

EXPONENTIAL DECAY IN THE SENSITIVITY ANALYSIS OF
NONLINEAR DYNAMIC PROGRAMMING*SEN NA[†] AND MIHAI ANITESCU[‡]

Abstract. In this paper, we study the sensitivity of discrete-time dynamic programs with nonlinear dynamics and objective to perturbations in the initial conditions and reference parameters. Under uniform controllability and boundedness assumptions for the problem data, we prove that the directional derivative of the optimal state and control at time k , \mathbf{x}_k^* , and \mathbf{u}_k^* , with respect to the reference signal at time i , \mathbf{d}_i , will have exponential decay in terms of $|k - i|$ with a decay rate ρ independent of the temporal horizon length. The key technical step is to prove that a version of the convexification approach proposed by Verschueren et al. can be applied to the KKT conditions and results in a convex quadratic program with uniformly bounded data. In turn, Riccati techniques can be further employed to obtain the sensitivity result, borne from the observation that the directional derivatives are solutions of quadratic programs with structure similar to the KKT conditions themselves. We validate our findings with numerical experiments on a small nonlinear, nonconvex, dynamic program.

Key words. sensitivity analysis, nonlinear dynamic programming, optimization

AMS subject classifications. 93B05, 90C39

DOI. 10.1137/19M1265065

1. Introduction. We consider the following discrete-time nonlinear dynamic programming (NLDP) problem with recursive equality constraints:

$$(1.1a) \quad \min_{\mathbf{x}, \mathbf{u}} \sum_{k=0}^{N-1} g_k(\mathbf{x}_k, \mathbf{u}_k, \mathbf{d}_k) + g_N(\mathbf{x}_N),$$

$$(1.1b) \quad \text{s.t. } \mathbf{x}_{k+1} = f_k(\mathbf{x}_k, \mathbf{u}_k, \mathbf{d}_k), \quad k = 0, 1, \dots, N-1,$$

$$(1.1c) \quad \mathbf{x}_0 = \bar{\mathbf{x}}_0.$$

Here $\mathbf{x}_k \in \mathbb{R}^{n_x}$ is the state variable, $\mathbf{u}_k \in \mathbb{R}^{n_u}$ is the control variable, and $\mathbf{d}_k \in \mathbb{R}^{n_d}$ is a reference variable which is externally provided data; $g_k : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_d} \rightarrow \mathbb{R}$, $f_k : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_d} \rightarrow \mathbb{R}^{n_x}$ are the cost function and constraint function for stage k , respectively; N is the temporal horizon length; and $\bar{\mathbf{x}}_0$ is the initial state variable. We also define $\mathbf{d}_{-1} = \bar{\mathbf{x}}_0$ for simplicity. In this paper, we let $\mathbf{d} = (\mathbf{d}_{-1}; \mathbf{d}_0; \dots; \mathbf{d}_{N-1})$ and assume g_k, f_k are both twice continuously differentiable.

Problem (1.1) includes a wide range of classical dynamic programs (DPs) as special cases. In our context, we call it quadratic dynamic programming (QDP) when-

*Received by the editors May 29, 2019; accepted for publication (in revised form) November 27, 2019; published electronically June 8, 2020. Preprint ANL/MCS-P9182-0520, Mathematics and Computer Science Division, Argonne National Laboratory. This work was performed by an employee of the U.S. Government or under U.S. Government contract. The U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U.S. Government purposes. Copyright is owned by SIAM to the extent not limited by these rights.

<https://doi.org/10.1137/19M1265065>

Funding: This material was based upon work supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research (ASCR), under contract DE-AC02-06CH11347, and by the NSF through award CNS-1545046.

[†]Department of Statistics, The University of Chicago, Chicago, IL 60637 (senna@uchicago.edu).

[‡]Mathematics and Computer Science Division, Argonne National Laboratory, Lemont, IL 60439 (anitescu@mcs.anl.gov).

ever, for any stage k , g_k is quadratic and f_k is affine. Since the reference vector is a parameter to the problem, we treat it as the input and aim to (locally) analyze the dependence of the optimal state and control variable on it. Specifically, we are interested in answering the following question:

For any two stages k and i , how much do optimal \mathbf{x}_k and \mathbf{u}_k change when we perturb \mathbf{d}_i ?

We can formulate this question using directional derivatives. Suppose we perturb \mathbf{d} along a parameterized path:

$$(1.2) \quad \mathbf{d}(\epsilon) = \mathbf{d}^0 + \epsilon \mathbf{l} + o(\epsilon),$$

where ϵ is the parameter of the path, \mathbf{d}^0 is the unperturbed variable, and the unit vector \mathbf{l} indicates the perturbation direction. Based on the path (1.2), we can define the following two directional derivatives (if the limits exist):

$$(1.3) \quad \mathbf{p}_k = \lim_{\epsilon \searrow 0} \frac{\mathbf{x}_k^*(\mathbf{d}(\epsilon)) - \mathbf{x}_k^*(\mathbf{d}(0))}{\epsilon}, \quad \mathbf{q}_k = \lim_{\epsilon \searrow 0} \frac{\mathbf{u}_k^*(\mathbf{d}(\epsilon)) - \mathbf{u}_k^*(\mathbf{d}(0))}{\epsilon},$$

where $\mathbf{x}_k^*(\mathbf{d}), \mathbf{u}_k^*(\mathbf{d})$ are the optimal solution of Problem (1.1) with specified input \mathbf{d} . For conciseness, we ignore the dependence of \mathbf{l} in $\mathbf{p}_k, \mathbf{q}_k$ in (1.3). Suppose $\mathbf{e}_q, q = 1, 2, \dots, Nn_d + n_x$, to be the q th canonical basis vector in $\mathbb{R}^{n_x + Nn_d}$. Then, if we let $\mathbf{l} = \mathbf{e}_{n_x + i \cdot n_d + j}$, $0 \leq i \leq N - 1$, $1 \leq j \leq n_d$, our perturbation occurs only at the temporal stage i , in entry j , and, particularly, $\{\mathbf{e}_j\}_{j=1}^{n_x}$ corresponds to the perturbation at the initial constraints. According to (1.3) we locally have $\|\mathbf{x}_k(\mathbf{d}(\epsilon)) - \mathbf{x}_k(\mathbf{d}(0))\| \approx \epsilon \|\mathbf{p}_k\|$ (the same for \mathbf{u}_k). With this perturbation structure, analyzing the magnitude of \mathbf{p}_k and \mathbf{q}_k measures the effect on stage k caused by the perturbations at stage i .

Some results have been proposed before for this type of local sensitivity analysis. For example, when $g_k(\mathbf{x}_k, \mathbf{u}_k, \mathbf{d}_k) = (\mathbf{x}_k - \mathbf{d}_k)^T Q_k (\mathbf{x}_k - \mathbf{d}_k) + \mathbf{u}_k^T R_k \mathbf{u}_k$ with $Q_k, R_k \succ 0$, and f_k is an affine function depending only on $(\mathbf{x}_k, \mathbf{u}_k)$, Xu and Anitescu [20] showed that the magnitude of \mathbf{p}_k and \mathbf{q}_k for this special case of QDP decays exponentially in $|k - i|$. This observation was crucial for designing a temporal decomposition approach for long-horizon linear-quadratic dynamic programming. A similar property, asymptotic decay, was assumed for the nonlinear case in [15] and also resulted in the ability to decompose long-horizon nonlinear dynamic programs and enable parallel computation. It also played a critical role in the analysis of the convergence of model predictive control and, in particular, characterizing its operational horizon [19]. The main objective of this research note is to extend these results to the case of nonconvex QDP and, more generally, nonconvex NLDP.

Dealing with indefinite quadratic matrices in QPs is an important issue in general NLP. Many high-performance solvers are built assuming the positive definiteness of the Hessian matrix, such as the QP solvers HPMPC [7] and qpOASES [6] and the NLP solver FORCES [4]. The usual way to accommodate the indefiniteness is to perturb it to a positive matrix, to make sure the calculated step is indeed a descent direction. For example, Levenberg and Marquardt added a multiple of the identity matrix to the Hessian (see [10, 8]) and showed that under some conditions it can achieve superlinear convergence for unconstrained NLP. Nocedal and Wright [11] described several methods that directly modify the factors of the Cholesky factorization or Bunch–Parlett–Kauffman factorization for indefinite Hessians. For the constrained case, regularization approaches targeting the KKT matrix have been proposed and applied in interior point methods such as IPOPT [18] and its variants [21, 2, 17]. Recently, targeting QDPs stemming from NLDP, Verschueren et al. proposed transforming the indefinite Hessian to a positive definite matrix, at a point where the

second-order conditions hold, while *keeping the reduced Hessian unchanged* [16]. We call this approach a *convexification procedure*. In [16] the researchers showed that this convexification can be conducted independently of the solvers, preserve the block sparsity structure of the problem, and have linear computational complexity with respect to the horizon length.

The convexification procedure from [16] plays an important role in this work, although for our ends we will need to enhance it in several directions. First, it depends on a certain regularization parameter whose selection is unclear because the analysis in [16] is centered on proving only its existence using a continuity argument. Second, although not crucial for convexification itself, that analysis does not include linear terms in the objective. In the present paper, we enhance their convexification procedure by addressing these limitations. In particular, we modify it to allow for linear terms in the objective, and we connect the range of the regularization parameter to quantities characterizing the optimality of the primal vector. In turn these enhancements allow us to prove that, when the system is controllable, the convexification results in the transformed data, including the linear terms, being bounded independent of the horizon. As a result, as was also the goal in [16], we can now apply a Riccati recursion to solve the QDP. Moreover, this recursion is exponentially stable, and we can then extend some of the results in [20] to prove the exponential decay of the sensitivity in the NLDP case, which is far more general than the QDP case studied in [20].

Notation. Throughout the paper, we use bold symbols to denote column vectors. We use I to denote the identity matrix whose dimension is always clear from the context. Given an integer k , we define $[k] = \{0, 1, \dots, k\}$ to be all integer indices from 0 to k . We let $i \wedge j = \min(i, j)$ and $i \vee j = \max(i, j)$ when i, j are two scalars. For any vectors \mathbf{a} and \mathbf{b} , we use $(\mathbf{a}; \mathbf{b})$ to denote a long column vector that stacks them together. We also have $\text{supp}(\mathbf{a})$ to be the support set of \mathbf{a} . Without the subscript, $\mathbf{x}, \mathbf{u}, \mathbf{d}, \mathbf{p}$, and \mathbf{q} denote the corresponding variable for the entire horizon. We also define $\mathbf{z} = (\mathbf{x}_0; \mathbf{u}_0; \dots; \mathbf{x}_{N-1}; \mathbf{u}_{N-1}; \mathbf{x}_N)$ to be the whole decision variable ordered by stages. But we may also interpret $\mathbf{z} = (\mathbf{x}, \mathbf{u})$ when connecting it to the state \mathbf{x} and control \mathbf{u} vectors. Let $n_z = (N+1)n_x + Nn_u$ be the dimension of \mathbf{z} . Similarly, we define $\mathbf{w} = (\mathbf{p}_0; \mathbf{q}_0; \dots; \mathbf{p}_{N-1}; \mathbf{q}_{N-1}; \mathbf{p}_N) \in \mathbb{R}^{n_z}$. Given a sequence of matrices A_i , we let $\text{diag}(A_1, A_2, \dots)$ be the block diagonal matrix, whose blocks from upper left to bottom right are specified by A_1, A_2, \dots . We also define $\prod_{i=m}^n A_i = A_n A_{n-1} \dots A_m$ if $n \geq m$ and I otherwise. Similarly, we let $\sum_{i=m}^n A_i = 0$ if $n < m$. For a positive definite matrix A and any vector \mathbf{x} , we define $\|\mathbf{x}\|_A = \sqrt{\mathbf{x}^T A \mathbf{x}}$. We also use $\|\cdot\|$ to denote the ℓ_2 norm for a vector or operator norm for a matrix. For the dynamics function $f_k(\mathbf{x}_k, \mathbf{u}_k, \mathbf{d}_k)$, we let $\nabla_{\mathbf{u}_k} f_k = \frac{\partial f_k}{\partial \mathbf{u}_k} \in \mathbb{R}^{n_u \times n_x}$.

Structure of the paper. In section 2, we introduce some preliminaries and derive the indefinite QP whose solution is the directional derivatives (\mathbf{p}, \mathbf{q}) . In section 3, we show the details of the convexification procedure and propose a method for setting the regularization parameter. In section 4, we probe the properties of the outputs of the algorithm, based on which in section 5 we then theoretically prove the exponential decay. In section 6 we describe our numerical simulations. Conclusions plus some potential future work are discussed in section 7.

2. Preliminaries.

Define the Lagrange function of Problem (1.1) as

$$(2.1) \quad \mathcal{L}(\mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}; \mathbf{d}) = \sum_{k=0}^{N-1} g_k(\mathbf{x}_k, \mathbf{u}_k, \mathbf{d}_k) + g_N(\mathbf{x}_N) \\ + \sum_{k=0}^{N-1} \boldsymbol{\lambda}_k^T (\mathbf{x}_{k+1} - f_k(\mathbf{x}_k, \mathbf{u}_k, \mathbf{d}_k)) + \boldsymbol{\lambda}_{-1}^T (\mathbf{x}_0 - \mathbf{d}_{-1}),$$

where $\boldsymbol{\lambda} = (\boldsymbol{\lambda}_{-1}; \boldsymbol{\lambda}_0; \dots; \boldsymbol{\lambda}_{N-1})$ is the Lagrange multiplier with $\boldsymbol{\lambda}_k \in \mathbb{R}^{n_x}$, $k \in \{-1\} \cup [N-1]$. An immediate but important observation we justify below is that any feasible point of Problem (1.1) satisfies the linear independence constraint qualification (LICQ). We now define several quantities of interest.

DEFINITION 2.1. *We introduce the notation $\tilde{f}_k(\mathbf{z}, \mathbf{d}) = \mathbf{x}_{k+1} - f_k(\mathbf{x}_k, \mathbf{u}_k, \mathbf{d}_k) \forall k \in [N-1]$ and $\tilde{f}_{-1}(\mathbf{z}, \mathbf{d}) = \mathbf{x}_0 - \mathbf{d}_{-1}$. For any prespecified \mathbf{d} , we define the Jacobian matrix as follows:*

$$G(\mathbf{z}; \mathbf{d}) = (\nabla_{\mathbf{z}} \tilde{f}_{-1}(\mathbf{z}, \mathbf{d}), \nabla_{\mathbf{z}} \tilde{f}_0(\mathbf{z}, \mathbf{d}), \dots, \nabla_{\mathbf{z}} \tilde{f}_{N-1}(\mathbf{z}, \mathbf{d}))^T \in \mathbb{R}^{(N+1)n_x \times n_z}.$$

(Note that $G(\mathbf{z}; \mathbf{d})$ has full row rank because $\frac{\partial \tilde{f}_k}{\partial x_{k+1}} = I_{n_x} \forall -1 \leq k \leq N-1$ and $\frac{\partial \tilde{f}_k}{\partial x_j} = 0 \forall -1 \leq k < j-1 \leq N-1$; this “staircase” structure ending with identity matrices on each row implies the full row rank property of $G(\mathbf{z}; \mathbf{d})$ and, thus, LICQ holds.) When \mathbf{z} is a local solution of Problem (1.1), the critical cone¹ at \mathbf{z} is the null space of $G(\mathbf{z}; \mathbf{d})$, denoted as $\mathcal{C}(\mathbf{z}; \mathbf{d})$. We use $Z(\mathbf{z}; \mathbf{d}) \in \mathbb{R}^{n_z \times Nn_u}$ to denote a full column rank matrix whose columns are orthonormal and span the space $\mathcal{C}(\mathbf{z}; \mathbf{d})$. We denote the Hessian matrix of Lagrangian (2.1) by $\forall k \in [N-1]$

$$H_k = \begin{pmatrix} Q_k & S_k^T \\ S_k & R_k \end{pmatrix} = \begin{pmatrix} \nabla_{\mathbf{x}_k}^2 \mathcal{L} & \nabla_{\mathbf{x}_k \mathbf{u}_k}^2 \mathcal{L} \\ \nabla_{\mathbf{u}_k \mathbf{x}_k}^2 \mathcal{L} & \nabla_{\mathbf{u}_k}^2 \mathcal{L} \end{pmatrix}, \quad D_k = (D_{k1}, D_{k2}) = (\nabla_{\mathbf{d}_k \mathbf{x}_k}^2 \mathcal{L}, \nabla_{\mathbf{d}_k \mathbf{u}_k}^2 \mathcal{L}),$$

and $H_N = Q_N = \nabla_{\mathbf{x}_N}^2 g_N(\mathbf{x}_N)$. We let

$$H = \text{diag}(H_0, \dots, H_N), \quad D = \text{diag}(D_0, \dots, D_{N-1}).$$

We now set the stage for the sensitivity analysis. Suppose the unperturbed reference variable is \mathbf{d}^0 and the unperturbed primal-dual optimal solution at $\mathbf{d} = \mathbf{d}^0$ is $(\mathbf{z}^0, \boldsymbol{\lambda}^0; \mathbf{d}^0) = (\mathbf{x}^0, \mathbf{u}^0, \boldsymbol{\lambda}^0; \mathbf{d}^0)$. In what follows, all matrices from Definition 2.1 are evaluated at this primal-dual solution, and their dependence on the primal-dual point is suppressed. We require the following second-order sufficient condition (SOSC).

Assumption 2.2 (SOSC). At $(\mathbf{x}^0, \mathbf{u}^0, \boldsymbol{\lambda}^0; \mathbf{d}^0)$, the reduced Hessian of Problem (1.1) is positive definite. That is,

$$Z^T H Z \succ 0.$$

We mention that Assumption 2.2 is standard in sensitivity analysis [3]. It implies that $\mathbf{z}^0 = (\mathbf{x}^0, \mathbf{u}^0)$ is a strict local solution and guarantees that directional derivatives of the solution with respect to the parameter \mathbf{d} are well defined. Under SOSC, we have the following sensitivity analysis theorem.

¹The detailed definition is in [11]. See (12.53).

THEOREM 2.3 (sensitivity analysis of Problem (1.1)). *Suppose we perturb \mathbf{d} along the path in (1.2) for Problem (1.1). Then, under Assumption 2.2, the directional derivative $\mathbf{w} = (\mathbf{p}, \mathbf{q})$ of the solution $(\mathbf{x}^*(\mathbf{d}(\epsilon)), \mathbf{u}^*(\mathbf{d}(\epsilon)))$ for positive ϵ at 0, defined in (1.3), exists and is the solution of the following quadratic program:*

$$(2.2a) \quad QP : \min_{\mathbf{w}} \mathbf{w}^T H \mathbf{w} + 2\mathbf{l}^T D \mathbf{w},$$

$$(2.2b) \quad \text{s.t. } G \mathbf{w} = \mathbf{y}.$$

Here, we denote $\mathbf{y} = (\mathbf{y}_{-1}; \mathbf{y}_0; \dots; \mathbf{y}_{N-1}) \in \mathbb{R}^{(N+1)n_x}$ with $\mathbf{y}_{-1} = \mathbf{l}_{-1}$ and $\mathbf{y}_k = (\nabla_{\mathbf{d}_k} f_k)^T \mathbf{l}_k$ for $k \in [N-1]$. The matrices H , D , G are from Definition 2.1 and are evaluated at $(\mathbf{z}^0, \boldsymbol{\lambda}^0; \mathbf{d}^0)$.

Proof. See Theorem 5.53 and Remark 5.55 in [3]. Note that Problem (1.1) satisfies Gollan's regularity condition at $(\mathbf{z}^0, \mathbf{d}^0)$ in any direction (as implied by LICQ). See also [12, 13, 14] and references therein for a detailed introduction of sensitivity analysis of nonlinear programming. \square

Under Assumption 2.2, the quadratic program in (2.2) has a unique global minimizer. Note that constraints (2.2b) come from linearizing (1.1b)–(1.1c) whereas the objective (2.2a) is a quadratic approximation of the Lagrange function (2.1). When H is indefinite, solving (2.2) in the context of Problem (1.1) becomes difficult. For example, Riccati-based approaches are not applicable since they all require positivity of H . As shown in [16], however, we can convexify (2.2), *without changing its solution space*, an endeavor we aim to refine in this work.

3. Convexification procedure with linear shifting. We write Problem (2.2) in a way that explicitly connects it to (1.1). We use A_k, B_k, C_k to denote $\nabla_{\mathbf{x}_k}^T f_k$, $\nabla_{\mathbf{u}_k}^T f_k$, $\nabla_{\mathbf{d}_k}^T f_k$ evaluated at $(\mathbf{z}^0; \mathbf{d}^0)$. Problem (2.2) can then be written as

$$(3.1a) \quad \min_{\mathbf{p}, \mathbf{q}} \sum_{k=0}^{N-1} \begin{pmatrix} \mathbf{p}_k \\ \mathbf{q}_k \\ \mathbf{l}_k \end{pmatrix}^T \begin{pmatrix} Q_k & S_k^T & D_{k1}^T \\ S_k & R_k & D_{k2}^T \\ D_{k1} & D_{k2} & 0 \end{pmatrix} \begin{pmatrix} \mathbf{p}_k \\ \mathbf{q}_k \\ \mathbf{l}_k \end{pmatrix} + \mathbf{p}_N^T Q_N \mathbf{p}_N,$$

$$(3.1b) \quad \text{s.t. } \mathbf{p}_{k+1} = A_k \mathbf{p}_k + B_k \mathbf{q}_k + C_k \mathbf{l}_k \quad \forall k \in [N-1],$$

$$(3.1c) \quad \mathbf{p}_0 = \mathbf{l}_{-1}.$$

If $C_k = 0$ and $D_k = 0$, one can directly use the convexification procedure from [16]. Since those terms are nonzero in our work, we need to slightly extend that procedure. To this end, we start by introducing the following uniform (in N) SOSC assumption.

Assumption 3.1 (uniform SOSC). We assume there exists $\gamma > 0$ independent from N such that the reduced Hessian of Problem (1.1), at its unperturbed solution $(\mathbf{z}^0, \boldsymbol{\lambda}^0; \mathbf{d}^0)$, satisfies

$$(3.2) \quad Z^T H Z \succeq \gamma I.$$

Comparing Assumption 3.1 with Assumption 2.2, we see that Assumption 3.1 requires the reduced Hessian matrix to have a positive lower bound uniformly with horizon length N . While our convexification results will also hold under general SOSC, this stronger SOSC will indicate that sensitivity to perturbation decays at a rate independent of N , which offers a theoretical guarantee on sensitivity analysis

of long-horizon nonconvex NLDP. We note that Assumption 3.1 holds for a wide range of dynamics because of the natural block structures in them. For example, the QDP studied in [20], which assumes $H_k = \text{diag}(Q_k, R_k) \succeq \gamma I$, satisfies (3.2). A detailed discussion is provided in the next remark, and a concrete example is tested in numerical experiments described in section 6.

We find it difficult to give a characterization of condition (3.2) that would convincingly identify how likely it is to occur in practice (though for fixed N , being equivalent to SOSC, it is a natural requirement). We give an example of a problem class (in the sense of N being allowed to vary) that exhibits this property.

Remark 3.2 (an example that satisfies Assumption 3.1). Suppose $n_x = n_u = 1$, $A_k = B_k = 1 \forall k \in [N-1]$. By Definition 2.1, we can formulate the Jacobian matrix $G \in \mathbb{R}^{(N+1) \times (2N+1)}$, and one of the matrices, $\tilde{Z} \in \mathbb{R}^{(2N+1) \times N}$, that spans its null space can be written as

$$\tilde{Z}^T = \begin{pmatrix} 0 & 1 & 1 & -1 & & & \\ & 1 & 1 & -1 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & 1 & 1 & -1 & 0 \\ & & & & 1 & 1 & \end{pmatrix}.$$

Note that $\tilde{Z}^T \tilde{Z}$ is a tridiagonal matrix with diagonal entry 3 (the last one is 2) and superdiagonal entry -1 . Thus, by the Gershgorin circle theorem (see [1]), we know $\tilde{Z}^T \tilde{Z}$ can be upper bounded uniformly with its dimension N . In particular, we have $I \preceq \tilde{Z}^T \tilde{Z} \preceq 4I$. Consider a simple Hessian matrix where $Q_k = a_k$, $R_k = b_k$, $S_k = 0$ for all stages k . Then one can easily see that $\tilde{Z}^T H \tilde{Z}$ is also a tridiagonal matrix, with diagonal entry $b_{k-1} + b_k + a_k$ (the last one is $b_{N-1} + a_N$) and superdiagonal entry $-b_k$. By the Gershgorin circle theorem again, we see that as long as $a_k \geq 2|b_k| + 2|b_{k-1}| + 4\gamma \forall k \in [1, N-1]$, and $a_N \geq 2|b_{N-1}| + 4\gamma$, we can get $\tilde{Z}^T H \tilde{Z} \succeq 4\gamma I$, which further implies (3.2), as can be seen from

$$\tilde{Z}^T H \tilde{Z} \succeq 4\gamma I \implies \tilde{Z}^T H \tilde{Z} \succeq \gamma \tilde{Z}^T \tilde{Z} = \gamma Z_2^T Z_2 \implies Z_1^T H Z_1 \succeq \gamma I,$$

where $\tilde{Z} = Z_1 Z_2$ is the QR decomposition. More generally, even if $n_x, n_u > 1$ and $S_k \neq 0$, $\tilde{Z}^T H \tilde{Z}$ still has a block tridiagonal structure, and Assumption 3.1 holds whenever Q_k is sufficiently large. We point out that R_k can be indefinite or negative definite under this setup. Thus, the problem is still nonconvex.

With Assumption 3.1, we now present our convexification procedure. While the algebra is similar to the one in [16], we describe it here for self-consistency of presentation.

DEFINITION 3.3. Given a sequence of matrices $\{\bar{Q}_k\}_{k=0}^N \in \mathbb{R}^{n_x \times n_x}$, we define the quadratic cost sequence as $\bar{r}_k(\mathbf{x}) = \mathbf{x}^T \bar{Q}_k \mathbf{x}, \forall k \in [N]$. Also, for $k \in [N-1]$ we denote the k th stage cost in objective (3.1a) as $L_k(\mathbf{p}_k, \mathbf{q}_k)$, which is a quadratic plus linear function of $\mathbf{p}_k, \mathbf{q}_k$, and $L_N(\mathbf{p}_N) = \mathbf{p}_N^T Q_N \mathbf{p}_N$.

The following lemma will show how to recursively use constraints to modify the objective matrices while keeping the primal solution of the QDP unchanged.

LEMMA 3.4. Given any sequence of matrices $\{\bar{Q}_k\}_{k=0}^N \in \mathbb{R}^{n_x \times n_x}$, we define the transformed cost function as $\bar{L}_N(\mathbf{p}_N) = L_N(\mathbf{p}_N) - \bar{r}_N(\mathbf{p}_N)$ and $\forall k \in [N-1]$

$$\bar{L}_k(\mathbf{p}_k, \mathbf{q}_k) = L_k(\mathbf{p}_k, \mathbf{q}_k) - \bar{r}_k(\mathbf{p}_k) + \bar{r}_{k+1}(A_k \mathbf{p}_k + B_k \mathbf{q}_k + C_k \mathbf{l}_k).$$

Then replacing the objective L_k in (3.1a) with \bar{L}_k will result in the same solution.

Proof. The argument can be derived from the following observation:

$$\begin{aligned}
 & \sum_{k=0}^{N-1} \bar{L}_k(\mathbf{p}_k, \mathbf{q}_k) + \bar{L}_N(\mathbf{p}_N) \\
 &= \sum_{k=0}^{N-1} \left(L_k(\mathbf{p}_k, \mathbf{q}_k) - \bar{r}_k(\mathbf{p}_k) + \bar{r}_{k+1}(A_k \mathbf{p}_k + B_k \mathbf{q}_k + C_k \mathbf{l}_k) \right) + \bar{L}_N(\mathbf{p}_N) \\
 &= \sum_{k=0}^{N-1} \left(L_k(\mathbf{p}_k, \mathbf{q}_k) - \bar{r}_k(\mathbf{p}_k) + \bar{r}_{k+1}(\mathbf{p}_{k+1}) \right) + \bar{L}_N(\mathbf{p}_N) \\
 &= \sum_{k=0}^{N-1} L_k(\mathbf{p}_k, \mathbf{q}_k) - \bar{r}_0(\mathbf{p}_0) + \bar{r}_N(\mathbf{p}_N) + \bar{L}_N(\mathbf{p}_N) \\
 &= \sum_{k=0}^{N-1} L_k(\mathbf{p}_k, \mathbf{q}_k) + L_N(\mathbf{p}_N) - \bar{r}_0(\mathbf{l}_{-1}),
 \end{aligned}$$

where the second and the last equalities are due to (3.1b) and (3.1c), respectively. Since the objective after the replacement stays the same, the conclusion follows. \square

Let us focus on the quadratic matrix in $\bar{L}_k(\mathbf{p}_k, \mathbf{q}_k)$. It can be written as

$$\begin{pmatrix} \tilde{Q}_k & \tilde{S}_k^T & \tilde{D}_{k1}^T \\ \tilde{S}_k & \tilde{R}_k & \tilde{D}_{k2}^T \\ \tilde{D}_{k1} & \tilde{D}_{k2} & * \end{pmatrix} := \begin{pmatrix} Q_k - \bar{Q}_k & S_k^T & D_{k1}^T \\ S_k & R_k & D_{k2}^T \\ D_{k1} & D_{k2} & 0 \end{pmatrix} + \begin{pmatrix} A_k^T \\ B_k^T \\ C_k^T \end{pmatrix} \bar{Q}_{k+1} (A_k \quad B_k \quad C_k).$$

We use “*” to denote the quadratic matrix for \mathbf{l}_k , which can be seen as constant. Now, the problem of obtaining a convexification transformation translates to finding $\{\bar{Q}_k\}_{k=0}^N$ such that the corner matrix $\tilde{H}_k := (\begin{smallmatrix} \tilde{Q}_k & \tilde{S}_k^T \\ \tilde{S}_k & \tilde{R}_k \end{smallmatrix})$ is positive definite. Our convexification procedure is shown in Algorithm 3.1.

Algorithm 3.1 convexification procedure with linear shifting.

- 1: **Input:** $\{H_k\}_{k=0}^N$, $\{A_k, B_k, C_k, D_k\}_{k=0}^{N-1}$, a scalar δ ;
 - 2: $\tilde{H}_N = \tilde{Q}_N = \delta I$;
 - 3: $\tilde{Q}_N = Q_N - \bar{Q}_N$;
 - 4: **for** $k = N-1, \dots, 0$ **do**
 - 5:
$$\begin{pmatrix} \hat{Q}_k & \tilde{S}_k^T & \tilde{D}_{k1}^T \\ \tilde{S}_k & \tilde{R}_k & \tilde{D}_{k2}^T \\ \tilde{D}_{k1} & \tilde{D}_{k2} & * \end{pmatrix} = \begin{pmatrix} Q_k & S_k^T & D_{k1}^T \\ S_k & R_k & D_{k2}^T \\ D_{k1} & D_{k2} & 0 \end{pmatrix} + \begin{pmatrix} A_k^T \\ B_k^T \\ C_k^T \end{pmatrix} \bar{Q}_{k+1} (A_k \quad B_k \quad C_k);$$
 - 6: $\tilde{Q}_k = \tilde{S}_k^T \tilde{R}_k^{-1} \tilde{S}_k + \delta I$;
 - 7: $\tilde{H}_k = \begin{pmatrix} \tilde{Q}_k & \tilde{S}_k^T \\ \tilde{S}_k & \tilde{R}_k \end{pmatrix}$;
 - 8: $\bar{Q}_k = \hat{Q}_k - \tilde{Q}_k$;
 - 9: **end for**
 - 10: **Output:** $\{\tilde{H}_k\}_{k=0}^N$, $\{\tilde{D}_k\}_{k=0}^{N-1}$.
-

This algorithm is an easy extension of [16], and some of its features (e.g., preserving sparsity structure) have been described in [16]. We will use $\tilde{H}(\delta)$ and $\tilde{D}(\delta)$ to denote the output \tilde{H} , \tilde{D} of the algorithm when choosing a specific δ .

From Lemma 3.4 we know that replacing H_k, D_k in (3.1a) with $\tilde{H}_k(\delta), \tilde{D}_k(\delta)$ for any δ will result in the same optimal solution; we subsequently aim to choose δ so that $\tilde{H}_k(\delta)$ is uniformly positive definite. In [16] such a δ was proved to exist but not otherwise estimated in terms of other elements of the problem. Because of the continuity argument used, it cannot be directly assessed whether $\tilde{H}_k(\delta)$ would grow or degrade its positive definiteness should N be increased. Since δ is of special interest in our paper, we propose a finer analysis of it by combining Algorithm 3.1 with a Riccati recursion. Further, as we later show, the uniform boundedness of $\tilde{H}_k(\delta), \tilde{D}_k(\delta)$ with feasible δ is implied by a controllability condition. In the remainder of this section, we first determine a *sufficient* interval for δ . Here *sufficient* means ensuring the positive definiteness (and thus, invertibility) of $\tilde{R}(\delta)$ and $\tilde{H}_k(\delta)$, so that the recursion in Algorithm 3.1 is well defined and the equivalent QDP is strongly convex.

The following two lemmas show the relationship between the convexification algorithm and the solution of Problem (3.1).

LEMMA 3.5 (cost-to-go function for QDP with linear shifting). *For Problem (3.1), let $K_N = Q_N$ be the terminal cost-to-go matrix. For any $k \in [N - 1]$, we define $W_k = R_k + B_k^T K_{k+1} B_k$, k th cost-to-go matrix*

$$(3.3) \quad K_k = Q_k + A_k^T K_{k+1} A_k - (B_k^T K_{k+1} A_k + S_k)^T W_k^{-1} (B_k^T K_{k+1} A_k + S_k),$$

feedback matrix $P_k = -W_k^{-1} (B_k^T K_{k+1} A_k + S_k)$, and closed-loop state transition matrix $E_k = A_k + B_k P_k$. Further, $\forall i, k \in [N - 1]$, we let $V_i^k = -K_{i+1} \prod_{j=k}^i E_j$ and $M_i^k = -(D_{i1} + D_{i2} P_i) \prod_{j=k}^{i-1} E_j$. Then under Assumption 2.2, we have the following.

(i) $W_k \succ 0 \ \forall k \in [N - 1]$.

(ii) *The optimal control variable at stage k for $k \in [N - 1]$ is*

$$(3.4) \quad \begin{aligned} \mathbf{q}_k(\mathbf{p}_k) &= P_k \mathbf{p}_k + W_k^{-1} B_k^T \sum_{i=k+1}^{N-1} (M_i^{k+1})^T \mathbf{l}_i + W_k^{-1} B_k^T \sum_{i=k+1}^{N-1} (V_i^{k+1})^T C_i \mathbf{l}_i \\ &\quad - W_k^{-1} B_k^T K_{k+1} C_k \mathbf{l}_k - W_k^{-1} D_{k2}^T \mathbf{l}_k. \end{aligned}$$

(iii) *The cost-to-go function is*

$$(3.5) \quad J_k(\mathbf{p}_k) = \mathbf{p}_k^T K_k \mathbf{p}_k - 2 \sum_{i=k}^{N-1} \mathbf{l}_i^T M_i^k \mathbf{p}_k - 2 \sum_{i=k}^{N-1} \mathbf{l}_i^T C_i^T V_i^k \mathbf{p}_k + T_k \quad \forall k \in [N],$$

where $T_N = 0$ and T_k , depending on $\{\mathbf{l}_i\}_{i \geq k}$, satisfies $T_k = 0$ if $\mathbf{l}_i = 0 \ \forall k \leq i \leq N - 1$.

Proof. We will use reverse induction to prove property (iii) and incidentally prove (i) and (ii) together. We know $J_N(\mathbf{p}_N) = \mathbf{p}_N^T Q_N \mathbf{p}_N = \mathbf{p}_N^T K_N \mathbf{p}_N$. Suppose $J_{k+1}(\mathbf{p}_{k+1})$ satisfies (3.5). At stage k , by definition of the cost-to-go function, we have

$$(3.6) \quad \begin{aligned} J_k(\mathbf{p}_k) &= \min_{\mathbf{q}_k} \left\{ L_k(\mathbf{p}_k, \mathbf{q}_k) + J_{k+1}(A_k \mathbf{p}_k + B_k \mathbf{q}_k + C_k \mathbf{l}_k) \right\} \\ &= \mathbf{p}_k^T Q_k \mathbf{p}_k + 2 \mathbf{l}_k^T D_{k1} \mathbf{p}_k + \min_{\mathbf{q}_k} \left\{ \mathbf{q}_k^T R_k \mathbf{q}_k + 2 \mathbf{p}_k^T S_k^T \mathbf{q}_k + 2 \mathbf{l}_k^T D_{k2} \mathbf{q}_k \right. \\ &\quad \left. + J_{k+1}(A_k \mathbf{p}_k + B_k \mathbf{q}_k + C_k \mathbf{l}_k) \right\}. \end{aligned}$$

Under Assumption 2.2, we know Problem (3.1) has a unique solution. So the minimization in (3.6) has a unique minimizer, whereas the quadratic matrix of its objective is W_k . Thus, we must have $W_k \succ 0$ (proving the induction step for (i)). In particular, for (3.6) we plug in the formula of $J_{k+1}(\cdot)$, shown in (3.5) at $k+1$, and have

$$\begin{aligned}
J_k(\mathbf{p}_k) &= \mathbf{p}_k^T Q_k \mathbf{p}_k + (A_k \mathbf{p}_k + C_k \mathbf{l}_k)^T K_{k+1} (A_k \mathbf{p}_k + C_k \mathbf{l}_k) + 2 \mathbf{l}_k^T D_{k1} \mathbf{p}_k \\
&\quad - 2 \sum_{i=k+1}^{N-1} \mathbf{l}_i^T M_i^{k+1} (A_k \mathbf{p}_k + C_k \mathbf{l}_k) - 2 \sum_{i=k+1}^{N-1} \mathbf{l}_i^T C_i^T V_i^{k+1} (A_k \mathbf{p}_k + C_k \mathbf{l}_k) + T_{k+1} \\
&\quad + \min_{\mathbf{q}_k} \left\{ \mathbf{q}_k^T W_k \mathbf{q}_k + 2 \mathbf{p}_k^T S_k^T \mathbf{q}_k + 2 \mathbf{l}_k^T D_{k2} \mathbf{q}_k + 2 (A_k \mathbf{p}_k + C_k \mathbf{l}_k)^T K_{k+1} B_k \mathbf{q}_k \right. \\
&\quad \left. - 2 \sum_{i=k+1}^{N-1} \mathbf{l}_i^T M_i^{k+1} B_k \mathbf{q}_k - 2 \sum_{i=k+1}^{N-1} \mathbf{l}_i^T C_i^T V_i^{k+1} B_k \mathbf{q}_k \right\} \\
(3.7) \quad &= \min_{\mathbf{q}_k} \left\{ \mathbf{q}_k^T W_k \mathbf{q}_k + 2 \mathbf{a}_k^T \mathbf{q}_k \right\} + \mathbf{p}_k^T (Q_k + A_k^T K_{k+1} A_k) \mathbf{p}_k + 2 \mathbf{b}_k^T \mathbf{p}_k \\
&\quad + T_{k+1} + \underbrace{\mathbf{l}_k^T C_k^T K_{k+1} C_k \mathbf{l}_k - 2 \sum_{i=k+1}^{N-1} \mathbf{l}_i^T M_i^{k+1} C_k \mathbf{l}_k - 2 \sum_{i=k+1}^{N-1} \mathbf{l}_i^T C_i^T V_i^{k+1} C_k \mathbf{l}_k}_{T'_k},
\end{aligned}$$

where $\mathbf{b}_k = (A_k^T K_{k+1} C_k + D_{k1}^T) \mathbf{l}_k - \sum_{i=k+1}^{N-1} A_k^T ((M_i^{k+1})^T + (V_i^{k+1})^T C_i) \mathbf{l}_i$ and

$$\begin{aligned}
\mathbf{a}_k &= S_k \mathbf{p}_k + D_{k2}^T \mathbf{l}_k + B_k^T K_{k+1} (A_k \mathbf{p}_k + C_k \mathbf{l}_k) \\
&\quad - \sum_{i=k+1}^{N-1} B_k^T (M_i^{k+1})^T \mathbf{l}_i - \sum_{i=k+1}^{N-1} B_k^T (V_i^{k+1})^T C_i \mathbf{l}_i.
\end{aligned}$$

From the definition of P_k , the unique solution \mathbf{q}_k of (3.7), at fixed \mathbf{p}_k , is then given by

$$\begin{aligned}
(3.8) \quad \mathbf{q}_k(\mathbf{p}_k) &= -W_k^{-1} \mathbf{a}_k \\
&= P_k \mathbf{p}_k + W_k^{-1} B_k^T \sum_{i=k+1}^{N-1} (M_i^{k+1})^T \mathbf{l}_i + W_k^{-1} B_k^T \sum_{i=k+1}^{N-1} (V_i^{k+1})^T C_i \mathbf{l}_i \\
&\quad - W_k^{-1} (D_{k2} + C_k^T K_{k+1} B_k)^T \mathbf{l}_k.
\end{aligned}$$

This verifies the induction step for claim (ii). Plugging (3.8) into (3.7) and noting that $M_i^{k+1} E_k = M_i^k$, $V_i^{k+1} E_k = V_i^k$, and the definitions in (3.3), by some extended but straightforward calculations we get

$$\begin{aligned}
J_k(\mathbf{p}_k) &= \mathbf{p}_k^T K_k \mathbf{p}_k + 2 \left(\mathbf{l}_k^T (D_{k1} + D_{k2} P_k) + \mathbf{l}_k^T C_k^T K_{k+1} E_k - \sum_{i=k+1}^{N-1} \mathbf{l}_i^T M_i^k \right. \\
&\quad \left. - \sum_{i=k+1}^{N-1} \mathbf{l}_i^T C_i^T V_i^k \right) \mathbf{p}_k + T_k \\
(3.9) \quad &= \mathbf{p}_k^T K_k \mathbf{p}_k - 2 \sum_{i=k}^{N-1} \mathbf{l}_i^T M_i^k \mathbf{p}_k - 2 \sum_{i=k}^{N-1} \mathbf{l}_i^T C_i^T V_i^k \mathbf{p}_k + T_k.
\end{aligned}$$

Here, $T_k = T'_k - \|(D_{k2}^T + B_k^T K_{k+1} C_k) \mathbf{l}_k - \sum_{i=k+1}^{N-1} B_k^T (M_i^{k+1} + C_i^T V_i^{k+1})^T \mathbf{l}_i\|_{W_k^{-1}}^2$.

We see that T_k depends only on $\{\mathbf{l}_i\}_{i=k}^{N-1}$; and if $\mathbf{l}_i = 0 \forall k \leq i \leq N-1$, we will have that $T_k = 0$. So we have proved claim (iii) and, thus, the statement. \square

Lemma 3.5 gives the explicit form for the optimal control variable \mathbf{q} for Problem (3.1). This lemma will also be useful in the next section when we analyze the optimal state variable. The next lemma shows how the convexification transformation relates to the cost-to-go matrices defined in the preceding lemma.

LEMMA 3.6. *When applying the convexification algorithm 3.1 with $\delta = 0$, we will have $\tilde{R}_k(0) = W_k \forall k \in [N-1]$ and $\bar{Q}_k(0) = K_k \forall k \in [N]$.*

Proof. We use reverse induction for $\bar{Q}_k(0)$, and $\tilde{R}_k(0)$ can be proved in concert. For the last stage N , we know that $\bar{Q}_N(0) = Q_N = K_N$. Suppose $\bar{Q}_{k+1}(0) = K_{k+1}$. Then using the definitions of W_k and K_k in (3.3), we can get

$$\begin{aligned}\tilde{R}_k(0) &= R_k + B_k^T \bar{Q}_{k+1}(0) B_k = R_k + B_k^T K_{k+1} B_k = W_k, \\ \bar{Q}_k(0) &= \hat{Q}_k(0) - \tilde{Q}_k(0) = Q_k + A_k^T \bar{Q}_{k+1}(0) A_k - \tilde{Q}_k(0) \\ &= Q_k + A_k^T \bar{Q}_{k+1}(0) A_k - \tilde{S}_k^T(0) \tilde{R}_k^{-1}(0) \tilde{S}_k(0) \\ &= Q_k + A_k^T K_{k+1} A_k - (B_k^T K_{k+1} A_k + S_k)^T W_k^{-1} (B_k^T K_{k+1} A_k + S_k) \\ &= K_k,\end{aligned}$$

where the second equality from the end is due to the induction assumption and the last equality is due to (3.3). This completes the proof. \square

COROLLARY 3.7. *Under Assumption 2.2, Algorithm 3.1 can be carried out successfully by setting $\delta = 0$. The output satisfies $\tilde{R}_k(0) \succ 0$ and $\tilde{H}_k(0) \succeq 0$.*

Proof. Without linear terms, Lemma 10 in [16] proved a similar result (with a different approach). For self-consistency, we present a brief proof for our more general convexification procedure. By Lemmas 3.5(i) and 3.6, we know that $\tilde{R}_k(0) = W_k \succ 0$ (and thus it is invertible) and

$$\begin{aligned}\tilde{H}_k(0) &= \begin{pmatrix} \tilde{S}_k^T(0) \tilde{R}_k^{-1}(0) \tilde{S}_k(0) & \tilde{S}_k^T(0) \\ \tilde{S}_k(0) & \tilde{R}_k(0) \end{pmatrix} \\ &= \begin{pmatrix} I & \tilde{S}_k^T(0) \tilde{R}_k^{-1}(0) \\ 0 & I \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & \tilde{R}_k(0) \end{pmatrix} \begin{pmatrix} I & 0 \\ \tilde{R}_k^{-1}(0) \tilde{S}_k(0) & I \end{pmatrix} \succeq 0.\end{aligned}\quad \square$$

Based on these two lemmas, we know that when increasing δ slightly, by continuity of $\tilde{R}_k(\delta)$, we still can get $\tilde{R}_k(\delta) \succ 0$. In this case, we can show $\tilde{H}_k(\delta) \succ 0$ because the Schur complement of $\tilde{R}_k(\delta)$ is $\delta I \succ 0$. However, since $\tilde{R}_k(\delta)$ is algebraically complicated, how large we can increase δ is not immediately clear. The following theorem presents a sufficient interval for δ in the sense previously discussed.

THEOREM 3.8 (sufficient bound for δ). *Under Assumption 3.1, executing Algorithm 3.1 with $\delta \in (0, \gamma)$ will result in $\tilde{R}_k(\delta) \succ 0$ and $\tilde{H}_k(\delta) \succ 0$, where γ is defined in (3.2).*

Proof. For any $\delta \in (0, \gamma)$, we define $Q_k^\delta = Q_k - \delta I$ and $H_k^\delta = \begin{pmatrix} Q_k^\delta & S_k^T \\ S_k & R_k \end{pmatrix}$ and let $H^\delta = \text{diag}(H_0^\delta, \dots, H_N^\delta)$. Moreover, we define $I^\delta = (H - H^\delta)/\delta = \text{diag}(I, \mathbf{0}, I, \mathbf{0}, \dots, I)$. Thus, we have $I - I^\delta = \text{diag}(\mathbf{0}, I, \mathbf{0}, I, \dots, \mathbf{0}) \succeq 0$ and $Z^T I^\delta Z = Z^T Z - Z^T (I - I^\delta) Z \preceq I$, where the last inequality comes from the fact that $Z^T Z = I$. (Note that Z has

orthonormal columns.) Furthermore, from Assumption 3.1 we can get

$$(3.10) \quad Z^T H^\delta Z = Z^T H Z - \delta Z^T I^\delta Z \succeq (\gamma - \delta) I \succ 0.$$

Consider applying Algorithm 3.1 on H^δ with the same input matrices $\{A_k, B_k, C_k, D_k\}_{k=0}^{N-1}$ and shifting parameter $\tilde{\delta} = 0$. Its output is denoted as $\tilde{H}_k^\delta(0) = (\tilde{Q}_k^\delta(0) \ \tilde{S}_k^{\delta T}(0))$. From (3.10) we know that SOSC stated in Assumption 2.2 is satisfied for H^δ . So, from Corollary 3.7, we can get $\tilde{R}_k^\delta(0) \succ 0$. Furthermore, we have the following claims.

CLAIM 1. *We have $\bar{Q}_k^\delta(0) = \bar{Q}_k(\delta) \ \forall k \in [N]$ and $\tilde{R}_k^\delta(0) = \tilde{R}_k(\delta) \ \forall k \in [N-1]$. We prove the statement by reverse induction. For $k = N$, we know $\bar{Q}_N^\delta(0) = Q_N^\delta = Q_N - \delta I = \bar{Q}_N(\delta)$. Suppose $\bar{Q}_{k+1}^\delta(0) = \bar{Q}_{k+1}(\delta)$. Then we have*

$$\begin{aligned} \tilde{R}_k^\delta(0) &= R_k + B_k^T \bar{Q}_{k+1}^\delta(0) B_k = R_k + B_k^T \bar{Q}_{k+1}(\delta) B_k = \tilde{R}_k(\delta), \\ \bar{Q}_k^\delta(0) &= \hat{Q}_k^\delta(0) - \tilde{Q}_k^\delta(0) \\ &= Q_k^\delta + A_k^T \bar{Q}_{k+1}^\delta(0) A_k - (S_k + B_k^T \bar{Q}_{k+1}^\delta(0) A_k)^T (\tilde{R}_k^\delta(0))^{-1} (S_k + B_k^T \bar{Q}_{k+1}^\delta(0) A_k) \\ &= Q_k + A_k^T \bar{Q}_{k+1}(\delta) A_k - (S_k + B_k^T \bar{Q}_{k+1}(\delta) A_k)^T \tilde{R}_k^{-1}(\delta) (S_k + B_k^T \bar{Q}_{k+1}(\delta) A_k) - \delta I \\ &= \bar{Q}_k(\delta), \end{aligned}$$

which proves Claim 1.

CLAIM 2. *We have $\tilde{S}_k^\delta(0) = \tilde{S}_k(\delta) \ \forall k \in [N-1]$ and $\tilde{Q}_k^\delta(0) = \tilde{Q}_k(\delta) - \delta I$, $\forall k \in [N]$. From Claim 1, we know that $\forall k \in [N-1]$*

$$\begin{aligned} \tilde{S}_k^\delta(0) &= S_k + B_k^T \bar{Q}_{k+1}^\delta(0) A_k = S_k + B_k^T \bar{Q}_{k+1}(\delta) A_k = \tilde{S}_k(\delta), \\ \tilde{Q}_k^\delta(0) &= (\tilde{S}_k^\delta(0))^T (\tilde{R}_k^\delta(0))^{-1} \tilde{S}_k^\delta(0) = \tilde{S}_k^T(\delta) \tilde{R}_k^{-1}(\delta) \tilde{S}_k(\delta) = \tilde{Q}_k(\delta) - \delta I. \end{aligned}$$

For $k = N$ we have $\tilde{Q}_N^\delta(0) = \mathbf{0} = \tilde{Q}_N(\delta) - \delta I$. This completes the proof of Claim 2.

From these two claims and Corollary 3.7, we can get that $\tilde{R}_k(\delta) = \tilde{R}_k^\delta(0) \succ 0$ and

$$\tilde{H}_k(\delta) = \begin{pmatrix} \tilde{Q}_k(\delta) & \tilde{S}_k^T(\delta) \\ \tilde{S}_k(\delta) & \tilde{R}_k(\delta) \end{pmatrix} = \begin{pmatrix} \tilde{S}_k^{\delta T}(0) \tilde{R}_k^\delta(0)^{-1} \tilde{S}_k^\delta(0) + \delta I & \tilde{S}_k^{\delta T}(0) \\ \tilde{S}_k^\delta(0) & \tilde{R}_k^\delta(0) \end{pmatrix} \succ 0. \quad \square$$

To summarize, in this section we propose a convexification procedure with linear shifting and derive a sufficient, explicit interval for the scalar δ that ensures that Algorithm 3.1 is well defined and results in positive definite matrices of the transformed quadratic program. In particular, we prove that under Assumption 3.1, δ can be set as large as the minimum eigenvalue of the reduced Hessian matrix (i.e., γ).

In the next section, some detailed properties of the matrices appearing in the transformed QDP are discussed in preparation for our main sensitivity analysis result.

4. Uniform boundedness of convexified QDP. We delve deeper into Algorithm 3.1 and analyze some properties of $\tilde{H}(\delta), \tilde{D}(\delta)$ that are critical for exponential decay results. Basically, we are going to analyze properties for the lower bound $\tilde{H}(\delta)$ and the upper bound $\tilde{D}(\delta)$ by some quantities independent of the time index that depend on both the parameter δ and the bounds on the problem data at the unperturbed solution $(\mathbf{z}^0, \boldsymbol{\lambda}^0; \mathbf{d}^0)$. Building on the boundedness results of this section, we further derive the decay rate of sensitivity to perturbations on \mathbf{d}^0 in the next section. We start by presenting boundedness assumptions for long-horizon dynamic programming. Similar assumptions have been discussed in [20].

DEFINITION 4.1 (controllability matrix). *For any starting stage $k \in [N - 1]$ and evolution length $t > 0$ such that $[k, k + t - 1] \subset [N - 1]$, we define the controllability matrix to be $\Xi_{k,t} = [B_{k+t-1} A_{k+t-1} B_{k+t-2} \dots, \prod_{l=1}^{t-1} A_{k+l} B_k]$.*

Assumption 4.2. We have following assumptions for Problem (3.1):

(i) There exists a constant $\Upsilon > 0$ independent of N such that $\|Q_k\| \vee \|R_k\| \vee \|S_k\| \vee \|D_{k1}\| \vee \|D_{k2}\| \vee \|A_k\| \vee \|B_k\| \vee \|C_k\| \leq \Upsilon \forall k$.

(ii) For any $k \in [N - 1]$, there exists $1 \leq t_k \leq N - k$ such that $\Xi_{k,t_k} \Xi_{k,t_k}^T \succeq \lambda_C I$, and there exists $1 \leq t \leq N$ such that $t_k \leq t$. Moreover, we assume that the positive parameters λ_C and t are independent from k and N .

Similar to our uniform SOSC assumptions, the parameters Υ , λ_C , and t are assumed independent of horizon length N . This setup will allow us to have the sensitivity decay rate depend only on these parameters and not on N . Thus, as N increases (i.e., we solve more and more stages), the sensitivity bounds remain unchanged. We note that Assumption 4.2(i) is also assumed in [20] (see Assumption 2.1), while the statement in Assumption 4.2(ii) is slightly different from its analogue (see Definition 2.2(b)), where the uniform controllability condition was borrowed from [9] and considered a truncation of an infinite-horizon problem. In that case $\Xi_{k,t}$ is well defined even when $k + t > N$. To avoid defining the perturbation of infinite-horizon problem, we define our controllability matrix $\Xi_{k,t}$ only for $k + t \leq N$, since we do not assume to have access to $\{A_k, B_k\}_{k=N}^\infty$.

From Assumption 4.2, we obtain that $\forall k \in [N - 1]$ and $\forall 1 \leq j \leq t_k$,

$$(4.1) \quad \|\Xi_{k,j}\| = \left\| \left[B_{k+j-1} \dots \prod_{l=1}^{j-1} A_{k+l} B_k \right] \right\| \\ \leq \sum_{i=0}^{j-1} \left\| \prod_{l=i+1}^{j-1} A_{k+l} B_{k+i} \right\| \leq \sum_{i=0}^{j-1} \Upsilon^{j-i} \leq \frac{\Upsilon(1 - \Upsilon^t)}{1 - \Upsilon} := \Psi.$$

Now we are ready to present some properties for $\tilde{H}(\delta)$. Similar to the proof of Theorem 3.8, we focus on $\tilde{H}(0)$ and then make use of the relationship between $\tilde{H}(\delta)$ and $\tilde{H}^\delta(0)$.

LEMMA 4.3. *Under Assumptions 2.2 and 4.2, we have that $\|\bar{Q}_k(0)\| \leq \Upsilon_Q \forall k \in [N]$ for some parameter Υ_Q independent of N .*

Proof. For any $k \in [N - 1]$ and $\bar{p}_k \in \mathbb{R}^{n_x}$, let us consider the following QDP without linear shift:

$$(4.2a) \quad \min_{\substack{\mathbf{p}_{k:N} \\ \mathbf{q}_{k:N-1}}} \sum_{i=k}^{N-1} \begin{pmatrix} \mathbf{p}_i \\ \mathbf{q}_i \end{pmatrix}^T \begin{pmatrix} Q_i & S_i^T \\ S_i & R_i \end{pmatrix} \begin{pmatrix} \mathbf{p}_i \\ \mathbf{q}_i \end{pmatrix} + \mathbf{p}_N^T Q_N \mathbf{p}_N,$$

$$(4.2b) \quad \text{s.t. } \mathbf{p}_{i+1} = A_i \mathbf{p}_i + B_i \mathbf{q}_i, \quad i = k, k + 1, \dots, N - 1,$$

$$(4.2c) \quad \mathbf{p}_k = \bar{p}_k.$$

Comparing Problem (4.2) with Problem (3.1), we see that D_k and C_k vanish, but we know that (4.2) still satisfies $Z^T H Z \succ 0$. From (3.5) in Lemma 3.5 with $\mathbf{l}_{k:N-1} = \mathbf{0}$, we know that Problem (4.2) has minimum value $\bar{p}_k K_k \bar{p}_k$, and by Lemma 3.6 we have the cost-to-go matrix $K_k = \bar{Q}_k(0)$. On the other hand, for $i \geq k$, successively applying

$\mathbf{p}_{i+1} = A_i \mathbf{p}_i + B_i \mathbf{q}_i$ returns that

$$(4.3) \quad \mathbf{p}_{k+j} - \left(\prod_{l=0}^{j-1} A_{k+l} \right) \bar{\mathbf{p}}_k = \Xi_{k,j} \begin{pmatrix} \mathbf{q}_{k+j-1} \\ \vdots \\ \mathbf{q}_k \end{pmatrix} := \Xi_{k,j} \mathbf{q}_{k:k+j-1} \quad \text{for } j \geq 0.$$

We now consider a special feasible point defined as follows. Let

$$(4.4) \quad \bar{\mathbf{q}}_{k:k+t_k-1} = -\Xi_{k,t_k}^T (\Xi_{k,t_k} \Xi_{k,t_k}^T)^{-1} \left(\prod_{l=0}^{t_k-1} A_{k+l} \right) \bar{\mathbf{p}}_k,$$

and $\bar{\mathbf{p}}_{k:k+t_k}$ is computed by using (4.3) for $1 \leq j \leq t_k$. Plugging (4.4) into (4.3) with $j = t_k$, we can get $\bar{\mathbf{p}}_{k+t_k} = 0$. Further, we let $\bar{\mathbf{q}}_{k+t_k:N-1} = 0$ and thus have $\bar{\mathbf{p}}_{k+t_k+1:N} = \mathbf{0}$ from (4.3). Our special feasible point can be summarized (in the reverse order) as

$$(4.5a) \quad \bar{\mathbf{q}}_{k:N-1} := \begin{pmatrix} \bar{\mathbf{q}}_{k+t_k:N-1} \\ \bar{\mathbf{q}}_{k:k+t_k-1} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ -\Xi_{k,t_k}^T (\Xi_{k,t_k} \Xi_{k,t_k}^T)^{-1} \left(\prod_{l=0}^{t_k-1} A_{k+l} \right) \bar{\mathbf{p}}_k \end{pmatrix},$$

$$(4.5b) \quad \bar{\mathbf{p}}_{k:N} := \begin{pmatrix} \bar{\mathbf{p}}_{k+t_k:N} \\ \bar{\mathbf{p}}_{k:k+t_k-1} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \bar{\mathbf{p}}_{k:k+t_k-1} \end{pmatrix}.$$

Combining (4.5), (4.3), and (4.1) and Assumption 4.2, we can get the following:

$$(4.6) \quad \|\bar{\mathbf{q}}_{k:N-1}\| \leq \frac{\Psi \Upsilon^t}{\lambda_C} \|\bar{\mathbf{p}}_k\|, \quad \|\bar{\mathbf{p}}_i\| \leq \left(\Upsilon^{i-k} + \frac{\Psi^2 \Upsilon^t}{\lambda_C} \right) \|\bar{\mathbf{p}}_k\| \quad \forall i \in [k+1, k+t_k-1].$$

We then have

$$\begin{aligned} \bar{\mathbf{p}}_k^T \bar{Q}_k(0) \bar{\mathbf{p}}_k &= \min_{\substack{(\mathbf{p}_{k:N}, \mathbf{q}_{k:N-1}) \\ \text{satisfy (4.2b)}}} \sum_{i=k}^{N-1} \begin{pmatrix} \mathbf{p}_i \\ \mathbf{q}_i \end{pmatrix}^T \begin{pmatrix} Q_i & S_i^T \\ S_i & R_i \end{pmatrix} \begin{pmatrix} \mathbf{p}_i \\ \mathbf{q}_i \end{pmatrix} + \mathbf{p}_N^T Q_N \mathbf{p}_N \\ &\leq \sum_{i=k}^{N-1} \begin{pmatrix} \bar{\mathbf{p}}_i \\ \bar{\mathbf{q}}_i \end{pmatrix}^T \begin{pmatrix} Q_i & S_i^T \\ S_i & R_i \end{pmatrix} \begin{pmatrix} \bar{\mathbf{p}}_i \\ \bar{\mathbf{q}}_i \end{pmatrix} + \bar{\mathbf{p}}_N^T Q_N \bar{\mathbf{p}}_N = \sum_{i=k}^{k+t_k-1} \begin{pmatrix} \bar{\mathbf{p}}_i \\ \bar{\mathbf{q}}_i \end{pmatrix}^T \begin{pmatrix} Q_i & S_i^T \\ S_i & R_i \end{pmatrix} \begin{pmatrix} \bar{\mathbf{p}}_i \\ \bar{\mathbf{q}}_i \end{pmatrix} \\ &\stackrel{(1)}{\leq} 2\Upsilon \sum_{i=k}^{k+t_k-1} \|\bar{\mathbf{p}}_i\|^2 + 2\Upsilon \sum_{i=k}^{k+t_k-1} \|\bar{\mathbf{q}}_i\|^2 \\ &\stackrel{(2)}{\leq} 2\Upsilon \underbrace{\left(1 + \frac{\Psi^2 \Upsilon^{2t}}{\lambda_C^2} + \sum_{i=1}^{t-1} \left(\Upsilon^i + \frac{\Psi^2 \Upsilon^t}{\lambda_C} \right)^2 \right)}_{\Upsilon_{\bar{Q}}} \|\bar{\mathbf{p}}_k\|^2 := \Upsilon_{\bar{Q}} \|\bar{\mathbf{p}}_k\|^2. \end{aligned}$$

Inequalities (1) and (2) hold by Assumptions 4.2(i) and (4.6), respectively. Since $\bar{\mathbf{p}}_k$ was chosen arbitrarily, we conclude the argument in the lemma by noting that $\Upsilon_{\bar{Q}} > \Upsilon \geq \|Q_N\| = \|\bar{Q}_N(0)\|$. \square

Note that the upper bound would be difficult to obtain from Algorithm 3.1, because of the challenge of bounding the terms $\tilde{S}_k^T(0) \tilde{R}_k^{-1}(0) \tilde{S}_k(0)$. We resolve the difficulty by making use of the controllability of the problem. A direct consequence of Lemma 4.3 is the boundedness of the cost-to-go matrix K_k . By Lemma 3.6, we have

$$(4.7) \quad \|K_k\| = \|\bar{Q}_k(0)\| \leq \Upsilon_{\bar{Q}}.$$

Another consequence is the boundedness of $\tilde{H}(0)$.

COROLLARY 4.4. *Under Assumptions 2.2 and 4.2, the matrices $\tilde{Q}_k(0)$, $\tilde{R}_k(0)$, and $\tilde{S}_k(0)$ have a global upper bound independent of N .*

Proof. From Algorithm 3.1, Assumption 4.2, and Lemma 4.3, we obtain the following:

$$\begin{aligned}\|\hat{Q}_k(0)\| &= \|Q_k + A_k^T \bar{Q}_{k+1}(0) A_k\| \leq \Upsilon + \Upsilon^2 \Upsilon_{\bar{Q}}, \\ \|\tilde{S}_k(0)\| &= \|S_k + B_k^T \bar{Q}_{k+1}(0) A_k\| \leq \Upsilon + \Upsilon^2 \Upsilon_{\bar{Q}}, \\ \|\tilde{R}_k(0)\| &= \|R_k + B_k^T \bar{Q}_{k+1}(0) B_k\| \leq \Upsilon + \Upsilon^2 \Upsilon_{\bar{Q}}.\end{aligned}$$

We also have $\|\tilde{Q}_k(0)\| = \|\hat{Q}_k(0) - \bar{Q}_k(0)\| \leq \Upsilon + \Upsilon_{\bar{Q}} + \Upsilon^2 \Upsilon_{\bar{Q}}$. \square

Based on the preceding results, we are ready to present the upper-bound-preserving property of Algorithm 3.1.

THEOREM 4.5 (upper-bound-preserving property). *For Problem (3.1), if Assumptions 3.1 and 4.2 hold, the output of Algorithm 3.1 for $\delta \in (0, \gamma)$ satisfies*

$$\max \{\|\tilde{Q}_k(\delta)\|, \|\tilde{R}_k(\delta)\|, \|\tilde{S}_k(\delta)\|, \|\tilde{D}_{k1}(\delta)\|, \|\tilde{D}_{k2}(\delta)\|\} \leq \tilde{\Upsilon} \quad \forall k,$$

for some constant $\tilde{\Upsilon}$ independent of N . Consequently, $\tilde{H}_k(\delta)$ is uniformly upper bounded in k .

Proof. For any $\delta \in (0, \gamma)$, similar to Theorem 3.8, we define $Q_k^\delta = Q_k - \delta I$ and $H_k^\delta = \begin{pmatrix} Q_k^\delta & S_k^T \\ S_k & R_k \end{pmatrix}$. We focus only on H^δ for which Assumption 2.2 still holds. Applying Algorithm 3.1 on input matrices $\{H_k^\delta, A_k, B_k, C_k, D_k\}_{k=0}^{N-1}$ with shifting parameter $\tilde{\delta} = 0$, we denote the output by $\tilde{H}_k^\delta(0)$. Also note that Assumption 4.2 still holds for H^δ . From Lemma 4.3 and Corollary 4.4 we thus have

$$(4.8) \quad \max \{\|\tilde{Q}_k^\delta(0)\|, \|\tilde{R}_k^\delta(0)\|, \|\tilde{S}_k^\delta(0)\|\} \leq \Upsilon_1$$

for some constant Υ_1 . By Theorem 3.8 (see Claims 1 and 2 in its proof), we know that

$$(4.9) \quad \tilde{Q}_k^\delta(0) = \tilde{Q}_k(\delta) - \delta I, \quad \tilde{R}_k^\delta(0) = \tilde{R}_k(\delta), \quad \text{and} \quad \tilde{S}_k^\delta(0) = \tilde{S}_k(\delta).$$

Combining (4.9) with (4.8), we obtain that $\tilde{Q}_k(\delta)$, $\tilde{R}_k(\delta)$, and $\tilde{S}_k(\delta)$ are globally upper bounded by $\Upsilon_1 + \delta$, implying $\tilde{H}_k(\delta)$ also has a global upper bound independent from N . Further, by Lemma 4.3 and Assumption 4.2(i), we can get

$$\begin{aligned}\|\tilde{D}_{k1}(\delta)\| &= \|D_{k1} + C_k^T \bar{Q}_k(\delta) A_k\| \leq \Upsilon + \Upsilon^2 \|\bar{Q}_k(\delta)\| \leq \Upsilon + \Upsilon^2 (\Upsilon_1 + \delta), \\ \|\tilde{D}_{k2}(\delta)\| &= \|D_{k2} + C_k^T \bar{Q}_k(\delta) B_k\| \leq \Upsilon + \Upsilon^2 \|\bar{Q}_k(\delta)\| \leq \Upsilon + \Upsilon^2 (\Upsilon_1 + \delta).\end{aligned}$$

We have thus shown that $\tilde{D}_{k1}(\delta)$ and $\tilde{D}_{k2}(\delta)$ have an upper bound uniform in k . We define $\tilde{\Upsilon} = \max(\Upsilon_1 + \delta, \Upsilon + \Upsilon^2 (\Upsilon_1 + \delta))$, which completes the proof. \square

As a side note, we see that if we let $\delta \in (0, \gamma \wedge 1)$, the upper bound $\tilde{\Upsilon}$ would be independent of δ and γ by replacing δ by 1 in the last terms of the above sequences of inequalities.

In the following, we will investigate the lower bound of the QDP obtained by convexification, Algorithm 3.1. Here, by the lower bound of a positive definite matrix, we mean its smallest eigenvalue. One can easily see that $\tilde{Q}_k(\delta) = \tilde{S}_k^T(\delta) \tilde{R}_k^{-1}(\delta) \tilde{S}_k(\delta) + \delta I \succeq \delta I$. On the other hand, there is no immediate guarantee that $\tilde{H}_k(\delta)$ has global

lower bound away from 0 without some conditions on $\tilde{R}_k(\delta)$. For example, if $\tilde{R}_k(\delta) \succ 0$ but $\tilde{R}_k(\delta) \rightarrow 0$ as $k \rightarrow \infty$ and $\tilde{S}_k(\delta) = 0$, then $\tilde{H}_k(\delta) = \text{diag}(\delta I, \tilde{R}_k(\delta)) \succ 0$ but does not have a lower bound away from 0 independently of N . Nevertheless, this adverse situation will not occur under our uniform SOSC (Assumption 3.1). We will illustrate this fact through the next several theorems.

First, we present examples in the next two remarks to illustrate that (i) $\delta \in (0, \gamma)$ is sufficient but not necessary and (ii) only SOSC cannot guarantee a lower bound for $\tilde{H}_k(\delta)$ uniformly in N .

Remark 4.6. Suppose $S_k = 0$, $Q_k = R_k = \gamma I$. Then we know $H_k = \text{diag}(Q_k, R_k) = \text{diag}(\gamma I, \gamma I)$ and uniform SOSC holds with $Z^T H Z \succeq \gamma I$ for any $\{A_k, B_k, C_k, D_k\}_{k=0}^{N-1}$. Let us apply Algorithm 3.1 on $\{H_k\}_{k=0}^N$ with any matrices $\{A_k, B_k, C_k, D_k\}_{k=0}^{N-1}$ and set parameter $\delta = \gamma$. Then we have the output $\tilde{H}_k(\gamma) = \text{diag}(\gamma I, \gamma I) \succeq \gamma I$. Thus, we see $\delta = \gamma$ is also feasible for this example although $\delta \in (0, \gamma)$ is crucial for us to prove Theorem 3.8, because of (3.10).

Remark 4.7. Let $n_x = n_u = 1$. For $k \in [N-1]$, suppose $A_k = B_k = C_k = S_k = 0$, $Q_k = 1$, and $R_k = 1/k$. We then have that $H_k = \text{diag}(Q_k, R_k) = \text{diag}(1, 1/k) \succ 0$, and it satisfies SOSC with $Z^T H Z \succeq 1/(N-1) \cdot I$ (here $Z = I_{2N+1}$). By applying Algorithm 3.1 with any $\delta > 0$, we get $\tilde{H}_k(\delta) = \text{diag}(\delta, 1/k)$. However, we see $\lambda_{\min}(H_k(\delta)) \leq 1/k \rightarrow 0$. So we cannot find a lower bound independent of N under only SOSC.

To make the proofs concise, we state some algebra preliminaries in the next lemma.

LEMMA 4.8. (i) For any matrix $A \in \mathbb{R}^{m \times n}$, suppose $\tilde{A} = (\tilde{A}_{ij}) \in \mathbb{R}^{\tilde{m} \times \tilde{n}}$ with $\tilde{m} \geq m$ and $\tilde{n} \geq n$ is its zero-padded extension (i.e., $\forall i \in [\tilde{m}] \setminus [m]$ or $j \in [\tilde{n}] \setminus [n]$, $\tilde{A}_{ij} = 0$). Then we have $\|\tilde{A}\| = \|A\|$.

(ii) For a sequence of symmetric positive definite matrices $H_k = \begin{pmatrix} A_k & B_k^T \\ B_k & C_k \end{pmatrix}$, $k = 1, 2, \dots$, where A_k, C_k are square matrices, suppose $C_k \succeq \beta_C I \succ 0$ and its Schur complement $A_k - B_k^T C_k^{-1} B_k \succeq \beta_S I \succ 0$, and $\|B_k\| \leq \beta_B$. Then we have $\forall k$

$$H_k \succeq \lambda_{BCS} I \succ 0, \quad \text{where } \lambda_{BCS} = \left(\frac{\beta_C}{\beta_C + \beta_B} \right)^2 \cdot (\beta_S \wedge \beta_C).$$

Proof. (i) Without loss of generality, we assume that $\tilde{A} = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}$; otherwise we just do permutations of \tilde{A} that preserve the operator norm. Let $\tilde{A} = (B, \mathbf{0})$, where $B^T = (A^T, \mathbf{0})$. We then have

$$\|\tilde{A}\|^2 = \lambda_{\max}(\tilde{A} \tilde{A}^T) = \lambda_{\max}(BB^T) = \|B\|^2 = \lambda_{\max}(B^T B) = \lambda_{\max}(A^T A) = \|A\|^2.$$

(ii) Note that if H_k is positive definite, $H_k \succeq \lambda_{BCS} I \Leftrightarrow \|H_k^{-1}\| \leq 1/\lambda_{BCS}$. Therefore, we need only to find an upper bound on H_k^{-1} . We know that

$$H_k^{-1} = \begin{pmatrix} I & 0 \\ -C_k^{-1} B_k & I \end{pmatrix} \begin{pmatrix} (A_k - B_k^T C_k^{-1} B_k)^{-1} & 0 \\ 0 & C_k^{-1} \end{pmatrix} \begin{pmatrix} I & -B_k^T C_k^{-1} \\ 0 & I \end{pmatrix} := \mathcal{I}_1 \mathcal{I}_2 \mathcal{I}_1^T.$$

Thus, we have $\|H_k^{-1}\| \leq \|\mathcal{I}_1\|^2 \|\mathcal{I}_2\|$. By part (i), we know

$$(4.10) \quad \|\mathcal{I}_1\| = \left\| \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ -C_k^{-1} B_k & 0 \end{pmatrix} \right\| \leq 1 + \|C_k^{-1} B_k\| \leq 1 + \beta_B / \beta_C.$$

For $\|\mathcal{I}_2\|$, we easily have that $\|\mathcal{I}_2\| \leq \max(\frac{1}{\beta_S}, \frac{1}{\beta_C})$. Using (4.10), we obtain that

$$\|H_k^{-1}\| \leq \left(1 + \frac{\beta_B}{\beta_C}\right)^2 \cdot \max\left(\frac{1}{\beta_S}, \frac{1}{\beta_C}\right) := \frac{1}{\lambda_{BCS}}.$$

This completes the proof. \square

From Theorem 4.5 we know that $\|\tilde{S}_k(\delta)\| \leq \tilde{\Upsilon}$, and the Schur complement of the bottom right matrix $\tilde{R}_k(\delta)$ in $\tilde{H}_k(\delta)$ in Algorithm 3.1 is δI . Thus, to show $\tilde{H}_k(\delta)$ is lower bounded, we need only a lower bound for $\tilde{R}_k(\delta)$, which we provide in the following lemma.

LEMMA 4.9. *Consider Problem (3.1). Under Assumption 3.1, we execute Algorithm 3.1 with $\delta \in (0, \gamma)$. Then we have that $\tilde{R}_k(\delta) \succeq \gamma I \succ 0 \quad \forall k \in [N - 1]$.*

Proof. We use a technique similar to that in the proof of Lemma 4.3 and Theorem 3.8. For any $k \in [N - 1]$ and $\delta \in (0, \gamma)$, we define $Q_i^\delta = Q_i - \delta I, i \geq k$. For any $\bar{\mathbf{q}}_k \in \mathbb{R}^{n_u}$, consider the following QDP:

$$(4.11a) \quad \min_{\substack{\mathbf{p}_{k+1:N} \\ \mathbf{q}_{k+1:N-1}}} \sum_{i=k+1}^{N-1} \begin{pmatrix} \mathbf{p}_i \\ \mathbf{q}_i \end{pmatrix}^T \begin{pmatrix} Q_i^\delta & S_i^T \\ S_i & R_i \end{pmatrix} \begin{pmatrix} \mathbf{p}_i \\ \mathbf{q}_i \end{pmatrix} + \mathbf{p}_N^T Q_N^\delta \mathbf{p}_N,$$

$$(4.11b) \quad \text{s.t. } \mathbf{p}_{i+1} = A_i \mathbf{p}_i + B_i \mathbf{q}_i \quad \text{for } k+1 \leq i \leq N-1,$$

$$(4.11c) \quad \mathbf{p}_{k+1} = B_k \bar{\mathbf{q}}_k.$$

The objective function (4.11a) can be rewritten as

$$(4.12) \quad \begin{aligned} & \sum_{i=k+1}^{N-1} \begin{pmatrix} \mathbf{p}_i \\ \mathbf{q}_i \end{pmatrix}^T \begin{pmatrix} Q_i^\delta & S_i^T \\ S_i & R_i \end{pmatrix} \begin{pmatrix} \mathbf{p}_i \\ \mathbf{q}_i \end{pmatrix} + \mathbf{p}_N^T Q_N^\delta \mathbf{p}_N \\ &= \sum_{i=k+1}^{N-1} \begin{pmatrix} \mathbf{p}_i \\ \mathbf{q}_i \end{pmatrix}^T \begin{pmatrix} Q_i & S_i^T \\ S_i & R_i \end{pmatrix} \begin{pmatrix} \mathbf{p}_i \\ \mathbf{q}_i \end{pmatrix} + \mathbf{p}_N^T Q_N \mathbf{p}_N - \delta \|\mathbf{p}_{k+1:N}\|^2. \end{aligned}$$

Let us define a special point $(\bar{\mathbf{p}}, \bar{\mathbf{q}})$ as follows: For $i \in [k - 1]$, let $(\bar{\mathbf{p}}_i, \bar{\mathbf{q}}_i) = (\mathbf{0}, \mathbf{0})$; let $(\bar{\mathbf{p}}_k, \bar{\mathbf{q}}_k) = (\mathbf{0}, \bar{\mathbf{q}}_k)$; and for $i > k$, let $\bar{\mathbf{p}}_i$ be computed recursively by using the constraint $\mathbf{p}_i = A_{i-1} \mathbf{p}_{i-1} + B_{i-1} \mathbf{q}_{i-1}$, and let $\bar{\mathbf{q}}_i = \mathbf{q}_i(\bar{\mathbf{p}}_i)$, where $\mathbf{q}_i(\cdot)$ is defined in (3.4) in Lemma 3.5 by replacing $\{Q_i\}_{i=0}^N$ by $\{Q_i^\delta\}_{i=0}^N$. Note that $\mathbf{q}_i(\cdot)$ is well defined because $Z^T H^\delta Z \succ 0$, shown in (3.10). By Assumption 3.1, we have

$$(4.13) \quad \begin{aligned} & \sum_{i=0}^{N-1} \begin{pmatrix} \bar{\mathbf{p}}_i \\ \bar{\mathbf{q}}_i \end{pmatrix}^T \begin{pmatrix} Q_i & S_i^T \\ S_i & R_i \end{pmatrix} \begin{pmatrix} \bar{\mathbf{p}}_i \\ \bar{\mathbf{q}}_i \end{pmatrix} + \bar{\mathbf{p}}_N^T Q_N \bar{\mathbf{p}}_N \\ &= \bar{\mathbf{q}}_k^T R_k \bar{\mathbf{q}}_k + \sum_{i=k+1}^{N-1} \begin{pmatrix} \bar{\mathbf{p}}_i \\ \bar{\mathbf{q}}_i \end{pmatrix}^T \begin{pmatrix} Q_i & S_i^T \\ S_i & R_i \end{pmatrix} \begin{pmatrix} \bar{\mathbf{p}}_i \\ \bar{\mathbf{q}}_i \end{pmatrix} + \bar{\mathbf{p}}_N^T Q_N \bar{\mathbf{p}}_N \\ &\geq \gamma(\|\bar{\mathbf{p}}\|^2 + \|\bar{\mathbf{q}}\|^2). \end{aligned}$$

On the other hand, under our setup for the special point $(\bar{\mathbf{p}}, \bar{\mathbf{q}})$ and from (3.5) in Lemma 3.5 (at stage $k + 1$ with $\mathbf{l}_i = \mathbf{0} \quad \forall i \geq k + 1$), we know that $(\bar{\mathbf{p}}_{k+1:N}, \bar{\mathbf{q}}_{k+1:N-1})$ has the objective value $\bar{\mathbf{p}}_{k+1}^T K_{k+1}^\delta \bar{\mathbf{p}}_{k+1}$ for Problem (4.11), where K_{k+1}^δ denotes K_{k+1} that is calculated by replacing $\{Q_i\}_{i=k+1}^N$ with $\{Q_i^\delta\}_{i=k+1}^N$. Further, from Lemma 3.6

and Claim 1 in Theorem 3.8, we have that $K_{k+1}^\delta = \bar{Q}_{k+1}^\delta(0) = \bar{Q}_{k+1}(\delta)$. Therefore, combining (4.12) and (4.13), we obtain the following:

$$\begin{aligned} \gamma(\|\bar{\mathbf{p}}\|^2 + \|\bar{\mathbf{q}}\|^2) &\stackrel{(4.13)}{\leq} \bar{\mathbf{q}}_k^T R_k \bar{\mathbf{q}}_k + \sum_{i=k+1}^{N-1} \begin{pmatrix} \bar{\mathbf{p}}_i \\ \bar{\mathbf{q}}_i \end{pmatrix}^T \begin{pmatrix} Q_i & S_i^T \\ S_i & R_i \end{pmatrix} \begin{pmatrix} \bar{\mathbf{p}}_i \\ \bar{\mathbf{q}}_i \end{pmatrix} + \bar{\mathbf{p}}_N^T Q_N \bar{\mathbf{p}}_N \\ &\stackrel{(4.12)}{=} \bar{\mathbf{q}}_k^T R_k \bar{\mathbf{q}}_k + \delta \|\bar{\mathbf{p}}_{k+1:N}\|^2 + \sum_{i=k+1}^{N-1} \begin{pmatrix} \bar{\mathbf{p}}_i \\ \bar{\mathbf{q}}_i \end{pmatrix}^T \begin{pmatrix} Q_i^\delta & S_i^T \\ S_i & R_i \end{pmatrix} \begin{pmatrix} \bar{\mathbf{p}}_i \\ \bar{\mathbf{q}}_i \end{pmatrix} + \bar{\mathbf{p}}_N^T Q_N^\delta \bar{\mathbf{p}}_N \\ &= \bar{\mathbf{q}}_k^T R_k \bar{\mathbf{q}}_k + \delta \|\bar{\mathbf{p}}_{k+1:N}\|^2 + \bar{\mathbf{p}}_{k+1}^T \bar{Q}_{k+1}(\delta) \bar{\mathbf{p}}_{k+1} \\ &= \bar{\mathbf{q}}_k^T R_k \bar{\mathbf{q}}_k + \delta \|\bar{\mathbf{p}}_{k+1:N}\|^2 + \bar{\mathbf{q}}_k^T B_k^T \bar{Q}_{k+1}(\delta) B_k \bar{\mathbf{q}}_k \\ &= \bar{\mathbf{q}}_k^T \tilde{R}_k(\delta) \bar{\mathbf{q}}_k + \delta \|\bar{\mathbf{p}}_{k+1:N}\|^2 \\ &= \bar{\mathbf{q}}_k^T \tilde{R}_k(\delta) \bar{\mathbf{q}}_k + \delta \|\bar{\mathbf{p}}\|^2. \end{aligned}$$

Thus, if $\delta \in (0, \gamma)$, we have

$$\bar{\mathbf{q}}_k^T \tilde{R}_k(\delta) \bar{\mathbf{q}}_k \geq \gamma \|\bar{\mathbf{q}}\|^2 + (\gamma - \delta) \|\bar{\mathbf{p}}\|^2 \geq \gamma \|\bar{\mathbf{q}}\|^2 \geq \gamma \|\bar{\mathbf{q}}_k\|^2.$$

This concludes the proof. \square

We summarize the lower-boundedness property of \tilde{H}_k in the next theorem.

THEOREM 4.10 (lower-bound-preserving property). *For Problem (3.1), if Assumptions 3.1 and 4.2 hold, we have, when executing Algorithm 3.1 with $\delta \in (0, \gamma)$, that $\forall k \in [N]$*

$$(4.14) \quad \tilde{H}_k(\delta) \succeq \lambda_H I \succ 0, \quad \text{where } \lambda_H = \left(\frac{\gamma}{\gamma + \tilde{\Upsilon}} \right)^2 \cdot \delta.$$

Proof. For any $k \in [N-1]$, we know

$$\tilde{H}_k(\delta) = \begin{pmatrix} \tilde{Q}_k(\delta) & \tilde{S}_k^T(\delta) \\ \tilde{S}_k(\delta) & \tilde{R}_k(\delta) \end{pmatrix} = \begin{pmatrix} \tilde{S}_k^T(\delta) (\tilde{R}_k(\delta))^{-1} \tilde{S}_k(\delta) + \delta I & \tilde{S}_k^T(\delta) \\ \tilde{S}_k(\delta) & \tilde{R}_k(\delta) \end{pmatrix}.$$

Combining Theorem 4.5 and Lemmas 4.8(ii) and 4.9 and using $\delta \in (0, \gamma)$, we have

$$\lambda_H = \left(\frac{\gamma}{\gamma + \tilde{\Upsilon}} \right)^2 \cdot \delta.$$

Note that $\lambda_H < \delta$. Thus, this bound also holds for $k = N$. \square

From Theorems 4.5 and 4.10, we obtained that Algorithm 3.1 has a uniform boundedness-preserving property: when the problem is uniform upper bounded (i.e., Assumption 4.2 holds), the convexified problem is also uniform upper bounded. Moreover, if Assumption 3.1 holds even if Hessian matrices H_k are indefinite, not only are the output Hessian matrices $\{\tilde{H}_k(\delta)\}_{k \in [N]}$ positive definite, as shown in Theorem 3.8, but they also have global lower bounds away from zero independent of N . This observation will have important consequences in sensitivity analysis in section 5. We further note that λ_H , the lower bound of $\tilde{H}(\delta)$ in (4.14), increases with both γ and δ . In section 5 we will see that the larger the lower bound, the faster the decay rate.

5. Exponential decay of NLDP sensitivities. In this section, we focus on bounding the magnitude of \mathbf{p}_k and \mathbf{q}_k , the optimal solution of Problem (3.1). From Lemma 3.4, we know that applying Algorithm 3.1 to convexify it will not change the primal solution. For a concise notation, we assume that Algorithm 3.1 has been applied and the resulting QDP is still denoted as (3.1); moreover, the quantities defining it also satisfy the various bounds proved in sections 3 and 4. Thus we will abuse the notation of H_k and D_k to truly indicate $\bar{H}_k(\delta)$ and $\bar{D}_k(\delta)$ for some $\delta \in (0, \gamma)$.

The main result of this paper is that the components of the sensitivity have exponential decay with respect to $|k-i|$ when setting $\mathbf{l} \in \Pi_i$ where $\forall i \in \{-1\} \cup [N-1]$

$$(5.1) \quad \begin{aligned} \Pi_{-1} &= \{\mathbf{l} \in \mathbb{R}^{n_x + N n_d} : \|\mathbf{l}\| = 1, \text{ supp}(\mathbf{l}) \subseteq [1, n_x]\}, \\ \Pi_i &= \{\mathbf{l} \in \mathbb{R}^{n_x + N n_d} : \|\mathbf{l}\| = 1, \text{ supp}(\mathbf{l}) \subseteq [n_x + i n_d + 1, n_x + (i+1)n_d]\}. \end{aligned}$$

Recall the perturbation path in (1.2). We know $\mathbf{l} = (\mathbf{l}_{-1}; \mathbf{l}_0; \dots; \mathbf{l}_{N-1}) \in \Pi_i$ indicates that the perturbation occurs only at stage i and, in particular, $i = -1$ indicates the perturbation at initial conditions. The key tool is the explicit solution to Problem (3.1) by combining (3.4) in Lemma 3.5 and the recursive (dynamics) constraints. Based on its closed form, we further take advantage of the boundedness property analyzed in section 4 to derive the decay rate. All matrices defined in Lemma 3.5 are computed from the QDP convexified by Algorithm 3.1.

The following lemma provides the closed form of optimal \mathbf{p} of Problem (3.1). We also mention that earlier [20] (see Proposition 2.6) Xu and Anitescu proposed similar results for the case that $D_k = 0$ and $C_k = I$. Our result is a straightforward extension.

LEMMA 5.1 (closed form of \mathbf{p}). *Define $O_k = B_k W_k^{-1} B_k^T \forall k \in [N-1]$. Then the optimal solution \mathbf{p} of Problem (3.1) at stage $k \in [N]$ is given (with notation from Lemma 3.5) by*

$$\mathbf{p}_k = \left(\prod_{i=0}^{k-1} E_i \right) \mathbf{l}_{-1} + \sum_{i=0}^{N-1} U_i^k \mathbf{l}_i + \sum_{i=0}^{N-1} F_i^k C_i \mathbf{l}_i,$$

where for $\forall i \in [N-1], k \in [N]$,

$$U_i^k = \sum_{s=0}^{i \wedge k-1} \left(\prod_{l=s+1}^{k-1} E_l \right) O_s (M_i^{s+1})^T - \left(\prod_{l=i+1}^{k-1} E_l \right) B_i W_i^{-1} D_{i2}^T \mathbf{1}_{i+1 \leq k},$$

$$F_i^k = \sum_{s=0}^{i \wedge k-1} \left(\prod_{l=s+1}^{k-1} E_l \right) O_s (V_i^{s+1})^T + \left(\prod_{l=i+1}^{k-1} E_l \right) (I - O_i K_{i+1}) \mathbf{1}_{i+1 \leq k}.$$

Proof. The argument holds for $k = 0$ trivially. For $k = 1$, using (3.4), we have

$$\begin{aligned} \mathbf{p}_1 &= A_0 \mathbf{p}_0 + B_0 \mathbf{q}_0(\mathbf{p}_0) + C_0 \mathbf{l}_0 \\ &= A_0 \mathbf{p}_0 + B_0 \left(P_0 \mathbf{p}_0 + W_0^{-1} B_0^T \sum_{i=1}^{N-1} (M_i^1)^T \mathbf{l}_i + W_0^{-1} B_0^T \sum_{i=1}^{N-1} (V_i^1)^T C_i \mathbf{l}_i \right. \\ &\quad \left. - W_0^{-1} B_0^T K_1 C_0 \mathbf{l}_0 - W_0^{-1} D_{02}^T \mathbf{l}_0 \right) + C_0 \mathbf{l}_0 \\ &= E_0 \mathbf{l}_{-1} + O_0 \sum_{i=1}^{N-1} (M_i^1)^T \mathbf{l}_i + O_0 \sum_{i=1}^{N-1} (V_i^1)^T C_i \mathbf{l}_i - B_0 W_0^{-1} D_{02}^T \mathbf{l}_0 + (I - O_0 K_1) C_0 \mathbf{l}_0, \end{aligned}$$

which is consistent with the argument at $k = 1$. Generally, we have

$$\begin{aligned}
 p_{k+1} &= A_k p_k + B_k q_k(p_k) + C_k l_k \\
 &= E_k p_k + B_k \left(W_k^{-1} B_k^T \sum_{i=k+1}^{N-1} (M_i^{k+1})^T l_i + W_k^{-1} B_k^T \sum_{i=k+1}^{N-1} (V_i^{k+1})^T C_i l_i \right. \\
 &\quad \left. - W_k^{-1} B_k^T K_{k+1} C_k l_k - W_k^{-1} D_{k2}^T l_k \right) + C_k l_k \\
 &= E_k p_k + O_k \underbrace{\sum_{i=k+1}^{N-1} (M_i^{k+1})^T l_i}_{\alpha_k^1} + O_k \underbrace{\sum_{i=k+1}^{N-1} (V_i^{k+1})^T C_i l_i}_{\alpha_k^2} - \underbrace{B_k W_k^{-1} D_{k2}^T l_k}_{\alpha_k^3} \\
 &\quad + \underbrace{(I - O_k K_{k+1}) C_k l_k}_{\alpha_k^4} \\
 &= E_k p_k + \alpha_k^1 + \alpha_k^2 + \alpha_k^3 + \alpha_k^4 = E_k E_{k-1} p_{k-1} + \sum_{j=1}^4 (\alpha_k^j + E_k \alpha_{k-1}^j) \\
 (5.2) \quad &= \dots = \left(\prod_{i=0}^k E_i \right) p_0 + \sum_{j=1}^4 \sum_{i=0}^k \left(\prod_{l=i+1}^k E_l \right) \alpha_i^j := \left(\prod_{i=0}^k E_i \right) p_0 + \sum_{j=1}^4 \mathcal{T}_j.
 \end{aligned}$$

For the terms \mathcal{T}_3 and \mathcal{T}_4 , we have

$$\begin{aligned}
 \mathcal{T}_3 &= \sum_{i=0}^k \left(\prod_{l=i+1}^k E_l \right) \alpha_i^3 = - \sum_{i=0}^k \left(\prod_{l=i+1}^k E_l \right) B_i^T W_i^{-1} D_{i2}^T l_i, \\
 (5.3) \quad \mathcal{T}_4 &= \sum_{i=0}^k \left(\prod_{l=i+1}^k E_l \right) \alpha_i^4 = \sum_{i=0}^k \left(\prod_{l=i+1}^k E_l \right) (I - O_i K_{i+1}) C_i l_i.
 \end{aligned}$$

The terms \mathcal{T}_1 and \mathcal{T}_2 can be simplified analogously. We take \mathcal{T}_1 as an example. We have

$$\begin{aligned}
 \mathcal{T}_1 &= \sum_{i=0}^k \left(\prod_{l=i+1}^k E_l \right) \alpha_i^1 = \alpha_k^1 + E_k \alpha_{k-1}^1 + \dots + E_k E_{k-1}, \dots, E_1 \alpha_0^1 \\
 &= O_k \sum_{i=k+1}^{N-1} (M_i^{k+1})^T l_i + E_k O_{k-1} \sum_{i=k}^{N-1} (M_i^k)^T l_i \\
 &\quad + \dots + E_k E_{k-1}, \dots, E_1 O_0 \sum_{i=1}^{N-1} (M_i^1)^T l_i \\
 &= E_k E_{k-1}, \dots, E_1 O_0 (M_1^1)^T l_1 + E_k E_{k-1}, \dots, E_1 O_0 (M_2^1)^T l_2 \\
 &\quad + E_k E_{k-1}, \dots, E_2 O_1 (M_2^2)^T l_2 \\
 &\quad + E_k E_{k-1}, \dots, E_1 O_0 (M_3^1)^T l_3 + E_k E_{k-1}, \dots, E_2 O_1 (M_3^2)^T l_3 \\
 &\quad + E_k E_{k-1}, \dots, E_3 O_2 (M_3^3)^T l_3 \\
 &\quad + \dots := \sum_{i=1}^{N-1} Z_i^{k+1} l_i,
 \end{aligned}$$

where Z_i^{k+1} is defined as follows: $Z_i^{k+1} = \sum_{s=0}^{i-1} (\prod_{l=s+1}^k E_l) O_s (M_i^{s+1})^T$ for $1 \leq i \leq k$; $Z_i^{k+1} = \sum_{s=0}^k (\prod_{l=s+1}^k E_l) O_s (M_i^{s+1})^T$ for $k+1 \leq i \leq N-1$. So we get

$$(5.4) \quad Z_i^{k+1} = \sum_{s=0}^{i \wedge (k+1)-1} \left(\prod_{l=s+1}^k E_l \right) O_s (M_i^{s+1})^T \text{ for } 1 \leq i \leq N-1.$$

Similarly, we have $\mathcal{T}_2 = \sum_{i=1}^{N-1} N_i^{k+1} C_i \mathbf{l}_i$, where

$$N_i^{k+1} = \sum_{s=0}^{i \wedge (k+1)-1} \left(\prod_{l=s+1}^k E_l \right) O_s (V_i^{s+1})^T \text{ for } 1 \leq i \leq N-1.$$

From (5.3) and (5.4), we get

$$(5.5) \quad \begin{aligned} \mathcal{T}_1 + \mathcal{T}_3 &= \sum_{i=1}^{N-1} Z_i^{k+1} \mathbf{l}_i - \sum_{i=0}^k \left(\prod_{l=i+1}^k E_l \right) B_i^T W_i^{-1} D_{i2}^T \mathbf{l}_i \\ &= \sum_{i=0}^{N-1} \left(Z_i^{k+1} - \left(\prod_{l=i+1}^k E_l \right) B_i^T W_i^{-1} D_{i2}^T \mathbf{1}_{i \leq k} \right) \mathbf{l}_i := \sum_{i=0}^{N-1} U_i^{k+1} \mathbf{l}_i \end{aligned}$$

and

$$(5.6) \quad \begin{aligned} \mathcal{T}_2 + \mathcal{T}_4 &= \sum_{i=1}^{N-1} N_i^{k+1} C_i \mathbf{l}_i + \sum_{i=0}^k \left(\prod_{l=i+1}^k E_l \right) (I - O_i K_{i+1}) C_i \mathbf{l}_i \\ &= \sum_{i=0}^{N-1} \left(N_i^{k+1} + \left(\prod_{l=i+1}^k E_l \right) (I - O_i K_{i+1}) \mathbf{1}_{i \leq k} \right) C_i \mathbf{l}_i := \sum_{i=0}^{N-1} F_i^{k+1} C_i \mathbf{l}_i. \end{aligned}$$

Taking (5.5) and (5.6) together and plugging back into (5.2), we have

$$\mathbf{p}_{k+1} = \left(\prod_{i=0}^k E_i \right) \mathbf{p}_0 + \sum_{i=0}^{N-1} U_i^{k+1} \mathbf{l}_i + \sum_{i=0}^{N-1} F_i^{k+1} C_i \mathbf{l}_i.$$

We replace $k+1$ with k and \mathbf{p}_0 with \mathbf{l}_{-1} (initial conditions) and conclude the proof. \square

From this lemma we see that \mathbf{p}_k , $k \in [N]$, is a linear combination of $\{\mathbf{l}_i\}_{i \in \{-1\} \cup [N-1]}$. To get the bound we seek for \mathbf{p}_k , we only need to analyze U_i^k and F_i^k . One of their common terms is the product of the closed-loop state transition matrix, $\prod_{l=i}^j E_l$, which we now bound. We first establish the relation of two consecutive cost-to-go matrices in next lemma.

LEMMA 5.2 (Riccati recursion). *Following Definition (3.3) in Lemma 3.5, we have $\forall k \in [N-1]$,*

$$K_k = E_k^T K_{k+1} E_k + (I - P_k^T) H_k \begin{pmatrix} I \\ P_k \end{pmatrix}.$$

Proof. From (3.3), we have

$$\begin{aligned}
K_k &\stackrel{(3.3)}{=} Q_k + A_k^T K_{k+1} A_k - (B_k^T K_{k+1} A_k + S_k)^T W_k^{-1} (B_k^T K_{k+1} A_k + S_k) \\
&= Q_k + (E_k - B_k P_k)^T K_{k+1} (E_k - B_k P_k) + A_k^T K_{k+1} B_k P_k + S_k^T P_k \\
&= Q_k + E_k^T K_{k+1} E_k - P_k^T B_k^T K_{k+1} E_k - E_k^T K_{k+1} B_k P_k \\
&\quad + P_k^T B_k^T K_{k+1} B_k P_k + A_k^T K_{k+1} B_k P_k + S_k^T P_k \\
&= Q_k + E_k^T K_{k+1} E_k - P_k^T B_k^T K_{k+1} E_k + S_k^T P_k \\
&= Q_k + E_k^T K_{k+1} E_k - P_k^T B_k^T K_{k+1} A_k - P_k^T B_k^T K_{k+1} B_k P_k + S_k^T P_k \\
&= Q_k + E_k^T K_{k+1} E_k - P_k^T (-W_k P_k - S_k) - P_k^T B_k^T K_{k+1} B_k P_k + S_k^T P_k \\
&= Q_k + E_k^T K_{k+1} E_k + P_k^T S_k + S_k^T P_k + P_k^T (W_k - B_k^T K_{k+1} B_k) P_k \\
&= Q_k + E_k^T K_{k+1} E_k + P_k^T S_k + S_k^T P_k + P_k^T R_k P_k \\
&= E_k^T K_{k+1} E_k + \begin{pmatrix} I & P_k^T \end{pmatrix} H_k \begin{pmatrix} I \\ P_k \end{pmatrix}.
\end{aligned}$$

Here the second, fifth, and sixth equalities are from the definition of E_k and P_k in Lemma 3.5, and the eighth equality is from the definition of W_k . \square

From (4.7), we know that $\|K_k\| \leq \tilde{\Upsilon}_{\bar{Q}}$ $\forall k \in [N]$ for some constant $\tilde{\Upsilon}_{\bar{Q}}$ independent of N . Note that we distinguish $\tilde{\Upsilon}_{\bar{Q}}$ from $\Upsilon_{\bar{Q}}$ in (4.7) deliberately since our cost-to-go matrix K_k in this section is calculated from the convexified problem, which still satisfies Assumptions 2.2 and 4.2, as proved in Theorems 4.5 and 4.10. Thus (4.7) holds for the convexified problem as well, although with a new bound. We use this result to bound the term $\prod_{l=i}^j E_l$.

LEMMA 5.3. *For any $0 \leq i \leq j \leq N-1$, we have that $\|\prod_{l=i}^j E_l\| \leq \Upsilon_E \rho^{j-i+1}$ for some Υ_E and $\rho \in (0, 1)$ independent from N .*

Proof. For any $\bar{p}_i \in \mathbb{R}^{n_x}$, we recursively define $\bar{p}_{l+1} = E_l \bar{p}_l$ for $i \leq l \leq j$. Based on Lemma 5.2, we have that

$$\begin{aligned}
(5.7) \quad \bar{p}_j^T K_j \bar{p}_j &= \bar{p}_j^T E_j^T K_{j+1} E_j \bar{p}_j + \bar{p}_j^T \begin{pmatrix} I & P_k^T \end{pmatrix} H_k \begin{pmatrix} I \\ P_k \end{pmatrix} \bar{p}_j \geq \bar{p}_j^T \begin{pmatrix} I & P_k^T \end{pmatrix} H_k \begin{pmatrix} I \\ P_k \end{pmatrix} \bar{p}_j \\
&\stackrel{(1)}{\geq} \lambda_H \left\| \begin{pmatrix} I \\ P_k \end{pmatrix} \bar{p}_j \right\|^2 \geq \lambda_H \|\bar{p}_j\|^2,
\end{aligned}$$

where inequality (1) is due to Theorem 4.10 and λ_H is defined in (4.14). Further, by Theorems 4.5, 4.10, we know the transformed QDP satisfies Assumptions 2.2 and 4.2. So we can apply Lemma 4.3, and consequently (4.7) holds for the transformed QDP. Therefore, we know that the cost-to-go matrices K_k (for transformed QDP) satisfy $\|K_k\| \leq \tilde{\Upsilon}_{\bar{Q}}$ for some constant $\tilde{\Upsilon}_{\bar{Q}}$. In particular, from the comment after the proof of Theorem 4.5 we know that the uniform upper bounds of the transformed problem do not depend on the lower bound γ , and as a result, neither does $\tilde{\Upsilon}_{\bar{Q}}$. From this observation and Lemma 5.2, we have

$$\begin{aligned}
(5.8) \quad \bar{p}_j^T K_j \bar{p}_j &= \bar{p}_{j+1}^T K_{j+1} \bar{p}_{j+1} + \bar{p}_j^T \begin{pmatrix} I & P_k^T \end{pmatrix} H_k \begin{pmatrix} I \\ P_k \end{pmatrix} \bar{p}_j \\
&\geq \bar{p}_{j+1}^T K_{j+1} \bar{p}_{j+1} + \lambda_H \|\bar{p}_j\|^2 \geq \left(1 + \frac{\lambda_H}{\tilde{\Upsilon}_{\bar{Q}}}\right) \bar{p}_{j+1}^T K_{j+1} \bar{p}_{j+1}.
\end{aligned}$$

Note that in the last inequality we use the fact that $\bar{\mathbf{p}}_j^T K_j \bar{\mathbf{p}}_j \geq \bar{\mathbf{p}}_{j+1}^T K_{j+1} \bar{\mathbf{p}}_{j+1}$. We obtain

$$\begin{aligned} \left\| \prod_{l=i}^j E_l \bar{\mathbf{p}}_l \right\|^2 &= \|\bar{\mathbf{p}}_{j+1}\|^2 \stackrel{(5.7)}{\leq} \frac{1}{\lambda_H} \bar{\mathbf{p}}_{j+1}^T K_{j+1} \bar{\mathbf{p}}_{j+1} \stackrel{(5.8)}{\leq} \frac{1}{\lambda_H(1 + \lambda_H/\tilde{\Upsilon}_{\bar{Q}})} \bar{\mathbf{p}}_j^T K_j \bar{\mathbf{p}}_j \\ &\leq \frac{1}{\lambda_H} \left(\frac{1}{1 + \lambda_H/\tilde{\Upsilon}_{\bar{Q}}} \right)^{j-i+1} \|\mathbf{p}_i^T K_i \mathbf{p}_i \leq \frac{\tilde{\Upsilon}_{\bar{Q}}}{\lambda_H} \left(\frac{1}{1 + \lambda_H/\tilde{\Upsilon}_{\bar{Q}}} \right)^{j-i+1} \|\mathbf{p}_i\|^2. \end{aligned}$$

Letting $\Upsilon_E = \sqrt{\tilde{\Upsilon}_{\bar{Q}}/\lambda_H}$ and $\rho = 1/\sqrt{1 + \lambda_H/\tilde{\Upsilon}_{\bar{Q}}}$, we conclude the proof. \square

From Lemma 3.5 we can immediately see that E_k is the closed loop matrix; thus the quantity in the preceding lemma is nothing but the relationship between \mathbf{p}_i and \mathbf{p}_j at optimality. From the proof, we have that ρ can be

$$(5.9) \quad \Upsilon_E = \frac{\sqrt{\tilde{\Upsilon}_{\bar{Q}}}}{\sqrt{\lambda_H}}, \quad \rho = \frac{\sqrt{\tilde{\Upsilon}_{\bar{Q}}}}{\sqrt{\tilde{\Upsilon}_{\bar{Q}} + \lambda_H}}.$$

Here λ_H is from Theorem 4.10 and $\tilde{\Upsilon}_{\bar{Q}}$ is the upper bound of cost-to-go matrices K_k for the transformed QP. Both of them are independent of the horizon length N . According to the discussions after the proof of Theorem 4.10, we see that when γ increases, the lower bound λ_H will also increase and Υ_E and ρ will hence decrease, which means that the larger the lower bound γ , the faster the decay. We will see this argument can also be applied for \mathbf{p}_k and \mathbf{q}_k .

Remark 5.4. Under SOSC, setting $\gamma = \lambda_{\min}(Z^T H Z) > 0$ works for finite horizons but will result in a rate depending on N . In practice, if we want to use the transformed QDP to solve the original Problem (3.1), we should make δ close to γ as much as possible because the lower bound of transformed QDP, if γ is fixed, is proportional to δ (see Theorem 4.10).

Using Lemma 5.3, we can show the decay rate for U_i^k and F_i^k , defined in Lemma 5.1.

LEMMA 5.5. *For any $i \in [N - 1]$, $k \in [N]$, we have*

$$\max(\|U_i^k\|, \|F_i^k\|) \leq \Upsilon_{uf} \rho^{|i-k|}$$

for some constant Υ_{uf} independent of N and $\rho \in (0, 1)$ from Lemma 5.3.

Proof. We first bound some terms appeared in U_i^k and F_i^k . From Assumption 4.2, (4.7), Theorem 4.5, and Lemma 4.9, we have $\forall s \in [N - 1]$

$$\begin{aligned} (5.10) \quad \|O_s\| &= \|B_s W_s^{-1} B_s^T\| \leq \Upsilon^2 \| (R_s + B_s^T K_{s+1} B_s)^{-1} \| \leq \Upsilon^2 \|R_s^{-1}\| \leq \Upsilon^2 / \gamma, \\ \|P_s\| &= \|W_s^{-1} (B_s^T K_{s+1} A_s + S_s)\| \leq \frac{1}{\gamma} (\Upsilon^2 \tilde{\Upsilon}_{\bar{Q}} + \tilde{\Upsilon}) := \Upsilon_P. \end{aligned}$$

Here all the matrices are calculated from the transformed problem as defined in Lemma 3.5, as stated in the beginning of this section. As with (5.9), we use $\tilde{\Upsilon}_{\bar{Q}}$ to denote the upper bound for K_k , to distinguish from the constant in (4.7). A direct result based on the boundedness of O_s and P_s is an upper bound for M_i^s and V_i^s . We have

$$(5.11) \quad \|M_i^s\| \leq (1 + \Upsilon_P) \Upsilon \Upsilon_E \rho^{i-s}, \quad \|V_i^s\| \leq \tilde{\Upsilon}_{\bar{Q}} \Upsilon \Upsilon_E \rho^{i-s+1} \quad \text{for } i \geq s.$$

Here Υ_E and ρ come from Lemma 5.3. So, by definition of U_i^k , F_i^k in Lemma 5.1, we can get

$$\begin{aligned} \|U_i^k\| &\leq \sum_{s=0}^{i \wedge k-1} \left(\Upsilon_E \rho^{k-s-1} \frac{\Upsilon^2}{\gamma} (1 + \Upsilon_P) \Upsilon \Upsilon_E \rho^{i-s-1} \right) + \Upsilon_E \rho^{k-i-1} \frac{\Upsilon^2}{\gamma} \mathbf{1}_{i+1 \leq k} \\ &= \frac{(1 + \Upsilon_P) \Upsilon_E^2 \Upsilon^3}{\gamma} \sum_{s=0}^{i \wedge k-1} \rho^{k+i-2s-2} + \frac{\Upsilon_E \Upsilon^2}{\gamma} \rho^{k-i-1} \mathbf{1}_{i+1 \leq k} \\ (5.12) \quad &= \frac{(1 + \Upsilon_P) \Upsilon_E^2 \Upsilon^3}{\gamma} \rho^{k+i-2} \frac{1 - (1/\rho^2)^{i \wedge k}}{1 - 1/\rho^2} + \frac{\Upsilon_E \Upsilon^2}{\gamma} \rho^{k-i-1} \mathbf{1}_{i+1 \leq k}. \end{aligned}$$

We separate (5.12) in different cases. If $i+1 \leq k$, we know

$$\begin{aligned} \|U_i^k\| &\leq \frac{(1 + \Upsilon_P) \Upsilon_E^2 \Upsilon^3}{\gamma} \rho^{k+i-2} \frac{1 - (1/\rho^2)^i}{1 - 1/\rho^2} + \frac{\Upsilon_E \Upsilon^2}{\gamma} \rho^{k-i-1} \\ &= \frac{(1 + \Upsilon_P) \Upsilon_E^2 \Upsilon^3}{\gamma} \cdot \frac{\rho^{k-i} - \rho^{k+i}}{1 - \rho^2} + \frac{\Upsilon_E \Upsilon^2}{\gamma \rho} \rho^{k-i} \\ (5.13) \quad &\leq \underbrace{\left(\frac{(1 + \Upsilon_P) \Upsilon_E^2 \Upsilon^3}{\gamma(1 - \rho^2)} + \frac{\Upsilon_E \Upsilon^2}{\gamma \rho} \right)}_{\Upsilon_u} \rho^{k-i}. \end{aligned}$$

If $k \leq i$, we know

$$(5.14) \quad \|U_i^k\| \leq \frac{(1 + \Upsilon_P) \Upsilon_E^2 \Upsilon^3}{\gamma} \rho^{k+i-2} \frac{1 - (1/\rho^2)^k}{1 - 1/\rho^2} \leq \Upsilon_u \rho^{i-k}.$$

Combining (5.13) with (5.14), we have

$$(5.15) \quad \|U_i^k\| \leq \Upsilon_u \rho^{|i-k|} \quad \forall i \in [N-1], k \in [N].$$

Analogously, for $\|F_i^k\|$, we have

$$\begin{aligned} \|F_i^k\| &\leq \sum_{s=0}^{i \wedge k-1} \left(\Upsilon_E \rho^{k-s-1} \frac{\Upsilon^2}{\gamma} \tilde{\Upsilon}_{\bar{Q}} \Upsilon_E \rho^{i-s} \right) + \Upsilon_E \rho^{k-i-1} \left(1 + \frac{\Upsilon^2}{\gamma} \tilde{\Upsilon}_{\bar{Q}} \right) \mathbf{1}_{i+1 \leq k} \\ &= \frac{\Upsilon^2 \Upsilon_E^2 \tilde{\Upsilon}_{\bar{Q}}}{\gamma} \rho^{k+i-1} \sum_{s=0}^{i \wedge k-1} \left(\frac{1}{\rho^2} \right)^s + \left(\Upsilon_E + \frac{\Upsilon^2 \Upsilon_E \tilde{\Upsilon}_{\bar{Q}}}{\gamma} \right) \rho^{k-i-1} \mathbf{1}_{i+1 \leq k} \\ &= \frac{\Upsilon^2 \Upsilon_E^2 \tilde{\Upsilon}_{\bar{Q}}}{\gamma} \rho^{k+i-1} \frac{1 - (1/\rho^2)^{i \wedge k}}{1 - 1/\rho^2} + \left(\Upsilon_E + \frac{\Upsilon^2 \Upsilon_E \tilde{\Upsilon}_{\bar{Q}}}{\gamma} \right) \rho^{k-i-1} \mathbf{1}_{i+1 \leq k}. \end{aligned}$$

If $i+1 \leq k$, we get

$$\begin{aligned} \|F_i^k\| &\leq \frac{\Upsilon^2 \Upsilon_E^2 \tilde{\Upsilon}_{\bar{Q}}}{\gamma} \rho^{k+i-1} \frac{1 - (1/\rho^2)^i}{1 - 1/\rho^2} + \left(\Upsilon_E + \frac{\Upsilon^2 \Upsilon_E \tilde{\Upsilon}_{\bar{Q}}}{\gamma} \right) \rho^{k-i-1} \\ (5.16) \quad &\leq \underbrace{\left(\frac{\Upsilon^2 \Upsilon_E^2 \tilde{\Upsilon}_{\bar{Q}} \rho}{\gamma(1 - \rho^2)} + \frac{\Upsilon_E}{\rho} + \frac{\Upsilon^2 \Upsilon_E \tilde{\Upsilon}_{\bar{Q}}}{\gamma \rho} \right)}_{\Upsilon_f} \rho^{k-i}. \end{aligned}$$

If $k \leq i$, we know

$$(5.17) \quad \|F_i^k\| \leq \frac{\Upsilon^2 \Upsilon_E^2 \tilde{\Upsilon}_{\bar{Q}}}{\gamma} \rho^{k+i-1} \frac{1 - (1/\rho^2)^k}{1 - 1/\rho^2} = \frac{\Upsilon^2 \Upsilon_E^2 \tilde{\Upsilon}_{\bar{Q}} \rho}{\gamma(1-\rho^2)} (\rho^{i-k} - \rho^{k+i}) \leq \Upsilon_f \rho^{i-k}.$$

Combining (5.16) and (5.17), we have

$$(5.18) \quad \|F_i^k\| \leq \Upsilon_f \rho^{|i-k|} \quad \forall i \in [N-1], k \in [N].$$

Thus, combining (5.15) and (5.18) and letting $\Upsilon_{uf} = \max(\Upsilon_u, \Upsilon_f)$, we complete the proof. \square

The next remark discusses the dependence of constant Υ_{uf} and ρ on γ .

Remark 5.6. Let us use “↗” (“↘”) to denote the increase (decrease) of parameters. Combining the expressions in (4.14), (5.9), (5.13), and (5.16), we have that $\frac{1}{\gamma(1-\rho^2)}, \frac{\Upsilon_E}{\rho}, \frac{1}{\gamma\rho}$ all decrease when γ increases. Therefore, we know

$$\gamma \nearrow \implies \lambda_H \nearrow \implies \rho, \Upsilon_E, \Upsilon_u, \Upsilon_f, \Upsilon_P \searrow \implies \Upsilon_{uf} \searrow.$$

All other upper bounds such as $\tilde{\Upsilon}_{\bar{Q}}, \Upsilon, \tilde{\Upsilon}$ are independent from γ .

Based on the previous analysis, we can show the exponential decay for directional derivatives $(\mathbf{p}_k, \mathbf{q}_k)$. We use their explicit form in Lemmas 3.5 and 5.1 combining with the upper bound in Lemmas 5.3 and 5.5. We summarize our main theoretical result in the next theorem.

THEOREM 5.7 (exponential decay for directional derivatives). *We consider the NLDP with only equality constraints as shown in Problem (1.1). Suppose we perturb the reference variable \mathbf{d} along direction \mathbf{l} as shown in (1.2), under Assumptions 3.1 and 4.2. Then there exist parameters $\Upsilon_{pq}, \rho \in (0, 1)$ independent from N such that $\forall k$*

$$\max(\|\mathbf{p}_k\|_2, \|\mathbf{q}_k\|_2) \leq \Upsilon_{pq} \rho^k \quad \text{if } \mathbf{l} \in \Pi_{-1}$$

and

$$\max(\|\mathbf{p}_k\|_2, \|\mathbf{q}_k\|_2) \leq \Upsilon_{pq} \rho^{|k-i|} \quad \text{if } \mathbf{l} \in \Pi_i, \forall i \in [N-1],$$

where set Π_i is defined in (5.1). Moreover, as γ increases, Υ_{pq} and ρ will decrease. Here $\mathbf{p}_k, \mathbf{q}_k$ are directional derivatives of $\mathbf{x}_k, \mathbf{u}_k$, respectively, defined as in (1.3).

Proof. When $\mathbf{l} \in \Pi_{-1} \ \forall i \in [N-1]$, by Lemma 5.1 and (3.4) in Lemma 3.5, we have

$$\mathbf{p}_k = \left(\prod_{i=0}^{k-1} E_i \right) \mathbf{l}_{-1}, \quad \mathbf{q}_k = P_k \mathbf{p}_k.$$

Using Lemma 5.3 and (5.10) in Lemma 5.5, we have

$$\|\mathbf{p}_k\| \leq \Upsilon_E \rho^k, \quad \|\mathbf{q}_k\| \leq \Upsilon_P \Upsilon_E \rho^k.$$

Without loss of generality, we assume $\Upsilon_P > 1$. So, we define $\Upsilon_{pq,1} := \Upsilon_P \Upsilon_E$ and finish the first part of proof. When $\mathbf{l} \in \Pi_i \ \forall i \in [N-1]$, by Lemma 5.1, we have

$$\mathbf{p}_k = U_i^k \mathbf{l}_i + F_i^k C_i \mathbf{l}_i \quad \forall k \in [N].$$

Then, from Lemma 5.5 and Assumption 4.2, we can get

$$(5.19) \quad \|\mathbf{p}_k\| = \|U_i^k \mathbf{l}_i + F_i^k C_i \mathbf{l}_i\| \leq \|U_i^k\| + \Upsilon \|F_i^k\| \leq (1 + \Upsilon) \Upsilon_{uf} \rho^{|k-i|} := \Upsilon_p \rho^{|k-i|}.$$

Moreover, we get the upper bound \mathbf{q}_k using (3.4) in Lemma 3.5. If $i \leq k-1$, we have

$$(5.20) \quad \|\mathbf{q}_k\| = \|P_k \mathbf{p}_k\| \leq \Upsilon_p \Upsilon_P \rho^{k-i}.$$

If $i = k$, we have

$$(5.21) \quad \|\mathbf{q}_k\| = \|P_k \mathbf{p}_k - W_k^{-1} (B_k^T K_{k+1} C_k + D_{k2}^T) \mathbf{l}_k\| \leq \Upsilon_p \Upsilon_P + \frac{\Upsilon^2 \tilde{\Upsilon}_{\bar{Q}} + \Upsilon}{\gamma},$$

where the inequality is due to the triangle inequality and $\|W_k^{-1}\| \leq \frac{1}{\gamma}$. If $i \geq k+1$, then

$$\begin{aligned} \|\mathbf{q}_k\| &= \|P_k \mathbf{p}_k + W_k^{-1} B_k^T (M_i^{k+1})^T \mathbf{l}_i + W_k^{-1} B_k^T (V_i^{k+1})^T C_i \mathbf{l}_i\| \\ &\stackrel{(5.11)}{\leq} \Upsilon_p \Upsilon_P \rho^{i-k} + \frac{\Upsilon}{\gamma} \left((1 + \Upsilon_P) \Upsilon \Upsilon_E \rho^{i-k-1} + \Upsilon \tilde{\Upsilon}_{\bar{Q}} \Upsilon_E \rho^{i-k} \right) \\ (5.22) \quad &= \left(\Upsilon_p \Upsilon_P + \frac{(1 + \Upsilon_P + \rho \tilde{\Upsilon}_{\bar{Q}}) \Upsilon_E \Upsilon^2}{\gamma \rho} \right) \rho^{i-k}. \end{aligned}$$

Combining (5.19), (5.20), (5.21), (5.22) and defining

$$(5.23) \quad \Upsilon_{pq,2} = \Upsilon_p \Upsilon_P + \frac{(1 + \Upsilon_P + \rho \tilde{\Upsilon}_{\bar{Q}}) \Upsilon_E \Upsilon^2}{\gamma \rho} + \frac{\Upsilon^2 \tilde{\Upsilon}_{\bar{Q}} + \Upsilon}{\gamma} > \Upsilon_p,$$

we get

$$\max(\|\mathbf{p}_k\|, \|\mathbf{q}_k\|) \leq \Upsilon_{pq,2} \rho^{|i-k|}.$$

This finishes the proof of the second part. To summarize, we complete the proof by defining $\Upsilon_{pq} = \max(\Upsilon_{pq,1}, \Upsilon_{pq,2})$. From Remark 5.6 and the definition of $\Upsilon_{pq,1}$ and $\Upsilon_{pq,2}$, we can see that Υ_{pq} and ρ will decrease when γ increases. \square

From the main theorem above, we see that the decay rate ρ is the same as the one in Lemma 5.3. From (4.14) and (5.9), we know

$$\rho = \frac{\sqrt{\tilde{\Upsilon}_{\bar{Q}}}}{\sqrt{\tilde{\Upsilon}_{\bar{Q}} + \left(\frac{\gamma}{\gamma + \Upsilon}\right)^2 \cdot \delta}}.$$

To clarify, although the sensitivity decay rate of an NLDP should not depend on δ , which is only a hyperparameter of the convexification algorithm, the above rate is our provable rate, and δ is treated as a parameter. Throughout the proof, we rely on setting $\delta \in (0, \gamma)$ to get uniform boundedness of the matrices. From a practical perspective, the suggested value is $\delta \approx \gamma$.

6. Numerical experiment. We conducted a numerical experiment to verify our theoretical result in Theorem 5.7. For simplicity, we chose a moderate temporal length and set $n_x = n_u = 1$. In particular, we studied the following nonconvex QDP:

$$(6.1a) \quad \min_{\mathbf{x}, \mathbf{u}} \sum_{k=0}^{N-1} \mu_1 (u_k - d_k)^2 - \mu_2 (x_k - d_k)^2 - \mu_2 x_N^2,$$

$$(6.1b) \quad \text{s.t. } x_{k+1} = u_k + f(d_k) \quad \forall k \in [N-1],$$

$$(6.1c) \quad x_0 = 0.$$

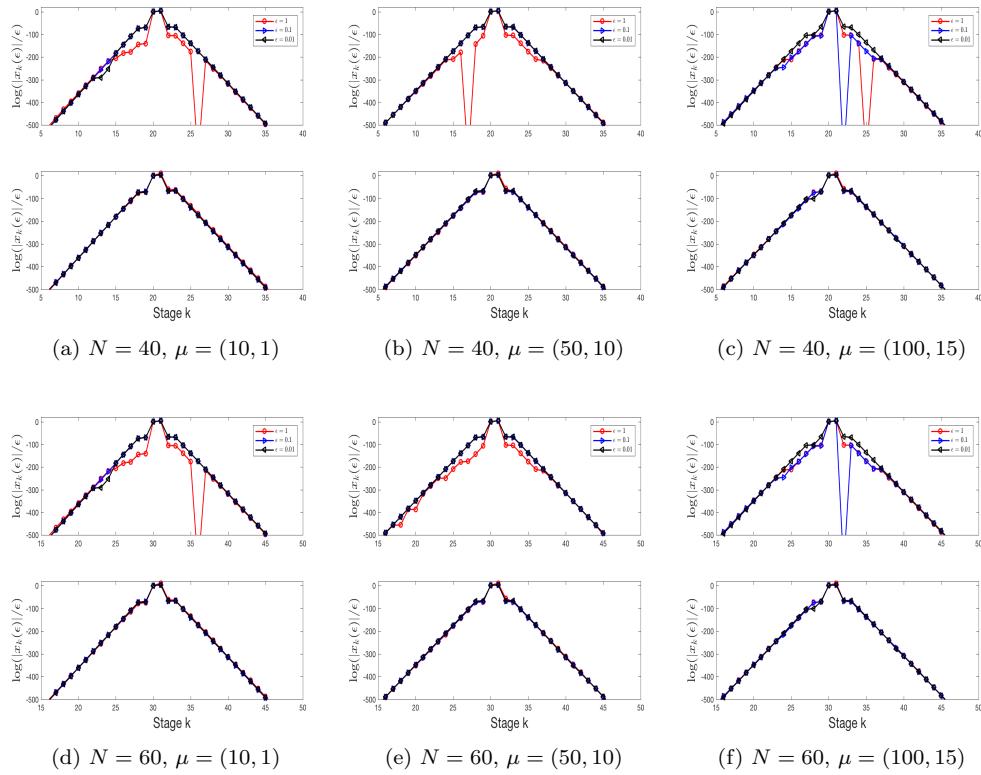


FIG. 1. Sensitivity plot. Each plot corresponds to a combination of N and $\mu = (\mu_1, \mu_2)$. Within each plot, the upper part corresponds to $f(x) = x$ while the bottom part corresponds to $f(x) = \exp(x) - 1$. We truncate the y -axis from below -500 . Except for some fluctuations, we do observe obvious linear decay trends in all settings, which is consistent with our theoretical result.

Here we assume that f is any function such that $f(0) = 0$. Note that we still have coupled KKT conditions for our toy example. Suppose the unperturbed reference variable is $\mathbf{d}^0 = (d_0, \dots, d_{N-1}) = \mathbf{0}$. Then the unperturbed problem has a unique minimizer that is $x_k = u_k = 0 \ \forall k$, provided $\mu_1 > \mu_2 > 0$. Furthermore, we can check that Assumption 3.1 holds with $\gamma = \mu_1 - \mu_2$ and that Assumption 4.2(i) holds with $\Upsilon = \max(1, \mu_1, \mu_2, f'(0))$, and therefore the controllability condition in Assumption 4.2(ii) holds with $t_k = t = \lambda_C = 1$.

Simulation setting. We set $N = 40, 60$, $\mu = (\mu_1, \mu_2) = (10, 1), (50, 10), (100, 15)$, and we let $f(x) = x$ or $\exp(x) - 1$. We perturb only \mathbf{d} at the middle stage but with three different scales: $\epsilon = 1, 0.1, 0.01$. Note that only if $\epsilon \rightarrow 0$ can we approximate the directional derivatives. Because the unperturbed solution is all zeros, we plot $\log(|x_k^*(\epsilon)|/\epsilon)$ v.s. k , for which we expect to observe a linear decay. All optimization problems are solved by using the JuMP package [5] in Julia 0.6.4.1.

Simulation results are shown in Figure 1. We see that the results all have an obvious linear decay trend, which is consistent with Theorem 5.7.

7. Conclusion and future work. In this paper, we study the sensitivity analysis for nonlinear dynamic programming with equality constraints. From the theoretical foundation for sensitivity analysis in [3] we derive a QDP whose solution is the sensi-

tivity vector but for which the quadratic objective matrix may be indefinite. Dealing with indefinite quadratic matrices has intrinsic difficulties in both practice and theory, which we address through the convexification procedure in [16]. To fit that procedure in our setup, we extend the procedure to a QDP with linear shift, and we delve deeper into this algorithm and propose a feasible interval for the regularization hyperparameter δ . Then we show that the algorithm can preserve the boundedness of the data as long as controllability holds, based on which we prove the exponential decay property of sensitivity. That is, we show that under mild Assumptions 3.1 and 4.2, when we perturb the reference variable at stage i along any direction, the magnitude of the directional derivative for the state and the control variable at stage k (i.e., \mathbf{p}_k and \mathbf{q}_k) will have exponential decay in terms of $|i - k|$. We also study the dependence of the decay rate with the minimum eigenvalue of the reduced Hessian matrix. We prove that as γ gets larger, our bounds modify to the ones of faster decay. Throughout our analysis, we assume a class of NLDP problems, which satisfy uniform SOSC, $\gamma = \lambda_{\min}(Z^T H Z) > 0$ independent of N , and we discuss the likelihood of this setup. Because of this uniform SOSC, all of the quantities in the theorems including decay rate are independent of N . By analogy with the importance of the exponential decay of the sensitivity in our previous work on convex QDP [20], we anticipate that this result will allow us prove the exponential convergence of temporal decomposition techniques for solving long-horizon NLDP, as was already demonstrated on an instance in [15], as well as allowing us to prove exponential convergence of nonlinear model predictive control, an extension of [19].

But we also point out that the analytical setup here may be useful in a general setting in the following sense. Our technique of proof relies on the observation that the sensitivity of the parameter-to-solution mapping of an NLDP, $(\mathbf{x}^*(\mathbf{d}), \mathbf{u}^*(\mathbf{d}))$, and thus, its approximate local behavior, is the same as of a nonconvex QDP (2.2). Under data boundedness, uniform SOSC, and uniform controllability, we proved that this nonconvex QDP is equivalent (in the sense of having identical parameter-to-solution mappings, which is a feature of this class of convexification approaches [16], stemming from Lemma 3.4) to a convex QDP built by the convexification Algorithm 3.1, whose data are now uniformly bounded above and whose objective matrices are bounded below. Therefore classical Riccati techniques can be applied to produce approximate solutions of the NLDP even if the horizon extends to infinity and bring with them stability and robustness-to-uncertainty features. This closes an important gap between the optimization point of view and control point of view of NLDP and may benefit other extensions and design objectives, such as the treatment of uncertainty.

Acknowledgment. We gratefully acknowledge discussions with the authors of [16].

REFERENCES

- [1] F. L. BAUER AND C. T. FIKE, *Norms and exclusion theorems*, Numer. Math., 2 (1960), pp. 137–141, <https://doi.org/10.1007/bf01386217>.
- [2] L. BIEGLER AND V. ZAVALA, *Large-scale nonlinear programming using IPOPT: An integrating framework for enterprise-wide dynamic optimization*, Comput. Chem. Eng., 33 (2009), pp. 575–582, <https://doi.org/10.1016/j.compchemeng.2008.08.006>.
- [3] J. F. BONNANS AND A. SHAPIRO, *Perturbation Analysis of Optimization Problems*, Springer, New York, 2000, <https://doi.org/10.1007/978-1-4612-1394-9>.
- [4] A. DOMAHIDI, A. U. ZGRAGGEN, M. N. ZEILINGER, M. MORARI, AND C. N. JONES, *Efficient interior point methods for multistage problems arising in receding horizon control*,

- in Proceedings of the 51st IEEE Conference on Decision and Control (CDC), IEEE, 2012, pp. 668–674, <https://doi.org/10.1109/cdc.2012.6426855>.
- [5] I. DUNNING, J. HUCHETTE, AND M. LUBIN, *JuMP: A modeling language for mathematical optimization*, SIAM Rev., 59 (2017), pp. 295–320, <https://doi.org/10.1137/15m1020575>.
 - [6] H. J. FERREAU, C. KIRCHES, A. POTSCHKA, H. G. BOCK, AND M. DIEHL, *qpOASES: A parametric active-set algorithm for quadratic programming*, Math. Program. Comput., 6 (2014), pp. 327–363, <https://doi.org/10.1007/s12532-014-0071-1>.
 - [7] G. FRISON, H. H. B. SØRENSEN, B. DAMMANN, AND J. B. JØRGENSEN, *High-performance small-scale solvers for linear model predictive control*, in Proceedings of the European Control Conference (ECC), IEEE, 2014, pp. 128–133, <https://doi.org/10.1109/ECC.2014.6862490>.
 - [8] M. HANKE, *A regularizing Levenberg–Marquardt scheme, with applications to inverse groundwater filtration problems*, Inverse Problems, 13 (1997), pp. 79–95, <https://doi.org/10.1088/0266-5611/13/1/007>.
 - [9] S. S. KEERTHI AND E. G. GILBERT, *Optimal infinite-horizon feedback laws for a general class of constrained discrete-time systems: Stability and moving-horizon approximations*, J. Optim. Theory Appl., 57 (1988), pp. 265–293, <https://doi.org/10.1007/bf00938540>.
 - [10] J. J. MORE, *The Levenberg–Marquardt algorithm: Implementation and theory*, Lecture Notes in Math., Springer, New York, 1978, pp. 105–116, <https://doi.org/10.1007/bfb0067700>.
 - [11] J. NOCEDAL AND S. J. WRIGHT, *Numerical Optimization*, 2nd ed., Springer Ser. Oper. Res. Financ. Eng., Springer, New York, 2006, <https://doi.org/10.1007/978-0-387-40065-5>.
 - [12] S. M. ROBINSON, *First order conditions for general nonlinear optimization*, SIAM J. Appl. Math., 30 (1976), pp. 597–607, <https://doi.org/10.1137/0130053>.
 - [13] S. M. ROBINSON, *Generalized equations and their solutions, part II: Applications to nonlinear programming*, in Optimality and Stability in Mathematical Programming, Springer, New York, 1982, pp. 200–221, <https://doi.org/10.1007/bfb0120989>.
 - [14] S. M. ROBINSON, *Local structure of feasible sets in nonlinear programming, part III: Stability and sensitivity*, in Nonlinear Analysis and Optimization, Math. Program. Stud., Springer, New York, 1987, pp. 45–66, <https://doi.org/10.1007/bfb0121154>.
 - [15] S. SHIN, T. FAULWASSER, M. ZANON, AND V. M. ZAVALA, *A Parallel Decomposition Scheme for Solving Long-Horizon Optimal Control Problems*, <https://arxiv.org/abs/1903.01055>, 2019.
 - [16] R. VERSCHUEREN, M. ZANON, R. QUIRYNEN, AND M. DIEHL, *A sparsity preserving convexification procedure for indefinite quadratic programs arising in direct optimal control*, SIAM J. Optim., 27 (2017), pp. 2085–2109, <https://doi.org/10.1137/16m1081543>.
 - [17] W. WAN AND L. T. BIEGLER, *Structured regularization for barrier NLP solvers*, Comput. Optim. Appl., 66 (2016), pp. 401–424, <https://doi.org/10.1007/s10589-016-9880-7>.
 - [18] A. WÄCHTER AND L. T. BIEGLER, *On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming*, Math. Program., 106 (2005), pp. 25–57, <https://doi.org/10.1007/s10107-004-0559-y>.
 - [19] Strategy for Constrained Optimal Control, preprint, ANL/MCS-P9015-1017, 2017.
 - [20] W. XU AND M. ANITESCU, *Exponentially accurate temporal decomposition for long-horizon linear-quadratic dynamic optimization*, SIAM J. Optim., 28 (2018), pp. 2541–2573, <https://doi.org/10.1137/16m1081993>.
 - [21] V. M. ZAVALA, C. D. LAIRD, AND L. T. BIEGLER, *Interior-point decomposition approaches for parallel solution of large-scale nonlinear parameter estimation problems*, Chem. Eng. Sci., 63 (2008), pp. 4834–4845, <https://doi.org/10.1016/j.ces.2007.05.022>.