On the Sample Complexity of Imitation Learning for Smoothed Model Predictive Control

by

Daniel Pfrommer

B.S., University of Pennsylvania (2022)

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2024

Authored by: Daniel Pfrommer
Department of Electrical Engineering and Computer Science
May 17, 2024

Certified by: Ali Jadbabaie
JR East Professor of Engineering
Head of Department of Civil and Environmental Engineering
Thesis Supervisor

Accepted by: Leslie A. Kolodziejski
Professor of Electrical Engineering and Computer Science
Chair, Department Committee on Graduate Students

On the Sample Complexity of Imitation Learning for Smoothed Model Predictive Control

by

Daniel Pfrommer

Submitted to the Department of Electrical Engineering and Computer Science
on May 17, 2024 in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

ABSTRACT

Recent work in imitation learning has shown that having an expert controller that is both suitably smooth and stable enables much stronger guarantees on the performance of the approximating learned controller. Constructing such smoothed expert controllers for arbitrary systems remains challenging, especially in the presence of input and state constraints. We show how such a smoothed expert can be designed for a general class of systems using a log-barrier-based relaxation of a standard Model Predictive Control (MPC) optimization problem. Our principal theoretical contributions include (1) demonstrating that the Jacobian of the barrier MPC controller can be written as a convex combination of pieces arising from the explicit MPC formulation, (2) bounding the Hessian of the barrier MPC as a function of the strength of the barrier function, and (3) presenting new results in both matrix and convex analysis for computing perturbed adjugate matrices and a tight (up to constant) lower bound on the distance of a solution with a self-concordant-barrier to the constraint set. We consider randomized smoothing as a point of comparison and show empirically that, unlike randomized smoothing, barrier MPC yields better performance while guaranteeing constraint satisfaction.

Thesis supervisor: Ali Jadbabaie
Title: JR East Professor of Engineering,
Head of Department of Civil and Environmental Engineering

Acknowledgments

I would like to thank my advisor, Professor Ali Jadbabaie, for the advice, guidance, and feedback he has provided over the last two years. This work would not have been possible without his support. I would also like to thank my undergraduate advisor, Professor Nikolai Matni, for helping me start my research career as well as my brother and father for inspiring me to continue onto a PhD.

Lastly, I would like to thank the many collaborators with whom I have had the fortune to work, including Max Simchowitz, Swati Padmanabhan, Stephen Tu, Adam Block, and Zakaria Mhammedi, among many others.

Contents

List of Figures

Chapter 1

Introduction

1.1 Motivation and Background

Imitation learning has emerged as a powerful tool in machine learning, enabling agents to learn complex behaviors by mimicking expert demonstrations acquired either from a human demonstrator or a policy computed offline [Pom88; RSB09; ACN10; RGB11]. Despite its significant success, imitation learning suffers from a compounding error problem: successive evaluations of the approximate policy accumulate error, resulting in out-of-distribution failures [Pom88]. Recent results in imitation learning [PZTM22; TRZM22; BJPST24] have identified *smoothness* (i.e. the derivative being Lipschitz) and *stability* of the expert as two key properties that circumvent this issue, thereby allowing for end-to-end performance guarantees for the final learned controller.

In this work, our focus is on enabling such guarantees when the expert being imitated is a Model Predictive Controller (MPC), a powerful class of control algorithms based on solving an optimization problem over a receding prediction horizon [AZ12]. In some cases, the solution to this multiparametric optimization problem, known as the explicit MPC representation [BMDP02], can be pre-computed. For our setup — linear systems with polytopic constraints — the optimal control input is a piecewise affine function of the state. However, the number of these pieces may grow exponentially with the time horizon and the state and input dimensions, making pre-computing and storing such a representation impractical in high dimensions.

While the approximation of MPC policies has garnered significant attention in prior works [Mor20; CSA+18; AMMHJ23], they are principally concerned with approximating the non-smooth explicit MPC with a neural network approximator and introduce schemes for enforcing stability under the learned policy. In contrast, we first construct a smoothed version of the expert and then apply stronger theoretical results derived from imitation of a smooth, stabilizing expert.

In this work, we demonstrate—both theoretically and empirically—that a log-barrier formulation of the underlying MPC optimization yields similar smoothness properties to its randomized-soothing-based counterpart while being easy to compute and satisfying the constraints of the original MPC. Our barrier MPC formulation replaces the constraints in the MPC optimization problem with "soft constraints" using the log-barrier (cf. Chapter 3). We show that, in conjunction with a black-box imitation learning algorithm, this enables end-to-end guarantees on the performance of the learned policy.

1.2 Review of Literature

There has been considerable prior work on both imitation learning and barrier MPC. We will discuss both below, and contrast them with our work.

Imitation learning attempts to mimic a policy by using demonstrations from an *expert* to learn an *imitator* policy. Previous advances in imitation learning [RB10; RGB11; LLFDG17] have generally focused on new data-collection techniques to prevent distribution shift. This approach, which includes widely-used algorithms such as DAGGER [RGB11] and DART [LLFDG17], consists of alternating between learning an approximate policy, rolling out new trajectories under the approximate policy, and then labeling the additional trajectories by evaluating the expert policy. This approach has the significant drawback that both the expert and environment must be continuously queried. Recent work, known as Taylor Series Imitation Learning (TaSIL) [PZTM22], overcomes these limitations. Instead of leveraging interactive rollouts to build theoretical guarantees, TaSIL leverages control-theoretic stability properties and expert smoothness to achieve high-probability guarantees from purely offline data. For imitating model predictive controllers, it is generally easier to query the controller and its corresponding derivatives than to collect more demonstration trajectories (as this may require physical hardware), making the offline setting more realistic. The TaSIL framework has also been extended in several other contexts to reason about multi-task imitation learning [ZKL+23], as well as stochastic policies [BJPST24]. As previously mentioned, the approximation of explicit MPC policies has previously been studied [Mor20; CSA+18; AMMHJ23], but prior work has tried to enforce the stability of the learned policy post-training, whereas we directly inherit the stability properties of the original MPC. Our presentation of the imitation learning problem and analysis in Section 3.4 builds on the TaSIL analysis framework, extended with novel smoothness bounds and known stability properties of barrier MPC.

The use of barriers to solve convex optimization problems generally has a long history, stretching back to Dikin [Dik67a] and Karmarkar [Kar84]. This class of methods, known as *interior-point methods*, solve constrained convex programs by adding a weighted *barrier* function to the cost, ensuring strict feasibility of the resulting solution. The weight on the barrier is gradually reduced until the objective is within the desired error tolerance of the original program. Interior-point methods were among the first efficient polynomial time methods and remain widely in use today [Wri05]. A notion of what constitutes a suitable barrier function, known as a *self-concordant-barrier*, was later developed by Nesterov and Nemirovskii [NN94]. We use the notion of self-concordance in many of our results (see Definition 3.0.1). For further treatment of interior point programming, we refer the reader to [Nes+18]. Our use case of barriers differs from that in interior point methods. Whereas interior point methods continuously reduce the barrier weight, we wish to fix the barrier weight away from zero and study the properties of the resulting solution. As a result, the standard duality gap analysis of self-concordant barriers, which bounds the suboptimality of the objective in terms of barrier weight, is insufficient. Namely, we are interested in both upper and lower bounding the solution approximation error (rather than the duality gap), in addition to understanding the second-order smoothness properties of how the solution varies with barrier strength.

Barrier-based methods for model predictive control [WH04; FE13; FE14; FE15a; FE16; PFDS20] have been extensively studied in the literature. However, prior work in barrier MPC has focused on satisfying particular safety or stability properties. Barrier MPC was first introduced by Wills and Heath [WH04], who show asymptotic stability of barrier MPC under an appropriately chosen ellipsoidal terminal constraint set. Additional work by Feller and Ebenbauer [FE15b] extends these results to show a more general exponential input-to-state stability property, using only properties of the terminal cost and for arbitrary constraint sets. Barrier-based methods have been used in other contexts for MPC, including as control-barrier functions for safety-critical control applications [ZZS21; WZ22]. Our work is orthogonal, yet complementary, to these directions. Our main result concerns the smoothness properties of quadratic programs (QPs) with respect to a problem parameter, where the cost and constraint set vary linearly with said parameter. In the context of barrier MPC, this sheds light on the smoothness properties of the control policy, an aspect that prior work does not explore.

1.3 Our Contributions

We separate our results into three categories: (1) theoretically optimal smoothness-to-error tradeoff bounds, (2) smoothness and approximation results for barrier MPC, and (3) general results for matrix analysis and interior point methods. We use $\mathcal{O}(\cdot)$ to denote that dependencies on other parameters have been suppressed.

Optimal Error to Smoothness Tradeoff. These results concern the smoothest possible approximation for a function with discontinuous gradients. In particular, we show that any smooth approximation of such a function which is $\mathcal{O}(\epsilon)$ close everywhere must have a smoothness constant (Lipschitzness of the gradient) of at least $\mathcal{O}(1/\epsilon)$. We note that randomized smoothing (convolution with a smoothing kernel) is optimal in this sense, but not well-suited for controls applications as it does not preserve the stability properties of the underlying controller. This motivates our central question: is it possible to smooth a function while matching the optimal error-to-smoothness tradeoff and preserving properties such as stability and constraints?

Log-Barrier MPC. Our main result is that log-barrier-based MPC is an optimal smoother along some direction and outperforms randomized smoothing for controls tasks. For a given MPC, let $\mathbf{u}^\star$ be the solution of the explicit MPC, and let $\mathbf{u}^\eta$ be the solution of the barrier-MPC formulation, with η being the weight applied to the barrier. Then, our main results for barrier MPC are as follows.

Our first result (Theorem 3.1.1) establishes that the distance of $\mathbf{u}^\eta$ from $\mathbf{u}^\star$ is bounded by $\mathcal{O}(\sqrt{\eta})$. While $\mathcal{O}(\sqrt{\eta})$ is a tight upper bound for arbitrary directions, Theorem 3.1.2 shows that there exists a direction $\mathbf{a}$ (independent of η or choice of barrier) along which the error $\mathbf{a}^\top(\mathbf{u}^\eta - \mathbf{u}^\star)$ is both upper and lower bounded by $\mathcal{O}(\sqrt{\eta + d^2} - d)$, where d is the distance of the unconstrained solution to the constraint polytope under an appropriate metric. We further show (Theorem 3.2.2) that the Jacobian of the log-barrier solution can be written as a convex combination of the Jacobian of the solution of the explicit MPC. In particular, this shows that the rate of change of $\mathbf{u}^\eta$ with respect to $\mathbf{x}_0$ is bounded independent of the weight η applied to the log barrier. We then provide in Theorem 3.3.1 a bound of $\mathcal{O}(1/(\sqrt{\eta + d^2} - d))$ on the spectral norm of the Hessian of $\mathbf{u}^\eta$ with respect to $\mathbf{x}_0$. In particular, we note that this matches the $\mathcal{O}(\sqrt{\eta + d^2} - d)$ directional error bound, meaning that barrier MPC acts locally like an optimal smoother along this direction. We demonstrate through experiments that barrier MPC outperforms randomized smoothing, affirming the merits of more sophisticated, controls-aware smoothing techniques.

Interior Point Optimization. A crucial technical component in obtaining our upper bound on the smoothness is a novel lower bound on the residual of the solution to a problem using a self-concordant barrier. Intuitively, the nature of the barrier function already suggests that the solution to a problem with a barrier function in the objective cannot be too close to the boundary of the constraint set. However, prior to our work (Theorem A.2.10), there does not seem to be an explicit lower bound quantifying this minimum residual. Furthermore, our smoothing analysis demonstrates that the lower bound we present is tight up to constants. We believe this result could have applications in optimization theory as well.

Chapter 2

Preliminaries

2.1 Notation

We first state our notation and setup that we use throughout. The notation $\|\cdot\|$ refers to the ℓ_2 norm $\|\cdot\|_2$ for vectors and, by extension, to the spectral norm (largest singular value) for square matrices and higher-order tensors. Unless transposed, all vectors are column vectors. We use uppercase boldfaced letters for matrices and lowercase boldfaced letters for vectors. For a vector $\mathbf{x}$, we use $\mathrm{Diag}(\mathbf{x})$ for the diagonal matrix with the entries of $\mathbf{x}$ along its diagonal. We use $[n]$ for the set $\{1, 2, \ldots, n\}$. Given a matrix $\mathbf{M} \in \mathbb{R}^{n\times n}$ and $\sigma \in \{0, 1\}^n$, we denote by $[\mathbf{M}]_\sigma$ the principal submatrix of $\mathbf{M}$ corresponding to the rows and columns i for which $\sigma_i = 1$. We use $\mathbf{M}_\sigma^{-1}$ to denote the matrix obtained by first computing the inverse of the matrix $[\mathbf{M}]_\sigma$ and then appropriately padding it with zeroes so that the resulting matrix $\mathbf{M}_\sigma^{-1}$ has the same size as $\mathbf{M}$. Similarly, we define $\mathrm{adj}(\mathbf{M})_\sigma$ to be the matrix obtained by first computing the adjugate (the transpose of the cofactor matrix) of $[\mathbf{M}]_\sigma$ and then appropriately padding it with zeroes so that $\mathrm{adj}(\mathbf{M})_\sigma$ has the same size as $\mathbf{M}$. Lastly, $\mathcal{O}(\cdot)$ denotes expressions where dependencies on other constants have been suppressed.

2.2 Problem Setup

We consider constrained discrete-time linear dynamical systems of the form,

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{B}\mathbf{u}_t, \quad \mathbf{x}_t \in \mathcal{X}_t, \mathbf{u}_t \in \mathcal{U}_t, \tag{2.2.1}$$

with state $\mathbf{x}_t \in \mathbb{R}^{d_x}$ and control-input $\mathbf{u}_t \in \mathbb{R}^{d_u}$ indexed by time step t, and state and input maps $\mathbf{A} \in \mathbb{R}^{d_x \times d_x}$, $\mathbf{B} \in \mathbb{R}^{d_x \times d_u}$. The sets $\mathcal{X}_t$ and $\mathcal{U}_t$, respectively, are the compact convex state and input constraint sets given by the polytopes,

$$\mathcal{X}_t := \{\mathbf{x} \in \mathbb{R}^{d_x} \mid \mathbf{A}_{x_t}\mathbf{x} \leq \mathbf{b}_x\}, \quad \mathcal{U}_t := \{\mathbf{u} \in \mathbb{R}^{d_u} \mid \mathbf{A}_{u_t}\mathbf{u} \leq \mathbf{b}_u\},$$

where $\mathbf{A}_{x_t} \in \mathbb{R}^{k_x \times d_x}$, $\mathbf{A}_{u_t} \in \mathbb{R}^{k_u \times d_x}$, $\mathbf{b}_{x_t} \in \mathbb{R}^{k_x}$, and $\mathbf{b}_{u_t} \in \mathbb{R}^{k_u}$. We use $\mathbf{A}_x \in \mathbb{R}^{(T \cdot k_x) \times d_x}, \mathbf{A}_u \in \mathbb{R}^{(T \cdot k_u) \times d_u}, \mathbf{b}_x \in \mathbb{R}^{T \cdot k_x}, \mathbf{b}_u \in \mathbb{R}^{T \cdot k_u}$ to denote the stacked constraints for the entire $x_{1:T}$ and $u_{0:T-1}$ sequences. A constraint $f(\mathbf{x}) \leq 0$ is said to be "active" at $\mathbf{y}$ if $f(\mathbf{y}) = 0$. For notational convenience, we overload ϕ to compactly denote the vector of constraint residuals for a state $\mathbf{x}$ and input $\mathbf{u}$ as well as for the sequences $\mathbf{x}_{1:T}$ and $\mathbf{u}_{0:T-1}$:

$$\phi_t(\mathbf{x}_t, \mathbf{u}_{t-1}) := \begin{bmatrix} \mathbf{b}_{x_t} - \mathbf{A}_{x_t}\mathbf{x}_t \\ \mathbf{b}_{u_{t-1}} - \mathbf{A}_{u_{t-1}}\mathbf{u}_{t-1} \end{bmatrix}, \ \phi(\mathbf{x}_0, \mathbf{u}_{0:T-1}) := \begin{bmatrix} \phi_1(\mathbf{x}_1, \mathbf{u}_0) \\ \vdots \\ \phi(\mathbf{x}_T, \mathbf{u}_{T-1}) \end{bmatrix}. \tag{2.2.2}$$

We consider deterministic state-feedback control policies which we denote by $\pi : \mathcal{X} \to \mathcal{U}$ for appropriate sets $\mathcal{X}$ and $\mathcal{U}$. In general, we use $\boldsymbol{\pi}^\star$ to refer to the expert policy and $\widehat{\pi}$ to refer to its learned approximation for the purpose of imitation learning.

In particular, our principal choice of $\boldsymbol{\pi}^\star$ in this paper is an MPC with quadratic cost and linear constraints. The MPC policy is obtained by solving the following minimization problem over future actions $\mathbf{u} := \mathbf{u}_{0:T-1}$ with quadratic cost in $\mathbf{u}$ and states $\mathbf{x} := \mathbf{x}_{1:T}$:

$$\begin{aligned} \text{minimize}_{\mathbf{u}} \quad & V(\mathbf{x}_0, \mathbf{u}) := \textstyle\sum_{t=1}^{T} \mathbf{x}_t^\top \mathbf{Q}_t \mathbf{x}_t + \sum_{t=0}^{T-1} \mathbf{u}_t^\top \mathbf{R}_t \mathbf{u}_t \\ \text{subject to} \quad & \mathbf{x}_{t+1} := \mathbf{A}\mathbf{x}_t + \mathbf{B}\mathbf{u}_t, \\ & \mathbf{x}_T \in \mathcal{X}_T, \mathbf{u}_0 \in \mathcal{U}_0, \\ & \mathbf{x}_t \in \mathcal{X}_t, \mathbf{u}_t \in \mathcal{U}_t, \ \forall t \in [T-1], \end{aligned} \tag{2.2.3}$$

where $\mathbf{Q}_t$ and $\mathbf{R}_{t-1}$ are positive definite for all $t \in [T]$. For a given state $\mathbf{x}$, the corresponding input $\boldsymbol{\pi}_{\text{mpc}}$ of the MPC is:

$$\boldsymbol{\pi}_{\text{mpc}}(\mathbf{x}) := \arg\min_{\mathbf{u}_0} \min_{\mathbf{u}_{1:T-1}} V(\mathbf{x}, \mathbf{u}_{0:T-1}), \tag{2.2.4}$$

where the minimization taken is over the feasible set defined in Problem 2.2.3. Note that $\boldsymbol{\pi}_{\text{mpc}}$ is well-defined, as $V(\mathbf{x}_0, \mathbf{u})$ has a unique global minimum in $\mathbf{u}$ for all feasible $\mathbf{x}_0$ due to strong convexity of V.

2.3 Explicit Solution to MPC

Explicit MPC [BMDP02] rewrites Problem 2.2.4 as a multi-parametric quadratic program with linear inequality constraints and solves it for every possible combination of active constraints, building an analytical solution to the control problem. Following the standard derivation (see [BMDP02, Section 4] and [BBM17, Chapter 11]), we rewrite Problem 2.2.4 as the optimization problem, in variable $\mathbf{u} := \mathbf{u}_{0:T-1} \in \mathbb{R}^{T \cdot d_u}$, as described below:

$$\begin{aligned} \text{minimize}_{\mathbf{u}} \quad & \mathcal{V}(\mathbf{x}_0, \mathbf{u}) := \tfrac{1}{2}\mathbf{u}^\top \mathbf{H}\mathbf{u} - \mathbf{x}_0^\top \mathbf{F}\mathbf{u} \\ \text{subject to} \quad & \mathbf{G}\mathbf{u} \le \mathbf{w} + \mathbf{P}\mathbf{x}_0, \end{aligned} \tag{2.3.1}$$

with matrices $\mathbf{H} \in \mathbb{R}^{T\cdot d_u \times T\cdot d_u}$, $\mathbf{F} \in \mathbb{R}^{d_x \times T\cdot d_u}$, $\mathbf{G} \in \mathbb{R}^{m \times T\cdot d_u}$, and $\mathbf{P} \in \mathbb{R}^{m\times d_x}$, and vector $\mathbf{w} \in \mathbb{R}^m$, all given by

$$\mathbf{H} = \mathbf{R}_{0:T-1} + \widehat{\mathbf{B}}^\top \mathbf{Q}_{1:T}\widehat{\mathbf{B}}, \quad \mathbf{F} = -2\widehat{\mathbf{A}}^\top \mathbf{Q}_{1:T}\widehat{\mathbf{B}},$$
$$\mathbf{G} = \begin{bmatrix} \mathbf{A}_u \\ \mathbf{A}_x\widehat{\mathbf{B}} \end{bmatrix}, \quad \mathbf{P} = \begin{bmatrix} 0 \\ -\mathbf{A}_x\widehat{\mathbf{A}} \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} \mathbf{b}_u \\ \mathbf{b}_x \end{bmatrix},$$

where $\mathbf{Q}_{1:T}$ and $\mathbf{R}_{0:T-1}$ are block diagonal with $\mathbf{Q}_1, \dots, \mathbf{Q}_T$ and $\mathbf{R}_0, \dots, \mathbf{R}_{T-1}$ on the diagonal, and $\hat{\mathbf{B}}$ and $\hat{\mathbf{A}}$ are

$$\hat{\mathbf{A}} = \begin{bmatrix} \mathbf{A} \\ \mathbf{A}^2 \\ \vdots \\ \mathbf{A}^T \end{bmatrix}, \quad \hat{\mathbf{B}} = \begin{bmatrix} \mathbf{B} & 0 & \dots & 0 \\ \mathbf{A}\mathbf{B} & \mathbf{B} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}^{T-1}\mathbf{B} & \mathbf{A}^{T-2}\mathbf{B} & \dots & \mathbf{B} \end{bmatrix}$$

so that $\mathbf{x}_{1:T} = \hat{\mathbf{A}}\mathbf{x}_0 + \hat{\mathbf{B}}\mathbf{u}$. We assume that the constraint polytope in Problem 2.3.1 contains a ball of radius r and is contained inside an origin-centered ball of radius R. As a consequence of the assumption that the problem variable is bounded in a ball of radius R, the objective in Problem 2.3.1 is L_V-Lipschitz for some constant L_V that depends on R and the operator norms of the cost matrices. We now state the solution to Problem 2.3.1 [AB09] and later (in Theorem 3.2.2) show how it appears in the smoothness of the *barrier* MPC solution.

Lemma 2.3.1. *Given a feasible initial state* $\mathbf{x}$*, let* $\boldsymbol{\sigma}(\mathbf{x}) \in \{0,1\}^m$ *denote the indicator of active constraints for Problem 2.3.1, with* $\sigma_i(\mathbf{x}) = 1$ *iff the ith constraint is active. For* $\boldsymbol{\sigma} \in \{0,1\}^m$*, let* $P_\sigma = \{\mathbf{x} | \sigma(\mathbf{x}) = \boldsymbol{\sigma}\}$ *be the set of initial states* $\mathbf{x}$ *for which the solution has active constraints determined by* $\boldsymbol{\sigma}$*. Then for* $\mathbf{x}_0 \in P_\sigma$*, the solution* $\mathbf{u}$ *of Problem 2.3.1 is expressed as* $\mathbf{u} = K_\sigma \mathbf{x}_0 + k_\sigma$*, where* K_σ *and* k_σ *are defined as:*

$$\begin{aligned} K_\sigma &:= \mathbf{H}^{-1}[\mathbf{F}^\top - \mathbf{G}^\top(\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top)_\sigma^{-1}(\mathbf{G}\mathbf{H}^{-1}\mathbf{F}^\top - \mathbf{P})], \\ k_\sigma &:= \mathbf{H}^{-1}\mathbf{G}^\top(\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top)_\sigma^\dagger \mathbf{w}. \end{aligned} \tag{2.3.2}$$

We omit the proof of Lemma 2.3.1 since it is standard (see, *e.g.*, [BMDP02, Theorem 2]).

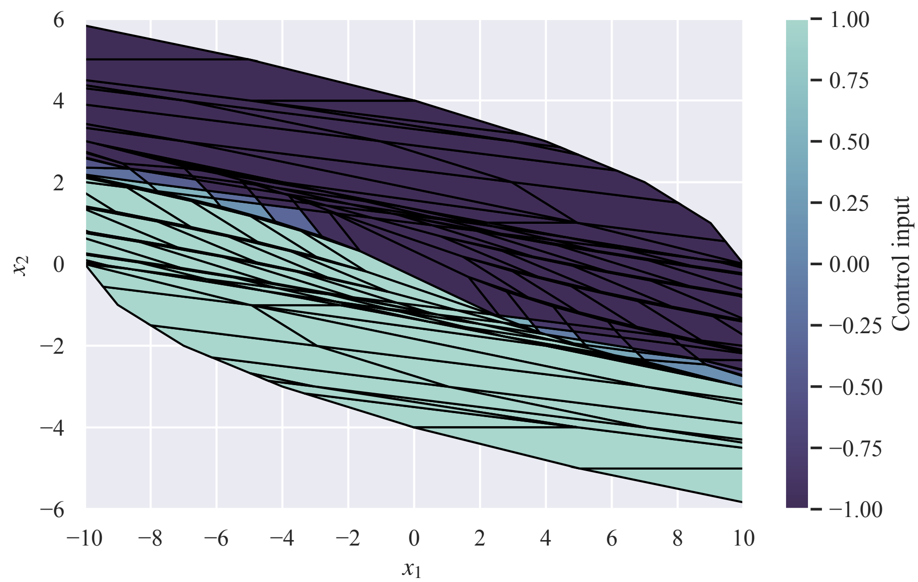

Figure 2.1: The explicit MPC controller for $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, Q = I, R = 0.01, T = 10$ with the constraints $\|x\|_\infty \leq 10, |u| \leq 1$. For this simple 2-dimensional system there are 261 K_σ. The pieces have been uniformly colored such that it is easier to see the boundaries of the individual pieces. See Figure 3.2 for a complete visualization of the policy.

Based on this lemma, one may pre-compute an efficient lookup structure mapping $\mathbf{x} \in P_\sigma$ to K_σ, k_σ. However, since every combination of active constraints may potentially yield a unique feedback law, the number of pieces to be computed may grow *exponentially* in the problem dimension or time horizon. For instance, even the simple two-dimensional toy system in Figure 2.1 has 261 pieces. In high dimensions or over long time horizons, merely enumerating all pieces of the explicit MPC may be computationally intractable.

This observation motivates us to consider a learning-based approach. In the spirit of imitation learning discussed in Chapter 1, we approximate explicit MPC using a polynomial number of sample trajectories, collected offline. We introduce this framework next.

2.4 Motivating Smoothness: Imitation Learning Frameworks

In this section, we instantiate the imitation learning framework to motivate our approach in the later sections. We specifically consider the Taylor series imitation learning framework of

[PZTM22], which gives high-probability guarantees on the quality of an approximation, as we introduce below.

2.4.1 Taylor Series Imitation Learning

We now formally introduce the setting for imitation learning considered in this paper. Suppose we are given an expert controller $\boldsymbol{\pi}^\star$, a policy class Π, a distribution of initial conditions $\mathcal{D}$, and N sample trajectories $\{\mathbf{x}_{0:K-1}^{(i)}\}_{i=1}^N$ of length K generated by $\boldsymbol{\pi}^\star$, with $\{\mathbf{x}_0^{(i)}\}_{i=1}^N$ sampled i.i.d from $\mathcal{D}$. Our goal is to find an approximate policy $\widehat{\boldsymbol{\pi}} \in \Pi$ such that given a suitably small accuracy parameter ϵ, the closed-loop states $\widehat{\mathbf{x}}_t$ and $\mathbf{x}_t^\star$ induced by $\widehat{\boldsymbol{\pi}}$ and $\boldsymbol{\pi}^\star$, respectively, satisfy, with high probability over $\mathbf{x}_0 \sim \mathcal{D}$,

$$\|\widehat{\mathbf{x}}_t - \mathbf{x}_t^\star\| \leq \epsilon, \forall t > 0.$$

This is formalized in Fact 2.4.6. To understand the sufficient conditions for such ag guarantee, we now introduce a few definitions.

We first assume through Assumption 2.4.1 that $\widehat{\boldsymbol{\pi}}$ has been chosen by a black-box supervised imitation learning algorithm which, given the input data, produces a $\widehat{\boldsymbol{\pi}} \in \Pi$ such that, with high probability over the distribution induced by $\mathcal{D}$, the policy and its Jacobian are close to the expert.

Assumption 2.4.1. *For some $\delta \in (0,1), \epsilon_0 > 0, \epsilon_1 > 0$ and given N trajectories $\{\mathbf{x}_{0:K-1}^{(i)}\}_{i=1}^{(N)}$ of length K with $\mathbf{x}_0$ sampled i.i.d. from $\mathcal{D}$ and rolled out under $\boldsymbol{\pi}^\star$, the learned approximating policy $\widehat{\boldsymbol{\pi}}$ satisfies:*

$$\mathbb{P}_{\mathbf{x}_0 \sim \mathcal{D}}\left[\sup_{k\geq 0} \|\widehat{\boldsymbol{\pi}}(\mathbf{x}_k) - \boldsymbol{\pi}^\star(\mathbf{x}_k)\| \leq \epsilon_0/N \;\; \wedge \;\; \sup_{k\geq 0}\left\|\frac{\partial \widehat{\boldsymbol{\pi}}}{\partial \mathbf{x}}(\mathbf{x}_k) - \frac{\partial \boldsymbol{\pi}^\star}{\partial \mathbf{x}}(\mathbf{x}_k)\right\| \leq \epsilon_1/N\right] \geq 1-\delta.$$

For instance, as shown in [PZTM22], Assumption 2.4.1 holds for $\widehat{\boldsymbol{\pi}}$ chosen as an empirical risk minimizer from a class of twice differentiable parametric functions with ℓ_2-bounded parameters, e.g. dense neural networks with smooth activation functions and trained with ℓ_2 weight regularization. We refer the reader to [PZTM22; TRZM22] for other possible examples of Π. Note the above definition requires only generalization on the state distribution induced by the expert, rather than the distribution induced by the learned policy, as in [AMMHJ23; CSA+18].

Next, we define a weaker variant of the standard *incremental input-to-state stability* (δISS) [VR20] and assume, in Assumption 2.4.3, that this property holds for the expert policy. We use $\mathcal{K}$ to denote the class of functions $[0, a) \to [0, \infty)$ which are zero at $x = 0$ and are monotonically increasing.

Definition 2.4.2 (Locally Incrementally Input-to-State Stabilizing Policy, cf. [PZTM22])**.** *Let $\tau > 0$ and $\gamma \in \mathcal{K}$. Consider any initial condition $\mathbf{x}_0 \in \mathcal{X}$ and bounded sequence of input perturbations $\{\mathbf{\Delta}_t\}_{t>0}$ such that $\|\mathbf{\Delta}\|_\infty < \tau$. Let $\overline{\mathbf{x}}_{t+1} = f(\overline{\mathbf{x}}_t, \pi(\overline{\mathbf{x}}_t))$, $\overline{\mathbf{x}}_0 = \mathbf{x}_0$ be the nominal trajectory, and $\mathbf{x}_{t+1} = f(\mathbf{x}_t, \pi(\mathbf{x})_t + \mathbf{\Delta}_t)$ be the perturbed trajectory. We say that π is (τ, γ)-locally-incrementally input-to-state stabilizing if,*

$$\|\mathbf{x}_t - \overline{\mathbf{x}}_t\| \leq \gamma \cdot \max_{k<t} \|\mathbf{\Delta}_k\|, \quad \forall t \geq 0.$$

Assumption 2.4.3. *The expert policy $\boldsymbol{\pi}^\star$ is (τ, γ)-locally incrementally stabilizing.*

As noted in [PZTM22], local incremental input-to-state stability (local δISS) is a much weaker criterion than even just regular incremental input-to-state stability (δISS), as local δISS uses the same initial condition and considers only bounded input perturbations. We will later show in Section 3.4 that under mild assumptions even input-to-state stabilizing policies (Definition 3.4.1) are also locally δ-ISS (Lemma 3.4.2). There is considerable prior work demonstrating that ISS holds under mild conditions for both the explicit MPC and the barrier-based MPC under consideration in this paper [PFDS20]. We refer the reader to [ZM11] for more details.

Having established some preliminaries for stability, we now move on to the smoothness property we consider.

Definition 2.4.4 (Smoothness). *We say that an MPC policy π is (L_0, L_1)-smooth if for all $\mathbf{x}, \mathbf{y} \in \mathcal{X}$,*

$$\|\pi(\mathbf{x}) - \pi(\mathbf{y})\| \leq L_0\|\mathbf{x} - \mathbf{y}\|,$$
$$\left\|\frac{\partial \pi}{\partial x}(\mathbf{x}) - \frac{\partial \pi}{\partial x}(\mathbf{y})\right\| \leq L_1\|\mathbf{x} - \mathbf{y}\|.$$

Assumption 2.4.5. *The expert policy $\boldsymbol{\pi}^\star$ and the learned policy $\widehat{\pi}$ are both (L_0, L_1)-smooth.*

At a high level, by assuming smoothness of the expert and the learned policy, we can implicitly ensure that the learned policy captures the stability of the expert in a neighborhood around the data distribution. If the expert or learned policy were to be only piecewise smooth, a transition from one piece to another in the expert, which is not replicated by the learned policy, could lead to unstable closed-loop behavior.

Having stated all the necessary assumptions, we are now ready to state below the main export of this section, guaranteeing closeness of the learned and expert policies.

Fact 2.4.6 (cf. [PZTM22], Corollary A.1). *Provided $\boldsymbol{\pi}^\star, \widehat{\pi}$ are (L_0, L_1)-smooth, (τ, γ)-locally incrementally stable, and $\widehat{\pi}$ satisfies Assumption 2.4.1 with $\delta > 0$ and N sufficiently large such that $\frac{\epsilon_0}{N} \leq \min\{\frac{1}{16\gamma^2 L_1}, \frac{1}{16\gamma}, \frac{\tau}{8\gamma}\}$ and $\frac{\epsilon_1}{N} \leq \frac{1}{4\gamma}$, then with probability $1 - \delta$ for $x_0 \sim \mathcal{D}$, we have*

$$\|\widehat{\mathbf{x}}_t - \mathbf{x}_t^\star\| \leq \frac{8\gamma\epsilon_0}{N} \quad \forall t \geq 0.$$

The upshot of this result is that to match the trajectory of the MPC policy $\boldsymbol{\pi}^\star$ with high probability, provided $\boldsymbol{\pi}^\star$ is (L_0, L_1)-smooth, we need to match the Jacobian and value of $\boldsymbol{\pi}^\star$ on *only* a fixed, finite number of points to get strong guarantees. This is in contrast to prior work such as [Mor20; KL20; CSA+18] on approximating explicit MPC, which require sampling new control inputs during training (in a reinforcement learning-like fashion) or post-training verification of the stability properties of the network.

However, as noted in Chapter 1, these strong guarantees require a smooth expert controller. We investigate two approaches for smoothing $\boldsymbol{\pi}_{\text{mpc}}$: randomized smoothing and barrier MPC.

We begin by first considering what constitutes an "optimal" smoothing approach in terms of the smallest possible Hessian norm for a given level of approximation error.

2.4.2 Optimal Smoothing

We begin by considering the properties of a general smoothing algorithm. For simplicity for this section we consider smoothing functions of the form $f : \mathbb{R} \to \mathbb{R}$, although we note that that this analysis can easily be extended to $f : \mathbb{R}^n \to \mathbb{R}^m$ by considering arbitrary paths $\mathbb{R} \to \mathbb{R}^n$ and projections $\mathbb{R}^m \to \mathbb{R}$. This motivates the following definition of a smoothing algorithm:

Definition 2.4.7 (ϵ-Smoothing Algorithm)**.** *Let $\epsilon > 0$. An ϵ-smoothing algorithm $\mathcal{S}$ for a function class $\mathcal{F}$ is a map $\mathcal{S} : \mathcal{F} \to C^1$ where C^1 is the class of functions $\mathbb{R} \to \mathbb{R}$ with continuous derivatives. Furthermore $\mathcal{S}$ satisfies,*

$$\sup_x \|\mathcal{S}(f) - f(x)\| \leq \epsilon \quad \forall f \in \mathcal{F}.$$

Analogously, we define a general smoothing algorithm as a map that can yield a smooth approximation for arbitrary choice of small ϵ.

Definition 2.4.8 (Smoothing Algorithm)**.** *A general smoothing algorithm $\mathcal{S}$ for a function class $\mathcal{F}$ is a map $\mathcal{S} : [0, a) \times \mathcal{F} \to C^1$ for $a > 0$ where $\mathcal{S}(\epsilon, \cdot)$ is an ϵ-smoothing algorithm.*

We begin by showing that for any ϵ-smoothing algorithm $\mathcal{S}$ for the class of L-Lipschitz functions (which we denote by $\mathcal{L}_L$), there exists $f \in \mathcal{L}_L$ such that the derivative of $g := \mathcal{S}(f)$ has Lipschitz constant at least $\mathcal{O}(\frac{1}{\epsilon})$. A simple example of such a function is given by the scaled absolute value function $f(x) = C|x|$.

Lemma 2.4.9. *Let $\mathcal{S}$ be any ϵ-smoothing algorithm $\mathcal{S} : \mathcal{L}_L \to C^1$ for $L, \epsilon > 0$ and let $f(x) := L|x|$, $g(x) := \mathcal{S}(f)$. Then there exists $x, y \in \mathbb{R}$ such that,*

$$|\nabla g(x) - \nabla g(y)| \geq \frac{L^2}{9\epsilon}|x - y|.$$

Proof. Consider the value of $g(x)$ at $x = -\frac{3\epsilon}{L}, 0, \frac{3\epsilon}{L}$. Since $\mathcal{S}$ is an ϵ-smoothing algorithm and $f(-\frac{3\epsilon}{L}) = f(\frac{3\epsilon}{L}) = 3\epsilon$ and $f(0) = 0$, we can conclude that $g(-\frac{3\epsilon}{L}), g(\frac{3\epsilon}{L}) \geq 2\epsilon$, $g(0) \leq \epsilon$. This implies that $g(\frac{3\epsilon}{L}) - g(0) \geq \epsilon$ and $g(0) - g(-\frac{3\epsilon}{L}) \leq -\epsilon$. By Rolle's theorem, there exists $y \in [-\frac{3\epsilon}{L}, 0], x \in [0, \frac{3\epsilon}{L}]$ such that,

$$\nabla g(y) \leq -\frac{L}{3}, \nabla g(x) \geq \frac{L}{3}.$$

Note that $|x - y| \leq \frac{6\epsilon}{L}$ or equivalently that $\frac{L^2}{9\epsilon}|x - y| \leq \frac{2L}{3}$. We conclude that,

$$|\nabla g(x) - \nabla g(y)| \geq \frac{2L}{3} \geq \frac{L^2}{9\epsilon}|x - y|.$$

This completes the proof. □

The above result suggests an inherent tradeoff between the approximation error ϵ and the Lipschitzness of the derivative of the smoothed function. Using intuition from Lemma 2.4.9, we state a more general bound for arbitrary, piecewise twice differentiable functions where the derivatives at the piece boundaries do not match.

Theorem 2.4.10. *Let $f : \mathbb{R} \to \mathbb{R}$ be a function with piecewise continuous derivative. Let $c \in \mathbb{R}$ be a point such that $\lim_{x \to c^-} \nabla f(x) = a$ and $\lim_{x \to c^+} \nabla f(x) = b$ where $a \neq b$ (i.e. the derivative is discontinuous). Then for sufficiently small ϵ and any ϵ-smoothing algorithm $\mathcal{S}$, we have that for $g := \mathcal{S}(f)$ there exists x, y such that,*

$$|\nabla g(x) - \nabla g(y)| \geq \frac{|a - b|^2}{144\epsilon}|x - y|.$$

Proof. WLOG we can shift f such that $c = 0, f(0) = 0$. Similarly, we can also subtract off $\frac{(a+b)}{2}x$ from both f and g as well as potentially perform the transformation $f(x) \to f(-x), g(x) \to g(-x)$ such that $\lim_{x \to 0^-} \nabla f(x) = -\frac{|a-b|}{2}$ and $\lim_{x \to 0^+} \nabla f(x) = \frac{|a-b|}{2}$. Let $d := \frac{|a-b|}{2}$.

Since f has piecewise continuous derivative, by definition there exists some radius $\delta > 0$ around 0 such that f is differentiable on $(-\delta, 0)$ and $(0, \delta)$ and that $\nabla f(x) \leq -\frac{d}{2}$ for $x \in (-\delta, 0)$

and $\nabla f(x) \geq \frac{d}{2}$ for $x \in (0, \delta)$. We can therefore lower bound

$$f(x) \geq -\frac{d}{2}x \quad \forall x \in (-\delta, 0), \qquad f(x) \geq \frac{d}{2}x \quad \forall x \in (0, \delta)$$

Similar to Lemma 2.4.9, we note that for $\epsilon \leq \frac{d}{6}\delta$, $f(\frac{6\epsilon}{d}), f(\frac{6\epsilon}{d}) > 3\epsilon$. Since $f(0) = 0$, $g(0) \leq \epsilon$ and therefore $g(0) - g(\frac{6\epsilon}{d}) < -\epsilon, g(\frac{6\epsilon}{d}) - g(0) \geq \epsilon$. Therefore for some $x, y \in \mathbb{R}$ such that $|x - y| \leq \frac{12\epsilon}{d}$, we have that,

$$|\nabla g(x) - \nabla g(y)| \geq \frac{d}{3} \geq \frac{d^2}{36\epsilon}|x - y| = \frac{|a - b|^2}{144\epsilon}|x - y|.$$

This completes the proof. □

The above result suggests that the derivative of an ϵ-smoothed function has a Lipschitz constant lower bounded by the square of the "discontinuity" in the derivatives times the inverse of the largest approximation error. If the smoothed function is twice differentiable, this is equivalent to a lower bound on the Hessian.

Guided by the above results, we now state our definition for an "optimal smoothing" algorithm, which is a smoothing algorithm such that the above bound is tight, up to a constant. For simplicity, we define optimal smoothing only for L-Lipschitz functions, although this definition could be extended to other function classes.

Definition 2.4.11. *A smoothing algorithm $\mathcal{S} : \mathcal{R} \times \mathcal{F} \to C^1$ for a function class $\mathcal{F}$ is worst-case optimal up to a constant if there exists $C > 0$ such that, for any sufficiently small $\epsilon > 0$, $L > 0$, and L-Lipschitz function $f \in \mathcal{L}_L \subset \mathcal{F}$, the following inequality holds with $g := \mathcal{S}(\epsilon, f)$,*

$$\|\nabla g(x) - \nabla g(y)\| \leq C\frac{L^2}{\epsilon}\|x - y\|.$$

Note that, by the above lemmas, an algorithm satisfying Definition 2.4.11 yields smoothed functions where the bound on the hessian which is at most a constant factor worse than the best possible bound for Lipschitz functions. Since the explicit model predictive control policy is always Lipschitz, for our purposes we will simply refer to smoothing algorithms satisfying Definition 2.4.11 as "optimal smoothers."

In the next two sections, we answer the question of whether an optimal smoothing algorithm can preserve the stability of an explicit MPC controller. We will show that while randomized smoothing is an optimal smoother, there exist systems for which randomized smoothing does not preserve the stability of the system. We will then introduce barrier MPC and prove that there exists a direction along which barrier MPC is an optimal smoother.

2.4.3 First Approach: Randomized Smoothing

We first consider randomized smoothing (see, *e.g.*, [DBW12]) as a baseline approach for smoothing $\boldsymbol{\pi}^\star$. Here, the imitator is learned with a loss function that randomly samples with noise drawn from a chosen probability distribution in order to smooth the policy, effectively convolving the controller with a smoothing kernel. This approach corresponds to the following controller.

Definition 2.4.12 (Randomized Smoothed MPC). *Given a control policy $\boldsymbol{\pi}_{\text{mpc}}$ of the form Problem 2.2.4, a desired zero-mean noise distribution $\mathcal{P}$, and magnitude $\sigma > 0$, the randomized-smoothing based MPC is defined as:*

$$\boldsymbol{\pi}^{\text{rs}}(\mathbf{x}) := \mathbb{E}_{\mathbf{w}\sim\mathcal{P}}[\boldsymbol{\pi}_{\text{mpc}}(\mathbf{x}+\sigma\mathbf{w})].$$

The distribution $\mathcal{P}$ in Definition 2.4.12 is usually chosen such that the following guarantees on error and smoothness hold.

Fact 2.4.13 (cf. [DBW12], Appendix E, Lemma 7-9). *Let L be the Lipschitz constant of $\boldsymbol{\pi}_{\text{mpc}}$. For each of $\mathcal{P} \in \{\text{Unif}(B_{\ell_2}(1)), \text{Unif}(B_{\ell_\infty}(1)), \mathcal{N}(0, I)\}$, there exist constants $C_0, C_1 > 0$ that* *depend on d_x such that,*

$$\|\boldsymbol{\pi}^{\text{rs}}(\mathbf{x}) - \boldsymbol{\pi}_{\text{mpc}}(\mathbf{x})\| \leq C_0\sigma \qquad \forall \mathbf{x} \in \mathcal{X},$$
$$\|\nabla\boldsymbol{\pi}^{\text{rs}}(\mathbf{x}) - \nabla\boldsymbol{\pi}^{\text{rs}}(\mathbf{y})\| \leq \frac{C_1 L^2}{\sigma}\|\mathbf{x} - \mathbf{y}\| \qquad \forall \mathbf{x}, \mathbf{y} \in \mathcal{X}.$$

This implies that randomized smoothing is an optimal smoother for the given choices of $\mathcal{P}$.

However, using randomized smoothing to obtain a smoothed policy has three key disadvantages: (1) the expectation $\mathbb{E}_{\mathbf{w}\sim\mathcal{P}}[\boldsymbol{\pi}_{\text{mpc}}(\mathbf{x} + \epsilon\mathbf{w})]$ is evaluated via sampling, which means the expert policy must be continuously re-evaluated during training in order to guarantee convergence to the smoothed policy. (2) smoothing in this manner may cause $\boldsymbol{\pi}^{\text{rs}}$ to violate state constraints. (3) simply smoothing the policy may not preserve the stability of $\boldsymbol{\pi}_{\text{mpc}}$.

Problems (2) and (3) ultimately arise as a result of randomized smoothing *oversmoothing* the underlying controller. Consider the following example:

Example 2.4.14. *Consider the system $f(x_t, u_t) = 1.5x + u_t$ and control policy given by $\pi^\star(x) = \min(\max(-x, -1), 1)$. Note that $\pi^\star$ is asymptotically exponentially stabilizing for $x_0 \in [-1, 1]$. We can see that for choices of $\mathcal{P}$ above, as $\sigma \to \infty$, $\pi^{\text{rs}}(x) \to 0$ for all x. We can conclude that π^{rs} does not stabilize the system for large σ.*

Ideally we would more aggressively smooth discontinuities that do not affect stability or constraint guarantees. This requires a smoothing technique that is aware of when more aggressive control inputs are being taken in order to more quickly stabilize versus in order to preserve constraint guarantees. As we shall show, barrier MPC is precisely such a method, and is both quick to compute while guaranteeing state constraints satisfaction. In the next section, we define barrier MPC and calculate bounds on the approximation error and smoothness of the resulting controller.

Chapter 3

Our Approach to Smoothing: Barrier MPC

Having described the guarantees obtained via randomized smoothing, we now consider smoothing via barrier functions. We begin by defining the notion of self-concordant barrier [NN94], upon which we base our analysis.

Definition 3.0.1 ([NN94])**.** *A convex, thrice differentiable function $\phi : \mathcal{Q} \mapsto \mathbb{R}$ is a ν-self-concordant barrier on an open convex set $\mathcal{Q} \subseteq \mathbb{R}^n$ if the following conditions hold.*

(i) *For all sequences $\mathbf{x}_i \in \mathcal{Q}$ converging to the boundary of $\mathcal{Q}$, we have $\lim_{i\to\infty} \phi(\mathbf{x}_i) \to \infty$.*

(ii) *For all $\mathbf{x} \in \mathcal{Q}$ and $\mathbf{h} \in \mathbb{R}^n$, we have the bound $|\mathcal{D}^3 f(\mathbf{x})[\mathbf{h}, \mathbf{h}, \mathbf{h}]| \leq 2(\mathcal{D}^2 f(\mathbf{x})[\mathbf{h}, \mathbf{h}])^{3/2}$, where $\mathcal{D}^k f(\mathbf{x})[\mathbf{h}_1, \ldots, \mathbf{h}_k]$ is the k-th derivative of f at $\mathbf{x}$ along directions $\mathbf{h}_1, \ldots, \mathbf{h}_k$,*

(iii) *For all $\mathbf{x} \in \mathcal{Q}$, we have $\nabla f(\mathbf{x})^\top (\nabla^2 f(\mathbf{x}))^{-1} \nabla f(\mathbf{x}) \leq \nu$. The parameter ν satisfies $\nu \geq 1$.*

The self-concordance property essentially says that locally, the Hessian does not change too fast — it has therefore proven extremely useful in interior-point methods to design fast algorithms for (constrained) convex programming [Dik67b; Kar84] and has also found use in model-predictive control [WH04; FE13; FE14; FE15a; FE16] in order to ensure strict feasibility of the control inputs.

In this work we consider barrier MPC as a naturally smooth alternative to randomized smoothing of Problem 2.2.4. In barrier MPC, the inequality constraints occurring in the optimal control problem are eliminated by incorporating them into the cost function via suitably scaled barrier terms. We principally consider the log-barrier, which turns a constraint $f(\mathbf{x}) \geq 0$ into the term $-\eta \log(f(\mathbf{x}))$ in the minimization objective, where η is the weight of the barrier, and is the standard choice of barrier on polytopes [NN94].

Concretely, starting from our MPC reformulation in Problem 2.3.1, the barrier MPC we work with is defined as follows.

Definition 3.0.2 (Barrier MPC)**.** *Given an MPC as in Problem 2.3.1 with associated matrices* $\mathbf{H} \in \mathbb{R}^{T\cdot d_u \times T\cdot d_u}, \mathbf{F} \in \mathbb{R}^{d_x \times T\cdot d_u}$*, and weight* $\eta > 0$*, the barrier MPC is defined by minimizing, over the input sequence* $\mathbf{u} \in \mathbb{R}^{T\cdot d_u}$*, the cost,*

$$\mathcal{V}^\eta(\mathbf{x}_0, \mathbf{u}) \;:= \tfrac{1}{2}\mathbf{u}^\top \mathbf{H}\mathbf{u} - \mathbf{x}_0^\top \mathbf{F}\mathbf{u} - \eta\left[\mathbf{1}^\top \log(\phi(\mathbf{x}_0, \mathbf{u})) - \mathbf{d}^\top \mathbf{u}\right], \tag{3.0.1}$$

where $\phi(\mathbf{x}_0, \mathbf{u}) = \mathbf{P}\mathbf{x}_0 + \mathbf{w} - \mathbf{G}\mathbf{u} \in \mathbb{R}^m$ *is the (vector) residual of constraints for* $\mathbf{x}_0$ *and* $\mathbf{u}$*, and the vector* $\mathbf{d}$ *is set to* $\mathbf{d} := \nabla_{\mathbf{u}} \sum_{i=1}^m \log(\phi_i(0, \mathbf{u}))|_{\mathbf{u}=0}$*. We denote by* $\mathbf{u}^\eta(\mathbf{x}_0)$ *the minimizer of Problem 3.0.1 for a given* $\mathbf{x}_0$ *and by* $\boldsymbol{\pi}^\eta_{\mathrm{mpc}}(\mathbf{x}) := \arg\min_{\mathbf{u}_0} \min_{\mathbf{u}_{1:T-1}} \mathcal{V}^\eta(\mathbf{x}, \mathbf{u})$ *the associated control policy.*

Some remarks are in order. First, the choice of $\mathbf{d}$ in Definition 3.0.2 is made so as to ensure that $\arg\min_{\mathbf{u}^\eta} \mathcal{V}^\eta(0, \mathbf{u}^\eta) = 0$, i.e. that $\boldsymbol{\pi}^\eta_{\mathrm{mpc}}$ satisfies $\boldsymbol{\pi}^\eta_{\mathrm{mpc}}(0) = \boldsymbol{\pi}_{\mathrm{mpc}}(0) = 0$, which is a necessary condition for the controller to be stabilizing at the origin. Further, note that $\|\mathbf{d}\|^2$ is a constant by construction, a fact that turns out to be useful in Theorem 3.3.1.

Secondly, the technical assumptions about the constraint polytope in Problem 2.3.1 containing a full-dimensional ball of radius r and being contained inside a ball of radius R around some point are both inherited by Problem 3.0.1.

The main export of this chapter is Theorem 3.3.1, which bounds the norm of the Hessian of $\mathbf{u}^\eta$ with respect to $\mathbf{x}_0$ and Theorem 3.4.4, which states our end-to-end result for barrier MPC. However, we will first begin by analyzing the approximation error.

3.1 Error Bound for Barrier MPC

To kick off our analysis of the barrier MPC, we first give the following upper bound on the distance between the optimal solution of Problem 3.0.1 and that of explicit MPC in Problem 2.3.1. Our result is based on standard techniques to analyze the sub-optimality gap in interior-point methods and crucially uses the strong convexity of our quadratic cost in Problem 3.0.1.

Theorem 3.1.1. *Suppose that $\mathbf{u}^\eta$ and $\mathbf{u}^\star$ are, respectively, the optimizers of Problem 3.0.1 and Problem 2.3.1. Then we have the following bound in terms of the barrier parameter η in Problem 3.0.1:*

$$\|\mathbf{u}^\eta - \mathbf{u}^\star\| \leq O(\sqrt{\eta}).$$

Proof. In this proof, we use $\mathcal{K}$ for the constraint polytope of Problem 2.3.1. First, Lemma A.2.3 shows that the recentered log-barrier $\phi_\mathcal{K}$ in Problem 3.0.1 is also a self-concordant barrier with some self-concordance parameter ν. Since $\mathbf{u}^\eta = \arg\min_\mathbf{u} q(\mathbf{u}) + \eta\phi_\mathcal{K}(\mathbf{u})$, where q is the quadratic cost function of Problem 3.0.1 and $\phi_\mathcal{K}$ the recentered log-barrier on $\mathcal{K}$, we have by first-order optimality:

$$\nabla q(\mathbf{u}^\eta) = -\eta \nabla \phi_\mathcal{K}(\mathbf{u}^\eta). \tag{3.1.1}$$

Denote by α the strong convexity parameter of the cost function in Problem 3.0.1 and by ν the self-concordance parameter of the barrier $\phi_\mathcal{K}$. Then,

$$\{q(\mathbf{u}^\eta) - q(\mathbf{u}^\star)\} + \frac{1}{2}\alpha\|\mathbf{u}^\eta - \mathbf{u}^\star\|^2 \leq \nabla q(\mathbf{u}^\eta)^\top(\mathbf{u}^\eta - \mathbf{u}^\star) = \eta \cdot \nabla\phi_\mathcal{K}(\mathbf{u}^\eta)^\top(\mathbf{u}^\star - \mathbf{u}^\eta) \leq \eta\nu,$$

where the first step is by α-strong convexity of q, the second step uses Equation (3.1.1), and the final step applies Fact A.2.1 at the points $\mathbf{u}^\eta$ and $\mathbf{u}^\star$. Since both $q(\mathbf{u}^\eta) - q(\mathbf{u}^\star)$ and $\frac{1}{2}\alpha\|\mathbf{u}^\eta - \mathbf{u}^\star\|^2$ are positive, we can bound the latter by $\eta\nu$. Finally, note that $\nu \geq 1$ to finish the proof.

□

We note that the above bound of $\mathcal{O}(\sqrt{\eta})$ holds for arbitrary directions and $\mathbf{x}_0$. However, provided that $\mathbf{u}^\star \neq \mathbf{K}_0\mathbf{x}_0$ (where $\mathbf{K}_0$ is the gain associated with the origin piece of the explicit MPC, i.e. the solution is not in the interior of the constraint set), we can show that there exists a direction (independent of η of the choice of barrier) along which the error scales with $\mathcal{O}(\eta/\|\mathbf{u}^\star - \mathbf{K}_0\mathbf{x}_0\|)$ for small η.

Theorem 3.1.2. *Suppose that $\mathbf{u}^\eta$ and $\mathbf{u}^\star$ are, respectively, the optimizers of Problem 3.0.1 and Problem 2.3.1. Consider the case where $\mathbf{u}^\star \neq \mathbf{K}_0\mathbf{x}_0$, for $\mathbf{K}_0 = \mathbf{H}^{-1}\mathbf{F}^\top$, i.e. $\mathbf{K}_0\mathbf{x}_0$ is the solution to the unconstrained problem. Let $\mathbf{a} = \mathbf{H}(\mathbf{u}^\star - \mathbf{K}_0\mathbf{x}_0)/\|\mathbf{H}(\mathbf{u}^\star - \mathbf{K}_0\mathbf{x}_0)\|$. Then we have the following upper and lower bounds in terms of the barrier parameter η in Problem 3.0.1:*

$$\mathbf{a}^\top(\mathbf{u}^\eta - \mathbf{u}^\star) \leq \frac{1}{2}\left(\sqrt{\frac{4m\eta}{\alpha_1} + \|\mathbf{u}^\star - \mathbf{K}_0\mathbf{x}_0\|_{\mathbf{H}}^2} - \|\mathbf{u}^\star - \mathbf{K}_0\mathbf{x}_0\|_{\mathbf{H}}\right),$$

$$\frac{r}{R}\min\left\{\frac{1}{\sqrt{m}}\left(\sqrt{\frac{\eta}{\alpha_2} + \|\mathbf{u}^\star - \mathbf{K}_0\mathbf{x}_0\|_{\mathbf{H}}^2} - \|\mathbf{u}^\star - \mathbf{K}_0\mathbf{x}_0\|_{\mathbf{H}}\right), \frac{r}{2m + 4\sqrt{m}}\right\} \leq \mathbf{a}^\top(\mathbf{u}^\eta - \mathbf{u}^\star),$$

where m is the number of constraints and $\alpha_1\mathbf{I} \preceq \mathbf{H} \preceq \alpha_2\mathbf{I}$.

Proof. This is a direct application of Theorem A.2.10, a result we prove for general ν-self-concordant barriers, using $\nu = m$, the self-concordant barrier parameter of our log barrier. □

Since by Theorem 2.4.10, the spectral norm of the hessian of $\mathbf{u}^\eta$ is lower-bounded by $\mathcal{O}(1/\epsilon)$ for $\mathcal{O}(\epsilon)$ error around a discontinuity, the upper and lower error bounds of Theorem 3.1.2 suggests the tightest-possible upper bound on the hessian that could be shown (and which would demonstrate barrier MPC is an optimal smoother along the $\mathbf{a}$ direction) is $\mathcal{O}\left(1/\left[\sqrt{\eta + \|\mathbf{u}^\star - \mathbf{K}_0\mathbf{x}_0\|^2} - \|\mathbf{u}^\star - \mathbf{K}_0\mathbf{x}_0\|\right]\right)$. This bound is later established in Theorem 3.3.1.

The above bound highlights that barrier MPC smooths in a manner which is shaped by problem constraints, unlike randomized smoothing, which generally smooths isotropically. Namely, note that the direction $\mathbf{a}$ is the direction of the gradient of the objective at $\mathbf{u}^\star$, which is a combination the directions of active constraints at $\mathbf{u}^\star$. This is indicative of barrier MPC smoothing less aggressively along directions associated with active constraints.

3.2 First-Derivative Bound for the Barrier MPC

To prove our main result on the spectral norm of the Hessian, Theorem 3.3.1, we first establish the following technical lemma bounding the first derivative of $\mathbf{u}^\eta$ with respect to $\mathbf{x}_0$. This result may be of independent interest, as it formulates the Jacobian of the log-barrier smoothed solution as a convex combination of derivatives associated with sets of active constraints from the original MPC problem. Our proof starts with the first-order optimality condition for $\mathbf{u}^\eta$ and obtains the desired simplification by applying the Sherman-Morrison-Woodbury identity (Fact A.1.3).

Lemma 3.2.1. *Consider Problem 3.0.1 with associated cost matrices* $\mathbf{H}$ *and* $\mathbf{F}$ *defined therein. Let* $\Phi := \mathrm{Diag}(\phi(\mathbf{x}_0, \mathbf{u}^\eta(x_0)))$ *be the diagonal matrix constructed via the (vector) residual* $\phi(x_0, \mathbf{u}^\eta(x_0)) = \mathbf{P}\mathbf{x}_0 + \mathbf{w} - \mathbf{G}\mathbf{u} \in \mathbb{R}^m$*. Then, the solution* $\mathbf{u}^\eta$ *to the barrier MPC in Problem 3.0.1 evolves as*

$$\frac{\partial \mathbf{u}^\eta}{\partial \mathbf{x}_0} = \mathbf{H}^{-1}[\mathbf{F}^\top - \mathbf{G}^\top(\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top + \eta^{-1}\Phi^2)^{-1}(\mathbf{G}\mathbf{H}^{-1}\mathbf{F}^\top - \mathbf{P})].$$

Proof. We first state the following first-order optimality condition for Problem 3.0.1:

$$\mathbf{H}\mathbf{u}^\eta(\mathbf{x}_0) - \mathbf{F}^\top\mathbf{x}_0 + \eta\sum_{i=1}^m\left(\frac{\mathbf{g}_i}{\phi_i(\mathbf{x}_0, \mathbf{u}^\eta(\mathbf{x}_0))} + \mathbf{d}_i\right) = 0.$$

Differentiating with respect to $\mathbf{x}_0$ and rearranging yields

$$\frac{\partial \mathbf{u}^\eta}{\partial \mathbf{x}_0} = (\mathbf{H} + \eta\mathbf{G}^\top\Phi^{-2}\mathbf{G})^{-1}(\mathbf{F}^\top + \eta\mathbf{G}^\top\Phi^{-2}\mathbf{P}). \tag{3.2.1}$$

For the rest of the proof, we introduce the notation $\mathbf{S} = \mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top + \eta^{-1}\Phi^2$. Then, we have by applying the Sherman-Morrison-Woodbury identity (Fact A.1.3) to the inverse in Equation (3.2.1) that

$$(\mathbf{H} + \eta\mathbf{G}^\top\Phi^{-2}\mathbf{G})^{-1} = \mathbf{H}^{-1} - \mathbf{H}^{-1}\mathbf{G}^\top\mathbf{S}^{-1}\mathbf{G}\mathbf{H}^{-1},$$

which simplifies our expression in Equation (3.2.1) to

$$\frac{\partial \mathbf{u}^\eta}{\partial \mathbf{x}_0} = (\mathbf{H}^{-1} - \mathbf{H}^{-1}\mathbf{G}^\top\mathbf{S}^{-1}\mathbf{G}\mathbf{H}^{-1}) \cdot (\mathbf{F}^\top + \eta\mathbf{G}^\top\Phi^{-2}\mathbf{P})$$
$$= \mathbf{H}^{-1}\mathbf{F}^\top - \mathbf{H}^{-1}\mathbf{G}^\top\mathbf{S}^{-1}\mathbf{G}\mathbf{H}^{-1}\mathbf{F}^\top + \underbrace{(\mathbf{H}^{-1} - \mathbf{H}^{-1}\mathbf{G}^\top\mathbf{S}^{-1}\mathbf{G}\mathbf{H}^{-1}) \cdot \eta\mathbf{G}^\top\Phi^{-2}\mathbf{P}}_{\text{Term 1}}. \quad (3.2.2)$$

We now show that "Term 1" may be simplified as follows.

$$(\mathbf{H}^{-1} - \mathbf{H}^{-1}\mathbf{G}^\top\mathbf{S}^{-1}\mathbf{G}\mathbf{H}^{-1}) \cdot \eta\mathbf{G}^\top\Phi^{-2}\mathbf{P} = \mathbf{H}^{-1}\mathbf{G}^\top\mathbf{S}^{-1}\mathbf{P}. \quad (3.2.3)$$

Once this is done, the claim is finished, since plugging the right-hand side from Equation (3.2.3) into "Term 1" from Equation (3.2.2) gives exactly the claimed expression in the statement of the lemma. Therefore, we now prove Equation (3.2.3). To this end, we observe that by factoring out $\mathbf{H}^{-1}\mathbf{G}^\top$ from the left and $\eta\Phi^{-2}\mathbf{P}$ from the right, we may re-write the left-hand side in Equation (3.2.3) as

$$\begin{aligned}(\mathbf{H}^{-1} - \mathbf{H}^{-1}\mathbf{G}^\top\mathbf{S}^{-1}\mathbf{G}\mathbf{H}^{-1}) \cdot \eta\mathbf{G}^\top\Phi^{-2}\mathbf{P} &= \mathbf{H}^{-1}\mathbf{G}^\top(\mathbf{I} - \mathbf{S}^{-1}\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top) \cdot \eta\Phi^{-2}\mathbf{P} \\ &= \mathbf{H}^{-1}\mathbf{G}^\top\mathbf{S}^{-1} \cdot (\mathbf{S} - \mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top) \cdot \eta\Phi^{-2} \cdot \mathbf{P} \\ &= \mathbf{H}^{-1}\mathbf{G}^\top\mathbf{S}^{-1}\mathbf{P},\end{aligned}$$

where the last step is by using our definition of $\mathbf{S}$ and cancelling $\eta^{-1}\Phi^2$ with $\eta\Phi^{-2}$. □

Equipped with Lemma 3.2.1, we are now ready to state Theorem 3.2.2, where we connect the rates of evolution of the solution Equation (2.3.2) to the constrained MPC and that of the barrier MPC (from Lemma 3.2.1). Put simply, Theorem 3.2.2 tells us that solving barrier MPC implicitly interpolates between a potentially exponential number of affine pieces from the original explicit MPC problem. This important connection helps us get a handle on the smoothness of barrier MPC as the rate at which this interpolation changes. The starting point for our proof for this result is the expression for $\frac{\partial \mathbf{u}^\eta}{\partial \mathbf{x}_0}$ from Lemma 3.2.1. To simplify this expression to be η-independent, we crucially use our linear algebraic result (Lemma A.1.8) on products of the form $\mathbf{L} \cdot \operatorname{adj}(\mathbf{L}\mathbf{L}^\top)_\sigma$ for which $\det(\mathbf{L}\mathbf{L}^\top)_\sigma = 0$.

Theorem 3.2.2. *Consider the setup in Problem 3.0.1 with associated cost matrices* $\mathbf{H}$ *and* $\mathbf{F}$*, constraint matrices* $\mathbf{P}$ *and* $\mathbf{G}$*, and barrier parameter* η*, all defined therein. We define the following quantities.*

(i) For any $\boldsymbol{\sigma} \in \{0,1\}^m$*, define the matrix* $K_{\boldsymbol{\sigma}} = \mathbf{H}^{-1}[\mathbf{F}^\top - \mathbf{G}^\top(\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top)_{\boldsymbol{\sigma}}^{-1}(\mathbf{G}\mathbf{H}^{-1}\mathbf{F}^\top - \mathbf{P})]$*, which, recall, in Lemma 2.3.1 describes the solution* $\mathbf{u}$ *to the constrained MPC.*

(ii) Recall from Problem 3.0.1 the residual $\phi(\mathbf{x}_0, \mathbf{u}) = \mathbf{P}\mathbf{x}_0 + \mathbf{w} - \mathbf{G}\mathbf{u} \in \mathbb{R}^m$*. Here, we denote the by* $\phi := \phi(\mathbf{x}_0, \mathbf{u})$*. For any* ϕ *and* $\boldsymbol{\sigma} = \{0,1\}^m$*, define the scaling factor* $h_{\boldsymbol{\sigma}} = \det([\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top]_{\boldsymbol{\sigma}}) \prod_{i=1}^m (\eta^{-1}\phi_i^2)^{1-\sigma_i}$*.*

(iii) We split the set $\boldsymbol{\sigma} \in \{0,1\}^m$ *into the following two sets:*

$$S := \left\{\boldsymbol{\sigma} \in \{0,1\}^m \mid \det([\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top]_{\boldsymbol{\sigma}}) > 0\right\},$$
$$S^\complement := \left\{\boldsymbol{\sigma} \in \{0,1\}^m \mid \det([\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top]_{\boldsymbol{\sigma}}) = 0\right\}.$$

Then the rate of evolution of the solution $\mathbf{u}^\eta$ *to the barrier MPC (in Lemma 3.2.1) is connected to that of the constrained MPC (in Equation (2.3.2)) as follows:*

$$\frac{\partial \mathbf{u}^\eta}{\partial \mathbf{x}_0} = \frac{1}{\sum_{\boldsymbol{\sigma}\in S} h_{\boldsymbol{\sigma}}} \sum_{\boldsymbol{\sigma}\in S} h_{\boldsymbol{\sigma}} \mathbf{K}_{\boldsymbol{\sigma}}.$$

Proof. Let $\Phi := \mathrm{Diag}(\phi)$. Then from Lemma 3.2.1, we have the following expression for $\frac{\partial \mathbf{u}^\eta}{\partial \mathbf{x}_0}$:

$$\frac{\partial \mathbf{u}^\eta}{\partial \mathbf{x}_0} = \mathbf{H}^{-1}[\mathbf{F}^\top - \mathbf{G}^\top(\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top + \eta^{-1}\Phi^2)^{-1}(\mathbf{G}\mathbf{H}^{-1}\mathbf{F}^\top - \mathbf{P})].$$

We now split $\mathbf{G}^\top(\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top + \eta^{-1}\Phi^2)$ above into the following two components via Lemma A.1.12.

$$\mathbf{G}^\top(\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top + \eta^{-1}\Phi^2)^{-1} = \frac{1}{h}\left(\sum_{\boldsymbol{\sigma}\in S} h_{\boldsymbol{\sigma}} \cdot \mathbf{G}^\top(\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top)_{\boldsymbol{\sigma}}^{-1} + \sum_{\boldsymbol{\sigma}\in S^\complement} c_{\boldsymbol{\sigma}} \cdot \mathbf{G}^\top \mathrm{adj}(\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top)_{\boldsymbol{\sigma}}\right),$$

where $c_{\boldsymbol{\sigma}} := \prod_{i=1}^m (\eta^{-1}\phi_i^2)^{1-\sigma_i}$ and $h := \sum_{\boldsymbol{\sigma}\in S} h_{\boldsymbol{\sigma}}$. By definition of $S^\complement$, the second sum in the preceding equation comprises those terms for which $\det(\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top)_{\boldsymbol{\sigma}} = 0$. We now invoke Lemma A.1.11, which states that $\mathbf{G}^\top \mathrm{adj}(\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top)_{\boldsymbol{\sigma}} = \mathbf{0}$ for all $\boldsymbol{\sigma} \in S^\complement$. Consequently, we may express $\frac{\partial \mathbf{u}^\eta}{\partial \mathbf{x}_0}$ in terms of only the first set of terms in the preceding equation (zeroing out the second set of terms):

$$\frac{\partial \mathbf{u}^\eta}{\partial \mathbf{x}_0} = \frac{1}{\sum_\sigma h_\sigma} \sum_{\sigma \in S} h_\sigma \mathbf{H}^{-1}[\mathbf{F}^\top - \mathbf{G}^\top(\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top)_\sigma^{-1}(\mathbf{G}\mathbf{H}^{-1}\mathbf{F}^\top - \mathbf{P})] = \frac{1}{\sum_\sigma h_\sigma} \sum_{\sigma \in S} h_\sigma \mathbf{K}_\sigma.$$

where we plugged in Equation (2.3.2) in the final step, thus concluding the proof. □

The above theorem immediately implies that $\left\|\frac{\partial \mathbf{u}^\eta}{\partial \mathbf{x}_0}\right\|$ is bounded from above as follows. Notably, this independence of the Lipschitz constant of $\mathbf{u}^\eta$ from η demonstrates that the log-barrier does not worsen the Lipschitz constant of the controller, only changing the interpolation between the different pieces.

Corollary 3.2.3. *In the setting of Theorem 3.2.2, we have,*

$$\left\|\frac{\partial \mathbf{u}^\eta}{\partial \mathbf{x}_0}\right\| \leq L := \max_{\sigma \in S} \|K_\sigma\|.$$

Proof. From Theorem 3.2.2, we can conclude that $\frac{\partial \mathbf{u}^\eta}{\partial x_0}$ lies in the convex hull of $\{K_\sigma\}_{\sigma \in S}$, and note that $|S| < \infty$. □

3.3 Our Main Result: Smoothness Bound for the Barrier MPC

We are now ready to state our main result, which effectively shows that $\mathbf{u}^\eta$ (and hence $\boldsymbol{\pi}^\eta_{\text{mpc}}$) satisfies the conditions of Assumption 2.4.5. Our proof of Theorem 3.3.1 starts with Lemma 3.2.1 and computes another derivative. To get an upper bound on the operator norm of the Hessian so obtained, our proof then crucially hinges on Lemma A.2.8 and Theorem A.2.10, which provide explicit lower bounds on residuals when minimizing a quadratic cost plus a self-concordant barrier over a polytope, a result we believe to be of independent interest to the optimization community.

Theorem 3.3.1. *Consider the setting of Problem 3.0.1 with associated cost matrices* $\mathbf{H}$ *and* $\mathbf{F}$*, constraint matrices* $\mathbf{P}$ *and* $\mathbf{G}$*, barrier parameter* η*, number of constraints* m*, the recentering* *vector* $\mathbf{d}$*, and the solution* $\mathbf{u}^\eta$*, all defined therein. We define the following quantities.*

(i) Denote by L *the Lipschitz constant of* $\mathbf{u}^\eta$ *from Corollary 3.2.3.*

(ii) We split the set $\boldsymbol{\sigma} \in \{0,1\}^m$ *into the following two sets:*

$$S := \left\{\boldsymbol{\sigma} \in \{0,1\}^m \mid \det([\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top]_\sigma) > 0\right\},$$
$$S^\complement := \left\{\boldsymbol{\sigma} \in \{0,1\}^m \mid \det([\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top]_\sigma) = 0\right\}.$$

(iii) For $\boldsymbol{\sigma} \in S$ *define the parameter* $C := \max_{\boldsymbol{\sigma}\in S} \|2\mathbf{H}^{-1}\mathbf{G}^\top(\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top)_{\boldsymbol{\sigma}}^\dagger\|$*.*

(iv) Denote by r*,* R*, and* L_V *the inner radius, outer radius, and Lipschitz constant, respectively, associated with Problem 3.0.1.*

(v) Define the residual lower bound,

$$res_{l.b.} = \min\left\{\frac{1}{\sqrt{\nu}}\left(\sqrt{\frac{\eta}{\lambda_{\max}(\mathbf{H})} + \|\mathbf{u}^\star - \mathbf{K}_0\mathbf{x}_0\|_{\mathbf{H}}^2} - \|\mathbf{u}^\star - \mathbf{K}_0\mathbf{x}_0\|_{\mathbf{H}}\right), \frac{r}{2\nu + 4\sqrt{\nu}}\right\}, \tag{3.3.1}$$

with $\nu = 20(m + R^2\|\mathbf{d}\|^2)$ *and* $\mathbf{K}_0\mathbf{x}_0 := \mathbf{H}^{-1}\mathbf{F}^\top\mathbf{x}_0$*, the solution to the unconstrained minimization of the quadratic objective. We denote* $\|\mathbf{u}\|_{\mathbf{H}} := \sqrt{\mathbf{u}^\top\mathbf{H}\mathbf{u}}$*.*

Then, the Hessian of $\mathbf{u}^\eta$ *with respect to* $\mathbf{x}_0$ *is bounded by:*

$$\left\|\frac{\partial^2\mathbf{u}^\eta}{\partial\mathbf{x}_0^2}\right\| \le \frac{C}{res_{l.b.}}(\|\mathbf{P}\| + \|\mathbf{G}\|L)^2.$$

where $\|\cdot\|$ *denotes the spectral norm of the* $\frac{\partial^2\mathbf{u}^\eta}{\partial\mathbf{x}_0^2}$ *third-order tensor.*

Proof. From Lemma 3.2.1, we have the following expression for $\frac{\partial\mathbf{u}^\eta}{\partial\mathbf{x}_0}$ evaluated at a particular $\mathbf{x}_0$:

$$\frac{\partial\mathbf{u}^\eta}{\partial\mathbf{x}_0}(\mathbf{x}_0) = \mathbf{H}^{-1}[\mathbf{F}^\top - \mathbf{G}^\top(\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top + \eta^{-1}\Phi(\mathbf{x}_0)^2)^{-1}(\mathbf{G}\mathbf{H}^{-1}\mathbf{F}^\top - \mathbf{P})],$$

where $\Phi(\mathbf{x}_0) := \mathrm{Diag}(\mathbf{P}\mathbf{x}_0 - \mathbf{w} + \mathbf{G}\mathbf{u}^\eta(\mathbf{x}_0))$. Let $\mathbf{y} \in \mathbb{R}^{d_x}$ be an arbitrary unit-norm vector, and define the univariate function

$$\mathbf{M}(t) := \mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top + \eta^{-1}\Phi(t)^2$$

where we overload Φ as $\Phi(t) := \mathrm{Diag}(\mathbf{P}(\mathbf{x}_0 + t\mathbf{y}) - \mathbf{w} + \mathbf{G}\mathbf{u}^\eta(\mathbf{x}_0 + t\mathbf{y}))$, the residual along the path $t \mapsto \mathbf{x}_0 + t\mathbf{y}$. We therefore have the following expression for $\frac{\partial \mathbf{u}^\eta}{\partial \mathbf{x}_0}$ evaluated at $\mathbf{x}_0 + t\mathbf{y}$:

$$\frac{\partial \mathbf{u}^\eta}{\partial \mathbf{x}_0}(\mathbf{x}_0 + t\mathbf{y}) = \mathbf{H}^{-1}[\mathbf{F}^\top - \mathbf{G}^\top \mathbf{M}(t)^{-1}(\mathbf{G}\mathbf{H}^{-1}\mathbf{F}^\top - \mathbf{P})].$$

Then by differentiating $\mathbf{M}(t)^{-1}$ and applying the chain rule, we get,

$$\begin{aligned}
\frac{d}{dt}\left(\frac{\partial \mathbf{u}^\eta}{\partial \mathbf{x}_0}(\mathbf{x}_0 + t\mathbf{y})\right) &= \mathbf{H}^{-1}\mathbf{G}^\top \mathbf{M}(t)^{-1}\frac{d\mathbf{M}(t)}{dt}\mathbf{M}(t)^{-1}(\mathbf{G}\mathbf{H}^{-1}\mathbf{F}^\top - \mathbf{P}) \\
&= 2\mathbf{H}^{-1}\mathbf{G}^\top \mathbf{M}(t)^{-1}\left(\frac{d\Phi(t)}{dt}\eta^{-1}\Phi(t)\right)\mathbf{M}(t)^{-1}(\mathbf{G}\mathbf{H}^{-1}\mathbf{F}^\top - \mathbf{P}) \\
&= 2\mathbf{H}^{-1}\mathbf{G}^\top \mathbf{M}(t)^{-1}\frac{d\Phi(t)}{dt}(\eta\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top\Phi^{-1}(t) + \Phi(t))^{-1}(\mathbf{G}\mathbf{H}^{-1}\mathbf{F}^\top - \mathbf{P}) \\
&= 2\mathbf{H}^{-1}\mathbf{G}^\top \mathbf{M}(t)^{-1}\frac{d\Phi(t)}{dt} \\
&\quad \cdot (\eta\Phi(t)^{-1}\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top\Phi(t)^{-1} + \mathbf{I})^{-1}\Phi(t)^{-1}(\mathbf{G}\mathbf{H}^{-1}\mathbf{F}^\top - \mathbf{P}),
\end{aligned}$$

where the third and fourth steps factor out $\Phi(t)$ from the right and left, respectively. We now bound groups of terms of the product on the right-hand side and then finish the bound by submultiplicativity of the spectral norm. First, since $\mathbf{M}(t)$ is a sum of a square matrix and a positive diagonal matrix, we may apply Lemma A.1.12 to express $\mathbf{G}^\top\mathbf{M}(t)^{-1}$ as follows with appropriate h_σ and c_σ:

$$\mathbf{G}^\top\mathbf{M}(t)^{-1} = \frac{1}{\sum_{\sigma\in S} h_\sigma}\left(\sum_{\sigma\in S} h_\sigma \mathbf{G}^\top(\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top)^\dagger_\sigma + \sum_{\sigma\in S^\complement} c_\sigma \mathbf{G}^\top \mathrm{adj}(\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top)_\sigma\right). \quad (3.3.2)$$

Now note that for $\boldsymbol{\sigma} \in S^\complement$, we have $\det(\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top)_\sigma = 0$. We may then invoke Lemma A.1.11 to infer that for $\boldsymbol{\sigma} \in S^\complement$, we have $\mathbf{G}^\top \mathrm{adj}(\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top)_\sigma = \mathbf{0}$. As a result, the second term on the right-hand side of Equation (3.3.2) vanishes, thereby affording us the following simplification:

$$\begin{aligned}
\|2\mathbf{H}^{-1}\mathbf{G}^\top\mathbf{M}(t)^{-1}\| &= \left\|\frac{1}{\sum_{\sigma\in S} h_\sigma}\sum_{\sigma\in S} 2h_\sigma \mathbf{H}^{-1}\mathbf{G}^\top(\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top)^\dagger_\sigma\right\| \\
&\leq C := \max_{\sigma\in S}\|2\mathbf{H}^{-1}\mathbf{G}^\top(\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top)^\dagger_\sigma\|,
\end{aligned}$$

where the second equality follows via Hölder's inequality. Next, from the definition of Φ, we have that $\frac{d\Phi}{dt} = \mathbf{P} + \mathbf{G}\left(\frac{\partial \mathbf{u}^\eta}{\partial \mathbf{x}_0}(\mathbf{x}_0 + t\mathbf{y})\mathbf{y}\right)$. By the triangle inequality, the Lipschitzness L of $\mathbf{u}^\eta$ (from Corollary 3.2.3), and the fact that $\mathbf{y}$ is unit norm, we have

$$\left\|\frac{d\Phi}{dt}\right\| \leq \|\mathbf{P}\| + \|\mathbf{G}\| \left\|\frac{\partial \mathbf{u}^\eta}{\partial \mathbf{x}_0}\right\| \leq \|\mathbf{P}\| + \|\mathbf{G}\| L.$$

To bound $\|(\eta\Phi^{-1}\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top\Phi^{-1}+\mathbf{I})^{-1}\Phi^{-1}\|$, we first note that because $\eta\Phi^{-1}\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top\Phi^{-1} \succeq \mathbf{0}$, we have $(\eta\Phi^{-1}\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top\Phi^{-1} + \mathbf{I})^{-1} \preceq \mathbf{I}$, which in turn implies that $\|(\eta\Phi^{-1}\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top\Phi^{-1} + \mathbf{I})^{-1}\| \leq 1$. Then, by submultiplicativity of the spectral norm, we have

$$\|(\eta\Phi^{-1}\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top\Phi^{-1} + \mathbf{I})^{-1}\Phi^{-1}\| \leq \|\Phi^{-1}\| \leq \frac{1}{\min_{i\in[m]} \phi_i}.$$

We may then plug in the lower bound on $\min_{i\in[m]} \phi_i$ from Theorem A.2.10 that uses $\nu = 20(m+R^2\|d\|^2)$, the self-concordance parameter (computed via Lemma A.2.3) of the recentered log-barrier in Problem 3.0.1. Finally, recognizing $\mathbf{H}^{-1}\mathbf{F}^\top$ as $K_{\boldsymbol{\sigma}}$ from Equation (2.3.2) (with $\boldsymbol{\sigma} = \mathbf{0}^m$) yields

$$\|\mathbf{G}\mathbf{H}^{-1}\mathbf{F}^\top - \mathbf{P}\| = \|\mathbf{G}K_0 - \mathbf{P}\| \leq \|\mathbf{P}\| + \|\mathbf{G}\| L.$$

Combining all the bounds obtained above, we may then finish the proof. □

Thus, Theorem 3.3.1 establishes bounds analogous to those in Fact 2.4.13 for randomized smoothing, demonstrating that the Jacobian of the smoothed expert policy is sufficiently Lipschitz. Indeed, in this case our result is stronger, showing that the Jacobian is differentiable and the Hessian tensor is bounded. This theoretically validates the core proposition of our paper: the barrier MPC policy in Problem 3.0.1 is suitably smooth, and therefore the guarantees in Section 2.4 hold.

We now briefly revisit our learning guarantees, applied to specifically to log-barrier MPC, before we demonstrate the efficacy of barrier MPC in experiments.

3.4 Learning Guarantees for Barrier MPC

We now revisit the learning guarantees discussed in Section 2.4, adapted specifically for a barrier MPC expert.

We begin by considering the stability properties of barrier MPC and note that since we are only interested in establishing that $\|\hat{\mathbf{x}} - \mathbf{x}^\star\| \leq \epsilon$ and we consider MPC controllers which stabilize to the origin, we can relax our incremental input-to-state stability requirements to simply input-to-state stability with minimal assumptions. Definition 3.4.1 introduces the weaker input-to-state stability property and Lemma 3.4.2 shows that ISS policies are locally δISS. We then observe that there is considerable prior work showing that ISS holds under minimal assumptions for barrier MPC, meaning Assumption 2.4.3 is satisfied for barrier MPC. Below we use $\mathcal{KL}$ to denote the class of functions which are class $\mathcal{K}$ in the first argument (monotonically increasing and zero at zero) and monotonically decreasing in the second.

Definition 3.4.1 (Input to State Stability, [Kha02])**.** *A system $\mathbf{x}_{t+1} = f(\mathbf{x}_t, \mathbf{u}_t)$ is input-to-state stable (equiv. a controller π is input-to-state stabilizing for $\mathbf{x}_{t+1} = f(\mathbf{x}_t, \pi(\mathbf{x}_t) + \mathbf{u}_t)$) if there exists $\mathcal{B} \in \mathcal{KL}$ and $\gamma \in \mathcal{K}$ such that,*

$$\|\mathbf{x}_t\| \leq \beta(\|\mathbf{x}_0\|, t) + \gamma(\|\mathbf{u}\|_\infty) \quad \forall t \geq 0.$$

Lemma 3.4.2. *Let π be an L-lipschitz controller which is input-to-state stabilizing for the dynamics $\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{B}\mathbf{u}_t$ with some gains $\mathcal{B} \in \mathcal{KL}, \gamma \in \mathcal{K}$. Define $\mathcal{B}^{-1}(\epsilon)$ such that $\mathcal{B}(\|\mathbf{x}_0\|, \mathcal{B}_{B_x}^{-1}(\epsilon)) \leq \epsilon$ for $\|\mathbf{x}_0\| \leq B_x$. As $\mathcal{B} \in \mathcal{KL}$, we know that $\mathcal{B}_{B_x}^{-1}$ exists and is montonically decreasing in ϵ. Define the gain,*

$$v(\epsilon) := \min\left\{\gamma^{-1}(\epsilon/2), \epsilon \cdot (1 + \|\mathbf{A}\| + (1+L)\|\mathbf{B}\|)^{-\beta^{-1}(\epsilon/4)}\right\}.$$

Then π is incrementally input-to-state stabilizing with gain $\gamma' := v^{-1} \in \mathcal{K}$, i.e. for $\|\mathbf{u}\|_\infty \leq v(\epsilon)$ we have that $\|\mathbf{x}_t - \bar{\mathbf{x}}_t\| \leq \epsilon$ where $\mathbf{x}_{t+1} = f(\mathbf{x}_t, \pi(\mathbf{x}_t) + \mathbf{u}_t)$ and $\bar{\mathbf{x}}_{t+1} = f(\bar{\mathbf{x}}_t, \pi(\mathbf{x}_t))$ with $\bar{\mathbf{x}}_0 = \mathbf{x}_0$.

Note that over a horizon of length K, we have,

$$\gamma'(\|\mathbf{u}\|_\infty) \leq \max\{2\gamma(\|\mathbf{u}\|_\infty), (1 + \|\mathbf{A}\| + (1+L)\|\mathbf{B}\|)^K \|\mathbf{u}\|_\infty\}.$$

Proof. We first show that $\gamma' \in \mathcal{K}$. Since $\gamma \in \mathcal{K}, \gamma^{-1} \in \mathcal{K}$. Furthermore, as $\mathcal{B}^{-1}$ is monotonically decreasing, $C^{-\mathcal{B}^{-1}(\epsilon)}$ is monotonically non-decreasing in ϵ for $C \geq 1$ and $\epsilon \cdot C^{-\mathcal{B}^{-1}(\epsilon)} \in \mathcal{K}$. Since v is the minimum of two class $\mathcal{K}$ gains, it follows that $v \in \mathcal{K}$ and therefore $\gamma' := v^{-1} \in \mathcal{K}$.

We now prove that π is incrementally input-to-state stabilizing with gain $\gamma' := v^{-1}$. Fix any $\epsilon > 0$. WTS that for $\|\mathbf{u}\|_\infty \leq v(\epsilon)$, $\|\mathbf{x} - \bar{\mathbf{x}}\|_\infty \leq \epsilon$. First, consider $t \leq \beta^{-1}(\epsilon/4)$. Note that,

$$\|\mathbf{x}_{t+1} - \bar{\mathbf{x}}_{t+1}\| \leq (\|\mathbf{A}\| + \|\mathbf{B}\|L)\|\mathbf{x}_t - \bar{\mathbf{x}}_t\| + \|\mathbf{B}\| \cdot \|\mathbf{u}\|_\infty.$$

By telescoping we can write,

$$\|\mathbf{x}_t - \bar{\mathbf{x}}_t\| \leq (1 + \|\mathbf{A}\| + (1+L)\|\mathbf{B}\|)^t \cdot \|\mathbf{u}\|_\infty \quad . \tag{3.4.1}$$

We then use that $t \leq \beta^{-1}(\epsilon/4)$ and that $\|\mathbf{u}\|_\infty \leq v(\epsilon) \leq \epsilon(1 + \|\mathbf{A}\| + (1+L)\|\mathbf{B}\|)^{-\beta^{-1}(\epsilon/4)}$. Combining with Equation (3.4.1), we get $\|\mathbf{x}_t - \bar{\mathbf{x}}_t\| \leq \epsilon$.

We now consider the case $t \geq \beta^{-1}(\epsilon/4)$. Here we use the input-to-state stability of π,

$$\begin{aligned}\|\mathbf{x}_t - \bar{\mathbf{x}}_t\| &\leq \|\mathbf{x}_t\| + \|\bar{\mathbf{x}}_t\| \leq 2\beta(\|\mathbf{x}_0\|, t) + \gamma(\|\mathbf{u}\|_\infty) \\ &\leq 2\beta(\|\mathbf{x}_0\|, \beta^{-1}(\epsilon/4)) + \gamma(\gamma^{-1}(\epsilon/2)) \\ &\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} \leq \epsilon.\end{aligned}$$

This concludes the proof. □

Assumption 3.4.3. *The barrier MPC controller π^η_{mpc} in Definition 3.0.2 is input-to-state stabilizing such that, by Lemma 3.4.2, is it incrementally input-to-state stablizing over $t \leq K$ for some with linear gain function γ.*

For stabilzable systems and proper choices of cost function and constraints, we note that barrier MPC is ISS, and therefore locally δISS. See [PFDS20; FE14] for treatment on the input-to-state stability properties of barrier MPC. This shows that π^{η}_{mpc} satisfies Assumption 2.4.3.

We now state our end-to-end learning guarantee, an extension of Fact 2.4.6.

Theorem 3.4.4. *Let π^{η}_{mpc} be a barrier MPC as in Definition 3.0.2 that satisfies Assumption 3.4.3 such that it is (τ, γ)-locally-δISS for linear gain γ over a horizon length K and maximum perturbation of κ. Let L be as defined in Corollary 3.2.3 and overload γ to denote the constant associated with $\gamma(\cdot)$. Let m be the number of constraints and r, R be the radii associated with the constraint polytope in Definition 3.0.2.*

Assume that $\|\mathbf{x}_0\| \le B_x$ and let $\widehat{\pi}$ be chosen such that, for some $\epsilon_0, \epsilon_1 > 0$ and given N sample trajectories of length K under π^{η}_{mpc} from an initial condition distribution $\mathcal{D}$,

$$\mathbb{P}_{\mathbf{x}_0 \sim \mathcal{D}}\left[\sup_{0 \le k \le K} \|\widehat{\pi}(\mathbf{x}_k) - \pi^{\eta}_{\text{mpc}}(\mathbf{x}_k)\| \le \epsilon_0 / N \right. \\ \left. \wedge \sup_{0 \le k \le K} \left\| \frac{\partial \widehat{\pi}}{\partial \mathbf{x}}(\mathbf{x}_k) - \frac{\partial \pi^{\eta}_{\text{mpc}}}{\partial \mathbf{x}}(\mathbf{x}_k) \right\| \le \epsilon_1 / N \right] \ge 1 - \delta.$$

Then, provided that $N \ge \mathcal{O}(1)\epsilon_0 \gamma^2 (1+L)^2 \frac{R}{r} \max\left\{\frac{LB_x}{\eta}, \frac{r}{2m+4\sqrt{m}}\right\}$, $N \ge \epsilon_0 \gamma \max\{16, 8/\tau\}$, and $N \ge 4\gamma\epsilon_1$, it follows that,

$$\|\hat{\mathbf{x}}_t - \mathbf{x}^{\eta}_t\| \le \frac{8\gamma\epsilon_0}{N} \quad \forall 0 \le t \le K.$$

Proof. This is an application of Fact 2.4.6, combined with our L_1 smoothness bound on the Hessian, where,

$$L_1 \le \mathcal{O}(1) \frac{(1+L)^2}{\frac{r}{R} \min\{\frac{\eta}{LB_x}, \frac{r}{2m+4\sqrt{m}}\}} \le \mathcal{O}(1)(1+L)^2 \frac{R}{r} \max\left\{\frac{LB_x}{\eta}, \frac{2m+4\sqrt{m}}{r}\right\}.$$

The bound on L_1 follows from Theorem 3.3.1 and noting that $\|\mathbf{K}_0 \mathbf{x}_0 - \mathbf{u}^\star\| \le 2LB_x$ in Equation (3.3.1). This yields that $\text{res}_{\text{l.b.}} \ge \mathcal{O}(1) \min\{\eta / LB_x, r/(2m + 4\sqrt{m})\}$. Plugging $\text{res}_{\text{l.b.}}$ into Theorem 3.3.1 yields the above bound on L_1. Substitution of this result into Fact 2.4.6 proves the final statement. □

The above theorem concludes our theoretical analysis of imitation learning with barrier

MPC. Due to the smoothness analysis we have performed, we note that Theorem 3.4.4's principal assumption is closed-loop incremental stability under $\boldsymbol{\pi}^{\eta}_{\mathrm{mpc}}$, and that the smoothness requirement has been removed.

We now confirm the efficacy of learning barrier MPC over randomized smoothing via experiments.

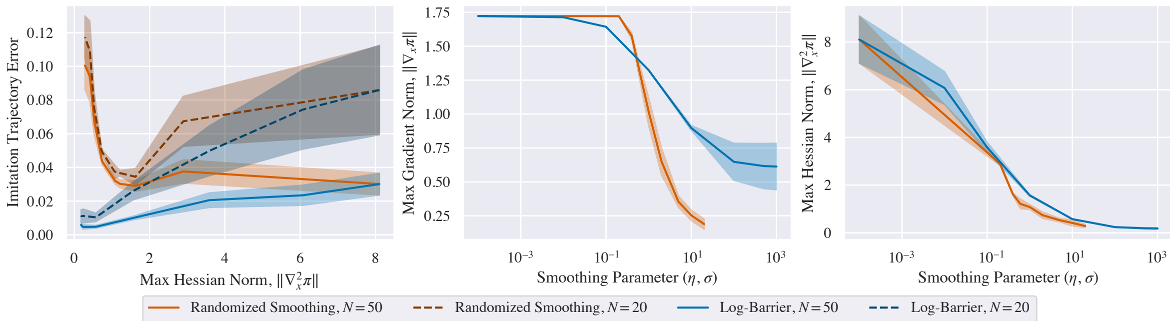

Figure 3.1: Left: The imitation error $\max_t \|\hat{x} - x^\star\|$ for the trained MLP over 5 seeds, as a function of the expert smoothness for both randomized smoothing and log-barrier MPC. Center, Right: The L_0 (gradient norm) and L_1 (hessian norm) smoothness of $\pi^\star$ as a function of the smoothing parameter.

3.5 Empirical Comparison of Imitation Learning of Smoothed MPC Policies

Experimental Setup. We compare barrier MPC with randomized smoothing for the same double integrator system visualized in Figure 2.1. The dynamics are given by,

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

with costs $\mathbf{Q}_t = \mathbf{I}$, $\mathbf{R}_t = 0.01\mathbf{I}$, and horizon length $T = 10$ in addition to the constraints $\|\mathbf{x}\|_\infty \leq 10, \|\mathbf{u}\|_\infty \leq 1$ chosen for Definition 3.0.2. Feasible initial conditions were chosen uniformly at random. This is the same setup as in [AMMHJ23].

We sample $N \in [20, 50]$ trajectories of length $K = 20$ using π^η_{mpc} and π^{rs} and smoothing parameters η and σ ranging from 10^{-4} to 10^3 and 10^{-4} to 20, respectively. We use $\mathcal{P} = \mathcal{N}(0, I)$ for the randomized smoothing distribution. For each parameter set, we trained a 4-layer multi-layer perceptron (MLP) using GELU activations [HG16] to ensure smoothness of Π. We used AdamW [LH18] with a learning rate of $3 \cdot 10^{-4}$ and weight decay of 10^{-3} in order to ensure boundedness of the weights (see [PZTM22]).

Results In Figure 3.1, we visualize the smoothness properties of the chosen expert $\pi^\star$ of each method (either π^η_{mpc} or π^{rs}) across the choices of η, σ. We note that for small Hessian norms,

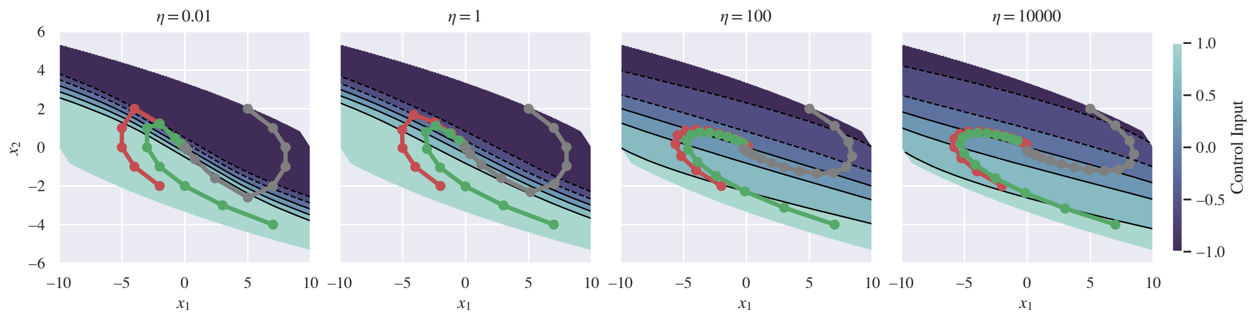

Figure 3.2: Visualizations of the log-barrier MPC control policy and several trajectories for the same system as Figure 2.1 and Figure 3.1 across different choices of η.

barrier MPC has larger gradient norm $\|\nabla \pi^\star\|$. This shows how π^η_{mpc} prevents oversmoothing in comparison to π^{rs}: whereas randomized smoothing reduces $\|\nabla^2 \pi^\star\|$ by essentially flattening the function, π^η_{mpc} achieves equally smooth functions while still maintaining control of the system. This effect can also be seen in Figure 3.2, where we visualize the barrier MPC controller for different choices of η and show that even for very large choices of η we successfully stabilize to the origin.

One interesting phenomena is that the maximum gradient of π^η_{mpc} begins decreasing much earlier than π^{rs}. This is due to the fact that π^{rs} only smooths locally, meaning that if the smoothing radius is sufficiently small, the gradient will not be affected. Meanwhile, π^η_{mpc} is always performing a *global* form of smoothing, so that even for small η the controller is being smoothed everywhere.

In Figure 3.1, we also compare the trajectory error when imitating trajectories from $\pi^{\mathrm{rs}}, \pi^\eta_{\mathrm{mpc}}$ for equivalent levels of smoothness. We can see that for $N = 20$ and $N = 50$, π^η_{mpc} significantly outperforms π^{rs} across all smoothness levels. This effect is particularly pronounced in the very smooth regime, where imitating π^{rs} proves unstable, leading to extremely large imitation errors. Meanwhile, π^η_{mpc} only performs better for higher levels of smoothing. Overall, these experiments confirm our hypothesis: that barrier MPC is an effective smoothing technique that outperforms randomized smoothing. This demonstrates that not all smoothing techniques are equal for the purposes of imitation learning.

3.6 Conclusion

This work shows that, for controls purposes, not all smoothing techniques are equivalent. We compared two methods for smoothing MPC policies for linear systems: randomized smoothing and barrier MPC. While we showed that both have theoretically optimal tradeoffs in terms of approximation error to Hessian norm, barrier MPC is aware of the underlying control problem. We show theoretically how this enables guarantees when learning barrier MPC and demonstrate in a set of experiments on a simple system that this difference can manifest in significantly improved performance of barrier MPC in comparison to a randomized smoothing baseline.

The key machinery we introduce to bound the smoothness properties of barrier MPC include novel bounds for general convex and quadratic optimization problems using self-concordant barrier functions, which we believe may be of independent interest to the broader optimization community.

Further development of the tools we present may yield techniques and guarantees for smoothing nonlinear MPC policies or other challenging scenarios, i.e. smoothing policies based on bi-level optimization using log-barriers. The development of better theoretical tools in this domain is potentially of great consequence. We conclude with several concrete proposals for future work:

1. While the optimal smoothing bound given in Theorem 2.4.10 suggests that the bound on the spectral norm of $\frac{\partial^2 \mathbf{u}^\eta}{\partial \mathbf{x}_0^2}$ (which is a third-order tensor) presented in Theorem 3.3.1 is tight by virtue of Theorem 3.1.2, a more careful analysis of the hessian of $\mathbf{v}_i^\top \mathbf{u}^\eta$ for a basis $\{\mathbf{v}_i\}$ may yield tighter bounds such that the dependence on η of the error of $\mathbf{v}_i^\top \mathbf{u}^\eta$ matches the Hessian of $\mathbf{v}_i^\top \mathbf{u}^\eta$, i.e. that barrier MPC smooths optimally along all directions.

2. Generalization of our results to MPC with arbitrary self-concordant barriers. Many of our results (notably our residual lower bound and directional error result, Theorem A.2.10)

are for general ν-self-concordant barriers. However, our analysis of MPC itself extends only to log-barrier-based MPC. It is unclear whether a result such as Theorem 3.2.2, which shows that the Jacobian can be written as a convex combination of pieces relating to the explicit MPC, can be developed for general self-concordant barriers.

3. Extension to nonlinear MPC. Many of these results may be applicable to barrier MPC with nonlinear dynamics (e.g., systems with piecewise affine dynamics) or more complex constraint sets. This could potentially be done by considering the appropriate local linearizations.

References

[Pom88] Dean A Pomerleau. Alvinn: an autonomous land vehicle in a neural network. *Advances in neural information processing systems*, 1988.

[RSB09] Nathan D Ratliff, David Silver, and J Andrew Bagnell. Learning to search: functional gradient techniques for imitation learning. *Autonomous Robots*, 2009.

[ACN10] Pieter Abbeel, Adam Coates, and Andrew Y Ng. Autonomous helicopter aerobatics through apprenticeship learning. *The International Journal of Robotics Research*, 2010.

[RGB11] Stéphane Ross, Geoffrey Gordon, and Drew Bagnell. A reduction of imitation learning and structured prediction to no-regret online learning. In *Proceedings of the fourteenth international conference on artificial intelligence and statistics*. JMLR Workshop and Conference Proceedings, 2011.

[PZTM22] Daniel Pfrommer, Thomas Zhang, Stephen Tu, and Nikolai Matni. Tasil: taylor series imitation learning. *Advances in Neural Information Processing Systems*, 2022.

[TRZM22] Stephen Tu, Alexander Robey, Tingnan Zhang, and Nikolai Matni. On the sample complexity of stability constrained imitation learning. In *Learning for Dynamics and Control Conference*. PMLR, 2022.

[BJPST24] Adam Block, Ali Jadbabaie, Daniel Pfrommer, Max Simchowitz, and Russ Tedrake. Provable guarantees for generative behavior cloning: bridging low-level stability and high-level behavior. *Advances in Neural Information Processing Systems*, 36, 2024.

[AZ12] Frank Allgöwer and Alex Zheng. *Nonlinear model predictive control*, volume 26. Birkhäuser, 2012.

[BMDP02] Alberto Bemporad, Manfred Morari, Vivek Dua, and Efstratios N Pistikopoulos. The explicit linear quadratic regulator for constrained systems. *Automatica*, 2002.

[Mor20] CG da S Moraes. A neural network architecture to learn explicit mpc controllers from data. *IFAC-PapersOnLine*, 2020.

[CSA+18] Steven Chen, Kelsey Saulnier, Nikolay Atanasov, Daniel D Lee, Vijay Kumar, George J Pappas, and Manfred Morari. Approximating explicit model predictive control using constrained neural networks. In *2018 Annual American control conference (ACC)*. IEEE, 2018.

[AMMHJ23] Kwangjun Ahn, Zakaria Mhammedi, Horia Mania, Zhang-Wei Hong, and Ali Jadbabaie. Model predictive control via on-policy imitation learning. In *Learning for Dynamics and Control Conference*. PMLR, 2023.

[RB10] Stéphane Ross and Drew Bagnell. Efficient reductions for imitation learning. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*. JMLR Workshop and Conference Proceedings, 2010.

[LLFDG17] Michael Laskey, Jonathan Lee, Roy Fox, Anca Dragan, and Ken Goldberg. Dart: noise injection for robust imitation learning. In *Conference on robot learning*. PMLR, 2017.

[ZKL+23] Thomas T Zhang, Katie Kang, Bruce D Lee, Claire Tomlin, Sergey Levine, Stephen Tu, and Nikolai Matni. Multi-task imitation learning for linear dynamical systems. In *Learning for Dynamics and Control Conference*, pages 586–599. PMLR, 2023.

[Dik67a] II Dikin. Iterative solution of problems of linear and quadratic programming. In *Doklady Akademii Nauk*, volume 174 of number 4, pages 747–748. Russian Academy of Sciences, 1967.

[Kar84] Narendra Karmarkar. A new polynomial-time algorithm for linear programming. In *Proceedings of the sixteenth annual ACM symposium on Theory of computing*, 1984.

[Wri05] Margaret Wright. The interior-point revolution in optimization: history, recent developments, and lasting consequences. *Bulletin of the American mathematical society*, 42(1):39–56, 2005.

[NN94] Yurii Nesterov and Arkadii Nemirovskii. *Interior-point polynomial algorithms in convex programming*. SIAM, 1994.

[Nes+18] Yurii Nesterov et al. *Lectures on convex optimization*, volume 137. Springer, 2018.

[WH04] Adrian G Wills and William P Heath. Barrier function based model predictive control. *Automatica*, 2004.

[FE13] Christian Feller and Christian Ebenbauer. A barrier function based continuous-time algorithm for linear model predictive control. In *2013 European Control Conference (ECC)*, pages 19–26. IEEE, 2013.

[FE14] Christian Feller and Christian Ebenbauer. Barrier function based linear model predictive control with polytopic terminal sets. In *53rd IEEE Conference on Decision and Control*, pages 6683–6688. IEEE, 2014.

[FE15a] Christian Feller and Christian Ebenbauer. Weight recentered barrier functions and smooth polytopic terminal set formulations for linear model predictive control. In *2015 American Control Conference (ACC)*, pages 1647–1652. IEEE, 2015.

[FE16] Christian Feller and Christian Ebenbauer. Relaxed logarithmic barrier function based model predictive control of linear systems. *IEEE Transactions on Automatic Control*, 2016.

[PFDS20] Maxime Pouilly-Cathelain, Philippe Feyel, Gilles Duc, and Guillaume Sandou. Stability of barrier model predictive control. In *17th International Conference on Informatics in Control, Automation and Robotics (ICINCO)*, 2020.

[FE15b] Christian Feller and Christian Ebenbauer. Input-to-state stability properties of relaxed barrier function based mpc. *IFAC-PapersOnLine*, 48(23):302–307, 2015.

[ZZS21] Jun Zeng, Bike Zhang, and Koushil Sreenath. Safety-critical model predictive control with discrete-time control barrier function. In *2021 American Control Conference (ACC)*. IEEE, 2021.

[WZ22] Kim Peter Wabersich and Melanie N Zeilinger. Predictive control barrier functions: enhanced safety mechanisms for learning-based control. *IEEE Transactions on Automatic Control*, 2022.

[BBM17] Francesco Borrelli, Alberto Bemporad, and Manfred Morari. *Predictive control for linear and hybrid systems*. Cambridge University Press, 2017.

[AB09] Alessandro Alessio and Alberto Bemporad. A survey on explicit model predictive control. *Nonlinear Model Predictive Control: Towards New Challenging Applications*, 2009.

[VR20] Rick Voßwinkel and Klaus Röbenack. Determining input-to-state and incremental input-to-state stability of nonpolynomial systems. *International Journal of Robust and Nonlinear Control*, 2020.

[ZM11] Majid Zamani and Rupak Majumdar. A lyapunov approach in incremental stability. In *2011 50th IEEE conference on decision and control and European control conference*. IEEE, 2011.

[KL20] Benjamin Karg and Sergio Lucia. Efficient representation and approximation of model predictive control laws via deep learning. *IEEE Transactions on Cybernetics*, 2020.

[DBW12] John C Duchi, Peter L Bartlett, and Martin J Wainwright. Randomized smoothing for stochastic optimization. *SIAM Journal on Optimization*, 2012.

[Dik67b] I. I. Dikin. Iterative solution of problems of linear and quadratic programming. English. *Sov. Math., Dokl.*, 1967.

[Kha02] Hassan K Khalil. *Control of nonlinear systems*. Prentice Hall, New York, NY, 2002.

[HG16] Dan Hendrycks and Kevin Gimpel. Gaussian error linear units (gelus). *arXiv preprint arXiv:1606.08415*, 2016.

[LH18] Ilya Loshchilov and Frank Hutter. Fixing weight decay regularization in adam, 2018.

[HJ12] Roger A Horn and Charles R Johnson. *Matrix analysis*. Cambridge university press, 2012.

[IR08] Ilse CF Ipsen and Rizwana Rehman. Perturbation bounds for determinants and characteristic polynomials. *SIAM Journal on Matrix Analysis and Applications*, 2008.

[Lau04] Alan J Laub. *Matrix analysis for scientists and engineers*. SIAM, 2004.

[GLP+24] Mehrdad Ghadiri, Yin Tat Lee, Swati Padmanabhan, William Swartworth, David Woodruff, and Guanghao Ye. Improving the bit complexity of communication for distributed convex optimization, 2024. arXiv: 2403.19146 [cs.DS].

[ZLY23] MR Zong, YT Lee, and MC Yue. Short-step methods are not strongly polynomial time. *Mathematical Programming*, 2023.

Appendix A

Proofs and Auxiliary Results

A.1 Technical Results in Matrix Analysis

We use the notation introduced in Section 2.2. Additionally, we use $\mathbf{e}_i$ to denote the vector with one at the i^{th} coordinate and zeroes at the remaining coordinates. We first collect the following relevant facts from matrix analysis before proving our technical results.

Fact A.1.1 (Definitions; [HJ12]). *Given a square matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$, its $(i,j)^{\text{th}}$ minor $\mathbf{M}_{i,j}$, is defined as the determinant of the $(n-1)\times(n-1)$ matrix resulting from deleting row i and column j of $\mathbf{A}$. Next, the $(i,j)^{\text{th}}$ cofactor is defined to be the $(i,j)^{\text{th}}$ minor scaled by $(-1)^{i+j}$:*

$$\mathbf{C}_{ij} = (-1)^{i+j}\mathbf{M}_{i,j}. \tag{A.1.1}$$

We then define the cofactor matrix $\mathbf{C} \in \mathbb{R}^{n\times n}$ of $\mathbf{A}$ as the matrix of cofactors of all entries of $\mathbf{A}$, i.e., $\mathbf{C} = ((-1)^{i+j}\mathbf{M}_{i,j})_{1\le i,j\le n}$. The adjugate of $\mathbf{A}$ is the transpose of the cofactor matrix $\mathbf{C}$ of $\mathbf{A}$, and hence its $(i,j)^{\text{th}}$ entry may be expressed as:

$$\text{adj}(\mathbf{A})_{ij} = (-1)^{i+j}\mathbf{M}_{j,i}. \tag{A.1.2}$$

In particular, if the matrix $\mathbf{A}$ is symmetric*, then the $(i,j)^{\text{th}}$ minor equals the $(j,i)^{\text{th}}$ minor,* *implying*

$$\mathrm{adj}(\mathbf{A})_{ij} = \mathrm{adj}(\mathbf{A})_{ji} \textit{ for all } i, j \in [n]. \tag{A.1.3}$$

The minors of a matrix are also useful in computing its determinant. Specifically, the Laplace expansion of a matrix $\mathbf{A}$ *along its column* j *is given as:*

$$\det(\mathbf{A}) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \mathbf{M}_{i,j}. \tag{A.1.4}$$

Finally, the adjugate $\mathrm{adj}(\mathbf{A})$ *also satisfies the following important property:*

$$\mathrm{adj}(\mathbf{A}) \cdot \mathbf{A} = \mathbf{A} \cdot \mathrm{adj}(\mathbf{A}) = \det(\mathbf{A}) \cdot \mathbf{I}. \tag{A.1.5}$$

Fact A.1.2 (Matrix determinant lemma, [HJ12])**.** *For any* $\mathbf{M}$*, the determinant for a unit-rank update may be expressed as:*

$$\det(\mathbf{M} + \mathbf{u}\mathbf{v}^\top) = \det(\mathbf{M}) + \mathbf{v}^\top \mathrm{adj}(\mathbf{M})\mathbf{u}. \tag{A.1.6}$$

Fact A.1.3 (Sherman-Morrison-Woodbury identity)**.** *Given conformable matrices* $\mathbf{A}, \mathbf{C}, \mathbf{U}$*, and* $\mathbf{V}$ *such that* $\mathbf{A}$ *and* $\mathbf{C}$ *are invertible, we have*

$$(\mathbf{A} + \mathbf{U}\mathbf{C}\mathbf{V})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{U}(\mathbf{C}^{-1} + \mathbf{V}\mathbf{A}^{-1}\mathbf{U})^{-}1\mathbf{V}\mathbf{A}^{-1}.$$

We crucially use the following expansion for determinants of perturbed matrices.

Fact A.1.4 (Theorem 2.3 of [IR08])**.** *Let* $\mathbf{D}$ *and* $\mathbf{F}$ *be* $n \times n$ *complex matrices. Denote by* $\mathbf{F}_{i_1 \dots i_k}$ *the principal submatrix of order* $n-k$ *obtained by deleting rows and columns* $i_1 \dots i_k$ *of the* $n \times n$ *matrix* $\mathbf{F}$*. If* $\mathbf{D} = \mathrm{Diag}(\delta_1, \dots, \delta_n)$*, then*

$$\det(\mathbf{D} + \mathbf{F}) = \det(\mathbf{D}) + \det(\mathbf{F}) + \mathbf{S}_1 + \dots + \mathbf{S}_{n-1},$$

where

$$\mathbf{S}_k := \sum_{1 \le i_1 < \dots < i_k \le n} \delta_{i_1} \dots \delta_{i_k} \det(\mathbf{F}_{i_1 \dots i_k}), \textit{ for } 1 \le k \le n-1.$$

Lemma A.1.5 (Theorem 2.3 of [IR08])**.** *Given* $\mathbf{A} \in \mathbb{R}^{m \times m}$ *as in Fact A.1.2, positive diagonal* *matrix* $\mathbf{\Lambda} = \mathrm{Diag}(\boldsymbol{\lambda}) \in \mathbb{R}^{m\times m}$*, and* $\mathbf{A}_{\boldsymbol{\sigma}}$ *denoting the principal submatrix formed by selecting* $\mathbf{A}$*'s rows and columns indexed by* $\sigma \in \{0,1\}^m$*, we have*

$$\det(\mathbf{A}+\mathbf{\Lambda}) = \sum_{\sigma\in\{0,1\}^n} \left(\prod_{i=1}^{m} \lambda_i^{1-\sigma_i}\right) \det(\mathbf{A}_{\boldsymbol{\sigma}})$$

We now state and prove a technical result that we build upon to prove Lemma A.1.12, which we in turn use in the proof of Theorem 3.2.2.

Lemma A.1.6. *Consider a matrix* $\mathbf{A} = \begin{bmatrix} a & \mathbf{b}^\top \\ \mathbf{b} & \mathbf{D} \end{bmatrix} \in \mathbb{R}^{n\times n}$*, where* $\mathbf{D} \in \mathbb{R}^{(n-1)\times(n-1)}$ *is a symmetric matrix. Then the adjugate* $\mathrm{adj}(\mathbf{A})$ *may be expressed as follows:*

$$\mathrm{adj}(\mathbf{A}) = \begin{bmatrix} \det(\mathbf{D}) & -\mathbf{b}^\top \mathrm{adj}(\mathbf{D}) \\ -\mathrm{adj}(\mathbf{D})\mathbf{b} & a\cdot \mathrm{adj}(\mathbf{D}) + \mathbf{K} \end{bmatrix},$$

for some matrix $\mathbf{K}$ *independent of* a*.*

Proof. Let $\widetilde{\mathbf{D}}_{ij}$ be the $(n-2)\times(n-2)$ matrix obtained by deleting the i^{th} row and j^{th} column of $\mathbf{D}$. Let $\widetilde{\mathbf{D}}_j$ be the $(n-1)\times(n-2)$ matrix formed by removing the j^{th} column of $\mathbf{D}$, and let $\widetilde{\mathbf{b}}_i \in \mathbb{R}^{(n-2)}$ be the vector obtained by deleting the i^{th} coordinate of $\mathbf{b}$. With this notation in hand, we now compute some relevant cofactors.

First, observe that $\mathbf{D}$ is the matrix obtained by deleting the first row and column of $\mathbf{A}$, and hence this fact along with Equation (A.1.1) yields the $(1,1)^{\text{th}}$ cofactor:

$$\mathbf{C}_{1,1} = \det(\mathbf{D}).$$

Second, for some $j>0$, observe that the matrix obtained by deleting the first row of $\mathbf{A}$ and the $(1+j)^{\text{th}}$ column of $\mathbf{A}$ is exactly the horizontal concatenation of $\mathbf{b}$ and $\widetilde{\mathbf{D}}_j$. Applying this observation in Equation (A.1.1) then gives the following expression for the $(1,1+j)^{\text{th}}$

cofactor:

$$\begin{aligned}\mathbf{C}_{1,1+j} &= (-1)^j \det\left(\begin{bmatrix}\mathbf{b} & \widetilde{\mathbf{D}}_j\end{bmatrix}\right) \\ &= (-1)^j \sum_{i=1}^{n-1} b_i (-1)^{i+1} \det(\widetilde{\mathbf{D}}_{ij}) \\ &= -\sum_{i=1}^{n-1} b_i \operatorname{adj}(\mathbf{D})_{ij} \\ &= -[\mathbf{b}^\top \operatorname{adj}(\mathbf{D})]_j,\end{aligned}$$

where the second step is by using Equation (A.1.4) to expand $\det\left(\begin{bmatrix}\mathbf{b} & \widetilde{\mathbf{D}}_j\end{bmatrix}\right)$ along the column vector $\mathbf{b}$, and the third step is by Equation (A.1.2) and Equation (A.1.3), which applies since $\mathbf{D}$ is assumed symmetric. Finally, to compute the $(1+i, 1+j)^{\text{th}}$ cofactor, we first construct the matrix obtained by deleting the $(1+i)^{\text{th}}$ row and $(1+j)^{\text{th}}$ column of $\mathbf{A}$. Based on the notation we introduced above, this may be expressed as $\begin{bmatrix} a & \widetilde{\mathbf{b}}_j^\top \\ \widetilde{\mathbf{b}}_i & \widetilde{\mathbf{D}}_{ij}\end{bmatrix}$, from which we have by Equation (A.1.1):

$$\mathbf{C}_{1+i,1+j} = (-1)^{i+j} \det\left(\begin{bmatrix} a & \widetilde{\mathbf{b}}_j^\top \\ \widetilde{\mathbf{b}}_i & \widetilde{\mathbf{D}}_{ij}\end{bmatrix}\right), \tag{A.1.7}$$

which we now simplify. To this end, we observe that

$$\begin{aligned}\det\left(\begin{bmatrix} a & \widetilde{\mathbf{b}}_j^\top \\ \widetilde{\mathbf{b}}_j & \widetilde{\mathbf{D}}_{ij}\end{bmatrix}\right) &= \det\left(\begin{bmatrix} a & \widetilde{\mathbf{b}}_j^\top \\ 0 & \widetilde{\mathbf{D}}_{ij}\end{bmatrix} + \begin{bmatrix} 0 \\ \widetilde{\mathbf{b}}_j\end{bmatrix} \cdot \mathbf{e}_1^\top\right) \\ &= \det\left(\begin{bmatrix} a & \widetilde{\mathbf{b}}_j^\top \\ 0 & \widetilde{\mathbf{D}}_{ij}\end{bmatrix}\right) + \mathbf{e}_1^\top \operatorname{adj}\left(\begin{bmatrix} a & \widetilde{\mathbf{b}}_j^\top \\ 0 & \widetilde{\mathbf{D}}_{ij}\end{bmatrix}\right) \begin{bmatrix} 0 \\ \widetilde{\mathbf{b}}_j\end{bmatrix},\end{aligned} \tag{A.1.8}$$

where we used Fact A.1.2 in the second step. The first term in the right-hand side of the preceding equation may be simplified to $a \cdot \det(\widetilde{\mathbf{D}}_{ij})$. To simplify the second term, we first observe that we wish to compute only the first row of $\operatorname{adj}\left(\begin{bmatrix} a & \widetilde{\mathbf{b}}_j^\top \\ 0 & \widetilde{\mathbf{D}}_{ij}\end{bmatrix}\right)$; to this end, we introduce the notation that $\mathbf{X} := \widetilde{\mathbf{D}}_{ij}$ and $\mathbf{y} = \widetilde{\mathbf{b}}_j$; we denote $\widetilde{\mathbf{X}}_{\ell j}$ to be the matrix obtained by deleting the ℓ^{th} row and j^{th} column of $\mathbf{X}$; we use $\widetilde{\mathbf{X}}_\ell$ for the matrix obtained by deleting the ℓ^{th} row of $\mathbf{X}$. Now observe that for $\ell > 0$, the $(1+\ell)^{\text{th}}$ entry of the desired first row may be computed as follows:

$$\begin{aligned}\operatorname{adj}\left(\begin{bmatrix} a & \widetilde{\mathbf{b}}_j^\top \\ 0 & \widetilde{\mathbf{D}}_{ij}\end{bmatrix}\right)_{1,1+\ell} &= (-1)^\ell \det \begin{bmatrix} \mathbf{y}^\top \\ \widetilde{\mathbf{X}}_\ell \end{bmatrix} = \sum_{j=1}^{n-1} y_j \cdot (-1)^{\ell+j} \det \widetilde{\mathbf{X}}_{\ell j} \\ &= \sum_{j=1}^{n-1} y_j \cdot (\operatorname{adj} \widetilde{\mathbf{X}})_{j\ell} = \widetilde{\mathbf{b}}_j^\top \operatorname{adj}(\widetilde{\mathbf{D}}_{ij}) \mathbf{e}_\ell,\end{aligned}$$

where the first step is by expressing the $(1, 1+\ell)^{\text{th}}$ entry of the adjugate in question in terms of its $(1+\ell, 1)^{\text{th}}$ minor (as per Equation (A.1.2)), the second step is by the Laplace expansion of the determinant along its first row (analogous to Equation (A.1.4)), the third step is by Equation (A.1.2), and the final step plugs back the newly introduced notation. Hence, we have

$$\mathbf{e}_1^\top \operatorname{adj}\left(\begin{bmatrix} a & \widetilde{\mathbf{b}}_j^\top \\ 0 & \widetilde{\mathbf{D}}_{ij}\end{bmatrix}\right) = \begin{bmatrix} a \cdot \det(\widetilde{\mathbf{D}}_{ij}) & \widetilde{\mathbf{b}}_j^\top \operatorname{adj}(\widetilde{\mathbf{D}}_{ij}) \end{bmatrix}. \tag{A.1.9}$$

Multiplying the right-hand side of Equation (A.1.9) by $\begin{bmatrix} 0 \\ \widetilde{\mathbf{b}}_j \end{bmatrix}$ and plugging the result back into Equation (A.1.8) and eventually into Equation (A.1.7) then gives

$$\mathbf{C}_{1+i,1+j} = a \cdot \operatorname{adj}(\mathbf{D})_{ij} + (-1)^{i+j} \widetilde{\mathbf{b}}_j^\top \operatorname{adj}(\widetilde{\mathbf{D}}_{ij}) \widetilde{\mathbf{b}}_j.$$

By mapping these cofactors back into the definition of the adjugate we want, one can then conclude the proof, with $\mathbf{K}$ collecting all the $(-1)^{i+j} \widetilde{\mathbf{b}}_i^\top \operatorname{adj}(\mathbf{M}_{ij}) \widetilde{\mathbf{b}}_j$ terms. □

Corollary A.1.7. *Let* $\mathbf{A} = \begin{bmatrix} a & \mathbf{b}^\top \\ \mathbf{b} & \mathbf{D} \end{bmatrix}$ *be a symmetric matrix. Then,*

$$\operatorname{adj}(\mathbf{A} + \lambda \mathbf{e}_1 \mathbf{e}_1^\top) = \operatorname{adj}(\mathbf{A}) + \lambda \begin{bmatrix} 0 & \mathbf{0}^\top \\ \mathbf{0} & \operatorname{adj}(\mathbf{D}) \end{bmatrix}$$

Proof. First, observe that by applying Lemma A.1.6, we have

$$\operatorname{adj}\left(\begin{bmatrix} a+\lambda & \mathbf{b}^\top \\ \mathbf{b} & \mathbf{D} \end{bmatrix}\right) = \begin{bmatrix} \det(\mathbf{D}) & -\mathbf{b}^\top \operatorname{adj}(\mathbf{D}) \\ -\operatorname{adj}(\mathbf{D})\mathbf{b} & (a+\lambda)\cdot \operatorname{adj}(\mathbf{D}) + \mathbf{K} \end{bmatrix}. \tag{A.1.10}$$

Next, observe that based on the definition of $\mathbf{A}$, the left-hand side of Equation (A.1.10) is precisely $\operatorname{adj}(\mathbf{A} + \lambda \mathbf{e}_1\mathbf{e}_1^\top)$; based on the expression for $\operatorname{adj}(\mathbf{A})$ from Lemma A.1.6, the right-hand side of Equation (A.1.10) may be split into $\operatorname{adj}(\mathbf{A}) + \lambda \begin{bmatrix} 0 & 0 \\ 0 & \operatorname{adj}(\mathbf{D}) \end{bmatrix}$, as desired. This concludes the proof. □

Lemma A.1.8. *Consider a matrix* $\mathbf{L} \in \mathbb{R}^{m\times n}$*, and define the matrix* $\mathbf{A} = \mathbf{L}\mathbf{L}^\top \in \mathbb{R}^{m\times m}$*. Suppose that* $\det(\mathbf{A}) = 0$*. Then, the following equation holds:*

$$\operatorname{adj}(\mathbf{A})\mathbf{L} = \mathbf{0}.$$

To prove Lemma A.1.8, we use the following two technical results from matrix analysis.

Fact A.1.9 (Theorem 4.18, [Lau04])**.** *Suppose* $\mathbf{A} \in \mathbb{R}^{n\times p}$ *and* $\mathbf{L} \in \mathbb{R}^{n\times m}$*. Then* $\mathbf{A}\mathbf{A}^\dagger\mathbf{L} = \mathbf{L}$ *if and only if the range spaces* $\mathcal{R}(\mathbf{L})$ *and* $\mathcal{R}(\mathbf{A})$ *satisfy the inclusion* $\mathcal{R}(\mathbf{L}) \subseteq \mathcal{R}(\mathbf{A})$*.*

Fact A.1.10 (Theorem 3.21, [Lau04])**.** *Let* $\mathbf{L} \in \mathbb{R}^{m\times n}$*. Then the range spaces* $\mathcal{R}(\mathbf{L})$ *and* $\mathcal{R}(\mathbf{L}\mathbf{L}^\top)$ *satisfy the property* $\mathcal{R}(\mathbf{L}) = \mathcal{R}(\mathbf{L}\mathbf{L}^\top)$*.*

Proof of Lemma A.1.8. We prove the claim by showing that

$$\operatorname{adj}(\mathbf{A})\mathbf{L} = [\operatorname{adj}(\mathbf{A})\mathbf{A}]\mathbf{A}^\dagger\mathbf{L} = \mathbf{0},$$

where the last equality follows from the property that $\operatorname{adj}(\mathbf{A})\mathbf{A} = \det(\mathbf{A})\mathbf{I} = 0$. All that remains is to prove that $\mathbf{L} = \mathbf{A}\mathbf{A}^\dagger\mathbf{L}$. By Fact A.1.9, this is true if and only if $\mathcal{R}(\mathbf{L}) \subseteq \mathcal{R}(\mathbf{A})$. From Fact A.1.10, we know that this is true. This concludes the proof.

□

Lemma A.1.11. *Given a binary vector* $\boldsymbol{\sigma} \in \{0,1\}^m$*, matrix* $\mathbf{G} \in \mathbb{R}^{m\times n}$ *and matrix* $\mathbf{H} \in \mathbb{R}^{n\times n}$

with the properties $\mathbf{H} \succ 0$ *and* $\det(\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top)_\sigma = 0$*, we have*

$$\mathbf{G}^\top \operatorname{adj}(\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top)_\sigma = 0.$$

Proof. Without loss of generality, let $\mathbf{G} = \begin{bmatrix} \mathbf{G}_1 \\ \mathbf{G}_2 \end{bmatrix}$ where $\sigma_i = 1$ for the rows and columns associated with $\mathbf{G}_2$. Then we may express $\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top$ in terms of $\mathbf{G}_1$ and $\mathbf{G}_2$ as follows:

$$\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top = \begin{bmatrix} \mathbf{G}_1\mathbf{H}^{-1}\mathbf{G}_1^\top & \mathbf{G}_1\mathbf{H}^{-1}\mathbf{G}_2^\top \\ \mathbf{G}_2\mathbf{H}^{-1}\mathbf{G}_1^\top & \mathbf{G}_2\mathbf{H}^{-1}\mathbf{G}_2^\top \end{bmatrix}.$$

Based on the expansion above, observe that $(\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top)_\sigma = \mathbf{G}_2\mathbf{H}^{-1}\mathbf{G}_2^\top$. As a result, we may express $\mathbf{G}^\top \operatorname{adj}(\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top)_\sigma$, our matrix product of interest, as follows:

$$\mathbf{G}^\top \operatorname{adj}(\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top)_\sigma = \begin{bmatrix} \mathbf{G}_1^\top & \mathbf{G}_2^\top \end{bmatrix} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \operatorname{adj}(\mathbf{G}_2\mathbf{H}^{-1}\mathbf{G}_2^\top) \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{G}_2^\top \operatorname{adj}(\mathbf{G}_2\mathbf{H}^{-1}\mathbf{G}_2^\top) \end{bmatrix}.$$

All that remains is to show that $\mathbf{G}_2^\top \operatorname{adj}(\mathbf{G}_2\mathbf{H}^{-1}\mathbf{G}_2^\top) = \mathbf{0}$. To this end, we note that

$$\mathbf{G}_2^\top \operatorname{adj}(\mathbf{G}_2\mathbf{H}^{-1}\mathbf{G}_2^\top) = \mathbf{H}^{1/2}\left[\mathbf{H}^{-1/2}\mathbf{G}_2^\top \operatorname{adj}(\mathbf{G}_2\mathbf{H}^{-1/2}\mathbf{H}^{-1/2}\mathbf{G}_2^\top)\right] = \mathbf{0},$$

where the last equality follows immeditely by applying Lemma A.1.8. □

Lemma A.1.12. *For a positive semi-definite matrix* $\mathbf{A} \in \mathbb{R}^{m\times m}$*, a diagonal matrix* $\mathbf{\Lambda} = \operatorname{Diag}(\lambda)$*, and the indicator vector* $\boldsymbol{\sigma} \in \{0,1\}^m$*, define the following parameters:*

(i) Scaling factor $h_\sigma = \det(\mathbf{A}_\sigma)\prod_{i=1}^m \lambda_i^{1-\sigma_i}$*,*

(ii) Normalizing factor $h = \sum_{\sigma\in\{0,1\}^m} h_\sigma$*,*

(iii) Adjugate scaling factor $c_\sigma = \prod_{i=1}^m \lambda_i^{1-\sigma_i}$*.*

(iv) We split the set $\boldsymbol{\sigma} \in \{0,1\}^m$ *into the following two sets:*

$$S := \left\{\boldsymbol{\sigma} \in \{0,1\}^m \mid \det([\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top]_\sigma) > 0\right\}$$
$$S^\complement := \left\{\boldsymbol{\sigma} \in \{0,1\}^m \mid \det([\mathbf{G}\mathbf{H}^{-1}\mathbf{G}^\top]_\sigma) = 0\right\}.$$

Assume that $\mathbf{A} + \mathbf{\Lambda}$ *is invertible. Then we have the following decomposition of* $(\mathbf{A}+\mathbf{\Lambda})^{-1}$ *in* *terms of inverses and adjugates of* $\mathbf{A}_\sigma$ *(the principal submatrices of* $\mathbf{A}$*), with the adjugate or inverse computed on the basis of whether or not* $\det(\mathbf{A}_\sigma) = 0$*, as follows:*

$$(\mathbf{A}+\mathbf{\Lambda})^{-1} = \sum_{\sigma\in S} \frac{h_\sigma}{h}\mathbf{A}_\sigma^{-1} + \sum_{\sigma\in S^{\complement}} \frac{c_\sigma}{h}\,\mathrm{adj}(\mathbf{A})_\sigma. \tag{A.1.11}$$

Proof. We begin by proving the following simpler statement for $c_\sigma := \prod_{i=1}^m \lambda_i^{1-\sigma_i}$:

$$\mathrm{adj}(\mathbf{A}+\mathbf{\Lambda}) = \sum_{\sigma\in\{0,1\}^m} c_\sigma\,\mathrm{adj}(\mathbf{A})_\sigma. \tag{A.1.12}$$

Once this statement is proven, Equation (A.1.11) is implied by the following argument: Per Fact A.1.4, we have that $\det(\mathbf{A}+\mathbf{\Lambda}) = \sum_{\sigma\in\{0,1\}^m} h_\sigma = h$, so dividing throughout by h yields $(\mathbf{A}+\mathbf{\Lambda})^{-1}$ on the left-hand side (by Equation (A.1.5)); the term $\sum_{\sigma\in\{0,1\}^m} \frac{c_\sigma}{h}\,\mathrm{adj}(\mathbf{A})_\sigma$ may be split into two sums of terms, one over those vectors $\boldsymbol{\sigma}\in\{0,1\}^m$ for which $\det(\mathbf{A}_\sigma)=0$ and the second over those choices of $\boldsymbol{\sigma}\in\{0,1\}^m$ for which $\det(\mathbf{A}_\sigma)\neq 0$. For terms such that $\det(\mathbf{A}_\sigma)\neq 0$, we have,

$$c_\sigma\,\mathrm{adj}(\mathbf{A})_\sigma = c_\sigma \det([\mathbf{A}]_\sigma)\frac{1}{\det([\mathbf{A}]_\sigma)}\,\mathrm{adj}(\mathbf{A})_\sigma = h_\sigma\cdot(\mathbf{A})_\sigma^{-1}.$$

Hence, Equation (A.1.12), when divided by h, gives Equation (A.1.11), as desired. We now prove Equation (A.1.12), proceeding via induction on $\mathbf{nnz}(\mathbf{\Lambda})$, the number of nonzero entries in $\mathbf{\Lambda}$.

Base case: When the number of non-zero entries $\mathbf{nnz}(\mathbf{\Lambda}) = 0$, by definition, $\mathbf{\Lambda} = \mathrm{Diag}(\mathbf{0})$, which implies that the left-hand side of Equation (A.1.12) is $\mathrm{adj}(\mathbf{A})$. Further, since by definition, $c_\sigma = \prod_{i=1}^m \lambda_i^{1-\sigma_i}$, for our choice of $\mathbf{\Lambda} = \mathrm{Diag}(\mathbf{0})$, this gives the following expression:

$$c_\sigma = \begin{cases} 0 & \text{if } \boldsymbol{\sigma}\neq\mathbf{1}, \\ 1 & \text{if } \boldsymbol{\sigma}=\mathbf{1}. \end{cases}.$$

With this choice of c_σ, the right-hand side of Equation (A.1.12) reduces to $\mathrm{adj}(\mathbf{A})$, which matches the left-hand side of Equation (A.1.12), thus implying that in this base case, Equation (A.1.12) is true.

Induction Step: Suppose Equation (A.1.12) holds for $\mathbf{nnz}(\mathbf{\Lambda}) = k$. We now show that Equation (A.1.12) holds for $\mathbf{nnz}(\mathbf{\Lambda}) = k+1$ as well with some scaling factor $c_{\boldsymbol{\sigma}}$. Without loss of generality, assume that $\lambda_i \neq 0$ for $i \in [k+1]$. Let $\widetilde{\mathbf{A}}_{11}$ be the $\mathbb{R}^{(m-1)\times(m-1)}$ matrix obtained by deleting the first row and first column of $\mathbf{A}$. By expressing $\mathbf{A}+\mathbf{\Lambda}$ as $(\mathbf{A}+\sum_{i=2}^{k+1}\lambda_i\mathbf{e}_i\mathbf{e}_i^\top)+\lambda_1\mathbf{e}_1\mathbf{e}_1^\top$, we may use Corollary A.1.7 to expand $\operatorname{adj}(\mathbf{A}+\mathbf{\Lambda})$ as follows:

$$\operatorname{adj}(\mathbf{A}+\mathbf{\Lambda}) = \operatorname{adj}\left(\mathbf{A}+\sum_{i=2}^{k+1}\lambda_i\mathbf{e}_i\mathbf{e}_i^\top\right)+\lambda_1\begin{bmatrix}0 & \mathbf{0}^\top \\ \mathbf{0} & \operatorname{adj}\left(\widetilde{\mathbf{A}}_{11}+\sum_{i=1}^{k}\lambda_{i+1}\widetilde{\mathbf{e}}_i\widetilde{\mathbf{e}}_i^\top\right)\end{bmatrix}, \tag{A.1.13}$$

where note that the $\mathbf{e}_i \in \mathbb{R}^m$ and $\widetilde{\mathbf{e}}_i \in \mathbb{R}^{m-1}$. We observe that both the terms on the right-hand side have $\mathbf{nnz}(\mathbf{\Lambda})-1=k$ nonzero entries in their respective diagonal components. Hence, by our assumption, the induction hypothesis is applicable; therefore, suppose that by Equation (A.1.12),

$$\begin{aligned}\operatorname{adj}(\mathbf{A}+\sum_{i=2}^{\ell}\lambda_i\mathbf{e}_i\mathbf{e}_i^\top) &= \sum_{\boldsymbol{\sigma}\in\{0,1\}^m}\hat{c}_{\boldsymbol{\sigma}}\operatorname{adj}(\mathbf{A})_{\boldsymbol{\sigma}}, \\ \operatorname{adj}(\widetilde{\mathbf{A}}_{11}+\sum_{i=1}^{\ell-1}\lambda_{i+1}\widetilde{\mathbf{e}}_i\widetilde{\mathbf{e}}_i^\top) &= \sum_{\boldsymbol{\sigma}'\in\{0,1\}^{m-1}}\tilde{c}_{\boldsymbol{\sigma}'}\operatorname{adj}(\widetilde{\mathbf{A}}_{11})_{\boldsymbol{\sigma}'},\end{aligned} \tag{A.1.14}$$

where, based on the diagonal components in each of the terms on the left-hand side, the scaling factors on the respective right-hand sides are $\hat{c}_{\boldsymbol{\sigma}} = 0^{(1-\sigma_1)}\prod_{i=2}^{m}\lambda_i^{1-\sigma_i}$, and $\tilde{c}_{\boldsymbol{\sigma}'} = \prod_{i=1}^{m-1}\lambda_{i+1}^{1-\sigma_i'}$. As a consequence of these definitions, we can re-write the terms in Equation (A.1.14) using $c_{\boldsymbol{\sigma}}$ as follows. First, observe that $\hat{c}_{\boldsymbol{\sigma}}=0$ when $\sigma_1=0$ and $\hat{c}_{\boldsymbol{\sigma}}=c_{\boldsymbol{\sigma}}$ otherwise. This implies:

$$\sum_{\boldsymbol{\sigma}\in\{0,1\}^m}\tilde{c}_{\boldsymbol{\sigma}}\operatorname{adj}(\mathbf{A})_{\boldsymbol{\sigma}} = \sum_{\substack{\boldsymbol{\sigma}\in\{0,1\}^m,\\ \sigma_1=1}}c_{\boldsymbol{\sigma}}\operatorname{adj}(\mathbf{A})_{\boldsymbol{\sigma}}. \tag{A.1.15}$$

Next, observe that for the vector $\boldsymbol{\sigma}=[0;\boldsymbol{\sigma}']$ formed by concatenating zero with $\boldsymbol{\sigma}'$, we have

$c_\sigma = \lambda_1 \tilde{c}_{\sigma'}$. This implies the following chain of equations:

$$\lambda_1 \begin{bmatrix} 0 & \mathbf{0}^\top \\ \mathbf{0} & \sum_{\sigma' \in \{0,1\}^{m-1}} \tilde{c}_{\sigma'} \operatorname{adj}(\widetilde{\mathbf{A}}_{11})_{\sigma'} \end{bmatrix} = \sum_{\sigma' \in \{0,1\}^{m-1}} \lambda_1 \tilde{c}_{\sigma'} \operatorname{adj}(\mathbf{A})_{\left[\begin{smallmatrix} 0 \\ \sigma' \end{smallmatrix}\right]} = \sum_{\substack{\sigma \in \{0,1\}^m, \\ \sigma_0 = 0}} c_\sigma \operatorname{adj}(\mathbf{A})_\sigma, \tag{A.1.16}$$

where in the first step, we used the fact that $\widetilde{\mathbf{A}}_{11}$ is, by definition, the principal submatrix of $\mathbf{A}$ obtained by deleting its first row and first column; the second step is by our prior observation connecting c_σ and $\tilde{c}_{\sigma'}$. Plugging the right-hand sides from Equation (A.1.14) into that of Equation (A.1.13) and then applying Equation (A.1.15) and Equation (A.1.16) gives

$$\begin{aligned}
\operatorname{adj}(\mathbf{A} + \mathbf{\Lambda}) &= \sum_{\sigma \in \{0,1\}^m} \hat{c}_\sigma \operatorname{adj}(\mathbf{A})_\sigma + \lambda_1 \begin{bmatrix} 0 & \mathbf{0}^\top \\ \mathbf{0} & \sum_{\sigma' \in \{0,1\}^{m-1}} \tilde{c}_{\sigma'} \operatorname{adj}(\widetilde{\mathbf{A}}_{11})_{\sigma'} \end{bmatrix} \\
&= \sum_{\substack{\sigma \in \{0,1\}^m, \\ \sigma_0 = 1}} c_\sigma \operatorname{adj}(\mathbf{A})_\sigma + \sum_{\substack{\sigma \in \{0,1\}^m, \\ \sigma_0 = 0}} c_\sigma \operatorname{adj}(\mathbf{A})_\sigma \\
&= \sum_{\sigma \in \{0,1\}^m} c_\sigma \operatorname{adj}(\mathbf{A})_\sigma.
\end{aligned}$$

Thus, we have shown Equation (A.1.12) for $\mathbf{nnz}(\mathbf{\Lambda}) = k+1$, thereby completing the induction and concluding the proof of Equation (A.1.12) and, consequently, of the stated lemma. □

A.2 Technical Lemmas in Convex Analysis

Fact A.2.1 ([NN94]). *Let Φ be a ν-self-concordant barrier. Then for any $\mathbf{x} \in$ dom(Φ) and $\mathbf{y} \in$ cl(dom)(Φ),*

$$\nabla \Phi(\mathbf{x})^\top (\mathbf{y} - \mathbf{x}) \leq \nu.$$

Lemma A.2.2 ([GLP+24]). *If f is a self-concordant barrier for a set $\mathcal{K} \subseteq \mathcal{B}(0, R)$, then for any $x \in \mathcal{K}$, we have*

$$\nabla^2 f(\mathbf{x}) \succeq \frac{1}{9R^2} \mathbf{I}.$$

Proof. For the sake of contradiction, suppose $\nabla^2 f \not\succeq \frac{1}{9R^2} \mathbf{I}$. This is equivalent to, for some

$x \in \mathcal{K}$ and unit vector $\mathbf{u}$,

$$(3R\mathbf{u})^\top \nabla^2 f(\mathbf{x})(3R\mathbf{u}) < 1. \tag{A.2.1}$$

Define the unit-radius Dikin ellipsoid around $\mathbf{x}$ as

$$\mathcal{E}_{\mathbf{x}}(\mathbf{x}, 1) = \left\{\mathbf{y} : (\mathbf{y} - \mathbf{x})^\top \nabla^2 f(\mathbf{x})(\mathbf{y} - \mathbf{x}) \leq 1\right\}.$$

Then, Inequality (A.2.1) is equivalent to the assertion that $\mathbf{x} + 3R\mathbf{u} \in \mathcal{E}_{\mathbf{x}}(\mathbf{x}, 1)$. Because f is self-concordant we have $\mathcal{E}_{\mathbf{x}}(\mathbf{x}, 1) \subseteq \mathcal{K}$ (see, e.g., [NN94, Theorem 2.1.1]). This, combined with $\mathbf{x} + 3R\mathbf{u} \in \mathcal{E}_{\mathbf{x}}(\mathbf{x}, 1)$, implies $\mathbf{x} + 3R\mathbf{u} \in \mathcal{K}$. However, since $\mathcal{K} \subseteq \mathcal{B}(0, R)$ and $x \in \mathcal{K}$ by construction, the inclusion $\mathbf{x} + 3R\mathbf{u} \in \mathcal{K}$ cannot hold for any unit vector $\mathbf{u}$, which implies that our initial assumption must be false, thus concluding the proof. □

Lemma A.2.3 ([GLP+24]). *If f is a ν-self-concordant barrier for a given convex set $\mathcal{K}$ then $g(\mathbf{x}) = c^\top \mathbf{x} + f(\mathbf{x})$ is a self-concordant barrier over $\mathcal{K}$. Further, if $\mathcal{K} \subseteq \mathcal{B}(0, R)$, then g has self-concordance parameter at most*

$$20(\nu + R^2\|\mathbf{c}\|^2).$$

Proof. Since $\nabla^2 g = \nabla^2 f$, we can conclude that g is also a self-concordant function. Since $\mathcal{K} \subseteq \mathcal{B}(0, R)$, Lemma A.2.2 applies, and we have $\nabla^2 f(\mathbf{x}) \succeq \frac{1}{9R^2}\mathbf{I}$ for all $x \in \mathcal{K}$. Equivalently,

$$\nabla^2 f(\mathbf{x})^{-1} \preceq 9R^2\mathbf{I} \text{ for all } x \in \mathcal{K}. \tag{A.2.2}$$

The self-concordance parameter (see Definition 3.0.1) of g is:

$$\|\nabla g(\mathbf{x})\|^2_{\nabla^2 g(\mathbf{x})^{-1}} = \|\mathbf{c} + \nabla f(\mathbf{x})\|^2_{\nabla^2 f(\mathbf{x})^{-1}} \leq 2\|\mathbf{c}\|^2_{\nabla^2 f(\mathbf{x})^{-1}} + 2\|\nabla f(\mathbf{x})\|^2_{\nabla^2 f(\mathbf{x})^{-1}}, \tag{A.2.3}$$

where the first step is by definition of self-concordance parameter of g. To finish the proof, we recall that $\|\mathbf{c}\|^2_{\nabla^2 f(\mathbf{x})^{-1}} \leq 9R^2\|\mathbf{c}\|_2^2$ by Inequality (A.2.2), and $\|\nabla f(\mathbf{x})\|^2_{\nabla^2 f(\mathbf{x})^{-1}} \leq \nu$ by the self-concordance parameter of f and put these bounds into Inequality (A.2.3). □

Our main result is based on the following result from [ZLY23].

Fact A.2.4 ([ZLY23], Theorem 2)**.** *Fix a vector* $\mathbf{c}$*, a polytope* $\mathcal{K}$*, and a point* $\mathbf{v}$*. We assume that the polytope* $\mathcal{K}$ *contains a full-dimensional ball of radius* r*. Let* $\mathbf{v}^\star = \arg\min_{\mathbf{u}\in\mathcal{K}} \mathbf{c}^\top \mathbf{u}$*. We define, for* $\mathbf{c}$*,*

$$\text{gap}(\mathbf{v}) = \mathbf{c}^\top(\mathbf{v} - \mathbf{v}^\star). \tag{A.2.4}$$

Further, define $\mathbf{v}_\eta = \arg\min_{\mathbf{v}} \mathbf{c}^\top \mathbf{v} + \eta\phi_{\mathcal{K}}(\mathbf{v})$*, where* $\phi_{\mathcal{K}}$ *is a self-concordant barrier on* $\mathcal{K}$*. Then we have the following lower bound on this suboptimality gap evaluated at* $\mathbf{v}_\eta$*:*

$$\min\left\{\frac{\eta}{2}, \frac{r\|\mathbf{c}\|}{2\nu + 4\sqrt{\nu}}\right\} \le \text{gap}(\mathbf{v}_\eta) = \mathbf{c}^\top(\mathbf{v}_\eta - \mathbf{v}^\star) \le \eta\nu. \tag{A.2.5}$$

A.2.1 Warmup: One-Dimensional Optimization

We begin with a lemma on optimizing quadratics in one dimension to motivate our later results for arbitrary polytopes in higher dimensions.

Lemma A.2.5. *Let* ϕ *be a* ν*-self-concordant barrier over* $(0, r)$ *and* q *be a convex function such that* $\nabla q(v) = 0$ *and* $0 < m \le \nabla^2 q(x) \le M$*. Define,*

$$x^\eta := \arg\min_x q(x) + \eta\phi(x).$$

Then,

$$\min\left\{\frac{1}{2}\left(\sqrt{\frac{2\eta}{M} + v^2} + v\right), \frac{r}{2\nu + 4\sqrt{\nu}}\right\} \le x^\eta \le \frac{1}{2}\left(\sqrt{\frac{4\eta\nu}{m} + v^2} + v\right). \tag{A.2.6}$$

Proof. Let $c := \nabla q(x^\eta)$. Note that we can equivalently write $x^\eta = \arg\min_x cx + \eta\phi(x)$. Let $x^\star := \arg\min_{x\in(0,r)} cx$ and $\tilde{x} := \arg\min_x \phi(x)$. We note that x^η must always be between $x^\star$ and $\tilde{x}$.

Case 1: $v < \tilde{x}$. Then $v < x^\eta < \tilde{x}$, meaning $c > 0$ and therefore $x^\star = 0$.

Applying Fact A.2.4, we have:

$$\min\left\{\frac{\eta}{2}, \frac{rc}{2\nu + 4\sqrt{\nu}}\right\} \leq cx^{\eta} \leq \eta\nu.$$

Using that $m(x^{\eta} - v) \leq c \leq M(x - v)$ we have,

$$\min\left\{\frac{\eta}{2M}, \frac{r(x^{\eta} - v)}{2\nu + 4\sqrt{\nu}}\right\} \leq (x^{\eta} - v)x^{\eta}$$
$$(x^{\eta} - v)x^{\eta} \leq \frac{\eta\nu}{m}.$$

Solving, $\frac{\eta}{2M} \leq (x^{\eta} - v)x^{\eta} \leq \frac{\nu\eta}{m}$ with the condition that $x^{\eta} > v$, we have,

$$\frac{1}{2}\left(\sqrt{\frac{2\eta}{M} + v^2} + v\right) \leq x^{\eta} \leq \frac{1}{2}\left(\sqrt{\frac{4\eta\nu}{m} + v^2} + v\right).$$

Combining with the minimum on the LHS bound, we arrive at,

$$\min\left\{\frac{1}{2}\left(\sqrt{\frac{2\eta}{M} + v^2} + v\right), \frac{r}{2\nu + 4\sqrt{\nu}}\right\} \leq x^{\eta} \leq \frac{1}{2}\left(\sqrt{\frac{4\eta\nu}{m} + v^2} + v\right).$$

Case 2: $v \geq \tilde{x}$. Then $\tilde{x} \leq x^{\eta} \leq v$. Note that by applying Fact A.2.4 with $c = 1$ and considering $\eta \to \infty$, we can deduce that $\tilde{x} \geq \frac{r}{2\nu + 4\sqrt{\nu}}$. We can see that Equation (A.2.6) still holds as,

$$\min\left\{\frac{1}{2}\left(\sqrt{\frac{\eta}{M} + v^2} + v\right), \frac{r}{(2\nu + 4\sqrt{\nu})}\right\} \leq \frac{r}{2\nu + 4\sqrt{\nu}} \leq \tilde{x} \leq x^{\eta} \leq v \leq \frac{1}{2}\left(\sqrt{\frac{4\eta\nu}{m} + v^2} + v\right).$$

□

The above result shows that if the minimizer of a strongly convex cost lies outside of the constraint set, we should expect to get a bound of the form $O(\sqrt{\eta + v^2} - v)$, where v is the distance to the constraint set.

A.2.2 Upper Bounds on Approximation Error for Interior Point Methods

Lemma A.2.6. *Let* $\mathcal{K} = \{\mathbf{x} : \mathbf{A}\mathbf{x} \geq \mathbf{b}\}$ *be a polytope such that each of* m *rows of* $\mathbf{A}$ *is normalized to the unit norm and* ϕ *be a* ν*-self-concordant barrier. Let* $\mathbf{x}^\eta := \arg\min_{\mathbf{x}} \frac{\alpha}{2}\|\mathbf{x} - \mathbf{v}\|^2 + \eta\phi(\mathbf{x})$, $\mathbf{x}^\star := \arg\min_{\mathbf{x}\in\mathcal{K}} \frac{\alpha}{2}\|\mathbf{x} - \mathbf{v}\|^2$*. Then,*

$$\|\mathbf{x}^\eta - \mathbf{x}^\star\| \leq \sqrt{\frac{\eta\nu}{\alpha}}.$$

Proof. We proceed similar to Theorem 3.1.1. Note that by α-strong-convexity of $q(\mathbf{x}) := \frac{\alpha}{2}\|\mathbf{x} - \mathbf{v}\|^2$, we have that,

$$[\nabla q(\mathbf{x}^\eta) - \nabla q(\mathbf{x}^\star)]^\top(\mathbf{x}^\eta - \mathbf{x}^\star) \geq m\|\mathbf{x} - \mathbf{v}\|^2.$$

Note that from the optimality condition $\nabla q(\mathbf{x}^\eta) + \eta\nabla\phi(\mathbf{x}^\eta) = 0$, and that $\nabla q(\mathbf{x}^\star)^\top[\mathbf{x}^\eta - \mathbf{x}^\star] \geq 0$, and by Fact A.2.1, it follows,

$$\alpha\|\mathbf{x}^\eta - \mathbf{x}^\star\|^2 \leq \eta\nabla\phi(\mathbf{x}^\eta)[\mathbf{x}^\star - \mathbf{x}^\eta] - \nabla q(\mathbf{x}^\star)^\top[\mathbf{x}^\star - \mathbf{x}^\star] \leq \phi(\mathbf{x}^\eta)^\top[\mathbf{x}^\star - \mathbf{x}^\eta] \leq \eta\nu.$$

Therefore we have that,

$$\|\mathbf{x}^\eta - \mathbf{x}^\star\| \leq \sqrt{\frac{\eta\nu}{\alpha}}.$$

□

Note that the above result can easily be generalized to α-strongly-convex functions.

In the next lemma, we show that we can make a similar bound along the gradient of the cost function at $\mathbf{x}^\star$:

Lemma A.2.7. *Let* $\mathcal{K} = \{\mathbf{x} : \mathbf{A}\mathbf{x} \geq \mathbf{b}\}$ *be a polytope such that each of* m *rows of* $\mathbf{A}$ *is normalized to the unit norm and* ϕ *be a* ν*-self-concordant barrier function. Let* $\mathbf{x}^\eta := \arg\min_{\mathbf{x}} \frac{\alpha}{2}\|\mathbf{x} - \mathbf{v}\|^2 + \eta\phi(\mathbf{x})$, $\mathbf{x}^\star := \arg\min_{\mathbf{x}\in\mathcal{K}} \frac{\alpha}{2}\|\mathbf{x} - \mathbf{v}\|^2$*. Assume that* $\mathbf{x}^\star \neq \mathbf{v}$ *and let* $\mathbf{a} = \frac{\mathbf{x}^\star - \mathbf{v}}{\|\mathbf{x}^\star - \mathbf{v}\|}$*. Then,*

$$0 \le \mathbf{a}^\top(\mathbf{x}^\eta - \mathbf{x}^\star) \le \frac{1}{2}\left(\sqrt{\frac{4\eta\nu}{\alpha} + \|\mathbf{x}^\star - \mathbf{v}\|^2} - \|\mathbf{x}^\star - \mathbf{v}\|\right).$$

Proof. Note that we can write:

$$\nabla q(\mathbf{x}^\eta) = \alpha \cdot \mathbf{a}^\top(\mathbf{x}^\eta - \mathbf{x}^\star)\mathbf{a} + \alpha \cdot \mathbf{b}^\top(\mathbf{x}^\eta - \mathbf{x}^\star)\mathbf{b} + \nabla q(\mathbf{x}^\star),$$

where $\|\mathbf{b}\| = 1$ and $\mathbf{b} \perp \mathbf{a}$. By Fact A.2.4 we have,

$$\nabla q(\mathbf{x}^\eta)^\top(\mathbf{x}^\eta - \mathbf{x}^\star) \le \eta\nu.$$

Then it follows that

$$\alpha \cdot [\mathbf{a}^\top(\mathbf{x}^\eta - \mathbf{x}^\star)]^2 + \alpha[\mathbf{b}^\top(\mathbf{x}^\eta - \mathbf{x}^\star)]^2 + \alpha\|\mathbf{x}^\star - \mathbf{v}\| \cdot \mathbf{a}^\top(\mathbf{x}^\eta - \mathbf{x}^\star) \le \eta\nu.$$

We can drop the $\alpha \cdot [b^\top(\mathbf{x}^\eta - \mathbf{x}^\star)]^2$ term and, as in Lemma A.2.5, solve for $\mathbf{a}^\top(\mathbf{x}^\eta - \mathbf{x}^\star)$ to arrive at our conclusion,

$$0 \le \mathbf{a}^\top(\mathbf{x}^\eta - \mathbf{x}^\star) \le \frac{1}{2}\left(\sqrt{\frac{4\eta\nu}{\alpha} + \|\mathbf{x}^\star - \mathbf{v}\|^2} - \|\mathbf{x}^\star - \mathbf{v}\|\right).$$

□

A.2.3 Lower Bounds on Residuals for Interior-Point Methods

Lemma A.2.8. *Fix a polytope $\mathcal{K}$, a convex function q, and a ν-self-concordant barrier ϕ over $\mathcal{K}$. Assume that the polytope $\mathcal{K}$ contains a full-dimensional ball of radius r and is contained within a ball of radius R around some point $\bar{\mathbf{x}}$, i.e. $\mathcal{B}(\bar{\mathbf{x}}, r) \subseteq \mathcal{K} \subseteq \mathcal{B}(\bar{\mathbf{x}}, R)$. Let $\mathbf{x}^\eta := \arg\min q(\mathbf{x}) + \eta\phi(\mathbf{x})$ for arbitrary $\eta > 0$,*

$$\mathcal{B}\left(\mathbf{x}^\eta, \frac{r}{R}\min\left\{\frac{\eta}{2\|\nabla q(\mathbf{x}^\eta)\|}, \frac{r}{2\nu + 4\sqrt{\nu}}\right\}\right) \subset \mathcal{K}.$$

Proof. Let $\bar{\mathbf{x}}$ be the center of the r-radius ball contained within $\mathcal{K}$. Consider the line passing through $\bar{\mathbf{x}}, \mathbf{x}^\eta$ given by $\mathcal{S} = \{\bar{\mathbf{x}}t + \mathbf{x}^\eta(1-t) : t\}$ and let $\mathbf{x}_1, \mathbf{x}_2$ be the endpoints of $\mathcal{K} \cap \mathcal{S}$. We will now show that,

$$\min(\|\mathbf{x}^\eta - \mathbf{x}_1\|, \|\mathbf{x}^\eta - \mathbf{x}_2\|) \geq \min\left\{\frac{\eta}{4\|\nabla q(\mathbf{x}^\eta)\|}, \frac{r}{2\nu + 4\sqrt{\nu}}\right\}.$$

WLOG let $\mathbf{x}_1$ be such that $\nabla q(\mathbf{x}^\eta)^\top \left(\frac{\bar{\mathbf{x}}-\mathbf{x}_1}{\|\bar{\mathbf{x}}-\mathbf{x}_1\|}\right) \geq 0$. Note that the analytic center along $\mathcal{S}$ is contained within a ball of radius $\frac{r}{2\nu+4\sqrt{\nu}}$ and that $\mathbf{x}^\eta$ must be further from $\mathbf{x}_2$ than the analytic center, implying that $\|\mathbf{x}^\eta - \mathbf{x}_2\| \geq \frac{r}{2\nu+4\sqrt{\nu}}$.

We then parameterize $\mathcal{S}$ with $\psi(t) = \mathbf{x}_1 + t\frac{\bar{\mathbf{x}}-\mathbf{x}_1}{\|\bar{\mathbf{x}}-\mathbf{x}_1\|}$. Note that $\psi([0, 2r]) \subset \mathcal{K}$. Consider the following optimization problem:

$$t^\eta = \arg\min_t \nabla q(\mathbf{x}^\eta)^\top \left(\frac{\bar{\mathbf{x}} - \mathbf{x}_1}{\|\bar{\mathbf{x}} - \mathbf{x}_1\|}\right) t + \eta\phi(\psi(t)).$$

We picked $\mathbf{x}_1$ such that $c := \nabla q(\mathbf{x}^\eta)^\top \left(\frac{\bar{\mathbf{x}}-\hat{\mathbf{x}}}{\|\bar{\mathbf{x}}-\hat{\mathbf{x}}\|}\right) \geq 0$. It follows that,

$$t^\star := \arg\min_{t, \phi(t)\in\mathcal{K}} c \cdot t = 0.$$

We then apply Fact A.2.4 to the above one-dimensional optimization problem to conclude that,

$$\min\left\{\frac{\eta}{2\|\nabla q(\mathbf{x}^\eta)\|}, \frac{r}{2\nu + 4\sqrt{\nu}}\right\} \leq \min\left\{\frac{\eta}{2c}, \frac{r}{2\nu + 4\sqrt{\nu}}\right\} \leq t^\eta = \|\mathbf{x}^\eta - \mathbf{x}_1\|.$$

Pick $\hat{\mathbf{x}} \in \{\mathbf{x}_1, \mathbf{x}_2\}$ such that $\mathbf{x}^\eta$ lies along the segment between $\hat{\mathbf{x}}$ and $\bar{\mathbf{x}}$. Consider any direction $\mathbf{c}$, $\|\mathbf{c}\| = 1$. We then argue by similarlity on the triangle $\hat{\mathbf{x}}, \bar{\mathbf{x}}, \bar{\mathbf{x}} + r\mathbf{c}$ and note that $\frac{\|r\mathbf{c}\|}{\|\bar{\mathbf{x}}-\hat{\mathbf{x}}\|} \leq \frac{r}{R}$. Since $\min\left\{\frac{\eta}{2\|\nabla q(\mathbf{x}^\eta)\|}, \frac{r}{2\nu+4\sqrt{\nu}}\right\} \leq \|\mathbf{x}^\eta - \hat{\mathbf{x}}\| \leq \|\bar{\mathbf{x}} - \hat{\mathbf{x}}\|$, we know that,

$$\mathbf{x}^\eta + \frac{r}{R}\min\left\{\frac{\eta}{2\|\nabla q(\mathbf{x}^\eta)\|}, \frac{r}{2\nu + 4\sqrt{\nu}}\right\} \cdot \mathbf{c} \in \mathcal{K}.$$

□

Similar to Lemma A.2.7, we now adapt this to get a lower bound for isotropic quadratics.

Lemma A.2.9. *Let $\mathcal{K} = \{\mathbf{x} : \mathbf{A}\mathbf{x} \geq \mathbf{b}\}$ be a polytope such that each of m rows of $\mathbf{A}$ is normalized to the unit norm and ϕ be a ν-self-concordant barrier function. Assume there* *exists* $\bar{\mathbf{x}} \in \mathcal{K}$ *such that* $\mathcal{B}(\bar{\mathbf{x}}, r) \subseteq \mathcal{K} \subseteq \mathcal{B}(\bar{\mathbf{x}}, R)$ *for some* $r, R > 0$*. Let* $\mathbf{x}^\eta := \arg\min_{\mathbf{x}} \frac{\alpha}{2}\|\mathbf{x} - \mathbf{v}\|^2 + \eta\phi(\mathbf{x})$*,* $\mathbf{x}^\star := \arg\min_{\mathbf{x}\in\mathcal{K}} \frac{\alpha}{2}\|\mathbf{x} - \mathbf{v}\|^2$*. Then we know the following ball centered around* $\mathbf{x}^\eta$ *is contained within* $\mathcal{K}$*.*

$$\mathcal{B}\left(\mathbf{x}^\eta, \frac{r}{R}\min\left\{\frac{1}{\sqrt{\nu}}\left(\sqrt{\frac{\eta}{\alpha} + \|\mathbf{x}^\star - \mathbf{v}\|^2} - \|\mathbf{x}^\star - \mathbf{v}\|\right), \frac{r}{2\nu + 4\sqrt{\nu}}\right\}\right) \subseteq \mathcal{K}.$$

Proof. To prove this we use Lemma A.2.8 and techniques similar to Lemma A.2.7. Note that,

$$\begin{aligned}\|\nabla q(\mathbf{x}^\eta)\| &= \alpha\|\mathbf{x}^\eta - \mathbf{v}\| \\ &\leq \alpha(\|\mathbf{x}^\eta - \mathbf{x}^\star\| + \|\mathbf{x}^\star - \mathbf{v}\|) \\ &\leq \sqrt{\alpha}\sqrt{\eta\nu} + \alpha\|\mathbf{x}^\star - \mathbf{v}\|.\end{aligned}$$

where the last inequality uses Lemma A.2.6. This implies that,

$$\mathcal{B}\left(\mathbf{x}^\eta, \frac{r}{R}\min\left\{\frac{\eta}{\sqrt{\alpha\eta}\sqrt{\nu} + \alpha\|\mathbf{x}^\star - \mathbf{v}\|}, \frac{r}{2\nu + 4\sqrt{\nu}}\right\}\right) \subseteq \mathcal{K}.$$

With some rearranging:

$$\mathcal{B}\left(\mathbf{x}^\eta, \frac{r}{R}\min\left\{\frac{1}{\sqrt{\nu}}\frac{\frac{\eta}{\alpha}}{\sqrt{\frac{\eta}{\alpha}} + \frac{1}{\sqrt{\nu}}\|\mathbf{x}^\star - \mathbf{v}\|}, \frac{r}{2\nu + 4\sqrt{\nu}}\right\}\right) \subseteq \mathcal{K}.$$

Observe that for any $x > 0, y \in \mathbb{R}$, we have that,

$$\sqrt{x + y^2} - y = \frac{x}{\sqrt{x + y^2} + y} \leq \frac{x}{\sqrt{x} + y}.$$

Since $\nu > 1$, $\frac{\eta/\alpha}{\sqrt{\eta/\alpha} + \frac{1}{\sqrt{\nu}}\|\mathbf{x}^\star - \mathbf{v}\|} \geq \frac{\eta/\alpha}{\sqrt{\eta/\alpha} + \alpha\|\mathbf{x}^\star + \mathbf{v}\|} \geq \sqrt{\frac{\eta}{\alpha} + \|\mathbf{x}^\star - \mathbf{v}\|^2} - \|\mathbf{x}^\star - \mathbf{v}\|$.

$$\mathcal{B}\left(\mathbf{x}^\eta, \frac{r}{R}\min\left\{\frac{1}{\sqrt{\nu}}\left(\sqrt{\frac{\eta}{\alpha} + \|\mathbf{x}^\star - \mathbf{v}\|^2} - \|\mathbf{x}^\star - \mathbf{v}\|\right), \frac{r}{2\nu + 4\sqrt{\nu}}\right\}\right) \subseteq \mathcal{K}.$$

□

A.2.4 Consolidated Upper and Lower Bounds

We now collect Lemma A.2.7, and Lemma A.2.9, performing a change of basis to provide bounds for arbitrary quadratic objective functions.

Theorem A.2.10. *Let $\mathcal{K} = \{\mathbf{x} : \mathbf{A}\mathbf{x} \geq \mathbf{b}\}$ be a polytope such that each of m rows of $\mathbf{A}$ is normalized to the unit norm and ϕ be a ν-self-concordant barrier over $\mathcal{K}$. Assume there exists $\bar{\mathbf{x}} \in \mathcal{K}$ such that $\mathcal{B}(\bar{\mathbf{x}}, r) \subseteq \mathcal{K} \subseteq \mathcal{B}(\bar{\mathbf{x}}, R)$ for some $r, R > 0$. Let,*

$$\mathbf{x}^\eta := \arg\min_{\mathbf{x}} \frac{1}{2}(\mathbf{x}-\mathbf{v})^\top \mathbf{H}(\mathbf{x}-\mathbf{v}) + \eta\phi(\mathbf{x}),$$
$$\mathbf{x}^\star := \arg\min_{\mathbf{x}\in\mathcal{K}} \frac{1}{2}(\mathbf{x}-\mathbf{v})^\top \mathbf{H}(\mathbf{x}-\mathbf{v}),$$

where $m\mathbf{I} \preceq \mathbf{H} \preceq M\mathbf{I}$. Then,

$$\|\mathbf{x}^\eta - \mathbf{x}^\star\| \leq \sqrt{\frac{\eta\nu}{m}},$$
$$0 \leq \mathbf{a}^\top(\mathbf{x}^\eta - \mathbf{x}^\star) \leq \frac{1}{2}\left(\sqrt{\frac{4\eta\nu}{m} + \|\mathbf{x}^\star - \mathbf{v}\|_{\mathbf{H}}^2} - \|\mathbf{x}^\star - \mathbf{v}\|_{\mathbf{H}}\right).$$

where the second inequality holds for $\mathbf{a} = \frac{\mathbf{H}(\mathbf{x}^\star - \mathbf{v})}{\|\mathbf{H}(\mathbf{x}^\star - \mathbf{v})\|}$ if $\|\mathbf{x}^\star - \mathbf{v}\| > 0$ and $\|\mathbf{x}\|_{\mathbf{H}} = \sqrt{\mathbf{x}^\top\mathbf{H}\mathbf{x}}$. Furthermore,

$$\mathcal{B}\left(\mathbf{x}^\eta, \frac{r}{R}\min\left\{\frac{1}{\sqrt{\nu}}\left(\sqrt{\frac{\eta}{M} + \|\mathbf{x}^\star - \mathbf{v}\|_{\mathbf{H}}^2} - \|\mathbf{x}^\star - \mathbf{v}\|_{\mathbf{H}}\right), \frac{r}{2\nu + 4\sqrt{\nu}}\right\}\right) \subseteq \mathcal{K}.$$

Note that this implies that, if $\mathbf{a}$ exists, since $\mathbf{x}^\star$ is on the boundary of $\mathcal{K}$

$$\frac{r}{R}\min\left\{\frac{1}{\sqrt{\nu}}\left(\sqrt{\frac{\eta}{M} + \|\mathbf{x}^\star - \mathbf{v}\|_{\mathbf{H}}^2} - \|\mathbf{x}^\star - \mathbf{v}\|_{\mathbf{H}}\right), \frac{r}{2\nu + 4\sqrt{\nu}}\right\} \leq \mathbf{a}^\top(\mathbf{x}^\eta - \mathbf{x}^\star).$$

Proof. For the upper bound we use the change of basis $\mathbf{y} = \frac{1}{\sqrt{m}}\mathbf{H}^{1/2}\mathbf{x}$, $\mathbf{z} = \frac{1}{\sqrt{m}}\mathbf{H}^{1/2}\mathbf{v}$. We then have the optimization problem in $\mathbf{y}$:

$$\mathbf{y}^\eta := \arg\min_{\mathbf{y}} \frac{m}{2}\|\mathbf{y}-\mathbf{z}\|^2 + \eta\phi(\mathbf{y}),$$
$$\mathbf{y}^\star := \arg\min_{\mathbf{y}\in\mathcal{K}} \frac{m}{2}\|\mathbf{y}-\mathbf{z}\|^2.$$

By Lemma A.2.6, and Lemma A.2.7 we have that,

$$\|\mathbf{x}^\eta - \mathbf{x}^\star\| \le \|\mathbf{y}^\eta - \mathbf{y}^\star\| \le \sqrt{\frac{\eta\nu}{m}},$$
$$0 \le \tilde{\mathbf{a}}^\top(\mathbf{y}^\eta - \mathbf{y}^\star) \le \frac{1}{2}\left(\sqrt{\frac{4\eta\nu}{m} + \|\mathbf{y}^\star - \mathbf{z}\|^2} - \|\mathbf{y}^\star - \mathbf{z}\|\right).$$

where $\tilde{\mathbf{a}} := \frac{\mathbf{y}^\star - \mathbf{z}}{\|\mathbf{y}^\star - \mathbf{z}\|}$. Let $\hat{\mathbf{a}} := \frac{1}{\sqrt{m}} \cdot \mathbf{H}^{1/2}\tilde{\mathbf{a}}$. Since $\sqrt{m}$ is the smallest singular value of $\mathbf{H}^{1/2}$, we have that $\|\hat{\mathbf{a}}\| \ge 1$. Let $\mathbf{a} = \hat{\mathbf{a}}/\|\hat{\mathbf{a}}\| = \frac{\mathbf{H}(\mathbf{x}^\star - \mathbf{v})}{\|\mathbf{H}(\mathbf{x}^\star - \mathbf{v})\|}$. By substitution,

$$0 \le \mathbf{a}^\top(\mathbf{x}^\eta - \mathbf{x}^\star) \le \hat{\mathbf{a}}^\top(\mathbf{x}^\eta - \mathbf{x}^\star) = \tilde{\mathbf{a}}^\top(\mathbf{y}^\eta - \mathbf{y}^\star) \le \frac{1}{2}\left(\sqrt{\frac{4\eta\nu}{m} + \|\mathbf{x}^\star - \mathbf{v}\|_{\mathbf{H}}^2} - \|\mathbf{x}^\star - \mathbf{v}\|_{\mathbf{H}}\right).$$

This shows the first two bounds.

For the lower bound we use the change of basis $\mathbf{y} = \frac{1}{\sqrt{M}}\mathbf{H}^{1/2}\mathbf{x}$, $\mathbf{z} = \frac{1}{\sqrt{M}}\mathbf{H}^{1/2}\mathbf{v}$. We then have the optimization problem in $\mathbf{y}$:

$$\mathbf{y}^\eta := \arg\min_{\mathbf{y}} \frac{M}{2}\|\mathbf{y} - \mathbf{z}\|^2 + \eta\phi(\mathbf{y}),$$
$$\mathbf{y}^\star := \arg\min_{\mathbf{y}\in\mathcal{K}} \frac{M}{2}\|\mathbf{y} - \mathbf{z}\|^2.$$

Since $\|\mathbf{H}^{1/2}\| \le \sqrt{M}$, the inverse transformation $\mathbf{x} = \sqrt{M}\mathbf{H}^{-1/2}\mathbf{y}$ has $\sigma_{\min}(\sqrt{M}\mathbf{H}^{-1/2}) \ge 1$, meaning by Lemma A.2.9

$$\mathcal{B}\left(\mathbf{x}^\eta, \frac{r}{R}\min\left\{\frac{1}{\sqrt{\nu}}\left(\sqrt{\frac{\eta}{M} + \|\mathbf{x}^\star - \mathbf{v}\|_{\mathbf{H}}^2} - \|\mathbf{x}^\star - \mathbf{v}\|_{\mathbf{H}}\right), \frac{r}{2\nu + 4\sqrt{\nu}}\right\}\right) \subseteq \mathcal{K}.$$

□