



HAL
open science

Convergence of Lasserre's hierarchy: the general case

Matteo Tacchi

► **To cite this version:**

Matteo Tacchi. Convergence of Lasserre's hierarchy: the general case. *Optimization Letters*, 2022, 16 (3), p. 1015-1033. 10.1007/s11590-021-01757-6 . hal-03008545v1

HAL Id: hal-03008545

<https://laas.hal.science/hal-03008545v1>

Submitted on 16 Nov 2020 (v1), last revised 4 Aug 2022 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Convergence of Lasserre's hierarchy: the general case

Matteo Tacchi^{1,2}

November 16, 2020

¹ LAAS-CNRS, 7 avenue du colonel Roche BP 54200 31031 Toulouse cedex 4, France.

² Réseau de Transport d'Electricité (RTE), 7c place du Dôme, Immeuble Window, 92073 Paris la Défense cedex, France.

✉ tacchi@laas.fr

orcid: 0000-0002-5517-5989

Abstract

Lasserre's moment-SOS hierarchy consists of approximating instances of the generalized moment problem (GMP) with moment relaxations and sums-of-squares (SOS) strengthening that boil down to convex semidefinite programming (SDP) problems. Due to the generality of the initial GMP, applications of this technology are countless, and one can cite among them the polynomial optimization problem (POP), the optimal control problem (OCP), the volume computation problem, stability sets approximation problems, and solving nonlinear partial differential equations (PDE). The solution to the original GMP is then approximated with finite truncatures of its moment sequence. For each application, proving convergence of these truncatures towards the optimal moment sequence gives valuable insight on the problem, including convergence of the relaxed values to the original GMP's optimal value. This note proposes a general proof of such convergence, regardless the problem one is faced with, under simple standard assumptions. As a byproduct of this proof, one also obtains strong duality properties both in the infinite dimensional GMP and its finite dimensional relaxations.

Keywords

Infinite dimensional optimization; semidefinite programming; Lasserre hierarchy; numerical analysis; convex optimization; duality theory.

Declarations

Funding: This work was funded by the French Company Réseau de Transport d'Électricité.

Competing interests: Not applicable.

Availability of data and material: Not applicable.

Code availability: Not applicable.

Acknowledgements

The author would like to thank Didier Henrion and Jean Bernard Lasserre for fruitful discussions.

1 Introduction

Let: $\bullet \mathbf{K} \subset \mathbb{R}^n$, $\bullet \varphi_\alpha \in \mathbb{R}[\mathbf{x}], \alpha \in \mathbb{A}$, $\bullet a_\alpha \in \mathbb{R}, \alpha \in \mathbb{A}$,
 $\bullet c \in \mathbb{R}[\mathbf{x}]$, $\bullet \psi_\beta \in \mathbb{R}[\mathbf{x}], \beta \in \mathbb{B}$, $\bullet b_\beta \in \mathbb{R}, \beta \in \mathbb{B}$,
 $\bullet \mathbb{A}, \mathbb{B}$ be index sets,

and consider the generalized moment problem (GMP), as stated in [7, (1.1)]:

$$\begin{aligned} p_{\text{GM}}^* &:= \sup \int c \, d\mu & (1) \\ \text{s.t.} \quad & \mu \in \mathcal{M}(\mathbf{K})_+ \\ & \int \varphi_\alpha \, d\mu = a_\alpha \quad \alpha \in \mathbb{A} \\ & \int \psi_\beta \, d\mu \leq b_\beta \quad \beta \in \mathbb{B}. \end{aligned}$$

Remark 1 (Generality of the problem). Problem (1) already covers the \mathbf{K} -moment problem [7, (3.1),(3.2)] ($c = 0$, $\mathbb{A} \subset \mathbb{N}^n$, $\mathbb{B} = \emptyset$, $\varphi_\alpha = \mathbf{x}^\alpha$), the polynomial optimization problem [6, (4.1)] ($\mathbb{A} = \{0\}$, $\mathbb{B} = \emptyset$, $\varphi_0 = 1$, $z_0 = 1$) and the volume computation problem [4] ($c = 1$, $\mathbb{A} = \emptyset$, $\mathbb{B} = \mathbb{R}[\mathbf{x}] \cap \mathcal{C}(\mathbf{B})_+$, $\psi_\beta = \beta$, $b_\beta = \int_{\mathbf{B}} \psi_\beta \, d\lambda$), for example.

In 2014, the authors of [5] proved that under specific conditions, the Lasserre hierarchy associated with the particular GMP instance known as the polynomial optimization problem (POP) has the strong duality property, which means that there is no duality gap between the moment relaxations and the sum-of-squares strengthenings that form the moment-SOS hierarchy.

This technical note is a follow-up that intends to close the subject of strong duality and convergence of the Lasserre hierarchy, by proving a general strong result at the abstraction level of the generic moment problem.

First, it is possible to slightly generalize Problem (1) by allowing for multiple decision variables:

$$\begin{aligned} p_{\text{GM}}^* &:= \sup \int \mathbf{c} \cdot d\boldsymbol{\mu} & (2) \\ \text{s.t.} \quad & \boldsymbol{\mu} \in \mathcal{M}(\mathbf{K}_1)_+ \times \dots \times \mathcal{M}(\mathbf{K}_N)_+ \\ & \int \boldsymbol{\varphi}_\alpha \cdot d\boldsymbol{\mu} = a_\alpha \quad \alpha \in \mathbb{A} \\ & \int \boldsymbol{\psi}_\beta \cdot d\boldsymbol{\mu} \leq b_\beta \quad \beta \in \mathbb{B}. \end{aligned}$$

where $\mathbf{c}, \boldsymbol{\varphi}_\alpha, \boldsymbol{\psi}_\beta \in \mathbb{R}[\mathbf{x}_1] \times \dots \times \mathbb{R}[\mathbf{x}_N]$ and

$$\int \mathbf{c} \cdot d\boldsymbol{\mu} := \sum_{i=1}^N \int c_i \, d\mu_i.$$

Remark 2 (A most general extension). Problem (2) covers all existing applications of the Lasserre hierarchy so far, including for instance:

- the optimal control problem [8] (with the notation of the paper, $\boldsymbol{\mu} = (\mu, \nu)$, $\mathbf{c} = (h, H)$, $\mathbb{A} = \mathbb{R}[t, \mathbf{x}]$, $\mathbb{B} = \emptyset$, $\boldsymbol{\varphi}_\alpha = (-\partial_t \alpha - \mathbf{f} \cdot \mathbf{grad} \alpha, \alpha(T, \cdot))$, $a_\alpha = \alpha(0, \mathbf{x}_0)$, where T is the terminal time),

- the controlled region of attraction problem [3] (using the reference notation, $\boldsymbol{\mu} = (\mu_0, \mu, \mu_T)$, $\mathbf{c} = (1, 0, 0)$, $\mathbb{A} = \mathbb{R}[t, \mathbf{x}]$, $\mathbb{B} = \mathbb{R}[\mathbf{x}] \cap \mathcal{C}(\mathbf{X})_+$, $\boldsymbol{\varphi}_\alpha = (\alpha(0, \cdot), \partial_t \alpha + \mathbf{f} \cdot \mathbf{grad} \alpha, -\alpha(T, \cdot))$, $a_\alpha = 0$, $\boldsymbol{\psi}_\beta = (\beta, 0, 0)$ and $b_\beta = \int \beta d\lambda$),
- the inner maximal positively invariant set problem [11] and the reachable set problem [9] (using the notations in [11], $\boldsymbol{\mu} = (\mu_0, \mu, \mu_\partial)$, $\mathbf{c} = (1, 0, 0)$, $\mathbb{A} = \mathbb{R}[\mathbf{x}]$, $\mathbb{B} = \mathbb{R}[\mathbf{x}] \cap \mathcal{C}(\mathbf{X})_+$, $\boldsymbol{\varphi}_\alpha = (\alpha, \mathbf{f} \cdot \mathbf{grad} \alpha, -\alpha)$, $a_\alpha = 0$, $\boldsymbol{\psi}_\beta = (1, 0, 0)$, $b_\beta = \int \beta d\lambda$),
- the nonlinear hyperbolic PDE problem [10, (45),(46)].

The aim of this technical note is to state and prove a most general convergence theorem for the hierarchy corresponding to Problem (2). The Lasserre hierarchy technology is most often deployed under the following standard working assumption (see [5, Assumptions 1 & 2]):

Assumption 1 (Ball constraints). Suppose that there exists $\mathbf{g}_1 \in \mathbb{R}[\mathbf{x}_1]^{m_1}, \dots, \mathbf{g}_N \in \mathbb{R}[\mathbf{x}_N]^{m_N}$ such that for $i \in \mathbb{N}_N := \{1, \dots, N\}$,

$$g_{i,1}(\mathbf{x}_i) = 1 \quad ; \quad g_{i,m_i}(\mathbf{x}_i) = 1 - |\mathbf{x}_i|^2$$

$$\mathbf{K}_i = \{\mathbf{x}_i \in \mathbb{R}^{n_i} : \mathbf{g}_i(\mathbf{x}_i) \geq \mathbf{0}\}.$$

Remark 3 (Compactness). Up to rescaling of the \mathbf{K}_i 's, it is always possible to enforce Assumption 1, as soon as they are compact basic semialgebraic sets.

Indeed let $\mathbf{g} \in \mathbb{R}[\mathbf{x}]^m$, $\mathbf{K} := \{\mathbf{x} \in \mathbb{R}^n : \mathbf{g}(\mathbf{x}) \geq \mathbf{0}\}$. For $R > 0$, define

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

where $\mathbf{B}_R := \{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x}| \leq R\}$ is the 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 to its description so that Putinar's Lemma [12, Lemma 3.2] holds for \mathbf{K}_R .

In order for the moment-SOS hierarchy to be well-defined, we need Assumption 1 to hold, as well as an assumption similar in spirit to Assumption 1:

Assumption 2 (SOS approximations). Suppose that there exists $(R_1, \dots, R_N) \in (0, +\infty)^N$, $\mathbf{h}_1 \in \mathbb{R}[\mathbf{x}_1]^{m'_1}, \dots, \mathbf{h}_N \in \mathbb{R}[\mathbf{x}_N]^{m'_N}$, such that for $i \in \mathbb{N}_N$,

$$h_{i,1}(\mathbf{x}_i) = 1 \quad ; \quad h_{i,m'_i}(\mathbf{x}_i) = R_i^2 - |\mathbf{x}_i|^2$$

$$\Sigma(\mathbf{h}_1) \times \dots \times \Sigma(\mathbf{h}_N) \text{ is dense in } \Psi_{\mathbb{B}} := \{\boldsymbol{\psi}_\beta : \beta \in \mathbb{B}\},$$

where $\Sigma[\mathbf{x}] := \{p_1^2 + \dots + p_k^2 : k \in \mathbb{N}^*, p_1, \dots, p_k \in \mathbb{R}[\mathbf{x}]\}$ is the convex cone of sums of squares of polynomials and $\Sigma(\mathbf{h}_i) := \{\boldsymbol{\sigma}_i \cdot \mathbf{h}_i : \boldsymbol{\sigma}_i \in \Sigma[\mathbf{x}_i]^{m'_i}\}$ is \mathbf{h}_i 's quadratic module.

Remark 4 (Link between assumptions 1 and 2). We say that Assumption 2 is similar in spirit to Assumption 1. To illustrate this point, we will be reasoning on the volume problem [4], for which $\Psi_{\mathbb{B}} = \mathbb{B} = \mathbb{R}[\mathbf{x}] \cap \mathcal{C}(\mathbf{B})_+$. In such setting, Assumption 2 is enforced through Putinar's Positivstellensatz [12, Theorem 1.3], assuming the existence of $\mathbf{h} \in \mathbb{R}[\mathbf{x}]^{m'}$ and $R > 0$ such that $h_1 = 1$, $h_{m'} = R^2 - |\cdot|^2$ and

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

so that one can replace $\Psi_{\mathbb{B}} = \mathbb{B}$ with $\Sigma(\mathbf{h})$ without changing the problem value.

2 Notations and theorem statement

We briefly recall the usual notations, definitions and results that are used in the Lasserre hierarchy framework. Then, we proceed to introduce the last assumptions that will be needed for the proof of our theorem.

The basic notions that are used to formulate the moment hierarchy are the Riesz functional and the moment an localizing matrices.

Definition (Riesz functional). Let $\mathbf{z} := (z_{\mathbf{k}})_{\mathbf{k} \in \mathbb{N}^n} \in \mathbb{R}^{\mathbb{N}^n}$ be a real sequence. We define, for $p(\mathbf{x}) := \sum_{|\mathbf{k}| \leq d} p_{\mathbf{k}} \mathbf{x}^{\mathbf{k}} \in \mathbb{R}[\mathbf{x}]$,

$$L_{\mathbf{z}}(p) := \sum_{|\mathbf{k}| \leq d} p_{\mathbf{k}} z_{\mathbf{k}}.$$

The linear map $L_{\mathbf{z}} : p \mapsto L_{\mathbf{z}}(p)$ is called the *Riesz functional* of \mathbf{z} .

Remark 5 (Link between $L_{\mathbf{z}}$ and integrals). If $\mathbf{z} = (z_{\mathbf{k}})_{\mathbf{k} \in \mathbb{N}^n} \in \mathbb{R}^{\mathbb{N}^n}$ and $\mu \in \mathcal{M}(\mathbf{K})_+$ are such that

$$\forall \mathbf{k} \in \mathbb{N}^n \quad z_{\mathbf{k}} = \int \mathbf{x}^{\mathbf{k}} d\mu,$$

then by definition for all $p \in \mathbb{R}[\mathbf{x}]$

$$\int p d\mu = L_{\mathbf{z}}(p).$$

Notation (Bounded multi-indices). For $n, d \in \mathbb{N}$, we use the following notations:

$$\mathbb{N}_d^n := \{\mathbf{k} \in \mathbb{N}^n : |\mathbf{k}| := k_1 + \dots + k_n \leq d\} \quad \& \quad N_{n,d} := |\mathbb{N}_d^n| = \binom{n+d}{n}$$

so that the space of polynomials in n variables with degree at most d satisfies

$$\mathbb{R}_d[\mathbf{x}] \simeq \mathbb{R}^{\mathbb{N}_d^n} \simeq \mathbb{R}^{N_{n,d}}.$$

Definition (Localizing matrix). Let $d, d_g \in \mathbb{N}$, $g \in \mathbb{R}_{d_g}[\mathbf{x}]$. Let $\mathbf{e}_d(\mathbf{x}) := (e_i(\mathbf{x}))_i$ be a base of $\mathbb{R}_d[\mathbf{x}]$. Let $\mathbf{z} = (z_{\mathbf{k}})_{|\mathbf{k}| \leq 2d+d_g} \in \mathbb{R}^{\mathbb{N}_{2d+d_g}^n}$.

The *degree d localizing matrix* $M_d(g, \mathbf{z})$ of \mathbf{z} in g is defined as the size $N_{n,d}$ matrix representation in base $\mathbf{e}_d(\mathbf{x})$ of the bilinear application

$$(p, q) \in \mathbb{R}_d[\mathbf{x}]^2 \mapsto L_{\mathbf{z}}(g p q).$$

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

$$L_{\mathbf{z}}(g p q) = \mathbf{p}^{\top} M_d(g, \mathbf{z}) \mathbf{q} \quad \& \quad L_{\mathbf{z}}(g p^2) = \mathbf{p}^{\top} M_d(g, \mathbf{z}) \mathbf{p}.$$

Notation (Positive semidefinite matrices). We denote by:

- $\mathbb{S}^n := \{M \in \mathbb{R}^{n \times n} : M^{\top} = M\}$ the space of symmetric matrices,
- $\mathbb{S}_+^n := \{M \in \mathbb{S}^n : \forall \mathbf{x} \in \mathbb{R}^n, \mathbf{x}^{\top} M \mathbf{x} \geq 0\}$ the closed convex cone of positive semidefinite matrices,
- if $M \in \mathbb{S}^n$, $M \succeq 0 \iff M \in \mathbb{S}_+^n$.

These definitions allow to formulate the moment hierarchy associated to Problem (2), under Assumption 1, thanks to Putinar's Lemma:

$$\begin{aligned}
p_{\text{GM}}^d &:= \sup \sum_{i=1}^N L_{\mathbf{z}_i}(c_i) & (3) \\
\text{s.t. } & \mathbf{z}_i \in \mathbb{R}^{\mathbb{N}_{2d}^{n_i}} & i \in \mathbb{N}_N \\
& M_{d-d_{ij}}(g_{i,j} \mathbf{z}_i) \succeq 0 & i \in \mathbb{N}_N, j \in \mathbb{N}_{m_i} \\
& \sum_{i=1}^N L_{\mathbf{z}_i}(\varphi_{\alpha,i}) = a_{\alpha} & \alpha \in \mathbb{A}_d \\
& \sum_{i=1}^N L_{\mathbf{z}_i}(\psi_{\beta,i}) \leq b_{\beta} & \beta \in \mathbb{B}_d,
\end{aligned}$$

where $d_{ij} = \lceil d^\circ g_{i,j}/2 \rceil$, and

$$\begin{aligned}
\mathbb{A}_d &:= \{\alpha \in \mathbb{A} : \forall i \in \mathbb{N}_N, d^\circ \varphi_{\alpha,i} \leq 2d\}, \\
\mathbb{B}_d &:= \{\beta \in \mathbb{B} : \forall i \in \mathbb{N}_N, d^\circ \psi_{\beta,i} \leq 2d\}
\end{aligned}$$

are finite (using linearity of the Riesz functional, Assumption 2 and finite dimensionality of $\mathbb{R}_d[\mathbf{x}] = \{p \in \mathbb{R}[\mathbf{x}] : d^\circ \leq d\}$).

Remark 6 (Minimal degree for the hierarchy). It appears in (3) that p_{GM}^d is only defined for $d \geq d_0$, where

$$d_0 := \max_{i \in \{1, \dots, N\}} \max_{j \in \{1, \dots, m_i\}} \max(\lceil d^\circ c_i/2 \rceil, d_{ij}).$$

We are finally able to formulate the last needed assumptions as well as our main theorem. First, convergence in the moment hierarchy is often obtained through an additional assumption on the mass of the involved measures:

Assumption 3 (Uniformly bounded mass). For $i \in \{1, \dots, N\}$, denote by $\mathbf{z}_i \in \mathbb{R}^{\mathbb{N}_{2d}^{n_i}}$ the pseudo-moment sequence that represents μ_i in the Lasserre hierarchy, and suppose that there exists $C_i > 0$ s.t. if \mathbf{z}_i is feasible for the degree d relaxation of problem (2), then

$$z_{i,0} \leq C.$$

Eventually, so as for the notion of "convergence of the hierarchy" to make sense, we need a unique candidate for the limit of our pseudo-moment sequences:

Assumption 4 (Existence of a unique optimal solution). Suppose that there exists a unique $\boldsymbol{\mu}^* \in \mathcal{M}(\mathbf{K}_1)_+ \times \dots \times \mathcal{M}(\mathbf{K}_N)$ feasible and optimal for problem (2):

$$\begin{cases} \forall \alpha \in \mathbb{A}, \int \varphi_{\alpha} \cdot d\boldsymbol{\mu}^* = a_{\alpha} \\ \forall \beta \in \mathbb{B}, \int \psi_{\beta} \cdot d\boldsymbol{\mu}^* \leq b_{\beta} \end{cases} \quad \& \quad \int \mathbf{c} \cdot d\boldsymbol{\mu}^* = p_{\text{GM}}^*.$$

Remark 7 (The question of uniqueness). Assumption 4 is crucial for the proof of convergence. However, if one removes it, existence of the optimal $\boldsymbol{\mu}^*$ could still be obtained from Assumption 3, through an infinite dimensional strong duality proof similar to what we display in the next section (which would bring no additional theoretical insight, so that we do not display it). The actual stake of this assumption is indeed the uniqueness assumption. Note that in most of the moment-SOS hierarchy applications, the GMP is designed in a way that enforces Assumption 4.

Our contribution consists of stating and proving a general theorem for convergence of the Lasserre hierarchy (3) corresponding to problem (2):

Theorem 5 (Convergence of the general Lasserre hierarchy). *Under Assumptions 1, 2, 3 and 4, there exist optimal sequences $(\mathbf{z}_i^d)_{d \geq d_0}$ of feasible pseudo-moment sequences for the moment hierarchy s.t. $\sum_{i=1}^N L_{z_i^d}(c_i) = p_{\text{GM}}^d$ and for all $\mathbf{k} \in \mathbb{N}^n$,*

$$z_{i,\mathbf{k}}^d \xrightarrow{d \rightarrow \infty} \int \mathbf{x}^{\mathbf{k}} d\mu_i^*(\mathbf{x}).$$

In particular, one has $p_{\text{GM}}^d \xrightarrow{d \rightarrow \infty} p_{\text{GM}}^*$.

Moreover, this automatically yields strong duality $p_{\text{GM}}^d = d_{\text{GM}}^d \ \& \ p_{\text{GM}}^* = d_{\text{GM}}^*$, where d_{GM}^* and d_{GM}^d are the values of the duals of (2) and (3), respectively, so that $d_{\text{GM}}^d \xrightarrow{d \rightarrow \infty} d_{\text{GM}}^*$.

3 Proof of Theorem 5

For the sake of simplicity and without loss of generality, we work on the simple case (1) where $N = 1$, with corresponding hierarchy:

$$\begin{aligned} p_{\text{GM}}^d &:= \sup L_{\mathbf{z}}(c) & (4) \\ \text{s.t. } \mathbf{z} &\in \mathbb{R}^{\mathbb{N}_{2d}^n} \\ M_{d-d_j}(g_j \mathbf{z}) &\succeq 0 & j \in \mathbb{N}_m \\ L_{\mathbf{z}}(\varphi_\alpha) &= a_\alpha & \alpha \in \mathbb{A}_d \\ L_{\mathbf{z}}(\psi_\beta) &\leq b_\beta & \beta \in \mathbb{B}_d. \end{aligned}$$

All the proofs are the same for $N > 1$, at the price of additional notations that do not bring any theoretical insight, so that we stick to the case $N = 1$. We first prove an easy lemma that gives all its importance to Assumption 1:

Lemma 6 (Pseudo moment sequences boundedness). *Let $d \in \mathbb{N}^*$, $R > 0$, $\mathbf{z} \in \mathbb{R}^{\mathbb{N}^n}$ s.t. $M_d(\mathbf{z}) \succeq 0$ & $M_{d-1}((R^2 - |\cdot|^2) \mathbf{z}) \succeq 0$. Then,*

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

Proof. $M_d(\mathbf{z}) \succeq 0$ is equivalent to

$$\forall p \in \mathbb{R}_d[\mathbf{x}], L_{\mathbf{z}}(p^2) \geq 0, \quad (a)$$

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

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

(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 $R^2 z_{2\mathbf{k}} \geq \sum_{|\mathbf{k}'|=1} z_{2(\mathbf{k}+\mathbf{k}')}^2$ with $z_{2(\mathbf{k}+\mathbf{k}')} \geq 0$ by (a), so that $R^4 z_{\mathbf{0}} \geq R^2 z_{2\mathbf{k}} \geq z_{2(\mathbf{k}+\mathbf{k}')}$ 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

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

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 $0 \leq L_{\mathbf{z}}(p^2) = z_{2\mathbf{k}} \pm 2z_{\mathbf{k}+\mathbf{k}'} + z_{2\mathbf{k}'}$ so that

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

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

This lemma proves several important facts, among which any nonzero feasible pseudo-moment vector \mathbf{z}_i for (3) satisfies $z_{i,\mathbf{0}} > 0$ if Assumption 1 holds, and under Assumptions 1 and 3, for all $d \in \mathbb{N}$, \mathbf{z}_i feasible for (3) satisfies $|z_{i,\mathbf{k}}| \leq C_i \forall \mathbf{k}$.

A first consequence of Lemma 6 is the strong duality property in the hierarchy.

Proposition 7 (Strong duality in the hierarchy). *Suppose that each relaxation of Problem (1) has a feasible solution. In that case, under Assumptions 1, 2 and 3,*

$$\forall d \geq d_0, \quad p_{\text{GM}}^d = d_{\text{GM}}^d.$$

Moreover, the degree d relaxation (4) of (1) has an optimal solution $\mathbf{z}_d \in \mathbb{R}^{\mathbb{N}_{2d}^n}$.

Proof. Assumption 2 is only needed for the hierarchy to be well-defined. We rely on [1, Theorem (7.2), Lemma (7.3)] to prove our result. Consider the cone

$$\mathcal{K}_d := \{(A_d \mathbf{z}, L_{\mathbf{z}}(c)) : \mathbf{z} \in \mathcal{X}_d\},$$

where $A_d \mathbf{z} := (L_{\mathbf{z}}(\varphi_\alpha), L_{\mathbf{z}}(\psi_\beta))_{\alpha \in A_d, \beta \in B_d}$ and

$$\mathcal{X}_d := \left\{ \mathbf{z} \in \mathbb{R}^{\mathbb{N}_{2d}^n} : \forall j \in \{1, \dots, m\}, M_{d-d_j}(g_j \cdot \mathbf{z}) \succeq 0 \right\}.$$

According to [1, Theorem (7.2)] a sufficient condition for our result to hold requires that $p_{\text{GM}}^d < \infty$ and \mathcal{K}_d is closed. Clearly,

$$p_{\text{GM}}^d \leq N_{n,d_0} \left(\max_{\mathbf{k} \in \mathbb{N}^n} c_{\mathbf{k}} \right) \left(\max_{\mathbf{z}} \max_{\mathbf{k} \in \mathbb{N}_{2d}^n} z_{\mathbf{k}} \right) \leq N_{n,d_0} \left(\max_{\mathbf{k} \in \mathbb{N}^n} c_{\mathbf{k}} \right) C < \infty.$$

Besides, [1, Lemma (7.3)] states that for \mathcal{K}_d to be closed, it is sufficient to prove that \mathcal{X}_d has a compact, convex base, and that

$$\forall \mathbf{z} \in \mathcal{X}_d, \quad (A_d \mathbf{z}, L_{\mathbf{z}}(c)) = (\mathbf{0}, 0) \implies \mathbf{z} = \mathbf{0}. \quad (*)$$

We first exhibit a compact convex base for \mathcal{X}_d . Let

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

\mathcal{P}_d is a base of \mathcal{X}_d in the sense that $\mathcal{X}_d \setminus \{\mathbf{0}\}$ is isomorphic to $(0, +\infty) \times \mathcal{P}_d$ 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 6 and Assumption 1, for any $M_d(\mathbf{z})$ and $M_{d-1}((1 - |\cdot|^2) \mathbf{z})$ 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 $\mathbf{p}_1, \mathbf{p}_2 \in \mathcal{P}_d$, $t \in [0, 1]$, $\tilde{\mathbf{p}} := t \mathbf{p}_1 + (1-t) \mathbf{p}_2$. Then, by linearity of the localizing matrix operator,

$$\forall g \in \mathbb{R}_d[\mathbf{x}] \quad M_{d-d_g}(g \tilde{\mathbf{p}}) = t M_{d-d_g}(g \mathbf{p}_1) + (1-t) M_{d-d_g}(g \mathbf{p}_2),$$

where $d_g = \lceil d^\circ g/2 \rceil$, so that its semidefinite positivity is preserved by convex combination, by convexity of \mathbb{S}_+^n for any $n \in \mathbb{N}$. Thus, $\tilde{\mathbf{p}} \in \mathcal{X}_d$. Besides,

$$\tilde{p}_0 = t p_{10} + (1-t) p_{20} = t + 1 - t = 1$$

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

Eventually, we move on to showing compactness of \mathcal{P}_d . According to Lemma 6, the ball constraint in the description of \mathbf{K} (Assumption 1) and the upper bound $z_0 \leq C$ (Assumption 3) yield boundedness of \mathcal{P}_d for the distance $\text{dist}(\mathbf{z}, \mathbf{z}') := \max_{|\mathbf{k}| \leq 2d} |z_{\mathbf{k}} - z'_{\mathbf{k}}|$.

Finally, \mathcal{P}_d is closed as the intersection between the level-1 set of the continuous function $\mathbf{z} \mapsto z_0$ and the closed cone \mathcal{X}_d . Indeed \mathcal{X}_d is closed as the pre-image of the closed cone $(\mathbb{S}_+^{N_{n,d}})^m$ by the (continuous) linear map

$$\mathbf{z} \mapsto (\mathbb{M}_{d-d_j}(g_j \mathbf{z}))_{j \in \{1, \dots, m\}}.$$

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

It remains to prove (*). Let $\mathbf{z} \in \mathcal{X}_d$ s.t. $\mathbb{A}_d \mathbf{z} = \mathbf{0}$ and $\mathbb{L}_{\mathbf{z}}(c) = 0$. We want to prove that $\mathbf{z} = \mathbf{0}$ so that (*) holds.

Let $\mathbf{z}_0 \in \mathcal{X}_d$ s.t. $\forall \alpha \in \mathbb{A}_d, \beta \in \mathbb{B}_d, \mathbb{L}_{\mathbf{z}_0}(\varphi_\alpha) = a_\alpha$ & $\mathbb{L}_{\mathbf{z}_0}(\psi_\beta) \leq b_\beta$. Define for $t \geq 0$ $\mathbf{z}_t := \mathbf{z}_0 + t \mathbf{z}$. Let $t \geq 0$.

Since \mathcal{X}_d is a convex cone, $\mathbf{z}_t \in \mathcal{X}_d$. In addition, by construction of \mathbf{z} and linearity of the operator $\mathbf{z} \mapsto \mathbb{L}_{\mathbf{z}}(\cdot)$, $\forall \alpha \in \mathbb{A}_d, \beta \in \mathbb{B}_d, \mathbb{L}_{\mathbf{z}_t}(\varphi_\alpha) = \mathbb{L}_{\mathbf{z}_0}(\varphi_\alpha) = a_\alpha$ & $\mathbb{L}_{\mathbf{z}_t}(\psi_\beta) = \mathbb{L}_{\mathbf{z}_0}(\psi_\beta) \leq b_\beta$, so that Assumption 3 ensures that $z_{t,0} \leq C$. However, $z_{t,0} = z_{0,0} + t z_0$.

Combined with the fact that $z_{0,0} \geq 0$ (with Lemma 6), this yields that for all $t \geq 0$,

$$t z_0 \leq C,$$

which is only possible if $z_0 = 0$, *i.e.* if $\mathbf{z} = \mathbf{0}$ using again Lemma 6. \square

Eventually, we can end this work by proving Theorem 5. As announced at the beginning of this section, w.l.o.g we actually only do the proof for the case $N = 1$:

Theorem 8 (Convergence of the pseudo-moment sequences). *Under Assumptions 1, 2, 3 and 4, there exists an optimal sequence $(\mathbf{z}_d)_{d \geq d_0}$ of feasible pseudo-moment sequences for the hierarchy (4) corresponding to Problem (1), s.t. $\mathbb{L}_{\mathbf{z}_d}(c) = p_{\text{GM}}^d$ and for all $\mathbf{k} \in \mathbb{N}^n$,*

$$z_{d,\mathbf{k}} \xrightarrow{d \rightarrow \infty} \int \mathbf{x}^{\mathbf{k}} d\mu^*(\mathbf{x}).$$

In particular, one has $p_{\text{GM}}^d \xrightarrow{d \rightarrow \infty} p_{\text{GM}}^$.*

Moreover, this automatically yields strong duality $p_{\text{GM}}^d = d_{\text{GM}}^d \stackrel{\mathcal{E}}{=} p_{\text{GM}}^ = d_{\text{GM}}^*$ and then dual convergence $d_{\text{GM}}^d \xrightarrow{d \rightarrow \infty} d_{\text{GM}}^*$.*

Proof. Existence of $(\mathbf{z}_d)_{d \geq d_0}$ follows from Proposition 7 (using Assumptions 1, 2 and 3), so we focus on the proof of convergence. Let $d \in \mathbb{N}$. For $\mathbf{k} \in \mathbb{N}^n$, define

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

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

$$\|\hat{\mathbf{z}}_d\|_{\ell^\infty(\mathbb{N}^n)} := \max_{\mathbf{k} \in \mathbb{N}^n} |\hat{z}_{d,\mathbf{k}}| = \max_{|\mathbf{k}| \leq 2d} |z_{d,\mathbf{k}}| \stackrel{\text{Lemma 6}}{\leq} z_{d,0} \leq C,$$

using again Assumptions 1 and 3. 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 $\ell^1(\mathbb{N}^n) := \left\{ \mathbf{u} \in \mathbb{R}^{\mathbb{N}^n} : \sum_{\mathbf{k} \in \mathbb{N}^n} |u_{\mathbf{k}}| < \infty \right\}$ has $\ell^\infty(\mathbb{N}^n)$ as its topological dual. Thus, the Banach-Alaoglu theorem [2, Theorem 3.16] yields a weak-* converging subsequence $(\hat{\mathbf{z}}_{d_r})_{r \in \mathbb{N}}: \exists \mathbf{z}_\infty \in \ell^\infty(\mathbb{N}^n); \forall \mathbf{u} \in \ell^1(\mathbb{N}^n),$

$$\sum_{\mathbf{k} \in \mathbb{N}^n} u_{\mathbf{k}} z_{d_r, \mathbf{k}} \xrightarrow{r \rightarrow \infty} \sum_{\mathbf{k} \in \mathbb{N}^n} u_{\mathbf{k}} z_{\infty, \mathbf{k}}.$$

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

Let $\alpha \in \mathbb{A}, \beta \in \mathbb{B}, j \in \mathbb{N}_m$. Then, for all $d \geq d_0$ and $r \in \mathbb{N}$ big enough, by feasibility of \mathbf{z}_{d_r} for the relaxation of degree d_r , one has

$$\begin{aligned} \bullet \quad 0 &\preceq M_{d-d_j}(g_j \mathbf{z}_{d_r}) = M_{d-d_j}(g_j \hat{\mathbf{z}}_{d_r}) \xrightarrow{r \rightarrow \infty} M_{d-d_j}(g_j \mathbf{z}_\infty) \\ \bullet \quad a_\alpha &= L_{\mathbf{z}_{d_r}}(\varphi_\alpha) = L_{\hat{\mathbf{z}}_{d_r}}(\varphi_\alpha) \xrightarrow{r \rightarrow \infty} L_{\mathbf{z}_\infty}(\varphi_\alpha) \\ \bullet \quad b_\beta &\succeq L_{\mathbf{z}_{d_r}}(\psi_\beta) = L_{\hat{\mathbf{z}}_{d_r}}(\psi_\beta) \xrightarrow{r \rightarrow \infty} L_{\mathbf{z}_\infty}(\psi_\beta) \end{aligned}$$

so that according to Putinar's Lemma, \mathbf{z}_∞ is the actual moment sequence of a measure μ_∞ that is feasible for problem (1). Then, one directly has

$$p_{\text{GM}}^* \geq \int c d\mu_\infty = L_{z_\infty}(c) = \lim_{r \rightarrow \infty} L_{\hat{\mathbf{z}}_{d_r}}(c) = \lim_{r \rightarrow \infty} p_{\text{GM}}^{d_r} \geq p_{\text{GM}}^*$$

since for any $d \geq d_0$ $p_{\text{GM}}^d \geq p_{\text{GM}}^*$ (by construction of the moment *relaxation*).

Hence, $\int c d\mu_\infty = p_{\text{GM}}^*$, *i.e.* μ_∞ is *optimal* for problem (1). By Assumption 4, this yields $\mu_\infty = \mu^*$, *i.e.* $\mathbf{z}_\infty = \mathbf{z}^*$. Thus, $(\hat{\mathbf{z}}_d)_d$ has a unique weak-* accumulation point \mathbf{z}^* , which means that for any $\mathbf{k} \in \mathbb{N}^n$,

$$z_{d, \mathbf{k}} \xrightarrow{d \rightarrow \infty} z_{\mathbf{k}}^* = \int \mathbf{x}^{\mathbf{k}} d\mu^*(\mathbf{x}),$$

Q.E.D.

Eventually, since Assumptions 1 and 3 hold, Proposition 7 ensures strong duality $p_{\text{GM}}^d = d_{\text{GM}}^d$, so that putting together weak GMP duality and the strenghtening property, one has

$$p_{\text{GM}}^* \stackrel{\text{weak duality}}{\leq} d_{\text{GM}}^* \stackrel{\text{strenghtening}}{\leq} d_{\text{GM}}^d \stackrel{\text{strong duality}}{=} p_{\text{GM}}^d \stackrel{\text{convergence}}{\xrightarrow{d \rightarrow \infty}} p_{\text{GM}}^*$$

and the sandwich rule yields strong GMP duality $p_{\text{GM}}^* = d_{\text{GM}}^*$. \square

4 Conclusion

While usually it is only proved that the optimal values of the GMP's moment relaxations monotonically converge towards the optimal value of the GMP, this note proves that the solutions of these moment relaxations converge to the actual solution of the GMP in ℓ^∞ 's weak-* topology, provided that some elementary assumptions hold. So far, such proof has only been provided for very specific instances of the GMP (see e.g. [4]), while here the generic case is completely dealt with. Up to rescaling the problem (which also precludes ill behaviours in the numerical implementations), one can usually enforce quite easily that these conditions are met. As a byproduct

of this proof, one also obtains strong duality both in the GMP and its corresponding hierarchy, generalizing the results of [5]. Among the open questions this note answers, one can cite the weak-* convergence of the moment sequences associated to the OCP [8] and set approximation problems in [3, 11, 9], as well as the strong duality between the GMP formulation in [10] and its dual, regardless of the formulation of such dual. Finally, this note should facilitate most proofs of convergence and strong duality related to future applications of Lasserre’s moment-SOS hierarchy.

References

- [1] Alexander Barvinok. A Course in Convexity. Number 54. American Mathematical Society, Providence, Rhode Island, 2002.
- [2] Haim Brezis. Functional Analysis, Sobolev Spaces and Partial Differential Equations. 2011.
- [3] 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.
- [4] Didier Henrion, Jean Bernard Lasserre, and Carlo Savorgnan. Approximate volume and integration for basic semialgebraic sets. Siam Review, 51:722–743, 2009.
- [5] Cédric Josz and Didier Henrion. Strong duality in lasserre’s hierarchy for polynomial optimization. Optimization Letters, 10:3–10, 2016.
- [6] Jean Bernard Lasserre. Global optimization with polynomials and the problem of moments. SIAM Journal on Optimization, 11:796–817, 2001.
- [7] Jean Bernard Lasserre. Moments, Positive Polynomials and Their Applications. 2010.
- [8] Jean-Bernard 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.
- [9] 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, 2019.
- [10] Swann Marx, Tillmann Weisser, Didier Henrion, and Jean Bernard Lasserre. A moment approach for entropy solutions to nonlinear hyperbolic partial differential equations. Mathematical Control and Related Fields, 10:113–140, 2020.
- [11] Antoine Oustry, Matteo Tacchi, and Didier Henrion. Inner approximations of the maximal positively invariant set for polynomial dynamical systems. IEEE Control System Letters, 2019.
- [12] Mihai Putinar. Positive polynomials on compact semi-algebraic sets. Indiana University Mathematics Journal, 42, 1993.