focuses on the first possibility, then the problem can be seen as the looking for $\widetilde{w} \in \mathcal{C}(\overline{\mathbf{\Omega}})_+$, $\mathbf{v} \in \mathcal{C}^1(\overline{\mathbf{\Omega}})^n$ such that:

$$\begin{array}{ll} -\mathrm{div}(h\mathbf{v}) = 1 - \widetilde{w} & \text{in } \mathbf{\Omega}, \\ \widetilde{w} = 0 & \text{on } \partial\mathbf{\Omega}, \end{array}$$

i.e. studying a degenerate linear PDE. While this PDE seems quite simple, so far we could not produce an analysis similar to the one presented in this chapter.

If one focuses on the second possibility, a key argument might be to keep the separation of $\partial\mathbf{K}$ as a union of basic semialgebraic sets

$$\partial\mathbf{K} = \bigcup_{i=1}^{m} \mathbf{K}_i \quad , \quad \mathbf{K}_i := \mathbf{K} \cap \{\mathbf{x} \in \mathbb{R}^n : g_i(\mathbf{x}) = 0\}$$

and to define a boundary measure ν_i for each one of them:

$$\begin{array}{lll} \mathrm{p}^\star_{\mathbf{K}} = \max \int 1\, d\mu & & \mathrm{d}''_{\mathbf{K}} = \inf \int w\, d\lambda \\ \quad \text{s.t. } \mu \in \mathcal{M}(\mathbf{K})_+ & & \quad \text{s.t. } w - \operatorname{div} \mathbf{u} - 1 \in \mathcal{C}(\mathbf{K})_+ \\ \quad\quad \nu_i \in \mathcal{M}(\mathbf{K}_i)_+, \quad i \in \mathbb{N}^\star_m & & \quad\quad -(\mathbf{u} \cdot \mathbf{grad}\ g_i) \in \mathcal{C}(\mathbf{K}_i)_+, \quad i \in \mathbb{N}^\star_m \\ \quad\quad \lambda - \mu \in \mathcal{M}(\mathbf{X})_+ & & \quad\quad w \in \mathcal{C}(\mathbf{X})_+ \\ \quad\quad \sum_{i=1}^{m} (\mathbf{grad}\ g_i)\nu_i - \mathbf{grad}\ \mu = 0, & & \quad\quad \mathbf{u} \in \mathcal{C}^1(\mathbf{K})^n. \end{array} \tag{4.16}$$

Then, extending Theorem 4.5 reduces to studying the Poisson PDE with Neumann boundary conditions on a non-smooth domain.

Conclusions

In this chapter we proposed a new primal-dual infinite-dimensional linear formulation of the problem of computing the volume of a smooth semi-algebraic set generated by a single polynomial, generalizing the approach of [44] while still allowing for the application of the moment-SOS hierarchy. The new dual formulation contains redundant linear constraints arising from Stokes' Theorem, generalizing the heuristic of [71]. A striking property of this new formulation is that the dual value is attained, contrary to the original formulation. As a consequence, the corresponding dual SOS hierarchy does not suffer from the Gibbs phenomenon, thereby accelerating the convergence.

Numerical experiments (not reported here) reveal that the values obtained with the new Stokes constraints (with a general vector field) are closely matching the values obtained with the original Stokes constraints of [71] (with the generating polynomial factoring the vector field). It may be then expected that the original and new Stokes constraints are equivalent, at least in some cases. However at this stage we have not been able to prove such equivalence. The crucial difference between Stokes formulations (2.8a′) and (4.7) is that the former, while allowing for a proof that the dual inf is attained, restricts to *simple* semialgebraic sets (defined by a single polynomial $g \in \mathbb{R}[\mathbf{x}]$), while the latter can be applied to general basic semialgebraic sets, but still lacks a proof of dual optimality attainment. Future works could focus on the extension of our proof to standard Stokes constraints (4.7) and (4.15), or to the most general Stokes constraints (4.16).

Eventually, the proof of dual attainment builds upon very classical tools from linear PDE analysis, thereby building up a new bridge between infinite-dimensional convex duality and PDE theory, in the context of the moment-SOS hierarchy. We

expect that these ideas can be exploited to prove regularity properties of linear reformulations of other problems in data science, beyond volume approximation. For example, it would be desirable to design Stokes constraints tailored to the infinite-dimensional linear reformulation of the region of attraction problem [42] or its sparse version [123].

5

Exploiting sparsity for volume computation

This chapter is based on contribution [126] which proposes a scheme to adapt the volume approximating moment-SOS hierarchy to a class of problem structures named sparsity patterns. Such sparsity patterns and their importance are presented in detail in Section 5.1, with two examples that illustrate both the stake of sparsity exploitation and the intuition that underlies our contribution. Then, Section 5.2 focuses on the simplest sparsity pattern we exploited to make the volume computation moment-SOS hierarchy more tractable, *i.e.* the path decomposition, with a variety of numerical examples, including high dimensional volume problems. This section is followed by a generalization in Section 5.3, which gives a computation scheme for general correlative sparsity patterns accompanied by another selection of numerical examples.

Contents

5.1 The importance of sparsity

5.1.1 Motivation

As stated in Section 3.1.2, the measure approach for the volume computation problem (2.8) is intimately linked to the use (3.7) of occupation measures, in dynamical systems theory, for computing the ROA of a given target set. Indeed, in [42], the problem of estimating the ROA is formulated as a GMP very similar to the volume computation problem. The idea is to maximize the volume of a set of initial conditions that yield trajectories ending in the target set after a given time.

This problem of estimating the ROA of a target set is crucial in power systems safety assessment, since the power grids must have good stability properties. The conservative, geometric characterization of the region of attraction as formulated in [42] is a very promising approach for this domain of application, see Section 3.1 and the results therein.

In both ROA estimation and volume computation, the main limitation of the moment-SOS method is that only problems of modest dimension can be handled by current solvers. Exploiting sparsity seems to be the best approach to allow scalability both in volume computation and ROA estimation. Since volume estimation is a simpler instance of the GMP than ROA estimation, we decided to address first the former problem.

In addition, volume computation with respect to a measure satisfying some conditions (e.g. compactly supported or Gaussian measure) also has applications in the fields of geometry and probability computation, which is the reason why many algorithms were already proposed for volume computation of convex polytopes and convex bodies in general.

5.1.2 Contribution

We design deterministic methods that provide approximations with strong asymptotic guarantees of convergence to the volume of $\mathbf{K}$. The methodology that we propose is similar in spirit to the one initially developed in [44] as described above and its extension to non-compact sets and Gaussian measures of [71]. However it is not a straightforward or direct extension of [44] or [71], and it has the following important distinguishing features:

(i) It can handle sets $\mathbf{K} \subset \mathbb{R}^n$ of potentially large dimension n provided that some sparsity pattern (namely: correlative sparsity, see section 5.1.4 as well as [132, 23] for details) is present in the description of $\mathbf{K}$. This is in sharp contrast with [44].

(ii) The computation of upper and lower bounds can be decomposed into smaller independent problems of the same type, and depending on the sparsity pattern, some of the computations can even be done in parallel. This fact alone is remarkable and unexpected.

To the best of our knowledge, this is the first deterministic method for volume computation that takes benefit from a correlative sparsity pattern in the description of $\mathbf{K}$ in the two directions of (a) decomposition into problems of smaller size and (b) parallel computation. Of course this sharp improvement is performed at some

price: our framework only works on semi-algebraic sets that present the appropriate correlative sparsity pattern (see Assumption 5.12 as well as Section 5.3.4 for detailed discussion on its applicability).

The key idea is to provide a new and very specific *sparse formulation* of the original problem in which one defines a set of marginal measures whose (small dimensional) support is in accordance with the correlative sparsity pattern present in the description of the set $\mathbf{K}$. However, those marginal measures are not similar to the ones used in the sparse formulation [69] of polynomial optimization problems over the same set $\mathbf{K}$. Indeed they are not expected to satisfy the consistency condition of [69][1].

Finally, in principle, our floating point volume computation in large dimension n is faced with a crucial numerical issue. Indeed as in Monte-Carlo methods, up to rescaling, one has to include the set $\mathbf{K}$ into a box $\mathbf{X}$ of unit volume. Therefore the volume of $\mathbf{K}$ is of the order ε^n for some $\varepsilon \in (0,1)$ and thus far beyond machine precision as soon as n is large. To handle this critical issue we develop a *sparsity-adapted rescaling* which allows us to compute very small volumes in potentially very high dimension with good precision.

5.1.3 A motivating example

Consider the following set

$$\mathbf{K} := \{\mathbf{x} \in [0,1]^{100} \ : \ \forall i \in \mathbb{N}^\star_{99},\ x_i x_{i+1} \leq 1/2\}.$$

This is a *high-dimensional non-convex sparse semi-algebraic set*. The precise definition of a sparse semi-algebraic set will be given later on, but so far notice that in the description of $\mathbf{K}$ each constraint involves only 2 variables out of 100. The volume of $\mathbf{K}$ is hard to compute, but thanks to the structured description of the set we are able to prove numerically that its volume is smaller than $2 \cdot 10^{-5}$ in less than 2 minutes on a standard computer.

For this we have implemented a specific version of the moment-SOS hierarchy of SDP relaxations to solve the GMP, in which we exploit the correlative sparsity pattern of the set $\mathbf{K}$. The basic idea is to replace the original GMP that involves an unknown measure on $\mathbb{R}^{100}$ (whose SDP relaxations are hence untractable) with a GMP involving 99 measures on $\mathbb{R}^2$ (hence tractable). In addition, this new GMP can be solved either in one shot (with the 99 unknown measures) or by solving sequentially 99 GMPs involving (i) one measure on $\mathbb{R}^2$ and (ii) some data obtained from the GMP solved at previous step. Our approach can be sketched as follows.

First, we rescale the problem so that the set $\mathbf{K}$ is included in the unit box $\mathbf{X} := [0,1]^n$ on which the moments of the Lebesgue measure are easily computed.

Next, we describe the volume problem on $\mathbf{K}$ as a chain of volume subproblems on the subspaces $\mathrm{Im}(\pi_i)$ where $\pi_i(x_1, \ldots, x_{100}) = (x_i, x_{i+1})$, with a link between the i-th and $(i+1)$-th sub-problems.

[1] If two measures share some variables then the consistency condition requires that their respective marginals coincide.

Finally, in this example, as $n = 100$ and $\mathbf{K} \subset \mathbf{X}$, the volume of $\mathbf{K}$ is very small and far below standard floating point machine precision. To handle this numerical issue, we have implemented a sparsity-adapted strategy which consists of rescaling each subproblem defined on the projections of $\mathbf{K}$ to obtain intermediate values all with the same order of magnitude. Once all computations (involving quantities of the same order of magnitude) have been performed, the correct value of the volume is obtained by a reverse scaling.

The sparse formulation stems from considering some measure marginals appropriately defined according to the correlative sparsity pattern present in the description of $\mathbf{K}$. It leads to a variety of algorithms to compute the volume of sparse semi-algebraic sets.

In this section we describe the method in the prototype case of linear clique trees. The more general case of branched clique trees is treated later on.

5.1.4 The correlative sparsity pattern and its graph representation

This work heavily relies on a specific notion of sparsity defined as follows.

Specific notation: Given a Euclidean space $\mathbb{X}$ and a subspace $\mathbb{Y} \subset \mathbb{X}$, the orthogonal projection map from $\mathbb{X}$ to $\mathbb{Y}$ is denoted by $\pi_{\mathbb{Y}}$. Let $\mathbf{1}_j$ denote the j-th vector of the canonical basis of $\mathbb{R}^n$ such that if $\mathbf{x} = (x_1, \ldots, x_n)$ then $x_j = \mathbf{x} \cdot \mathbf{1}_j$. The N-dimensional subspace spanned by vectors $\mathbf{1}_{i_1}, \ldots, \mathbf{1}_{i_N}$ is denoted $\langle x_{i_1}, \ldots, x_{i_N} \rangle$ or $\langle x_{i_j} \rangle_{1 \leq j \leq N}$. Given a measure $\mu \in \mathcal{M}(\mathbb{X})$, its marginal with respect to $\mathbb{Y}$ is denoted by $\mu^{\mathbb{Y}} \in \mathcal{M}(\mathbb{Y})$. It is equal to the image or push-forward measure of μ through the map $\pi_{\mathbb{Y}}$.

Definition 5.1: Sparse polynomials

A scalar polynomial $p \in \mathbb{R}[\mathbf{x}]$ is said to be *sparse* when its vector of coefficients $\mathbf{p}$ (such that $p(\mathbf{x}) = \mathbf{p} \cdot \mathbf{e}(\mathbf{x})$ where $\mathbf{e}(\mathbf{x}) = (\mathbf{x}^{\mathbf{k}})_{|\mathbf{k}| \leq \mathrm{d}^\circ p}$) is sparse. In other words, p is a linear combination of a small number of monomials.

Definition 5.2: Correlative sparsity

A family of polynomial vectors $(\mathbf{g}_1, \ldots, \mathbf{g}_N)$ is said to be *correlatively sparse* whenever its correlative sparsity pattern matrix $\mathrm{R} := (R_{ij})_{1 \leq i,j \leq n}$, defined by

$$R_{ij} := \delta_{ij} + \sum_{k=1}^{N} \left\| \frac{\partial}{\partial x_i} \mathbf{g}_k \right\| \left\| \frac{\partial}{\partial x_j} \mathbf{g}_k \right\|$$

(where $\delta_{ij} = 1$ if $i = j$ and 0 otherwise, and $\| \cdot \|$ is any norm on polynomial vectors), is sparse. In other words, for many pairs of indices $i \neq j$, the variables x_i and x_j do not appear simultaneously in any element of $\{\mathbf{g}_1, \ldots, \mathbf{g}_N\}$.

Definition 5.3: Correlation graph

The *correlation graph* $\mathbb{G} = (\mathbf{V}, \mathbf{E})$ of $(\mathbf{g}_1, \ldots, \mathbf{g}_N)$ is defined by vertices $\mathbf{V} = \mathbb{N}_n^\star$ and edges $\mathbf{E} = \{(i,j) \in \mathbf{V}^2 : i \neq j \ \wedge \ R_{ij} \neq 0\}$.
The *correlative sparsity* CS of $(\mathbf{g}_1, \ldots, \mathbf{g}_N)$ is the treewidth of its correlation graph[a].

[a] Intuitively, the treewidth quantifies how "far" a graph is from being a tree. A proper definition is given in Section 5.3.4.

This chapter proposes a method to reduce the size of the degree d volume computation SDP to $\binom{\mathrm{CS}+d+1}{d}$ instead of $\binom{n+d}{d}$, under appropriate assumptions. To give an idea of the gain in computational complexity, we illustrate it on our motivating example of section 5.1.3, where $n = 100$ while $\mathrm{CS} = 1$.

Definition 5.4: Support of a polynomial

The *support* of $\mathbf{g}_i$ is the set $\mathbb{I}(\mathbf{g}_i) := \left\{ j \in \mathbb{N}_n^\star : \frac{\partial}{\partial x_j} \mathbf{g}_i \neq \mathbf{0} \right\}$.
The *support subspace* of $\mathbf{g}_i$ is the set $\mathbb{X}_i := \langle x_j \rangle_{j \in \mathbb{I}(\mathbf{g}_i)}$, whose dimension is smaller than n. Since by definition $\mathbf{g}_i = \mathbf{g}_i \circ \pi_{\mathbb{X}_i}$, we use both notations with the same meaning. Then $\mathbb{X} := \sum_{i=1}^N \mathbb{X}_i$ is called the *coordinate subspace decomposition* associated to $(\mathbf{g}_1, \ldots, \mathbf{g}_N)$.

Without loss of generality, we can suppose that $\mathbb{X} = \mathbb{R}^n$ (otherwise, there would be variables that appear in none of the $\mathbf{g}_i$s).

Definition 5.5: Sparse semialgebraic set

A *sparse basic semi-algebraic set* has a description

$$\mathbf{K} := \{\mathbf{x} \in \mathbb{X} \ : \ \forall i \in \mathbb{N}_N^\star, \ \mathbf{g}_i(\pi_{\mathbb{X}_i}(\mathbf{x})) \geq \mathbf{0}\}$$

where $(\mathbf{g}_i)_{1 \leq i \leq N}$ is a correlatively sparse family of polynomial vectors (inequalities are meant entrywise) and $\mathbb{X} = \sum_{i=1}^N \mathbb{X}_i$ is the coordinate subspace decomposition associated to $(\mathbf{g}_i)_{1 \leq i \leq N}$ (and, by extension, to $\mathbf{K}$).
A *sparse semi-algebraic set* is a finite union of sparse basic semi-algebraic sets that share the same coordinate subspace decomposition.

A simple example of a sparse basic semi-algebraic set is

$$\mathbf{K} := \{\mathbf{x} \in \mathbb{R}^4 \ : \ (\mathbf{g}_1(x_1, x_2) \geq \mathbf{0}) \wedge (\mathbf{g}_2(x_2, x_3) \geq \mathbf{0}) \wedge (\mathbf{g}_3(x_3, x_4) \geq \mathbf{0})\} \tag{5.1}$$

for $\mathbb{X} = \mathbb{R}^4$, $\mathbb{X}_1 = \langle x_1, x_2 \rangle$, $\mathbb{X}_2 = \langle x_2, x_3 \rangle$, $\mathbb{X}_3 = \langle x_3, x_4 \rangle$ and the projection maps are $\pi_{\mathbb{X}_1}(\mathbf{x}) = (x_1, x_2)$, $\pi_{\mathbb{X}_2}(\mathbf{x}) = (x_2, x_3)$, $\pi_{\mathbb{X}_3}(\mathbf{x}) = (x_3, x_4)$.

Our methodology is based on the classical theory of clique trees. Up to a chordal extension (which is equivalent to slightly weakening the correlative sparsity pattern), we suppose that the correlation graph $\mathbb{G} = (\mathbf{V}, \mathbf{E})$ is *chordal* (i.e. every cycle of length greater than 3 has a chord, that is, an edge linking two nonconsecutive vertices). Then we construct the maximal *cliques* of the graph (a clique $\mathbf{C}$ is a subset of $\mathbf{V}$ such that every vertex of $\mathbf{C}$ is connected to all the other vertices of $\mathbf{C}$ or, in

other words, such that $\mathbf{C}^2 \subset \mathbf{E}$; a clique is maximal when its cardinal is maximal). Most of the time, up to concatenation of some of the $\mathbf{g}_i$, the maximal cliques of a chordal correlation graph are exactly the supports of the $\mathbf{g}_i$: $\mathbf{C}_i = \mathbb{I}(\mathbf{g}_i)$, so in the following we will consider only such case[2]. Figure 5.1 illustrates this constuction for the sparse set (5.1), the vertices are denoted by x_i and our maximal cliques are denoted by $\mathbf{C}_j$.

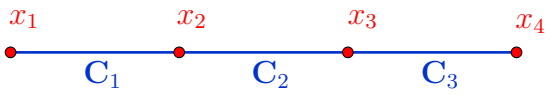


Figure 5.1 – Graph associated to the sparse set (5.1).

Then, we construct a *clique tree* which is instrumental to the computer implementation. It is proved in [13] that if the graph is chordal, then its maximal cliques can be organized within a tree satisfying the clique intersection property: for two maximal cliques $\mathbf{C}$ and $\mathbf{C}'$ the intersection $\mathbf{C} \cap \mathbf{C}'$ is contained in every maximal clique on the path from $\mathbf{C}$ to $\mathbf{C}'$. Figure 5.2 represents the clique tree associated to the sparse set (5.1).

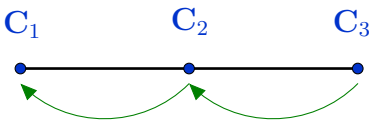


Figure 5.2 – Linear clique tree associated to the sparse set (5.1).

For a slightly more sophisticated illustration, consider the sparse set

$$\mathbf{K} := \{\mathbf{x} \in \mathbb{R}^6 \ : \ \mathbf{g}_1(x_1, x_2), \mathbf{g}_2(x_2, x_3, x_4), \mathbf{g}_3(x_3, x_5), \mathbf{g}_4(x_4, x_6) \geq \mathbf{0}\} \tag{5.2}$$

whose correlation graph is represented on Figure 5.3 and whose clique tree is represented on Figure 5.4. The clique tree of Figure 5.2 is called *linear* because all maximal cliques form a single chain (i.e. they are in a sequence) with no branching. In contrast, the clique tree of Figure 5.4 is called *branched*.

Our method consists of conveniently rooting the clique tree and splitting the volume computation problem into lower-dimensional subproblems that are in correspondence with the maximal cliques of the graph. The subproblem associated with a maximal clique $\mathbf{C}$ takes as only input the solutions of the subproblems associated with the children of $\mathbf{C}$ in the clique tree. This way, one can compute in parallel the solutions of all the subproblems of a given generation, and then use their results to solve the subproblems of the parent generation. This is the meaning of the arrows in Figures 5.2 and 5.4. The volume of $\mathbf{K}$ is the optimal value of the (last) sub-problem associated with the root $\mathbf{C}_1$ of the tree.

[2]The only exception would be cliques forming a triangle and is tackled in detail in section 5.3.4.

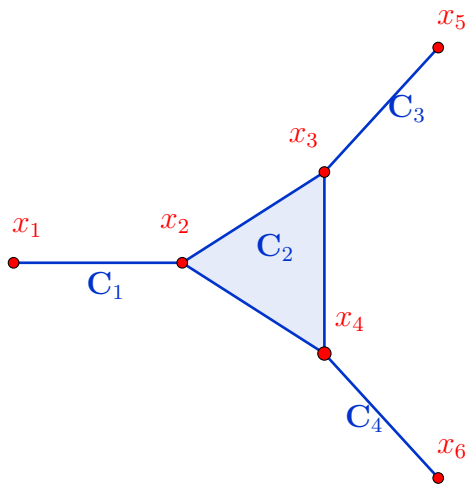


Figure 5.3 – Graph associated to the sparse set (5.2).

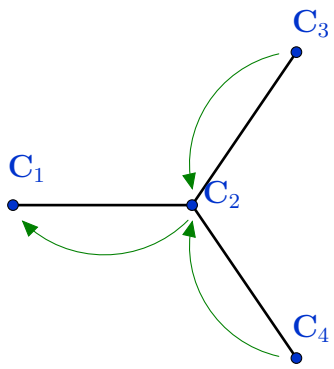


Figure 5.4 – Branched clique tree associated to the sparse set (5.2).

5.1.5 An illustrative example: the bicylinder

A sparse moment-SOS hierarchy

Before describing the methodology in the general case, we briefly explain the general underlying idea on a simple illustrative example. Consider the sparse semi-algebraic set

$$\mathbf{K} := \left\{ \mathbf{x} \in \mathbb{R}^3 \ : \ \begin{array}{l} g_1(x_1, x_2) := 1 - x_1^2 - x_2^2 \geq 0 \\ g_2(x_2, x_3) := 1 - x_2^2 - x_3^2 \geq 0 \end{array} \right\} \subset \mathbf{X} := [-1, 1]^3 \qquad (5.3)$$

modelling the intersection of two cylinders $\mathbf{U}_1 := \{\mathbf{x} \in \mathbb{R}^3 \ : \ x_1^2 + x_2^2 \leq 1\}$ and $\mathbf{U}_2 := \{\mathbf{x} \in \mathbb{R}^3 \ : \ x_2^2 + x_3^2 \leq 1\}$, see Figure 5.5. The subspaces are $\mathbb{X}_1 = \langle x_1, x_2 \rangle$ and $\mathbb{X}_2 = \langle x_2, x_3 \rangle$ and the projection maps are $\pi_{\mathbb{X}_1}(\mathbf{x}) = (x_1, x_2)$ and $\pi_{\mathbb{X}_2}(\mathbf{x}) = (x_2, x_3)$. Let $\mathbf{K}_i := \pi_{\mathbb{X}_i}(\mathbf{U}_i)$ and $\mathbf{X}_i := \pi_{\mathbb{X}_i}(\mathbf{X}) = [-1, 1]^2$ for $i = 1, 2$.

Following [44], computing $\operatorname{vol} \mathbf{K}$ is equivalent to solving the infinite-dimensional

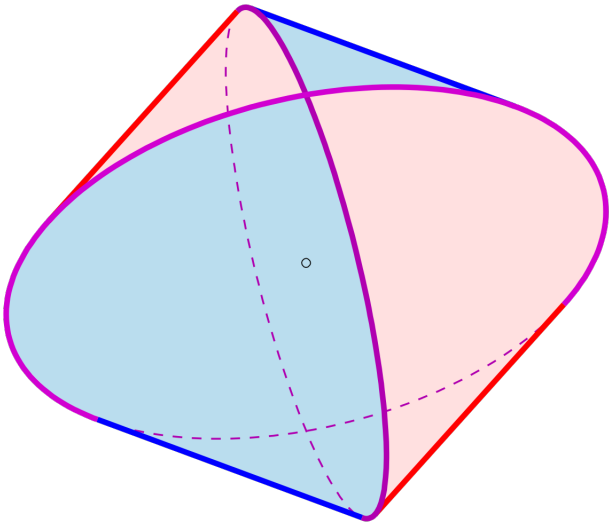


Figure 5.5 – A representation of the bicylinder.

LP (2.8a). Next observe that in the description (5.3) of $\mathbf{K}$ there is no direct interaction between variables x_1 and x_3, but this is neither exploited in the LP formulation (2.8a) nor in the SDP relaxations (2.12) to solve (2.8a). To exploit this correlative sparsity pattern we propose the following alternative formulation

$$\begin{aligned} \operatorname{vol}\mathbf{K} \;=\; \mathrm{p}^\star_{\mathrm{sp}\mathbf{K}} := \max_{\mu_1,\mu_2} & \int 1\, d\mu_1 \\ \text{s.t. } & \mu_i \in \mathcal{M}_+(\mathbf{K}_i) \qquad i = 1,2 \\ & \mu_2 \preceq \lambda \otimes \lambda \\ & \mu_1 \preceq \lambda \otimes \mu_2^{\langle x_2\rangle} \end{aligned} \tag{5.4}$$

where $\mu_2^{\langle x_2\rangle}$ denotes the marginal of μ_2 in the variable x_2, $\mu \preceq \nu$ means that $\nu - \mu$ is a nonnegative measure and $\otimes$ denotes the tensor product between independent measures. Notice that here (as well as in the rest of this chapter) for disambiguation purposes we denote by λ the *one-dimensional* Lebesgue measure on $[-1,1]$ (whose moments are well known). Then, we will use λ^n to denote the n-dimensional Lebesgue measure on $[-1,1]^n$. For the sake of readability, we only use this notation in the present chapter, where it is necessary.

In the sparse case, the basic idea behind our reformulation of the volume problem is as follows. We are interested in $\operatorname{vol}\mathbf{K}$. However, as the marginal of a measure has the same mass as the measure itself, instead of looking for the full measure μ in problem (2.8a), we only look for its marginal on $\mathbb{X}_1$.

This marginal $\mu^{\mathbb{X}_1}$ is modeled by μ_1 in (5.4). In order to compute it, we need some additional information on μ captured by the measure μ_2 in (5.4). The unique optimal solution μ of (2.8a) is

$$\mu = \lambda^3_{\mathbf{K}} = \mathbb{1}_{\mathbf{U}_1}\, \mathbb{1}_{\mathbf{U}_2}\, \lambda^3$$

and therefore its marginal $\mu_1 := \mu^{\mathbb{X}_1}$ in (x_1, x_2) is given by

$$\begin{aligned} d\mu_1(x_1, x_2) &= \int_0^1 d\mu(x_1, x_2, x_3) \\ &= \mathbb{1}_{\mathbf{K}_1}(x_1, x_2)\, dx_1 \underbrace{\left(\int_0^1 \mathbb{1}_{\mathbf{K}_2}(x_2, x_3)\, dx_3 \right)}_{d\mu_2^{\langle x_2 \rangle}(x_2)} dx_2 \end{aligned} \tag{5.5}$$

where

$$\mu_2 = \lambda^2_{\mathbf{K}_2}. \tag{5.6}$$

What is the gain in solving (5.4) when compared to solving (2.8a) ? Observe that in (5.4) we have two unknown measures μ_1 and μ_2 on $\mathbb{R}^2$, instead of a single measure μ on $\mathbb{R}^3$ in (2.8a). In the resulting SDP relaxations associated with (5.4) this translates into SDP constraints of potentially much smaller size. For instance, and to fix ideas, for the same relaxation degree d:

- The SDP relaxation $\mathrm{p}^d_{\mathbf{K}}$ corresponding to (2.8a) contains a moment matrix (associated with μ in (2.8a)) of size $\binom{3+d}{d}$;
- The SDP relaxation $\mathrm{p}^d_{\mathrm{sp}\mathbf{K}}$ corresponding to (5.4) contains two moment matrices, one associated with μ_1 of size $\binom{2+d}{d}$, and one associated with μ_2 of size $\binom{2+d}{d}$, where μ_1 and μ_2 are as in (5.4).

As the size of those matrices is the crucial limiting parameter for all SDP solvers, one can immediately appreciate the computational gain that can be expected from the formulation (5.4) versus the formulation (2.8a) when the dimension is high or the relaxation order increases. Next it is not difficult to extrapolate that the gain can be even more impressive in the case where the correlative sparsity pattern is of the form

$$\mathbf{K} = \{(\mathbf{x}_0, \ldots \mathbf{x}_N) \in \mathbb{X} : \ \forall i \in \mathbb{N}^\star_N, \quad g_i(\mathbf{x}_{i-1}, \mathbf{x}_i) \geq 0 \}, \tag{5.7}$$

with $\mathbf{x}_i \in \mathbb{R}^{n_i}$ and $n_i \ll n$ for $i = 0, \ldots, N$. In fact, it is straightforward to define examples of sets $\mathbf{K}$ of the form (5.7) where the first SDP relaxation associated with the original dense LP formulation (2.8a) cannot be even implemented on state-of-the-art computers, whereas the SDP relaxations associated with a generalization of the sparse LP formulation (5.4) can be easily implemented, at least for reasonable values of d.

Sparse Stokes constraints: the bicylinder

In Section 4.4.4 we designed efficient general Stokes constraints for the dense formulation of problem (2.8a), at the price of introducing a polynomial h vanishing on the boundary of $\mathbf{K}$ to obtain problem (4.15). However, in the sparse case (5.4), the polynomial h would destroy the sparsity structure, as it is the product of all

polynomials defining $\mathbf{K}$ and as such it depends on all components of $\mathbf{x}$. Thus, we must adapt our strategy to introduce sparse Stokes constraints[3].

In this section, to keep the notations simple, we illustrate the ideas on our introductive bicylinder example of Section 5.1.5. Considering the optimal measures μ_1 and μ_2 defined in (5.5),(5.6), we can apply Stokes constraints derived from the Gauss formula (4.4), in the directions in which they are Lebesgue: for μ_1 in the x_1 direction and for μ_2 in the remaining directions. To see this, define

$$\begin{aligned}\mathbf{h}_1(x_1,x_2) &= g_1(x_1,x_2)\,\mathbf{e}_1,\\ \mathbf{h}_2(x_2,x_3) &= g_2(x_2,x_3)\,\mathbf{e}_2,\\ \mathbf{h}_3(x_2,x_3) &= g_2(x_2,x_3)\,\mathbf{e}_3\end{aligned}$$

where $g_i(x_i,x_{i+1}) = 1-x_i^2-x_{i+1}^2$, such that $\mathbf{h}_1\cdot\mathbf{n}_{\mathbf{K}_1}$ vanishes on the boundary of $\mathbf{K}_1$ and $\mathbf{h}_2\cdot\mathbf{n}_{\mathbf{K}_2}$ and $\mathbf{h}_3\cdot\mathbf{n}_{\mathbf{K}_2}$ vanish on the boundary of $\mathbf{K}_2$, where $\mathbf{n}_{\mathbf{K}_i}$ is the outward point vector orthogonal to the boundary of $\mathbf{K}_i$. For $i,j,k\in\mathbb{N}$, the Gauss formula (4.4) yields

$$\begin{aligned}\int_{\mathbf{K}_1}\frac{\partial}{\partial x_1}(g_1(x_1,x_2)x_1^ix_2^j)\,d\mu_1 &= 0,\\ \int_{\mathbf{K}_2}\frac{\partial}{\partial x_2}(g_2(x_2,x_3)x_2^jx_3^k)\,d\mu_2 &= 0,\\ \int_{\mathbf{K}_2}\frac{\partial}{\partial x_3}(g_2(x_2,x_3)x_2^jx_3^k)\,d\mu_2 &= 0.\end{aligned}$$

Hence, adding these constraints does not change the optimal value of the LP problem (5.4).

5.2 Exploiting path decomposition sparsity

5.2.1 Path computation theorem

We consider a sparse basic semialgebraic set

$$\mathbf{K} := \{\mathbf{x}\in\mathbb{X} : \forall i\in\mathbb{N}_N^\star, \mathbf{g}_i(\pi_{\mathbb{X}_i}(\mathbf{x}))\geq\mathbf{0}\}$$

with $\mathbf{g}_i\in\mathbb{R}[\mathbf{x}_i]^{m_i}$ and its coordinate subspace decomposition $\mathbb{R}^n =: \mathbb{X} = \sum_{i=1}^N \mathbb{X}_i$. Let $\mathbf{U}_i := \{\mathbf{x}\in\mathbb{X} \;:\; \mathbf{g}_i(\pi_{\mathbb{X}_i}(\mathbf{x}))\geq\mathbf{0}\}$ so that our sparse semi-algebraic set can be written

$$\mathbf{K} = \bigcap_{i=1}^N \mathbf{U}_i.$$

[3]Actually, Stokes formulation (4.16) can be made sparse more directly, by enforcing the fact that the boundary measure ν_i only depends on the variables of the polynomial g_i. However, during the PhD, we tackled sparse volume computation before working on problem (4.16). As a result, in this chapter we only use Stokes constraints under the form of (4.15).

Up to translation and rescaling, we suppose that $\mathbf{K} \subset \mathbf{X} := [0,1]^n$. Moreover, let

$$\mathbf{K}_i := \{\mathbf{x}_i \in \mathbb{X}_i \ : \ \mathbf{g}_i(\mathbf{x}_i) \geq \mathbf{0}\} = \pi_{\mathbb{X}_i}(\mathbf{U}_i) \quad ; \quad \mathbf{X}_i := \pi_{\mathbb{X}_i}(\mathbf{X})$$

and let

$$\mathbb{Y}_i := \mathbb{X}_i \cap \mathbb{X}_{i+1}^{\perp} = \langle x_j \rangle_{j \in \mathbf{C}_i \cap \mathbf{C}_{i+1}^c}$$

be a subspace of dimension $n_i := |\mathbf{C}_i \cap \mathbf{C}_{i+1}^c|$ for $i \in \mathbb{N}_{N-1}^\star$ with $\mathbb{Y}_N = \mathbb{X}_N$. The superscript $\perp$ denotes the orthogonal complement. We work under the following assumptions.

Assumption 5.6: Connected correlation graph

For all $i \in \{2, \ldots, N\}$ one has $\mathbb{X}_i \cap \sum_{j=1}^{i-1} \mathbb{X}_j \neq \{\mathbf{0}\}$.

Assumption 5.7: Path decomposition

For all $i \in \{2, \ldots, N\}$ one has $\mathbb{X}_i \cap \sum_{j=1}^{i-1} \mathbb{X}_j \subset \mathbb{X}_{i-1}$.

If Assumption 5.6 is violated then $\mathbf{K}$ can be decomposed as a Cartesian product, and one should just apply our methodology to each one of its factors. Assumption 5.7 ensures that the associated clique tree is a path decomposition as in Figure 5.2, *i.e.* that it does not contain branchings.

Theorem 5.8: Path computation

If Assumptions 5.6 and 5.7 hold, then

$$\begin{aligned} \operatorname{vol} \mathbf{K} = \mathrm{p}_{\mathrm{sp}\mathbf{K}}^\star := \max_{(\mu_i)_i} & \int 1 \, d\mu_1 \\ \text{s.t. } & \mu_i \in \mathcal{M}(\mathbf{K}_i)_+ && i \in \mathbb{N}_N^\star && \text{(5.8a)} \\ & \mu_i \preceq \mu_{i+1}^{\mathbb{X}_i \cap \mathbb{X}_{i+1}} \otimes \lambda^{n_i} && i \in \mathbb{N}_{N-1}^\star && \text{(5.8b)} \\ & \mu_N \preceq \lambda^{n_N}. && && \text{(5.8c)} \end{aligned}$$

Proof : Let us first prove that $\mathrm{p}_{\mathrm{sp}\mathbf{K}} \geq \operatorname{vol}\mathbf{K}$. For $i \in \mathbb{N}_N^\star$, let $\mathbb{W}_i := \mathbb{X}_i^{\perp} \cap \sum_{j=i+1}^{N} \mathbb{X}_j$ so that $\sum_{j=i}^{N} \mathbb{X}_j = \mathbb{X}_i \oplus \mathbb{W}_i$. Our working assumptions ensure that $\mathbb{W}_i$ is never reduced to $\{\mathbf{0}\}$. For $\mathbf{x}_i \in \mathbb{X}_i$ define

$$d\mu_i(\mathbf{x}_i) := \mathbb{1}_{\mathbf{K}_i}(\mathbf{x}_i) \left(\int_{\mathbb{W}_i} \prod_{j=i+1}^{N} \mathbb{1}_{\mathbf{K}_j} \circ \pi_{\mathbb{X}_j}(\mathbf{x}_i + \mathbf{w}_i) \, \mathrm{d}\mathbf{w}_i \right) \mathrm{d}\mathbf{x}_i. \tag{5.9}$$

By construction $\mu_i \in \mathcal{M}_+(\mathbf{K}_i)$ and constraints (5.8a) are enforced. In addition, one

can check that, if $\mathbf{x}_{i,i+1} \in \mathbb{X}_i \cap \mathbb{X}_{i+1}$, then

$$\begin{aligned} d\mu_{i+1}^{\mathbb{X}_i \cap \mathbb{X}_{i+1}}(\mathbf{x}_{i,i+1}) &\stackrel{\text{def}}{=} \int_{\mathbf{y}_{i,i+1} \in \mathbb{X}_i^\perp \cap \mathbb{X}_{i+1}} d\mu_{i+1}(\mathbf{x}_{i,i+1} + \mathbf{y}_{i,i+1}) \\ &\stackrel{(5.9)}{=} \left(\int_{\mathbb{X}_i^\perp \cap \mathbb{X}_{i+1}} \mathbb{1}_{\mathbf{K}_{i+1}}(\mathbf{x}_{i,i+1} + \mathbf{y}_{i,i+1}) \dots \right. \\ &\dots \left. \left(\int_{\mathbb{W}_{i+1}} \prod_{j=i+2}^{N} \mathbb{1}_{\mathbf{K}_j} \circ \pi_{\mathbb{X}_j}(\mathbf{x}_{i,i+1} + \mathbf{y}_{i,i+1} + \mathbf{w}_{i+1}) \, \mathrm{d}\mathbf{w}_{i+1} \right) \mathrm{d}\mathbf{y}_{i,i+1} \right) \mathrm{d}\mathbf{x}_{i,i+1} \\ &= \left(\int_{\mathbb{W}_i} \prod_{j=i+1}^{N} \mathbb{1}_{\mathbf{K}_j} \circ \pi_{\mathbb{X}_j}(\mathbf{x}_{i,i+1} + \mathbf{w}_i) \, \mathrm{d}\mathbf{w}_i \right) \mathrm{d}\mathbf{x}_{i,i+1} \end{aligned}$$

since by construction of the $\mathbb{W}_j$, $(\mathbb{X}_i^\perp \cap \mathbb{X}_{i+1}) \oplus \mathbb{W}_{i+1} = \mathbb{W}_i$.

Thus, comparing this to (5.9), one can see that constraints (5.8b) are satisfied. Moreover, they are saturated on $\mathbf{K}_i$. Eventually, one has

$$\mathbb{X} = \mathbb{X}_1 \oplus \mathbb{W}_1$$

and thus

$$\begin{aligned} \int 1 \, d\mu_1 &= \int_{\mathbb{X}_1} \mathbb{1}_{\mathbf{K}_1}(\mathbf{x}_1) \left(\int_{\mathbb{W}_1} \prod_{j=2}^{N} \mathbb{1}_{\mathbf{U}_j}(\mathbf{x}_1 + \mathbf{w}_1) \, \mathrm{d}\mathbf{w}_1 \right) \mathrm{d}\mathbf{x}_1 \\ &= \int_{\mathbb{X}} \left(\prod_{j=1}^{N} \mathbb{1}_{\mathbf{U}_j}(\mathbf{x}) \right) \mathrm{d}\mathbf{x} \\ &= \int_{\mathbb{X}} \mathbb{1}_{\mathbf{K}}(\mathbf{x}) \, \mathrm{d}\mathbf{x} \\ &= \operatorname{vol} \mathbf{K}, \end{aligned}$$

that is, we have just proved that $\mathrm{p}^\star_{\mathrm{sp}\mathbf{K}} \geq \operatorname{vol} \mathbf{K}$.

To prove the converse inequality, observe that our previous choice $\mu_1, \dots, \mu_N$ saturates the constraints (5.8b) while enforcing the constraints (5.8a). Any other feasible solution $\tilde{\mu}_1, \dots, \tilde{\mu}_N$ directly satisfies the inequality $\tilde{\mu}_i \preceq \mu_i$. In particular, $\tilde{\mu}_1 \preceq \mu_1$ and thus

$$\int 1 \, d\tilde{\mu}_1 \leq \int 1 \, d\mu_1 = \operatorname{vol} \mathbf{K}.$$

Taking the maximum over all feasible $(\tilde{\mu}_i)_{i \in \mathbb{N}^\star_N}$ yields that $\mathrm{p}^\star_{\mathrm{sp}\mathbf{K}} \leq \operatorname{vol} \mathbf{K}$. ♢

Remark 5.1 (Duality in the sparse scheme)
The dual of the LP problem of Theorem 5.8 is the LP problem

$$\begin{aligned} \mathrm{d}^\star_{\mathrm{sp}\mathbf{K}} := \inf_{(w_i)_i} & \int w_N \, d\lambda^{n_N} && && (5.10) \\ \text{s.t. } & w_1(\mathbf{x}_1) \geq 1 && \mathbf{x}_1 \in \mathbf{K}_1 && \\ & w_{i+1}(\mathbf{x}_{i+1}) \geq \int_{\mathbb{Y}_i} w_i(\mathbf{y}_i + \pi_{\mathbb{X}_i}(\mathbf{x}_{i+1})) \, \mathrm{d}\mathbf{y}_i && \mathbf{x}_{i+1} \in \mathbf{K}_{i+1} && i \in \mathbb{N}^\star_{N-1} \\ & w_i \in \mathcal{C}(\mathbf{X}_i)_+ && && i \in \mathbb{N}^\star_N. \end{aligned}$$

It is straightforward to prove that if $\mu_1, \ldots, \mu_N$ is feasible for problem (5.8)*, then for any $i \in \mathbb{N}_N^\star$, one has $\int 1\, d\mu_i \leq \int 1\, d\mu_N \leq C := \operatorname{vol} \mathbf{X}_N = 1$, so that Theorem 2.6 holds, and there is no duality gap, i.e. $\mathrm{d}^\star_{\mathrm{sp}\mathbf{K}} = \mathrm{p}^\star_{\mathrm{sp}\mathbf{K}} = \operatorname{vol} \mathbf{K}$. For example, in the case of the bicylinder treated in Section 5.1.5, the dual of* (5.4) *reads:*

$$\begin{aligned} \mathrm{d}^\star_{\mathrm{sp}\mathbf{K}} = \inf_{w_1, w_2} & \int_{\mathbf{X}_2} w_2(x_2, x_3)\, \mathrm{d}x_2 \mathrm{d}x_3 \\ \text{s.t. } & w_1(x_1, x_2) \geq 1 && (x_1, x_2) \in \mathbf{K}_1 \\ & w_2(x_2, x_3) \geq \int_{-1}^{1} w_1(x_1, x_2)\, \mathrm{d}x_1 && (x_2, x_3) \in \mathbf{K}_2 \\ & w_i \in \mathcal{C}(\mathbf{X}_i)_+ && i \in \mathbb{N}_N^\star. \end{aligned}$$

Thus, if $(w_1^{(d)}, w_2^{(d)})_{k \in \mathbb{N}}$ is a minimizing sequence for this dual LP, then the sets

$$\hat{\mathbf{K}}_d := \left\{ (x_1, x_2, x_3) \in [0,1]^3 \ : \ (w_1^{(d)}(x_1, x_2) \geq 1) \wedge \left(w_2^{(d)}(x_2, x_3) \geq \int_0^1 w_1^{(d)}(x_1, x_2)\, \mathrm{d}x_1 \right) \right\}$$

are outer approximations of the set $\mathbf{K}$ and the sequences $(\operatorname{vol} \hat{\mathbf{K}}_d)_k$ and $(\int w_2^{(d)}\, d\lambda^2)$ decrease to $\operatorname{vol} \mathbf{K}$. Similar statements can be made for the general dual problem.

Corollary 5.9: Convergence of the path computation scheme

Under Assumption 2.7 on the $\mathbf{K}_i$ and $\mathbf{X}_i = \pi_{\mathbb{X}_i}([-1,1]^n)$, $i \in \mathbb{N}_N^\star$, the Moment-SOS hierarchy corresponding to problem (5.8) converges to $\operatorname{vol} \mathbf{K}$.

Proof : Since strong duality $\mathrm{p}^\star_{\mathrm{sp}\mathbf{K}} = \mathrm{d}^\star_{\mathrm{sp}\mathbf{K}}$ holds, this is a direct consequence of Theorems 5.8, 2.14 and 2.15. ♢

Remark 5.2 (Sequential Moment-SOS hierarchy)

The LP (5.8) *is formulated as a single problem on N unknown measures. However, it is possible to split it in small chained subproblems to be solved in sequence. Each subproblem is associated with a maximal clique (in the linear clique tree) and it takes as input the results of the subproblem associated with its parent clique. This way, the sparse volume computation is split into N linked low-dimension problems and solved sequentially. This may prove useful when N is large because when solving the SDP relaxations associated with the single LP* (5.8)*, the SDP solver may encounter difficulties in handling a high number of measures simultaneously. It should be easier to sequentially solve a high number of low-dimensional problems with only one unknown measure. Both formulations being strictly equivalent, this would not change the convergence properties of the sparse scheme.*

As explained in Chapter 2, the hierarchy of SDP relaxations associated with our infinite-dimensional LP provides us with a sequence of upper bounds on $\operatorname{vol} \mathbf{K}$. One may also be interested in computing lower bounds on $\operatorname{vol} \mathbf{K}$. In principle it suffices to apply the same methodology and approximate from above the volume of $\mathbf{X} \setminus \mathbf{K}$ since $\mathbf{K}$ is included in the unit box $\mathbf{X}$. However, it is unclear whether $\mathbf{X} \setminus \mathbf{K}$ has

also a sparse description. We show that this is actually the case and so one may exploit correlative sparsity to compute lower bounds although it is more technical. The following result is a consequence of Theorem 5.8:

Corollary 5.10: Lower bounds sparse computation

If $\mathbf{K}$ is sparse, then $\widetilde{\mathbf{K}} := \mathbf{X} \setminus \mathbf{K}$ is sparse as well, and $\operatorname{vol} \widetilde{\mathbf{K}}$ is the value of the LP problem

$$\begin{aligned} \mathrm{p}^{\star}_{\mathrm{sp}\widetilde{\mathbf{K}}} := \max_{(\mu_{i,j})_{i,j}} & \sum_{j=1}^{N} \int 1 \, d\mu_{1,j} \\ \text{s.t. } & \mu_{j,j} \in \mathcal{M}(\operatorname{cl} \widetilde{\mathbf{K}}_j)_+ && 1 \le j \le N \\ & \mu_{i,j} \in \mathcal{M}(\mathbf{K}_i)_+ && 1 \le i < j \le N \\ & \mu_{j,j} \preceq \lambda^{p_j} && 1 \le j \le N \\ & \mu_{i,j} \preceq \mu_{i+1,j}^{\mathbb{X}_i \cap \mathbb{X}_{i+1}} \otimes \lambda^{n_i} && 1 \le i < j \le N \end{aligned}$$

where $p_j := \dim \mathbb{X}_j$, $n_i := \dim \mathbb{X}_{i+1}^{\perp} \cap \mathbb{X}_i$, $\widetilde{\mathbf{K}}_j := [0,1]^{p_j} \setminus \mathbf{K}_j$ is open and $\operatorname{cl} \widetilde{\mathbf{K}}_j$ denotes its closure[a].

[a] This is necessary to ensure that our measures are defined on compact sets so that all representation theorems hold.

Proof : The following description

$$\widetilde{\mathbf{K}} = \bigsqcup_{j=1}^{N} \left[\bigcap_{i=1}^{j-1} \mathbf{U}_i \cap \widetilde{\mathbf{U}}_j \right],$$

where $\bigsqcup$ stands for disjoint union and $\widetilde{\mathbf{U}}_j := \{\mathbf{x} \in \mathbb{R}^n : \pi_{\mathbb{X}_j}(\mathbf{x}) \in \widetilde{\mathbf{K}}_j\}$, is sparse. Indeed the description of the basic semi-algebraic set

$$\mathbf{L}_j := \bigcap_{i=1}^{j-1} \mathbf{U}_i \cap \widetilde{\mathbf{U}}_j$$

is sparse. In addition, by σ-additivity of the Lebesgue measure, one has

$$\operatorname{vol} \widetilde{\mathbf{K}} = \sum_{j=1}^{N} \operatorname{vol} \mathbf{L}_j.$$

Finally, by using Theorem 5.8 we conclude that $\operatorname{vol} \mathbf{L}_j$ is the value of LP consisting of maximizing $\int_{\mathbb{X}_j} d\mu_{1,j}$ subject to the same constraints as in the above LP problem. Summing up yields the correct value. ♢

5.2.2 General sparse Stokes constraints

Consider the sequential decomposition of problem (5.8) in Theorem 5.8:

$$\begin{aligned} \mathrm{p}^\star_{\mathrm{sp}\mathbf{K},i} := \max_{\mu_i} & \int 1 \, d\mu_i \\ & \text{s.t. } \mu_i \in \mathcal{M}(\mathbf{K}_i)_+ \\ & \quad \mu_i \preceq \mu_{i+1}^{\star \mathbb{X}_i \cap \mathbb{X}_{i+1}} \otimes \lambda^{n_i} \end{aligned}$$

for $i \in \mathbb{N}^\star_{N-1}$, and

$$\begin{aligned} \mathrm{p}^\star_{\mathrm{sp}\mathbf{K},N} := \max_{\mu_N} & \int 1 \, d\mu_N \\ & \text{s.t. } \mu_N \in \mathcal{M}(\mathbf{K}_N)_+ \\ & \quad \mu_N \preceq \lambda^{n_N} \end{aligned}$$

Our algorithm consists of sequentially solving these problems, starting with determining μ_N, then μ_{N-1}, and so on until μ_1, whose mass will be $\mathrm{vol}(\mathbf{K})$. We implement Stokes constraints on each one of these problems. For the problem in μ_N, we implement regular Stokes constraints as in Section 4.4.4:

$$\begin{aligned} \mathrm{p}^\star_{\mathrm{sp}\mathbf{K},N} := \max_{\mu_N} & \int 1 \, d\mu_N \\ & \text{s.t. } \mu_N \in \mathcal{M}(\mathbf{K}_N)_+ \\ & \quad \mu_N \preceq \lambda^{n_N} \\ & \quad (\mathbf{grad}\, h_N)\mu_N - (\mathbf{grad}\, h_N \, \mu_N) = \mathbf{0} \end{aligned}$$

where h_N is a polynomial vanishing on $\partial\mathbf{K}_N$, for example $h_N = g_{N,1} \cdots g_{N,m_N}$.

Then, let $i \in \mathbb{N}^\star_{N-1}$ and suppose that μ_{i+1} is known, such that determining μ_i is reduced to solving the LP problem of computing $\mathrm{p}^\star_{\mathrm{sp}\mathbf{K},i}$. From the arguments of Section 5.2.1, we know that the optimal measure μ_i is supported on $\mathbf{K}_i$ and that on this set it is the product measure between $\mu_{i+1}^{\mathbb{X}_i \cap \mathbb{X}_{i+1}}$ and the uniform measure on $\mathbb{Y}_i = \langle x_j \rangle_{j \in \mathbf{C}_i \cap \mathbf{C}_{i+1}^{\mathrm{c}}}$. Since Stokes' theorem is only valid for uniform measures, it will only apply to $\mu_i^{\mathbb{Y}_i}$.

Let $\mathbb{J} \subset \mathbb{N}^\star_n$. For $f \in \mathcal{C}^1(\mathbb{R}^n)$ we define

$$\mathbf{grad}_{\mathbb{J}} f := \left(\frac{\partial f}{\partial x_j} \right)_{j \in \mathbb{J}}$$

such that $\mathbf{grad}_{\mathbb{N}^\star_n} f = \mathbf{grad} f$ and $\mathbf{grad}_{\{j\}} f = \frac{\partial f}{\partial x_j}$ for example. This notation allows us to define Stokes constraints exactly in the directions we are interested in and to formulate the general sparse Stokes constraints:

$$\begin{aligned} \mathrm{p}^\star_{\mathrm{sp}\mathbf{K},i} := \max_{\mu_i} & \int 1 \, d\mu_i \\ & \text{s.t. } \mu_i \in \mathcal{M}(\mathbf{K}_i)_+ \\ & \quad \mu_i \preceq \mu_{i+1}^{\star \mathbb{X}_i \cap \mathbb{X}_{i+1}} \otimes \lambda^{n_i} \\ & \quad (\mathbf{grad}_{\mathbf{C}_i \cap \mathbf{C}_{i+1}^{\mathrm{c}}} h_i)\mu_i - \mathbf{grad}_{\mathbf{C}_i \cap \mathbf{C}_{i+1}^{\mathrm{c}}}(h_i \, \mu_i) = \mathbf{0} \end{aligned}$$

where h_i is a polynomial vanishing on $\partial\mathbf{K}_i$, for example $h_i = g_{i,1} \cdots g_{i,m_i}$.

Remark 5.3 (Refining Stokes constraints)

In some cases, in both dense and sparse contexts, these Stokes constraints can be slightly improved by choosing a different polynomial $h_{(i,)j}$ for each basis vector $\mathbf{e}_j$ when applying the Gauss formula (4.4) to $\mathbf{u} = h_j\,\mathbf{e}_j$, such that h_j can be taken with the lowest possible degree, allowing for a better implementation of the hierarchy. For example, if one is looking for the volume of $\mathbf{K} := [0,1]^2$, the polynomial vanishing on $\partial\mathbf{K}$ with the lowest degree is $h(x_1,x_2) := x_1(1-x_1)x_2(1-x_2)$, but one can formulate Stokes constraints by applying the Gauss formula (4.4) to $x_1(1-x_1)\,\mathbf{e}_1$ and $x_2(1-x_2)\,\mathbf{e}_2$, instead of $h(x_1,x_2)\,\mathbf{e}_1$ and $h(x_1,x_2)\,\mathbf{e}_2$. By doing so, one would replace the constraint $(\mathbf{grad}\,h)\,\mu - \mathbf{grad}(h\,\mu) = \mathbf{0}$ with $\left(\frac{\partial h_j}{\partial x_j}\right)\mu - \frac{\partial}{\partial x_j}\,(h_j\,\mu) = 0$ for every possible j. This is what we actually implemented in our numerical examples, but we presented the Stokes constraints in the restrictive case of $h_j = h$ for all j for the sake of readability.

5.2.3 Path computation examples

Bicylinder revisited

We refer to (2.8a) as the dense problem and to (5.4) as the sparse problem. For both problems, we consider instances with and without additional Stokes constraints. Note that for the bicylinder example of Section 5.1.5 the optimal value for both the dense and the sparse problem is

$$\operatorname{vol}\mathbf{K} = \frac{16}{3} \approx 5.3333$$

since adding Stokes constraints does not change the optimal value.

We solve the SDP relaxations with Mosek on a standard laptop, for various relaxation orders and we report the bounds and the computation times in Table 5.1. We observe a slow convergence for the dense and the sparse versions without Stokes constraints, and a much faster convergence with Stokes constraints. *We also observe significantly smaller computation times when using the sparse formulation.*

A nonconvex set

Let $\mathbb{X} := \mathbb{R}^5$, $\mathbb{X}_1 = \langle x_1, x_2\rangle$, $\mathbb{X}_2 = \langle x_1, x_3\rangle$, $\mathbb{X}_3 = \langle x_1, x_4\rangle$, $\mathbb{X}_4 = \langle x_1, x_5\rangle$ and for $i = 1,2,3,4$

- $\mathbf{g}_i(x_1, x_{i+1}) := (2x_1^2 - x_{1+i}^2 - 1\ ,\ x_1\,(1-x_1)\ ,\ x_{i+1}\,(1-x_{i+1}))$,
- $\mathbf{K}_i := \mathbf{g}_i^{-1}\left((\mathbb{R}_+)^3\right) = \{(x_1, x_{i+1}) \in [0,1]^2 : 2x_1^2 - x_{i+1}^2 \geq 1\}$.

Let us approximate the volume of the sparse set

$$\mathbf{K} := \left\{(x_1,x_2,x_3,x_4,x_5) \in [0,1]^5 : \forall i \in \mathbb{N}_4^\star,\ 2x_1^2 - x_{i+1}^2 \geq 1\right\} = \bigcap_{i=1}^{4} \pi_{\mathbb{X}_i}^{-1}\left(\mathbf{K}_i\right).$$

Here the coordinates x_2, x_3, x_4 and x_5 do not interact: they are only linked with the coordinate x_1. The proper way to apply our linear computation Theorem 5.8 is to define a linear clique tree as shown in Figure 5.6.

	full		sparse	
d	without Stokes	with Stokes	without Stokes	with Stokes
2	7.8232 (1.0s)	5,828 (1.1s)	7,7424 (1.1s)	5,4984 (1.1s)
3	7.2368 (0.9s)	5,4200 (1.3s)	7,1920 (0.9s)	5,3488 (1.1s)
4	7.0496 (1.4s)	5,3520 (2.2s)	7,0040 (1.2s)	5,3376 (1.2s)
5	6,8136 (3.1s)	5,3400 (4.4s)	6,7944 (1.8s)	5,3352 (1.8s)
6	6,7376 (7.2s)	5,3376 (8.2s)	6,6960 (2.1s)	5,3344 (2.3s)
7	6,6336 (12.8s)	5,3360 (18.3s)	6,6168 (2.6s)	5,3344 (3.2s)

Table 5.1 – Performance of sparse computation of the bicylinder's volume.

Bounds on the volume (and computation times in seconds) vs relaxation order for the bicylinder.

One can verify that such organization of the cliques satisfies Assumptions 5.6 & 5.7, allowing for the application of Theorem 5.8 and the implementation of the Moment hierarchy corresponding to problem (5.8) with sparse Stokes constraints.

This yields the following formulation

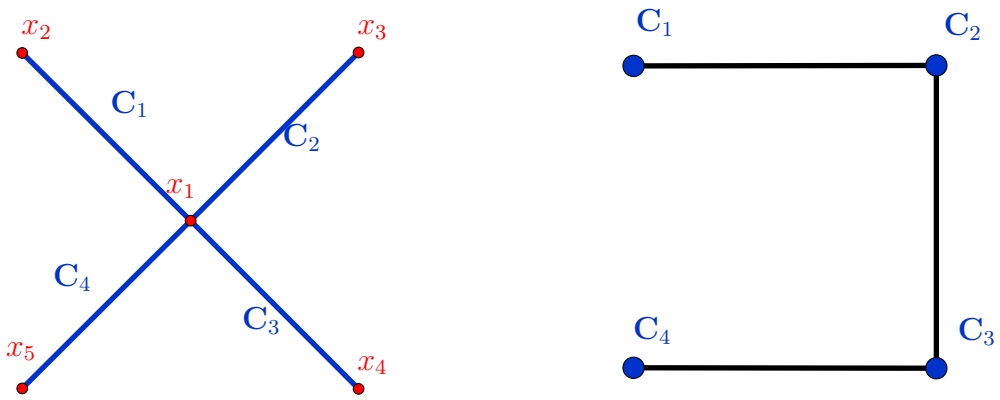
$$\begin{aligned}
\operatorname{vol}\mathbf{K} = \mathrm{p}^{\star}_{\mathrm{sp}\mathbf{K}} = \max_{(\mu_i)_i} & \int 1 \, d\mu_1 && (5.11)\\
\text{s.t. } & \mu_i \in \mathcal{M}(\mathbf{K}_i)_+ && i = 1,2,3,4\\
& d\mu_1(x_1,x_2) \preceq d\mu_2^{\langle x_1\rangle}(x_1)\,\mathrm{d}x_2\\
& d\mu_2(x_1,x_3) \preceq d\mu_3^{\langle x_1\rangle}(x_1)\,\mathrm{d}x_3\\
& d\mu_3(x_1,x_4) \preceq d\mu_4^{\langle x_1\rangle}(x_1)\,\mathrm{d}x_4\\
& d\mu_4(x_1,x_5) \preceq \mathrm{d}x_1\,\mathrm{d}x_5
\end{aligned}$$

with Stokes constraints

$$\begin{aligned}
\frac{\partial}{\partial x_2}\Big[(2x_1^2 - x_2^2 - 1)\,x_2\,(1-x_2)\Big]\, d\mu_1(x_1,x_2) &= \frac{\partial}{\partial x_2}\Big[(2x_1^2 - x_2^2 - 1)\,x_2\,(1-x_2)\, d\mu_1(x_1,x_2)\Big]\\
\frac{\partial}{\partial x_3}\Big[(2x_1^2 - x_3^2 - 1)\,x_3\,(1-x_3)\Big]\, d\mu_2(x_1,x_3) &= \frac{\partial}{\partial x_3}\Big[(2x_1^2 - x_3^2 - 1)\,x_3\,(1-x_3)\, d\mu_2(x_1,x_3)\Big]\\
\frac{\partial}{\partial x_4}\Big[(2x_1^2 - x_4^2 - 1)\,x_4\,(1-x_4)\Big]\, d\mu_3(x_1,x_4) &= \frac{\partial}{\partial x_4}\Big[(2x_1^2 - x_4^2 - 1)\,x_4\,(1-x_4)\, d\mu_3(x_1,x_4)\Big]\\
\frac{\partial}{\partial x_5}\Big[(2x_1^2 - x_5^2 - 1)\,x_5\,(1-x_5)\Big]\, d\mu_4(x_1,x_5) &= \frac{\partial}{\partial x_5}\Big[(2x_1^2 - x_5^2 - 1)\,x_5\,(1-x_5)\, d\mu_4(x_1,x_5)\Big]\\
\frac{\partial}{\partial x_1}\Big[(2x_1^2 - x_5^2 - 1)\,x_1\,(1-x_1)\Big]\, d\mu_4(x_1,x_5) &= \frac{\partial}{\partial x_1}\Big[(2x_1^2 - x_5^2 - 1)\,x_1\,(1-x_1)\, d\mu_4(x_1,x_5)\Big].
\end{aligned}$$

We can compute analytically

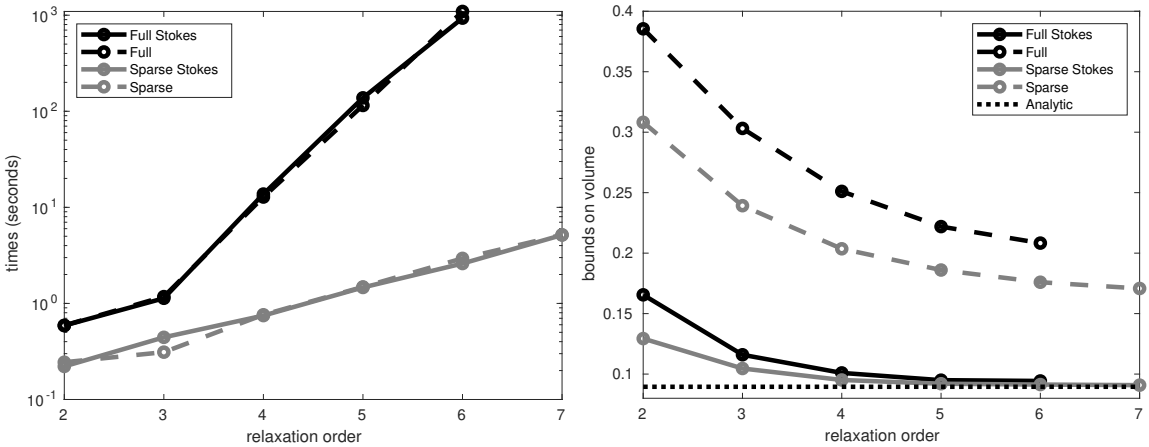
$$\operatorname{vol}\mathbf{K} = \frac{1}{15}\left(7 - 4\sqrt{2}\right) \simeq 0.0895.$$



(a) Variable graph. (b) Clique tree.

Figure 5.6 – Graph with linear clique tree for the nonconvex set.

On Figure 5.7 we show results from solving several relaxations via the dense and the sparse approach, with and without Stokes constraints. While solving with Mosek the degree 12 dense relaxation took about 1000 seconds, solving the degree 12 sparse relaxation took less than 10 seconds. With the sparse relaxations it was possible to go much higher in the hierarchy. Figure 5.7b shows convincingly how Stokes constraints accelerate the convergence of the hierarchy. We can also observe that the nonconvexity of $\mathbf{K}$ poses no special difficulty for the volume computation.



(a) Computation time vs relaxation order. (b) Bounds on the volume vs relaxation order.

Figure 5.7 – Performance for the nonconvex set.

A high dimensional polytope

Consider

$$\mathbf{K}^{(n)} := \{\mathbf{x} \in [0,1]^n \ : \ \forall i \in \mathbb{N}^{\star}_{n-1},\ x_i + x_{i+1} \leq 1\}.$$

According to [118], for any $\theta \in (-\frac{\pi}{2}, \frac{\pi}{2})$, one has the elegant formula :

$$\tan\theta + \sec\theta = 1 + \sum_{n=1}^{\infty} \operatorname{vol} \mathbf{K}^{(n)}\, \theta^n$$

which gives access to the volume of $\mathbf{K}^{(n)}$ for n arbitrarily large. For example when $n = 20$ we obtain

$$\operatorname{vol} \mathbf{K}^{(20)} = \frac{14814847529501}{97316080327065600} \approx 1.522 \cdot 10^{-4}.$$

From the SDP viewpoint, $\operatorname{vol} \mathbf{K}^{(n)}$ is computed by solving relaxations of the LP problem given in Theorem 5.8 where $N = n - 1$, $\mathbb{X}_i = \langle x_i, x_{i+1} \rangle$ and $\mathbf{g}_i(x_i, x_{i+1}) = (x_i, x_{i+1}, 1 - x_i - x_{i+1})$, $i = 1, \ldots, n-1$.

We implemented the volume computation algorithm for $n = 20$, with Stokes contraints. This cannot be achieved without resorting to sparse computation as the dimension is too high for regular SDP solvers. With the sparse formulation however we could solve relaxations up to degree 28 in less than 100 seconds, see Figure 5.8. Note however, that the analytic volume is of the order of 10^{-4}. In consequence we observe a non monotonicity of the relaxation values which contradicts the theory. This issue is surprising as the Mosek SDP solver terminates without reporting issues. This indicates that computing small volumes in large dimension can be numerically sensitive.

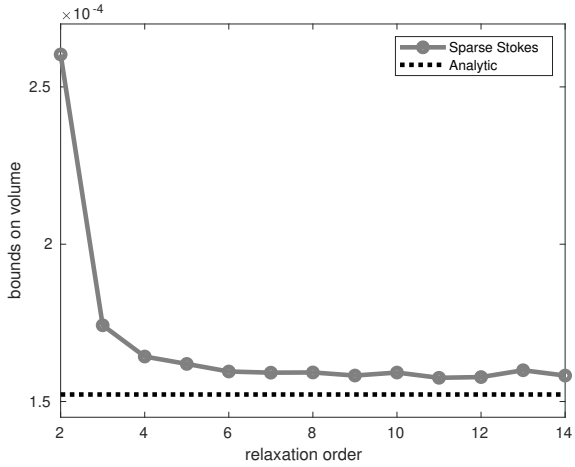


Figure 5.8 – Performance for the high dimensional polytope.

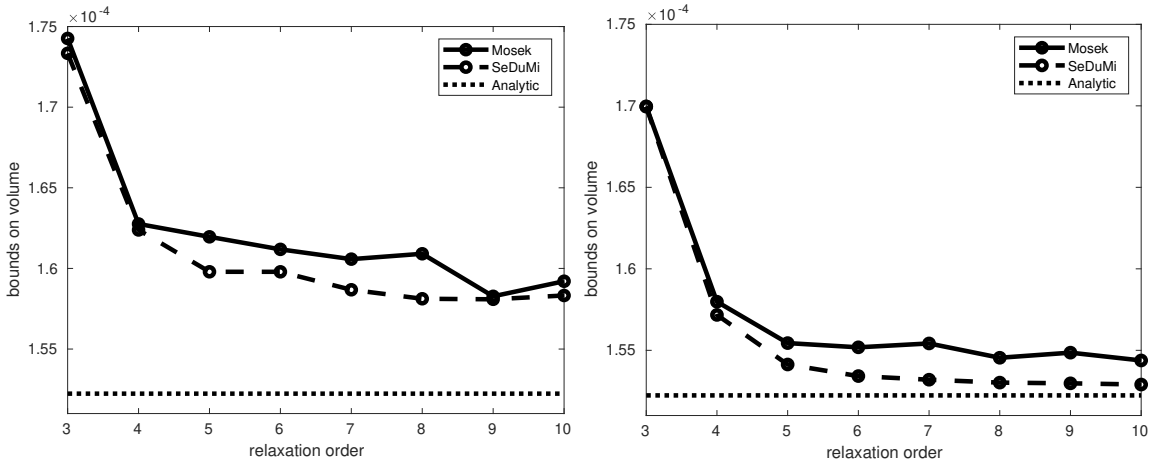
We represent the obtained bounds on the volume vs relaxation order.

In order to fix the monotonicity issue, we added a sparse rescaling to our problem. The idea is the following: at each step of the algorithm, the mass of the measure μ_i is less than the mass of the reference measure

$$\rho_i := \mu_{i+1}^{\mathbb{X}_i \cap \mathbb{X}_{i+1}} \otimes \lambda^{n_i}.$$

Defining

$$\varepsilon_i := \frac{|\mu_i|}{|\rho_i|} \in (0, 1),$$



(a) Without rescaling.

(b) With rescaling.

Figure 5.9 – Sparse rescaling performance for the high dimensional polytope.

we obtain that

$$\operatorname{vol} \mathbf{K} = \prod_{i=1}^{N} \varepsilon_i$$

as a telescoping product, since $|\rho_N| = \operatorname{vol} \mathbf{X}_N = 1$ (recall that $\mathbf{X} = [0,1]^n$ & $\mathbf{X}_N = [0,1]^{n_N}$). As a result, if N is large and the ε_i are small, one can expect the volume to be very small, which explains why the SDP solver encounters difficulties. Thus, a solution is to multiply each domination constraint by a well-chosen rescaling factor ε such that the mass of μ_i does not decrease too much. As a result, one obtains $\operatorname{vol} \mathbf{K} = \varepsilon^{N-1} \mathrm{p}^{\star\varepsilon}_{\mathrm{sp}\mathbf{K}}$, where

$$\begin{aligned} \mathrm{p}^{\star\varepsilon}_{\mathrm{sp}\mathbf{K}} := \max_{(\mu_i)_i} & \int 1 \, d\mu_1 \\ \text{s.t. } & \mu_i \in \mathcal{M}(\mathbf{K}_i)_+ && i \in \mathbb{N}^\star_N \\ & \varepsilon\, \mu_i \preceq \mu_{i+1}^{\mathbb{X}_i \cap \mathbb{X}_{i+1}} \otimes \lambda^{n_i} && i \in \mathbb{N}^\star_{N-1} \\ & \mu_N \preceq \lambda^{n_N}. \end{aligned} \tag{5.12}$$

Figure 5.9 gives a comparison between the results obtained with and without sparse rescaling, using the SDP Solvers SeDuMi and Mosek, for the choice $\varepsilon = \frac{1}{2}$.

First, one can see that without rescaling (Figure 5.9a), both SeDuMi and Mosek have accuracy issues that make them lose monotonicity, while the rescaling (Figure 5.9b) allows recovering monotonicity at least when using SeDuMi (which is slower but more accurate than Mosek to our general experience). Second, it is clear that the relative approximation error is much smaller with scaling. This, combined to the fact that the error is relative (after rescaling, the error is much smaller), demonstrates the power of our rescaling method.

A nonconvex high dimensional set

Finally, we consider the set already mentioned in Section 5.1.3, which is both nonconvex and high dimensional. Let

$$\mathbf{K}^{(n)} := \left\{ \mathbf{x} \in [0,1]^n \ : \ \forall i \in \mathbb{N}^\star_{n-1},\ x_{i+1}\ x_i \leq 1/2 \right\}$$

whose analytic volume is a function of the dimension n. For $n = 3$ the analytic volume is 0.75, for $n = 4$ it is 0.6566, approximately. In higher dimensions we do not have an analytic expression for the volume. However, in order to get a feeling for its value for bigger n, we ran a Monte Carlo simulation with one million samples for $n = 10, 20, 50$, and 100.

Remark 5.4 (On Monte Carlo simulations)

We describe the very basic Monte Carlo approach used here. Let $\mathrm{X}_1, \ldots, \mathrm{X}_N$ *be i.i.d. samples from some law* μ*. In our case* μ *is the uniform distribution on* $[0,1]^n$*. Further let* f *be a function from the probability space into* $\{0,1\}$*. Again, in our case,* f *would return* 1 *if the sample* X_i *is in the set* $\mathbf{K}^{(n)}$*, and* 0 *else. By the strong law of large numbers*

$$\hat{\mathrm{F}}_N := \frac{1}{N} \sum_{i=1}^{N} f(\mathrm{X}_i) \underset{N\to\infty}{\longrightarrow} \int f \ d\mu = \operatorname{vol} \mathbf{K}^{(n)}.$$

This makes $\hat{\mathrm{F}}_N$ *a reasonable guess if* N *is large. However,* $\hat{\mathrm{F}}_N$ *is still a guess, and it might happen that* $\hat{\mathrm{F}}_N$ *is actually far away from the approximated quantity.*

As a consequence of the Central Limit Theorem, the difference $\hat{\mathrm{F}}_N - \int f \ d\mu$ *behaves (almost) like a normal distributed random variable with zero mean and variance* σ^2/N *where* $\sigma^2 = \int (f - \int f d\ \mu)^2 \ d\mu$*. Note that the variance* σ^2 *can also be estimated based on the i.i.d. samples* $\mathrm{X}_1, \ldots, \mathrm{X}_N$*:*

$$\hat{\mathrm{S}}_N^2 := \frac{1}{N-1} \sum_{i=1}^{N} (\mathrm{X}_i - \hat{\mathrm{F}}_N)^2.$$

This allows one to define a confidence interval for the approximated volume. Indeed, say we are interested in a 99%-confidence interval. Then for G *a standard normal distributed random variable, we have* $\mathbb{P}(|\mathrm{G}| < 2.58) \approx 0.99$ *and consequently,*

$$\mathbb{P}\left(\hat{\mathrm{F}}_N - \frac{2.58 \hat{\mathrm{S}}_N}{\sqrt{N}} \leq \int f d\mu \leq \hat{\mathrm{F}}_N + \frac{2.58 \hat{\mathrm{S}}_N}{\sqrt{N}} \right) \approx 0.99.$$

Remark 5.5 (Moment-SOS hierarchy VS Monte-Carlo)

Before we go on, let us emphasize that the method proposed in this chapter is not in competition with the Monte Carlo approach. While the Monte Carlo method gives a probabilistic estimate of the volume, our method provides a guaranteed upper bound. Nonetheless, it would be disconcerting if the computed upper bound were much smaller than the lower bound of the confidence interval, and we consider our approximation valid, when it returns something in the order of the Monte Carlo approximation. The results for different dimensions n *and solved with the Mosek SDP solver are summarized in Table 5.2.*

As in the previous section we experience accuracy issues for the relaxations of order 14 and $n = 20, 100$, as well as for order 16 and $n = 50$. Otherwise, the approximations provide better upper bounds for increased relaxation orders as expected. For $n = 3, 4$ the approximation is reasonably close to the analytic value. For $n = 10, 20, 50$ our scheme provides an upper bound in the same order of magnitude as the 99%-confidence interval of the Monte Carlo simulation. We interpret this as a validation for both the Monte Carlo approach and our own approach. For $n = 100$ we could not derive a meaningful confidence interval. Indeed, as our approximation shows, the volume for $n = 100$ is less than $9 \cdot 10^{-6}$. In order to get an accuracy of $\varepsilon = 10^{-6}$ one would have to draw approximately $N = \frac{1}{\varepsilon^2} = 10^{12}$ samples. With our non-sophisticated implementation, the Monte Carlo simulation for one million points took about 5 seconds. Extending this linearly to a simulation with 10^{12} samples would therefore take a little less than 2 months ($5 \cdot 10^6$ $s \simeq 1389$ $h \simeq 58$ d). With more sophisticated methods, this time could certainly be reduced dramatically. However, it sets the 44 minutes it took to solve relaxation order 16 for $n = 100$ into perspective.

5.3 Exploiting correlative sparsity

5.3.1 General correlative sparsity pattern

Let us describe a general method to compute the volume of

$$\mathbf{K} := \bigcap_{i=1}^{N} \mathbf{U}_i$$

where $\mathbf{U}_i = \{\mathbf{x} \in \mathbb{X} : \mathbf{g}_i(\mathbf{x}) \geq \mathbf{0}\}$ and $(\mathbf{g}_1, \ldots, \mathbf{g}_N)$ is a correlatively sparse family of polynomial vectors with associated coordinate subspace decomposition $\mathbb{X} = \sum_{i=1}^{N} \mathbb{X}_i$. For this we construct the correlation graph $\mathbb{G} = (\mathbf{V}, \mathbf{E})$ as follows:

- $\mathbf{V} = \{1 \ldots, n\}$ represents the canonical basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ of $\mathbb{X} = \mathbb{R}^n$;
- $\mathbf{E} = \{(i, j) \in \{1, \ldots, n\}^2 \ : \ i \neq j \ \& \ \mathbf{e}_i, \mathbf{e}_j \in \mathbb{X}_k \text{ for some } k \in \mathbb{N}_N^\star$.

As stated in Section 5.1.4, we suppose that the correlation graph of $(\mathbf{g}_i)_{1 \leq i \leq N}$ has exactly N maximal cliques (see Section 5.3.4 for discussions when it is not the case) that are in correspondence with the $\mathbb{X}_i$.

Let $\mathcal{K}$ be the set of maximal cliques of $\mathbb{G}$. We will use the following property of graphs:

Definition 5.11: Clique Intersection Property (CIP)

The graph $\mathbb{G} = (\mathbf{V}, \mathbf{E})$ is said to satisfy the *clique intersection property (CIP)* *iff* there is a clique tree $\mathbb{T} = (\mathcal{K}, \mathcal{E})$, $\mathcal{E} \subset \mathcal{K}^2$, such that for all $\mathbf{C}, \mathbf{C}' \in \mathcal{K}$, $\mathbf{C} \cap \mathbf{C}' \subset \mathbf{C}''$ for any $\mathbf{C}''$ on the path connecting $\mathbf{C}$ and $\mathbf{C}'$ in the tree $\mathbb{T}$.

	n=3		n = 4		n=10	
d	value	time (s)	value	time (s)	value	time (s)
4	7.86E-01	0.95	7.09E-01	0.61	3.93E-01	1.24
5	7.73E-01	2.87	6.90E-01	0.78	3.57E-01	1.84
6	7.69E-01	2.74	6.84E-01	0.86	3.45E-01	4.21
7	7.66E-01	4.58	6.79E-01	1.83	3.38E-01	4.55
8	7.63E-01	5.00	6.77E-01	2.29	3.34E-01	5.97
9	7.63E-01	6.11	6.74E-01	3.33	3.30E-01	11.56
10	7.62E-01	9.83	6.73E-01	6.86	3.26E-01	18.21
11	7.61E-01	18.16	6.72E-01	8.57	3.26E-01	22.24
12	7.60E-01	19.45	6.71E-01	10.43	3.23E-01	33.78
13	7.60E-01	22.49	6.70E-01	17.89	3.22E-01	74.00
14	7.60E-01	27.02	6.69E-01	26.84	3.21E-01	79.68
15	7.59E-01	32.90	6.69E-01	39.25	3.20E-01	119.7
16	7.58E-01	78.20	6.68E-01	61.32	3.19E-01	176.6
ana/mc	7.50E-01	-	6.57E-01	-	[2.99e-01, 3.03e-01]	

	n = 20		n=50		n = 100	
d	value	time (s)	value	time (s)	value	time (s)
4	1.47E-01	5.11	7.68E-03	10.64	9.49E-05	15.19
5	1.20E-01	3.58	4.78E-03	15.10	4.80E-05	26.87
6	1.11E-01	8.13	3.86E-03	21.75	2.84E-05	49.89
7	1.07E-01	11.06	3.42E-03	48.31	2.27E-05	77.07
8	1.03E-01	17.93	3.20E-03	72.78	1.91E-05	135.08
9	1.00E-01	33.31	2.99E-03	120.49	1.63E-05	202.12
10	9.81E-02	41.02	2.99E-03	103.61	1.44E-05	299.44
11	9.70E-02	85.83	2.89E-03	165.56	1.22E-05	441.67
12	9.59E-02	117.08	2.77E-03	220.84	1.19E-05	623.24
13	9.51E-02	138.38	2.67E-03	314.00	1.08E-05	850.92
14	9.57E-02	156.32	2.60E-03	457.92	1.10E-05	1175.02
15	9.39E-02	249.82	2.54E-03	685.64	9.86E-06	1589.49
16	9.36E-02	357.87	2.56E-03	859.60	9.46E-06	2623.02
ana/mc	[8.09e-02, 8.24e-02]		[1.48e-03, 1.68e-03]		-	

Table 5.2 – Performances on a nonconvex high dimensional set.

ana/mc refers to the analytic value and the 99%-confidence interval, respectively.

Such a property is equivalent to the fact that $\mathbb{G}$ is chordal[4], see [13]. We then replace Assumption 5.7 with the following strong correlative sparsity assumption:

Assumption 5.12: Disjoint Intersection Property (DIP)

We suppose that there is a clique tree $\mathbb{T} = (\mathcal{K}, \mathcal{E})$, rooted in some $\mathbf{C}_1$, that simultaneously satisfies the CIP and the following *disjoint intersection property (DIP):* $\forall \mathbf{C}, \mathbf{C}', \mathbf{C}'' \in \mathcal{K}$, if $(\mathbf{C}, \mathbf{C}') \in \mathcal{E}$ and $(\mathbf{C}, \mathbf{C}'') \in \mathcal{E}$ then $\mathbf{C}' = \mathbf{C}''$ or $\mathbf{C}' \cap \mathbf{C}'' = \varnothing$.

In words, each clique has an empty intersection with all its siblings. See Section 5.3.4 for details on how to check this assumption and construct such a tree when it exists, as well as possible solutions when Assumption 5.12 does not hold. Figure 5.10 illustrates the meaning of this assumption for $n = 12$ and $N = 8$. One can check that Assumptions 5.6 and 5.12 hold.

Remark 5.6 (On Assumption 5.12)

With these assumptions, the only possible clique trees for applying our method to the nonconvex example illustrated in Figure 5.6 are linear clique trees. Indeed, any branched clique tree would imply sibling cliques containing x_1.

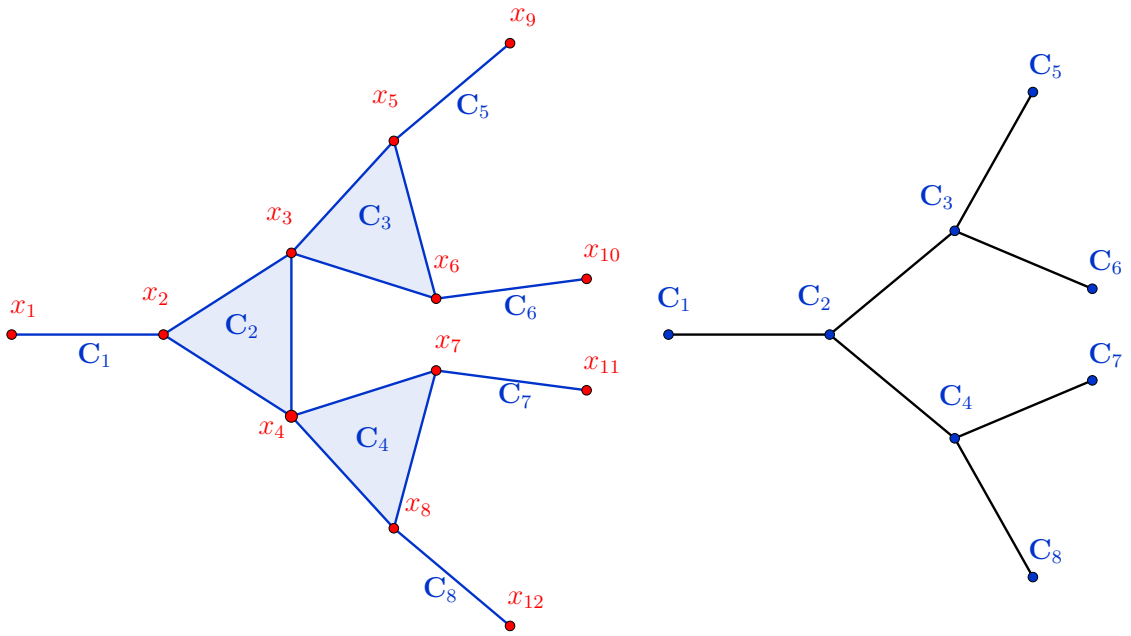


Figure 5.10 – Chordal graph (left) with its clique tree (right).

[4]Thus the CIP always holds, up to a chordal extension. In particular, cyclic graphs can be handled with empty interactions between well-chosen variables.

5.3.2 Distributed computation theorem

One can formulate a simple generalization of the sequential implementation of Theorem 5.8 to our general correlative sparsity pattern.

Theorem 5.13: Distributed computation

Let Assumptions 5.6 and 5.12 hold. Let $\mathbb{T} = (\mathcal{K}, \mathcal{E})$ be a clique tree as in Assumption 5.12. Then $\operatorname{vol} \mathbf{K} = \int 1 \, d\mu_1^\star$ where for $i \in \mathbb{N}_N^\star$, $\mu_i^\star$ is an optimal solution to the following GMP:

$$\begin{aligned} \mathrm{p}_{\mathrm{sp}\mathbf{K},i}^\star := \max_{\mu_i} & \int 1 \, d\mu_i \\ \text{s.t. } & \mu_i \in \mathcal{M}(\mathbf{K}_i)_+ && \text{(5.13a)} \\ & \mu_i \preceq \left(\bigotimes_{(\mathbf{C}_i, \mathbf{C}_j) \in \mathcal{E}} \mu_j^{\star \mathbb{X}_i \cap \mathbb{X}_j} \right) \otimes \lambda^{n_i} && \text{(5.13b)} \end{aligned}$$

and $n_i = \dim \mathbb{X}_i \cap \left(\sum_{(\mathbf{C}_i, \mathbf{C}_j) \in \mathcal{E}} \mathbb{X}_j \right)^\perp = \left| \mathbf{C}_i \cap \left(\bigcup_{(\mathbf{C}_i, \mathbf{C}_j) \in \mathcal{E}} \mathbf{C}_j \right)^c \right|$.

Proof : We define, for $i \in \mathbb{N}_N^\star$,

$$\mathbb{Y}_i := \mathbb{X}_i \cap \left(\sum_{(\mathbf{C}_i, \mathbf{C}_j) \in \mathcal{E}} \mathbb{X}_j \right)^\perp = \langle x_k \rangle_{k \in \mathbf{C}_i \ ; \ (\mathbf{C}_i, \mathbf{C}_j) \in \mathcal{E} \Rightarrow k \notin \mathbf{C}_j}$$

and we observe that, according to Assumption 5.12, for any $i \in \mathbb{N}_N^\star$

$$\mathbb{X}_i = \left(\bigoplus_{(\mathbf{C}_i, \mathbf{C}_j) \in \mathcal{E}} \mathbb{X}_i \cap \mathbb{X}_j \right) \oplus \mathbb{Y}_i.$$

Thus, constraint (5.13b) is well-posed.

For $i \in \mathbb{N}_N^\star$, let $\mathcal{D}(i) := \{ j > i \ : \ \exists$ an oriented path from $\mathbf{C}_i$ to $\mathbf{C}_j$ in $\mathbb{T} \}$ be the set of descendants of $\mathbf{C}_i$, as well as $\mathbb{W}_i := \sum_{j \in \mathcal{D}(i)} \mathbb{X}_j = \langle x_k \rangle_{k \in \mathbf{C}_j \ ; \ j \in \mathcal{D}(i)}$ and $r_i := \dim \mathbb{W}_i = \left| \bigcup_{j \in \mathcal{D}(i)} \mathbf{C}_j \right|$.

We are going to show by induction that for $i \in \mathbb{N}_N^\star$,

$$\mu_i^\star = \mathbb{1}_{\mathbf{K}_i} \left(\prod_{j \in \mathcal{D}(i)} \mathbb{1}_{\mathbf{K}_j} \circ \pi_{\mathbb{X}_j} \, \lambda^{r_i} \right)^{\mathbb{X}_i \cap \mathbb{W}_i} \otimes \lambda^{n_i}.$$

Our base cases are the leaves of $\mathbb{T}$, i.e. the i such that $\mathcal{D}(i) = \varnothing$. Then, problem (5.13) is reduced to the classical problem of computing the volume of $\mathbf{K}_i$, whose optimal solution is exactly $\mu_i^\star = \lambda_{\mathbf{K}_i} = \mathbb{1}_{\mathbf{K}_i} \, \lambda^{n_i}$ (because $D(i) = \varnothing \Rightarrow n_i = \dim \mathbb{X}_i$), which is the expected result.

Then we can proceed to the induction: let i be a node of $\mathbb{T}$ that is not a leaf: $\mathcal{D}(i) \neq \varnothing$; and suppose that for $j \in \mathcal{D}(i)$ such that $(\mathbf{C}_i, \mathbf{C}_j) \in \mathcal{E}$,

$$\mu_j^\star = \mathbb{1}_{\mathbf{K}_j} \left(\prod_{k \in \mathcal{D}(j)} \mathbb{1}_{\mathbf{K}_k} \circ \pi_{\mathbb{X}_k} \, \lambda^{q_j} \right)^{\mathbb{X}_j \cap \mathbb{W}_j} \otimes \lambda^{n_j}.$$

Then, constraint (5.13b) can be rewritten as

$$\mu_i \preceq \left(\bigotimes_{(\mathbf{C}_i, \mathbf{C}_j) \in \mathcal{E}} \left(\mathbb{1}_{\mathbf{K}_j} \left(\prod_{k \in \mathcal{D}(j)} \mathbb{1}_{\mathbf{K}_k} \circ \pi_{\mathbb{X}_k} \, \lambda^{q_j} \right)^{\mathbb{X}_j \cap \mathbb{W}_j} \otimes \lambda^{n_j} \right)^{\mathbb{X}_i \cap \mathbb{X}_j} \right) \otimes \lambda^{n_i}$$

which in turn is simplified into

$$\mu_i \preceq \left(\bigotimes_{(\mathbf{C}_i, \mathbf{C}_j) \in \mathcal{E}} \left(\mathbb{1}_{\mathbf{K}_j} \left(\prod_{k \in \mathcal{D}(j)} \mathbb{1}_{\mathbf{K}_k} \circ \pi_{\mathbb{X}_k} \right) \lambda^{n_j + q_j} \right)^{\mathbb{X}_i \cap \mathbb{W}_j} \right) \otimes \lambda^{n_i}$$

since the CIP ensures that $\mathbb{X}_i \cap \mathbb{X}_j \cap \mathbb{W}_j = \mathbb{X}_i \cap \mathbb{W}_j$: indeed $\mathbf{C}_j$ is on the path between $\mathbf{C}_i$ and any $\mathbf{C}_k$ with $k \in \mathcal{D}(j)$, so that $\mathbf{C}_i \cap \mathbf{C}_k \subset \mathbf{C}_j$ and thus $\mathbb{X}_i \cap \mathbb{X}_k \subset \mathbb{X}_j$, yielding $\mathbb{X}_i \cap \mathbb{W}_j \subset \mathbb{X}_j$. At this point one can notice that $n_j + q_j = \dim(\mathbb{X}_j + \mathbb{W}_j)$.

Then, using the DIP, we know that if $(\mathbf{C}_i, \mathbf{C}_j), (\mathbf{C}_i, \mathbf{C}_k) \in \mathcal{E}$ with $j \neq k$ then $\mathbf{C}_j \cap \mathbf{C}_k = \varnothing$ and with the CIP $\mathbf{C}_j \cap \mathbf{C}_l = \varnothing$ for any $l \in \mathcal{D}(k)$. This yields that $(\mathbb{X}_j + \mathbb{W}_j) \cap (\mathbb{X}_k + \mathbb{W}_k) = \{\mathbf{0}\}$ and thus $\mathbb{W}_i = \bigoplus_{(\mathbf{C}_i, \mathbf{C}_j) \in \mathcal{E}} (\mathbb{X}_j + \mathbb{W}_j)$, allowing us to rewrite constraint (5.13b) as

$$\mu_i \preceq \left(\bigotimes_{(\mathbf{C}_i, \mathbf{C}_j) \in \mathcal{E}} \mathbb{1}_{\mathbf{K}_j} \left(\prod_{k \in \mathcal{D}(j)} \mathbb{1}_{\mathbf{K}_k} \circ \pi_{\mathbb{X}_k} \right) \lambda^{n_j + q_j} \right)^{\mathbb{X}_i \cap \mathbb{W}_i} \otimes \lambda^{n_i},$$

which simplifies into

$$\mu_i \preceq \left(\prod_{j \in \mathcal{D}(i)} \mathbb{1}_{\mathbf{K}_j} \circ \pi_{\mathbb{X}_j} \, \lambda^{r_i} \right)^{\mathbb{X}_i \cap \mathbb{W}_i} \otimes \lambda^{n_i}.$$

Eventually, we are again faced to a classical instance of the dense volume problem for $\mathbf{K}_i$, with $\left(\prod_{j \in \mathcal{D}(i)} \mathbb{1}_{\mathbf{K}_j} \circ \pi_{\mathbb{X}_j} \, \lambda^{r_i} \right)^{\mathbb{X}_i \cap \mathbb{W}_i} \otimes \lambda^{n_i}$ instead of only the uniform Lebesgue measure, and we know that the optimal solution is obtained by multiplying this non-negative dominating measure with the indicator of $\mathbf{K}_i$, yielding

$$\mu_i^\star = \mathbb{1}_{\mathbf{K}_i} \left(\prod_{j \in \mathcal{D}(i)} \mathbb{1}_{\mathbf{K}_j} \circ \pi_{\mathbb{X}_j} \, \lambda^{r_i} \right)^{\mathbb{X}_i \cap \mathbb{W}_i} \otimes \lambda^{n_i}$$

which is the announced result.

We conclude by using the fact that $\mathcal{D}(1) = \{2, \dots, N\}$ and $\mathbb{R}^n = \mathbb{X}_1 + \mathbb{W}_1 = \mathbb{X}_1 \oplus (\mathbb{X}_1^\perp \cap \mathbb{W}_1)$ to compute the value:

$$\begin{aligned}
\int 1 \, d\mu_1^\star &= \int \mathbb{1}_{\mathbf{K}_1} \left(\prod_{j=2}^{N} \mathbb{1}_{\mathbf{K}_j} \circ \pi_{\mathbb{X}_j} \, \lambda^{q_1} \right)^{\mathbb{X}_1 \cap \mathbb{W}_1} d\lambda^{n_1} \\
&= \int_{\mathbb{X}_1} \mathbb{1}_{\mathbf{K}_1}(\mathbf{x}_1) \left(\int_{\mathbb{X}_1^\perp \cap \mathbb{W}_1} \prod_{j=2}^{N} \mathbb{1}_{\mathbf{U}_j}(\mathbf{x}_1 + \mathbf{w}_1) \, \mathrm{d}\mathbf{w}_1 \right) \mathrm{d}\mathbf{x}_1 \\
&= \int_{\mathbb{R}^n} \left(\prod_{i=1}^{N} \mathbb{1}_{\mathbf{U}_i}(\mathbf{x}) \right) \mathrm{d}\mathbf{x} \\
&= \int_{\mathbb{R}^n} \mathbb{1}_{\mathbf{K}}(\mathbf{x}) \, \mathrm{d}\mathbf{x} \\
&= \text{vol } \mathbf{K}.
\end{aligned}$$

♢

Therefore one obtains a sequence of infinite dimensional LPs on measures that can be algorithmically addressed using the usual SDP relaxations. The computations start from the leaves of the clique tree and proceed down to the root. It is worth noting that all the maximal cliques of the same generation in the tree are totally independent, which allows treating them simultaneously, *i.e.* to partially parallelize the computations. Let $d \in \mathbb{N}$. We consider the solutions $\mathbf{z}_i^{(d)}$ to the degree d moment relaxations corresponding to problems (5.13), for $i \in \mathbb{N}_N^\star$, as well as the relaxation values $\mathrm{p}_{\mathrm{sp}\mathbf{K},i}^d$. We are going to study the convergence of the sequence $\mathrm{p}_{\mathrm{sp}\mathbf{K},1}^d$ to vol $\mathbf{K}$.

Theorem 5.14: Convergence of the branched Moment-SOS hierarchy

Suppose that Assumption 2.7 holds for the $\mathbf{K}_i$ and the $\mathbf{X}_i$. Let $\mathbf{C}_i \in \mathcal{K}$ such that for $\mathbf{C}_j \in \mathcal{K}$ satisfying $(\mathbf{C}_i, \mathbf{C}_j) \in \mathcal{E}$, we have a converging sequence of pseudo-moment vectors $(\mathbf{z}_j^{(d)})_d$: for any appropriate multi-index $\mathbf{k}$, $z_{j,\mathbf{k}}^{(d)} \underset{d\to\infty}{\longrightarrow} \int_{\mathbb{X}_j} \mathbf{x}_j^{\mathbf{k}} \, d\mu_j^\star(\mathbf{x}_j)$.

In this setting, if each relaxation of the i-th LP problem (5.13) has at least one feasible solution, then the corresponding Moment hierarchy converges: for any appropriate multi-index $\mathbf{k}$, $z_{i,\mathbf{k}}^{(d)} \underset{d\to\infty}{\longrightarrow} \int \mathbf{x}_i^{\mathbf{k}} \, d\mu_i^\star(\mathbf{x}_i)$.

Thus, by induction, if at all nodes of $\mathbb{T}$ the moment relaxations *remain feasible at all degrees* of relaxation, then the branched Moment hierarchy converges, namely $\mathrm{p}_{\mathrm{sp}\mathbf{K},1}^d \underset{d\to\infty}{\longrightarrow} \mathrm{p}_{\mathrm{sp}\mathbf{K},1}^\star = \text{vol } \mathbf{K}$.

Proof : The feasibility assumption ensures that the $\mathbf{z}_i^{(d)}$ are properly defined at all degrees d. Then, pointwise convergence of the $(\mathbf{z}_j^{(d)})_d$ yields that it is bounded for the weak-$*$ topology on $\mathbb{R}[\mathbf{x}]'$. The moment relaxation of constraint (5.13b) combined with Lemma 2.9 then yield that $(\mathbf{z}_i^{(d)})_d$ is bounded for the weak-$*$ topology on $\mathbb{R}[\mathbf{x}]'$, which means, according to the Banach-Alaoglu theorem, that it has an accumulation point. Finally, by uniqueness of the solution to the infinite dimensional LP problem

(5.13), the convergence of $(\mathbf{z}_j^{(d)})_d$ to the moment sequence of $\mu_j^\star$ ensures that this accumulation point is none other than the moment sequence of $\mu_i^\star$. This proves existence and uniqueness of the accumulation point of $(\mathbf{z}_i^{(d)})_d$. Then, we get for any appropriate multi-index $\mathbf{k}$

$$z_{i,\mathbf{k}}^{(d)} \underset{d\to\infty}{\longrightarrow} \int \mathbf{x}^{\mathbf{k}}\, d\mu_i^\star(\mathbf{x}).$$

To conclude for the global convergence of the sparse scheme, we just need to check the base case of this induction. Here again the base case is the leaves of the tree at which we are faced to standard instances of the volume problem, whose associated Moment-SOS hierarchy is already proved to converge. Thus, our convergence assumption is satisfied, which means that as long as all the relaxations are feasible, their solutions converge weakly-$*$ to the infinite dimensional optimal measures, and in particular

$$\mathrm{p}_{\mathrm{sp}\mathbf{K},1}^{d} \underset{d\to\infty}{\longrightarrow} \mathrm{p}_{\mathrm{sp}\mathbf{K},1}^{\star} = \mathrm{vol}\ \mathbf{K}.$$

♢

Remark 5.7 *Stokes constraints can be implemented similarly to the linear case.*

5.3.3 Distributed computation examples

6D polytope

Let $\mathbb{X} := \mathbb{R}^6$ and $\mathbb{X}_1 = \langle x_1, x_2\rangle$, $\mathbb{X}_2 = \langle x_2, x_3, x_4\rangle$, $\mathbb{X}_3 = \langle x_3, x_5\rangle$, $\mathbb{X}_4 = \langle x_4, x_6\rangle$. For $i = 1, 3, 4$ let $\mathbf{g}_i(x, y) := (x, y, 1 - x - y)$ and

$$\mathbf{K}_i := \mathbf{g}_i^{-1}\left(\mathbb{R}_+^2\right) = \{(u, v) \in [0, 1]^2 \ :\ u + v \le 1\}.$$

Let $\mathbf{g}_2(x, y, z) := (x, y, z, 1 - x - y - z)$ and

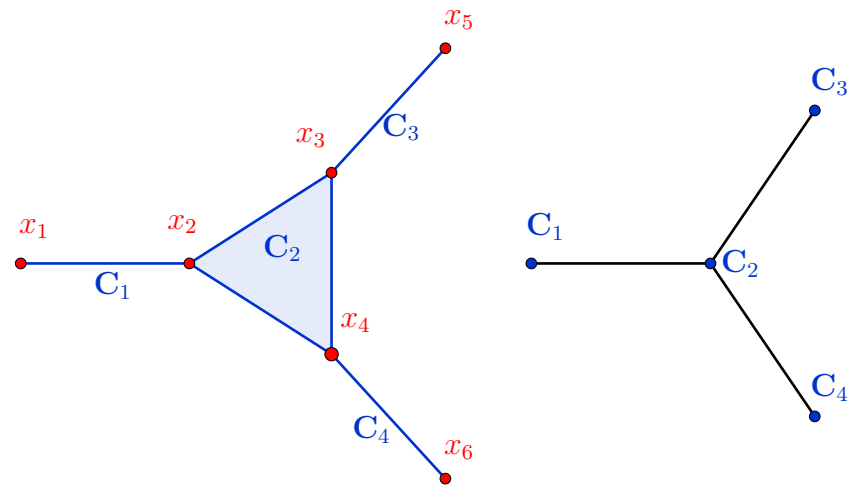
$$\mathbf{K}_2 := \mathbf{g}_2^{-1}\left(\mathbb{R}_+^3\right) = \{(x, y, z) \in [0, 1]^3 \ :\ x + y + z \le 1\}.$$

Let us approximate the volume of the 6D polytope

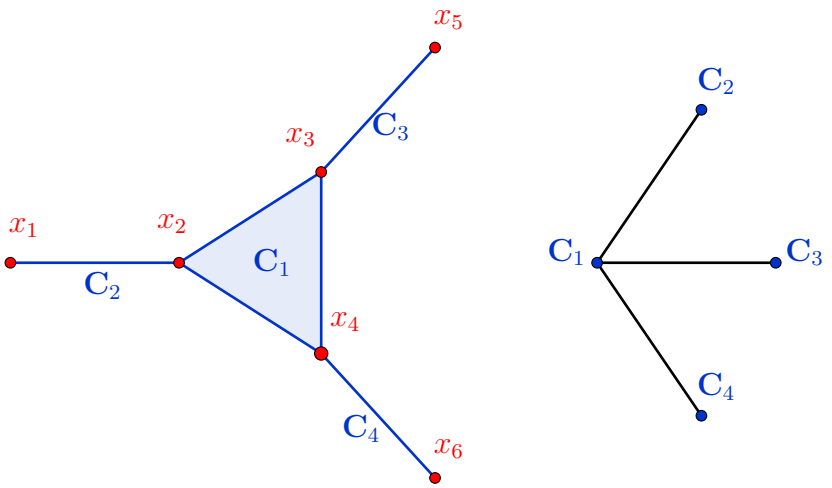
$$\mathbf{K} := \left\{ \mathbf{x} \in \mathbb{R}_+^6 \ :\ \begin{array}{l} x_1 + x_2 \le 1 \\ x_2 + x_3 + x_4 \le 1 \\ x_3 + x_5 \le 1 \\ x_4 + x_6 \le 1 \end{array} \right\} = \bigcap_{i=1}^{4} \pi_{\mathbb{X}_i}^{-1}(\mathbf{K}_i)\,.$$

No linear clique tree is associated to this problem through Proposition 5.12. The only possible clique trees for applying our method are the two branched clique trees of Figure 5.11.

Let us compare the performance of the algorithms derived from the two possible clique tree configurations and with the dense problem. For that, we first write the



(a) 3 step clique tree.



(b) 2 step clique tree.

Figure 5.11 – Two possible branched clique trees for the 6D polytope.

problem associated with the 3 step clique tree configuration of the top of Figure 5.11:

$$\text{vol}\,\mathbf{K} = \int 1\, d\mu_1^\star \tag{5.14}$$

where

$$\begin{aligned}
\mu_1^\star = \underset{\mu_1}{\operatorname{argmax}} & \int 1 \, d\mu_1 \\
& \text{s.t. } \mu_1 \in \mathcal{M}(\mathbf{K}_1)_+ \\
& \quad d\mu_1(x_1, x_2) \preceq \mathrm{d}x_1 \, d\mu_2^{\star\langle x_2 \rangle}(x_2) \\
\mu_2^\star = \underset{\mu_2}{\operatorname{argmax}} & \int 1 \, d\mu_2 \\
& \text{s.t. } \mu_2 \in \mathcal{M}(\mathbf{K}_2)_+ \\
& \quad d\mu_2(x_2, x_3, x_4) \preceq \mathrm{d}x_2 \, d\mu_3^{\star\langle x_3 \rangle}(x_3) \, d\mu_4^{\star\langle x_4 \rangle}(x_4) \\
\mu_i^\star = \underset{\mu_3}{\operatorname{argmax}} & \int 1 \, d\mu_i, \qquad\qquad i = 3, 4 \\
& \text{s.t. } \mu_i \in \mathcal{M}(\mathbf{K}_i)_+ \\
& \quad \mu_i \preceq \lambda^2.
\end{aligned}$$

This problem can be complemented with the following Stokes constraints:

$$\begin{aligned}
\frac{\partial}{\partial x_1}\Big[x_1\,(1-x_1-x_2)\Big]\; d\mu_1(x_1,x_2) &= \frac{\partial}{\partial x_1}\left[x_1\,(1-x_1-x_2)\; d\mu_1(x_1,x_2)\right] \\
\frac{\partial}{\partial x_2}\Big[x_2\,(1-x_2-x_3-x_4)\Big]\; d\mu_2(x_2,x_3,x_4) &= \frac{\partial}{\partial x_2}\left[x_2\,(1-x_2-x_3-x_4)\; d\mu_2(x_2,x_3,x_4)\right] \\
\frac{\partial}{\partial x_i}\Big[x_i\,(1-x_i-x_{i+2})\Big]\; d\mu_i(x_i,x_{i+2}) &= \frac{\partial}{\partial x_i}\left[x_i\,(1-x_i-x_{i+2})\; d\mu_i(x_i,x_{i+2})\right] \\
\frac{\partial}{\partial x_{i+2}}\Big[x_{i+2}\,(1-x_i-x_{i+2})\Big]\; d\mu_i(x_i,x_{i+2}) &= \frac{\partial}{\partial x_{i+2}}\left[x_{i+2}\,(1-x_i-x_{i+2})\; d\mu_i(x_i,x_{i+2})\right] \quad i = 3, 4.
\end{aligned}$$

The 2 step clique tree of the bottom of Figure 5.11 yields the following formulation

$$\operatorname{vol} \mathbf{K} = \int 1 \, d\mu_2^\star \tag{5.15}$$

where

$$\begin{aligned}
\mu_2^\star = \underset{\mu_2}{\operatorname{argmax}} & \int 1 \, d\mu_2 \\
& \text{s.t. } \mu_2 \in \mathcal{M}(\mathbf{K}_2)_+ \\
& \quad d\mu_2(x_2, x_3, x_4) \preceq d\mu_1^{\star\langle x_2 \rangle}(x_2) \, d\mu_3^{\star\langle x_3 \rangle}(x_3) \, d\mu_4^{\star\langle x_4 \rangle}(x_4) \\
\mu_i^\star = \underset{\mu_i}{\operatorname{argmax}} & \int 1 \, d\mu_i, \qquad\qquad i = 1, 3, 4 \\
& \text{s.t. } \mu_i \in \mathcal{M}(\mathbf{K}_i)_+ \\
& \quad \mu_i \preceq \lambda^2,
\end{aligned}$$

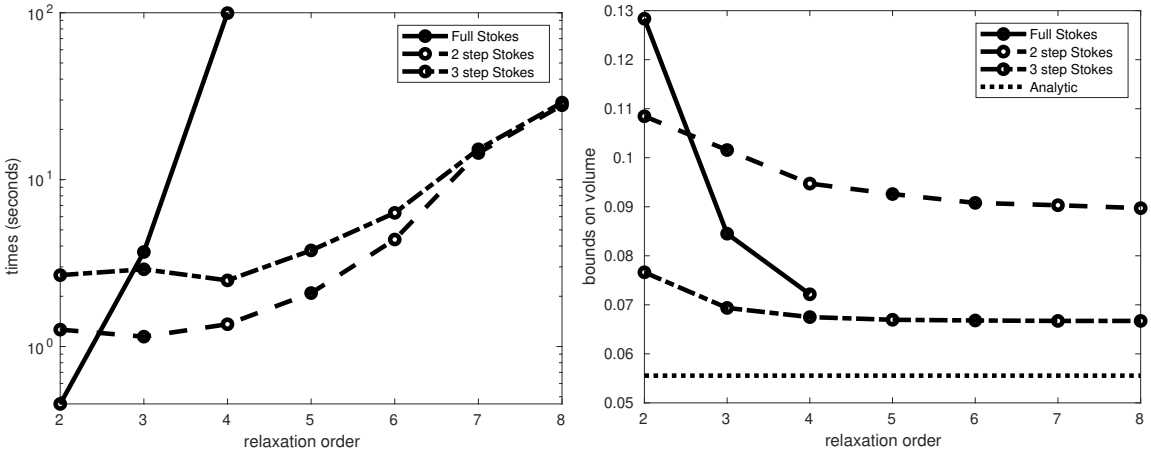
with Stokes constraints

$$\begin{aligned}
\frac{\partial}{\partial x_i}\Big[x_i\,(1-x_i-x_j)\Big]\; d\mu_i(x_i,x_j) &= \frac{\partial}{\partial x_i}\left[x_i\,(1-x_i-x_j)\; d\mu_i(x_i,x_j)\right] \\
\frac{\partial}{\partial x_j}\Big[x_j\,(1-x_i-x_j)\Big]\; d\mu_i(x_i,x_j) &= \frac{\partial}{\partial x_j}\left[x_j\,(1-x_i-x_j)\; d\mu_i(x_i,x_j)\right]
\end{aligned}$$

for $(i,j) = (1,2),(3,5),(4,6)$. However, no Stokes constraints can be applied for the computation of μ_2 (there is no Lebesgue measure in the domination constraint, so the optimal measure is not uniform). For this reason one can expect a slower convergence than in the linear configuration.

We implement the hierarchies associated to the 2 and 3 step sparse formulations, as well as the dense problem hierarchy, and compare their performance in Figure 5.12. We can compute analytically

$$\operatorname{vol} \mathbf{K} = \frac{1}{18} \simeq 0.0556.$$



(a) Computation time vs relaxation order. (b) Bounds on the volume vs relaxation order.

Figure 5.12 – Performance for the 6D polytope.

Both sparse formulations outperform the dense one in terms of computational time needed to solve the corresponding SDPs (Figure 5.12a). On the accuracy side however (Figure 5.12b), we observe that the 2 step formulation seems to be less efficient than the 3 step formulation. In particular when considering the accuracy/time effort relation at order 3 the dense formulation provides a better value in almost the same time.

We believe that this can be explained by the way Stokes constraints are added to the program. Indeed, at a given clique, Stokes constraints can only be implemented in the variables that are not shared with the input measure. In the fully (2 step) branched configuration, the last step of the optimization program cannot be accelerated by Stokes constraints at all, while in the 3 step configuration, step 1 includes Stokes constraints in x_3, x_4, x_5, x_6, step 2 includes Stokes constraints in x_2 and step 3 includes Stokes constraints on x_1, which explains the gap between the optimal values of these two configurations.

Moreover, it seems that even the least branched (3 step) configuration still presents a gap between its optimal value and the analytic solution. This might also happen with a non-sparse instance of the Moment-SOS hierarchy (which converges theoretically) and it is likely due to the choice of the monomial basis to represent

polynomials. Indeed, most of the Moment-SOS parsers generate SDP problems with the basis of monomials, while sometimes other bases (e.g. Chebyshev or Legendre polynomials) are more appropriate. However, in this precise case, it might also be linked again with the sparse Stokes constraints implementation. Indeed, in step 2 of the scheme, the unknown measure measures x_2, x_3 and x_4 but Stokes constraints are implemented only in x_2, leaving a possible Gibbs effect in x_3, x_4. Unlike most of our numerical examples, this one still includes an optimization step in which most of the variables are not controlled through Stokes constraints. The gap between the optimal value and the analytic value for the 3 step branched formulation in Figure 5.12b could be explained by a Gibbs effect in the second optimization step.

As a consequence, in the following, one should avoid the branched hierarchies that cannot be accelerated at each step at least partially by Stokes constraints. Such a hierarchy appears when the root of the chosen clique tree shares all its vertices with its children cliques. It can be proved that such a configuration can always be avoided while implementing sparse volume computation, by choosing a leaf as the new root of the tree.

4D polytope

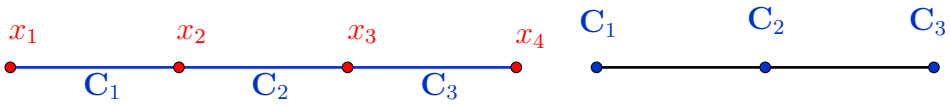
Let $\mathbb{X} := \mathbb{R}^4$, $\mathbb{X}_1 := \langle x_1, x_2 \rangle$, $\mathbb{X}_2 := \langle x_2, x_3 \rangle$, $\mathbb{X}_3 := \langle x_3, x_4 \rangle$, $\mathbf{g}_i(a, b) := (a, b, 1-a-b)$, $i = 1, 2, 3$ and $\mathbf{K}_i := \mathbf{g}_i^{-1}\left(\mathbb{R}^3_+\right) = \{(u, v) \in [0, 1]^2 \ : \ u + v \leq 1\}$. Let us approximate the volume of the 4D polytope

$$\mathbf{K} := \left\{ (x_1, x_2, x_3, x_4) \in \mathbb{R}^4_+ \ : \ \begin{array}{r} x_1 + x_2 \leq 1 \\ x_2 + x_3 \leq 1 \\ x_3 + x_4 \leq 1 \end{array} \right\} = \bigcap_{i=1}^{3} \pi_{\mathbb{X}_i}^{-1}(\mathbf{K}_i) \, .$$

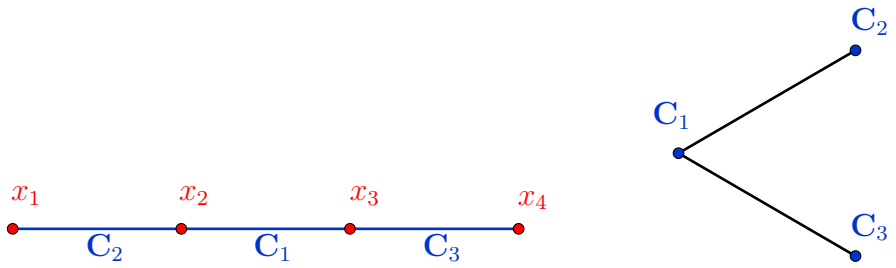
In such a case, there are two possible configurations for the associated clique tree of Proposition 5.12, see Figure 5.13. Accordingly, we can compute $\operatorname{vol} \mathbf{K}$ in two different ways. The first way

$$\begin{aligned} \operatorname{vol} \mathbf{K} = \max_{(\mu_i)_i} & \int 1 \, d\mu_1 \\ \text{s.t. } & \mu_i \in \mathcal{M}(\mathbf{K}_i)_+ \\ & d\mu_1(x_1, x_2) \preceq \mathrm{d}x_1 \, d\mu_2^{\langle x_2 \rangle}(x_2) \\ & d\mu_2(x_2, x_3) \preceq \mathrm{d}x_2 \, d\mu_3^{\langle x_3 \rangle}(x_2) \\ & d\mu_3(x_3, x_4) \preceq \mathrm{d}x_3 \, \mathrm{d}x_4 \end{aligned} \tag{5.16}$$

is the linear formulation given by Corollary 5.8, which is under the form of a non-parallelizable single linear problem. The following additional Stokes constraints can



(a) linear clique tree.



(b) branched clique tree.

Figure 5.13 – Two possible clique trees for the 4D polytope.

be added:

$$\begin{aligned}
\frac{\partial}{\partial x_1}\Big[x_1\,(1-x_1-x_2)\Big]\;d\mu_1(x_1,x_2) &= \frac{\partial}{\partial x_1}\left[x_1\,(1-x_1-x_2)\;d\mu_1(x_1,x_2)\right]\\
\frac{\partial}{\partial x_2}\Big[x_2\,(1-x_2-x_3)\Big]\;d\mu_2(x_2,x_3) &= \frac{\partial}{\partial x_2}\left[x_2\,(1-x_2-x_3)\;d\mu_2(x_2,x_3)\right]\\
\frac{\partial}{\partial x_3}\Big[x_3\,(1-x_3-x_4)\Big]\;d\mu_3(x_3,x_4) &= \frac{\partial}{\partial x_3}\left[x_3\,(1-x_3-x_4)\;d\mu_3(x_3,x_4)\right]\\
\frac{\partial}{\partial x_4}\Big[x_4\,(1-x_3-x_4)\Big]\;d\mu_3(x_3,x_4) &= \frac{\partial}{\partial x_4}\left[x_4\,(1-x_3-x_4)\;d\mu_3(x_3,x_4)\right].
\end{aligned}$$

On the other hand, if one associates the maximal clique $\mathbf{C}_1$ to the subspace $\mathbb{X}_2$ and the maximal clique $\mathbf{C}_2$ to the subspace $\mathbb{X}_1$, one also has

$$\operatorname{vol}\mathbf{K} = \int 1\;d\mu_2^\star \tag{5.17}$$

where

$$\begin{aligned}
\mu_2^\star = \underset{\mu_2}{\operatorname{argmax}} & \int 1 \, d\mu_2 \\
\text{s.t. } & \mu_2 \in \mathcal{M}(\mathbf{K}_2)_+ \\
& d\mu_2(x_2, x_3) \preceq \mu_1^{\star\langle x_2\rangle}(dx_2)\, d\mu_3^{\star\langle x_3\rangle}(x_3) \\
\mu_1^\star = \underset{\mu_1}{\operatorname{argmax}} & \int 1 \, d\mu_1 \\
\text{s.t. } & \mu_1 \in \mathcal{M}(\mathbf{K}_1)_+ \\
& d\mu_1(x_1, x_2) \preceq \mathrm{d}x_1\, \mathrm{d}x_2 \\
\mu_3^\star = \underset{\mu_3}{\operatorname{argmax}} & \int 1 \, d\mu_3 \\
\text{s.t. } & \mu_3 \in \mathcal{M}(\mathbf{K}_3)_+ \\
& d\mu_3(x_3, x_4) \preceq \mathrm{d}x_3\, \mathrm{d}x_4
\end{aligned}$$

which is the branched formulation associated to Theorem 5.13. Here one can see that $\mu_1^\star$ and $\mu_3^\star$ can be computed independently in parallel, and then re-injected in the problem to which $\mu_2^\star$ is the solution. One can add the following Stokes constraints:

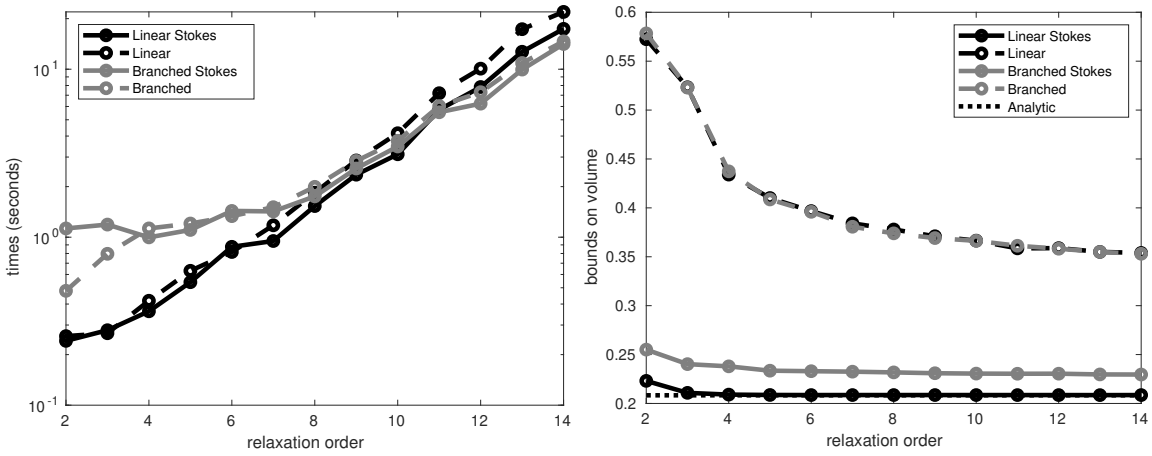
$$\begin{aligned}
\frac{\partial}{\partial x_1}\Big[x_1\,(1-x_1-x_2)\Big]\; d\mu_1(x_1,x_2) &= \frac{\partial}{\partial x_1}\left[x_1\,(1-x_1-x_2)\; d\mu_1(x_1,x_2)\right] \\
\frac{\partial}{\partial x_2}\Big[x_2\,(1-x_1-x_2)\Big]\; d\mu_1(x_1,x_2) &= \frac{\partial}{\partial x_2}\left[x_2\,(1-x_1-x_2)\; d\mu_1(x_1,x_2)\right] \\
\frac{\partial}{\partial x_3}\Big[x_3\,(1-x_3-x_4)\Big]\; d\mu_3(x_3,x_4) &= \frac{\partial}{\partial x_3}\left[x_3\,(1-x_3-x_4)\; d\mu_3(x_3,x_4)\right] \\
\frac{\partial}{\partial x_4}\Big[x_4\,(1-x_3-x_4)\Big]\; d\mu_3(x_3,x_4) &= \frac{\partial}{\partial x_4}\left[x_4\,(1-x_3-x_4)\; d\mu_3(x_3,x_4)\right].
\end{aligned}$$

We can compute analytically

$$\operatorname{vol}\mathbf{K} = \frac{5}{24} \simeq 0.2083.$$

In Figure 5.14 we compare the two sparse formulations with and without Stokes constraints. Surprisingly the linear formulation is faster than the branched one for small relaxation degrees, most probably because at this level of precision the branching costs more in terms of constructing and parsing the LMIs than it saves in computational time. When going deeper in the hierarchy we see the advantage of the branched formulation where more computations are done in parallel. As observed in the previous example however, the branched formulation seems to have problems to converge to the optimal value on an early relaxation. While the values of both formulations without Stokes constraints almost coincide, the values of the linear formulation with Stokes are strictly better than the ones of the accelerated branched formulation. This further supports our conjecture that formulations where Stokes constraints can be added at every step of the optimization program are to be preferred: the fact that both configurations behave equally without Stokes constraints and that the branched configuration keeps a relaxation gap when implementing Stokes constraints suggests that these Stokes constraints behave better

in linear configurations than in branched configurations. For this reason, in the 6D case where all possible configurations are branched, we could not completely eliminate the relaxation gap, while in this case where there is a linear configuration, the relaxation gap vanishes.



(a) Computation time vs relaxation order. (b) Bounds on the volume vs relaxation order.

Figure 5.14 – Performance for the 4D polytope.

5.3.4 The disjoint intersection hypothesis

It may happen that Assumption 5.12 does not hold, in which case all the above results would not apply. For example one could think of the following set:

$$\mathbf{K} := \left\{ \mathbf{x} \in \mathbb{R}^6 : x_1^2 + x_2^2 + x_3^2 \leq 1 \ \wedge \ \forall i \in \{2,3\}, x_i^2 + x_{i+2}^2 + x_{i+3}^2 \leq 1 \right\}$$

whose correlative graph is represented on Figure 5.15. Here the DIP and CIP cannot be simultaneously enforced: the CIP would only be satisfied by a branched clique tree, but since all the cliques share common vertices, in such a branched tree there would automatically be sibling cliques with nonempty intersection. Also, one can notice that in this case (and, as far as we know, only in similar configurations where Assumption 5.12 is violated), the clique $\mathbf{C}_2$ does not correspond to a polynomial appearing in the description of $\mathbf{K}$.

First, we would like to emphasize that the core of this chapter is the linear computation theorem, for which the working assumption always holds. The branched generalizations are only consequences of this linear computation theorem.

Second, the fact that Assumption 5.12 does not hold is not a dead-end for using our scheme. In fact, even the simpler CIP might not hold, in which case one would need to perform a *chordal extension*, which consists of adding virtual links between variables to construct a chordal graph. Basically, a chordal extension would make the graph chordal at the price of slightly weakening the correlative sparsity pattern. In general, the same can be performed to enforce Assumption 5.12: one could add virtual links between variables to enforce the assumption to hold. For example, if

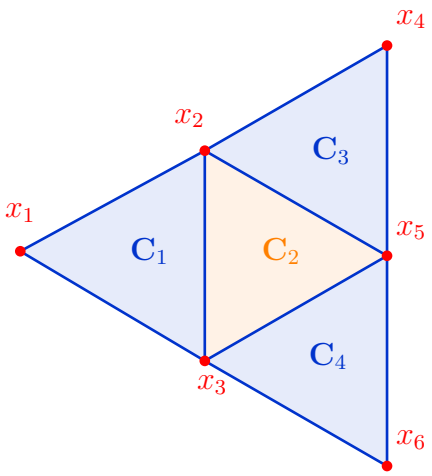


Figure 5.15 – A correlation graph that violates Assumption 5.12.

one artificially links variables x_3 and x_4 in the correlation graph of $\mathbf{K}$, one obtains a new correlation graph, with an associated clique tree satisfying all our working assumptions (see Figure 5.16). This manipulation results in increasing the correlative sparsity CS from 3 to 4, which is a weakening of the correlative sparsity pattern. However, our framework still allows reducing the dimensionality of the problem from 6 to 4.

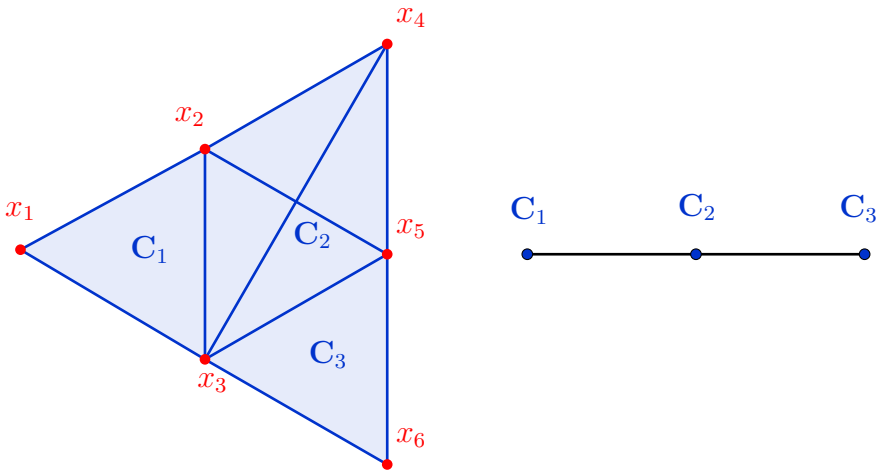


Figure 5.16 – A way to fix our counterexample.

Remark 5.8 (Alternative to graph extension)

There might exist examples for which the DIP would only be obtained by completely destroying the sparsity pattern one wants to exploit. However, this could also happen with the more common CIP. In either case, an option might be to

consider the standard dual volume computation problem (2.8b) *as a way to find minimizing sequences* $(w_d)_d$ *that approximate* $\mathbb{1}_{\mathbf{K}}$*, and to apply it to each one of the* $\mathbf{K}_i := \{\mathbf{x}_i \in \mathbb{X}_i : \mathbf{g}_i(\mathbf{x}_i) \geq \mathbf{0}\}$ *to obtain a sequence* $\left(w_1^{(d)}, \ldots, w_N^{(d)}\right)_d$ *such that* $\left(w_i^{(d)}\right)_d$ *approximates* $\mathbb{1}_{\mathbf{K}_i}$*. Then, one would still have to prove that the convergence of the hierarchy is stable by product (which is nontrivial) to conclude that*

$$\int \prod_{i=1}^{N} w_i^{(d)} \circ \pi_{\mathbb{X}_i}(\mathbf{x}) \, \mathrm{d}\mathbf{x}$$

converges to vol $\mathbf{K}$*. The major drawback of this solution is that in order for* $(w_d)_d$ *to approximate* $\mathbb{1}_{\mathbf{K}}$*, we cannot implement Stokes constraints, as they would modify problem* (2.8b) *(see Chapter 4) in a way that makes us lose the convergence to an indicator (this is precisely the point of Stokes constraints: they allow their user to obtain the volume without trying to approximate discontinuous functions with polynomials). However, we know that the volume approximation hierarchy has a bad convergence rate without Stokes constraints. In general, we should not expect any method that approximates indicators with polynomials to yield satisfactory results in terms of volume computation. Such a method should be considered only as a last resort if any trial to apply the above scheme fails. Finally, this option represents a framework that would be completely independent from the above and thus it remains out of the scope of this chapter.*

We now move on to proposing a detailed method to enforce Assumption 5.12. To do so, we first introduce a more detailed glossary for graph theory. Let $\mathbb{G} = (\mathbf{V}, \mathbf{E})$ be a graph with vertices set $\mathbf{V}$ and edges set $\mathbf{E} \subset \mathbf{V}^2$. The following definitions can be found in e.g. [13]:

- The *degree* $\deg \mathrm{v}$ of $\mathrm{v} \in \mathbf{V}$ is the cardial of the set $\{\mathrm{v}' \in \mathbf{V} : (\mathrm{v}, \mathrm{v}') \in \mathbf{E}\}$ *i.e.* the number of vertices connected to v.
- A *clique* of $\mathbb{G}$ is a subset of vertices $\mathbf{C} \subset \mathbf{V}$ such that $\mathrm{v}, \mathrm{v}' \in \mathbf{C}$ implies $(\mathrm{v}, \mathrm{v}') \in \mathbf{E}$.
- A graph $\mathbb{G}$ is *chordal* if every cycle of length greater than 3 has a chord, i.e. an edge connecting two nonconsecutive vertices on the cycle.
- A *tree* $\mathbb{T} = (\mathcal{K}, \mathcal{E})$ is a graph without cycle.
- The *treewidth* of a chordal graph is the size of its biggest clique minus 1. Thus the treewidth of a tree is 1 and the treewidth of a complete graph ($\mathbf{E} = \mathbf{V}^2$) of size n is $n - 1$.
- A *rooted tree* is a tree in which one vertex has been designated the *root*.
- In a rooted tree, the *parent* of a vertex v is the vertex v connected to it on the path to the root; v is then called a *child* of v; two vertices that have the same parent are called *siblings*; a *descendant* of a vertex v is any vertex which is either the child of v or is (recursively) the descendant of any of the children of v; v is then called an *ancestor* of itself and any of its descendants.

- The vertices of a rooted tree can be partitioned between the root, the *leaves* (the vertices that have parents but no children) and the *branches* (that have children and parents).

- Let $\mathcal{K}$ be the set of maximal cliques of $\mathbb{G}$. A *clique tree* $\mathbb{T} = (\mathcal{K}, \mathcal{E})$ of $\mathbb{G}$ is a tree whose vertices are the maximal cliques of $\mathbb{G}$.

- A clique tree satisfies the clique intersection property (CIP) if for every pair of distinct cliques $\mathbf{C}, \mathbf{C}' \in \mathcal{K}$, the set $\mathbf{C} \cap \mathbf{C}'$ is contained in every clique on the path connecting $\mathbf{C}$ and $\mathbf{C}'$ in the tree. We denote by $\mathcal{T}^{ct}$ the set of clique trees of $\mathbb{G}$ that satisfy the CIP.

- The ordering $\mathcal{K} = \{\mathbf{C}_1, \dots, \mathbf{C}_N\}$ satisfies the *Running Intersection Property* (RIP) if $\forall i \in \{2, \dots, N\}$, $\exists\, k_i \in \{1, \dots, i-1\}$ such that $\mathbf{C}_i \cap \bigcup_{j=1}^{i-1} \mathbf{C}_j \subset \mathbf{C}_{k_i}$.

- Let $\mathbf{A} \subset \mathbf{V}$. Then, the *subgraph of* $\mathbb{G}$ *generated by* $\mathbf{A}$ is given by

$$\langle \mathbf{A} \rangle_{\mathbb{G}} := (\mathbf{A}, \mathbf{E} \cap \mathbf{A}^2).$$

Theorem 5.15: Chordality & RIP

A connected graph $\mathbb{G}$ is chordal if and only if $\mathcal{T}^{ct} \neq \varnothing$ if and only if $\mathcal{K}$ admits an ordering that satisfies the RIP.

Definition 5.16: Disjoint Intersection Property

Let $\mathbb{G} = (\mathbf{V}, \mathbf{E})$ be chordal and connected. Let $\mathbb{T} = (\mathcal{K}, \mathcal{E}) \in \mathcal{T}^{ct}$ be a clique tree rooted in $\mathbf{C}_1 \in \mathcal{K}$. $\mathbb{T}$ satisfies the *Disjoint Intersection Property* (DIP) if $\forall \mathbf{C}, \mathbf{C}', \mathbf{C}'' \in \mathcal{K}$, if $(\mathbf{C}, \mathbf{C}') \in \mathcal{E}$ and $(\mathbf{C}, \mathbf{C}'') \in \mathcal{E}$ then $\mathbf{C}' = \mathbf{C}''$ or $\mathbf{C}' \cap \mathbf{C}'' = \varnothing$. In words, each clique has an empty intersection with all its siblings.

We are now going to give a systematic way to enforce Assumption 5.12 and generate the associated clique trees. Let $(\mathbf{g}_1, \dots, \mathbf{g}_N)$ be a correlatively sparse family of polynomial vectors with a connected chordal correlation graph $\mathbb{G} = (\mathbf{V}, \mathbf{E})$. Let $\mathcal{K}$ be the set of maximal cliques of $\mathbb{G}$. We construct the *clique graph* $\mathbb{G}_{\mathcal{K}} = (\mathcal{K}, \mathcal{F})$ such that $(\mathbf{C}, \mathbf{C}') \in \mathcal{F}$ *iff* $\mathbf{C} \cap \mathbf{C}' \neq \varnothing$. One can in turn define cliques (called *meta-cliques*) for this new graph, and its correlative sparsity $\mathrm{CS}_{\mathcal{K}}$ is the size of its biggest maximal meta-clique minus 1. One can note that any clique tree is a subtree of $\mathbb{G}_{\mathcal{K}}$ including all its vertices.

Remark 5.9 *If $\mathbb{G}_{\mathcal{K}}$ itself is a tree (as in section 5.3.3), then it trivially satisfies the DIP and CIP, and Assumption 5.12 automatically holds.*

Lemma 5.17: Genealogy

Let $\mathbb{T} = (\mathcal{K}, \mathcal{E})$ be a clique tree satisfying Assumption 5.12.

1) Let $\mathbf{C}, \mathbf{C}' \in \mathcal{K}$ such that $\mathbf{C} \neq \mathbf{C}'$ and $\mathbf{C} \cap \mathbf{C}' \neq \varnothing$. Then, up to permuting them, $\mathbf{C}$ is a descendant of $\mathbf{C}'$.

2) Let $\mathcal{Q}$ be a meta-clique. Then, $\langle \mathcal{Q} \rangle_{\mathbb{T}}$ is an oriented path of $\mathbb{T}$: the elements of $\mathcal{Q}$ are ancestor to one another and each $\mathbf{C} \in \mathcal{Q}$ has its parent in $\mathcal{Q}$ except one of them.

Proof :

1) Let $\mathbf{C}''$ be the last common ancestor of $\mathbf{C}$ and $\mathbf{C}'$, meaning that $\mathbf{C}''$ is an ancestor of both $\mathbf{C}$ and $\mathbf{C}'$ but any child of $\mathbf{C}''$ is the ancestor of at most one of them. Such ancestor exists since the root $\mathbf{C}_1$ is a common ancestor to $\mathbf{C}$ and $\mathbf{C}'$. Then, $\mathbf{C}''$ is on the path between $\mathbf{C}$ and $\mathbf{C}'$. Since $\mathbf{C} \neq \mathbf{C}'$, up to permuting them, we can suppose that $\mathbf{C} \neq \mathbf{C}''$.

   By contradiction, we suppose that $\mathbf{C}' \neq \mathbf{C}''$. Then, let $\hat{\mathbf{C}}$ be the child of $\mathbf{C}''$ that is also the ancestor of $\mathbf{C}$, and $\tilde{\mathbf{C}}$ the child of $\mathbf{C}''$ that is also the ancestor of $\mathbf{C}'$. Both exist since $\mathbf{C}''$ is an ancestor of $\mathbf{C}$ and $\mathbf{C}'$ and $\mathbf{C} \neq \mathbf{C}' \neq \mathbf{C}''$. $\mathbf{C}''$ being the latest common ancestor of $\mathbf{C}$ and $\mathbf{C}'$, we deduce that $\hat{\mathbf{C}} \neq \tilde{\mathbf{C}}$ so that $\hat{\mathbf{C}}$ and $\tilde{\mathbf{C}}$ are siblings. Then, the DIP ensures that $\hat{\mathbf{C}} \cap \tilde{\mathbf{C}} = \varnothing$. However, $\hat{\mathbf{C}}$ and $\tilde{\mathbf{C}}$ are on the path between $\mathbf{C}$ and $\mathbf{C}'$, so according to the CIP they both contain $\mathbf{C} \cap \mathbf{C}'$ which is nonempty. This is a contradiction.

2) According to point 1), all elements of $\mathcal{Q}$ are descendants of one another, so that they are all on the same path in $\mathbb{T}$. We only have to show that any $\mathbf{C}$ between two elements $\mathbf{C}', \mathbf{C}''$ of $\mathcal{Q}$ on this path is also an element of $\mathcal{Q}$. Indeed, let $\mathbf{C}''' \in \mathcal{Q}$. Then, up to a permutation on $\{\mathbf{C}', \mathbf{C}'', \mathbf{C}'''\}$ we can suppose that the unoriented path includes in this order: $(\mathbf{C}', \mathbf{C}, \mathbf{C}'', \mathbf{C}''')$. Then, $\mathbf{C}$ is on the path between $\mathbf{C}'$ and $\mathbf{C}'''$, so the CIP implies that $\mathbf{C} \supset \mathbf{C}' \cap \mathbf{C}'''$ is nonempty (since $\mathbf{C}'$ and $\mathbf{C}'''$ belong to the same clique $\mathcal{Q}$), and then $\mathbf{C}$ has a nonempty intersection with $\mathbf{C}'''$. This shows that $\mathbf{C}$ has a nonempty intersection with any element of $\mathcal{Q}$, which by maximality of $\mathcal{Q}$ is the definition of $\mathbf{C} \in \mathcal{Q}$.

♢

Corollary 5.18: Path decomposition

If $\mathbb{G}_{\mathcal{K}}$ is a complete graph (all pairs of maximal cliques have nonempty intersection as in section 5.3.3), then the only candidates for our clique tree are linear clique trees (*i.e.* path decompositions). In such case, Assumption 5.12 is equivalent to the existence of a reordering of $(\mathbf{g}_1, \ldots, \mathbf{g}_N)$ such that Assumption 5.7 holds.

We now give an alogrithm to generate a clique tree $\mathbb{T} = (\mathcal{K}, \mathcal{E})$ that satisfies the DIP and the CIP. In case Assumption 5.12 does not hold, this algorithm automatically adds edges to $\mathbf{E}$ until it finds an appropriate clique tree (see Algorithm 1).

Remark 5.10 (Explanation of Algorithm 1)

- *Minimizing* $\deg \mathbf{C}_1$ *is a way to ensure Stokes constraints will be fully implementable, in contrast to the 2 step implementation in section 5.3.3.*
- *Index i denotes a clique that has already been added to the tree; the algorithm adds to the tree every clique that shares vertices with $\mathbf{C}_i$, and then increments i.*
- *Index j denotes the meta-clique in which we are working; according to Lemma 5.17, in such meta-clique the cliques should be added in line.*
- *Index k denotes the number of elements that have already been added to the tree; when k is equal to the number of cliques, our clique tree is complete.*
- *Index l denotes the lates clique of the meta-clique $\mathcal{Q}_j$ that has been added to the tree; according to Lemma 5.17, $\mathbf{C}_l$ should then be the parent of $\mathbf{C}_{k+1}$.*
- *At line 1 we maximize $|\mathcal{Q}_j|$ to favor linear configurations as they are the most compatible with Stokes constraints.*
- *The if loop at line 1 checks whether it is possible to add the remaining cliques of $\mathcal{Q}_j$ to our tree without destroying the CIP.*
- *The if loop at line 1 checks whether the clique we want to add destroys the DIP or not.*
- *At line 1 we maximize $|\mathbf{C}_{k+1} \cap \mathbf{C}_l|$ so that it is less likely to pose problems with CIP and DIP in the future iterations.*
- *The if loops at lines 1 and 1 are meant to minimize the correlative sparsity of the new graph $\mathbb{G} = (\mathbf{V}, \mathbf{E})$, since it is the limiting factor for the tractability of our algorithm.*

Theorem 5.19: Convergence of Algorithm 1

Any clique tree returned by Algorithm 1 satisfies the DIP and the CIP.

Proof : We are going to show by induction that at any step k, the graph $\mathbb{T}_k := (\mathcal{P}_k, \mathcal{E}_k)$ is a tree that satisfies the CIP and the DIP. First, it is trivial that $\mathbb{T}_1 = (\{\mathbf{C}_1\}, \varnothing)$ is a tree and satisfies the CIP and the DIP. Next, we suppose that we have constructed a tree $\mathbb{T}_k$ satisfying CIP and DIP through iterations of our algorithm, and that the next iteration leads us to define a $\mathbb{T}_{k+1} := (\mathcal{P}_k \cup \{\mathbf{C}_{k+1}\}, \mathcal{E}_k \cup \{(\mathbf{C}_l, \mathbf{C}_{k+1}\})$. Since the only edge we added connected $\mathbf{C}_l$ to a new vertex that was not in $\mathbb{T}_k$, it did not introduce any cycle, thus $\mathbb{T}_{k+1}$ is still a tree. We are now going to check the CIP and DIP.

Algorithm 1: How to build an appropriate clique tree

Data: $\mathbb{G} = (\mathbf{V}, \mathbf{E})$ and its clique graph $\mathbb{G}_{\mathcal{K}} = (\mathcal{K}, \mathcal{F})$.
Result: $\mathbb{T} = (\mathcal{K}, \mathcal{E})$ satisfying CIP & DIP if Assumption 5.12 holds,
$\mathbb{G} = (\mathbf{V}, \mathbf{E})$ with additional edges else.
Initialization: Choose $\mathbf{C}_1 \in \mathcal{K}$ with minimal degree in $\mathbb{G}_{\mathcal{K}}$;
Initialize $i = j = k = 1$, $\mathcal{P}_1 := \{\mathbf{C}_1\}$, $\mathcal{E}_1 := \varnothing$;
while $k < |\mathcal{K}|$ **do**
  while $\mathbb{M}_{ik} := \{\mathcal{Q}$ *maximal meta-clique* $: \mathcal{Q} \cap \mathcal{P}_k^c \neq \varnothing, \mathbf{C}_i \in \mathcal{Q}\} \neq \varnothing$ **do**
    Choose $\mathcal{Q}_j \in \underset{\mathcal{Q} \in \mathbb{M}_{ik}}{\operatorname{argmax}} |\mathcal{Q}|$;
    while $\mathcal{Q}_j \cap \mathcal{P}_k^c \neq \varnothing$ **do**
      $l := \max\{r \leq k : \mathbf{C}_r \in \mathcal{Q}_j\}$;
      if
       $\exists (\mathbf{C}, \mathbf{C}') \in \operatorname{argmax}\{|\hat{\mathbf{C}} \cap \mathbf{C}_l| : (\tilde{\mathbf{C}}, \hat{\mathbf{C}}) \in \mathcal{P}_k \times \mathcal{Q}_j \cap \mathcal{P}_k^c, \tilde{\mathbf{C}} \cap \hat{\mathbf{C}} \not\subseteq \mathbf{C}_l\}$
       then
        if $|\mathbf{C}_l \cup \mathbf{C}| > |\mathbf{C}_l \cup \mathbf{C}'|$ **then**
          foreach $\mathrm{v} \in \mathbf{C}', \mathrm{v}' \in \mathbf{C}_l$ **do** $\mathbf{E} \leftarrow \mathbf{E} \cup \{(\mathrm{v}, \mathrm{v}'), (\mathrm{v}', \mathrm{v})\}$;
          return $\mathbb{G} = (\mathbf{V}, \mathbf{E})$;
        else
          foreach $\mathrm{v} \in \mathbf{C}, \mathrm{v}' \in \mathbf{C}_l$ **do** $\mathbf{E} \leftarrow \mathbf{E} \cup \{(\mathrm{v}, \mathrm{v}'), (\mathrm{v}', \mathrm{v})\}$;
          return $\mathbb{G} = (\mathbf{V}, \mathbf{E})$;
        end
      else
        Choose $\mathbf{C}_{k+1} \in \underset{\mathbf{C} \in \mathcal{Q}_j \cap \mathcal{P}_k^c}{\operatorname{argmax}} |\mathbf{C} \cap \mathbf{C}_l|$
      end
      if $\exists \mathbf{C} \in \mathcal{P}_k$ *s.t.* $(\mathbf{C}_l, \mathbf{C}) \in \mathcal{E}_k$ *&* $\mathbf{C} \cap \mathbf{C}_{k+1} \neq \varnothing$ **then**
        if $|\mathbf{C} \cup \mathbf{C}_{k+1} < |\mathbf{C}_l \cup \mathbf{C}_{k+1}| \wedge |\mathbf{C}_l \cup \mathbf{C}|$ **then**
          foreach $\mathrm{v} \in \mathbf{C}, \mathrm{v}' \in \mathbf{C}_{k+1}$ **do** $\mathbf{E} \leftarrow \mathbf{E} \cup \{(\mathrm{v}, \mathrm{v}'), (\mathrm{v}', \mathrm{v})\}$;
          return $\mathbb{G} = (\mathbf{V}, \mathbf{E})$;
        else if $|\mathbf{C}_l \cup \mathbf{C}| > |\mathbf{C}_l \cup \mathbf{C}_{k+1}|$ **then**
          foreach $\mathrm{v} \in \mathbf{C}_{k+1}, \mathrm{v}' \in \mathbf{C}_l$ **do** $\mathbf{E} \leftarrow \mathbf{E} \cup \{(\mathrm{v}, \mathrm{v}'), (\mathrm{v}', \mathrm{v})\}$;
          return $\mathbb{G} = (\mathbf{V}, \mathbf{E})$;
        else
          foreach $\mathrm{v} \in \mathbf{C}, \mathrm{v}' \in \mathbf{C}_l$ **do** $\mathbf{E} \leftarrow \mathbf{E} \cup \{(\mathrm{v}, \mathrm{v}'), (\mathrm{v}', \mathrm{v})\}$;
          return $\mathbb{G} = (\mathbf{V}, \mathbf{E})$;
        end
      else
        $\mathcal{P}_{k+1} := \mathcal{P}_k \cup \{\mathbf{C}_{k+1}\}$;
        $\mathcal{E}_{k+1} := \mathcal{E}_k \cup \{(\mathbf{C}_l, \mathbf{C}_{k+1})\}$;
        $k \leftarrow k + 1$;
      end
    end
    $j \leftarrow j + 1$;
  end
  $i \leftarrow i + 1$;
end
return $\mathbb{T} := (\mathcal{P}_{|\mathcal{K}|}, \mathcal{E}_{|\mathcal{K}|})$;

- Let $\mathbf{C}, \mathbf{C}'' \in \mathcal{P}_{k+1}; \mathbf{C} \cap \mathbf{C}'' \neq \varnothing$. Let $\mathbf{C}' \in \mathcal{P}_k$ be on the path between $\mathbf{C}$ and $\mathbf{C}''$ in $\mathbb{T}_{k+1}$ ($\mathbf{C}' \neq \mathbf{C}_{k+1}$ because $\mathbf{C}_{k+1}$ is on no path in $\mathbb{T}_{k+1}$).
  - If $\mathbf{C}, \mathbf{C}'' \in \mathcal{P}_k$ then by our induction hypothesis $\mathbf{C} \cap \mathbf{C}'' \subset \mathbf{C}'$.
  - Else, without loss of generality we have $\mathbf{C}'' = \mathbf{C}_{k+1}$ and $\mathbf{C} \in \mathcal{P}_k$.
    * Since we successfully passed through the if loop of line 1, we have $\mathbf{C} \cap \mathbf{C}_{k+1} \subset \mathbf{C}_l$.
    * By our induction assumption ($\mathbb{T}_k$ satisfies the CIP), we have $\mathbf{C} \cap \mathbf{C}_l \subset \mathbf{C}'$ (because either $\mathbf{C}' = \mathbf{C}_l$ or $\mathbf{C}'$ is on the path between $\mathbf{C}$ and $\mathbf{C}_l$, the parent of $\mathbf{C}_{k+1}$).

    This yields $\mathbf{C}' \supset \mathbf{C} \cap \mathbf{C}_l \supset \mathbf{C} \cap (\mathbf{C} \cap \mathbf{C}_{k+1}) = \mathbf{C} \cap \mathbf{C}_{k+1} = \mathbf{C} \cap \mathbf{C}''$.

  Then, $\mathbb{T}_{k+1}$ satisfies the CIP.

- Let $\mathbf{C} \in \mathcal{P}_k, \mathbf{C}', \mathbf{C}'' \in \mathcal{P}_{k+1}$ such that $(\mathbf{C}, \mathbf{C}'), (\mathbf{C}, \mathbf{C}'') \in \mathcal{E}_{k+1}$.
  - If $\mathbf{C}', \mathbf{C}'' \in \mathcal{P}_k$ then by our induction hypothesis $\mathbf{C}' \cap \mathbf{C}'' = \varnothing$.
  - Else, without loss of generality we have $\mathbf{C}'' = \mathbf{C}_{k+1}, \mathbf{C} = \mathbf{C}_l$, and since we successfully passed through the if loop of line 1, we have $\mathbf{C}' \cap \mathbf{C}'' = \mathbf{C}' \cap \mathbf{C}_{k+1} = \varnothing$.

  Then, $\mathbb{T}_{k+1}$ satisfies the DIP.

♢

Finally, we conjecture that if Assumption 5.12 holds, then Algorithm 1 will directly return a clique tree satisfying the CIP and the DIP without adding any edge to $\mathbb{G}$.

Conclusion

Our results

In this chapter we addressed the problem of approximating the volume of sparse semi-algebraic sets with the Moment-SOS hierarchy of SDP relaxations. As illustrated by our examples, our sparse formulation allows one to dramatically decrease the computational time for each relaxation, and to tackle high dimensional volume computation problems that are not tractable with the usual SDP methods. By splitting the problems into low dimensional subproblems, one drastically reduces the dimension of each relaxation, without loss of precision. This reduction of complexity is due to the correspondance between the structure of our algorithm and the correlative sparsity pattern in the description of the semi-algebraic set.

We also showed that additional Stokes constraints have a huge effect on convergence and precision for volume computation, and that they can successfully be adapted to our sparse formulations. This yields a much better rate of convergence for the corresponding hierarchy. However, implementing these Stokes constraints leads

to subtle constraints that have to be enforced if one wants to efficiently compute the volume:

- First, one should always prefer the linear formulation of Theorem 5.8 whenever possible, since this ensures that Stokes constraints can always be efficiently implemented.
- Then, in the more general case of Theorem 5.13, one should always avoid formulations in which the root of the computation tree has no Stokes constraint; fortunately, such configurations can always be avoided by chosing a leaf as the root of the clique tree.

Furthermore, in the branched case, one should be aware of the fact that each step of the algorithm introduces an approximation error, and the errors accumulate until the root is reached. Consequently, a formulation in which the clique tree has too many generations will lead to a larger global error than a formulation with less generations. For this reason, one should minimize the number of generations in the clique tree, which is equivalent to parallelizing as much as possible. In addition to that, when the problem has many dimensions and branches, parallelization can obviously drastically increase the speed of the computations.

Applications and future work

To the best of our knowledge, this sparse method for solving volume problems is new and full of promises for future applications. For instance, the problem of computing the mass of any compactly supported measure absolutely continuous (with respect to the Lebesgue measure) can be adressed using this sparsity method. Also, measures that are not compactly supported but have some decay properties (e.g. Gaussian measures) can also be handled by our method, which may prove useful in computations for probability and statistics. Also, specific constraints could probably be used in addition to Stokes constraints when the semi-algebraic set presents a specific structure (e.g. a polytope, a convex body).

Furthermore, the framework of exploiting correlative sparsity can be applied to any method that relies on computations on measures, whether these measures are represented by their moments (as it is done in this chapter), or by samples (as in the stochastic volume computation methods). In particular, we believe that this formalism could easily be extended to Monte-Carlo-based volume computations.

Finally, we also believe that this method can be adapted to the computation of regions of attraction, through the formalism developed in [42], for high dimensional differential-algebraic systems that present a network structure, such as power grids, distribution networks in general and possibly other problems. The main difficulty resides in taking sparsity into account when formulating the Liouville equation, and keeping uniqueness of the solution as in the non-controlled non-sparse framework.

6

Theoretical contributions to stability analysis

This final chapter gathers original contributions to the particular problem of differential systems stability analysis, which is the goal of this thesis. While Chapter 3 merely was an extension of existing frameworks to power systems TSA, this chapter draws on most previously presented results to build two new frameworks for stability analysis of differential systems.

Section 6.1, which is based on the work [100], uses the results of Chapter 2 to extend the existing frameworks presented in Chapter 3, to the inner approximation of the Maximal Positively Invariant (MPI) set of a differential system, whose relevance for power systems security and slightly reduced computational complexity will be discussed hereafter.

Section 6.2, published as [123], then extends the results from Chapters 3 and 5 to approximating regions of attraction of large scale, sparsely coupled, differential systems, in a first attempt to fill the gap between stability regions approximation theory and its application to large scale power grids.

Contents

6.1 Inner approximation of maximal positively invariant sets

This section is an effort along a research line initiated in [42] for developing convex optimization techniques to approximate sets relevant to non-linear control systems subject to non-linear constraints, with rigorous proofs of convergence in volume. The approximations are obtained by numerically solving a hierarchy of semidefinite programming or linear matrix inequality (LMI) relaxations, as proposed originally by Lasserre in the context of polynomial optimization [70]. Convergence proofs are achieved by exploiting duality between non-negative continuous functions and Borel measures, approximated respectively with sums of squares (SOS) of polynomials and moments, justifying the terminology moment-SOS or Lasserre hierarchy. In the context of control systems, the primal moment formulation builds upon the notion of occupation measures [72] and the dual SOS formulation can be classified under Hamilton-Jacobi techniques [9].

Previous works along this line include inner approximations of the region of attraction [58], outer approximations of the MPI set [57], as well as outer approximations of the reachability set [83]. Besides their use for power systems stability analysis (see Section 3.1), these techniques were applied e.g. in robotics [86] and biological systems [109]. In [42, 58] the regions of attraction are defined for a finite time horizon, which is a technical convenient framework since the occupation measures have then finite mass. To cope with an infinite time horizon and MPI sets, a discount factor was added in [57] so that the mass of the occupation measure decreases fast enough when time increases. In [83], the mass was controlled by enforcing a growth condition on the volume of complement sets. This condition, difficult to check a priori, can be validated a posteriori using duality theory.

It must be emphasized here that, in general, the infinite time horizon setup is more convenient for the classical Lyapunov framework and asymptotic stability, see e.g. [20] and references therein, whereas the finite time horizon setup is more convenient for approaches based on occupation measures. In the current section, we make efforts to adapt the occupation measure framework to an infinite time horizon setup, at the price of technical difficulties similar to the ones already encountered in [83]. Contrary to the outer approximations derived in [57], we have not been able to use discounted occupation measures for constructing inner approximations. Instead, the technical device on which we relied is a growth condition on the average exit time of trajectories.

The main contributions of this section are:

1. A moment-SOS hierarchy for constructing *inner* approximations of the MPI set for a polynomial dynamical system with semialgebraic constraints;

2. A detailed, self-contained, rigorous proof of convergence of the hierarchy, under an assumption on the average exit time of trajectories.

Section 6.1.1 presents the problem statement. In section 6.1.2 we introduce a modified version of occupation measures which allow for a new GMP formulation. Section 6.1.3 describes the MPI set inner approximation method with proof of convergence under appropriate assumptions. Numerical results are analyzed in Section 6.1.4.

6.1.1 MPI set

Consider the autonomous system

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}), \quad \mathbf{x} \in \mathbf{\Omega} \subset \mathbb{R}^n, \quad t \in [0, +\infty) \tag{6.1}$$

with a given polynomial vector field $\mathbf{f} \in \mathbb{R}[\mathbf{x}]^n$ of degree d_0. The state trajectory $\mathbf{x}(.)$ is constrained to the interior $\mathbf{\Omega}$ of a nonempty compact basic semi-algebraic set

$$\mathbf{X} := \{\mathbf{x} \in \mathbb{R}^n : \forall i \in \mathbb{N}^\star_{m_\mathbf{X}},\ g_i(\mathbf{x}) \geq 0\}$$

where the g_i are given polynomials of degree d_i. We define $\partial\mathbf{X} := \mathbf{X} \setminus \mathbf{\Omega}$.

The vector field $\mathbf{f}$ is polynomial and therefore Lipschitz on the compact set $\mathbf{X}$. As a result, for any $\mathbf{x}_0 \in \mathbf{\Omega}$, there exists a unique maximal solution $\mathbf{x}(\cdot|\mathbf{x}_0)$ to ordinary differential equation (6.1) with initial condition $\mathbf{x}(0|\mathbf{x}_0) = \mathbf{x}_0$. The time interval on which this solution is defined contains the time interval on which $\mathbf{x}(\cdot|\mathbf{x}_0) \in \mathbf{\Omega}$.

For any $T \in [0, +\infty]$, we define the following set.

Definition 6.1: Time T secure initializing set

$$\mathbf{X}_T := \{\mathbf{x}_0 \in \mathbf{\Omega} : \forall t \in [0, T],\ \mathbf{x}(t|\mathbf{x}_0) \in \mathbf{\Omega}\}$$

is the set of all initial states generating trajectories staying in $\mathbf{\Omega}$ between time $t = 0$ and $t = T$.

Remark 6.1 (Meaning of the name "*secure* initializing set")

If one considers the set $\mathbf{X}$ as a set of polynomial security constraints that should be enforced (e.g. in the field of power systems: bound specifications for current and voltage magnitudes as well as active and reactive power or phase shifting), then $\mathbf{X}_T$ is the set of initial conditions generating trajectories that satisfy these security constraints between time $t = 0$ and $t = T$, hence the name of time T "secure" initializing set.

Remark 6.2 (Link with constrained regions of attraction)

Using the notation of Definition 1.5, one has

$$\mathbf{X}_T \overset{\text{def}}{=} \mathbf{A}^{\mathbf{\Omega}}_T(\mathbf{\Omega}).$$

To avoid heavy notations, we use the simpler $\mathbf{X}_T$ naming.

On another hand, the maximal positively invariant set included in $\mathbf{\Omega}$ is defined as follows

Definition 6.2: Maximal Positively Invariant (MPI) set

Let $\mathbf{P} \subset \mathbb{R}^n$. $\mathbf{P}$ is said to be *positively invariant (PI)* for the system (6.1) *iff*

$$\forall \mathbf{x}_0 \in \mathbf{P}, \forall t \geq 0, \mathbf{x}(t|\mathbf{x}_0) \in \mathbf{P}.$$

Then, the *maximal positively invariant set* included in $\mathbf{X}$ is defined as its name suggests:

$$\text{MPI}_{\mathbf{X}} := \bigcup_{\substack{\mathbf{P} \subset \mathbf{X} \\ \mathbf{P} \text{ PI}}} \mathbf{P}.$$

Remark 6.3 (Positive invariance and boundary)

Defining as in Chapter 4, for any $\mathbf{x} \in \partial\mathbf{P}$, the outward pointing normal vector $\mathbf{n}(\mathbf{x})$, one immediately has that $\mathbf{P}$ is PI for (6.1) iff

$$\forall \mathbf{x} \in \partial\mathbf{P}, \mathbf{f}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) \leq 0.$$

It is then straightforward to draw a link between the MPI set and infinite time secure initializing sets, under the form of the following lemma.

Lemma 6.3: MPI set characterization

The infinite time secure initializing set is the MPI set included in $\mathbf{\Omega}$:

$$\mathbf{X}_\infty = \text{MPI}_{\mathbf{X}}.$$

Proof : Let $\mathbf{x}_0 \in \mathbf{X}_\infty$, $t, t' \in [0, \infty)$. Then

$$\mathbf{x}(t'|\mathbf{x}(t|\mathbf{x}_0)) = \mathbf{x}(t + t'|\mathbf{x}_0) \in \mathbf{\Omega},$$

so that $\mathbf{x}(t|\mathbf{x}_0) \in \mathbf{X}_\infty$: $\mathbf{X}_\infty$ is PI, and thus $\mathbf{X}_\infty \subset \text{MPI}_{\mathbf{X}}$.

Let $\mathbf{x}_0 \in \text{MPI}_{\mathbf{X}}$, $t \geq 0$. By definition, $\mathbf{x}(t|\mathbf{x}_0) \in \text{MPI}_{\mathbf{X}} \subset \mathbf{X}$, and thus $\mathbf{x}_0 \in \mathbf{X}_\infty$, which yields $\text{MPI}_{\mathbf{X}} \subset \mathbf{X}_\infty$. ♢

Remark 6.4 (Relevance of the MPI set)

In terms of stability analysis, computing the infinite time secure initializing set is very interesting. Indeed, this set is defined as the set of initial (or post-fault) conditions for which the system is bound to "eternally" satisfy all specified security constraints that define the set $\mathbf{X}$. In other words, a trajectory initialized in $\mathbf{X}_\infty$ is sure not to lose synchronism, and the values of the current, voltage and power are prevented from reaching dangerous heights.

We make the following assumption implying that $\mathbf{X}_\infty$ has non-empty interior:

Assumption 6.4: Lyapunov stability

$\mathbf{\Omega}$ contains a Lyapunov-stable equilibrium point $\overline{\mathbf{x}}$ for $\mathbf{f}$.

Indeed in such case if $\varepsilon > 0$ is such that $\mathbf{B}_\varepsilon(\overline{\mathbf{x}}) \subset \mathbf{\Omega}$, then Definition 3.7 provides us with a $\eta > 0$ such that $\mathbf{B}_\eta(\overline{\mathbf{x}}) \subset \mathbf{X}_\infty$, where $\mathbf{B}_R(\mathbf{x}) = \{\mathbf{y} \in \mathbb{R}^n : |\mathbf{x} - \mathbf{y}| \leq R\}$ is the radius R euclidean ball around $\mathbf{x}$.

The complementary set $\mathbf{X}_T^{\mathrm{c}} := \mathbf{\Omega} \setminus \mathbf{X}_T$ is the set of initial conditions generating trajectories reaching the target set $\partial\mathbf{X}$ at any time before T: this is the region of attraction of $\partial\mathbf{X}$ with free final time lower than T

$$\mathbf{X}_T^{\mathrm{c}} = \bigcup_{\tau \in [0,T)} \mathbf{A}_\tau^{\mathbf{\Omega}}(\partial\mathbf{X}). \tag{6.2a}$$

The complementary set $\mathbf{X}_\infty^{\mathrm{c}}$ is the region of attraction of $\partial\mathbf{X}$ with free and unbounded final time

$$\mathbf{X}_\infty^{\mathrm{c}} = \bigcup_{\tau \geq 0} \mathbf{A}_\tau^{\mathbf{\Omega}}(\partial\mathbf{X}). \tag{6.2b}$$

In this section we want to approximate the MPI set $\mathbf{X}_\infty$ from inside as closely as possible.

The next section presents a new instance for the GMP, which is as usual accompanied with a hierarchy of convex linear matrix inequality (LMI) relaxations yielding a converging sequence (in the sense of the Lebesgue measure) of inner approximations of the MPI set.

However, due to the infinite time horizon, such a strong result is available only under some assumptions, based on the notion of exit time.

Definition 6.5: Exit time

For a given $\mathbf{x}_0 \in \mathbf{X}$, we define the *exit time* as the smallest time at which the trajectory initialized in $\mathbf{x}_0$ hits the boundary $\partial\mathbf{X}$:

$$\tau(\mathbf{x}_0) := \inf\{t \geq 0 : \mathbf{x}(t|\mathbf{x}_0) \notin \mathbf{\Omega}\}.$$

Definition 6.5 allows for straightforward characterizations of the previously introduced sets of interest: for $T \geq 0$,

$$\mathbf{X}_T = \{\mathbf{x} \in \mathbf{\Omega} : \tau(\mathbf{x}) > T\}, \quad \mathbf{X}_T^{\mathrm{c}} = \{\mathbf{x} \in \mathbf{\Omega} : \tau(\mathbf{x}) \leq T\}$$
$$\mathbf{X}_\infty = \{\mathbf{x} \in \mathbf{\Omega} : \tau(\mathbf{x}) = \infty\}, \quad \mathbf{X}_\infty^{\mathrm{c}} = \{\mathbf{x} \in \mathbf{\Omega} : \tau(\mathbf{x}) < \infty\}.$$

In the rest of this section we make the assumption that the average exit time of trajectories leaving $\mathbf{\Omega}$ is finite:

Assumption 6.6: Exit time integrability

$$\overline{\tau} := \frac{1}{\lambda(\mathbf{X})} \int_{\mathbf{X}_\infty^{\mathrm{c}}} \tau(\mathbf{x}) \, \mathrm{d}\mathbf{x} < +\infty.$$

Remark 6.5 (Motivation for Assumption 6.6) *This assumption is necessary for the rigorous proof of convergence of the sequence of approximations of $\mathbf{X}_\infty$. It is difficult to check a priori. We will show however that, independently of this assumption, the validity of our approximations can be checked numerically a posteriori.*

6.1.2 Primal approximation problem and its value

As seen in the literature, the standard set approximation hierarchies are designed for *outer* approximations: see Section 3.1.2, [42], [57]. Then, inner approximations of a set $\mathbf{K}$ are obtained as a byproduct of the complementary outer approximation of $\mathbf{K}^{\mathrm{c}}$: see Corollary 5.10 and [58]. In the present case, we thus aim at outer approximations of $\mathbf{X}_\infty^{\mathrm{c}}$. We take our inspiration in the concept of reachable set. Consider the continuous time dynamical system (6.1) as well as the discrete time dynamical system

$$\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k) \tag{6.3}$$

so that $\mathbf{x}_k = \mathbf{f}^k(\mathbf{x}_0)$ for any $k \in \mathbb{N}$, $\mathbf{x}_0 \in \mathbf{X}$, where $\mathbf{f}^k = \underbrace{\mathbf{f} \circ \mathbf{f} \ldots \circ \mathbf{f}}_{k \text{ times}}$.

Definition 6.7: Reachable sets

Let $\mathbf{K}_0 \subset \mathbf{X}$ and inductively define $\mathbf{K}_{r+1} := \mathbf{f}(\mathbf{K}_r) \cap \mathbf{X}$.

- The *(forward) reachable set (RS)* of $\mathbf{K}_0$ for system (6.1) is defined as

$$\mathbf{R}_{\mathbf{K}_0}^{\mathbf{X}} := \{\mathbf{x}(T|\mathbf{x}_0) : (T \geq 0) \wedge \mathbf{x}_0 \in \mathbf{K}_0 \text{ s.t. } \forall t \in [0,T], \mathbf{x}(t|\mathbf{x}_0) \in \mathbf{X}\}$$

- The *(forward) reachable set (RS)* of $\mathbf{K}_0$ for system (6.3) is defined as

$$\mathbf{R}_{\mathbf{K}_0}^{\mathbf{X}} := \left\{\mathbf{f}^K(\mathbf{x}_0) : K \in \mathbb{N} \wedge \mathbf{x}_0 \in \mathbf{K}_0 \text{ s.t. } \forall k \in \mathbb{N}_K^\star, \mathbf{f}^k(\mathbf{x}_0) \in \mathbf{X}\right\} = \bigcup_{r=0}^{\infty} \mathbf{K}_r$$

- The *backward reachable set (BRS)* of $\mathbf{K}_0$ for a discrete (resp. continuous) time system is the RS if the discrete (resp. continuous) time system with $\mathbf{f}$ replaced with $-\mathbf{f}$ (*i.e.* time flowing backwards).

On the one hand, in [83], the authors dealt with the problem of approximating by outside the forward reachable set for a given initializing set, in the setting of discrete time dynamical systems (6.3). On the other hand, $\mathbf{X}_\infty^{\mathrm{c}}$ can be seen as the *continuous time, backward* reachable set of the boundary $\partial\mathbf{X}$: indeed, it is constituted of all the initial conditions generating trajectories that hit the boundary $\partial\mathbf{X}$. Thus, our contribution here mostly consists of a (nontrivial) adaptation of the work found in [83] to continuous time systems.

Again, such a method resorts to occupation measures, in a slightly adapted form: instead of working with the occupation measure $\nu_\mu = \mathbf{I}\times\mathbf{A} \mapsto \int_{\mathbf{X}} \int_{\mathbf{I}} \mathbb{1}_{\mathbf{A}}(\mathbf{x}(t|\mathbf{x}_0))\, \mathrm{d}t\, d\mu$ (see Section 3.1.1), consider its time average counterpart

$$\overline{\nu}_\mu := \mathbf{A} \mapsto \int_{\mathbf{X}} \left(\int_0^{\tau(\mathbf{x}_0)} \mathbb{1}_{\mathbf{A}}(\mathbf{x}(t|\mathbf{x}_0))\, \mathrm{d}t \right) d\mu.$$

Also, we replace the terminal measure $\xi_\mu \in \mathcal{M}(\mathbf{K})_+$ with a boundary measure $\gamma_\mu \in \mathcal{M}(\partial\mathbf{X})_+$ which will measure where the trajectories leave $\mathbf{X}$. Then, if $\mu \in \mathcal{M}(\mathbf{X}_\infty^{\mathrm{c}})_+$, the Liouville equation (3.3) can be integrated with respect to time to yield

$$\mathrm{div}(\overline{\nu}_\mu\, \mathbf{f}) + \gamma_\mu = \mu, \tag{3.3$'$}$$

which we incorporate into our traditional set approximation problem. Eventually, we enforce "by hand" an upper bound on the mass of our average occupation measure.

Problem 17: Primal inner MPI approximation LP

For a given $\tau_+ \in \mathbb{R}_+$, we define the following instance of the GMP

$$\begin{aligned} \mathrm{p}^\star_{\mathrm{MPI}}(\tau_+) := \sup_{\mu,\overline{\nu},\gamma} & \int 1 \, d\mu \\ \text{s.t. } & \mu \in \mathcal{M}(\mathbf{X})_+ \\ & \overline{\nu} \in \mathcal{M}(\mathbf{X})_+ \\ & \gamma \in \mathcal{M}(\partial\mathbf{X})_+ \\ & \mu \preceq \lambda & \text{(6.4a)} \\ & \mathrm{div}(\overline{\nu}\,\mathbf{f}) + \gamma - \mu = 0 & \text{(6.4b)} \\ & \int 1 \, d\overline{\nu} \leq \tau_+ \, \lambda(\mathbf{X}) & \text{(6.4c)} \end{aligned}$$

Remark 6.6 (Loss of uniqueness)

While the classical Liouville equation (3.3) *has a unique solution* $(\nu, \xi) = (\nu_\mu, \xi_\mu)$*, it is not the case anymore with its integral counterpart: there can be several couples* $(\overline{\nu}, \gamma)$*, different from the expected* $(\overline{\nu}_\mu, \gamma_\mu)$ *and satisfying* (3.3′)*. For example, if* $\mu = 0$*, then any* $(\overline{\nu}, \gamma) = (t\,\delta_{\overline{\mathbf{x}}}, 0)$ *is a solution to* (3.3′)*, where* $\overline{\mathbf{x}}$ *is the L-S equilibrium point for* $\mathbf{f}$ *and* $t \in \mathbb{R}$*. Thus, Theorem 3.5 for the finite time ROA does not have an MPI counterpart, making the analysis of Problem 17 harder to carry out.*

Remark 6.7 (Strong duality condition)

Here, τ_+ *is introduced to ensure that all the feasible measures have a finite mass. Otherwise, the strong duality Theorem 2.6 would not hold. We give more details on this subject in Section 6.1.3.*

Note that problem (6.4) is linear in the decision variables which are the three measures $\mu, \overline{\nu}, \gamma$. The two following lemmata link the infinite-dimensional LP (6.4) and the MPI set $\mathbf{X}_\infty$.

Lemma 6.8: Upper bound on $\mathrm{p}^\star_{\mathrm{MPI}}$

Assuming that $\tau_+ \geq \overline{\tau}$, we have $\mathrm{p}^\star_{\mathrm{MPI}}(\tau_+) \geq \lambda(\mathbf{X}^{\mathrm{c}}_\infty)$.

Proof :

- $\mu^\star := \lambda_{\mathbf{X}^{\mathrm{c}}_\infty}$
- $\overline{\nu}^\star := \mathbf{A} \mapsto \int_{\mathbf{X}^{\mathrm{c}}_\infty} \int_0^{\tau(\mathbf{x}_0)} \mathbb{1}_{\mathbf{A}}(\mathbf{x}(t|\mathbf{x}_0)) \, \mathrm{d}t \, \mathrm{d}\mathbf{x}_0$
- $\gamma^\star := \mathbf{A} \mapsto \int_{\mathbf{X}^{\mathrm{c}}_\infty} \mathbb{1}_{\mathbf{A}}(\mathbf{x}(\tau(\mathbf{x}_0)|\mathbf{x}_0)) \, \mathrm{d}\mathbf{x}_0$

define a feasible triplet. Indeed, constraint (6.4a) is automatically satisfied and one has :

- $\int 1 \, d\overline{\nu}^\star = \int_{\mathbf{X}_\infty^c} \left(\int_0^{\tau(\mathbf{x}_0)} \mathrm{d}t \right) \mathrm{d}\mathbf{x}_0 = \overline{\tau} \, \lambda(\mathbf{X}) \leq \tau_+ \lambda(\mathbf{X})$ so that (6.4c) holds
- constraint (6.4b) is satisfied, since $\forall v \in \mathcal{C}^1(\mathbf{X})$,

$$\begin{aligned} \langle \mathrm{div}(\overline{\nu}^\star \, \mathbf{f}), v \rangle &= - \int_{\mathbf{X}_\infty^c} \int_0^{\tau(\mathbf{x}_0)} \mathbf{grad}\, v(\mathbf{x}(t|\mathbf{x}_0)) \cdot \mathbf{f}(\mathbf{x}(t|\mathbf{x}_0)) \, \mathrm{d}t \, \mathrm{d}\mathbf{x}_0 \\ &= - \int_{\mathbf{X}_\infty^c} (v(\mathbf{x}(\tau(\mathbf{x}_0)|\mathbf{x}_0)) - v(\mathbf{x}_0)) \, \mathrm{d}\mathbf{x}_0 \\ &= \langle \mu^\star - \gamma^\star, v \rangle. \end{aligned}$$

then, $\mathrm{p}^\star_{\mathrm{MPI}}(\tau_+) \geq \int 1 \, d\mu^\star = \int 1 \, d\lambda_{\mathbf{X}_\infty^c} = \lambda(\mathbf{X}_\infty^c)$. ♢

Lemma 6.9: Lower bound for $\mathrm{p}^\star_{\mathrm{MPI}}$

For any triplet $(\mu, \overline{\nu}, \gamma)$ feasible in (6.4), μ is supported on $\mathbf{X}_\infty^c$, i.e.

$$\int \mathbb{1}_{\mathbf{X}_\infty} \, d\mu = 0.$$

The proof of this lemma uses the following assumption on the MPI set:

Assumption 6.10: Choice of X

$\forall \mathbf{x} \in \partial \mathbf{X}_\infty \cap \partial \mathbf{X}$, $\mathbf{f}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) < 0$. In words, at all points where $\partial \mathbf{X}_\infty$ is tangent to $\partial \mathbf{X}$, the trajectories strictly enter $\mathbf{X}$. Up to the choice of $\mathbf{X}$, this assumption is reasonable for any physical system.

Proof : Let $(\mu, \overline{\nu}, \gamma)$ be a feasible triplet for (6.4). Let $\chi := \mathrm{div}(\overline{\nu} \, \mathbf{f}) \overset{(6.4b)}{=} \mu - \gamma$. For $\mathbf{x} \in \mathbb{R}^n$, let

$$\varphi(\mathbf{x}) := \begin{cases} K \exp\left(-\frac{1}{1-|\mathbf{x}|^2}\right) & \text{if } |\mathbf{x}| < 1 \\ 0 & \text{else} \end{cases}$$

where $K > 0$ is such that $\int \varphi \, d\lambda = 1$. Then, for $\varepsilon > 0$ and $\mathbf{x} \in \mathbb{R}^n$, let:

- $\varphi_\varepsilon(\mathbf{x}) := \frac{1}{\varepsilon} \varphi\left(\frac{\mathbf{x}}{\varepsilon}\right) \geq 0$
- $\overline{\nu}_\varepsilon(\mathbf{x}) := \int_{\mathbf{X}} \varphi_\varepsilon(\mathbf{y} - \mathbf{x}) \, d\mu(\mathbf{y}) \geq 0$
- $\chi_\varepsilon(\mathbf{x}) := \mathrm{div}(\overline{\nu}_\varepsilon \, \mathbf{f})(\mathbf{x}) = \mathbf{grad}\, \overline{\nu}_\varepsilon(\mathbf{x}) \cdot \mathbf{f}(\mathbf{x}) + \overline{\nu}_\varepsilon(\mathbf{x}) \, \mathrm{div}\, \mathbf{f}(\mathbf{x})$.

According to the theory of mollifiers (see [15, Section 4.4]), φ, φ_ε, $\overline{\nu}_\varepsilon$ and χ_ε are smooth compactly supported functions, and for any $w \in \mathcal{C}_c^0(\mathbb{R}^n)$,

$$\int_{\mathbb{R}^n} w(\mathbf{x}) \, \overline{\nu}_\varepsilon(\mathbf{x}) \, dx \underset{\varepsilon \to 0}{\longrightarrow} \int_{\mathbf{X}} w(\mathbf{x}) \, d\mu(\mathbf{x})$$

from which it directly follows that for $v \in \mathcal{C}^1_c(\mathbb{R}^n)$

$$\begin{aligned}\int_{\mathbb{R}^n} v(\mathbf{x})\,\chi_\varepsilon(\mathbf{x})\,dx &= \int_{\mathbb{R}^n} v(\mathbf{x})\,\mathrm{div}(\overline{\nu}_\epsilon\,\mathbf{f})(\mathbf{x})\,\mathrm{d}\mathbf{x}\\ &= -\int_{\mathbb{R}^n} \mathbf{grad}\,v(\mathbf{x})\cdot\mathbf{f}(\mathbf{x})\,\overline{\nu}_\varepsilon(\mathbf{x})\,\mathrm{d}\mathbf{x}\\ &\underset{\varepsilon\to 0}{\longrightarrow} -\int_{\mathbb{R}^n} \mathbf{grad}\,v(\mathbf{x})\cdot\mathbf{f}(\mathbf{x})\,d\mu(\mathbf{x})\\ &= \int_{\mathbb{R}^n} v(\mathbf{x})\,d\chi(\mathbf{x}).\end{aligned}$$

By density of $\mathcal{C}^1_c(\mathbb{R}^n)$ in $\mathcal{C}^0_c(\mathbb{R})$ with respect to the supremum norm $\|.\|_{L^\infty(\mathbb{R}^n)}$, this implies that $\chi_\varepsilon\ \lambda$ weak-$*$ converges (in the sense of measures) to χ.

For a given $\eta > 0$ consider the set

$$\mathbf{X}^\eta_\infty := \left\{\mathbf{x}\in\mathbf{X}_\infty : \inf_{\mathbf{y}\in\partial\mathbf{X}} |\mathbf{x}-\mathbf{y}| > \eta\right\}.$$

By definition, $\mathbf{X}^\eta_\infty \cap \partial\mathbf{X} = \varnothing$, and then for any Borel set $\mathbf{A}\subset\mathbf{X}^\eta_\infty$, one has $\chi(\mathbf{A}) = \mu(\mathbf{A})$. In particular, $\chi(\partial\mathbf{X}^\eta_\infty) = \mu(\partial\mathbf{X}^\eta_\infty) = 0$ since $\mu \preceq \lambda$. Then, we can apply the Portemanteau theorem (equality marked with a $*$, see [56, Theorem 13.16]) to $\chi(\mathbf{X}^\eta_\infty)$:

$$\begin{aligned}\mu(\mathbf{X}^\eta_\infty) &= \chi(\mathbf{X}^\eta_\infty)\\ &\overset{*}{=} \lim_{\varepsilon\to 0}\int_{\mathbf{X}^\eta_\infty} \chi_\varepsilon(\mathbf{x})\,\mathrm{d}\mathbf{x}\\ &\overset{\text{def}}{=} \lim_{\varepsilon\to 0}\int_{\mathbf{X}^\eta_\infty} \mathrm{div}(\overline{\nu}_\varepsilon\,\mathbf{f})(\mathbf{x})\,\mathrm{d}\mathbf{x}\\ &= \lim_{\varepsilon\to 0}\int_{\partial\mathbf{X}^\eta_\infty} \mathbf{f}(\mathbf{x})\cdot\mathbf{n}_\eta(\mathbf{x})\,\overline{\nu}_\varepsilon(\mathbf{x})\,d\sigma(\mathbf{x})\end{aligned} \tag{6.5}$$

where $\mathbf{n}_\eta$ stands for the unit normal vector to $\partial\mathbf{X}^\eta_\infty$ pointing towards $\mathbf{X}^{\eta\mathrm{c}}_\infty$, according to the Gauss formula (4.4). Now, let R_+ be the function

$$\mathrm{R}_+ : \left\{\begin{array}{ccc}\partial\mathbf{X}_\infty\cap\partial\mathbf{X} & \longrightarrow & \mathbb{R}\\ \mathbf{x} & \longmapsto & \sup\left\{\begin{array}{c} R>0 : \forall\eta\in(0,R), \forall\mathbf{y}\in\partial\mathbf{X}^\eta_\infty\\ |\mathbf{x}-\mathbf{y}|<R \Longrightarrow \mathbf{f}(\mathbf{y})\cdot\mathbf{n}_\eta(\mathbf{y})\le 0\end{array}\right\}\end{array}\right..$$

In words, $\mathrm{R}_+(\mathbf{x})$ is the largest range around $\mathbf{x}$ in which the $\mathbf{f}\cdot\mathbf{n}_\eta$ are non-positive. According to Assumption 6.10, $\mathbf{f}$ being continuous, R_+ takes only positive values. Moreover, due to the regularity of $\mathbf{f}$, $\mathbf{X}$ and $\mathbf{X}_\infty$, R_+ is continuous on the compact set $\partial\mathbf{X}_\infty\cap\partial\mathbf{X}$, therefore it attains a minimum $\mathrm{R}^\star_+ > 0$.

Let $\eta\in(0,\mathrm{R}^\star_+)$, $\mathbf{x}\in\partial\mathbf{X}^\eta_\infty$. Then, there are two possibilities:

- either $\mathbf{x}\in\partial\mathbf{X}_\infty$, and then by positive invariance of $\mathbf{X}_\infty$, $\mathbf{f}(\mathbf{x})\cdot\mathbf{n}_\eta(\mathbf{x})\le 0$;
- or $\inf_{\mathbf{y}\in\partial\mathbf{X}} |x-\mathbf{y}| = \eta < \mathrm{R}^\star_+$, and by definition of $\mathrm{R}^\star_+$, $\mathbf{f}(\mathbf{x})\cdot\mathbf{n}_\eta(\mathbf{x})\le 0$.

It follows that for any $\mathbf{x} \in \partial\mathbf{X}_\infty^\eta$, $\mathbf{f}(\mathbf{x}) \cdot \mathbf{n}_\eta(\mathbf{x}) \leq 0$. Thus, one obtains

$$\int_{\partial\mathbf{X}_\infty^\eta} \mathbf{f}(\mathbf{x}) \cdot \mathbf{n}_\eta(\mathbf{x})\, \overline{\nu}_\varepsilon(\mathbf{x})\, \mathrm{d}\mathbf{x} \leq 0$$

and after letting ε tend to 0, using equation (6.5), we have $\mu(\mathbf{X}_\infty^\eta) \leq 0$, which means, by non-negativity of μ, that $\mu(\mathbf{X}_\infty^\eta) = 0$.

Eventually, since $\mathbf{X}_\infty^\eta \subset \mathbf{X}_\infty$ and $\mu \preceq \lambda$, one has

$$\begin{aligned}
\mu(\mathbf{X}_\infty) &= \mu(\mathbf{X}_\infty) - \mu(\mathbf{X}_\infty^\eta) \\
&= \mu(\mathbf{X}_\infty \setminus \mathbf{X}_\infty^\eta) \\
&\leq \lambda(\mathbf{X}_\infty \setminus \mathbf{X}_\infty^\eta) \\
&= \lambda\left(\left\{\mathbf{x} \in \mathbf{X}_\infty \;;\; \inf_{\mathbf{y}\in\partial\mathbf{X}} |\mathbf{x}-\mathbf{y}| \leq \eta\right\}\right) \\
&\underset{\eta\to 0}{\longrightarrow} \lambda(\mathbf{X}_\infty \cap \partial\mathbf{X}) = 0
\end{aligned}$$

which leads to the conclusion that $\mu(\mathbf{X}_\infty) = 0$. ♢

Theorem 6.11: Value of Problem 17

Assuming that $\tau_+ \geq \overline{\tau}$, the infinite-dimensional LP (6.4) has a value $\mathrm{p}^\star_{\mathrm{MPI}}(\tau_+) = \lambda(\mathbf{X}_\infty^{\mathrm{c}})$. Moreover the supremum is attained, and the $\mu^\star$ component of any solution is necessarily the measure $\lambda_{\mathbf{X}_\infty^{\mathrm{c}}}$.

Proof : This is a straightforward consequence of Lemmata 6.8 and 6.9. ♢

6.1.3 Dual approximation problem and its value

According to the developments in Section 2.1, the dual LP of problem (6.4) reads as follows.

Problem 18: Primal inner MPI approximation LP

For a given $\tau_+ \in \mathbb{R}_+$, we define the following infinite-dimensional LP

$$\begin{aligned}
\mathrm{d}^\star_{\mathrm{MPI}}(\tau_+) := \inf_{u,v,w} & \int_{\mathbf{X}} (w(\mathbf{x}) + \tau_+\, u)\, \mathrm{d}\mathbf{x} \\
\text{s.t. } & w - v - 1 \in \mathcal{C}(\mathbf{X})_+ && \text{(6.6a)} \\
& u - \mathbf{f}\cdot\mathbf{grad}\, v \in \mathcal{C}(\mathbf{X})_+ && \text{(6.6b)} \\
& v \in \mathcal{C}(\partial\mathbf{X})_+ && \text{(6.6c)} \\
& w \in \mathcal{C}(\mathbf{X})_+ \\
& v \in \mathcal{C}^1(\mathbf{X}) \\
& u \in \mathbb{R}_+.
\end{aligned}$$

Lemma 6.12: Positive invariance

Let $(0, v, w)$ be a feasible triplet for problem (6.6).
Then, the set $\hat{\mathbf{X}}_\infty := \{\mathbf{x} \in \mathbf{\Omega} : v(\mathbf{x}) < 0\}$ is a PI subset of $\mathbf{X}_\infty$.

Proof : Since $\mathbf{X}_\infty$ is the MPI set included in $\mathbf{X}$ and $\hat{\mathbf{X}}_\infty \subset \mathbf{X}$ by definition, it is sufficient to prove that $\hat{\mathbf{X}}_\infty$ is positively invariant.

Let $\mathbf{x}_0 \in \hat{\mathbf{X}}_\infty$. Then, for any $t > 0$, one has

$$v(\mathbf{x}(t|\mathbf{x}_0)) = v(\mathbf{x}_0) + \int_0^t \nabla v \cdot \mathbf{f}(\mathbf{x}(s|\mathbf{x}_0))\, ds \leq v(\mathbf{x}_0) < 0$$

using constraint (6.6b).

We still have to show that $\mathbf{x}(t|\mathbf{x}_0)$ remains in $\mathbf{\Omega}$ at all times $t \geq 0$. If not, then there exists a $t_\partial > 0$ such that $\mathbf{x}(t_\partial|\mathbf{x}_0) \in \partial\mathbf{X}$ according to the intermediate value theorem, the trajectory being of course continuous in time. However, by feasibility of $(0, v, w)$, one then has $v(\mathbf{x}(t_\partial|\mathbf{x}_0)) \geq 0$, which is in contradiction with the fact that $v(\mathbf{x}(t|\mathbf{x}_0)) < 0$ for all $t > 0$ which we just proved.

Thus, we obtain that for all $t > 0$, $\mathbf{x}(t|\mathbf{x}_0) \in \mathbf{\Omega}$ and $v(\mathbf{x}(t|\mathbf{x}_0)) < 0$, i.e. $\mathbf{x}(t|\mathbf{x}_0) \in \hat{\mathbf{X}}_\infty$. ♢

Remark 6.8 (Inner MPI approximation)

For a feasible triplet (u, v, w), if $u \neq 0$, then there is no guarantee that the solution of (6.6) yields an inner approximation of $\mathbf{X}_\infty$. However, it still gives access to inner approximations of the $\mathbf{X}_T$, $T \in \mathbb{R}_+$, and we will show that under Assumption 6.6, these approximations converge to $\mathbf{X}_\infty$.

Lemma 6.13: Inner approximation of $\mathbf{X}_T$

For any triplet (u, v, w) feasible in (6.6), for any $T > 0$,

$$\hat{\mathbf{X}}_T := \{\mathbf{x}_0 \in \mathbf{\Omega} : v(\mathbf{x}_0) + u\,T < 0\} \subset \mathbf{X}_T.$$

Proof : Let (u, v, w) be a feasible triplet in (6.6) and let $\mathbf{x}_0$ be an element of $\mathbf{X}_T^{\mathrm{c}}$ for a given $T > 0$.

By definition of $\mathbf{X}_T$ we know that $T \geq \tau(\mathbf{x}_0)$, where τ is the exit time, and that for any $t \in [0, \tau(\mathbf{x}_0)], \mathbf{x}(t|\mathbf{x}_0) \in \mathbf{X}$. Thanks to constraint (6.6b), we can therefore say that for any $t \in [0, \tau(\mathbf{x}_0)], (\nabla v \cdot \mathbf{f})(\mathbf{x}(t|\mathbf{x}_0)) \leq u$. Hence for any $t \in [0, \tau(\mathbf{x}_0)], v(\mathbf{x}(t|\mathbf{x}_0)) \leq v(\mathbf{x}_0) + u\,t$. In particular, we deduce that

$$v(\mathbf{x}(\tau(\mathbf{x}_0)|\mathbf{x}_0)) \leq v(\mathbf{x}_0) + u\tau(\mathbf{x}_0) \leq v(\mathbf{x}_0) + u\,T.$$

As $\mathbf{x}(\tau(\mathbf{x}_0)|\mathbf{x}_0) \in \partial\mathbf{X}$, (6.6c) yields that $v(\mathbf{x}(\tau(\mathbf{x}_0)|\mathbf{x}_0)) \geq 0$ and thus $v(\mathbf{x}_0) \geq -u\,T$. This proves that

$$\mathbf{X}_T^{\mathrm{c}} \subset \{\mathbf{x}_0 \in \mathbf{\Omega} : v(\mathbf{x}_0) \geq -u\,T\}$$

hence $\hat{\mathbf{X}}_T \subset \mathbf{X}_T$. ♢

Theorem 6.14: Strong duality

There is no duality gap between primal LP problem (6.4) on measures and dual LP problem (6.6) on functions in the sense that $\mathrm{p}^\star_{\mathrm{MPI}}(\tau_+) = \mathrm{d}^\star_{\mathrm{MPI}}(\tau_+)$.

Proof : This is a direct consequence of Theorem 2.6: the only thing that we have to prove is the boundedness of the masses of μ, $\overline{\nu}$ and γ. Boundedness of μ and $\overline{\nu}$ respectively follow from constraints (6.4a) and (6.4c). Then,

$$\int 1 \, d\gamma \overset{(6.4b)}{=} \int 1 \, d\mu$$

so that the mass of γ admits the same bound as the mass of μ. ♢

Remark 6.9 (The importance of τ_+)

Taking τ_+ to infinity or, equivalently, removing constraint (6.4c) *boils down to imposing $u = 0$ in the dual problem* (6.6)*. Although this would indeed ensure that our approximating set $\hat{\mathbf{X}}_\infty$ is included in the actual MPI set $\mathbf{X}_\infty$, it would also destroy our proof of strong duality.*

But looking in detail leads to understand that things are even worse: without constraint (6.4c)*, any feasible $\overline{\nu}$ generates an infinity of feasible candidates $\overline{\nu}_t := \overline{\nu} + t\, \delta_{\overline{\mathbf{x}}}$, where $\mathbf{x}$ is the L-S equilibrium point of system* (6.1) *and $\delta_{\mathbf{x}}$ is the Dirac measure in $\mathbf{x}$. Actually, this $\delta_{\overline{\mathbf{x}}}$ would be a nontrivial element of the kernel of $(\mathcal{A}, \mathbf{c})$ (using the notations of Chapter 2), i.e. a counterexample to equation $(*)$ which is asked to hold in the proof of Theorem 2.6. In other words, constraint* (6.4c) *is absolutely necessary to enforce strong duality.*

Without this constraint, the results in Section 6.1.2 would still hold, yielding that $\mathrm{p}^\star_{\mathrm{MPI}}(\tau_+) = \mathrm{vol}\, \mathbf{X}^{\mathrm{c}}_\infty$, but one would only have $\mathrm{d}^\star_{\mathrm{MPI}}(\tau_+) \geq \mathrm{p}^\star_{\mathrm{MPI}}(\tau_+)$. From a "Slater" viewpoint (see [70][pp. 8,128,313] or [11][p. 171]) problem (6.6) *might not admit any interior point. Finding a way to bypass the artificial use of constraint* (6.4c) *is actually an open question for inner MPI computation.*

6.1.4 Numerical implementation and its convergence

Following the developments of Section 2.2.2, problem (6.6) admits a SOS tightening which can be written as follows:

$$\begin{aligned}
\mathrm{d}^d_{\mathrm{MPI}}(\tau_+) := \inf_{\substack{u,v,w,p \\ (s_{ij})_{i,j}}} \; & \mathbf{w} \cdot \mathbf{l} + u\, \tau_+\, l_{\mathbf{0}} \\
\text{s.t.} \; & w - v - 1 = s_{10} + \mathbf{s}_1 \cdot \mathbf{g} && \text{(6.7a)} \\
& u - \mathbf{f} \cdot \mathbf{grad}\, v = s_{20} + \mathbf{s}_2 \cdot \mathbf{g} && \text{(6.7b)} \\
& v = s_{30} + \mathbf{s}_3 \cdot \mathbf{g} + p\, h && \text{(6.7c)} \\
& w = s_{40} + \mathbf{s}_4 \cdot \mathbf{g} \\
& v, w \in \mathbb{R}_{2d}[\mathbf{x}] \\
& u \in \mathbb{R}_+
\end{aligned}$$

$$
\begin{aligned}
&s_{10}, s_{10}, s_{10}, s_{10} \in \Sigma_d[\mathbf{x}] \\
&s_{1i}, s_{1i}, s_{1i}, s_{1i} \in \Sigma_{d-\lceil d_i/2 \rceil}[\mathbf{x}] \\
&p \in \mathbb{R}_{2d-\sum_i d_i}[\mathbf{x}],
\end{aligned}
$$

where $h = g_1 \cdots g_{m_\mathbf{X}}$. Vector $\mathbf{l}$ denotes the Lebesgue moments over $\mathbf{X}$ indexed in the same basis in which the polynomial w with vector of coefficients $\mathbf{w}$ is expressed.

SOS problem (6.7) is a tightening of problem (6.6) in the sense that any feasible solution in (6.7) gives a triplet (u, v, w) feasible in (6.6).

Theorem 6.15: Inner MPI approximation

Problem (6.7) is an LMI problem and any feasible solution (u_d, v_d, w_d) gives inner approximations $\hat{\mathbf{X}}_t^d := \{\mathbf{x} \in \mathbf{\Omega} : v_d(\mathbf{x}) + u_k t < 0\}$ of the $\mathbf{X}_T$s. In particular, if $u_d = 0$, $\hat{\mathbf{X}}_\infty^d := \{\mathbf{x} \in \mathbf{\Omega} : v_d(\mathbf{x}) < 0\}$ is an inner approximation of $\mathbf{X}_\infty$.

Proof : This is a direct consequence of Proposition 2.13 as well as Lemmata 6.12 and 6.13. ♢

This SOS tightening is a finite dimension convex optimization, and as such it admits a primal formulation derived from Lagrangian theory, which can be seen as an LMI relaxation of infinite dimensional LP (6.4) (see Section 2.2 for details).

Theorem 6.16: Convergence of the hierarchy

Let $\tau_+ > \overline{\tau}$. Then, under Assumption 2.7 for $\mathbf{X}$,

1. The sequence $(\mathrm{d}_{\mathrm{MPI}}^d(\tau_+))$ is monotonically decreasing and converging to $\lambda(\mathbf{X}_\infty^{\mathrm{c}})$

For every $d \geq d_{\min} := 1/2 \sum_i d_i$, let $\psi_d := (u_d, v_d, w_d)$ denote a $\frac{1}{d}$-optimal solution of the dual tightening of order d. One has then :

2. $u_d \underset{d\to\infty}{\longrightarrow} 0$,
3. $w_d \overset{L^1(\mathbf{X})}{\underset{d\to\infty}{\longrightarrow}} \mathbb{1}_{\mathbf{X}_\infty^{\mathrm{c}}}$.

Proof :

1. is a direct consequence of Theorems 2.14 and 6.14.
2. We define $\phi^\star := (\mu^\star, \overline{\nu}^\star, \gamma^\star)$ feasible for (6.4) as in the proof of Lemma 6.8. Then,

$$
\begin{aligned}
u_d \int 1 \, d\overline{\nu}^\star &= \int u_d \, d\overline{\nu}^\star \\
&\overset{(6.6b)}{\geq} \int \mathbf{f} \cdot \mathbf{grad}\, v_d \, d\overline{\nu}^\star \\
&\overset{(6.4b)}{=} \int v_d \, d\gamma^\star - \int v_d \, d\mu^\star
\end{aligned}
$$

$$\begin{aligned}
&\overset{(6.6c)}{\geq} -\int v_d \, d\mu^\star \\
&\overset{(6.6a)}{\geq} \int (1 - w_d) \, d\mu^\star \\
&\overset{\text{def}}{=} \lambda(\mathbf{X}_\infty^{\mathrm{c}}) - \int w_d \, d\mu^\star \\
&\overset{(6.4a)}{\geq} \lambda(\mathbf{X}_\infty^{\mathrm{c}}) - \int_{\mathbf{X}} w_d(\mathbf{x}) \, \mathrm{d}\mathbf{x}
\end{aligned} \tag{6.8}$$

so that

$$\begin{aligned}
0 &\overset{(6.4c)}{\leq} (\tau_+ \, \lambda(\mathbf{X}) - \overline{\nu}^\star(\mathbf{X})) \, u_d \\
&\overset{(6.8)}{\leq} \int_{\mathbf{X}} (w_d(\mathbf{x}) + \tau_+ \, u_d) \, \mathrm{d}\mathbf{x} - \lambda(\mathbf{X}_\infty^{\mathrm{c}}) \\
&\overset{\frac{1}{d}\text{-optim.}}{\leq} \mathrm{d}_{\mathrm{MPI}}^d(\tau_+) + \frac{1}{d} - \lambda(\mathbf{X}_\infty^{\mathrm{c}}) \\
&\xrightarrow[d\to\infty]{1.} 0.
\end{aligned}$$

Since by assumption $\tau_+ > \bar{\tau} = \frac{\overline{\nu}^\star(\mathbf{X})}{\lambda(\mathbf{X})}$, this means that $u_d \underset{d\to\infty}{\longrightarrow} 0$.

3. Let $\varepsilon > 0$. Let $T > 0$ such that $\lambda(\mathbf{X}_T \setminus \mathbf{X}_\infty) \leq \varepsilon$. Let $\bar{d} \geq d_{\min}$ such that for all $d \geq \bar{d}$ one has that $\|u_d \, T\|_{L^1(\mathbf{X})} \leq \varepsilon$ and $|\int_{\mathbf{X}} w_d(\mathbf{x}) \, \mathrm{d}\mathbf{x} - \lambda(\mathbf{X}_\infty^{\mathrm{c}})| \leq \varepsilon$. Such an integer exists from points 1 and 2. Using the triangle inequality and the fact that $\|u_d \, T\|_{L^1(\mathbf{X})} \leq \varepsilon$ one has

$$\|w_d - \mathbb{1}_{\mathbf{X}_\infty^{\mathrm{c}}}\|_{L^1(\mathbf{X})} \leq \|w_d + u_d \, T - \mathbb{1}_{\mathbf{X}_\infty^{\mathrm{c}}}\|_{L^1(\mathbf{X})} + \varepsilon. \tag{6.9}$$

With the notation $A = \|w_d + u_d \, T - \mathbb{1}_{\mathbf{X}_\infty^{\mathrm{c}}}\|_{L^1(\mathbf{X})}$, one has

$$A = \int_{\mathbf{X}_T^{\mathrm{c}}} |w_d(\mathbf{x}) + u_d \, T - \mathbb{1}_{\mathbf{X}_\infty^{\mathrm{c}}}(\mathbf{x})| \, \mathrm{d}\mathbf{x} + \int_{\mathbf{X}_T} |w_d(\mathbf{x}) + u_d \, T - \mathbb{1}_{\mathbf{X}_\infty^{\mathrm{c}}}(\mathbf{x})| \, \mathrm{d}\mathbf{x}.$$

We denote by B and C these two terms, respectively. Since $\mathbf{X}_T^{\mathrm{c}} \subset \mathbf{X}_\infty^{\mathrm{c}}$, one can replace $\mathbb{1}_{\mathbf{X}_\infty^{\mathrm{c}}}$ with 1 in B. Then, constraint (6.6a) along with Theorem 6.15 allow us to remove the absolute value to obtain

$$B = \int_{\mathbf{X}_T^{\mathrm{c}}} (w_d(\mathbf{x}) + u_d \, T - 1) \, \mathrm{d}\mathbf{x} = \int_{\mathbf{X}_T^{\mathrm{c}}} w_d(\mathbf{x}) \, \mathrm{d}\mathbf{x} - \lambda(\mathbf{X}_T^{\mathrm{c}}) + \lambda(\mathbf{X}_T^{\mathrm{c}}) u_d \, T$$

and since $\lambda(\mathbf{X}_T^{\mathrm{c}}) u_d \, T \leq \|u_d \, T\|_{L^1(\mathbf{X})} \leq \varepsilon$,

$$B \leq \int_{\mathbf{X}_T^{\mathrm{c}}} w_d(\mathbf{x}) \, \mathrm{d}\mathbf{x} - \lambda(\mathbf{X}_T^{\mathrm{c}}) + \varepsilon. \tag{6.10}$$

Moreover, we have that $C \leq \int_{\mathbf{X}_T} (|w_d(\mathbf{x})| + |u_d \, T| + |\mathbb{1}_{\mathbf{X}_\infty^{\mathrm{c}}}(\mathbf{x})|) \, \mathrm{d}\mathbf{x}$ and therefore, using the nonnegativity of w_d and the fact that $\|u_d \, T\|_{L^1(\mathbf{X})} \leq \varepsilon$, one has $C \leq \int_{\mathbf{X}_T} w_d(\mathbf{x}) \, \mathrm{d}\mathbf{x} + \varepsilon + \lambda(\mathbf{X}_T \setminus \mathbf{X}_\infty)$. Since we have $\lambda(\mathbf{X}_T \setminus \mathbf{X}_\infty) \leq \varepsilon$ by choice

of T, we deduce that $C \leq \int_{\mathbf{X}_T} w_d(\mathbf{x})\mathrm{d}\mathbf{x} + 2\varepsilon$. Combining this inequality with (6.10), we have :

$$A = B + C \leq \int_{\mathbf{X}} w_d(\mathbf{x})\ \mathrm{d}\mathbf{x} - \lambda(\mathbf{X}_T^{\mathrm{c}}) + 3\varepsilon$$

from which we deduce that $A \leq 5\varepsilon$, using that $|\int_{\mathbf{X}} w_d(\mathbf{x})\ \mathrm{d}\mathbf{x} - \lambda(\mathbf{X}_\infty^{\mathrm{c}})| \leq \varepsilon$ and $\lambda(\mathbf{X}_\infty^{\mathrm{c}} \setminus \mathbf{X}_T^{\mathrm{c}}) \leq \varepsilon$. Combining this with (6.9), we have that

$$\|w_d - \mathbb{1}_{\mathbf{X}_\infty^{\mathrm{c}}}\|_{L^1(\mathbf{X})} \leq 6\varepsilon.$$

♢

Remark 6.10 (No free lunch rule) *Despite this convergence result, one should be aware of the fact that the computational burden increases sharply with the dimension of the state space and the degree of the relaxations. Indeed, the involved polynomials have $D_n^{2d} = \binom{n+2d}{n}$ coefficients. Consequently, high values of n and d might be intractable. A possible way to handle this consists in exploiting the structure of the considered problems, such as sparsity. The key is to split the state space into low dimensional subspaces and distribute the problem over the obtained partitioning (see Chapter 5 as a first example of what can be done in practice for volume computation).*

For this section, we chose to focus on the simple example of the Van der Pol oscillator, as was done in [42]:

$$\begin{cases} \dot{\mathbf{x}}_1 = -2\ \mathbf{x}_2 \\ \dot{\mathbf{x}}_2 = 0.8\ \mathbf{x}_1 + 10\ (1.02^2 x_1^2 - 0.2)\ \mathbf{x}_2. \end{cases} \tag{6.11}$$

Let $\mathbf{X} = \{\mathbf{x} \in \mathbb{R}^2 : \mathbf{x}_1^2 + \mathbf{x}_2^2 \leq 1\}$ and $\tau_+ = \frac{100}{\pi}$.

We implemented the hierarchy of SOS problems (6.7) in MATLAB, using the toolbox YALMIP interfaced with the SDP solver MOSEK. For $d = 6$ and 7 (SOS degrees 12 and 14 respectively), we compared the obtained regions to the outer approximations computed using the framework presented in [57], see Figure 6.1. In this implementation, we checked at each relaxation whether u was near to zero: for $d = 6$, we had $u \sim 10^{-7}$, and for $d = 7$ we obtained $u \sim 10^{-6}$, which is satisfactory. Moreover, we also ran the hierarchy with constraint $u_d = 0$ (to enforce inner approximations) and obtained the same results.

However, we observed some difficulties:

- For low degrees, the only solution v found by the solver is very close to the zero polynomial: the coefficients are of the order 10^{-5}, therefore the plots are irrelevant; one loses conservativeness and several constraints are violated (namely the positivity constraint on v on $\partial\mathbf{X}$).

- For higher degrees, the basis of monomials is not adapted since for example in dimension 1 $\mathbf{x}^\alpha$ is close to the indicator of $\{-1, 1\}^n$. As a result, the coefficients are of the order 10^5 or more, and again the plots make little sense.

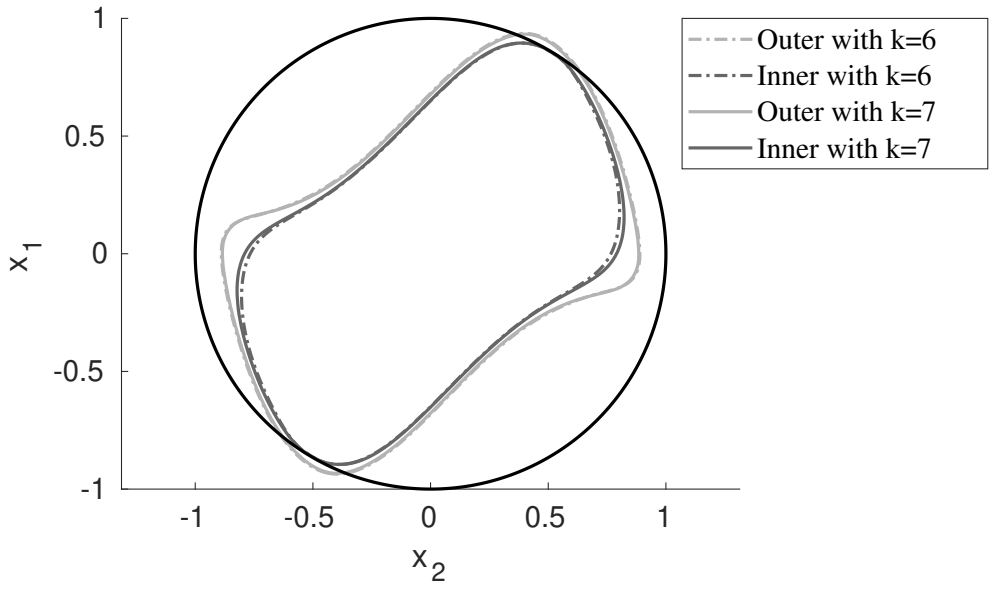


Figure 6.1 – Outer and inner MPI set approximations.

Carried out for the Van der Pol oscillator with the unit disk as the admissible state set $\mathbf{X}$. The reported k are the relaxation orders (named d in the rest of the section).

One can also find numerical applications of this method to actual electrical engineering problems in [99] with very promising results.

Our original motivation is the study of transient phenomena in large-scale electrical power systems, see Section 3.1 and references therein. Our objective is to design a hierarchy of approximations of the MPI set for large-scale systems described by non-linear differential equations. A first step towards non-polynomial dynamics can be found in [99]. Since the initial work [42] relied on the mathematical technology behind the approximation of the volume of semi-algebraic sets, we already studied in Chapter 5 the problem of approximating the volume of a large-scale sparse semi-algebraic set. We are now investigating extensions of the techniques for approximating the MPI set of large-scale sparse dynamical systems, and the current section contributes to a better understanding of its inner approximations, in the small-scale non-sparse case. Our next step consists of combining the ideas of Chapter 5 with those of the current section, so as to design a Lasserre hierarchy of inner approximations of the MPI set in the large-scale case, and apply it to electrical power system models.

As it was done at the end of Chapter 3 we display an updated comparison table of the various approaches to stability analysis, depending on the set one intends to approximate. The need for parameter τ_+ is the main drawback of this method, for two reasons:

- It should be greater than the average exit time $\overline{\tau}$ of the studied system, which most of the time is completely unknown;

- When τ_+ is too big, numerical issues arise in the hierarchy, mostly because the integral $\mathbf{w} \cdot \mathbf{l}$ in the objective of the dual (6.7) becomes negligible with respect to $\tau_+ \, u \, l_{\mathbf{0}}$.

However, the advantages of approximating the MPI set are multiple:

- The approximating scheme presents the classical convexity and convergence properties of the Lasserre hierarchy, as for the finite time ROA approximation schemes
- The involved measures and polynomials do not depend on time, so that e.g. for $d = 10$ and in state dimension 5 (which is a reasonable setting in practice), the LMI size is $D_5^{10} = 3003$ instead of $D_6^{10} = 8008$;
- From the power systems stability viewpoint, it is very interesting since it exactly computes the set of post-fault conditions that will not lead to violation of the given security constraints that make up $\mathbf{X}$.

	Lyapunov ROA	finite time ROA	MPI set
Nb of variables	n $(\mathbf{x})$	$n+1$ $(t, \mathbf{x})$	n $(\mathbf{x})$
Constraints	BMI	LMI	LMI
Convexity	no	yes	yes
Scheme convergence	local	global	global
Parameters	none	$\mathbf{X}, \mathbf{K}, T$	$\mathbf{X}, \tau_+$
Time horizon	infinite	finite	infinite
Target	$\overline{\mathbf{x}}$	$\mathbf{K}$	$\mathbf{X}$

Table 6.1 – Comparison between ROA, finite time ROA & MPI set schemes.

6.2 Sparsity-based approximation for finite time ROA

This section describes a computational technique for generating outer approximations of finite time ROA of sparse polynomial ODEs, with the purpose of assessing the stability of large scale power systems in the future. These outer approximations contain all initial conditions for which the dynamical systems can operate safely. Indeed, power networks are usually modeled by an interconnection of weakly coupled nodes, while the dynamic behaviour of the system is mainly driven by generators, which are modeled by (closed-loop controlled) ordinary differential equations.

Most of the technical literature on stability analysis for power networks focuses on the construction of Lyapunov functions computed by nonconvex optimization, and more specifically a bilinear variant of polynomial SOS optimization, as in e.g. [6, 50] and Section 3.2. An inner approximation of the infinite time ROA is then modeled as a sublevel set of the Lyapunov function, and various heuristics are used to enlarge this sublevel set as much as possible, see e.g. [20] and references therein. It can be enforced that the Lyapunov functions have the same sparsity structure as the system to be analyzed, see e.g. [147] and references therein, but to our knowldge, it was never applied to ROA approximation. Regardless of the application, some efforts have been made to improve the scalability of the costly SOS programming techniques (see [2, 3, 85] and the references therein), at the price of accuracy of the computed solutions. The work of [63, 64] is a first step towards the application of Lyapunov techniques to ROA estimation for interconnected systems. Another way of exploiting the system's structure is to rely on sparsity in terms of time scales instead of sparsity in terms of variables, see e.g. [121].

The main contribution of the current section is to identify a sparsity structure that allows us to apply the moment-SOS hierarchy for sparse ROA approximation. We construct a hierarchy of outer approximations of increasing degree, though the sparsity is introduced at the price of the convergence proof that no longer holds. For this, we rely heavily on Chapter 5 which focused on the approximation of the volume of a sparse semi-algebraic set.

In the context of differential systems stability analysis, our work can be seen an extension to large-scale systems of results of Sections 3.1 and 6.1 as well as [99]. It can be interpreted as well as a finite time dual approach to the standard Lyapunov approach of Section 3.2 and [20, 6, 64, 147]. We prefer however to see the Lyapunov approach as a dual to an infinite time occupation measure approach, in the sense that Lyapunov functions are obtained as a (dual Lagrangian) certificate of a property (stability) of the system's trajectories (modeled by occupation measures in a primal problem). The advantage of considering finite time ROA instead of standard Lyapunov ROA is the linearity of its characterization (which leads to solving convex LMIs instead of nonconvex bilinear matrix inequalities as in the Lyapunov framework), as well as the proof of convergence in volume (in the non-sparse case).

6.2.1 A path decomposition sparsity pattern

Let $N \in \mathbb{N}$. We consider the following system of sparsely coupled polynomial ODEs:

$$\dot{\mathbf{x}}_i = \mathbf{f}_i(\mathbf{x}_i, \mathbf{x}_{i+1}) \qquad \mathbf{x}_i \in \mathbf{X}_i \qquad i \in \mathbb{N}^\star_{N-1} \tag{6.12a}$$

$$\dot{\mathbf{x}}_N = \mathbf{f}_N(\mathbf{x}_{N-1}, \mathbf{x}_N) \qquad \mathbf{x}_N \in \mathbf{X}_N \tag{6.12b}$$

where for all $i \in \mathbb{N}^\star_N$, $\mathbf{X}_i := \{\mathbf{x}_i \in \mathbb{R}^{n_i} : \mathbf{g}_{\mathbf{X}_i}(\mathbf{x}_i) \geq \mathbf{0}\}$, are finite dimensional compact basic semialgebraic sets and $\mathbf{f}_i \in \mathbb{R}[\mathbf{x}_i, \mathbf{x}_{i+1}]^{n_i}$, $i \in \mathbb{N}^\star_{N-1}$, $\mathbf{f}_N \in \mathbb{R}[\mathbf{x}_{N-1}, \mathbf{x}_N]^{n_N}$ are polynomial maps. We define $\mathbb{X}_i := \mathbb{R}^{n_i}$, $\mathbb{X} := \mathbb{X}_1 \times \ldots \times \mathbb{X}_N$, $\mathbf{X} := \mathbf{X}_1 \times \ldots \times \mathbf{X}_N$, $\mathbf{f} := (\mathbf{f}_1, \ldots, \mathbf{f}_N)$ and $n := n_1 + \ldots + n_N$. Note that $\mathbf{X}$ is also a compact semialgebraic set described by the polynomial vector $\mathbf{g}_{\mathbf{X}} := (\mathbf{g}_{\mathbf{X}_1} \circ \pi_{\mathbb{X}_1}, \ldots, \mathbf{g}_{\mathbf{X}_N} \circ \pi_{\mathbb{X}_N})$.

Given a finite time horizon $T > 0$ and compact basic semialgebraic target sets $\mathbf{K}_i = \{\mathbf{x}_i \in \mathbb{X}_i : \mathbf{g}_{\mathbf{K}_i}(\mathbf{x}_i) \geq \mathbf{0}\} \subset \mathbf{X}_i$, $\mathbf{K} := \mathbf{K}_1 \times \ldots \times \mathbf{K}_N$, we aim at computing outer approximations of the finite time ROA, defined as

$$\mathbf{A}_T^{\mathbf{X}}(\mathbf{K}) := \left\{ \mathbf{x}_0 \in \mathbb{R}^n : \begin{array}{l} \forall t \in [0,T], \mathbf{x}(t|\mathbf{x}_0) \in \mathbf{X} \\ \mathbf{x}(T|\mathbf{x}_0) \in \mathbf{K} \end{array} \right\}. \tag{6.13}$$

Remark 6.11 (Link with Chapter 5)

This is the path decomposition special case of correlative sparsity that we studied in Section 5.2. The only difference here is that instead of appearing in the description of the set $\mathbf{K}$ *we indended to measure, the sparsity pattern is found in the dynamics of the differential system* $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$.

Here there are $N-1$ *cliques* $\mathbf{C}_i = \mathbb{I}(\mathbf{g}_{\mathbf{X}_i}) \cup \mathbb{I}(\mathbf{g}_{\mathbf{X}_{i+1}})$, $i \in \mathbb{N}^\star_{N-1}$ *(using the notations of Section 5.1.4), see Figure 6.2.*

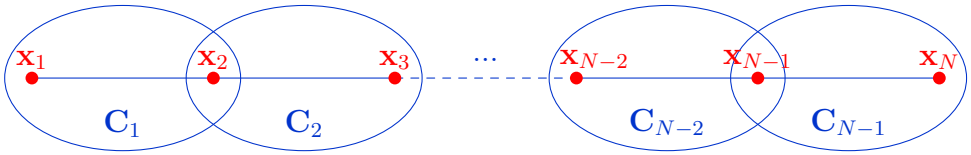


Figure 6.2 – Illustration of the studied sparsity pattern.

Remark 6.12 (Information transfer)

A way to look at this sparse differential system consists of building the exact same clique tree as in Chapter 5 (in the present context, we call it the clique path), and then look at the clique $\mathbf{C}_{N-1}$*, which is the leaf of the tree. Indeed, the dynamics of this clique do not depend on any exogenous variable, and it can be viewed as a subsystem of the form*

$$\dot{\mathbf{x}}_{N-1} = \mathbf{f}_{N-1}(\mathbf{x}_{N-1}, \mathbf{x}_N)$$
$$\dot{\mathbf{x}}_N = \mathbf{f}_N(\mathbf{x}_{N-1}, \mathbf{x}_N).$$

Then, solving this system provides us with trajectories $\mathbf{x}_{N-1}(t|\mathbf{x}_0)$, $\mathbf{x}_N(t|\mathbf{x}_0)$, *and one can consider* $\mathbf{u}_{N-2}(t) := \mathbf{x}_{N-1}(t|\mathbf{x}_0)$ *as a control law for the dynamics of* $\mathbf{x}_{N-1}$*:*

$$\dot{\mathbf{x}}_{N-2} = \mathbf{f}_{N-2}(\mathbf{x}_{N-2}, \mathbf{u}_{N-2}).$$

Eventually, iterating until one reaches the root of the clique path yields the total trajectory. However, this is only the problem of finding a trajectory given some initial condition $\mathbf{x}_0 \in \mathbf{X}$*, while our interest is in the inverse problem of fighting suitable initial conditions* $\mathbf{A}_T^{\mathbf{X}}(\mathbf{K})$ *such that the trajectory hits the target set* $\mathbf{K}$ *at time* T.

Remark 6.13 (Specificity of set approximation)
Contrary to the volume problem where our interest is only in the volume of a given set, i.e. the mass (or first moment) of a measure, in set approximation we are interested in determining the set, i.e. computing the support of a measure. The reason why classical schemes for sparse polynomial optimization (see e.g. [132, 69]) do not directly apply to volume computation is precisely that they are rather adapted to retrieving the support of a measure (i.e. determining minimizers of a polynomial), but if implemented for the volume computation problem, they do not converge to the volume of the considered set, hence the necessity for the original contribution in Chapter 5. However, again the scheme developed in Chapter 5 is specifically adapted for volume computation, which is not relevant for approximating regions of attraction. In the next sections we develop a new method, inspired from the existing ones, specifically designed for the study of dynamical systems.

With the future application to electrical power systems in mind, we focus on exploiting the network-like structure in our computations. A power network model has the particularity that not all the variables directly interact in the equations. Especially, nodes that are geographically far from each other are not connected together in the dynamics of the system. This corresponds to a sparse structure, which motivates this work.

6.2.2 A sparse infinite dimensional formulation

Following the inspiration given by both Chapter 5 and [147], we derive an LP problem that can be split into small dimensional subproblems, and thus is a lot more scalable than the dense formulation (3.7). To that end, we introduce the number of cliques $K := N - 1$ as well as the sets $\mathbb{X}'_i := \mathbb{X}_i \times \mathbb{X}_{i+1}$, $\mathbf{X}'_i := \mathbf{X}_i \times \mathbf{X}_{i+1}$, $\mathbf{K}'_i := \mathbf{K}_i \times \mathbf{K}_{i+1}$, states $\mathbf{x}'_i := (\mathbf{x}_i, \mathbf{x}_{i+1})$, maps $\mathbf{f}'_i := (\mathbf{f}_i, \mathbf{f}_{i+1})$ and dimensions $n'_i := \dim \mathbf{X}'_i = n_i + n_{i+1}$, $i \in \mathbb{N}^\star_K$.

We recall that the dense formulation for the finite time ROA outer approximation (3.7) reads:

$$
\begin{aligned}
\mathrm{p}^\star_{\mathrm{ROA}} := \sup_{\mu,\nu,\xi} \ & \int 1 \, d\mu && \text{(3.7a)} \\
\text{s.t. } & \mu \in \mathcal{M}(\mathbf{X})_+ \\
& \nu \in \mathcal{M}(\mathbf{I} \times \mathbf{X})_+ \\
& \xi \in \mathcal{M}(\mathbf{K})_+ \\
& \lambda - \mu \in \mathcal{M}(\mathbf{X})_+ \\
& \partial_t \nu + \operatorname{div}(\nu\, \mathbf{f}) = \delta_0 \, \mu - \delta_T \, \xi,
\end{aligned}
\qquad
\begin{aligned}
\mathrm{d}^\star_{\mathrm{ROA}} := \inf_{v,w} \ & \int w \, d\lambda && \text{(3.7b)} \\
\text{s.t. } & w - v(0,\cdot) - 1 \in \mathcal{C}(\mathbf{X})_+ \\
& -\partial_t v - \mathbf{f} \cdot \mathbf{grad}\, v \in \mathcal{C}(\mathbf{I} \times \mathbf{X})_+ \\
& v(T,\cdot) \in \mathcal{C}(\mathbf{K})_+ \\
& w \in \mathcal{C}(\mathbf{X})_+ \\
& v \in \mathcal{C}^1(\mathbf{I} \times \mathbf{X}).
\end{aligned}
$$

Now, as in Chapter 5 the idea is to split the measures along the cliques distribution. However, as we are not looking for a volume, we do not need any of the resulting measures to be marginals of the original ones, which gives us more slack. The main difficulty lies in the Liouville equation (3.3). Indeed, if we split our measures μ, ν, ξ into μ_i, ν_i, ξ_i, then the following happens in the Liouville equation

applied to μ_1, ν_1, ξ_1(extending derivative notation of Section 5.2.2) :

$$\partial_t \nu_1 + \mathrm{div}_{\mathbf{x}_1}(\nu_1\, \mathbf{f}_1') + \mathrm{div}_{\mathbf{x}_2}(\nu_1\, \mathbf{f}_2) = \delta_0\, \mu_1 - \delta_T\, \xi_1,$$

which makes no sense if the measures μ_1, ν_1, ξ_1 only depend on states $\mathbf{x}_1, \mathbf{x}_2$, as $\mathbf{f}_2$ depends on $\mathbf{x}_2$ *and* $\mathbf{x}_3$. Enlarging the number of states modelled in a measure is not an option, since one would always have to deal with the term $\mathrm{div}_{\mathbf{x}_j}(\nu_1\, \mathbf{f}_j)$, $\mathbf{f}_j$ depending on both $\mathbf{x}_j$ and $\mathbf{x}_{j+1}$. For this reason and to preserve sparsity, we extend ν_1 with a measure $\phi_1 \in \mathcal{M}(\mathbf{I} \times \mathbf{X}_2')_+$ and introduce a *consistency condition* (see [69]) between them:

$$\nu_1^{\langle t, \mathbf{x}_2 \rangle} = \phi_1^{\langle t, \mathbf{x}_2 \rangle}$$

so that ν_1 and ϕ_1 can be seen as the two marginals of an occupation measure, subject to a hybrid "Liouville consistency" constraint:

$$\begin{aligned}(\delta_0\, \mu_1 - \delta_T\, \xi_1 - \partial_t \nu_1)^{\langle t, \mathbf{x}_1 \rangle} &= \mathrm{div}_{\mathbf{x}_1}(\nu_1\, \mathbf{f}_1)^{\langle t, \mathbf{x}_1 \rangle} \\ (\delta_0\, \mu_1 - \delta_T\, \xi_1 - \partial_t \nu_1)^{\langle t, \mathbf{x}_2 \rangle} &= \mathrm{div}_{\mathbf{x}_2}(\phi_1\, \mathbf{f}_2)^{\langle t, \mathbf{x}_2 \rangle}.\end{aligned}$$

Then, one can write the following GMP:

Problem 19: Primal sparse ROA approximation

$$\mathrm{p}^\star_{\mathrm{spROA}} := \sup_{\substack{(\mu_i)_i, (\xi_i)_i \\ (\nu_i)_i, (\phi_i)_i}} \sum_{i=1}^{K} \int 1\, d\mu_i \tag{6.15}$$

$$\begin{aligned}
\text{s.t.} \quad & \mu_i \in \mathcal{M}(\mathbf{X}_i')_+ && i \in \mathbb{N}^\star_K \\
& \phi_i \in \mathcal{M}(\mathbf{I} \times \mathbf{X}_{i+1}')_+ && i \in \mathbb{N}^\star_{K-1} \\
& \nu_i \in \mathcal{M}(\mathbf{I} \times \mathbf{X}_i')_+ && i \in \mathbb{N}^\star_K \\
& \xi_i \in \mathcal{M}(\mathbf{K}_i')_+ && i \in \mathbb{N}^\star_K \\
& \lambda^{n_i'} - \mu_i \in \mathcal{M}(\mathbf{X}_i')_+ && i \in \mathbb{N}^\star_K \\
& (\delta_0\, \mu_i - \delta_T\, \xi_i - \partial_t \nu_i)^{\langle t, \mathbf{x}_1 \rangle} = \mathrm{div}_{\mathbf{x}_i}(\nu_i\, \mathbf{f}_i)^{\langle t, \mathbf{x}_i \rangle} && i \in \mathbb{N}^\star_{K-1} \\
& (\delta_0\, \mu_i - \delta_T\, \xi_i - \partial_t \nu_i)^{\langle t, \mathbf{x}_{i+1} \rangle} = \mathrm{div}_{\mathbf{x}_{i+1}}(\phi_i\, \mathbf{f}_{i+1})^{\langle t, \mathbf{x}_{i+1} \rangle} && i \in \mathbb{N}^\star_{K-1} \\
& \partial_t \nu_K + \mathrm{div}(\nu_K\ \mathbf{f}_K') = \delta_0\ \mu_K - \delta_T\ \xi_K && \\
& \nu_i^{\langle t, \mathbf{x}_{i+1} \rangle} = \phi_i^{\langle t, \mathbf{x}_{i+1} \rangle} && i \in \mathbb{N}^\star_{K-1}
\end{aligned}$$

whose dual writes

$$\mathrm{d}^\star_{\mathrm{spROA}} := \inf_{\substack{(u_i)_i, (w_i)_i \\ (v_{ij})_{i,j}, v_K}} \sum_{i=1}^{K} \int w_i(\mathbf{x}_i, \mathbf{x}_{i+1})\, \mathrm{d}\mathbf{x}_i\, \mathrm{d}\mathbf{x}_{i+1} \tag{6.16a}$$

$$\text{s.t. } w_i(\mathbf{x}_i, \mathbf{x}_{i+1}) - v_{i1}(0, \mathbf{x}_i) - v_{i2}(0, \mathbf{x}_{i+1}) \geq 1 \tag{6.16b}$$

$$(\mathbf{x}_i, \mathbf{x}_{i+1}) \in \mathbf{X}_i', \qquad i \in \mathbb{N}^\star_{K-1}$$

$$w_K(\mathbf{x}_K, \mathbf{x}_N) - v_K(0, \mathbf{x}_K, \mathbf{x}_N) \geq 1 \tag{6.16c}$$

$$
\begin{aligned}
&(\mathbf{x}_K, \mathbf{x}_N) \in \mathbf{X}'_K \\
&\mathbf{f}_{i+1}(\mathbf{x}_{i+1}, \mathbf{x}_{i+2}) \cdot \mathbf{grad}\, v_{i2}(t, \mathbf{x}_{i+1}) \leq -u(t, \mathbf{x}_{i+1}) && \text{(6.16d)} \\
&(t, \mathbf{x}_{i+1}, \mathbf{x}_{i+2}) \in \mathbf{I} \times \mathbf{X}'_{i+1}, \qquad i \in \mathbb{N}^\star_{K-1} \\
&\partial_t v_{i1}(t, \mathbf{x}_i) + \partial_t v_{i2}(t, \mathbf{x}_{i+1}) + \mathbf{f}_i(\mathbf{x}_i, \mathbf{x}_{i+1}) \cdot \mathbf{grad}\, v_{i1}(t, \mathbf{x}_i) \leq u(t, \mathbf{x}_{i+1}) && \text{(6.16e)} \\
&(t, \mathbf{x}_i, \mathbf{x}_{i+1}) \in \mathbf{I} \times \mathbf{X}'_i, \qquad i \in \mathbb{N}^\star_{K-1} \\
&\partial_t v_K(t, \mathbf{x}_K, \mathbf{x}_N) + \mathbf{f}'_K(\mathbf{x}_K, \mathbf{x}_N) \cdot \mathbf{grad}\, v_K(t, \mathbf{x}_K, \mathbf{x}_N) \leq 0 && \text{(6.16f)} \\
&(\mathbf{x}_K, \mathbf{x}_N) \in \mathbf{X}'_K \\
&v_{i1}(T, \mathbf{x}_i) + v_{i2}(T, \mathbf{x}_{i+1}) \geq 0 && \text{(6.16g)} \\
&(\mathbf{x}_i, \mathbf{x}_{i+1}) \in \mathbf{K}'_i \qquad i \in \mathbb{N}^\star_{K-1} \\
&v_K(T, \mathbf{x}_K, \mathbf{x}_N) \geq 0 && \text{(6.16h)} \\
&(\mathbf{x}_K, \mathbf{x}_N) \in \mathbf{X}'_K \\
&w_i \in \mathcal{C}(\mathbf{X}'_i)_+, \qquad i \in \mathbb{N}^\star_K \\
&v_{i1} \in \mathcal{C}^1(\mathbf{I} \times \mathbf{X}_i), \qquad i \in \mathbb{N}^\star_{K-1} \\
&v_{i2} \in \mathcal{C}^1(\mathbf{I} \times \mathbf{X}_{i+1}), \qquad i \in \mathbb{N}^\star_{K-1} \\
&v_K \in \mathcal{C}^1(\mathbf{I} \times \mathbf{X}'_K) \\
&u_i \in \mathcal{C}(\mathbf{I} \times \mathbf{X}_{i+1}), \qquad i \in \mathbb{N}^\star_{K-1}
\end{aligned}
$$

Here the idea is to split the decision variables v and w of problem (3.7b) and distribute them along the components of our sparse system. The decision variables u_i are added to take into account the interconnection between the components: summing (6.16d) and (6.16e) yields a regular Lyapunov-like inequality on $v_i = v_{i1} + v_{i2}$. Thus, we do not simply compute an uncertified intersection of regions of attraction of smaller subsystems, but rather a sparsely defined outer approximation of the global region of attraction. By doing so, we end up with inequality constraints involving only the variables of one of the considered subsystems at a time, which drastically reduces the dimension of the decision space in the SOS hierarchy.

Our main result is the numerical certification that can be stated as follows:

Theorem 6.17: Sparse outer ROA approximation

Let $(\mathbf{u}, \mathbf{v}, \mathbf{w})$ be feasible for problem (6.16), and consider the set

$$
\hat{\mathbf{A}}_{\mathbf{v}} := \left\{ \mathbf{x} \in \mathbb{R}^n : \begin{array}{l} \forall i \in \mathbb{N}^\star_{K-1}, v_{i1}(0, \pi_{\mathbb{X}_i}(\mathbf{x})) + v_{i2}(0, \pi_{\mathbb{X}_{i+1}}(\mathbf{x})) \geq 0 \\ v_K(0, \pi_{\mathbb{X}'_K}(\mathbf{x})) \geq 0 \end{array} \right\}. \quad \text{(6.17)}
$$

Then, one has $\mathbf{A}_T^{\mathbf{X}}(\mathbf{K}) \subset \hat{\mathbf{A}}_{\mathbf{v}}$.

Proof : Let $\mathbf{x}_0 \in \mathbf{A}_T^{\mathbf{X}}(\mathbf{K})$. Then, by definition, $\mathbf{x}(T|\mathbf{x}_0) \in \mathbf{K}$, and according to constraint (6.16g) one has for $i \in \mathbb{N}^\star_{K-1}$

$$
v_{i1}(T, \pi_{\mathbb{X}_i}(\mathbf{x}(T|\mathbf{x}_0))) + v_{i2}(T, \pi_{\mathbb{X}_{i+1}}(\mathbf{x}_{i+1}(T|\mathbf{x}_0))) \geq 0. \quad \text{(6.18)}
$$

Moreover we know that

$$\begin{aligned} v_{i1}(T, \pi_{\mathbb{X}_i}(\mathbf{x}(T|\mathbf{x}_0))) - v_{i1}(0, \pi_{\mathbb{X}_i}(\mathbf{x}_0)) &= \int_0^T \frac{\mathrm{d}}{\mathrm{d}t}(v_{i1}(t, \pi_{\mathbb{X}_i}(\mathbf{x}(t|\mathbf{x}_0)))) \, \mathrm{d}t \\ &= \int_0^T \partial_t v_{i1}(t, \pi_{\mathbb{X}_i}(\mathbf{x}(t|\mathbf{x}_0))) + \\ &\qquad \mathbf{f}_i(\pi_{\mathbb{X}_i'}(\mathbf{x})) \cdot \mathbf{grad}\, v_{i1}(t, \pi_{\mathbb{X}_i}(\mathbf{x}(t|\mathbf{x}_0))) \, \mathrm{d}t \\ &\overset{(6.16e)}{\leq} \int_0^T u_i(t, \pi_{\mathbb{X}_{i+1}}(\mathbf{x}(t|\mathbf{x}_0))) - \partial_t v_{i2}(t, \pi_{\mathbb{X}_{i+1}}(\mathbf{x}(t|\mathbf{x}_0))) \, \mathrm{d}t. \end{aligned}$$

The same reasoning on v_{i2} yields

$$\begin{aligned} v_{i2}(T, \pi_{\mathbb{X}_{i+1}}(\mathbf{x}(T|\mathbf{x}_0))) - v_{i2}(0, \pi_{\mathbb{X}_{i+1}}(\mathbf{x}_0)) &\overset{(6.16d)}{\leq} \\ \int_0^T \partial_t v_{i2}(t, \pi_{\mathbb{X}_{i+1}}(\mathbf{x}(t|\mathbf{x}_0))) - u_i(t, \pi_{\mathbb{X}_{i+1}}(\mathbf{x}(t|\mathbf{x}_0))) \, \mathrm{d}t. \end{aligned}$$

Finally, adding both inequalities, one obtains

$$\begin{aligned} 0 &\overset{(6.18)}{\leq} v_{i1}(T, \pi_{\mathbb{X}_i}(\mathbf{x}(T|\mathbf{x}_0))) + v_{i2}(T, \pi_{\mathbb{X}_{i+1}}(\mathbf{x}(T|\mathbf{x}_0))) \\ &\leq v_{i1}(0, \pi_{\mathbb{X}_i}(\mathbf{x}_0)) + v_{i2}(0, \pi_{\mathbb{X}_{i+1}}(\mathbf{x}_0)). \end{aligned}$$

Since v_K is nonnegative at time T in $\mathbf{K}_K$ in virtue of (6.16h), and decreasing along trajectories in virtue of (6.16f), the last required inequality is also satisfied. Thus, $\mathbf{x}_0 \in \hat{\mathbf{A}}_{\mathbf{v}}$. ♢

With this formulation, we design a method to compute sparse outer approximations of the ROA, using only convex semidefinite programming, while all existing methods resort only to nonconvex optimization, namely bilinear matrix inequalities. However, the constraint that the approximation should be sparse is a significant restriction that prevents us from proving convergence to the actual ROA. Indeed, with the following elementary exemple we show that in the case of sparse dynamics and target set, the ROA has no reason to be sparse.

6.2.3 Sparsity of the actual finite time ROA

Consider the simple case where $N = 3$ and the dynamics are:

$$\begin{aligned} \dot{x}_1 &= (x_1^2 + x_2^2 - 0.25)x_1 && \text{(6.19a)} \\ \dot{x}_2 &= (x_2^2 + x_3^2 - 0.25)x_2 && \text{(6.19b)} \\ \dot{x}_3 &= (x_2^2 + x_3^2 - 0.25)x_3. && \text{(6.19c)} \end{aligned}$$

Here, it is clear that the bicylinder $\mathbf{C} := \{x \in \mathbb{R}^3 : x_1^2 + x_2^2 \leq 0.25,\ x_2^2 + x_3^2 \leq 0.25\}$ is contained in the infinite time ROA of the equilibrium point $\mathbf{0}$.

However, this ROA is strictly larger than our sparsely defined $\mathbf{C}$, and it intricates all variables, which means that it cannot be sparsely described.

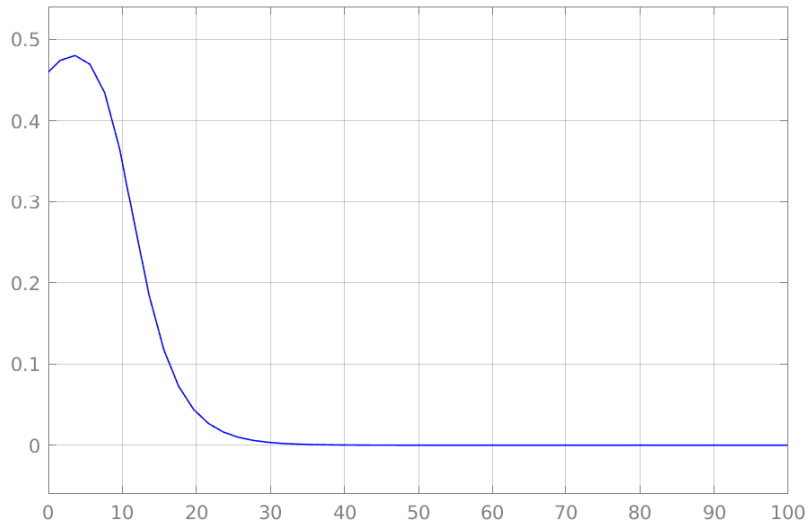


Figure 6.3 – A stable trajectory $x_1(t|\mathbf{x}_0)$.

Here the initial conditions are $x_{01} = 0.46$, $x_{02} = x_{03} = 0.25$.

To illustrate this fact, we plotted the evolution of $x_1(t|\mathbf{x}_0)$ with different initial conditions $\mathbf{x}_0 = (x_{01}, x_{02}, x_{03})$ outside $\mathbf{C}$ (see Figures 6.3 and 6.4). In the three cases, (x_{02}, x_{03}) is in the disk of radius 0.5 such that $x_2(t|\mathbf{x}_0)$ and $x_3(t|\mathbf{x}_0)$ go to 0 quickly.

However, depending on both x_{02} and x_{03}, the trajectory of $x_1(t|\mathbf{x}_0)$ is either stable (with quick convergence to 0) or unstable (with finite time explosion).

This example highlights the non-sparsity of the infinite time ROA. The same observation carries over for any finite time ROA (say for $T = 100$, $\mathbf{K} = [-0.1, 0.1]^3$) which is very close to the infinite time ROA.

From this we can deduce that exploiting sparsity prevents from directly proving the convergence of our ROA estimations towards the actual ROA, the former being sparsely defined while the latter is not. However, we can still obtain good outer approximations of the ROA using this technique. The advantages that one gains while giving up convergence are twofold :

- The computational time is drastically reduced for systems that were tractable using the converging dense framework;
- This framework allows one to handle systems that are intractable with the standard dense framework, as shown experimentally below.

Besides, two possibilities exist that might allow improving this contribution into a converging scheme:

- First, the fact that the exact ROA itself is not sparse does not mean that it cannot be approximated with sparsely defined sets; thus, looking into sparse approximation of dense sets, especially emphasizing the sparse structure of the dynamics, might be an interesting option;

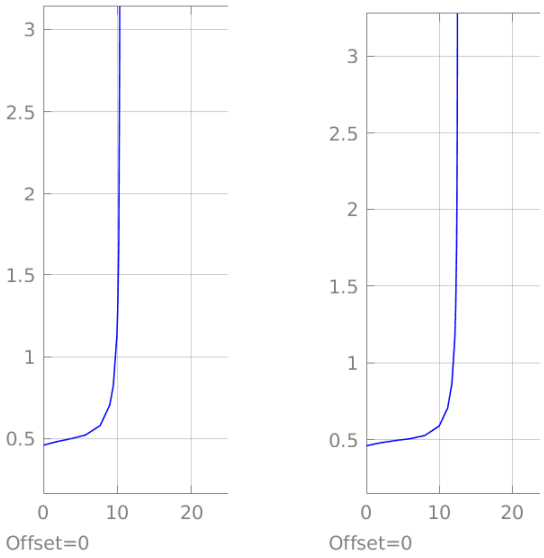


Figure 6.4 – Two unstable trajectories $x_1(t|\mathbf{x}_0)$.

Here the initial conditions are $x_{01} = 0.46$, $x_{02} = 0.26$, $x_{03} = 0.25$ (left) and $x_{01} = 0.46$, $x_{02} = 0.25$, $x_{03} = 0.3$ (right).

- Second, looking for sparse polynomials does not mean that our ROA approximations have to be correlatively sparse; for example, one could also look at the set $\tilde{\mathbf{A}}_\mathbf{v} := \{\mathbf{x} \in \mathbb{R}^n : v(\mathbf{x}) := v_K(0, \pi_{\mathbb{X}'_K}(\mathbf{x})) + \sum_{i=1}^{K-1} v_i(0, \pi_{\mathbb{X}_i}(\mathbf{x})) \geq 0\}$, whose description is not correlatively sparse as v depends on all variables; however it is obvious that $\mathbf{A}_T^\mathbf{X}(\mathbf{K}) \subset \hat{\mathbf{A}}_\mathbf{v} \subset \tilde{\mathbf{A}}_\mathbf{v}$ so that the sparse approximation is tighter; nevertheless, other non-sparse ROA approximations with sparse schemes could also be investigated.

6.2.4 Computing sparse ROA approximations

We tested our formulation (6.16) on two numerical examples: the first one is the example that we mentioned in section 6.2.3, and the second one is a dimension 20 chain constituted by 10 interconnected Van der Pol oscillators.

Reducing computational time: a toy example

To check that our sparse method is relevant, we implemented it on system (6.19) and compared its performances to those of the dense formulation of [42], with SOS polynomials of degrees 8 (Figure 6.5a) and 10 (Figure 6.5b), with state constraint set $\mathbf{X} = [-1, 1]^3$, time horizon $T = 100$ and target set $\mathbf{K} = [-0.1, 0.1]^3$.

On these figures we also plot the bicylinder (that should be inside the infinite time ROA and the finite time ROA $\mathbf{A}_T^\mathbf{X}(\mathbf{K})$ for T large enough and $\mathbf{K}$ small enough).

We gathered the computational times in Table 6.2.

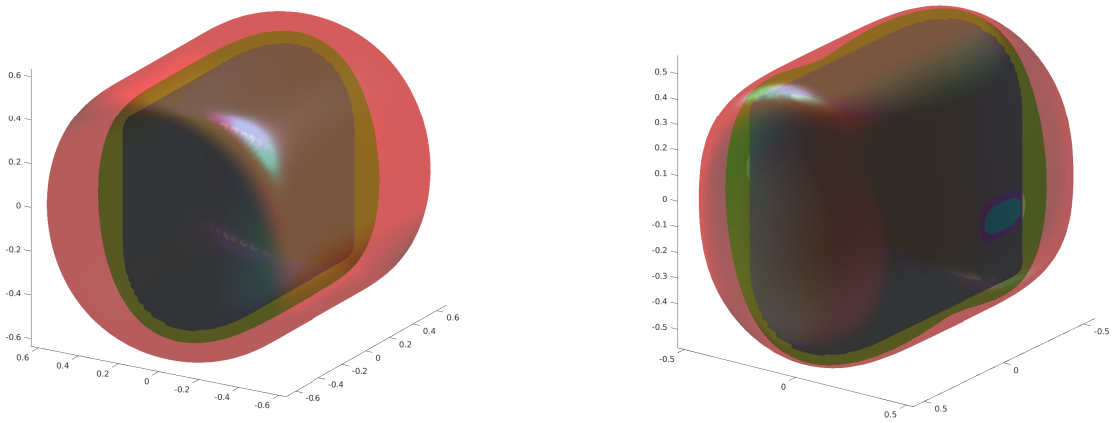
The first thing one can note is the important gain in computational time: our sparse formulation is by far less costly than the standard dense formulation, and the

degree	dense	sparse
4	4	4
6	24	10
8	334	83
10	5542	440
12	-	1865

Table 6.2 – Performances of the sparse ROA approximation scheme.

We report and compare the computation times (in seconds) for the dense and sparse formulations.

gap increases with the degree (at degree 10 the sparse formulation is more than 10 times faster).



(a) Degree 8.

(b) Degree 10.

Figure 6.5 – Comparing the sparse and dense ROA approximation schemes. We compare the sparse (red) and dense (green) ROA approximations and the bicylinder (brown).

Second, one can see that while at degree 8 the dense approximation is tighter than the sparse one, at degree 10 this does not hold anymore: the sparse approximation is actually tighter around $\mathbf{x} = 0$ (resulting in the blue-green spot on the side of the surface), and more generally both approximations are close one to another.

High dimension: a chain of Van der Pol oscillators

To test our method on large scale systems, we take the same example as in [63], but adapted to our first sparsity pattern: we consider a chain of Van der Pol oscillators linked with random couplings. The general framework is as follows:

$$\begin{aligned} \dot{y}_i &= -2z_i && \text{(6.20a)} \\ \dot{z}_j &= 0.8y_j + 10(1.2^2y_j^2 - 0.21)z_j + \epsilon_j z_{j+1}y_j && \text{(6.20b)} \\ \dot{z}_N &= 0.8y_N + 10(1.2^2y_N^2 - 0.21)z_N && \text{(6.20c)} \end{aligned}$$

with $i \in \mathbb{N}_N^\star$ and $j \in \mathbb{N}_K^\star, K = N-1$. This corresponds to our sparse polynomial ODE (6.12) with $n_i = 2$ and $\mathbf{x}_i = (y_i, z_i)$ for $i = 1, \ldots, N$ (thus $n = 2N$). One can notice that the sparse structure is even more specific than stated in our general framework since

$$\mathbf{f}_j(\mathbf{x}_j, \mathbf{x}_{j+1}) = \begin{pmatrix} f_{j1}(z_j) \\ f_{j2}(\mathbf{x}_j, z_{j+1}) \end{pmatrix} \text{ for } j = 1, \ldots, K.$$

Here ϵ_j is a random variable that follows the uniform law on $[-0.5, 0.5]$, modelling a weak interaction between the oscillators. For reporting our results, we let $N = 10$, $\mathbf{X} = [-1, 1]^{20}$, $T = 30$ and $\mathbf{K} = [-0.1, 0.1]^{20}$ and we use a particular sample ϵ. We report on degree 12 certificates, which takes approximately 23', among which 11'35" for declaring the decision variables with the YALMIP interface, 10'46" for solving the SDP problem with MOSEK and 41" for plotting the results with Matlab.

For $j = 1, \ldots, K$ we plot the sets

$$\hat{\mathbf{A}}_j := \{\mathbf{x}_j \in \mathbf{X}_j : v_{j1}(0, \mathbf{x}_j) \geq 0\}$$

which correspond to $T = 30$, $\mathbf{K}_j = [-0.1, 0.1]^2$ for the j-th Van der Pol oscillator with perturbation $\epsilon_j z_{j+1} y_j$ where z_{j+1} is a trajectory from the $(j+1)$-th Van der Pol oscillator, starting in 0 at $t = 0$, as well as

$$\hat{\mathbf{A}}_N := \{\mathbf{x}_N \in \mathbf{X}_N : v_N(0, \mathbf{x}_N) \geq 0\}$$

which corresponds to $T = 30$, $\mathbf{K}_K = [-0.1, 0.1]^2$ for the N-th (non perturbed) Van der Pol oscillator see Figure 6.6.

As expected considering the low magnitude of the interactions, on Figure 6.6 one can identify shapes similar to the ROA of a standard Van der Pol oscillator. However, the shapes are perturbed: their respective sizes differ slightly. The standard framework of [42] cannot be used here, due to the high dimension of the state space. It is also important to note that an important part of the computational time was spent for modelling the SDP problem, while the SDP solver was quite fast, once the decision variables were properly declared. We believe that these results are quite encouraging for future works on sparse ROA approximation.

Further comments

This work is a first step towards convex computation of large scale stability regions for sparse systems. Like Lyapunov-based methods, this framework gives no convergence guarantee for the polynomial approximations when the degree tends to infinity,

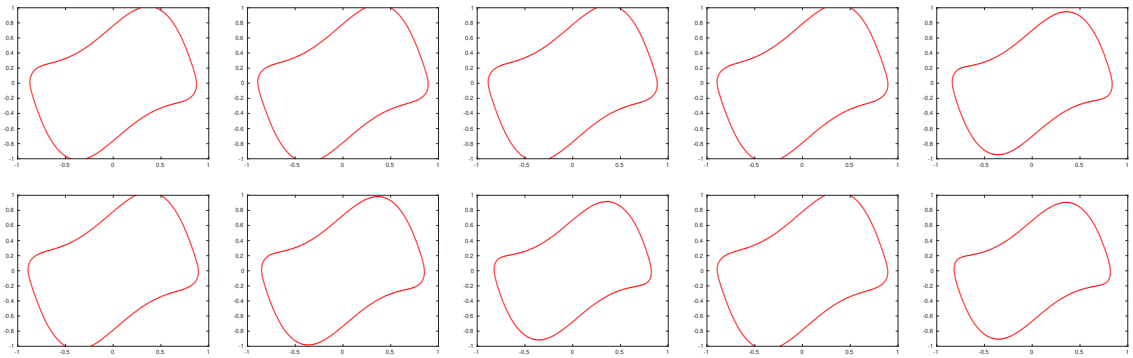


Figure 6.6 – 2D representations of finite time ROA approximations. We represented $\hat{\mathbf{A}}_1$ to $\hat{\mathbf{A}}_{10}$ (from left to right and top to bottom). These are not actual projections of the ROA approximations but can be visually compared with an actual Van der Pol oscillator ROA (due to the small magnitude of the random coupling ϵ).

due to the strong sparsity constraints imposed to the SOS certificates. However, we have been able to reduce the problem of assessing stability of a large scale sparse system into a tractable convex problem. In our opinion this is a complete novelty, since previous works resulted into nonconvex bilinear problems.

This framework is valid for any chain of coupled ODEs, and it can readily be extended to other sparsity patterns, as highlighted in [126]. The presentation of the results is however more complicated, which is the reason why we only presented chained ODEs in this chapter.

For now, we applied it only for outer approximations of the finite time ROA, while inner approximations of the finite time ROA and maximal positively invariant sets remain to be studied. Future work will also include the transient stability assessment of a meshed multi-machine system as in [6, 53], and the stability analysis of different converter grid-forming controls as in [7, 127].

The present contribution seems to outperform the direct application of [69, 132, 147] on example (6.19), although it seems very similar to them. At this point, we do not have a theoretical explanation of why this heuristic works better in practice than previously established works on sparsity. Similarly, we expect our framework to be more accurate than DSOS and SDSOS approaches such as [84] (where similar scales are tackled) as sparsity keeping full semidefinite constraints usually is.

Synthesis of sections 6.1 & 6.2

In this chapter we developed two new GMP instances for power systems stability analysis. The MPI set had already been studied in [57] where an outer approximation scheme was given. However, inner approximations required a significant improvement of the method, through a non-trivial extension of the results proposed in [83], which was the object of Section 6.1. Such a contribution yields a very interesting tradeoff between the time-independent Lyapunov formulation of Section 3.2 and the convex Lasserre hierarchy strong convergence guarantees associated to Sec-

tion 3.1. Then, an original, general heuristic was developed in Section 6.2 in order to exploit correlative sparsity patterns that may appear in the dynamics of power grids, due to their network structure. Exploiting network topology for moment-SOS ROA approximation is a new feature, since so far sparsity was mostly used for polynomial optimization and Lyapunov function computation only.

Naturally, while the scheme is proposed and illustrated for the finite time ROA outer approximation, it can as well be applied to inner ROA approximation as well as inner and outer MPI set approximation. We chose to use the finite time ROA outer approximation framework to illustrate this heuristic, as it is historically the first ever application of moment-SOS hierarchies to stability analysis. Combining the schemes of sections 6.1 and 6.2 means combining the advantages of the two methods: in particular, the size of each clique decreases by 1, as time disappears from the problem. This combination, which we postponed to future developments, might yield very interesting results.

Eventually, we want to highlight the fact that both our MPI approximation framework and our correlatively sparse scheme also have the potential to combine with other sparsity methods, such as sparse BSOS [138], term sparsity [108, 134, 135, 136] or time sparsity as it is exploited in the case of grid-forming power converters in [121]. Especially, we have an ongoing collaboration with the authors of this last promising paper, which might lead to time sparsity exploiting, inner MPI set approximation schemes, and thus allow for improved stability analyses of grid-forming power converters, combining the benefits of these approaches.

7 Conclusions and perspectives

7.1 General conclusions

Although the motivation for this thesis is the industrial need for innovative methods in the field of large scale stability analysis, it was the occasion for several contributions in various domains. Indeed, a particular focus was put on methods resorting to set approximating moment-SOS hierarchies, and more precisely on the expected computational gain that exploiting network sparsity structures can yield. This led to a variety of related problems, that can be listed in chronological order:

1. Since moment-SOS hierarchies take only polynomials as input, is it possible to apply them to AC power systems, which feature non-polynomial functions such as sine, cosine, modulus and saturations? Based on the existing work [6], the answer was yes, but a systematic proof was missing.

2. While sparsity had already been used for polynomial optimization (see [132, 69, 54]), its extension to set approximating hierarchies was nontrivial. The problem of formulating a sparse scheme for the simplest set approximating hierarchies (*i.e.* volume computation) then naturally arose.

3. A side problem to this was the question of theoretical and practical convergence of the moment-SOS hierarchies, especially for set approximation, which led us to look for systematic proofs of theoretical convergence as well as techniques to speed up practical convergence, and to investigate more in depth the connexion between volume computation and set approximation.

4. Eventually, in parallel to solving all these problem, the fundamental question of designing new moment-SOS hierarchies, allowing for a fast, tractable stability analysis in the case of high dimensional, network-like differential systems, needed to be addressed.

Each of these subproblems was then addressed within one or several chapters of this thesis, whose technical contributions are listed in what follows:

- Although **Chapter 2** was mainly dedicated to formally introducing Lasserre's moment-SOS hierarchies that are at the core of the thesis, it represented an opportunity for several original contributions regarding general features of the GMP and its numerical solution using moment-SOS hierarchies. Indeed, in

most previous works on the subject, the features of the involved hierarchies had to be studied "from scratch", while always invoking the same set of mathematical theorems. For the first time, in this Chapter we formulated general results *independent from the specific problem at hand*: Theorem 2.6 states explicit conditions for strong duality to hold in any instance of the GMP, Lemma 2.9 gives a new bound on pseudo moment sequences magnitudes, Theorem 2.16 is a finite dimensional version of Theorem 2.6 for strong duality in general moment-SOS hierarchies, and Theorem 2.17 is an "all inclusive" theorem which proposes standard conditions one can check to ensure convergence of the moment-hierarchy in terms of pseudo-moment sequences, which is one of the strongest possible convergence properties.

- In **Chapter 3** we continued to introduce key concepts for moment-SOS stability analysis, such as occupation measures, Liouville PDEs and set approximating properties, as well as mentioning classical stability notions such as local asymptotic stability and Lyapunov functions. However, these standard objects were again accompanied with some contributions w.r.t. their application: Section 3.1 studied the application of moment-SOS hierarchies to non-polynomial differential systems, through the use of changes of variables and the corresponding equality constraints. The core contribution was the observation that these equality constraints do not have to be implemented in the moment-SOS hierarchy (which would require the use of Hausdorff measures and Stokes theorem), but can instead be added *after* solving the corresponding SDP relaxations. This was the first application of moment-SOS hierarchies to the problem of power systems stability analysis. Then, Section 3.2 extended these results to a Lyapunov-based hierarchy of SOS programming problems, identifying exactly the class of dynamics that could be studied through polynomial optimization (which we named *algebraic dynamics*), and pushing standard solvers to their limits in terms of problem dimension ($n = 6$).

- **Chapter 4** can be seen as a practical follow up to Chapter 2 in the sense that, after studying theoretical convergence of general moment-SOS hierarchies, we focused on practical convergence of the volume computation hierarchies through the use of so-called Stokes constraints. Classical results from the theory of PDEs could be invoked to prove the existence of optimal solutions to both primal and dual infinite dimensional GMP instances (main Theorem 4.5), so that the dual optimal solution could be uniformly approximated with polynomials, which is the most favorable setting for hierarchical polynomial optimization, and in practice drastically improves the convergence of the volume computation hierarchy. However, we also observed that this convergence improvement mostly concerns the scalar value that approximates the actual volume of the considered set, and that it destroys the set approximation property of the volume computing problem, which is a crucial property for stability analysis (as our strategy consists of applying the set approximation property to stability sets such as regions of attraction). This is a fundamental difference between the original volume computation problem and the derived stability analysis problems.

- In **Chapter 5** we started studying the possibility of exploiting sparsity in set approximating moment-SOS hierarchies. Focusing on the fundamental volume problem, we designed two original distributed computation schemes (summed up in Theorems 5.8 and 5.13) that exploit correlative sparsity patterns to solve previously intractable high-dimensional volume computation problems. Additional graph theory assumptions were required in contrast to sparsity in polynomial optimization [69], so that a computation propagating strategy could be implemented, based on the notion of marginal measures, allowing for partial parallelization, distributed rescaling and again Stokes-based convergence acceleration. This original scheme allowed for unprecedented numerical computation of volumes in dimension 100, outperforming by far the standard moment-SOS hierarchy and even competing with non-certified randomized methods. However, we later found out that as for sparse polynomial optimization schemes, this strategy is very specific to volume computation where the target is the scalar value of the volume, approximated with marginals of measures, and does not directly extends to stability analysis, for which another sparsity exploiting strategy is required.
- Eventually, **Chapter 6** groups two original contributions to the domain of hierarchical stability analysis: Section 6.1 is a contribution on the inner approximation of maximal positively invariant sets, which can be understood as the biggest secure sets regarding security conditions formulated as polynomial constraints of type $g(\mathbf{x}) \geq 0$. Strongly relying on the results of Chapter 2, a new moment-SOS hierarchy is proposed to inner-approximate MPI sets, with convergence guarantees. Then, Section 6.2 is a first attempt to formulate a sparsity exploiting moment-SOS hierarchy for the problem of stability analysis. Here we focused on outer, finite time ROA approximation, and highlighted the fact that even sparse dynamics do not ensure sparsity of the description of the actual ROA. However, we were still able to compute quite accurate outer approximations in settings where the state space dimension exceeded by far what we used to be able to tackle with non-sparse moment-SOS hierarchies.

To conclude, in this thesis, we studied a variety of aspects of moment-SOS hierarchies, from theoretical general scheme convergence properties to pratical application to power systems, including various subtleties around original set approximating hierarchy formulations. The possible routes for next research works are detailed in the following section.

7.2 Perspectives

In this section we give our view of possible future works related to the contributions that we presented in this thesis.

7.2.1 Exploiting time sparsity

As a follow-up to the work in Section 6.2, a collaboration was starded at ETH Zürich with Florian Dörfler and Irina Subotić, on the exploitation of time sparsity, based

on [121]. The aim of this project is twofold.

- Apply moment-SOS related methods to the stability analysis of grid forming power converters, which interface renewable energy sources to global AC power grids, and using the results to compare different grid forming control strategies, such as droop control, distributed virtual oscillator control, matching control, and more. See [127] for details on these control strategies.
- Combine the heuristics developed in [121], that allow looking for sparse Lyapunov functions in the setting of nested differential systems with multiple time scale, with the sparsity exploiting methods developed in this thesis, mostly in Chapter 6, to be able to tackle accurate power converter models, which scale at a minimum of 12 state variables.

Briefly, the focus is on a system of the form:

$$\dot{\mathbf{x}}_1 = \mathbf{f}_1(\mathbf{x}_1, \mathbf{x}_2) = \mathbf{p}(\mathbf{x}_1) + \mathrm{M}(\mathbf{x}_1)(\mathrm{P}\ \mathbf{x}_1 + \mathrm{Q}\ \mathbf{x}_2) \tag{7.1a}$$

$$\dot{\mathbf{x}}_i = \mathbf{f}_i(\mathbf{x}_1, \ldots, \mathbf{x}_{i+1}) = \sum_{j=1}^{i+1} \mathrm{A}_{ij}\ \mathbf{x}_j \qquad 2 \leq i \leq N-1 \tag{7.1b}$$

$$\dot{\mathbf{x}}_N = \mathbf{f}_N(\mathbf{x}_1, \ldots, \mathbf{x}_N) = \sum_{j=1}^{N} \mathrm{A}_{Nj}\ \mathbf{x}_j \tag{7.1c}$$

where $\mathbf{x}_1 \in \mathbb{R}^{n_1}, \ldots, \mathbf{x}_N \in \mathbb{R}^{n_N}$ are ordered from slower to faster convergence towards steady-state, $\mathbf{p} \in \mathbb{R}[\mathbf{x}_1]^{n_1}$ is a polynomial vector field, $\mathrm{M} : \mathbb{R}^{n_1} \to \mathbb{R}^{n_1 \times m}$ (with $m \in \mathbb{N}^\star$) is a linear map, $\mathrm{P} \in \mathbb{R}^{m \times n_1}$, $\mathrm{Q} \in \mathbb{R}^{m \times n_3}$ and $\mathrm{A}_{ij} \in \mathbb{R}^{n_i \times n_j}$ are matrices. Intuitively, each state variable's dynamics linearly depend on all the slower variables as well as the first faster variable (except for $\mathbf{x}_1$).

In such setting, and under some additional technical assumptions, [121] proposes to look for Lyapunov functions under the form

$$V(\mathbf{x}_1, \ldots, \mathbf{x}_N) = \sum_{i=1}^{N} V_i(\mathbf{y}_i), \tag{7.2}$$

where $\mathbf{y}_i = \mathbf{x}_i - \mathbf{x}_i^s$ is the difference between $\mathbf{x}_i$ and its steady-state map $\mathbf{x}_i^s$. Thus, reproducing their developments, it is possible to adapt the moment-SOS hierarchies corresponding to finite time ROA and MPI set approximation schemes, and look for dual variables under the same form, drastically reducing the computational time, as the sparse schemes presented in this thesis allowed to do.

Theoretical developments are near to complete, and mostly the numerical experiments remain to be carried out, with challenging model complexity and parameter tunings.

7.2.2 Combining Stokes & Christoffel-Darboux

The issue with Stokes constraints presented in Chapter 4 is that while they address the Gibbs phenomenon, they also lead to losing the set approximating property,

which is crucial to deduce any stability region from the super-level set of a polynomial, as was done in Sections 3.1, 6.1 and 6.2. However, Stokes constraints consist of adding linear constraints on the involved measures, that are redundant in the GMP formulation but become active in its moment relaxations. Thus, even though our contribution focused on the dual side of Stokes constraints, strong duality in the hierarchy implies that they also have a positive impact on the convergence of the pseudo moment sequences corresponding to the SDP relaxations. Then, instead of looking for a description of the approximated set in terms of super-level set of a function from the SOS hierarchy, one could directly use methods to deduce set descriptions from moments, such as the Christoffel-Darboux polynomial method.

Definition 7.1: Christoffel Darboux (CD) polynomial

Let $\mu \in \mathcal{M}(\mathbf{X})_+$, and let $z_{\mathbf{k}} := \int \mathbf{x}^{\mathbf{k}}\, d\mu$ define its moment sequence $\mathbf{z}$. Let $d \in \mathbb{N}$ and define a basis $\mathbf{e}_d(\mathbf{x})$ of $\mathbb{R}_d[\mathbf{x}]$ as in Definition 2.8. We recall that μ then has a moment matrix

$$\mathrm{M}_{\mu,d} := \mathrm{M}_d(\mathbf{z}) = \int \mathbf{e}_d(\mathbf{x}) \cdot \mathbf{e}_d(\mathbf{x})\, d\mu(\mathbf{x})$$

which represents the bilinear functional $(p, q) \mapsto \int p\, q\, d\mu$.
Assuming that the support of μ contains a ball $\mathbf{B}$, for all $p := \mathbf{p} \cdot \mathbf{e}_d(\mathbf{x}) \in \mathbb{R}_d[\mathbf{x}]$ with $\mathbf{p} \in \mathbb{R}^{D_n^d} \setminus \{\mathbf{0}\}$, one has

$$\mathbf{p}^\top \mathrm{M}_{\mu,d}\, \mathbf{p} = \int p^2\, d\mu \geq \int_{\mathbf{B}} p^2\, d\mu > 0,$$

and $\mathrm{M}_{\mu,d}$ is positive definite. In such setting ($\mathrm{M}_{\mu,d} \succ 0$), the *Christoffel-Darboux (CD) polynomial* is defined by

$$q_{\mu,d}(\mathbf{x}) := \mathbf{e}_d(\mathbf{x})^\top \mathrm{M}_{\mu,d}^{-1}\, \mathbf{e}_d(\mathbf{x}).$$

The work in [74] shows how sublevel sets of $q_{\mu,d}$ can be used to recover the support of μ for large d, and its extension [87] gives a method to do the same when the support of μ has an empty interior. Our hope is that Stokes constraints can improve the accuracy of the CD semialgebraic set approximations, and we believe that it would be very interesting to compare the results obtained using the set approximation property highlighted in this thesis with those obtained using the Stokes-augmented, CD approximation. The question of conservativeness, which is guaranteed by the set approximation property but does not hold in general in the CD framework, should be studied with care.

7.2.3 Studying sparsity for general lift-and-project methods

As highlighted in Remark 5.5, although our sparsity exploiting schemes were designed to be applied with moment-SOS hierarchies, they are compatible with other computational methods, such as Monte-Carlo schemes. Also, moment-SOS hierarchies can be seen as lift-and-project methods, which consist in taking a difficult problem, lifting into a higher dimensional linear problem, and solving projections

of the obtained high dimensional linear problem. Indeed, in the context of power systems moment-SOS stability analysis, one is looking for a stability region for a finite dimensional differential system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, which one represents using occupation measures and the Liouville equation (*lift*), and then one solves a hierarchy of corresponding finite dimensional, SDP relaxations (*project*).

However, other lift-and-project methods exist, among which one can cite the framework of *Reproducing Kernel Hilbert Spaces* (RKHS).

Definition 7.2: Reproducing Kernel Hilbert Spaces

If $\mathcal{X}$ is a set and $\mathcal{H} \subset \mathbb{R}^{\mathcal{X}}$ is a Hilbert space of functions on $\mathcal{X}$, define the evaluation functional

$$\mathrm{L}_x : \begin{cases} \mathcal{H} & \longrightarrow & \mathbb{R} \\ f & \longmapsto & f(x) \end{cases}$$

- $\mathcal{H}$ is called a *Reproducing Kernel Hilbert Space* (RKHS) if $\forall x \in \mathcal{X}$, operator L_x is continuous: $\exists\, C_x > 0$ s.t. $|\mathrm{L}_x(f)| = |f(x)| \leq C_x \|f\|_{\mathcal{H}}$.
- Then the Riesz representation theorem ensures the existence of $\kappa_x \in \mathcal{H}$ s.t. $\mathrm{L}_x = \langle \cdot, \kappa_x \rangle_{\mathcal{H}}$, which allows us to define the *reproducing kernel*

$$\kappa : (x, y) \mapsto \langle \kappa_x, \kappa_y \rangle_{\mathcal{H}}, \text{ so that } \mathrm{L}_x = f \mapsto \langle f, \kappa(x, \cdot) \rangle_{\mathcal{H}}.$$

Remark 7.1 (Christoffel-Darboux kernel)
Using the notations of Definition 7.2.2, let

$$\kappa_{\mu,d}(\mathbf{x}, \mathbf{y}) := \mathbf{e}_d(\mathbf{x})^\top \, \mathrm{M}_{\mu,d}^{-1} \, \mathbf{e}_d(\mathbf{y})$$

(note that one then has $q_{\mu,d}(\mathbf{x}) = \kappa_{\mu,d}(\mathbf{x}, \mathbf{x})$, and the vector of coefficients of $\kappa_{\mu,d}(\mathbf{x}, \cdot)$ in the basis $\mathbf{e}_d(\mathbf{y})$ is exactly $\mathrm{M}_{\mu,d}^{-1} \mathbf{e}_d(\mathbf{x})$). Consider the set $\mathcal{X} = \mathbb{R}^n$ as well as the Hilbert space $\mathcal{H} := L^2(\mu) = \{f \in \mathbb{R}^{\mathcal{X}} : \int f^2 \, d\mu < \infty\}$ and let $p(\mathbf{y}) := \mathbf{p} \cdot \mathbf{e}_d(\mathbf{y}) \in \mathbb{R}_d[\mathbf{y}]$. One then has, by construction of $\mathrm{M}_{\mu,d}$:

$$\langle p, \kappa_{\mu,d}(\mathbf{x}, \cdot) \rangle_{\mathcal{H}} = \int p(\mathbf{y}) \, \kappa_{\mu,d}(\mathbf{x}, \mathbf{y}) \, d\mu(\mathbf{y}) = \mathbf{p}^\top \mathrm{M}_{\mu,d} \left(\mathrm{M}_{\mu,d}^{-1} \mathbf{e}_d(\mathbf{x}) \right) = \mathbf{p} \cdot \mathbf{e}_d(\mathbf{x}) = p(\mathbf{x}),$$

so that $\kappa_{\mu,d}$ is a reproducing kernel for the Hilbert space $\mathbb{R}_d[\mathbf{x}] \subset \mathcal{H}$.

Then, moment-SOS hierarchies and RKHS intersect in approximating $\mathrm{M}_{\mu,d}$. Although RKHS have their own sparsity exploiting heuristics, it could be interesting to compare both fields and see whether moment-SOS sparse schemes could be transposed to RKHS, or conversely if RKHS sparse schemes could have applications in moment-SOS hierarchies.

7.2.4 Studying Active electricity Distribution Networks

There exists different types of power grids, such as high voltage electricity transmission networks that exist at the national / continental scale and are operated by

transmission system operators (TSO) such as Réseau de Transport d'Électricité in France. At a more local scale, power supply is managed through active distribution networks (ADN) that are at the interface between the high voltage network and most power generators and consumers. Such networks have a tree structure very similar to the clique tree structures exploited in Chapter 5, which gives us the intuition that sparse schemes should be quite compatible with the study of ADNs.

Thus, from the power systems point of view, all previously mentioned theoretical options could be implemented for the study of ADNs. The idea, from the TSO viewpoint, is that having access to stability regions of ADNs makes it possible to add stability constraints at the interface between the transmission grid and the ADN, and these constraints can be taken into account by the TSO when planning the next operations, e.g. as additional constraints in the ACOPF problem.

In addition to the possibilities mentioned in the previous sections, new approaches can also be considered for ADN (as well as power systems in general) stability analysis, such as data-driven control barrier functions, that are a generalization of the dual variables v and w in (3.7). For example, the contributions found in [12] could prove to be useful; indeed, in this paper, an uncertainty set and a control policy are designed to fit an error bound on the trajectory of a linear discrete time system, and these outputs (uncertainty set and control policy) might be integrated as constraints or additional information for the transmission grid side OPF problem. Here one could aim at extending from discrete time to continuous time or from linear dynamics to polynomial dynamics if electrical powers are involved.

Back to RKHS, recent developments make possible to combine data-driven analysis and robust computation (see e.g. [82], which computes *deterministic* error bounds on the predictions provided by RKHS methods), allowing for robust kernel-based control and certified data-driven stability assessment. Here the idea would be to see whether it is possible to obtain certified (conservative) approximations of the ADN stability regions at a lower cost than the hierarchy's relaxations.

All these possible approaches pave the way for new computational methods in the field of power systems stability analysis and control, in the continuation of this thesis: with appropriate effort and investment, one of them might be tomorrow's industrial standard.

Bibliography

[1] Amir A. Ahmadi. *Algebraic Relaxations and Hardness Results in Polynomial Optimization and Lyapunov Analysis*. PhD thesis, Massachusetts Institute of Technology, 2011.

[2] Amir A. Ahmadi, Georgina Hall, Antonis Papachristodoulou, James Saunderson, and Yang Zheng. Improving efficiency and scalability of sum of squares optimization: recent advances and limitations. In *Proceedings of the* 56^{th} *IEEE Conference on Decision and Control*, 2017.

[3] Amir A. Ahmadi and Pablo A. Parrilo. Some recent directions in algebraic methods for optimization and Lyapunov analysis. In Jean-Paul Laumond, Nicolas Mansard, and Jean B. Lasserre, editors, *Geometric and Numerical Foundations of Movements. Springer Tracts in Advanced Robotics*, volume 117, pages 89–112. Springer, Cham, 2017.

[4] Aaron D. Ames, Samuel Coogan, Magnus Egerstedt, Gennaro Notomista, Sreenath Koushil, and Paulo Tabuada. Control barrier functions: theory and applications. In *Proceedings of the* 18^{th} *IEEE European Control Conference*, 2019.

[5] George J. Anders. *Probability concepts in electric power systems*. Wiley, 1989.

[6] Marian Anghel, Federico Milano, and Antonis Papachristodoulou. Algorithmic construction of Lyapunov functions for power system stability analysis. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 60(9):2533–2546, 2013.

[7] Catalin Arghir, Taouba Jouini, and Florian Dörfler. Grid-forming control for power converters based on matching of synchronous machines. *Automatica*, 95:273–282, 2018.

[8] Emil Artin. Über die Zerlegung definiter Funktionene in Qadrate. *Abhandlungen aus dem Mathematischen Seminar des Universität Hamburg*, 5:100–115, 1927.

[9] Somil Bansal, Mo Chen, Sylvia Herbert, and Clair J. Tomlin. Hamilton-jacobi reachability: A brief overview and recent advances. In *Proceedings of the* 56^{th} *IEEE Conference on Decision and Control*, 2017.

[10] Evgeny A. Barbashin and Nikolai N. Krasovsky. On the stability of motion as a whole. *Doklady Akademii Nauk SSSR*, 86:453–456, 1952.

[11] Alexander Barvinok. *A Course in Convexity*, volume 54 of *Graduate Studies in Mathematics*. American Mathematical Society, 2002.

[12] Atluğ Bitlisoğlu, Thomasz T. Gorecki, and Colin N. Jones. Robust tracking commitment. *IEEE Transactions on Automatic Control*, 62(9):4451–4466, 2017.

[13] Jean R. S. Blair and Barry Peyton. An introduction to chordal graphs and clique trees. In Alan George, John R. Gilbert, and Joseph W.H. Liu, editors, *Graph Theory and Sparse Matrix Computation*, pages 1–29. Springer, 1993. Part of the IMA Volumes in Mathematics and its Applicationd book series (IMA, volume 56).

[14] Béla Bollobás. Volume estimates and rapid mixing. In *Flavors of geometry (MSRI Publications 31)*, pages 151–180. Cambridge University Press, 1997.

[15] Haim Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Springer, 2011.

[16] Benno Büeler, Andreas Enge, and Komei Fukuda. Exact volume computation for polytopes: A practical study. In Gil Kalai and Günter M. Ziegler, editors, *Polytopes - Combinatorics and Computation (DMV Seminar:29)*, pages 131–154. Birhäuser, 2000.

[17] Claude Bélisle. Slow hit-and-run sampling. *Statistics & Probability Letters*, 47:33–43, 2000.

[18] C. Cardozo, A. Diaz Garcia, G. Giannuzzi, G. Torresan, F. Xavier, A. Cordón, L. Coronado, R. Zaottini, and C. Pisani. Small signal stability analyis of the angle difference control on a HVDC interconnection embedded in the CE synchronous power system. Preprint (ResearchGate), 2019.

[19] C.S. Chang. Online transient stability evaluation of interconnected power systems using pattern recognition strategy. *IEE Proceedings C (Generation, Transmission and Distribution)*, 140(2):115–122, 1993.

[20] Graziano Chesi. *Domain of Attraction. Analysis and Control via SOS Programming*, volume 415 of *Notes in Control and Information Sciences*. Springer, 2011.

[21] Hsiao-Dong Chiang. *Direct Methods for Stability Analysis of Electric Power Systems*. Wiley, 2011.

[22] Hsiao-Dong Chiang, Chia-Chi Chu, and G. Cauley. Direct stability analysis of electric power systems using energy functions: theory, applications, and perspective. *Proceedings of the IEEE*, 83(11):1497–1529, 1995.

[23] Diego Cifuentes and Pablo A. Parrilo. Exploiting chordal structure in polynomial ideals: A Gröbner bases approach. *SIAM Journal on Discrete Mathematics*, 30(3):1534–1570, 2016.

[24] Ben Cousins and Santosh Vempala. A practical volume algorithm. *Mathematical Programming Computation*, 8:133–160, 2016.

[25] Daniel Dadush and Santosh S. Vempala. Near-optimal deterministic algorithms for volume computation via m-ellipsoids. *Proceedings of the National Academy of Sciences*, 110(48):19237–19245, 2013.

[26] Joachim Dahl. Semidefinite optimization using MOSEK. In *Proceedings of the 21st International Symposium on Mathematical Programming*, Berlin, 2012.

[27] John P. D'Angelo and Mihai Putinar. Polynomial Optimization on Odd-Dimensional Spheres. In *Emerging Applications of Algebraic Geometry*. Springer, 2008.

[28] Matthieu Dussaule. Pour une recherche kropotkinienne. In *Propriétés asymptotiques des marches aléatoires dans les groupes relativement hyperboliques*. 2020.

[29] Martin E. Dyer and M. Frieze Alan. On the complexity of computing the volume of a polyhedron. *SIAM Journal on Computing*, 17:967–974, 1988.

[30] Martin E. Dyer, M. Frieze Alan, and Ravi Kannan. A random polynomial-time algorithm for approximating the volume of convex bodies. *Journal of the Association for Computing Machinery*, 38:1–17, 1991.

[31] Mohamed A. El-Sharkawi, Robert J. Marks, and Siri Weerasooriya. Neural networks and their application to power engineering. *Control and Dynamic Systems*, 41:359–461, 1991.

[32] G. Elekes. A geometric inequality and the complexity of measuring the volume. *Discrete and Computational Geometry*, 1:289–292, 1986.

[33] Lawrence C. Evans. *Partial Differential Equations (2nd edition)*. American Mathematical Society, 2010.

[34] Michel Fliess, Jean Levine, Philippe Martin, and Pierre Rouchon. A lie-bäcklund approach to equivalence and flatness of nonlinear systems. *IEEE Transactions on Automatic Control*, 44(5):922–937, 1999.

[35] Stephen Frank, Irina Steponavice, and Steffen Rebennack. Optimal power flow: a bibliographic survey II. *Energy Systems*, 3:259–289, 2012.

[36] Guido Fubini. *Opere scelte*, volume 2. Cremonese, 1958.

[37] G. Garcia, J. Bernussou, and M. Berbiche. Pattern recognition applied to transient stability analysis of power systems with modelling including voltage and speed regulation. *IEE Proceedings B (Electric Power Applications)*, 139(4):321–335, 1992.

[38] Francisco M. Gonzalez-Longatt, Martha N. Acosta, Harold R. Chamorro, and José Luis Rueda Torres. Power converter dominated power systems. In Francisco M. Gonzalez-Longatt and José Luis Rueda Torres, editors, *Modelling and simulation of power electronic converter dominated power systems in PowerFactory*, pages 1–35. Springer, 2020.

[39] S. Gopinath, Hassan L. Hijazi, Tillmann Weisser, Harsha Nagarajan, Mertcan Yetkin, Kaarthik Sundar, and Russel W. Bent. Proving global optimality of ACOPF solutions. Preprint; arXiv: 1910.03716 [math.OC], 2020.

[40] Trevor Hastie, Robert Tibshirani, and J.H. Friedman. *The elements of statistical learning: data mining, inference, and prediction*. Springer, 2001.

[41] Nikos D. Hatziargyriou, Jovica V. Milanović, Claudia Rahmann, Venkataramana Ajjarapu, Claudio Cañizares, István Erlich, Dustin Hill, Ian Hiskens, Innocent Kamwa, Bikash Pal, Pouyan Pourbeik, Juan J. Sanchez-Gasca, Aleksandar M. Stanković, Thierry Van Custem, Vijay Vittal, and Constantine Vournas. Definition and classification of power system stability – revisited & extended. *IEEE Transactions on Power Systems*. To appear after editing.

[42] Didier Henrion and Milan Korda. Convex computation of the region of attraction of polynomial control systems. *IEEE Transactions on Automatic Control*, 59(2):297–312, 2014.

[43] Didier Henrion, Jean B. Lasserre, and Johan Löfberg. Gloptipoly 3: moments, optimization and semidefinite programming. *Optimization Methods and Software*, 24(4-5):761–779, 2009.

[44] Didier Henrion, Jean B. Lasserre, and Carlo Savorgnan. Approximate volume and integration for basic semialgebraic sets. *SIAM Review*, 51:722–743, 2009.

[45] David Hilbert. Über die Darstellung definiter Formen als Summe von Formenquadraten. *Mathematische Annalen*, 32(3):342–350, 1888.

[46] David Hilbert. Mathematical problems. *Bulletin of the American Mathematical Society*, 8(10):437–479, 1901.

[47] John H. Hubbard and Barbara Burke Hubbard. *Vector Calculus, Linear Algebra and Differential Forms (A Unified Approach), 3rd edition*. Pearson, 2007.

[48] M. Huneault and Francisco D. Galiana. A survey of the optimal power flow literature. *IEEE Transactions on Power Systems*, 6(2):762–770, 1991.

[49] Semich Impram, Secil Varbak Nese, and Bülent Oral. Challenges of renewable energy penetration on power system flexibility: A survey. *Energy Strategy Reviews*, 31:100539, 2020.

[50] Shinsaku Izumi, Hiroki Somekawa, Xin Xin, and Taiga Yamasaki. Estimating regions of attraction of power systems by using sum of squares programming. *Electrical Engineering*, 100:2205–2216, 2018.

[51] Zachary W. Jarvis-Wloszek. *Lyapunov based analysis and controller synthesis for polynomial systems using sum-of-squares optimization*. PhD thesis, University of California, Berkeley, CA, 2003.

[52] Cédric Josz. *Application of Polynomial Optimization to Electricity Transmission Networks*. PhD thesis, Université Pierre et Maris Curie, 2016.

[53] Cédric Josz, Daniel K. Molzahn, Matteo Tacchi, and Soumayeh Sojoudi. Transient stability analysis of power systems via occupation measures. In *Proceedings of the 10th IEEE Innovative Smart Grid Technologies Conference*, 2019.

[54] Cédric Josz and Daniel K. Molzahn. Lasserre Hierarchy for Large-Scale Polynomial Optimization in Real and Complex Variables. *SIAM Journal on Optimization*, 28(2):1017–1048, 2018.

[55] I. Kamwa, S. R. Samantaray, and Geza Joos. Development of rule-based classifiers for rapid stability assessment of wide-area post-disturbance records. *IEEE Transactions on Power Systems*, 24(1):258–270, 2009.

[56] Alchim Klenke. *Probability Theory*. Springer-Verlag, 2 edition, 2014.

[57] Milan Korda, Didier Henrion, and Colin Jones. Convex computation of the maximum controlled invariant set for polynomial control systems. *SIAM Journal on Control and Optimization*, 52(5):2944–2969, 2014.

[58] Milan Korda, Didier Henrion, and Colin N. Jones. Inner approximations of the region of attraction for polynomial dynamical systems. In *Proceedings of the IFAC Symposium on Nonlinear Control Systems*, Toulouse, France, 2013.

[59] Soorya Krishna and K. R. Padiyar. Transient stability assessment using artificial neural networks. In *Proceedings of the IEEE International Conference on Industrial Technology*, pages 627–632, 2000.

[60] Jean-Louis Krivine. Anneaux préordonnés. *Journal d'Analyse Mathématique*, 12:307–326, 1964.

[61] Soumya Kundu and Marian Anghel. Computation of Linear Comparison Equations for Stability Analysis of Interconnected Systems. In *Proceedings of the 54th IEEE Conference on Decision and Control*, 2015.

[62] Soumya Kundu and Marian Anghel. Stability and Control of Power Systems using Vector Lyapunov Functions and Sum-of-Squares Methods. In *Proceedings of the European Control Conference*, 2015.

[63] Soumya Kundu and Marian Anghel. A sum-of-squares approach to the stability and control of interconnected systems using vector Lyapunov functions. In *Proceedings of the American Control Conference*, 2015.

[64] Soumya Kundu and Marian Anghel. A multiple-comparison-systems method for distributed stability analysis of large-scale nonlinear systems. *Automatica*, 78:25–33, 2017.

[65] Prabha Kundur, John Paserba, Venkat Ajjarapu, Göran Andersson, Anjan Bose, Claudio Canizares, Nikos Hatziargyriou, David Hill, Alex Stankivoc, Carson Taylor, and Thierry Van Custem. Definition and classification of power system stability. *IEEE Transactions on Power Systems*, 19(3):1387–1401, 2004.

[66] P.R.S. Kuruganty and Roy Billinton. A probabilistic assessment of transient stability. *International Journal of Electrical Power & Energy Systems*, 2(2):115–119, 1980.

[67] Joseph Pierre LaSalle. Some extensions of Liapunov's second method. *IRE Transactions on Circuit Theory*, 7:520–527, 1960.

[68] Jean B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, 11(3):796–817, 2001.

[69] Jean B. Lasserre. Convergent SDP relaxations in polynomial optimization with sparsity. *SIAM Journal on Optimization*, 17:822–843, 2006.

[70] Jean B. Lasserre. *Moments, Positive Polynomials and Their Applications*. Imperial College Press, 2010.

[71] Jean B. Lasserre. Computing Gaussian and exponential measures of semi-algebraic sets. *Advances in Applied Mathematics*, 91:137–163, 2017.

[72] Jean B. Lasserre, Didier Henrion, Christophe Prieur, and Emmanuel Trélat. Nonlinear optimal control via occupation measures and LMI-relaxations. *SIAM Journal on Control and Optimization*, 47(4):1643–1666, 2008.

[73] Jean B. Lasserre and Victor Magron. Computing the Hausdorff boundary measure of semi-algebraic sets. *SIAM Journal on Applied Algebra and Geometry*, 4(3):441–469, 2020.

[74] Jean B. Lasserre and Édouard Pauwels. The empirical Christoffel function with applications in data analysis. *Advances in Computational Mathematics*, 45:1439–1468, 2020.

[75] Javad Lavaei and Steven H. Low. Zero duality gap in optimal power flow problem. *IEEE Transactions on Power Systems*, 27(1):92–107, 2012.

[76] Richard M. Lewis and Richard B. Vinter. Relaxation of optimal control problems to equivalent convex programs. *Journal of Mathematical Analysis and Applications*, 74(2):475–493, 1980.

[77] Ernest Lindelöf. Sur l'application de la méthode des approximations successives aux équations différentielles ordinaires du premier ordre. *Comptes rendus hebdomadaires des séances de l'Académie des Sciences*, 118:454–457, 1894.

[78] J. Löfberg. YALMIP: A Toolbox for Modeling and Optimization in MATLAB. In *Proceedings of the IEEE International Symposium on Computer Aided Control System Design*, 2004.

[79] Steven H. Low. Convex relaxation of Optimal Power Flow – Part I: Formulations and equivalence. *IEEE Transactions on Control of Network Systems*, 1(1):15–27, 2014.

[80] Steven H. Low. Convex relaxation of Optimal Power Flow – Part II: Exactness. *IEEE Transactions on Control of Network Systems*, 1(2):177–189, 2014.

[81] Alexandr M. Lyapunov. *The general problem of stability of motion*. PhD thesis, University of Kharkov, Kharkov Mathematical Society, 1892.

[82] Emilio T. Maddalena, Paul Scharnhorst, and Colin N. Jones. Deterministic error bounds for kernel-based learning techniques under bounded noise. Preprint; arXiv: 2008.04005 [eess.SY], 2020.

[83] Victor Magron, Pierre-Loïc Garoche, Didier Henrion, and Xavier Thirioux. Semidefinite approximations of reachable sets for discrete-time polynomial systems. *SIAM Journal on Control and Optimization*, 57(4):2799–2820, 2019.

[84] Anirudha Majumdar, Amir A. Ahmadi, and Russ Tedrake. Control and verification of high-dimensional systems with DSOS and SDSOS programming. In *Proceedings of the* 53^{rd} *IEEE Conference on Decision and Control*, pages 394–401, 2014.

[85] Anirudha Majumdar, Georgine Hall, and Amir A. Ahmadi. Recent scalability improvements for semidefinite programming with applications in machine learning, control and robotics. *Annual Review of Control, Robotics and Autonomous Systems*, 3:331–360, 2020.

[86] Anirudha Majumdar, Ram Vasudevan, Mark M. Tobenkin, and Russ Tedrake. Convex optimization of nonlinear feedback controllers via occupation measures. *International Journal on Robotics Research*, 33(9):1209–1230, 2014.

[87] Swann Marx, Édouard Pauwels, Tillmann Weisser, Didier Henrion, and Jean B. Lasserre. Semi-algebraic approximation using Christoffel-Darboux kernel. *To appear in Constructive Approximations*, 2021. arXiv: 1904.01833 [math.OC].

[88] José Luis Massera. On Liapounoff's conditions of stability. *Annals of Mathematics*, 50(2):705–721, 1949.

[89] Dieter H. Mayer. *The Ruelle-Araki transfer operator in classical statistical mechanics*. Springer-Verlag, 1978.

[90] M. Moechtar, T.C. Cheng, and L. Hu. Transient stability of power system – a survey. In *Proceedings of IEEE WESCON'95*, 1995.

[91] Daniel K. Molzahn and Ian A. Hiskens. A survey of relaxations and approximations of the power flow equations. *Foundations and Trends® in Electric Energy Systems*, 4(1–2):1–221, 2019.

[92] Daniel K. Molzahn, Cédric Josz, Ian Hiskens, and Patrick Panciatici. Solution of optimal power flow problems using moment relaxations augmented with objective function penalization. In *Proceedings of the 54th IEEE Conference on Decision and Control*, 2015.

[93] T.S. Motzkin. The arithmetic-geometric inequality. In Oved Shisha, editor, *Inequalities*, pages 205–224. Academic Press, 1967.

[94] L. S. Moulin, A. P. A. da Silva, Mohamed A. El-Sharkawi, and Robert J. Marks. Support vector machines for transient stability analysis of large-scale power systems. *IEEE Transactions on Power Systems*, 19(2):818–825, 2004.

[95] D. Subbaram Naidu. The Hamilton-Jacobi-Bellman Equation. In *Optimal Control Systems*, pages 277–283. Taylor & Francis, 2003.

[96] Yurii Nesterov and Arkadii Nemirovskii. *Interior-Point Polynomial Algorithms in Convex Programming*. SIAM, 1994.

[97] P. K. Olulope, K. A. Folly, and S. P. Chowdhury. Transient stability assessment using artificial neural network considering fault location. In *Proceedings of the 1st IEEE International Conference on Energy, Power and Control*, 2010.

[98] Dejan R. Ostojic and G. Thomas Heydt. Transient stability assessment by pattern recognition in the frequency domain. *IEEE Transactions on Power Systems*, 6(1):231–237, 1991.

[99] Antoine Oustry, Carmen Cardozo, Patrick Panciatici, and Didier Henrion. Maximal positively invariant set determination for transient stability assessment in power systems. In *Proceedings of the 58th IEEE Conference on Decision and Control*, 2019.

[100] Antoine Oustry, Matteo Tacchi, and Didier Henrion. Inner approximations of the maximal positively invariant set for polynomial dynamical systems. *IEEE Control System Letters*, 3(3):733–738, 2019.

[101] Kartik Pandya and Satish K. Joshi. A survey of optimal power flow methods. *Journal of Theoretical and Applied Information Technology*, 4(5):450–458, 2008.

[102] Panagiotis Papadopoulos and Jovica V. Milanović. Probabilistic framework for transient stability assessment of power systems with high penetration of renewable generation. *IEEE Transactions on Power Systems*, 32(4):3078–3088, 2017.

[103] Pablo A. Parrilo. *Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization*. PhD thesis, California Institute of Technology, 2000.

[104] K.P. Persidsky. On the theory of stability of solutions of differential equations. *Uspekhi Matematicheskikh Nauk*, 1(5-6 / 15-16):250–255, 1946.

[105] Émile Picard. Sur l'application des méthodes d'approximations successives à l'étude de certaines équations différentielles ordinaires. *Journal de mathématiques pures et appliquées*, 4(9):217–272, 1893.

[106] Stephen Prajna, Antonis Papachristodoulou, and Pablo A. Parrilo. *SOSTOOLS: Sum of squares optimization toolbox for Matlab*, 2002. http://www.cds.caltech.edu/sostools.

[107] Mihai Putinar. Positive polynomials on compact semi-algebraic sets. *Indiana University Mathematics Journal*, 42(3):969–984, 1993.

[108] Bruce Reznick. Extremal PSD forms with few terms. *Duke Mathematical Journal*, 45(2):363–374, 1978.

[109] Alexandre Rocca, Marcelo Forets, Victor Magron, Eric Fanchon, and Thao Dang. Occupation measure methods for modelling and analysis of biological hybrid systems. In *Proceedings of the IFAC Conference on Analysis and Design of Hybrid Systems*, Oxford, UK, 2018.

[110] Halsey L. Royden and Patrick M. Fitzpatrick. *Real Analysis*. Pearson, fourth edition, 2010.

[111] J.E. Rubio. Generalized curves and extremal points. *SIAM Journal on Control*, 13(1):28–47, 1975.

[112] J.E. Rubio. Extremal points and optimal control theory. *Annali di Matematica Pura ed Applicata*, 109(1):165–176, 1976.

[113] David Ruelle. *Thermodynamic formalism: the mathematical structures of classical equilibrium statistical mechanics*. Addison-Wesley, 1978.

[114] Peter W. Sauer and M.A. Pai. *Power System Dynamics and Stability*. Prentice Hall, 1998.

[115] Konrad Schmüdgen. The K-moment problem for compact semi-algebraic sets. *Mathematische Annalen*, 289(1):203–206, 1991.

[116] Robert L. Smith. The hit-and-run sampler: A globally reaching Markov chain sampler for generating arbitrary multivariate distribution. In J. M. Charnes, D. J. Morrice, D. T. Brunner, and J. J Swain, editors, *Proceedings of the Winter Simulation Conference*, pages 260–264, 1996.

[117] Jyoti Sohoni and S. K. Joshi. A survey on transient stability studies for electrical power systems. In *Proceedings of the Clemson University Power Systems Conference*, 2018.

[118] Richard P. Stanley, Ian G. Macdonald, and Roger B. Nelsen. Solution of elementary problem e2701. *The American Mathematical Monthly*, 86(5):396, 1979.

[119] Gilbert Stengle. A Nullstellensatz and a Positivstellensatz in semialgebraic geometry. *Mathematische Annalen*, 207(2):87–97, 1974.

[120] Jos F. Sturm. Using Sedumi 1.02, a Matlab toolbox for optimizaton over symmetric cones. *Optimization Methods and Software*, 11(1–4):625–653, 1999.

[121] Irina Subotić, Dominic Groß, Marcello Colombino, and Florian Dörfler. A Lyapunov framework for nested dynamical systems on multiple time scales with application to converter-based power systems. *To appear in IEEE Transactions on Automatic Control*, 2021. arXiv:1911.08945 [math.OC].

[122] Matteo Tacchi. Convergence of Lasserre's hierarchy: the general case. *To appear in Optimization Letters*, 2021. arXiv: 2011.08139 [math.OC].

[123] Matteo Tacchi, Carmen Cardozo, Didier Henrion, and Jean B. Lasserre. Approximating regions of attraction of a sparse polynomial differential system. In *Proceedings of the 21st IFAC World Congress*, 2020.

[124] Matteo Tacchi, Jean B. Lasserre, and Didier Henrion. Stokes, Gibbs and volume computation of semialgebraic sets. Preprint; arXiv: 2009.12139 [math.OC], 2020.

[125] Matteo Tacchi, Bogdan Marinescu, Marian Anghel, Soumya Kundu, and Siffedine Benahmed. Power system transient stability analysis using sum-of-squares programming. In *Proceedings of the IEEE Power Systems Computation Conference*, 2018.

[126] Matteo Tacchi, Tillmann Weisser, Jean B. Lasserre, and Didier Henrion. Exploiting sparsity in semialgebraic set volume computation. *To appear in Foundations of Computational Mathematics*, 2021. arXiv: 1902.02976 [math.OC].

[127] Ali Tayyebi, Dominic Groß, Adolfo Anta, Friederich Kupzog, and Florian Dörfler. Interactions of grid-forming power converters and synchronous machines. *To appear in IEEE Transactions on Power Systems*, 2021. arXiv:1902.10750 [math.OC].

[128] Lloyd N. Trefethen. *Approximation Theory and Approximation Practice*. SIAM, 2013.

[129] Cédric Villani. *Topics in Optimal Transportation*. American Mathematical Society, 2003.

[130] Richard B. Vinter and Richard M. Lewis. The equivalence of strong and weak formulations for certain problems in optimal control. *SIAM Journal on Control and Optimization*, 16(4):546–570, 1978.

[131] V. Vittal, A.N. Michel, and A.A. Fouad. Power system transient stability analysis: formulation as nearly hamiltonian systems. *Circuits System Signal Process*, 3(1):105–122, 1984.

[132] Hayato Waki, Sunyoung Kim, Masakazu Kojima, and Masakazu Muramatsu. Sums of squares and semidefinite program relaxations for polynomial optimization problems with structured sparsity. *SIAM Journal on Optimization*, 17(1):218–242, 2006.

[133] Bo Wang, Biwu Fang, Yajun Wang, Hesen Liu, and Yilu Liu. Power system transient stability assessment based on big data and the core vector machine. *IEEE Transactions on Smart Grid*, 7(5):2561–2570, 2016.

[134] Jie Wang, Victor Magron, and Jean B. Lasserre. TSSOS: a moment-SOS hierarchy that exploits term sparsity. *SIAM Journal on Optimization*, 31(1):30–58, 2021.

[135] Jie Wang, Victor Magron, Jean B. Lasserre, and N. Hoang A. Mai. CS–TSSOS: Correlative and term sparsity for large-scale polynomial optimization. Preprint; arXiv: 2005.02828 [math.OC], 2020.

[136] Jie Wang, Victor Magron, and Martina Maggio. SparseJSR: a fast algorithm to compute joint spectral radius via sparse SOS decompositions. In *Proceedings of the Americal Control Conference*, 2021.

[137] Tillmann Weisser. *Computing Approximations and Generalized Solutions using Moments and Positive Polynomials*. PhD thesis, University of Toulouse, 2018.

[138] Tillmann Weisser, Jean B. Lasserre, and Kim-Chuan Toh. Sparse-BSOS: a bounded degree SOS hierarchy for large scale polynomial optimization with sparsity. *Mathematical Programming Computation*, 10:1–32, 2018.

[139] Tillmann Weisser, Line A. Roald, and Sidhant Misra. Chance-constrained optimization for non-linear network flow problems. Preprint; arXiv: 1803.02696, 2018.

[140] Yusheng Xue, Louis Wehenkel, Régine Belhomme, Patricia Rousseaux, Mania Pavella, Éric Euxibie, Bertrand Heilbronn, and Jean-François Lesigne. Extended equal area criterion revisited. *IEEE Transactions on Power Systems*, 7(3):1012–1022, 1992.

[141] Shengyong Ye, Xin Li, Xiaoru Wang, and Qingquan Qian. Power system transient stability assessment based on Adaboost and support vector machines. In *Proceedings of the IEEE Asia-Pacific and Energy Engineering Conference*, 2012.

[142] Dahai You, Ke Wang, Lei Ye, Junchun Wu, and Ruoyin Huang. Transient stability assessment of power system using support vector machine with generator combinatorial trajectories inputs. *International Journal of Electrical Power & Energy Systems*, 44(1):318–325, 2013.

[143] Laurence C. Young. *Lectures on the Calculus of Variations and Optimal Control Theory*. W.B. Saunders Co., 1969.

[144] James J. Q. Yu, David J. Hill, Albert Y. S. Lam, Jiatao Gu, and Victor O. K. Li. Intelligent time-adaptive transient stability assessment system. *IEEE Transactions on Power Systems*, 33(1):1049–1058, 2018.

[145] Zelda B. Zabinsky and Robert L. Smith. Hit-and-run methods. In S.I. Gass and M.C. Fu, editors, *Encyclopedia of Operations Research and Management Science*. Springer, 2013.

[146] Harry Zhang. The optimality of naive Bayes. In *Proceedings of the 17th International Florida Artificial Intelligence Research Society Conference*, 2004.

[147] Yang Zheng. *Chordal sparsity in control and optimization of large-scale systems*. PhD thesis, Oxford University, 2019.

[148] Fariba Zohrizadeh, Cédric Josz, Ming Jin, Ramtin Madani, Javad Lavaei, and Somayeh Sojoudi. A survey on conic relaxations of optimal power flow problem. *European Journal of Operational Research*, 287(2):391–409, 2020.