2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

Natural Policy Gradient Preserves Spatial Decay Properties for Control of Networked Dynamical Systems

Eric Xu[1] and Guannan Qu[1]

Abstract— **We consider the distributed control of networked linear time-invariant systems. Previous work has established the spatial decay property of the centralized controller, which allows truncating the centralized controller to obtain a κ-hop distributed controller with small performance loss. This paper makes a step further by showing a policy optimization approach, Natural Policy Gradient (NPG), preserves the spatial decay property of controllers. This enables "truncating" Natural Policy Gradient to directly learn a κ-hop distributed controller.**

I. Introduction

Control of networked dynamical systems is a fundamental problem with applications spreading across power networks [1], vehicle platoons [2], traffic networks [3], and sensor networks [4]. In particular, there has been tremendous interest in designing distributed controllers such that the controller at each node only has access to the state information of itself and possibly nearby nodes [5]. Such interest in distributed controllers can be attributed to the fact that they are simple to implement, require only local communication, and are more scalable and practical compared to centralized controllers.

Despite the interest in distributed controllers, even for the simplest setting with linear dynamics and quadratic costs, designing the optimal distributed controller is extremely challenging in general. On one hand, computational intractability results have been established in the worst case [5]; on the other hand, many approaches have been proposed to design the optimal distributed controller under specific problem settings, e.g. the nested information structure [6], finite-dimensional linear policy [7], [8], quadratic invariance [9], sparsity invariance [10], system level synthesis [11], [12].

Recently, a promising line of work approaches the distributed control problem from a spatial decay perspective [8], [13]–[16]. The main idea is that when the dynamics follow a network structure, it is shown the optimal centralized controller satisfies the *spatial decay* property, namely, the control gain between the control action at node i and the state at node j decays when the distance between i and j increases. Here, the distance is measured with respect to the network graph that underlies the system dynamics. The spatial decay rate varies by the specific setting, e.g. [8] proves exponential decay rates for spatially invariant systems; [13], [14] proves exponential decay rates for networked systems, [16] proves sub-exponential decay rates for a general class of exponentially decaying dynamical systems, and [15] proves polynomial decay rates for a networked Markov Decision Process.

The significance of these spatial decaying results is that it allows truncating the optimal centralized controller to a *κ-hop distributed controller*, i.e. setting the gains between nodes with distance larger than κ to 0. Such a κ-hop distributed controller can be implemented in a distributed manner as it only requires communication with κ-hop neighbors. Further, the performance loss resulting from truncating is (sub)-exponentially small in κ. In other words, using a small κ will already have near-optimal performance when compared with the optimal centralized controller.

However, the aforementioned work only studies the spatial decay property of the optimal centralized controller and its truncation to the κ-hop distributed controller; there is a lack of study on how to learn the κ-hop distributed controller when the system matrices are unknown. Parallel to this effort, a recent line of work [17]–[23] (see [24] for a review) studies policy optimization for dynamical systems, which searches for the optimal centralized controller when the system dynamics are unknown. Specific methods inlcude policy gradient, natural policy gradient, etc [17]. Given the ability of the policy optimization methods to learn controllers, it is natural to ask: *is it possible to use policy optimization to learn near-optimal κ-hop distributed controllers?*

Contribution. In this paper, we make a step towards understanding the above question. Our key result is that if we start from a spatially decaying controller, we show that Natural Policy Gradient (NPG) preserves the spatially decaying property of the controller (Theorem 8). This means that Truncated NPG (TNPG), which directly learns a κ-hop distributed controller, has error decaying in κ in each step of the iteration (Theorem 9) when compared with the original NPG. Our results also lead to open questions regarding the uniform boundedness of the decay constant throughout the TNPG iterations and the overall convergence of TNPG, which we leave as future directions. Beyond the theoretical analysis, we show numerical simulations that validate our results and conjectures.

We note that the work in [25] establishes similar partial results as ours, particularly the decay property of the P matrix (cf. Theorem 8.1). What differentiates our work from [25] is that our goal is to search for near-optimal truncated controllers with convergence guarantee, using the spatial decay property as an intermediate tool; whereas [25] focuses specifically on the spatial decay property of spatially-decaying systems.

Notation. For a set S, its cardinality is written as $|S|$. The

[1]Department of Electrical and Computer Engineering, Carnegie Mellon University. Emails: ex@andrew.cmu.edu and gqu@andrew.cmu.edu

set of positive integers is denoted as $\mathbb{Z}_+$, the reals: $\mathbb{R}$, and the set of real $m \times n$ matrices: $\mathbb{R}^{m\times n}$. We then denote the spectral norm of a matrix A as $\|A\|$, the minimum singular value of A as $\sigma_{min}(A)$, $\mathrm{Tr}(A)$ the trace of A (assuming A is square), and denote $\|v\|_p$ for the L_p-norm of vector v. $A \succ 0$ denotes that A is a positive definite matrix. For a matrix A, we use lowercase letters $a(i,j)$ to denote A's sub-blocks which are indexed by i, j, hence we may write the matrix A as a sequence of sub-blocks $A = (a(i,j))_{i,j\in\mathcal{V}}$ for some index set $\mathcal{V}$. Similarly, for a vector v, we may write $v = (v(i))_{i\in\mathcal{V}}$. For a distribution $\mathcal{D}$ we denote $x \sim \mathcal{D}$ for a random variable x to be distributed as $\mathcal{D}$; we also denote its expectation as $\mathbb{E}[x]$. We further denote $O(\cdot)$ as the usual big-O notation and $poly(x)$ as a polynomial of x.

II. Preliminaries

A. Problem Formulation

We consider a graph $\mathcal{G} := \{\mathcal{V}, \mathcal{E}\}$, where $\mathcal{V} := \{1, \ldots, N\}$ is the set of nodes and $\mathcal{E}$ is the set of edges. We let $\mathcal{N}_\mathcal{G}(i)$ denote the set of neighbors of i in the graph (including i itself), and $d_\mathcal{G}(i,j)$ denote the shortest path distance between node i and j in the graph. We also make the following assumption on the polynomial growth of the cardinality of τ-hop neighborhoods, which is standard in the literature [14].

Assumption 1. *For any τ-hop neighbourhood of a node, that is $\mathcal{N}_\mathcal{G}(i,\tau) := \{j \in \mathcal{V} : d_\mathcal{G}(i,j) < \tau\}$, there exists positive constants $C_P \geq 1$ and d s.t.*

$$|\mathcal{N}_\mathcal{G}(i,\tau)| \leq C_P \tau^d, \quad \forall i \in \mathcal{V} \text{ and } \tau \geq 1. \tag{1}$$

Given the graph $\mathcal{G}$, we consider the following networked Linear Time Invariant (LTI) system:

$$x_{t+1}(i) = \sum_{j\in\mathcal{N}_\mathcal{G}(i)} a(i,j)x_t(j) + b(i,j)u_t(j), \quad i \in \mathcal{V}. \tag{2}$$

Here, $x_t(i) \in \mathbb{R}^{n_{x_i}}$, and $u_t(i) \in \mathbb{R}^{n_{u_i}}$ are the state and control at node $i \in \mathcal{V}$ at time $t \in \mathbb{Z}_+$; furthermore, $a(i,j) \in \mathbb{R}^{n_{x_i}\times n_{x_j}}$ and $b(i,j) \in \mathbb{R}^{n_{x_i}\times n_{u_j}}$. We also denote $n_x = \sum_{i\in\mathcal{V}} n_{x_i}$ and $n_u = \sum_{i\in\mathcal{V}} n_{u_i}$. Given $x(i), u(i)$ at node i, the stage cost at node i is given by,

$$\ell_i(x(i), u(i)) := x(i)^\top q(i,i)x(i) + u(i)^\top r(i,i)u(i),$$

where $q(i,i) \in \mathbb{R}^{n_{x_i}\times n_{x_i}}$ and $r(i,i) \in \mathbb{R}^{n_{u_i}\times n_{u_i}}$ are symmetric and positive definite matrices. Given the stage costs, the nodes seek to minimize the sum of their stage costs for an infinite horizon:

$$\begin{aligned} &\min \mathbb{E}\left[\sum_{t=0}^{\infty} x_t^\top Q x_t + u_t^\top R u_t\right] \\ &\text{s.t. } x_{t+1} = Ax_t + Bu_t, \quad x_0 \sim \mathcal{D}, \end{aligned} \tag{3}$$

where $x_t := (x_t(i))_{i\in\mathcal{V}}$, $u_t := (u_t(i))_{i\in\mathcal{V}}$, $A := (a(i,j))_{i,j\in\mathcal{V}}$, $B := (b(i,j))_{i,j\in\mathcal{V}}$, $Q := (q(i,j))_{i,j\in\mathcal{V}}$, and $R := (r(i,j))_{i,j\in\mathcal{V}}$ (we note that $a(i,j) = 0$ and $b(i,j) = 0$ for $\{i,j\} \notin \mathcal{E}$, and $q(i,j) = 0$ and $r(i,j) = 0$ for $i \neq j$). In addition, $\mathcal{D}$ is the distribution for the initial state, which we assume has correlation matrix $\Sigma_0 := \mathbb{E}\left[x_0 x_0^\top\right] \succ 0$.

An important class of controllers for (3) is the class of linear controllers $u_t = -Kx_t$, where K is the control gain. An important concept for these controllers is stability, defined below.

Definition 2. *The controller K is (L,γ)-stable for $L > 0, \gamma \in (0,1)$ if $\|K\| \leq L$ and $\|(A - BK)^t\| \leq L\gamma^t$ for all $t \in \mathbb{Z}_+$.*

Throughout the paper, we also make the following assumptions.

Assumption 3. *We assume $Q \succ 0$, $R \succ 0$, (A, B) is controllable, and $(A, Q^{1/2})$ is detectable.*

It is well known that under the above assumption, the optimal policy is the linear state feedback law: $u^*(x) = -K^*x$ where $K^* = (k^*(i,j))_{i,j\in\mathcal{V}}$ is the optimal gain matrix. However, the optimal controller, K^*, is usually dense and impractical to implement. Therefore, it is preferred to focus on distributed controllers. Specifically, we consider the class of κ-hop distributed controllers defined as follows.

Definition 4. *A controller $K = (k(i,j))_{i,j\in\mathcal{V}}$ is said to be κ-hop distributed if $k(i,j) = 0$ when $d_\mathcal{G}(i,j) \geq \kappa$.*

Thus, a κ-hop distributed controller is structured such that the control action at node i has a non-zero gain from the state at node j only when i and j are within distance κ. Therefore, a κ-hop distributed controller is more practical to implement than a centralized controller, at least for small κ.

The goal of our paper is to search for a κ-hop distributed controller that minimizes (3). Towards this end, we utilize policy optimization, which we review next:

B. Policy Optimization

Restricting to the class of linear stabilizing state feedback control policies: $u_t = -Kx_t$, we may rewrite the cost in (3) as

$$C(K) := \mathbb{E}[x_0^\top P_K x_0], \tag{4}$$

where

$$P_K = Q + K^\top RK + (A - BK)^\top P_K(A - BK). \tag{5}$$

It has been shown in [17] the gradient of (4) is

$$\nabla C(K) = 2\underbrace{\left((R + B^\top P_K B)K - B^\top P_K A\right)}_{:=E_K}\Sigma_K, \tag{6}$$

where Σ_K is the state correlation matrix defined as

$$\Sigma_K := \mathbb{E}_K\left[\sum_{t=0}^{\infty} x_t x_t^\top\right],$$

and $\mathbb{E}_K$ is taken w.r.t. the randomness on x_0 and assumes the state trajectory is generated by the controller $u_t = -Kx_t$. [17] also considers the Natural Policy Gradient (NPG),

$$\nabla_N C(K) = \nabla C(K)\Sigma_K^{-1} = 2E_K. \tag{7}$$

The concept of NPG naturally leads to the NPG algorithm:

$$K_{t+1} = K_t - \eta E_{K_t}, \tag{8}$$

for some step size η. The key contribution of [17] is that (4) satisfies the so-called gradient dominance property:

$$C(K) - C(K^*) \leq \frac{\|\Sigma_{K^*}\|}{\sigma_{min}(R)} \|E_K\|_F^2, \tag{9}$$

where $\|\cdot\|_F$ denotes the Frobenius norm: $\|A\|_F^2 = \mathrm{Tr}(A^\top A)$ which is induced by the inner product: $\langle A, B\rangle_F = \mathrm{Tr}(A^\top B)$.

A significant consequence of the gradient dominance property is that the NPG algorithm (8) converges linearly to the optimal LQR cost.

Lemma 5 (Lemma 15 in [17])**.** *Applying a single NPG step $K' = K - \eta E_K$ with $\eta \leq \frac{1}{\|R+B^\top P_K B\|}$, we have for $\mu := \sigma_{min}(\Sigma_0)$ and $\lambda := \frac{\sigma_{min}(R)}{\|\Sigma_{K^*}\|}$,*

$$C(K') - C(K^*) \leq (1 - \eta\lambda\mu)\left(C(K) - C(K^*)\right).$$

As can be seen from above, after applying one step of NPG, the optimality gap $(C(K) - C(K^*))$ shrinks by a fixed constant, and hence the NPG converges to the optimal controller with a linear convergence rate. It is natural to ask whether NPG can learn κ-hop distributed controllers, which is the main focus of Section III. Before we present the main results, we review one final concept, spatially decaying matrices, which will be the key to our analysis later.

C. Spatially Decaying Matrices

From [14], it is known that the optimal centralized controller K^* to (3) is *spatially exponentially decaying*, i.e. $\|k^*(i,j)\|$ decays exponentially as the graph distance between i and j increases. This motivates us to restrict the class of policies we search through to such controllers. To formalize this, we define the general spatial decay property below.

Definition 6. *Given a matrix $A = [a(i,j)]_{i,j\in\mathcal{V}}$ and a weight $w = [w(i,j)]_{i,j\in\mathcal{V}}$ where each $w(i,j) \in \mathbb{R}_+$, the decay constant of A under weight w is defined as*

$$\alpha_w(A) := \max_{i\in\mathcal{V}} \sum_{j\in\mathcal{V}} \|a(i,j)\| w(i,j) + \max_{j\in\mathcal{V}} \sum_{i\in\mathcal{V}} \|a(i,j)\| w(i,j). \tag{10}$$

Throughout this paper, we will be using the class of (D,δ)-sub-exponential decay weights defined as follows:

$$w(i,j) = \exp(D d_{\mathcal{G}}(i,j)^\delta) \tag{11}$$

for $D \in (0,\infty)$ and $\delta \in (0,1)$. Under this weight, we call a matrix A a *sub-exponentially decaying matrix*. Furthermore, one implication of Definition 6 is as follows:

$$\|a(i,j)\| = \frac{1}{w(i,j)} \|a(i,j)\| w(i,j) \leq \alpha_w(A) \exp(-D d_{\mathcal{G}}(i,j)^\delta). \tag{12}$$

In other words, the block $a(i,j)$ decays sub-exponentially when $d_{\mathcal{G}}(i,j)$ increases. We note that for the above definition to be not vacuous, we need $\alpha_w(A)$ to not depend on N, the size of the network; as otherwise, setting $\alpha_w(A) = \exp(DN^\delta)$ will render (12) degenerate.

Further, it is straightforward to prove the following properties whose proofs are given in Appendix A.

Lemma 7. *Given (D,δ)-sub-exponential decay weights w, matrices A and B, we have*

$$\alpha_w(A+B) \leq \alpha_w(A) + \alpha_w(B), \tag{13}$$

$$\alpha_w(AB) \leq \alpha_w(A)\alpha_w(B). \tag{14}$$

III. Main Results

Our first key result is that, the NPG update equation preserves the spatial decaying structure. This result is formally stated below.

Theorem 8. *Suppose K is (L,γ)-stable and is a sub-exponentially decaying matrix in the sense of Definition 6 with weights* (11) *and decay constant $\alpha_w(K)$. Then, the matrix E_K in NPG* (7) *is also sub-exponentially decaying with decay constant*

$$\alpha_w(E_K) \leq (\alpha_w(R) + \alpha_w(B)^2\alpha_w(P_K))\alpha_w(K) + \alpha_w(B)\alpha_w(P_K)\alpha_w(A),$$

with

$$\alpha_w(P_K) \leq \alpha_w(K)^2 e^{O\left((\log(\alpha_w(K)))^{\frac{1}{1-\log_2(1+\theta)}}\right)},$$

where $\theta = \frac{1+\frac{1}{3-2^\delta}}{2} \in (0,1)$ and the $O(\cdot)$ notation hides everything that does not depend on $\alpha_w(K)$; the precise form of $\alpha_w(P_K)$ is given in (19). *As a result, a single step of NPG, $K' = K - \eta E_K$, will produce a controller K' that is also sub-exponentially decaying with decay constant,*

$$\alpha_w(K') \leq \alpha_w(K)^3 e^{O\left((\log(\alpha_w(K)))^{\frac{1}{1-\log_2(1+\theta)}}\right)}. \tag{15}$$

The precise form showing dependence on all constants is given in (20).

The consequence of the above result is that, if we start from a sub-exponentially decaying stable controller, the NPG algorithm will produce controllers that are also sub-exponentially decaying. Further, the NPG E_K is also sub-exponentially decaying. As a result, "truncating" the NPG to κ-hop distributed controllers will lead to only a small error. Such a truncated NPG directly learns a κ-hop distributed controller, and since the truncation error is small, we expect it to have similar convergence properties as the original NPG (Lemma 5).

More specifically, the truncated NPG is described as follows

$$K_{t+1} = K_t - \eta \mathcal{T}_{\mathcal{G}}(E_{K_t}, \kappa), \tag{16}$$

for some step size η. We can define the truncation operator in terms of a (dense) gain matrix K; its truncation is $K^\kappa := (k^\kappa(i,j))_{i,j\in\mathcal{V}}$ defined as follows:

$$k^\kappa(i,j) = \mathcal{T}_{\mathcal{G}}(k(i,j),\kappa) := \begin{cases} k(i,j), & d_{\mathcal{G}}(i,j) < \kappa, \\ 0, & d_{\mathcal{G}}(i,j) \geq \kappa. \end{cases} \tag{17}$$

Overloading notation, we may write $K^\kappa = \mathcal{T}_\mathcal{G}(K, \kappa)$.

Leveraging Theorem 8, we also show the following single-step error guarantee for the truncated NPG.

Theorem 9. *Suppose K is (L, γ)-stable and is a κ-hop distributed controller with decay constant $\alpha_w(K)$. Then, for $\eta \leq \frac{3}{8\|R+B^\top P_K B\|}$, one step of truncated NPG, $K' = K - \eta \mathcal{T}_\mathcal{G}(E_K, \kappa)$, will produce another κ-hop distributed controller K' that satisfies:*

$$C(K') - C(K^*) \leq (1 - \eta\lambda\mu)\,(C(K) - C(K^*)) + 5\eta \frac{C(K')}{\sigma_{min}(Q)} \epsilon(\kappa), \quad (18)$$

where $n_x^* := \max_{i \in \mathcal{V}} n_{x_i}$, and the truncation error is: $\epsilon(\kappa) := \alpha_w(E_K)^2 N n_x^* e^{-2D\kappa^\delta}$ which is sub-exponentially decaying in κ. $\alpha_w(E_K)$ is given as in Theorem 8.

The first term in Theorem 9 is similar to Lemma 5 in that the optimality gap shrinks. The second term is a result of the truncation and is small due to the sub-exponential decay property, provided κ is large.

While we can in principle apply Theorem 8 and Theorem 9 multiple times to analyze the convergence of TNPG across multiple iterations, the decay constant $\alpha_w(K)$ will increase after each NPG step and will remain unbounded. This can be seen from the results in Theorem 8, and one can check $\alpha(K') \sim e^{poly(\log \alpha(K))}$. Therefore, our existing results cannot directly lead to the convergence of TNPG. However, we conjecture that our bound in Theorem 8 is not tight. Our numerical simulations show that $\alpha_w(K)$ will remain uniformly bounded throughout the NPG iterations. How to show the boundedness of $\alpha_w(K)$ throughout the iterations, and how to prove the overall convergence of TNPG and its model-free, distributed implementation, remains our ongoing work. Finally, we note that while the results above hold for NPG, we conjecture that similar results (i.e., Theorems 8 and 9) should also hold for vanilla policy gradient. In particular, we expect that the decay constant of the gradient in (6) will be larger due to the extra factor of Σ_K.

IV. Proof

A. Proof of Theorem 8

By Lemma 7, it is easy to see that

$$\alpha_w(K') \leq \alpha_w(K) + \eta\alpha(E_K),$$

and further,

$$\alpha_w(E_K) \leq (\alpha_w(R) + \alpha_w(B)^2\alpha_w(P_K))\alpha_w(K) + \alpha_w(B)\alpha_w(P_K)\alpha_w(A).$$

Therefore, it is critical to bound $\alpha_w(P_K)$, that is P_K preserves the sub-exponential decay property. Note that P_K can be written as,

$$P_K = \sum_{n=0}^{\infty} ((A - BK)^\top)^n (Q + K^\top RK)(A - BK)^n.$$

Therefore, to show P_K is sub-exponentially decaying, we need to show that each $(A - BK)^n$ is sub-exponentially decaying with decay constant $\alpha_w((A-BK)^n)$ converging to 0 as $n \to \infty$. This is expected as we know $(A-BK)^n$ converges to the zero matrix since K is (L, γ)-stable. However, we note that directly applying the sub-multiplicative bounds in Lemma 7 is not sufficient for proving $\alpha_w((A - BK)^n)$ converges to 0, as the sub-multiplicative bound leads to $\alpha_w((A-BK)^n) \leq (\alpha_w(A-BK))^n$, yet $\alpha_w(A-BK)$ may not be smaller than 1. Therefore, to bound $\alpha_w((A-BK)^n)$, we need to relate it to the fact that $(A-BK)$ is (L, γ)-stable. More specifically, we prove a general result that shows how the decay constant for the power of a matrix changes when the power increases.

Lemma 10. *If M is sub-exponentially decaying with decay constant $\alpha_w(M)$, the weight w is as in (11), and $\|M^n\| \leq L\gamma^n$ for some $L > 0, \gamma \in (0, 1)$, then*

$$\alpha_w(M^n) \leq C_L^{-1/\theta} \gamma^n \left(\frac{C_L^{1/\theta} \alpha_w(M)}{\gamma} \right)^{\frac{1+\theta}{\theta} n^{\log_2(1+\theta)}},$$

where $C_L := \max(1, 2CL^{1-\theta})$ with C and θ defined in (23).

The proof of Lemma 10 can be found in Section IV-B. From Lemma 10, we can see that $\alpha_w(A^n)$ can be controlled by two factors: the first of which is γ^n which decays to 0 exponentially fast; the second factor increases with n, though in a sub-exponential manner (note $\log_2(1+\theta) < 1$). Therefore, $\alpha_w(A^n)$ decays to 0 exponentially.

With the help of Lemma 10, we now proceed to prove Theorem 8. Applying Lemma 10 to $A_K := A - BK$, we have that

$$\alpha_w(A_K^n) \leq C_L^{-1/\theta} \gamma^n \left(\frac{C_L^{1/\theta} \alpha_w(A_K)}{\gamma} \right)^{\frac{1+\theta}{\theta} n^{\log_2(1+\theta)}}.$$

Then from the definition of P_K, we have that,

$$\begin{aligned}
\alpha_w(P_K) &\leq (\alpha_w(Q) + \alpha_w(K)^2\alpha_w(R)) \sum_{n=0}^{\infty} \alpha_w(A_K^n)^2 \\
&\leq \alpha_{Q_K} C_L^{-2/\theta} \sum_{n=0}^{\infty} \gamma^{2n} \left(\frac{C_L^{1/\theta} \alpha_w(A_K)}{\gamma} \right)^{\frac{2(1+\theta)}{\theta} n^{\log_2(1+\theta)}} \\
&= \alpha_{Q_K} C_L^{-2/\theta} \sum_{n=0}^{\infty} \gamma^n e^{n \log\gamma + n^{\log_2(1+\theta)} \log\zeta(A_K)} \\
&\leq \frac{\alpha_{Q_K} C_L^{-2/\theta}}{1-\gamma} e^{(\log\zeta(A_K))^{\frac{1}{1-\xi}} \left[\left| \frac{\xi}{\log\gamma} \right|^{\frac{\xi}{1-\xi}} - (\xi|\log\gamma|^\xi)^{\frac{1}{1-\xi}} \right]},
\end{aligned} \quad (19)$$

where $\zeta(A_K) := \left(\frac{C_L^{1/\theta}\alpha_w(A_K)}{\gamma} \right)^{\frac{2(1+\theta)}{\theta}}$, $\alpha_{Q_K} = \alpha_w(Q) + \alpha_w(K)^2\alpha_w(R)$, and $\xi = \log_2(1+\theta)$.

Therefore, we have that for NPG: $K' = K - \eta E_K$

$$\begin{aligned}
\alpha_w(K') &\leq \alpha_w(K)\,(1 + \eta\alpha_w(R)) \\
&+ \eta(\alpha_w(B)^2\alpha_w(K) + \alpha_w(B)\alpha_w(A)) \\
&\cdot \left(\frac{\alpha_w(Q) + \alpha_w(R)\alpha_w(K)^2}{(1-\gamma)} \right)
\end{aligned}$$

$$\cdot \max\big\{1, 4C_P\big[1+\big(\frac{2d}{D\delta(2-2^\delta)}\big)^{d/\delta}\big]L^{1-\theta}\big\}^{-2/\theta}$$
$$\cdot \exp\left[\left(\frac{2(1+\theta)}{\theta}\right)^{\frac{1}{1-\log_2(1+\theta)}}\right.$$
$$\cdot\Bigg(\log\alpha_w(A_K)-\log\gamma+\frac{1}{\theta}\log\max\Bigg\{1,4C_P$$
$$\cdot(1+\left(\frac{2d}{D\delta(2-2^\delta)}\right)^{\frac{d}{\delta}}L^{1-\theta})\Bigg)\Bigg\}^{\frac{1}{1-\log_2(1+\theta)}}$$
$$\cdot\Bigg(\left|\frac{\log_2(1+\theta)}{\log\gamma}\right|^{\frac{1}{1-\log_2(1+\theta)}-1}-\Bigg(\log_2(1+\theta)$$
$$\left.\cdot|\log\gamma|^{\log_2(1+\theta)}\Bigg)^{1/(1-\log_2(1+\theta))}\Bigg).\right] \tag{20}$$

This concludes the proof of Theorem 8. □

B. Proof of Lemma 10

We acknowledge that our proof of Lemma 10 is based on a modification of the analysis in [26]. To start, we prove the following Lemma 11, which reveals a few important properties of the sub-exponential weight (11).

Lemma 11. *Consider the weight in (11) and define an auxiliary weight* $v(i,j) := \exp(D(2^\delta-1)d_{\mathcal{G}}(i,j)^\delta)$. *Then, the following conditions hold*

$$w(i,j)\le w(i,k)v(k,j)+v(i,k)w(k,j),\quad \forall i,j,k\in\mathcal{V} \tag{21}$$

$$\max_{i\in\mathcal{V}}\|(vw^{-1})(i,\cdot)\|_\infty+\max_{j\in\mathcal{V}}\|(vw^{-1})(\cdot,j)\|_\infty\le 2, \tag{22}$$

and

$$\inf_{\tau>0} a(\tau)+b(\tau)t\le Ct^\theta,\quad \forall\, t\ge 1 \tag{23}$$

where the notation $(vw^{-1})(i,\cdot)$ *is the vector containing the products* $v(i,j)w^{-1}(i,j)$ *for all* $j\in\mathcal{V}$ *and similarly for* $(vw^{-1})(\cdot,j)$, $C = 2C_P\left(1+\left(\frac{2d}{D\delta(2-2^\delta)}\right)^{d/\delta}\right)$, *and* $\theta = \frac{2-2^{\delta-1}}{3-2^\delta}$; *the functions* $a(\cdot)$ *and* $b(\cdot)$ *are defined as*

$$a(\tau) := \max_{i\in\mathcal{V}}\sum_{j\in\mathcal{N}_{\mathcal{G}}(i,\tau)}v(i,j)+\max_{j\in\mathcal{V}}\sum_{i\in\mathcal{N}_{\mathcal{G}}(j,\tau)}v(i,j) \tag{24}$$

and

$$b(\tau) := \max_{i\in\mathcal{V}}\max_{j\in\mathcal{V}\setminus\mathcal{N}_{\mathcal{G}}(i,\tau)}vw^{-1}(i,j)+\max_{j\in\mathcal{V}}\max_{i\in\mathcal{V}\setminus\mathcal{N}_{\mathcal{G}}(j,\tau)}vw^{-1}(i,j). \tag{25}$$

The proof of Lemma 11 can be found in Appendix A. In Lemma 11, we can see that (21) is a stronger condition than the triangle inequality in the view that the auxiliary weight v, is a slower growing weight than w, as seen in (22). The final condition (23) will be useful later when we split the summation over $\mathcal{V}$ into $\mathcal{N}_{\mathcal{G}}(i,\tau)$ and $\mathcal{V}\setminus\mathcal{N}_{\mathcal{G}}(j,\tau)$.

With the help of Lemma 11, we are now ready to prove Lemma 10.

Proof of Lemma 10. We denote $M^n = (m^n(i,j))_{i,j\in\mathcal{V}}$. Given a positive integer τ, we note that

$$\sum_{j\in\mathcal{V}}\|m^n(i,j)\|v(i,j)$$
$$=\left(\sum_{j:d_{\mathcal{G}}(i,j)<\tau}+\sum_{j:d_{\mathcal{G}}(i,j)\ge\tau}\right)\|m^n(i,j)\|v(i,j)$$
$$\le\max_{j\in\mathcal{V}}(\|m^n(i,j)\|)\sum_{j\in\mathcal{N}_{\mathcal{G}}(i,\tau)}v(i,j)$$
$$+\left(\sum_{j\in\mathcal{V}\setminus\mathcal{N}_{\mathcal{G}}(i,\tau)}\|m^n(i,j)\|w(i,j)\right)\max_{j\in\mathcal{V}\setminus\mathcal{N}_G(i,\tau)}vw^{-1}(i,j)$$
$$\le\|M^n\|\sum_{j\in\mathcal{N}_{\mathcal{G}}(i,\tau)}v(i,j)+\alpha_w(M^n)\max_{j\in\mathcal{V}\setminus\mathcal{N}_G(i,\tau)}vw^{-1}(i,j),$$

where in the last inequality, we have used $\|m^n(i,j)\|\le\|M^n\|$, i.e. the norm of the submatrix is upper bounded by the norm of the entire matrix. Repeating the same bound for $\sum_{i\in\mathcal{V}}\|m^n(i,j)\|v(i,j)$ and by the definition of $\alpha_v(\cdot)$, $a(\tau)$, and $b(\tau)$ (in Lemma 11), we have

$$\begin{aligned}\alpha_v(M^n)&\le a(\tau)\|M^n\|+b(\tau)\alpha_w(M^n)\\&\le\inf_{\tau>0}a(\tau)\|M^n\|+b(\tau)\alpha_w(M^n)\\&\le C\|M^n\|^{1-\theta}\alpha_w(M^n)^\theta,\end{aligned} \tag{26}$$

where the third inequality is from (23).

Since w satisfies (21), we have that for any $Z = XY$, where $Z = (z(i,j))_{i,j\in\mathcal{V}}$, $X = (x(i,j))_{i,j\in\mathcal{V}}$, $Y = (y(i,j))_{i,j\in\mathcal{V}}$,

$$\begin{aligned}\|z(i,j)\|w(i,j)&\le\sum_{k\in\mathcal{V}}\|x(i,k)y(k,j)\|w(i,j)\\&\le\sum_{k\in\mathcal{V}}\|x(i,k)\|w(i,k)\|y(k,j)\|v(k,j)\\&\quad+\sum_{k\in\mathcal{V}}\|x(i,k)\|v(i,k)\|y(k,j)\|w(k,j).\end{aligned}$$

Taking the summation over $j\in\mathcal{V}$, we get that

$$\alpha_w(Z)\le\alpha_w(X)\alpha_v(Y)+\alpha_v(X)\alpha_w(Y). \tag{27}$$

Then using (26) and (27), we get

$$\begin{aligned}\alpha_w(M^{2n})&\le 2\alpha_w(M^n)\alpha_v(M^n)\\&\le 2\alpha_w(M^n)^{1+\theta}C\|M^n\|^{1-\theta}\\&\le 2C\alpha_w(M^n)^{1+\theta}L^{1-\theta}\gamma^{n(1-\theta)}.\end{aligned} \tag{28}$$

Further, using Lemma 7, we have,

$$\begin{aligned}\alpha_w(M^{2n+1})&\le\alpha_w(M^{2n})\alpha_w(M)\\&\le 2C\alpha_w(M)\alpha_w(M^n)^{1+\theta}L^{1-\theta}\gamma^{n(1-\theta)}.\end{aligned} \tag{29}$$

Next, we will use an induction argument based on (28) and (29) to finish the proof. We first define

$$b_n := (2CL^{1-\theta})^{1/\theta}\alpha_w(M^n)\gamma^{-n},$$

and rewrite (28) and (29) as

$$b_{2n}\le b_n^{1+\theta},\qquad b_{2n+1}\le c_0 b_n^{1+\theta}, \tag{30}$$

for all $n\ge 1$ and $c_0 := \max(1,(2CL^{1-\theta})^{1/\theta})\alpha_w(M)\gamma^{-1}$. Then we show by induction our sequence b_n is bounded above by

$$b_n\le c_0^{\sum_{i=0}^k e_i(1+\theta)^i},$$

for $n = \sum_{i=0}^{k} e_i 2^i$ where $e_i \in \{0,1\}$, $e_k = 1$. We first note that $b_1 = (2CL^{1-\theta})^{1/\theta}\alpha_w(M)\gamma^{-1} \le c_0$, so the base case is satisfied. For even $n+1$, we let $n+1 = \sum_{i=1}^{k} e_i 2^i$, and so $\frac{n+1}{2} = \sum_{i=1}^{k} e_i 2^{i-1} = \sum_{j=0}^{k-1} e_{j+1}2^j$. Therefore, from the induction hypothesis and (30),

$$b_{n+1} \le b_{\frac{n+1}{2}}^{1+\theta} \le c_0^{(1+\theta)\sum_{i=0}^{k-1} e_{i+1}(1+\theta)^i} = c_0^{\sum_{i=1}^{k} e_i(1+\theta)^i},$$

so the induction hypothesis is true for $n+1$. The case when $n+1$ is odd is similar so we omit the details here. Thus, the induction is concluded.

Then, applying the inequality

$$\begin{aligned}\sum_{i=0}^{k} e_i(1+\theta)^i &\le \sum_{i=0}^{k}(1+\theta)^i = \frac{(1+\theta)^{k+1}-1}{\theta}\\ &\le \frac{1+\theta}{\theta}(1+\theta)^k = \frac{1+\theta}{\theta}(2^k)^{\log_2(1+\theta)}\\ &\le \frac{1+\theta}{\theta} n^{\log_2(1+\theta)},\end{aligned}$$

we have that

$$\alpha_w(M^n) \le C_L^{-1/\theta}\gamma^n \left(C_L^{1/\theta}\alpha_w(M)\gamma^{-1}\right)^{\frac{1+\theta}{\theta}n^{\log_2(1+\theta)}}, \tag{31}$$

where $C_L := \max(1, 2CL^{1-\theta})$. This concludes the proof. □

C. Proof of Theorem 9

Denoting $G_K := \mathcal{T}_{\mathcal{G}}(E_K, \kappa)$ and $G_K^{\perp} := E_K - G_K$, our proof starts with a bound on the error caused by truncation $G_K^{\perp}$, which is a direct consequence of Theorem 8.

Lemma 12. *We have the following bound on $G_K^{\perp}$:*

$$\mathrm{Tr}((G_K^{\perp})^\top G_K^{\perp}) \le \epsilon(\kappa) := \alpha_w(E_K)^2 N n_{x^*} e^{-2D\kappa^\delta}.$$

Proof. We denote $E_K = (e_K(i,j))_{i,j\in\mathcal{V}}$, then

$$\begin{aligned}\|G_K^{\perp}\|_F^2 &= \mathrm{Tr}\left((G_K^{\perp})^\top G_K^{\perp}\right)\\ &= \sum_{i\in\mathcal{V}}\sum_{k:d_{\mathcal{G}}(i,k)\ge\kappa} \mathrm{Tr}\left(e_K(i,k)^\top e_K(i,k)\right)\\ &\le \sum_{i\in\mathcal{V}}\sum_{k:d_{\mathcal{G}}(i,k)\ge\kappa} n_{x_i}\|e_K(i,k)\|^2\\ &\le n_x^*\sum_{i\in\mathcal{V}}\left(\sum_{k:d_{\mathcal{G}}(i,k)\ge\kappa}\|e_K(i,k)\|\right)^2\\ &\le N n_x^* \alpha_w(E_K)^2 e^{-2D\kappa^\delta},\end{aligned}$$

where $n_x^* = \max_{i\in\mathcal{V}} n_{x_i}$ and the last inequality can be shown from the definition of the decay constant in (10):

$$\begin{aligned}\sum_{k:d_{\mathcal{G}}(i,k)\ge\kappa}\|e_K(i,k)\| &= \sum_{k:d_{\mathcal{G}}(i,k)\ge\kappa}\|e_K(i,k)\|w(i,k)\frac{1}{w(i,k)}\\ &\le e^{-D\kappa^\delta}\alpha_w(E_K).\end{aligned}$$

□

Our proof also uses the following result in [17].

Lemma 13 (Lemma 6 in [17])**.** *The cost $C(K)$ in (4) satisfies*

$$\begin{aligned}C(K') - C(K) &= -2\mathrm{Tr}(\Sigma_{K'}(K-K')^\top E_K)\\ &\quad + \mathrm{Tr}(\Sigma_{K'}(K-K')^\top(R + B^\top P_K B)(K-K')).\end{aligned} \tag{32}$$

We begin with a κ-hop policy K. We can write our Truncated NPG update as $K' = K - \eta G_K$. Then from (32), we get

$$\begin{aligned}C(K') - C(K) &= -2\eta\mathrm{Tr}(\Sigma_{K'}G_K^\top E_K)\\ &\quad + \eta^2\mathrm{Tr}(\Sigma_{K'}G_K^\top(R + B^\top P_K B)G_K).\end{aligned} \tag{33}$$

Analyzing the first term in (33), we have that

$$\mathrm{Tr}(\Sigma_{K'}G_K^\top E_K) = \mathrm{Tr}(\Sigma_{K'}E_K^\top E_K) - \mathrm{Tr}(\Sigma_{K'}(G_K^{\perp})^\top E_K),$$

for which the second term can be bounded as

$$\begin{aligned}&2\eta\mathrm{Tr}(\Sigma_{K'}(G_K^{\perp})^\top E_K) = 2\eta\langle\Sigma_{K'}^{1/2}E_K^\top, \Sigma_{K'}^{1/2}(G_K^{\perp})^\top\rangle_F\\ &\le \frac{\eta}{4}\|\Sigma_{K'}^{1/2}E_K^\top\|_F^2 + 4\eta\|\Sigma_{K'}^{1/2}(G_K^{\perp})^\top\|_F^2\\ &= \frac{\eta}{4}\mathrm{Tr}(\Sigma_{K'}E_K^\top E_K) + 4\eta\mathrm{Tr}(\Sigma_{K'}(G_K^{\perp})^\top G_K^{\perp}).\end{aligned}$$

Denoting $R_K := R + B^\top P_K B$, for the second term in (33), we have

$$\begin{aligned}&\mathrm{Tr}(\Sigma_{K'}G_K^\top R_K G_K) =\\ &\mathrm{Tr}(\Sigma_{K'}^{1/2}G_K^\top R_K^{1/2}R_K^{1/2}G_K\Sigma_{K'}^{1/2}) = \|R_K^{1/2}G_K\Sigma_{K'}^{1/2}\|_F^2\\ &= \|R_K^{1/2}(E_K + G_K - E_K)\Sigma_{K'}^{1/2}\|_F^2\\ &\le 2\|R_K^{1/2}E_K\Sigma_{K'}^{1/2}\|_F^2 + 2\|R_K^{1/2}G_K^{\perp}\Sigma_{K'}^{1/2}\|_F^2\\ &\le 2\|R_K\|\left(\mathrm{Tr}(\Sigma_{K'}E_K^\top E_K) + \mathrm{Tr}(\Sigma_{K'}(G_K^{\perp})^\top G_K^{\perp})\right).\end{aligned}$$

Then taking $\eta \le \frac{3}{8\|R_K\|}$, and combining the above upper bounds into (33), we have for $\lambda := \frac{\sigma_{min}(R)}{\|\Sigma_{K^*}\|}$,

$$\begin{aligned}C(K') - C(K) &\le -\eta\mathrm{Tr}(\Sigma_{K'}E_K^\top E_K) + 5\eta\mathrm{Tr}(\Sigma_{K'}(G_K^{\perp})^\top G_K^{\perp})\\ &\le -\lambda\eta\mu(C(K) - C(K^*)) + 5\eta\|\Sigma_{K'}\|\epsilon(\kappa)\\ &\le -\lambda\eta\mu(C(K) - C(K^*)) + 5\eta\frac{C(K')}{\sigma_{min}(Q)}\epsilon(\kappa),\end{aligned}$$

where we have used (9), Lemma 12, and the fact that $C(K') = \mathrm{Tr}((Q + K'^\top RK')\Sigma_{K'}) \ge \|\Sigma_{K'}\|\sigma_{min}(Q)$.

V. Numerical Simulations

For our simulations, we consider the AC network in [14], described below.

Example 14. *The frequency control of a DC-approximated AC power system can be written as a networked control problem. The graph of this model is $\mathcal{G}_{DC} := \{\mathcal{V}_{DC}, \mathcal{E}_{DC}\}$ with N_{DC} nodes (busses) and weighted edges with weights $(k_{DC})_{ij}$ (normalized line susceptances). Forward Euler with sampling time $t_{s,DC}$ gives*

$$\begin{bmatrix}\theta_{t+1}\\ \omega_{t+1}\end{bmatrix} = \begin{bmatrix} I & (t_{s,DC})I\\ -(t_{s,DC})L_{DC} & I\end{bmatrix}\begin{bmatrix}\theta_t\\ \omega_t\end{bmatrix} + \begin{bmatrix}0\\ 0.5(t_{s,DC})I\end{bmatrix}u_t,$$

where θ_t denotes the vector of phase angles of each bus at time t, $\theta_t = (\theta_t(i))_{i\in\mathcal{V}_{DC}}$, and ω_t is the vector of frequencies of each bus at time t, $\omega_t = (\omega_t(i))_{i\in\mathcal{V}_{DC}}$. L_{DC} is the

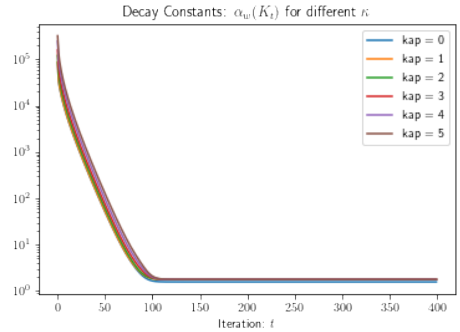

Fig. 1. Subexponential Decay Constants $\alpha_w(K_t)$ for the AC network.

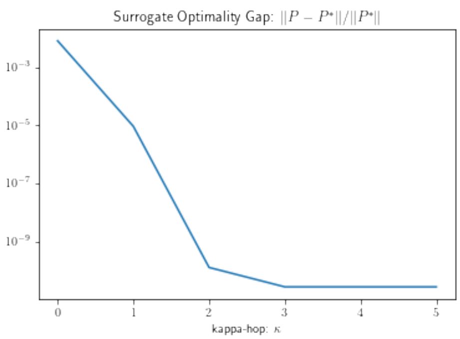

Fig. 2. Performance of truncated NPG on the AC network.

weighted laplacian of $\mathcal{G}_{DC}$. For $Q = (0.5^2)I$ and $R = I$, the discrete system above is stabilizable and detectable.

Setup. We use the graph and susceptance data from the IEEE 118-bus dataset described in [27]. We set $t_{s,DC} = .005$ms, $(k_{DC})_{ij} = B_{ij}V_{ref}^2/M$ where B_{ij} is the line susceptance, $M = 10^{-5}\text{kgm}^2$, $V_{ref} = 132$kV, and simulate the model via its discretization. We run (16) for $T = 400$ time steps with stepsize $\eta = 0.1$ on the AC network with K_0 generated via pole placement. The poles of $A - BK_0$ were sampled from the standard normal distribution and were ensured to be stabilizing.

Results. We calculate the decay constants $\alpha_w(K_t)$ under the weight in (11) with constant $D = \log 2$ and the sub-exponential power $\delta = 0.75$. As noted previously, our simulation results in Fig. 1 suggest that the decay constant bound of $\alpha_w(K_{t+1}) \sim e^{\text{poly}(\log \alpha(K_t))}$ (Theorem 8) is not tight and support our conjecture that the decay constant should be uniformly bounded throughout the Truncated NPG iterations. The overall convergence of the truncated NPG is shown in Fig. 2, where we plot $\|P - P^*\|/\|P^*\|$ against parameter κ, where $P = P_{K_T}$ is the solution to (5) for K_T (the controller at the final time step T), and P^* is the corresponding LQR solution. This quantity $\|P - P^*\|/\|P^*\|$ can be viewed as a surrogate of the optimality gap $C(K_T) - C(K^*)$ [14]. As κ increases, we see a decrease in the surrogate optimality gap, which is consistent with the single-step error bound in Theorem 9.

VI. Conclusion

In this work, we have shown that an NPG update from a spatially decaying controller preserves its spatial decay property. Furthermore, we have also shown that the truncation of the updated controller to a κ-hop distributed controller exhibits an error that decays with κ. As for future work, we are interested in analytically showing boundedness on the spatial decay of NPG iterations beyond one step and ultimately, near-optimal convergence guarantees for the truncated NPG algorithm and its model-free and distributed implementation.

References

[1] C. Zhao, U. Topcu, N. Li, and S. Low, "Design and stability of load-side primary frequency control in power systems," *IEEE Transactions on Automatic Control*, vol. 59, no. 5, pp. 1177–1189, 2014.
[2] S. E. Li, Y. Zheng, K. Li, Y. Wu, J. K. Hedrick, F. Gao, and H. Zhang, "Dynamical modeling and distributed control of connected and automated vehicles: Challenges and opportunities," *IEEE Intelligent Transportation Systems Magazine*, vol. 9, no. 3, pp. 46–58, 2017.
[3] P. Varaiya, "Max pressure control of a network of signalized intersections," *Transportation Research Part C: Emerging Technologies*, vol. 36, pp. 177–195, 2013.
[4] B. Sinopoli, C. Sharp, L. Schenato, S. Schaffert, and S. S. Sastry, "Distributed control applications within sensor networks," *Proceedings of the IEEE*, vol. 91, no. 8, pp. 1235–1246, 2003.
[5] V. D. Blondel and J. N. Tsitsiklis, "A survey of computational complexity results in systems and control," *Automatica*, vol. 36, no. 9, pp. 1249–1274, 2000.
[6] Y.-C. Ho *et al.*, "Team decision theory and information structures in optimal control problems–Part I," *IEEE Transactions on Automatic control*, vol. 17, no. 1, pp. 15–22, 1972.
[7] K. Mårtensson and A. Rantzer, "Gradient methods for iterative distributed control synthesis," in *Proceedings of the 48h IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference*, pp. 549–554, IEEE, 2009.
[8] B. Bamieh, F. Paganini, and M. A. Dahleh, "Distributed control of spatially invariant systems," *IEEE Transactions on automatic control*, vol. 47, no. 7, pp. 1091–1107, 2002.
[9] M. Rotkowitz and S. Lall, "A characterization of convex problems in decentralized control," *IEEE transactions on Automatic Control*, vol. 50, no. 12, pp. 1984–1996, 2005.
[10] L. Furieri, Y. Zheng, A. Papachristodoulou, and M. Kamgarpour, "Sparsity invariance for convex design of distributed controllers," *IEEE Transactions on Control of Network Systems*, vol. 7, no. 4, pp. 1836–1847, 2020.
[11] Y.-S. Wang, N. Matni, and J. C. Doyle, "A system-level approach to controller synthesis," *IEEE Transactions on Automatic Control*, vol. 64, no. 10, pp. 4079–4093, 2019.
[12] J. Anderson, J. C. Doyle, S. H. Low, and N. Matni, "System level synthesis," *Annual Reviews in Control*, vol. 47, pp. 364–393, 2019.
[13] N. Motee and A. Jadbabaie, "Optimal control of spatially distributed systems," *IEEE Transactions on Automatic Control*, vol. 53, no. 7, pp. 1616–1629, 2008.
[14] S. Shin, Y. Lin, G. Qu, A. Wierman, and M. Anitescu, "Near-optimal distributed linear-quadratic regulator for networked systems," *SIAM Journal on Control and Optimization*, vol. 61, no. 3, pp. 1113–1135, 2023.
[15] Y. Zhang, G. Qu, P. Xu, Y. Lin, Z. Chen, and A. Wierman, "Global convergence of localized policy iteration in networked multi-agent reinforcement learning," *Proc. ACM Meas. Anal. Comput. Syst.*, vol. 7, mar 2023.
[16] R. C. Zhang, W. Li, and N. Li, "On the optimal control of network lqr with spatially-exponential decaying structure," in *2023 American Control Conference (ACC)*, pp. 1775–1780, 2023.
[17] M. Fazel, R. Ge, S. Kakade, and M. Mesbahi, "Global convergence of policy gradient methods for the linear quadratic regulator," in *International Conference on Machine Learning*, pp. 1467–1476, PMLR, 2018.

[18] D. Malik, A. Pananjady, K. Bhatia, K. Khamaru, P. L. Bartlett, and M. J. Wainwright, "Derivative-free methods for policy optimization: Guarantees for linear quadratic systems," *J. Mach. Learn. Res.*, vol. 21, jan 2020.
[19] J. Bu, A. Mesbahi, M. Fazel, and M. Mesbahi, "LQR through the lens of first order methods: Discrete-time case," *arXiv preprint arXiv:1907.08921*, 2019.
[20] H. Mohammadi, A. Zare, M. Soltanolkotabi, and M. R. Jovanović, "Convergence and sample complexity of gradient methods for the model-free linear–quadratic regulator problem," *IEEE Transactions on Automatic Control*, vol. 67, no. 5, pp. 2435–2450, 2022.
[21] Y. Li, Y. Tang, R. Zhang, and N. Li, "Distributed reinforcement learning for decentralized linear quadratic control: A derivative-free policy optimization approach," *IEEE Transactions on Automatic Control*, vol. 67, no. 12, pp. 6429–6444, 2022.
[22] Z. Yang, Y. Chen, M. Hong, and Z. Wang, *Provably Global Convergence of Actor-Critic: A Case for Linear Quadratic Regulator with Ergodic Cost*, p. 8353–8365. Red Hook, NY, USA: Curran Associates Inc., 2019.
[23] K. Zhang, B. Hu, and T. Basar, "Policy optimization for H_2 linear control with H_∞ robustness guarantee: Implicit regularization and global convergence," in *Learning for Dynamics and Control*, pp. 179–190, 2020.
[24] B. Hu, K. Zhang, N. Li, M. Mesbahi, M. Fazel, and T. Başar, "Toward a theoretical foundation of policy optimization for learning control policies," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 6, no. 1, pp. 123–158, 2023.
[25] N. Motee and Q. Sun, "Sparsity and spatial localization measures for spatially distributed systems," *SIAM Journal on Control and Optimization*, vol. 55, no. 1, pp. 200–235, 2017.
[26] Q. Sun, "Wiener's lemma for infinite matrices," *Transactions of the American Mathematical Society*, vol. 359, no. 7, pp. 3099–3123, 2007.
[27] I. Peña, C. B. Martinez-Anido, and B.-M. Hodge, "An extended ieee 118-bus test system with high renewable penetration," *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 281–289, 2018.

APPENDIX

A. Proof of Weight Lemmas

Proof of Lemma 7: (13) follows directly from the triangle inequality on norms. As for (14), take $AB = (ab(i,j))_{i,j\in\mathcal{V}}$

$$\begin{aligned}\sum_{j\in\mathcal{V}}\|ab(i,j)\|w(i,j) &= \sum_{j\in\mathcal{V}}\left\|\sum_{k\in\mathcal{V}} a(i,k)b(k,j)\right\| w(i,j)\\ &\le \sum_{k\in\mathcal{V}}\|a(i,k)\|w(i,k)\sum_{j\in\mathcal{V}}\|b(i,k)\|w(j,k),\end{aligned}$$

where in the second line we apply the triangle inequality, property of the metric in w, and swap the sums. The weight w must obey the triangle inequality as it is an increasing function of the metric $d_\mathcal{G}$ with $\delta \in (0,1)$. This proves (14) □

Proof of Lemma 11: **Proof of (21).** For the condition in (21) to hold, we will use the inequality

$$1 \le s^\delta + (2^\delta - 1)(1-s)^\delta, \text{ for } s \in [1/2, 1]. \quad (34)$$

Denote the RHS of (34) $f(s)$. The inequality then comes from the fact that $f(1/2) = f(1) = 1$ and f is a continuous concave function of s on $[1/2, 1]$.

We divide the proof of (21) into two cases. In the first case, we assume $d_\mathcal{G}(i,k) \ge d_\mathcal{G}(k,j)$. Then, using triangle inequality, we have,

$$\begin{aligned}&w(i,j)\\ &\le \exp\left(D(d_\mathcal{G}(i,k)+d_\mathcal{G}(k,j))^\delta\left(\frac{d_\mathcal{G}(i,k)+d_\mathcal{G}(k,j)}{d_\mathcal{G}(i,k)+d_\mathcal{G}(k,j)}\right)^\delta\right)\\ &\le \exp\left(D(d_\mathcal{G}(i,k)^\delta + (2^\delta-1)d_\mathcal{G}(k,j)^\delta)\right)\\ &= w(i,k)v(k,j),\end{aligned}$$

where the last inequality uses (34) and the assumption. By symmetry, when $d_\mathcal{G}(k,j) \ge d_\mathcal{G}(i,k)$, we must have that $w(i,j) \le v(i,k)w(k,j)$. As a result, we can upper-bound $w(i,j)$ by combining the two cases

$$w(i,j) \le w(i,k)v(k,j) + v(i,k)w(k,j).$$

Proof of (22). Immediately, we have that

$$vw^{-1}(i,j) = \exp\left(D(2^\delta - 2)d_\mathcal{G}(i,j)^\delta\right) \le 1.$$

Proof of (23). From the definition of $a(\tau)$, if we consider the first term, we have

$$\begin{aligned}\max_{i\in\mathcal{V}} &\sum_{j\in\mathcal{N}_\mathcal{G}(i,\tau)} v(i,j)\\ &\le \max_{i\in\mathcal{V}}\sum_{j\in\mathcal{N}_\mathcal{G}(i,\tau)}\exp(D(2^\delta-1)d_\mathcal{G}(i,j)^\delta)\\ &\le C_P\tau^d\exp\left(D(2^\delta-1)\tau^\delta\right)\end{aligned}$$

where we have used the assumption (1). Doing the same for the second term in $a(\tau)$, we have that

$$a(\tau) \le 2C_P\tau^d\exp\left(D(2^\delta-1)\tau^\delta\right) \le 2C_P\tau^d e^{D\tau^\delta},$$

Similarly, for $b(\tau)$, we have, from $vw^{-1}(i,j) = e^{D(2^\delta-2)d_\mathcal{G}(i,j)^\delta}$ and $2^\delta - 2 \le 0$, that $b(\tau) \le 2\exp(D(2^\delta - 2)\tau^\delta)$. Then the optimization problem on the LHS of (23) can be upperbounded by the following:

$$2C_P\inf_{\tau\ge1}\tau^d e^{D\tau^\delta} + e^{D(2^\delta-2)\tau^\delta}t,$$

where we have used $C_P \ge 1$. Then picking $\tau^\delta = \frac{1}{(3-2^\delta)D}\log t$, we get the upper bound

$$2C_P\left(\left(\frac{1}{D(3-2^\delta)}\log t\right)^{d/\delta}+1\right)t^{1/(3-2^\delta)} \le Ct^\theta,$$

where we have set $\theta = \frac{1+\frac{1}{3-2^\delta}}{2} \in (0,1)$ and also used the following simple bound:

$$\log t \le C_b t^{\delta(1-\theta)/d}$$

which holds for $C_b = \frac{d}{(1-\theta)\delta}$, and set $C = 2C_P\left(1+\left(\frac{C_b}{D(3-2^\delta)}\right)^{d/\delta}\right)$. Hence w in (11) satisfies (23).