PredSLS: A Unified Framework for Distributed Predictive Control

Yifei Wu, Jing Yu, Tongxin Li

***Abstract*— Distributed control of large-scale systems is challenging due to the need for scalable and localized communication and computation. In this work, we introduce a PREDictive System-Level Synthesis (PredSLS) framework that designs closed-loop controllers by integrating prediction information into an affine feedback structure. Rather than focusing on worst-case uncertainty, PredSLS leverages both current state feedback and future predictions to achieve effective control in distributed settings. A core feature of the PredSLS framework is its temporal decaying property, which ensures that the influence of prediction errors diminishes exponentially over time, effectively localizing the effect of disturbances and enabling the use of finite impulse response approximations. Moreover, the PredSLS framework can be naturally decomposed into spatial and temporal components for efficient and parallelizable computation across the network.**

I. Introduction

Modern large-scale and networked systems pose significant challenges for traditional control design. These applications have motivated the development of control synthesis at the system level, as known as system-level synthesis (SLS) as a new framework that shifts the design focus from crafting an individual controller to designing the entire closed loop system response. This framework has been shown to enable the systematic incorporation of structural constraints such as locality and sparsity while delivering scalable and robust performance in distributed control architectures [1], [2], [3].

Despite its proven potential, the standard SLS framework typically assumes static or worst-case uncertainty models, thus overlooking the value of predictive information that is available in many modern applications. In contexts of networked control systems, short term forecasts generated by artificial intelligence tools or human inputs can be exploited to anticipate disturbances and adapt control actions in advance. The benefits of prediction in control have long been recognized in the literature on model predictive control [4], [5], and regret-optimal control [6], [7]. However, a unified SLS framework that fully integrates future disturbance predictions remains absent.

Incorporating prediction into distributed control scenarios, however, introduces new challenges. First, constructing system-level responses compatible with predictive control is nontrivial. The classic predictive synthesis problem typically considers only causal closed-loop mappings, which are not directly generalizable to incorporate prediction. Thus, the problem of designing controllers that not only meet general structural communication constraints such as locality and sparsity but also seamlessly integrate predictions in the system-level remains unexplored. Second, in distributed control systems, a prevalent approach involves designing an optimal controller and then truncating its gains to fit within communication constraints, as explored in recent studies [8], [9], [10]. While this truncation method simplifies implementation by adapting precomputed gains to limited communication topologies, e.g., κ-hop networks and sparse connections, it often sacrifices optimality and the flexibility to incorporate general communication constraints such as communication delays. Achieving optimal control requires a unified approach that simultaneously optimizes the controller and accounts for communication constraints, rather than relying on post hoc truncation. Consequently, the control parameters themselves need to be optimized in light of the communication constraints, and a unified distributed control framework for predictive control thus becomes essential.

To address these challenges, we introduce a PREDictive SLS (termed PredSLS) framework that integrates predictive information into the synthesis of system-level responses for an affine controller. This framework is designed so that it synthesizes a class of non-causal affine predictive controllers that contain the optimal offline controller [7]. Our main results are two-fold:

1) **The PredSLS framework** We propose the PREDictive System-Level Synthesis (PredSLS) framework, n this section, we introduce the framework of predictive system-level synthesis of the form $\mathbf{u} = \mathbf{K}\mathbf{x} + \widehat{\mathbf{L}}\widehat{\mathbf{w}}$ (see Equation (3) for more details), where $\mathbf{u}$, $\mathbf{x}$, and $\widehat{\mathbf{w}}$ represent action, state, and disturbance predictions respectively; $\mathbf{K}$ and $\widehat{\mathbf{L}}$ are proper and improper transfer matrices (Theorem III.1). This enables optimal controller synthesis via a convex quadratic program over system responses, matching the offline optimal policy for arbitrary disturbances when predictions are error-free (Proposition 1).
2) **Distributed predictive control** We incorporate general structural communication constraints into PredSLS, and propose a scalable decomposition of the centralized control problem for large-scale networks. When the communication constraints are column-decomposable, spatial and temporal separability is used to split the distributed PredSLS into N agent-wise and T temporally independent sub-problems. To ensure computational tractability, we introduce a finite-horizon approximation in Section III-B using finite impulse response (FIR) mappings. This is made possible by leveraging the expo-

Wu and Li are with School of Data Science, The Chinese University of Hong Kong, Shenzhen, Guangdong, 518172, China (e-mails: `yifeiwu1@link.cuhk.edu.cn`, `litongxin@cuhk.edu.cn`). Yu is with the Department of Electrical and Computer Engineering, University of Washington, Seattle, WA, USA.(e-mail: `jing5@uw.edu`).

nentially decaying property established in Theorem IV.1. For κ-hop networks, a spatial decaying property is guaranteed by Theorem IV.2. Unlike the results in [8], which require the system to be open-loop stable, and approaches that do not integrate predictions [9], PREDSLS leverages system-level parameterization to directly incorporate any communication constraints that can be expressed as convex constraints into the controller synthesis.

Related Work. Our work builds upon and extends several recent advancements in distributed control for networked linear systems. Distributed model predictive control is proposed in [11], whose stability is satisfied with time-varying local terminal cost functions and sets. Zhang et al. [8] investigate optimal control for networked linear quadratic regulators (LQR) with spatially-exponentially decaying (SED) structures. They propose a distributed disturbance-action control (DAC) controller, which relies on conditions such as open-loop stability and SED system matrices (A, B, Q, R). In contrast, our framework applies to a broader class of stabilizable systems (A, B, Q, R) without requiring SED properties. Xu and Qu [10] explore distributed truncated predictive control within a model predictive control (MPC) framework for linear systems under limited communication. Their work emphasizes stability and performance for systems with specific communication structures. Shin et al. [9] propose a near-optimal truncated distributed LQR for disturbance-free networked systems. Our method differs by introducing a more general functional closed-loop control law with disturbance predictions, which can be applied to a wider range of networked linear systems beyond those constrained by specific communication patterns.

The System Level Synthesis (SLS) framework [12], [13], [14], [15], [16], provides an effective approach for designing distributed controllers that satisfy locality constraints in networked systems. Recently, SLS has been applied to broader contexts, including the design of predictive safety filters using linear causal controllers, as demonstrated by Leeman et al. [17]. However, the integration of future disturbance predictions into the SLS framework remains largely unaddressed. In contrast, our proposed Predictive SLS (PREDSLS) framework extends SLS by incorporating predictions of future disturbances, offering a more general approach to optimal control in networked linear systems, while still preserving the ability to handle general communication constraints such as locality and delay.

Notational Conventions. Throughout this paper, $\|\cdot\|$ denotes the ℓ_2-norm for vectors and ℓ_2-induced norm for matrices, $\|\cdot\|_F$ denotes the the Frobenius norm. Scalars, vectors and matrices are denoted by lowercase italic font like "x", lowercase normal font like "x", and uppercase font like "M", respectively. Let $[T] := \{0, 1, \cdots, T-1\}$ denote the discrete control time indices and $[N] := \{1, \cdots, N\}$ denote a set of N subsystems. We write a sequence of vectors as $\mathrm{w}_{0:T} := (\mathrm{w}_0, \ldots, \mathrm{w}_T)$. We denote the (i, j)-th entry of a matrix M as M^{ij} with superscripts and use $M(j)$ for the j-th column of M. Similarly, we use single and double subscripts to represent time-dependent sub-matrices or sub-vectors arranged in order, such as M_t and $M_{t,k}$. Let boldface font $\mathbf{x}$ denote a collection of states $\mathbf{x} = \{\mathrm{x}_t\}_{t=0}^{\infty}$ with $\mathrm{x}_t \in \mathbb{R}^n$, $\mathbf{M}$ and $\widehat{\mathbf{L}}$ denote proper and improper transfer matrices $\mathbf{M}(z) := \sum_{t=0}^{\infty} z^{-t} M_t$ and $\widehat{\mathbf{L}}(z) := \sum_{t=-\infty}^{\infty} z^{-t} L_t$ with kernels $M_t, L_t \in \mathbb{R}^{m \times n}$.

II. Preliminaries and Problem Setup

We study a networked system with N subsystems over a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ where the nodes $\mathcal{V} := [N]$ are the subsystems and $\mathcal{E} \subseteq [N] \times [N]$. We denote $\mathcal{N}_{\mathcal{G}}(i) := \{i\} \cup \{j \in [N] | (i, j) \in \mathcal{E}\}$ as the *neighboring set* of subsystem i and $d_{\mathcal{G}}(i, j) : [N] \times [N] \to \mathbb{N}_+$ as the shortest distance from subsystem i to j. The dynamics for each subsystem $i \in [N]$ is governed by

$$\mathrm{x}_{t+1}^i = \sum_{j \in \mathcal{N}_{\mathcal{G}}(i)} \left(A^{ij} \mathrm{x}_t^j + B^{ij} \mathrm{u}_t^j \right) + \mathrm{w}_t^i \, , \, t \in [T], \tag{1}$$

where $\mathrm{x}_t^i \in \mathbb{R}^{n_i}, \mathrm{w}_t^i \in \mathbb{R}^{n_i}$, and $\mathrm{u}_t^i \in \mathbb{R}^{m_i}$ are the local state, disturbance, and action for the i-th subsystem, respectively. Let $n = \sum_{i \in [N]} n_i$, $m = \sum_{i \in [N]} m_i$. We define $\mathrm{x}_t \in \mathbb{R}^n$, $\mathrm{u}_t \in \mathbb{R}^m$, and $\mathrm{w}_t \in \mathbb{R}^n$ to denote the joint vectors of the N agents with

$$\mathrm{x}_t = \left[\left(\mathrm{x}_t^1\right)^\top ; \left(\mathrm{x}_t^2\right)^\top ; \cdots ; \left(\mathrm{x}_t^N\right)^\top \right]^\top .$$

The joint vectors u_t and w_t are similarly defined. Let $A := [A^{ij}] \in \mathbb{R}^{n \times n}$ and $B := [B^{ij}] \in \mathbb{R}^{n \times m}$ be the concatenated global system dynamic matrices with $A^{ij} \equiv 0$, $B^{ij} \equiv 0$ for all $j \notin \mathcal{N}_{\mathcal{G}}(i)$. Dynamics (1) can be equivalently represented as

$$\mathrm{x}_{t+1} = A\mathrm{x}_t + B\mathrm{u}_t + \mathrm{w}_t. \tag{2}$$

Without loss of generality, we assume (2) is initialized with $\mathrm{x}_0 = \mathbf{0}$. Furthermore, we assume the system (A, B) is stabilizable so that there exists $K \in \mathbb{R}^{m \times n}$ such that $A - BK$ has a spectral radius ρ less than 1 [18]. Thus, the Gelfand's formula implies that there exist $L > 0$, $\rho \in (0, 1)$ such that $\|A - BK\|^t \leq L\rho^t$ for all $t \geq 0$.

Problem Setup: We aim to design a distributed predictive controller to address the finite horizon *centralized linear quadratic regulator* (CLQR) problem, formulated as:

$$J^\star := \min_{\mathrm{x}_{0:T}, \mathrm{u}_{0:T-1}} \sum_{t=0}^{T-1} \left(\mathrm{x}_t^\top Q \mathrm{x}_t \; + \mathrm{u}_t^\top R \mathrm{u}_t \; \right) + \mathrm{x}_T^\top Q_T \mathrm{x}_T,$$
$$\text{subject to (2).} \tag{CLQR}$$

Here, $Q, Q_T \in \mathbb{R}^{n \times n}$ and $R \in \mathbb{R}^{m \times m}$ are positive definite matrices, with Q_T representing the terminal state cost.

In this work, we consider a distributed predictive controller that operates with predictions on the disturbances. In particular, each subsystem i has access to its own local predictions of future disturbances $\widehat{\mathrm{w}}_{0:T-1}^i$ and shares them with its neighbors subject to general communication constraints, which are specified in Section IV after we introduce the system-level parameterization. Notably, these predictions are neither global nor guaranteed to be accurate, as they may not span the entire graph $\mathcal{G}$.

III. Predictive System-Level Synthesis

In this section, we introduce the framework of predictive system-level synthesis (PredSLS). PredSLS parameterizes all achievable closed-loop system responses under affine feedback predictive controllers of the form

$$\mathrm{u}_t = \sum_{\tau \le t} K_{t,\tau} \mathrm{x}_t + \sum_{\tau=0}^{\infty} L_{t,\tau} \widehat{\mathrm{w}}_{t+\tau}, \tag{3}$$

where $K_{t,\tau} \in \mathbb{R}^{m \times n}$ and $L_{t,\tau} \in \mathbb{R}^{n \times n}$ for $t, \tau \ge 0$ parameterizes the affine predictive controller (3). Furthermore, PredSLS provides a dynamic predictive controller that realizes the prescribed closed-loop system responses (Theorem III.1). We show that PredSLS naturally lends itself to distributed and localized synthesis and implementation thanks to the spatio-temporally exponential decay properties of the parameterization, as detailed in Theorem IV.1.

A. System-Level Parameterization of Predictive Controllers

Consider the closed loop of Equation (2) under the (dynamic) state-feedback non-causal controller (3), which can be compactly written as $\mathbf{u} = \mathbf{K}\mathbf{x} + \widehat{\mathbf{L}}\widehat{\mathbf{w}}$ where $\mathbf{K}$ and $\widehat{\mathbf{L}}$ are proper and improper transfer matrices.[1] Denote the closed-loop mappings from exogenous inputs $\mathbf{w}$ and $\widehat{\mathbf{w}}$ to the state and control actions as follows

$$\begin{bmatrix} \mathbf{x} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} \mathbf{\Phi}^{\mathrm{x}} & \widehat{\mathbf{\Phi}}^{\mathrm{x}} \\ \mathbf{\Phi}^{\mathrm{u}} & \widehat{\mathbf{\Phi}}^{\mathrm{u}} \end{bmatrix} \begin{bmatrix} \mathbf{w} \\ \widehat{\mathbf{w}} \end{bmatrix}, \tag{4}$$

where the mappings $\mathbf{\Phi}^{\mathrm{x}} : \mathbf{w} \to \mathbf{x}$, $\widehat{\mathbf{\Phi}}^{\mathrm{x}} : \widehat{\mathbf{w}} \to \mathbf{x}$, $\mathbf{\Phi}^{\mathrm{u}} : \mathbf{w} \to \mathbf{u}$, and $\widehat{\mathbf{\Phi}}^{\mathrm{u}} : \widehat{\mathbf{w}} \to \mathbf{u}$ are computed via simple algebra such that

$$\begin{bmatrix} \mathbf{\Phi}^{\mathrm{x}} & \widehat{\mathbf{\Phi}}^{\mathrm{x}} \\ \mathbf{\Phi}^{\mathrm{u}} & \widehat{\mathbf{\Phi}}^{\mathrm{u}} \end{bmatrix} = \begin{bmatrix} U & UB\widehat{\mathbf{L}} \\ \mathbf{K}U & (\mathbf{K}UB + I)\widehat{\mathbf{L}} \end{bmatrix} \tag{5}$$

with $U := (zI - (A + B\mathbf{K}))^{-1}$. We refer to $\mathbf{\Phi}$ as a group of these four *closed-loop mappings* (transfer functions), and define as $\mathbf{\Phi} := [\mathbf{\Phi}^{\mathrm{x}}, \widehat{\mathbf{\Phi}}^{\mathrm{x}}; \mathbf{\Phi}^{\mathrm{u}}, \widehat{\mathbf{\Phi}}^{\mathrm{u}}]$. The following theorem parameterizes all achievable closed-loop mappings by controllers of the form (3).

Theorem III.1 (Predictive System-Level Parameterization)**.** *The affine subspace defined by*

$$\begin{bmatrix} zI - A & -B \end{bmatrix} \begin{bmatrix} \mathbf{\Phi}^{\mathrm{x}} & \widehat{\mathbf{\Phi}}^{\mathrm{x}} \\ \mathbf{\Phi}^{\mathrm{u}} & \widehat{\mathbf{\Phi}}^{\mathrm{u}} \end{bmatrix} = \begin{bmatrix} I & 0 \end{bmatrix} \tag{6}$$

parameterizes all possible system responses of (2) *under predictive controllers of the form* (3)*. Furthermore, given any* $\mathbf{\Phi}$ *that satisfies* (6)*, the predictive controller* (3) *with*

$$\mathbf{K} = \mathbf{\Phi}^{\mathrm{u}}(\mathbf{\Phi}^{\mathrm{x}})^{-1}, \quad \widehat{\mathbf{L}} = \widehat{\mathbf{\Phi}}^{\mathrm{u}} - \mathbf{\Phi}^{\mathrm{u}}(\mathbf{\Phi}^{\mathrm{x}})^{-1}\widehat{\mathbf{\Phi}}^{\mathrm{x}} \tag{7}$$

achieves the closed-loop response prescribed by $\mathbf{\Phi}$ *in* (5)*.*

The proof of Theorem III.1 can be found in Appendix A. If we impose the constraints $\mathbf{\Phi}^{\mathrm{x}}, \mathbf{\Phi}^{\mathrm{u}} \in \frac{1}{z}\mathcal{RH}_\infty$ and $\widehat{\mathbf{\Phi}}^{\mathrm{x}}, \widehat{\mathbf{\Phi}}^{\mathrm{u}} \in \frac{1}{z}\mathcal{RH}_\infty^i$ in addition to (6), where $\mathcal{RH}_\infty$ denotes the space of improper and real rational stable transfer matrices and $\mathcal{RH}_\infty^i$ denotes the space of all improper and real rational stable transfer matrices, then the corresponding controller (7) is internally stabilizing.

[1] Note that the optimal non-causal (offline) controller for Equation (PSLS) follows this form, as discussed in [6], [7].

The parameterization in Theorem III.1 enables the formulation of a system-level predictive controller optimization with quadratic costs, which we term PSLS:

$$\min_{(\mathbf{\Phi}^{\mathrm{x}}, \widehat{\mathbf{\Phi}}^{\mathrm{x}}, \mathbf{\Phi}^{\mathrm{u}}, \widehat{\mathbf{\Phi}}^{\mathrm{u}})} \left\| \begin{bmatrix} Q^{\frac{1}{2}} & \mathbf{0} \\ \mathbf{0} & R^{\frac{1}{2}} \end{bmatrix} \begin{bmatrix} \mathbf{\Phi}^{\mathrm{x}} + \widehat{\mathbf{\Phi}}^{\mathrm{x}} \\ \mathbf{\Phi}^{\mathrm{u}} + \widehat{\mathbf{\Phi}}^{\mathrm{u}} \end{bmatrix} \right\|_{\mathcal{H}_2}^2 \quad \text{subject to (6)}, \tag{PSLS}$$

where for $\mathbf{G} = \sum_{t=0}^{\infty} z^{-t} G_t$, $\|\mathbf{G}\|_{\mathcal{H}_2} := \left(\sum_{t=0}^{\infty} \|G_t\|_F^2\right)^{1/2}$ is the $\mathcal{H}_2$-norm of a transfer function $\mathbf{G}$.

It turns out that the optimal closed-loop mappings obtained by solving (PSLS) are also optimal for the optimal control problem (CLQR) in Section II, as established by the theorem below, whose proof can be found in Appendix B.

Proposition 1 (Optimal PredSLS Controller)**.** *Consider an optimal solution* $\mathbf{\Phi}^\star = (\mathbf{\Phi}^{\mathrm{x},\star}, \widehat{\mathbf{\Phi}}^{\mathrm{x},\star}, \mathbf{\Phi}^{\mathrm{u},\star}, \widehat{\mathbf{\Phi}}^{\mathrm{u},\star})$ *of* (PSLS)*. Then for any disturbance sequence* $\mathrm{w}_{0:T-1}$*, the predictive controller* (3) *instantiated with*

$$\mathbf{K} = \mathbf{\Phi}^{\mathrm{u},\star}(\mathbf{\Phi}^{\mathrm{x},\star})^{-1} \textit{ and } \widehat{\mathbf{L}} = \widehat{\mathbf{\Phi}}^{\mathrm{u},\star} - \mathbf{\Phi}^{\mathrm{u},\star}(\mathbf{\Phi}^{\mathrm{x},\star})^{-1}\widehat{\mathbf{\Phi}}^{\mathrm{x},\star}$$

and implemented using exact predictions such that $\widehat{\mathrm{w}}_t = \mathrm{w}_t$ *for all* $t \in [T]$ *achieves the optimal cost of* (CLQR) *with a terminal cost* $Q_T = P$ *where* P *is the solution to the algebraic Riccati equation* $P = A^\top PA - (A^\top PB)(R + B^\top PB)^{-1}(B^\top PA) + Q$*.*

B. Finite Horizon Approximation and Implementation

The framework introduced in the previous subsection is formulated for closed-loop mappings that are transfer matrices. Consequently, (PSLS) contains infinitely many optimization variables, making it impractical for direct computation using standard off-the-shelf convex optimization tools. To address this limitation, we introduce a finite horizon approximation of PredSLS by restricting the closed-loop mappings to have finite impulse response (FIR) of horizon H, e.g., $\mathbf{\Phi} \in \mathcal{F}_H := \left\{\mathbf{\Phi} : \mathbf{\Phi} = \sum_{t=0}^{H} z^{-t}\Phi_t\right\}$. In particular, the FIR closed-loop mappings are of the form:

$$\mathbf{\Phi}^{\mathrm{x}}_{\mathcal{F}_H} := \begin{bmatrix} \Phi^{\mathrm{x}}_{0,0} & \mathbf{0} & \cdots & \cdots & \cdots & \cdots & \mathbf{0} \\ \vdots & \ddots & \ddots & & & & \vdots \\ \Phi^{\mathrm{x}}_{H,0} & \cdots & \Phi^{\mathrm{x}}_{H,H} & \ddots & & & \vdots \\ \mathbf{0} & \ddots & & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & & \Phi^{\mathrm{x}}_{T-H,T-H} & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \vdots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \cdots & \mathbf{0} & \Phi^{\mathrm{x}}_{T,T-H} & \cdots & \Phi^{\mathrm{x}}_{T,T} \end{bmatrix},$$

and similarly,

$$\widehat{\mathbf{\Phi}}^{\mathrm{x}}_{\mathcal{F}_H} := \begin{bmatrix} \widehat{\Phi}^{\mathrm{x}}_{0,0} \cdots \widehat{\Phi}^{\mathrm{x}}_{0,H} & \mathbf{0} & \cdots & \cdots & \mathbf{0} \\ \vdots \quad \ddots \quad \vdots & \ddots & \ddots & & \vdots \\ \widehat{\Phi}^{\mathrm{x}}_{H,0} \cdots \widehat{\Phi}^{\mathrm{x}}_{H,H} & \cdots & \widehat{\Phi}^{\mathrm{x}}_{H,2H} & \ddots & \vdots \\ \mathbf{0} \quad \ddots & \ddots & & \ddots & \mathbf{0} \\ \vdots \quad \ddots \quad \ddots & & \widehat{\Phi}^{\mathrm{x}}_{T-H,T-H} & \cdots & \widehat{\Phi}^{\mathrm{x}}_{T-H,T} \\ \vdots \quad\quad \ddots & \ddots & \vdots & \ddots & \vdots \\ \mathbf{0} \cdots \cdots & \mathbf{0} & \widehat{\Phi}^{\mathrm{x}}_{T,T-H} & \cdots & \widehat{\Phi}^{\mathrm{x}}_{T,T} \end{bmatrix}, \tag{8}$$

with component matrices $\Phi^{\mathrm{x}}_{t,k}, \widehat{\Phi}^{\mathrm{x}}_{t,k} \in \mathbb{R}^{n\times n}$. The component matrices of $\mathbf{\Phi}^{\mathrm{u}}_{\mathcal{F}_H}$ and $\widehat{\mathbf{\Phi}}^{\mathrm{u}}_{\mathcal{F}_H}$ are similarly defined. Restricting the closed-loop mappings to have FIR, the synthesis problem (PSLS) can be equivalently expressed in terms of the component matrices as

$$\min_{\mathbf{\Phi}_{\mathcal{F}_T}} \sum_{t=0}^{T} \sum_{k=\underline{t}}^{\bar{t}} \left\| \begin{bmatrix} Q^{\frac{1}{2}} & \mathbf{0} \\ \mathbf{0} & R^{\frac{1}{2}} \end{bmatrix} \begin{bmatrix} \Phi^{\mathrm{x}}_{t,k} + \widehat{\Phi}^{\mathrm{x}}_{t,k} \\ \Phi^{\mathrm{u}}_{t,k} + \widehat{\Phi}^{\mathrm{u}}_{t,k} \end{bmatrix} \right\|_F^2$$
$$\text{subject to (10a)-(10d), for all } k \in [T], \tag{9}$$

where $\underline{t} := \max(0, t-H)$, and $\bar{t} := \min(T, t+H)$ for all $t \in [T]$. The constraints (6) are also transformed into equivalent time-domain constraints on the component matrices:

$$\Phi^{\mathrm{x}}_{t,k} = I, \text{for } t = k, \tag{10a}$$

$$\Phi^{\mathrm{x}}_{t+1,k} = \begin{cases} A\Phi^{\mathrm{x}}_{t,k} + B\Phi^{\mathrm{u}}_{t,k}, & \text{if } t \in \{k, \dots, \bar{k}\}, \\ \mathbf{0}, & \text{otherwise}, \end{cases} \tag{10b}$$

$$\widehat{\Phi}^{\mathrm{x}}_{t,k} = \mathbf{0}, \text{for } t = 0, \tag{10c}$$

$$\widehat{\Phi}^{\mathrm{x}}_{t+1,k} = \begin{cases} A\widehat{\Phi}^{\mathrm{x}}_{t,k} + B\widehat{\Phi}^{\mathrm{u}}_{t,k}, & \text{if } t \in \{\underline{k}, \dots, \bar{k}\}, \\ \mathbf{0}, & \text{otherwise}, \end{cases} \tag{10d}$$

where $\bar{k} := \min(T-1, k+H-1)$, $\underline{k} := \max(0, k-H+1)$ for all $k \in [T]$. Correspondingly, the time-domain realization of (7) is as follows

$$\widetilde{\mathrm{w}}_t := \mathrm{x}_t - \sum_{k=\underline{t}}^{t-1} \Phi^{\mathrm{x}}_{t,k} \widetilde{\mathrm{w}}_k, \quad \widehat{\mathrm{x}}_t := \sum_{k=0}^{T} \widehat{\Phi}^{\mathrm{x}}_{t,k} \widehat{\mathrm{w}}_{k-1}, \tag{11a}$$

$$\bar{\mathrm{w}}_t := \widehat{\mathrm{x}}_t - \sum_{k=\underline{t}}^{t-1} \Phi^{\mathrm{x}}_{t,k} \bar{\mathrm{w}}_k, \tag{11b}$$

$$\mathrm{u}_t = \sum_{k=\underline{t}}^{t} \Phi^{\mathrm{u}}_{t,k} \left(\widetilde{\mathrm{w}}_k - \bar{\mathrm{w}}_k\right) + \sum_{k=\underline{t}}^{\bar{t}} \widehat{\Phi}^{\mathrm{u}}_{t,k} \widehat{\mathrm{w}}_{k-1}, \tag{11c}$$

where $\bar{\mathrm{w}}$ and $\widetilde{\mathrm{w}}$ are controller internal states and can be interpreted as internal disturbance and prediction estimations. With simple algebra, it is straightforward to verify that $\widetilde{\mathrm{w}}_t - \bar{\mathrm{w}}_t = \mathrm{w}_{t-1}$ and $\widetilde{\mathrm{w}}_0 := \mathrm{x}_0, \bar{\mathrm{w}}_0 := 0$.

Remark 1 (Approximation Error)**.** *Due to the additional FIR constraint on the closed-loop mappings, the solution to (9) is sub-optimal. As a first step towards quantifying the sub-optimality, in Theorem IV.1 we show that the closed-loop mappings present exponential decay structure over the FIR horizon, e.g., $\|\widehat{\Phi}^{\mathrm{x}}_{t,k}\|$ decays exponentially with respect to $|t-k|$.*

IV. Distributed and Localized Synthesis

For large-scale networks, solving (PSLS) using a centralized controller imposes significant communication and computation burdens. In this section, we describe a decomposition of the centralized problem (PSLS) via agent-wise parallel synthesis, thus enabling scalable computation. In what follows, we let the FIR horizon $H = T$ where T is the problem horizon defined in (CLQR) for simplicity.[2]

Specifically, PredSLS described above can be viewed as an instantiation of the general optimal control problem:

$$\min_{\mathbf{\Phi}} C\left(\mathbf{\Phi}^{\mathrm{x}}, \widehat{\mathbf{\Phi}}^{\mathrm{x}}, \mathbf{\Phi}^{\mathrm{u}}, \widehat{\mathbf{\Phi}}^{\mathrm{u}}\right),$$
$$\text{subject to (6)}, \ \ \mathbf{\Phi} \in \mathcal{L}_d \cap \mathcal{F}_T \cap \mathcal{X}. \tag{12}$$

Function $C(\cdot)$ denotes a convex system-level objective, subspace $\mathcal{L}_d$ represents any convex spatial-locality constraints, and $\mathcal{X}$ is a convex sparsity subspace that can be used to model communication delay. All these definitions follow the classic SLS literature (see [12], [14]).

Let $\mathbf{\Phi}(i) := [\mathbf{\Phi}^{\mathrm{x}}(i), \widehat{\mathbf{\Phi}}^{\mathrm{x}}(i); \mathbf{\Phi}^{\mathrm{u}}(i), \widehat{\mathbf{\Phi}}^{\mathrm{u}}(i)]$ for distributed control. If (12) is *column-wise separable* [13], then it can be partitioned into N parallel sub-problems in the form:

$$\min_{\mathbf{\Phi}(i)} C_i \left(\mathbf{\Phi}^{\mathrm{x}}(i), \widehat{\mathbf{\Phi}}^{\mathrm{x}}(i), \mathbf{\Phi}^{\mathrm{u}}(i), \widehat{\mathbf{\Phi}}^{\mathrm{u}}(i)\right),$$
$$\text{subject to } \begin{bmatrix} zI - A & -B \end{bmatrix} \mathbf{\Phi}(i) = \begin{bmatrix} I(i) & 0 \end{bmatrix},$$
$$\mathbf{\Phi}(i) \in \mathcal{L}_d(i) \cap \mathcal{F}_T \cap \mathcal{X}_i, \tag{13}$$

where C_i, $\mathcal{L}_d(i)$, and $\mathcal{X}_i$ are column-wise objective function and subspaces.

A. κ-distributed control

Based on the communication topology $\mathcal{G}$, a controller is κ-*distributed*[3], if for any subsystem i, the computation of u^i_t only depends on the information available at all subsystems $j \in \mathcal{N}^{\kappa}(i) := \{j \in [N] \,|\, d_{\mathcal{G}}(j,i) \le \kappa\}$ with some *communication distance* κ.

We now turn to illustrate a specific case of (13), called i-PSLS, which naturally leads to an optimal κ-distributed controller. We first introduce an equivalent representation of the communication constraints for κ-distributed control, in terms of the closed-loop mappings $\mathbf{\Phi}$. For simplicity,[4] consider a communication graph that shares the same topology as the dynamics $\mathcal{G} = ([N], \mathcal{E})$. The κ-adjacency matrix $\mathcal{C}^{\kappa} \in \{0,1\}^{N\times N}$ is a binary matrix such that $\mathcal{C}^{\kappa}(i,j) = 1$ for $j \in \mathcal{N}^{\kappa}(i)$ with any agent $i \in [N]$, otherwise $\mathcal{C}^{\kappa}(i,j) = 0$. $\mathrm{sp}(\cdot)$ is the support of a matrix.

[2] It is straightforward to verify that (9) is equivalent to the H-horizon linear problem (CLQR) with the termination condition with $Q_T = P$.

[3] Unlike the κ-distributed controller in [9], which is a linear state feedback controller in x_t, the affine controller in this work may depend not only on the state x_t, but also predictions of future disturbances $(w_\tau : \tau \ge t)$.

[4] For SLS, the communication graph in general does not have to match the system dynamics [19].

$$\min_{\varphi^{\mathrm{x},i},\widehat{\varphi}^{\mathrm{x},i},\varphi^{\mathrm{u},i},\widehat{\varphi}^{\mathrm{u},i}} \sum_{t=0}^{T}\sum_{k=0}^{T}\left\|\begin{bmatrix}Q^{\frac{1}{2}} & \mathbf{0}\\ \mathbf{0} & R^{\frac{1}{2}}\end{bmatrix}\begin{bmatrix}\varphi_{t,k}^{\mathrm{x},i}+\widehat{\varphi}_{t,k}^{\mathrm{x},i}\\ \varphi_{t,k}^{\mathrm{u},i}+\widehat{\varphi}_{t,k}^{\mathrm{u},i}\end{bmatrix}\right\|^2 \tag{i-PSLS}$$

subject to for all $k \in [T]$, $\varphi_{t,k}^{\mathrm{x},i} = \mathrm{e}_i$, for $t = k$,

$$\varphi_{t+1,k}^{\mathrm{x},i} = \begin{cases} A\varphi_{t,k}^{\mathrm{x},i} + B\varphi_{t,k}^{\mathrm{u},i}, & \text{if } t \in \{k,\cdots,T-1\},\\ \mathbf{0}, & \text{otherwise.}\end{cases} \tag{causal dynamics}$$

$$\widehat{\varphi}_{t,k}^{\mathrm{x},i} = \mathbf{0}, \text{for } t = 0,$$

$$\widehat{\varphi}_{t+1,k}^{\mathrm{x},i} = \begin{cases} A\widehat{\varphi}_{t,k}^{\mathrm{x},i} + B\widehat{\varphi}_{t,k}^{\mathrm{u},i}, & \text{if } t \in [T],\\ \mathbf{0}, & \text{otherwise,}\end{cases} \tag{non-causal dynamics}$$

$$\left(\varphi_{t,k}^{\mathrm{x},i}, \varphi_{t,k}^{\mathrm{u},i}\right), \left(\widehat{\varphi}_{t,k}^{\mathrm{x},i}, \widehat{\varphi}_{t,k}^{\mathrm{u},i}\right) \in \mathcal{L}_\kappa(i), \text{ for } t \in [T]. \tag{locality constraints}$$

Definition 1 (κ-Locality Constraint)**.** *The subspace $\mathcal{L}_\kappa$ is a κ-locality constraint for mappings $(\Phi^{\mathrm{x}}, \Phi^{\mathrm{u}})$ if and only if the support matrices of the mappings satisfy $\mathrm{sp}(\Phi^{\mathrm{x}}) = \mathcal{C}^\kappa$ and $\mathrm{sp}(\Phi^{\mathrm{u}}) = \mathcal{C}^{\kappa+1}$. In addition, $\mathcal{L}^\kappa$ is column-wise separable with respect to $[N]$. Moreover, for any subsystem $i \in [N]$, $\mathcal{L}^\kappa(i)$ denotes a column-wise constraint for $(\varphi^{\mathrm{x}}, \varphi^{\mathrm{u}})$ with $\mathrm{sp}(\varphi^{\mathrm{x}}) = \mathcal{C}^\kappa(i)$ and $\mathrm{sp}(\varphi^{\mathrm{u}}) = \mathcal{C}^{\kappa+1}(i)$.*

Imposing localization on the closed-loop mappings enforces that in the closed loop, disturbances are prevented from propagating outside of the localized region. Thanks to Theorem III.1, a key feature of the system-level approach of PREDSLS is that any localization constraint on the closed-loop mappings directly translates to the controller realization, e.g., (11), enabling localized controller implementation.

Recall the general problem (12), and it is straightforward to see that (PSLS) is a column-wise separable problem [13], which is easily extended to the parallel alternatives.

B. Spatial and Temporal Decompositions

This section considers the finite PREDSLS problem (9), separating it into parallel column-wise sub-problems through two (spatial decomposition in Section IV-B.1 and temporal decomposition in Section IV-B.2) forms.

1) Spatial Decomposition: We decompose (PSLS) into N column sub-problems where each column sub-problem performs the following optimization on the i-th column of the matrices $(\Phi_{t,k}^{\mathrm{x}}, \widehat{\Phi}_{t,k}^{\mathrm{x}}, \Phi_{t,k}^{\mathrm{u}}, \widehat{\Phi}_{t,k}^{\mathrm{u}})_{t,k=1}^T$. Similarly, we denote the corresponding column variables associated with the i-th sub-problem as $(\varphi_{t,k}^{\mathrm{x},i}, \widehat{\varphi}_{t,k}^{\mathrm{x},i}, \varphi_{t,k}^{\mathrm{u},i}, \widehat{\varphi}_{t,k}^{\mathrm{u},i})_{t,k=1}^T$. The i-th column sub-problem is then given by (i-PSLS). Note that the constraints in (i-PSLS) are the column-wise components of the constraints in (PSLS).

2) Temporal Decomposition: We further decompose the (i-PSLS) problem temporally along the FIR horizon for each $k \in [T]$ as follows:

$$\min \sum_{t=0}^{T}\left\|\begin{bmatrix}Q^{\frac{1}{2}} & \mathbf{0}\\ \mathbf{0} & R^{\frac{1}{2}}\end{bmatrix}\begin{bmatrix}\varphi_{t,k}^{\mathrm{x},i}+\widehat{\varphi}_{t,k}^{\mathrm{x},i}\\ \varphi_{t,k}^{\mathrm{u},i}+\widehat{\varphi}_{t,k}^{\mathrm{u},i}\end{bmatrix}\right\|^2 \tag{k-DECOMP}$$

subject to: (causal dynamics), (non-causal dynamics), (locality constraints).

It is straightforward to verify that (i-PSLS) is separable induced by temporal axis $k \in [T]$, as neither the objective function nor the constraints contain coupling terms with respected to k.

3) Temporally Decaying Property: We now introduce a critical property of the closed-loop mapping $\mathbf{\Phi}$. The solution synthesized from the PSLS problem satisfies the following decaying property:

Theorem IV.1 (Temporally Decaying Property)**.** *Consider the closed-loop mappings $(\varphi_{t,k}^{\mathrm{x},i}, \widehat{\varphi}_{t,k}^{\mathrm{x},i}, \varphi_{t,k}^{\mathrm{u},i}, \widehat{\varphi}_{t,k}^{\mathrm{u},i})_{t,k\in[T]}$ as the optimal solutions to the decomposed PREDSLS problem (k-DECOMP) for subsystem $i \in [N]$. Then, there exist constants $C > 0$ and $\rho \in (0,1)$ dependent on the system matrices A, B, Q, R in (CLQR) such that*

$$\left\|\begin{bmatrix}\varphi_{t,k}^{\mathrm{x},i} & \widehat{\varphi}_{t,k}^{\mathrm{x},i}\\ \varphi_{t,k}^{\mathrm{u},i} & \widehat{\varphi}_{t,k}^{\mathrm{u},i}\end{bmatrix}\right\|^2 \le C\rho^{2|t-k|}. \tag{14}$$

We defer the proof of Theorem IV.1 to Appendix C. Intuitively, Theorem IV.1 implies that when temporal index $k \in [T]$ and time step $t \in [T]$ are far away from each other, the k temporal component of the closed-loop mappings is negligible. Therefore, using the FIR constraint to truncate the closed loop responses as illustrated in (8) results in approximations that are exponentially close to their infinite-horizon counterparts. Moreover, Theorem IV.1 provides a characterization of the finite-time stability of the closed loop of (2) under the predictive SLS controller (11) with explicit decay rate and bounding constants, which have been shown to be essential for downstream controller performance analysis [20], [21].

4) Spatially Decaying Property: Next, we confirm that the gap between the localized and centralized distributed closed-loop mappings forms an exponential decay with respect to the distance κ. The following assumption regulates the expansion of the communication graph $\mathcal{G}$.

Assumption 1 (Sub-exponential Expansion)**.** *Given a topology $\mathcal{G} = (\mathcal{E}, \mathcal{V})$ where $\mathcal{E}, \mathcal{V}$ are the sets of nodes and* *communication edges and a distance metric function $d_{\mathcal{G}}(i,j)$ for nodes i,j, there exists a sub-exponential function $g(\cdot)$ such that $|\{j \in \mathcal{V} : d_{\mathcal{G}}(i,j) = d\}| \le g(d),\ \forall i \in \mathcal{V}$.*

This sub-exponentially growing requirement of $\mathcal{G}$ is normal in the distributed control literature (c.f. Assumption 3.4 in [9] and Assumption 3 in [10]). Next, we formally present the problem considered in this work.

Theorem IV.2 (Spatial Decaying Property)**.** *Let the distributed closed-loop mapping $(\varphi^{\mathrm{x},i}_{t,k}, \widehat{\varphi}^{\mathrm{x},i}_{t,k}, \varphi^{\mathrm{u},i}_{t,k}, \widehat{\varphi}^{\mathrm{u},i}_{t,k})$ and $(\phi^{\mathrm{x},i}_{t,k}, \widehat{\phi}^{\mathrm{x},i}_{t,k}, \phi^{\mathrm{u},i}_{t,k}, \widehat{\phi}^{\mathrm{u},i}_{t,k})$ denote the optimal solutions to the problem k-*DECOMP *corresponding to the κ-distributed and centralized settings, respectively. Suppose the system (A,B) is controllable. Under Assumption 1, there exist constants $D > 0$ and $\varrho \in (0,1)$ such that*

$$\sum_{t=0}^{T}\sum_{k=0}^{T}\left\|\begin{bmatrix}\phi^{\mathrm{x},i}_{t,k}(i)-\varphi^{\mathrm{x},i}_{t,k}(i) & \widehat{\phi}^{\mathrm{x},i}_{t,k}(i)-\widehat{\varphi}^{\mathrm{x},i}_{t,k}(i)\\ \phi^{\mathrm{u},i}_{t,k}(i)-\varphi^{\mathrm{u},i}_{t,k}(i) & \widehat{\phi}^{\mathrm{u},i}_{t,k}(i)-\widehat{\varphi}^{\mathrm{u},i}_{t,k}(i)\end{bmatrix}\right\|^2 \le D\varrho^{\kappa}.$$

Note that here D, ϱ depend on the system parameters (A,B,Q,R) and the communication structure $\mathcal{G}$. Due to space limitations, we defer the proof of the spatially decaying property to a subsequent full version of this paper, and instead verify the property empirically in Section V (see Figure 2).

Theorem IV.2 implies that the localized closed-loop mapping $\mathbf{\Phi}$ approximates centralized predictive controller at an exponential rate as the communication distance κ increases. Therefore, the performance of the localized and distributed PREDSLS controller can be near-optimal by setting an appropriately chosen $\kappa \ll N$ in large-scale networked systems. Compared to a similar property in [8], Theorem IV.2 does not require the system to be open-loop stable. Our result generalizes that of [9], which considers only a disturbance-free model. Beyond κ-localization, thanks to the system-level parameterization, any communication constraints that can be expressed as convex constraints in $\mathbf{\Phi}$ can be directly incorporated into the PREDSLS controller synthesis.

V. NUMERICAL EXPERIMENTS

In this section, we validate the properties and highlight the advantage of the proposed PREDSLS through an illustrative example. Under the model assumptions, this example shares the same system matrices A, B and cost matrices Q, R as the following sparsity structure

$$A = B = \begin{bmatrix} 1 & \frac{1}{2} & 0 & \cdots & 0 & \frac{1}{2}\\ \frac{1}{2} & 1 & \frac{1}{2} & 0 & \cdots & 0\\ 0 & \frac{1}{2} & 1 & \frac{1}{2} & \cdots & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & \cdots & 0 & \frac{1}{2} & 1 & \frac{1}{2}\\ \frac{1}{2} & 0 & \cdots & 0 & \frac{1}{2} & 1 \end{bmatrix}_{30\times 30}, Q = R = I_{30\times 30}.$$

Note that here we consider a scalar system, that is, each subsystem exists only as a single entry. We use a Gaussian generator to randomly generate eligible system matrices (A,B) and consider them over 50 samplings $(\{(A_n,B_n) = \eta_n(A,B) : \eta_n \sim \mathcal{N}(0, 0.1), n \in [50]\})$ and 40 time steps $(T = 40)$. Figure 1 plots the heatmap of the average norm values of closed-loop mappings with respected to $t, k \in [T]$, visually demonstrating the temporal decaying property in Theorem IV.1. Similarly, Figure 2 shows the spatial decaying property of the closed-loop mappings with respected to κ.

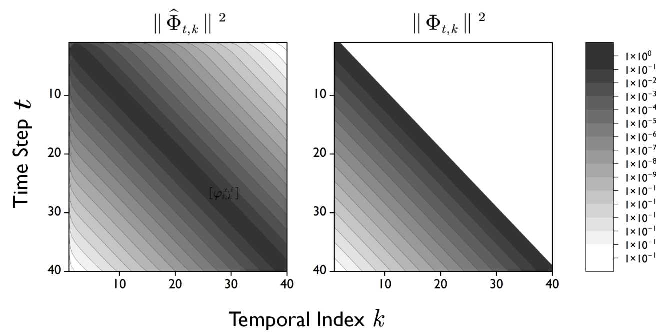

Fig. 1: Heatmap of norms of closed-loop mappings with respected to $t, k \in [T]$.

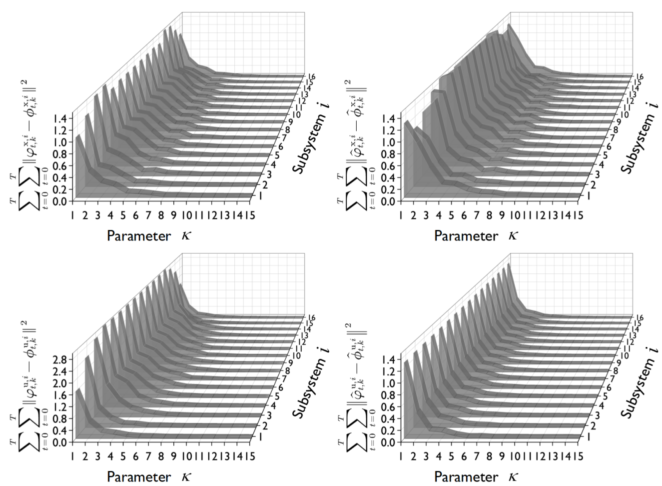

Fig. 2: Spatially decaying property in Theorem IV.2 with respect to κ. Each $\phi^i_{t,k}$ denotes an optimal solution to PREDSLS$_i$ without the (locality constraints) in (i-PSLS) (equivalently, with the maximal κ).

In addition, PREDSLS is compared to a truncated control policy with limited communication. First, we define predictive truncated control as follows

$$\mathrm{u}_i(t) = \sum_{j\in\mathcal{N}^{\kappa}(i)}\left(K^{ij,\star}\mathrm{x}_j(t) + \sum_{\tau=0}^{T-t-1} L^{ij,\star}_{\tau}\widehat{\mathrm{w}}_{\tau+t}\right),$$

where the matrices $K^{\star} := (R+B^{\top}PB)^{-1}B^{\top}PA$ and $L^{\star}_{\tau} := (R+B^{\top}PB)^{-1}[(A+BK)^{\top}]^{\tau}P$. We call this controller a *predictive truncated controller*, with similar forms considered in [9], [8].

Figure 3 presents the performance with varying communication distances where $\varepsilon := \sum_{t=0}^{T-1}\|\mathrm{w}_t - \widehat{\mathrm{w}}_t\|$ denotes the prediction error. In this setting, a constant perturbation of $w^i_t = 0.174$ is applied to each node $i \in [N]$ in $[1, 10]$ time steps, and two cases of $\varepsilon \in \{0, 10\}$ are considered.

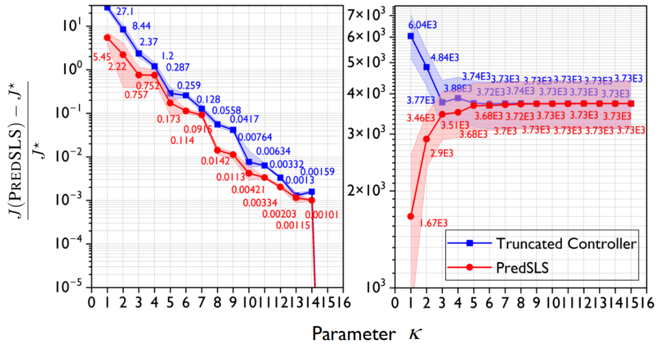

Fig. 3: Performance comparison between PREDSLS in this work and standard truncated controllers [8], [9], [10] with varying communication distance κ. **Left**: prediction error $\varepsilon = 0$; **Right**: prediction error $\varepsilon = 10$. The y-axis shows the ratio $(J(\text{PREDSLS}) - J^\star)/J^\star$ using a logarithmic scale.

Furthermore, denote by $J(\text{PREDSLS})$ the quadratic cost defined in Problem (CLQR), induced by the state-action trajectory of the closed-loop mapping $\mathbf{\Phi}$ of PREDSLS. It is evident that PREDSLS performs better than the predictive truncated control in all scenarios, in large part because PREDSLS with locality constraints maintains optimality. Under sufficiently large κ, the PREDSLS algorithm is also a perfect substitute for offline optimal controller, and we will analyze this in the extended version.

VI. Concluding Remarks

The PREDSLS framework presented in this work enables scalable, localized control of distributed systems by integrating state feedback and prediction information into an affine feedback structure. Two key directions for future research emerge. First, while the temporal decaying property mitigates prediction errors, a formal analysis of performance degradation under varying prediction inaccuracies would strengthen robustness guarantees. Second, integrating online optimization (such as in [22]) to adaptively tune prediction confidence levels could enhance the framework's resilience, dynamically balancing reliance on current states and future predictions in response to real-time uncertainty.

References

[1] S. Fattahi, N. Matni, and S. Sojoudi, "Efficient learning of distributed linear-quadratic control policies," *SIAM Journal on Control and Optimization*, vol. 58, no. 5, pp. 2927–2951, 2020.

[2] J. Yu, D. Ho, and A. Wierman, "Online adversarial stabilization of unknown networked systems," *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, vol. 7, no. 1, pp. 1–43, 2023.

[3] S. Chen, V. M. Preciado, M. Morari, and N. Matni, "Robust model predictive control with polytopic model uncertainty through system level synthesis," *Automatica*, vol. 162, p. 111431, 2024.

[4] A. Bemporad and M. Morari, "Control of systems integrating logic, dynamics, and constraints," *Automatica*, vol. 35, no. 3, pp. 407–427, 1999.

[5] E. F. Camacho, C. Bordons, E. F. Camacho, and C. Bordons, *Constrained model predictive control*. Springer, 2007.

[6] C. Yu, G. Shi, S.-J. Chung, Y. Yue, and A. Wierman, "The power of predictions in online control," *Advances in Neural Information Processing Systems*, vol. 33, pp. 1994–2004, 2020.

[7] G. Goel and B. Hassibi, "The power of linear controllers in lqr control," in *2022 IEEE 61st Conference on Decision and Control (CDC)*, pp. 6652–6657, IEEE, 2022.

[8] 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, IEEE, 2023.

[9] 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.

[10] E. Xu and G. Qu, "Stability and regret bounds on distributed truncated predictive control for networked dynamical systems," in *2024 American Control Conference (ACC)*, pp. 2604–2611, IEEE, 2024.

[11] C. Conte, C. N. Jones, M. Morari, and M. N. Zeilinger, "Distributed synthesis and stability of cooperative distributed model predictive control for linear systems," *Automatica*, vol. 69, pp. 117–125, 2016.

[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] Y.-S. Wang, N. Matni, and J. C. Doyle, "Separable and localized system-level synthesis for large-scale systems," *IEEE Transactions on Automatic Control*, vol. 63, no. 12, pp. 4234–4249, 2018.

[14] 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.

[15] C. A. Alonso, J. S. Li, J. Anderson, and N. Matni, "Distributed and localized model-predictive control–part i: Synthesis and implementation," *IEEE Transactions on Control of Network Systems*, vol. 10, no. 2, pp. 1058–1068, 2022.

[16] J. Yu, Y.-S. Wang, and J. Anderson, "Localized and distributed h2 state feedback control," in *2021 American Control Conference (ACC)*, pp. 2732–2738, IEEE, 2021.

[17] A. Leeman, J. Köhler, S. Bennani, and M. Zeilinger, "Predictive safety filter using system level synthesis," in *Learning for Dynamics and Control Conference*, pp. 1180–1192, PMLR, 2023.

[18] G. E. Dullerud and F. Paganini, *A course in robust control theory: a convex approach*, vol. 36. Springer Science & Business Media, 2013.

[19] Y.-S. Wang, N. Matni, and J. C. Doyle, "Localized lqr optimal control," in *53rd IEEE Conference on Decision and Control*, pp. 1661–1668, IEEE, 2014.

[20] S. Dean, H. Mania, N. Matni, B. Recht, and S. Tu, "On the sample complexity of the linear quadratic regulator," *Foundations of Computational Mathematics*, vol. 20, no. 4, pp. 633–679, 2020.

[21] S. Dean, H. Mania, N. Matni, B. Recht, and S. Tu, "Regret bounds for robust adaptive control of the linear quadratic regulator," in *Advances in Neural Information Processing Systems* (S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, eds.), vol. 31, Curran Associates, Inc., 2018.

[22] T. Li, H. Liu, and Y. Yue, "Disentangling linear quadratic control with untrusted ml predictions," in *The Thirty-eighth Annual Conference on Neural Information Processing Systems*.

Appendix

A. Proof of Theorem III.1

Proof. Consider arbitrary linear predictive controller $\mathbf{u} := \mathbf{K}\mathbf{x} + \widehat{\mathbf{L}}\widehat{\mathbf{w}}$. Then the closed loop response of (2) is the right hand side of (5). It can verified that all closed-loop responses generated by a linear predictive controller satisfy (6) as follows

$$\begin{aligned}
&\begin{bmatrix} zI - A & -B \end{bmatrix} \begin{bmatrix} U & UB\widehat{\mathbf{L}} \\ \mathbf{K}U & (\mathbf{K}UB + I)\widehat{\mathbf{L}} \end{bmatrix} \\
&= \begin{bmatrix} (zI - A - B\mathbf{K})U & (zI - A - B\mathbf{K})UB\widehat{\mathbf{L}} - \mathcal{B}\widehat{\mathbf{L}} \end{bmatrix} \\
&= \begin{bmatrix} I & 0 \end{bmatrix}.
\end{aligned}$$

For the other direction, consider $\mathbf{\Phi}$ satisfying constraint (6). First, there exists $(\mathbf{\Phi}^{\mathrm{x}})^{-1}$ since (6) implies that $\mathbf{\Phi}$ is a block lower triangular matrix with identity matrices on its diagonal blocks. Now let $\widehat{\mathbf{L}} := \widehat{\mathbf{\Phi}}^{\mathrm{u}} - \mathbf{\Phi}^{\mathrm{u}}(\mathbf{\Phi}^{\mathrm{x}})^{-1}\widehat{\mathbf{\Phi}}^{\mathrm{x}}$ and

$\mathbf{K} := \mathbf{\Phi}^{\mathrm{u}}(\mathbf{\Phi}^{\mathrm{x}})^{-1}$, we have

$$
\begin{aligned}
\mathbf{x} &= U\mathbf{w} + U\mathcal{B}\widehat{\mathbf{L}}\widehat{\mathbf{w}} \\
&= (zI - (A + B\mathbf{\Phi}^{\mathrm{u}}(\mathbf{\Phi}^{\mathrm{x}})^{-1}))^{-1}\mathbf{w} \\
&\quad + (zI - (A + B\mathbf{\Phi}^{\mathrm{u}}(\mathbf{\Phi}^{\mathrm{x}})^{-1}))^{-1}B(\widehat{\mathbf{\Phi}}^{\mathrm{u}} - \mathbf{\Phi}^{\mathrm{u}}(\mathbf{\Phi}^{\mathrm{x}})^{-1}\widehat{\mathbf{\Phi}}^{\mathrm{x}})\widehat{\mathbf{w}} \\
&= ((zI - A)\mathbf{\Phi}^{\mathrm{x}} + B\mathbf{\Phi}^{\mathrm{u}})^{-1}\mathbf{\Phi}^{\mathrm{x}}\mathbf{w} \\
&\quad + (zI - (A + B\mathbf{\Phi}^{\mathrm{u}}(\mathbf{\Phi}^{\mathrm{x}})^{-1}))^{-1} \\
&\qquad (zI - (\mathcal{A} + \mathcal{B}\mathbf{\Phi}^{\mathrm{u}}(\mathbf{\Phi}^{\mathrm{x}})^{-1}))\widehat{\mathbf{\Phi}}^{\mathrm{x}}\widehat{\mathbf{w}} \\
&= \mathbf{\Phi}^{\mathrm{x}}\mathbf{w} + \widehat{\mathbf{\Phi}}^{\mathrm{x}}\widehat{\mathbf{w}},
\end{aligned}
$$

where the third equality utilizes the property of constraint (6). Furthermore,

$$
\begin{aligned}
\mathbf{u} &= \mathbf{K}\mathbf{x} + \widehat{\mathbf{L}}\widehat{\mathbf{w}} \\
&= \mathbf{\Phi}^{\mathrm{u}}(\mathbf{\Phi}^{\mathrm{x}})^{-1}(\mathbf{\Phi}^{\mathrm{x}}\mathbf{w} + \widehat{\mathbf{\Phi}}^{\mathrm{x}}\widehat{\mathbf{w}}) + (\widehat{\mathbf{\Phi}}^{\mathrm{u}} - \mathbf{\Phi}^{\mathrm{u}}(\mathbf{\Phi}^{\mathrm{x}})^{-1}\widehat{\mathbf{\Phi}}^{\mathrm{x}})\widehat{\mathbf{w}} \\
&= \mathbf{\Phi}^{\mathrm{u}}\mathbf{w} + \widehat{\mathbf{\Phi}}^{\mathrm{u}}\widehat{\mathbf{w}},
\end{aligned}
$$

where the third equality is derived from the state trajectory in the last step. Therefore, there exists linear predictive controller that realizes the prescribed closed-loop mappings $\mathbf{\Phi}$. This concludes the proof. □

B. Proof of Proposition 1

Proof. Firstly, we consider a special case where the disturbances w_t are independent and identically distributed Gaussian with zero mean and identity variance and the predictions are modeled as random variables. Then, the following equivalence holds:

$$
\begin{aligned}
(\text{CLQR}) &\xlongequal{Q_T = P} \min_{\mathbf{x},\mathbf{u}} \sum_{t=0}^{\infty}\left(\mathrm{x}_t^\top Q \mathrm{x}_t + \mathrm{u}_t^\top R \mathrm{u}_t\right), \text{s.t. } (2) \\
&\xlongequal{\text{Theorem III.1}} \min_{\mathbf{\Phi}} \left\| \begin{bmatrix} Q^{\frac{1}{2}} & \mathbf{0} \\ \mathbf{0} & R^{\frac{1}{2}} \end{bmatrix} \begin{bmatrix} \mathbf{\Phi}^{\mathrm{x}} & \widehat{\mathbf{\Phi}}^{\mathrm{x}} \\ \mathbf{\Phi}^{\mathrm{u}} & \widehat{\mathbf{\Phi}}^{\mathrm{u}} \end{bmatrix} \begin{bmatrix} I & \Sigma_{\mathbf{w}\widehat{\mathbf{w}}} \\ \Sigma_{\widehat{\mathbf{w}}\mathbf{w}} & I \end{bmatrix}^{\frac{1}{2}} \right\|_F^2, \text{s.t. } (10) \\
&\xlongequal{\widehat{\mathbf{w}} = \mathbf{w}} \min_{\mathbf{\Phi}} \left\| \begin{bmatrix} Q^{\frac{1}{2}} & \mathbf{0} \\ \mathbf{0} & R^{\frac{1}{2}} \end{bmatrix} \begin{bmatrix} \mathbf{\Phi}^{\mathrm{x}} + \widehat{\mathbf{\Phi}}^{\mathrm{x}} \\ \mathbf{\Phi}^{\mathrm{u}} + \widehat{\mathbf{\Phi}}^{\mathrm{u}} \end{bmatrix} \right\|_F^2, \text{s.t. } (10),
\end{aligned}
$$

where $\Sigma_{\mathbf{w}\widehat{\mathbf{w}}}$ denotes the covariance matrix of $\mathrm{w}_t, \widehat{\mathrm{w}}_t$. Let $(\mathbf{\Phi}^{\mathrm{x}\star}, \widehat{\mathbf{\Phi}}^{\mathrm{x}\star}, \mathbf{\Phi}^{\mathrm{u}\star}, \widehat{\mathbf{\Phi}}^{\mathrm{u}\star})$ be the optimal solution to (PSLS), we know that by Theorem III.1, $\mathbf{K}^\star = \mathbf{\Phi}^{\mathrm{u}\star}(\mathbf{\Phi}^{\mathrm{x}\star})^{-1}$ and $\widehat{\mathbf{L}}^\star = \widehat{\mathbf{\Phi}}^{\mathrm{u}\star} - \mathbf{\Phi}^{\mathrm{u}\star}(\mathbf{\Phi}^{\mathrm{x}\star})^{-1}\widehat{\mathbf{\Phi}}^{\mathrm{x}\star}$ form an optimal predictive controller to (CLQR).

In the case of general disturbances, by Theorem 4.1 in [6], the optimal control policy for (CLQR) is indeed the optimal controller for all general disturbances when the predictions match the disturbances exactly. Hence, we can extend the optimality of PREDSLS to all other disturbance sequences. This completes the proof. □

C. Proof of Theorem IV.1

The proof leverages the characterization of the optimal solution to (k-DECOMP), formalized as follows:

Lemma 1. *There exist $\bar{K} \in \mathbb{R}^{m\times n}$, $\bar{M}^{[\tau]} \in \mathbb{R}^{m\times n}$, $C_1 > 0$, $L > 0$, and $\rho \in (0, 1)$ such that $\|\bar{K}\| \leq L$, $\|\bar{M}^{[\tau]}\| \leq C_1\rho^\tau$, and $\|A + B\bar{K}\|^\tau \leq L\rho^\tau$ for all $\tau \in \mathbb{Z}_+$. In particular, the optimal solution to (k-DECOMP) can be written as*

$$
\begin{bmatrix} \varphi_{t,k}^{\mathrm{x},i} \\ \varphi_{t,k}^{\mathrm{u},i} \\ \widehat{\varphi}_{t,k}^{\mathrm{x},i} \\ \widehat{\varphi}_{t,k}^{\mathrm{u},i} \end{bmatrix} = \begin{bmatrix} (A + B\bar{K})^{t-k}(\mathrm{e}_i \circ \mathbf{1}(t-k)) \\ \bar{K}(A + B\bar{K})^{t-k}(\mathrm{e}_i \circ \mathbf{1}(t-k)) \\ \sum_{\tau=0}^{t-1}(A + B\bar{K})^\tau B\bar{M}^{[k-t+\tau+1]}\mathrm{e}_i \\ \left(\sum_{\tau=0}^{t-1}\bar{K}(A + B\bar{K})^\tau B\bar{M}^{[k-t+\tau+1]} + \bar{M}^{[k-t]}\right)\mathrm{e}_i \end{bmatrix},
$$

where e_i denotes the standard basis vector, indices $t, k \in \mathbb{Z}_+$; $\mathbf{1}(\tau)$ denotes the Heaviside step function, and $\bar{M}^{[\tau]} = 0$ if $\tau < 0$.

Due to space limit, we refer the proof of Lemma 1 to the full version of this paper. Now consider

$$
\left\| \begin{bmatrix} \varphi_{t,k}^{\mathrm{x},i} & \widehat{\varphi}_{t,k}^{\mathrm{x},i} \\ \varphi_{t,k}^{\mathrm{u},i} & \widehat{\varphi}_{t,k}^{\mathrm{u},i} \end{bmatrix} \right\|^2 \leq \underbrace{\left(\left\|\varphi_{t,k}^{\mathrm{x},i}\right\| + \left\|\varphi_{t,k}^{\mathrm{u},i}\right\|\right)^2}_{\text{(a)}} + \underbrace{\left(\left\|\widehat{\varphi}_{t,k}^{\mathrm{x},i}\right\| + \left\|\widehat{\varphi}_{t,k}^{\mathrm{u},i}\right\|\right)^2}_{\text{(b)}},
$$

where we have used the fact that $\left\|\begin{bmatrix} M_1 \ M_2 \end{bmatrix}\right\|^2 \leq \|M_1\|^2 + \|M_2\|^2$. We now bound (a) and (b) separately. Applying Lemma 1, it follows that

$$
\begin{aligned}
\text{(a)} &\leq \left(\|(A + B\bar{K})^{t-k}\| + \|\bar{K}(A + B\bar{K})^{t-k}\|\right)^2 \qquad (16) \\
&\leq (1 + \|\bar{K}\|)^2\|A + B\bar{K}\|^{t-k} \\
&\leq (1 + L)^2L^2\rho^{2|t-k|}. \qquad (17)
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
\text{(b)} &\leq \Bigg(\left\|\sum_{\tau=0}^{t-1}(A + B\bar{K})^\tau B\bar{M}^{[k-t+\tau+1]}\right\| \\
&\quad + \left\|\sum_{\tau=0}^{t-1}\bar{K}(A + B\bar{K})^\tau B\bar{M}^{[k-t+\tau+1]} + \bar{M}^{[k-t]}\right\|\Bigg)^2 \\
&\leq \Bigg(\sum_{\tau=0}^{t-1}(1 + \|\bar{K}\|)\|A + B\bar{K}\|^\tau\|B\|\|\bar{M}^{[k-t+\tau+1]}\| \\
&\quad + \|\bar{M}^{[k-t]}\|\Bigg)^2 \\
&\leq \left(\sum_{\tau=0}^{t-1}(1 + L)LC_1\rho^{2\tau+1} + 1\right)^2 C_1^2\rho^{2|t-k|} \qquad (18) \\
&\leq \frac{((1 + L)LC_1\rho + (1 - \rho^2))^2C_1^2\rho^{2|t-k|}}{(1 - \rho^2)^2}, \qquad (19)
\end{aligned}
$$

where (18) holds because by assumption the system is stabilizable. Combining the upper bounds in (17) and (19) on (a) and (b), we obtain that $\left\| \begin{bmatrix} \varphi_{t,k}^{\mathrm{x},i} & \widehat{\varphi}_{t,k}^{\mathrm{x},i} \\ \varphi_{t,k}^{\mathrm{u},i} & \widehat{\varphi}_{t,k}^{\mathrm{u},i} \end{bmatrix} \right\|^2$

$$
\leq \left((1 + L)^2L^2 + \frac{((1 + L)LC_1\rho + (1 - \rho^2))^2C_1^2}{(1 - \rho^2)^2}\right)\rho^{2|t-k|}.
$$

□