Fast Approximate Dynamic Programming for Input-Affine Dynamics

M. A. S. Kolarijani and P. Mohajerin Esfahani

Abstract. We propose two novel numerical schemes for approximate implementation of the Dynamic Programming (DP) operation concerned with finite-horizon optimal control of discrete-time, stochastic systems with input-affine dynamics. The proposed algorithms involve discretization of the state and input spaces, and are based on an alternative path that solves the dual problem corresponding to the DP operation. We provide error bounds for the proposed algorithms, along with a detailed analyses of their computational complexity. In particular, for a specific class of problems with separable data in the state and input variables, the proposed approach can reduce the typical time complexity of the DP operation from $\mathcal{O}(XU)$ to $\mathcal{O}(X+U)$ where X and U denote the size of the discrete state and input spaces, respectively. In a broader perspective, the key contribution here can be viewed as an algorithmic transformation of the minimization in DP operation to addition via discrete conjugation. This bridge enables us to utilize any complexity reduction on the discrete conjugation front within the proposed algorithms. In particular, motivated by the recent development of quantum algorithms for computing the discrete conjugate transform, we discuss the possibility of a quantum mechanical implementation of the proposed algorithms.

Keywords: approximate dynamic programming, conjugate duality, input-affine dynamics, computational complexity

1. Introduction

Motivation. Since its introduction, Dynamic Programming (DP) has been used for solving sequential decision problems in various applications emerging in several communities including Operation Research, Optimal Control, and Machine Learning. The basic idea of DP is to solve the Bellman equation

$$J_t(x_t) = \min_{u_t} \big\{ C(x_t, u_t) + J_{t+1}(x_{t+1}) \big\}, \tag{1}$$

backward in time t for the costs-to-go J_t, where $C(x_t, u_t)$ is the cost of taking the control action u_t at the state x_t (value iteration). Arguably, the most important drawback of DP is in its high computational cost in solving problems with a large scale *finite* state-action space, which are usually described as Markov Decision Processes (MDPs). Indeed, in [4], the authors show that for a finite-horizon MDP, the problem of determining whether a control action u_0 is an optimal action at a given initial state x_0 using value iteration is EXPTIME-complete. For problems with a *continuous* state space, which is commonly the case in engineering applications, solving the Bellman equation

Date: November 25, 2020.

The authors are with Delft Center for Systems and Control, Delft University of Technology, Delft, The Netherlands. Email: {M.A.SharifiKolarijani, P.MohajerinEsfahani}@tudelft.nl.

The authors would like to thank G. F. Max for several fruitful discussions.

amounts to solving infinite number of optimization problems. This usually renders the exact implementation of the DP operation impossible, except for a few cases with an available closed form solution, e.g., Linear Quadratic Regulator [7, Sec. 4.1]. To address this issue, various schemes have been introduced, commonly known as Approximate Dynamic Programming; see, e.g., [9, 28]. A straightforward approximation is to deploy a brute force search over the discretizations of the state and input spaces. Such an approximation scheme leads to a time complexity of at least $\mathcal{O}(XU)$, where X and U are the cardinality of the discrete state and input spaces, respectively.

For some DP problems, it is possible to reduce this complexity by using duality, i.e., approaching the minimization problem in (1) in the conjugate domain. For instance, for the deterministic linear dynamics $x_{t+1} = Ax_t + Bu_t$ with the separable cost $C(x_t, u_t) = C_{\mathrm{s}}(x_t) + C_{\mathrm{i}}(u_t)$, we have

$$J_t(x_t) \geq C_{\mathrm{s}}(x_t) + \left[C_{\mathrm{i}}^*(-B^\top \cdot) + J_{t+1}^*\right]^*(Ax_t), \tag{2}$$

where the operator $[\cdot]^*$ denotes the Legendre-Fenchel Transform, also known as (convex) conjugate transform. Under some technical assumptions (e.g., convexity of the functions C_{i} and J_{t+1}), we have equality in (2); see, e.g., [8, Prop. 5.3.1]. Notice how the minimization operator in (1) transforms to a simple addition in (2). This observation signals the possibility of a significant reduction in the complexity of solving the Bellman equation, at least for particular classes of DP problems.

Related works. Approaching the DP problem through the lens of the conjugate duality goes back to Bellman [5]. Further applications of this idea for reducing the computational complexity were later explored in [16] and [20]. Fundamentally, these approaches exploit the operational duality of infimal convolution and addition with respect to (w.r.t.) the conjugate transform [29]: For two functions $f_1, f_2 : \mathbb{R}^n \to [-\infty, \infty]$, we have $(f_1 \square f_2)^* = f_1^* + f_2^*$, where $f_1 \square f_2(x) := \inf\{f_1(x_1) + f_2(x_2) : x_1 + x_2 = x\}$ is the infimal convolution of f_1 and f_2. This is analogous to the well-known operational duality of convolution and multiplication w.r.t. the Fourier Transform. Actually, Legendre-Fenchel Transform plays a similar role as Fourier Transform when the underlying algebra is the *max-plus* algebra, as opposed to the conventional plus-times algebra. Much like the extensive application of the latter operational duality upon introduction of the Fast Fourier Transform, "fast" numerical algorithms for conjugate transform facilitate efficient applications of the former one. Interestingly, the first fast algorithm for computing (discrete) conjugate functions, known as Fast Legendre Transform, designed inspired by Fast Fourier Transform, enjoys the same *log-linear* complexity in the number of data points; see [13, 22] and the references therein. Later, this complexity was reduced by introducing a *linear-time* algorithm known as Linear-time Legendre Transform (LLT) [23]. We refer the interested reader to [25] for an extensive review of these algorithms (and other similar algorithms) and their applications. In this regard, we also note that recently, in [33], the authors introduced a quantum algorithm for computing the (discrete) conjugate of convex functions, which achieves a *poly-logarithmic* time complexity in the number of data points.

One of the first and most widespread applications of these fast algorithms has been in solving Hamilton-Jacobi Equation [1, 13, 14]. Another interesting area of application is image processing, where the Legendre-Fenchel Transform is commonly known as "distance transform"[17, 24]. Recently, in [18], the authors used these algorithms to tackle the Optimal Transport problem with strictly convex costs, with applications in image processing and in numerical methods for solving partial differential equations. However, surprisingly, the application of these fast algorithms in solving discrete-time optimal control problems seems to remain largely unexplored. An exception is [12], where the authors deploy LLT to propose the "Fast Value Iteration" algorithm for computing the fixed-point of the Bellman operator arising from a specific class of infinite-horizon, discrete-time DP problems. In the current article, we describe what we believe to be the most general class of discrete-time optimal control problems for which the operational duality of addition and infimal convolution can be used to reduce the time complexity of the DP operation. Indeed, the setup in [12] corresponds to a subclass of problems considered in our study, which allows for a "perfect" transformation of the minimization in the DP operation in the primal domain to an addition in the dual (conjugate) domain; this connection will be discussed in detail in Remark 5.5. Let us also note that the algorithms developed in [17, 24] for distance transform can also potentially tackle the (dicretized) optimal control problems of interest in the current study. In particular, these algorithms require the stage cost to be reformulated as a distance function of the current and next states, i.e., $C(x_t, u_t) = D(x_t - x_{t+1})$, where D is a convex distance function. While the this property might arise naturally, it can generally be restrictive as it is in our case.

Another line of work, closely related to ours, involves algorithms that utilize max-plus algebra in solving deterministic optimal control problems; see, e.g., [2, 26]. These works exploit the compatibility of the Bellman operation with max-plus operations, and approximate the value function as a max-plus linear combination. Similarly, in [3, 6], the authors this idea to propose an approximate value iteration algorithm for deteministic MDPs with continuous state space. In this regard, we note that the proposed algorithms in the current study also involve representing cost functions as max-plus linear combinations, yielding piece-wise affine approximation of the value functions. The key property of the proposed algorithms is, however, to utilize the linear-time complexity of LLT in order to reduce the computational cost when a grid-like (factorized) set of slopes is chosen in the dual space; the formal assertion is provided in Proposition 4.4 along with a follow-up discussion.

Overview of main results. In this study, we consider the approximate implementation of the DP operation arising in the finite-horizon optimal control of discrete-time systems with continuous state-input space. The proposed approach involves discretization of the state-input space, and is based on an alternative path that solves the dual problem corresponding to the DP operation by utilizing the LLT algorithm for discrete conjugation. Our contributions can be summarized as follows:

(i) **From minimization in primal domain to addition in dual domain:** We show that the *linearity* of the dynamics in the input is the key factor that allows us to use the operational duality of addition and infimal convolution. We use this property to introduce the discrete Conjugate Dynamic Programming (d-CDP) algorithm for problems with deterministic *input-affine* dynamics (Algorithm 1); see Figure 1a for the sketch of the algorithm. The d-CDP algorithm essentially involves transforming the minimization in the DP operation to addition at the expense of two discrete conjugate transforms. This, in turn, leads to transferring the computational cost from the input domain $\mathbb{U}$ to the state dual domain $\mathbb{Y}$. In particular, the time complexity of computing the optimal costs-to-go at each step is $\mathcal{O}(XY)$ using the d-CDP algorithm, compared to the benchmark $\mathcal{O}(XU)$, where X, U and Y are the sizes of the discrete state, input, and state dual spaces, respectively (Theorem 4.3). The extension of the this algorithm for stochastic dynamics is also discussed in Section 4.3.1.

Primal domain

$$\begin{tikzcd}
J(x^+) \arrow[rr, "\min\limits_{u} \{C(x,u) + J(x^+)\}"] \arrow[dd, "{[\cdot]^*}"] & & \mathcal{T}[J](x) \arrow[d, dashed, leftrightarrow, "\text{Thm. } 4.5"'] \\
& \phi_x^*(f_\mathrm{s}(x)) \arrow[r, phantom, "="] & \widehat{\mathcal{T}}[J](x) \\
J^*(y) \arrow[r, "+C_x^*(-f_\mathrm{i}(x)^\top y)"] & \phi_x(y) \arrow[u, "{[\cdot]^*}"'] &
\end{tikzcd}$$

Dual domain

(a) Setting 1 (Alg. 1, complexity in Thm. 4.3):
– dynamics $x^+ = f_\mathrm{s}(x) + f_\mathrm{i}(x) \cdot u$,
– generic convex cost $C(x,u)$.

Primal domain

$$\begin{tikzcd}
J(x^+) \arrow[rr, "\min\limits_{u} \{C(x,u) + J(x^+)\}"] \arrow[dd, "{[\cdot]^*}"] & & \mathcal{T}[J](x) \arrow[d, dashed, leftrightarrow, "\text{Thm. } 5.3"'] \\
& \phi^*(f_\mathrm{s}(x)) \arrow[r, "+C_\mathrm{s}(x)"'] & \widehat{\mathcal{T}}[J](x) \\
J^*(y) \arrow[r, "+C_\mathrm{i}^*(-B^\top y)"] & \phi(y) \arrow[u, "{[\cdot]^*}"'] &
\end{tikzcd}$$

Dual domain

(b) Setting 2 (Alg. 2, complexity in Thm. 5.2):
– dynamics $x^+ = f_\mathrm{s}(x) + B \cdot u$,
– separable convex cost $C(x,u) = C_\mathrm{s}(x) + C_\mathrm{i}(u)$.

FIGURE 1. Sketch of proposed algorithms: the standard DP operation in the primal domain (upper red paths) and the Conjugate DP (CDP) operation through the dual domain (bottom blue paths).

(ii) **From quadratic to linear time complexity:** We identify a subclass of problems with *separable* data in the state and input variables for which the proposed d-CDP algorithm can be modified to achieve a better time complexity (Algorithm 2); see Figure 1b for the sketch of the algorithm. In particular, for this class, the time complexity of computing the optimal costs-to-go at each step is $\mathcal{O}(X+Y)$ using the modified d-CDP algorithm (Theorem 5.2). This result, in turn, implies that the complexity of finding the optimal control sequence for a given initial condition via DP can be reduced to $\mathcal{O}(X+U)$ from $\mathcal{O}(XU)$.

(iii) **Error bounds and construction of discrete dual domains:** We analyze the error of the proposed d-CDP algorithm and its modification (Theorems 4.5 and 5.3, respectively). The error analysis is based on two preliminary results on the error of discrete conjugation (Lemma 2.6) and approximate conjugation (Lemma 2.7 and Corollary 2.8). Morevoer, we use the results of our error analysis to provide concrete *guidelines* for the construction of the discrete dual spaces in the proposed algorithms.

(iv) **The d-CDP MATLAB package:** We provide a MATLAB package [21], accompanied by lists of developed routines and subroutines for the implementation of the proposed d-CDP algorithms (Appendix D). The theoretical results of this study are validated through several numerical examples using this package (Section 6 and Appendix C). All the presented numerical examples are reproducible and included in the package.

(v) **Towards quantum dynamic programming:** This last item is more of a potential significance of the conjugate dynamic programming framework proposed in this study. The proposed d-CDP algorithms are developed in way that any reduction in the computational complexity of computing the discrete conjugate transform translates to a reduced computational cost of these algorithms. In particular, motivated by the recent quantum speedup for discrete conjugation [33], we envision that the proposed d-CDP Algorithm 2 paves the way for developing a quantum DP algorithm that can potentially address the infamous "curse of dimensinality" in DP. We will discuss this possibility in further details in Section 5.3.1.

To the best of our knowledge, this work constitutes the first attempt on providing explicit, quantitative analysis of application of fast algorithms for conjugate transform in solving finite-horizon discrete-time planning/control problems with continuous state–action space via DP.

Paper organization. After presenting some preliminaries in Section 2, we provide the problem statement and its standard solution via the d-DP algorithm (in the primal domain) in Section 3. Sections 4 and 5 contain our main results on the the proposed alternative approach for solving the DP problem in the conjugate domain. Section 4 presents the d-CDP algorithm for a general class of problems with input-affine dynamics. There, we also discuss the possible extensions of the proposed d-CDP algorithm. In Section 5, we focus on problems with separable data in order to modify the d-CDP algorithm for a better time complexity. In Section 6, we compare the performance of the proposed algorithms with the benchmark d-DP algorithm via multiple numerical examples. Section 7 concludes the paper by providing some final remarks on the implementation of the proposed algorithms, their limitations, and possible extensions.

2. Notations and Preliminaries

2.1. General notations

We use $\mathbb{R}$ to denote the real line and $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$, $\overline{\underline{\mathbb{R}}} = \mathbb{R} \cup \{\pm\infty\}$ to denote its extensions. The standard inner product in $\mathbb{R}^n$ and the corresponding induced 2-norm are denoted by $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$, respectively. We also use $\|\cdot\|$ to denote the operator norm (w.r.t. the 2-norm) of a matrix; i.e., for $A \in \mathbb{R}^{m\times n}$, we denote $\|A\| = \sup\{\|Ax\| : \|x\| = 1\}$. We use the common convention in optimization whereby the optimal value of an infeasible minimization (resp. maximization) problem is set to $+\infty$ (resp. $-\infty$). Arbitrary sets (finite/infinite, countable/uncountable) are denoted as $\mathbb{X}, \mathbb{Y}, \ldots$. For *finite* (discrete) sets, we use the subscript d as in $\mathbb{X}_{\mathrm{d}}, \mathbb{Y}_{\mathrm{d}}, \ldots$ to differentiate them form infinite sets. Moreover, we use the subscript g to differentiate *grid-like* (factorized) finite sets. Precisely, a grid $\mathbb{X}_{\mathrm{g}} \subset \mathbb{R}^n$ is the Cartesian product $\mathbb{X}_{\mathrm{g}} = \Pi_{i=1}^n \mathbb{X}_{\mathrm{g}_i} = \mathbb{X}_{\mathrm{g}_1} \times \ldots \times \mathbb{X}_{\mathrm{g}_n}$, where $\mathbb{X}_{\mathrm{g}_i}$ is a finite subset of $\mathbb{R}$. The cardinality of a finite set $\mathbb{X}_{\mathrm{d}}$ (or $\mathbb{X}_{\mathrm{g}}$) is denoted by X. Let $\mathbb{X}, \mathbb{Y}$ be two arbitrary sets in $\mathbb{R}^n$. The convex hull of $\mathbb{X}$ is denoted by $\mathrm{co}(\mathbb{X})$. The diameter of $\mathbb{X}$ is defined as $\Delta_{\mathbb{X}} := \sup_{x,x'\in\mathbb{X}} \|x - x'\|$. We use $\mathrm{d}(\mathbb{X}, \mathbb{Y}) := \inf_{x\in\mathbb{X},y\in\mathbb{Y}} \|x - y\|$ to denote the distance between $\mathbb{X}$ and $\mathbb{Y}$. The one-sided Hausdorff distance *from* $\mathbb{X}$ *to* $\mathbb{Y}$ is defined as $\mathrm{d}_{\mathrm{H}}(\mathbb{X}, \mathbb{Y}) := \sup_{x\in\mathbb{X}} \inf_{y\in\mathbb{Y}} \|x - y\|$. We use $\delta_{\mathbb{X}_{\mathrm{g}}} = 2 \cdot \mathrm{d}_{\mathrm{H}}\left(\mathrm{co}(\mathbb{X}_{\mathrm{g}}), \mathbb{X}_{\mathrm{g}}\right)$ to denote the granularity of the grid $\mathbb{X}_{\mathrm{g}}$, that is, the maximum distance between any two vertices of the hyper-rectangular cells of the grid $\mathbb{X}_{\mathrm{g}}$. For an extended real-valued function $h : \mathbb{R}^n \to \overline{\mathbb{R}}$, the effective domain of h is defined by $\mathrm{dom}(h) := \{x \in \mathbb{R}^n : h(x) < +\infty\}$. The Lipschtiz constant of h over a set $\mathbb{Y} \subset \mathbb{R}^n$ is denoted by

$$\mathrm{L}(h; \mathbb{Y}) := \sup_{x_1,x_2\in\mathbb{Y}} \frac{|h(x_1) - h(x_2)|}{\|x_1 - x_2\|}.$$

We also denote $\mathrm{L}(h) := \mathrm{L}\left(h; \mathrm{dom}(h)\right)$. The subdifferential of h at a point $x \in \mathbb{R}^n$ is defined as

$$\partial h(x) := \left\{y \in \mathbb{R}^n : h(\tilde{x}) \geq h(x) + \langle y, \tilde{x} - x\rangle, \forall \tilde{x} \in \mathbb{X}\right\}.$$

For reader's convenience, we also provide the list of the most important objects used throughout this article in the following table.

Notation	& Description	Definition
h^{d}	Discretization of the function h	–
$\overline{h^{\mathrm{d}}}$	Extension of the discrete function h^{d}	–
h^*	Conjugate of h	(3)
$h^{\mathrm{d}*}$	Discrete conjugate of h (conjugate of h^{d})	(4)
h^{**}	Biconjugate of h	(5)
$h^{\mathrm{d}*\mathrm{d}*}$	Discrete biconjugate of h	(6)
LERP	Multilinear interpolation & extrapolation	–
LLT	Linear-time Legendre Transform	–
$\mathcal{T}$	Dynamic Programming (DP) operator	(14) & (26)
$\mathcal{T}_{\mathrm{d}}$	Discrete DP (d-DP) operator [for Setting 1]	(16)
$\widehat{\mathcal{T}}$	Conjugate DP (CDP) operator [for Setting 1]	(21)
$\widehat{\mathcal{T}}_{\mathrm{d}}$	Discrete CDP (d-CDP) operator	(22) & (27)
$\widehat{\mathcal{T}}_{\mathrm{d}}^{\mathrm{m}}$	Modified d-CDP operator [for Setting 2]	(28)

2.2. Extension of discrete functions

Consider an extended real-valued function $h : \mathbb{R}^n \to \overline{\mathbb{R}}$, and its discretization $h : \mathbb{X}_{\mathrm{d}} \to \overline{\mathbb{R}}$, where $\mathbb{X}_{\mathrm{d}}$ is a finite subset of $\mathbb{R}^n$. Whether a function is discrete is usually clarified by providing its domain explicitly. Nevertheless, occasionally, we use the superscript d, as in h^{d}, to denote the discretization of h. We particularly use this notation in combination with a second operation to emphasize that the second operation is applied on the discretized version of the operand. In particular, we use $\overline{h^{\mathrm{d}}} : \mathbb{R}^n \to \overline{\mathbb{R}}$ to denote the *extension* of the discrete function $h : \mathbb{X}_{\mathrm{d}} \to \overline{\mathbb{R}}$.

Remark 2.1 (Extension operation)**.** *Throughout this article, we almost exclusively assume that the extension operation* $\overline{[\cdot]}$ *preserves the value of the function at the discrete points. That is, for a discrete function* $h : \mathbb{X}_{\mathrm{d}} \to \overline{\mathbb{R}}$*, its extension* $\overline{h^{\mathrm{d}}} : \mathbb{R}^n \to \overline{\mathbb{R}}$ *is such that* $\overline{h^{\mathrm{d}}}(x) = h(x)$ *for all* $x \in \mathbb{X}_{\mathrm{d}}$*. Moreover, we use* E *to denote the complexity of the extension operation. That is, for each* $x \in \mathbb{R}^n$*, the time complexity of the single evaluation* $\overline{h^{\mathrm{d}}}(x)$ *is assumed to be of* $\mathcal{O}(E)$*, with* E *(possibly) being a function of* X *and* n*.*

For example, *multilinear interpolation & extrapolation* (LERP) of a discrete function with a *grid-like* domain, preserves the value of the function at the discrete points. In this study, we particularly use LERP extension of discrete functions for approximate conjugation, which we will discuss in Section 2.4. In order to facilitate our complexity analysis, we discusses the computational complexity of LERP in the following remark.

Remark 2.2 (Complexity of LERP)**.** *Given a discrete function* $h : \mathbb{X}_{\mathrm{g}} \to \mathbb{R}$ *with a* grid-like *domain* $\mathbb{X}_{\mathrm{g}} \subset \mathbb{R}^n$*, the time complexity of a single evaluation of the* LERP *extension* $\overline{h^{\mathrm{d}}}$ *at a point* $x \in \mathbb{R}^n$ *is of* $\mathcal{O}(2^n + \log X)$ *if* $\mathbb{X}_{\mathrm{g}}$ *is non-uniform, and of* $\mathcal{O}(2^n)$ *if* $\mathbb{X}_{\mathrm{g}}$ *is uniform – according to Remark 2.1, that is to say,* $E = 2^n + \log X$ *(resp.* $E = 2^n$*) if* $\mathbb{X}_{\mathrm{g}}$ *is non-uniform (resp. uniform), when the extension operation uses LERP. To see this, note that, in the case that* $\mathbb{X}_{\mathrm{g}}$ *is non-uniform, LERP requires* $\mathcal{O}(\log X)$ *operations to find the position of* x *w.r.t. the grid points, using binary search. If* $\mathbb{X}_{\mathrm{g}}$ *is a uniform grid, this can be done in* $\mathcal{O}(n)$ *time. Upon finding the position of* x *w.r.t. the grid* *points, LERP then involves a series of one-dimensional linear interpolations or extrapolations along each dimension, which takes $\mathcal{O}(2^n)$ operations.*

For a convex function $h : \mathbb{R}^n \to \overline{\mathbb{R}}$, we have $\partial h(x) \neq \emptyset$ for all x in the relative interior of $\mathbb{X}$ [8, Prop. 5.4.1]. This characterization of convexity can be extended to discrete functions. A discrete function $h : \mathbb{X}_\mathrm{d} \to \mathbb{R}$ is called *convex-extensible* if $\partial h(x) \neq \emptyset$ for all $x \in \mathrm{dom}(h) = \mathbb{X}_\mathrm{d}$. Equivalently, h is convex-extensible, if it can be extended to a convex function $\overline{h^\mathrm{d}} : \mathbb{R}^n \to \overline{\mathbb{R}}$ such that $\overline{h^\mathrm{d}}(x) = h(x)$ for all $x \in \mathbb{X}_\mathrm{d}$; we refer the reader to, e.g., [27] for different extensions of the notion of convexity to discrete functions. We finish with a remark on the reported time complexities in this article.

Remark 2.3 (On the reported complexities)**.** *(i) In this study, we are mainly concerned with the dependence of the computational complexities on* the size of the finite sets *involved (discretization of the primal and dual domains). In particular, we ignore the possible dependence of the computational complexities on the dimension of the variables, unless they appear in the power of the size of those discrete sets. (ii) We occasionally report the complexities using the notation $\widetilde{\mathcal{O}}$, which hides the logarithmic factors in the size of the discrete sets.*

As a consequence of Remark 2.3, hereafter, the time complexity of LERP, reported in Remark 2.2, will be taken to be of $\widetilde{\mathcal{O}}(1)$. Similarly, a single evaluation of an analytically available function is taken to be of $\mathcal{O}(1)$, regardless of the dimension of its input and output arguments.

2.3. **Legendre-Fenchel Transform**

Consider an extended-real-valued function $h : \mathbb{R}^n \to \overline{\mathbb{R}}$, with a nonempty effective domain $\mathrm{dom}(h) = \mathbb{X}$. The Legendre-Fenchel Transform (convex conjugate) of h is the function $h^* : \mathbb{R}^n \to \underline{\overline{\mathbb{R}}}$, defined by

$$h^*(y) = \sup_{x \in \mathbb{X}} \left\{ \langle y, x \rangle - h(x) \right\}. \tag{3}$$

Note that the conjugate function h^* is convex by construction. In this study, we particularly consider *discrete* conjugation, which involves computing the conjugate function using the discretized version $h^\mathrm{d} : \mathbb{X}_\mathrm{d} \to \mathbb{R}$ of the function h, where $\mathbb{X}_\mathrm{d}$ is a *finite* subset of $\mathbb{X}$. We use the notation $[\cdot]^{\mathrm{d}*}$, as opposed the standard notation $[\cdot]^*$, for discrete conjugation; that is, $h^{\mathrm{d}*} = [h^\mathrm{d}]^* : \mathbb{R}^n \to \mathbb{R}$ is defined by

$$h^{\mathrm{d}*}(y) = \max_{x \in \mathbb{X}_\mathrm{d}} \left\{ \langle y, x \rangle - h(x) \right\}. \tag{4}$$

The biconjugate of h is the function $h^{**} = [h^*]^* : \mathbb{R}^n \to \underline{\overline{\mathbb{R}}}$, explicitly given by

$$h^{**}(x) = \sup_{y \in \mathbb{R}^n} \left\{ \langle x, y \rangle - h^*(y) \right\} = \sup_{y \in \mathbb{R}^n} \inf_{z \in \mathbb{X}} \left\{ \langle x - z, y \rangle + h(z) \right\}. \tag{5}$$

Using the notion of discrete conjugation $[\cdot]^{\mathrm{d}*}$, we also define the (doubly) discrete biconjugate operation $[\cdot]^{\mathrm{d}*\mathrm{d}*} = [[\cdot]^{\mathrm{d}*}]^{\mathrm{d}*}$. Explicitly, the discrete biconjugate $h^{\mathrm{d}*\mathrm{d}*} : \mathbb{R}^n \to \mathbb{R}$ is defined by

$$h^{\mathrm{d}*\mathrm{d}*}(x) = \max_{y \in \mathbb{Y}_\mathrm{d}} \left\{ \langle x, y \rangle - h^{\mathrm{d}*}(y) \right\} = \max_{y \in \mathbb{Y}_\mathrm{d}} \min_{z \in \mathbb{X}_\mathrm{d}} \left\{ \langle x - z, y \rangle + h(z) \right\}, \tag{6}$$

where $\mathbb{X}_\mathrm{d}$ and $\mathbb{Y}_\mathrm{d}$ are finite subsets of $\mathbb{R}^n$ such that $\mathbb{X}_\mathrm{d} \subset \mathbb{X}$.

The Linear-time Legendre Transform (LLT) is an efficient algorithm for computing the *discrete* conjugate over a finite *grid-like* dual domain. Precisely, to compute the conjugate of the function h :

$\mathbb{X} \to \mathbb{R}$, LLT takes its discretization $h^{\mathrm{d}} : \mathbb{X}_{\mathrm{d}} \to \mathbb{R}$ as an input, and outputs $h^{\mathrm{d}*\mathrm{d}} : \mathbb{Y}_{\mathrm{g}} \to \mathbb{R}$, for the grid-like dual domain $\mathbb{Y}_{\mathrm{g}}$. That is, the LLT operation is equivalent to the operation $[\cdot]^{\mathrm{d}*\mathrm{d}}$. We refer the interested reader to [23] for a detailed description of the LLT algorithm. We will use the following result for analyzing the computational complexity of the algorithms developed in this study.

Remark 2.4 (Complexity of LLT)**.** *Consider a function $h : \mathbb{R}^n \to \overline{\mathbb{R}}$ and its discretization over a grid-like set $\mathbb{X}_{\mathrm{g}} \subset \mathbb{R}^n$ such that $\mathbb{X}_{\mathrm{g}} \cap \operatorname{dom}(h) \neq \emptyset$. LLT computes the discrete conjugate function $h^{\mathrm{d}*} : \mathbb{Y}_{\mathrm{g}} \to \mathbb{R}$ using the data points $h : \mathbb{X}_{\mathrm{g}} \to \overline{\mathbb{R}}$, with a time complexity of $\mathcal{O}\left(\Pi_{i=1}^{n}(X_i + Y_i)\right)$, where X_i (resp. Y_i) is the cardinality of the i-th dimension of the grid $\mathbb{X}_{\mathrm{g}}$ (resp. $\mathbb{Y}_{\mathrm{g}}$). In particular, if the grids $\mathbb{X}_{\mathrm{g}}$ and $\mathbb{Y}_{\mathrm{g}}$ have approximately the same cardinality in each dimension, then the time complexity of LLT is of $\mathcal{O}(X + Y)$ [23, Cor. 5].*

Hereafter, in order to simplify the exposition, we consider the following assumption.

Assumption 2.5 (Grid sizes in LLT)**.** *The primal and dual grids used for LLT operation have approximately the same cardinality in each dimension.*

2.4. Preliminary results on conjugate transform

In what follows, we provide two preliminary lemmas on the error of discrete conjugate transform (LLT) and its approximate version. Although tailored for the error analysis of the proposed algorithms, we present these results in a generic format in order to facilitate their possible application/extension beyond the optimal control problems considered in this study. To this end, we recall some of the notations introduced so far. For a function $h : \mathbb{R}^n \to \overline{\mathbb{R}}$ with effective domain $\mathbb{X} = \operatorname{dom}(h)$, let $h^{\mathrm{d}} : \mathbb{X}_{\mathrm{d}} \to \mathbb{R}$ be the discretization of h, where $\mathbb{X}_{\mathrm{d}}$ is a finite subset of $\mathbb{X}$. Also, let $h^* : \mathbb{R}^n \to \overline{\underline{\mathbb{R}}}$ be the conjugate (3) of h. We use $h^{\mathrm{d}*} : \mathbb{R}^n \to \mathbb{R}$ to denote the discrete conjugate (4) of h, using the primal discrete domain $\mathbb{X}_{\mathrm{d}}$.

Lemma 2.6 (Conjugate vs. discrete conjugate)**.** *Let $h : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a* proper, closed, convex *function with a* nonempty *domain $\mathbb{X} = \operatorname{dom}(h)$. For each $y \in \mathbb{R}^n$, it holds that*

$$0 \leq h^*(y) - h^{\mathrm{d}*}(y) \leq \min_{x \in \partial h^*(y)} \left\{ \left[\, \|y\| + \mathrm{L}\left(h; \{x\} \cup \mathbb{X}_{\mathrm{d}}\right) \right] \cdot \mathrm{d}(x, \mathbb{X}_{\mathrm{d}}) \right\} =: \widetilde{e}_1(y, h, \mathbb{X}_{\mathrm{d}}). \tag{7}$$

If, moreover, $\mathbb{X}$ is compact *and h is* Lipschitz-continuous*, then*

$$0 \leq h^*(y) - h^{\mathrm{d}*}(y) \leq \left[\, \|y\| + \mathrm{L}(h) \right] \cdot \mathrm{d}_{\mathrm{H}}(\mathbb{X}, \mathbb{X}_{\mathrm{d}}) =: \widetilde{e}_2(y, h, \mathbb{X}_{\mathrm{d}}), \quad \forall y \in \mathbb{R}^n. \tag{8}$$

Proof. See Appendix B.1.1. □

The preceding lemma indicates that discrete conjugation leads to an under-approximation of the conjugate function, with the error depending on the "quality" of discrete representation $\mathbb{X}_{\mathrm{d}}$ of the primal domain $\mathbb{X}$. In particular, the inequality (7) implies that for $y \in \mathbb{R}^n$, if $\mathbb{X}_{\mathrm{d}}$ contains $x \in \partial h^*(y)$ – which is equivalent to $y \in \partial h(x)$ by the assumptions, then $h^{\mathrm{d}*}(y) = h^*(y)$.

We next present another preliminary however vital result on *approximate conjugation*. Let $h^{*\mathrm{d}} : \mathbb{Y}_{\mathrm{g}} \to \mathbb{R}$ be the discretization of h^* over the grid-like dual domain $\mathbb{Y}_{\mathrm{g}} = \Pi_{i=1}^{n} \mathbb{Y}_{\mathrm{g}i} \subset \mathbb{R}^n$, where $\mathbb{Y}_{\mathrm{g}i}$ is a finite set of real numbers $y_i^1 < y_i^2 < \ldots < y_i^{N_i}$. Also, let $\overline{h^{*\mathrm{d}}} : \mathbb{R}^n \to \mathbb{R}$ be the extension of $h^{*\mathrm{d}}$

using *LERP*. The approximate conjugation is then simply the approximation of $h^*(y)$ via $\overline{h^{*\mathrm{d}}}(y)$ for $y \in \mathbb{R}^n$. This approximation introduces a one-sided error which is the focus of our next result. Let us fix some more notations. Assume $N_i \geq 5$ for all $i = 1, \dots, n$, and denote $\mathbb{Y}'_{\mathrm{g}} := \Pi_{i=1}^n \mathbb{Y}'_{\mathrm{g}i}$, where $\mathbb{Y}'_{\mathrm{g}i} = \mathbb{Y}_{\mathrm{g}i} \setminus \{y_i^1, y_i^2, y_i^{N_i-1}, y_i^{N_i}\}$; that is, $\mathbb{Y}'_{\mathrm{g}}$ *is the sub-grid of* $\mathbb{Y}_{\mathrm{g}}$ *derived by omitting the first two elements and the last two elements of* $\mathbb{Y}_{\mathrm{g}}$ *in each dimension.* For each dimension $i \in \{1, \dots, n\}$, let x_i denote the i-th component of the vector $x \in \mathbb{R}^n$, and define

$$\mathrm{L}_i^+(h) := \max\left\{\frac{h(x) - h(\tilde{x})}{x_i - \tilde{x}_i} : x, \tilde{x} \in \mathbb{X},\ x_i > \tilde{x}_i,\ x_j = \tilde{x}_j\ (j \neq i)\right\},$$
$$\mathrm{L}_i^-(h) := \min\left\{\frac{h(x) - h(\tilde{x})}{x_i - \tilde{x}_i} : x, \tilde{x} \in \mathbb{X},\ x_i > \tilde{x}_i,\ x_j = \tilde{x}_j\ (j \neq i)\right\}.$$

Notice that $\mathrm{L}_i^+(h)$ (resp. $\mathrm{L}_i^-(h)$) is the maximum (resp. minimum) slope of the function h along the dimension i. In particular, for the hyper-rectangle $\mathbb{Y}(h) := \Pi_{i=1}^n \left[\mathrm{L}_i^-(h), \mathrm{L}_i^+(h)\right]$, we have $\mathbb{Y}(h) \cap \partial h(x) \neq \emptyset$ for all x in the (relative) interior of $\mathbb{X}$.

Lemma 2.7 (Approximate conjugation using LERP)**.** *Consider a function h with a* compact *domain $\mathbb{X} = \mathrm{dom}(h)$. For all $y \in \mathrm{co}(\mathbb{Y}_{\mathrm{g}})$, it holds that*

$$0 \leq \overline{h^{*\mathrm{d}}}(y) - h^*(y) \leq \delta_{\mathbb{Y}_{\mathrm{g}}} \cdot \Delta_{\mathbb{X}}. \tag{9}$$

If, moreover, the dual grid $\mathbb{Y}_{\mathrm{g}}$ is such that $\mathrm{co}(\mathbb{Y}'_{\mathrm{g}}) \supseteq \mathbb{Y}(h)$, then (9) *holds for all $y \in \mathbb{R}^n$.*

Proof. See Appendix B.1.2. □

As expected, the error due to the discretization $\mathbb{Y}_{\mathrm{g}}$ of the dual domain $\mathbb{Y}$ depends on the granularity $\delta_{\mathbb{Y}_{\mathrm{g}}}$ of the discrete dual domain. We also note that the condition $\mathrm{co}(\mathbb{Y}'_{\mathrm{g}}) \supseteq \mathbb{Y}(h)$ in the second part of the preceding lemma (which implies that h is Lipschitz-continuous), essentially requires the dual grid $\mathbb{Y}_{\mathrm{g}}$ to *more than cover* the range of slopes of the function h.

The algorithms developed in this study use LLT to compute discrete conjugate functions. However, as we will see, we often require the value of the conjugate function at points other than the dual grid points used in LLT. To solve this issue, we use the same approximation described above, but now for discrete conjugation. In this regard, we note that the result of Lemme 2.7 also holds for discrete conjugation. To be precise, let $h^{\mathrm{d}*\mathrm{d}} : \mathbb{Y}_{\mathrm{g}} \to \mathbb{R}$ be the discretization of $h^{\mathrm{d}*}$ over the grid-like dual domain $\mathbb{Y}_{\mathrm{g}} \subset \mathbb{R}^n$, and $\overline{h^{\mathrm{d}*\mathrm{d}}} : \mathbb{R}^n \to \mathbb{R}$ be the extension of $h^{\mathrm{d}*\mathrm{d}}$ using LERP. Define the set $\mathbb{Y}(h^{\mathrm{d}}) := \Pi_{i=1}^n \left[\mathrm{L}_i^-(h^{\mathrm{d}}), \mathrm{L}_i^+(h^{\mathrm{d}})\right]$, where

$$\mathrm{L}_i^+(h^{\mathrm{d}}) := \max\left\{\frac{h(x) - h(\tilde{x})}{x_i - \tilde{x}_i} : x, \tilde{x} \in \mathbb{X}_{\mathrm{d}};\ x_i > \tilde{x}_i\right\},$$
$$\mathrm{L}_i^-(h^{\mathrm{d}}) := \min\left\{\frac{h(x) - h(\tilde{x})}{x_i - \tilde{x}_i} : x, \tilde{x} \in \mathbb{X}_{\mathrm{d}};\ x_i > \tilde{x}_i\right\}.$$

Corollary 2.8 (Approximate discrete conjugation using LERP)**.** *Consider a discrete function $h^{\mathrm{d}} : \mathbb{X}_{\mathrm{d}} \to \mathbb{R}$. For all $y \in \mathrm{co}(\mathbb{Y}_{\mathrm{g}})$, it holds that*

$$0 \leq \overline{h^{\mathrm{d}*\mathrm{d}}}(y) - h^{\mathrm{d}*}(y) \leq \delta_{\mathbb{Y}_{\mathrm{g}}} \cdot \Delta_{\mathbb{X}_{\mathrm{d}}}. \tag{10}$$

If, moreover, the dual grid $\mathbb{Y}_{\mathrm{g}}$ is such that $\mathrm{co}(\mathbb{Y}'_{\mathrm{g}}) \supseteq \mathbb{Y}(h)$, then (10) *holds for all $y \in \mathbb{R}^n$.*

Proof. See Appendix B.1.3. □

3. Problem Statement and Standard Solution

In this study, we consider the optimal control of discrete-time systems

$$x_{t+1} = f(x_t, u_t), \quad t = 0, \dots, T-1, \tag{11}$$

where $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ describes the dynamics, and T is the finite horizon. We also consider state and input constraints of the form

$$\begin{cases} x_t \in \mathbb{X} \subset \mathbb{R}^n & \text{for} \quad t \in \{0, \dots, T\}, \\ u_t \in \mathbb{U} \subset \mathbb{R}^m & \text{for} \quad t \in \{0, \dots, T-1\}. \end{cases} \tag{12}$$

Let $C : \mathbb{X} \times \mathbb{U} \to \overline{\mathbb{R}}$ and $C_T : \mathbb{X} \to \mathbb{R}$ be the stage and terminal cost functions, respectively. In particular, notice that we let the stage cost C take ∞ for $(x, u) \in \mathbb{X} \times \mathbb{U}$ so that it can embed the *state-dependent input constraints*. For an initial state $x_0 \in \mathbb{X}$, the cost incurred by the state trajectory $\mathbf{x} = (x_0, \dots, x_T)$ in response to the input sequence $\mathbf{u} = (u_0, \dots, u_{T-1})$ is given by

$$J(x_0, \mathbf{u}) = \textstyle\sum_{t=0}^{T-1} C(x_t, u_t) + C_T(x_T).$$

The problem of interest is then to find the optimal control sequence $\mathbf{u}^\star(x_0)$, that is, a solution to the minimization problem

$$J^\star(x_0) = \min_{\mathbf{u}} \left\{ J(x_0, \mathbf{u}) : (11) \ \& \ (12) \right\}. \tag{13}$$

In order to solve this problem using DP, we have to solve the Bellman equation

$$J_t(x_t) = \min_u \left\{ C(x_t, u_T) + J_{t+1}(x_{t+1}) : (11) \ \& \ (12) \right\}, \quad x \in \mathbb{X},$$

backward in time $t = T-1, \dots, 0$, initialized by $J_T = C_T$. The iteration finally outputs $J_0 = J^\star$ [7, Prop. 1.3.1]. Let us take J_{t+1} to be the extended real-valued function with effective domain $\mathrm{dom}(J_{t+1}) = \mathbb{X}$ so that it embeds the constraint $x_{t+1} \in \mathbb{X}$. In order to simplify the exposition, let us also drop the time subscript t and focus on a single step of the recursion by defining the DP operator

$$\mathcal{T}[J](x) := \min_u \left\{ C(x, u) + J\big(f(x, u)\big) \right\}, \quad x \in \mathbb{X}, \tag{14}$$

so that $J_t = \mathcal{T}[J_{t+1}] = \mathcal{T}^{(T-t)}[J_T]$ for $t = T-1, \dots, 0$. *The main focus of this study is to use conjugate duality in order to develop "fast" schemes for the numerical implementation of the DP operator* (14) *for particular classes of optimal control problems.*

In what follows, we lay out a standard numerical scheme which involves solving the DP operator via a brute force search over the dicretized state-input space. In this regard, note that the DP operation (14) requires solving an infinite number of optimization problems for all $x \in \mathbb{X}$. Except for a few cases with an available closed form solution, the exact implementation of DP operation is impossible. In this study, we consider an approximate implementation of DP operation which addresss this issue by discretizing the state space. Precisely, we consider solving the optimization in (14) for a finite number of $x \in \mathbb{X}_{\mathrm{g}}$, where $\mathbb{X}_{\mathrm{g}} \subset \mathbb{X}$ is a *grid-like* discretization of the state space.

Problem 3.1 (Value Iteration)**.** *Given the discretization $\mathbb{X}_{\mathrm{g}} \subset \mathbb{X}$ of the state space, find the discrete optimal costs-to-go $J_t : \mathbb{X}_{\mathrm{g}} \to \mathbb{R}$ for $t = 0, 1, \dots, T-1$.*

Note that the DP operator $\mathcal{T}$ now takes the discrete function $J : \mathbb{X}_\mathrm{g} \to \mathbb{R}$ as an input, and outputs another discrete function $\mathcal{T}[J] : \mathbb{X}_\mathrm{g} \to \mathbb{R}$. Then, the computation of $\mathcal{T}[J]$ requires evaluating J at points $f(x,u)$ for $(x,u) \in \mathbb{X}_\mathrm{g} \times \mathbb{U}$, which do not necessarily belong to the discrete state space $\mathbb{X}_\mathrm{g}$. Hence, along with the discretization of the state space, we also need to consider some form of function approximation for the cost-to-go function, that is, an *extension* $\overline{J^\mathrm{d}} : \mathbb{X} \to \mathbb{R}$ of the function $J : \mathbb{X}_\mathrm{g} \to \mathbb{R}$. What remains to be addressed is the issue of solving the minimization

$$\min_{u \in \mathbb{U}} \left\{ C(x,u) + \overline{J^\mathrm{d}}\big(f(x,u)\big) \right\}, \tag{15}$$

for each $x \in \mathbb{X}_\mathrm{g}$, where the next step cost-to-go is approximated by the extension $\overline{J^\mathrm{d}}$. Here, again, we consider an approximation that involves enumeration over a proper discretization $\mathbb{U}_\mathrm{g} \subset \mathbb{U}$ of the inputs space. These approximations introduce some error which, under some regularity assumptions, depends on the "quality" of the discretization of the state and input spaces; see Proposition A.1. Incorporating these approximations, we can introduce the *discrete* DP operator (d-DP in short) as follows

$$\mathcal{T}_\mathrm{d}[J](x) := \min_{u \in \mathbb{U}_\mathrm{g}} \left\{ C(x,u) + \overline{J^\mathrm{d}}\big(f(x,u)\big) \right\}, \quad x \in \mathbb{X}_\mathrm{g}. \tag{16}$$

The d-DP operator/algorithm and its multistep implementation will be our *benchmark* for evaluating the performance of the alternative algorithms developed in this study. To this end, we discuss the time complexity of the d-DP operation in the following remark.

Remark 3.2 (Complexity of d-DP)**.** *Let the time complexity of a single evaluation of the extension* $\overline{J^\mathrm{d}} : \mathbb{R}^n \to \overline{\mathbb{R}}$ *with* $\mathrm{dom}(\overline{J^\mathrm{d}}) = \mathbb{X}$*, at a given point* $x \in \mathbb{R}^n$*, be of* $\mathcal{O}(E)$*. Then, the time complexity of the d-DP operation* (16) *is of* $\mathcal{O}\big(XUE\big)$*. Moreover, for solving the* T*-step Value Iteration Problem 3.1, the time complexity of d-DP algorithm increases linearly with the horizon* T*.*

Let us clarify that the scheme described above essentially involves approximating a continuous state/action MDP with a finite state/action MDP, and then applying the value iteration algorithm. In this regard, we note that $\mathcal{O}(XU)$ is the best existing time-complexity in the literature for finite MDPs; see, e.g., [3, 31]. Indeed, regardless of the problem data, the d-DP algorithm involves solving a minimization problem for each $x \in \mathbb{X}_\mathrm{g}$, via enumeration over $u \in \mathbb{U}_\mathrm{g}$. We note that, alternatively, one can consider another method for solving the minimization (15), i.e., a method other than enumeration over the discretization $\mathbb{U}_\mathrm{g}$ of $\mathbb{U}$ (as in d-DP). In that case, the computational complexity is $\mathcal{O}(X\mathfrak{U})$, where $\mathfrak{U}$ denotes the complexity of solving the minimization (15) for each $x \in \mathbb{X}_\mathrm{g}$. However, as we will see, for certain classes of optimal control problems, it is possible to exploit the structure of the underlying continuous setup to avoid the minimization over the input and achieve a lower time complexity.

We finish this section with some remarks on using the output of (the multistep implementation of) the d-DP algorithm, for finding a suboptimal control sequence $\widetilde{\mathbf{u}}^\star(x_0)$ for a given instance of the optimal control problem with initial state x_0. Upon solving the Value Iteration Problem 3.1, we have the optimal costs-to-go $J_t : \mathbb{X}_\mathrm{g} \to \mathbb{R},\ t = 0,1,\dots,T-1$, at our disposal. Then, in order to find $\widetilde{\mathbf{u}}^\star(x_0)$, we need to solve T minimization problems in the form of (15) forward in time, i.e.,

$$\widetilde{u}_t^\star \in \operatorname*{argmin}_{u_t \in \mathbb{U}} \left\{ C(x,u) + \overline{J_{t+1}^\mathrm{d}}\big(f(x_t,u_t)\big) \right\}, \quad t = 0,1,\dots,T-1. \tag{17}$$

Assuming these extra minimization problems are handled via enumeration over the same discretization $\mathbb{U}_{\mathrm{g}}$ of the input space $\mathbb{U}$, they lead to an additional computational burden of $\mathcal{O}(TUE)$ in solving a T-step optimal control problem using the d-DP algorithm, where E represents the complexity of the extension operation used in (17). On the other hand, the backward value iteration using the d-DP algorithm also provides us with optimal control laws $\mu_t : \mathbb{X}_{\mathrm{g}} \to \mathbb{U}_{\mathrm{g}},\ t = 0, 1, \dots, T-1$. Hence, alternatively, we can use these laws, accompanied by a proper extension operator, to produce a suboptimal control sequence $\widetilde{\mathbf{u}}^\star(x_0)$ for the initial state x_0 forward in time, i.e.,

$$\widetilde{u}_t^\star(x_t) = \overline{\mu_t^{\mathrm{d}}}(x_t), \quad t = 0, 1, \dots, T-1. \tag{18}$$

The second method (using optimal control laws) then has a time complexity of $\mathcal{O}(TE)$, where E represents the complexity of the extension operation used in (18). This complexity is particularly lower than $\mathcal{O}(TUE)$ of the first method (using optimal costs). However, generating control actions using optimal control laws has a higher memory complexity for systems with multiple inputs, and is also usually more sensitive to modelling errors due to its completely open-loop nature. Moreover, we note that the *total* time complexity of solving an instance of the optimal control problem, i.e., backward iteration for computing of optimal costs J_t and control laws μ_t, and forward iteration for computing of suboptimal control sequence $\widetilde{\mathbf{u}}^\star(x_0)$, is in both mehtods of $\mathcal{O}(TXUE)$. That is, computationally, the backward value iteration is the dominating factor. We will see this effect in our numerical study in Section 6, where we will consider both of the methods described above.

4. DP in Conjugate Domain: Alternative Solution

In this section, we introduce a general class of problems that allow us to employ conjugate duality for the DP problem and hence propose an alternative path for implementing the corresponding operator. In particular, we analyze the performance of the proposed algorithm by considering its error and time complexity. We also discuss the extensions of the proposed algorithm, which address two of the assumptions in our development.

4.1. The d-CDP algorithm

For now, we focus on deterministic systems described by (11) and (12), and postpone the discussion on the extension of the proposed scheme for stochastic systems to Section 4.3. Throughout this study, we assume that the problem data satisfy the following conditions.

Setting 1. *The dynamics, constraints and costs have the following properties:*

(i) **Dynamics.** *The dynamics is input-affine, that is, $f(x,u) = f_{\mathrm{s}}(x) + f_{\mathrm{i}}(x) \cdot u$, where $f_{\mathrm{s}} : \mathbb{R}^n \to \mathbb{R}^n$ is the "state" dynamics, and $f_{\mathrm{i}} : \mathbb{R}^n \to \mathbb{R}^{n \times m}$ is the "input" dynamics.*

(ii) **Constraints.** *The sets $\mathbb{X}$ and $\mathbb{U}$ are compact and convex. Moreover, for each $x \in \mathbb{X}$, the set of admissible inputs $\mathbb{U}(x) := \{u \in \mathbb{U} : C(x,u) < \infty,\ f(x,u) \in \mathbb{X}\}$ is* nonempty.

(iii) **Cost functions.** *The stage cost C is jointly convex in the state and input variables, with a compact effective domain. The terminal cost C_T is also convex.*

Note that the properties laid out in Setting 1 imply that the set of admissible inputs $\mathbb{U}(x)$ is a compact, convex set for each $x \in \mathbb{X}$. Hence, the optimal value in (13) is achieved. Hereafter, we also assume that the joint discretization of the state-input space is "proper" in the sense that *the feasibility condition of Setting 1-(ii) holds for the discrete state-input space, that is,* $\mathbb{U}_{\mathrm{g}}(x) := \mathbb{U}(x) \cap \mathbb{U}_{\mathrm{g}} \neq \emptyset$ *for all* $x \in \mathbb{X}_{\mathrm{g}}$.

Alternatively, we can approach the optimization problem in the DP operation (14) in the dual domain. To this end, let us fix $x \in \mathbb{X}_{\mathrm{g}}$, and consider the following reformulation of the problem (14),

$$\mathcal{T}[J](x) = \min_{u,z} \left\{ C(x,u) + J(z) : z = f(x,u) \right\}.$$

Notice how for input-affine dynamics, this formulation resembles the infimal convolution. In this regard, consider the corresponding dual problem

$$\widehat{\mathcal{T}}[J](x) := \max_{y} \min_{u,z} \{ C(x,u) + J(z) + \langle y, f(x,u) - z \rangle \}, \tag{19}$$

where $y \in \mathbb{R}^n$ is the dual variable. Indeed, for input-affine dynamics, we can derive an equivalent formulation for the dual problem (19), which forms the basis for the alternative algorithm for solving the DP operation.

Lemma 4.1 (CDP operator)**.** *Let*

$$C_x^*(v) := \max_u \big\{ \langle v, u \rangle - C(x,u) \big\}, \quad v \in \mathbb{R}^m, \tag{20}$$

denote the partial conjugate of the stage cost w.r.t. the input variable u*. Then, for the input-affine dynamics of Setting 1, the operator* $\widehat{\mathcal{T}}$ (19) *equivalently reads as*

$$\widehat{\mathcal{T}}[J](x) = \phi_x^*\big(f_{\mathrm{s}}(x)\big), \quad x \in \mathbb{X}_{\mathrm{g}}, \tag{21a}$$

$$\phi_x(y) := C_x^*(-f_{\mathrm{i}}(x)^\top y) + J^*(y), \quad y \in \mathbb{R}^n. \tag{21b}$$

Proof. See Appendix B.2.1. □

As we mentioned, the construction above suggests an alternative path for computing the output of the DP operator through the conjugate domain. We call this alternative approach Conjugate DP (CDP in short). Figure 1a characterizes this alternative path schematically. Numerical implementation of CDP operation requires computation of conjugate functions. In particular, as shown in Figure 1a, CDP operation involves three conjugate transforms. For now, we assume that the partial conjugate of the stage cost C_x^* (20) is analytically available. We will discuss the possibility of computing this object numerically in Section 4.3.

Assumption 4.2 (Conjugate of stage cost)**.** *For each* $x \in \mathbb{X}_{\mathrm{g}}$*, the conjugate function* C_x^* (20) *is analytically available. That is, the time complexity of evaluating* $C_x^*(v)$ *for each* $v \in \mathbb{R}^m$ *is of* $\mathcal{O}(1)$*; see also Remark 2.3.*

The two remaining conjugate operations of the CDP path in Figure 1a are handled numerically. To be precise, we first consider a *grid-like* discretization $\mathbb{Y}_{\mathrm{g}}$ of the dual domain, and employ LLT to compute $J^{\mathrm{d}*} : \mathbb{Y}_{\mathrm{g}} \to \mathbb{R}$ using the data points $J : \mathbb{X}_{\mathrm{g}} \to \mathbb{R}$. Now, let

$$\varphi_x(y) = C_x^*(-f_{\mathrm{i}}(x)^\top y) + J^{\mathrm{d}*}(y), \quad y \in \mathbb{Y}_{\mathrm{g}}.$$

Algorithm 1 Implementation of the d-CDP operator (22) for Setting 1.

```
Input: dynamics $f_s : \mathbb{R}^n \to \mathbb{R}^n$, $f_i : \mathbb{R}^n \to \mathbb{R}^{n \times m}$; cost-to-go (at $t+1$) $J : \mathbb{X}_g \to \mathbb{R}$; conjugate of stage
    cost $C_x^* : \mathbb{X}_g \times \mathbb{R}^m \to \mathbb{R}$.
Output: cost-to-go (at $t$) $\widehat{\mathcal{T}}_d[J](x) : \mathbb{X}_g \to \mathbb{R}$.
1: construct the grid $\mathbb{Y}_g$;
2: use LLT to compute $J^{d*} : \mathbb{Y}_g \to \mathbb{R}$ from $J^d : \mathbb{X}_g \to \mathbb{R}$;
3: for each $x \in \mathbb{X}_g$ do
4:     $\varphi_x(y) \leftarrow C_x^*(-f_i(x)^\top y) + J^{d*}(y)$ for $y \in \mathbb{Y}_g$;
5:     $\widehat{\mathcal{T}}_d[J](x) \leftarrow \max_{y \in \mathbb{Y}_g} \{\langle f_s(x), y\rangle - \varphi_x(y)\}$.
6: end for
```

Notice that φ_x is an approximation of ϕ_x (21b), where we used the discrete conjugate J^{d*} instead of the conjugate J^*. With φ_x at hand, we can also approximate the last conjugation $\phi_x^*(f_s(x))$ by

$$\varphi_x^{d*}\big(f_s(x)\big) = \max_{y \in \mathbb{Y}_g} \{\langle f_s(x), y\rangle - \varphi_x(y)\},$$

via enumeration over $y \in \mathbb{Y}_g$. Based on the construction described above, we can introduce the *discrete* CDP (d-CDP in short) operator as follows

$$\widehat{\mathcal{T}}_d[J](x) := \varphi_x^{d*}\big(f_s(x)\big), \quad x \in \mathbb{X}_g, \tag{22a}$$

$$\varphi_x(y) := C_x^*(-f_i(x)^\top y) + J^{d*}(y), \quad y \in \mathbb{Y}_g. \tag{22b}$$

Algorithm 1 provides the pseudo-code for the numerical implementation of the d-CDP operation (22). In the next section, we analyze the performance of Algorithm 1 by considering its error and complexity.

4.2. Analysis of d-CDP algorithm

4.2.1. *Complexity analysis.* We begin with the computational complexity of Algorithm 1. To this end, we assume that the construction of the grid $\mathbb{Y}_g$ in Algorithm 1:1 has a time complexity at most linear in the size of $\mathbb{X}_g$. We will discuss construction of $\mathbb{Y}_g$ and propose a method with such complexity in the next subsection.

Theorem 4.3 (Complexity of d-CDP – Algorithm 1)**.** *Let Assumptions 2.5 and 4.2 hold, and also assume that the construction of the grid $\mathbb{Y}_g$ has a time complexity at most linear in the size X of the discrete state space $\mathbb{X}_g$. Then, the implementation of the d-CDP operator* (22) *via Algorithm 1 requires $\mathcal{O}(XY)$ operations. Moreover, in the T-step application of Algorithm 1 for solving the Value Iteration Problem 3.1, the complexity increases linearly with the horizon T.*

Proof. See Appendix B.2.2. □

Recall that the time complexity of the d-DP operator (16) is of $\mathcal{O}(XUE)$; see Remark 3.2. Comparing this complexity to the one reported in Theorem 4.3, points to a basic characteristic of CDP w.r.t. DP: CDP avoids the minimization over control input in DP and casts it as simple addition in the dual domain at the expense of two conjugate transforms. Consequently, the time complexity is transferred form the primal input domain $\mathbb{U}_{\mathrm{g}}$ into the state dual domain $\mathbb{Y}_{\mathrm{g}}$. This observation implies that if $Y < UE$, then d-CDP is expected to computationally outperform d-DP.

We finish with some remarks on the application of the d-CDP algorithm for finding a suboptimal control sequence $\widetilde{\mathbf{u}}^{\star}(x_0)$ for a given initial state x_0. In this regard, note that, unlike the d-DP algorithm, the backward value iteration using the d-CDP algorithm *only* provides us with the optimal costs-to-go $J_t : \mathbb{X}_{\mathrm{g}} \to \mathbb{R}$, $t = 0, 1, \ldots, T-1$. Hence, in order to find $\widetilde{\mathbf{u}}^{\star}(x_0)$ using d-CDP, we also need to solve T minimization problems in the form of (17) forward in time, where at each time step, J_t is approximated by some extension $\overline{J_t^{\mathrm{d}}}$ of the output of the d-CDP algorithm. Assuming these extra minimization problems are handled via enumeration over the same discretization $\mathbb{U}_{\mathrm{g}}$ of the input space $\mathbb{U}$ (as in d-DP), an additional burden of $\mathcal{O}(TUE)$ must considered for finding a suboptimal control sequence $\widetilde{\mathbf{u}}^{\star}(x_0)$ using the d-CDP algorithm. As a result, the *total* complexity of computing $\widetilde{\mathbf{u}}^{\star}(x_0)$ using the d-CDP Algorithm 1 is of $\mathcal{O}\big(T(XY + UE)\big)$.

4.2.2. *Error analysis.* We now consider the error introduced by Algorithm 1 w.r.t. the DP operator (14). Let us begin with presenting an alternative representation for the d-CDP operator that sheds some light on the main sources of error in the d-CDP operation.

Proposition 4.4 (d-CDP reformulation)**.** *The d-CDP operator* (22) *equivalently reads as*

$$\widehat{\mathcal{T}}_{\mathrm{d}}[J](x) = \min_{u} \left\{ C(x,u) + J^{\mathrm{d}*\mathrm{d}*}\big(f(x,u)\big) \right\}, \quad x \in \mathbb{X}_{\mathrm{g}}, \tag{23}$$

where $J^{\mathrm{d}*\mathrm{d}*}$ is the discrete biconjugate (6) *of J, using the finite primal and dual sets $\mathbb{X}_{\mathrm{g}}$ and $\mathbb{Y}_{\mathrm{g}}$.*

Proof. See Appendix B.2.3. □

Comparing the representations (14) and (23), we note that the d-CDP operator $\widehat{\mathcal{T}}_{\mathrm{d}}$ differs from the DP operator $\mathcal{T}$ in that it uses $J^{\mathrm{d}*\mathrm{d}*}$ as an approximation of J. Also, note that

$$J^{\mathrm{d}*\mathrm{d}*}(x) = \max_{y \in \mathbb{Y}_{\mathrm{g}}} \left\{ \langle x, y \rangle - J^{\mathrm{d}*}(y) \right\},$$

which is a max-plus linear combination using the dictionary (set of basis functions) $\{\langle \cdot, y \rangle : y \in \mathbb{Y}_{\mathrm{g}}\}$ and coefficients $\{J^{\mathrm{d}*}(y) : y \in \mathbb{Y}_{\mathrm{g}}\}$. That is, d-CDP algorithm, similarly to the approximate value iteration algorithms in [3, 6], employs a max-plus approximation of J. This scheme particularly results in a piece-wise affine approximation, and the key property of the proposed algorithm is that by choosing a grid-like finite dual domain $\mathbb{Y}_{\mathrm{g}}$ (i.e., set of slopes for the basis functions used for approximations), we can incorporate the linear-time complexity of the LLT in our advantage.

More importantly, the representation (23) of the d-CDP operator, points to two main sources of error in the proposed approach, namely, dualization and discretization. Indeed, $\widehat{\mathcal{T}}_{\mathrm{d}}$ is a discretized version (in the variables z and y) of the dual problem (19). Regarding the dualization error, we note that the d-CDP operator is "blind" to non-convexity; that is, it essentially replaces the cost-to-go J by its convex envelop (the greatest convex function that supports J from below). The discretization error, on the other hand, depends on the choice of the finite primal and dual domains $\mathbb{X}_{\mathrm{g}}$ and $\mathbb{Y}_{\mathrm{g}}$. The following result captures this dependence.

Theorem 4.5 (Error of d-CDP – Algorithm 1)**.** *Consider Setting 1. Also, consider the DP operator $\mathcal{T}$* (14) *and the implementation of the d-CDP operator $\widehat{\mathcal{T}}_{\mathrm{d}}$* (22) *via Algorithm 1. Assume that*

$J : \mathbb{X} \to \mathbb{R}$ is a Lipschitz-continuous, convex function. Then, for each $x \in \mathbb{X}_{\mathrm{g}}$, it holds that

$$-e_2 \leq \mathcal{T}[J](x) - \widehat{\mathcal{T}}_{\mathrm{d}}[J](x) \leq e_1(x), \tag{24}$$

where

$$\begin{aligned} e_1(x) &= \left[\, \|f_{\mathrm{s}}(x)\| + \|f_{\mathrm{i}}(x)\| \cdot \Delta_{\mathbb{U}} + \Delta_{\mathbb{X}} \right] \cdot \mathrm{d}\left(\partial\mathcal{T}[J](x), \mathbb{Y}_{\mathrm{g}}\right), \\ e_2 &= \left[\Delta_{\mathbb{Y}_{\mathrm{g}}} + \mathrm{L}(J)\right] \cdot \mathrm{d}_{\mathrm{H}}(\mathbb{X}, \mathbb{X}_{\mathrm{g}}). \end{aligned} \tag{25}$$

Proof. See Appendix B.2.4. □

Some remarks are in order regarding the error bounds of the preceding theorem. First, notice how the two terms e_1 and e_2 capture the errors due to the discretization of the state dual space ($\mathbb{Y}$) and the primal state space ($\mathbb{X}$), respectively. In particular, by a proper choice of the discrete dual domain, one can eliminate the first error term e_1. To illustrate, let J be a one-dimensional, discrete convex function with the domain $\mathbb{X}_{\mathrm{g}} = \{x_i\}_{i=1}^{N} \subset \mathbb{R}$, where $x_i < x_{i+1}$. Also, choose $\mathbb{Y}_{\mathrm{g}} = \{y_i\}_{i=1}^{N-1} \subset \mathbb{R}$ with $y_i = \frac{J(x_{i+1})-J(x_i)}{x_{i+1}-x_i}$ as the discrete dual domain. Then, for all $x \in \mathrm{co}(\mathbb{X}_{\mathrm{g}}) = [x_1, x_N]$, we have $J^{\mathrm{d}*\mathrm{d}*}(x) = \overline{J^{\mathrm{d}}}(x)$, where $\overline{[\cdot]}$ is linear interpolation. This, in turn, implies that $e_1 = 0$; indeed, by Proposition 4.4, the only source of error under such construction is the state space discretization. However, we note that satisfying a similar condition in dimensions $n \geq 2$ can lead to dual grids of size $Y = \mathcal{O}(X^n)$, which renders the proposed algorithm impractical. We finish this section by providing some guidelines on the construction of the dual grid $\mathbb{Y}_{\mathrm{g}}$ using the result of our error analysis.

Construction of $\mathbb{Y}_{\mathrm{g}}$. Note that the first error term e_1 in (25) mainly depends on the dual grid $\mathbb{Y}_{\mathrm{g}}$. In particular, by choosing a "rich enough" grid $\mathbb{Y}_{\mathrm{g}}$ such that $\mathbb{Y}_{\mathrm{g}} \cap \partial\mathcal{T}[J](x) \neq \emptyset$ for all $x \in \mathbb{X}_{\mathrm{g}}$, we have $e_1 = 0$. However, since we require the dual domain $\mathbb{Y}_{\mathrm{g}}$ to be *grid-like*, satisfying such a condition can potentially lead to grids of size $\mathcal{O}(X^n)$ for the dual domain. This basically renders the proposed approach computationally impractical for systems of dimension $n \geq 2$; see Theorem 4.3. A more realistic objective is to choose $\mathbb{Y}_{\mathrm{g}}$ such that $\mathrm{co}(\mathbb{Y}_{\mathrm{g}}) \cap \partial\mathcal{T}[J](x) \neq \emptyset$ for all $x \in \mathbb{X}_{\mathrm{g}}$. With such a construction, we have $\mathrm{d}\left(\partial\mathcal{T}[J](x), \mathbb{Y}_{\mathrm{g}}\right) \leq \delta_{\mathbb{Y}_{\mathrm{g}}}$, and hence the first error term is expected to decrease by using finer grids for the dual domain. To this end, we need to approximate "the range of slopes" of the function $\mathcal{T}[J]$. Notice, however, that we do not have access to $\mathcal{T}[J]$ since it is the *output* of the d-CDP operation in Algorithm 1. What we have at our disposal are the stage cost C and the next step cost-to-go J, i.e., *inputs* to Algorithm 1. A coarse way to approximate the range of slopes of $\mathcal{T}[J]$ is to use the extrema of the functions C and J for all $x \in \mathbb{X}_{\mathrm{g}}$, and the diameter of $\mathbb{X}_{\mathrm{g}}$ in each dimension for the construction of $\mathbb{Y}_{\mathrm{g}}$. To be precise, we first compute

$$C^M = \max_{x\in\mathbb{X}_{\mathrm{g}}} \max_{u\in\mathbb{U}(x)} C(x,u), \;\; C^m = \min_{x\in\mathbb{X}_{\mathrm{g}}} \min_{u\in\mathbb{U}(x)} C(x,u), \;\; J^M = \max_{x\in\mathbb{X}_{\mathrm{g}}} J, \;\; J^m = \min_{x\in\mathbb{X}_{\mathrm{g}}} J,$$

and then choose $\mathbb{Y}_{\mathrm{g}} = \Pi_{i=1}^{n} \mathbb{Y}_{\mathrm{g}_i} \subset \mathbb{R}^n$ such that for each dimension $i \in \{1, \ldots, n\}$, we have

$$\pm\alpha \cdot \frac{C^M + J^M - C^m - J^m}{\Delta_{\mathbb{X}_{\mathrm{g}_i}}} \in \mathrm{co}(\mathbb{Y}_{\mathrm{g}_i}),$$

where $\alpha > 0$ is a scaling factor (mainly depending on the dimension n of the state space). The construction described above has a linear time complexity w.r.t. the size X of the state space.

4.3. **Extensions of d-CDP algorithm**

In this section, we consider the extensions of the proposed d-CDP algorithm and their implications on its complexity. In particular, extension to stochastic systems with additive disturbance and the numerical computation of the conjugate of the stage cost are discussed. The pseudo-code for the multistep implementation of the extended d-CDP algorithm is provided in Appendix C.1.

4.3.1. *Stochastic systems.* Consider the stochastic version of the dynamics (11) described by

$$x_{t+1} = f(x_t, u_t) + w_t,$$

where $w_t,\ t = 0, \dots, T-1$, are *additive* disturbances. The extension of the d-CDP algorithm for handling this type of stochasticity involves applying the d-CDP operator $\widehat{\mathcal{T}}_{\mathrm{d}}$ (22) to $\mathbb{E}_w \overline{J^{\mathrm{d}}}(x+w)$, where $\mathbb{E}_w$ is the expectation operator w.r.t. w, and $\overline{[\cdot]}$ is an extension operator. This scheme essentially comes to first passing the discrete cost-to-go J from the "expectation filter", and then feeding it to the d-CDP operator. Let us illustrate this procedure through the following case. Assume that the disturbances are i.i.d. and belong to a *finite* set $\mathbb{W}_{\mathrm{d}} \subset \mathbb{R}^n$, with a known probability mass function $p : \mathbb{W}_{\mathrm{d}} \to [0, 1]$. The set $\mathbb{W}_{\mathrm{d}}$ can indeed be considered as a discretization of a bounded set of disturbances. Then, the extension of the d-CDP algorithm involves applying the d-CDP operator $\widehat{\mathcal{T}}_{\mathrm{d}}$ (22) to $J^{\mathrm{w}}(x) : \mathbb{X}_{\mathrm{g}} \to \mathbb{R}$ defined by

$$J^{\mathrm{w}}(x) := \mathbb{E}_w \left[\overline{J^{\mathrm{d}}}(x+w)\right] = \sum_{w \in \mathbb{W}_{\mathrm{d}}} p(w) \cdot \overline{J^{\mathrm{d}}}(x+w),$$

where $\overline{[\cdot]}$ is an extension operator such as LERP.

Assuming that a single evaluation of the employed extension operator requires $\mathcal{O}(E)$ operations, and the time complexity of computing the expectation w.r.t. w is of $\mathcal{O}(W)$, the stochastic version of the d-CDP operation that utilizes the scheme described above requires $\mathcal{O}\big(X(WE+Y)\big)$ operations. On the other hand, the stochastic version of the d-DP operation, described by

$$\mathcal{T}^{\mathrm{s}}_{\mathrm{d}}[J](x) := \min_{u \in \mathbb{U}_{\mathrm{g}}} \left\{ C(x,u) + \mathbb{E}_w \left[\overline{J^{\mathrm{d}}}\big(f(x,u)+w\big)\right]\right\}, \quad x \in \mathbb{X}_{\mathrm{g}},$$

has a time complexity of $\mathcal{O}(XUWE)$. However, we note that a similar scheme can be used to introduce a *fast approximation* of the stochastic d-DP operation. To be precise, in the stochastic d-DP operator, we can swap the order of expectation and extension operations, and approximate $\mathcal{T}^{\mathrm{s}}_{\mathrm{d}}[J]$ via $\mathcal{T}_{\mathrm{d}}[J^{\mathrm{w}}]$. This approximation scheme reduces the time complexity of the stochastic d-DP operation to $\mathcal{O}(X(W+U)E)$, at the cost of introducing error.

4.3.2. *Numerical computation of C^*_x.* Assumption 4.2 on the availability of the conjugate function C^*_x (20) of the cost function is indeed quite restrictive. Alternatively, we can use the approximate discrete conjugation to compute C^*_x numerically as described by the following scheme:

- **Step 1.** For each $x \in \mathbb{X}_{\mathrm{g}}$:
  1.a. compute/evaluate $C(x,\cdot) : \mathbb{U}_{\mathrm{g}} \to \overline{\mathbb{R}}$, where $\mathbb{U}_{\mathrm{g}}$ is a discretization of the space $\mathbb{U}$;
  1.b. construct the dual grid $\mathbb{V}_{\mathrm{g}}(x)$ using the method described below; and,
  1.c. apply LLT to compute $C^{\mathrm{d}*}_x : \mathbb{V}_{\mathrm{g}}(x) \to \mathbb{R}$ using the data points $C(x,\cdot) : \mathbb{U}_{\mathrm{g}} \to \overline{\mathbb{R}}$.
- **Step 2.** For each $y \in \mathbb{Y}_{\mathrm{g}}$: use LERP to compute $\overline{C^{\mathrm{d}*\mathrm{d}}_x}(-f_{\mathrm{i}}(x)^\top y)$ from the data points $C^{\mathrm{d}*}_x : \mathbb{V}_{\mathrm{g}}(x) \to \mathbb{R}$, and use it in Algorithm 1:4 as an approximation of $C^*_x(-f_{\mathrm{i}}(x)^\top y)$.

The proposed numerical scheme increases the computational cost of the d-CDP algorithm. However, note that the first step of the scheme laid out above is carried out *once* in a multistep implementation of the d-CDP algorithm. In particular, assuming the grids $\mathbb{V}_{\rm g}(x),\ x \in \mathbb{X}_{\rm g}$, are all of the same size V, for the T-step implementation of d-CDP algorithm which uses the above scheme to compute C^*_x numerically,

- Step 1 introduces an extra computational cost of $\mathcal{O}(X(U+V))$, independent of T; and,
- Step 2 increases the computational cost of the backward iterations to $\widetilde{\mathcal{O}}(TXY)$.

This scheme also introduces some error that mainly depends on the grids $\mathbb{U}_{\rm g}$ and $\mathbb{V}_{\rm g}(x)$ used for the discretization of the input space and its dual domain, respectively. Indeed, we can use Lemmas 2.6 and Corollary 2.8 to bound this error. Below, we use those results, particularly Corollary 2.8, to provide some guidelines on the construction of the dual grids $\mathbb{V}_{\rm g}(x)$ for $x \in \mathbb{X}_{\rm g}$.

Construction of $\mathbb{V}_{\rm g}(x)$. Let us fix $x \in \mathbb{X}_{\rm g}$, and consider the construction of $\mathbb{V}_{\rm g}(x)$. Notice that the scheme describe above essentially involves *approximate discrete conjugation using LERP*. Then, by Corollary 2.8, we need to construct $\mathbb{V}_{\rm g}(x)$ such that $\operatorname{co}(\mathbb{V}_{\rm g}(x))$ *more than covers* the range of slopes of the function $C(x,\cdot)$. The following remark specifies this condition.

Remark 4.6 (Construction of dual grid for approximate discrete conjugation)**.** *Given the discrete function $C(x,\cdot) : \mathbb{U}_{\rm g} \to \overline{\mathbb{R}}$, construct the dual grid $\mathbb{V}_{\rm g}(x) = \Pi_{i=1}^{m} \mathbb{V}_{{\rm g}_i}(x) \subset \mathbb{R}^m$ such that for each dimension $i \in \{1,\dots,m\}$, the set $\mathbb{V}_{{\rm g}_i}(x) \subset \mathbb{R}$ contains* at least two distinct elements *that are less (resp. greater) than the minimum (resp. maximum) finite slope of $C(x,\cdot)$ along the dimension i.*

What remains to be addressed is how to compute the range of slopes of $C(x,\cdot) : \mathbb{U}_{\rm g} \to \overline{\mathbb{R}}$. In this regard, notice that the function $C(x,\cdot)$ is convex (see Setting 1), and hence its discretization is convex-extensible. Indeed, for a convex-extensible function with a grid-like domain, it is possible to compute its range of slopes with an acceptable cost as explained in the following remark.

Remark 4.7 (Range of slopes of a convex-extensible function)**.** *Given a convex-extensible function $h : \mathbb{U}_{\rm g} \to \mathbb{R}$, where $\mathbb{U}_{\rm g} \subset \mathbb{R}^m$, the minimum (resp. maximum) value of the slopes of h along each dimension is equal to the minimum first forward (resp. maximum last backward) difference of h along that dimension. Computing the range slopes of h using this method has a complexity of $\mathcal{O}(U)$.*

5. Reducing Complexity by Exploiting Structure

In this section, we focus on a specific subclass of the optimal control problems considered in this study. In particular, we exploit the problem structure in this subclass in order to reduce the computational cost of the d-CDP algorithm. In this regard, a closer look to Algorithm 1 reveals a computational bottleneck in its numerical implementation: computing $\varphi_x(y), y \in \mathbb{Y}_{\rm g}$, and its conjugate within the `for loop` over $x \in \mathbb{X}_{\rm g}$. This step is indeed the dominating factor in the time complexity of $\mathcal{O}(XY)$ of Algorithm 1; see Appendix B.2.2 for the proof of Theorem 4.3. Hence, if the structure of the problem allows for computing φ and its conjugate out of the `for loop` over $x \in \mathbb{X}_{\rm g}$, then a significant reduction in the time complexity is achievable. This is indeed possible for problems with separable data in the state and input variables. We again focus on deterministic systems in our development, and postpone the discussion on the extension of the proposed scheme for stochastic systems to Section 5.3.

5.1. **Modified d-CDP algorithm**

Consider the following subclass of the problems described in Setting 1.

Setting 2. *In addition to Setting 1, the dynamics and costs have the following properties:*

(i) **Dynamics.** *The input dynamics is state-independent, i.e., $f_{\mathrm{i}}(\cdot) = B \in \mathbb{R}^{n\times m}$.*

(ii) **Cost functions.** *The stage cost is separable in the state and input variables, that is, $C(x,u) = C_{\mathrm{s}}(x) + C_{\mathrm{i}}(u)$, where $C_{\mathrm{s}} : \mathbb{X} \to \mathbb{R}$ and $C_{\mathrm{i}} : \mathbb{U} \to \mathbb{R}$.*

Note that the separability of the stage cost C implies that the constraints are also separable, i.e, there are no state-dependent input constraints. Under the conditions above, the state-dependent part of the stage cost (C_{s}) can be taken out of the minimization in the DP operator (14) as follows

$$\mathcal{T}[J](x) = C_{\mathrm{s}}(x) + \min_{u} \left\{ C_{\mathrm{i}}(u) + J\big(f(x,u)\big) \right\}, \quad x \in \mathbb{X}_{\mathrm{g}}. \tag{26}$$

Following the same dualization and then discretization procedure described in Section 4.1, we can derive the corresponding d-CDP operator

$$\widehat{\mathcal{T}}_{\mathrm{d}}[J](x) = C_{\mathrm{s}}(x) + \varphi^{\mathrm{d}*}\big(f_{\mathrm{s}}(x)\big), \quad x \in \mathbb{X}_{\mathrm{g}}, \tag{27a}$$

$$\varphi(y) := C_{\mathrm{i}}^{*}(-B^{\top}y) + J^{\mathrm{d}*}(y), \quad y \in \mathbb{Y}_{\mathrm{g}}. \tag{27b}$$

Here, again, we assume that the conjugate of the (input-dependent) stage cost is analytically available (similar to Assumption 4.2, now in the context posed by Setting 2). We discuss a the possibility of handling this conjugation numerically in Section 5.3.

Assumption 5.1 (Conjugate of input-dependent stage cost)**.** *The conjugate function $C_{\mathrm{i}}^{*}(v) = \max_u\{\langle v,u\rangle - C_{\mathrm{i}}(u)\}$ is analytically available; i.e., the complexity of evaluating $C_{\mathrm{i}}^{*}(v)$ for each $v \in \mathbb{R}^m$ is of $\mathcal{O}(1)$; see also Remark 2.3.*

Notice how the function φ in (27b) is now independent of the state variable x. This, in particular, allows us to compute φ outside the `for loop` over $x \in \mathbb{X}_{\mathrm{g}}$; cf. the computation of φ_x in Algorithm 1. What remains to be addressed is the computation of the conjugate function $\varphi^{\mathrm{d}*}\big(f_{\mathrm{s}}(x)\big) = \max_{y\in\mathbb{Y}_{\mathrm{g}}}\{\langle f_{\mathrm{s}}(x), y\rangle - \varphi(y)\}$ for $x \in \mathbb{X}_{\mathrm{g}}$ in (27a). The straightforward maximization via enumeration over $y \in \mathbb{Y}_{\mathrm{g}}$ for each $x \in \mathbb{X}_{\mathrm{g}}$ (as in Algorithm 1) again leads to a time complexity of $\mathcal{O}(XY)$. In order to circumvent this issue, we deploy the following scheme of approximate discrete conjugatation:

- **Step 1.** Fix a grid $\mathbb{Z}_{\mathrm{g}}$ and use LLT to compute $\varphi^{\mathrm{d}*} : \mathbb{Z}_{\mathrm{g}} \to \mathbb{R}$, from the data points $\varphi : \mathbb{Y}_{\mathrm{g}} \to \mathbb{R}$.
- **Step 2.** For each $x \in \mathbb{X}_{\mathrm{g}}$: use LERP to compute $\overline{\varphi^{\mathrm{d}*\mathrm{d}}}\big(f_{\mathrm{s}}(x)\big)$ as an approximation of $\varphi^{\mathrm{d}*}\big(f_{\mathrm{s}}(x)\big)$, from the data points $\varphi^{\mathrm{d}*} : \mathbb{Z}_{\mathrm{g}} \to \mathbb{R}$.

With such an approximation, the d-CDP operator (27) *modifies* to

$$\widehat{\mathcal{T}}_{\mathrm{d}}^{\mathrm{m}}[J](x) := C_{\mathrm{s}}(x) + \overline{\varphi^{\mathrm{d}*\mathrm{d}}}\big(f_{\mathrm{s}}(x)\big), \quad x \in \mathbb{X}_{\mathrm{g}}, \tag{28a}$$

$$\varphi^{\mathrm{d}*}(z) := \max_{y\in\mathbb{Y}_{\mathrm{g}}}\{\langle z,y\rangle - \varphi(y)\}, \quad z \in \mathbb{Z}_{\mathrm{g}}, \tag{28b}$$

$$\varphi(y) := C_{\mathrm{i}}^{*}(-B^{\top}y) + J^{\mathrm{d}*}(y), \quad y \in \mathbb{Y}_{\mathrm{g}}. \tag{28c}$$

Algorithm 2 Implementation of the modified d-CDP operator (28) for Setting 2.

Input: dynamics $f_{\mathrm{s}} : \mathbb{R}^n \to \mathbb{R}^n$, $B \in \mathbb{R}^{n\times m}$; cost-to-go (at $t+1$) $J : \mathbb{X}_{\mathrm{g}} \to \mathbb{R}$; stage cost (state) $C_{\mathrm{s}}(x) : \mathbb{R}^n \to \mathbb{R}$; conjugate of stage cost (input) $C_{\mathrm{i}}^* : \mathbb{R}^m \to \mathbb{R}$; grid $\mathbb{Z}_{\mathrm{g}} \subset \mathbb{R}^n$.
Output: cost-to-go (at t) $\widehat{\mathcal{T}}_{\mathrm{d}}^{\mathrm{m}}[J](x) : \mathbb{X}_{\mathrm{g}} \to \mathbb{R}$.

```
1: construct the grid $\mathbb{Y}_{\mathrm{g}}$;
2: use LLT to compute $J^{\mathrm{d}*} : \mathbb{Y}_{\mathrm{g}} \to \mathbb{R}$ from $J : \mathbb{X}_{\mathrm{g}} \to \mathbb{R}$;
3: $\varphi(y) \leftarrow C_{\mathrm{i}}^*(-B^\top y) + J^{\mathrm{d}*}(y)$ for $y \in \mathbb{Y}_{\mathrm{g}}$;
4: use LLT to compute $\varphi^{\mathrm{d}*} : \mathbb{Z}_{\mathrm{g}} \to \mathbb{R}$ from $\varphi : \mathbb{Y}_{\mathrm{g}} \to \mathbb{R}$;
5: for each $x \in \mathbb{X}_{\mathrm{d}}$ do
6:     use LERP to compute $\overline{\varphi^{\mathrm{d}*\mathrm{d}}}\big(f_{\mathrm{s}}(x)\big)$ from $\varphi^{\mathrm{d}*} : \mathbb{Z}_{\mathrm{g}} \to \mathbb{R}$;
7:     $\widehat{\mathcal{T}}_{\mathrm{d}}^{\mathrm{m}}[J](x) \leftarrow C_{\mathrm{s}}(x) + \overline{\varphi^{\mathrm{d}*\mathrm{d}}}\big(f_{\mathrm{s}}(x)\big)$;
8: end for
```

Algorithm 2 provides the pseudo-code for the numerical scheme described above. In particular, notice that the gird $\mathbb{Z}_{\mathrm{g}}$ is treated as an input; we will discuss this issue in the next section.

5.2. **Analysis of modified d-CDP algorithm**

5.2.1. *Complexity analysis.* We again begin with the time complexity of the proposed algorithm.

Theorem 5.2 (Complexity of modified d-CDP – Algorithm 2)**.** *Let Assumptions 2.5 and 5.1 hold. Also, assume that the construction of the grid $\mathbb{Y}_{\mathrm{g}}$ has a time complexity at most linear in the size X of the discrete state space $\mathbb{X}_{\mathrm{g}}$. Then, the computation of the modified d-CDP operator (28) via Algorithm 2 has a time complexity of $\widetilde{\mathcal{O}}(X+Y+Z)$. Moreover, in the T-step application of Algorithm 2 for solving the Value Iteration Problem 3.1, the time complexity increases linearly with the horizon T.*

Proof. See Appendix B.3.1. □

Notice how the time complexity of Algorithm 2 compares with that of Algorithm 1, provided in Theorem 4.3. Particularly, if all of the involved grids ($\mathbb{X}_{\mathrm{g}}$, $\mathbb{Y}_{\mathrm{g}}$, and $\mathbb{Z}_{\mathrm{g}}$) are of the same size, then the complexity of Algorithm 1 is of $\mathcal{O}(X^2)$, while the complexity of Algorithm 2 is of $\mathcal{O}(X)$. Hence, for "separable" problems, application of Algorithm 2 can lead to a significant reduction in the complexity.

5.2.2. *Error analysis.* We next consider the error of the modified d-CDP algorithm. We begin with providing a bound on the difference between the outputs of the modified d-CDP operator (28) and the d-CDP operator (27).

Theorem 5.3 (Error of modified d-CDP – Algorithm 2)**.** *Consider Setting 2. Also, consider the implementation of the modified d-CDP operator $\widehat{\mathcal{T}}_{\mathrm{d}}^{\mathrm{m}}$ (28) via Algorithm 2. Let $\widehat{\mathcal{T}}_{\mathrm{d}}$ denote the output of the implementation of the d-CDP operator (27) via Algorithm 1. Assume that $J : \mathbb{X} \to \mathbb{R}$ is a Lipschitz-continuous, convex function, and that the grid $\mathbb{Z}_{\mathrm{g}}$ in Algorthim 2 is such that* $\mathrm{co}(\mathbb{Z}_{\mathrm{g}}) \supset f_{\mathrm{s}}(\mathbb{X}_{\mathrm{g}})$*. Then, for each $x \in \mathbb{X}_{\mathrm{g}}$, we have*

$$0 \le \widehat{\mathcal{T}}_{\mathrm{d}}^{\mathrm{m}}[J](x) - \widehat{\mathcal{T}}_{\mathrm{d}}[J](x) \le \delta_{\mathbb{Z}_{\mathrm{g}}} \cdot \Delta_{\mathbb{Y}_{\mathrm{g}}} =: e_3.$$

Proof. See Appendix B.3.2. □

Combining this result with the error analysis of $\widehat{\mathcal{T}}_{\mathrm{d}}$ in Section 4.2 for Algorithms 1, we can derive the bounds for the difference between the modified d-CDP operator (28) and the DP operator (26).

Corollary 5.4 (Error of modified d-CDP – Algorithm 2)**.** *Let the assumptions of Theorem 5.3 hold, and consider the DP operator* $\mathcal{T}$ (26)*. We have*

$$-\big(e_2+e_3\big) \leq \mathcal{T}[J](x) - \widehat{\mathcal{T}}^{\mathrm{m}}_{\mathrm{d}}[J](x) \leq e^m_1(x), \quad \forall x \in \mathbb{X}_{\mathrm{g}}, \tag{29}$$

where

$$\begin{aligned} e^m_1(x) &:= \big[\, \|f_{\mathrm{s}}(x)\| + \|B\| \cdot \Delta_{\mathbb{U}} + \Delta_{\mathbb{X}}\big] \cdot \mathrm{d}\left(\partial\big(\mathcal{T}[J] - C_{\mathrm{s}}\big)(x), \mathbb{Y}_{\mathrm{g}}\right), \\ e_2 &= \big[\Delta_{\mathbb{Y}_{\mathrm{g}}} + \mathrm{L}(J)\big] \cdot \mathrm{d}_{\mathrm{H}}(\mathbb{X}, \mathbb{X}_{\mathrm{g}}), \\ e_3 &= \delta_{\mathbb{Z}_{\mathrm{g}}} \cdot \Delta_{\mathbb{Y}_{\mathrm{g}}}. \end{aligned} \tag{30}$$

Proof. See Appendix B.3.3. □

Once again, the three terms capture the errors due to discretization in y, x, and z, respectively. We now use this result to provide some guidelines on the construction of the required grids.

Construction of $\mathbb{Y}_{\mathrm{g}}$ and $\mathbb{Z}_{\mathrm{g}}$. Concerning the grid $\mathbb{Y}_{\mathrm{g}}$, because of the error term e^m_1, similar guidelines to the ones provided in Section 4.2.2 apply here. In particular, notice that the first error term e^m_1 (30) now depends on $\mathrm{d}\left(\partial\big(\mathcal{T}[J]-C_{\mathrm{s}}\big)(x), \mathbb{Y}_{\mathrm{g}}\right)$, and hence in the construction of $\mathbb{Y}_{\mathrm{g}}$, we need to consider the range of slopes of $\mathcal{T}[J]-C_{\mathrm{s}}$. Next to be addressed is the construction of the grid $\mathbb{Z}_{\mathrm{g}}$. Here, again, we are dealing with the issue of constructing the dual grid for approximate discrete conjugation; see Corollary 2.8 for more details. Note that Theorem 5.3 assumes that $\mathrm{co}(\mathbb{Z}_{\mathrm{g}}) \supset f_{\mathrm{s}}(\mathbb{X}_{\mathrm{g}})$. This condition particularly guarantees that the application of LERP in Algorithm 2:6 only leads to *interpolative* approximate discrete conjugations. Also, notice that the state dynamics f_{s} and the state discretization $\mathbb{X}_{\mathrm{g}}$ are assumed to be time-invariant. This implies that the set $f_{\mathrm{s}}(\mathbb{X}_{\mathrm{g}})$ is time-invariant. This, in turn, allows us to choose a *fixed* grid $\mathbb{Z}_{\mathrm{g}}$ such that $\mathrm{co}(\mathbb{Z}_{\mathrm{g}}) \supset f_{\mathrm{s}}(\mathbb{X}_{\mathrm{g}})$. Construction of such a grid $\mathbb{Z}_{\mathrm{g}}$ then has a *one-time* (independent of the horizon T) computational cost of $\mathcal{O}(X)$. Alternatively, we can apply the method of Remark 4.6 to the convex-extensible function $\varphi : \mathbb{Y}_{\mathrm{g}} \to \mathbb{R}$ at each step, and still achieve the same error bound in Theorem 5.3; see Corollary 2.8 for more details. This scheme requires $\mathcal{O}(Y)$ operations *at each step.* Since the error term e_3 in Theorem 5.3 depends on the granularity $\delta_{\mathbb{Z}_{\mathrm{g}}}$ of the grid $\mathbb{Z}_{\mathrm{g}}$, in order to achieve the minimum error, we can apply both of the schemes described above and choose the grid $\mathbb{Z}_{\mathrm{g}}$ with smaller diameter $\Delta_{\mathbb{Z}_{\mathrm{g}}}$. This way, assuming that the size (number of the grid points) in $\mathbb{Z}_{\mathrm{g}}$ is fixed, we can have a smaller granularity $\delta_{\mathbb{Z}_{\mathrm{g}}}$, and hence a smaller error e_3.

5.3. **Extensions of modified d-CDP algorithm**

In this section, we discuss the possible extensions of the proposed modified d-CDP algorithm for problems with separable data. Let us first note that the same scheme described in Section 4.3.1 can be used for handling additive disturbances. Also, regarding Assumption 5.1, a similar scheme to the one provided in Section 4.3.2 can be used to compute the conjugate of input-dependent stage cost numerically. The pseudo-code for the multistep implementation of the modified d-CDP algorithm which also include these extensions is provided in Appendix C.1. In what follows we discuss two other extensions of the proposed d-CDP Algorithm 2 which can possibly lead to reduced computational complexity.

5.3.1. *Towards quantum dynamic programming.* Application of quantum computing for solving Optimal Control problems has attracted a lot of attentions recently. In particular, in [30], a quantum algorithm is proposed for solving the finite-horizon DP problem with a time complexity of $\mathcal{O}(X^{1/2} \cdot U^{9/2})$. Such a complexity is particularly attractive for problems with a huge state space and a relatively small action space, such as the Travelling Salesman Problem. More related to our work is the recent introduction of the quantum mechanical implementation of the LLT algorithm for the discrete conjugate transform in [33] which enjoys a poly-logarithmic complexity in the size of the discrete primal and dual domains. An interesting feature of the d-CDP Algorithm 2 is that one can readily leverage the quantum development of [33] and develop a quantum mechanical version of the modified d-CDP algorithm for problems of Setting 2. In this regard, we note that Algorithm 2 consists of three main operations: (i) LLT (lines 2 and 4), (ii) addition (lines 3 and 7), and (iii) composite extended function query (line 6). In particular, by choosing $Z, Y = X$, all these operations can be handled with a log-linear complexity in the size of the discrete state space X, leading to a log-linear time complexity for the d-CDP Algorithm 2 (see Theorem 5.2). The quantum algorithm proposed in [33], on the other hand, reduces the complexity of the LLT operations to $\mathcal{O}\big(\mathrm{poly}(\log X, \log Y)\big)$. To the best of our knowledge, the quantum-mechanical implementation of addition also has a poly-logarithmic complexity in the size of the input vectors. Thus, assuming that composite function query can be also be handled quantum-mechanically with a similar logarithmic complexity, we envision a quantum version of the d-CDP Algorithm 2 with a poly-logarithmic complexity $\mathcal{O}\big(\mathrm{poly}(\log X)\big)$ in this size of the discrete state space. Such a reduction in the time complexity is particularly interesting since it can address the infamous "curse of dimensionality" in the DP literature, to our knowledge, for the first time.

5.3.2. *Value iteration in the conjugate domain.* We finish this section with some remarks on the possibility of "perfect" transformation of the minimization in the primal domain to a simple addition in the conjugate domain, that is, performing the value iteration completely in the conjugate domain for the *conjugate* of the costs-to-go. In this regard, notice that the algorithms developed in this study involve two LLT transforms (of the cost functions) at the beginning and end of each step (see, e.g., lines 2 and 4 in Algorithm 2). A perfect transformation then allows us to stay in the conjugate domain over multiple steps in time, and is particularly interesting since the conjugate operation at the beginning of the intermediate steps can be avoided. This, in turn, leads to a lower computational cost in multistep implementations. In the following remark, we describe the class of problems for which such a perfect transformation is possible. As expected, we need to impose further restrictions on the problem class.

Remark 5.5 (Value iteration in the conjugate domain)**.** *Consider Setting 2. Moreover, let*

(i) the dynamics be linear*, i.e., $f(x,u) = Ax + Bu$, where the state matrix A is* invertible*,*
(ii) and the stage cost be state-independent *($C_{\mathrm{s}} = 0$), i.e., $C(x,u) = C_{\mathrm{i}}(u)$.*

For systems satisfying these conditions, the DP operator reads as

$$\mathcal{T}[J](x) = \min_u \{C_{\mathrm{i}}(u) + J(Ax + Bu)\}, \quad x \in \mathbb{X},$$

and the conjugate of the output of the DP operation is given by

$$\mathcal{T}[J]^*(y) = C_{\mathrm{i}}^*(-B^\top A^{-\top} y) + J^*(A^{-\top} y), \quad y \in \mathbb{R}^n.$$

Notice how the minimization in the DP operator in the primal domain perfectly *transforms to an addition in the dual domain. This property indeed allows us to stay in the dual domain over multiple steps in time, while only computing the* conjugate *of the costs in the intermediate steps.*

We note that the possibility of such a perfect transformation, accompanied by application of LLT for better time complexity, was first noticed in [12]. Indeed, in that paper, the authors introduced the Fast Value Iteration algorithm for a class of DP problems that satisfy the conditions of Remark 5.5 (and some other condition, e.g., the state matrix A being non-negative and monotone). In this regard, we also note that, as in [12], the possibility of staying in the conjugate domain over multiple steps is particularly interesting for infinite-horizon problems.

6. Numerical Experiments

In this section, we examine the performance of the proposed d-CDP algorithms in comparison with the d-DP algorithm through two synthetic numerical examples. In particular, we use these numerical examples to verify our theoretical results on the complexity and error of the proposed algorithms. In this section, we present the results of our numerical simulations for the two algorithms introduced in Sections 4 and 5. Here, we focus on the performance of the basic algorithms, introduced in Sections 4 and 5, for deterministic systems for which the conjugate of the (input-dependent) stage cost is analytically available (see Assumptions 4.2 and 5.1). The results of the numerical simulations of the extended versions of these algorithms are provided in Appendix C.1. In Appendix C.2, we also showcase the application of these algorithms in solving the optimal control problem for a simple epidemic model and a noisy inverted pendulum. In order to facilitate the application of the proposed algorithms, we also provide a d-CDP MATLAB package accompanied by the implementation of all the numerical examples presented in this article; see Appendix D for further details. Finally, we note that all the simulations presented in this article were implemented via MATLAB version R2017b, on a PC with Intel Xeon 3.60 GHz processor and 16 GB RAM.

Throughout this section, we consider a linear system with two states and two inputs described by

$$x_{t+1} = \begin{bmatrix} -0.5 & 2 \\ 1 & 3 \end{bmatrix} x_t + \begin{bmatrix} 1 & 0.5 \\ 1 & 1 \end{bmatrix} u_t, \tag{31}$$

over the finite horizon $T = 10$, with the following state and input constraints

$$x_t \in \mathbb{X} = [-1, 1]^2 \subset \mathbb{R}^2, \quad u_t \in \mathbb{U} = [-2, 2]^2 \subset \mathbb{R}^2. \tag{32}$$

Moreover, we consider *quadratic state cost* and *exponential input cost* as follows

$$C_{\mathrm{s}}(x) = C_T(x) = \|x\|^2, \quad C_{\mathrm{i}}(u) = e^{|u_1|} + e^{|u_2|} - 2. \tag{33}$$

We use *uniform* grid-like discretizations $\mathbb{X}_{\mathrm{g}} \subset \mathbb{X}$ and $\mathbb{U}_{\mathrm{g}} \subset \mathbb{U}$ for the state and input spaces, respectively. All of the other grids ($\mathbb{Y}_{\mathrm{g}}$, $\mathbb{Z}_{\mathrm{g}}$, and $\mathbb{V}_{\mathrm{g}}$) involved in the proposed d-CDP algorithms are also constructed *uniformly*, according to the guidelines provided in the respective sections. We are particularly interested in the performance (error and time complexity) of the d-CDP algorithms in comparison with the d-DP algorithm, as the size of these discrete sets increases. Considering the fact that all the involved grids in the following examples are constructed *uniformly*, and all the extension operations are handled via *LERP*, we can summarize the computational complexities as follows.

Remark 6.1 (Comparison of complexities)**.** *Assume that all the finite sets involved in the d-DP and d-CDP algorithms* ($\mathbb{X}_{\rm g}, \mathbb{U}_{\rm g}, \mathbb{Y}_{\rm g}, \mathbb{Z}_{\rm g}$) *are* uniform *grids, and that the extension of the cost functions in these algorithms (for solving the minimization 15 over the control input) is handled via* LERP *so that $E = 1$. Then, the complexity of finding the optimal input sequence in a T-step optimal control problem for a given initial state is of*

(i) $\mathcal{O}(TXU)$ *for d-DP algorithm,*
(ii) $\mathcal{O}\big(T(XY + U)\big)$ *for d-CDP Algorithm 1,*
(iii) $\mathcal{O}\big(T(X + Y + Z + U)\big)$ *for d-CDP Algorithm 2.*

6.1. Numerical study of Algorithm 1

In this section, in addition to the constraints (32), we also consider the following constraint

$$h(x_t, u_t) = x_t + u_t - (2, 2)^\top \leq 0,$$

where $h : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ describes *the state-dependent input constraints.* Corresponding to the notation of Section 4, the stage and terminal costs are respectively given by (see (33))

$$C(x, u) = \|x\|^2 + e^{|u_1|} + e^{|u_2|} - 2, \quad C_T(x) = \|x\|^2.$$

Then, the conjugate of the stage cost is indeed analytically available and given by

$$C^*_x(v) = 1 - \|x\|^2 + \langle \hat{u}, v\rangle - e^{|\hat{u}_1|} - e^{|\hat{u}_2|}, \quad v \in \mathbb{R}^2,$$

where

$$\hat{u}_i = \begin{cases} \max\big\{-2,\ \min\{2 - x_i,\ \operatorname{sgn}(v_i)\ln|v_i|\}\big\}, & v_i \neq 0, \\ 0, & v_i = 0, \end{cases} \qquad i = 1, 2.$$

We note that, for construction of $\mathbb{Y}_{\rm g}$, we use the method described in Section 4.2.2, with $\alpha = 1$.

We begin with examining the error in the d-DP and d-CDP algorithms w.r.t. the "reference" optimal costs-to-go $J^\star_t : \mathbb{X} \to \mathbb{R}$. *These reference costs $J^\star_t$ are computed numerically via a high resolution application of the d-DP algorithm with $X, U = 81^2$.* Figure 2 depicts the maximum absolute error in the cost functions J_t computed using these algorithms over the horizon. As expected and in line with our error analysis (Theorem 4.5 and Proposition A.1), using a finer discretization scheme with larger $N = X, Y, U$, leads to a smaller error. Moreover, for each gird size, a general increase is seen in the error, over the time steps in the backward iteration, due to accumulation of error. For further illustration, Figure 3 shows the corresponding costs-to-go at $t = 0$ and $t = 9$, with $N = 21^2$. In particular, notice how $J^{\rm DP}_0$ is not a convex function, while $J^{\rm CDP}_0$ is convex. Indeed, since the costs are convex and the dynamics is linear, the d-CDP algorithm preserves the convexity, while, due to the application of LERP extension, the d-DP algorithm can output non-convex cost functions.

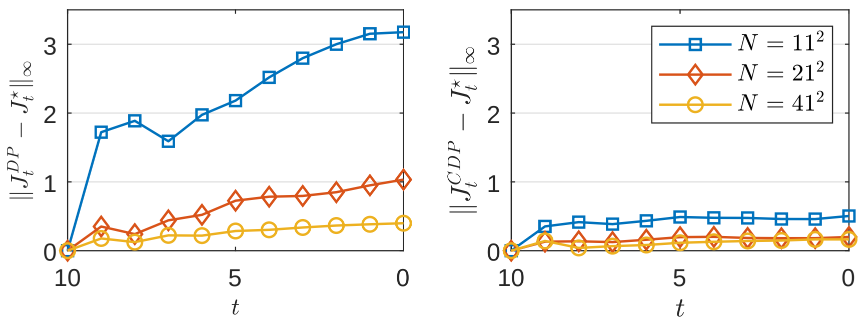


FIGURE 2. Error of Algorithm 1: the maximum absolute error (over $x \in \mathbb{X}_g$) in the computed costs-to-go from backward value iteration for different grid sizes ($X, Y, U = N$). Notice that the time axis is also backward.

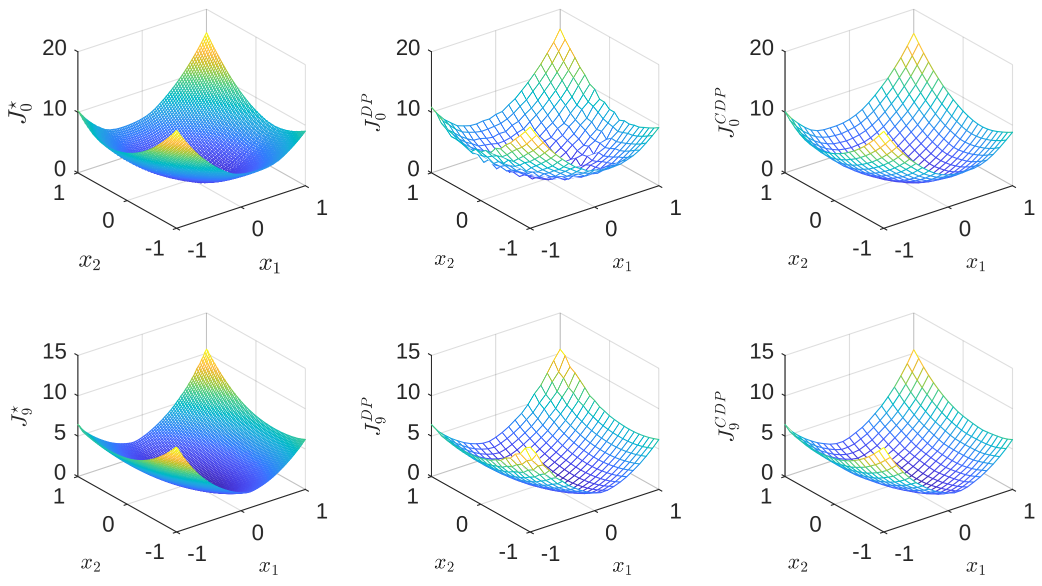


FIGURE 3. Error of Algorithm 1: the costs-to-go with $N = 21^2$ at $t = 0$ (top) and $t = 9$ (bottom).

We next compare the performance of the d-DP and d-CDP algorithms for solving ten instances of the optimal control problem, using the cost functions derived from the backward value iteration. In this regards, recall that the d-DP algorithm also provides us with the optimal laws $\mu_t : \mathbb{X}_g \to \mathbb{U}_g$. Hence, we report the performance of the d-DP algorithm for the two following cases:

(i) d-DP (J) denoting the case where the control sequence $\mathbf{u}^\star(x_0)$ is derived via minimization (over the discrete input space) of the computed costs-to-go $J_t : \mathbb{X}_g \to \mathbb{R}$; see also (17).

(ii) d-DP (π) denoting the case where the control sequence $\mathbf{u}^\star(x_0)$ is derived via LERP extension of the optimal control laws $\mu_t : \mathbb{X}_g \to \mathbb{U}_g$; see also (18).

Table 1 reports the average relative cost (w.r.t. the trajectory cost of d-DP (π) with $N = 81^2$) and the average total run-time (the running time of the backward value iteration plus the running time of the forward computation of control sequence for each initial condition) for different grid sizes. The average is taken over *ten* instances of the optimal control problem with random initial conditions, chosen uniformly from $\mathbb{X} = [-1,1]^2$. Comparing the first two rows of Table 1, we see that d-CDP has a slightly better performance compared to d-DP (J), w.r.t. both the trajectory cost and the running time. Indeed, by Remark 6.1, the time complexity of the both algorithms is of $\mathcal{O}(TN^2)$, which matches the reported running times. In this regard, we also note that the backward value iteration is the absolutely dominant factor in the reported running times. (Effectively, the reported numbers can be taken to be the run-time of backward iteration).

On the other hand, as reported in the last row of Table 1, d-DP (π) achieves a significantly better performance with regard to the cost of the controlled trajectories. This is because the control input computed using this approach is smooth, while the control input computed using the optimal cost functions (in the first two rows of Table 1) is limited to the discrete input space considered in the forward minimization problems. The d-CDP algorithm, however, gives us an extra degree of freedom for the size Y of the dual grid. Then, if the the cost functions are "compactly representable" in the dual domain via their slopes, we can reduce the time complexity by using a more coarse grid $\mathbb{Y}_{\mathrm{g}}$, with a limited effect on the "quality" of computed cost functions. Indeed, as reported in the first three cases in Table 2, for solving the same ten instances of the considered optimal control problem, we can reduce the size of the dual grid by a factor of 4, and hence reduce the running time of d-CDP algorithm, while achieving approximately the same average relative cost in the controlled trajectories. This, in turn, allows us to increase the size U of the discrete input space $\mathbb{U}_{\mathrm{g}}$ in order to generate a finer control input (and hence a smaller cost) in the minimizations of the forward iteration, while keeping the *total* running time at approximately the same level. As reported in Table 2, doing so, the d-CDP algorithm can achieve a similar performance to that of d-DP (π); cf.

Table 1. Performance of the d-CDP Algorithm 1 and the d-DP algorithm for different grid sizes $(X, Y, U = N)$: The reported numbers are the *average relative* trajectory cost (left – blue), and the *average total* running time (right – red); see the text for more details.

Relative trajectory cost / Running time (seconds)				
Alg. \ N	11^2	21^2	41^2	81^2
d-CDP	2.06 / 3.4e0	1.49 / 3.8e1	1.13 / 5.5e2	1.04 / 8.1e3
d-DP (J)	2.30 / 6.3e0	1.56 / 8.0e1	1.15 / 1.2e3	1.04 / 1.8e4
d-DP (π)	1.55 / 6.2e0	1.20 / 8.0e1	1.04 / 1.2e3	1 / 1.8e4

Table 2. Performance of the d-CDP Algorithm 1 using modified grid sizes (X, Y, U); cf. Table 1.

Relative trajectory cost / Running time (seconds)			
(X, Y, U)	$(11^2, 7^2, 11^2)$	$(21^2, 11^2, 21^2)$	$(41^2, 21^2, 41^2)$
d-CDP	2.07 / 1.9e0	1.52 / 1.5e1	1.13 / 1.9e2
(X, Y, U)	$(11^2, 7^2, 21^2)$	$(21^2, 11^2, 41^2)$	$(41^2, 21^2, 81^2)$
d-CDP	1.48 / 2.6e0	1.12 / 2.5e1	1.02 / 3.5e2

the last row of Table 1. However, we note that such a modification in the size of the grids leads to a significant increase in the running time of the forward iteration.

6.2. **Numerical study of Algorithm 2**

We now focus on the performance of the d-CDP Algorithm 2 for solving the optimal control problem described by the dynamics (31), constriants (32), and costs (33). The conjugate of the *input-dependent* stage cost now reads as

$$C_{\mathrm{i}}^{*}(v) = 1 + \langle \hat{u}, v\rangle - e^{|\hat{u}_1|} - e^{|\hat{u}_2|}, \quad v \in \mathbb{R}^2, \tag{34}$$

where

$$\hat{u}_i = \begin{cases} \max\left\{ -2,\ \min\left\{2,\ \mathrm{sgn}(v_i)\ln|v_i|\right\}\right\}, & v_i \neq 0, \\ 0, & v_i = 0, \end{cases} \qquad i = 1, 2.$$

For construction of $\mathbb{Y}_{\mathrm{g}}$, we again use the scheme described in Section 4.2.2, with $\alpha = 1$ and the required modification mentioned in Section 5.2.2. Also, the grid $\mathbb{Z}_{\mathrm{g}}$ is constructed such that $\mathrm{co}(\mathbb{Z}_{\mathrm{g}}) \supset A\mathbb{X}_{\mathrm{g}}$ according to guidelines of Section 5.2.2. We use the same set up as before for examining the performance of the d-CDP algorithm in comparison with the d-DP algorithm. Figure 4 depicts the maximum absolute error in the costs-to-go J_t computed using these algorithms over the horizon w.r.t. the reference $J_t^{\star}$ (i.e., the output of the d-DP algorithm with $X, U = 81^2$). In Table 3, we report the average relative cost and the average total run-time for different grid sizes (the average is taken over the same ten instances of the optimal control problem). Once again, in line with our error analysis (Theorem 5.3 and Proposition A.1), using finer grids leads to smaller errors. Also, the running times reported in Table 3 correspond to the the complexities given in Remark 6.1, i.e., $\mathcal{O}(TN^2)$ for d-DP algorithm and $\mathcal{O}(TN)$ for d-CDP Algorithm 2.

Comparing the first two rows of Table 3, we see that the d-CDP algorithm achieves a lower average cost compared to the d-DP algorithm, with a *significant* reduction in the running time. In particular, notice how the lower complexity of d-CDP allows us to increase the size of the grids to $N = 81^2$, while keeping the running time at the same order as the d-DP algorithm with $N = 11^2$. Moreover, the lower time complexity of the d-CDP algorithm also allows us to achieve a lower average cost in comparison with the d-DP (π) by using larger grid sizes in the d-CDP algorithm; ; e.g., compare the performance of d-CDP with $N = 41^2$ with that of d-DP (π) with $N = 21^2$ in

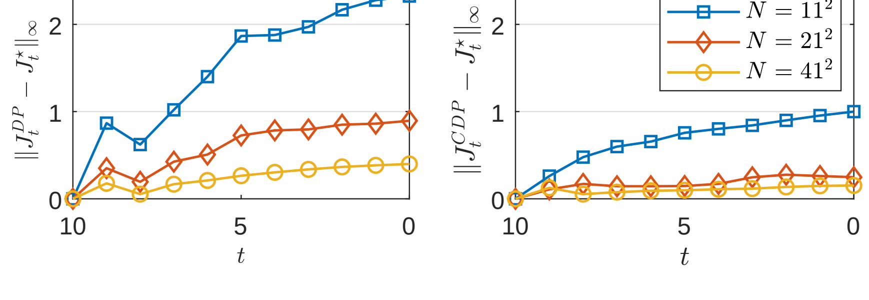


FIGURE 4. Error of Algorithm 2: the maximum absolute error (over $x \in \mathbb{X}_{\mathrm{g}}$) in the computed costs-to-go from backward value iteration for different grid sizes ($X, Y, Z, U = N$). Notice that the time axis is also backward.

Table 3. Performance of the d-CDP Algorithm 2 and the d-DP algorithm for different grid sizes $(X, Y, Z, U = N)$; see the caption of Table 1 and the text for more details.

Relative trajectory cost / Running time (seconds)				
Alg. \ N	11^2	21^2	41^2	81^2
d-CDP	2.02 / $1.8\text{e}-1$	1.47 / $3.9\text{e}-1$	1.13 / $1.5\text{e}+0$	1.02 / $7.5\text{e}+0$
d-DP (J)	2.30 / $3.3\text{e}+0$	1.55 / $4.2\text{e}+1$	1.15 / $6.2\text{e}+2$	1.04 / $1.0\text{e}+4$
d-DP (π)	1.56 / $3.3\text{e}+0$	1.20 / $4.2\text{e}+1$	1.03 / $6.2\text{e}+2$	1 / $1.0\text{e}+4$

Table 3. However, such an increase in the grid sizes in the d-CDP algorithm implies an increase in the running time of the forward iteration, and also a higher memory requirement in comparison with the d-DP algorithm.

7. Final Remarks

In this final section, we provide some remarks on the limitations of the proposed d-CDP algorithms and possible remedies to alleviate them. We also mention possible extensions of the current work as future research directions.

Preserving convexity in multistep implementation. Recall that the proposed alternative approach involves solving the dual program corresponding the minimization of the Bellman equation. As discussed before, due to this dualization, the d-CDP algorithm is essentially blind to non-convexity; see Proposition 4.4 and the discussion after. In this regard, note that the convexity of the stage and terminal costs (as assumed in this article) is necessary but not sufficient for the costs-to-go J_t to be convex for all $t = 1, \ldots, T$. In particular, for preservation of convexity in the proposed d-CDP algorithms, we also need the composition $J_t\big(f(x, u)\big)$ to be jointly convex in x and u, given that J_t is convex. This is, for example, the case when (the extended-value extension of) J_t is monotonic, and the dynamics f is convex in each argument [10, Sec. 3.2.4]. For Algorithm 2, there is another issue that further complicates the preservation of convexity in the multistep implementation: the approximate conjugation in the last step, particularly, the LERP extension in line 6. To see this, note that the mapping $x \mapsto \overline{\varphi^{\mathrm{d}*\mathrm{d}}}\big(f_{\mathrm{s}}(x)\big)$ is not necessarily convex, despite the fact that the underlying discrete function $\varphi^{\mathrm{d}*} : \mathbb{Z}_{\mathrm{g}} \to \mathbb{R}$ is convex-extensible.

Separability assumption in Setting 2. The significant reduction in the time complexity of solving the optimal control problem in Setting 2 via the proposed d-CDP Algorithm 2 is achieved by exploiting the separability of the problem data in the state and input variables. While the separability of cost ($C(x, u) = C_{\mathrm{s}}(x) + C_{\mathrm{i}}(u)$) is typical in Control applications, the separability of the state and input constraints is indeed quite restrictive. In particular, the d-CDP Algorithm 2 (and its extension), in its current format, cannot handle state-dependent input constraints (e.g., of the form $Cx + Du \leq d$). An interesting and impactful research direction is to look at possible extensions of the proposed algorithm that can handle interdependent constraints on the state and input variables.

The optimizer map in LLT. Consider a discrete function $h : \mathbb{X}_{\mathrm{d}} \to \mathbb{R}$ and its discrete conjugate $h^{\mathrm{d}*} : \mathbb{Y}_{\mathrm{d}} \to \mathbb{R}$ computed using LLT for some finite set $\mathbb{Y}_{\mathrm{d}}$. LLT is, in principle, capable of providing us with the optimizer mapping $x^\star : \mathbb{Y}_\mathrm{d} \to \mathbb{X}_\mathrm{d} : y \mapsto \mathrm{argmax}\{\langle x, y\rangle - h(x)\}$, where for each $y \in \mathbb{Y}_\mathrm{d}$, we have $h^{\mathrm{d}*}(y) = \langle x^\star(y), y\rangle - h\big(x^\star(y)\big)$. This capability of LLT can be employed to address some of the drawbacks of the proposed d-CDP algorithm:

(i) Avoiding approximate conjugation: Let us first recall that by approximate (discrete) conjugation we mean that we first compute the conjugate function $h^{\mathrm{d}*} : \mathbb{Y}_\mathrm{d} \to \mathbb{R}$ for some finite set $\mathbb{Y}_\mathrm{d}$, using the data points $h : \mathbb{X}_\mathrm{d} \to \mathbb{R}$; and then for any $\tilde{y}$ (not necessarily belonging to $\mathbb{Y}_\mathrm{d}$), we use the LERP extension $\overline{h^{\mathrm{d}*\mathrm{d}}}(\tilde{y})$ as an approximation for $h^{\mathrm{d}*}(\tilde{y})$. This approximation scheme is used in Algorithm 2 (and all the extended algorithms in Appendix C.1 in computing the conjugate of the stage cost numerically). Indeed, it is possible to avoid this approximation and compute $h^{\mathrm{d}*}(\tilde{y})$ exactly by using the data points $h^{\mathrm{d}*} : \mathbb{Y}_\mathrm{d} \to \mathbb{R}$, and by incorporating a smart search for the corresponding optimizer $\tilde{x} \in \mathbb{X}_\mathrm{d}$ for which $h^{\mathrm{d}*}(\tilde{y}) = \langle \tilde{x}, \tilde{y}\rangle - h(\tilde{x})$. To be precise, if $\tilde{y} \in \mathrm{co}(\widetilde{\mathbb{Y}}_\mathrm{d})$ for some subset $\widetilde{\mathbb{Y}}_\mathrm{d}$ of $\mathbb{Y}_\mathrm{d}$, then $\tilde{x} \in \mathrm{co}\left(x^\star(\widetilde{\mathbb{Y}}_\mathrm{d})\right)$, where $x^\star : \mathbb{Y}_\mathrm{d} \to \mathbb{X}_\mathrm{d}$ is the corresponding optimizer mapping; that is, in order to find the optimizer $\tilde{x} \in \mathbb{X}_\mathrm{d}$ corresponding to $\tilde{y}$, it suffices to search in the set $\mathbb{X}_\mathrm{d} \cap \mathrm{co}\left(x^\star(\widetilde{\mathbb{Y}}_\mathrm{d})\right)$, instead of the whole discrete primal domain $\mathbb{X}_\mathrm{d}$. This, in turn, can lead to lower time requirement for computing the *exact* discrete conjugate function.

(ii) Extracting the optimal policy within the d-CDP algorithm: As shown by our theoretical results and also confirmed by the numerical examples of Section 6, the d-CDP algorithms computationally outperforms the d-DP algorithm in solving the value iteration problem, i.e., computing the costs-to-go J_t backward in time. However, as discussed before, the d-CDP algorithms require solving minimization problems for finding the optimal control sequence $\mathbf{u}^\star(x_0)$ for a given initial condition x_0 forward in time; while, by using the d-DP algorithm, we have access to optimal control control laws μ_t, which can render finding $\mathbf{u}^\star(x_0)$ less costly. We note that, thanks to the significant reduction in the computational cost in the backward steps, the d-CDP algorithm can achieve a lower total running time compared to the d-DP algorithm, while keeping the cost of the generated controlled trajectories at the same level; see Tables 1 and 2. However, this is usually achieved by increasing the size of the discrete input space, which leads to a higher computational cost in the forward iteration. In order to address this issue, we have to look at the possibility of extracting the optimal policy within the d-CDP algorithm. A promising approach is to keep track of the dual pairs in each conjugate transform, i.e., the pairs (x, y) for which $\langle x, y\rangle = h(x) + h^*(y)$. This indeed seems possible considering the capability of LLT in providing the optimizer mapping $x^\star : \mathbb{Y}_\mathrm{d} \to \mathbb{X}_\mathrm{d}$.

Gird-like discretization and curse of dimensionality. In this study, we exclusively considered grid-like discretizations of both primal and dual domains. This is particularly suitable for problem with (almost) boxed constraints on the state and input spaces (as illustrated in the numerical examples in Section 6). More importantly, such discretizations suffer from the so-called "curse of dimensionality"; that is, the size of the finite representations of the corresponding spaces increases exponentially with the dimension of those spaces. In this regard, we note that in order to enjoy the linear-time complexity of LLT, we are *only* required to choose a grid-like *dual* grid [23, Rem. 5]; that is, the discretization of the state and input spaces in the primal domain need not be grid-like. However, since the grid-like dual domain is usually chosen to include the same number of points as the primal domain in each dimension (see Assumption 2.5), we still face the curse of dimensionality. This, in particular, impairs the performance of the d-CDP Algorithm 1 for problems in which the dimension of the state space is greater than that of the input space; see also the numerical example of Appendix C.2.1. Proper exploitation of the aforementioned property of LLT in such cases calls for a more efficient construction of the dual grid based on the provided data points in the primal domain.

Appendix A. Error of d-DP

In this section, we consider the error in the d-DP operator w.r.t. the DP operator.

Proposition A.1 (Error of d-DP)**.** *Consider the DP operator $\mathcal{T}$ (14) and the d-DP operator $\mathcal{T}_{\mathrm{d}}$ (16). Assume that the functions J, $\overline{J^{\mathrm{d}}}$, and C are* Lipschtiz-continuous*, and $\overline{J^{\mathrm{d}}}(x) = J(x)$ for all $x \in \mathbb{X}_{\mathrm{g}}$. Also, assume that the set of admissible inputs $\mathbb{U}(x)$ is compact for each $x \in \mathbb{X}_{\mathrm{g}}$, and denote $\mathbb{U}_{\mathrm{g}}(x) = \mathbb{U}(x) \cap \mathbb{U}_{\mathrm{g}}$. Then, for each $x \in \mathbb{X}_{\mathrm{g}}$, it holds that*

$$-e_1 \leq \mathcal{T}_{\mathrm{d}}[J](x) - \mathcal{T}[J](x) \leq e_1 + e_2(x),$$

where

$$\begin{aligned} e_1 &= \big[\mathrm{L}(J) + \mathrm{L}(\overline{J^{\mathrm{d}}})\big] \cdot \mathrm{d_H}(\mathbb{X}, \mathbb{X}_{\mathrm{g}}), \\ e_2(x) &= \big[\mathrm{L}(J) + \mathrm{L}(C)\big] \cdot \mathrm{d_H}\big(\mathbb{U}(x), \mathbb{U}_{\mathrm{g}}(x)\big). \end{aligned}$$

Proof. Define $Q_x(u) := C(x,u) + J\big(f(x,u)\big)$ and $\overline{Q}_x(u) := C(x,u) + \overline{J^{\mathrm{d}}}\big(f(x,u)\big)$. Let us fix $x \in \mathbb{X}_{\mathrm{g}}$. In what follows, we consider the effect of (i) replacing J with $\overline{J^{\mathrm{d}}}$, and (ii) minimizing over $\mathbb{U}_{\mathrm{g}}$ instead of $\mathbb{U}(x)$, separately. To this end, we define the *intermediate* DP operator

$$\mathcal{T}_{\mathrm{i}}[J](x) := \min_{u} \ \overline{Q}_x(u), \quad x \in \mathbb{X}_{\mathrm{g}}.$$

(i) Difference between $\mathcal{T}$ and $\mathcal{T}_{\mathrm{i}}$: Let $u^\star \in \operatorname*{argmin}_u Q(x,u) \subseteq \mathbb{U}(x)$, so that $\mathcal{T}[J](x) = Q(x,u^\star)$ and $\mathcal{T}_{\mathrm{i}}[J](x) \leq \overline{Q}(x,u^\star)$. Also, let $z^\star \in \operatorname*{argmin}_{z \in \mathbb{X}_{\mathrm{g}}} \|z - f(x,u^\star)\|$. Then,

$$\begin{aligned} \mathcal{T}_{\mathrm{i}}[J](x) - \mathcal{T}[J](x) &\leq \overline{Q}(x,u^\star) - Q(x,u^\star) \\ &= \overline{J^{\mathrm{d}}}\big(f(x,u^\star)\big) - \overline{J^{\mathrm{d}}}(z^\star) + J(z^\star) - J\big(f(x,u^\star)\big), \end{aligned}$$

where we used the assumption that $\overline{J^{\mathrm{d}}}(z^\star) = J(z^\star)$ for $z^\star \in \mathbb{X}_{\mathrm{g}}$. Hence,

$$\begin{aligned} \mathcal{T}_{\mathrm{i}}[J](x) - \mathcal{T}[J](x) &\leq \big[\mathrm{L}(J) + \mathrm{L}(\overline{J^{\mathrm{d}}})\big] \cdot \|z^\star - f(x,u^\star)\| \\ &= \big[\mathrm{L}(J) + \mathrm{L}(\overline{J^{\mathrm{d}}})\big] \cdot \min_{z \in \mathbb{X}_{\mathrm{g}}} \|z - f(x,u^\star)\| \\ &\leq \big[\mathrm{L}(J) + \mathrm{L}(\overline{J^{\mathrm{d}}})\big] \cdot \max_{z' \in \mathbb{X}} \ \min_{z \in \mathbb{X}_{\mathrm{g}}} \|z - z'\| \\ &= \big[\mathrm{L}(J) + \mathrm{L}(\overline{J^{\mathrm{d}}})\big] \cdot \mathrm{d_H}(\mathbb{X}, \mathbb{X}_{\mathrm{g}}) = e_1, \end{aligned}$$

where for the second inequality we used the fact that $f(x,u^\star) \in \mathbb{X}$. We can use the same line of arguments by defining $\overline{u}^\star \in \operatorname*{argmin}_u \overline{Q}(x,u)$, and $\overline{z}^\star \in \operatorname*{argmin}_{z \in \mathbb{X}_{\mathrm{g}}} \|z - f(x,\overline{u}^\star)\|$ to show that $\mathcal{T}_{\mathrm{i}}[J](x) - \mathcal{T}[J](x) \leq e_1$. Combining these results, we have

$$-e_1 \leq \mathcal{T}_{\mathrm{i}}[J](x) - \mathcal{T}[J](x) \leq e_1. \tag{35}$$

(ii) Difference between $\mathcal{T}_{\mathrm{i}}$ and $\mathcal{T}_{\mathrm{d}}$: First note that, by construction, we have $\mathcal{T}_{\mathrm{i}}[J](x) \leq \mathcal{T}_{\mathrm{d}}[J](x)$. Now, let $\overline{u}^{\star} \in \operatorname{argmin}_{u} \overline{Q}(x,u) \subseteq \mathbb{U}(x)$, so that $\mathcal{T}_{\mathrm{i}}[J](x) = \overline{Q}(x,\overline{u}^{\star})$. Also, let $\widetilde{u}^{\star} \in \operatorname{argmin}_{u\in\mathbb{U}_{\mathrm{g}}(x)} \|u-\overline{u}^{\star}\|$, and note that $\mathcal{T}_{\mathrm{d}}[J](x) \leq \overline{Q}(x,\widetilde{u}^{\star})$. Then, using the fact that $\overline{Q}$ is Lipschitz-continuous, we have

$$\begin{aligned}
0 \leq \mathcal{T}_{\mathrm{d}}[J](x) - \mathcal{T}_{\mathrm{i}}[J](x) &\leq \overline{Q}(x,\widetilde{u}^{\star}) - \overline{Q}(x,\overline{u}^{\star}) \leq \mathrm{L}(\overline{Q}_x)\cdot\|\widetilde{u}^{\star}-\overline{u}^{\star}\| \\
&\leq \big[\mathrm{L}(J)+\mathrm{L}(C)\big]\cdot \min_{u\in\mathbb{U}_{\mathrm{g}}(x)} \|u-\overline{u}^{\star}\| \\
&\leq \big[\mathrm{L}(J)+\mathrm{L}(C)\big]\cdot \max_{u'\in\mathbb{U}(x)} \min_{u\in\mathbb{U}_{\mathrm{g}}(x)} \|u-u'\| \\
&= \big[\mathrm{L}(J)+\mathrm{L}(C)\big]\cdot \mathrm{d}_{\mathrm{H}}\left(\mathbb{U}(x),\mathbb{U}_{\mathrm{g}}(x)\right) = e_2(x),
\end{aligned}$$

Combining this last result with the inequality (35), we derive the bounds in the statement of the proposition. □

Appendix B. Technical Proofs

B.1. Proofs of Section 2

B.1.1. *Proof of Lemma 2.6.* Let $y\in\mathbb{R}^n$, and observe that

$$h^{\mathrm{d}*}(y) = \max_{x\in\mathbb{X}_{\mathrm{d}}}\{\langle y,x\rangle - h(x)\} \leq \max_{x\in\mathbb{R}^n}\{\langle y,x\rangle - h(x)\} = h^*(y).$$

This settles the first inequality in (7) and (8). Now, assume that $\partial h^*(y)\neq\emptyset$, and let $x\in\partial h^*(y)$ so that $h(x)+h^*(y)=\langle y,x\rangle$ [8, Prop. 5.4.3]. Also, let $\widetilde{x}\in\operatorname{argmin}_{z\in\mathbb{X}_{\mathrm{d}}}\|x-z\|$, and note that $h^{\mathrm{d}*}(y)\geq \langle y,\widetilde{x}\rangle - h(\widetilde{x})$. Then,

$$\begin{aligned}
h^*(y)-h^{\mathrm{d}*}(y) &\leq \langle y, x-\widetilde{x}\rangle - h(x) + h(\widetilde{x}) \\
&\leq \big[\,\|y\| + \mathrm{L}\left(h;\{x\}\cup\mathbb{X}_{\mathrm{d}}\right)\big]\cdot\|x-\widetilde{x}\|\,.
\end{aligned}$$

Hence, by minimizing over the choice $x\in\partial h^*(y)$, we derive the upper bound provided in (7). In particular, note that if $\partial h^*(y)=\emptyset$, then the upper bound becomes trivial, i.e., $\widetilde{e}_1=\infty$. Finally, the additional constraints of compactness of $\mathbb{X}=\mathrm{dom}(h)$ implies that $\partial h^*(y)\cap\mathbb{X}\neq\emptyset$. Hence, we can choose $x\in\partial h^*(y)\cap\mathbb{X}$ and use Lipschitz-continuity of h to write

$$\begin{aligned}
h^*(y)-h^{\mathrm{d}*}(y) &\leq \big[\,\|y\| + \mathrm{L}\left(h;\{x\}\cup\mathbb{X}_{\mathrm{d}}\right)\big]\cdot \mathrm{d}(x,\mathbb{X}_{\mathrm{d}}) \\
&\leq \big[\,\|y\|+\mathrm{L}(h)\big]\cdot\max_{z\in\mathbb{X}}\mathrm{d}(z,\mathbb{X}_{\mathrm{d}}) = \widetilde{e}_2(y,h,\mathbb{X}_{\mathrm{d}}).
\end{aligned}$$

B.1.2. *Proof of Lemma 2.7.* Let us first consider the case $y\in\mathrm{co}(\mathbb{Y}_{\mathrm{g}})$. The value of the multilinear interpolation $\overline{h^{*\mathrm{d}}}(y)$ is a convex combination of $h^{*\mathrm{d}}(y^{(k)}) = h^*(y^{(k)})$ over the grid points $y^{(k)}\in\mathbb{Y}_{\mathrm{g}},\ k\in 1,\dots,2^n$, located at the vertices of the hyper-rectangular cell that contains y. That is, $\overline{h^{*\mathrm{d}}}(y) = \sum_k \alpha^{(k)}\ h^*(y^{(k)})$, where $\sum_k\alpha^{(k)}=1$ and $\alpha^{(k)}\in[0,1]$. Note that, since we are using LERP, we also have $y=\sum_k\alpha^{(k)}\ y^{(k)}$. Then,

$$h^*(y) = h^*\left(\textstyle\sum_k \alpha^{(k)}\ y^{(k)}\right) \leq \textstyle\sum_k\alpha^{(k)}\ h^*(y^{(k)}) = \overline{h^{*\mathrm{d}}}(y), \tag{36}$$

where the inequality follows from the convexity of h^*. Also, notice that

$$\overline{h^{*\mathrm{d}}}(y) = \textstyle\sum_k\alpha^{(k)}\ h^*(y^{(k)}) = \sum_k\alpha^{(k)}\ \displaystyle\max_{x\in\mathbb{X}}\left\{\langle y^{(k)},x\rangle - h(x)\right\}$$

$$
\begin{aligned}
&= \textstyle\sum_k \alpha^{(k)} \max_{x\in\mathbb{X}} \left\{ \langle y, x\rangle - h(x) + \langle y^{(k)} - y, x\rangle \right\} \\
&\leq \textstyle\sum_k \alpha^{(k)} \max_{x\in\mathbb{X}} \left\{ \langle y, x\rangle - h(x) + \left\| y^{(k)} - y\right\| \cdot \|x\| \right\} \\
&\leq \textstyle\sum_k \alpha^{(k)} \max_{x\in\mathbb{X}} \left\{ \langle y, x\rangle - h(x) + \delta_{\mathbb{Y}_{\mathrm{g}}} \cdot \Delta_{\mathbb{X}} \right\},
\end{aligned}
$$

where for the last inequality we used the fact that y is contained in the cell with $y^{(k)}$'s as its vertices, and hence the granularity $\delta_{\mathbb{Y}_{\mathrm{g}}}$ provides an upper bound for $\left\| y^{(k)} - y \right\|$ for all k. Then, using $\sum_k \alpha^k = 1$, we have

$$
\overline{h^{*\mathrm{d}}}(y) \leq \max_{x\in\mathbb{X}} \left\{ \langle y, x\rangle - h(x) \right\} + \delta_{\mathbb{Y}_{\mathrm{g}}} \cdot \Delta_{\mathbb{X}} \leq h^*(y) + \delta_{\mathbb{Y}_{\mathrm{g}}} \cdot \Delta_{\mathbb{X}}. \tag{37}
$$

Combining the two inequalities (36) and (37) gives us the inequality (9) in the statement of the lemma.

We next consider the case $y \notin \mathrm{co}(\mathbb{Y}_{\mathrm{g}})$ under the extra assumption $\mathrm{co}(\mathbb{Y}'_{\mathrm{g}}) \supset \mathbb{Y}(h)$. Note that this assumption implies that (consult the notation preceding the lemma):

- $\mathbb{Y}(h)$ is bounded (h is Lipschitz-continuous); and,
- $y_i^1 < y_i^2 < \mathrm{L}_i^-(h)$ and $\mathrm{L}_i^+(h) < y_i^{N_i-1} < y_i^{N_i}$ for all $i \in \{1, \ldots, n\}$.

In order to simplify the exposition, we consider the two-dimensional case ($n = 2$), while noting that the provided arguments can be generalized to higher dimensions. So, let $\mathbb{Y}_{\mathrm{g}} = \mathbb{Y}_{\mathrm{g}1} \times \mathbb{Y}_{\mathrm{g}2}$, where $\mathbb{Y}_{\mathrm{g}i}$ ($i = 1, 2$) is the finite set of real numbers $y_i^1 < y_i^2 < \ldots < y_i^{N_i}$ with $N_i \geq 5$. Let us further simplify the argument by letting $y = (y_1, y_2) \notin \mathrm{co}(\mathbb{Y}_{\mathrm{g}})$ be such that $y_1 < y_1^1$ and $y_2^1 \leq y_2 \leq y_2^2$, so that computing $\overline{h^{*\mathrm{d}}}(y)$ involves extrapolation in the first dimension and interpolation in the second dimension. See Figure 5a for a visualization of this instantiation. Since the extension uses LERP, using the points depicted in Figure 5a, we can write

$$
\overline{h^{*\mathrm{d}}}(y) = \alpha \; \overline{h^{*\mathrm{d}}}(y') + (1-\alpha) \; \overline{h^{*\mathrm{d}}}(y''), \tag{38}
$$

where $\alpha = (y_1^2 - y_1)/(y_1^2 - y_1^1)$, and

$$
\overline{h^{*\mathrm{d}}}(y') = \beta \; h^*(y^{1,1}) + (1-\beta) \; h^*(y^{1,2}), \quad \overline{h^{*\mathrm{d}}}(y'') = \beta \; h^*(y^{1,2}) + (1-\beta) \; h^*(y^{2,2}), \tag{39}
$$

where $\beta = (y_2^2 - y_2)/(y_2^2 - y_2^1)$. In Figure 5a, we have also paired each of the points of interest in the dual domain with its corresponding maximizer in the primal domain. That is, for $\xi = y, y', y'', y^{1,1}, y^{1,2}, y^{1,2}, y^{2,2}$, we have respectively identified $\eta = x, x', x'', x^{1,1}, x^{1,2}, x^{1,2}, x^{2,2} \in \mathbb{X}$, where $\xi \in \partial h(\eta)$ so that

$$
h^*(\xi) = \langle \eta, \xi \rangle - h(\eta). \tag{40}
$$

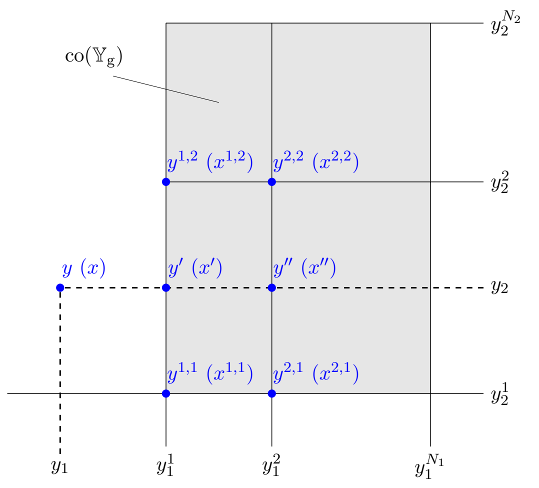
We now list the *implications* of the assumption $y_1^1 < y_1^2 < \mathrm{L}_1^-(h)$ – Figure 5b illustrates these implications in the one-dimensional case:

I.1. We have $h^*(y) = \alpha \; h^*(y') + (1-\alpha) \; h^*(y'')$.

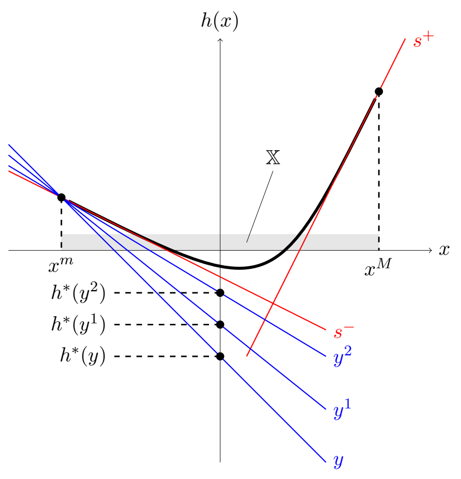
I.2. We can choose the maximizers in the primal domain such that

  I.2.1. $x^{1,1} = x^{2,1}$, $x^{1,2} = x^{2,2}$, and $x = x' = x''$;

  I.2.2. $x_1^{1,1} = x_1^{1,2} = x_1 = \min_{(z_1,z_2)\in\mathbb{X}} z_1$.



(a) Position of the point y w.r.t. the grid $\mathbb{Y}_\mathrm{g}$



(b) Implications of the assumption

FIGURE 5. Illustration of the proof of Lemma 2.7. (a) The dual grid $\mathbb{Y}_\mathrm{g}$ and the position of the point y w.r.t. the grid. The blue dots show the points of interest and their corresponding maximizer in the primal domain. E.g., "$y\ (x)$" implies that $y \in \partial h(x)$, where $x \in \mathbb{X}$, so that $\langle x, y\rangle = h(x) + h^*(y)$. (b) Illustration of the implications of the assumption $y^1 < y^2 < s^- = \mathrm{L}^-(h)$ in the one-dimensional case. The colored (red and blue) variables denote the slope of the corresponding lines. Note that $\{y, y^1, y^2\} \subset \partial h(x^m)$, where $x^m = \min_{x\in\mathbb{X}} x$. Indeed, for all $y \leq s^-$, the conjugate $h^*(y) = \langle x^m, y\rangle - h(x^m)$ is a linear function with slope x^m. In particular, for $y < y^1$, we have $h^*(y) = \alpha h^*(y^1) + (1-\alpha)h^*(y^2)$, where $\alpha = (y^2 - y)/(y^2 - y^1)$.

With these preparatory discussions, we can now consider the error of extrapolative discrete conjugation at the point y. In this regard, first note that $\{y', y''\} \subset \mathrm{co}(\mathbb{Y}_\mathrm{g})$, and hence we can use the result of first part of the lemma to write

$$\overline{h^{*\mathrm{d}}}(y') = h^*(y') + e', \quad \overline{h^{*\mathrm{d}}}(y'') = h^*(y'') + e'', \tag{41}$$

where $\{e', e''\} \subset [0, \delta_{\mathbb{Y}_\mathrm{g}} \cdot \Delta_\mathbb{X}]$. We claim that these error terms are equal. Indeed, from (39) and (41), we have

$$e' - e'' = \beta\left[h^*(y^{1,1}) - h^*(y^{2,1})\right] + (1-\beta)\left[h^*(y^{1,2}) - h^*(y^{2,2})\right] + h^*(y'') - h^*(y').$$

Then, using the pairings in (40) and the implication I.2, we can write

$$\begin{aligned} e' - e'' &\overset{(I.2.1)}{=} \beta\left\langle x^{1,1}, y^{1,1} - y^{2,1}\right\rangle + (1-\beta)\left\langle x^{1,2}, y^{1,2} - y^{2,2}\right\rangle + \left\langle x, y'' - y'\right\rangle \\ &= \beta\left\langle x^{1,1}, (y_1^1 - y_1^2, 0)\right\rangle + (1-\beta)\left\langle x^{1,2}, (y_1^1 - y_1^2, 0)\right\rangle + \left\langle x, (y_1^2 - y_1^1, 0)\right\rangle \\ &= \left(\beta x_1^{1,1} + (1-\beta)x_1^{1,2} - x_1\right)(y_1^1 - y_1^2) \overset{(I.2.2)}{=} 0. \end{aligned}$$

With this result at hand, we can employ the equality (38) and the implication I.1 to write

$$\overline{h^{*\mathrm{d}}}(y) - h^*(y) = \alpha\left[\overline{h^{*\mathrm{d}}}(y') - h^*(y')\right] + (1-\alpha)\left[\overline{h^{*\mathrm{d}}}(y'') - h^*(y'')\right] = \alpha e' + (1-\alpha)e'' = e'.$$

That is,

$$0 \leq \overline{h^{*\mathrm{d}}}(y) - h^*(y) \leq \delta_{\mathbb{Y}_\mathrm{g}} \cdot \Delta_\mathbb{X}.$$

B.1.3. *Proof of Corollary 2.8.* The first state immediately follows from Lemma 2.7 since the finite set $\mathbb{X}_{\mathrm{g}}$ is compact. For the second statement, the extra condition $\mathrm{co}(\mathbb{Y}'_{\mathrm{g}}) \supseteq \mathbb{Y}(h)$ has the same implications as the ones provided for the proof of Lemma 2.7 in Appendix B.1.2. Hence, following the same arguments, we can show that provided bounds hold for all $y \in \mathbb{R}^n$ under the aforementioned condition.

B.2. **Proofs of Section 4**

B.2.1. *Proof of Lemma 4.1.* Using the definition of conjugate transform, we have

$$\begin{aligned}
\widehat{\mathcal{T}}[J](x) &= \max_{y\in\mathbb{R}^n} \min_{u,z\in\mathbb{R}^n} \{C(x,u) + J(z) + \langle y, f_{\mathrm{s}}(x) + f_{\mathrm{i}}(x)u - z\rangle\} \\
&= \max_{y} \Big\{ \langle y, f_{\mathrm{s}}(x)\rangle - \max_{u} \Big[\Big\langle -f_{\mathrm{i}}(x)^\top y, u\Big\rangle - C(x,u)\Big] - \max_{z} [\langle y, z\rangle - J(z)] \Big\} \\
&= \max_{y} \Big\{ \langle y, f_{\mathrm{s}}(x)\rangle - C_x^*(-f_{\mathrm{i}}(x)^\top y) - J^*(y) \Big\} \\
&= \max_{y} \{\langle y, f_{\mathrm{s}}(x)\rangle - \phi_x(y)\} = \phi_x^*\big(f_{\mathrm{s}}(x)\big).
\end{aligned}$$

B.2.2. *Proof of Theorem 4.3.* In what follows, we provide the time complexity of each line of Algorithm 1. By assumption, the time complexity of construction of $\mathbb{Y}_{\mathrm{g}}$ in line 1 is at most of $\mathcal{O}(X)$. The LLT of line 2 requires $\mathcal{O}(X+Y)$ operations; see Remark 2.4. By Assumption 4.2, computing φ_x in line 4 has a complexity of $\mathcal{O}(Y)$. The enumeration in line 5 also has a complexity of $\mathcal{O}(Y)$. This, in turn, implies that the `for loop` over $x \in \mathbb{X}_{\mathrm{g}}$ requires $\mathcal{O}(XY)$ operations. Hence, the time complexity of the whole algorithm is of $\mathcal{O}(XY)$.

B.2.3. *Proof of Proposition 4.4.* We can use the representation (22) and the definition (20) to obtain

$$\begin{aligned}
\widehat{\mathcal{T}}_{\mathrm{d}}[J](x) &= \max_{y\in\mathbb{Y}_{\mathrm{g}}} \{\langle f_{\mathrm{s}}(x), y\rangle - \varphi_x(y)\} = \max_{y\in\mathbb{Y}_{\mathrm{g}}} \Big\{ \langle f_{\mathrm{s}}(x), y\rangle - C_x^*(-f_{\mathrm{i}}(x)^\top y) - J^{\mathrm{d}*}(y) \Big\} \\
&= \max_{y\in\mathbb{Y}_{\mathrm{g}}} \Big\{ \langle f_{\mathrm{s}}(x), y\rangle - \max_{u\in\mathrm{dom}\, C(x,\cdot)} \Big[\Big\langle -f_{\mathrm{i}}(x)^\top y, u\Big\rangle - C(x,u)\Big] - J^{\mathrm{d}*}(y) \Big\} \\
&= \max_{y\in\mathbb{Y}_{\mathrm{g}}} \min_{u\in\mathrm{dom}\, C(x,\cdot)} \Big\{ C(x,u) + \langle y, f(x,u)\rangle - J^{\mathrm{d}*}(y) \Big\}.
\end{aligned}$$

By the properties laid out in Setting 1, the objective function of this maximin problem is convex in u, with $\mathrm{dom}\,\big(C(x,\cdot)\big)$ being compact. Also, the objective function is concave in y, which follows from the convexity of $J^{\mathrm{d}*}$. Then, by the Ky Fan's Minimax Theorem (see, e.g., [19, Thm. A]), we can swap the maximization and minimization operators, without affecting the optimal value, to obtain

$$\begin{aligned}
\widehat{\mathcal{T}}_{\mathrm{d}}[J](x) &= \min_{u\in\mathrm{dom}\, C(x,\cdot)} \max_{y\in\mathbb{Y}_{\mathrm{g}}} \Big\{ C(x,u) + \langle y, f(x,u)\rangle - J^{\mathrm{d}*}(y) \Big\} \\
&= \min_{u} \Big\{ C(x,u) + J^{\mathrm{d}*\mathrm{d}*}\big(f(x,u)\big) \Big\}.
\end{aligned}$$

B.2.4. *Proof of Theorem 4.5.* Let us first note that the convexity of $J : \mathbb{X} \to \mathbb{R}$ implies that the duality gap is zero. Indeed, following a similar argument as the one provided in the proof of

Proposition 4.4 in Appendix B.2.3, and using Sion's Minimax Theorem (see, e.g., [32, Thm. 3]), we can show that the CDP operator (21) equivalently reads as

$$\widehat{\mathcal{T}}[J](x) = \min_{u} \left\{ C(x,u) + J^{**}\big(f(x,u)\big) \right\}, \quad x \in \mathbb{X}_{\mathrm{g}}.$$

Then, since J is a proper, closed, convex function, we have $J^{**} = J$, and hence $\widehat{\mathcal{T}}[J] = \mathcal{T}[J]$.

We next consider the disctretization error in $\widehat{\mathcal{T}}_{\mathrm{d}}$ (22) w.r.t. $\widehat{\mathcal{T}}$ (21). First, we can use Lemma 2.6, and the fact that $\mathrm{dom}(J) = \mathbb{X}$ is compact, to write

$$0 \le \phi_x(y) - \varphi_x(y) = J^*(y) - J^{\mathrm{d}*}(y) \le \widetilde{e}_2(y, J, \mathbb{X}_{\mathrm{g}}) \le \max_{y \in \mathbb{Y}_{\mathrm{g}}} \widetilde{e}_2(y, J, \mathbb{X}_{\mathrm{g}}) = e_2, \quad \forall y \in \mathbb{Y}_{\mathrm{g}}.$$

The preceding inequality captures the error due to discretization of the primal domain $\mathbb{X}$, i.e., using $J^{\mathrm{d}*}$ in (22b) instead of J^* in (21b). Using this inequality and the definition of discrete conjugate, we can write

$$0 \le \varphi_x^{\mathrm{d}*}\big(f_{\mathrm{s}}(x)\big) - \phi_x^{\mathrm{d}*}\big(f_{\mathrm{s}}(x)\big) \le e_2, , \quad \forall x \in \mathbb{X}_{\mathrm{g}}. \tag{42}$$

We can also use Lemma 2.6, to write

$$0 \le \phi_x^{*}\big(f_{\mathrm{s}}(x)\big) - \phi_x^{\mathrm{d}*}\big(f_{\mathrm{s}}(x)\big) \le \widetilde{e}_1(f_{\mathrm{s}}(x), \phi_x, \mathbb{Y}_{\mathrm{g}}), \quad \forall x \in \mathbb{X}_{\mathrm{g}}.$$

This captures the error due to discretization of the dual domain $\mathbb{Y} = \mathbb{R}^n$, i.e., approximating ϕ_x^* in (21a) via $\varphi_x^{\mathrm{d}*}$ in (22a). Now, observe that

$$\begin{aligned} \widetilde{e}_1(f_{\mathrm{s}}(x), \phi_x, \mathbb{Y}_{\mathrm{g}}) &= \min_{y \in \partial \phi_x^*(f_{\mathrm{s}}(x))} \left\{ \big[\, \|f_{\mathrm{s}}(x)\| + \mathrm{L}\left(\phi_x; \{y\} \cup \mathbb{Y}_{\mathrm{g}}\right) \big] \cdot \mathrm{d}(y, \mathbb{Y}_{\mathrm{g}}) \right\} \\ &\le \min_{y \in \partial \mathcal{T}[J](x)} \left\{ \big[\, \|f_{\mathrm{s}}(x)\| + \|f_{\mathrm{i}}(x)\| \cdot \Delta_{\mathbb{U}} + \Delta_{\mathbb{X}} \big] \cdot \mathrm{d}(y, \mathbb{Y}_{\mathrm{g}}) \right\}, \end{aligned}$$

where we used the fact that $\phi_x^*\big(f_{\mathrm{s}}(\cdot)\big) = \widehat{\mathcal{T}}[J](\cdot) = \mathcal{T}[J](\cdot)$, and

$$\begin{aligned} \mathrm{L}\left(\phi_x(\cdot)\right) &\le \mathrm{L}\left(C_x^*(-f_{\mathrm{i}}(x)^\top \cdot)\right) + \mathrm{L}\left(J^*(\cdot)\right) \le \|f_{\mathrm{i}}(x)\| \cdot \mathrm{L}(C_x^*) + \mathrm{L}(J^*) \\ &\le \|f_{\mathrm{i}}(x)\| \cdot \Delta_{\mathrm{dom}(C(x,\cdot))} + \Delta_{\mathrm{dom}(J)} \le \|f_{\mathrm{i}}(x)\| \cdot \Delta_{\mathbb{U}} + \Delta_{\mathbb{X}}. \end{aligned}$$

Hence, for each $x \in \mathbb{X}_{\mathrm{g}}$ we have

$$\begin{aligned} 0 \le \phi_x^{*}\big(f_{\mathrm{s}}(x)\big) - \phi_x^{\mathrm{d}*}\big(f_{\mathrm{s}}(x)\big) &\le \widetilde{e}_1(f_{\mathrm{s}}(x), \phi_x, \mathbb{Y}_{\mathrm{g}}) \\ &\le \big[\, \|f_{\mathrm{s}}(x)\| + \|f_{\mathrm{i}}(x)\| \cdot \Delta_{\mathbb{U}} + \Delta_{\mathbb{X}} \big] \cdot \mathrm{d}\left(\partial \mathcal{T}[J](x), \mathbb{Y}_{\mathrm{g}}\right) = e_1(x). \end{aligned}$$

Combining the last inequality with the inequality (42) provides the bounds in the statement of the theorem.

B.3. **Proofs of Section 5**

B.3.1. *Proof of Theorem 5.2.* In what follows, we provide the time complexity of each line of Algorithm 2. By assumption, the time complexity of construction of $\mathbb{Y}_{\mathrm{g}}$ in line 1 is at most of $\mathcal{O}(X)$. The LLT of line 2 requires $\mathcal{O}(X+Y)$ operations; see Remark 2.4. By assumption, computing φ in line 3 has a complexity of $\mathcal{O}(Y)$. The LLT of line 4 requires $\mathcal{O}(Y+Z)$ operations. The approximation of line 6 using LERP has a complexity of $\mathcal{O}(\log Z)$; see Remark 2.2. Hence, the `for loop` over $x \in \mathbb{X}_{\mathrm{g}}$ requires $\mathcal{O}(X \log Z)$ operations. The time complexity of the whole algorithm can then be computed by adding all the aforementioned complexities.

B.3.2. *Proof of Theorem 5.3.* Let us first note that the computation of the modified d-CDP operator $\widehat{\mathcal{T}}_{\mathrm{d}}^{\mathrm{m}}$ (28) via Algorithm 2 differs form that of the d-CDP operator $\widehat{\mathcal{T}}_{\mathrm{d}}$ (27) via Algorithm 1 only in the last step. To see this, note that computation of $\widehat{\mathcal{T}}_{\mathrm{d}}$ *exactly* computes $\varphi^{\mathrm{d}*}\big(f_{\mathrm{s}}(x)\big)$ for $x \in \mathbb{X}_{\mathrm{d}}$ (see Algorithm 1:5). However, in $\widehat{\mathcal{T}}_{\mathrm{d}}^{\mathrm{m}}$, the approximation $\overline{\varphi^{\mathrm{d}*\mathrm{d}}}\big(f_{\mathrm{s}}(x)\big)$ is used (see Algorithm 2:6), where the approximation uses LERP over the data points $\varphi^{\mathrm{d}*} : \mathbb{Z}_{\mathrm{g}} \to \mathbb{R}$. By Corollary 2.8, this leads to an over-approximation of $\varphi^{\mathrm{d}*}$, with the upper bound $e_3 = \delta_{\mathbb{Z}_{\mathrm{g}}} \cdot \Delta_{\mathbb{Y}_{\mathrm{g}}}$ on the error. Hence, compared to $\widehat{\mathcal{T}}_{\mathrm{d}}$, the operator $\widehat{\mathcal{T}}_{\mathrm{d}}^{\mathrm{m}}$ is an over-approximation with the difference bounded by e_3, that is,

$$0 \le \widehat{\mathcal{T}}_{\mathrm{d}}^{\mathrm{m}}[J](x) - \widehat{\mathcal{T}}_{\mathrm{d}}[J](x) \le e_3, \quad \forall x \in \mathbb{X}_{\mathrm{g}}.$$

B.3.3. *Proof of Corollary 5.4.* The result follows from Theorem 4.5 and 5.3. Indeed, form the definition of the d-CDP operator (27), we have

$$\begin{aligned} \widehat{\mathcal{I}}_{\mathrm{d}}[J](x) &:= \widehat{\mathcal{T}}_{\mathrm{d}}[J](x) - C_{\mathrm{s}}(x) = \varphi^{\mathrm{d}*}\big(f_{\mathrm{s}}(x)\big), \quad x \in \mathbb{X}_{\mathrm{g}}, \\ \varphi(y) &:= C_{\mathrm{i}}^{*}(-B^{\top} y) + J^{\mathrm{d}*}(y), \quad y \in \mathbb{Y}_{\mathrm{g}}. \end{aligned}$$

Similarly, for the DP operator (26), we can write

$$\mathcal{I}[J](x) := \mathcal{T}[J](x) - C_{\mathrm{s}}(x) = \min_{u} \left\{ C_{\mathrm{i}}(u) + J\big(f(x,u)\big) \right\}.$$

Then, by Theorem 4.5, it holds that

$$-e_2 \le \mathcal{I}[J](x) - \widehat{\mathcal{I}}_{\mathrm{d}}[J](x) = \mathcal{T}[J](x) - \widehat{\mathcal{T}}_{\mathrm{d}}[J](x) \le e_1^{m}(x), \quad \forall x \in \mathbb{X}_{\mathrm{g}}, \tag{43}$$

where e_2 is as in (25), and

$$\begin{aligned} e_1^{m}(x) &= \big[\, \|f_{\mathrm{s}}(x)\| + \|B\| \cdot \Delta_{\mathbb{U}} + \Delta_{\mathbb{X}} \big] \cdot \mathrm{d}\left(\partial \mathcal{I}[J](x), \mathbb{Y}_{\mathrm{g}}\right) \\ &= \big[\, \|f_{\mathrm{s}}(x)\| + \|B\| \cdot \Delta_{\mathbb{U}} + \Delta_{\mathbb{X}} \big] \cdot \mathrm{d}\left(\partial\big(\mathcal{T}[J] - C_{\mathrm{s}}\big)(x), \mathbb{Y}_{\mathrm{g}}\right). \end{aligned}$$

Combining the inequality (43) with the error bound of Theorem 5.3 completes the proof.

Appendix C. Extended Algorithms & Further Numerical Examples

C.1. Extended algorithms and their numerical study

In this section, we provide the multistep version of d-CDP algorithms developed in this study that also take into account the extensions discussed in Section 4.3, that is, additive disturbance in the dynamics and numerical computation of the conjugate of the (input-dependent) stage cost. The provided algorithms are

(i) Algorithm 3: multistep implementation of the extended version of Algorithm 1;
(ii) Algorithm 4: multistep implementation of the extended version of Algorithm 2.

We note that all the functions involved in these extended algorithms are discrete. Assuming that all the extension operations $\overline{[\cdot]}$ in these algorithms are handled via LERP, the corresponding time complexities are as follows

(i) Algorithm 3: $\widetilde{\mathcal{O}}\big(X(U+V) + TX(W+Y)\big)$ – assuming all the grids $\mathbb{V}_{\mathrm{g}}(x)$ are of size V;
(ii) Algorithm 4: $\widetilde{\mathcal{O}}\big(U + V + T(XW + Y + Z)\big)$.

Algorithm 3 Multistep implementation of the extended d-CDP Algorithm 1

```
Input: dynamics $f_{\mathrm{s}} : \mathbb{R}^n \to \mathbb{R}^n,\ f_{\mathrm{i}} : \mathbb{R}^n \to \mathbb{R}^{n \times m}$;
    discrete stage cost $C(x, \cdot) : \mathbb{U}_{\mathrm{g}} \to \overline{\mathbb{R}}$ for $x \in \mathbb{X}_{\mathrm{g}}$;
    discrete terminal cost $C_T : \mathbb{X}_{\mathrm{g}} \to \mathbb{R}$;
    discrete disturbance $\mathbb{W}_{\mathrm{d}}$ and its p.m.f. $p : \mathbb{W}_{\mathrm{d}} \to [0, 1]$.
Output: discrete cost-to-go $J_t : \mathbb{X}_{\mathrm{g}} \to \mathbb{R},\ t = 0, 1, \dots, T$.
 1: for each $x \in \mathbb{X}_{\mathrm{d}}$ do
 2:   construct the grid $\mathbb{V}_{\mathrm{g}}(x)$;
 3:   use LLT to compute $C_x^{\mathrm{d}*} : \mathbb{V}_{\mathrm{g}}(x) \to \mathbb{R}$ from $C(x, \cdot) : \mathbb{U}_{\mathrm{g}} \to \overline{\mathbb{R}}$;
 4: end for
 5: $J_T(x) \leftarrow g_T(x)$ for $x \in \mathbb{X}_{\mathrm{g}}$;
 6: for $t = T, \dots, 1$ do
 7:   $J_t^{\mathrm{w}}(x) \leftarrow \sum_{w \in \mathbb{W}_{\mathrm{d}}} p(w) \cdot \overline{J}(x + w)$ for $x \in \mathbb{X}_{\mathrm{g}}$;
 8:   construct the grid $\mathbb{Y}_{\mathrm{g}}$;
 9:   use LLT to compute $[J_t^{\mathrm{w}}]^{\mathrm{d}*} : \mathbb{Y}_{\mathrm{g}} \to \mathbb{R}$ from $J_t^{\mathrm{w}} : \mathbb{X}_{\mathrm{g}} \to \mathbb{R}$;
10:   for each $x \in \mathbb{X}_{\mathrm{g}}$ do
11:     for each $y \in \mathbb{Y}_{\mathrm{g}}$ do
12:       use LERP to compute $\overline{C_x^{\mathrm{d}*\mathrm{d}}}(-f_{\mathrm{i}}(x)^\top y)$ from $C_x^{\mathrm{d}*} : \mathbb{V}_{\mathrm{g}}(x) \to \mathbb{R}$;
13:       $\varphi_x(y) \leftarrow \overline{C_x^{\mathrm{d}*\mathrm{d}}}(-f_{\mathrm{i}}(x)^\top y) + [J_t^{\mathrm{w}}]^{\mathrm{d}*}(y)$;
14:     end for
15:     $J_{t-1}(x) \leftarrow \max_{y \in \mathbb{Y}_{\mathrm{g}}} \{\langle f_{\mathrm{s}}(x), y \rangle - \varphi_x(y)\}$.
16:   end for
17: end for
```

For the numerical implementation of the extended algorithms, we consider the setup of Section 6.1 for Algorithm 3 and the setup of Section 6.2 for Algorithm 4. However, we now consider stochastic dynamics by introducing an i.i.d. additive disturbance belonging to the (grid-like) finite set $\mathbb{W}_{\mathrm{d}} = \{-0.1, 0, 0.1\}^2$ with a uniform p.m.f., that is, $p(w) = \frac{1}{9}$ for all $w \in \mathbb{W}_{\mathrm{d}}$. Moreover, the conjugate of the stage cost (although analytically available) is computed numerically. In this regard, we note that the dual grids $\mathbb{V}_{\mathrm{g}}(x)$ of the input space are constructed following the guidelines described in Section 4.3.2. Through these numerical simulations, we compare the performance of the d-DP and d-CDP algorithms for solving ten instances of the optimal control problem for random initial conditions, chosen uniformly from $\mathbb{X} = [-1, 1]^2$. To this end, and similar to the setup of Section 6, we report the average of the relative trajectory cost and the average of the total running time in seconds. The results of our numerical simulations are reported in Table 4. We note that for the d-DP algorithm, we are reporting the performance of the *fast approximation* of the corresponding stochastic d-DP algorithm; see Section 4.3.1 for more details.

C.2. **Echt examples**

In this section, we showcase the application of the proposed d-CDP algorithms in solving the optimal control problem for two typical systems. In particular, we use the extended versions of these algorithms for the optimal control of an SIR (Susceptible–Infected–Recovered) model for epidemics and a noisy inverted pendulum. Once again, in order to show the effectiveness of the proposed

Algorithm 4 Multistep implementation of the extended d-CDP Algorithm 2

```
Input: dynamics $f_{\mathrm{s}} : \mathbb{R}^n \to \mathbb{R}^n,\ B \in \mathbb{R}^{n \times m}$;
    discrete stage cost (state) $C_{\mathrm{s}} : \mathbb{X}_{\mathrm{g}} \to \mathbb{R}$;
    discrete stage cost (input) $C_{\mathrm{i}} : \mathbb{U}_{\mathrm{g}} \to \mathbb{R}$;
    discrete terminal cost $C_T : \mathbb{X}_{\mathrm{g}} \to \mathbb{R}$;
    discrete disturbance $\mathbb{W}_{\mathrm{d}}$ and its p.m.f. $p : \mathbb{W}_{\mathrm{d}} \to [0,1]$.
Output: discrete cost-to-go $J_t : \mathbb{X}_{\mathrm{g}} \to \mathbb{R},\ t = 0,1,\dots,T$.
 1: construct the grid $\mathbb{V}_{\mathrm{g}}$;
 2: use LLT to compute $C_{\mathrm{i}}^{\mathrm{d}*} : \mathbb{V}_{\mathrm{g}} \to \mathbb{R}$ from $C_{\mathrm{i}} : \mathbb{U}_{\mathrm{g}} \to \mathbb{R}$;
 3: construct the grid $\mathbb{Z}_{\mathrm{g}}$;
 4: $J_T(x) \leftarrow g_T(x)$ for $x \in \mathbb{X}_{\mathrm{g}}$;
 5: for $t = T,\dots,1$ do
 6:     $J_t^{\mathrm{w}}(x) \leftarrow \sum_{w \in \mathbb{W}_{\mathrm{d}}} p(w) \cdot \overline{J}(x+w)$ for $x \in \mathbb{X}_{\mathrm{g}}$;
 7:     construct the grid $\mathbb{Y}_{\mathrm{g}}$;
 8:     use LLT to compute $[J_t^{\mathrm{w}}]^{\mathrm{d}*} : \mathbb{Y}_{\mathrm{g}} \to \mathbb{R}$ from $J_t^{\mathrm{w}} : \mathbb{X}_{\mathrm{g}} \to \mathbb{R}$;
 9:     for each $y \in \mathbb{Y}_{\mathrm{g}}$ do
10:         use LERP to compute $\overline{C_{\mathrm{i}}^{\mathrm{d}*\mathrm{d}}}(-B^\top y)$ from $C_{\mathrm{i}}^{\mathrm{d}*} : \mathbb{V}_{\mathrm{g}} \to \mathbb{R}$;
11:         $\varphi(y) \leftarrow \overline{C_{\mathrm{i}}^{\mathrm{d}*\mathrm{d}}}(-B^\top y) + [J_t^{\mathrm{w}}]^{\mathrm{d}*}(y)$;
12:     end for
13:     use LLT to compute $\varphi^{\mathrm{d}*} : \mathbb{Z}_{\mathrm{g}} \to \mathbb{R}$ from $\varphi : \mathbb{Y}_{\mathrm{g}} \to \mathbb{R}$;
14:     for each $x \in \mathbb{X}_{\mathrm{d}}$ do
15:         use LERP to compute $\overline{\varphi^{\mathrm{d}*\mathrm{d}}}\big(f_{\mathrm{s}}(x)\big)$ from $\varphi^{\mathrm{d}*} : \mathbb{Z}_{\mathrm{g}} \to \mathbb{R}$;
16:         $J_{t-1}(x) \leftarrow C_{\mathrm{s}}(x) + \overline{\varphi^{\mathrm{d}*\mathrm{d}}}\big(f_{\mathrm{s}}(x)\big)$;
17:     end for
18: end for
```

TABLE 4. Comparison of the performance of the d-DP algorithm and the extended d-CDP Algorithms 3 and 4 for different grid sizes ($X, Y, U, Z, V(x) = N$): The reported numbers are the average of the relative trajectory cost (w.r.t. the trajectory cost of d-DP (π) with $N = 81^2$) (left – blue), and the average of the total running time (right – red). See the setup described in Section 6.1 for more details.

Relative trajectory cost / Running time (seconds)				
Alg. \ N	11^2	21^2	41^2	81^2
d-CDP Alg. 3	1.50 / 6.2e + 0	0.99 / 6.9e + 1	1.03 / 9.7e + 2	1.00 / 1.5e + 4
d-DP (J)	1.55 / 7.2e + 0	0.98 / 9.0e + 1	1.03 / 1.3e + 3	1.00 / 2.0e + 4
d-DP (π)	1.03 / 7.0e + 0	1.05 / 8.9e + 1	1.03 / 1.3e + 3	1 / 2.0e + 4
d-CDP Alg. 4	1.52 / 5.4e − 1	0.96 / 1.8e + 0	1.01 / 7.0e + 0	1.01 / 2.6e + 1
d-DP (J)	1.55 / 4.0e + 0	0.98 / 4.5e + 1	1.02 / 6.7e + 2	1.00 / 1.0e + 4
d-DP (π)	1.03 / 3.7e + 0	1.05 / 4.4e + 1	1.03 / 6.7e + 2	1 / 1.0e + 4

algorithms, we compare their performance with the benchmark d-DP algorithm. More importantly, through these examples, we highlight some issues that can arise in the real world application of the proposed algorithms.

C.2.1. *SIR model.* We consider the application of the extended version of the d-CDP Algorithm 1 for computing the optimal vaccination plan in a simple epidemic model. To this end, we consider an SIR system described by [15, Sec. 4]

$$\begin{cases} s_{t+1} = s_t(1-u_t) - \alpha i_t s_t(1-u_t) \\ i_{t+1} = i_t + \alpha i_t s_t(1-u_t) - \beta i_t \\ r_{t+1} = r_t + u_t s_t, \end{cases}$$

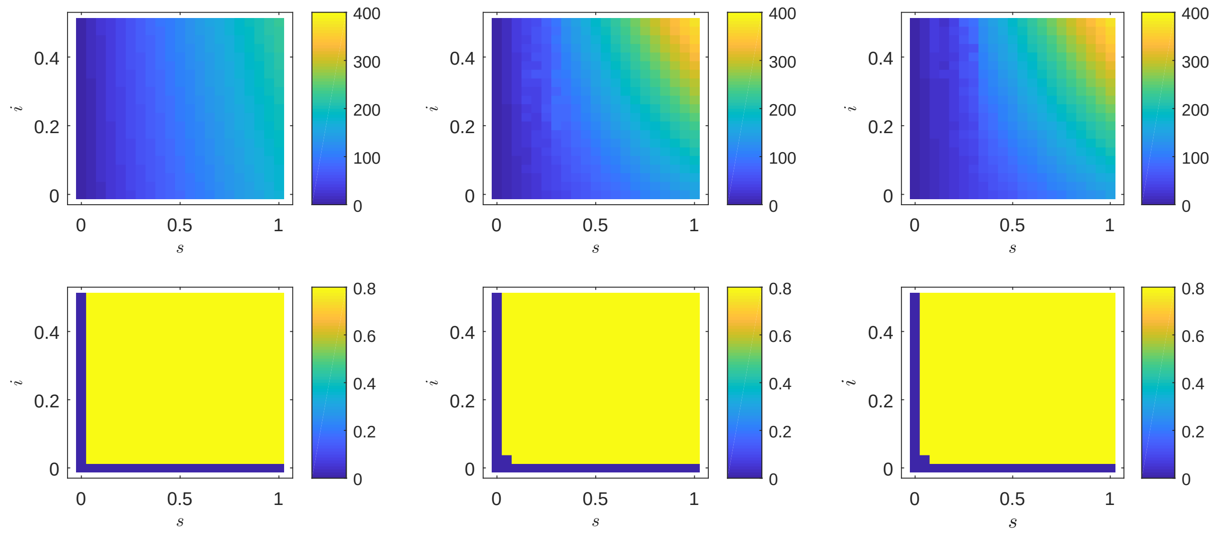
where $s_t, i_t, r_t \geq 0$ are, respectively, the (normalized) number of susceptible, infected, and immune individuals in the population, and $u_t \in [0, u_{\max}]$ is the control input which can be interpreted as the proportion of the susceptibles to be vaccinated ($u_{\max} \leq 1$). We are interested in computing the optimal vaccination policy with linear cost $\sum_{t=0}^{T-1}(\gamma i_t + u_t) + \gamma i_T$, over $T = 3$ steps ($\gamma > 0$). The model parameters are the transmission rate $\alpha = 2$, the death rate $\beta = 0.1$, the maximum vaccination capacity $u_{\max} = 0.8$, and the cost coefficient $\gamma = 100$ (corresponding to the values in [15, Sec. 4.2]).

We now provide the formulation of this problem w.r.t. Setting 1. In this regard, note that the variable r_t (number of immune individuals) can be safely ignored as it affects neither the evolution of the other two variables, nor the cost to be minimized. Hence, we can take $x_t = (s_t, i_t) \in \mathbb{R}^2$ and $u_t \in \mathbb{R}$ as the state and input variables, respectively. The dynamics of the system is then described by $x_{t+1} = f_{\mathrm{s}}(x_t) + f_{\mathrm{i}}(x_t) \cdot u_t$, where

$$f_{\mathrm{s}}(s,i) = \begin{bmatrix} s - \alpha s i \\ (1-\beta)i + \alpha s i \end{bmatrix}, \quad f_{\mathrm{i}}(s,i) = \begin{bmatrix} -s + \alpha s i \\ -\alpha s i \end{bmatrix}.$$

We consider the state constraint $x_t \in \mathbb{X} = [0,1] \times [0, 0.5]$, and the input constraint $u_t \in \mathbb{U} = [0, 0.8]$. In particular, the constraint $i_t \in [0, 0.5]$ is chosen so that the feasibility condition of Setting 1-(ii) is satisfied. Also, the corresponding stage and terminal costs read as $C(s,i,u) = \gamma i + u$, and $C_T(s,i) = i$, respectively. We note that, although the conjugate of the stage cost (C_x^*) is analytically available, we use the scheme provided in Section 4.3.2 to compute C_x^* numerically; ; see also Algorithm 3 in Appendix C.1. In order to deploy the d-DP algorithm and the extended d-CDP algorithm, we use uniform grid-like discretizations of the state and input spaces and the their dual spaces ($\mathbb{X}_{\mathrm{g}}$, $\mathbb{U}_{\mathrm{g}}$, $\mathbb{Y}_{\mathrm{g}}$, and $\mathbb{V}_{\mathrm{g}}(x)$ for $x \in \mathbb{X}_{\mathrm{g}}$). The dual grids $\mathbb{Y}_{\mathrm{g}}$ and $\mathbb{V}_{\mathrm{g}}(x)$ are constructed following the guidelines provided in Section 4.2.2 and 4.3.2, respectively.

Figure 6 depicts the computed optimal cost $J_0 : \mathbb{X}_{\mathrm{g}} \to \mathbb{R}$ and control law $\mu_0 : \mathbb{X}_{\mathrm{g}} \to \mathbb{U}_{\mathrm{g}}$, using the d-DP and d-CDP algorithms. In particular, for the d-CDP algorithm, we are reporting the simulation results for two configurations of the dual grids. Table 5 reports the corresponding grid sizes and the running times of the deployed algorithms for solving the Value Iteration Problem 3.1. In particular, notice how d-DP algorithm significantly outperforms the d-CDP algorithm with the discretization scheme of configuration 1, where $X = Y$ and $U = V$. In this regard, recall that the time complexity of d-DP algorithm is of $\mathcal{O}(TXU)$, while that of d-CDP algorithm is of $\mathcal{O}\big(X(U+V)+TXY\big) = \mathcal{O}\big(XU+TX^2\big)$, when $X = Y$ and $U = V$. Hence, what we observe is indeed expected since the number of input channels is less than the dimension of the state space. For such problems, we should be cautious when using the d-CDP algorithm, particularly, in choosing the sizes Y and V of the dual grids. For instance, for the problem at hand, as reported in Table 5, we can reduce the size of the dual grids as in configuration 2 and hence reduce the running time of the d-CDP



(a) d-DP (b) d-CDP (config. 1) (c) d-CDP (config. 2)

FIGURE 6. Optimal control of SIR model: optimal cost J_0 (top) and control law μ_0 (bottom).

TABLE 5. Optimal control of SIR model: Grid sizes and running times.

Alg.	Grid size	Running time
d-DP	$X = 21^2,\ U = 21$	2.00 sec
d-CDP (config. 1)*	$Y = 21^2,\ V = 21$	18.26 sec
d-CDP (config. 2)*	$Y = 11^2,\ V = 11$	6.19 sec

*X and U are the same as in d-DP.

algorithm. However, as shown in Figure 6, this reduction in the size of the dual grids does not affect the quality of the computed costs and hence the corresponding control laws.

C.2.2. *Inverted pendulum.* We now consider an application of the extension of the d-CDP Algorithm 2, which handles additive disturbance in the dynamics; see Algorithm 4 in Appendix C.1. To this end, we consider the optimal control of a noisy inverted pendulum with quadratic stage and terminal costs, over a finite horizon. The deterministic continuous-time dynamics of the system is described by [11, Sec. 4.5.3]

$$\ddot{\theta} = \alpha \sin\theta + \beta\dot{\theta} + \gamma u,$$

where θ is the angle (with $\theta = 0$ corresponding to upward position), and u is the control input. The values of the parameters are $\alpha = 118.6445$, $\beta = -1.599$, and $\gamma = 29.5398$ (corresponding to the values of the physical parameters in [11, Sec. 4.5.3]). Here, we consider the corresponding discrete-time dynamics, by using forward Euler method, with sampling time $\tau = 0.05$. We also introduce stochasticity by considering an additive disturbance in the dynamics. The discrete-time dynamics then reads as $x_{t+1} = f_{\mathrm{s}}(x_t) + Bu_t + w_t$, where $x_t = (\theta_t, \dot{\theta}_t) \in \mathbb{R}^2$ is the state variable (angle and angular velocity), $w_t \subset \mathbb{R}^2$ is the disturbance, and

$$f_{\mathrm{s}}(\theta,\dot{\theta}) = \begin{bmatrix} \theta \\ \dot{\theta} \end{bmatrix} + \tau \cdot \begin{bmatrix} \dot{\theta} \\ \alpha \sin\theta + \beta\dot{\theta} \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ \gamma \end{bmatrix}.$$

We consider the state constraints $x_t \in \mathbb{X} = [-\frac{\pi}{4}, \frac{\pi}{4}] \times [\pi, \pi]$, and the input constraints $u_t \in \mathbb{U} = [-3, 3]$. Moreover, we assume that the disturbances w_t are i.i.d., with a *uniform* distribution over the compact support $\mathbb{W} = \frac{\pi}{4} \cdot [-0.05, 0.05] \times \pi \cdot [-0.05, 0.05]$. The problem of interest is then to compute the costs-to-go $J^{\mathrm{d}}_t : \mathbb{X}_{\mathrm{g}} \to \overline{\mathbb{R}}$ for $t = T-1, \dots, 0$, over the horizon $T = 50$, with quadratic costs $C_\nu(\cdot) = \|\cdot\|^2, \ \nu \in \{\mathrm{s}, \mathrm{i}, T\}$. We recall that $\mathbb{X}_{\mathrm{g}}$ is a grid-like discretization of the state space $\mathbb{X}$. Also, we note that the conjugate of the input-dependent stage cost C^*_{i} is analytically available, and given by $C^*_{\mathrm{i}}(v) = \hat{u}v - \hat{u}^2, \ v \in \mathbb{R}$, where $\hat{u} = \max\left\{-3, \ \min\left\{\frac{v}{2}, \ 3\right\}\right\}$.

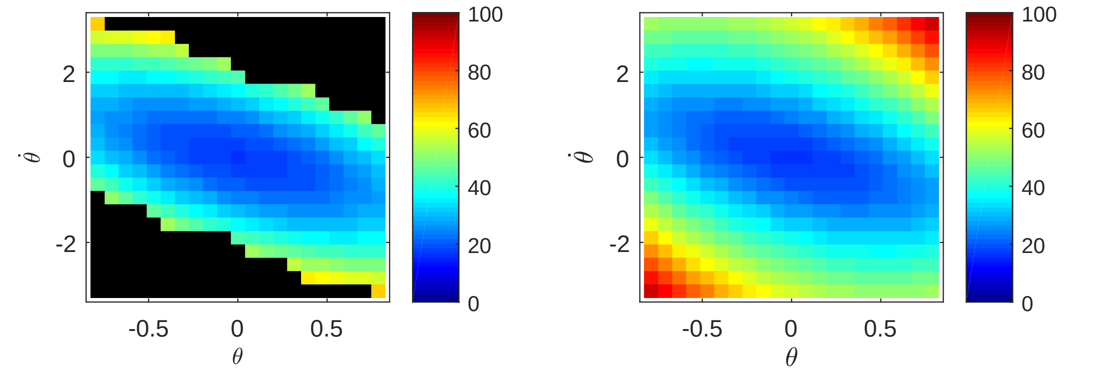
The extension of the d-CDP algorithm for handling additive disturbance involves applying the d-CDP operation to $J^{\mathrm{w}}(x) := \mathbb{E}_w \overline{J^{\mathrm{d}}}(x+w)$, where $\mathbb{E}$ is the expectation operator, and $\overline{[\cdot]}$ is an extension operator. For the extension operation, we simply use LERP. More importantly, for the expectation operation, we consider the approximation scheme described in Section 4.3, involving discretization of the disturbance set. Precisely, we assume that $w_t \in \mathbb{W}_{\mathrm{d}} = \mathbb{W}_{\mathrm{d}1} \times \mathbb{W}_{\mathrm{d}2} \subset \mathbb{W}$ with a uniform probability mass function, where

$$\mathbb{W}_{\mathrm{d}1} = \frac{\pi}{4} \cdot \{-0.05, -0.025, 0, 0.025, 0.05\}, \ \mathbb{W}_{\mathrm{d}2} = \pi \cdot \{-0.05, -0.025, 0, 0.025, 0.05\}.$$

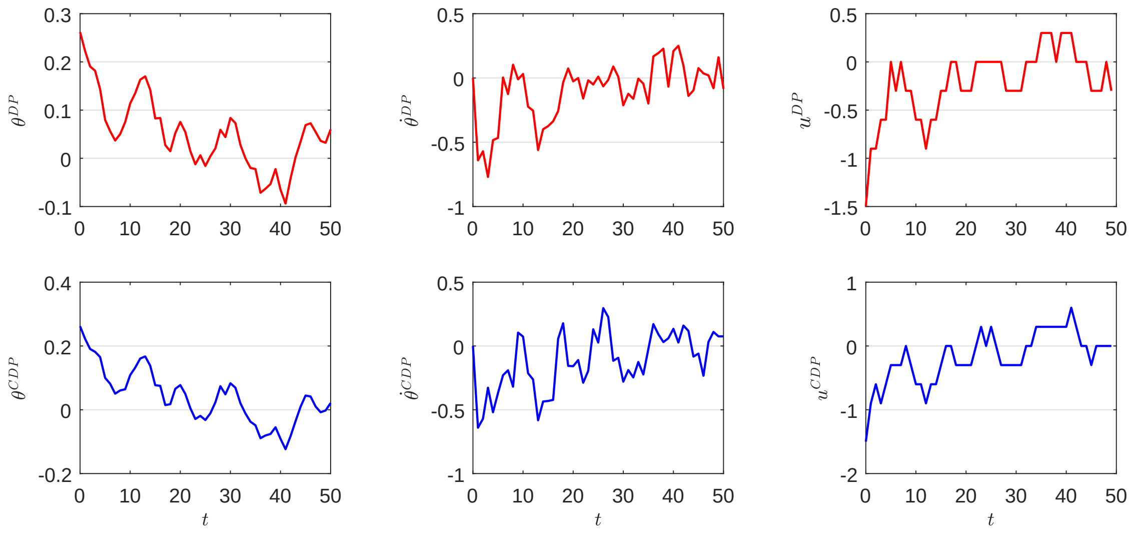
Under such assumption, we have $J^{\mathrm{w}}(x) = \frac{1}{W} \sum_{w \in \mathbb{W}_{\mathrm{d}}} \overline{J^{\mathrm{d}}}(x+w), \ x \in \mathbb{X}_{\mathrm{g}}$.

In order to deploy the (extended) d-DP and d-CDP algorithms for the optimal control problem described above, we use uniform discretizations of the state, input, and state dual spaces, with $N_i = 21$ discrete points in each dimension, i.e., $X = Y = 21^2$ and $U = 21$. The grid $\mathbb{Z}_{\mathrm{g}}$ is also constructed with the same size ($Z = 21^2$). For the construction of the grids $\mathbb{Y}_{\mathrm{g}}$ and $\mathbb{Z}_{\mathrm{g}}$, we follow the guidelines provided in Section 5.2. We also note that for the d-DP algorithm, we are reporting the performance of the *fast approximation* of the corresponding stochastic d-DP algorithm; see Section 4.3 for more details.

Figure 7a shows the computed cost-to-go at $t = 0$, using the described extensions of the d-DP and d-CDP algorithms. We note that the optimal control problem at hand does not satisfy the feasibility condition assumed in this study. That is, there exist $x \in \mathbb{X}_{\mathrm{g}}$ for which there is no $u \in \mathbb{U}_{\mathrm{g}}$ such that $x^+ = f(x) + Bu \in \mathbb{X}_{\mathrm{g}}$. This explains the black areas in the left panel of Figure 7a with $J_0 = \infty$, computed using d-DP algorithm. Notice, however, that in the right panel of Figure 7a, the d-CDP algorithm assigns finite values for these states. This does not contradict our error analysis, as the assumption on the optimal control problem to be feasible for all $x \in \mathbb{X}$ is violated. Indeed, for feasible initial states in the state space, our theoretical results still hold true. We also note that *the running times of the two algorithms for solving the Value Iteration Problem 3.1 were approximately* 22 *seconds for the d-DP algorithm, and* 9 *seconds for the d-CDP algorithm.* As a further illustration, Figure 7b depicts a sample state trajectory of the system, where the control input sequence is derived via minimization of the costs-to-go computed using the d-DP and d-CDP algorithms.



(a) Cost-to-go at $t = 0$: d-DP (left) and d-CDP (right). The black areas correspond to $J_0 = \infty$.



(b) The state trajectory and input sequence for $x_0 = (\frac{\pi}{12}, 0)^\top$ using d-DP (top) and d-CDP (bottom).

FIGURE 7. Optimal control of noisy inverted pendulum.

Appendix D. The d-CDP MATLAB package

The MATLAB package [21] concerns the implementation of the two d-CDP algorithms developed in this study. The provided codes include detailed instructions/comments on how to use them. Also provided is the implementation of the numerical examples of Section 6 and Appendix C.1. In what follows we highlight the most important aspects of the developed package with a list of available routines.

Recall that, in this study, we exclusively considered *grid-like* discretizations of both primal and dual domains. This allows us to use the MATALB function `griddedInterpolant` for all the extension operations. We also note that the interpolation and extrapolation methods of this fucntion are all set to `linear`, hence leading to multilinear interpolation & extrapolation (LERP). However, this need not be the case in general, and the user can choose other options available in the `griddedInterpolant` routine, by modifying the corresponding parts of the provided codes; see the comments in the codes for more details. We also note that for the for the discrete conjugation (LLT), we used the MATLAB package (the `LLTd` routine and two other subroutines, specifically) provided in [23] to develop an n-dimensional LLT routine via factorization (the function `LLT` in the package). Table 6 lists other routines that are available in the developed package. In particular, there are four high level functions (functions (1-4) in Table 6) that are developed separately for the two settings considered in this article. We also note that the provided implementations do not require the discretization of the state and input spaces to satisfy the state and input constraints (particularly, the feasibility condition of Setting 1-(ii)). Nevertheless, the function `feasibility_check_`$*$ ($* = 1, 2$) is developed to provide the user with a warning if that case. Finally, we note that the conjugate of four extended real-valued convex functions are also provided in the package as listed in Table 7.

Table 6. List of routines available in the d-CDP MATLAB package.

MATLAB Function	Description
(1) `d_CDP_Alg_`$*$	Backward value iteration for finding optimal costs using d-CDP
(2) `d_DP_Alg_`$*$	Backward value iteration for finding optimal costs and control laws using d-DP
(3) `forward_iter_J_`$*$	Forward iteration for finding optimal control sequence for a given initial condition using optimal costs (derived via d-DP or d-CDP)
(4) `forward_iter_Pi_`$*$	Forward iteration for finding optimal control sequence for a given initial condition using optimal control laws (derived via d-DP)
(5) `feasibility_check_`$*$	For checking if the discrete state-input space satisfy the constraints
(6) `eval_func`	For discretization of an analytically available function over a given grid
(7) `eval_func_constr`	An extension of `eval_func` that also checks given constraints
(8) `ext_constr`	For extension of a discrete function while checking a given set of constraints
(9) `ext_constr_expect`	For computing expectation of a discrete function subjected to additive noise
(10) `slope_range`	For computing the range of slopes of a convex-extensible discrete function with a grid-like domain

$* = 1, 2$, corresponding to Settings 1 and 2, respectively.

Table 7. List of analytically available conjugate functions in the d-CDP MATLAB package.

Function	Effective Domain	MATLAB Func. for Conj.
$g : \mathbb{R}^n \to \overline{\mathbb{R}} : u \mapsto u^\top Ru \ (R \succ 0)$	Ball centered at the origin	`conj_Quad_ball`
$g : \mathbb{R}^n \to \overline{\mathbb{R}} : u \mapsto u^\top Ru \ (R \succ 0)$	Box containing the origin	`conj_Quad_box`
$g : \mathbb{R}^n \to \overline{\mathbb{R}} : u \mapsto \sum_{i=1}^{n} \lvert u_i \rvert$	Box containing the origin	`conj_L1_box`
$g : \mathbb{R}^n \to \overline{\mathbb{R}} : u \mapsto \sum_{i=1}^{n} e^{\lvert u_i \rvert} - n$	Box containing the origin	`conj_ExpL1_box`

References

[1] Achdou, Y., Camilli, F., and Corrias, L. (2014). On numerical approximation of the Hamilton-Jacobi-transport system arising in high frequency approximations. *Discrete & Continuous Dynamical Systems-Series B*, 19(3).

[2] Akian, M., Gaubert, S., and Lakhoua, A. (2008). The max-plus finite element method for solving deterministic optimal control problems: Basic properties and convergence analysis. *SIAM Journal on Control and Optimization*, 47(2):817–848.

[3] Bach, F. (2019). Max-plus matching pursuit for deterministic Markov decision processes. *arXiv preprint arXiv:1906.08524*.

[4] Balaji, N., Kiefer, S., Novotnỳ, P., Pérez, G. A., and Shirmohammadi, M. (2018). On the complexity of value iteration. *preprint arXiv:1807.04920*.

[5] Bellman, R. and Karush, W. (1962). Mathematical programming and the maximum transform. *Journal of the Society for Industrial and Applied Mathematics*, 10(3):550–567.

[6] Berthier, E. and Bach, F. (2020). Max-plus linear approximations for deterministic continuous-state markov decision processes. *IEEE Control Systems Letters*, pages 1–1.

[7] Bertsekas, D. P. (2005). *Dynamic Programming and Optimal Control, Vol. I.* Athena Scientific, Belmont, MA, 3rd edition.

[8] Bertsekas, D. P. (2009). *Convex Optimization Theory.* Athena Scientific, Belmont, MA.

[9] Bertsekas, D. P. (2019). *Reinforcement Learning and Optimal Control.* Athena Scientific, Belmont, MA.

[10] Boyd, S. and Vandenberghe, L. (2004). *Convex Optimization.* Cambridge University Press.

[11] Busoniu, L., Babuska, R., De Schutter, B., and Ernst, D. (2017). *Reinforcement learning and dynamic programming using function approximators.* CRC press.

[12] Carpio, R. and Kamihigashi, T. (2020). Fast value iteration: an application of Legendre-Fenchel duality to a class of deterministic dynamic programming problems in discrete time. *Journal of Difference Equations and Applications*, 26(2):209–222.

[13] Corrias, L. (1996). Fast Legendre-Fenchel transform and applications to Hamilton-Jacobi equations and conservation laws. *SIAM Journal on Numerical Analysis*, 33(4):1534–1558.

[14] Costeseque, G. and Lebacque, J.-P. (2014). A variational formulation for higher order macroscopic traffic flow models: Numerical investigation. *Transportation Research Part B: Methodological*, 70:112 – 133.

[15] Ding, W. and Lenhart, S. (2010). Introduction to optimal control for discrete time models with an application to disease modeling. In Gumel, A. B. and Lenhart, S., editors, *Modeling Paradigms and Analysis of Disease Transmission Models*, pages 109–120. American Mathematical Society.

[16] Esogbue, A. O. and Ahn, C. W. (1990). Computational experiments with a class of dynamic programming algorithms of higher dimensions. *Computers & Mathematics with Applications*, 19(11):3 – 23.

[17] Felzenszwalb, P. F. and Huttenlocher, D. P. (2012). Distance transforms of sampled functions. *Theory of computing*, 8(1):415–428.

[18] Jacobs, M. and Léger, F. (2019). A fast approach to optimal transport: the back-and-forth method. *arXiv preprint arXiv:1905.12154*.

[19] Joó, I. and Stachó, L. L. (1982). A note on Ky Fan's minimax theorem. *Acta Mathematica Academiae Scientiarum Hungarica*, 39(4):401–407.

[20] Klein, C. M. and Morin, T. L. (1991). Conjugate duality and the curse of dimensionality. *European Journal of Operational Research*, 50(2):220 – 228.

[21] Kolarijani, M. A. S. and Mohajerin Esfahani, P. (2020). Discrete conjugate dynamic programming (d-CDP) MATLAB package. Available online at https://github.com/AminKolarijani/d-CDP.

[22] Lucet, Y. (1996). A fast computational algorithm for the Legendre-Fenchel transform. *Computational Optimization and Applications*, 6(1):27–57.

[23] Lucet, Y. (1997). Faster than the fast Legendre transform, the linear-time Legendre transform. *Numerical Algorithms*, 16(2):171–185.

[24] Lucet, Y. (2009). New sequential exact Euclidean distance transform algorithms based on convex analysis. *Image and Vision Computing*, 27(1):37 – 44.

[25] Lucet, Y. (2010). What shape is your conjugate? A survey of computational convex analysis and its applications. *SIAM Review*, 52(3):505–542.

[26] McEneaney, W. M. (2003). Max-plus eigenvector representations for solution of nonlinear H_∞ problems: basic concepts. *IEEE Transactions on Automatic Control*, 48(7):1150–1163.

[27] Murota, K. (2003). *Discrete Convex Analysis.* Society for Industrial and Applied Mathematics.

[28] Powell, W. B. (2011). *Approximate Dynamic Programming: Solving the Curses of Dimensionality.* John Wiley & Sons, Hoboken, NJ, 2nd edition.

[29] Rockafellar, R. (1974). *Conjugate Duality and Optimization.* Philadelphia: Society for Industrial and Applied Mathematics.

[30] Ronagh, P. (2019). Quantum algorithms for solving dynamic programming problems. *preprint arXiv:1906.02229.*

[31] Sidford, A., Wang, M., Wu, X., and Ye, Y. (2018). Variance reduced value iteration and faster algorithms for solving Markov decision processes. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 770–787. SIAM.

[32] Simons, S. (1995). Minimax theorems and their proofs. In Du, D.-Z. and Pardalos, P. M., editors, *Minimax and Applications*, pages 1–23. Springer US, Boston, MA.

[33] Sutter, D., Nannicini, G., Sutter, T., and Woerner, S. (2020). Quantum Legendre-Fenchel transform. *preprint arXiv:2006.04823.*