



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

## A feasibility-driven approach to control-limited DDP

**Citation for published version:**

Mastalli, C, Merkt, W, Marti-Saumell, J, Ferrolho, H, Solà, J, Mansard, N & Vijayakumar, S 2022, 'A feasibility-driven approach to control-limited DDP', *Autonomous Robots*, vol. 46, no. 8, pp. 985-1005. <https://doi.org/10.1007/s10514-022-10061-w>

**Digital Object Identifier (DOI):**

[10.1007/s10514-022-10061-w](https://doi.org/10.1007/s10514-022-10061-w)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Publisher's PDF, also known as Version of record

**Published In:**

Autonomous Robots

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.





# A feasibility-driven approach to control-limited DDP

Carlos Mastalli<sup>1</sup> · Wolfgang Merkt<sup>3</sup> · Josep Marti-Saumell<sup>4</sup> · Henrique Ferrolho<sup>2</sup> · Joan Solà<sup>4</sup> · Nicolas Mansard<sup>5</sup> · Sethu Vijayakumar<sup>2</sup>

Received: 23 March 2022 / Accepted: 22 August 2022 / Published online: 21 September 2022  
© The Author(s) 2022

## Abstract

Differential dynamic programming (DDP) is a direct single shooting method for trajectory optimization. Its efficiency derives from the exploitation of temporal structure (inherent to optimal control problems) and explicit roll-out/integration of the system dynamics. However, it suffers from numerical instability and, when compared to direct multiple shooting methods, it has limited initialization options (allows initialization of controls, but not of states) and lacks proper handling of control constraints. In this work, we tackle these issues with a feasibility-driven approach that regulates the dynamic feasibility during the numerical optimization and ensures control limits. Our feasibility search emulates the numerical resolution of a direct multiple shooting problem with only dynamics constraints. We show that our approach (named BOX-FDDP) has better numerical convergence than BOX-DDP<sup>+</sup> (a single shooting method), and that its convergence rate and runtime performance are competitive with state-of-the-art direct transcription formulations solved using the interior point and active set algorithms available in KNITRO. We further show that BOX-FDDP decreases the dynamic feasibility error monotonically—as in state-of-the-art nonlinear programming algorithms. We demonstrate the benefits of our approach by generating complex and athletic motions for quadruped and humanoid robots. Finally, we highlight that BOX-FDDP is suitable for model predictive control in legged robots.

**Keywords** Optimal control · Differential dynamic programming · Feasibility · Direct multiple shooting · Control limits

---

This research was supported by (1) the European Commission under the Horizon 2020 Project Memory of Motion (MEMMO, Project ID: 780684), (2) the Engineering and Physical Sciences Research Council (EPSRC) UK RAI Hub for Offshore Robotics for Certification of Assets (ORCA, Grant reference EP/R026173/1), and (3) the Alan Turing Institute.

---

✉ Carlos Mastalli  
c.mastalli@hw.ac.uk

Wolfgang Merkt  
wolfgang@robots.ox.ac.uk

Josep Marti-Saumell  
jmarti@iri.upc.edu

Henrique Ferrolho  
henrique.ferrolho@ed.ac.uk

Joan Solà  
jsola@iri.upc.edu

Nicolas Mansard  
nicolas.mansard@laas.fr

Sethu Vijayakumar  
sethu.vijayakumar@ed.ac.uk

## 1 Introduction

Optimal control is a powerful tool to synthesize motions and controls through task goals (cost/optimality) and constraints (e.g., system dynamics, interaction constraints). We can formulate such problems using *direct methods* (Betts, 2009), which first discretize over both state and controls, and then optimize using sparse general-purpose nonlinear programming (NLP) software such as SNOPT (Gill et al., 2005), KNITRO (Byrd et al., 2006), and IPOPT (Wächter & Biegler, 2006). To ensure that the system state evolves as

- <sup>1</sup> Institute of Sensors, Signals and Systems, School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh, UK
- <sup>2</sup> School of Informatics, University of Edinburgh, Edinburgh, UK
- <sup>3</sup> Oxford Robotics Institute, University of Oxford, Oxford, UK
- <sup>4</sup> Institut de Robòtica i Informàtica Industrial, Universitat Politècnica de Catalunya, Barcelona, Spain
- <sup>5</sup> LAAS-CNRS, Toulouse, France

described by the equations of motion, we define equality constraints in the NLP problem. However, common algorithms for nonlinear programming can only guarantee the constraint satisfaction at their convergence. When this occurs, we say that the discretized states are *dynamically* feasible. Furthermore, despite the advantage of using advanced software for nonlinear programming, these algorithms perform very large matrix factorizations during the computation of the search direction, i.e., the resolution of the Karush-Kuhn-Tucker (KKT) problem. To do so, they use sparse linear solvers such as MA27, MA57, and MA97 (see <http://www.hsl.rl.ac.uk/>) that do not exploit the temporal/Markovian structure of optimal control problems efficiently. Indeed, the expensive factorizations of these linear solvers limit their practical use to realtime control on reduced models [e.g., Wieber (2006), Pardo et al. (2016), Mastalli et al. (2017), Di Carlo et al. (2018)] or motion planning [e.g., Carpentier et al. (2016), Aceituno-Cabezas et al. (2017), Winkler et al. (2018), Merkt et al. (2018), Mastalli et al. (2020)] in robotics. Furthermore, classical line search methods used in general-purpose NLP solvers are less effective than the nonlinear roll-out of the dynamics used in shooting methods as their use increases the number of iterations [cf. Liao and Shoemaker (1992)], which is used in recent method for multiple shooting (Gifftthaler et al., 2018).

*Dynamic programming methods*, which have their foundations in the calculus of variations as indirect methods, have once again attracted attention due to recent results on fast nonlinear model predictive control based on DDP [e.g., Tassa et al. (2012), Koenemann et al. (2015), Neunert et al. (2018), Farshidian et al. (2017)]. In particular, there is a significant interest in the iterative linear-quadratic regulator (iLQR) algorithm (Li & Todorov, 2004) as its Gauss-Newton (GN) approximation reduces the computation time while having super-linear convergence. Both iLQR and DDP algorithms perform a Riccati sweep in the backward pass, which incorporates elements that are reminiscent of Pontryagin's maximum principle (PMP). For instance, at convergence, the gradient of the value function in the backward pass represents the *costate*; instead, the roll-out of the system dynamics describes the *state integration* step. This connection was recognized by Bellman's groundbreaking work (Bellman, 1954) that established the so-called Hamilton-Jacobi-Bellman (HJB) equation in the continuous-time domain. In contrast to classical direct collocation approaches, these approaches exploit the temporal/Markovian structure of the optimal control problem by solving a sequence of smaller sub-problems derived from Bellman's principle of optimality (Mayne, 1966). This leads to fast and cheap computations due to very small matrix factorizations and effective data cache accesses. Despite these advantages, both algorithms are unable to handle equality and inequality constraints efficiently. Furthermore, they have a poor basin

of attraction for a good local optimum as it requires a good initialization in order to converge and are prone to numerical instability—commonly recognized challenges for single shooting approaches (Betts, 2009). These undesirable properties are mainly due to the fact that the iLQR/DDP algorithms implicitly enforce the dynamic feasibility through the system roll-out.

## 1.1 Related work

Trade-offs between feasibility and optimality appear in most of the state-of-the-art nonlinear programming software. For instance, IPOPT includes a feasibility restoration phase which aims at reducing the constraint violation (Wächter & Biegler, 2006). In KNITRO, the progress on both feasibility and optimality is achieved by adding an  $\ell^1$ -norm penalty term for the constraints in the *merit* function (Byrd et al., 2006). In fact, by changing the merit function or the line search procedure, we can put emphasis on obtaining feasible solutions *before* trying to optimize them. Instead, the iLQR/DDP algorithms do not make this trade-off, as the backward and forward passes do not accept infeasible iterations. However, recent work on *multiple shooting DDP* (Gifftthaler et al., 2018; Mastalli et al., 2020) has provided ways of handling dynamically infeasible iterations, which we elaborate below.

The multiple shooting variants in Gifftthaler et al. (2018), Mastalli et al. (2020) are rooted in dynamic programming. For instance, Gifftthaler et al. (2018) introduced a *lifted*<sup>1</sup> version of the algebraic Riccati equation that allows initialization of both state and control trajectories; it further accounts for the relinearization required by the dynamics gaps in the backward pass and uses a merit function to balance feasibility and optimality. In turn, in our previous work (Mastalli et al., 2020), we proposed a modification of the forward pass that numerically matches the gap contraction expected by a *direct multiple shooting* method subject to equality constraints only. It factorizes the KKT matrix via a Riccati recursion and defines the behavior of the defect constraints based on the first-order necessary condition (FONC) of optimality.<sup>2</sup> These approaches improve numerical robustness against poor initialization, as they are able to use an initial guess for the state trajectory. Unfortunately, none of these methods handle inequality constraints such as control limits, with the exception of a recent work that computes squashed control sequences (Marti-Saumell et al., 2020).

There are two main strategies for incorporating arbitrary constraints: active set and penalization methods (as extensively described in Nocedal and Wright (2006)). In the

<sup>1</sup> This name is coined by Albersmeyer and Diehl (2010), and we refer to *gaps* or *defects* produced between multiple shooting nodes.

<sup>2</sup> For more details about the FONC of optimality see Nocedal and Wright (2006).

robotics community, one of the first successful attempts to incorporate inequality constraints in DDP used an active set approach (Tassa et al., 2014), which is based on Ohno (1978)—a pioneering work in the control community. Concretely, this approach focused on handling control limits during the computation of the backward pass, i.e., in the minimization of the action-value function ( $Q$ -function),<sup>3</sup> which resembles the control Hamiltonian at convergence [see Kirk (1998), Section 3.11]. The method is popularly named BOX-DDP, and the authors also showed a better convergence rate when compared with a squashing function approach. Later, Xie et al. (2017) included general inequality constraints into the  $Q$ -function and the forward pass. The method sacrifices the computational performance by including a second quadratic program, which is solved in the forward pass. However, it still remains faster than solving the same problem using direct collocation with SNOPT as reported in Hargraves and Paris (1987).

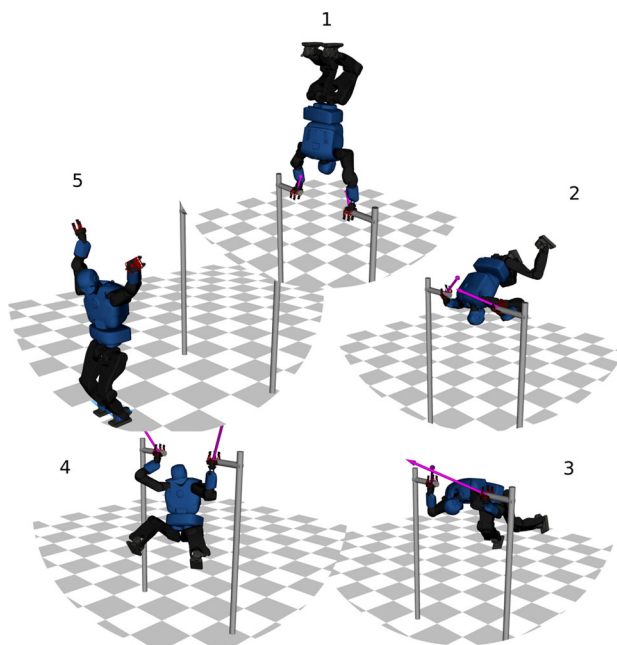
Generally speaking, active set methods are suitable for small-size problems (such as minimizing the  $Q$ -function described above) as their accuracy and speed often outperform other methods. However, the combinatorial complexity of finding the active set is prohibitive in large-scale optimization problems. This motivates the development of penalty-based methods, despite their numerical difficulties: ill-conditioning and slow convergence. To overcome these difficulties, Lantoine and Russell (2008) proposed a method that incorporates an augmented Lagrangian term. This method was studied in the context of robust thrust optimization, in which the dynamical system has fewer degrees of freedom compared to complex legged robots. Later, Howell et al. (2019) extended the augmented Lagrangian approach to handle arbitrary inequality constraints for aerial navigation and manipulation problems. Additionally, the algorithm incorporates an active set projection for solution polishing and is often faster than direct collocation solved with IPOPT or SNOPT.

Our work proposes a feasibility-driven search for nonlinear optimal control problems with control limits. The main motivation of our approach is to increase the algorithm's basins of attraction, by focusing on feasibility instead of focusing solely on efficiency and optimality. Apart from the control limits and dynamics, we handle all remaining constraints (e.g., state and friction cone) through quadratic penalization, as described in the results section.

## 1.2 Contribution

The main contribution of this work is the first complete study of the numerical properties, behaviors, and guarantees of

<sup>3</sup> In the following section we formally describe the action-value function (i.e.,  $Q$ -function).



**Fig. 1** Snapshots of an athletic behavior computed by our feasibility-driven approach with explicit control limits: BOX-FDDP. The optimized motion considers the robot's full-body dynamics, joint limits, and friction cone constraints. We initialize our algorithm using the default posture and quasi-static torques (as described in the results section). We describe three desired actions: grasping the bar at a specific point, raising the feet up-ward a given location (Seq. 1), and landing over a specific placement (Seq. 5). The shown behavior emerges without further specification from our problem formulation. The arrows describe the magnitude and direction of the contact forces. To watch the video, click the figure or see [https://youtu.be/bOGBPTh\\_lsU](https://youtu.be/bOGBPTh_lsU)(Color figure online)

feasibility-driven search in differential dynamic programming. It relies on three technical contributions:

- (i) An original and efficient optimal control algorithm that directly handles control limits (BOX-FDDP),
- (ii) Extensive comparisons against direct transcription and BOX-DDP<sup>+</sup> (a single shooting method),
- (iii) A tutorial that connects the different branches of theory in optimal control, and
- (iv) An experimental validation of the dynamic feasibility evolution against interior point and active set algorithms for nonlinear programming.

Our approach builds on top of our previous results on feasibility-driven search (Mastalli et al., 2020), for which we hereby propose to define two modes in our algorithm: feasibility-driven and control-bounded. It considers the dynamic feasibility in the forward pass and explicitly incorporates control limits, which does not require a merit function as in Gifftthaler et al. (2018). Additionally, our approach has outstanding numerical capabilities, which allow us to generate motions that go beyond state-of-the-

art methods on optimal control or trajectory optimization in robotics, e.g., the athletic maneuver of a humanoid robot shown in Fig. 1.

## 2 Direct multiple shooting and differential dynamic programming

Before describing our approach, we introduce direct multiple shooting, and explain its numerical advantages when compared to single shooting methods such as DDP (Sect. 2.1). Then, in Sect. 2.2 we present a unique tutorial that connects the various branches of theory: KKT, PMP, and HJB. Additionally, in Sect. 2.3 we describe the salient aspects of original BOX-DDP proposed by Tassa et al. (2014), and our variant BOX-DDP<sup>+</sup>. This section contains known material, although we believe it contributes (i) to unveil the underlying problems of differential dynamic programming, and (ii) to understand the theoretical foundations of our feasibility-driven approach.

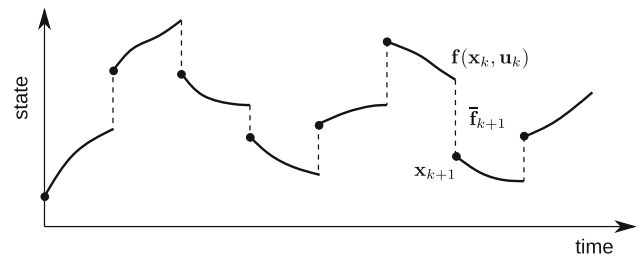
### 2.1 Direct multiple shooting for optimal control

Without loss of generality, we consider a direct multiple shooting approach for the nonlinear optimal control problem with control bounds in which each shooting segment defines a single timestep:s

$$\begin{aligned}
 \min_{\mathbf{x}_s, \mathbf{u}_s} \quad & \ell_N(\mathbf{x}_N) + \sum_{k=0}^{N-1} \ell_k(\mathbf{x}_k, \mathbf{u}_k) \\
 \text{s.t.} \quad & \bar{\mathbf{f}}_0 := \mathbf{x}_0 \ominus \tilde{\mathbf{x}}_0 = \mathbf{0}, \\
 & \bar{\mathbf{f}}_{k+1} := \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k) \ominus \mathbf{x}_{k+1} = \mathbf{0}, \quad \forall k=\{0,1,\dots,N-1\} \\
 & \underline{\mathbf{u}} \leq \mathbf{u}_k \leq \bar{\mathbf{u}}, \quad \forall k=\{0,1,\dots,N-1\}
 \end{aligned} \tag{1}$$

where the state  $\mathbf{x} \in X$  lies in a differential manifold (with dimension  $n_x$ ); the control  $\mathbf{u} \in \mathbb{R}^{n_u}$  defines the input commands;  $\underline{\mathbf{u}}, \bar{\mathbf{u}}$  are the lower and upper control bounds;  $\tilde{\mathbf{x}}_0$  is the initial state of the system;  $\ominus$  describes the *difference* operator of the state manifold [notation inspired by Frese (2008) that is needed to optimize over manifolds (Gabay, 1982)];  $N \in \mathbb{N}$  describes the number of nodes (horizon);  $\ell_N, \ell_k$  are the terminal and running cost functions; and  $\bar{\mathbf{f}}_0, \bar{\mathbf{f}}_{k+1} \in T_{\mathbf{x}}X$  are the gap residual functions that impose the dynamic feasibility; and  $T_{\mathbf{x}}X$  describes the tangent space of the state manifold at the state  $\mathbf{x}$ .

Equation (1) describes a nonlinear program as the system dynamics are transcribed into a set of algebraic equations with defects in each timestep. It is possible to extend this notation for cases where shooting segments contain multiple timesteps; however, as seen later, this does not provide any computational benefit, i.e., reduction in the computation time or better distribution of nonlinearities of the dynamics.



**Fig. 2** A schematic of a direct multiple shooting formulation. Different intermediate states are introduced as decision variables. A set of equality constraints enforces the dynamics feasibility. The black dots represent the initial states used for the system roll-out. The dashed lines represent the gaps between the roll-outs. Dynamic feasibility is achieved once these gaps are closed

Fig. 2 depicts the transcription process incorporating state and control trajectories ( $\mathbf{x}_s, \mathbf{u}_s$ ) as decision variables. This is in contrast to differential dynamic programming, which only transcribe the control sequence  $\mathbf{u}_s$  and obtain  $\mathbf{x}_s$  by integrating the system dynamics (i.e., a single shoot).

#### 2.1.1 Numerical behavior of direct multiple shooting

Algorithms for nonlinear programming aim at finding the Karush–Kuhn–Tucker (KKT) conditions defined by the FONC of optimality (Nocedal & Wright, 2006). This process involves iteratively solving a KKT problem (i.e., linear system of equations) until satisfaction of a stopping criterion. In the *line search* strategy, the solution of this KKT problem provides a search direction  $\delta \mathbf{w}_k$ , and the selected step length  $\alpha$  defines how much the current guess  $\mathbf{w}_k$  moves along that direction, i.e.,  $\mathbf{w}_{k+1} = \mathbf{w}_k \oplus \alpha \delta \mathbf{w}_k$ . Note that the *integrator* operator  $\oplus$  enables us to optimize over the manifold [as in Frese (2008), Gabay (1982)], however, it is a feature that general-purpose nonlinear programming libraries often does not have.

We can easily analyze the numerical behavior of direct multiple shooting formulations by focusing on the KKT problem for the shooting interval  $k$  only. This is possible because of the temporal/Markovian structure of optimal control problems. Therefore, when we apply the Newton method on this KKT sub-problem together with the Bellman’s principle of optimality, we obtain:

$$\begin{bmatrix} \ell_{\mathbf{x}\mathbf{x}k} & \ell_{\mathbf{u}\mathbf{x}k}^\top & \mathbf{f}_{\mathbf{x}k}^\top \\ \ell_{\mathbf{u}\mathbf{x}k} & \ell_{\mathbf{u}\mathbf{u}k} & \mathbf{f}_{\mathbf{u}k}^\top \\ \mathbf{f}_{\mathbf{x}k} & \mathbf{f}_{\mathbf{u}k} & -\mathbf{I} \\ & & -\mathbf{I} \mathcal{V}_{\mathbf{x}\mathbf{x}k+1} \end{bmatrix} \begin{bmatrix} \delta \mathbf{x}_k \\ \delta \mathbf{u}_k \\ \lambda_{k+1}^+ \\ \delta \mathbf{x}_{k+1} \end{bmatrix} = - \begin{bmatrix} \ell_{\mathbf{x}k} \\ \ell_{\mathbf{u}k} \\ \bar{\mathbf{f}}_{k+1} \\ \mathcal{V}_{\mathbf{x}k+1} \end{bmatrix}, \tag{2}$$

where Eq. (2) defines the stationary condition (first and second rows) and the primal feasibility (third row) of the FONC of optimality, respectively;  $\delta \mathbf{x}_k, \delta \mathbf{u}_k$  define the search direction for the primal variables;  $\lambda_{k+1}^+$  is the updated Lagrangian

multipliers;  $\ell_{\mathbf{x}_k}$ ,  $\ell_{\mathbf{u}_k}$ , and  $\ell_{\mathbf{x}\mathbf{x}_k}$ ,  $\ell_{\mathbf{x}\mathbf{u}_k}$ ,  $\ell_{\mathbf{u}\mathbf{u}_k}$  are the Jacobians and Hessians of the cost function;  $\mathbf{f}_{\mathbf{x}_k}$ ,  $\mathbf{f}_{\mathbf{u}_k}$  are the Jacobians of the system dynamics; and  $\mathcal{V}_{\mathbf{x}_k}$ ,  $\mathcal{V}_{\mathbf{x}\mathbf{x}_k}$  are the gradient and Hessian of the value function. Note that we apply the Gauss-Newton (GN) approximation as we ignore the Hessian of the system dynamics to avoid expensive tensor-vector multiplications.

When we factorize this system of equations, the resulting search direction always satisfies the dynamics constraints if the Jacobians and Hessians are constant (i.e., a LQR problem). However, if we apply an  $\alpha$ -step, the gap of the dynamics closes by a factor of  $(1 - \alpha)$ . We observe this by inspecting the primal feasibility at the next iteration:

$$\begin{aligned} \bar{\mathbf{f}}_{k+1}^{i+1} &= \bar{\mathbf{f}}_{k+1}^i - (\delta \mathbf{x}_{k+1} - \mathbf{f}_{\mathbf{x}_k} \delta \mathbf{x}_k - \mathbf{f}_{\mathbf{u}_k} \delta \mathbf{u}_k) \\ &= (1 - \alpha)(\mathbf{f}(\mathbf{x}_k^i, \mathbf{u}_k^i) \ominus \mathbf{x}_{k+1}^i), \end{aligned} \tag{3}$$

where, by definition in Eq. (2) (third row), we have that

$$\alpha \bar{\mathbf{f}}_{k+1}^i = \delta \mathbf{x}_{k+1} - \mathbf{f}_{\mathbf{x}_k} \delta \mathbf{x}_k - \mathbf{f}_{\mathbf{u}_k} \delta \mathbf{u}_k,$$

with the gap defined as  $\bar{\mathbf{f}}_{k+1}^i = \mathbf{f}(\mathbf{x}_k^i, \mathbf{u}_k^i) \ominus \mathbf{x}_{k+1}^i$ , and  $i$  describes the iteration number.

As described later in Sect. 3, injecting this numerical behavior can be interpreted as a feasibility-driven approach for multiple shooting. However, our approach operates quite differently from classical multiple shooting approaches. For instance, it does not increase the computation time by defining extra state (decision) variables. But there is no such thing as a free lunch as our approach cannot temporarily increase the defects (e.g., to reduce the cost value) after taking its first full step ( $\alpha = 1$ ).

### 2.1.2 Advantages of direct multiple shooting

The rationale for a direct multiple shooting approach (namely, adding  $\mathbf{x}_s$  as decision variables) is to distribute the nonlinearities of the dynamics over the entire horizon (Diehl et al., 2006). To illustrate this statement, we recognize that integrating over a horizon implies recursively calling *integrator* functions, i.e.,

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k) \\ &= \mathbf{f}(\mathbf{f}(\mathbf{x}_{k-1}, \mathbf{u}_{k-1}), \mathbf{u}_k) \\ &= \vdots \\ &= \mathbf{f}(\mathbf{f}(\dots \mathbf{f}(\mathbf{x}_0, \mathbf{u}_0), \mathbf{u}_{k-1}), \mathbf{u}_k), \end{aligned} \tag{4}$$

in which the nonlinearity increases along the horizon. This means that the local prediction, constructed by the derivatives of the nonlinear system in the KKT problem, becomes more inaccurate as the horizon increases. This is a well recognized drawback of single shooting approaches, and certainly a

numerical limitation of the BOX-DDP<sup>+</sup> algorithm described below.

### 2.2 Connection between KKT, PMP and HJB branches

The fourth row of Eq. (2) connects the value function with the Lagrange multipliers associated with the state equations. By definition, this multiplier corresponds to the next costate value at node  $k$ , which reveals an interesting connection with the PMP used in indirect multiple shooting methods and the KKT approach, i.e.

$$\lambda_k^+ = \mathcal{V}_{\mathbf{x}_k} + \mathcal{V}_{\mathbf{x}\mathbf{x}_k} \delta \mathbf{x}_k.$$

This might be not surprising if we realize that the PMP or KKT approach write the optimal control in term of the costate, while HJB expresses it in terms of the value function. Equation (2) is also at the heart of direct single shooting approaches such as DDP if each dynamics gap  $\bar{\mathbf{f}}_{k+1}$  vanishes, therefore this connection holds for single shooting settings as well.

Different interpretations can arise from this connection. For instance, DDP or our approach can be interpreted as iterative methods for solving the PMP in discrete-time optimal control problems under single and multiple shooting settings, respectively. Furthermore, under the context of linear dynamics and quadratic cost, DDP or our approach can be classified as global methods as they compute an optimal policy (i.e., a closed-loop solution).

### 2.3 Differential dynamic programming with control limits

As proposed by Tassa et al. (2014), BOX-DDP locally approximates the value function at node  $k$  as

$$\begin{aligned} \mathcal{V}_k(\delta \mathbf{x}_k) &= \min_{\delta \mathbf{u}_k} \ell_k(\delta \mathbf{x}_k, \delta \mathbf{u}_k) + \mathcal{V}_{k+1}(\mathbf{f}(\delta \mathbf{x}_k, \delta \mathbf{u}_k)), \\ \text{s.t. } \quad \underline{\mathbf{u}} &\leq \mathbf{u}_k + \delta \mathbf{u}_k \leq \bar{\mathbf{u}}, \end{aligned} \tag{5}$$

which breaks the optimal control problem into a sequence of simpler sub-problems with control bounds. Then, a local search direction is computed through a linear quadratic (LQ) approximation of the value function:

$$\begin{aligned} \delta \mathbf{u}_k^*(\delta \mathbf{x}_k) &= \arg \min_{\delta \mathbf{u}_k} \frac{1}{2} \begin{bmatrix} 1 \\ \delta \mathbf{x}_k \\ \delta \mathbf{u}_k \end{bmatrix}^\top \overbrace{\begin{bmatrix} 0 & \mathbf{Q}_{\mathbf{x}_k}^\top & \mathbf{Q}_{\mathbf{u}_k}^\top \\ \mathbf{Q}_{\mathbf{x}_k} & \mathbf{Q}_{\mathbf{x}\mathbf{x}_k} & \mathbf{Q}_{\mathbf{x}\mathbf{u}_k} \\ \mathbf{Q}_{\mathbf{u}_k} & \mathbf{Q}_{\mathbf{x}\mathbf{u}_k}^\top & \mathbf{Q}_{\mathbf{u}\mathbf{u}_k} \end{bmatrix}}^{\mathcal{Q}_k(\delta \mathbf{x}_k, \delta \mathbf{u}_k, \bar{\mathcal{V}}_k)} \begin{bmatrix} 1 \\ \delta \mathbf{x}_k \\ \delta \mathbf{u}_k \end{bmatrix}, \\ \text{s.t. } \quad \underline{\mathbf{u}} &\leq \mathbf{u}_k + \delta \mathbf{u}_k \leq \bar{\mathbf{u}}, \end{aligned} \tag{6}$$

where the  $\mathbf{Q}_k$  terms describe the LQ approximation of the action-value function  $\mathcal{Q}_k(\cdot)$  that can be seen as a function of the derivatives of the value function  $\bar{\mathcal{V}}_k = (\mathcal{V}_{\mathbf{x}_k}, \mathcal{V}_{\mathbf{x}\mathbf{x}_k})$ .

Solving Eq. (6) results in a local feedback control law  $\delta \mathbf{u}_k = \mathbf{K}_k + \mathbf{K}_k \delta \mathbf{x}_k$  consisting of a feed-forward term  $\mathbf{k}_k$  and a state feedback gain  $\mathbf{K}_k$  for each discretization point  $k$ . Below, we describe how to compute the  $\mathbf{Q}_k$  terms in the so-called Riccati sweep step.

### 2.3.1 Riccati sweep

The LQ approximation of the action-value function  $Q_k$  is computed recursively, backwards in time, as follows

$$\begin{aligned} \mathbf{Q}_{\mathbf{x}_k} &= \ell_{\mathbf{x}_k} + \mathbf{f}_{\mathbf{x}_k}^\top \mathcal{V}_{\mathbf{x}_{k+1}}, \\ \mathbf{Q}_{\mathbf{u}_k} &= \ell_{\mathbf{u}_k} + \mathbf{f}_{\mathbf{u}_k}^\top \mathcal{V}_{\mathbf{x}_{k+1}}, \\ \mathbf{Q}_{\mathbf{xx}_k} &= \ell_{\mathbf{xx}_k} + \mathbf{f}_{\mathbf{x}_k}^\top \mathcal{V}_{\mathbf{xx}_{k+1}} \mathbf{f}_{\mathbf{x}_k}, \\ \mathbf{Q}_{\mathbf{xu}_k} &= \ell_{\mathbf{xu}_k} + \mathbf{f}_{\mathbf{x}_k}^\top \mathcal{V}_{\mathbf{xx}_{k+1}} \mathbf{f}_{\mathbf{u}_k}, \\ \mathbf{Q}_{\mathbf{uu}_k} &= \ell_{\mathbf{uu}_k} + \mathbf{f}_{\mathbf{u}_k}^\top \mathcal{V}_{\mathbf{xx}_{k+1}} \mathbf{f}_{\mathbf{u}_k}, \end{aligned} \tag{7}$$

where  $\mathcal{V}_{\mathbf{x}_{k+1}}, \mathcal{V}_{\mathbf{xx}_{k+1}}$  are obtained by solving the following algebraic Riccati equations at  $k + 1$ :

$$\begin{aligned} \mathcal{V}_{\mathbf{x}_{k+1}} &= \mathbf{Q}_{\mathbf{x}_{k+1}} - \mathbf{Q}_{\mathbf{xu}_{k+1}} \hat{\mathbf{Q}}_{\mathbf{uu}, \mathbf{f}_{k+1}}^{-1} \mathbf{Q}_{\mathbf{u}_{k+1}}, \\ \mathcal{V}_{\mathbf{xx}_{k+1}} &= \mathbf{Q}_{\mathbf{xx}_{k+1}} - \mathbf{Q}_{\mathbf{xu}_{k+1}} \hat{\mathbf{Q}}_{\mathbf{uu}, \mathbf{f}_{k+1}}^{-1} \mathbf{Q}_{\mathbf{xu}_{k+1}}^\top, \end{aligned} \tag{8}$$

with  $\hat{\mathbf{Q}}_{\mathbf{uu}, \mathbf{f}_{k+1}}$  as the control Hessian in the free space, which we will describe below. Additionally, we use the gradient and Hessian of the value function to find a local search direction as described below. In the case of BOX-DDP+, our adaptation of BOX-DDP to allow initialization with state trajectories, the gradient of the value function in the first iteration is relinearized by the *initialization infeasibility*  $\bar{\mathbf{f}}_{k+1}^0$  as  $\mathcal{V}_{\mathbf{x}_{k+1}} + \mathcal{V}_{\mathbf{xx}_{k+1}} \bar{\mathbf{f}}_{k+1}^0$ .

### 2.3.2 Control-bounded direction

We compute the control-bounded direction, defined in Eq. (6), by breaking it down into the so-called *feed-forward* and *feedback* sub-problems (Tassa et al., 2014). We first compute the feed-forward term by solving the following quadratic programming (QP) program with box constraints:

$$\begin{aligned} \mathbf{k}_k &= \arg \min_{\delta \mathbf{u}_k} \frac{1}{2} \delta \mathbf{u}_k^\top \mathbf{Q}_{\mathbf{uu}_k} \delta \mathbf{u}_k + \mathbf{Q}_{\mathbf{u}_k}^\top \delta \mathbf{u}_k, \\ \text{s.t.} \quad \underline{\mathbf{u}} &\leq \mathbf{u}_k + \delta \mathbf{u}_k \leq \bar{\mathbf{u}}, \end{aligned} \tag{9}$$

and then the feedback gain as

$$\mathbf{K}_k = -\hat{\mathbf{Q}}_{\mathbf{uu}, \mathbf{f}_k}^{-1} \mathbf{Q}_{\mathbf{ux}_k}, \tag{10}$$

where  $\hat{\mathbf{Q}}_{\mathbf{uu}, \mathbf{f}_k}$  is the control Hessian of the free subspace obtained in Eq. (9) in which  $\mathbf{f}_k$  describes the free subspace

at node  $k$ , i.e., the indexes of the inactive bounds. With these indexes, we sort and partition the control Hessian as:

$$\mathbf{Q}_{\mathbf{uu}_k} = \begin{bmatrix} \mathbf{Q}_{\mathbf{uu}, \mathbf{f}_k} & \mathbf{Q}_{\mathbf{uu}, \mathbf{fc}_k} \\ \mathbf{Q}_{\mathbf{uu}, \mathbf{cf}_k} & \mathbf{Q}_{\mathbf{uu}, \mathbf{c}_k} \end{bmatrix}, \tag{11}$$

and compute  $\hat{\mathbf{Q}}_{\mathbf{uu}, \mathbf{f}_k}^{-1}$  internally based on the factorization of  $\mathbf{Q}_{\mathbf{uu}, \mathbf{f}_k}^{-1}$ . This is what our BOX-QP program does to solve the feed-forward sub-problem efficiently via the Projected-Newton QP algorithm (Bertsekas, 1982). This algorithm quickly identifies the active set and moves along the free subspace of the Newton step. It also has a similar computational cost to the unconstrained QP if the active set remains unchanged. Thus, the runtime performance is similar to the DDP algorithm. However, it requires a feasible initialization  $\delta \mathbf{u}_k^0$ .

Again, by using a Projected-Newton QP algorithm, we further efficiently obtain the control Hessian of the free subspace  $\mathbf{Q}_{\mathbf{uu}, \mathbf{f}_k}^{-1}$  as the algorithm computes it internally when it moves along the free subspace of the Newton step. With it, we compute a state feedback gain that generates corrections within the control limits. This is an important feature for controlling the robot as well as for rolling-out the nonlinear dynamics in the forward pass. For more details about the Projected-Newton QP algorithm see Bertsekas (1982).

### 2.3.3 State integration

In any DDP algorithm such as BOX-DDP, we perform a state integration using the locally-linear policy as

$$\begin{aligned} \hat{\mathbf{x}}_0 &= \mathbf{x}_0, \\ \hat{\mathbf{u}}_k &= \mathbf{u}_k + \alpha \mathbf{k}_k + \mathbf{K}_k (\hat{\mathbf{x}}_k \ominus \mathbf{x}_k), \quad \forall k = \{0, 1, \dots, N-1\} \\ \hat{\mathbf{x}}_{k+1} &= \mathbf{f}(\hat{\mathbf{x}}_k, \hat{\mathbf{u}}_k), \quad \forall k = \{0, 1, \dots, N-1\} \end{aligned} \tag{12}$$

where  $\hat{\mathbf{x}}_k, \hat{\mathbf{u}}_k$  are the new state and control at node  $k$  generated using a step length  $\alpha$ . The feedback gain helps to distribute the nonlinearities; however, as seen in the previous section, it does not resemble the numerical behavior described by the FONC of optimality in direct multiple shooting. This different numerical behavior stems from the state integration procedure closing the gaps, i.e.,  $\bar{\mathbf{f}}_k = \mathbf{0}, \forall k = \{0, 1, \dots, N\}$ .

### 2.3.4 Expected improvement

When solving the algebraic Riccati equations, we obtain the expected improvement as

$$\Delta \mathcal{V}_k = -\frac{1}{2} \mathbf{Q}_{\mathbf{u}_k}^\top \hat{\mathbf{Q}}_{\mathbf{uu}, \mathbf{f}_k}^{-1} \mathbf{Q}_{\mathbf{u}_k}. \tag{13}$$

Below, we elaborate the proposed algorithm based on the aforementioned description.

### 3 Box-FDDP: a feasibility-driven approach for multiple shooting

We now introduce a novel algorithm that combines a feasibility-driven search (Sect. 2.1) with an active set treatment of the control limits (Sect. 2.3) named BOX-FDDP. The BOX-FDDP algorithm comprises two modes: *feasibility-driven* and *control-bounded* modes, one of which is chosen for a given iteration (Algorithm 1). The feasibility-driven<sup>4</sup> mode mimics the numerical resolution of a direct multiple shooting problem with only dynamics constraints when computing the search direction and step length (lines 4–10 and 11–21, respectively). This mode neglects the control limits of the system as it focuses on dynamic feasibility only. In contrast, the control-bounded mode projects the search direction onto the feasible control region whenever the dynamics constraint is feasible (line 10). In both modes, the applied controls in the forward pass are projected onto their feasible box (line 13), causing dynamically-infeasible iterations to reach the control box. With this strategy, our solver focuses on feasibility early on, which increases its basins of attraction, and later on optimality. Technical descriptions of both modes are elaborated in Sects. 3.1 and 3.2. Note that the existence of feasible descent directions are introduced later in Sect. 3.3.

---

**Algorithm 1:** Control-limited FDDP (BOX-FDDP)

---

```

1  compute LQ approximation of the cost and dynamics
2  if infeasible iterate then
3    | compute the gaps, Eq. (14)
4  for  $k \leftarrow N - 1$  to  $0$  do
5    | update the feasibility-driven  $Q$ -function, Eq. (17)
6    | if infeasible iterate then
7      | compute feasibility-driven direction, Eq. (19)
8    | else
9      | project BOX-QP warm start, Eq. (20)
10     | compute control-bounded direction, Eq. (9)-(10)
11  for  $\alpha \in \{1, \frac{1}{2}, \dots, \frac{1}{2^n}\}$  do
12    | for  $k \leftarrow 0$  to  $N$  do
13      | project control onto the feasible box, Eq. (21)
14      | if infeasible iterate or  $\alpha \neq 1$  then
15        | update the gaps, Eq. (22)
16      | else
17        | close the gaps,  $\mathbf{f}_k = \mathbf{0} \ \forall k \in \{0, \dots, N - 1\}$ 
18      | perform step, Eq. (23)
19    | compute the expected improvement, Eq. (24)
20    | if success step then
21      | break

```

---

<sup>4</sup> Here, *feasibility* concerns the dynamics of the system, not the feasibility of other problem constraints.

### 3.1 Search direction

In the standard BOX-DDP algorithm, an initial forward pass is performed to obtain the initial state trajectory  $\mathbf{x}_s$ . This trajectory enforces the dynamics explicitly; thus, the gaps are zero, i.e.,  $\bar{\mathbf{f}}_k = \mathbf{0}$  for all  $k = \{0, 1, \dots, N - 1\}$ . Instead, our multiple shooting variant, BOX-FDDP, computes the gaps once at each iteration (line 3), which are used to find the search direction and to compute the expected improvement. However, if the iteration is dynamically feasible, then the search direction procedure is the same as in the standard BOX-DDP (Tassa et al., 2014). Below, we describe the steps performed to compute the search direction.

#### 3.1.1 Computing the gaps

Given a current iterate  $(\mathbf{x}_s, \mathbf{u}_s)$ , we perform a nonlinear roll-out to compute the gaps as

$$\bar{\mathbf{f}}_{k+1} := \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k) \ominus \mathbf{x}_{k+1}, \quad \forall k = \{0, 1, \dots, N-1\} \tag{14}$$

where  $\mathbf{f}(\mathbf{x}_k, \mathbf{u}_k)$  is the roll-out state at interval  $k + 1$ ,  $\mathbf{x}_{k+1}$  is the next shooting state, and  $\ominus$  is the difference operator.

#### 3.1.2 Action-value function of direct multiple shooting formulation

Without loss of generality, we use the Gauss-Newton (GN) approximation (Li & Todorov, 2004) to write the action-value function of our algorithm as

$$Q_k(\cdot) = \frac{1}{2} \begin{bmatrix} 1 \\ \delta \mathbf{x}_{k+1} \end{bmatrix}^\top \begin{bmatrix} 0 & \mathcal{V}_{\mathbf{x}_{k+1}}^\top \\ \mathcal{V}_{\mathbf{x}_{k+1}} & \mathcal{V}_{\mathbf{x}\mathbf{x}_{k+1}} \end{bmatrix} \begin{bmatrix} 1 \\ \delta \mathbf{x}_{k+1} \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 1 \\ \delta \mathbf{x}_k \\ \delta \mathbf{u}_k \end{bmatrix}^\top \begin{bmatrix} 0 & \ell_{\mathbf{x}_k}^\top & \ell_{\mathbf{u}_k}^\top \\ \ell_{\mathbf{x}_k} & \ell_{\mathbf{x}\mathbf{x}_k} & \ell_{\mathbf{x}\mathbf{u}_k} \\ \ell_{\mathbf{u}_k} & \ell_{\mathbf{x}\mathbf{u}_k}^\top & \ell_{\mathbf{u}\mathbf{u}_k} \end{bmatrix} \begin{bmatrix} 1 \\ \delta \mathbf{x}_k \\ \delta \mathbf{u}_k \end{bmatrix}, \tag{15}$$

where again  $\ell_{\mathbf{x}_k}$ ,  $\ell_{\mathbf{u}_k}$  and  $\ell_{\mathbf{x}\mathbf{x}_k}$ ,  $\ell_{\mathbf{x}\mathbf{u}_k}$ ,  $\ell_{\mathbf{u}\mathbf{u}_k}$  describe the gradient and Hessian of the cost function, respectively;  $\delta \mathbf{x}_{k+1} = \mathbf{f}_{\mathbf{x}_k} \delta \mathbf{x}_k + \mathbf{f}_{\mathbf{u}_k} \delta \mathbf{u}_k$  is the linearized dynamics; and  $\mathbf{f}_{\mathbf{x}_k}$ ,  $\mathbf{f}_{\mathbf{u}_k}$  are its Jacobians.

In direct multiple shooting settings, linearization of the system dynamics includes a drift term

$$\delta \mathbf{x}_{k+1} = \mathbf{f}_{\mathbf{x}_k} \delta \mathbf{x}_k + \mathbf{f}_{\mathbf{u}_k} \delta \mathbf{u}_k + \bar{\mathbf{f}}_{k+1}, \tag{16}$$

as there are gaps in the dynamics term  $\bar{\mathbf{f}}_{k+1}$  produced between subsequent shooting segments (q.v. Fig. 2). Then, the Riccati sweep needs to be adapted as follows:

$$\begin{aligned} \mathbf{Q}_{\mathbf{x}_k} &= \ell_{\mathbf{x}_k} + \mathbf{f}_{\mathbf{x}_k}^\top \mathcal{V}_{\mathbf{x}_{k+1}}^+, \\ \mathbf{Q}_{\mathbf{u}_k} &= \ell_{\mathbf{u}_k} + \mathbf{f}_{\mathbf{u}_k}^\top \mathcal{V}_{\mathbf{x}_{k+1}}^+, \end{aligned}$$

$$\begin{aligned}
 \mathbf{Q}_{xx_k} &= \ell_{xx_k} + \mathbf{f}_{x_k}^\top \mathcal{V}_{xx_{k+1}} \mathbf{f}_{x_k}, \\
 \mathbf{Q}_{xu_k} &= \ell_{xu_k} + \mathbf{f}_{x_k}^\top \mathcal{V}_{xx_{k+1}} \mathbf{f}_{u_k}, \\
 \mathbf{Q}_{uu_k} &= \ell_{uu_k} + \mathbf{f}_{u_k}^\top \mathcal{V}_{xx_{k+1}} \mathbf{f}_{u_k}
 \end{aligned} \tag{17}$$

in which

$$\mathcal{V}_{x_{k+1}}^+ = \mathcal{V}_{x_{k+1}} + \mathcal{V}_{xx_{k+1}} \bar{\mathbf{f}}_{k+1} \tag{18}$$

is the gradient of the value function after the deflection produced by  $\bar{\mathbf{f}}_{k+1}$  (also described above as relinearization). Note that the Hessian of the value function remains unchanged as DDP approximates the value function through a LQ model. Additionally, this modification affects the values of the Riccati equations, Eq. (8), and expected improvement, Eq. (13), as they depend on the gradient of the value function.

### 3.1.3 Feasibility-driven direction

For *dynamically-infeasible* iterates (line 7), we ignore the control constraints and compute a *control-unbounded* direction:

$$\begin{aligned}
 \mathbf{k}_k &= -\mathbf{Q}_{uu_k}^{-1} \mathbf{Q}_{u_k}, \\
 \mathbf{K}_k &= -\mathbf{Q}_{uu_k}^{-1} \mathbf{Q}_{ux_k}.
 \end{aligned} \tag{19}$$

We do this because we cannot quantify the effect of the gaps on the control bounds, which are needed to solve the feed-forward sub-problem Eq. (9). Our approach is equivalent to opening the control bounds during dynamically-infeasible iterates.

### 3.1.4 Control-bounded direction

We warm-start the BOX-QP using the feed-forward term  $\mathbf{k}_k$  computed in the previous iteration. However, if the algorithm is switching from feasibility to control-bounded mode (i.e., the previous iteration is infeasible), then  $\mathbf{k}_k$  might fall outside the feasible box and  $\underline{\mathbf{u}} - \mathbf{u}_k \leq \mathbf{k}_k \leq \bar{\mathbf{u}} - \mathbf{u}_k$  do not hold. This violates the assumption of the previously-described BOX-QP, for which a feasible initial point needs to be provided.

To handle infeasible iterates, we propose to project the warm-start of the BOX-QP (line 9) as

$$\llbracket \mathbf{k}_k \rrbracket_{\underline{\mathbf{u}}, \bar{\mathbf{u}}} = \min(\max(\mathbf{k}_k, \underline{\mathbf{u}} - \mathbf{u}_k), \bar{\mathbf{u}} - \mathbf{u}_k), \tag{20}$$

where  $\underline{\mathbf{u}}, \bar{\mathbf{u}}$  are the lower and upper bounds of the feed-forward sub-problem, Eq. (9), respectively.

Once we project the warm-start  $\mathbf{k}_k$ , we solve the feed-forward and feedback sub-problems as explained in Sect. 2.3.2. Furthermore, we solve the BOX-QP using a Projected-Newton method (Bertsekas, 1982), which handles box constraints efficiently as described above.

## 3.2 Step length

As far as we know, the standard BOX-DDP (Tassa et al., 2014) modifies only the search direction (i.e., backward pass) to handle the control limits. However, it is also important to project the controls onto the feasible box during the forward pass. We do this by finding a step length that minimizes the cost (Nocedal & Wright, 2006).

### 3.2.1 Projecting the roll-out towards the feasible box

We propose to project the controls onto the feasible box in the nonlinear roll-out (line 13), i.e.,

$$\hat{\mathbf{u}}_k \leftarrow \min(\max(\hat{\mathbf{u}}_k, \underline{\mathbf{u}}), \bar{\mathbf{u}}), \tag{21}$$

where  $\hat{\mathbf{u}}_k$  is the updated control from the control policy. Our method does not require to solve another QP problem (Xie et al., 2017) or to project the linear search direction given the gaps on the dynamics (Howell et al., 2019). Furthermore, the control policy considers a gap prediction that guarantees a feasible descent direction. We formally describe the technical details of this procedure in Sect. 3.2.3.

### 3.2.2 Updating the gaps

As analyzed earlier, the evolution of the gaps in direct multiple shooting is affected by the selected step length. For an optimal control problem without control limits, this evolution is defined as

$$\bar{\mathbf{f}}_k \leftarrow (1 - \alpha) \bar{\mathbf{f}}_k, \tag{22}$$

where  $\alpha$  is the step-length found by the line-search procedure (line 11–21). Note that a full step ( $\alpha = 1$ ) closes the gaps completely. We described this gap contraction rate in Sect. 2.1.1.

### 3.2.3 Nonlinear step

With a nonlinear roll-out<sup>5</sup> (line 18), we avoid the linear prediction error of the dynamics that is typically handled by a *merit* function in general-purpose NLP algorithms, as explained in Mastalli et al. (2020). If we keep the gap-contraction rate of Eq. (22), then we obtain

$$\begin{aligned}
 \hat{\mathbf{x}}_k &= \mathbf{f}(\hat{\mathbf{x}}_{k-1}, \hat{\mathbf{u}}_{k-1}) \oplus (\alpha - 1) \bar{\mathbf{f}}_{k-1}, \\
 \hat{\mathbf{u}}_k &= \mathbf{u}_k + \alpha \mathbf{k}_k + \mathbf{K}_k (\hat{\mathbf{x}}_k \ominus \mathbf{x}_k),
 \end{aligned} \tag{23}$$

<sup>5</sup> In this work, a nonlinear *roll-out* is also referred to as a nonlinear step.

where  $\hat{\mathbf{x}}_k$ ,  $\hat{\mathbf{u}}_k$  are the next state and control along an  $\alpha$ -step;  $\mathbf{K}_k$  and  $\bar{\mathbf{K}}_k$  are the feed-forward term and feedback gains computed by Eq. (19) or Eq. (9)–(10). Furthermore, the initial condition of the roll-out is defined as  $\hat{\mathbf{x}}_0 = \mathbf{Q}_0 \oplus (\alpha - 1)\bar{\mathbf{f}}_0$ . Note that this is in contrast to the standard BOX-DDP, in which the gaps are always closed, even for  $\alpha < 1$ .

Avoiding the use of a merit function helps the algorithm to check the search direction more accurately. Indeed, it has been shown that the nonlinear roll-out is more effective than a standard line search procedure as it reduces the number of iterations (Liao & Shoemaker, 1992).

### 3.2.4 Expected Improvement

It is critical to properly evaluate the success of a trial step. Given the current dynamics gaps  $\bar{\mathbf{f}}_k$ , BOX-FDDP computes the expected improvement of a computed search direction as

$$\Delta J(\alpha) = \Delta_1 \alpha + \frac{1}{2} \Delta_2 \alpha^2, \quad (24)$$

with

$$\begin{aligned} \Delta_1 &= \sum_{k=0}^N \mathbf{k}_k^\top \mathbf{Q}_{\mathbf{u}k} + \bar{\mathbf{f}}_k^\top (\mathcal{V}_{\mathbf{x}k} - \mathcal{V}_{\mathbf{x}\mathbf{x}k} \delta \hat{\mathbf{x}}_k), \\ \Delta_2 &= \sum_{k=0}^N \mathbf{k}_k^\top \hat{\mathbf{Q}}_{\mathbf{u}\mathbf{u}, \mathbf{f}_k} \mathbf{k}_k + \bar{\mathbf{f}}_k^\top (2\mathcal{V}_{\mathbf{x}\mathbf{x}k} \delta \hat{\mathbf{x}}_k - \mathcal{V}_{\mathbf{x}\mathbf{x}k} \bar{\mathbf{f}}_k), \end{aligned} \quad (25)$$

where  $\hat{\mathbf{Q}}_{\mathbf{u}\mathbf{u}, \mathbf{f}_k}$  is the control Hessian of the free space,  $\delta \hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k \ominus \mathbf{x}_k$ , and  $J$  is the total cost of a given state-control trajectory  $(\mathbf{x}_s, \mathbf{u}_s)$ . We use this expected improvement model for both modes. Note that, in the feasibility-driven mode, the free space spans the entire control space; instead, in the control-bounded mode, the gaps are zero.

We obtain this expression by computing the cost from a linear roll-out of the current control policy as described in Eq. (23). We also accept ascent directions when evaluating the trial step, our approach is inspired by the Goldstein condition (Nocedal & Wright, 2006, Chapter 3):

$$\ell' - \ell \leq \begin{cases} b_1 \Delta J(\alpha) & \text{if } \Delta J(\alpha) \leq 0 \\ b_2 \Delta J(\alpha) & \text{otherwise} \end{cases}, \quad (26)$$

where  $b_1, b_2$  are adjustable parameters, we used in this paper  $b_1 = 0.1$  and  $b_2 = 2$ . Ascent directions improve the algorithm convergence as it helps to reduce the feasibility error through a moderate increment in the cost.

### 3.2.5 Regularization

We regularize the  $\mathbf{Q}_{\mathbf{u}\mathbf{u}}$  and  $\mathcal{V}_{\mathbf{x}\mathbf{x}}$  terms through a Levenberg-Marquardt scheme (Fletcher, 1971). Concretely, we increase

the damping value  $\mu$  when the computation of the feed-forward sub-problem—formulated in Eq. (9)—fails, or when the forward pass accepts a step length smaller than  $\alpha_0 = 0.01$ . Moreover, we decrease the damping value if the iteration accepts a step larger than  $\alpha_1 = 0.5$ . Both regularization procedures modify the values of the  $\mathbf{Q}_{\mathbf{u}\mathbf{u}}$  and  $\mathcal{V}_{\mathbf{x}\mathbf{x}}$  terms during the Riccati sweep computation as

$$\begin{aligned} \mu' &\leftarrow \beta_{i,d} \mu, \\ \mathbf{Q}_{\mathbf{u}\mathbf{u}} &\leftarrow \mathbf{Q}_{\mathbf{u}\mathbf{u}} + \mu' \mathbf{I}, \\ \mathcal{V}_{\mathbf{x}\mathbf{x}} &\leftarrow \mathcal{V}_{\mathbf{x}\mathbf{x}} + \mu' \mathbf{I}, \end{aligned}$$

where  $\beta_i$  and  $\beta_d$  are the factors<sup>6</sup> used to increase or decrease the current damping value  $\mu$ , respectively;  $\mu'$  is the newly-computed damping value; and  $\mathbf{I}$  is the identity matrix. Additionally, we start the regularization procedure with an initial, and user-defined, damping value.<sup>7</sup> We also define minimum and maximum damping values to avoid increasing or decreasing the damping value unnecessarily.<sup>8</sup>

Both regularizations significantly increase the robustness of the algorithm and ensures convergence, as it moves from Newton direction to steepest-descent, and vice versa. The Newton direction, which occurs with  $\mu = 0$ , provides fast convergence and is robust against scaling because it exploits the Hessian of the problem. However, it does not always produce valid descent directions as  $\mathbf{Q}_{\mathbf{u}\mathbf{u}}$  might be indefinite and the problem nonconvex. In such cases, increasing the damping value guarantees that  $\mathbf{Q}_{\mathbf{u}\mathbf{u}}$  is positive-definite which, in turn, computes a search direction closer to the steepest-descent one. Instead,  $\mathcal{V}_{\mathbf{x}\mathbf{x}}$  enforces the state trajectory to be closer to the one previously computed (Tassa, 2011). It will also not result in vanishing feedback gains even for large damping values.

### 3.3 Existence of feasible descent directions

As described above, our approach has two main modes: feasibility-driven and control-bounded. During the feasibility-driven phase, we compute a search direction to drive the next guess towards dynamic feasibility and try a step while keeping the control within the box constraints. This projection procedure can be seen as a nonlinear term in our dynamics, but we assume its effect is negligible for finding a feasible direction. On the other hand, our algorithm computes a search direction that considers the box constraints after the dynamic feasibility has been achieved. This is needed to improve the

<sup>6</sup>  $\beta_{i,d}$  commonly range between 2–10. We set  $\beta_{i,d} = 10$  in this work.

<sup>7</sup> We use  $10^{-9}$  as the initial regularization value.

<sup>8</sup> We use  $10^{-16}$  and  $10^{12}$  as the minimum and maximum damping values, respectively. Note that  $10^{-16}$  is approximately the resolution of a double number.

next current guess by taking control constraints into account when computing the feedback gains along the free subspace.

As analyzed in Sect. 2.1.1, the feasibility-driven direction is computed by mimicking the numerical behavior of a nonlinear program during the resolution of a direct multiple shooting problem with only dynamics constraints. It implies that the feasibility-driven search produces a descent direction, and eventually the algorithm converges, if the cost Hessian is a positive definite matrix. Indeed, the positiveness is always guaranteed by our regularization procedure as described before. Furthermore, the feasibility-driven step aims at reducing the nonlinearities produced by a single shooting formulation (e.g., DDP algorithm). When the dynamics are feasible, we apply a control-bounded search which also produces a descent direction as it boils down to a QP program.

In the next section, we present a series of results that demonstrate the benefits of our feasibility-driven approach.

## 4 Results

We analyze BOX-FDDP across a wide range of optimal control problems (briefly introduced in Sect. 4.1) as follows. First, in Sect. 4.2, we show the benefits of the feasibility mode by analyzing the gap contraction and its connection with the nonlinearities in the dynamics. Second, we compare our algorithm against a direct transcription formulation in Sect. 4.4. Concretely, we compare the dynamic feasibility and optimality evolutions, runtime performance, and robustness to different initial guesses against the interior point and active set algorithms available in KNITRO. Finally, in Sect. 4.5, we report the results of a squashing approach for solving the control bounds as it demonstrates the numerical performance of having two modes.

We benchmark the algorithms using an Intel Core i9-9900KF CPU with eight cores @ 3.60 GHz and 16 MB cache. Our implementation of BOX-FDDP supports code generation, but all runtime performances reported henceforth do not employ it for fair comparison with other approaches. We use 8 threads to compute the cost and dynamics derivatives for the experiments with the BOX-FDDP algorithm and the legged robots only. We use the same initial regularization and stopping criteria values for each problem, and the values are  $10^{-9}$  and  $5 \times 10^{-5}$ , respectively.

### 4.1 Optimal control problems

To provide empirical evidence on the benefits of the feasibility-driven approach, we developed a range of different optimal control problems: (1) an under-actuated double pendulum (PEND); (2) a quadcopter navigating towards a goal (GOAL) or through a narrow passage (NARROW) and looping (LOOP); (3)

various gaits, aggressive jumps (JUMP) and unstable hopping (HOP) in a quadruped robot; (4) whole-body manipulation (MAN), hand control while balancing in single leg (TAICHI), dip on parallel bars (DIP) and a pull-up bar task (PULLUP) in a humanoid robot. Figure 3 shows snapshots of motions computed by BOX-FDDP for some of these problems, and the accompanying video shows the entire motion sequences.<sup>9</sup>

We describe the cost functions, dynamics, control limits, penalization terms, and initialization of each optimal control problem in Appendix A. Finally, some of these problems, as well as our implementation of the BOX-FDDP algorithm, are publicly available in the CROCODDYL repository (Mastalli et al., 2019).

### 4.2 Advantages of the feasibility-driven mode

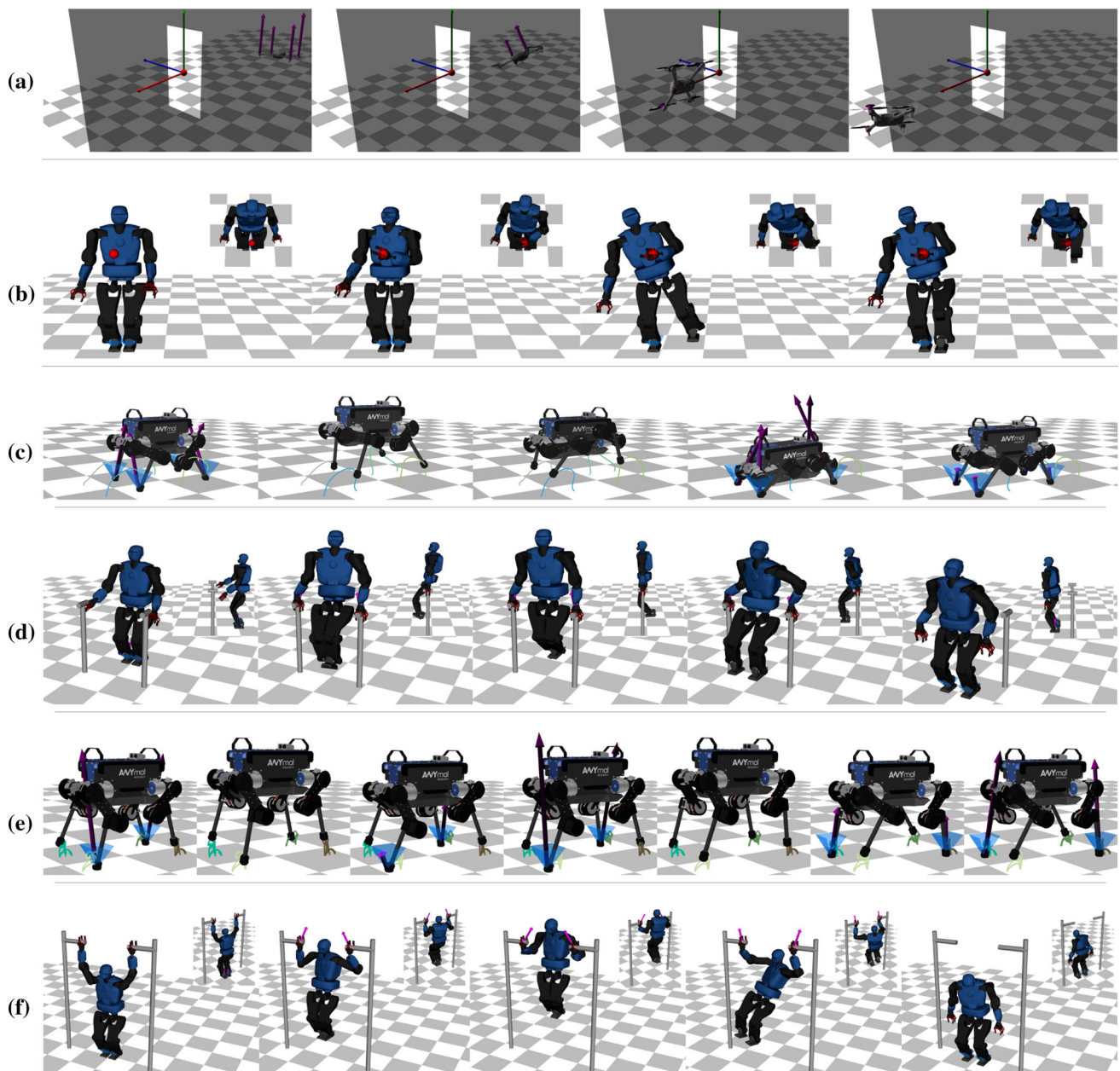
To understand the benefits of the feasibility-driven mode, we analyze the resulting total cost, number of iterations and total computation time obtained in both algorithms: BOX-FDDP and BOX-DDP<sup>+</sup> using the same initial guess. BOX-DDP<sup>+</sup> is an improved version of the standard BOX-DDP proposed in Tassa et al. (2014), which it accepts initialization for both: state and control trajectories as described above. This version accepts infeasible warm-starts as in Eq. (17), and it is available in the CROCODDYL repository. Without this modification, the standard BOX-DDP could easily diverge (and not converge at all) when we initialize it using quasi-static torques in problems with medium to longer horizons.

#### 4.2.1 Larger basin of attraction and convergence

In our experiments, BOX-DDP<sup>+</sup> was unable to generate jumping and hopping motions for quadrupeds, as well as pull-ups for humanoids, i.e., it failed to solve our JUMP, HOP, and PULLUP task specifications (marked by the  $\times$  in Table 1). BOX-DDP<sup>+</sup> failed to generate such aggressive motions because trajectories satisfying all the specifications of those tasks are significantly distant from the initial guess provided to the solver. BOX-DDP<sup>+</sup> behaves poorly as it computes solution with higher cost value and computation time, which is critical drawback in model predictive control applications.

In contrast, our approach (BOX-FDDP) was able to solve all of the tasks, and it did so with fewer iterations and lower total cost (see Table 1 and Fig. 4). Furthermore, BOX-FDDP and BOX-DDP<sup>+</sup> have the same algorithmic complexity, but since our approach requires fewer iterations than BOX-DDP<sup>+</sup>, the total computation time of our approach is lower. These results are a direct consequence of the feasibility-driven mode in our approach, which is able to find control sequences even when faced with poor initial guesses. The

<sup>9</sup> Supplementary video: [https://youtu.be/bOGBPTh\\_IsU](https://youtu.be/bOGBPTh_IsU).



**Fig. 3** Snapshots of different robot maneuvers computed by BOX-FDDP. **a** traversing a narrow passage with a quadcopter (GOAL); **b** Talos balancing on a single leg (TAICHI); **c** aggressive jumping of 30 cm that reaches ANYmal limits (JUMP); **d** Talos dipping on a parallel bars (DIP);

**e** ANYmal hopping with two legs (HOP); **f** Talos performing a pull-up bar task (PULLUP). To watch the video, click its respective figure or see [https://youtu.be/BOGBPTh\\_IsU](https://youtu.be/BOGBPTh_IsU) (Color figure online)

infeasible iterations ensure convergence from remote initial guesses through a balance between optimality and feasibility.

Finally, we would like to emphasize that the feasibility-driven mode of BOX-FDDP can not only solve tasks that BOX-DDP<sup>+</sup> is unable to solve, but also *improve* the solutions of tasks that BOX-DDP<sup>+</sup> is able to solve. For instance, consider the quadcopter tasks GOAL, NARROW, and LOOP: In the accompanying video, we show that our approach gen-

erates concise and smooth quadcopter trajectories, whereas BOX-DDP<sup>+</sup> generates jerky motions and with unnecessary loops, due to early projection of the control commands.

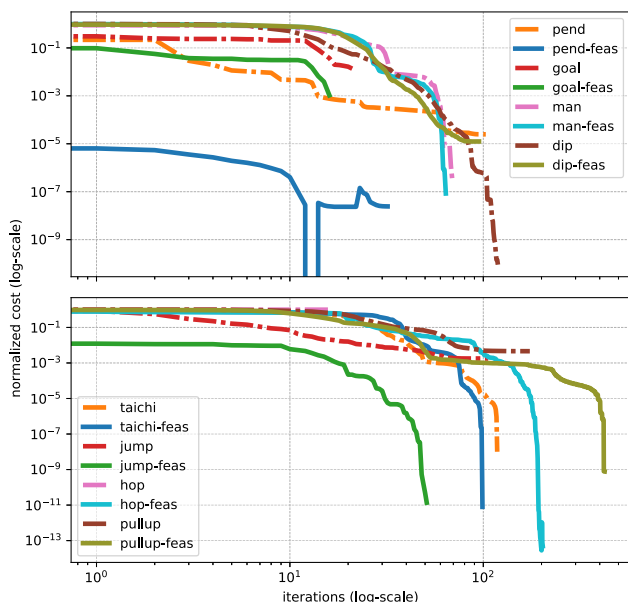
#### 4.2.2 Gap contraction and nonlinearities

We observed that the gap contraction rate is highly influenced by the nonlinearities of the system dynamics (see Fig. 5).

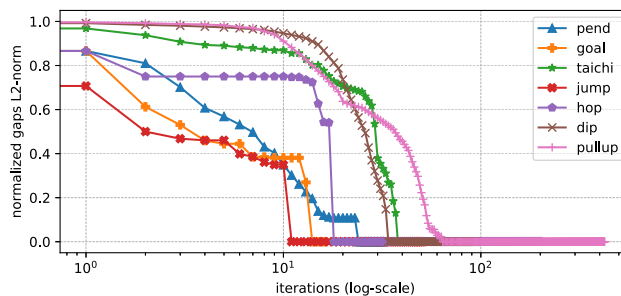
**Table 1** Number of iterations, total cost, and average total computation time over 100 trials

Problems	BOX-DDP <sup>+</sup>			BOX-FDDP (feas)		
	Iter.	Cost	Time (s.)	Iter.	Cost	Time (s.)
PEND	105	4.4292	0.2473	34	0.4997	0.0721
GOAL	23	0.0764	0.0913	18	0.0072	0.0775
LOOP	133	6.7211	0.9144	56	0.6444	0.3982
NARROW	70	1.9492	0.6781	35	0.4577	0.3136
MAN	71	4.6193	11.428	66	4.6193	7.7417
TAICHI	120	6.8184	38.482	101	6.8184	21.909
JUMP	108	1.34e5	✗	53	6.67e4	0.8226
HOP	17	1.3e12	✗	205	1.91e4	19.844
DIP	121	34.2	106.9	97	34.2	107.8
PULLUP	176	276.87	✗	426	146.29	422.5

✗ algorithm does not find a solution



**Fig. 4** Cost and convergence comparison for different optimal control problems. BOX-FDDP outperformed BOX-DDP<sup>+</sup> in all the cases: (top) double pendulum (PEND), quadcopter navigation (QUAD), and whole-body manipulation (MAN), Talos dipping on parallel bars (DIP); and (bottom) whole-body balance (TAICHI), quadrupedal jumping (JUMP), quadrupedal hopping (HOP), Talos’ pull-up workout (PULLUP). BOX-FDDP (\*-feas) solved the problem with fewer iterations and often with lower cost than BOX-DDP<sup>+</sup>. Furthermore, BOX-DDP<sup>+</sup> failed to solve some of the hardest problems: i.e., quadrupedal jumping and hopping, Talos’ pull-up task. Our algorithm showed a large basin of attraction for local optimum as it is less sensitive to poor initialization compared with BOX-DDP<sup>+</sup>. We use the same initial guess for both cases: BOX-FDDP (\*-feas) and BOX-DDP<sup>+</sup> (Color figure online)



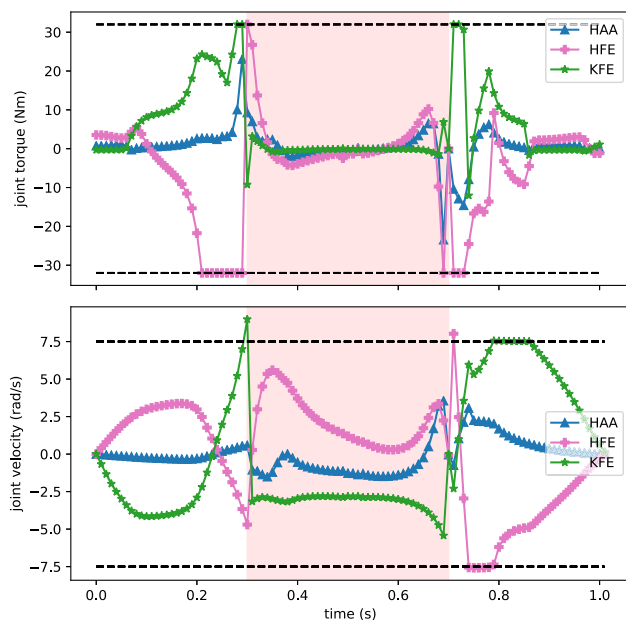
**Fig. 5** Gap contraction of BOX-FDDP for different optimal control problems. For all the cases, the gaps were open for the first several iterations. The gap contraction rate varies according to the accepted step-length. Smaller contraction rates, during the first iterations, appeared in very nonlinear problems (TAICHI, MAN, HOP, and JUMP), because of the larger error of the search direction (Color figure online)

When compared to the dynamics, the nonlinearities of the task often have a smaller effect (e.g., DIP vs PULLUP). Indeed, the gap contraction speed followed the order: humanoid, quadruped, double pendulum, and quadcopter.

Propagation errors due to the dynamics linearization have an important effect on the algorithm progress as Riccati recursions maintain a local quadratic approximation of the value function. The prediction of the expected improvement is indeed more accurate for systems with less nonlinearities, which is the reason why the algorithm tends to accept larger steps that result in higher gap contractions. Indeed, the effect of having a feasibility search is more significant in problems with very nonlinear dynamics as it reduces the total cost faster due to the nonlinearity distribution motivated in Sect. 2.1.2 (see Figs. 4 and 5). It also produces a low cost reduction rate during the gap contraction phase as the algorithm is first focusing on achieving dynamic feasibility.

### 4.2.3 Highly-dynamic and complex maneuvers

The BOX-FDDP algorithm can solve a wide range of motions: from unstable and consecutive hops to aggressive and complex motor behaviors. In Fig. 6, we show the joint torques and velocities of a single leg for the ANYmal’s jumping problem (depicted in Fig. 3c). The motion consisted of three phases: jumping (0–300 ms), flying (300–700 ms), and landing (700–1000 ms). We used 0.7 as a friction coefficient and reduced the real joint limits of the ANYmal robot: from 40 to 32 N m (torque limits) and from 15 to 7.5 rad/s (velocity limits). Thus, generating a 30 cm jump becomes a very challenging task. Furthermore, in this experiment, the velocity limit violations appeared since we used quadratic penalization (with constant weights) to enforce them, and the swing phases are likely too short for such a large jump. Note that if we use constant weights, it might turn out that, for some cases, these weights are not big enough. Nonetheless, we only encountered these violations in very constrained prob-



**Fig. 6** Joint torques and velocities for the ANYmal jumping maneuver. (top) Generated torques of the LF joints and its limits (32 N m); (bottom) Generated velocities of the LF joint and its limits (7.5 rad/s). The red region describes the flight phase. Note that HAA, HFE, and KFE are the abduction/adduction, hip flexion/extension and knee flexion/extension joints, respectively (Color figure online)

lems. For instance, we did not find velocity violations for the walking, trotting, pacing, and bounding gaits (reported in the accompanying video). For these cases, BOX-FDDP converged approximately with the same number of iterations achieved by the FDDP solver (i.e., a fully unconstrained case). This is due to the fact that the robot could generate those gaits without reaching its torque limits.

Surprisingly, naturally looking behaviors emerged during the computation of the DIP and PULLUP problems on the Talos humanoid robot. We did not include any heuristic that could have helped the algorithm to generate these undefined behaviors. For instance, the balancing and leg-crossing on the bars emerge if we allocate a significant amount of time in that motion phase. Similarly, the pull-up motion emerges if we significantly increase the maximum torque limits on the arms.

### 4.3 Experimental trials

We demonstrated the capabilities of our BOX-FDDP algorithm in a model predictive control application for the ANYmal C quadruped robot. The BOX-FDDP algorithm computed reference motions in real-time and its efficiency enables our predictive controller to run in excess of 100 Hz computation frequency with a horizon of 1.25 s. Figure 7 shows snapshots of the forward trotting gait computed with the BOX-FDDP algorithm. It also displays the contact-force tracking and updates of the swing-foot reference trajectories.

In this experiment, we predefined the timings for the swing-foot motions and the foothold locations.

### 4.4 Box-FDDP vs a direct transcription formulation

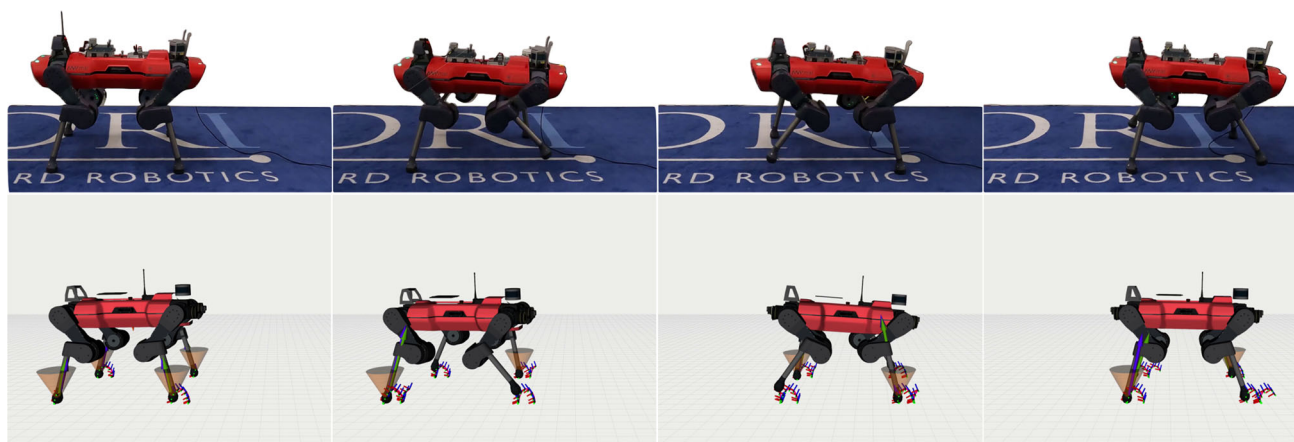
We formulated an efficient direct transcription problem with dynamics defects constraints in each node. The formulation is conceptually similar to our feasible descent direction introduced in Sect. 3.3. We transcribed the dynamics using a symplectic Euler scheme, the same integration scheme also used in BOX-FDDP. We solved the direct transcription problem using different optimization algorithms provided by KNITRO (Byrd et al., 2006). These available algorithms are: Interior/Direct (Waltz et al., 2006), Interior/CG (Byrd et al., 1999), active set (Byrd et al., 2003) and sequential quadratic programming (SQP) algorithms. Below we briefly describe each of KNITRO’s algorithms, and then report the comparison results with BOX-FDDP.

The Interior/Direct (KNITRO-IDIR) algorithm replaces the NLP problem with a series of barrier sub-problems. In each iteration, it solves the primal-dual KKT problem using a line search procedure.<sup>10</sup> Instead, Interior/CG (KNITRO-ICG) solves the primal-dual KKT problem using a projected conjugate gradient method. This method uses exact second derivatives, without explicitly storing the Hessian matrix, through a tailored trust region procedure. Interior/Direct also invokes this trust region procedure if the line search iteration converges to a non-stationary point (Waltz et al., 2006). In contrast to the interior point methods, the active set algorithm (KNITRO-SLQP) replaces the NLP problem with a sequence of quadratic programs to form a sequential linear-quadratic programming algorithm. This algorithm selects a set of active constraints in each iteration, and produces a more exterior path (i.e., along the constraints) to the solution. Finally, KNITRO’s SQP algorithm (KNITRO-SQP) is also an active set method designed for small to medium scale problems with expensive function evaluations. Both active set approaches are often preferable to interior point methods on small- to medium-sized problems when we can provide a good initial guess (Nocedal & Wright, 2006). However, the problems in robotics are often large with many inequality constraints. Indeed, the benefits of interior point methods have been pointed out in the context of direct methods (Pardo et al., 2016).

#### 4.4.1 Optimality versus feasibility

We compared the total cost, number of iterations, and total computation time against the different algorithms implemented in KNITRO over 100 trials. For the comparison, we

<sup>10</sup> For more details about interior point methods, the authors suggest the reader to see Nocedal and Wright (2006).



**Fig. 7** Snapshots of hardware experiments using BOX-FDDP on the ANYmal C quadruped: Here, the robot completes a trotting gait in a model predictive control fashion. Our method is able to resolve the optimal control problem for a 1.25 s horizon in  $\approx 10$  ms on laptop hardware. The top figure shows snapshots of the ANYmal robot trotting for-

ward, while the bottom figure displays the measured and desired contact forces and reference swing trajectories sent to the controller. To watch the video, click on the figure or see [https://youtu.be/bOGBPTb\\_IsU](https://youtu.be/bOGBPTb_IsU). For more details about the predictive control algorithm, please refer to Mastalli et al. (2022)(Color figure online)

solved the double pendulum problem (PEND), as it requires discovery of a swing-up maneuver. With this, we can clearly compare the trade-off between optimality and feasibility across the different algorithms. Note that, as described earlier, we used a single-thread for both KNITRO and BOX-FDDP despite our algorithm supporting multithreading.

Table 2 reports three different formulations used in the KNITRO algorithms. The first one (PEN) emulates exactly the optimal control formulation used in BOX-FDDP, i.e., control constraints, regularization terms, and a terminal quadratic cost. The second case (REGCONST) uses a terminal constraint to impose the desired up-ward position together with the regularization terms. The third case (CONST) uses only the ter-

minal constraints. Below we summarize the obtained results for each formulation.

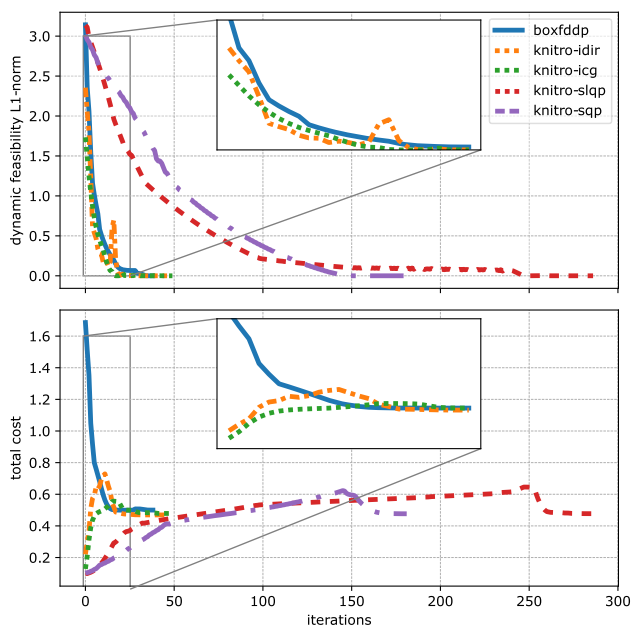
BOX-FDDP converges faster (w.r.t. time) than KNITRO algorithms in all of the above formulations. However, KNITRO-ICG is as fast as our approach with the CONST formulation. On the other hand, when it comes to optimality, KNITRO produces more optimal solutions if we use the REGCONST formulation. Indeed, in our experience, KNITRO generally has a better behavior when the formulation is dominated by constraint functions. Note that we do not report the cost values for the CONST formulation as this boils down to a feasibility problem, i.e., a problem with only constraints.

We used the REGCONST formulation to be able to compare both: cost and feasibility evolution. The dynamic infea-

**Table 2** Box-FDDP vs Knitro in double pendulum problem

Case	Algorithms	Cost	Iteration	Total time (sec.)
	BOXFDDP	0.4997	34	<b>0.0721</b>
PEN	KNITRO-IDIR	2.1094	1064	8.6704
REGCONST	KNITRO-IDIR	<b>0.4693</b>	47	0.2766
CONST	KNITRO-IDIR	–	20	0.1387
PEN	KNITRO-ICG	0.5441	59	0.2390
REGCONST	KNITRO-ICG	0.4787	50	0.1875
CONST	KNITRO-ICG	–	<b>16</b>	0.0733
PEN	KNITRO-SLQP	0.5978	165	1.2595
REGCONST	KNITRO-SLQP	0.4773	287	1.6704
CONST	KNITRO-SLQP	–	27	0.2359
PEN	KNITRO-SQP	0.5889	114	46.88
REGCONST	KNITRO-SQP	0.4767	182	12.233
CONST	KNITRO-SQP	–	22	8.7736

Lowest cost, iteration and total time are given in bold



**Fig. 8** Dynamic feasibility and cost evolution for the double pendulum problem. BOX-FDDP had a similar evolution (dynamic feasibility and cost) to the interior point algorithms implemented in KNITRO. KNITRO produced lower cost solutions, but the computational burden is often much higher

sibility decreased monotonically for all the algorithms as plotted in Fig. 8 (top). BOX-FDDP shows a fast resolution of the dynamics feasibility as interior point algorithms that, generally speaking, require fewer iterations than active set approaches (Hei et al., 2008). The cost evolution is also similar to the interior point algorithms, where the total costs are reported as REGCONST cases in Table 2.

#### 4.4.2 Computation time

BOX-FDDP had a better runtime performance than the KNITRO algorithms for the double pendulum problem (cf. Table 2). However, to answer the runtime performance scalability to higher-dimensional optimal control problems, we analyzed the problem of generating a forward jumping maneuver with the ANYmal robot (i.e., JUMP).

We used the same phase timings, joint limits and friction coefficient reported in Sect. 4.2.3. The results reported with KNITRO and BOX-FDDP cases are based on slightly different optimal control formulations. The idea is to define the most suitable formulation for each algorithm. For instance, we use quadratic penalization terms to impose the desired foothold placement, joint velocity limits and friction cone constraints for the BOX-FDDP algorithm. Instead, for the KNITRO algorithms, we substitute these penalization terms by general equality and inequality constraints. To further reduce the computation time of KNITRO cases, we also impose a constraint for the terminal position of the trunk. Note that we did

**Table 3** Runtime performance for a forward jumping maneuver over 100 trials

Trunk height (m.)	KNITRO-IDIR		BOX-FDDP	
	Iter.	Time (sec.)	Iter.	Time (sec.)
0.25	<b>11</b>	1.5039	59	<b>0.9061</b>
0.30	<b>10</b>	1.3900	72	<b>1.1409</b>
0.35	<b>15</b>	1.9565	59	<b>0.9030</b>
0.40	<b>18</b>	2.3864	49	<b>0.7849</b>
0.45	<b>16</b>	2.1320	51	<b>0.8574</b>

Bold values indicate Knitro has fewer iterations but our Box-FDDP is faster

not include any cost term since it negatively affects the convergence rate of KNITRO, i.e., we treated it as a feasibility problem.

Table 3 reports the runtime performance over 100 trials for the KNITRO-IDIR algorithm only. The other methods (i.e., KNITRO-ICG, KNITRO-SLQP and KNITRO-SQP) were unable to solve this problem. We used 5 different initial trunk heights, and we initialized the algorithms using their corresponding joint posture (as described in Appendix A-C) and no controls (i.e.,  $\mathbf{u}_s^0 = \{\mathbf{0}, \dots, \mathbf{0}\}$ ). As in the double pendulum case, BOX-FDDP also solved this problem faster than KNITRO algorithms, even though it required a significant number of extra—computationally inexpensive—iterations. We suspect that this increment in the number of iterations is due to the use of penalization terms in the contact placement, friction cone and state limits constraints.

#### 4.4.3 Robustness against different initial guesses

We compared the robustness against different initial guesses for the double pendulum (PEND) and quadrupedal jump (JUMP) problems. In both problems, we generated random joint postures—around the nominal state—and used them to define an initial guess for the state trajectory  $\mathbf{x}_s^0$ . In addition to the robot’s joint postures, we also generated random joint velocities around the *zero-velocity* condition for the double pendulum case only. We used this single random posture and velocity for each node in  $\mathbf{x}_s^0$ , and initialized the control sequence with zeros. We used the most suitable formulations for BOX-FDDP and KNITRO algorithms as justified above.

Table 4 reports the number of successful resolutions over 100 trials. We considered a problem to be successfully solved if the gradient of the BOX-FDDP or the feasibility of the KNITRO algorithms are lower than  $5 \times 10^{-5}$  (absolute feasibility tolerance). Note that this includes the cases where KNITRO found a feasible approximate solution.<sup>11</sup> Furthermore, we considered a problem resolution unsuccessful if the prob-

<sup>11</sup> For further detail, we suggest the reader to consult the KNITRO manual: [https://www.artelys.com/docs/knitro/3\\_referenceManual.html](https://www.artelys.com/docs/knitro/3_referenceManual.html).

**Table 4** Percentage of successful resolutions from random initial guesses

Algorithms	PEND		JUMP	
	$\ 1\ _\infty$	$\ 100\ _\infty$	$\ 0.0005\ _\infty$	$\ 0.005\ _\infty$
BOX-FDDP	74 %	<b>44 %</b>	<b>99 %</b>	<b>99 %</b>
KNITRO-IDIR	<b>100 %</b>	1 %	51 %	13 %
KNITRO-ICG	<b>100 %</b>	0 %	<b>X</b>	<b>X</b>
KNITRO-SLQP	95 %	0 %	<b>X</b>	<b>X</b>
KNITRO-SQP	92 %	0 %	<b>X</b>	<b>X</b>

**X** algorithm does not find a solution

lem does not converge within 70 s, which is enough time as we can see above. For each problem, we used two different maximum values of the random initialization, which their maximum magnitude are described using the  $\ell^\infty$  norm (i.e.,  $\|\cdot\|_\infty$ ). We added this additive noise to the default initial guess (described in Appendix A) used for the state trajectory.

As expected, the KNITRO interior point methods performed better than the active set ones. For the double pendulum problem, the interior point algorithms (i.e., KNITRO-IDIR and KNITRO-ICG) perform better than BOX-FDDP if the warm-starting point is close to the initial condition. Despite that, BOX-FDDP shows more robustness to initial guesses as its percentage of successful resolutions is consistent. Furthermore, we observed a significant increment in the number of successful resolutions for the JUMP problem. Indeed, KNITRO was not able to solve this problem at all for random magnitudes bigger than  $\|0.01\|_\infty$ .

### 4.5 Box-FDDP, Box-DDP, and squashing approach in nonlinear problems

To evaluate the numerical performance of having two modes, we compared the BOX-FDDP (with two modes depending on the dynamics feasibility), Box-DDP<sup>+</sup> (using a single mode) and DDP<sup>+</sup> with a squashing function (using a single mode) for three scenarios with the IRIS quadcopter: reaching goal (GOAL), looping maneuver (LOOP), and traversing a narrow passage (NARROW). We used a sigmoidal element-wise squashing function of the form:

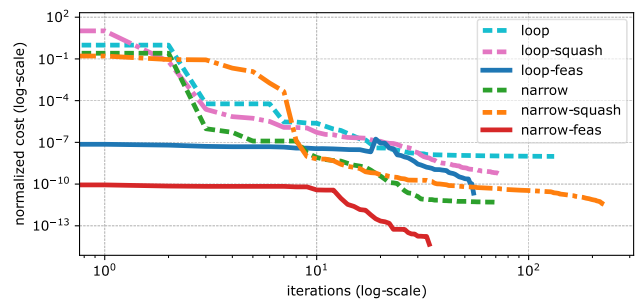
$$s^i(\mathbf{u}^i) = \frac{1}{2} \left( \mathbf{u}^i + \sqrt{\gamma^2 + (\mathbf{u}^i - \underline{\mathbf{u}}^i)^2} \right) + \frac{1}{2} \left( \bar{\mathbf{u}}^i - \sqrt{\gamma^2 + (\mathbf{u}^i - \bar{\mathbf{u}}^i)^2} \right)$$

in which the sigmoid is approximated through two smooth-abs functions,  $\gamma$  defines its smoothness, and  $\underline{\mathbf{u}}^i, \bar{\mathbf{u}}^i$  are the element-wise lower and upper control bounds, respectively. We introduced this squashing function on the system con-

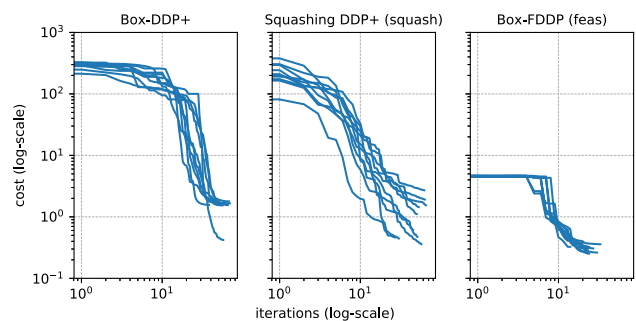
trols as:  $\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k, \mathbf{s}(\mathbf{u}_k))$ . We used  $\gamma = 2$  for all the experiments presented in this section.

Figure 9 shows that BOX-FDDP converged faster than the other approaches. As reported in the accompanying video, BOX-FDDP did not generate undesired loops and jerky motions as in the other cases. Indeed, the solutions with BOX-FDDP have the lowest cost values (cf. Table 1). We also observed that the squashing approach often converges sooner compared to BOX-DDP<sup>+</sup>. The main reason is due to the early saturation of the controls performed by BOX-DDP<sup>+</sup>.

In Fig. 10, we show the cost evolution for 10 different initial configurations of the reaching goal task. The target and initial configurations are  $(3, 0, 1)$  and  $(-0.3 \pm 0.6, 0, 0)$  m, respectively. Infeasible iterations, in BOX-FDDP, produce a very low cost in the first iterations. The squashing approach is the most sensitive to initial conditions. However, on average, it produces slightly better solutions than BOX-DDP<sup>+</sup>. This is in contrast to the reported results in Tassa et al. (2014), where the performance was analyzed only for the linear-quadratic regulator problem.



**Fig. 9** Cost and convergence comparison for different quadcopter maneuvers: looping (LOOP) and narrow passage traversing (NARROW). BOX-FDDP (\*-feas) outperformed both BOX-DDP<sup>+</sup> and DDP<sup>+</sup> with squashing function (\*-squash)(Color figure online)



**Fig. 10** Costs associated for 10 different initial configurations of reaching goal tasks. BOX-FDDP converges earlier and with lower total cost than Box-DDP<sup>+</sup> and DDP<sup>+</sup> with squashing function. The performance of the squashing function approach exhibits a high dependency on the initial condition(Color figure online)

## 5 Conclusion

We proposed a feasibility-driven approach whose search is primarily driven by the dynamics feasibility of the optimal control problem. The dynamically-infeasible iterations, which mimic a direct multiple shooting approach, allowed us to solve a wide range of optimal control problems despite it being provided with a poor initialization. The benefits of our approach are crystallized over a set of athletic and highly-dynamic maneuvers computed for the Talos humanoid and the ANYmal quadruped robots, respectively. Its improvement on the basin of attraction for a good local optimum has been a key factor to optimize such kind of complex maneuvers while considering the robot's full rigid body dynamics, joint limits and friction cone constraints. Indeed, BOX-FDDP has shown an increment in the robustness against different initial guesses compared with advanced KNITRO algorithms.

We have provided evidence that our algorithm produces descent search directions. For instance, we have observed that the feasibility contraction decreases monotonically as often happens in the advanced nonlinear programming algorithms available in KNITRO. Our approach has also shown to quickly reduce the dynamic infeasibility as observed in the most competitive KNITRO algorithms. A similar effect is observed in the cost evolution as well. Our results suggest that the gap contraction rate is influenced by the nonlinearities of the system dynamics.

The runtime performance of BOX-FDDP is often superior to direct transcription solved using state-of-the-art KNITRO algorithms despite the increase in number of iterations due to the use of quadratic penalization terms. However, when comparing the computation time per iteration, our approach is between 2 to 10 times faster, which makes it suitable for model predictive control applications. Indeed, we demonstrated that our BOX-FDDP algorithm can generate trotting gaits on the ANYmal C robot in a predictive control fashion. One additional remark is that we have not considered the runtime reduction due to code generation support in CROCODDYL via CPPADCODEGEN (Leal, 2011) and CPPAD (Bell, 2012). According to our experience, code generation can lead to a computation time reduction between 30% to 60% as can be seen in our public benchmarks (Mastalli et al., 2019).

We have developed and reported the results for a wide range of optimal control problems in robotics. These problems cover an important part of the spectrum of robotics applications. The comparison across all these problems is unusual, and it represents an important contribution to the research community since we have open sourced many of these examples, as well as the BOX-FDDP algorithm, in the CROCODDYL repository (Mastalli et al., 2019). Our feasibility-driven approach enabled model predictive control applications on the ANYmal C robot, however, it can

potentially be used in other applications such as in humanoid robotics (Eßer et al., 2021), robot co-design (IROS, 2022), and learning (Lembono et al., 2020).

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s10514-022-10061-w>.

**Acknowledgements** The authors are grateful to Matt Timmons-Brown and Vladimir Ivan, from the University of Edinburgh, for the production of the audio material of our video and the fruitful discussions on the study of the effects of our feasibility-driven approach, respectively.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## Appendix A Optimal control problems

We divide the optimal control problems into four subsections: double pendulum, quadcopter, quadruped and humanoid robots.

### A Double pendulum (pend)

The goal is to swing from the stable to the unstable equilibrium points, i.e. from down-ward to up-ward positions, respectively. To increase the problem complexity, the double pendulum (with weight of  $\approx 4.5$  N) has a single actuated joint with small range of control (from  $-5$  to  $5$  N m, largely insufficient for a static execution of the trajectory). The time horizon is 1 s with 100 nodes. We define a quadratic cost function, for each node, that aims to reach the up-ward position. For the running and terminal nodes, we use the weight values of  $2 \times 10^{-4}$  and  $2 \times 10^4$ , respectively. Additionally, we provide state and control regularization terms. To inspect the algorithm capabilities, we *do not provide* an initial guess (i.e., zeros for the states and controls), thus the swing-up strategy is discovered by the solver itself. Finally, we implement the cost, dynamics, and their analytical derivatives.

### B Quadcopter

We consider three tasks for the IRIS quadcopter: reaching goal (GOAL), looping maneuver (LOOP), and traversing a narrow passage (NARROW). We define different way-points to

describe the tasks, where each way-point specifies the desired pose and velocity. The way-points are described through cost functions in the robot placement and velocity at specific instants of the motion. These cost functions are quadratic term with  $10^2$  as the weight value for both: pose and velocity. We also include quadratic regularization terms for the state and control, their weight values are  $10^{-5}$  and  $10^{-1}$ , respectively. The vehicle pose is described as a  $\mathbb{SE}(3)$  element, which allows us to consider any kind of motion such as *looping* maneuvers. Control inputs are considered to be the thrust produced by the propellers, which can vary within a range from 0.1 to 10.3 N each. We compute the dynamics using the articulated body algorithm (ABA) algorithm<sup>12</sup>, and the analytical derivatives are calculated as described in Carpentier and Mansard (2018). We integrate the dynamics with a fixed time step of 10 ms. The solution is computed from a cold-start of the algorithm. As in the double pendulum case, we do not provide an initial guess to the solvers.

### C Aggressive jump, unstable hopping, and various gaits

We use the ANYmal quadruped robot to generate a wide range of motions—jumping, hopping, walking, trotting, pacing, and bounding. We deliberately reduce the torque and velocity limits to 32 N m and 7.5 rad/s, respectively, which intentionally increases the complexity of the jumping task (JUMP). We use a quadratic barrier to penalize the joint velocities and the contact forces that are outside the limits. The unstable hopping problem (HOP) has a long horizon: 7.14 s with 714 nodes. It includes 10 hops in total with a phase that switches the feet in contact. Instead, the other problems have a horizon between 0.7–1 s with 70 – 100 nodes. We define quadratic terms, with a weight value of  $10^6$ , to track the desired swing-foot placements for each case. We regularize the state trajectory around the robot's default configuration. We use a friction coefficient of 0.7. Furthermore, we formulate a multi-phase optimal control problem using rigid contact dynamics, and their analytical derivatives, as described in Budhiraja et al. (2018), Mastalli et al. (2020), respectively. During a contact transition, we employ the impulse dynamics with analytical derivatives as also described in Mastalli et al. (2020). Our contact dynamics model also includes the Baumgarte gains, which are needed to numerically stabilize differential algebraic equations (Baumgarte, 1972); we use the values: (0, 50). We initialize the solver using the default posture and the quasi-static torques for each node of the initial guess trajectories. The default posture defines the standing position of the robot; it does not provide any relevant information for a specific maneuver (e.g., JUMP). The quasi-static

torques describe the forces required to cancel the effects of gravity subject to the robot's default posture.

### D Whole-body manipulation and balance

We consider four problems for the Talos humanoid robot: whole-body manipulation (MAN), hand control while balancing in single leg (TAICHI), dip on parallel bars (DIP) and a pull-up bar task (PULLUP). For the dip and pull-up workouts, we define the real torque limits for the leg and torso joints.<sup>13</sup> Additionally, we consider joint position and velocity limits in each scenario through a quadratic barrier. Both TAICHI, DIP and PULLUP tasks are divided into three phases; for the TAICHI task: manipulation, standing on one foot, and balancing; for the DIP and PULLUP bar task: grasping the bar, workout, and landing on ground. DIP and PULLUP problems have a horizon of 12 s with 400 nodes. The horizon of the whole-body manipulation and TAICHI problem are 3 and 6 s with 60 and 120 nodes, respectively. We define different Baumgarte gains for the hands and feet, i.e. (0, 20) and (0, 40), respectively. We use the contact-placement cost functions for both: feet and hands. Indeed, behaviors such as dip/pull-up balancing, pull-up motion, or the leg-crossing strategy emerge from our optimization algorithm since we do not describe the center of mass and swing-contact motions. Finally, the regularization terms, dynamics and initialization strategy are the same that we use for the ANYmal cases. For the details about the ANYmal problems see above (Section A–C).

### References

- A versatile co-design approach for dynamic legged robots, In *IEEE/RSJ international conference on intelligence robotics and systems. (IROS)*.
- Aceituno-Cabezas, B., Mastalli, C., Dai, H., Focchi, M., Radulescu, A., Caldwell, D. G., Cappelletto, J., Grieco, J. C., Fernandez-Lopez, G., & Semini, C. (2017). Simultaneous contact, gait and motion planning for robust multi-legged locomotion via mixed-integer convex optimization, In *IEEE robotics and automation letters (RA-L)*.
- Albersmeyer, J., & Diehl, M. (2010). The lifted Newton method and its application in optimization, *SIAM Journal on Optimization*.
- Baumgarte, J. (1972). Stabilization of constraints and integrals of motion in dynamical systems. *Computer Methods in Applied Mechanics and Engineering*, 1, 1–16.
- Bell, B. (2012). CppAD: A package for differentiation of C++ algorithms, In *Computational infrastructure for operations research*. [Online]. Available: <http://www.coin-or.org/CppAD/>
- Bellman, R. E. (1954). *The theory of dynamic programming*, *Bulletin of the American Mathematical Society*.
- Bertsekas, D. P. (1982). Projected newton methods for optimization problems with simple constraints. *SIAM Journal on Control and Optimization*, 20, 221–246.
- Betts, J. T. (2009). *Practical methods for optimal control and estimation using nonlinear programming* (2nd ed.). Cambridge University Press.

<sup>12</sup> For more details about the ABA algorithm see Featherstone (2007).

<sup>13</sup> Talos' arms are not strong enough to support its own weight.

- Budhiraja, R., Carpentier, J., Mastalli, C., & Mansard, N. (2018). Differential dynamic programming for multi-phase rigid contact dynamics, In *IEEE international conference on human and robotics (ICHR)*.
- Byrd, R. H., Gould, N. I. M., Nocedal, J., & Waltz, R. A. (2003). An algorithm for nonlinear optimization using linear programming and equality constrained subproblems. *Mathematical of Programming*, 100, 27–48.
- Byrd, R. H., Hribar, M. E., & Nocedal, J. (1999). An interior point algorithm for large-scale nonlinear programming. *SIAM Journal on Optimization*, 9, 877–900.
- Byrd, R. H., Nocedal, J., & Waltz, R. A. (2006). KNITRO: An integrated package for nonlinear optimization, In *Large scale nonlinear optimization*, (pp. 35–59).
- Carpentier, J., & Mansard, N. (2018). Analytical derivatives of rigid body dynamics algorithms, In *Robotics: science and systems (RSS)*.
- Carpentier, J., Tonneau, S., Naveau, M., Stasse, O., & Mansard, N. (2016). A versatile and efficient pattern generator for generalized legged locomotion, In *International conference on robotics and automation (ICRA)*.
- Di Carlo, J., Wensing, P. M., Katz, B., Bledt, G., & Kim, S. (2018). Dynamic locomotion in the MIT cheetah 3 through convex model-predictive control, In *IEEE/RSJ international conference on intelligent robots and systems (IROS)*.
- Diehl, M., Bock, H. G., Diedam, H., & Wieber, P. -B. (2006). Fast direct multiple shooting algorithms for optimal robot control, In *Proceedings of Berlin Heidelberg: Fast Motions in Biomechanics Robot*. Springer.
- Eßer, J., Kumar, S., Peters, H., Bargsten, V., Fernandez, J. d. G., Mastalli, C., Stasse, O., & Kirchner, F. (2021). Design, analysis and control of the series-parallel hybrid RH5 humanoid robot, In *IEEE International conference on human and robotics (ICHR)*.
- Farshidian, F., Jelavic, E., Satapathy, A., Gifftaler, M., & Buchli, J. (2017). Real-time motion planning of legged robots: A model predictive control approach, In *IEEE international conference on human and robotics (ICHR)*.
- Featherstone, R. (2007). *Rigid body dynamics algorithms*. Springer-Verlag.
- Fletcher, R. (1971). *A modified Marquardt subroutine for non-linear least squares*. Technical Report.
- Frese, U. (2008). *A framework for sparse non-linear least squares problems on manifolds*, Ph.D. dissertation, Universität Bremen.
- Gabay, D. (1982). Minimizing a differentiable function over a differential manifold. *Journal of Optimization Theory and Applications*, 37, 177–219.
- Gifftaler, M., Neunert, M., Stäubel, M., Buchli, J., & Diehl, M. (2018). A family of iterative gauss-newton shooting methods for nonlinear optimal control, In *IEEE/RSJ international conference on intelligent robots and systems (IROS)*.
- Gill, P. E., Murray, W., & Saunders, M. A. (2005). *SNOPT: An SQP algorithm for large-scale constrained optimization*, SIAM Rev.
- Hargraves, C. R., & Paris, S. W. (1987). Direct trajectory optimization using nonlinear programming and collocation. *Journal of Guidance*, 10(4), 338–342.
- Harwell Subroutine Library, AEA Technology, Harwell, Oxfordshire, England. A catalogue of subroutines, <http://www.hsl.rl.ac.uk/>.
- Hei, L., Nocedal, J., & Waltz, R. A. (2008). A numerical study of active-set and interior-point methods for bound constrained optimization, In *Modeling, Simulation and Optimization of Complex Processes*.
- Howell, T. A., Jackson, B., & Manchester, Z. (2019). ALTRO: a fast solver for constrained trajectory optimization, In *IEEE/RSJ international conference on intelligent robots and systems (IROS)*.
- Kirk, D. E. (1998). *Optimal control theory - An introduction*. Dover Publications.
- Koenemann, J., Prete, A. Del, Tassa, Y., Todorov, E., Stasse, O., Bennewitz, M., & Mansard, N. (2015). Whole-body model-predictive control applied to the HRP-2 humanoid, In *IEEE/RSJ international conference on intelligent robots and systems (IROS)*.
- Lantoine, G., & Russell, R. (2008). A hybrid differential dynamic programming algorithm for robust low-thrust optimization, In *AIAA/AAS astrodyn special conference and exhibition*.
- Leal, J. “CppADCodeGen,” <https://github.com/joaleal/CppADCodeGen/>, 2011–2020.
- Lombono, T. S., Mastalli, C., Fernbach, P., Mansard, N., & Calinon, S. (2020). Learning how to walk: warm-starting optimal control solver with memory of motion, In *IEEE international conference on robotics and automation (ICRA)*.
- Li, W., & Todorov, E. (2004). Iterative linear quadratic regulator design for nonlinear biological movement systems, In *ICINCO*.
- Liao, L. zhi, & Shoemaker, C. A. (1992). *Advantages of differential dynamic programming over Newton’s method for discrete-time optimal control problems*, Cornell University, Technical Report.
- Marti-Saumell, J., Solà, J., Mastalli, C., & Santamaria-Navarro, A. (2020). Squash-box feasibility driven differential dynamic programming, In *IEEE/RSJ international conference on intelligent robots and systems (IROS)*.
- Mastalli, C., Budhiraja, R., & Mansard, N. (2019). Crocodyl: a fast and flexible optimal control library for robot control under contact sequence, <https://github.com/loco-3d/crocodyl>.
- Mastalli, C., Budhiraja, R., Merkt, W., Saurel, G., Hammoud, B., Naveau, M., Carpentier, J., Righetti, L., Vijayakumar, S., & Mansard, N. (2020). Crocodyl: An efficient and versatile framework for multi-contact optimal control, In *IEEE international conference on robotics and automation (ICRA)*.
- Mastalli, C., Focchi, M., Havoutis, I., Radulescu, A., Calinon, S., Buchli, J., Caldwell, D. G., & Semini, C. (2017). Trajectory and foothold optimization using low-dimensional models for rough terrain locomotion, In *IEEE robotics and automation letters (ICRA)*.
- Mastalli, C., Havoutis, I., Focchi, M., Caldwell, D. G., & Semini, C. (2020). Motion planning for quadrupedal locomotion: coupled planning, terrain mapping and whole-body control, *IEEE Transactions on Robotics (T-RO)*.
- Mastalli, C., Merkt, W., Xin, G., Shim, J., Mistry, M., Havoutis, I., & Vijayakumar S. (2022). Agile maneuvers in legged robots: a predictive control approach, [arXiv:2203.07554](https://arxiv.org/abs/2203.07554).
- Mayne, D. (1966). A second-order gradient method for determining optimal trajectories of non-linear discrete-time systems, *International Journal of Control*.
- Merkt, W., Ivan, V., & Vijayakumar, S. (2018). Leveraging pre-computation with problem encoding for warm-starting trajectory optimization in complex environments, In *IEEE/RSJ international conference on intelligent robots and systems (IROS)*.
- Neunert, M., Stauble, M., Gifftaler, M., Bellicoso, C. D., Carius, J., Gehring, C., Hutter, M., & Buchli, J. (2018). Whole-body nonlinear model predictive control through contacts for quadrupeds, *IEEE Robotics and Automation Letters (RA-L)*.
- Nocedal, J., & Wright, S. J. (2006). *Numerical optimization* (2nd ed.). Springer.
- Ohno, K. (1978). A new approach to differential dynamic programming for discrete time systems. *IEEE Transactions on Automatic Control*, 23(1), 37–47.
- Pardo, D., Moller, L., Neunert, M., Winkler, A. W., & Buchli, J. (2016). Evaluating direct transcription and nonlinear optimization methods for robot motion planning, *IEEE Robotics and Automation Letters (RA-L)*.
- Tassa, Y. (2011). *Theory and implementation of biomimetic motor controllers*, Ph.D. dissertation, Hebrew University of Jerusalem.
- Tassa, Y., Erez, T., & Todorov, E. (2012). Synthesis and stabilization of complex behaviors through online trajectory optimization,” In

*IEEE/RSJ international conference on intelligent robots and systems (IROS).*

- Tassa, Y., Mansard, N., & Todorov, E. (2014). Control-limited differential dynamic programming. In *IEEE international conference on robotics and automation (ICRA)*.
- Wächter, A., & Biegler, L. T. (2006). On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical Programming*.
- Waltz, R. A., Morales, J. L., Nocedal, J., & Orban, D. (2006). An interior algorithm for nonlinear optimization that combines line search and trust region steps. *Mathematical of Programming*, 107, 391–408.
- Wieber, P. -B. (2006). Trajectory free linear model predictive control for stable walking in the presence of strong perturbations, In *IEEE international conference on human robotics (ICHR)*.
- Winkler, A. W., Bellicoso, D. C., Hutter, M., & Buchli, J. (2018). Gait and trajectory optimization for legged systems through phase-based end-effector parameterization, In *IEEE robotics and automation letters (RA-L)*.
- Xie, Z., Liu, C. K., & Hauser, K. (2017). Differential dynamic programming with nonlinear constraints, In *IEEE international conference on robotics and automation (ICRA)*.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Carlos Mastalli** received the M.Sc. degree in mechatronics engineering from the Simón Bolívar University, Caracas, Venezuela, in 2013 and the Ph.D. degree in bio-engineering and robotics from the Istituto Italiano di Tecnologia, Genoa, Italy, in 2017. He is currently an Assistant Professor at Heriot-Watt University, Edinburgh, U.K. He is the Head of the Robot Motor Intelligence (RoMI) Lab affiliated to the National Robotarium and Edinburgh Centre for Robotics. He is also appointed as

Research Scientist at IHMC, USA. Previously, he conducted cutting-edge research in several world-leading labs: Istituto Italiano di Tecnologia (Italy), LAAS-CNRS (France), ETH Zürich (Switzerland), and the University of Edinburgh (UK). His research focuses on building athletic intelligence for robots with legs and arms. To do so, he is working at the intersection of model predictive control, numerical optimization, deep learning, machine learning, whole-body control, and robot co-design.



**Wolfgang Merkt** received the B.Eng.(Hns) degree in mechanical engineering with management and the M.Sc.(R) and Ph.D. degrees in robotics and autonomous systems from the University of Edinburgh, Edinburgh, U.K., in 2014, 2015 and 2019, respectively. He is currently a Postdoctoral Researcher at the Oxford Robotics Institute, University of Oxford with I. Havoutis. During his Ph.D., he worked on trajectory optimization and warm starting optimal control for high-dimensional sys-

tems and humanoid robots under the supervision of S. Vijayakumar. His research interests include fast optimization-based methods for planning and control, loco-manipulation, and legged robots.



**Josep Marti-Saumell** received the B.Sc. degree in industrial engineering (majoring in mechanics) and the M.Sc. degree in automatic control and robotics from the Universitat Politècnica de Catalunya, Barcelona, Spain, in 2013 and 2018, respectively. He is currently a Ph.D. candidate at the Institut de Robòtica i Informàtica Industrial, CSIC-UPC, Barcelona, Spain. His current research interests include optimal control applied to mobile robotics, numerical optimization, and unmanned aerial manipulators.



**Henrique Ferrolho** received his M.Sc. degree in informatics and computing engineering from the University of Porto, Porto, Portugal in 2017. He is currently pursuing a Ph.D. degree in robotics and autonomous systems at the University of Edinburgh under the supervision of S. Vijayakumar. His research interests include robust motion planning, and optimal control of agile robotic systems.



**Joan Solà** received the M.Sc. degree in telecommunication and electronics from the Universitat Politècnica de Catalunya, Barcelona, Spain, and the Ph.D. degree in robotics from the University of Toulouse, Toulouse, France, in 2007. He is currently a CSIC Researcher with the Institut de Robòtica i Informàtica Industrial, CSIC-UPC, Barcelona, Spain. He has also worked in the industry in the renewable energies sector, and was involved in the construction of a manned submarine for depths

up to 1200 m. He has contributed to monocular SLAM, especially in the undelayed initialization of landmarks, and is interested in state estimation for robots with particularly large dynamics and degrees of freedom, such as humanoids and aerial manipulators. His current projects turn around whole-body estimation and control, including multisensor fusion, localization and mapping, machine learning, and model predictive control.



**Nicolas Mansard** received the M.Sc. degree in computer science from the University of Grenoble, Grenoble, France, in 2003 and the Ph.D. degree in robotics from the University of Rennes, Rennes, France, in 2006. He has been a CNRS Researcher since 2009. He was then Postdoctoral Researcher at Stanford University, Stanford, CA, USA with O. Khatib in 2007 and in JRL-Japan with A. Kheddar in 2008. He was Invited Resea-

cher at the University of Washington with E. Todorov in 2014. He received the CNRS Bronze Medal in 2015 (one medal is awarded in France in automatic/robotic/signal-processing every year). His main research interests include the motion generation, planning and control of complex robots, with a special regard in humanoid robotics. His expertise covers sensor-based (vision and force) control, numerical mathematics for control, bipedal locomotion and locomotion planning. He published more than 70 papers in international journals and conferences and supervised 10 Ph.D. theses. Dr. Mansard is currently an Associate Editor of the IEEE TRANSACTIONS ON ROBOTICS.



**Sethu Vijayakumar** received the Ph.D. degree in computer science and engineering from the Tokyo Institute of Technology, Tokyo, Japan, in 1998. He is Professor of Robotics and Founding Director of the Edinburgh Centre for Robotics, where he holds the Royal Academy of Engineering Microsoft Research Chair in Learning Robotics within the School of Informatics at the University of Edinburgh, U.K. He also has additional appointments as an Adjunct Faculty with the University of

Southern California, Los Angeles, CA, USA and a Visiting Research Scientist with the RIKEN Brain Science Institute, Tokyo. His research interests include statistical machine learning, whole body motion planning and optimal control in robotics, optimization in autonomous systems as well as optimality in human motor control and prosthetics and exoskeletons. Professor Vijayakumar is a Fellow of the Royal Society of Edinburgh. In his recent role as the Programme Director for Artificial Intelligence and Robotics at The Alan Turing Institute, Sethu helps shape and drive the UK national agenda in Robotics and Autonomous Systems.