Whole-body Trajectory Optimization for Non-periodic Dynamic Motions on Quadrupedal Systems

Andreea Radulescu[1] Ioannis Havoutis[2,3] Darwin G. Caldwell[1] Claudio Semini[1]

Abstract— Autonomous legged robots will be required to handle a wide range of tasks in complex environments. While a lot of research has focused on developing their abilities for periodic locomotion tasks, less effort has been invested in devising generalized strategies for dynamic, non-periodic movements. Motion design approaches are frequently enlisted in the form of teleoperation or predefined heuristics in such scenarios. We employ a realistic simulation of the hydraulically actuated HyQ2Max quadrupedal system for investigations on two distinctive tasks: rearing and posture recovery. We present a whole-body optimization methodology for non-periodic tasks on quadrupedal systems. This approach delivers solutions involving multiple contacts without the need for predefined feet placements. The results obtained show the potential of optimization approaches for motion synthesis in the context of complex tasks.

Index Terms— *optimization, parametrized policy, multi-legged systems, switching contacts, non-periodic movements, quadruped, posture recovery, whole-body trajectory*

I. INTRODUCTION

Modern robotic systems come in diverse configurations depending on their function. As a consequence of the wide range of applications, complex designs emerged to meet their demands. Controlling such complex robotic systems is a challenging task, due to the kinematic and actuation redundancies and due to discontinuities in the dynamics, introduced by switching contact conditions with the environment.

Biological legged systems can achieve a variety of whole body movements, in order to manipulate and traverse their environment. They exercise control over their limbs with significant versatility, compliance and energy efficiency. This performance is achieved despite noticeable levels of delay and noise affecting the biological motor systems [1]. When transferring these skills to their robotic counterparts, most research has focused on periodic tasks, frequently designed with respect to a stability criterion, such as trotting and walking.

A fully autonomous locomotion system will have to complete a heterogeneous range of tasks. Some of these frequent tasks in complex environments will require non-periodic solutions which can be described as single-shot movements. Examples in quadrupedal locomotion include

rearing, overcoming an obstacle or gap, squat-jumping in place, posture and fall recovery.

Currently, the majority of robotic systems operating in an unsafe, disorganized and cluttered environment (e.g., search and rescue missions, disaster response, nuclear decommissioning) have to rely heavily on teleoperation in order to achieve their objectives. Extending the autonomy level of legged systems with such dynamic capabilities would ease the workload of the human operators. Taking into consideration the time-sensitive nature of some of these tasks, the presence of a large motion library would facilitate the improvement of the overall performance of the system.

Optimization and learning methodologies could deliver solutions for such scenarios by using high-level task specifications, in the form of an evaluation criterion of the overall performance of the emerging behavior.

In this work we present a whole-body optimization methodology for non-periodic tasks on quadrupedal systems. We encode the high-level goals using a task specific cost function, which provides an intuitive way of defining the desired outcome. Although tuning the relative weights of such cost functions is a manual process, a heuristic encoding of the same behavior would require a significantly higher effort. The resultant trajectories and their performance indicate the capability of the approach to deliver diverse sets of motions, without prior definitions for the feet placements.

II. STATE OF THE ART

The relationship between learning and optimization has been under analysis for a long time [2], [3]. However, it is only in the recent past that their use has been extended to high dimensional problems, common to modern multi-degree-of-freedom robotics applications. The use of policy based formulations, rather than value function based ones, allows the integration of task/domain specific knowledge in the pre-structure of the policy. This allows the optimization methods to focus the search in promising areas of the space.

Various optimization approaches were proposed for dealing with multiple contact events. In [4], on-line Model Predictive Control (MPC) is combined with offline trajectory optimization. The limit cycle of a periodic movement is found by offline optimization with an infinite-horizon average cost, while on-line MPC is used to obtain an optimal feedback control law. Another study [5], generalizes MPC from the usual finite horizon (or receding horizon) setup to a first-exit setting (i.e., a solution is found based on the assumption of an exit state), which avoids dealing with discontinuities in the on-line optimization phase.

The work in [6] proposes a locally optimal solution based on MPC with smooth approximation of contact forces without the need of switching dynamics. In [7], the contact forces are explicitly included as constraints (using complementarity conditions) and directly optimized, together with the trajectories and control commands, using sequential quadratic programming.

Approaches such as Policy Improvement with Path Integrals (PI^2) [8], based on stochastic optimal control principles, have been successful in generating optimal manipulation solutions for compliant robotic arms [9]. In [10], PI^2 is used on a combination of simulation and hardware based optimizations, to synthesize a periodic hopping behavior for a one-legged system, in a number of scenarios. Using iterative optimal control, [11] delivers locally optimal solutions for both periodic and non-periodic tasks on a quadrupedal system.

The Covariance Matrix Adaptation (CMA) algorithm [12] has been similarly used to generate whole body movements. The study in [13] employs the CMA Evolution Strategy (CMA-ES) to obtain an optimal fast walking solution for both forward and sideways stepping. A preliminary study on the potential of CMA-ES was presented in [14], where the algorithm was used to obtain a rearing solution for the HyQ quadrupedal robotic system. In [15] we conduct a preliminary study for the HyQ2Max platform for similar tasks. Likewise, in [16] the method is employed to achieve a squat-jump movement, as well as various periodic gaits. The CMA method was shown to provide improved performance, when compared with state-of-the-art global search methods [17], thus making it a method of choice for such investigations.

Another major challenge is solving such high-level tasks without the use of pre-defined heuristics such as hard-coded sequences or feet placements [18]. Trajectory optimization for multi-legged robotic systems that operate in complex environments is challenging, due to the varying number of contacts. By allowing predefined elements, part of the burden is alleviated, but at the same time the solutions obtained might be suboptimal and their generality can also be affected. Recently, the work in [19] has combined optimal control and learning of movement primitives to generate gaits for a humanoid system, but the work is restricted to open-loop control.

In spite of the significant efforts in the area of fall avoidance, comparatively little research has focused on developing generalized self-righting techniques. Most work has revolved around devising specific solutions for particular systems, either at hardware design level (low center of mass, invertible robots [20]) or defining embodiment specific strategies [21]. The work in [22] is attempting to develop a generalized method for self-righting strategies, by analyzing and exploiting the given robot structure. However, the results are still restricted to small dimensional designs and do not address multi-legged systems (the study focuses on a tracked one-arm manipulator).

III. OUR METHODOLOGY

In our work we try to address this issue by providing a generalized approach of delivering whole body movement solutions. The high-level tasks are encoded using a cost function, while trying to avoid pre-specifying how these tasks should be solved (i.e., no predefined contact points or sequences). As the accuracy of the model is a crucial requirement, we choose to perform the optimization using the whole body dynamics.

A. Robot description and System model

We use a realistic simulation of the 80 kg HyQ2Max [23] quadrupedal robotic system with contacts. The platform was designed as a light-weight, robust locomotion vehicle and features 12 hydraulically actuated joints (controlled by a hydraulic valve). Each limb has 3 actuators, defined as: HAA (hip abduction/adduction), HFE (hip flexion/extension) and KFE (knee flexion/extension). The corresponding kinematic ranges are $90°$, $270°$, $165°$, respectively.

We model the behavior of the platform as a rigid body system defined as:

$$\mathbf{M}\left(\mathbf{q}\right)\ddot{\mathbf{q}}+\mathbf{C}\left(\mathbf{q},\dot{\mathbf{q}}\right)\dot{\mathbf{q}}+\mathbf{g}\left(\mathbf{q}\right)+\mathbf{D}\dot{\mathbf{q}} = \mathbf{J}_c^T\lambda\left(\mathbf{q},\dot{\mathbf{q}}\right)+\mathbf{S}^T\boldsymbol{\tau}, \quad (1)$$

where $\mathbf{q} = [\mathbf{q}_B, \mathbf{x}_B, \mathbf{q}_J]^T$ is the vector containing the 6 DoF base state (orientation $\mathbf{q}_B$, position $\mathbf{x}_B$ of the trunk's Center of Mass (COM)) and the joint angles $\mathbf{q}_J$, with $\dot{\mathbf{q}}, \ddot{\mathbf{q}}$ the corresponding velocities and accelerations. The remaining notations are defined as following: $\mathbf{M}$ is a symmetric positive definite inertia mass matrix, $\mathbf{C}$ represents centrifugal and Coriolis forces, $\mathbf{g}$ are the gravitational forces, $\mathbf{D}$ is the viscous damping matrix. The contact forces λ exert their effect on the system via the Jacobian $\mathbf{J}_c$, while the joint torques from the actuators $\boldsymbol{\tau}$ are mapped via the selection matrix $\mathbf{S}$ onto the states.

The state vector of the HyQ2Max system is defined as:

$$\mathbf{x} = [\mathbf{q}, \dot{\mathbf{q}}]^T = [\mathbf{q}_B, \mathbf{x}_B, \mathbf{q}_J, \dot{\mathbf{q}}_B, \dot{\mathbf{x}}_B, \dot{\mathbf{q}}_J]^T \quad (2)$$

where $\mathbf{q}, \dot{\mathbf{q}}$ are the vectors defined for (1), expressed in the world frame. The definitions of the axes of the robot base frame are depicted in Fig. 2 (left).

The choice of contact model is crucial for ensuring the validity of the resultant trajectory. When considering the choice of models, there is always a trade-off between numerical stability (smooth contact dynamics) and physical accuracy.

In the Open Dynamics Engine (ODE) [24] contact model, any potential penetration errors are corrected at each time step (some penetration threshold is tolerated in order to remove jitter) which provides a good gradient of the dynamics. We employ Gazebo for our experiments, a well established simulation environment for robotic applications, which combines all the functionalities of ODE with an increased user accessibility [25].

B. Optimization

In our investigation we use a CMA-ES based approach to address dynamic non-periodic tasks for a quadrupedal robotic system: rearing and posture recovery. This technique proved

effective in handling nonlinear, high-dimensional problems. The method operates by generating and evaluating a set of solutions at each iteration. The samples are extracted from a multivariate Gaussian distribution. After their evaluation the covariance matrix of the search distribution is adapted.

For complex robotic systems such as HyQ2Max, the main challenge consists of determining the suitable joint motions which achieve the desired movement for each given task. Using a direct optimization approach on the time-parametrized joint or torque trajectories would involve an inconveniently large search space. Hence, we use a parametrized policy to encode these profiles, represented as a weighted average of Gaussian kernels:

$$f(t) = \sum_{i=1}^{M} w_i \phi_i(t) / \sum_{i=1}^{M} \phi_i(t), \quad (3)$$

where w_i are the weights associated with each kernel ϕ_i, $i \in [1, M]$, defined by mean μ_i and variance σ^2 as in:

$$\phi_i(t) = exp(-\frac{1}{2\sigma^2}(t - \mu_i)). \quad (4)$$

The CMA algorithm is then used to optimize the weights of all policies according to a task specific cost function (5), applied to the whole body trajectory generated through (3) :

$$J = p(\mathbf{s}_T, \dot{\mathbf{s}}_T) + \int_{t=0}^{T} r(\tau_t, \mathbf{q}_J) \, dt, \; t \in [0, T], \quad (5)$$

where $\mathbf{s}_t = [\mathbf{q}_B, \mathbf{x}_B]'$ is the trunk state (i.e., $\mathbf{s} = [x_{COM}, y_{COM}, z_{COM}, roll, pitch, yaw]'$) and $\boldsymbol{\tau}_t$ is the set of 12 torque actuation commands at time $t \in [0, T]$.

The cost function consists of a running term r that seeks to minimize the effort used for producing the motion and a final cost p that evaluates the success of the motion according to the task goals (desired final states $(\mathbf{q}_B^*, \mathbf{x}_B^*)$) :

$$r(\mathbf{s}_t, \dot{\mathbf{s}}_t, \tau_t) = Q_3 \tau_t^2 + Q_6[(\mathbf{q}_J - \mathbf{q}_{J max}))^2 \big|_{\mathbf{q}_J > \mathbf{q}_{J max}} \quad (6)$$

$$+ (\mathbf{q}_J - \mathbf{q}_{J min}))^2 \big|_{\mathbf{q}_J < \mathbf{q}_{J min}}]$$

$$p(\mathbf{s}_T, \dot{\mathbf{s}}_T) = (\mathbf{q}_B(T) - \mathbf{q}_B^*)' Q_1 (\mathbf{q}_B(T) - \mathbf{q}_B^*)$$

$$+ (\mathbf{x}_B(T) - \mathbf{x}_B^*)' Q_2 (\mathbf{x}_B(T) - \mathbf{x}_B^*) \quad (7)$$

where $\mathbf{q}_{J min}$, $\mathbf{q}_{J max}$ are the upper and lower bounds on the joint limits $\mathbf{q}_J$, and Q_i, $i \in \{1, 2, 3, 4\}$ are the relative weights associated with the terms for each individual task.

Since the CMA-ES is a local method, a starting policy is required. In our experiments the policies are initialized with values that maintain an initial pose (in the case of rearing tasks) or as a direct interpolation between the initial configuration and desired final pose (for posture recovery). No specific footstep sequence information is enforced. We apply a staged optimization approach, where preliminary solutions are obtained using a relaxed cost (e.g., lower weights for the joint limits term). The complexity of the cost function is increased gradually, until the solution meets all the desired requirements[1]. The procedure is based on

---

[1] We note that the robot model used incorporates a model of the actuator dynamics, thus increasing the feasibility of the resultant behavior.
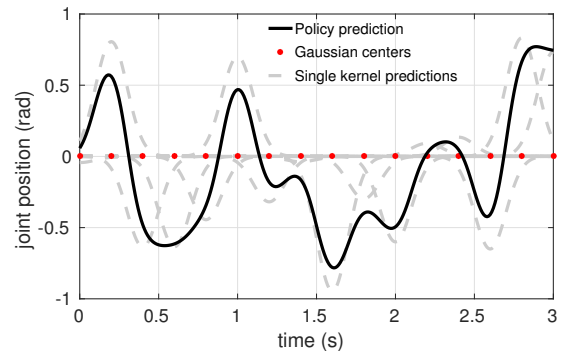

Fig. 1: Example of a policy encoded as a weighted average of Gaussian kernels: the means μ_i are equally spaced and the variances are all fixed to 0.01 and the weights have been sampled from $[-1, 1]$.

optimization techniques, but the policy always starts from a general non-specific point. Thus, the method can be viewed as a learning approach, where a strategy is learned from scratch, rather than an existing solution is improved.

An example of such a resultant policy is depicted in Fig. 1 where the $M = 16$ Gaussian kernels' means are equally spaced, the variances σ^2 are all fixed to 0.01 and the weights w_i have been sampled from the interval $[-1, 1]$. In our experiments we use 12 such representations, one for each degree of freedom (DoF) of the quadrupedal system.

IV. SIMULATION RESULTS

We present the results obtained on a realistic simulation of the HyQ2Max robotic platform, as detailed in the previous section. All policies are encoded with a fixed set of $M = 16$ Gaussian kernels evenly spaced in the time interval between 0 to T seconds, while their variance σ^2 is fixed to 0.01. The number of samples for each task is determined by the CMAS-ES algorithm, based on the number of decision variables.

A. Rearing

In the rearing scenario, the quadrupedal system starts from a neutral, four legged support pose (Fig. 2 (left)). During the task execution the front legs push the torso upwards, while the lower legs are supporting the body. This behavior can serve as a preliminary stage for much more complex maneuvers (e.g., obstacle traversing, transition to bipedal posture). We note that in general such postures cannot be reached in a static manner. To further reduce the complexity of the problem, we exploit the task structure and impose the same policy for the front and hind leg pairs, respectively.

The policy is initialized to values that maintain the default pose of the system, in four legged support (Fig. 2, left). The policy converged to a feasible solution within approximately 3002 trials (177 evaluation episodes with 17 samples per episode), for an allocated time horizon of $T = 0.3 \; s$. The values of the weights used for this task were $Q_1 = 10^3$, $Q_2 = 10^2$, $Q_3 = 0.1$, $Q_4 = 1$.

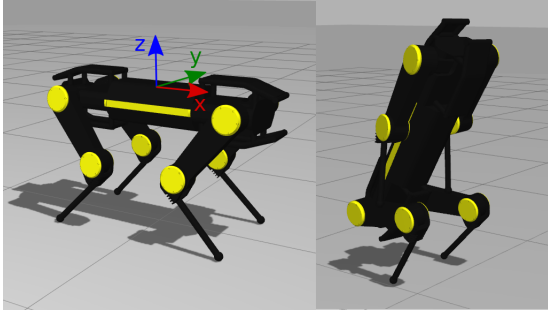Fig. 2 (right) shows the final pose reached by the system under the resultant policy, as imposed by the requirements

Fig. 2: HyQ2Max performing the rearing task in simulation under the resultant policy. *Left:* initial pose. *Right:* final pose.

on the position and orientation of the trunk, encoded in the terminal cost p. In the case of rearing we designated these only as the desired torso pitch ($-1.0472 \ rad$) and CoM (center of mass) height ($90 \ cm$), leaving the other dimensions free for the optimization. The target pose was achieved with an accuracy of $0.006 \ rad$ and $10 \ cm$, the performance is reflecting the relative ratios of the cost function weights.
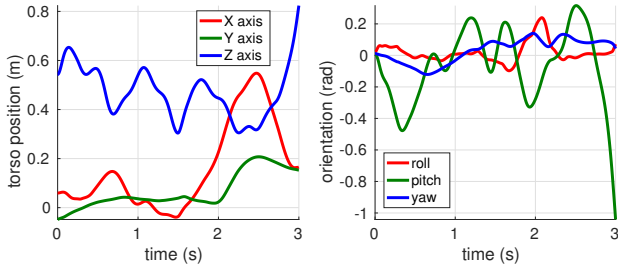
(a) Trunk trajectories (trunk CoM) of HyQ2Max performing therearing task under the resultant policy.
Body position (*left*) and orientation (*right*) in the world frame.
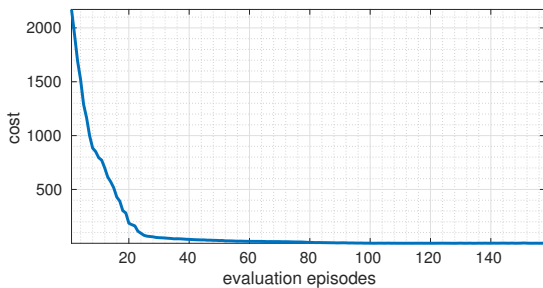
(b) Evolution of the cost for the rearing experiment
(158 evaluation episodes with 19 samples per episode)

Fig. 3: Joint trajectory solution for the HyQ2Max platform in the rearing task: (a) Resultant trunk trajectory. (b) Cost evolution.

The evolution of the position and orientation of the robot's trunk for the final trial are depicted in Fig. 3, alongside the evolution of the cost throughout the CMAS-ES (Fig. 3-(b)). The resultant individual joint policies (with the hind and front leg pairs sharing the same policies) are displayed in Fig. 4.
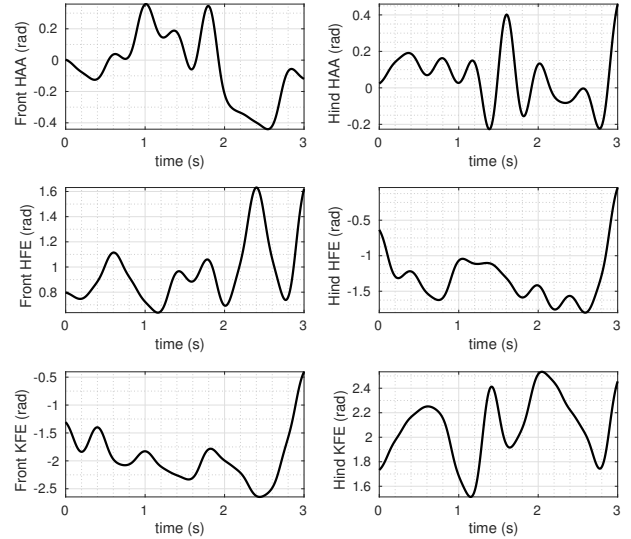

Fig. 4: The optimized joint trajectories for the rearing task on the HyQ2Max platform. We impose the same policy for the front and hind leg pairs. Each limb has 3 actuators : HAA (hip abduction/adduction), HFE (hip flexion extension) and KFE (knee flexion/extension).

B. Posture recovery

For the task of posture recovery we introduce a scenario where the regular locomotion task fails, due to an unexpected obstacle, and the robot finds itself in an unforeseen, fallen state (Fig. 5, left). The task consists of returning the system to an upright position, that allows the resuming of the locomotion task. Unlike in the rearing scenario, due to the nature of the task, the individual leg policies are independent of each other.

The final cost p penalizes deviations from a target final state of the robot's trunk (in both linear and angular DoF). The values of the weights used for this task were $Q_1 = 10^3$, $Q_2 = 10^3$, $Q_3 = 0.1$, $Q_4 = 1$. The desired final pose was defined as a neutral orientation ($\mathbf{q}_B^* = (roll, \ pitch, \ yaw) = (0, 0, 0)$), while the position was specified only by CoM (center of mass) height ($61 \ cm$), leaving the other dimensions free for the optimization.

The policy converged to a feasible solution within approximately 3002 trials (158 evaluation episodes with 19 samples per episode), with a time horizon $T = 0.3 \ s$. The desired pose was achieved with an accuracy of $0.001 \ rad$ and $2 \ cm$. As in the previous scenario, the performance is reflecting the relative ratios of the cost function weights.

Fig. 5 (middle) shows an intermediary pose reached by the system under the resultant policy. Fig. 5 (right) depicts the final pose after the policy is executed and the current goal is maintaining the pose. The trajectories of the position and orientation of the robot's trunk in the final solution are depicted in Fig. 6 (a). The evolution of the cost function from initial plan to the delivered solution is depicted in Fig. 6 (b). The resultant individual joint policies are displayed in Fig. 7.
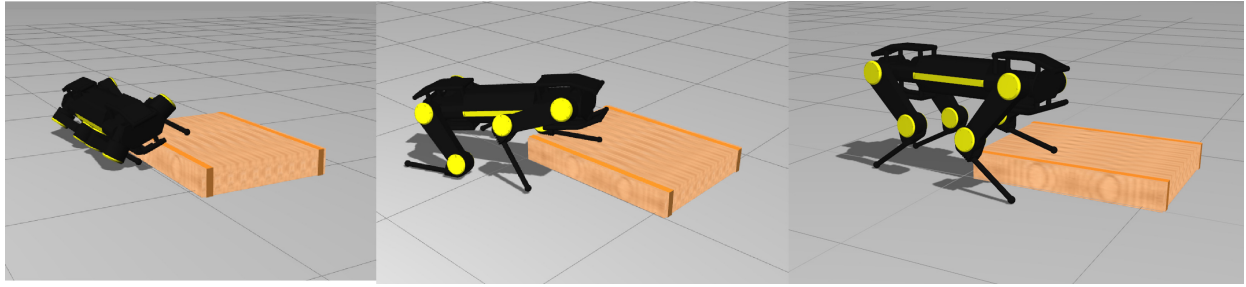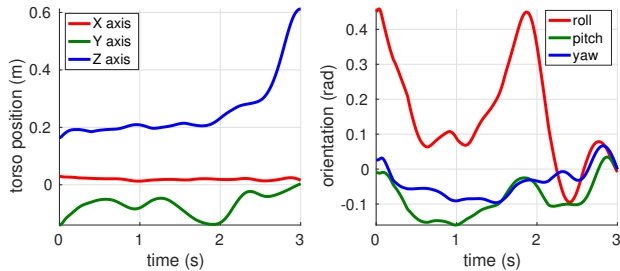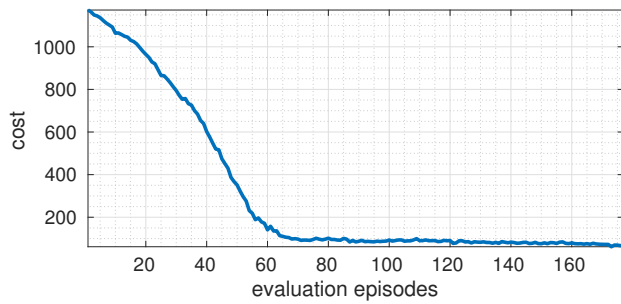
Fig. 5: HyQ2Max performing the posture recovery task in simulation. *Left:* initial pose (unexpected fall, ending up partially on an obstacle). *Middle:* intermediary pose under the resultant policy. *Right:* final pose (after the policy is executed and the current goal is maintaining the pose).

(a) Trunk trajectories (trunk CoM) of HyQ2Max performing the posture recovery task under the resultant policy.
Left: absolute body position. *Right:* body orientation.

(b) Evolution of the cost for the posture recovery experiment
(177 evaluation episodes with 17 samples per episode)

Fig. 6: Joint trajectory solution for the HyQ2Max platform in the posture recovery task: (a) Resultant trunk trajectory (position and orientation). (b) Cost evolution.

The presented results suggest that the obtained behaviors could be transferred to the real hardware, which is what we are focusing on implementing in the near future. By enforcing high gains on the torque and joint limits' costs, the feasibility of the whole body trajectory generated can be achieved. At the same time we are aiming to expand the range of posture recovery and extend the approach to full body self-righting. A video of the presented results and behaviors is available in the additional material or at the following link.[2]

V. CONCLUSION

In this work we present a whole body optimization methodology for non-periodic dynamic movements on
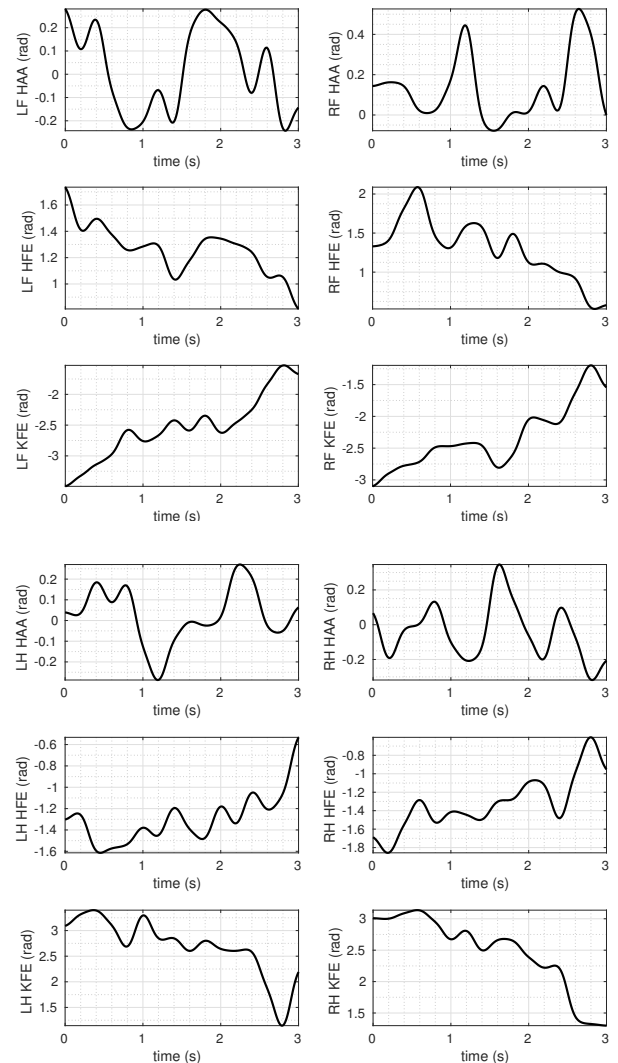
Fig. 7: The optimized joint trajectories for the posture recovery task on the HyQ2Max platform. (LF - left front, RF - right front, LH - left hind, RH - right hind).

quadrupedal systems. The approach is able to deliver trajectory solutions which involve multiple contacts, without any predefined feet placement heuristics (e.g., contact points, timing or order of succession).

Using a realistic simulation of the hydraulically actuated HyQ2Max quadrupedal system we investigate two distinctive tasks: rearing and posture recovery. Task specific cost functions encode the high-level goals in each scenario. We employ a generalized form for the cost function, while adjusting the relative gains of each term according to the current task. Although tuning the relative weights of such cost functions is a manual process, a heuristic encoding of the same behavior would require a significantly higher effort. By exploiting the whole body model in order to obtain the solution, the optimization does not have to depend upon heuristics and can overcome errors caused by the use of simplified models. The resultant trajectories and the accuracy with which the user defined goals are achieved are reflecting the relative ratios of the weights on the individual cost function terms.

VI. FUTURE DIRECTIONS

The results depicted in this work showcase the possibilities that optimization can offer to motion synthesis for complex, non-periodic tasks. For example, the goal of a rearing motion can be to reach the basin of attraction of a balancing controller, keeping the quadruped upright. In future work we aim to transfer the obtained behaviors onto the hardware system, while also expanding the range of scenarios addressed.

As the future task of the system was beyond the scope of this investigation, certain elements, (e.g., the allocated time and final velocities of the trunk) were not included in the optimization goals. We aim to extend the approach to deliver an optimal duration for the given task, while the final pose is fully determined, based on subsequent requirements. To increase the autonomy of the system under the suggested approach, real-time sensory information from the environment and a methodology for determining the ideal final pose of the policy will have to be integrated. In the long term we aim to develop a general tool for generating optimal dynamic whole-body motions that are not necessarily periodic in nature.

The speed of computing such solutions might not always allow for on-line optimization using conventional approaches. Machine learning methods (in conjunction with an adequate motion library) could be employed to speed up the solution delivery time. In [26] the solutions of an optimization task are used to guide the learning of neural network controllers, for a variety of locomotion tasks on generic robotic models in simulation. Such dynamic movements will serve in complementing and extending the capabilities of robots with arms and legs.

REFERENCES

[1] A. Faisal, L. Selen, and D. Wolpert, "Noise in the nervous system," *Nature Reviews Neuroscience*, vol. 9, no. 4, 2008.

[2] K. P. Bennett and E. Parrado-Hernández, "The interplay of optimization and machine learning research," *Journal of Machine Learning Research*, vol. 7, 2006.

[3] J. Kober, J. A. Bagnell, and J. Peters, "Reinforcement learning in robotics: A survey," *The International Journal of Robotics Research*, 2013.

[4] T. Erez, Y. Tassa, and E. Todorov, "Infinite-horizon model predictive control for periodic tasks with contacts," *Robotics: Science and systems VII*, p. 73, 2012.

[5] P. Kulchenko and E. Todorov, "First-exit model predictive control of fast discontinuous dynamics: Application to ball bouncing," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2011.

[6] Y. Tassa, T. Erez, and E. Todorov, "Synthesis and stabilization of complex behaviors through online trajectory optimization," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2012.

[7] M. Posa, C. Cantu, and R. Tedrake, "A direct method for trajectory optimization of rigid bodies through contact," *The International Journal of Robotics Research*, vol. 33, no. 1, 2014.

[8] E. Theodorou, J. Buchli, and S. Schaal, "A generalized path integral control approach to reinforcement learning," *Journal of Machine Learning Research*, vol. 11, 2010.

[9] M. Kalakrishnan, L. Righetti, P. Pastor, and S. Schaal, "Learning force control policies for compliant manipulation," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2011.

[10] P. Fankhauser, M. Hutter, C. Gehring, M. Bloesch, M. A. Hoepflinger, and R. Siegwart, "Reinforcement learning of single legged locomotion," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2013.

[11] M. Neunert, F. Farshidian, A. W. Winkler, and J. Buchli, "Trajectory optimization through contacts and automatic gait discovery for quadrupeds," *IEEE Robotics and Automation Letters*, no. 99, 2016.

[12] N. Hansen and A. Ostermeier, "Completely derandomized self-adaptation in evolution strategies," *Evolutionary computation*, vol. 9, no. 2, 2001.

[13] N. Shafii, N. Lau, and L. P. Reis, "Learning to walk fast: Optimized hip height movement for simulated and real humanoid robots," *Journal of Intelligent & Robotic Systems*, vol. 80, no. 3-4, 2015.

[14] I. Havoutis, J. Buchli, D. G. Caldwell, and C. Semini, "Optimization of dynamic motions for legged robots," in *Dynamic Walking*, 2014.

[15] A. Radulescu, I. Havoutis, D. G. Caldwell, and C. Semini, "Optimization for non-periodic dynamic motions of legged systems," in *The 9th International Workshop on Human Friendly Robotics*, 2016.

[16] C. Gehring, S. Coros, M. Hutler, C. D. Bellicoso, H. Heijnen, R. Diethelm, M. Bloesch, P. Fankhauser, J. Hwangbo, M. Hoepflinger *et al.*, "Practice makes perfect: An optimization-based approach to controlling agile motions for a quadruped robot," *IEEE Robotics & Automation Magazine*, vol. 23, no. 1, 2016.

[17] N. Hansen and S. Kern, "Evaluating the cma evolution strategy on multimodal test functions," in *International Conference on Parallel Problem Solving from Nature*, 2004.

[18] A. Winkler, C. Mastalli, I. Havoutis, M. Focchi, D. G. Caldwell, and C. Semini, "Planning and Execution of Dynamic Whole-Body Locomotion for a Hydraulic Quadruped on Challenging Terrain," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2015.

[19] D. Clever, M. Harant, H. Koch, K. Mombaur, and D. Endres, "A novel approach for the generation of complex humanoid walking sequences based on a combination of optimal control and learning of movement primitives," *Robotics and Autonomous Systems*, vol. 83, 2016.

[20] P. Ben-Tzvi, A. A. Goldenberg, and J. W. Zu, "Design, simulations and optimization of a tracked mobile robot manipulator with hybrid locomotion and manipulation capabilities," in *IEEE International Conference on Robotics and Automation,(ICRA)*, 2008.

[21] U. Saranli, A. A. Rizzi, and D. E. Koditschek, "Model-based dynamic self-righting maneuvers for a hexapedal robot," *The International Journal of Robotics Research*, vol. 23, no. 9, 2004.

[22] C. C. Kessens, C. T. Lennon, and J. Collins, "A metric for self-rightability and understanding its relationship to simple morphologies," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2014.

[23] C. Semini, V. Barasuol, J. Goldsmith, M. Frigerio, M. Focchi, Y. Gao, and D. G. Caldwell, "Design of the hydraulically-actuated, torque-controlled quadruped robot HyQ2Max," *IEEE/ASME Transactions on Mechatronics*, 2016.

[24] R. Smith, "Open dynamics engine-home," 2002.

[25] N. Koenig and A. Howard, "Design and use paradigms for gazebo, an open-source multi-robot simulator," in *IEEE/RSJ International Conference on Intelligent Robots and Systems, (IROS)*, vol. 3, 2004.

[26] I. Mordatch and E. Todorov, "Combining the benefits of function approximation and trajectory optimization," in *Robotics: Science and Systems (RSS)*, 2014.