

A Family of Iterative Gauss-Newton Shooting Methods for Nonlinear Optimal Control

Markus Gifftthaler¹, Michael Neunert¹, Markus Stäubli¹, Jonas Buchli¹ and Moritz Diehl²

Abstract—This paper introduces a family of iterative algorithms for unconstrained nonlinear optimal control. We generalize the well-known iLQR algorithm to different multiple shooting variants, combining advantages like straightforward initialization and a closed-loop forward integration. All algorithms have similar computational complexity, i.e. linear complexity in the time horizon, and can be derived in the same computational framework. We compare the full-step variants of our algorithms and present several simulation examples, including a high-dimensional underactuated robot subject to contact switches. Simulation results show that our multiple shooting algorithms can achieve faster convergence, better local contraction rates and much shorter runtimes than classical iLQR, which makes them a superior choice for nonlinear model predictive control applications.

Index Terms—Numerical Optimal Control, Trajectory Optimization, Multiple Shooting, Quadrupedal Robots, Nonlinear Model Predictive Control, Differential Dynamic Programming

I. INTRODUCTION

A. Overview and Motivation

In this paper, we discuss a family of iterative Gauss-Newton shooting methods for numerically solving unconstrained optimal control problems, and illustrate the effectiveness of our algorithms with various robotics examples. We outline the connection between a number of ‘direct’ optimal control methods and Gauss-Newton methods from the class of Differential Dynamic Programming (DDP) [1] algorithms. Additionally, we present a natural extension arising from this connection and introduce a family of hybrid Gauss-Newton Multiple Shooting methods.

In direct approaches to optimal control, infinite-dimensional optimal control problems are transcribed into finite dimensional Nonlinear Programs (NLPs). Two prominent ways of transcription by ‘shooting’ are direct single shooting (SS) and direct multiple shooting (MS) [2], which differ in the choice of decision variables. In single shooting, solely the control inputs are the decision variables. Generally speaking, the control trajectory is discretized in a piece-wise polynomial fashion (for simplicity, we focus on piece-wise constant controls in this paper, c.f. Fig. 1a). A corresponding state trajectory is obtained by means of numerical forward integration of the system dynamics, starting at a given initial state. SS is often called a ‘sequential’ approach. In multiple shooting, the same discretization scheme is employed for the control inputs, but additionally, intermediate states are added to the

decision variables. This provides several advantages [3], but requires the introduction of additional matching constraints to ensure continuity of the state trajectory. The technique of introducing these additional degrees of freedom into the original problem, combined with adding matching constraints, is called *lifting* [4], and results in a ‘simultaneous’ method.

The formulation of both SS and MS as standard NLPs is straightforward and any state-of-the-art NLP solvers can be used to solve them. It is important to note that under the assumption of having a piece-wise polynomial control parameterization, the intrinsic sparsity structure of the underlying optimal control problem of both SS and MS allow them to achieve linear time complexity by performing a Riccati recursion [5].

Classical single shooting often does not perform well for unstable systems due to the pure open-loop forward integration of the system dynamics. In DDP, this is handled by doing a closed-loop forward integration, using a feedforward plus a time-varying state-feedback control law. The Riccati backward sweep designs time-varying feedback gains on the fly without additional computational cost. DDP is an exact-Hessian method, requiring the computation of second derivatives of the dynamics. While this gives the algorithm quadratic convergence, this can be impractical for use in systems with complex dynamics. For that reason, Hessian-approximating variants of DDP have become popular in the robotics community [6]–[10].

An important Hessian-approximating variant of DDP is the *iterative Linear-Quadratic Regulator* (iLQR) [11], which is also known as *Sequential Linear Quadratic Optimal Control* [12]. This method can be classified as closed-loop single shooting using a Gauss-Newton Hessian approximation and a Riccati backward sweep to solve linear-quadratic (LQ) subproblems. The Gauss-Newton Hessian approximation is based on the assumption that the objective function can be locally approximated as a sum of quadratic terms, and requires only first-order derivatives of the system dynamics. This comes at the cost of giving only linear convergence, however. The Gauss-Newton approach can be lifted, too, which has for example been shown in [13]. While it initially appears to be a drawback to increase the number of decision variables, it is important to emphasize that the lifted problem can be solved at approximately the same computational cost as the original non-lifted problem, and can lead to a significant increase of convergence speed [4]. Therefore, the fundamental motivation for this paper is to combine the benefits of iLQR with a multiple shooting approach.

¹Agile & Dexterous Robotics Lab, ETH Zürich, Switzerland. {mgifftthaler, neunertm, markusta, buchli} @ethz.ch

²Systems Control and Optimization Laboratory, Department of Microsystems Engineering (IMTEK), University of Freiburg, Germany. moritz.diehl@imtek.uni-freiburg.de

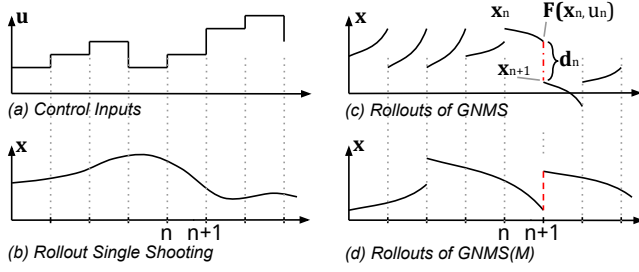


Fig. 1. Intuition about different shooting variants. (a) A zero-order hold control parameterization, with a constant control input at each stage. (b) single shooting, where the state trajectory is obtained through a single forward integration over the whole problem horizon. (c) In GNMS, intermediate states at every time-index n are introduced as additional decision variables. (d) In hybrid versions of GNMS, the multiple shooting intervals span several control intervals. The intermediate states at the beginning of the multiple shooting intervals are decision variables, and the states in between are obtained by forward integration.

B. Contribution

In this work, we derive a lifted equivalent of iLQR, called *Gauss-Newton Multiple Shooting* (GNMS), which introduces the intermediate states as additional decision variables. Next, we extend this relationship to form an entire family of open-loop multiple shooting algorithms, denoted GNMS(M), and closed-loop multiple shooting algorithms, denoted as iLQR-GNMS(M). The latter is shown to be a generalization of iLQR and can be considered multiple shooting iLQR. We outline the relationship between these algorithms and existing methods. We give simulation examples including a complex underactuated robot and compare the performance of the full-step algorithms using data gained from hardware experiments. Furthermore, we show the benefits of iLQR-GNMS(M) methods for nonlinear model predictive control.

C. Outline

This paper is structured as follows. In Section II, we derive GNMS, and present the basic update routine for state and control trajectories. Using these update equations, we generalize iLQR and GNMS to a family of algorithms in Section III. Section IV showcases several simulation results, based on data gained from hardware experiments. A discussion and outlook concludes the paper in Section V.

II. GAUSS-NEWTON MULTIPLE SHOOTING

In the following, a brief derivation of the unconstrained Gauss-Newton Multiple Shooting method is presented. We show the derivation using an intuitive value-function approach in the style of [11] in order to highlight the close relationship between GNMS and iLQR. However, from the beginning, we lift the optimization problem and introduce intermediate states as additional decision variables besides the controls. In that sense, having each control decision variable accompanied with a state-decision variable, GNMS is closely related to the original multiple shooting algorithm [2].

Consider the following discrete-time, finite-horizon, non-linear optimal control problem

$$\min_{\mathbf{u}_n, \mathbf{x}_n} \left\{ \Phi(\mathbf{x}_N) + \sum_{n=0}^{N-1} L_n(\mathbf{x}_n, \mathbf{u}_n, n) \right\} \quad (1)$$

$$\text{s.t.} \quad \mathbf{x}_{n+1} - \mathbf{f}_n(\mathbf{x}_n, \mathbf{u}_n, n) = 0, \quad \mathbf{x}_0 = \mathbf{x}_{init} \quad (2)$$

with state-vector $\mathbf{x}_n \in \mathbb{R}^m$ and control input vector $\mathbf{u}_n \in \mathbb{R}^p$. Let L_n be the intermediate cost at time-step n and $\Phi(\mathbf{x}_N)$ the terminal cost at the time horizon N .

A. Forming LQ subproblems

The optimal control law is computed in an iterative way. In each iteration $k = 0, 1, 2, \dots$, we construct a LQ optimal control problem around the state trajectory $\mathbf{X}^{[k]} = \{\mathbf{x}_0^{[k]}, \mathbf{x}_1^{[k]}, \dots, \mathbf{x}_N^{[k]}\}$, with $\mathbf{x}_0^{[k]} = \mathbf{x}_{init}$ and the control trajectory $\mathbf{U}^{[k]} = \{\mathbf{u}_0^{[k]}, \mathbf{u}_1^{[k]}, \dots, \mathbf{u}_{N-1}^{[k]}\}$. In the first iteration, $k = 0$, the LQ problem is hence constructed around the initial guesses for $\mathbf{X}^{[0]}$, $\mathbf{U}^{[0]}$. Possible initialization strategies are summarized in Section III-D.

At each iteration, we numerically forward integrate all multiple shooting intervals using the respective control inputs $\mathbf{u}_n^{[k]}$, starting at every state $\mathbf{x}_n^{[k]} \forall n = 0, 1, \dots, N-1$. Fig. 1c shows a sketch of the multiple shooting intervals in GNMS, where the resulting state at the end of each interval is denoted $\mathbf{F}^{[k]}(\mathbf{x}_n^{[k]}, \mathbf{u}_n^{[k]})$. Accordingly, we define the ‘defect’ between the integrated trajectory segment and the next intermediate state $\mathbf{x}_{n+1}^{[k]}$ as

$$\mathbf{d}_n^{[k]} = \mathbf{F}^{[k]}(\mathbf{x}_n^{[k]}, \mathbf{u}_n^{[k]}) - \mathbf{x}_{n+1}^{[k]}. \quad (3)$$

Defining state and control increments $\delta \mathbf{x}_n^{[k]}$ and $\delta \mathbf{u}_n^{[k]}$ for every single time-stage n , we can write the nonlinear system dynamics constraint (2) in terms of the simulated interval as

$$\mathbf{x}_{n+1}^{[k]} + \delta \mathbf{x}_{n+1}^{[k]} - \mathbf{F}^{[k]}(\mathbf{x}_n^{[k]} + \delta \mathbf{x}_n^{[k]}, \mathbf{u}_n^{[k]} + \delta \mathbf{u}_n^{[k]}) = 0 \quad (4)$$

which can also be considered a matching condition which ensures the continuity of the state trajectory w.r.t. state and control increments. Performing a first-order Taylor expansion of Equation (4) w.r.t. $\delta \mathbf{x}_n^{[k]}$ and $\delta \mathbf{u}_n^{[k]}$, denoting the sensitivities w.r.t state and control \mathbf{A}_n and \mathbf{B}_n and taking into account the defects as defined by Equation (3), results in the following affine system dynamics constraint

$$\delta \mathbf{x}_{n+1}^{[k]} - \mathbf{A}_n^{[k]} \delta \mathbf{x}_n^{[k]} - \mathbf{B}_n^{[k]} \delta \mathbf{u}_n^{[k]} - \mathbf{d}_n^{[k]} = 0. \quad (5)$$

Analogously performing a second-order Taylor expansion of the nonlinear cost function (1) gives rise to the following LQ optimal control problem

$$\begin{aligned} \min_{\delta \mathbf{u}_n, \delta \mathbf{x}_n} \left\{ q_N + \delta \mathbf{x}_N^\top \mathbf{q}_N + \frac{1}{2} \delta \mathbf{x}_N^\top \mathbf{Q}_N \delta \mathbf{x}_N \right. \\ \left. + \sum_{n=0}^{N-1} q_n + \delta \mathbf{x}_n^\top \mathbf{q}_n + \delta \mathbf{u}_n^\top \mathbf{r}_n + \frac{1}{2} \delta \mathbf{x}_n^\top \mathbf{Q}_n \delta \mathbf{x}_n \right. \\ \left. + \frac{1}{2} \delta \mathbf{u}_n^\top \mathbf{R}_n \delta \mathbf{u}_n + \delta \mathbf{u}_n^\top \mathbf{P}_n \delta \mathbf{x}_n \right\} \end{aligned} \quad (6)$$

$$\text{s.t.} \quad \delta \mathbf{x}_{n+1} = \mathbf{A}_n \delta \mathbf{x}_n + \mathbf{B}_n \delta \mathbf{u}_n + \mathbf{d}_n \quad (7)$$

$$\mathbf{x}_0 = \mathbf{x}_{init} \quad (8)$$

where we assume $\mathbf{Q}_n, \mathbf{Q}_N \geq 0$ and $\mathbf{R}_n > 0$. Here, and in the following subsection, we drop the superscript indices $[k]$ for better readability.

B. Computing the Optimal Control by Riccati Recursion

Considering the LQ subproblem (6)-(8), the optimal control and state updates can be computed using a value-function approach. Assume a quadratic value function of the form

$$V_n(\delta \mathbf{x}_n) = s_n + \delta \mathbf{x}_n^\top \mathbf{s}_n + \frac{1}{2} \delta \mathbf{x}_n^\top \mathbf{S}_n \delta \mathbf{x}_n \quad (9)$$

with weighting matrices $\mathbf{S}_n \in \mathbb{R}^{m \times m}$, $\mathbf{s}_n \in \mathbb{R}^{m \times 1}$ and $s_n \in \mathbb{R}$. The optimal control update can be derived by minimizing the value function V_n as a function of $\delta \mathbf{x}_n$.

As Equation (9) is quadratic in $\delta \mathbf{x}_{n+1}$ at time $n+1$, it remains quadratic during back-propagation in time, given the affine system dynamics and the linear-quadratic cost in Equations (6)-(8). Due to Bellman's Principle of Optimality, the optimal control $\delta \mathbf{u}_n^*$ at time n can be computed from

$$V_n^*(\delta \mathbf{x}_n) = \min_{\delta \mathbf{u}_n} \left[q_n + \delta \mathbf{x}^\top (\mathbf{q}_n + \frac{1}{2} \mathbf{Q}_n \delta \mathbf{x}_n) + \delta \mathbf{u}_n^\top \mathbf{P}_n \delta \mathbf{x}_n + \delta \mathbf{u}_n^\top (\mathbf{r}_n + \frac{1}{2} \mathbf{R}_n \delta \mathbf{u}_n) + V_{n+1}^*(\mathbf{A}_n \delta \mathbf{x}_n + \mathbf{B}_n \delta \mathbf{u}_n + \mathbf{d}_n) \right]$$

Inserting Equation (9) and the affine system dynamics (7) and minimizing the overall expression w.r.t. $\delta \mathbf{u}_n$ leads to an optimal control update of the form

$$\delta \mathbf{u}_n^* = -\mathbf{H}_n^{-1} \mathbf{h}_n - \mathbf{H}_n^{-1} \mathbf{G}_n \delta \mathbf{x}_n \quad (10)$$

where we have defined

$$\mathbf{h}_n = \mathbf{r}_n + \mathbf{B}_n^\top (\mathbf{s}_{n+1} + \mathbf{S}_{n+1} \mathbf{d}_n)$$

$$\mathbf{G}_n = \mathbf{P}_n + \mathbf{B}_n^\top \mathbf{S}_{n+1} \mathbf{A}_n$$

$$\mathbf{H}_n = \mathbf{R}_n + \mathbf{B}_n^\top \mathbf{S}_{n+1} \mathbf{B}_n$$

and $\mathbf{B}_n^\top \mathbf{S}_{n+1} \mathbf{B}_n + \mathbf{R}_n > 0$. After equating coefficients with a quadratic value function ansatz (9) for $\delta \mathbf{x}_n$, we define $\mathbf{l}_n = -\mathbf{H}_n^{-1} \mathbf{h}_n$ and $\mathbf{L}_n = -\mathbf{H}_n^{-1} \mathbf{G}_n$ and obtain the following recursive Riccati difference equations for \mathbf{S}_n , \mathbf{s}_n and s_n

$$\mathbf{S}_n = \mathbf{Q}_n + \mathbf{A}_n^\top \mathbf{S}_{n+1} \mathbf{A}_n - \mathbf{L}_n^\top \mathbf{H}_n \mathbf{L}_n \quad (11)$$

$$\mathbf{s}_n = \mathbf{q}_n + \mathbf{A}_n^\top (\mathbf{s}_{n+1} + \mathbf{S}_{n+1} \mathbf{d}_n) + \mathbf{G}_n^\top \mathbf{l}_n + \mathbf{L}_n^\top (\mathbf{h}_n + \mathbf{H}_n \mathbf{l}_n) \quad (12)$$

$$s_n = q_n + s_{n+1} + \mathbf{d}_n^\top \mathbf{s}_{n+1} + \frac{1}{2} \mathbf{d}_n^\top \mathbf{S}_{n+1} \mathbf{d}_n + \mathbf{l}_n^\top (\mathbf{h}_n + \frac{1}{2} \mathbf{H}_n \mathbf{l}_n) \quad (13)$$

for $n \in 0, \dots, N-1$. For the final time-step N we obtain the terminal conditions $\mathbf{S}_N = \mathbf{Q}_N$, $\mathbf{s}_N = \mathbf{q}_N$ and $s_N = q_N$, and the recursion is subsequently swept backwards. Note that Equation (13) does not contribute to the control update and can therefore be omitted in practice.

C. Updating State and Control Trajectories

Finally, using Equations (7) and (10), and readopting the superscript indices $[k]$ for the iteration count, we obtain

equations for a forward sweep resulting in a full-step update for the control and state decision variables $\mathbf{X}^{[k+1]}$, $\mathbf{U}^{[k+1]}$

$$\mathbf{u}_n^{[k+1]} = \mathbf{u}_n^{[k]} + \mathbf{l}_n^{[k]} + \mathbf{L}_n^{[k]} (\mathbf{x}_n^{[k+1]} - \mathbf{x}_n^{[k]}) \quad (14)$$

$$\mathbf{x}_{n+1}^{[k+1]} = \mathbf{x}_{n+1}^{[k]} + (\mathbf{A}_n^{[k]} + \mathbf{B}_n^{[k]} \mathbf{L}_n^{[k]}) (\mathbf{x}_n^{[k+1]} - \mathbf{x}_n^{[k]}) + \mathbf{B}_n^{[k]} \mathbf{l}_n^{[k]} + \mathbf{d}_n^{[k]} \quad (15)$$

with initial condition $\mathbf{x}_0^{[k+1]} = \mathbf{x}_{init}$. The updated decision variables are dynamically consistent w.r.t. the LQ subproblem dynamics. The nonlinear optimal control problem is solved iteratively, starting from Section II-A and solving LQ subproblems at each iteration, until convergence.

III. A FAMILY OF iLQR-GNMS ALGORITHMS

Equations (14) and (15) present the GNMS update rule where all states and controls (except for \mathbf{x}_{init}) are decision variables. For every time-step, both states and controls are updated using a linear forward sweep. Considering Equations (14) and (15), we can draw connections between GNMS and other existing algorithms and extend them to a family of 'hybrid' variants.

A. Connection to iLQR and Single Shooting

Interestingly, full-step iLQR employs the very same control update rule as in Equation (14). In fact, GNMS can be transcribed into iLQR by substituting the state update equation (15) with a numeric forward integration of the nonlinear system (2) using the time-varying state-feedback control law provided by Equation (14). In this case, the forward integration naturally results in a dynamically consistent state trajectory, all defects \mathbf{d}_n become zero and the formulation from section II-B drops back to the well-known iLQR Riccati recursion. Moreover, standard unconstrained single shooting can be recovered by additionally ignoring the state feedback gains and running the forward-integration purely open-loop.

B. Hybrid Algorithms

Consider a case where the overall time horizon N is split into an integer number of multiple shooting intervals M with length l and $1 < M < N$, while the control input discretization is kept at its original resolution. Without loss of generality, let us assume that the MS integration intervals start at time indices $i \in \mathcal{I}$, with $\mathcal{I} = \{0, l, 2l, \dots\}$. Fig. 1d sketches an example of such a hybrid case with $l = 3$. Every interval is simulated using the nonlinear system dynamics (2) and the initial states and controls $\mathbf{x}_i^{[k]}$ and $\mathbf{u}_i^{[k]}$. All $\mathbf{x}_j^{[k]}$ with $j \notin \mathcal{I}$ are *overwritten* by the integration. For an open-loop forward integration, $\mathbf{U}^{[k]}$ remains as is, but for a closed-loop forward integration, we additionally overwrite all $\mathbf{u}_j^{[k]}$ with $j \notin \mathcal{I}$ using the given feedback control law. Note that in this case, the defect equation (3) remains valid, but is zero along the multiple shooting intervals. The only non-zero defects occur at $\mathbf{d}_{i+l-1}^{[k]}$, $i \in \mathcal{I}$. In this setting, the LQ approximation, Riccati recursion and state- and control updates (14)-(15) can be performed as described before. This gives rise to two 'hybrid' GNMS variants:

TABLE I
AN OVERVIEW OF DIFFERENT GNMS-TYPE METHODS.

	SS	iLQR	GNMS	GNMS(M)	iLQR-GNMS(M)
closed-loop	–	✓	–	–	✓
No. intervals	1	1	N	M	M
overwrite states by integration	✓	✓	–	✓	✓
need stable initial policy	✓	✓	–	depends	depends

(✓ = true, – = false)

- *GNMS(M)*, using solely the feedforward control and thus performing an open-loop forward integration on each of the M multiple shooting intervals, which themselves are multiples of the control interval. Herewith, standard single shooting is the limit case GNMS(1).
- *iLQR-GNMS(M)*, using the full state feedback controller (14) and a closed-loop forward integration of each multiple shooting interval. In other words, this is equivalent to a multiple shooting variant of iLQR. The standard iLQR algorithm is the limit case of iLQR-GNMS(1), with only one multiple shooting interval.

Note that both GNMS(N) and iLQR-GNMS(N), with the number of multiple shooting-intervals being equal to the number of stages, revert to the standard GNMS formulation as introduced in Section II. Table I provides a compact overview of the algorithmic variants and compares their features.

C. Main Iteration and Implementation

We emphasize that all algorithmic variants feature linear complexity in the time horizon, $O(N)$. All algorithms execute almost identical linear algebra operations during one major iteration and therefore have very similar computational effort. Since the discussed family of GNMS algorithms only differs in a few features, it can be summarized in one framework, given in Algorithm 1. From a software-engineering perspective, the algorithmic variants are easy to implement and can all be treated at once, given a proper design of classes and interfaces. We provide an open-source C++ implementation of all discussed algorithms in [14].

D. Initialization

The GNMS variants listed in Table I differ in their requirements for initialization. For iLQR and SS, the nominal state (and control) input trajectories are first updated through a forward integration. This implies that for unstable systems, an initialization with a stabilizing initial control policy, which keeps the first rollout in the vicinity of the expected optimum, is essential. For iLQR, the initially provided state trajectory $\mathbf{X}^{[0]}$ serves as state reference trajectory for the feedback controller. For SS, it is irrelevant, except for the initial state. Common choices for SS and iLQR initial guesses are policies that stabilize the given initial state or draw the system towards the goal state, for example simple LQR or PD controllers. Generally, the increased efforts for initial guess design for SS and iLQR can be a significant disadvantage. In the worst

Algorithm 1 Generalized iLQR-GNMS(M) Algorithm

Given

- nonlinear dynamics, cost function and initial state \mathbf{x}_{init} as given in (1)-(2)
- initial state trajectory $\mathbf{X}^{[0]}$
- initial feedforward trajectory $\mathbf{U}^{[0]}$
- (initial feedback law, if applicable)
- maximum total constraint violation d_{max}
- minimum relative cost change J_{min}^{rel}

Prepare

- set iteration count $k = 0$
- split time horizon N into M MS integration intervals of length l , each starting at an index $i \in \mathcal{I}$, $\mathcal{I} = \{0, l, 2l, \dots\}$

Initial multiple shooting rollouts

- simulate M intervals using the nonlinear system dynamics (2) and initial states and controls $\mathbf{x}_i^{[0]}$ and $\mathbf{u}_i^{[0]}$, overwrite all $\mathbf{x}_j^{[0]}$ and $\mathbf{u}_j^{[0]}$ for $j \notin \mathcal{I}$
- compute defects $\mathbf{d}_n^{[0]}$ according to Equation (3).

Repeat (main iteration)

LQ approximation

- Linearize dynamics along the trajectories, obtain the affine constraint (5)
- Quadratize cost function along the trajectories to obtain (6)

Riccati backward sweep

- Backwards solve the Riccati-like difference equations (11)-(12) with boundary conditions $\mathbf{S}_N = \mathbf{Q}_N$ and $\mathbf{s}_N = \mathbf{q}_N$

Linear forward sweep

- compute state and control solution candidates $\mathbf{X}^{[k+1]}$ and $\mathbf{U}^{[k+1]}$ by forward sweeping Equations (14) and (15).

Rollout multiple shooting intervals

IF (open loop shooting)

- set feedback gain in Equation (14) to zero.

ENDIF

- simulate M shooting intervals using nonlin. dynamics (2), controller (14) and initial states and controls $\mathbf{x}_i^{[k+1]}$ and $\mathbf{u}_i^{[k+1]} \in \mathcal{I}$, overwrite all $\mathbf{x}_j^{[k+1]}$ and $\mathbf{u}_j^{[k+1]}$ for $j \notin \mathcal{I}$
- compute defects $\mathbf{d}_n^{[k+1]}$ according to Equation (3)
- compute cost $J^{[k+1]}$ by evaluating Equation (1)
- increment iteration count k

until $|J^{[k]} - J^{[k-1]}|/J^{[k]} < J_{min}^{rel}$ and $\sum |\mathbf{d}_n^{[k]}| < d_{max}$

case, a poor initial guess can lead to a local minimum with a solution far from desired behavior.

Multiple shooting algorithms, by contrast, offer greater flexibility and simplicity at initial guess design, and are often more robust w.r.t. bad initial policies. It is a well known fact that the convergence of multiple shooting methods can be accelerated through an ‘educated’ initial guess, such as direct interpolation between initial and desired final state. For the hybrid algorithms iLQR-GNMS(M) and GNMS(M) it often depends on the system characteristics if a stabilizing control policy is required, or if the multiple shooting intervals are short enough to prevent significant divergence during integration. In the video attachment [15], we show two simulation examples where initialization with a bad state-feedback controller significantly extends the runtime of iLQR compared to GNMS, or even causes iLQR to fail.

Note that, when all possible GNMS variants are initialized with dynamically consistent state and control trajectories, the defects for the first iteration are zero and the feedforward control updates are identical.

IV. RESULTS AND COMPARISON

A. An Illustrative, One-Dimensional System

As an illustrative example, we present a simple one-dimensional system, which is slightly nonlinear, unstable and constructed to help the reader build an intuition about the methods. The system dynamics are $\dot{x} = (1+x)x + u$, $x(0) = 1.5$ and discretized with $\Delta t = 0.01$ s, $N = 300$. The cost function is defined as quadratic cost of form (6) with

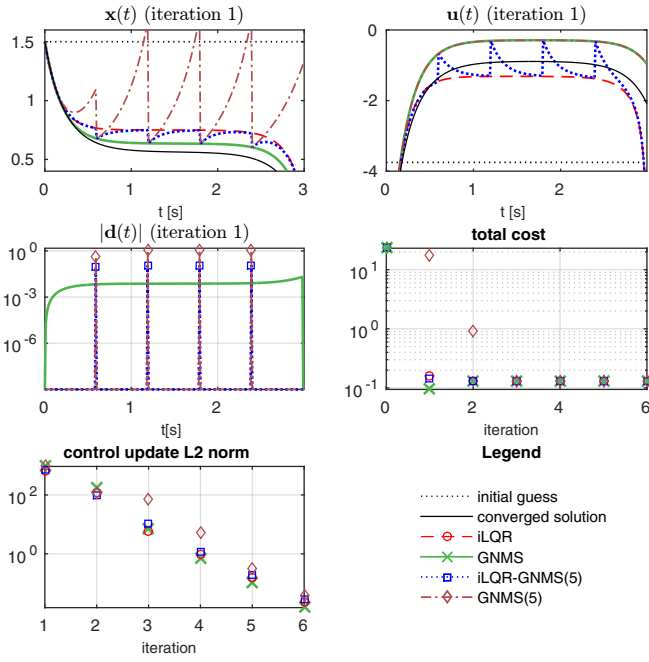


Fig. 2. Results for a one-dimensional illustrative system, main text for detailed description.

desired terminal state $\mathbf{x}_N^{des} = 0$, $\mathbf{Q}_N = 10$ and $\mathbf{R}_n = 0.01$. Fig. 2 shows results for iLQR, GNMS and the hybrid variants with five multiple shooting intervals, GNMS(5) and iLQR-GNMS(5). We plot the state, control and defect trajectories for the first iteration of the algorithms, along with the initial guess and converged solution.

The state and control trajectories illustrate the relationship between the algorithms: the multiple shooting intervals of GNMS(5) and iLQR-GNMS(5) start with states and controls lying on the respective GNMS trajectories. For GNMS(5), the system is simulated open-loop, the controls are identical to GNMS, and the state trajectories on the multiple shooting intervals start to diverge. By contrast, for iLQR-GNMS(5), in every multiple shooting interval both state and control trajectories converge asymptotically towards the simulated iLQR state and control trajectories. For the hybrid variants, a defect occurs every 0.6s, for GNMS, the defect is evenly distributed across all time intervals. Due to the long shooting intervals, GNMS(5) requires one iteration more to catch up with the other algorithms in terms of overall cost. Importantly, the control update plot shows that the asymptotic contraction rates, which are defined as

$$C = \lim_{k \rightarrow \infty} \frac{\|\mathbf{U}^{[k+1]} - \mathbf{U}^*\|_2}{\|\mathbf{U}^{[k]} - \mathbf{U}^*\|_2} \quad (16)$$

are not the same. In this example, GNMS and GNMS(5) show better contraction than iLQR. Asymptotic contraction rates are investigated in more detail in Section IV-C.

B. Quadruped Trot Optimization Example

The quadrupedal robot ‘HyQ’ [16] is an 18 DoF, floating-base underactuated robot subject to contacts with the environment, c.f. Fig. 3. In this paper, the contacts are not

incorporated as constraints, but added to the system dynamics using an explicit contact model. We employ a static, plain environment and a ‘soft’ contact model, consisting of a nonlinear spring in surface-normal direction and a nonlinear damping term. The contact model is detailed in [17]. Using this formulation, the contact force is a function of the current robot state only. It is clear that such a soft contact model presents only a rough approximation of the complicated physics of contact, and also introduces a number of potential disadvantages such as increased stiffness and nonlinearity of the combined system dynamics. However, the contact model allows a straight-forward computation of derivatives [17], which creates an ideal test-bed for comparing our shooting algorithms. We obtain exact discrete sensitivities \mathbf{A}_n and \mathbf{B}_n through evaluating a sensitivity differential equation on the multiple shooting intervals [18].

The example task considered is the optimization of a periodic trotting gait. To achieve the trotting gait, we impose a time-varying quadratic penalty on the leg joint positions. Furthermore, we penalize the intermediate and final position of the robot’s trunk and the intermediate and final velocities of the leg joints. For an in-depth description of the cost modelling to achieve different gait patterns the reader is referred to [8]. In the following, the trotting gait optimization is used to compare the algorithms developed in this paper. For a meaningful comparison, we initialize all algorithms with identical state trajectories and control policies. The initial guess corresponds to standing still in a steady state. We optimize over 36 states, 12 control inputs and a total time horizon of 2.5 seconds with $N = 2500$.

Fig. 4 compares iLQR, GNMS, and iLQR-GNMS(M) with three different numbers of multiple shooting intervals in terms of cost descent, control update norms and total defects. Note that SS and GNMS(M) are unstable due to the strong instability of the system. All remaining algorithms converge to the same minimum within 20 iterations. iLQR and iLQR-GNMS(25) show short phases of increasing cost, which we accept in this simulation example. Since the provided initial guess is dynamically consistent, the initial defects are zero. GNMS, having the largest number of multiple shooting intervals (2500), also shows the largest total defect sum after the first iteration. The multiple shooting iLQR methods, all having significantly fewer continuity constraints to enforce, feature a lower total defect.

As expected for a Gauss-Newton method, all approaches show linear convergence. Considering the control update norms, we see that GNMS and iLQR feature a similar contraction rate for this example. In fact, the contraction rate of GNMS is slightly better, which is visually hard to distinguish here, but is detailed in following example. For the hybrid multiple shooting iLQR variants, we observe a significantly better contraction rate than for both iLQR and

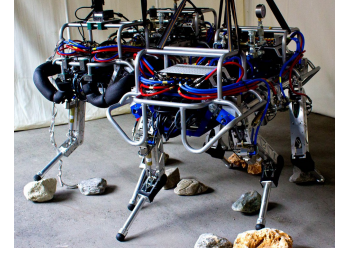


Fig. 3. The quadruped HyQ

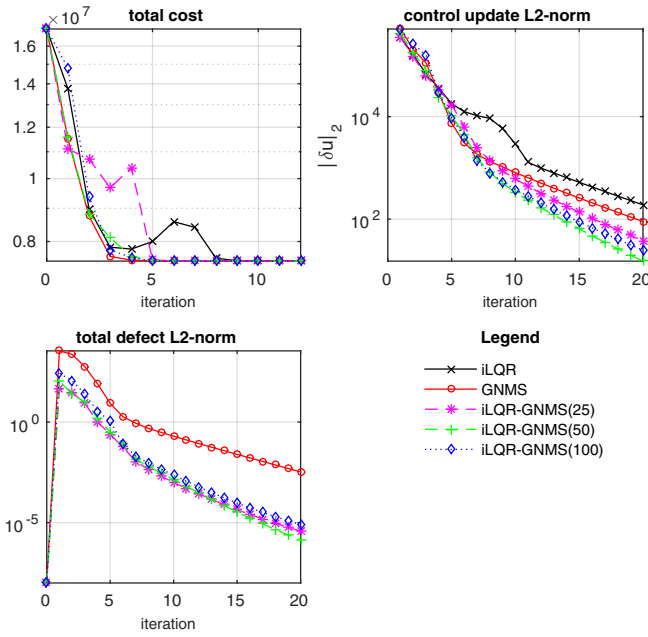


Fig. 4. Comparing different algorithms from the family of Gauss-Newton shooting methods w.r.t. cost descent, total defects and control update norm, illustrated with the example of a quadruped trot optimization problem.

GNMS. When applying an identical termination criterion based on the relative change of the cost function and a defect threshold to all algorithms, all lifted methods converge in fewer iterations than iLQR. Furthermore, all displayed iLQR-GNMS(M) variants converge notably faster than GNMS. Screen recordings of the optimized trotting motions are provided in the video attachment [15].

C. Local Contraction Rates for Quadruped Trot Tracking

While Section IV-B gives an optimization example for a single motion, starting with an initial guess far from the optimal solution, we now show a comparison based on statistical data from 1000 runs: the trotting gait from Section IV-B is now considered in a tracking MPC problem. All algorithms are initialized with an optimal, dynamically consistent solution, but the initial state is locally perturbed. The state perturbations are sampled from the hardware-experiments detailed in [19]. For every perturbation, we let different algorithms iterate until convergence. Fig. 5 compares average asymptotic contraction rates for four different algorithms. It shows the normalized difference between a fully converged optimal feedforward trajectory and trajectories obtained at previous iterations. Furthermore it shows first-order regressions approximating the local contraction rates, in terms of the slopes of the difference norms in the semi-logarithmic plot. It can be seen that GNMS outperforms iLQR in terms of local contraction rate. GNMS(50) shows a contraction rate similar to GNMS. The example indicates better local convergence for iLQR-GNMS(50) than for classical iLQR, GNMS and GNMS(50).

Fig. 6 generalizes the result from Fig. 5 for a range of multiple shooting intervals M , showing numerically approximated asymptotic contraction rates, Equation (16), as a

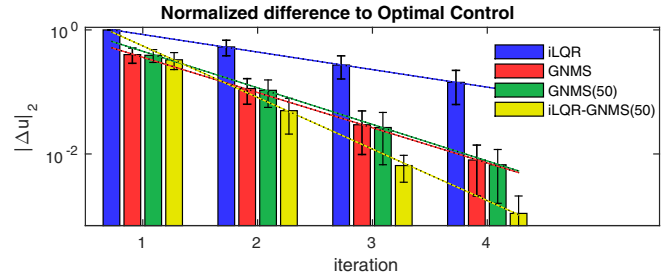


Fig. 5. A statistical comparison of local contraction rates for four different algorithms. This plot shows the normalized difference between solutions to perturbed optimal control problems and a fully converged reference solution. The normalization is w.r.t. the iLQR control trajectory. The data is averaged from 1000 samples, the corresponding standard deviations are plotted as error-bars. Estimates for the contraction rates are indicated by straight lines. GNMS outperforms iLQR in terms of local contraction rate, and gets closer to the true optimal solution in fewer iterations. In contrast to the example in Section IV-B, GNMS(50) is stable and shows a contraction rate similar to GNMS. iLQR-GNMS(50) clearly outperforms all other algorithms with significantly better local contraction rate. After 4 iterations, it is on average 0.1% away from the fully converged reference solution, while iLQR is on average 14% away.

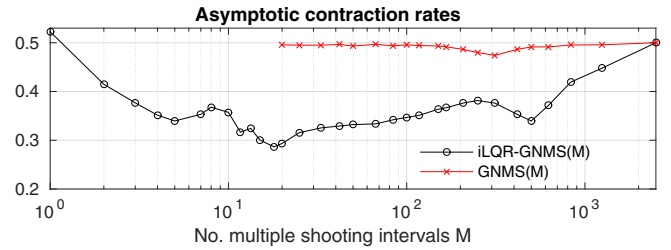


Fig. 6. Comparing asymptotic contraction rates as a function of the number of multiple shooting intervals M . For GNMS(M), the integration on the multiple shooting intervals is only stable for $M \geq 20$. For this range of 'stable' multiple shooting intervals, the contraction rates are almost constant and similar to the limiting case GNMS. The iLQR multiple shooting methods are stable for all M . Here, the contraction rates for intermediate numbers of multiple shooting intervals are distinctively better than for the limit cases iLQR and GNMS.

function of M . Again, GNMS(M) is unstable for overly long multiple shooting intervals, similar to the limiting case open-loop single shooting. For closed-loop shooting, the asymptotic contraction rates for all multiple shooting variants are better than for iLQR, and the contraction rates for the hybrid variants outperform the limiting case GNMS. In this example, the relative improvement over iLQR is up to a factor two. Note that the resulting iLQR-GNMS(M) contraction rates differ slightly from the ones in Fig. 4 for $M = 25, 50, 100$, which is due to the different problem setting. However, both experiments exhibit the same trend.

D. Nonlinear MPC on HyQ

The suitability of iLQR for nonlinear model-predictive control (NMPC) in robotics applications has been shown many times before, in [6], [20], [21]. In this section, we show that GNMS and its hybrid variants are even more promising for NMPC applications. First, they converge faster to the optimal solution, c.f. Section IV-C. A second advantage of the multiple shooting variants of the presented algorithms is that the forward integrations can be parallelized. Therefore, the

Algorithm 2 iLQR-GNMS(M)-NMPC Algorithm

Given

- cost function (1) and system dynamics (2).
- receding MPC time horizon N
- number of multiple shooting intervals M with length l
- initial state and control trajectories $\mathbf{X} = \{\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_N\}$
 $\mathbf{U} = \{\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_{N-1}\}$, state-feedback controller $\mathbf{u}_n(\mathbf{x})$ of form (10)

Repeat Online:

Feedback phase

- get state measurement \mathbf{x}_{meas} .
- forward integrate system dynamics (2) with $\mathbf{x}_0 = \mathbf{x}_{meas}$ on the first multiple shooting interval, compute $\mathbf{A}_{0,\dots,l-1}$, $\mathbf{B}_{0,\dots,l-1}$, defect \mathbf{d}_{l-1} .
- quadratize cost function (1) around \mathbf{X} and \mathbf{U} for control stages $1, \dots, l-1$.
- solve LQ optimal control problem using a Riccati backward sweep
- retrieve updated control policy $\mathbf{u}_n^+(\mathbf{x})$ and updated trajectories \mathbf{U}^+ , \mathbf{X}^+ .
- send policy $\mathbf{u}_n^+(\mathbf{x})$ and \mathbf{X}^+ to the control system

Preparation phase

- update: $\mathbf{u}_n(\mathbf{x}) \leftarrow \mathbf{u}_n^+(\mathbf{x})$, $\mathbf{X} \leftarrow \mathbf{X}^+$, $\mathbf{U} \leftarrow \mathbf{U}^+$
- forward integrate system dynamics (2) for the multiple shooting intervals 1 to M , obtain sensitivities $\mathbf{A}_1, \dots, \mathbf{A}_{N-1}$, $\mathbf{B}_1, \dots, \mathbf{B}_{N-1}$ and defects $\mathbf{d}_1, \dots, \mathbf{d}_{N-1}$
- quadratize cost function (2) around \mathbf{X} , \mathbf{U} for multiple shooting intervals l to N .

achievable MPC cycle time decreases approximately linearly with the number of available CPU cores. By combining faster update rates with better contraction rate, our multiple shooting algorithms outperform classical iLQR-NMPC.

In the following simulation example, we compare the NMPC performance of our Gauss-Newton shooting algorithms against iLQR-NMPC in a HyQ simulation environment. In each NMPC cycle, we run an adapted version of the main iteration in Algorithm 1 and ‘warm-start’ it with the previous solution. In such a setting, we can separate an NMPC iteration into a ‘preparation’ and a ‘feedback’ phase [22], thus minimizing the latency between receiving a state-measurement and sending an updated policy to the control system. Our NMPC loop is described in Algorithm 2.

In this experiment, we run a trotting gait on HyQ, in closed-loop MPC in a simulation environment [23]. For the NMPC optimal control problem, we choose a time-step size of 4 ms and $N = 125$. We parallelize the integration of all multiple shooting intervals and the sensitivity computation on four cores, and run both simulator and MPC controller on the same notebook equipped with an Intel Core i7 (2.8 GHz) processor. For four different algorithmic combinations, we record the executed trot under identical conditions for 18 seconds and compute the resulting accumulated intermediate cost. A summary of the achieved average cost and NMPC frequencies is given in Fig. 7. In these experiments, iLQR results in the highest accumulated cost. The multiple shooting variants outperform iLQR, with relative cost differences up to 5%. At the same time, the multiple shooting variants achieve up to 40% higher MPC frequencies, with a maximum of 103 Hz.

In our simulation, all four algorithm variants run in a stable and robust fashion. The relatively small cost difference is an indicator of better convergence, but the main reason why the multiple shooting variants should be preferred over iLQR in real-world applications, is the superior control bandwidth. The algorithms in this paper have been validated in hardware experiments on two different quadruped platforms, where a variety of motions and gaits was implemented. However, a

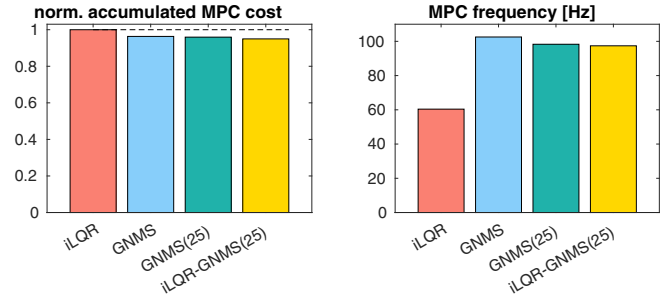


Fig. 7. Comparing four different solvers for HyQ-MPC. The left plot compares costs accumulated over 18.0 seconds of trotting-MPC execution. iLQR shows the highest accumulated cost. The multiple shooting variants outperform iLQR, but the relative cost difference is small. GNMS runs at 96.3% of the iLQR cost, GNMS(25) at 95.9% and iLQR-GNMS(25) at 95.0%. Regarding the achieved average MPC frequencies, shown on the right, we find up to 40% higher frequencies for the multiple shooting variants. The maximum average MPC frequency, 103 Hz, is achieved using GNMS.

in-depth description of the experimental setups, the optimized computational framework and practical tuning considerations are beyond the scope of this paper. The interested reader is therefore referred to [19], where we apply the GNMS-algorithm for full-body NMPC on the quadrupeds HyQ and ANYmal [24], explain the robotic setup in detail and present a variety of hardware experiments. As outlook, a video sequence of GNMS-NMPC running on hardware is provided in the attachment [15].

V. CONCLUSION AND OUTLOOK

In this paper, we have shown how the well-known iLQR algorithm can be lifted and transformed into Gauss-Newton Multiple Shooting, GNMS. We have generalized the concept to form a family of Gauss-Newton shooting algorithms, which can be distinguished into sequential and simultaneous algorithms and closed and open-loop algorithms. Some algorithms partially overwrite decision variables by means of a numeric forward integration. All presented variants have approximately the same computational cost and feature linear time-complexity. Furthermore, all discussed algorithms share a large number of computational routines, and it is not difficult to implement all of the presented variants in a single software framework.

We have compared the performance of the algorithms in different simulation experiments, which indicate that the lifted algorithms can outperform classical iLQR. While not included in this paper for reasons of compactness, similar results were obtained for other rigid-body dynamic systems including a 6 DoF fixed-base arm model. A more fundamental investigation for formalizing the conditions that result in improved convergence rates for GNMS(M) and iLQR-GNMS(M) is subject to ongoing work.

In the application examples, we limited the comparison to full-step variants of all considered algorithms. However, for even more nonlinear dynamics or cost functions, where the LQ optimal control problem is a bad approximation to the nonlinear problem, a globalization strategy may be required. For single shooting methods, a straight-forward solution is to

employ a line-search scheme. This is simple to implement, as it is sufficient to search over the cost for different control update step-sizes. For multiple shooting approaches, however, there are additional continuity constraints, and we need to line-search over a merit-function which trades off the costs and defects. It is typically required to introduce additional tuning variables or to implement a complex filter-scheme [25]. Our open-source reference implementation [14] provides a line-search scheme using a simple merit function.

For complex robot trajectory optimization problems, we do not recommend to generally prioritize one of the presented algorithms over another. While the multiple shooting algorithms allow for advanced initialization strategies and are more robust w.r.t. bad initial guesses, they may require slightly more tuning efforts when the full-step algorithm is not sufficient. By contrast, in NMPC applications with well-defined cost functions and using warm-starting, additional globalization steps are rarely required at all. Here, our multiple shooting algorithms offer significant advantages, better local contraction rates and much shorter runtimes.

The focus of this paper is on unconstrained optimal control problems without general (in)equality path constraints. It is obvious that the lifting approach naturally transfers to equality-constrained variants of iLQR, such as [26]. The inclusion of general (in)equality path constraints is part of ongoing work. One option to include them in the existing framework is to replace the standard Riccati backward sweep with a dedicated solver for constrained LQ optimal control problems [27]. In this way, general (in)equality path constraints can be included while keeping linear time-complexity.

While this work treats algorithms using a Gauss-Newton Hessian approximation, it similarly transfers to exact-Hessian approaches, resulting in a multiple shooting DDP algorithm combining the advantages of simultaneous methods, quadratic convergence and closed-loop integration.

ACKNOWLEDGEMENTS

The authors would like to thank Dimitris Kouzoupis, Gianluca Frison, Mario Zanon and Timothy Sandy for fruitful discussions. This research was supported by the Swiss National Science Foundation through the NCCR Digital Fabrication, the NCCR Robotics and a Professorship Award to Jonas Buchli. Further, this research was supported by the EU via FP7-ITN-TEMPO (607 957) and H2020-ITN-AWESCO (642 682), by the Federal Ministry for Economic Affairs and Energy (BMWi) via eco4wind and DyConPV, and by DFG via Research Unit FOR 2401.

REFERENCES

- [1] D. Mayne, "A second-order gradient method for determining optimal trajectories of non-linear discrete-time systems," *International Journal of Control*, vol. 3, no. 1, pp. 85–95, 1966.
- [2] H. G. Bock and K.-J. Plitt, "A multiple shooting algorithm for direct solution of optimal control problems," in *Proceedings of the IFAC World Congress*, 1984.
- [3] M. Diehl, H. G. Bock, H. Diedam, and P.-B. Wieber, "Fast direct multiple shooting algorithms for optimal robot control," in *Fast motions in biomechanics and robotics*, pp. 65–93, Springer, 2006.
- [4] J. Albersmeyer and M. Diehl, "The Lifted Newton Method and its Application in Optimization," *SIAM Journal on Optimization*, vol. 20, no. 3, pp. 1655–1684, 2010.
- [5] C. V. Rao, S. J. Wright, and J. B. Rawlings, "Application of interior-point methods to model predictive control," *Journal of Optimization Theory and Applications*, vol. 99, pp. 723–757, Dec 1998.
- [6] J. Koenemann, A. Del Prete, Y. Tassa, E. Todorov, O. Stasse, M. Bennewitz, and N. Mansard, "Whole-body model-predictive control applied to the HRP-2 humanoid," in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 3346–3351, 2015.
- [7] Y. Tassa, T. Erez, and E. Todorov, "Synthesis and stabilization of complex behaviors through online trajectory optimization," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 4906–4913, Oct 2012.
- [8] 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 (RA-L)*, vol. 2, pp. 1502–1509, July 2017.
- [9] B. Ponton, S. Schaal, and L. Righetti, "Risk sensitive nonlinear optimal control with measurement uncertainty," *CoRR*, 2016.
- [10] F. Farshidian, M. Neunert, A. W. Winkler, G. Rey, and J. Buchli, "An efficient optimal planning and control framework for quadrupedal locomotion," in *IEEE International Conference on Robotics and Automation (ICRA)*, pp. 93–100, May 2017.
- [11] E. Todorov and W. Li, "A generalized iterative LQG method for locally-optimal feedback control of constrained nonlinear stochastic systems," in *American Control Conference, 2005. Proceedings of the 2005*, pp. 300–306, IEEE, 2005.
- [12] A. Sideris and J. E. Bobrow, "An efficient sequential linear quadratic algorithm for solving nonlinear optimal control problems," *Transactions on Automatic Control*, vol. 50, no. 12, pp. 2043–2047, 2005.
- [13] A. Schäfer, *Efficient Reduced Newton-Type Methods for Solution of Large-Scale Structured Optimization Problems with Application to Biological and Chemical Processes*. PhD thesis, University of Heidelberg, 2005.
- [14] M. Gifftthaler, M. Neunert, M. Stäuble, and J. Buchli, "The Control Toolbox - an open-source C++ library for robotics, optimal and model predictive control," in *2018 IEEE International Conference on Simulation, Modeling, and Programming for Autonomous Robots (SIMPAP)*, pp. 123–129, May 2018.
- [15] "Video attachment," https://youtu.be/423BTKn_cjQ.
- [16] C. Semini, N. G. Tsagarakis, E. Guglielmino, M. Focchi, F. Cannella, and D. G. Caldwell, "Design of HyQ—a hydraulically and electrically actuated quadruped robot," *Institution of Mechanical Engineers, Journal of Systems and Control Engineering*, vol. 225, pp. 831–849, 2011.
- [17] M. Gifftthaler, M. Neunert, M. Stäuble, M. Frigerio, C. Semini, and J. Buchli, "Automatic Differentiation of Rigid Body Dynamics for Optimal Control and Estimation," *Advanced Robotics*, vol. 31, pp. 1225–1237, November 2017.
- [18] R. P. Dickinson and R. J. Gelinas, "Sensitivity analysis of ordinary differential equation systems - a direct method," *Journal of Computational Physics*, vol. 21, no. 2, pp. 123 – 143, 1976.
- [19] M. Neunert, M. Stuble, M. Gifftthaler, C. D. Bellicoso, J. Carius, C. Gehring, M. Hutter, and J. Buchli, "Whole-body nonlinear model predictive control through contacts for quadrupeds," *IEEE Robotics and Automation Letters*, vol. 3, pp. 1458–1465, July 2018.
- [20] M. Neunert, C. de Crousaz, F. Furrer, M. Kamel, F. Farshidian, R. Siegwart, and J. Buchli, "Fast nonlinear model predictive control for unified trajectory optimization and tracking," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2016.
- [21] M. Gifftthaler, F. Farshidian, T. Sandy, L. Stadelmann, and J. Buchli, "Efficient Kinematic Planning for Mobile Manipulators with Non-holonomic Constraints Using Optimal Control," in *IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3411–3417, May 2017.
- [22] M. Diehl, H. Bock, J. P. Schlöder, R. Findeisen, Z. Nagy, and F. Allgöwer, "Real-time optimization and nonlinear model predictive control of processes governed by differential-algebraic equations," *Journal of Process Control*, vol. 12, no. 4, pp. 577 – 585, 2002.
- [23] S. Schaal, "The SL simulation and real-time control software package," tech. rep., Los Angeles, CA, 2009.
- [24] M. Hutter, C. Gehring, D. Jud, A. Lauber, C. D. Bellicoso, V. Tsounis, J. Hwangbo, K. Bodie, P. Fankhauser, M. Bloesch, R. Diethelm, S. Bachmann, A. Melzer, and M. Hoepflinger, "AnyMal - a highly mobile and dynamic quadrupedal robot," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 38–44, Oct 2016.
- [25] J. Nocedal and S. J. Wright, *Numerical Optimization*. Springer, 1999.
- [26] A. Sideris and L. A. Rodriguez, "A Riccati approach for constrained linear quadratic optimal control," *International Journal of Control*, vol. 84, no. 2, pp. 370–380, 2011.
- [27] G. Frison, *Algorithms and Methods for Fast Model Predictive Control*. PhD thesis, Technical University of Denmark, 2015.