



HAL
open science

Learning Feasibility Constraints for Multi-contact Locomotion of Legged Robots

Justin Carpentier, Rohan Budhiraja, Nicolas Mansard

► **To cite this version:**

Justin Carpentier, Rohan Budhiraja, Nicolas Mansard. Learning Feasibility Constraints for Multi-contact Locomotion of Legged Robots. Robotics: Science and Systems, Jul 2017, Cambridge, MA, United States. 9p. hal-01526200

HAL Id: hal-01526200

<https://laas.hal.science/hal-01526200v1>

Submitted on 22 May 2017

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.



Distributed under a Creative Commons Attribution - ShareAlike 4.0 International License

Learning Feasibility Constraints for Multi-contact Locomotion of Legged Robots

Justin Carpentier, Rohan Budhiraja and Nicolas Mansard

Laboratoire d'Analyse et d'Architecture des Systèmes

Université de Toulouse

7 avenue du Colonel Roche

Toulouse, FRANCE

Email: justin.carpentier@laas.fr

Abstract—Relying on reduced models is nowadays a standard cunning to tackle the computational complexity of multi-contact locomotion. To be really effective, reduced models must respect some feasibility constraints in regards to the full model. However, such kind of constraints are either partially considered or just neglected inside the existing reduced problem formulation. This work presents a systematic approach to incorporate feasibility constraints inside trajectory optimization problems. In particular, we show how to learn the kinematic feasibility of the centre of mass to be achievable by the whole-body model. We validate the proposed method in the context of multi-contact locomotion: we perform two stairs climbing experiments on two humanoid robots, namely the HRP-2 robot and the new TALOS platform.

I. INTRODUCTION

Multi-contact locomotion of legged robots inside heterogeneous non-flat environments is a hard and challenging problem. One major difficulty of this problem lies in its computational complexity: finding the whole-body trajectory solving a locomotion task corresponds to a high dimensional mathematical problem with a complex structure, hardly tractable by modern computers and particularly the ones embedded in modern legged robots [27, 17].

1) *Reduced models*: To tackle those limitations, various strategies have been proposed in the literature. Most of them are based upon using reduced models: instead of working with the full dynamics, only a subpart is considered, covering the essential properties of the whole dynamics. In the context of bipedal locomotion, the most famous reduced model is the linear inverted pendulum model (LIPM) [15].

The locomotion is then reduced to the problem of finding a trajectory for the reduced model which will then drive the whole-body system. Starting with [16], various optimal control formulations have been proposed by the community, to either tackle the robustness problem [30], include viability conditions [25], allow altitude variations of the center of mass (CoM) [2], or also include foot placements as parameters of the problem [11].

However, LIPM-based methods are restricted to basic environments (flat ground, no obstacles) and cannot deal with more complex scenarios as non-coplanar contact cases, climbing stairs using handrail, etc. Considering non-coplanar contacts breaks the nice linearization leading to the LIPM model. A first approach to handle the non-linear dynamics

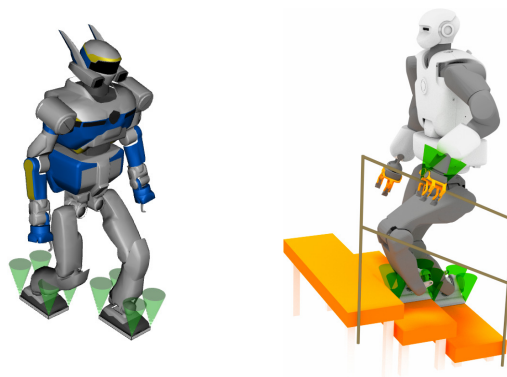


Fig. 1. Illustration of HRP-2 robot and TALOS robot making contacts with their environment. The green “ice-cream” cones are dispatched on the 4 vertices of the feet, symbolizing the friction cones with friction coefficient of value 0.3.

was proposed in [14], however requiring technical and dedicated developments based on limiting assumption (e.g. prior knowledge of force distribution). In quite another vein, it has been proposed to simplify the whole-body optimization problem by e.g. assuming unconstrained torque capabilities [6]. Both approaches indeed boil down to optimizing the so-called centroidal dynamics [20] as reduced model. Direct resolution of the underlying optimal control problem based on multiple-shooting approach has been recently proposed [4], leading to real-time performances. Other contributions have also been suggested that exhibit approximate dynamics (with possibly bounded approximations) leading to convex optimization problems, thus ensuring global optimality [13, 5, 2]. In most cases, the footstep sequence is assumed given, although some solvers are also able to discover it while optimizing the centroidal dynamics [19, 7], to the price of heavier computational costs.

2) *Feasibility constraints*: The reduced model (either LIPM or centroidal) is subject to feasibility constraints implied by the whole body (e.g. kinematic or torque limits, footstep length). Such constraints are difficult to express as solely function of the reduced model. For instance, the CoM trajectory must be achievable by the whole-body kinematics. These constraints can be tackled explicitly, by adding the corresponding whole-body variables in the optimization scheme [19, 6].

However, this direct representation is also the most expensive in terms of computation.

Such constraints can also be represented at the level of the reduced model by using so-called proxy constraints [32]. In most of previous works, proxy constraints are defined by some rough approximations (box constraints, elliptic bounds, etc) leading to a certain conservatism; or it is simply ignored inside the reduced problem formulation. Footstep limits have been encoded by hyper-plane based on a dataset of robot success and failure inside a dynamic simulator [23]. Similar constraints can be obtained by training a training a neural network [21]. In [32], similar bounds are obtained by trial and errors based on stability analysis of the whole-body system.

Constraining only the CoM position is not sufficient. It is also necessary to consider the constraints related to the contact forces [29] which must lie inside so-called friction cones, the capacity of robot to generate such value of angular momentum, etc. The main problem lies in the fact that it is hard to find analytic formulas to represent and express those constraints.

3) *Contribution*: In this paper, we propose a systematic approach to handle feasibility constraints in the context of trajectory optimization for reduced models, leading to efficient resolution on the real robot. The resulting constraint formulation could be employed in most of the optimal control solver based on centroidal dynamics [19, 7, 13], although we implemented it inside a multiple-shooting solver [4].

We first recall the equations of motion for both the whole-body and the centroidal models and define the resulting optimal control problem (OCP) for multi-contact locomotion in Section II. These definitions logically introduce the need of proxy constraint. Our main contribution, described in Section III, is to represent proxy constraints by occupancy measures, whose corresponding cost of transport is optimized in the OCP. In Section IV, we propose a complete solution to learn the CoM feasibility constraint by off-line sampling the robot motion capabilities. Finally, the effectiveness of the approach is highlighted with two real experiments on the HRP-2 robot climbing stairs with or without using handrail and one in simulation with the TALOS humanoid robot [26] climbing stairs using handrail, reported in Section V.

II. MULTI-CONTACT LOCOMOTION

This section reviews the centroidal dynamics and its links with the whole-body model. We then describe the generic optimal control problem used to compute the centroidal trajectories, based on [4]. Although this section contains mostly known materials, we believe that the clean formulation of the dynamics decoupling is a contribution per se. It is indeed a prerequisite to the introduction of proxy constraints in the next section.

A. Contact model

The interaction between a robot and the environment is defined through a set of contact points $\{\mathbf{p}_k \in \mathbb{R}^3, k = 1 \dots K\}$. For instance, for a humanoid robot equipped with rectangular feet, the contact points correspond to the four vertices of the

rectangular shape. At each contact point \mathbf{p}_k is defined a contact force \mathbf{f}_k . In the case of unilateral contacts, \mathbf{f}_k must lie inside a 3-dimensional quadratic friction cone \mathcal{K}_k^3 , characterized by a positive friction coefficient μ_k . Fig. 1 depicts humanoid robots making contact with the environment.

We only consider here rigid contact interaction (contacting bodies are fixed) which is a reasonable assumption for modern legged robots which are mostly equipped with rigid soles.

A *contact phase* is defined by a constant set of contact points. In the context of bipedal walking, two examples of contact phases are the single and double support phases. As soon as a creation or a rupture of contact point occurs, the contact set is modified, defining a new contact phase. The concatenation of contact phases describes what we name a *contact sequence*, inside which all the contact phases have their own duration.

Computing automatically the contact sequence is a difficult problem [3], but efficient contact planners now exist to compute it in a short amount of time [10, 28].

B. Whole-body dynamics and centroidal dynamics

A legged robot is by nature a free-floating base system composed of $6 + n$ degrees of freedom (DoF). Its dynamics is governed by the Lagrange equation of motion, which links the joint configuration \mathbf{q} and its time derivatives $\dot{\mathbf{q}}, \ddot{\mathbf{q}}$ to the torque actuation $\boldsymbol{\tau}_a$ and the contact forces \mathbf{f}_k :

$$\begin{bmatrix} H_u \\ H_a \end{bmatrix} \ddot{\mathbf{q}} + \begin{bmatrix} \mathbf{b}_u \\ \mathbf{b}_a \end{bmatrix} = \begin{bmatrix} \mathbf{g}_u \\ \mathbf{g}_a \end{bmatrix} + \begin{bmatrix} \mathbf{0}_6 \\ \boldsymbol{\tau}_a \end{bmatrix} + \sum_{k=1}^K \begin{bmatrix} J_{k,u}^\top \\ J_{k,a}^\top \end{bmatrix} \mathbf{f}_k \quad (1)$$

where subscripts u and a stands for the under-actuated and actuated parts respectively, H is the generalized mass matrix, \mathbf{b} covers centrifugal and Coriolis effects, \mathbf{g} is generalized gravity vector and J_k is the Jacobian of the contact k .

The 6 first rows of (1) represent the under-actuated dynamics of the robot, also called the centroidal dynamics [20]. It coincides with the Newton-Euler equations of motion which links the variation of the whole-system linear and angular momentum to the contact forces. Denoting by \mathbf{L}_c the angular momentum, the centroidal dynamics is:

$$m\dot{\mathbf{c}} = \sum_k \mathbf{f}_k + m\mathbf{g}$$

$$\dot{\mathbf{L}}_c = \sum_k (\mathbf{p}_k - \mathbf{c}) \times \mathbf{f}_k$$

where $\mathbf{g} \stackrel{\text{def}}{=} (0, 0, -9.81)$ is the gravity vector and the operator \times denotes the cross product.

The n last rows of (1) are the classic Lagrange dynamics of a fixed robot manipulator making contacts, which is both linear in the torque and force inputs.

C. Partial decoupling of centroidal and manipulator dynamics

Under the assumption of sufficient torque, it is possible to split up the problem of whole-body locomotion into two successive (hierarchical) stages: a) find the contact force trajectories driving the the centroidal dynamics; b) recover the required joint torque trajectory from the centroidal trajectory and the end-effector trajectories using for instance a second

order inverse kinematics/dynamics solver. This assumption is reasonable given motorization of current legged robots, that are sufficient to generate high torques. We discuss the resulting limits in the conclusion.

To ensure the feasibility of the first stage with respect to the second stage, some conditions must be met by the first stage:

- i) the forces must remain inside the friction cones in case of unilateral contacts;
- ii) the centroidal dynamics must be kinematically and dynamically feasible by the system;

The first constraint stems directly from the contact model introduced in Sec. II-A. The second constraint comes from the fact that the centroidal dynamics is linked to the joint configuration and its derivatives through the centroidal mapping:

$$\begin{bmatrix} \mathbf{h} \\ \mathbf{L}_c \end{bmatrix} = A_g(\mathbf{q}) \dot{\mathbf{q}}, \quad (2)$$

with A_g the so-called centroidal momentum matrix [20].

D. Centroidal optimal control formulation

From the previous assumptions, the problem of multi-contact locomotion can be reduced to the following formulation:

From a given contact sequence and an initial centroidal state, find a feasible centroidal trajectory, satisfying the Newton-Euler dynamics, respecting the contact constraints and leading to a viable state.

This formulation can be directly transcribed as an optimal control problem of the form:

$$\min_{\underline{\mathbf{x}}, \underline{\mathbf{u}}} \sum_{s=1}^S \int_{t_s}^{t_s + \Delta t_s} \ell_s(\mathbf{x}, \mathbf{u}) dt \quad (3a)$$

$$\text{s.t. } \forall t \quad \dot{\mathbf{x}} = f(\mathbf{x}, \mathbf{u}) \quad (3b)$$

$$\forall t \quad \mathbf{u} \in \mathcal{K} \quad (3c)$$

$$\forall t \quad \exists (\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}}) \text{ s.t. } \mathbf{x}, \dot{\mathbf{x}} \text{ is feasible} \quad (3d)$$

$$\mathbf{x}(0) = (\mathbf{c}_0, \mathbf{0}, \mathbf{0}) \quad (3e)$$

$$\mathbf{x}(T) \text{ is viable} \quad (3f)$$

where s is the index of the contact phase, $\underline{\mathbf{x}}$ and $\underline{\mathbf{u}}$ are the state and control trajectories, t_s is the start time of the contact phase s with $t_1 = 0$ and $t_{s+1} = t_s + \Delta t_s$. Constraints (3b) and (3c) enforce consistent dynamics with respect to the contact model. Eq. (3d) is the constraint enforcing the feasibility of the centroidal dynamics with respect to the whole-body problem: it handles kinematics limits, bounds on the angular momentum quantity, etc. We will show how it can be replaced by an automatic proxy constraint. Constraint (3e) constrains the trajectory to start with a given state (typically estimated by the sensor of the real robot). Terminal constraint (3f) is difficult to exactly represent [31] and is replaced in practice by zero terminal movement $\ddot{\mathbf{c}}(T) = \dot{\mathbf{L}}_c(T) = 0$ and $\mathbf{x}(T) = (\mathbf{c}^*, \mathbf{0}, \mathbf{0})$. Finally, ℓ_s is the cost function which enforces the smoothness of both the state and control trajectories. Various

ℓ_s can be discussed and implemented. The one used in the experiments requires some additional definitions and is given in Section IV.

III. FEASIBILITY OF THE CENTROIDAL PROBLEM

This section reports the main contribution of the paper. We present a mathematical coding of the feasibility constraints as probability measures. We then discuss the interest of this representation with respect to more-classical set-membership and show how it can be used to efficiently implement (3d) in the OCP. This section introduces the abstract definitions, that next section uses to build the complete implementation.

A. Mathematical representation of feasibility constraints

Our objective is to efficiently implement the feasibility constraint (3d) in our OCP. This constraint explicitly depends on the robot configuration, which is not a variable of the centroidal OCP. A straight-forward implementation is to add the robot configuration in the variables of the OCP [6]. However, this would surely lead the OCP to optimize the whole-body trajectory in order to handle all the robot constraints, which is yet not tractable especially if targeting real-time performances. We rather believe that it is possible to represent this constraint by an equivalent ‘‘proxy’’ constraint not dependent on the robot configuration.

Various ways to encode proxy constraints have been proposed in the literature. Most of them rely on set-membership. Denoting by γ the centroidal projection function:

$$\gamma : (\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}}) \rightarrow (\mathbf{x}, \dot{\mathbf{x}}) = \gamma(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}})$$

the proxy can be written as the constraint to have the state variables in the range space of γ . Set-membership proxies are used for instance in [11, 7] to encode maximal step size in biped walking, or in [5] to bound the CoM position by simple geometric shape. In all these cases, the set boundaries are represented by very simple mathematical structures (typically linear inequalities) in order not to burden the OCP solver. Remarkably, there are few papers about the automatic synthesis of the set boundaries [23, 21, 32].

Despite its popularity, the set-membership representation has important drawbacks. First, it is often difficult to handle by the OCP solver, in particular when the feasible set is not convex. The boundary, which is a singular mathematical object, is also complex to describe or numerically approximate. Finally, the OCP solver often tends to saturate the set boundary, where the inverse kinematics γ^{-1} is likely to fail. Consequently, the set is often arbitrarily reduced to improve the robustness of the whole-body solution.

B. Proxy as occupation measure

In this paper, we rather state that the proxy is best implemented by the occupation measure over $\mathbf{x}, \dot{\mathbf{x}}$.

Consider a state trajectory $\underline{\mathbf{x}}$. With (3d), we want to maximize the likelihood that the inverse-kinematics solver converges on a trajectory $\underline{\mathbf{q}}$ such that $\underline{\mathbf{x}}$ is the image of

\mathbf{q} by γ . For that purpose, it is desirable that any state \mathbf{x} corresponds to as many robot configurations as possible, so that the inverse kinematics is likely to converge to a solution \mathbf{q} meeting continuity constraints.

We define the occupation measure as the image of the uniform distribution in configuration space through the centroidal projection γ :

$$\mu_o(\tilde{\mathbf{x}}) \stackrel{\text{def}}{=} \int_{\tilde{\mathbf{q}} \text{ s.t. } \gamma(\tilde{\mathbf{q}})=\tilde{\mathbf{x}}} d\tilde{\mathbf{q}} = \int_{\mathcal{Q}} \mathbb{1}_{\gamma(\tilde{\mathbf{q}})=\tilde{\mathbf{x}}} d\mu_{\mathcal{Q}}$$

where $\tilde{\mathbf{x}} \stackrel{\text{def}}{=} (\mathbf{x}, \dot{\mathbf{x}})$, $\tilde{\mathbf{q}} \stackrel{\text{def}}{=} (\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}})$, \mathcal{Q} is the whole-body motion range, $\mathbb{1}_a$ is the indicator function (i.e. 1 when the assertion a is true, 0 otherwise) and $\mu_{\mathcal{Q}}$ is the uniform distribution on \mathcal{Q} .

Measure μ_o has several properties of the set-membership representation. First, its support is the feasibility set, which means that μ_o contains at least as much information as the set boundaries. It indeed contains more information, as for example the level sets of μ_o can be used as boundaries of the inner of the feasibility set, used to improve the robustness.

In practice, it is desirable that OCP (3) promotes centroidal states $\tilde{\mathbf{x}}$ where μ_o is the highest. First, it makes it easier to then compute a corresponding configuration $\tilde{\mathbf{q}}$. Second, the configuration is well inside the kinematic feasibility set, where redundancy will help the robot to handle disturbances. We will see in the experimental results that the resulting whole-body trajectory corresponds to configurations with large manipulability.

Finally, the measure also eases the life of the OCP solver, compared to handling directly the feasibility set membership, as explained next.

C. Maximizing the occupation measure

Before deriving an effective solution to represent μ_o for the specific case of the kinematic feasibility, we quickly show how μ_o can be integrated in the OCP (3).

In practice, the measure can be normalized and represented by the corresponding probability density function (PDF), denoted by $p(\mathbf{x}, \dot{\mathbf{x}})$. It is then possible to directly exploit the measure to represent the set-membership constraint (by imposing the integral of the measure to be positive on any small neighborhood around the trajectory). In addition, we could use the PDF to directly optimize the robustness, either by optimizing over a level set of the PDF, or by maximizing the neighborhood around the trajectory where the measure is nonzero.

However, adding a PDF as a constraint of an OCP is not straightforward. Therefore, we propose to remove the hard constraint (3d) and penalize the OCP cost with the log PDF.

In practice, the logarithm prevents the solver from selecting non-feasible \mathbf{x} states. Constraints (3d) is always satisfied. It also penalizes non-robust behavior where no redundancy \mathbf{q} is available, and avoids saturation of the hard constraint. Finally, the OCP solver is gently pushed away from the constraint, instead of searching for a solution living on the boundaries, which greatly improves its efficiency. Furthermore, it is unlikely

that the OCP solver is trapped in local minima of μ_o , as it manipulates a full trajectory $\underline{\mathbf{x}}$ and not a single state \mathbf{x} . Experimentally, we observed that our OCP solver robustly computes a good local minimum when optimizing over a cost penalizing the log-PDF, while it is unlikely to converge to a solution when optimizing over set-membership.

IV. LEARNING THE CoM REACHABILITY PROXY

We now present a complete solution to efficiently approximate the CoM feasibility, i.e. for any time t , there exists a joint configuration $\mathbf{q}(t)$ such that (i) the contact placements are respected and (ii) the CoM of the poly-articulated system matches $\mathbf{c}(t)$. Handling this sole constraint first is a proper way of validating our approach. It is also interesting in practice, as the feasibility of the CoM is the most limiting constraint for humanoid robots. Generalization to velocity and acceleration of the CoM with respect to joint velocity and acceleration limits would be straight-forward. Extension to the construction of the proxy on the torque limits is left as a perspective.

A. Probabilistic model

The geometric condition can be stated as the *conditional probability* of the CoM to be at the position \mathbf{c} given the current set of K contact points $\{\mathbf{p}_k \in \mathbb{R}^3, k = 1 \dots K\}$. This probability is denoted by $p(\mathbf{c}|\mathbf{p}_k, k = 1 \dots K)$. It lives in the high dimensionality domain $\mathbb{R}^{3(K+1)}$ and it is hard to compute in general.

The probability domain can be exactly reduced by gathering together the contact points belonging to the same rigid end-effector (e.g., the 4 vertices of the humanoid foot belongs to the same end-effector). We denote by $M_i = (R_i, \mathbf{p}_i) \in SE(3)$ the placement (position and orientation) of the contact body i . The conditional probability is then reduced to $p(\mathbf{c}|M_i, i = 1 \dots K_c)$ where K_c is the number of end-effectors in contact.

We now assume that variables M_i are all independent. This assumption is clearly abusive, however is a reasonable approximation under knowledge of \mathbf{c} . It is later discussed. Under this assumption, the conditional probability reads:

$$p(\mathbf{c}|M_i, i = 1 \dots K_c) \propto \prod_{i=1}^{K_c} p_i(\mathbf{c}) \quad (4)$$

where $p_i(\mathbf{c})$ stands for $p(\mathbf{c}|M_i)$ and \propto stands for ‘‘is proportional to’’. $p_i(\mathbf{c})$ is nothing more than the probability distribution of the CoM to be at position \mathbf{c} w.r.t the frame defined by M_i .

The assumption of independence of the M_i is commonly employed inside the machine-learning community as a trick to make the problem numerically tractable. In this particular case, it greatly simplifies the learning process: instead of working in a high dimensional space, the problem is restricted to a subset of \mathbb{R}^3 . In addition, the independence of end-effector placements plays the role of an upper-bound for the real probability: if a CoM is not feasible for at least one of the

end-effectors (i.e. one of the $p_i(\mathbf{c})$ is equal to 0), then the joint probability is also zero. The converse is not true.

B. Kernel density estimation by CoM sampling

There is in general no closed form to encode $p_i(\mathbf{c})$ for a particular legged robot. Nevertheless, this conditional probability can be easily approximated by extensive sampling of the CoM position expressed in the end-effector frames.

Sampling N_{samples} of the CoM position expressed in the frame M_i does not raise particular difficulties. For each sample, a configuration q_a of the actuated joints is randomly sampled and the CoM position is computed (expressed in placement frame) by forward kinematics. The sample is rejected if joint limits or self collision are violated.

The probability distribution can be approximation from the cloud of CoM points by the kernel density estimators (KDE) [22]. KDE are in some sense the analogues of histograms but for continuous domains: for each point of the data set, it associates one kernel centered on the point and all kernels share the same parameters. In the present work, we use isotropic Gaussian kernel.

C. Reduction of dimension

One drawback of the KDE representation is its computational complexity: evaluating the exponential function contained in the Gaussian kernel takes around 10 ns on modern CPU. So, roughly speaking, evaluating the PDF of the KDE takes approximately $10 \times N_{\text{samples}}$ ns which becomes rapidly a bottleneck when the number of points is huge (N_{samples} greater than 100 points).

We propose to then approximate the KDE by a Gaussian mixture model (GMM) [1]. GMMs are particularly suited to approximate a PDF with only few Gaussians in the mixture. The GMMs are learned for each end-effector from the corresponding cloud of samples by means of the expectation-maximization (EM) algorithm [8].

The quality of the GMM approximation can be estimated using the Kullback-Leibler (KL) divergence between the KDE (ground-truth) and the learned GMM (approximation) using the Monte Carlo estimator proposed in [12]. Depending on the number of Gaussians in the mixture, the divergence can reveal under or over fitting effects. The optimal number of Gaussians is easily selected for each end effector by dichotomy, as exemplified in next section.

D. Summary of the learning procedure

In summary, for each end effector, N_{samples} configurations are sampled and the corresponding CoM is computed in the end-effector frame. The resulting KDE is approximated by fitting a GMM using EM. Finally, the probability of CoM occupancy is approximated as the product of $p_i(\mathbf{c})$, for i the end effectors in contact with the environment.

E. Proposed optimal control formulation

We can now express the complete formulation of the cost function ℓ_s .

$$\ell_s(\mathbf{x}, \mathbf{u}) = w_x \|\dot{\mathbf{x}}\|^2 - \sum_{i=1}^{K_c} \log(p_i(\mathbf{c})) \quad (5)$$

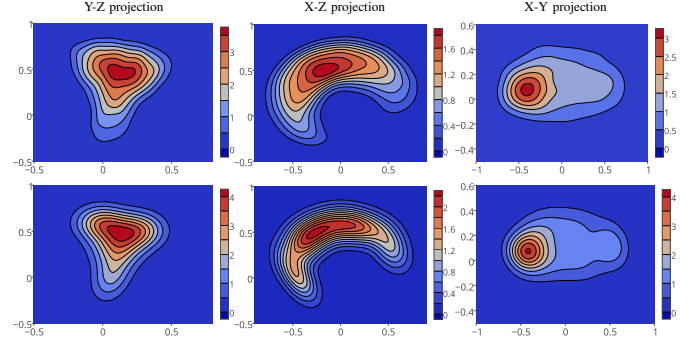


Fig. 2. Illustration of the probability density distribution of the CoM w.r.t. the right foot frame of HRP-2. The PDF are projected along the three axis X,Y,Z and represented by the means of colormap: the low values are closed to the blue colour while the high values tend to be more red. The first row corresponds to the ground truth distribution estimated through KDE. The KDE is composed of 20000 points. The second row is the colour map of the GMM used in the OCP and composed of 7 Gaussian kernels.

where the first term¹ enforces a smooth trajectory, the second term is the cost of transport for the approximate CoM occupancy measure, and w_x weights their relative importance. The first term can similarly be interpreted as a weak formulation of the occupancy measures for the second order terms ($\dot{\mathbf{c}}$ and \mathbf{L}_c) and their derivatives, through centred Gaussian measures (i.e. no prior on occupancy distribution). If the complete occupancy measure μ_o is available, the first term would become useless.

V. RESULTS

We first illustrate the learning procedure exposed in Sec. IV on the HRP-2 robot. We then present the complete pipeline we used to achieve both real and simulated motions on the two different robots. After that, we show two real experiments of multi-contact locomotion with the HRP-2 robot inside an environment similar to what can be found in the industry. Finally, we end this section by showing the versatility of the approach with the simulation of a multi-contact motion on another humanoid robot. We refer the reader to the accompanying video for a better viewing of the multi-contact experiments².

A. Illustration of the learning procedure

We validate the proposed learning approach on the HRP-2 robot. For that purpose, we only expose for space reasons the learning of the accessibility space of the CoM w.r.t. the right foot (RF). A similar study can be conducted on the three other end-effectors.

The learning process is made from a set of 20000 points sampled uniformly in the configuration space. The KDE of this set is represented on the first row of Fig. 2. The first observation is that the PDF of the RF is not convex and follows a kind of banana distribution on the X-Z plane. In other words, this means that the distribution cannot be approximated by a

¹ $\|\dot{\mathbf{x}}\|$ is a function of \mathbf{x} and \mathbf{u} through $\dot{\mathbf{x}} = f(\mathbf{x}, \mathbf{u})$

²available at <https://youtu.be/7hiLf6DpMAA>

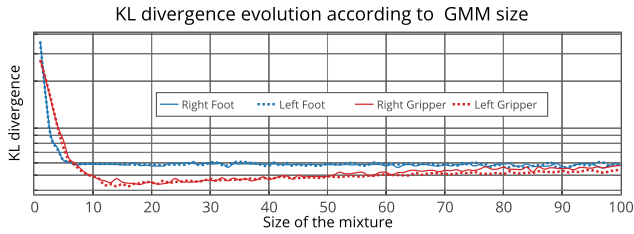


Fig. 3. Evolution of the KL divergence between the KDE distribution and GMMs of different sizes for the four end-effectors of the HRP-2 robot.

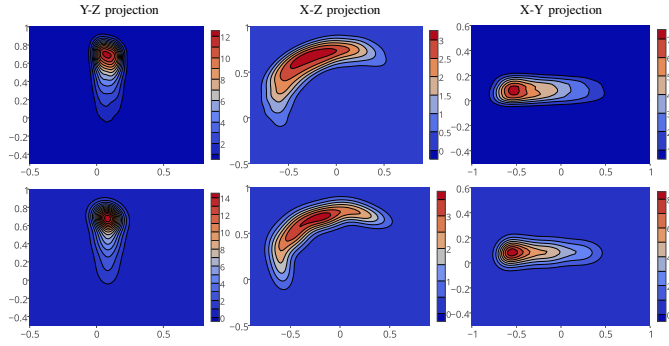


Fig. 4. Illustration of the probability density distribution of the CoM w.r.t. the right foot frame of TALOS robot. The PDF are projected along the three axis X,Y,Z and represented by the means of colormap: the low values are closed to the blue colour while the high values tend to be more red. The first row corresponds to the ground truth distribution estimated through KDE. The KDE is composed of 20000 points. The second row is the colour map of the GMM used in the OCP and composed of 4 Gaussian kernels. The axes have the same scale than in Fig 2.

single normal distribution but must be composed of several distributions. The second row of Fig. 2 represents the colour map of the GMM used inside the OCP. At this stage, it is important to notice that the approximation with GMMs does not fit perfectly the maximal values of the real distribution. However, this approximation is conservative with respect to the support and the level sets of the original distribution.

Fig. 3 highlights the experimental procedure suggested in Sec. IV-C and shows the evolution of the KL-divergence with respect to the size of the GMMs. For the right and left feet, the KL-divergence stagnates from 7 kernels in the mixture. In other words, it is sufficient to takes a GMM of size 7 to represent the CoM distribution in the foot frames. For the right and left grippers, it is a little bit different. The KL-divergence first decreases and then increases from 14 kernels. This behaviour can be explained by the fact that the EM algorithm does not optimize the KL divergence but the likelihood of observation (expectation). We chose to represent the CoM distribution w.r.t. the grippers with a GMM of size 14.

A similar study has been done on the TALOS humanoid robot, which is larger than HRP-2 and as different leg and arm kinematics. The distributions for the right foot of TALOS is depicted in Fig. 4.

B. Complete pipeline

Our locomotion framework is composed of three stages:

a) *Contact sequence planning*: For all the experiments, the contact sequences are manually designed with fixed durations for each contact phase depending on the scenario. We also manually design the end-effector trajectories by using splines with zero acceleration and velocity at take off and landing instants. This information could have been automatically generated using a contact planner [28].

b) *Centroidal resolution*: From the contact sequence and the learned CoM feasibility constraints, we solve the optimal control formulation (3) with the tailored cost function (5). For all the experiments and robots, we use the same weighting in the cost function: $w_x = 10$. This weight is a trade-off between the smoothness of the dynamics trajectory and the enforcement of the feasibility constraint. We instantiate the OCP (3) inside the MUSCOD-II framework [18] which implements an efficient multiple-shooting algorithm [9] particularly suited for multi-staged problems as the one we propose.

c) *Whole-body resolution*: From the OCP, we obtain a reference trajectory for the centroidal dynamics that we follow using a second-order inverse kinematics (IK) solver similar to [24]. In addition, the IK must track the end-effector trajectories. Optimal forces are also extracted from the OCP and can be used as references to control the robot.

C. Experimental results

This part reports the experiments achieved on the HRP-2 robot in real conditions and on TALOS in simulation.

1) *Experiment 1 - climbing up 10-cm high steps*: The experimental setup is an industrial stairs made of six 10-cm high steps. The steps have a length of 30 cm. The durations of the single and double support phases are 1.4 s and 0.2 s respectively. The resulting motion is depicted in Fig. 6. During execution, the reference posture is tracked as well as the reference foot forces using the robot low-level control system (named HRP “stabilizer”).

Computing the 25s of motion takes 42 iterations of the multiple-shooting algorithm, costing about 8s in total. In average, each iteration takes approximately 0.2s for 25s of motion. About 70% of the computation time is spent solving the underlying quadratic program of the multiple-shooting algorithm and other 20% are dedicated to the numerical integration of the dynamics together with the computations of sensitivities (derivatives).

Fig. 5 shows two trajectories of the CoM projected in the right foot frame: the black curve takes into account the log-pdf term in the cost function, while the green one does not. The figure also includes the level sets of the GMM of right foot (depicted in Fig. 2). It appears that the OCP tends to maximize the inclination of the CoM to stay in the most feasible region, i.e. closed to the maxima of the PDF. On the contrary, if we do not add the log-pdf term, the CoM tends to be infeasible.

It is worth to mention that maximizing the stay of the CoM in the most feasible region leads to a crouching walk. This may be not desirable for aesthetics. However, in the case of HRP-2, this crouch of the legs must exist to avoid singularities which are not tolerated by the low-level controller.

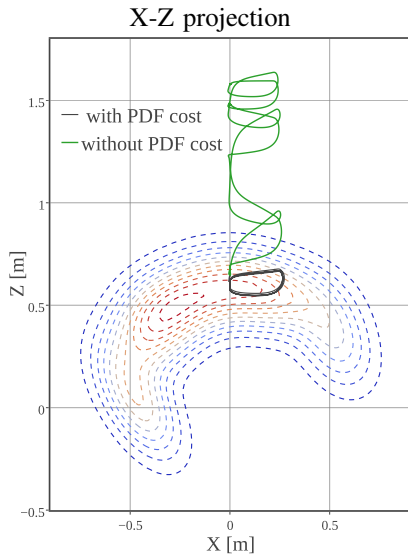


Fig. 5. Projection of the CoM trajectory inside the right foot frame with and without taking into account the log-pdf term in the cost function. The level set corresponds to the GMM distribution used in our OCP.

2) *Experiment 2 - climbing up 15-cm high steps with guardrail support*: The experimental setup is another industrial stairs made of four 15-cm high steps and equipped with a guardrail. The steps have a length of 30 cm too. The durations of the double and triple support phases are 1.8 s and 0.4 s respectively. Here, the double support phases correspond either to the case of two feet on the steps or one foot plus the right gripper on the handrail. Snapshots of the entire motion are shown in Fig. 7.

We reproduce the climbing stairs with guardrail scenario, but this time with the TALOS robot in simulation. Compared to HRP-2, TALOS is a 1.78m high humanoid robot weighting around 100kg. For this experiment, only the end-effector trajectories and the GMMs are different: the cost function remains the same. The complete motion is depicted in Fig. 8.

VI. CONCLUSION AND FUTURE WORKS

In this paper, we introduce a systematic approach to include feasibility constraints inside the optimal control formulation as occupancy measure. In particular, we propose an effective way to learn the CoM feasibility constraint by learning the probability density of the CoM positions with respect to the end-effector locations. We demonstrate the validity of the methods with two real experiments on the HRP-2 which was asked to climb industrial stairs with or without handrails and one experiment in simulation with the TALOS platform which was asked to achieve multi-contact stairs climbing.

The methodology requires a systematic learning procedure to be executed off-line in simulation. On-line, the resulting optimal control is solved in a very efficient way (about 100 times faster than execution time) and leads to smooth centroidal trajectory easily tracked by the robot whole body.

We have defined our proxy to be an occupancy measure over the whole centroidal state and contact forces, although only the

measure over the CoM was approximated. Learning the CoM kinematic feasibility is only a first stage. We plan to extend the approach by learning the constraints related to the other centroidal variables. This may severely improve the quality of the whole-body behavior. The methodology would also nicely apply to learning the terminal viability constraint (3f). As discussed at the end of Section IV, we believe that introducing the occupancy measure over all centroidal variables would reduce the locomotion problem to a simple optimal control problem composed of a single cost function, with only initial constraints.

Finally, our methodology goes beyond trajectory optimization. We have introduced proxies as a way to ensure consistency between two sets of motion models (centroidal model and whole-body model). The same consistency is desirable between the centroidal pattern generator and the upstream contact planner, whose motion models could be similarly learned by statistical sampling.

ACKNOWLEDGMENTS

We warmly thank the Simulation and Optimization of Heidelberg University for providing the optimal control framework MUSCOD-II. This work is supported by the European Research Council (ERC) through the Actanthrope project (ERC grant agreement 340050), the RoboCom++ FLAG-ERA JTC 2016 proposal and the French National Research Agency through the Loco3D project (ANR grant agreement ANR-16-CE33-0003).

REFERENCES

- [1] Christopher M Bishop. Pattern recognition and machine learning, 2006.
- [2] Camille Brasseur, Alexander Sherikov, Cyrille Collette, Dimitar Dimitrov, and Pierre-Brice Wieber. A robust linear mpc approach to online generation of 3D biped walking motion. In *IEEE-RAS Int. Conf. on Humanoid Robotics (ICHR)*, 2015.
- [3] Timothy Bretl. Motion planning of multi-limbed robots subject to equilibrium constraints: The free-climbing robot problem. *The International Journal of Robotics Research (IJRR)*, 25(4):317–342, 2006.
- [4] Justin Carpentier, Steve Tonneau, Maximilien Naveau, Olivier Stasse, and Nicolas Mansard. A versatile and efficient pattern generator for generalized legged locomotion. In *IEEE-RAS Int. Conf. on Robotics and Automation (ICRA)*, 2016.
- [5] Hongkai Dai and Russ Tedrake. Planning robust walking motion on uneven terrain via convex optimization. In *IEEE-RAS Int. Conf. on Humanoid Robotics (ICHR)*, 2016.
- [6] Hongkai Dai, Andrés Valenzuela, and Russ Tedrake. Whole-body motion planning with centroidal dynamics and full kinematics. In *IEEE-RAS Int. Conf. on Humanoid Robotics (ICHR)*, 2014.
- [7] Robin Deits and Russ Tedrake. Footstep planning on uneven terrain with mixed-integer convex optimization.

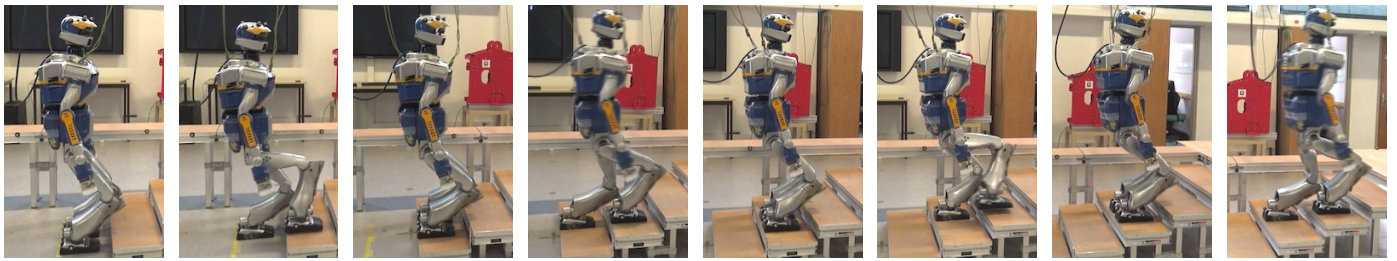


Fig. 6. Snapshots of the climbing up 10-cm high steps motion with the HRP-2 robot.

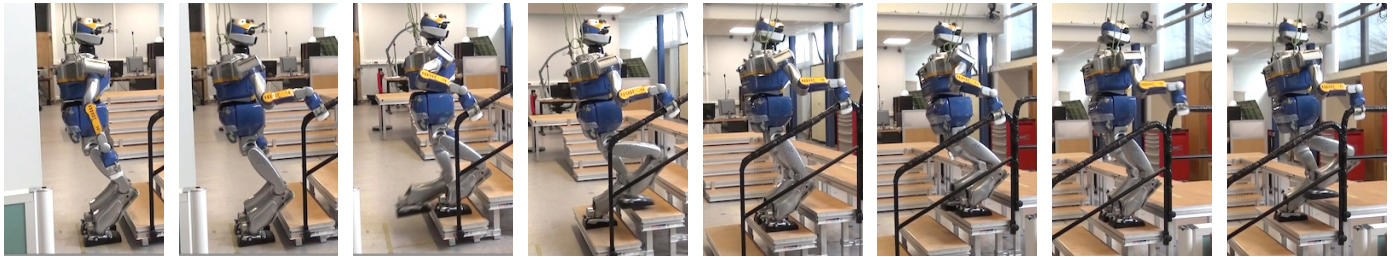


Fig. 7. Snapshots of the climbing up 15-cm high steps motion with the HRP-2 using the guardrail.



Fig. 8. Snapshots of the climbing 15-cm high steps motion with guardrail by the TALOS robot in simulation.

- In *IEEE-RAS Int. Conf. on Humanoid Robotics (ICHR)*, 2014.
- [8] Arthur Dempster, Nan Laird, and Donald Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistical society. Series B (methodological)*, pages 1–38, 1977.
- [9] Moritz Diehl, Hans Georg Bock, Holger Diedam, and Pierre-Brice Wieber. Fast direct multiple shooting algorithms for optimal robot control. In *Fast motions in biomechanics and robotics*, pages 65–93. Springer, 2006.
- [10] Adrien Escande, Abderrahmanne Kheddar, and Sylvain Miossec. Planning support contact-points for humanoid robots and experiments on HRP-2. In *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2006.
- [11] Andrei Herdt, Nicolas Perrin, and Pierre-Brice Wieber. Walking without thinking about it. In *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2010.
- [12] John Hershey and Peder Olsen. Approximating the kullback leibler divergence between gaussian mixture models. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2007.
- [13] Alexander Herzog, Nicholas Rotella, Stefan Schaal, and Ludovic Righetti. Trajectory generation for multi-contact momentum control. In *IEEE-RAS Int. Conf. on Humanoid Robotics (ICHR)*, 2015.
- [14] Hirohisa Hirukawa, Shizuko Hattori, Kensuke Harada, Shuuji Kajita, Kenji Kaneko, Fumio Kanehiro, Kiyoshi Fujiwara, and Mitsuharu Morisawa. A universal stability criterion of the foot contact of legged robots-adios zmp. In *IEEE-RAS Int. Conf. on Robotics and Automation (ICRA)*, 2006.
- [15] Shuuji Kajita, Fumio Kanehiro, Kenji Kaneko, Kazuhito Yokoi, and Hirohisa Hirukawa. The 3D linear inverted pendulum mode: A simple modeling for a biped walking pattern generation. In *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2001.
- [16] Shuuji Kajita, Fumio Kanehiro, Kenji Kaneko, Kiyoshi Fujiwara, Kensuke Harada, Kazuhito Yokoi, and Hirohisa Hirukawa. Biped walking pattern generation by using preview control of zero-moment point. In *IEEE-RAS Int. Conf. on Robotics and Automation (ICRA)*, 2003.
- [17] Kai Henning Koch, Katja Mombaur, and Philippe Souères. Optimization-based walking generation for humanoid robot. In *IFAC Symposium on Robot Control (SYROCO)*, 2012.
- [18] Daniel Leineweber, Irene Bauer, Hans Georg Bock, and Johannes P Schlöder. An efficient multiple shooting based reduced sqp strategy for large-scale dynamic process optimization. part 1: theoretical aspects. *Computers & Chemical Engineering*, 27(2):157–166,

- 2003.
- [19] Igor Mordatch, Emanuel Todorov, and Zoran Popović. Discovery of complex behaviors through contact-invariant optimization. *ACM Transactions on Graphics (TOG)*, 31(4):43, 2012.
 - [20] David E. Orin, Ambarish Goswami, and Sung-Hee Lee. Centroidal dynamics of a humanoid robot. *Autonomous Robots*, 35(2):161–176, 2013.
 - [21] Andreas Orthey and Olivier Stasse. Towards reactive whole-body motion planning in cluttered environments by precomputing feasible motion spaces. In *IEEE-RAS Int. Conf. on Humanoid Robotics (ICHR)*, 2013.
 - [22] Emanuel Parzen. On estimation of a probability density function and mode. *The annals of mathematical statistics*, 33(3):1065–1076, 1962.
 - [23] Nicolas Perrin, Olivier Stasse, Léo Baudouin, Florent Lamiroux, and Eiichi Yoshida. Fast humanoid robot collision-free footstep planning using swept volume approximations. *IEEE Transactions on Robotics (T-RO)*, 28(2):427–439, 2012.
 - [24] Layale Saab, Oscar E. Ramos, François Keith, Nicolas Mansard, Philippe Souères, and Jean-Yves Fourquet. Dynamic whole-body motion generation under rigid contacts and other unilateral constraints. *Transactions on Robotics (TRO)*, 29(2):346–362, 2013.
 - [25] Alexander Sherikov, Dimitar Dimitrov, and Pierre-Brice Wieber. Whole body motion controller with long-term balance constraints. In *IEEE-RAS Int. Conf. on Humanoid Robotics (ICHR)*, 2014.
 - [26] Olivier Stasse, Thomas Flayols, Rohan Budhiraja, Kevin Giraud-Esclasse, Justin Carpentier, Andrea Del Prete, Philippe Souères, Nicolas Mansard, Florent Lamiroux, Jean-Paul Laumond, Luca Marchionni, Hilario Tome, and Francesco Ferro. TALOS: A new humanoid research platform targeted for industrial applications. Submitted to IROS, March 2017.
 - [27] Yuval Tassa, Tom Erez, and Emanuel Todorov. Synthesis and stabilization of complex behaviors through online trajectory optimization. In *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2012.
 - [28] Steve Tonneau, Nicolas Mansard, Chonhyon Park, Dinesh Manocha, Franck Multon, and Julien Pettré. A reachability-based planner for sequences of acyclic contacts in cluttered environments. In *The International Symposium on Robotics Research (ISRR)*, 2015.
 - [29] Pierre-Brice Wieber. On the stability of walking systems. In *Int. Ws. on Human Friendly Robotics (HFR)*, 2002.
 - [30] Pierre-Brice Wieber. Trajectory free linear model predictive control for stable walking in the presence of strong perturbations. In *IEEE-RAS Int. Conf. on Humanoid Robotics (ICHR)*, Dec 2006.
 - [31] Pierre-Brice Wieber. Viability and predictive control for safe locomotion. In *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2008.
 - [32] Petr Zaytsev. *Using controllability of simple models to generate maximally robust walking-robot controllers*. PhD thesis, Cornell University, 2015.