A Proof of Theorem 1

Recall that by assumption, the population function $\mathcal{C}_{\mathsf{init},\gamma}$ satisfies the following properties:

Assumption A:

(a) It has (ϕ_0, ρ_0)-locally Lipschitz gradients,

(b) It is (λ_0, ρ_0)-locally Lipschitz, and

(c) It is globally μ-PL.

Recall the values of the step-size η, smoothing radius r, and iteration complexity T posited by Theorem 1. For ease of exposition, it is helpful to run our stochastic zero-order method on this problem for $2T$ iterations; we thus obtain a (random) sequence of iterates $\{K_t\}_{t=0}^{2T}$. For each $t = 0, 1, 2, \ldots$, we define the cost error $\Delta_t = \mathcal{C}_{\mathsf{init},\gamma}(K_t) - \mathcal{C}_{\mathsf{init},\gamma}(K^*)$, as well as the stopping time

$$\tau := \min\Big\{t \mid \Delta_t > 10\Delta_0\Big\}. \tag{15}$$

In words, the time τ is the index of the first iterate that exits the bounded region $\mathcal{G}^0$. The gradient estimate g at the point $K \in \mathcal{G}^0$ (cf. equation (12) for defintion of the set $\mathcal{G}^0$) is assumed to satisfy the bounds

$$\mathsf{var}(\mathrm{g}(K)) \le G_2 \quad \text{and} \quad \|\mathrm{g}(K)\|_2 \le G_\infty \text{ almost surely.}$$

We provide the proof of the above bounds in Appendix A.2

With this set up in place, we now state and prove a proposition that is stronger than the assertion of Theorem 1.

Proposition 1. *With the parameter settings of Theorem 1, we have*

$$\mathbb{E}[\Delta_T 1_{\tau>T}] \le \epsilon/20,$$

and furthermore, the event $\{\tau > T\}$ occurs with probability greater than $4/5$.

Let us verify that Proposition 1 implies the claim of Theorem 1. We have

$$\begin{aligned}\mathbb{P}\{\Delta_T \ge \epsilon\} &\le \mathbb{P}\{\Delta_T 1_{\tau>T} \ge \epsilon\} + \mathbb{P}\{1_{\tau\le T}\} \\ &\overset{(i)}{\le} \frac{1}{\epsilon}\mathbb{E}[\Delta_T 1_{\tau>T}] + \mathbb{P}\{1_{\tau\le T}\} \\ &\overset{(ii)}{\le} 1/20 + 1/5 \\ &\le 1/4,\end{aligned}$$

where step (i) follows from Markov's inequality, and step (ii) from Proposition 1. Thus, Theorem 1 follows as a direct consequence of Proposition 1, and we dedicate the rest of the proof to establishing Proposition 1.

Let $\mathbb{E}^t$ to represent the expectation conditioned on the randomness up to time t. The following lemma bounds the progress of one step of the algorithm:

Lemma 4. *Given any function satisfying the previously stated properties, suppose that we run Algorithm 1 with smoothing radius $r \le \rho_0$, and with a step-size η such that $\|\eta g^t\|_2 \le \rho_0$ almost surely. Then for any $t = 0, 1, \ldots$ such that $K_t \in \mathcal{G}^0$, we have*

$$\mathbb{E}^t\left[\Delta_{t+1}\right] \le \Big(1 - \frac{\eta\mu}{4}\Big)\Delta_t + \frac{\phi_0\eta^2}{2}G_2 + \eta\mu\frac{\epsilon}{120}. \tag{16}$$

The proof of the lemma is postponed to Section B. Taking it as given, let us now establish Proposition 1.

Proposition 1 has two natural parts; let us focus first on proving the bound on the expectation. Let $\mathcal{F}_t$ denote the σ-field containing all the randomness in the first t iterates. Conditioning on this σ-field yields

$$\mathbb{E}[\Delta_{t+1}1_{\tau>t+1} \mid \mathcal{F}_t] \le \mathbb{E}[\Delta_{t+1}1_{\tau>t} \mid \mathcal{F}_t] \overset{(i)}{=} [\mathbb{E}[\Delta_{t+1} \mid \mathcal{F}_t]1_{\tau>t}],$$

where step (i) follows since τ is a stopping time, and so the random variable $1_{\tau>t}$ is determined completely by the sigma-field $\mathcal{F}_t$.

We now split the proof into two cases.

Case 1: Assume that $\tau > t$, so that we have the inclusion $K_t \in \mathcal{G}^0$. In addition, note that the iterate K_{t+1} is obtained after a stochastic zero-order step whose size is bounded as

$$\|\eta g^t\|_2 \;\le\; \eta G_\infty \le \rho_0,$$

where we have used the fact that $\eta \le \frac{\rho_0}{G_\infty}$.

We may thus apply Lemma 4 to obtain

$$\mathbb{E}[\Delta_{t+1} \;\mid\; \mathcal{F}_t] \le \Big(1 - \frac{\eta\mu}{4}\Big)\Delta_t + \frac{\phi_0\eta^2}{2} G_2 + \eta\mu\frac{\epsilon}{120}. \tag{17a}$$

Case 2: In this case, we have $\tau \le t$, so that

$$\mathbb{E}[\Delta_{t+1} \;\mid\; \mathcal{F}_t]1_{\tau>t} = 0. \tag{17b}$$

Now combining the bounds (17a) and (17b) from the the two cases yields the inequality

$$\begin{aligned}\mathbb{E}[\Delta_{t+1} \;\mid\; \mathcal{F}_t]1_{\tau>t} &\le \left\{\Big(1 - \frac{\eta\mu}{4}\Big)\Delta_t + \frac{\phi_0\eta^2}{2} G_2 + \eta\mu\frac{\epsilon}{120}\right\}1_{\tau>t} \\ &\le \left(1 - \frac{\eta\mu}{4}\right)\Delta_t 1_{\tau>t} + \frac{\phi_0\eta^2}{2} G_2 + \eta\mu\frac{\epsilon}{120}.\end{aligned} \tag{18}$$

Taking expectations over the sigma-field $\mathcal{F}_t$ and then arguing inductively yields

$$\begin{aligned}\mathbb{E}[\Delta_{t+1}1_{\tau>t+1}] &\le \Big(1 - \frac{\eta\mu}{4}\Big)^{t+1}\Delta_0 + \left(\frac{\phi_0\eta^2}{2} G_2 + \eta\mu\frac{\epsilon}{120}\right)\sum_{i=0}^{t}\Big(1 - \frac{\eta\mu}{4}\Big)^i \\ &\le \Big(1 - \frac{\eta\mu}{4}\Big)^{t+1}\Delta_0 + 2\frac{\eta}{\mu}\phi_0 G_2 + \frac{4\epsilon}{120}.\end{aligned}$$

Setting $t+1 = T$ then establishes the first part of the proposition with substitutions of the various parameters.

We now turn to establishing that $\mathbb{P}\{\tau > T\} \ge 4/5$. We do so by setting up a suitable super-martingale on our iterate sequence and appealing to classical maximal inequalities. Recall that we run the algorithm for $2T$ steps for convenience, and thereby obtain a set of $2T$ random variables $\{\Delta_1, \ldots, \Delta_{2T}\}$. With the stopping time τ defined as before (22), define the stopped process

$$Y_t := \Delta_{\tau\wedge t} + (2T - t)\left(\frac{\phi_0\eta^2}{2} G_2 + \eta\mu\frac{\epsilon}{120}\right) \qquad \text{for each } t \in [2T].$$

Note that by construction, each random variable Y_t is non-negative by definition and invoking the local Lipschitz property of the function $\mathcal{C}(\cdot, s_0)$ at time $\tau - 1$ we can ensure that the random variable Y_t is almost surely bounded.

We claim that $\{Y_t\}_{t=0}^{2T}$ is a super-martingale. In order to prove this claim, we first write

$$\mathbb{E}[Y_{t+1} \mid \mathcal{F}_t] = \mathbb{E}[\Delta_{\tau\wedge(t+1)}1_{\tau\le t} \mid \mathcal{F}_t] + \mathbb{E}[\Delta_{\tau\wedge(t+1)}1_{\tau>t} \mid \mathcal{F}_t] + (2T - (t+1))\left(\frac{\phi_0\eta^2}{2} G_2 + \eta\mu\frac{\epsilon}{120}\right). \tag{19}$$

Beginning by bounding the first term on the right-hand side, we have

$$\mathbb{E}[\Delta_{\tau\wedge(t+1)}1_{\tau\le t} \mid \mathcal{F}_t] = \mathbb{E}[\Delta_{\tau\wedge t}1_{\tau\le t} \mid \mathcal{F}_t] = \Delta_{\tau\wedge t}1_{\tau\le t}. \tag{20a}$$

As for the second term, we have

$$\begin{aligned}\mathbb{E}[\Delta_{\tau\wedge(t+1)}1_{\tau>t} \mid \mathcal{F}_t] &= \mathbb{E}[\Delta_{t+1}1_{\tau>t} \mid \mathcal{F}_t] \\ &= \mathbb{E}[\Delta_{t+1} \mid \mathcal{F}_t]1_{\tau>t} \\ &\overset{(iii)}{\le} \Big(1 - \frac{\eta\mu}{4}\Big)\Delta_t 1_{\tau>t} + \left(\frac{\phi_0\eta^2}{2} G_2 + \eta\mu\frac{\epsilon}{120}\right)1_{\tau>t} \\ &\le \Big(1 - \frac{\eta\mu}{4}\Big)\Delta_{\tau\wedge t}1_{\tau>t} + \frac{\phi_0\eta^2}{2} G_2 + \eta\mu\frac{\epsilon}{120},\end{aligned} \tag{20b}$$

where step (iii) follows from using Inequality (18).

Substituting the bounds (20a) and (20b) into our original inequality (19), we find that

$$\begin{aligned}
\mathbb{E}[Y_{t+1} \mid \mathcal{F}_t] &= \mathbb{E}[\Delta_{\tau\wedge(t+1)} 1_{\tau\leq t} \mid \mathcal{F}_t] + \mathbb{E}[\Delta_{\tau\wedge(t+1)} 1_{\tau>t} \mid \mathcal{F}_t] + (2T-(t+1))\left(\frac{\phi_0\eta^2}{2}G_2 + \eta\mu\frac{\epsilon}{120}\right) \\
&\leq \Delta_{\tau\wedge t}1_{\tau\leq t} + (1-\eta\mu/4)\Delta_{\tau\wedge t}1_{\tau>t} + \left(\frac{\phi_0\eta^2}{2}G_2 + \eta\mu\frac{\epsilon}{120}\right) + (2T-(t+1))\left(\frac{\phi_0\eta^2}{2}G_2 + \eta\mu\frac{\epsilon}{120}\right) \\
&\overset{(iv)}{\leq} \Delta_{\tau\wedge t} + (2T-t)\left(\frac{\phi_0\eta^2}{2}G_2 + \eta\mu\frac{\epsilon}{120}\right) \\
&= Y_t,
\end{aligned}$$

where step (iv) follows from the inequality $\eta\mu\Delta_{\tau\wedge t} \geq 0$. We have thus verified the super-martingale property.

Finally, applying Doob's maximal inequality for super-martingales (see, e.g., Durrett [15]) yields

$$\begin{aligned}
\Pr\{\max_{t\in[2T]} Y_t \geq \nu\} &\leq \frac{\mathbb{E}[Y_0]}{\nu} \\
&= \frac{1}{\nu}\left(\Delta_0 + 2T\left\{\frac{\phi_0\eta^2}{2}G_2 + \eta\mu\frac{\epsilon}{120}\right\}\right) \\
&\overset{(v)}{\leq} \frac{1}{\nu}\left(\Delta_0 + \frac{\epsilon}{5}\log(120\Delta_0/\epsilon)\right),
\end{aligned}$$

where step (v) follows from the substitutions $T = \frac{2}{\eta\mu}\log(120\Delta_0/\epsilon)$, and $\eta \leq \frac{\epsilon\mu}{120 D\phi_0\lambda_0^2}$. As long as ϵ is sufficiently small so as to ensure that $\epsilon\log(120\Delta_0/\epsilon) < 5\Delta_0$, setting $\nu = 10\Delta_0$ completes the proof.

A.1 Stochastic zeroth-order rate for general non-convex functions:

It is worth pointing out the proof of Theorem 1 is only uses the local smoothness properties and the fact that the function is globally PL. Consequently, the same proof provides us an analogous result for Theorem 1 for any non-convex function f satisfying Assumption A, stated at the beginning of the proof of Theorem 1.

A.2 Bounds on G_2 and G_∞ for LQR:

In this section we provide bounds in equation (13). In particular, let us establish bounds on these quantities for general optimization of a function with a two-point gradient estimate. The following computations closely follow those of Shamir [40]. While proving upper bounds for G_2 and G_∞, we use $u \in \mathrm{Unif}(\mathbb{S}^{D-1})$, where D is the dimension of the matrix K.

Second moment control: Using the law of iterated expectations, we have

$$\mathbb{E}\left[\left\|D\frac{\mathcal{C}(K+ru;s_0)-\mathcal{C}(K-ru,s_0)}{2r}u\right\|_2^2\right] = \mathbb{E}\left[\mathbb{E}\left[\left\|D\frac{\mathcal{C}(K+ru;s_0)-\mathcal{C}(K-ru;s_0)}{2r}u\right\|_2^2\middle| s_0\right]\right].$$

Define the placeholder variable q and now evaluate:

$$\begin{aligned}
\mathbb{E}\left[\left\|D\frac{\mathcal{C}(K+ru;s_0)-\mathcal{C}(K-ru,s_0)}{2r}u\right\|_2^2\middle| s_0\right] &= \frac{D^2}{4r^2}\mathbb{E}\left[(\mathcal{C}(K+ru;s_0)-\mathcal{C}(K-ru;s_0))^2\|u\|_2^2\middle| s_0\right]. \\
&\overset{(i)}{=} \frac{D^2}{4r^2}\mathbb{E}\left[(\mathcal{C}(K+ru;s_0)-\mathcal{C}(K-ru;s_0))^2\middle| s_0\right] \\
&= \frac{D^2}{4r^2}\mathbb{E}\left[(\mathcal{C}(K+ru;s_0)-q-\mathcal{C}(K-ru;s_0)+q)^2\middle| s_0\right] \\
&\overset{(ii)}{\leq} \frac{D^2}{2r^2}\mathbb{E}\left[(\mathcal{C}(K+ru;s_0)-q)^2 + (\mathcal{C}(x-ru;s_0)-q)^2\middle| s_0\right],
\end{aligned}$$

where equality (i) follows from the fact that u is a unit vector and inequality (ii) follows from the inequality $(a-b)^2 \leq 2(a^2+b^2)$. We further simplify this to obtain:

$$\begin{aligned}
\mathbb{E}\left[\left\|D\frac{\mathcal{C}(K+ru;s_0)-\mathcal{C}(K-ru,s_0)}{2r}u\right\|_2^2 \middle| s_0\right] &\overset{(i)}{\leq} \frac{D^2}{r^2}\mathbb{E}\left[(\mathcal{C}(K+ru;s_0)-q)^2 \middle| s_0\right] \\
&\overset{(ii)}{\leq} \frac{D^2}{r^2}\sqrt{\mathbb{E}\left[(\mathcal{C}(K+ru;s_0)-q)^4 \middle| s_0\right]},
\end{aligned}$$

where inequality (i) follows from the symmetry of the uniform distribution on the sphere, and inequality (ii) follows from Jensen's Inequality. For a fixed s_0, we now define $q = \mathbb{E}[\mathcal{C}(K+ru;s_0)|s_0]$. Substituting this expression yields

$$\begin{aligned}
\mathbb{E}\left[\left\|D\frac{\mathcal{C}(K+ru;s_0)-\mathcal{C}(K-ru,s_0)}{2r}u\right\|_2^2 \middle| s_0\right] &\leq \frac{D^2}{r^2}\sqrt{\mathbb{E}\left[(\mathcal{C}(K+ru;s_0)-\mathbb{E}[\mathcal{C}(K+ru;s_0)|s_0])^4 \middle| s_0\right]} \\
&\overset{(i)}{\leq} \frac{D^2}{r^2}\frac{(\lambda_0 r)^2}{D} \\
&= D\lambda_0^2,
\end{aligned}$$

where inequality (i) follows directly from Lemma 9 in Shamir [40]. The lemma can be applied since we are conditioning on s_0, and all the randomness lies in the selection of u.

Gradient estimates are bounded: Note that smoothing radius r satisfies $r \leq \rho_0$, where ρ_0 is the radius within which the function is Lipschitz. Consequently, the local Lipschitz property of the function $\mathcal{C}(\cdot)$ implies that

$$\begin{aligned}
\|\mathrm{g}_t\|_2 &:= \left\|D\frac{\mathcal{C}(K_t+ru_t,s_0)-\mathcal{C}(K_t-ru_t,s_0)}{2r}u_t\right\| \\
&\leq \left\|D\frac{\mathcal{C}(K_t+ru_t;s_0)-\mathcal{C}(K_t;s_0)}{2r}u_t\right\| + \left\|D\frac{\mathcal{C}(K_t;s_0)-\mathcal{C}(K_t-ru_t;s_0)}{2r}u_t\right\| \\
&\leq D\lambda_0\frac{2\|ru_t\|_2}{2r} \leq D\lambda_0.
\end{aligned}$$

B Auxiliary results for Theorem 1

In order to emphasize the generality of Theorem 1, we prove the auxiliary results used in the proof of Theorem 1 for general non-convex function.

In what follows, we use f to denote a general non-convex function, and use F to denote the noisy version of the function f. We assume that the function f that satisfies the following properties:

(a) The function f has (ϕ_0,ρ_0)-locally Lipschitz gradients,

(b) The function $F(\cdot,\xi)$ is (λ_0,ρ_0)-locally Lipschitz for all ξ.

(c) The function f is globally μ-PL.

We use the step-size η, smoothing radius r, and iteration complexity T posited by Theorem 1. In particular, we assume

$$\eta \leq \min\left\{\frac{\epsilon\mu}{240\phi_0 G_2},\ \frac{1}{2\phi_0},\ \frac{\rho_0}{G_\infty}\right\}, \qquad \text{and} \tag{21a}$$

$$r \leq \min\left\{\frac{\theta_0\mu}{8\phi_0}\sqrt{\frac{\epsilon}{15}},\ \frac{1}{2\phi_0}\sqrt{\frac{\epsilon\mu}{30}},\ \rho_0\right\}. \tag{21b}$$

For each $t = 0,1,2,\ldots$, we define the cost error $\Delta_t = f(x_t) - f(x^*)$, as well as the stopping time

$$\tau := \min\Big\{t \mid \Delta_t > 10\Delta_0\Big\}. \tag{22}$$

The quantities $\mathcal{G}^0$ and Δ_0 are defined in equation 12 with the function C replaced by f. The gradient estimate g at the point $x \in \mathcal{G}^0$ is assumed to satisfy the bounds

$$\mathsf{var}(\mathrm{g}(x)) \leq G_2 \quad \text{ and } \quad \|\mathrm{g}(x)\|_2 \leq G_\infty \text{ almost surely.}$$

With this set up in place, we are now ready to prove the auxiliary results used in the proof of Theorem 1.

B.1 Proof of Lemma 4

For a scalar $r > 0$, the smoothed version $f_r(x)$ is given by $f_r(x) := \mathbb{E}\left[f(x + rv)\right]$, where the expectation above is taken with respect to the randomness in v, and v has uniform distribution on a d-dimensional ball $\mathbb{B}^d$ of unit radius. The estimate g of the gradient ∇f_r at x is given by

$$\mathrm{g}(x) = \left[F(x + ru, \xi) - F(x - ru, \xi)\right] \frac{d}{2r} u$$

where u has a uniform distribution on the shell of the sphere $\mathbb{S}^{d-1}$ of unit radius, and ξ is sampled at random from $\mathcal{D}$. The following result summarizes some useful properties of the smoothed version of f, and relates it to the gradient estimate g.

Lemma 5. *The smoothed version f_r of f with smoothing radius r has the following properties:*

(a) $\nabla f_r(x) = \mathbb{E}\left[\mathrm{g}(x)\right]$.

(b) $\|\nabla f_r(x) - \nabla f(x)\|_2 \leq \phi_0 r$.

Versions of these properties have appeared in past work [19, 4, 40], but we provide proofs in Appendix B.2 for completeness.

Taking Lemma 5 as given, we now prove Lemma 4. Let $\mathcal{F}_t$ denote the sigma field generated by the randomness up to iteration t, and $\mathbb{E}$ denote the total expectation operator. We define $\mathbb{E}^t := \mathbb{E}\left[\cdot \mid \mathcal{F}_t\right]$ as the expectation operator conditioned on the sigma field $\mathcal{F}_t$. Recall that the function f is smooth with smoothness parameter ϕ_0, and we have

$$\begin{aligned}
\mathbb{E}^t\left[f(x_{t+1}) - f(x_t)\right] &\leq \mathbb{E}^t\left[\langle \nabla f(x_t),\, x_{t+1} - x_t\rangle + \frac{\phi_0}{2}\|x_{t+1} - x_t\|_2^2\right] \\
&\overset{(i)}{=} -\langle \eta \nabla f(x_t),\, \nabla f_r(x_t)\rangle + \frac{\phi_0 \eta^2}{2}\mathbb{E}^t\left[\|\mathrm{g}(x_t)\|_2^2\right] \\
&\overset{(ii)}{=} -\eta\|\nabla f(x_t)\|_2^2 + \eta\phi_0 r\|\nabla f(x_t)\|_2 + \frac{\phi_0\eta^2}{2}\mathbb{E}^t\left[\|\mathrm{g}(x_t)\|_2^2\right].
\end{aligned}$$

Steps (i) and (ii) above follow from parts (a) and (b), respectively, of Lemma 5. Now make the observation that

$$\begin{aligned}
\mathbb{E}^t\left[\|\mathrm{g}(x_t)\|_2^2\right] &= \mathsf{var}(\mathrm{g}(x_t)) + \|\nabla f_r(x_t)\|_2^2 \\
&\leq \mathsf{var}(\mathrm{g}(x_t)) + 2\|\nabla f(x_t)\|_2^2 + 2\|\nabla f_r(x_t) - \nabla f(x_t)\|_2^2 \\
&\leq G_2 + 2\|\nabla f(x_t)\|_2^2 + 2(\phi_0 r)^2.
\end{aligned}$$

In addition, since the function is locally smooth at the point x_t, we have

$$\begin{aligned}
(\theta - \theta^2\phi_0/2)\|\nabla f(x_t)\|_2^2 &\leq f(x_t) - f(x_t - \theta\nabla f(x_t)) \\
&\leq f(x_t) - f(x^*),
\end{aligned}$$

for some parameter θ chosen small enough such that the relation $\theta\|\nabla f(x_t)\|_2 \leq \rho_0$ holds. We may thus set $\theta = \theta_0 = \min\left\{\frac{1}{2\phi_0}, \frac{\rho_0}{\lambda_0}\right\}$ and recall the notation $\Delta_t = f(x_t) - f(x^*)$ to obtain

$$\begin{aligned}
\mathbb{E}^t\left[\Delta_{t+1}\right] &\leq -\eta\|\nabla f(x_t)\|_2^2 + \eta\phi_0 r\frac{2}{\theta_0}\Delta_t^{1/2} + \frac{\phi_0\eta^2}{2}G_2 + \phi_0\eta^2\left(\|\nabla f(x_t)\|_2^2 + (\phi_0 r)^2\right) \\
&\overset{(iii)}{\leq} -\frac{\eta\mu}{2}\Delta_t + 2\frac{\eta\phi_0 r}{\theta_0}\Delta_t^{1/2} + \frac{\phi_0\eta^2}{2}G_2 + \phi_0\eta^2(\phi_0 r)^2, \\
&\overset{(iv)}{\leq} -\frac{\eta\mu}{2}\Delta_t + \frac{\eta\mu}{4}\Delta_t + 4\frac{\eta(\phi_0 r)^2}{\mu\theta_0^2} + \frac{\phi_0\eta^2}{2}G_2 + \phi_0\eta^2(\phi_0 r)^2,
\end{aligned}$$

where step (iii) follows from applying the PL inequality and using the fact that $\eta \leq \frac{1}{2\phi_0}$, and step (iv) from the inequality $2ab \leq a^2 + b^2$ which holds for any pair of scalars (a, b).

Recall the assumed bounds on our parameters, namely

$$\eta \leq \min\left\{\frac{\epsilon\mu}{240\phi_0}, \frac{1}{2\phi_0}\right\}, \quad \text{and} \quad r \leq \frac{1}{2\phi_0}\min\left\{\theta_0\mu\sqrt{\frac{\epsilon}{240}}, \frac{1}{\phi_0}\sqrt{\frac{\epsilon\mu}{30}}\right\}.$$

Using these bounds, we have

$$\mathbb{E}^t\left[\Delta_{t+1}\right] \leq -\frac{\eta\mu}{4}\Delta_t + \frac{\phi_0\eta^2}{2}G_2 + \eta\mu\frac{\epsilon}{120}.$$

Finally, rearranging yields

$$\mathbb{E}^t\left[\Delta_{t+1}\right] \leq \left(1 - \frac{\eta\mu}{4}\right)\Delta_t + \frac{\phi_0\eta^2}{2}G_2 + \eta\mu\frac{\epsilon}{120}, \tag{23}$$

which completes the proof of Lemma 4.

B.2 Proof of Lemma 5

We now provide the proof of Lemma 5, splitting our analysis into the two separate claims.

Proof of part (a): Unwrapping the definition of $\nabla f_r(x)$ yields

$$\begin{aligned}
\nabla f_r(x) &\overset{(i)}{=} \frac{d}{r}\mathbb{E}[f(x+ru)u] \\
&= \frac{d}{2r}(\mathbb{E}[f(x+ru)u] + \mathbb{E}[f(x+ru)u]) \\
&\overset{(ii)}{=} \frac{d}{2r}(\mathbb{E}[f(x+ru)u] - \mathbb{E}[f(x-ru)u]) \\
&= \frac{d}{2r}\mathbb{E}[f(x+ru)u - f(x-ru)u],
\end{aligned}$$

where equality (i) follows from Lemma 1 in Flaxman et al. [19], and equality (ii) follows from the symmetry of the uniform distribution on the shell $\mathbb{S}^{d-1}$. Now observe that

$$\begin{aligned}
\mathbb{E}[F(x+ru,\xi)u - F(x-ru,\xi)u] &= \mathbb{E}\Big[\mathbb{E}[F(x+ru,\xi) - F(x-ru,\xi)u|u]\Big] \\
&\overset{(i)}{=} \mathbb{E}\Big[f(x+ru)u - f(x-ru)u\Big],
\end{aligned}$$

where equality (i) follows from the assumption that $f(x) = \mathbb{E}_{\xi\sim\mathcal{D}}[F(x,\xi)]$. Putting the equations together establishes the claim in part (a). □

Proof of part (b): Observe that

$$\begin{aligned}
\|\nabla f_r(x) - \nabla f(x)\|_2 &= \|\nabla\mathbb{E}[f(x+rv)] - \nabla f(x)\|_2 \\
&= \|\mathbb{E}[\nabla[f(x+rv) - \nabla f(x)]\|_2 \\
&\overset{(i)}{\leq} \mathbb{E}[\|\nabla[f(x+rv) - \nabla f(x)\|_2] \\
&\overset{(ii)}{\leq} \phi_0 r,
\end{aligned}$$

where inequality (i) above follows from Jensen's inequality, whereas step (ii) follows since $r \leq \rho$ and ∇f is locally Lipschitz continuous with parameter ϕ_0. □

C LQR

In this section, we establish some fundamental properties of the cost function $\mathcal{C}$, and provide proofs of Lemmas 1 and 2. As part of these proofs, we provide explicit bounds for the local curvature parameters $(\lambda_0, \rho_0, \phi_0)$. We make frequent use of results established by Fazel et al. [17], and as mentioned before, Lemmas 1 and 2 are refinements of their results.

Notation: In this section, we introduce some shorthand to reduce notational overhead. Much (but not all) of the notation we use overlaps with the notation used in Fazel et al. [17].

We define the matrix P_K as the solution to the following fixed point equation:

$$P_K = Q + K^\top RK + (A - BK)^\top P_K(A - BK),$$

and we define the state correlation matrix Σ_K as:

$$\Sigma_K = \mathbb{E}\left[\sum_{t=0}^{\infty} s_t s_t^\top\right] \quad \text{such that} \quad s_t = (A - BK)s_{t-1}. \tag{24}$$

It is straightforward to see that we have

$$\mathcal{C}(K) = \mathbb{E}[s_0^\top P_K s_0], \tag{25}$$

and we make frequent use of this representation in the sequel.

Recall that we have $\mathbb{E}[s_0 s_0^\top] = I$, so that

$$\mathcal{C}(K) = \operatorname{tr}(P_K). \tag{26}$$

Moreover, under this assumption, the cost function $\mathcal{C}$ satisfies the PL Inequality with PL constant $\frac{\|\Sigma_{K^*}\|_2}{\sigma_{min}(R)}$ (see Lemma 3 in the paper [17]).

Also define the *natural gradient* of the cost function as

$$E_K := 2(R + B^\top P_K B)K - B^\top P_K A,$$

so that we have $\nabla\mathcal{C}(K) = E_K\Sigma_K$. For any symmetric matrix X, the perturbation operators $\mathcal{T}_K(\cdot)$ and $\mathcal{F}_K(\cdot)$ are defined as

$$\mathcal{T}_K(X) = \sum_{t=0}^{\infty}(A - BK)^t X[(A - BK)^\top]^t, \quad \text{and} \quad \mathcal{F}_K(X) = (A - BK)X(A - BK)^\top.$$

Finally, the operator norms of the operators $\mathcal{T}_K(\cdot)$ and $\mathcal{F}_K(\cdot)$ are defined as

$$\|\mathcal{T}_K\|_2 = \sup_X \frac{\|\mathcal{T}_K(X)\|_2}{\|X\|_2} \text{ and}$$
$$\|\mathcal{F}_K\|_2 = \sup_X \frac{\|\mathcal{F}_K(X)\|_2}{\|X\|_2}.$$

Useful constants:

We now define several polynomials of $\mathcal{C}(K)$, which are useful in various proofs in this section.

- $c_{K_1} = \frac{\mathcal{C}(K)}{\sigma_{min}(Q)}\sqrt{(\|R\|_2 + \|B\|_2^2\mathcal{C}(K))(\mathcal{C}(K) - \mathcal{C}(K^*))}$
- $c_{K_2} = 4\left(\frac{\mathcal{C}(K)}{\sigma_{min}(Q)}\right)^2\|Q\|_2\|B\|_2(\|A\|_2 + \|B\|_2 c_{K_1} + 1)$
- $c_{K_3} = 8\left(\frac{\mathcal{C}(K)}{\sigma_{min}(Q)}\right)^2(c_{K_1})^2\|R\|_2\|B\|_2(\|A\|_2 + \|B\|_2 c_{K_1} + 1)$

- $c_{K_4} = 2\left(\frac{\mathcal{C}(K)}{\sigma_{min}(Q)}\right)^2 (c_{K_1}+1)\|R\|_2$
- $c_{K_5} = \sqrt{(\|R\|_2 + \|B\|_2^2 \mathcal{C}(K))(\mathcal{C}(K) - \mathcal{C}(K^*))}$
- $c_{K_6} = \|R\|_{\mathrm{F}} + \|B\|_{\mathrm{F}}^2 (c_{K_1}+1)(c_{K_2}+c_{K_3}+c_{K_4}) + \|B\|_{\mathrm{F}}^2 \mathcal{C}(K) + \|B\|_{\mathrm{F}}\|A\|_2(c_{K_2}+c_{K_3}+c_{K_4})$
- $c_{K_7} = 5c_{K_6}\frac{\mathcal{C}(K)}{\sigma_{min}(Q)} + 4c_{K_5}\left(\frac{\mathcal{C}(K)}{\sigma_{min}(Q)}\right)^2 \|B\|_2(\|A\|_2 + \|B\|_2 c_{K_1})$.
- $c_{K_8} = C_m(c_{K_2}+c_{K_3}+c_{K_4})$.
- $c_{K_9} = \min\left\{\frac{\sigma_{min}(Q)}{4\mathcal{C}(K)\|B\|_2(\|A\|_2+\|B\|_2 c_{K_1}+1)}, 1\right\}$.

With these definitions at hand, we are now in a position to establish Lemmas 1 and 2.

C.1 Proof of Lemma 1

Let us restate a precise version of the lemma for convenience.

Lemma 6. *For any pair (K', K) such that $\|K'-K\|_F \le c_{K_9}$, we have*

$$|\mathcal{C}(K', s_0) - \mathcal{C}(K, s_0)| \le c_{K_8}\|K'-K\|_F.$$

Comparing Lemma 6 with the statement of Lemma 1, we have therefore established the relations

$$\zeta_K \ge c_{K_9} \quad \text{and} \quad \lambda_K \le c_{K_8}.$$

Proof. The sample cost satisfies the relation

$$\begin{aligned}
|\mathcal{C}(K', s_0) - \mathcal{C}(K, s_0)| &= |s_0^\top P_{K'} s_0 - s_0^\top P_K s_0| \\
&= |\operatorname{tr}(s_0^\top (P_{K'} - P_K) s_0)| \\
&\le \|P_{K'} - P_K\|_2 \, \|s_0\|_2^2 \\
&\le \|P_{K'} - P_K\|_2 C_m.
\end{aligned} \tag{27}$$

Hence, it remains to bound $\|P_{K'} - P_K\|_2$. To this end, substituting the definition of the linear operator $\mathcal{T}_K$, we have

$$\begin{aligned}
\|P_{K'} - P_K\|_2 &= \|\mathcal{T}_{K'}(Q + (K')^\top R K') - \mathcal{T}_K(Q + K^\top R K)\|_2 \\
&= \|(\mathcal{T}_{K'} - \mathcal{T}_K)(Q + (K')^\top R K') - \mathcal{T}_K(K^\top R K - (K')^\top R K')\|_2 \\
&\le \|(\mathcal{T}_{K'} - \mathcal{T}_K)Q\|_2 + \|(\mathcal{T}_{K'} - \mathcal{T}_K)(K')^\top R K')\|_2 \\
&\qquad + \|\mathcal{T}_K\|_2 \|K^\top R K - (K')^\top R K'\|_2.
\end{aligned} \tag{28}$$

We provide upper bounds for the three terms above as follows:

$$\|(\mathcal{T}_{K'} - \mathcal{T}_K)(K')^\top R K')\|_2 \le c_{K_3}\|K - K'\|_2 \tag{29a}$$

$$\|(\mathcal{T}_{K'} - \mathcal{T}_K)Q\|_2 \le c_{K_2}\|K - K'\|_2 \tag{29b}$$

$$\|\mathcal{T}_K\|_2 \|K^\top R K - (K')^\top R K'\|_2 \le c_{K_4}\|K - K'\|_2. \tag{29c}$$

Taking the above bounds as given at the moment, we have from Equation (28) that

$$\|P_{K'} - P_K\|_2 \le (c_{K_2} + c_{K_3} + c_{K_4})\|K' - K\|_2, \tag{30}$$

Putting together the pieces completes the proof of Lemma 1. □

It remains to prove the upper bounds (29a)- (29c).

Auxiliary bounds: Proofs of the bounds (29a) through (29c) are based on the following intermediate bounds:

$$\|(K')^\top RK' - K^\top RK\|_2 \le (c_{K_1}+1)\|R\|_2\|K'-K\|_2. \tag{31a}$$

$$\|\mathcal{F}_{K'} - \mathcal{F}_K\|_2 \le 2\|B\|_2(\|A\|_2 + \|B\|_2 c_{K_1} + 1)\|K'-K\|_2 \tag{31b}$$

$$\|\mathcal{T}_K\|_2 \le \frac{\mathcal{C}(K)}{\sigma_{min}(Q)} \tag{31c}$$

$$\|K^\top RK\|_2 \le c_{K_1}^2\|R\|_2 \tag{31d}$$

We prove these bounds at the end, but let us complete the rest of the proofs assuming these auxiliary bounds.

Proof of the bound (29a)**:** The proof of this upper bound is based on Lemma 20 from the paper [17]. Accordingly, we start by verifying the following condition for Lemma 20:

$$\|\mathcal{F}_K - \mathcal{F}_{K'}\|_2\|(K')^\top RK'\|_2 \le \frac{1}{2}. \tag{32}$$

Observe that our assumption $\|K'-K\|_{\mathrm{F}} \le c_{K_9}$, satisfies the assumption of Lemma 10 in the paper [17], whence we have

$$\begin{aligned}
\|B\|_2\|K'-K\|_2 &\overset{(i)}{\le} \|B\|_2\frac{\sigma_{min}(Q)}{4\mathcal{C}(K)\|B\|_2(\|A\|_2+\|B\|_2c_{K_1}+1)} \\
&\overset{(ii)}{\le} \frac{\sigma_{min}(Q)}{4\mathcal{C}(K)(\|A-BK\|_2+1)} \\
&\overset{(iii)}{\le} \frac{1}{4}.
\end{aligned} \tag{33}$$

where step (i) follows by substituting the value of c_{K_9}, and step (ii) follows since $\|A-BK\|_2 \le \|A\|_2+\|B\|_2c_{K_1}+1$. Step (iii) above follows since $\mathcal{C}(K) \ge \sigma_{min}(Q)$. Combining the last inequality with Lemma 16 in the paper [17] yields

$$\begin{aligned}
\|\mathcal{F}_{K'} - \mathcal{F}_K\|_2 &\le 2\|A-BK\|_2\|B\|_2\|K'-K\|_2 + \|B\|_2^2\|K'-K\|_2^2 \\
&\le 2\|B\|_2(\|A-BK\|_2+1)\|K'-K\|_2
\end{aligned}$$

Finally, invoking Lemma 14 from the paper [17] guarantees that $\|\mathcal{T}_K\|_2 \le \frac{\mathcal{C}(K)}{\sigma_{min}(Q)}$, and we deduce that

$$\begin{aligned}
\|\mathcal{T}_K\|_2\|\mathcal{F}_{K'} - \mathcal{F}_K\|_2 &\le \frac{\mathcal{C}(K)}{\sigma_{min}(Q)}2\|B\|_2(\|A-BK\|_2+1)\|K'-K\|_2 \\
&\le \frac{1}{2},
\end{aligned}$$

where the last inequality follows from the assumption $\|K'-K\|_{\mathrm{F}} \le c_{K_9}$.

Now that we have verified that condition 32, invoking Lemma 20 in the paper [17] yields

$$\begin{aligned}
\|(\mathcal{T}_{K'} - \mathcal{T}_K)(K')^\top RK'\|_2 &\le 2\|\mathcal{T}_K\|_2^2\|\mathcal{F}_K - \mathcal{F}_{K'}\|_2\|(K')^\top RK'\|_2 \\
&\le 2\|\mathcal{T}_K\|_2^2\|\mathcal{F}_K - \mathcal{F}_{K'}\|_2\|K^\top RK\|_2 \\
&\qquad + 2\|\mathcal{T}_K\|_2^2\|\mathcal{F}_K - \mathcal{F}_{K'}\|_2\|(K')^\top RK' - K^\top RK\|_2 \\
&\le c_{K_3}\|K-K'\|_2,
\end{aligned}$$

where the last step above follows by substituting the bounds (31a)- (31d).

Proof of the bounds (29b) **and** (29c)**:** The proof of the bound (29b) is similar to the part (29a) and is based on Lemma 20 from the paper [17]. More concretely, we have

$$\|(\mathcal{T}_{K'} - \mathcal{T}_K)Q\|_2 \le 2\|\mathcal{T}_K\|_2^2\|\mathcal{F}_K - \mathcal{F}_{K'}\|_2\|Q\|_2 \le c_{K_2}\|K-K'\|_2$$

where the last step above follows from the bounds (31b) and (31c). The proof of the bound (29c) is a direct consequence of the bounds (31a) and (31c).

C.1.1 Proofs of the auxiliary bounds

In this section we prove the auxiliary bounds (31a) through to (31d)

Bound (31a)*:* Observe that

$$\begin{aligned}
\|K^\top RK - (K')^\top RK'\|_2 &= \|(K'-K)^\top R(K'-K) + (K')^\top RK + K^\top R(K') - 2K^\top RK\|_2 \\
&\leq (2\|R\|_2\|K\|_2\|K'-K\|_2 + \|R\|_2\|K'-K\|_2^2) \\
&\overset{(i)}{\leq} (2\|K\|_2 + 1)\|R\|_2\|K'-K\|_2 \\
&\overset{(ii)}{\leq} (2c_{K_1} + 1)\|R\|_2\|K'-K\|_2.
\end{aligned}$$

where step (i) follows since $\|K - K'\|_2 \leq 1$ by assumption, and step (ii) follows since $\|K\|_2 \leq c_{K_1}$ (see Lemma 22 in the paper [17]). This completes the proof of bound (31a).

Bound (31b)*:* In order to prove bound (31b), we invoke Lemma 19 in the paper [17] to obtain

$$\begin{aligned}
\|\mathcal{F}_{K'} - \mathcal{F}_K\|_2 &\leq 2\|A - BK\|_2\|B\|_2\|K'-K\|_2 + \|B\|_2^2\|K'-K\|_2^2 \\
&\overset{(iii)}{\leq} 2\|A - BK\|_2\|B\|_2\|K'-K\|_2 + \frac{1}{4}\|B\|_2\|K'-K\|_2 \\
&\leq 2\|B\|_2(\|A\|_2 + \|B\|_2 c_{K_1} + 1)\|K'-K\|_2
\end{aligned}$$

where step (iii) above follows from the upper bound (33). This completes the proof of the bound (31b).

Bound (31c) *and* (31d)*:* The bound (31c) above follows from Lemma 17 in the paper [17], whereas the bound (31c) follows from the fact that $\|K\|_2 \leq c_{K_1}$ (see Lemma 22 in the paper [17]).

Having established all of our auxiliary bounds, let us now proceed to a proof of Lemma 2.

C.2 Proof of Lemma 2

Lemma 2 is a consequence of the following result.

Lemma 7. *If* $\|K' - K\|_F \leq c_{K_9}$*, then*

$$\|\nabla\mathcal{C}(K') - \mathcal{C}(K)\|_F \leq c_{K_7}\|K' - K\|_F.$$

Indeed, comparing Lemmas 7 and 2, we have the bounds

$$\beta_K \geq c_{K_9} \quad \text{and} \quad \phi_K \leq c_{K_7}.$$

Let us now prove Lemma 7.

Proof. We start by noting that from Lemma 1 we have that the cost function $\mathcal{C}(K)$ is locally Lipschitz in a ball of ζ_K around the point K. Before moving into the main argument, we mention a few auxiliary results that are helpful in the sequel. We start by invoking Lemma 13 from the paper [17], whence we have

$$\|P_K\|_2 \leq \mathcal{C}(K) \quad \text{and} \quad \|\Sigma_K\|_2 \leq \frac{\mathcal{C}(K)}{\sigma_{min}(Q)}.$$

We also have

$$\|A - BK\|_2 \leq \|A\|_2 + \|B\|_2\|K\|_2 \overset{(i)}{\leq} \|A\|_2 + \|B\|_2 c_{K_1} \quad \text{and} \tag{34a}$$

$$\|\Sigma_{K'}\|_2 \leq \|\Sigma_K\|_2 + \|\Sigma_{K'} - \Sigma_K\|_2 \overset{(ii)}{\leq} 5\frac{\mathcal{C}(K)}{\sigma_{min}(Q)}. \tag{34b}$$

Step (i) above follows since $\|K\|_2 \leq c_{K_1}$ (see Lemma 22 in the paper [17]), whereas step (ii) follows since $\|\Sigma_{K'} - \Sigma_K\|_2 \leq 4\frac{\mathcal{C}(K)}{\sigma_{min}(Q)}$ (see Lemma 16 in the paper [17]).

Recalling the gradient expression $\nabla\mathcal{C}(K) = E_K\Sigma_K$. Let K' be a policy such that $\|K' - K\|_{\mathrm{F}} \leq c_{K_9}$. We have

$$\begin{aligned}
\|\nabla\mathcal{C}(K') - \nabla\mathcal{C}(K)\|_{\mathrm{F}} &= \|(E_{K'} - E_K)\Sigma_{K'} + E_K(\Sigma_{K'} - \Sigma_K)\|_{\mathrm{F}} \\
&\leq \|(E_{K'} - E_K)\|_{\mathrm{F}}\|\Sigma_{K'}\|_2 + \|E_K\|_{\mathrm{F}}\|(\Sigma_{K'} - \Sigma_K)\|_2 \\
&\overset{(iii)}{\leq} 5c_{K_6}\frac{\mathcal{C}(K)}{\sigma_{min}(Q)}\|K' - K\|_{\mathrm{F}} \\
&\quad + 4c_{K_5}\left(\frac{\mathcal{C}(K)}{\sigma_{min}(Q)}\right)^2 \frac{\|B\|_2(\|A\|_2 + \|B\|_2 c_{K_1})}{\sigma_{min}(\Sigma_0)}\|K' - K\|_{\mathrm{F}}.
\end{aligned}$$

The upper bound in step (iii) on the term $\|(E_{K'} - E_K)\|_{\mathrm{F}}\|\Sigma_{K'}\|_2$ follows from Equation (34b) and from the following upper bound which we prove later:

$$\|E_{K'} - E_K\|_{\mathrm{F}} \leq c_{K_6}\|K' - K\|_{\mathrm{F}} \text{ provided } \|K' - K\|_{\mathrm{F}} \leq c_{K_9}. \tag{35}$$

The upper bound on the term $\|E_K\|_{\mathrm{F}}\|(\Sigma_{K'} - \Sigma_K)\|_2$ in step (iii) follows from the fact that $\|E_K\|_{\mathrm{F}} \leq c_{K_5}$ (see Lemma 11 in the paper [17]) and from the fact that

$$\begin{aligned}
\|(\Sigma_{K'} - \Sigma_K)\|_2 &\overset{(iv)}{\leq} 4\left(\frac{\mathcal{C}(K)}{\sigma_{min}(Q)}\right)^2 \frac{\|B\|_2(\|A - BK\|_2 + 1)}{\sigma_{min}(\Sigma_0)}\|K' - K\|_{\mathrm{F}} \\
&\overset{(v)}{\leq} 4\left(\frac{\mathcal{C}(K)}{\sigma_{min}(Q)}\right)^2 \frac{\|B\|_2(\|A\|_2 + \|B\|_2 c_{K_1} + 1)}{\sigma_{min}(\Sigma_0)}\|K' - K\|_{\mathrm{F}},
\end{aligned}$$

where step (iv) follows from Lemma 16 in the paper [17], and step (v) follows from Inequality (34a).

Putting together the pieces, we conclude that the function $\nabla\mathcal{C}(K)$ is Lipschitz with constant ϕ_K, where ϕ_K is given by

$$\phi_K = 5c_{K_6}\frac{\mathcal{C}(K)}{\sigma_{min}(Q)} + 4c_{K_5}\left(\frac{\mathcal{C}(K)}{\sigma_{min}(Q)}\right)^2 \|B\|_2(\|A\|_2 + \|B\|_2 c_{K_1} + 1) = c_{K_7}.$$

□

It remains to prove Inequality (35).

Proof of Inequality (35)**:** From the definition of E_K, we have

$$\begin{aligned}
\|E_{K'} - E_K\|_{\mathrm{F}} &= 2\|(R + B^\top P_{K'}B)K' - B^\top P_{K'}A - (R + B^\top P_K B)K + B^\top P_K A\|_{\mathrm{F}} \\
&= 2\|R(K' - K) + B^\top(P_{K'} - P_K)BK' + B^\top P_K B(K' - K) - B^\top(P_{K'} - P_K)A\|_{\mathrm{F}} \\
&\leq 2\|R\|_{\mathrm{F}}\|K' - K\|_{\mathrm{F}} + 2\|B^\top(P_{K'} - P_K)BK'\|_{\mathrm{F}} \\
&\quad + 2\|B^\top P_K B(K' - K)\|_{\mathrm{F}} + 2\|B^\top(P_{K'} - P_K)A\|_{\mathrm{F}}
\end{aligned} \tag{36}$$

We provide upper bounds for the three terms above as follows. First, we have

$$\|B^\top(P_{K'} - P_K)BK'\|_{\mathrm{F}} \leq \|B\|_{\mathrm{F}}^2(c_{K_1} + 1)(c_{K_2} + c_{K_3} + c_{K_4})\|K' - K\|_{\mathrm{F}},$$

which follows from the bound (30), since $\|K' - K\|_{\mathrm{F}} \leq c_{K_9}$, and the relation $\|K'\|_2 \leq \|K\|_2 + \|K' - K\|_2 \leq c_{K_1} + 1$. The same reasoning also yields the bound

$$\|B^\top(P_{K'} - P_K)A\|_{\mathrm{F}} \leq \|B\|_{\mathrm{F}}\|A\|_2(c_{K_2} + c_{K_3} + c_{K_4})\|K' - K\|_{\mathrm{F}},$$

Finally, since $\|P_K\|_2 \leq \mathcal{C}(K)$, we have

$$\|B^\top P_K B(K' - K)\|_{\mathrm{F}} \leq \|B\|_{\mathrm{F}}^2\mathcal{C}(K)\|K' - K\|_{\mathrm{F}}.$$

Combining the above upper bounds with the upper bound (36) we conclude that

$$\|E_{K'} - E_K\|_{\mathrm{F}} \le c_{K_6}\|K' - K\|_{\mathrm{F}},$$

where c_{K_6} is given by

$$c_{K_6} = 2\left[\|R\|_{\mathrm{F}} + \|B\|_{\mathrm{F}}\|A\|_2(c_{K_2} + c_{K_3} + c_{K_4}) + \|B\|_{\mathrm{F}}^2\left((c_{K_1}+1)(c_{K_2} + c_{K_3} + c_{K_4}) + \mathcal{C}(K)\right)\right].$$

C.3 Explicit bounds on the parameters (ρ_0, λ_0, ϕ_0)

In order to ease notation, we define constants $\widetilde{c}_{K_7}$, $\widetilde{c}_{K_8}$ and $\widetilde{c}_{K_9}$ by replacing the scalar $\mathcal{C}(K)$ by $10\mathcal{C}(K_0) - 9\mathcal{C}(K^*)$ in the definitions of c_{K_7}, c_{K_8} and c_{K_9} respectively (see Section C).

Lemma 8. *The parameters ρ_0, λ_0, ϕ_0 satisfy the following bounds*

$$\rho_0 \ge \widetilde{c_{K_9}}, \quad \phi_0 \le \widetilde{c_{K_7}} \quad \textit{and} \quad \lambda_0 \le \widetilde{c_{K_8}}.$$

Proof. Observe that from the definition of the set $\mathcal{G}^0$ we have that for all $K \in \mathcal{G}^0$, the function value $\mathcal{C}(K)$ is upper bounded as $\mathcal{C}(K) \le 10\mathcal{C}(K_0) - 9\mathcal{C}(K^*)$. Consequently, for any $K \in \mathcal{G}^0$ and any K' such that $\|K' - K\|_{\mathrm{F}} \le \widetilde{c_{K_9}}$, we can use Lemma 2 and Lemma 1 respectively to show that the cost function $\mathcal{C}(K)$ has locally Lipschitz gradients with parameter $\widetilde{c_{K_8}}$ and the function $\mathcal{C}(K)$ has locally Lipschitz function values parameter $\widetilde{c_{K_7}}$. Combining the last observation with the definitions of ρ_0, λ_0 and ϕ_0 we have that $\rho_0 \ge \widetilde{c_{K_9}}$, $\phi_0 \le \widetilde{c_{K_7}}$ and $\lambda_0 \le \widetilde{c_{K_8}}$. This completes the proof. □

D Experimental Details & Additional Experiments

For all experiments in this paper, the initial K_0 was generated by randomly perturbing the entries of K^* by a Gaussian random matrix with independent entries. Since we operate in the setting where we get noisy evaluations of the true infinite horizon cost, the length of the rollout used was manually tuned until the truncated cost converged arbitrarily close to the true infinite horizon cost. The step size was also tuned manually, and the smoothing radius was always chosen to be the minimum of $\sqrt{\epsilon}$ and the largest value required to ensure stability.

We now present the LQR problem we used to generate the plots in Figure 1 and experimental results in Section 4:

$$A = \begin{bmatrix} 1 & 0 & -10 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -10 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}, \quad Q = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}, \quad R = \begin{bmatrix} 5 & -3 & 0 \\ -3 & 5 & -2 \\ 0 & -2 & 5 \end{bmatrix},$$

The initial state was sampled uniformly at random from the canonical basis vectors.

We also performed experiments on several additional LQR instances to test the robustness of the behavior observed in Figures 1(b) and 1(c). Note that for all figures shown in this section, each dotted line represents the line of best fit for its corresponding data points, as in Figure 1. Using the same LQR problem shown above, we tested the performance of our two-point algorithm with different values of ϵ and $\mathcal{C}(K_0)$.

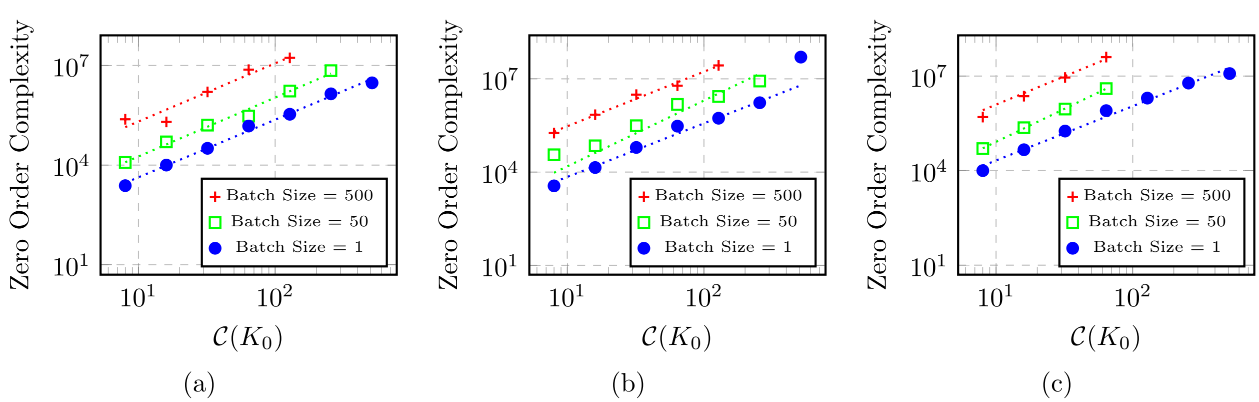

Figure 2. Scaling of complexity vs. $\mathcal{C}(K_0)$ while using mini-batches of size 1, 50 and 500, to achieve an error tolerance of (a) $\epsilon = 0.1$, (b) $\epsilon = 0.05$ and (c) $\epsilon = 0.01$. Due to the prohibitive complexity when using batches of size 50 and 500, we omit data points for large values of $\mathcal{C}(K_0)$.

In Figure 2 (a) (b) and (c), we plot the scaling of the zero-order complexity with $\mathcal{C}(K_0)$ for different values of the tolerance ϵ, and each figure additionally contains plots for different values of the batch-size. We observe that the scaling of our algorithm with respect to $\mathcal{C}(K_0)$ is approximately on the order of $\mathcal{O}(\mathcal{C}(K_0)^2)$, suggesting that our bounds for the Lipschitz and smoothness constants are not sharp in this respect. The same plots also demonstrate that using larger batch sizes, such as the algorithm from Fazel et al. [17], is often suboptimal: while the step size can be increased with increasing batch-size, it eventually plateaus due to stability considerations, leading to higher overall zero-order complexity.

We also ran our algorithm on the following problem introduced by Dean et al. [11], who used this example in their study of model based control methods for the LQR problem. Consider the LQR problem defined by:

$$A = \begin{bmatrix} 1.01 & 0.01 & 0 \\ 0.01 & 1.01 & 0.01 \\ 0 & 0.01 & 1.01 \end{bmatrix}, \quad B = I, \quad Q = 10^{-3} \times I, \quad R = I.$$

For three different values of $\mathcal{C}(K_0)$, we picked 8 evenly spaced (logarithmic scale) values of ϵ in the interval $(0.005, 1)$. The initial state was sampled uniformly at random from $\{[5, 0, 0], [5, 5, 5], [0, 0, 5]\}$. The cost of the optimal policy in our example was $\mathcal{C}(K^*) = 2.36$. We then measured the total zero order complexity required to attain ϵ convergence. These results are plotted in Figure 3, and confirm the prediction of Theorem 1.

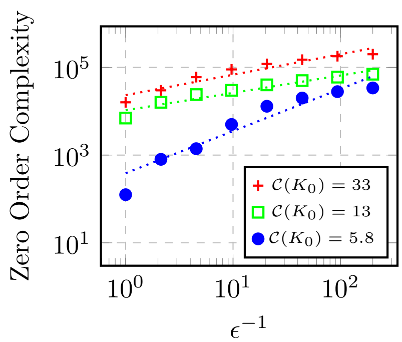

Figure 3. Scaling of complexity vs. ϵ^{-1} in LQR instance from Dean et al. [11]

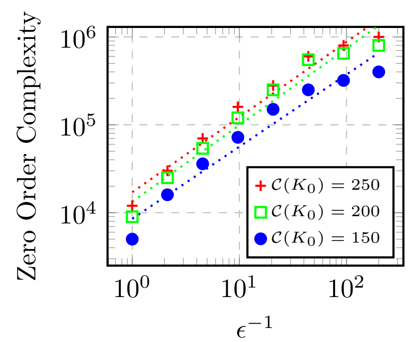

Figure 4. Scaling of complexity vs. ϵ^{-1} in randomly generated 8×8 example.

Finally, we also obtained data for the scaling with respect to ϵ on an example in slightly higher dimensions, to empirically verify the fact that our algorithm can be used for LQR problems larger than 3×3. We randomly generated A, B, Q and R as 8×8 matrices. Each entry of A was independently sampled from the Gaussian distribution $\mathcal{N}(2, 1)$, and each entry of B was independently sampled from the Gaussian distribution $\mathcal{N}(0, 1)$. To generate each of Q and R, we generated a matrix where each entry was independently sampled from the Gaussian distribution $\mathcal{N}(5, 1)$, then symmetrized the matrix by adding it to its transpose, finally adding $10I$ to ensure positive definiteness. The initial states were sampled uniformly at random from the canonical basis vectors. For three different values of $\mathcal{C}(K_0)$, we picked 8 evenly spaced (logarithmic scale) values of ϵ in the interval $(0.005, 1)$. We then measured the total zero order complexity required to attain ϵ convergence. These results are plotted in Figure 4, and confirm the prediction of Theorem 1.