Stochastic Data-Driven Predictive Control with Equivalence to Stochastic MPC

Ruiqi Li, *Student Member, IEEE*, John W. Simpson-Porco, *Senior Member, IEEE*, and Stephen L. Smith, *Senior Member, IEEE*

***Abstract*— We propose a data-driven receding-horizon control method for the chance-constrained output-tracking problem in unknown stochastic LTI systems. The method accounts for the statistics of process noise, measurement noise, and the uncertain initial condition, following a framework analogous to stochastic MPC (SMPC), but without a parametric system model. Accordingly, the algorithm generates a sequence of closed-loop control policies over the prediction horizon, rather than open-loop control actions. Under suitable conditions, we show that the proposed data-driven method produces the same control inputs as the corresponding model-based SMPC. Simulation results on a grid-connected power converter are provided to illustrate the performance benefits of our methodology.**

I. Introduction

Model predictive control (MPC) is a standard multivariable control method [1] that handles hard constraints on inputs, states, and outputs, and accommodates complex performance objectives. Constraints capture actuator limits and safety requirements in safety-critical settings. For stochastic systems, *Stochastic MPC (SMPC)* [2]–[4] models system uncertainty probabilistically. SMPC optimizes over feedback policies rather than open-loop actions, yielding better performance than naïvely applying deterministic MPC [5]. SMPC also supports probabilistic constraints, enabling paradigms such as risk-aware controller design. A related approach is *Robust MPC (RMPC)* [6], which guards against worst-case deterministic uncertainty; here we focus on the stochastic setting.

MPC and SMPC use system models, derived either from first principles or via identification, to predict the effect of candidate inputs, and are therefore *indirect* design approachs [7], [8]. In contrast, *direct* data-driven methods aim to design controllers directly from input–output data, and are attractive for systems that are difficult to model [9], [10]. The best achievable performance obtainable via data-driven control is clearly that of model-based control with a perfect model. For deterministic linear time-invariant (LTI) systems, recent work has demonstrated that the data-driven control methods can indeed produce controls that are equivalent to their model-based counterparts [11], [12]. However, for stochastic systems, equivalence between a data-based and model-based method have not been established, except in a few special cases which will be discussed shortly. Thus, the focus of this work is to develop a stochastic data-driven control framework with provable equivalence to its model-based SMPC counterpart.

Related Work: Early data-driven methods often ignored input/output constraints [10], which motivated *data-driven predictive control (DDPC)* methods that handle such constraints. Two prominent DDPC schemes are Data-enabled Predictive Control (DeePC) [12]–[14] and Subspace Predictive Control (SPC) [11], both validated experimentally across several applications [15]–[20]. For deterministic LTI systems, DeePC and SPC are theoretically equivalent to model-based MPC, producing the same control actions [11], [12].

Outside the idealized deterministic linear setting, real systems are often stochastic or nonlinear, and measured data are noisy. Practical data-driven methods must therefore handle such perturbations. Many classical approaches are designed to be robust [10], so performance is not highly sensitive to noise. In SPC with noisy data, the predictor matrix is typically denoised using methods such as prediction error techniques [18], [19] or by truncating a SVD of a data matrix [16].

Robust DeePC variants for stochastic settings have been proposed, including norm-regularized DeePC [12], [13], whose regularization admits a worst-case robust interpretation [21], [22], and distributionally robust DeePC [13], [14]. Other DeePC extensions address closed-loop stability [23]–[26] and robustness to nonlinear systems [27]. Although these stochastic DeePC/SPC variants are experimentally validated, they lack a theoretical equivalence to stochastic MPC or other model-based schemes. Related stochastic DDPC approaches include tube-based [28], sampling-based [29], innovation-based [30], and constraint-tightening [31] formulations, but none establish performance equivalence with model-based SMPC.

This disconnect between data-driven and model-based methods in the stochastic case has been noticed by some researchers, and some recent DDPC methods were developed for stochastic systems that have provable equivalence to model-based MPC methods. The works in [32]–[35] proposed data-driven control frameworks for stochastic systems applying Polynomial Chaos Expansion (PCE); the use of PCE enables modeling of arbitrary random variables of finite mean and variance. Their methods have equivalent performance to SMPC when disturbances are known and when stochastic signals are exactly represented by finite PCE terms [32, Thm. 1] [33, Cor. 1]. In practice, disturbances should be estimated using input-output data, which

This research is supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC).

Ruiqi Li and Stephen L. Smith are with the Electrical and Computer Engineering at the University of Waterloo, Waterloo, ON, Canada {r298li,stephen.smith}@uwaterloo.ca

John W. Simpson-Porco is with the Department of Electrical and Computer Engineering at the University of Toronto, Toronto, ON, Canada jwsimpson@ece.utoronto.ca

requires heavier computation with larger amounts of data. Their frameworks have considered systems without sensor noise and systems in the Auto-Regressive form with eXogenous input (ARX), which are special cases of systems in the state-space representation. Thus, the gap addressed in this paper is to develop an alternative data-driven stochastic control method that has provably equivalent performance to the model-based SMPC, where we only estimate a fixed number of parameters regardless of data amount, and we consider general systems in the state-space form with separate process and sensor noise.

Contributions: We develop a DDPC control method for stochastic LTI systems. Our technical approach is based on the construction of an auxiliary state model directly parameterized by input-output data. Mirroring SMPC, we formulate a stochastic control problem using this data-based auxiliary model, and establish equivalence between the proposed data-driven approach and its model-based SMPC counterpart. The method retains three main SMPC features: (i) it explicitly accounts for process and measurement noise, allowing performance to be analyzed under different noise levels, (ii) it computes a feedback policy at every step, so inputs are chosen after real-time measurements are obtained, and (iii) it enforces safety chance constraints, consistent with the SMPC framework. Because the construction closely parallels SMPC, it can be adapted to data-driven versions of variants such as distributionally robust SMPC and SMPC with correlated noise.

Organization: Section II presents the formal problem statement, with a brief overview of SMPC in Section II-A. Our control method is introduced in Section III, where we show the formulation and the theoretical performance guarantee, i.e., equivalence to SMPC. Simulation results are displayed in Section IV, comparing our proposed method and some benchmark control methods, and Section V offers conclusions.

Notation: Let $M^\dagger$ be the pseudo-inverse of a matrix M. Let $\otimes$ denote the Kronecker product. Let $\mathbb{S}^q_+$ and $\mathbb{S}^q_{++}$ be the sets of $q\times q$ positive semi-definite and positive definite matrices respectively. Let $\mathrm{col}(M_1,\dots,M_k)$ denote the column concatenation, and $\mathrm{Diag}(M_1,\dots,M_k)$ the block-diagonal concatenation, of matrices/vectors $M_1,\dots,M_k$. Let $\mathbb{Z}_{[a,b]} := [a,b]\cap\mathbb{Z}$ denote a set of consecutive integers from a to b. Let $\mathbb{Z}_{[a,b)} := \mathbb{Z}_{[a,b-1]}$. For a $\mathbb{R}^q$-valued discrete-time signal z_t with integer index t, let $z_{[t_1,t_2]}$ denote either a sequence $\{z_t\}_{t=t_1}^{t_2}$ or a concatenated vector $\mathrm{col}(z_{t_1},\dots,z_{t_2})\in\mathbb{R}^{q(t_2-t_1+1)}$ where the usage is clear from the context. Similarly, let $z_{[t_1,t_2)} := z_{[t_1,t_2-1]}$. A matrix sequence $\{M_t\}_{t=t_1}^{t_2}$ and a function sequence $\{\pi_t(\cdot)\}_{t=t_1}^{t_2}$ are denoted by $M_{[t_1,t_2]}$ and $\pi_{[t_1,t_2]}$ respectively.

II. Problem Statement

We consider a stochastic linear time-invariant (LTI) system

$$x_{t+1} = Ax_t + Bu_t + w_t, \tag{1a}$$
$$y_t = Cx_t + Du_t + v_t, \tag{1b}$$

with input $u_t\in\mathbb{R}^m$, state $x_t\in\mathbb{R}^n$, output $y_t\in\mathbb{R}^p$, process noise $w_t\in\mathbb{R}^n$, and measurement noise $v_t\in\mathbb{R}^p$, all of which are random variables. The initial state x_0 is uncertain with given mean $\mu^{\mathsf{x}}_{\mathsf{ini}}$ and with variance to be specified by a steady-state Kalman filter. The system matrices A,B,C,D are *unknown* and the state x_t is *unmeasured*; we have access only to the input u_t and output y_t in (1). The disturbances w_t and v_t in (1) are independent of each other and of x_0, and are independently and identically distributed (i.i.d.) normally with zero mean and with variances $\Sigma^{\mathsf{w}}\in\mathbb{S}^n_+$ and $\Sigma^{\mathsf{v}}\in\mathbb{S}^p_{++}$ respectively, i.e.,

$$w_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0_{n\times 1},\Sigma^{\mathsf{w}}),\quad v_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0_{p\times 1},\Sigma^{\mathsf{v}}). \tag{2}$$

We assume the system (A,B,C,D) is controllable and observable (i.e., a minimal realization), where observability is assumed without loss of generality for an unknown system [7, Sec. 2.4]. Let $L\in\mathbb{N}$ be such that the extended observability matrix $\mathcal{O} := \mathrm{col}(C,CA,\dots,CA^{L-1})$ has full column rank; such smallest L is the *lag* of the system [7], [8]. Finally, we assume the pair (A,Σ^{w}) is stabilizable (or equivalently, $(A,(\Sigma^{\mathsf{w}})^{1/2})$ is stabilizable), which will subsequently ensure uniqueness of the state variance by the Kalman filter [36].

The objective is for the output y_t to follow a specified reference signal $r_t\in\mathbb{R}^p$. The trade-off between tracking error and control effort is encoded in the stage cost

$$J_t(u_t,y_t) := \|y_t - r_t\|^2_Q + \|u_t\|^2_R \tag{3}$$

where $Q\in\mathbb{S}^p_{++}$ and $R\in\mathbb{S}^m_{++}$ are user-selected parameters. This tracking should be achieved subject to constraints on the inputs and outputs. We consider a polytopic constraint in the form $E\,\mathrm{col}(u_t,y_t)\le f$, modeled in the stochastic setting as a probabilistic *chance constraint*

$$\mathbb{P}\{E\,\mathrm{col}(u_t,y_t)\le f\}\ge 1-p \tag{4}$$

for $t\in\mathbb{N}_{\ge 0}$, where $E\in\mathbb{R}^{q\times(m+p)}$ is a fixed matrix, $f\in\mathbb{R}^q$ is a fixed vector, with some $q\in\mathbb{N}$, and $p\in(0,1)$ is a probability bound of constraint violation. One can similarly impose multiple chance constraints, e.g., separate input and output chance constraints, in the form of (4). We remark that the input-output constraints (4) are more natural than state constraints (e.g., [2]), as our focus later will be on input-output control in a data-driven, where there is no unique physical internal state [12]. In a model-based setting where A,B,C,D are known, the general control problem above can be addressed by SMPC, as will be reviewed in Section II-A. Our broad objective is to construct a direct data-driven method that addresses the same stochastic control problem and is equivalent, under certain tuning conditions, to SMPC.

A. Stochastic MPC: A Benchmark Model-Based Design

Several formulations of SMPC methods have been developed in the literature [2, Table 2]. Our focus is on output-feedback SMPC [37]–[41], which is typically approached by enforcing a separation principle within the design, augmenting full-state-feedback SMPC with state estimation. Our formulation here is based on an affine feedback-policy ([38], [39]), integrates an interpolation of the initial condition [42], [43], which is required for recursive feasibility with unbounded noise, and approximates chance constraints as in [44].

1) Initial Condition and State Estimation: At control step $t = k$, the initial condition of the state x_k is modelled as

$$x_k \sim \mathcal{N}(\mu_k^{\mathsf{x}}, \Sigma^{\mathsf{x}}), \tag{5}$$

where the mean $\mu_k^{\mathsf{x}} \in \mathbb{R}^n$ depends on a decision variable $\theta \in [0,1]$, according to an interpolation technique to be introduced in Section II-A-2. The state variance $\Sigma^{\mathsf{x}} \in \mathbb{S}_+^n$ in (5) is fixed and induced by the steady-state Kalman filter. Specifically, Σ^{x} is the *unique* positive semidefinite solution to the associated discrete-time algebraic Riccatti equation (DARE) [36]

$$\Sigma^{\mathsf{x}} = (A - L_{\mathsf{L}} C)\Sigma^{\mathsf{x}} A^{\mathsf{T}} + \Sigma^{\mathsf{w}} \tag{6a}$$

$$L_{\mathsf{L}} := A L_{\mathsf{K}}, \quad L_{\mathsf{K}} := \Sigma^{\mathsf{x}} C^{\mathsf{T}} (C\Sigma^{\mathsf{x}} C^{\mathsf{T}} + \Sigma^{\mathsf{v}})^{-1} \tag{6b}$$

where $L_{\mathsf{K}} \in \mathbb{R}^{n\times p}$ denotes the steady-state Kalman gain and $L_{\mathsf{L}} \in \mathbb{R}^{n\times p}$ the associated Luenberger observer gain. With the initial condition (5), we simulate the noise-free model for future N time steps,

$$\bar{x}_{t+1} := A\bar{x}_t + B\bar{u}_t, \qquad t \in \mathbb{Z}_{[k,k+N)} \tag{7a}$$

$$\bar{y}_t := C\bar{x}_t + D\bar{u}_t, \qquad t \in \mathbb{Z}_{[k,k+N)} \tag{7b}$$

$$\bar{x}_k := \mu_k^{\mathsf{x}} \tag{7c}$$

where the *nominal inputs* $\bar{u}_t \in \mathbb{R}^m$ for $t \in \mathbb{Z}_{[k,k+N)}$ will be decision variables in optimization, with resulting *nominal states* $\bar{x}_t \in \mathbb{R}^n$ and *nominal outputs* $\bar{y}_t \in \mathbb{R}^p$.

As measurements are taken, estimates of the future states over the desired horizon will be computed through the steady-state Kalman filter, with L_{K} in (6b),

$$\hat{x}_t^+ := \hat{x}_t^- + L_{\mathsf{K}}(y_t - C\hat{x}_t^- - Du_t), \quad t \in \mathbb{Z}_{[k,k+N)} \tag{8a}$$

$$\hat{x}_{t+1}^- := A\hat{x}_t^+ + Bu_t, \qquad t \in \mathbb{Z}_{[k,k+N)} \tag{8b}$$

$$\hat{x}_k^- := \mu_k^{\mathsf{x}} \tag{8c}$$

where $\hat{x}_t^+$ and $\hat{x}_t^-$ denote the posterior and prior estimates of x_t, respectively.

2) Interpolation of Initial Condition: Two common choices of μ_k^{x} in (5) are (i) the prior state estimate $\hat{x}_k^-$ produced by the estimator (8) in the previous control step, denoted by $\mu_k^{\hat{\mathsf{x}}}$ [39]–[41], and (ii) the deterministic prediction $\bar{x}_k$ of state the x_k, obtained via (7) at last control step [38]; we denote this choice by $\mu_k^{\bar{\mathsf{x}}}$. Here, we let the initial condition μ_k^{x} in (5) *interpolate* between $\mu_k^{\hat{\mathsf{x}}}$ and $\mu_k^{\bar{\mathsf{x}}}$ [42], [43] as

$$\mu_k^{\mathsf{x}} := (1-\theta)\,\mu_k^{\hat{\mathsf{x}}} + \theta\,\mu_k^{\bar{\mathsf{x}}}, \tag{9}$$

where $\theta \in [0,1]$ is a decision variable, and both $\mu_k^{\hat{\mathsf{x}}}, \mu_k^{\bar{\mathsf{x}}} \in \mathbb{R}^n$ are fixed and known at time $t = k$. At initial control step $k = 0$, μ_0^{x} is equal to a given parameter $\mu_{\mathsf{ini}}^{\mathsf{x}}$, i.e., we let $\mu_0^{\hat{\mathsf{x}}} := \mu_{\mathsf{ini}}^{\mathsf{x}}$ and $\mu_0^{\bar{\mathsf{x}}} := \mu_{\mathsf{ini}}^{\mathsf{x}}$.

3) Feedback Control Policies: Stochastic control requires the determination of (causal) feedback policies π_t which map the observation history into control actions. We restrict our attention here to affine feedback policies, written as (cf. [37]–[39])

$$u_t = \pi_t(\hat{x}_t^-) := \bar{u}_t - K(\hat{x}_t^- - \bar{x}_t), \tag{10}$$

where $K \in \mathbb{R}^{m\times n}$ is the infinite-horizon LQR gain of system (1) with LQR stage cost $\|Cx_t + Du_t\|_Q^2 + \|u_t\|_R^2$, i.e.,

$$K := (R + B^{\mathsf{T}} P B + D^{\mathsf{T}} Q D)^{-1}(B^{\mathsf{T}} P A + D^{\mathsf{T}} Q C) \tag{11}$$

where $P \in \mathbb{S}_+^n$ is the *unique* positive semidefinite solution to the discrete-time algebraic Riccati equation (DARE) [36]

$$P = A^{\mathsf{T}} P (A - BK) + C^{\mathsf{T}} Q (C - DK). \tag{12}$$

See [2], [37]–[40], [45] for related policy parameterizations, and [38], [41] for methods which can account for input constraints.

4) Propagation of Input-Output Distribution: With (1), (2), (5), (7), (8) and (10), at control step $t = k$, the resulting future inputs u_t and outputs y_t for $t \in \mathbb{Z}_{[k,k+N)}$ are distributed as

$$\begin{bmatrix} u_t \\ y_t \end{bmatrix} \sim \mathcal{N}\Big(\begin{bmatrix} \bar{u}_t \\ \bar{y}_t \end{bmatrix}, \Delta_{t-k}\Big), \tag{13}$$

where the covariance matrix $\Delta_s \in \mathbb{S}_+^{m+p}$ for $s \in \mathbb{Z}_{[0,N)}$ can be computed as (14a) using $\Lambda_s \in \mathbb{S}_+^n$ defined by (14b),

$$\Delta_s := \begin{bmatrix} -K \\ C - DK \end{bmatrix} \Lambda_s \begin{bmatrix} -K \\ C - DK \end{bmatrix}^{\mathsf{T}} + \begin{bmatrix} 0_{m\times m} & \\ & C\Sigma^{\mathsf{x}} C^{\mathsf{T}} + \Sigma^{\mathsf{v}} \end{bmatrix} \tag{14a}$$

$$\Lambda_s := \textstyle\sum_{r=0}^{s-1} (A - BK)^r L_{\mathsf{L}} \,(C\Sigma^{\mathsf{x}} C^{\mathsf{T}} + \Sigma^{\mathsf{v}})\, L_{\mathsf{L}}^{\mathsf{T}} (A - BK)^{r\mathsf{T}} \tag{14b}$$

with L_{L} in (6b) and K in (11). The distribution (13) is derived analogously to [38] and a complete derivation can be found in the extended version [46, Appendix A]. Note that the matrices $\Delta_0, \Delta_1, \ldots, \Delta_{N-1}$ are fixed and can be computed offline.

SMPC problems typically consider the expectation of cost (3) to be summed over the desired horizon. Given distribution (13), the expected cost is known as a deterministic value

$$\textstyle\sum_{t=k}^{k+N-1} \mathbb{E}[J_t(u_t, y_t)] = \sum_{t=k}^{k+N-1} \big[J_t(\bar{u}_t, \bar{y}_t) + J_{t-k}^{\mathsf{var}}\big], \tag{15}$$

where $J_s^{\mathsf{var}} := \mathrm{Trace}(\Delta_s \,\mathrm{Diag}(R, Q))$ is a constant independent of the decision variables $\bar{u}$ and θ.

5) Chance-Constraint Approximation: The joint chance constraint (4) is challenging to accommodate computationally [4, Sec. 2.2]. Following [44], we approximate it via individual chance constraints for each time $t \in \mathbb{Z}_{[k,k+N)}$,

$$\mathbb{P}\big\{e_i^{\mathsf{T}} \begin{bmatrix} u_t \\ y_t \end{bmatrix} \le f_i\big\} \ge 1 - p_{i,t}, \quad i \in \mathbb{Z}_{[1,q]} \tag{16}$$

where $e_i \in \mathbb{R}^{m+p}$ is the transposed i-th row of E, and $f_i \in \mathbb{R}$ is the i-th entry of f. The allocated risk probabilities $p_{i,t} > 0$ in (16) are introduced as additional decision variables, such that $p_{1,t}, p_{2,t}, \ldots, p_{q,t}$ sum up to the total risk p for each time t. Note that (16) is a conservative approximation (or a sufficient condition) of (4), due to subadditivity of probabilities. Given distribution (13), the chance constraints (16) are converted into an equivalent deterministic form, cf. [38], [44],

$$e_i^{\mathsf{T}} \begin{bmatrix} \bar{u}_t \\ \bar{y}_t \end{bmatrix} \le f_i - \sqrt{e_i^{\mathsf{T}} \Delta_{t-k}\, e_i}\,\mathrm{icdfn}(1 - p_{i,t}),\ i \in \mathbb{Z}_{[1,q]} \tag{17a}$$

$$\textstyle\sum_{i=1}^q p_{i,t} = p, \qquad p_{i,t} > 0, \quad i \in \mathbb{Z}_{[1,q]} \tag{17b}$$

for $t \in \mathbb{Z}_{[k,k+N)}$, where $\mathrm{icdfn}(z) := \sqrt{2}\,\mathrm{erf}^{-1}(2z-1)$ is the inverse cumulative distribution function (inverse c.d.f.) or the z-quantile of the standard normal distribution, with erf^{-1} the inverse error function. The constraints (17) are convex when we require $p \in (0, \frac{1}{2}]$ [44, Thm. 1].

Remark 1 (**Non-Gaussian Noise**). In the case where random signals are non-Gaussian but with the same means and variances in (2) and (5), the resulting inputs u_t and outputs y_t still possess the mean and variance in (13), and thus the expected

cost is still (15). However, the inverse c.d.f. in (17a) should change correspondingly to the actual distribution (if known), or be replaced into an upper bound $\sqrt{(1-p_{i,t})/p_{i,t}}$ via Chebyshev–Cantelli inequality that guarantees the worst case over all distributions [29], [31]. ■

6) Terminal Condition: Terminal constraints are considered in (S)MPC frameworks to ensure recursive feasibility and closed-loop stability. Assume $N \geq L$ going forward. Here, we impose a *terminal equality constraint* [23]–[26],

$$\begin{aligned} \bar{u}_{k+N-L} &= \bar{u}_{k+N-L+1} = \cdots = \bar{u}_{k+N-1} \\ \bar{y}_{k+N-L} &= \bar{y}_{k+N-L+1} = \cdots = \bar{y}_{k+N-1} \end{aligned} \tag{18}$$

that requires the nominal input-output trajectory to stay at some setpoint for final L steps in the prediction horizon.

7) SMPC Optimization Problem and Implementation: With the expected cost (15), the approximation (17) of the constraint (4), the interpolation (9) and the terminal constraint (18), the SMPC problem is formulated as

$$\begin{aligned} \underset{\bar{u},\,\theta,\,p_{i,t}}{\text{minimize}} \quad & \textstyle\sum_{t=k}^{k+N-1} J_t(\bar{u}_t, \bar{y}_t) + \lambda_\theta\, \theta \\ \text{subject to} \quad & (17) \text{ for } t \in \mathbb{Z}_{[k,k+N)}, (7), (9), (18), \end{aligned} \tag{19}$$

with an interpolation penalty term of parameter $\lambda_\theta > 0$ [43]. With $R \succ 0$ and $\lambda_\theta > 0$, the cost in (19) is jointly strongly convex in $\bar{u}$ and θ, and thus problem (19) possesses a unique optimal $(\bar{u}, \theta)$ if feasible, although optimal $p_{i,t}$ may not be unique. Problem (19) can be efficiently solved by the Iterative Risk Allocation method [44]; see [46, Appendix B] for more details of our implementation.

The nominal inputs $\bar{u}_{[k,k+N)}$ and interpolation variable θ determined from problem (19) complete the parameterization of the control policies $\pi_{[k,k+N)}$ in (10). The upcoming N_c control inputs $u_{[k,k+N_\mathrm{c})}$ are decided by the first N_c policies $\pi_{[k,k+N_\mathrm{c})}$ respectively, with a parameter $N_\mathrm{c} \in \mathbb{Z}_{[1,N]}$. Then, the next control step is set as $t = k + N_\mathrm{c}$. At the new control step, the initial condition $\mu^{\mathsf{x}}_{k+N_\mathrm{c}}$ interpolates between two fixed options $\mu^{\hat{\mathsf{x}}}_{k+N_\mathrm{c}}$ and $\mu^{\bar{\mathsf{x}}}_{k+N_\mathrm{c}}$ which are decided by

$$\mu^{\hat{\mathsf{x}}}_{k+N_\mathrm{c}} := \hat{x}^{-}_{k+N_\mathrm{c}}, \qquad \mu^{\bar{\mathsf{x}}}_{k+N_\mathrm{c}} := \bar{x}_{k+N_\mathrm{c}}, \tag{20}$$

as described in Section II-A-2. The entire SMPC control process is shown in Algorithm 1.

8) Closed-loop Properties: The investigated SMPC framework possesses recursive feasibility and closed-loop stability.

Lemma 1 (SMPC Recursive Feasibility)**.** *Assume $p \in (0, \frac{1}{2}]$. In Algorithm 1, if the problem* (19) *is feasible at control step $k = \kappa$, then it is feasible at next control step $k = \kappa + N_\mathrm{c}$.*

In Lemma 1, initial feasibility can be achieved via proper choice of the parameters $\mu^{\mathsf{x}}_{\mathsf{ini}}, E, f, p$. Closed-loop stability is reflected in the decrease of the optimal cost value and in a finite asymptotic cost [33], [34], [47], [48]. Let $\mathbb{E}_k[\cdot]$ denote the expectation given the initial condition (5) at control step k. Define $V_k^* := \Sigma_{t=k}^{k+N-1} J_t(u_t, y_t)$ the stochastic cost with the optimal nominal trajectory solved from problem (19).

Lemma 2 (SMPC Closed-loop Stability)**.** *Assume $\{z : Ez \leq f\}$ is a bounded set. Let system* (1) *be controlled by Algorithm 1, where problem* (19) *is assumed feasible at all control steps, and the reference signal $r_t = r$ is fixed. Then, the expectation of optimal cost values at consecutive control steps differ as*

$$\mathbb{E}_\kappa[V^*_{\kappa+N_\mathrm{c}} - V^*_\kappa] \leq -\textstyle\sum_{t=\kappa}^{\kappa+N_\mathrm{c}-1} \mathbb{E}_\kappa[J_t(u_t, y_t)] + N_\mathrm{c} c, \tag{21}$$

and the asymptotic expected cost is upper bounded as

$$\lim_{T\to\infty} \tfrac{1}{T} \textstyle\sum_{t=0}^{T-1} \mathbb{E}_0[J_t(u_t, y_t)] \leq c \tag{22}$$

for some $c > 0$.

The proofs are analogous to the proofs in [38] and available in the extended version [46, Appendices C and D].

Algorithm 1 a Framework of Stochastic MPC (SMPC)

Input: horizon lengths L, N, N_c, system matrices A, B, C, D, noise variances $\Sigma^{\mathsf{w}}, \Sigma^{\mathsf{v}}$, initial state mean $\mu^{\mathsf{x}}_{\mathsf{ini}}$, cost matrices Q, R, constraint coefficients E, f, probability bound p, interpolation penalty coefficient λ_θ.

Compute Kalman gain L_{K} via (6b), feedback gain K via (11), and covariance matrices $\Delta_{[0,N)}$ via (14).
Initialize the control step $k \leftarrow 0$ and set the initial condition $\mu^{\hat{\mathsf{x}}}_0 \leftarrow \mu^{\mathsf{x}}_{\mathsf{ini}}$ and $\mu^{\bar{\mathsf{x}}}_0 \leftarrow \mu^{\mathsf{x}}_{\mathsf{ini}}$.
while `true` **do**
  Solve $\bar{u}_{[k,k+N)}$ and θ from problem (19).
  Obtain μ^{x}_k via (9) and obtain $\bar{x}_{[k,k+N]}$ via (7).
  Obtain policies $\pi_{[k,k+N)}$ from (10).
  for t **from** k **to** $k + N_\mathrm{c} - 1$ **do**
    Compute $\hat{x}^{-}_t$ via (8).
    Input $u_t \leftarrow \pi_t(\hat{x}^{-}_t)$ to the system (1).
    Measure y_t from the system (1).
  Set $\mu^{\hat{\mathsf{x}}}_{k+N_\mathrm{c}} \leftarrow \hat{x}^{-}_{k+N_\mathrm{c}}$ and $\mu^{\bar{\mathsf{x}}}_{k+N_\mathrm{c}} \leftarrow \bar{x}_{k+N_\mathrm{c}}$ as (20).
  Set $k \leftarrow k + N_\mathrm{c}$.

B. Our Objective: An Equivalent Data-Driven Method

Direct data-driven methods such as DeePC and SPC for deterministic systems collect a sufficiently long, persistently exciting, noise-free input–output dataset. Under standard conditions, this dataset represents the system dynamics and can replace the parametric model in predictive control, yielding controllers that are equivalent to model-based MPC [11], [12]. Our aim is to construct an analogous direct data-driven method that, for the same system (1), initial condition x_0, and noise realizations w_t, v_t, generates exactly the same input–state-output trajectories as Algorithm 1; in short, a direct data-driven counterpart of SMPC. As in DeePC and SPC, we will establish this equivalence in an idealized setting with access to noise-free offline data. This assumption is only to facilitate a rigorous theory, and is not an intrinsic requirement of the design method. Although such an assumption may appear at first unusual in a stochastic framework, it is the natural target when the prediction model must be replaced using only a finite amount of recorded data. Noisy offline data can instead be treated robustly using regularized least-squares (Section III-A), as illustrated in Section IV, and the online stochastic controller still fully accounts for process and measurement noise.

III. Stochastic Data-Driven Predictive Control

This section develops a data-driven control method whose performance will be shown to be equivalent to SMPC under certain tuning conditions. In the spirit of DeePC and SPC, the control procedure consists of an offline process, where data is collected and used for system representation, and an online process which controls the system. The technical approach has three steps. First, we collect offline input-output data (Section III-A), and use this offline data to parameterize an auxiliary model (Section III-B-1). Second, we formulate a stochastic predictive control method using the auxiliary model (Section III-B, Section III-C-1, Section III-D-1). Third and finally, we establish theoretical equivalences between the model-based and data-based control methods (Section III-C-2, Section III-D-2).

A. Use of Offline Data

1) Deterministic Offline Data: Consider the deterministic version of system (1), reproduced for convenience as

$$x_{t+1} = Ax_t + Bu_t, \qquad y_t = Cx_t + Du_t, \tag{23}$$

where with a slight abuse of notation, we temporarily in this section let x_t and y_t denote the state and output of system (23). By assumption, (23) is minimal; recall $L \in \mathbb{N}$ in Section II such that $\mathcal{O} := \mathrm{col}(C, CA, \dots, CA^{L-1})$ has full column rank. Let $u^{\mathrm{d}}_{[1,T_{\mathrm{d}}]}, y^{\mathrm{d}}_{[1,T_{\mathrm{d}}]}$ be a T_{d}-length trajectory of input-output data collected from (23). The input sequence $u^{\mathrm{d}}_{[1,T_{\mathrm{d}}]}$ is assumed to be *persistently exciting* of order $K_{\mathrm{d}} := L + n + 1$, i.e., its associated K_{d}-depth block-Hankel matrix $\mathcal{H}_{K_{\mathrm{d}}}(u^{\mathrm{d}}_{[1,T_{\mathrm{d}}]})$, defined as

$$\mathcal{H}_{K_{\mathrm{d}}}(u^{\mathrm{d}}_{[1,T_{\mathrm{d}}]}) := \begin{bmatrix} u^{\mathrm{d}}_1 & u^{\mathrm{d}}_2 & \cdots & u^{\mathrm{d}}_{T_{\mathrm{d}}-K_{\mathrm{d}}+1} \\ u^{\mathrm{d}}_2 & u^{\mathrm{d}}_3 & \cdots & u^{\mathrm{d}}_{T_{\mathrm{d}}-K_{\mathrm{d}}+2} \\ \vdots & \vdots & \ddots & \vdots \\ u^{\mathrm{d}}_{K_{\mathrm{d}}} & u^{\mathrm{d}}_{K_{\mathrm{d}}+1} & \cdots & u^{\mathrm{d}}_{T_{\mathrm{d}}} \end{bmatrix},$$

has full row rank. To achieve persistent excitation, one must collect at least $T_{\mathrm{d}} \geq (m+1)K_{\mathrm{d}} - 1$ data samples [12]. We formulate data matrices $U_1 \in \mathbb{R}^{mL \times h}$, $U_2 \in \mathbb{R}^{m \times h}$, $Y_1 \in \mathbb{R}^{pL \times h}$ and $Y_2 \in \mathbb{R}^{p \times h}$ of a common width $h := T_{\mathrm{d}} - L$ by partitioning associated Hankel matrices as

$$\begin{aligned} \mathrm{col}(U_1, U_2) &:= \mathcal{H}_{L+1}\big(u^{\mathrm{d}}_{[1,T_{\mathrm{d}}]}\big), \\ \mathrm{col}(Y_1, Y_2) &:= \mathcal{H}_{L+1}\big(y^{\mathrm{d}}_{[1,T_{\mathrm{d}}]}\big). \end{aligned} \tag{24}$$

The data matrices in (24) will now be used to represent some quantities related to the system (23). Before stating the result, we introduce some additional notation. Define a system-related matrix $\mathbf{\Gamma} \in \mathbb{R}^{p \times (m+p)L}$ as

$$\mathbf{\Gamma} = \begin{bmatrix} \mathbf{\Gamma}_{\mathsf{U}} & \mathbf{\Gamma}_{\mathsf{Y}} \end{bmatrix} := \begin{bmatrix} C\mathcal{C} & CA^L \end{bmatrix} \begin{bmatrix} I_{mL} & \\ \mathcal{G} & \mathcal{O} \end{bmatrix}^{\dagger}. \tag{25}$$

with sub-blocks $\mathbf{\Gamma}_{\mathsf{U}} \in \mathbb{R}^{p \times mL}$ and $\mathbf{\Gamma}_{\mathsf{Y}} \in \mathbb{R}^{p \times pL}$, where $\mathcal{C} := [A^{L-1}B, \dots, AB, B]$ is the extended (reversed) controllability matrix, and $\mathcal{G} \in \mathbb{R}^{pL \times mL}$ is an impulse-response matrix

$$\mathcal{G} := \begin{bmatrix} D & & & \\ CB & D & & \\ \vdots & \ddots & \ddots & \\ CA^{L-2}B & \cdots & CB & D \end{bmatrix}. \tag{26}$$

The following result provides expressions for the quantity $\mathbf{\Gamma}$ and the system matrix D in terms of raw data.

Lemma 3 (Data Representation of Model Quantities). *Given the data matrices in* (24)*, if system* (23) *is controllable and the input data $u^{\mathrm{d}}_{[1,T_{\mathrm{d}}]}$ is persistently exciting of order $L+n+1$, then the matrix $\mathbf{\Gamma}$ defined in* (25) *and the matrix D in the model* (23) *can be expressed as*

$$[\mathbf{\Gamma}_{\mathsf{U}}, \mathbf{\Gamma}_{\mathsf{Y}}, D] = Y_2 \,\mathrm{col}(U_1, Y_1, U_2)^{\dagger}.$$

Proof. See Appendix A. ∎

The data-expression of impulse response, e.g., D and $\mathcal{G}$, is present in SPC literature [11]. The novelty of Lemma 3 is the data-based representation of $\mathbf{\Gamma}$, which will be used as part of the construction for our data-driven control method.

2) The Case of Stochastic Offline Data: Lemma 3 applies to noise-free data. With noisy measurements, the pseudoinverse computations in Lemma 3 become fragile and do not recover the desired matrices $\mathbf{\Gamma}$ and D. A standard technique to robustify these computations is to replace the pseudoinverse $W^{\dagger}$ of $W := \mathrm{col}(U_1, Y_1, U_2)$ in Lemma 3 with its Tikhonov regularization $W^{\mathsf{tik}} := (W^{\mathsf{T}}W + \lambda I_h)^{-1}W^{\mathsf{T}}$ where $\lambda > 0$ is the regularization parameter. To interpret this, recall that $\mathcal{P} := Y_2 W^{\dagger}$ is a least-square solution to $\mathrm{argmin}_{\mathcal{P}} \|Y_2 - \mathcal{P}W\|_{\mathsf{F}}^2$. Correspondingly, the regularization $Y_2 W^{\mathsf{tik}}$ is the solution to a ridge-regression problem $\mathrm{argmin}_{\mathcal{P}} \|Y_2 - \mathcal{P}W\|_{\mathsf{F}}^2 + \lambda \|\mathcal{P}\|_{\mathsf{F}}^2$, which gives a maximum-likelihood or worst-case robust solution to the original least-square problem $\mathrm{argmin}_{\mathcal{P}} \|Y_2 - \mathcal{P}W\|_{\mathsf{F}}^2$ whose multiplicative parameter W has uncertain entries; see [8] sidebar "Roles of Regularization" and references therein for more details. Following this standard technique, in the stochastic case, we estimate matrices $\mathbf{\Gamma}$ and D by applying Lemma 3 with $\mathcal{P} = Y_2 W^{\dagger}$ replaced by $\widehat{\mathcal{P}} := Y_2 W^{\mathsf{tik}}$.

B. Data-Driven State Estimation and Output Feedback

The SMPC approach of Section II-A uses as sub-components a state estimator and an affine feedback law. We now leverage the offline data as described in Section III-A to directly design analogs of these components based on data.

1) Auxiliary State-Space Model: We begin by constructing an auxiliary state-space model which has equivalent input-output behavior to (1), but is parameterized only by the recorded data sequences of Section III-A. Define auxiliary signals $\mathbf{x}_t, \mathbf{w}_t \in \mathbb{R}^{n_{\mathsf{aux}}}$ of dimension $n_{\mathsf{aux}} := mL + pL + pL^2$ for system (1) by

$$\mathbf{x}_t := \begin{bmatrix} u_{[t-L,t)} \\ \hline y^{\circ}_{[t-L,t)} \\ \hline \rho_{[t-L,t)} \end{bmatrix}, \qquad \mathbf{w}_t := \begin{bmatrix} 0_{mL \times 1} \\ \hline 0_{pL \times 1} \\ \hline 0_{pL(L-1) \times 1} \\ \rho_t \end{bmatrix} \tag{27}$$

where $y^{\circ}_t := y_t - v_t \in \mathbb{R}^p$ is the output excluding measurement noise, and $\rho_t := \mathcal{O} w_t \in \mathbb{R}^{pL}$ stacks the system's response to process noise w_t on time interval $[t+1, t+L]$. The construction of the auxiliary state $\mathbf{x}_t$ was inspired by [49]. The auxiliary signals $\mathbf{x}_t, \mathbf{w}_t$ together with u_t, y_t, v_t then satisfy the relations given by Lemma 4.

Lemma 4 (Auxiliary Model)**.** *For system* (1)*, the signals* u_t, y_t, v_t *and the auxiliary signals* $\mathbf{x}_t, \mathbf{w}_t$ *in* (27) *satisfy*

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{B}u_t + \mathbf{w}_t, \tag{28a}$$
$$y_t = \mathbf{C}\mathbf{x}_t + Du_t + v_t, \tag{28b}$$

with $\mathbf{A} \in \mathbb{R}^{n_{\mathsf{aux}} \times n_{\mathsf{aux}}}$, $\mathbf{B} \in \mathbb{R}^{n_{\mathsf{aux}} \times m}$, $\mathbf{C} \in \mathbb{R}^{p \times n_{\mathsf{aux}}}$ *given by*

$$\mathbf{A} := \left[\begin{array}{c|c|c} \begin{matrix} 0_{m\times m} \\ 0 \\ \mathbf{\Gamma}_{\mathsf{U}} \end{matrix} & \begin{matrix} I_{m(L-1)} \\ 0 \quad 0 \\ \mathbf{\Gamma}_{\mathsf{Y}} \end{matrix} & \begin{matrix} 0 \\ 0 \\ \mathbf{F} - \mathbf{\Gamma}_{\mathsf{Y}}\mathbf{E} \end{matrix} \\ \hline 0 & 0 & \begin{matrix} I_{pL(L-1)} \\ 0_{pL\times pL} \end{matrix} \end{array}\right]$$

$$\mathbf{B} := \left[\begin{array}{c} 0_{m(L-1)\times m} \\ \hline I_m \\ \hline 0_{p(L-1)\times m} \\ \hline D \\ \hline 0_{pL^2\times m} \end{array}\right], \quad \mathbf{C} := \left[\begin{array}{c|c|c} \mathbf{\Gamma}_{\mathsf{U}} & \mathbf{\Gamma}_{\mathsf{Y}} & \mathbf{F} - \mathbf{\Gamma}_{\mathsf{Y}}\mathbf{E} \end{array}\right]$$

with matrices $\mathbf{\Gamma}_{\mathsf{U}}, \mathbf{\Gamma}_{\mathsf{Y}}$ *in* (25)*, the same matrix* D *in* (1)*, and zero-one matrices* $\mathbf{E} \in \mathbb{R}^{pL\times pL^2}$ *and* $\mathbf{F} \in \mathbb{R}^{p\times pL^2}$ *composed by selection matrices* $S_j := [0_{p\times(j-1)p}, I_p, 0_{p\times(L-j)p}] \in \mathbb{R}^{p\times pL}$ *for* $j \in \{1, \dots, L\}$ *as*

$$\left[\begin{array}{c} \mathbf{E} \\ \hline \mathbf{F} \end{array}\right] := \left[\begin{array}{cccc} 0_{p\times pL} & & & \\ S_1 & 0_{p\times pL} & & \\ \vdots & \ddots & \ddots & \\ S_{L-1} & \cdots & S_1 & 0_{p\times pL} \\ \hline S_L & \cdots & S_2 & S_L \end{array}\right].$$

Proof. See Appendix B. ∎

The output noise signal v_t in (28) is precisely the same as in (1); the signal $\mathbf{w}_t$ appears now as a new disturbance; $\mathbf{w}_t$ and v_t are independent and follow the i.i.d. zero-mean normal distributions

$$\mathbf{w}_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0_{n_{\mathsf{aux}}\times 1}, \mathbf{\Sigma}^{\mathsf{w}}), \quad v_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0_{p\times 1}, \Sigma^{\mathsf{v}}) \tag{29}$$

with variances $\mathbf{\Sigma}^{\mathsf{w}} \in \mathbb{S}^{n_{\mathsf{aux}}}_{+}$ and $\Sigma^{\mathsf{v}} \in \mathbb{S}^{p}_{++}$,

$$\mathbf{\Sigma}^{\mathsf{w}} := \mathrm{Diag}(0_{(n_{\mathsf{aux}}-pL)\times(n_{\mathsf{aux}}-pL)}, \Sigma^{\rho}) \tag{30}$$

where $\Sigma^{\rho} := \mathcal{O}\Sigma^{\mathsf{w}}\mathcal{O}^{\mathsf{T}} \in \mathbb{S}^{pL}_{+}$ is the variance of ρ_t. The matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}, D$ are known given offline data described in Section III-A, since they by definition only depend on matrices $\mathbf{\Gamma}$ and D which are data-representable via Lemma 3. Hence, the auxiliary model (28) can be interpreted as a non-minimal data-representable realization of system (1). Nonetheless, the model is indeed stabilizable and detectable.

Lemma 5. *For the auxiliary model* (28) *and matrix* $\mathbf{\Sigma}^{\mathsf{w}}$ *in* (30)*, the pairs* $(\mathbf{A}, \mathbf{B})$ *and* $(\mathbf{A}, \mathbf{\Sigma}^{\mathsf{w}})$ *are stabilizable and the pair* $(\mathbf{A}, \mathbf{C})$ *is detectable.*

Proof. See Appendix C. ∎

2) Auxiliary State Initial Condition: Suppose we are at a control step $t = k$ in a receding-horizon process. Similar to (5), we model the auxiliary state $\mathbf{x}_k$ from (28) following a prior distribution

$$\mathbf{x}_k \sim \mathcal{N}(\boldsymbol{\mu}^{\mathsf{x}}_k, \mathbf{\Sigma}^{\mathsf{x}}). \tag{31}$$

The mean $\boldsymbol{\mu}^{\mathsf{x}}_k \in \mathbb{R}^{n_{\mathsf{aux}}}$ in (31) interpolates between two fixed vectors $\boldsymbol{\mu}^{\hat{\mathsf{x}}}_k, \boldsymbol{\mu}^{\bar{\mathsf{x}}}_k \in \mathbb{R}^{n_{\mathsf{aux}}}$ with a decision variable $\theta \in [0, 1]$,

$$\boldsymbol{\mu}^{\mathsf{x}}_k := (1-\theta)\,\boldsymbol{\mu}^{\hat{\mathsf{x}}}_k + \theta\,\boldsymbol{\mu}^{\bar{\mathsf{x}}}_k \tag{32}$$

wherein $\boldsymbol{\mu}^{\hat{\mathsf{x}}}_k$ and $\boldsymbol{\mu}^{\bar{\mathsf{x}}}_k$ are produced by a state estimator (see (35)) and a noise-free model (see (36)), respectively, of last control step. At initial time $k = 0$, the initial state mean $\boldsymbol{\mu}^{\mathsf{x}}_0$ is given as a parameter $\boldsymbol{\mu}^{\mathsf{x}}_{\mathsf{ini}} \in \mathbb{R}^{n_{\mathsf{aux}}}$, i.e., we let $\boldsymbol{\mu}^{\hat{\mathsf{x}}}_0 := \boldsymbol{\mu}^{\mathsf{x}}_{\mathsf{ini}}$ and $\boldsymbol{\mu}^{\bar{\mathsf{x}}}_0 := \boldsymbol{\mu}^{\mathsf{x}}_{\mathsf{ini}}$. The variance $\mathbf{\Sigma}^{\mathsf{x}} \in \mathbb{S}^{n_{\mathsf{aux}}}_{+}$ in (31) is fixed as the unique positive semidefinite solution to the DARE (33a),

$$\mathbf{\Sigma}^{\mathsf{x}} = (\mathbf{A} - \mathbf{L}_{\mathsf{L}}\mathbf{C})\mathbf{\Sigma}^{\mathsf{x}}\mathbf{A}^{\mathsf{T}} + \mathbf{\Sigma}^{\mathsf{w}} \tag{33a}$$
$$\mathbf{L}_{\mathsf{L}} := \mathbf{A}\mathbf{L}_{\mathsf{K}}, \quad \mathbf{L}_{\mathsf{K}} := \mathbf{\Sigma}^{\mathsf{x}}\mathbf{C}^{\mathsf{T}}(\mathbf{C}\mathbf{\Sigma}^{\mathsf{x}}\mathbf{C}^{\mathsf{T}} + \Sigma^{\mathsf{v}})^{-1} \tag{33b}$$

given detectable $(\mathbf{A}, \mathbf{C})$ and stabilizable $(\mathbf{A}, \mathbf{\Sigma}^{\mathsf{w}})$ via Lemma 5, where we define Kalman gain $\mathbf{L}_{\mathsf{K}} \in \mathbb{R}^{n_{\mathsf{aux}}\times p}$ and Luenberger observer gain $\mathbf{L}_{\mathsf{L}} \in \mathbb{R}^{n_{\mathsf{aux}}\times p}$. Not surprisingly, there is a close relationship between the distributions of x_k and $\mathbf{x}_k$.

Lemma 6 (Related Means of x_k and $\mathbf{x}_k$)**.** *For system* (1) *and auxiliary state* $\mathbf{x}_t$ *in* (27)*, if* μ_k *is the mean of* x_k *and* $\boldsymbol{\mu}_k$ *is the mean of* $\mathbf{x}_k$*, then we have*

$$\mu_k = \Phi_{\mathsf{orig}}\,\widetilde{\mu}_k, \qquad \boldsymbol{\mu}_k = \Phi_{\mathsf{aux}}\,\widetilde{\mu}_k, \tag{34}$$

for some $\widetilde{\mu}_k \in \mathbb{R}^{mL+n(L+1)}$*, with the matrices* Φ_{orig} *and* Φ_{aux} *defined in Appendix B.*

Proof. Given $x_k = \Phi_{\mathsf{orig}}\,\xi_k$ and $\mathbf{x}_k = \Phi_{\mathsf{aux}}\,\xi_k$ via Claim 4.1, we have (34) by choosing $\widetilde{\mu}_k$ as the mean of ξ_k. ∎

The result (34) will be leveraged in establishing equivalence between SMPC and our proposed method, as we will see in (44) in Proposition 7 and (d) in Assumption 8.

3) Auxiliary State Estimation and Feedback: The auxiliary model (28) will now be used for both estimation and control purposes. Analogous to the Kalman filter (8), we formulate a Kalman filter for the auxiliary model (28) as

$$\hat{\mathbf{x}}^{+}_t := \hat{\mathbf{x}}^{-}_t + \mathbf{L}_{\mathsf{K}}(y_t - \mathbf{C}\hat{\mathbf{x}}^{-}_t - Du_t), \quad t \in \mathbb{Z}_{[k,k+N)} \tag{35a}$$
$$\hat{\mathbf{x}}^{-}_{t+1} := \mathbf{A}\hat{\mathbf{x}}^{+}_t + \mathbf{B}u_t, \quad t \in \mathbb{Z}_{[k,k+N)} \tag{35b}$$
$$\hat{\mathbf{x}}^{-}_k := \boldsymbol{\mu}^{\mathsf{x}}_k \tag{35c}$$

where $\hat{\mathbf{x}}^{+}_t$ and $\hat{\mathbf{x}}^{-}_t$ are the posterior and prior estimates of $\mathbf{x}_t$, respectively, with $\mathbf{L}_{\mathsf{K}} \in \mathbb{R}^{n_{\mathsf{aux}}\times p}$ in (33b). A noise-free model can be formed similarly as (7), given initial condition (31),

$$\bar{\mathbf{x}}_{t+1} := \mathbf{A}\bar{\mathbf{x}}_t + \mathbf{B}\bar{u}_t, \quad t \in \mathbb{Z}_{[k,k+N)} \tag{36a}$$
$$\bar{\mathbf{y}}_t := \mathbf{C}\bar{\mathbf{x}}_t + D\bar{u}_t, \quad t \in \mathbb{Z}_{[k,k+N)} \tag{36b}$$
$$\bar{\mathbf{x}}_k := \boldsymbol{\mu}^{\mathsf{x}}_k, \tag{36c}$$

where $\bar{u}_t \in \mathbb{R}^m$ is the nominal input decided through optimization, and $\bar{\mathbf{x}}_t \in \mathbb{R}^{n_{\mathsf{aux}}}$ and $\bar{\mathbf{y}}_t \in \mathbb{R}^p$ are the resulting nominal state and output, respectively. The affine output feedback policy (10) from SMPC is now extended as $\boldsymbol{\pi}_t(\cdot)$,

$$u_t \leftarrow \boldsymbol{\pi}_t(\hat{\mathbf{x}}^{-}_t) := \bar{u}_t - \mathbf{K}(\hat{\mathbf{x}}^{-}_t - \bar{\mathbf{x}}_t) \tag{37}$$

where as in (11) the feedback gain $\mathbf{K} \in \mathbb{R}^{m\times n_{\mathsf{aux}}}$ is determined through LQR-based design, where

$$\mathbf{K} := (R + \mathbf{B}^{\mathsf{T}}\mathbf{P}\mathbf{B} + D^{\mathsf{T}}QD)^{-1}(\mathbf{B}^{\mathsf{T}}\mathbf{P}\mathbf{A} + D^{\mathsf{T}}Q\mathbf{C}), \tag{38}$$

and where $\mathbf{P} \in \mathbb{S}^{n_{\mathsf{aux}}}_{+}$ is the unique positive semidefinite solution, given the stabilizability of $(\mathbf{A}, \mathbf{B})$ and detectability of $(\mathbf{A}, \mathbf{C})$ by Lemma 5, to the DARE

$$\mathbf{P} = \mathbf{A}^{\mathsf{T}}\mathbf{P}(\mathbf{A} - \mathbf{B}\mathbf{K}) + \mathbf{C}^{\mathsf{T}}Q(\mathbf{C} - D\mathbf{K}). \tag{39}$$

The state estimator and feedback policy designs (35) and (37) directly parallel the SMPC framework, providing a clear bridge between the model-based and data-driven control settings. While ultimately these state estimates are internal variables within the subsequent optimization problem and may be eliminated, retaining this structure provides conceptual clarity and illustrates the potential to similarly obtain data-driven versions of other estimator-based SMPC schemes.

C. Optimization Problem

1) SDDPC Optimization Problem: With results of Section III-B, we are now ready to mirror the steps which led to (19) and formulate a Stochastic Data-Driven Predictive Control (SDDPC) optimization problem. First, following a similar process as that which led to (13), we may combine (28), (29), (31), (35), (36) and (37), to conclude that the input-output trajectory (u_t, y_t) for $t \in \mathbb{Z}_{[k,k+N)}$ is normally distributed as $\mathcal{N}(\mathrm{col}(\bar{u}_t, \bar{\mathbf{y}}_t), \boldsymbol{\Delta}_{t-k})$, where the covariance matrices $\boldsymbol{\Delta}_s \in \mathbb{S}_+^{m+p}$ for $s \in \mathbb{Z}_{[0,N)}$ are computed as (40a) using $\boldsymbol{\Lambda}_s \in \mathbb{S}_+^{n_{\mathsf{aux}}}$ defined as (40b),

$$\boldsymbol{\Delta}_s := \begin{bmatrix} -\mathbf{K} \\ \mathbf{C}-D\mathbf{K} \end{bmatrix} \boldsymbol{\Lambda}_s \begin{bmatrix} -\mathbf{K} \\ \mathbf{C}-D\mathbf{K} \end{bmatrix}^{\mathsf{T}} + \begin{bmatrix} 0_{m\times m} & \\ & \mathbf{C}\boldsymbol{\Sigma}^{\mathsf{x}}\mathbf{C}^{\mathsf{T}}+\Sigma^{\mathsf{v}} \end{bmatrix} \tag{40a}$$

$$\boldsymbol{\Lambda}_s := \textstyle\sum_{r=0}^{s-1} (\mathbf{A}-\mathbf{B}\mathbf{K})^r \mathbf{L}_{\mathsf{L}} (\mathbf{C}\boldsymbol{\Sigma}^{\mathsf{x}}\mathbf{C}^{\mathsf{T}}+\Sigma^{\mathsf{v}}) \mathbf{L}_{\mathsf{L}}^{\mathsf{T}} (\mathbf{A}-\mathbf{B}\mathbf{K})^{r\mathsf{T}} \tag{40b}$$

with $\mathbf{L}_{\mathsf{L}}$ in (33b) and $\mathbf{K}$ in (38). Then, the SDDPC problem for computing $\bar{u}$ and θ at control step $t = k$ is written as

$$\begin{aligned} &\underset{\bar{u},\, \theta,\, p_{i,t}}{\text{minimize}} \quad \textstyle\sum_{t=k}^{k+N-1} J_t(\bar{u}_t, \bar{\mathbf{y}}_t) + \lambda_\theta\, \theta \\ &\text{subject to} \quad (42) \text{ for } t \in \mathbb{Z}_{[k,k+N)},\ (32), (36), (43), \end{aligned} \tag{41}$$

with the safety constraint for $t \in \mathbb{Z}_{[k,k+N)}$,

$$\begin{aligned} &e_i^{\mathsf{T}} \begin{bmatrix} \bar{u}_t \\ \bar{\mathbf{y}}_t \end{bmatrix} \le f_i - \sqrt{e_i^{\mathsf{T}} \boldsymbol{\Delta}_{t-k}\, e_i}\, \mathrm{icdfn}(1-p_{i,t}), \quad i \in \mathbb{Z}_{[1,q]} \\ &\textstyle\sum_{i=1}^{q} p_{i,t} = p, \qquad\qquad p_{i,t} > 0, \quad i \in \mathbb{Z}_{[1,q]} \end{aligned} \tag{42}$$

and with the terminal equality constraint

$$\begin{aligned} \bar{u}_{k+N-L} &= \bar{u}_{k+N-L+1} = \cdots = \bar{u}_{k+N-1}, \\ \bar{\mathbf{y}}_{k+N-L} &= \bar{\mathbf{y}}_{k+N-L+1} = \cdots = \bar{\mathbf{y}}_{k+N-1}. \end{aligned} \tag{43}$$

Problem (41) not only mirrors problem (19) through the use of auxiliary model (28), but also introduces a novel formulation that explicitly delineates which quantities are replaced by their data-driven counterparts and which remain unchanged.

2) Equivalence to SMPC Optimization Problem: We now establish that the SDDPC problem (41) and the SMPC problem (19) have equal feasible sets and equal optimal sets, when the initial-condition parameters are related in the form of (34).

Proposition 7 (Equivalence of Optimization Problems)**.** *If the parameters $\mu_k^{\hat{\mathsf{x}}}, \mu_k^{\bar{\mathsf{x}}}, \boldsymbol{\mu}_k^{\hat{\mathsf{x}}}, \boldsymbol{\mu}_k^{\bar{\mathsf{x}}}$ satisfy*

$$\begin{aligned} \mu_k^{\hat{\mathsf{x}}} &= \Phi_{\mathsf{orig}}\, \widetilde{\mu}_k^{\hat{\mathsf{x}}}, \qquad \boldsymbol{\mu}_k^{\hat{\mathsf{x}}} = \Phi_{\mathsf{aux}}\, \widetilde{\mu}_k^{\hat{\mathsf{x}}}, \\ \mu_k^{\bar{\mathsf{x}}} &= \Phi_{\mathsf{orig}}\, \widetilde{\mu}_k^{\bar{\mathsf{x}}}, \qquad \boldsymbol{\mu}_k^{\bar{\mathsf{x}}} = \Phi_{\mathsf{aux}}\, \widetilde{\mu}_k^{\bar{\mathsf{x}}}, \end{aligned} \tag{44}$$

with the matrices $\Phi_{\mathsf{orig}}, \Phi_{\mathsf{aux}}$ defined in Appendix B, for some vectors $\widetilde{\mu}_k^{\hat{\mathsf{x}}}, \widetilde{\mu}_k^{\bar{\mathsf{x}}} \in \mathbb{R}^{mL+n(L+1)}$, then the optimal (resp. feasible) solution set of SDDPC problem (41) *is equal to the optimal (resp. feasible) solution set of SMPC problem* (19)*.*

Proof. See Appendix D. ∎

We conclude by noting that problem (41) produces a unique optimal $(\bar{u}, \theta)$ when feasible, following from Proposition 7 and the fact that problem (19) gives a unique optimal $(\bar{u}, \theta)$ when it is feasible, as mentioned in Section II-A.

D. Online Control Algorithm

1) SDDPC Control Algorithm: We now describe the online implementation of our SDDPC. At control step $t = k$, the nominal input $\bar{u}_{[k,k+N)}$ and interpolation variable θ are solved from (41), and the policies $\boldsymbol{\pi}_{[k,k+N)}$ are constructed via (37), where the first N_{c} policies are implemented. At the next control step $t = k + N_{\mathrm{c}}$, the initial condition interpolates as in (32) between vectors $\boldsymbol{\mu}_{k+N_{\mathrm{c}}}^{\hat{\mathsf{x}}}$ and $\boldsymbol{\mu}_{k+N_{\mathrm{c}}}^{\bar{\mathsf{x}}}$ decided as

$$\boldsymbol{\mu}_{k+N_{\mathrm{c}}}^{\hat{\mathsf{x}}} := \hat{\mathbf{x}}_{k+N_{\mathrm{c}}}^{-}, \qquad \boldsymbol{\mu}_{k+N_{\mathrm{c}}}^{\bar{\mathsf{x}}} := \bar{\mathbf{x}}_{k+N_{\mathrm{c}}}. \tag{45}$$

The method is formally summarized in Algorithm 2. Note that Algorithm 2 is for control of system (1), although the auxiliary model is used in its design. Algorithm 2 presents a novel control scheme, with analogy to Algorithm 1, where some components are replaced by data-driven counterparts.

Algorithm 2 Stochastic Data-Driven Predictive Control (SDDPC)

Input: horizon lengths L, N, N_{c}, offline data $u^{\mathrm{d}}, y^{\mathrm{d}}$, noise variances $\Sigma^{\rho}, \Sigma^{\mathsf{v}}$, initial-state mean $\boldsymbol{\mu}_{\mathsf{ini}}^{\mathsf{x}}$, cost matrices Q, R, constraint coefficients E, f, probability bound p, interpolation penalty coefficient λ_θ.

Compute matrices $\boldsymbol{\Gamma}$ and D as in Section III-A using data $u^{\mathrm{d}}, y^{\mathrm{d}}$, and formulate matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}$ as in Section III-B.
Compute Kalman gain $\mathbf{L}_{\mathsf{K}}$ via (33b), feedback gain $\mathbf{K}$ via (38), and covariance matrices $\boldsymbol{\Delta}_{[0,N)}$ via (40).
Initialize the control step $k \leftarrow 0$ and set the initial condition $\boldsymbol{\mu}_0^{\hat{\mathsf{x}}} \leftarrow \boldsymbol{\mu}_{\mathsf{ini}}^{\mathsf{x}}$ and $\boldsymbol{\mu}_0^{\bar{\mathsf{x}}} \leftarrow \boldsymbol{\mu}_{\mathsf{ini}}^{\mathsf{x}}$.
while `true` **do**
    Solve $\bar{u}_{[k,k+N)}$ and θ from problem (41).
    Obtain $\boldsymbol{\mu}_k^{\mathsf{x}}$ via (32) and obtain $\bar{\mathbf{x}}_{[k,k+N]}$ via (36).
    Obtain policies $\boldsymbol{\pi}_{[k,k+N)}$ from (37).
    for t **from** k **to** $k + N_{\mathrm{c}} - 1$ **do**
        Compute $\hat{\mathbf{x}}_t^{-}$ via (35).
        Input $u_t \leftarrow \boldsymbol{\pi}_t(\hat{\mathbf{x}}_t^{-})$ to the system (1).
        Measure y_t from the system (1).
    Set $\boldsymbol{\mu}_{k+N_{\mathrm{c}}}^{\hat{\mathsf{x}}} \leftarrow \hat{\mathbf{x}}_{k+N_{\mathrm{c}}}^{-}$ and $\boldsymbol{\mu}_{k+N_{\mathrm{c}}}^{\bar{\mathsf{x}}} \leftarrow \bar{\mathbf{x}}_{k+N_{\mathrm{c}}}$ as (45).
    Set $k \leftarrow k + N_{\mathrm{c}}$.

2) Closed-loop Properties of SDDPC: Similar to Lemma 1 and Lemma 2, Algorithm 2 possesses recursive feasibility and closed-loop stability, as formally stated below.

Corollary 1.1 (SDDPC Recursive feasibility)**.** *Assume $p \in (0, \frac{1}{2}]$. In Algorithm 2, if the problem* (41) *is feasible at control step $k = \kappa$, then it is feasible at next control step $k = \kappa + N_{\mathrm{c}}$.*

Corollary 2.1 (SDDPC Closed-loop Stability)**.** *Assume $\{z : Ez \le f\}$ is a bounded set. Let system* (1) *be controlled by Algorithm 2, where problem* (41) *is assumed feasible at all control steps, and the reference signal $r_t = r$ is fixed. Then, the expectation of optimal cost values at consecutive control steps differ as* (21)*, and the asymptotic expected cost is upper bounded as* (22) *with some $c \ge 0$.*

The proofs of the above corollaries are analogies to the proofs of Lemma 1 and Lemma 2, respectively, where the auxiliary model (28) is considered in place of model (1).

3) Equivalence to SMPC Algorithm: We present in Theorem 9 our main result, which says that under idealized conditions, our proposed SDDPC control method and the benchmark SMPC method will result in identical control actions.

Assumption 8 (SDDPC Parameter Choice w.r.t. SMPC)**.** Given the parameters in Algorithm 1, we assume the parameters in Algorithm 2 satisfy the following.

(a) L is sufficiently large so that $\mathcal{O}$ has full column rank.
(b) Data $u^{\mathrm{d}}, y^{\mathrm{d}}$ comes from the deterministic system (23), and input data u^{d} is persistently exciting of order $L+n+1$.
(c) Given Σ^{w} in Algorithm 1, the parameter Σ^{ρ} in Algorithm 2 is set equal to $\mathcal{O}\Sigma^{\mathsf{w}}\mathcal{O}^{\mathsf{T}}$.
(d) Given $\mu^{\mathsf{x}}_{\mathsf{ini}}$ in Algorithm 1, the parameter $\boldsymbol{\mu}^{\mathsf{x}}_{\mathsf{ini}}$ in Algorithm 2 is selected as $\Phi_{\mathsf{aux}}\widetilde{\mu}^{\mathsf{x}}_{\mathsf{ini}}$ for some $\widetilde{\mu}^{\mathsf{x}}_{\mathsf{ini}} \in \mathbb{R}^{mL+(n+1)L}$ satisfying $\mu^{\mathsf{x}}_{\mathsf{ini}} = \Phi_{\mathsf{orig}}\widetilde{\mu}^{\mathsf{x}}_{\mathsf{ini}}$, with matrices $\Phi_{\mathsf{orig}}, \Phi_{\mathsf{aux}}$ defined in Appendix B. (Such $\widetilde{\mu}^{\mathsf{x}}_{\mathsf{ini}}$ always exists because Φ_{orig} has full row rank.)

Theorem 9 (Equivalence of SMPC and SDDPC). *Consider the stochastic system* (1) *with a specific initial state* x_0 *and a specific noise realization* $\{w_t, v_t\}_{t=0}^{\infty}$*, and consider the following two control processes:*

a) decide control actions $\{u_t\}_{t=0}^{\infty}$ *by executing Algorithm 1;*
b) decide control actions $\{u_t\}_{t=0}^{\infty}$ *by executing Algorithm 2, where the parameters satisfy Assumption 8.*

Then, the state-input-output trajectories $\{x_t, u_t, y_t\}_{t=0}^{\infty}$ *resulting from process a) and from process b) are the same.*

Proof. See Appendix E. ∎

Theorem 9 should be interpreted as equivalence between SMPC and SDDPC in the idealized setting. Specifically, it establishes that if the proposed SDDPC algorithm is provided with noise-free offline data, if the initial conditions set within SMPC and SDDPC match, and if the process noise variance Σ^{ρ} in the algorithm is set in a specific idealized fashion relative to the original process noise variance Σ^{w}, then the method will produce identical results to those obtained by applying SMPC. While in practice these assumptions will not hold, noisy offline data can be accommodated as discussed in Section III-A, and Σ^{ρ} becomes a tuning parameter of our SDDPC method.

IV. Numerical Case Study

We numerically test our proposed method on the nonlinear grid-connected power converter system from [21], shown in Fig. 1, and compare the results with those of several benchmark model-based and data-based techniques.

The AC grid in the power part of Fig. 1 is modeled as an infinite bus with fixed voltage (1 p.u.) and fixed frequency (1 p.u.). This model has $n = 6$ states, $m = 3$ inputs and $p = 3$ outputs. The inputs are the angular frequency correction $\Delta\omega$ and current references $I_{\mathrm{d}}^{\mathrm{ref}}$ and $I_{\mathrm{q}}^{\mathrm{ref}}$ of d- and q-axes, respectively. The outputs to be controlled are the q-axis voltage V_{q}, the active power P_{E} and the reactive power Q_{E}. The LCL-filter parameters and the PI parameters in Fig. 1 are consistent

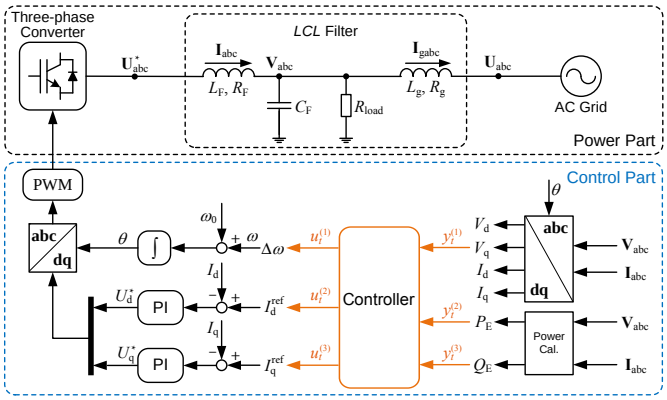

Fig. 1. The one-line diagram of a grid-connected power converter [21, Fig. 1].

with [21], whereas we choose the load resistance R_{load} as a Gaussian signal with mean 4 p.u. and noise power 10^{-3} p.u., which introduces process noise. The measurement noise on each output is Gaussian with noise power 10^{-7} p.u.

A. Benchmark Control Methods

1) Stochastic MPC and (Deterministic) MPC: We test two model-based methods: SMPC and (deterministic) MPC. For both SMPC and MPC, we use an identified system model in place of the true model (A, B, C, D), through N4SID system identification method [50] using offline data $u^{\mathrm{d}}, y^{\mathrm{d}}$ collected from the system. MPC follows a similar receding-horizon control process as SMPC, but the control action u_t is the decision $\bar{u}_t$, and the deterministic safety constraint

$$E\,\mathrm{col}(\bar{u}_t, \bar{y}_t) \le f \tag{46}$$

is used in place of (17).

2) DeePC and SPC: We test ℓ_2-regularized DeePC [21] and regularized SPC [11] as benchmark data-driven methods. In DeePC and SPC, the decisions $\bar{u}_t$ of optimization are applied as control actions u_t, and the deterministic constraint (46) is considered. Using offline data $u^{\mathrm{d}}, y^{\mathrm{d}}$, we formulate data Hankel matrices $U_{\mathrm{p}}, U_{\mathrm{f}}, Y_{\mathrm{p}}, Y_{\mathrm{f}}$ similar to matrices U_1, U_2, Y_1, Y_2 in (24), but matrices $U_{\mathrm{p}}, U_{\mathrm{f}}, Y_{\mathrm{p}}, Y_{\mathrm{f}}$ have mL, mN, pL, pN rows respectively. The DeePC problem at control step $t = k$ is

$$\begin{aligned} \underset{g,\,\sigma_{\mathrm{y}}}{\text{minimize}} \quad & \textstyle\sum_{t=k}^{k+N-1} J_t(\bar{u}_t, \bar{y}_t) + \lambda_{\mathrm{y}}\|\sigma_{\mathrm{y}}\|_2^2 + \lambda_{\mathrm{g}}\|g\|_2^2 \\ \text{subject to} \quad & \begin{bmatrix} U_{\mathrm{p}} \\ Y_{\mathrm{p}} \end{bmatrix} g = \begin{bmatrix} u_{\mathrm{ini}} \\ y_{\mathrm{ini}} + \sigma_{\mathrm{y}} \end{bmatrix}, \quad \begin{bmatrix} \bar{u}_{[k,k+N)} \\ \bar{y}_{[k,k+N)} \end{bmatrix} := \begin{bmatrix} U_{\mathrm{f}} \\ Y_{\mathrm{f}} \end{bmatrix} g \\ & (46) \text{ for } t \in \mathbb{Z}_{[k,k+N)} \end{aligned}$$

where $u_{\mathrm{ini}} := u_{[k-L,k)}$ and $y_{\mathrm{ini}} := y_{[k-L,k)}$ are past inputs and outputs, and $\lambda_{\mathrm{y}} > 0$ and $\lambda_{\mathrm{g}} > 0$ are regularization parameters. The SPC optimization problem at control step $t = k$,

$$\begin{aligned} \underset{\bar{u}}{\text{minimize}} \quad & \textstyle\sum_{t=k}^{k+N-1} J_t(\bar{u}_t, \bar{y}_t) \\ \text{subject to} \quad & \bar{y}_{[k,k+N)} := \widehat{\mathcal{P}}_{\mathsf{spc}}\,\mathrm{col}(u_{\mathrm{ini}}, y_{\mathrm{ini}}, \bar{u}_{[k,k+N)}) \\ & (46) \text{ for } t \in \mathbb{Z}_{[k,k+N)} \end{aligned}$$

where $\widehat{\mathcal{P}}_{\mathsf{spc}}$ is the Tikhonov regularization of the prediction matrix $\mathcal{P}_{\mathsf{spc}} := Y_{\mathrm{f}}\,\mathrm{col}(U_{\mathrm{p}}, Y_{\mathrm{p}}, U_{\mathrm{f}})^{\dagger}$, obtained similarly as $\widehat{\mathcal{P}}$ in Section III-A, with a regularization parameter $\lambda > 0$.

TABLE I
CONTROL PARAMETERS

Time Horizon Lengths	
Initial-condition horizon length	$L = 5$
Prediction horizon length	$N = 30$
Control horizon length	$N_\mathrm{c} = 10$
Problem Setup Parameters	
Sampling Period	$T_\mathrm{s} = 1\mathrm{ms}$
Cost matrices	$Q = 10^4 I_p$, $R = I_m$
Constraint coefficients	$E = I_{m+p} \otimes \left[\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right]$, $f = \left[\begin{smallmatrix} 0.6\times\mathbf{1}_{2m\times 1} \\ 0.4\times\mathbf{1}_{2p\times 1} \end{smallmatrix}\right]$
Risk probability bound	$p = 0.2$
Interpolation penalty	$\lambda_\theta = 10$
Variance of v_t for SMPC/SDDPC	$\Sigma^\mathrm{v} = 10^{-8} I_p$
Variance of ρ_t for SDDPC	$\Sigma^\rho = 10^{-4} I_{pL}$
Variance of w_t for SMPC[a]	$\Sigma^\mathrm{w} = \mathcal{O}^\dagger \Sigma^\rho \mathcal{O}^{\dagger\mathsf{T}}$
Regularization Parameters	
Regularization in DeePC	$\lambda_\mathrm{y} = 10^6$, $\lambda_\mathrm{g} = 10^3$
Regularization of $\mathcal{P}$ in SDDPC	$\lambda = 10^{-3}$
Regularization of $\mathcal{P}_\mathrm{spc}$ in SPC	$\lambda = 10^{-3}$

[a]In computation of Σ^w, matrix $\mathcal{O}$ is obtained given the identified system (A, B, C, D) in SMPC.

B. Offline Data Collection

Offline data is required in all our investigated control methods, for use in either data matrices (SDDPC, DeePC and SPC) or for system identification (MPC and SMPC). In our simulation, the data collection process lasted for 1 second and produced a data trajectory of length $T_\mathrm{d} = 1000$ with a sampling period of 1ms. The input data was generated as follows: $\Delta\omega$ (input 1) was set as the phase-locked loop (PLL) control action (see e.g. [18]) plus a white-noise signal, $I_\mathrm{d}^\mathrm{ref}$ (input 2) was set as 0.4 p.u. plus a white-noise signal, and $I_\mathrm{q}^\mathrm{ref}$ (input 3) was set at 0 p.u. plus a white-noise signal. Each white noise signal had noise power of 10^{-6} p.u..

C. Results

All controller parameters are reported in Table I. Our simulation consists of two parts. In the first part, we compare the tracking performances of the different controllers. In the second part, we examine the ability of the controllers to maintain safety constraints.

1) Tracking Performance: For each controller, we perform the following control process. From time 0s to time 0.2s, the controller is switched off, and the inputs $I_\mathrm{d}^\mathrm{ref}$ and $I_\mathrm{q}^\mathrm{ref}$ are set to zero, with $\Delta\omega$ generated from the PLL. After time 0.2s, the controller is switched on, and the output reference signal is $r_t = [0, 0, 0]^\mathsf{T}$ before time 0.5s and $r_t = [0, 0.3, 0]^\mathsf{T}$ after time 0.5s. Fig. 2 shows the stage cost accumulated over the first two seconds for each controller. The stochastic control methods (SMPC and SDDPC) outperformed the deterministic control methods (DeePC, SPC and MPC) in terms of their cumulative costs; this is expected, as they make use of measurement feedback at each time step, rather as opposed to only at each control step, i.e., every $N_\mathrm{c} = 10$ time steps. However, this benefit vanishes when we select shorter control horizons. Fig. 3 shows the cumulative stage costs when the control horizon has length $N_\mathrm{c} = 1$, where we no longer observe a performance gap between all stochastic methods and all deterministic methods; SDDPC and SPC outperform the other controllers in this particular test. The benefit of the $N_\mathrm{c} = 10$ setting is that it requires less computation, since the optimization problems are solved less frequently.

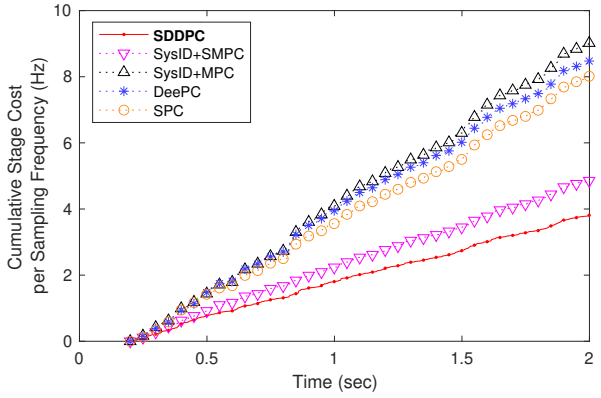

Fig. 2. Cumulative stage cost with different controllers, $N_\mathrm{c} = 10$.

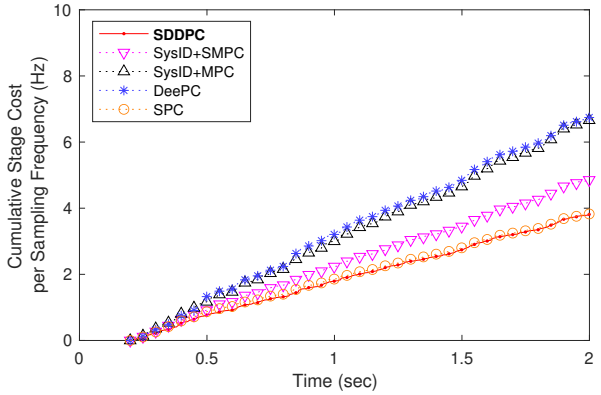

Fig. 3. Cumulative stage cost with different controllers, $N_\mathrm{c} = 1$.

2) Output Constraint Satisfaction: We next evaluate constraint violation for each controller. We repeat the control process above, but the reference signal becomes $r_t = [0, 0, 0]^\mathsf{T}$ before time 0.5s and $r_t = [0, 0.5, 0]^\mathsf{T}$ after time 0.5s. The reference value 0.5 for the second output channel after time 0.5s is beyond the range of output safety constraint $[-0.4, 0.4]$ (TABLE I). The results for SPC and SDDPC are plotted in Fig. 4, where the second output channel remained close to the upper safety bound of 0.4 after time 0.5s.

To quantify the constraint satisfaction with each controller, from time 0.5s to time 2.0s (1500 time steps), we count the number and compute the rate of time steps where the measurement of the second output channel violates the safety constraint. As a second metric, we sum the amount of constraint violation that occurs between 0.5s to 2.0s for each controller. The results are displayed in TABLE II, where we also displayed the results of SMPC and SDDPC with parameter p changed from 0.2 (as in TABLE I) to 0.05. As the result shows, both violation rates of SMPC and SDDPC declined as we decrease p, while the violation rate of SDDPC shrank more effectively than that of SMPC. The total violation amounts of SMPC and

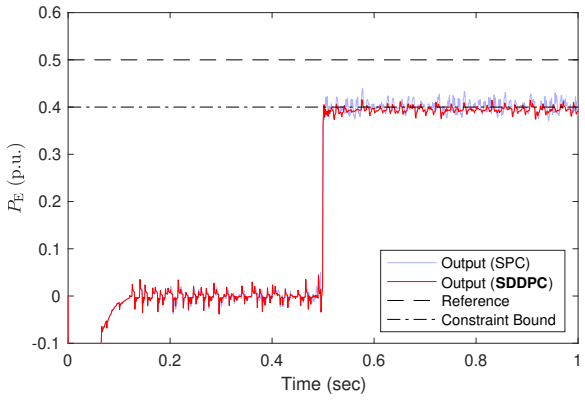

Fig. 4. The second output signals with SPC (light blue) and SDDPC (red) in the constraint satisfaction test.

TABLE II
STATISTICS OF CONSTRAINT VIOLATION OF THE SECOND OUTPUT CHANNEL FROM 0.5S TO 2.0S

Controller	Violation Rate	Total Violation Amount
SDDPC ($p = 0.2$)	0.15	1.10
SDDPC ($p = 0.05$)	0.03	0.05
SysID+SMPC ($p = 0.2$)	0.19	1.55
SysID+SMPC ($p = 0.05$)	0.11	0.52
SysID+MPC	0.57	6.79
DeePC	0.20	1.46
SPC	0.49	8.42

SDDPC also reduced when we decrease p. Among the methods using deterministic safety constraint, DeePC outperformed both MPC and SPC.

V. CONCLUSIONS

We proposed a direct data-driven control framework, Stochastic Data-Driven Predictive Control (SDDPC), which parallels stochastic MPC by explicitly treating process and measurement noise and by computing closed-loop control policies via optimization. Under suitable conditions, SDDPC is equivalent to SMPC. Simulation results demonstrate improvements in cumulative stage cost and reduced output-constraint violations compared to deterministic DDPC methods. Future work includes improving computational efficiency, strengthening robustness to noisy offline data, handling non-Gaussian noise, optimizing the feedback gain K, and enforcing tighter risk constraints (e.g., via CVaR safety constraints).

APPENDIX A. PROOF OF LEMMA 3

Proof. Let $(x^{\mathrm{d}}, u^{\mathrm{d}}, y^{\mathrm{d}})$ be the state-input-output trajectory of (23), and define $X_1, X_2 \in \mathbb{R}^{n\times h}$ as

$$X_1 := \begin{bmatrix} x_1^{\mathrm{d}}, x_2^{\mathrm{d}}, \dots, x_h^{\mathrm{d}} \end{bmatrix}, \quad X_2 := \begin{bmatrix} x_{1+L}^{\mathrm{d}}, x_{2+L}^{\mathrm{d}}, \dots, x_{h+L}^{\mathrm{d}} \end{bmatrix}.$$

It follows by straightforward algebra that data matrices satisfy

$$X_2 = A^L X_1 + \mathcal{C} U_1, \tag{47a}$$
$$Y_1 = \mathcal{O} X_1 + \mathcal{G} U_1, \tag{47b}$$
$$Y_2 = C X_2 + D U_2. \tag{47c}$$

Under our assumptions of controllability and persistent excitation, it follows from [51, Corollary 2(iii)] that the matrix $\mathrm{col}(X_1, U_1, U_2)$ has full row rank. Moreover, $\left[\begin{smallmatrix} I_{mL} \\ \mathcal{G} \end{smallmatrix} \begin{smallmatrix} 0 \\ \mathcal{O} \end{smallmatrix}\right]$ has full column rank, as it is block lower triangular and its diagonal blocks each has full column rank (Section III-A).

First, the matrix Y_2 can be represented in terms of X_1, U_1, U_2 by combining (47a) and (47c) and eliminating X_2, i.e.,

$$Y_2 = \begin{bmatrix} C\mathcal{C}, CA^L, D \end{bmatrix} \mathrm{col}(U_1, X_1, U_2). \tag{48}$$

We can also express $\mathrm{col}(U_1, Y_1, U_2)$ in terms of X_1, U_1, U_2 as

$$\mathrm{col}(U_1, Y_1, U_2) = \mathrm{Diag}\Big(\left[\begin{smallmatrix} I_{mL} \\ \mathcal{G} \end{smallmatrix} \begin{smallmatrix} 0 \\ \mathcal{O} \end{smallmatrix}\right], I_m\Big) \mathrm{col}(U_1, X_1, U_2)$$

using (47b). As we know that $\mathrm{Diag}\Big(\left[\begin{smallmatrix} I_{mL} \\ \mathcal{G} \end{smallmatrix} \begin{smallmatrix} 0 \\ \mathcal{O} \end{smallmatrix}\right], I_m\Big)$ has full column rank and $\mathrm{col}(U_1, X_1, U_2)$ has full row rank, the pseudo-inverse of above is [52]

$$\mathrm{col}(U_1, Y_1, U_2)^\dagger = \mathrm{col}(U_1, X_1, U_2)^\dagger \, \mathrm{Diag}\Big(\left[\begin{smallmatrix} I_{mL} \\ \mathcal{G} \end{smallmatrix} \begin{smallmatrix} 0 \\ \mathcal{O} \end{smallmatrix}\right], I_m\Big)^\dagger.$$

By multiplying (48) and the relation above, we find the result

$$\begin{aligned} Y_2 \, \mathrm{col}(U_1, Y_1, U_2)^\dagger &= \begin{bmatrix} C\mathcal{C}, CA^L, D \end{bmatrix} \mathrm{Diag}\Big(\left[\begin{smallmatrix} I_{mL} \\ \mathcal{G} \end{smallmatrix} \begin{smallmatrix} 0 \\ \mathcal{O} \end{smallmatrix}\right], I_m\Big)^\dagger \\ &= \Big[\begin{bmatrix} C\mathcal{C}, CA^L \end{bmatrix} \left[\begin{smallmatrix} I_{mL} \\ \mathcal{G} \end{smallmatrix} \begin{smallmatrix} 0 \\ \mathcal{O} \end{smallmatrix}\right]^\dagger \quad D \Big] \overset{\text{via (25)}}{=} [\mathbf{\Gamma}_{\mathsf{U}}, \mathbf{\Gamma}_{\mathsf{Y}}, D]. \end{aligned}$$

■

APPENDIX B. PROOF OF LEMMA 4

Proof. We start with intermediate results Claim 4.1 and Claim 4.2. Define matrices $\Phi_{\mathsf{orig}} \in \mathbb{R}^{n\times n_\xi}$ and $\Phi_{\mathsf{aux}} \in \mathbb{R}^{n_{\mathsf{aux}}\times n_\xi}$

$$\Phi_{\mathsf{orig}} := [\mathcal{C}, A^L, \mathcal{C}_{\mathsf{w}}], \quad \Phi_{\mathsf{aux}} := \begin{bmatrix} I_{mL} & & \\ \mathcal{G} & \mathcal{O} & \mathcal{G}_{\mathsf{w}} \\ & & I_L \otimes \mathcal{O} \end{bmatrix}$$

with matrix $\mathcal{O}$ in Section II, matrices $\mathcal{C}, \mathcal{G}$ in Section III-A and $\mathcal{C}_{\mathsf{w}} := [A^{L-1}, \dots, A, I_n]$ and $\mathcal{G}_{\mathsf{w}} := \begin{bmatrix} 0_{p\times n} & & & \\ C & 0_{p\times n} & & \\ \vdots & \ddots & \ddots & \\ CA^{L-2} & \cdots & C & 0_{p\times n} \end{bmatrix}$.

Claim 4.1. *For system* (1) *and the auxiliary state* $\mathbf{x}_t$ *in* (27)*, we have* $x_t = \Phi_{\mathsf{orig}}\, \xi_t$ *and* $\mathbf{x}_t = \Phi_{\mathsf{aux}}\, \xi_t$*, where we let* $\xi_t := \mathrm{col}(u_{[t-L,t)}, x_{t-L}, w_{[t-L,t)}) \in \mathbb{R}^{n_\xi}$ *with* $n_\xi := mL + n(L+1)$.

Proof. Given the system model (1), the state x_t and noise-free outputs $y^\circ_{[t-L,t)}$ can be expressed in terms of the previous state x_{t-L}, inputs $u_{[t-L,t)}$ and disturbances $w_{[t-L,t)}$ via

$$x_t = A^L\, x_{t-L} + \mathcal{C}\, u_{[t-L,t)} + \mathcal{C}_{\mathsf{w}}\, w_{[t-L,t)}, \tag{49a}$$
$$y^\circ_{[t-L,t)} = \mathcal{O}\, x_{t-L} + \mathcal{G}\, u_{[t-L,t)} + \mathcal{G}_{\mathsf{w}}\, w_{[t-L,t)}. \tag{49b}$$

Thus, we have $x_t = \Phi_{\mathsf{orig}}\, \xi_t$ given (49a) and the definitions of ξ_t and Φ_{orig}. Given the definition of $\mathbf{x}_t$ in (27) with $\rho_t := \mathcal{O} w_t$, we have $\mathbf{x}_t = \Phi_{\mathsf{aux}}\, \xi_t$ implied by (49b). ♦

Claim 4.2. *For system* (1) *and the auxiliary state* $\mathbf{x}_t$ *in* (27)*, we have* $Cx_t = \mathbf{C}\mathbf{x}_t$*. Moreover,* $C\Phi_{\mathsf{orig}} = \mathbf{C}\Phi_{\mathsf{aux}}$.

Proof. With Claim 4.1, it suffices to show $C\Phi_{\mathsf{orig}} = \mathbf{C}\Phi_{\mathsf{aux}}$. Given the definitions of $\Phi_{\mathsf{orig}}, \Phi_{\mathsf{aux}}, \mathbf{C}$, we compute $C\Phi_{\mathsf{orig}}$ as

$$C\Phi_{\mathsf{orig}} = [C\mathcal{C}, CA^L, C\mathcal{C}_{\mathsf{w}}]$$

and calculate $\mathbf{C}\Phi_{\mathsf{aux}}$ as

$$\begin{aligned} \mathbf{C}\Phi_{\mathsf{aux}} &= [\mathbf{\Gamma}_{\mathsf{U}} + \mathbf{\Gamma}_{\mathsf{Y}}\mathcal{G},\ \mathbf{\Gamma}_{\mathsf{Y}}\mathcal{O},\ \mathbf{\Gamma}_{\mathsf{Y}}\mathcal{G}_{\mathsf{w}} + (\mathbf{F} - \mathbf{\Gamma}_{\mathsf{Y}}\mathbf{E})(I_L \otimes \mathcal{O})] \\ &= [\mathbf{\Gamma}_{\mathsf{U}} + \mathbf{\Gamma}_{\mathsf{Y}}\mathcal{G},\ \mathbf{\Gamma}_{\mathsf{Y}}\mathcal{O},\ C\mathcal{C}_{\mathsf{w}}] = [C\mathcal{C}, CA^L, C\mathcal{C}_{\mathsf{w}}], \end{aligned}$$

where the second equality used the facts that $C\mathcal{C}_{\mathsf{w}} = \mathbf{F}(I_L \otimes \mathcal{O})$ and $\mathcal{G}_{\mathsf{w}} = \mathbf{E}(I_L \otimes \mathcal{O})$ which can be verified from the definitions of $\mathbf{E}, \mathbf{F}, \mathcal{C}_{\mathsf{w}}, \mathcal{G}_{\mathsf{w}}$, and the last equality above used the relation

$$[\mathbf{\Gamma}_{\mathsf{U}} + \mathbf{\Gamma}_{\mathsf{Y}}\mathcal{G},\, \mathbf{\Gamma}_{\mathsf{Y}}\mathcal{O}] = [\mathbf{\Gamma}_{\mathsf{U}},\, \mathbf{\Gamma}_{\mathsf{Y}}]\begin{bmatrix} I_{mL} & \\ \mathcal{G} & \mathcal{O} \end{bmatrix} = [C\mathcal{C}, CA^L]$$

where the last equality is due to the definition $[\mathbf{\Gamma}_{\mathsf{U}},\, \mathbf{\Gamma}_{\mathsf{Y}}] := [C\mathcal{C}, CA^L]\begin{bmatrix} I_{mL} & \\ \mathcal{G} & \mathcal{O} \end{bmatrix}^{\dagger}$ where $\begin{bmatrix} I_{mL} & \\ \mathcal{G} & \mathcal{O} \end{bmatrix}$ has full column rank. Comparing the above results of calculation, we have $C\Phi_{\mathsf{orig}} = \mathbf{C}\Phi_{\mathsf{aux}}$, and thus the result follows from Claim 4.1. ♦

We directly have (28b) from Claim 4.2 and (1b). To show (28a), we know by substitution that $\mathbf{A}\mathbf{x}_t + \mathbf{B}u_t + \mathbf{w}_t$ equals

$$\mathrm{col}\Big(\begin{bmatrix} u_{[t-L+1,t)} \\ u_t \end{bmatrix}, \begin{bmatrix} y^{\circ}_{[t-L+1,t)} \\ \mathbf{C}\mathbf{x}_t + Du_t \end{bmatrix}, \begin{bmatrix} \rho_{[t-L+1,t)} \\ \rho_t \end{bmatrix}\Big),$$

given the definitions of $\mathbf{x}_t, \mathbf{w}_t, \mathbf{A}, \mathbf{B}$ with matrix $\mathbf{A}$ consisting of upper-shift matrices and the matrix $\mathbf{C}$. The above is $\mathbf{x}_{t+1}$ by definition, given the fact $\mathbf{C}\mathbf{x}_t + Du_t = y_t - v_t = y^{\circ}_t$ via (28b), and thereby (28a) is obtained. ■

Appendix C. Proof of Lemma 5

Proof. The pair $(\mathbf{A}, \mathbf{C})$ is detectable by definition since there exists a matrix $\mathbf{L}^* := \mathrm{col}(0_{mL\times p}, 0_{p(L-1)\times p}, I_p, 0_{pL^2\times p})$ such that $\mathbf{A} - \mathbf{L}^*\mathbf{C}$ equal to $\mathrm{Diag}(\mathcal{D}_m, \mathcal{D}_p, \mathcal{D}_{pL})$ is Schur stable, where $\mathcal{D}_q := \begin{bmatrix} 0_{q\times q} & I_{q(L-1)} \end{bmatrix} \in \mathbb{R}^{qL\times qL}$.

The rest of the proof shows stabilizability of $(\mathbf{A}, \mathbf{B})$ and $(\mathbf{A}, \mathbf{\Sigma}^{\mathsf{w}})$. We start with basic results Claim 5.1 and Claim 5.2. Recall matrices $\Phi_{\mathsf{orig}} \in \mathbb{R}^{n\times n_\xi}$ and $\Phi_{\mathsf{aux}} \in \mathbb{R}^{n_{\mathsf{aux}}\times n_\xi}$ defined in Appendix B, with $n_{\mathsf{aux}} := mL + pL + pL^2$ and $n_\xi := mL + n + nL$. Define matrix $\Phi := [\Phi_{\mathsf{U}}, \Phi_{\mathsf{Y}}, \Phi_{\mathsf{P}}] \in \mathbb{R}^{n\times n_{\mathsf{aux}}}$ whose sub-blocks are defined as

$$[\Phi_{\mathsf{U}}, \Phi_{\mathsf{Y}}] := [\mathcal{C}, A^L]\begin{bmatrix} I_{mL} & \\ \mathcal{G} & \mathcal{O} \end{bmatrix}^{\dagger} \in \mathbb{R}^{n\times(mL+pL)},$$
$$\Phi_{\mathsf{P}} := (\mathcal{C}_{\mathsf{w}} - \Phi_{\mathsf{Y}}\mathcal{G}_{\mathsf{w}})(I_L \otimes \mathcal{O})^{\dagger} \in \mathbb{R}^{n\times pL^2},$$

and we have the following relation regarding Φ, Φ_{aux} and Φ_{orig}.

Claim 5.1. $\Phi_{\mathsf{orig}} = \Phi\Phi_{\mathsf{aux}}$.

Proof. Given the definitions of Φ and Φ_{aux}, with both $\begin{bmatrix} I_{mL} & \\ \mathcal{G} & \mathcal{O} \end{bmatrix}$ and $I_L \otimes \mathcal{O}$ having full column rank, the product $\Phi\Phi_{\mathsf{aux}}$ is calculated as $[\mathcal{C}, A^L, \mathcal{C}_{\mathsf{w}}]$, equal to Φ_{orig} by definition. ♦

Define matrices $\widetilde{A} \in \mathbb{R}^{n_\xi\times n_\xi}$, $\widetilde{B} \in \mathbb{R}^{n_\xi\times m}$, $\widetilde{\Sigma}^{\mathsf{w}} \in \mathbb{S}^{n_\xi}_{+}$

$$\widetilde{A} := \left[\begin{array}{cc|c|cc} & I_{m(L-1)} & & & \\ 0_{m\times m} & & & & \\ \hline B & 0_{n\times m(L-1)} & A & I_n & 0_{n\times n(L-1)} \\ \hline & & & & I_{n(L-1)} \\ & & & 0_{n\times n} & \end{array}\right], \quad (50)$$

$$\widetilde{B} := \left[\begin{array}{c} 0_{m(L-1)\times m} \\ \hline I_m \\ \hline 0_{n\times m} \\ \hline 0_{nL\times m} \end{array}\right], \quad \widetilde{\Sigma}^{\mathsf{w}} := \begin{bmatrix} 0_{(n_\xi-n)\times(n_\xi-n)} & \\ & \Sigma^{\mathsf{w}} \end{bmatrix},$$

and the following result relates $\widetilde{A}, \widetilde{B}, \widetilde{\Sigma}^{\mathsf{w}}$ to $\mathbf{A}, \mathbf{B}, \mathbf{\Sigma}^{\mathsf{w}}$.

Claim 5.2. *For matrices* $\mathbf{A}$,$\mathbf{B}$ *in* (28)*,* $\mathbf{\Sigma}^{\mathsf{w}}$ *in* (30) *and* $\widetilde{A}, \widetilde{B}, \widetilde{\Sigma}^{\mathsf{w}}$ *in* (50)*, we have*

$$\mathbf{A}\Phi_{\mathsf{aux}} = \Phi_{\mathsf{aux}}\widetilde{A}, \quad \mathbf{B} = \Phi_{\mathsf{aux}}\widetilde{B}, \quad \mathbf{\Sigma}^{\mathsf{w}} = \Phi_{\mathsf{aux}}\widetilde{\Sigma}^{\mathsf{w}}\Phi_{\mathsf{aux}}^{\mathsf{T}}.$$

Proof. $\mathbf{B} = \Phi_{\mathsf{aux}}\widetilde{B}$ and $\mathbf{\Sigma}^{\mathsf{w}} = \Phi_{\mathsf{aux}}\widetilde{\Sigma}^{\mathsf{w}}\Phi_{\mathsf{aux}}^{\mathsf{T}}$ follow directly by expressing $\Phi_{\mathsf{aux}}, \mathbf{B}, \mathbf{\Sigma}^{\mathsf{w}}, \widetilde{B}, \widetilde{\Sigma}^{\mathsf{w}}$ and underlying $\mathcal{G}, \mathcal{O}, \mathcal{G}_{\mathsf{w}}$ in terms of A, B, C, D by definition. To show $\mathbf{A}\Phi_{\mathsf{aux}} = \Phi_{\mathsf{aux}}\widetilde{A}$, we first replace a subblock $\mathbf{C}\Phi_{\mathsf{aux}}$ of $\mathbf{A}\Phi_{\mathsf{aux}}$ (since $\mathbf{C}$ is a subblock of $\mathbf{A}$) using the relation $\mathbf{C}\Phi_{\mathsf{aux}} = C\Phi_{\mathsf{orig}}$ (shown in the proof of Claim 4.2), and then similarly express all the matrices in terms of A, B, C, D by definition. ♦

The next result shows that the pairs $(\widetilde{A}, \widetilde{B})$ and $(\widetilde{A}, \widetilde{\Sigma}^{\mathsf{w}})$ are stabilizable. Define gain matrices

$$\mathbf{K}^* := K\Phi, \qquad \mathbf{K}^{\mathsf{w}} := \Phi_{\mathsf{aux}}^{\dagger\mathsf{T}}\,\mathrm{col}(0_{(n_\xi-n)\times n}, K^{\mathsf{w}})\,\Phi \quad (51a)$$
$$\widetilde{K}^* := K\Phi_{\mathsf{orig}}, \quad \widetilde{K}^{\mathsf{w}} := \mathrm{col}(0_{(n_\xi-n)\times n}, K^{\mathsf{w}})\,\Phi_{\mathsf{orig}} \quad (51b)$$

where K is the feedback gain from (11) and $K^{\mathsf{w}} \in \mathbb{R}^{n\times n}$ is an arbitrary matrix such that $A - \Sigma^{\mathsf{w}}K^{\mathsf{w}}$ is Schur stable.

Claim 5.3. *For matrices* $\widetilde{A}, \widetilde{B}, \widetilde{\Sigma}^{\mathsf{w}}$ *in* (50) *and* $\widetilde{K}^*, \widetilde{K}^{\mathsf{w}}$ *in* (51b)*, both* $\widetilde{A} - \widetilde{B}\widetilde{K}^*$ *and* $\widetilde{A} - \widetilde{\Sigma}^{\mathsf{w}}\widetilde{K}^{\mathsf{w}}$ *are Schur stable.*

Proof. Define $\xi_t := \mathrm{col}(u_{[t-L,t)}, x_{t-L}, w_{[t-L,t)}) \in \mathbb{R}^{n_\xi}$ and $\delta_t := \mathrm{col}(0_{(n_\xi-n)\times 1}, w_t) \in \mathbb{R}^{n_\xi}$. We have the relation

$$\xi_{t+1} = \widetilde{A}\xi_t + \widetilde{B}u_t + \delta_t \quad (52)$$

which can be verified given the system model (1a) and the definition (50) of $\widetilde{A}, \widetilde{B}, \widetilde{\Sigma}^{\mathsf{w}}$.

To show that $\widetilde{A} - \widetilde{B}\widetilde{K}^*$ is stable, consider the following process of system (1a) starting at time $t = -L$: the initial state x_{-L}, the inputs $u_{[-L,0)}$ and the noises $w_{[-L,0)}$ are arbitrarily chosen (i.e., ξ_0 is arbitrary), the noise is $w_t = 0$ for $t \geq 0$, and the inputs u_t for $t \geq 0$ are generated by state feedback $u_t = -Kx_t$. With this process, we have $x_{t+1} = (A - BK)x_t$ for $t \geq 0$, and hence $x_t \to 0$ as $t \to \infty$ because $A - BK$ is Schur stable. We therefore have $u_t, w_t \to 0$ and thus $\xi_t \to 0$ as $t \to \infty$, given the relations $u_t = -Kx_t$ and $w_t = 0$ for $t \geq 0$ and the definition of ξ_t. On the other hand, with the process, we have $\delta_t = 0$ since $w_t = 0$ for $t \geq 0$, and the state feedback $u_t = -Kx_t$ can be written as $u_t = -K\Phi_{\mathsf{orig}}\xi_t$ given the relation $x_t = \Phi_{\mathsf{orig}}\xi_t$ from Claim 4.1, so we have $u_t = -\widetilde{K}^*\xi_t$ with $\widetilde{K}^*$ defined in (51b). Therefore, the evolution (52) is reduced as $\xi_{t+1} = (\widetilde{A} - \widetilde{B}\widetilde{K}^*)\xi_t$ for $t \geq 0$, which implies that $\xi_t = (\widetilde{A} - \widetilde{B}\widetilde{K}^*)^t\xi_0$ for $t \geq 0$. Since $\xi_t \to 0$ as $t \to \infty$ and ξ_0 is arbitrarily chosen, we conclude that $(\widetilde{A} - \widetilde{B}\widetilde{K}^*)^t \to 0$ as $t \to \infty$, i.e., $\widetilde{A} - \widetilde{B}\widetilde{K}^*$ is Schur stable.

To show that $\widetilde{A} - \widetilde{\Sigma}^{\mathsf{w}}\widetilde{K}^{\mathsf{w}}$ is stable, consider a similar process of system (1a) from initial time $t = -L$: the initial state x_{-L}, the inputs $u_{[-L,0)}$ and the noises $w_{[-L,0)}$ are arbitrarily chosen (i.e., ξ_0 is arbitrary), the input is $u_t = 0$ for $t \geq 0$, and the disturbances w_t for $t \geq 0$ are realized as $w_t = -\Sigma^{\mathsf{w}}K^{\mathsf{w}}x_t$. With the process, we have $x_{t+1} = (A - \Sigma^{\mathsf{w}}K^{\mathsf{w}})x_t$ for $t \geq 0$, and hence $x_t \to 0$ as $t \to \infty$ because $A - \Sigma^{\mathsf{w}}K^{\mathsf{w}}$ is Schur stable. We therefore have $u_t, w_t \to 0$ and thus $\xi_t \to 0$ as $t \to \infty$, given the relations $u_t = 0$ and $w_t = -\Sigma^{\mathsf{w}}K^{\mathsf{w}}x_t$ for $t \geq 0$ and the definition of ξ_t. On the other hand, with the process, we have $\delta_t = -\widetilde{\Sigma}^{\mathsf{w}}\widetilde{K}^{\mathsf{w}}\xi_t$ for $t \geq 0$, given the definition of δ_t, the choice of noise $w_t = -\Sigma^{\mathsf{w}}K^{\mathsf{w}}x_t$, the relation $x_t = \Phi_{\mathsf{orig}}\xi_t$ from Claim 4.1 and the definitions of $\widetilde{K}^{\mathsf{w}}$ in (51b) and $\widetilde{\Sigma}^{\mathsf{w}}$ in Claim 5.2. Therefore, the evolution (52) is reduced as $\xi_{t+1} = (\widetilde{A} - \widetilde{\Sigma}^{\mathsf{w}}\widetilde{K}^{\mathsf{w}})\xi_t$ for $t \geq 0$, which implies that $\xi_t = (\widetilde{A} - \widetilde{\Sigma}^{\mathsf{w}}\widetilde{K}^{\mathsf{w}})^t\xi_0$ for $t \geq 0$. Since $\xi_t \to 0$ as $t \to \infty$ and

ξ_0 is arbitrarily chosen, we conclude that $(\widetilde{A}-\widetilde{\Sigma}^{\mathsf{w}}\widetilde{K}^{\mathsf{w}})^t \to 0$ as $t\to\infty$, i.e., $\widetilde{A}-\widetilde{\Sigma}^{\mathsf{w}}\widetilde{K}^{\mathsf{w}}$ is Schur stable. ♦

We then obtain the next result regarding Schur stability of $\mathbf{A}-\mathbf{B}\mathbf{K}^*$ and $\mathbf{A}-\mathbf{\Sigma}^{\mathsf{w}}\mathbf{K}^{\mathsf{w}}$.

Claim 5.4. *For matrices* $\mathbf{A},\mathbf{B}$ *in* (28), $\mathbf{\Sigma}^{\mathsf{w}}$ *in* (30) *and* $\mathbf{K}^*,\mathbf{K}^{\mathsf{w}}$ *in* (51a), $\mathbf{A}-\mathbf{B}\mathbf{K}^*$ *is Schur stable if, and only if,* $\mathbf{A}-\mathbf{\Sigma}^{\mathsf{w}}\mathbf{K}^{\mathsf{w}}$ *is Schur stable.*

Proof. Since $\Phi_{\mathsf{aux}}\in\mathbb{R}^{n_{\mathsf{aux}}\times n_\xi}$ by definition has full column rank, there exists a matrix $\Phi_{\mathsf{orth}}\in\mathbb{R}^{n_{\mathsf{aux}}\times(n_{\mathsf{aux}}-n_\xi)}$ such that $\mathrm{Range}(\Phi_{\mathsf{orth}})=\mathrm{Null}(\Phi_{\mathsf{aux}}^{\mathsf{T}})$; it follows that

$$\Phi_{\mathsf{aux}}\Phi_{\mathsf{aux}}^{\dagger}+\Phi_{\mathsf{orth}}\Phi_{\mathsf{orth}}^{\dagger}=I_{n_{\mathsf{aux}}}. \tag{53}$$

It follows from the definitions in (51) that

$$(\mathbf{A}-\mathbf{B}\mathbf{K}^*)\Phi_{\mathsf{aux}}=\Phi_{\mathsf{aux}}(\widetilde{A}-\widetilde{B}\widetilde{K}^*), \tag{54a}$$
$$(\mathbf{A}-\mathbf{\Sigma}^{\mathsf{w}}\mathbf{K}^{\mathsf{w}})\Phi_{\mathsf{aux}}=\Phi_{\mathsf{aux}}(\widetilde{A}-\widetilde{\Sigma}^{\mathsf{w}}\widetilde{K}^{\mathsf{w}}), \tag{54b}$$

given Claim 5.1, Claim 5.2 and $\Phi_{\mathsf{aux}}^{\mathsf{T}}\Phi_{\mathsf{aux}}^{\dagger\mathsf{T}}=I_{n_\xi}$ for Φ_{aux} of full column rank. Define matrices $\mathcal{S}^*,\mathcal{S}^{\mathsf{w}},\mathcal{R}^*,\mathcal{R}^{\mathsf{w}}$,

$$\mathcal{S}^*:=\Phi_{\mathsf{orth}}^{\dagger}(\mathbf{A}-\mathbf{B}\mathbf{K}^*)\Phi_{\mathsf{orth}},\ \ \mathcal{S}^{\mathsf{w}}:=\Phi_{\mathsf{orth}}^{\dagger}(\mathbf{A}-\mathbf{\Sigma}^{\mathsf{w}}\mathbf{K}^{\mathsf{w}})\Phi_{\mathsf{orth}}$$
$$\mathcal{R}^*:=\Phi_{\mathsf{aux}}^{\dagger}(\mathbf{A}-\mathbf{B}\mathbf{K}^*)\Phi_{\mathsf{orth}},\ \ \mathcal{R}^{\mathsf{w}}:=\Phi_{\mathsf{aux}}^{\dagger}(\mathbf{A}-\mathbf{\Sigma}^{\mathsf{w}}\mathbf{K}^{\mathsf{w}})\Phi_{\mathsf{orth}}$$

and it follows from (53) that

$$\begin{aligned}(\mathbf{A}-\mathbf{B}\mathbf{K}^*)\Phi_{\mathsf{orth}}&=\Phi_{\mathsf{aux}}\mathcal{R}^*+\Phi_{\mathsf{orth}}\mathcal{S}^*,\\ (\mathbf{A}-\mathbf{\Sigma}^{\mathsf{w}}\mathbf{K}^{\mathsf{w}})\Phi_{\mathsf{orth}}&=\Phi_{\mathsf{aux}}\mathcal{R}^{\mathsf{w}}+\Phi_{\mathsf{orth}}\mathcal{S}^{\mathsf{w}}.\end{aligned} \tag{55}$$

We moreover notice that $\mathcal{S}^*=\mathcal{S}^{\mathsf{w}}=\Phi_{\mathsf{orth}}^{\dagger}\mathbf{A}\Phi_{\mathsf{orth}}$ given the definitions of $\mathcal{S}^*,\mathcal{S}^{\mathsf{w}}$ and the facts $\Phi_{\mathsf{orth}}^{\dagger}\mathbf{B}=0$ and $\Phi_{\mathsf{orth}}^{\dagger}\mathbf{\Sigma}^{\mathsf{w}}=0$ which follow from the fact $\Phi_{\mathsf{orth}}^{\dagger}\Phi_{\mathsf{aux}}=0$ via (53) and the relations $\mathbf{B}=\Phi_{\mathsf{aux}}\widetilde{B}$ and $\mathbf{\Sigma}^{\mathsf{w}}=\Phi_{\mathsf{aux}}\widetilde{\Sigma}^{\mathsf{w}}\Phi_{\mathsf{aux}}^{\mathsf{T}}$ from Claim 5.2.

Define $\Phi_{\mathsf{full}}:=[\Phi_{\mathsf{aux}},\Phi_{\mathsf{orth}}]\in\mathbb{R}^{n_{\mathsf{aux}}\times n_{\mathsf{aux}}}$ which is non-singular given (53); the horizontal stack of (54) and (55) yields

$$\begin{aligned}(\mathbf{A}-\mathbf{B}\mathbf{K}^*)\Phi_{\mathsf{full}}&=\Phi_{\mathsf{full}}\begin{bmatrix}\widetilde{A}-\widetilde{B}\widetilde{K}^* & \mathcal{R}^*\\ 0 & \mathcal{S}^*\end{bmatrix},\\ (\mathbf{A}-\mathbf{\Sigma}^{\mathsf{w}}\mathbf{K}^{\mathsf{w}})\Phi_{\mathsf{full}}&=\Phi_{\mathsf{full}}\begin{bmatrix}\widetilde{A}-\widetilde{\Sigma}^{\mathsf{w}}\widetilde{K}^{\mathsf{w}} & \mathcal{R}^{\mathsf{w}}\\ 0 & \mathcal{S}^{\mathsf{w}}\end{bmatrix}.\end{aligned} \tag{56}$$

Since $\widetilde{A}-\widetilde{B}\widetilde{K}^*$ and $\widetilde{A}-\widetilde{\Sigma}^{\mathsf{w}}\widetilde{K}^{\mathsf{w}}$ are Schur stable (Claim 5.3), the matrix similarity relations (56) imply that $\mathbf{A}-\mathbf{B}\mathbf{K}^*$ (resp. $\mathbf{A}-\mathbf{\Sigma}^{\mathsf{w}}\mathbf{K}^{\mathsf{w}}$) is Schur stable if, and only if, $\mathcal{S}^*$ (resp. $\mathcal{S}^{\mathsf{w}}$) is Schur stable. Hence, the result follows from the fact $\mathcal{S}^*=\mathcal{S}^{\mathsf{w}}$. ♦

We need the following intermediate result to show that $\mathbf{A}-\mathbf{B}\mathbf{K}^*$ is Schur stable.

Claim 5.5. *For matrices* $\mathbf{A},\mathbf{B}$ *in* (28) *and* $\mathbf{K}^*$ *in* (51a), *if*

$$(\mathbf{A}-\mathbf{B}\mathbf{K}^*)^t\Phi_{\mathsf{aux}}\to 0 \quad \textit{as} \quad t\to\infty, \tag{57}$$

then $\mathbf{A}-\mathbf{B}\mathbf{K}^*$ *is Schur stable.*

Proof. We calculate $\mathbf{A}-\mathbf{B}\mathbf{K}^*$ as

$$\left[\begin{array}{cc|c}\overbrace{\begin{array}{cc}0 & I_{m(L-1)}\\ -K\Phi_{\mathsf{U}} & \end{array}}^{}\ \overbrace{\begin{array}{cc}0 & \\ -K\Phi_{\mathsf{Y}} & \end{array}}^{=:\mathcal{A}} & & \begin{array}{c}0\\ -K\Phi_{\mathsf{P}}\end{array}\\ \begin{array}{cc}0 & \\ (C-DK)\Phi_{\mathsf{U}} & \end{array}\ \begin{array}{cc}0 & I_{p(L-1)}\\ (C-DK)\Phi_{\mathsf{Y}} & \end{array} & & \begin{array}{c}0\\ \mathbf{F}-\mathbf{\Gamma}_{\mathsf{Y}}\mathbf{E}\end{array}\\ \hline 0 & & \begin{array}{c}I_{pL(L-1)}\\ 0_{pL\times pL}\end{array}\end{array}\right],$$

which is Schur stable if, and only if, its sub-matrix $\mathcal{A}$ is Schur stable. Moreover, since both $\mathbf{A}-\mathbf{B}\mathbf{K}^*=\left[\begin{smallmatrix}\mathcal{A} & *\\ 0 & *\end{smallmatrix}\right]$ and $\Phi_{\mathsf{aux}}=\left[\begin{smallmatrix}\mathcal{S} & *\\ 0 & *\end{smallmatrix}\right]$ are upper block-triangular, $(\mathbf{A}-\mathbf{B}\mathbf{K}^*)^t\Phi_{\mathsf{aux}}=\left[\begin{smallmatrix}\mathcal{A}^t\mathcal{S} & *\\ 0 & *\end{smallmatrix}\right]$ is also upper block-triangular. Since $(\mathbf{A}-\mathbf{B}\mathbf{K}^*)^t\Phi_{\mathsf{aux}}\to 0$ as $t\to\infty$ via (57), its sub-matrix yields $\mathcal{A}^t\mathcal{S}\to 0$ as $t\to\infty$.

Let $\mathcal{L}:=\lim_{t\to\infty}\mathcal{A}^t$ denote the limiting value. Given the definition $[\Phi_{\mathsf{U}},\Phi_{\mathsf{Y}}]:=[\mathcal{C},A^L]\mathcal{S}^{\dagger}$ where $\mathcal{S}$ denotes $\left[\begin{smallmatrix}I_{mL} & \\ \mathcal{G} & \mathcal{O}\end{smallmatrix}\right]$, $\mathcal{A}$ can be written as $\mathcal{A}=\mathcal{D}+\mathcal{E}\mathcal{S}^{\dagger}$ where

$$\mathcal{D}:=\mathrm{Diag}\Big(\begin{bmatrix}0_{m\times m} & I_{m(L-1)}\end{bmatrix},\begin{bmatrix}0_{p\times p} & I_{p(L-1)}\end{bmatrix}\Big),$$
$$\mathcal{E}:=\mathrm{col}(0_{m(L-1)\times n},-K,0_{p(L-1)\times n},C-DK)\,[\mathcal{C},A^L].$$

Define $\mathcal{P}:=I-\mathcal{S}\mathcal{S}^{\dagger}$ as a projection matrix. With the fact $\mathcal{S}^{\dagger}\mathcal{P}=\mathcal{S}^{\dagger}(I-\mathcal{S}\mathcal{S}^{\dagger})=0$, it follows that

$$\mathcal{A}\mathcal{S}\mathcal{S}^{\dagger}=\mathcal{A}-\mathcal{A}\mathcal{P}=\mathcal{A}-(\mathcal{D}+\mathcal{E}\mathcal{S}^{\dagger})\mathcal{P}=\mathcal{A}-\mathcal{D}\mathcal{P}.$$

Left-multiplying the above by $\mathcal{A}^{t-1}$ and taking the limit as $t\to\infty$, we find that

$$\lim_{t\to\infty}\mathcal{A}^t\mathcal{S}\mathcal{S}^{\dagger}=\underbrace{\lim_{t\to\infty}\mathcal{A}^t}_{=\mathcal{L}}-\underbrace{\lim_{t\to\infty}\mathcal{A}^{t-1}}_{=\mathcal{L}}\mathcal{D}\mathcal{P}$$

Since $\mathcal{A}^t\mathcal{S}\to 0$ as $t\to\infty$, the left-hand side of above is zero, so the above further reduces to $0=\mathcal{L}(I-\mathcal{D}\mathcal{P})$. Therefore, to show $\mathcal{L}=0$, it suffices to show that $I-\mathcal{D}\mathcal{P}$ is non-singular. Suppose a vector z in $\mathrm{Null}(I-\mathcal{D}\mathcal{P})$. If $z\notin\mathrm{Range}(\mathcal{P})$, then $\|\mathcal{P}z\|_2<\|z\|_2$ for a projection matrix $\mathcal{P}$, and then we have

$$\|z\|_2=\|\mathcal{D}\mathcal{P}z\|_2\le\underbrace{\|\mathcal{D}\|_2}_{=1}\underbrace{\|\mathcal{P}z\|_2}_{<\|z\|_2}<\|z\|_2,$$

which is a contradiction. Hence, we know that $z\in\mathrm{Range}(\mathcal{P})$, which implies that $\mathcal{P}z=z$ because $\mathcal{P}$ is projection. Combining $z=\mathcal{D}\mathcal{P}z$ and $\mathcal{P}z=z$, we have $(I-\mathcal{D})z=0$, which implies $z=0$ since $I-\mathcal{D}$ is non-singular. Therefore, we conclude that $\mathrm{Null}(I-\mathcal{D}\mathcal{P})=\{0\}$ and $I-\mathcal{D}\mathcal{P}$ is non-singular, so we have $\mathcal{L}=0$, which implies that $\mathcal{A}$ is Schur stable. Thus, $\mathbf{A}-\mathbf{B}\mathbf{K}^*$ is Schur stable. ♦

We finally show stabilizability of $(\mathbf{A},\mathbf{B})$ and $(\mathbf{A},\mathbf{\Sigma}^{\mathsf{w}})$. By applying (54a) repeatedly, we have $(\mathbf{A}-\mathbf{B}\mathbf{K}^*)^t\Phi_{\mathsf{aux}}=\Phi_{\mathsf{aux}}(\widetilde{A}-\widetilde{B}\widetilde{K}^*)^t$ for all $t\in\mathbb{N}$. Combining this relation with the fact $(\widetilde{A}-\widetilde{B}\widetilde{K}^*)^t\to 0$ as $t\to\infty$ via Schur stability in Claim 5.3, we have (57), which implies Schur stability of $\mathbf{A}-\mathbf{B}\mathbf{K}^*$ through Claim 5.5. Given Claim 5.4, both $\mathbf{A}-\mathbf{B}\mathbf{K}^*$ and $\mathbf{A}-\mathbf{\Sigma}^{\mathsf{w}}\mathbf{K}^{\mathsf{w}}$ are Schur stable, which indicates that both pairs $(\mathbf{A},\mathbf{B})$ and $(\mathbf{A},\mathbf{\Sigma}^{\mathsf{w}})$ are stabilizable. ■

Appendix D. Proof of Proposition 7

We present preliminary results in Subsection A and prove Proposition 7 in Subsection B.

A. Preliminary Results

We begin by establishing useful identities in Claim 7.1 that will be leveraged in the remainder of the proof. Recall the matrices $\Phi_{\mathsf{orig}}\in\mathbb{R}^{n\times n_\xi},\Phi_{\mathsf{aux}}\in\mathbb{R}^{n_{\mathsf{aux}}\times n_\xi}$ defined in Appendix B and matrix $\Phi\in\mathbb{R}^{n\times n_{\mathsf{aux}}}$ defined in Appendix C, with $n_{\mathsf{aux}}:=mL+pL+pL^2$ and $n_\xi:=mL+n+nL$.

Claim 7.1. *For the system* (1) *and auxiliary model* (28)*, it holds for all* $t \in \mathbb{N}_{\geq 0}$ *that*

$$x_t = \Phi \mathbf{x}_t \qquad A\Phi\Phi_{\mathsf{aux}} = \Phi\mathbf{A}\Phi_{\mathsf{aux}} \qquad B = \Phi\mathbf{B}$$
$$w_t = \Phi \mathbf{w}_t \qquad C\Phi\Phi_{\mathsf{aux}} = \mathbf{C}\Phi_{\mathsf{aux}} \qquad \Sigma^{\mathsf{w}} = \Phi\mathbf{\Sigma}^{\mathsf{w}}\Phi^{\mathsf{T}}.$$

Proof. The relation $x_t = \Phi \mathbf{x}_t$ follows from Claim 4.1 and Claim 5.1. We have $C\Phi\Phi_{\mathsf{aux}} = \mathbf{C}\Phi_{\mathsf{aux}}$ from Claim 4.2. To show $w_t = \Phi \mathbf{w}_t$ and $\Sigma^{\mathsf{w}} = \Phi\mathbf{\Sigma}^{\mathsf{w}}\Phi^{\mathsf{T}}$, recall from the definition that $\mathbf{w}_t = J_0 w_t$ and $\mathbf{\Sigma}^{\mathsf{w}} = J_0 \Sigma^{\mathsf{w}} J_0^{\mathsf{T}}$ where $J_0 := \mathrm{col}(0_{(n_{\mathsf{aux}}-pL)\times n}, \mathcal{O})$. By direct calculation one can verify that $\Phi J_0 = I_n$, using which we obtain $w_t = \Phi \mathbf{w}_t$ given $\mathbf{w}_t = J_0 w_t$ and obtain $\Sigma^{\mathsf{w}} = \Phi\mathbf{\Sigma}^{\mathsf{w}}\Phi^{\mathsf{T}}$ given $\mathbf{\Sigma}^{\mathsf{w}} = J_0\Sigma^{\mathsf{w}}J_0^{\mathsf{T}}$. We have $B = \Phi_{\mathsf{orig}}\widetilde{B} = \Phi\Phi_{\mathsf{aux}}\widetilde{B} = \Phi\mathbf{B}$, using $\Phi_{\mathsf{orig}} = \Phi\Phi_{\mathsf{aux}}$ as Claim 5.1, $\Phi_{\mathsf{aux}}\widetilde{B} = \mathbf{B}$ from Claim 5.2 and $B = \Phi_{\mathsf{orig}}\widetilde{B}$ which can be verified by definitions of Φ_{orig} and $\widetilde{B}$. We finally have $A\Phi\Phi_{\mathsf{aux}} = A\Phi_{\mathsf{orig}} = \Phi_{\mathsf{orig}}\widetilde{A} = \Phi\Phi_{\mathsf{aux}}\widetilde{A} = \Phi\mathbf{A}\Phi_{\mathsf{aux}}$, where we used $\Phi_{\mathsf{orig}} = \Phi\Phi_{\mathsf{aux}}$ as Claim 5.1, $\mathbf{A}\Phi_{\mathsf{aux}} = \Phi_{\mathsf{aux}}\widetilde{A}$ in Claim 5.2 and $A\Phi_{\mathsf{orig}} = \Phi_{\mathsf{orig}}\widetilde{A}$ which can be verified given the definitions of Φ_{orig} and $\widetilde{A}$. ♦

Next, we relate the LQR feedback gains K and $\mathbf{K}$.

Claim 7.2. *For matrices* K *in* (11) *and* $\mathbf{K}$ *in* (38)*, it holds that* $K\Phi\Phi_{\mathsf{aux}} = \mathbf{K}\Phi_{\mathsf{aux}}$.

Proof. Let $\widetilde{C} := C\Phi_{\mathsf{orig}}$ and let $\widetilde{A}, \widetilde{B}$ be as in (50). We first show the pair $(\widetilde{A}, \widetilde{C})$ is detectable. For $\lambda \in \mathbb{C}$, define $H_{\mathsf{obs}} := \mathrm{col}(\lambda I_{n_\xi} - \widetilde{A}, \widetilde{C})$, which can be permuted into the form

$$\left[\begin{array}{cc|cc|c} \multicolumn{2}{c|}{\lambda I_{mL} - \mathcal{D}_m} & & & \\ \hline & & \multicolumn{2}{c|}{\lambda I_{nL} - \mathcal{D}_n} & \\ \hline -B & 0_{n\times m(L-1)} & -I_n & 0_{n\times n(L-1)} & \lambda I_n - A \\ \hline \multicolumn{2}{c|}{C\mathcal{C}} & \multicolumn{2}{c|}{C\mathcal{C}_{\mathsf{w}}} & CA^L \end{array}\right], \quad (58)$$

wherein $\mathcal{D}_q := \begin{bmatrix} I_{q(L-1)} \\ 0_{q\times q} \end{bmatrix}$. Since the blocks $\lambda I_{mL} - \mathcal{D}_m$ and $\lambda I_{nL} - \mathcal{D}_n$ in (58) are non-singular for all $\lambda \neq 0$, to show that (58) has full column rank when $|\lambda| \geq 1$, we only need to verify the rank of the last block column in (58). Since (A, C) is observable, $\mathcal{O}_n := \mathrm{col}(C, CA, \ldots, CA^{n-1})$ has full column rank, so we have $\mathrm{Null}(\mathcal{O}_n A^L) = \mathrm{Null}(A^L)$ where Null denotes the null space. Note that $\mathcal{O}_n A^L$ is the observability matrix of the pair (A, CA^L), and thus $\mathrm{Null}(\mathcal{O}_n A^L)$ is the unobservable space of the pair (A, CA^L). Given $\mathrm{Null}(\mathcal{O}_n A^L) = \mathrm{Null}(A^L)$, all unobservable states x_{nobs} of (A, CA^L) satisfy $A^L x_{\mathsf{nobs}} = 0$ and hence are strictly stable, which implies that (A, CA^L) is detectable. From the Hautus lemma, $\mathrm{col}(\lambda I_n - A, CA^L)$ has full column rank for all λ that $|\lambda| \geq 1$. With diagonal blocks $\lambda I_{mL} - \mathcal{D}_m$, $\lambda I_{nL} - \mathcal{D}_n$ and $\mathrm{col}(\lambda I_n - A, CA^L)$ having full column rank, the matrix (58) has full column rank when $|\lambda| \geq 1$, and so does the pre-permutational matrix H_{obs}, which implies that $(\widetilde{A}, \widetilde{C})$ is detectable through Hautus lemma.

Next, we show that $\widetilde{P}_1 = \widetilde{P}_2$ with $\widetilde{P}_1 := \Phi_{\mathsf{orig}}^{\mathsf{T}} P \Phi_{\mathsf{orig}}$ and $\widetilde{P}_2 := \Phi_{\mathsf{aux}}^{\mathsf{T}} \mathbf{P} \Phi_{\mathsf{aux}}$, where P is the solution to (12) and $\mathbf{P}$ the solution to (39). By left- and right-multiplying (12) by Φ and Φ^{T} respectively, the resulting equation can be written as

$$\widetilde{P}_1 = \widetilde{A}^{\mathsf{T}} \widetilde{P}_1 (\widetilde{A} - \widetilde{B}\widetilde{K}_1) + \widetilde{C}^{\mathsf{T}} Q (\widetilde{C} - D\widetilde{K}_1) \quad (59)$$

wherein $\widetilde{K}_1 := (R + D^{\mathsf{T}}QD + \widetilde{B}^{\mathsf{T}}\widetilde{P}_1\widetilde{B})^{-1}(\widetilde{B}^{\mathsf{T}}\widetilde{P}_1\widetilde{A} + D^{\mathsf{T}}Q\widetilde{C})$, provided the definitions $\widetilde{C} := C\Phi_{\mathsf{orig}}$, $\widetilde{P}_1 := \Phi_{\mathsf{orig}}^{\mathsf{T}} P \Phi_{\mathsf{orig}}$ and the relations $A\Phi_{\mathsf{orig}} = \Phi_{\mathsf{orig}}\widetilde{A}$, $B = \Phi_{\mathsf{orig}}\widetilde{B}$ implied by Claim 5.1, Claim 5.2 and Claim 7.1. Similarly, left- and right-multiply (39) by $\Phi_{\mathsf{aux}}^{\mathsf{T}}$ and Φ_{aux} respectively, and the resulting equation can be written in the form

$$\widetilde{P}_2 = \widetilde{A}^{\mathsf{T}} \widetilde{P}_2 (\widetilde{A} - \widetilde{B}\widetilde{K}_2) + \widetilde{C}^{\mathsf{T}} Q (\widetilde{C} - D\widetilde{K}_2) \quad (60)$$

with $\widetilde{K}_2 := (R + D^{\mathsf{T}}QD + \widetilde{B}^{\mathsf{T}}\widetilde{P}_2\widetilde{B})^{-1}(\widetilde{B}^{\mathsf{T}}\widetilde{P}_2\widetilde{A} + D^{\mathsf{T}}Q\widetilde{C})$, given the definitions $\widetilde{C} := C\Phi_{\mathsf{orig}}$, $\widetilde{P}_2 := \Phi_{\mathsf{aux}}^{\mathsf{T}}\mathbf{P}\Phi_{\mathsf{aux}}$ and the relations $\mathbf{A}\Phi_{\mathsf{aux}} = \Phi_{\mathsf{aux}}\widetilde{A}$, $\mathbf{B} = \Phi_{\mathsf{aux}}\widetilde{B}$, $\mathbf{C}\Phi_{\mathsf{aux}} = C\Phi_{\mathsf{orig}}$ according to Claim 5.1 and Claim 7.1. Observing (59) and (60), we know that both $\widetilde{P}_1$ and $\widetilde{P}_2$ are (positive semi-definite) solutions to a similar DARE to (12) and (39), for dynamical system $(\widetilde{A}, \widetilde{B}, \widetilde{C}, D)$. In fact, this DARE has a unique positive semi-definite solution, given stabilizable $(\widetilde{A}, \widetilde{B})$ via Claim 5.3, detectable $(\widetilde{A}, \widetilde{C})$ as proved before and $Q \succ 0$. Hence, the solutions $\widetilde{P}_1, \widetilde{P}_2$ are equal.

Finally, we obtain the result by noting the relations $B^{\mathsf{T}}PA\Phi_{\mathsf{orig}} = \mathbf{B}^{\mathsf{T}}\mathbf{P}\mathbf{A}\Phi_{\mathsf{aux}}$ and $B^{\mathsf{T}}PB = \mathbf{B}^{\mathsf{T}}\mathbf{P}\mathbf{B}$, which can be verified given $\Phi_{\mathsf{orig}}^{\mathsf{T}} P \Phi_{\mathsf{orig}} = \Phi_{\mathsf{aux}}^{\mathsf{T}}\mathbf{P}\Phi_{\mathsf{aux}}$ (as $\widetilde{P}_1 = \widetilde{P}_2$) and given Claim 5.1, Claim 5.2 and Claim 7.1. It follows from the definitions (11), (38) of K and $\mathbf{K}$ that $K\Phi_{\mathsf{orig}} = \mathbf{K}\Phi_{\mathsf{aux}}$, which is the result given $\Phi_{\mathsf{orig}} = \Phi\Phi_{\mathsf{aux}}$ (Claim 5.1). ♦

We mention in Claim 7.3 some identities which will be used multiple times in the rest of the proof.

Claim 7.3. *If* $v \in \mathbb{R}^n$*,* $\boldsymbol{v} \in \mathbb{R}^{n_{\mathsf{aux}}}$ *and* $\tilde{v} \in \mathbb{R}^{n_\xi}$ *are such that* $v = \Phi\boldsymbol{v}$ *and* $\boldsymbol{v} = \Phi_{\mathsf{aux}}\tilde{v}$*, then*

$$Cv = \mathbf{C}\boldsymbol{v}, \quad Kv = \mathbf{K}\boldsymbol{v}, \quad Av = \Phi\mathbf{A}\boldsymbol{v}, \quad \mathbf{A}\boldsymbol{v} = \Phi_{\mathsf{aux}}\widetilde{A}\tilde{v}.$$

If $M \in \mathbb{S}_+^n$*,* $\mathbf{M} \in \mathbb{S}_+^{n_{\mathsf{aux}}}$ *and* $\widetilde{M} \in \mathbb{S}_+^{n_\xi}$ *are such that* $M = \Phi\mathbf{M}\Phi^{\mathsf{T}}$ *and* $\mathbf{M} = \Phi_{\mathsf{aux}}\widetilde{M}\Phi_{\mathsf{aux}}^{\mathsf{T}}$*, then*

$$CM = \mathbf{C}\mathbf{M}\Phi^{\mathsf{T}}, \ CMC^{\mathsf{T}} = \mathbf{C}\mathbf{M}\mathbf{C}^{\mathsf{T}}, \ CMK^{\mathsf{T}} = \mathbf{C}\mathbf{M}\mathbf{K}^{\mathsf{T}},$$
$$KM = \mathbf{K}\mathbf{M}\Phi^{\mathsf{T}}, \ KMK^{\mathsf{T}} = \mathbf{K}\mathbf{M}\mathbf{K}^{\mathsf{T}}.$$

Proof. Using $C\Phi\Phi_{\mathsf{aux}} = \mathbf{C}\Phi_{\mathsf{aux}}$ (Claim 7.1), we have $Cv = C\Phi\boldsymbol{v} = C\Phi\Phi_{\mathsf{aux}}\tilde{v} = \mathbf{C}\Phi_{\mathsf{aux}}\tilde{v} = \mathbf{C}\boldsymbol{v}$, and one can show $CM = \mathbf{C}\mathbf{M}\Phi^{\mathsf{T}}$ and $CMC^{\mathsf{T}} = \mathbf{C}\mathbf{M}\mathbf{C}^{\mathsf{T}}$ given the facts $M = \Phi\mathbf{M}\Phi^{\mathsf{T}} = \Phi\Phi_{\mathsf{aux}}\widetilde{M}\Phi_{\mathsf{aux}}^{\mathsf{T}}\Phi^{\mathsf{T}}$ and $\mathbf{M} = \Phi_{\mathsf{aux}}\widetilde{M}\Phi_{\mathsf{aux}}^{\mathsf{T}}$. Similarly, using $K\Phi\Phi_{\mathsf{aux}} = \mathbf{K}\Phi_{\mathsf{aux}}$ (Claim 7.2), we prove $Kv = \mathbf{K}\boldsymbol{v}$, $KM = \mathbf{K}\mathbf{M}\Phi^{\mathsf{T}}$, $KMK^{\mathsf{T}} = \mathbf{K}\mathbf{M}\mathbf{K}^{\mathsf{T}}$ and $CMK^{\mathsf{T}} = \mathbf{C}\mathbf{M}\mathbf{K}^{\mathsf{T}}$ in the same way by replacing $(C, \mathbf{C})$ into $(K, \mathbf{K})$. Using $A\Phi\Phi_{\mathsf{aux}} = \Phi\mathbf{A}\Phi_{\mathsf{aux}}$ (Claim 7.1) and $\mathbf{A}\Phi_{\mathsf{aux}} = \Phi_{\mathsf{aux}}\widetilde{A}$ (Claim 5.2), we show that $Av = A\Phi\boldsymbol{v} = A\Phi\Phi_{\mathsf{aux}}\tilde{v} = \Phi\mathbf{A}\Phi_{\mathsf{aux}}\tilde{v} = \Phi\mathbf{A}\boldsymbol{v}$ and also $\mathbf{A}\boldsymbol{v} = \mathbf{A}\Phi_{\mathsf{aux}}\tilde{v} = \Phi_{\mathsf{aux}}\widetilde{A}\tilde{v}$. ♦

In the following claim, the state variances $\Sigma^{\mathsf{x}}, \mathbf{\Sigma}^{\mathsf{x}}$, Kalman gains $L_{\mathsf{K}}, \mathbf{L}_{\mathsf{K}}$ and Luenberger gains $L_{\mathsf{L}}, \mathbf{L}_{\mathsf{L}}$ are related.

Claim 7.4. *For matrices* $\Sigma^{\mathsf{x}}, L_{\mathsf{K}}, L_{\mathsf{L}}$ *in* (6) *and* $\mathbf{\Sigma}^{\mathsf{x}}, \mathbf{L}_{\mathsf{K}}, \mathbf{L}_{\mathsf{L}}$ *in* (33)*, it holds that*

(a) $\Sigma^{\mathsf{x}} = \Phi\mathbf{\Sigma}^{\mathsf{x}}\Phi^{\mathsf{T}}$ *and* $\mathbf{\Sigma}^{\mathsf{x}} = \Phi_{\mathsf{aux}}\widetilde{\Sigma}^{\mathsf{x}}\Phi_{\mathsf{aux}}^{\mathsf{T}}$ *for some* $\widetilde{\Sigma}^{\mathsf{x}} \in \mathbb{S}_+^{n_\xi}$*;*
(b) $L_{\mathsf{K}} = \Phi\mathbf{L}_{\mathsf{K}}$ *and* $\mathbf{L}_{\mathsf{K}} = \Phi_{\mathsf{aux}}\widetilde{L}_{\mathsf{K}}$ *for some* $\widetilde{L}_{\mathsf{K}} \in \mathbb{R}^{n_\xi \times p}$*;*
(c) $L_{\mathsf{L}} = \Phi\mathbf{L}_{\mathsf{L}}$ *and* $\mathbf{L}_{\mathsf{L}} = \Phi_{\mathsf{aux}}\widetilde{L}_{\mathsf{L}}$ *for some* $\widetilde{L}_{\mathsf{L}} \in \mathbb{R}^{n_\xi \times p}$*.*

Proof. We first show $\mathbf{\Sigma}^{\mathsf{x}} = \mathbf{\Sigma}' := \Phi_{\mathsf{aux}}\widetilde{\Sigma}^{\mathsf{x}}\Phi_{\mathsf{aux}}^{\mathsf{T}}$ in (a). Let $\widetilde{C} := C\Phi_{\mathsf{orig}}$ and let $\widetilde{A}, \widetilde{B}$ be as in Claim 5.2. Since $(\widetilde{A}, \widetilde{\Sigma}^{\mathsf{w}})$

is stabilizable through Claim 5.3 and $(\widetilde{A}, \widetilde{C})$ is detectable as shown in the proof of Claim 7.2, the DARE

$$\widetilde{\Sigma}^{\mathsf{x}} = \widetilde{A}\widetilde{\Sigma}^{\mathsf{x}}\widetilde{A}^{\mathsf{T}} + \widetilde{\Sigma}^{\mathsf{w}} - \widetilde{A}\widetilde{\Sigma}^{\mathsf{x}}\widetilde{C}^{\mathsf{T}}(\widetilde{C}\widetilde{\Sigma}^{\mathsf{x}}\widetilde{C}^{\mathsf{T}} + \Sigma^{\mathsf{v}})^{-1}\widetilde{C}\widetilde{\Sigma}^{\mathsf{x}}\widetilde{A}^{\mathsf{T}} \quad (61)$$

has a unique positive semi-definite solution $\widetilde{\Sigma}^{\mathsf{x}}$. Left- and right-multiply (61) by Φ_{aux} and by $\Phi_{\mathsf{aux}}^{\mathsf{T}}$ respectively, and the resulting equation can be written in the form

$$\mathbf{\Sigma}' = \mathbf{A}\mathbf{\Sigma}'\mathbf{A}^{\mathsf{T}} + \mathbf{\Sigma}^{\mathsf{w}} - \mathbf{A}\mathbf{\Sigma}'\mathbf{C}^{\mathsf{T}}(\mathbf{C}\mathbf{\Sigma}'\mathbf{C}^{\mathsf{T}} + \Sigma^{\mathsf{v}})^{-1}\mathbf{C}\mathbf{\Sigma}'\widetilde{A}^{\mathsf{T}} \quad (62)$$

with substitutions $\Phi_{\mathsf{aux}}\widetilde{A} = \mathbf{A}\Phi_{\mathsf{aux}}$, $\mathbf{\Sigma}^{\mathsf{w}} = \Phi_{\mathsf{aux}}\widetilde{\Sigma}^{\mathsf{w}}\Phi_{\mathsf{aux}}^{\mathsf{T}}$ and $\widetilde{C} = C\Phi_{\mathsf{orig}} = C\Phi\Phi_{\mathsf{aux}} = \mathbf{C}\Phi_{\mathsf{aux}}$ via Claim 5.1 and Claim 7.1. Due to (62), $\mathbf{\Sigma}'$ is a positive semi-definite solution to the DARE (33a). Since (33a) has a unique positive semi-definite solution $\mathbf{\Sigma}^{\mathsf{x}}$, we have $\mathbf{\Sigma}^{\mathsf{x}} = \mathbf{\Sigma}'$.

Next, we show $\Sigma^{\mathsf{x}} = \Sigma' := \Phi\mathbf{\Sigma}^{\mathsf{x}}\Phi^{\mathsf{T}}$ in (a). Left- and right-multiply (33a) by Φ and by Φ^{T} respectively, and the resulting equality can be written as

$$\Sigma' = A\Sigma'A^{\mathsf{T}} + \Sigma^{\mathsf{w}} - A\Sigma'C^{\mathsf{T}}(C\Sigma'C^{\mathsf{T}} + \Sigma^{\mathsf{v}})^{-1}C\Sigma'A^{\mathsf{T}} \quad (63)$$

given $\Phi\mathbf{\Sigma}^{\mathsf{w}}\Phi^{\mathsf{T}} = \Sigma^{\mathsf{w}}$ (Claim 7.1) and the substitutions $\Phi\mathbf{A}\mathbf{\Sigma}^{\mathsf{x}} = A\Phi\mathbf{\Sigma}^{\mathsf{x}}$ and $\mathbf{C}\mathbf{\Sigma}^{\mathsf{x}} = C\Phi\mathbf{\Sigma}^{\mathsf{x}}$, which are implied by $\Phi\mathbf{A}\Phi_{\mathsf{aux}} = A\Phi\Phi_{\mathsf{aux}}$ and $\mathbf{C}\Phi_{\mathsf{aux}} = C\Phi\Phi_{\mathsf{aux}}$ (Claim 7.1) respectively, provided $\mathbf{\Sigma}^{\mathsf{x}} = \Phi_{\mathsf{aux}}\widetilde{\Sigma}^{\mathsf{x}}\Phi_{\mathsf{aux}}^{\mathsf{T}}$. Due to (63), Σ' is a positive semi-definite solution to the DARE (6a). Since (6a) has a unique positive definite solution Σ^{x}, we have $\Sigma^{\mathsf{x}} = \Sigma'$.

We finally show (b) and (c). Given the definitions (33b) of $\mathbf{L}_{\mathsf{K}}$ and $\mathbf{L}_{\mathsf{L}}$, we obtain $\mathbf{L}_{\mathsf{K}} = \Phi_{\mathsf{aux}}\widetilde{L}_{\mathsf{K}}$ and $\mathbf{L}_{\mathsf{L}} = \Phi_{\mathsf{aux}}\widetilde{L}_{\mathsf{L}}$ as

$$\begin{aligned}\mathbf{L}_{\mathsf{K}} &:= \mathbf{\Sigma}^{\mathsf{x}}\mathbf{C}^{\mathsf{T}}(\mathbf{C}\mathbf{\Sigma}^{\mathsf{x}}\mathbf{C}^{\mathsf{T}} + \Sigma^{\mathsf{v}})^{-1} \\ &= \Phi_{\mathsf{aux}}\widetilde{\Sigma}^{\mathsf{x}}\Phi_{\mathsf{aux}}^{\mathsf{T}}\mathbf{C}^{\mathsf{T}}(\mathbf{C}\mathbf{\Sigma}^{\mathsf{x}}\mathbf{C}^{\mathsf{T}} + \Sigma^{\mathsf{v}})^{-1} = \Phi_{\mathsf{aux}}\widetilde{L}_{\mathsf{K}} \\ \mathbf{L}_{\mathsf{L}} &:= \mathbf{A}\mathbf{L}_{\mathsf{K}} = \mathbf{A}\Phi_{\mathsf{aux}}\widetilde{L}_{\mathsf{K}} = \Phi_{\mathsf{aux}}\widetilde{A}\widetilde{L}_{\mathsf{K}} = \Phi_{\mathsf{aux}}\widetilde{L}_{\mathsf{L}}\end{aligned}$$

with $\widetilde{L}_{\mathsf{K}} := \widetilde{\Sigma}^{\mathsf{x}}\Phi_{\mathsf{aux}}^{\mathsf{T}}\mathbf{C}^{\mathsf{T}}(\mathbf{C}\mathbf{\Sigma}^{\mathsf{x}}\mathbf{C}^{\mathsf{T}} + \Sigma^{\mathsf{v}})^{-1}$ and $\widetilde{L}_{\mathsf{L}} := \widetilde{A}\widetilde{L}_{\mathsf{K}}$, where we used $\mathbf{\Sigma}^{\mathsf{x}} = \Phi_{\mathsf{aux}}\widetilde{\Sigma}^{\mathsf{x}}\Phi_{\mathsf{aux}}^{\mathsf{T}}$ in (a) and $\mathbf{A}\Phi_{\mathsf{aux}} = \Phi_{\mathsf{aux}}\widetilde{A}$ in Claim 5.2. With definitions (6b), (33b) of $L_{\mathsf{K}}, L_{\mathsf{L}}, \mathbf{L}_{\mathsf{K}}, \mathbf{L}_{\mathsf{L}}$, we have

$$\begin{aligned}L_{\mathsf{K}} &:= \Sigma^{\mathsf{x}}C^{\mathsf{T}}(C\Sigma^{\mathsf{x}}C^{\mathsf{T}} + \Sigma^{\mathsf{v}})^{-1} \\ &= \Phi\mathbf{\Sigma}^{\mathsf{x}}\mathbf{C}^{\mathsf{T}}(\mathbf{C}\mathbf{\Sigma}^{\mathsf{x}}\mathbf{C}^{\mathsf{T}} + \Sigma^{\mathsf{v}})^{-1} = \Phi\mathbf{L}_{\mathsf{K}} \\ L_{\mathsf{L}} &:= AL_{\mathsf{K}} = \Phi\mathbf{A}\mathbf{L}_{\mathsf{K}} = \Phi\mathbf{L}_{\mathsf{L}},\end{aligned}$$

where we used $C\Sigma^{\mathsf{x}} = \mathbf{C}\mathbf{\Sigma}^{\mathsf{x}}\Phi^{\mathsf{T}}$ and $C\widetilde{\Sigma}^{\mathsf{x}}C^{\mathsf{T}} = \mathbf{C}\mathbf{\Sigma}^{\mathsf{x}}\mathbf{C}^{\mathsf{T}}$ through Claim 7.3 with selection $(M, \mathbf{M}, \widetilde{M}) \leftarrow (\Sigma^{\mathsf{x}}, \mathbf{\Sigma}^{\mathsf{x}}, \widetilde{\Sigma}^{\mathsf{x}})$ given (a), and used $AL_{\mathsf{K}} = \Phi\mathbf{A}\mathbf{L}_{\mathsf{K}}$ implied by $Av = \Phi\mathbf{A}\boldsymbol{v}$ from Claim 7.3 where $v, \boldsymbol{v}, \tilde{v}$ are chosen as the i-th columns of $L_{\mathsf{K}}, \mathbf{L}_{\mathsf{K}}, \widetilde{L}_{\mathsf{K}}$, respectively, for $i \in \{1, \ldots, p\}$. ♦

B. Proof of Proposition 7

Proof. We first show in Claim 7.5 that the matrices Δ_s and $\mathbf{\Delta}_s$ are identical, and then in Claim 7.6(c) that the nominal outputs $\bar{y}_t$ and $\bar{\mathbf{y}}_t$ are equal.

Claim 7.5. *For matrices Δ_s in (14a) and $\mathbf{\Delta}_s$ in (40a), we have $\Delta_s = \mathbf{\Delta}_s$ for $s \in \mathbb{Z}_{[0,N)}$.*

Proof. We first show $\Lambda_s = \Phi\mathbf{\Lambda}_s\Phi^{\mathsf{T}}$ and $\mathbf{\Lambda}_s = \Phi_{\mathsf{aux}}\widetilde{\Lambda}_s\Phi_{\mathsf{aux}}^{\mathsf{T}}$ for $s \in \mathbb{Z}_{[0,N)}$, where Λ_s and $\mathbf{\Lambda}_s$ are defined in (14b) and (40b), and let $\widetilde{\Lambda}_s := \sum_{r=0}^{s}(\widetilde{A} - \widetilde{B}\widetilde{K})^r\,\widetilde{L}_{\mathsf{L}}(\mathbf{C}\mathbf{\Sigma}^{\mathsf{x}}\mathbf{C}^{\mathsf{T}} + \Sigma^{\mathsf{v}})\widetilde{L}_{\mathsf{L}}(\widetilde{A} - \widetilde{B}\widetilde{K})^{r\mathsf{T}}$, with $\widetilde{L}_{\mathsf{L}}$ in Claim 7.4 and $\widetilde{K} := \mathbf{K}\Phi_{\mathsf{aux}}$. Given Claim 7.4(c), the definitions of $\Lambda_s, \mathbf{\Lambda}_s$, and the identity $C\Sigma^{\mathsf{x}}C^{\mathsf{T}} = \mathbf{C}\mathbf{\Sigma}^{\mathsf{x}}\mathbf{C}^{\mathsf{T}}$ shown in the proof Claim 7.4(b-c), it suffices to show

$$\begin{aligned}(A - BK)^r\,\Phi\Phi_{\mathsf{aux}} &= \Phi\,(\mathbf{A} - \mathbf{B}\mathbf{K})^r\,\Phi_{\mathsf{aux}} \\ (\mathbf{A} - \mathbf{B}\mathbf{K})^r\,\Phi_{\mathsf{aux}} &= \Phi_{\mathsf{aux}}\,(\widetilde{A} - \widetilde{B}\widetilde{K})^r\end{aligned}$$

for all $r \in \mathbb{N}_{\geq 0}$, which can be obtained by repeatedly applying $(A - BK)\Phi\Phi_{\mathsf{aux}} = \Phi(\mathbf{A} - \mathbf{B}\mathbf{K})\Phi_{\mathsf{aux}}$ and $(\mathbf{A} - \mathbf{B}\mathbf{K})\Phi_{\mathsf{aux}} = \Phi_{\mathsf{aux}}(\widetilde{A} - \widetilde{B}\widetilde{K})$ respectively, which follow from $\mathbf{A} = \Phi_{\mathsf{aux}}\widetilde{A}$ and $\mathbf{B} = \Phi_{\mathsf{aux}}\widetilde{B}$ in Claim 5.2, $A\Phi\Phi_{\mathsf{aux}} = \Phi\mathbf{A}\Phi_{\mathsf{aux}}$ and $B = \Phi\mathbf{B}$ in Claim 7.1, and $K\Phi\Phi_{\mathsf{aux}} = \mathbf{K}\Phi_{\mathsf{aux}}$ in Claim 7.2.

We finally show $\Delta_s = \mathbf{\Delta}_s$ for $s \in \mathbb{Z}_{[0,N)}$. Given the definitions of Δ_s and $\mathbf{\Delta}_s$ and the relation $C\Sigma^{\mathsf{x}}C^{\mathsf{T}} = \mathbf{C}\mathbf{\Sigma}^{\mathsf{x}}\mathbf{C}^{\mathsf{T}}$, it suffices to show the relations $C\Lambda_sC^{\mathsf{T}} = \mathbf{C}\mathbf{\Lambda}_s\mathbf{C}^{\mathsf{T}}$, $K\Lambda_sK^{\mathsf{T}} = \mathbf{K}\mathbf{\Lambda}_s\mathbf{K}^{\mathsf{T}}$ and $C\Lambda_sK^{\mathsf{T}} = \mathbf{C}\mathbf{\Lambda}_s\mathbf{K}^{\mathsf{T}}$, which are obtained through Claim 7.3 with selection $(M, \mathbf{M}, \widetilde{M}) \leftarrow (\Lambda_s, \mathbf{\Lambda}_s, \widetilde{\Lambda}_s)$ given $\Lambda_s = \Phi\mathbf{\Lambda}_s\Phi^{\mathsf{T}}$ and $\mathbf{\Lambda}_s = \Phi_{\mathsf{aux}}\widetilde{\Lambda}_s\Phi_{\mathsf{aux}}^{\mathsf{T}}$ as proved. ♦

Claim 7.6. *If parameters $\mu_k^{\hat{\mathsf{x}}}, \mu_k^{\tilde{\mathsf{x}}}, \boldsymbol{\mu}_k^{\tilde{\mathsf{x}}}, \boldsymbol{\mu}_k^{\hat{\mathsf{x}}}$ satisfy (44) for some $\widetilde{\mu}_k^{\hat{\mathsf{x}}}, \widetilde{\mu}_k^{\tilde{\mathsf{x}}} \in \mathbb{R}^{n_\xi}$, then, for all $\bar{u}_{[k,k+N)}$ and θ, we have*

(a) $\mu_k^{\mathsf{x}} = \Phi\boldsymbol{\mu}_k^{\mathsf{x}}$ and $\boldsymbol{\mu}_k^{\mathsf{x}} = \Phi_{\mathsf{aux}}\widetilde{\mu}_k^{\mathsf{x}}$ with some $\widetilde{\mu}_k^{\mathsf{x}} \in \mathbb{R}^{n_\xi}$,

and, for all $t \in \mathbb{Z}_{[k,k+N)}$, we have

(b) $\bar{x}_t = \Phi\,\bar{\mathbf{x}}_t$ and $\bar{\mathbf{x}}_t = \Phi_{\mathsf{aux}}\,\widetilde{\bar{x}}_t$ with some $\widetilde{\bar{x}}_t \in \mathbb{R}^{n_\xi}$,

(c) $\bar{y}_t = \bar{\mathbf{y}}_t$.

Proof. To prove (a), we obtain $\mu_k^{\mathsf{x}} = \Phi_{\mathsf{orig}}\widetilde{\mu}_k^{\mathsf{x}}$ by combining (9) and (44), and obtain $\boldsymbol{\mu}_k^{\mathsf{x}} = \Phi_{\mathsf{aux}}\widetilde{\mu}_k^{\mathsf{x}}$ by combining (32) and (44), where we let $\widetilde{\mu}_k^{\mathsf{x}} := (1-\theta)\widetilde{\mu}_k^{\hat{\mathsf{x}}} + \theta\widetilde{\mu}_k^{\tilde{\mathsf{x}}}$. Then, $\mu_k^{\mathsf{x}} = \Phi\boldsymbol{\mu}_k^{\mathsf{x}}$ follows from $\Phi_{\mathsf{orig}} = \Phi\Phi_{\mathsf{aux}}$.

(b) is proved by induction. **Base Case.** Select $\widetilde{\bar{x}}_k := \widetilde{\mu}_k^{\mathsf{x}}$. The $t = k$ case of (b) follows from (a) and relations $\bar{x}_k := \mu_k^{\mathsf{x}}$ as (7c) and $\bar{\mathbf{x}}_k := \boldsymbol{\mu}_k^{\mathsf{x}}$ as (36c). **Inductive Step.** Assume the $t = \tau$ case of (b) for some $\tau \in \mathbb{Z}_{[k,k+N-2]}$, and thus we have

$$\bar{x}_{\tau+1} \overset{\text{via (7a)}}{=} A\bar{x}_\tau + B\bar{u}_\tau = \Phi\mathbf{A}\bar{\mathbf{x}}_\tau + \Phi\mathbf{B}\bar{u}_\tau \overset{\text{via (36a)}}{=} \Phi\bar{\mathbf{x}}_{\tau+1},$$

where the second equality used $B = \Phi\mathbf{B}$ in Claim 7.1 and $A\bar{x}_\tau = \Phi\mathbf{A}\bar{\mathbf{x}}_\tau$ through Claim 7.3 with selection $(v, \boldsymbol{v}, \tilde{v}) \leftarrow (\bar{x}_\tau, \bar{\mathbf{x}}_\tau, \widetilde{\bar{x}}_\tau)$ given (b) of $t = \tau$. Moreover, we have

$$\bar{\mathbf{x}}_{\tau+1} \overset{\text{via (36a)}}{=} \mathbf{A}\bar{\mathbf{x}}_\tau + \mathbf{B}\bar{u}_\tau = \Phi_{\mathsf{aux}}\widetilde{A}\,\widetilde{\bar{x}}_\tau + \Phi_{\mathsf{aux}}\widetilde{B}\bar{u}_\tau = \Phi_{\mathsf{aux}}\widetilde{\bar{x}}_{\tau+1}$$

by choosing $\widetilde{\bar{x}}_{\tau+1} := \widetilde{A}\,\widetilde{\bar{x}}_\tau + \widetilde{B}\bar{u}_\tau$, where the second equality used $\mathbf{B} = \Phi_{\mathsf{aux}}\widetilde{B}$ in Claim 5.2 and $\mathbf{A}\bar{\mathbf{x}}_\tau = \Phi_{\mathsf{aux}}\widetilde{A}\,\widetilde{\bar{x}}_\tau$ through Claim 7.3 with $(v, \boldsymbol{v}, \tilde{v}) \leftarrow (\bar{x}_\tau, \bar{\mathbf{x}}_\tau, \widetilde{\bar{x}}_\tau)$ given (b) of $t = \tau$. Thus, we have the $t = \tau + 1$ case of (b). This shows (b).

Last, we have (c) $\bar{y}_t \overset{\text{via (7b)}}{=} C\bar{x}_t + D\bar{u}_t = \mathbf{C}\bar{\mathbf{x}}_t + D\bar{u}_t \overset{\text{via (36b)}}{=} \bar{\mathbf{y}}_t$ using $C\bar{x}_t = \mathbf{C}\bar{\mathbf{x}}_t$ through Claim 7.3 with selection $(v, \boldsymbol{v}, \tilde{v}) \leftarrow (\bar{x}_t, \bar{\mathbf{x}}_t, \widetilde{\bar{x}}_t)$ given (b). ♦

With Claim 7.5 and Claim 7.6(c), the objective functions of problems (19) and (41) are equal, and the constraint (17) in problem (19) and the constraint (42) in problem (41) are equivalent. Thus, problems (19) and (41) have the same objective function and constraints, and the result follows. ■

APPENDIX E. PROOF OF THEOREM 9

Proof. Let $\{x_t^{\mathsf{a}}, u_t^{\mathsf{a}}, y_t^{\mathsf{a}}\}$ denote the trajectory produced by process a), and $\{x_t^{\mathsf{b}}, u_t^{\mathsf{b}}, y_t^{\mathsf{b}}\}$ the trajectory from process b). We have the following intermediate result.

Claim 9.1. *Consider a control step $k = \kappa$ in both processes a) and b). Assume that*

i) the states $x_\kappa^{\mathsf{a}}, x_\kappa^{\mathsf{b}}$ in processes a) and b) are equal, and

ii) the parameters $\mu_\kappa^{\hat{\mathsf{x}}}, \mu_\kappa^{\bar{\mathsf{x}}}$ in process a) and parameters $\boldsymbol{\mu}_\kappa^{\hat{\mathsf{x}}}, \boldsymbol{\mu}_\kappa^{\bar{\mathsf{x}}}$ in process b) satisfy (44) with $k = \kappa$.

Let $\kappa^+ := \kappa + N_{\mathrm{c}}$. Then, for $t \in \mathbb{Z}_{[\kappa,\kappa^+]}$, we have

(a) the states $x_t^{\mathsf{a}}, x_t^{\mathsf{b}}$ in processes a) and b) are equal,

(b) the variable $\hat{x}_t^-$ in process a) and variable $\hat{\mathbf{x}}_t^-$ in process b) satisfy $\hat{x}_t^- = \Phi\hat{\mathbf{x}}_t^-$ and $\hat{\mathbf{x}}_t^- = \Phi_{\mathsf{aux}}\widetilde{\hat{x}}_t^-$ for some $\widetilde{\hat{x}}_t^- \in \mathbb{R}^{n_\xi}$,

and, for $t \in \mathbb{Z}_{[\kappa,\kappa^+)}$, we have

(c) the inputs $u_t^{\mathsf{a}}, u_t^{\mathsf{b}}$ in processes a) and b) are equal,

(d) the outputs $y_t^{\mathsf{a}}, y_t^{\mathsf{b}}$ in processes a) and b) are equal.

Moreover, at the next control step $k = \kappa^+$, we have

(e) the parameters $\mu_{\kappa^+}^{\hat{\mathsf{x}}}, \mu_{\kappa^+}^{\bar{\mathsf{x}}}$ in process a) and parameters $\boldsymbol{\mu}_{\kappa^+}^{\hat{\mathsf{x}}}, \boldsymbol{\mu}_{\kappa^+}^{\bar{\mathsf{x}}}$ in process b) satisfy (44) with $k = \kappa^+$.

Proof. We prove (a)-(d) by induction. **Base Case:** we show (a) and (b) for $t = \kappa$. Result (a) of $t = \kappa$ is exactly as condition i). Through Proposition 7 and the fact that both problems (19) and (41) produce unique optimal θ, the values of θ are the same in processes a) and b). Given condition ii), μ_κ^{x} in process a) and $\boldsymbol{\mu}_\kappa^{\mathsf{x}}$ in process b) satisfy $\mu_\kappa^{\mathsf{x}} = \Phi\boldsymbol{\mu}_k^{\mathsf{x}}$ and $\boldsymbol{\mu}_\kappa^{\mathsf{x}} = \Phi_{\mathsf{aux}}\widetilde{\mu}_\kappa^{\mathsf{x}}$ for some $\widetilde{\mu}_\kappa^{\mathsf{x}}$ according to Claim 7.6. Combining these relations with $\hat{x}_\kappa^- := \mu_\kappa^{\mathsf{x}}$ as (8c) and $\hat{\mathbf{x}}_\kappa^- := \boldsymbol{\mu}_\kappa^{\mathsf{x}}$ as (35c), we obtain (b) of $t = \kappa$ by choosing $\widetilde{\hat{x}}_\kappa^- := \widetilde{\mu}_\kappa^{\mathsf{x}}$, as

$$\hat{x}_\kappa^- = \mu_\kappa^{\mathsf{x}} = \Phi\boldsymbol{\mu}_\kappa^{\mathsf{x}} = \Phi\hat{\mathbf{x}}_\kappa^-, \quad \hat{\mathbf{x}}_\kappa^- = \boldsymbol{\mu}_\kappa^{\mathsf{x}} = \Phi_{\mathsf{aux}}\widetilde{\mu}_\kappa^{\mathsf{x}} = \Phi_{\mathsf{aux}}\widetilde{\hat{x}}_\kappa^-.$$

Inductive Step: we assume (a) and (b) for $t = \tau \in \mathbb{Z}_{[\kappa,\kappa^+)}$, and then prove (c), (d) for $t = \tau$ and (a), (b) for $t = \tau + 1$. The control inputs u_τ^{a}, u_τ^{b} are obtained through (10) and (37) respectively, where the nominal inputs $\bar{u}_\tau$ are the same according to Proposition 7 and the fact that both problems (19), (41) produce a unique optimal $\bar{u}$, i.e.,

$$u_\tau^{\mathsf{a}} = \bar{u}_\tau - K(\hat{x}_\tau - \bar{x}_\tau), \quad u_\tau^{\mathsf{b}} = \bar{u}_\tau - \mathbf{K}(\hat{\mathbf{x}}_\tau - \bar{\mathbf{x}}_\tau).$$

Thus, we have (c) $u_\tau^{\mathsf{a}} = u_\tau^{\mathsf{b}}$ of $t = \tau$, because of $K\hat{x}_\tau = \mathbf{K}\hat{\mathbf{x}}_\tau$ and $K\bar{x}_\tau = \mathbf{K}\bar{\mathbf{x}}_\tau$ through Claim 7.3 where we choose $(v, \boldsymbol{v}, \tilde{v})$ as $(\hat{x}_\tau^-, \hat{\mathbf{x}}_\tau^-, \widetilde{\hat{x}}_\tau^-)$ and $(\bar{x}_\tau, \bar{\mathbf{x}}_\tau, \widetilde{\bar{x}}_\tau)$, given (b) of $t = \tau$ and Claim 7.6(b) of $t = \tau$. We then have (d) $y_\tau^{\mathsf{a}} = y_\tau^{\mathsf{b}}$ for $t = \tau$ and (a) $x_{\tau+1}^{\mathsf{a}} = x_{\tau+1}^{\mathsf{b}}$ for $t = \tau + 1$, given the system model $y_\tau^{\mathsf{z}} = Cx_\tau^{\mathsf{z}} + Du_\tau^{\mathsf{z}} + v_t$ as (1b) and $x_{\tau+1}^{\mathsf{z}} = Ax_\tau^{\mathsf{z}} + Bu_\tau^{\mathsf{z}} + w_t$ as (1a), for $\mathsf{z} \in \{\mathsf{a}, \mathsf{b}\}$. Finally, we prove (b) for $t = \tau + 1$ as

$$\begin{aligned}
&\hat{x}_{\tau+1}^- \overset{\text{via (8)}}{=} A\hat{x}_\tau^- + Bu_\tau^{\mathsf{a}} + L_{\mathsf{L}}(y_\tau^{\mathsf{a}} - C\hat{x}_\tau^-) \\
&= \Phi\mathbf{A}\hat{\mathbf{x}}_\tau + \Phi\mathbf{B}u_\tau^{\mathsf{b}} + \Phi\mathbf{L}_{\mathsf{L}}(y_\tau^{\mathsf{b}} - \mathbf{C}\hat{\mathbf{x}}_\tau^-) \overset{\text{via (35)}}{=} \Phi\hat{\mathbf{x}}_{\tau+1}^- \\
&\hat{\mathbf{x}}_{\tau+1}^- \overset{\text{via (35)}}{=} \mathbf{A}\hat{\mathbf{x}}_\tau + \mathbf{B}u_\tau^{\mathsf{b}} + \mathbf{L}_{\mathsf{L}}(y_\tau^{\mathsf{b}} - \mathbf{C}\hat{\mathbf{x}}_\tau^-) \\
&= \Phi_{\mathsf{aux}}\widetilde{A}\widetilde{\hat{x}}_\tau + \Phi_{\mathsf{aux}}\widetilde{B}u_\tau^{\mathsf{b}} + \Phi_{\mathsf{aux}}\widetilde{L}_{\mathsf{L}}(y_\tau^{\mathsf{b}} - \mathbf{C}\hat{\mathbf{x}}_\tau^-) = \Phi_{\mathsf{aux}}\widetilde{\hat{x}}_{\tau+1}^-
\end{aligned}$$

by choosing $\widetilde{\hat{x}}_{\tau+1}^- := \widetilde{A}\widetilde{\hat{x}}_\tau + \widetilde{B}u_\tau^{\mathsf{b}} + \widetilde{L}_{\mathsf{L}}(y_\tau^{\mathsf{b}} - \mathbf{C}\hat{\mathbf{x}}_\tau^-)$, where we used $B = \Phi\mathbf{B}$ in Claim 7.1, $\mathbf{B} = \Phi_{\mathsf{aux}}\widetilde{B}$ in Claim 5.2, $L_{\mathsf{L}} = \Phi\mathbf{L}_{\mathsf{L}}$ and $\mathbf{L}_{\mathsf{L}} = \Phi\widetilde{L}_{\mathsf{L}}$ in Claim 7.4, and $A\hat{x}_\tau^- = \Phi\mathbf{A}\hat{\mathbf{x}}_\tau^-$ and $\mathbf{A}\hat{\mathbf{x}}_\tau^- = \Phi_{\mathsf{aux}}\widetilde{A}\widetilde{\hat{x}}_\tau^-$ by applying Claim 7.3 with $(v, \boldsymbol{v}, \tilde{v}) \leftarrow (\hat{x}_\tau^-, \hat{\mathbf{x}}_\tau^-, \widetilde{\hat{x}}_\tau^-)$ given (b) of $t = \tau$. By induction on t, (a) and (b) hold for $t \in \mathbb{Z}_{[\kappa,\kappa^+]}$, and (c) and (d) hold for $t \in \mathbb{Z}_{[\kappa,\kappa^+)}$.

We finally show (e). Notice the following relations,

$$\hat{x}_{\kappa^+}^- = \Phi\Phi_{\mathsf{aux}}\widetilde{\hat{x}}_{\kappa^+}^-, \qquad \hat{\mathbf{x}}_{\kappa^+}^- = \Phi_{\mathsf{aux}}\widetilde{\hat{x}}_{\kappa^+}^- \tag{64a}$$

$$\bar{x}_{\kappa^+} = \Phi\Phi_{\mathsf{aux}}\widetilde{\bar{x}}_{\kappa^+}, \qquad \bar{\mathbf{x}}_{\kappa^+} = \Phi_{\mathsf{aux}}\widetilde{\bar{x}}_{\kappa^+} \tag{64b}$$

where (64a) is due to (b) with $t = \kappa^+$, and (64b) follows from Claim 7.6 with $k = \kappa$ and $t = \kappa^+$. According to (20) applied in Algorithm 1 and (45) applied in Algorithm 2, we have

$$\mu_{\kappa^+}^{\hat{\mathsf{x}}} = \hat{x}_{\kappa^+}^-, \quad \mu_{\kappa^+}^{\bar{\mathsf{x}}} = \bar{x}_{\kappa^+}, \quad \boldsymbol{\mu}_{\kappa^+}^{\hat{\mathsf{x}}} = \hat{\mathbf{x}}_{\kappa^+}^-, \quad \boldsymbol{\mu}_{\kappa^+}^{\bar{\mathsf{x}}} = \bar{\mathbf{x}}_{\kappa^+}. \tag{65}$$

Combining (64) and (65), with $\Phi_{\mathsf{orig}} = \Phi\Phi_{\mathsf{aux}}$ via Claim 5.1, we obtain (44) with $k = \kappa^+$ where we select $\widetilde{\mu}_{\kappa^+}^{\hat{\mathsf{x}}} := \widetilde{\hat{x}}_{\kappa^+}^-$ and $\widetilde{\mu}_{\kappa^+}^{\bar{\mathsf{x}}} := \widetilde{\bar{x}}_{\kappa^+}$. This shows (e). ♦

We finish the proof by showing that the results (a)-(e) in Claim 9.1 are true for all control steps $\kappa \in \{0, N_{\mathrm{c}}, 2N_{\mathrm{c}}, \ldots\}$, by induction on κ. **Base Case.** For $\kappa = 0$, condition i) of Claim 9.1 holds given that both processes start with a common initial state x_0, and condition ii) of Claim 9.1 holds due to Assumption 8(d) and due to the selections $(\mu_0^{\hat{\mathsf{x}}}, \mu_0^{\bar{\mathsf{x}}}) \leftarrow (\mu_{\mathsf{ini}}^{\mathsf{x}}, \mu_{\mathsf{ini}}^{\mathsf{x}})$ in Algorithm 1 and $(\boldsymbol{\mu}_0^{\hat{\mathsf{x}}}, \boldsymbol{\mu}_0^{\bar{\mathsf{x}}}) \leftarrow (\boldsymbol{\mu}_{\mathsf{ini}}^{\mathsf{x}}, \boldsymbol{\mu}_{\mathsf{ini}}^{\mathsf{x}})$ in Algorithm 2. With both conditions i) and ii) satisfied, the results (a)-(e) of Claim 9.1 are true for $\kappa = 0$. **Inductive Step.** Assume for $\kappa = \overline{\kappa}$ that results (a)-(e) of Claim 9.1 are true. Due to (a) and (e) of Claim 9.1 for $\kappa = \overline{\kappa}$, the assumptions i) and ii) in Claim 9.1 for $\kappa = \overline{\kappa} + N_{\mathrm{c}}$ are satisfied, thereby ensuring that the results (a)-(e) of Claim 9.1 with $\kappa = \overline{\kappa} + N_{\mathrm{c}}$ are true. By induction on κ, we have the results (a)-(e) of Claim 9.1 for all control steps $\kappa \in \{0, N_{\mathrm{c}}, 2N_{\mathrm{c}}, \ldots\}$. The results (a), (c), (d) for all κ suffices to prove the theorem. ■

REFERENCES

[1] D. Q. Mayne, "Model predictive control: Recent developments and future promise," *Automatica*, vol. 50, no. 12, pp. 2967–2986, 2014.

[2] A. Mesbah, "Stochastic model predictive control: An overview and perspectives for future research," *IEEE Control Syst. Mag.*, vol. 36, no. 6, pp. 30–44, 2016.

[3] T. A. N. Heirung, J. A. Paulson, J. O'Leary, and A. Mesbah, "Stochastic model predictive control–how does it work?" *Comput Chem Eng*, vol. 114, pp. 158–170, 2018.

[4] M. Farina, L. Giulioni, and R. Scattolini, "Stochastic linear model predictive control with chance constraints–a review," *J. Process Control*, vol. 44, pp. 53–67, 2016.

[5] R. Kumar, J. Jalving, M. J. Wenzel, M. J. Ellis, M. N. ElBsat, K. H. Drees, and V. M. Zavala, "Benchmarking stochastic and deterministic mpc: A case study in stationary battery systems," *AIChE Journal*, vol. 65, no. 7, p. e16551, 2019.

[6] A. Bemporad and M. Morari, "Robust model predictive control: A survey," in *Robustness in identification and control*. Springer, 2007, pp. 207–226.

[7] I. Markovsky and F. Dörfler, "Behavioral systems theory in data-driven analysis, signal processing, and control," *Annu. Rev. Control*, vol. 52, pp. 42–64, 2021.

[8] I. Markovsky, L. Huang, and F. Dörfler, "Data-driven control based on the behavioral approach: From theory to applications in power systems," *IEEE Control Syst.*, vol. 43, no. 5, pp. 28–68, 2023.

[9] F. Dörfler, "Data-driven control: Part two of two: Hot take: Why not go with models?" *IEEE Control Syst. Mag.*, vol. 43, no. 6, pp. 27–31, 2023.

[10] Z.-S. Hou and Z. Wang, "From model-based control to data-driven control: Survey, classification and perspective," *Inf. Sci.*, vol. 235, pp. 3–35, 2013.

[11] B. Huang and R. Kadali, *Dynamic modeling, predictive control and performance monitoring: a data-driven subspace approach*. Springer, 2008.

[12] J. Coulson, J. Lygeros, and F. Dörfler, "Data-enabled predictive control: In the shallows of the DeePC," in *Proc. ECC*. IEEE, 2019, pp. 307–312.
[13] ——, "Regularized and distributionally robust data-enabled predictive control," in *Proc. IEEE CDC*. IEEE, 2019, pp. 2696–2701.
[14] ——, "Distributionally robust chance constrained data-enabled predictive control," *IEEE Trans. Autom. Control*, vol. 67, no. 7, pp. 3289–3304, 2021.
[15] E. Elokda, J. Coulson, P. N. Beuchat, J. Lygeros, and F. Dörfler, "Data-enabled predictive control for quadcopters," *Int. J. Robust & Nonlinear Control*, vol. 31, no. 18, pp. 8916–8936, 2021.
[16] P. G. Carlet, A. Favato, S. Bolognani, and F. Dörfler, "Data-driven predictive current control for synchronous motor drives," in *ECCE*. IEEE, 2020, pp. 5148–5154.
[17] P. Mahdavipour, C. Wieland, and H. Spliethoff, "Optimal control of combined-cycle power plants: A data-enabled predictive control perspective," *IFAC-PapersOnLine*, vol. 55, no. 13, pp. 91–96, 2022.
[18] L. Huang, J. Coulson, J. Lygeros, and F. Dörfler, "Data-enabled predictive control for grid-connected power converters," in *Proc. IEEE CDC*. IEEE, 2019, pp. 8130–8135.
[19] ——, "Decentralized data-enabled predictive control for power system oscillation damping," *IEEE Trans. Control Syst. Tech.*, vol. 30, no. 3, pp. 1065–1077, 2021.
[20] Y. Zhao, T. Liu, and D. J. Hill, "A data-enabled predictive control method for frequency regulation of power systems," in *ISGT Europe*. IEEE, 2021, pp. 01–06.
[21] L. Huang, J. Zhen, J. Lygeros, and F. Dörfler, "Quadratic regularization of data-enabled predictive control: Theory and application to power converter experiments," *IFAC-PapersOnLine*, vol. 54, no. 7, pp. 192–197, 2021.
[22] ——, "Robust data-enabled predictive control: Tractable formulations and performance guarantees," *IEEE Trans. Autom. Control*, vol. 68, no. 5, pp. 3163–3170, 2023.
[23] J. Berberich, J. Köhler, M. A. Müller, and F. Allgöwer, "Data-driven model predictive control with stability and robustness guarantees," *IEEE Trans. Autom. Control*, vol. 66, no. 4, pp. 1702–1717, 2020.
[24] ——, "Robust constraint satisfaction in data-driven MPC," in *Proc. IEEE CDC*. IEEE, 2020, pp. 1260–1267.
[25] ——, "Data-driven tracking MPC for changing setpoints," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 6923–6930, 2020.
[26] ——, "On the design of terminal ingredients for data-driven MPC," *IFAC-PapersOnLine*, vol. 54, no. 6, pp. 257–263, 2021.
[27] L. Huang, J. Lygeros, and F. Dörfler, "Robust and kernelized data-enabled predictive control for nonlinear systems," *IEEE Trans. Control Syst. Tech.*, vol. 32, no. 2, pp. 611–624, 2023.
[28] S. Kerz, J. Teutsch, T. Brüdigam, M. Leibold, and D. Wollherr, "Data-driven tube-based stochastic predictive control," *IEEE Open J. Control Syst.*, vol. 2, pp. 185–199, 2023.
[29] J. Teutsch, S. Kerz, D. Wollherr, and M. Leibold, "Sampling-based stochastic data-driven predictive control under data uncertainty," *IEEE Trans. Autom. Control*, 2025.
[30] Y. Wang, K. You, D. Huang, and C. Shang, "Data-driven output prediction and control of stochastic systems: An innovation-based approach," *Automatica*, vol. 171, p. 111897, 2025.
[31] M. Yin, A. Iannelli, and R. S. Smith, "Stochastic data-driven predictive control: Regularization, estimation, and constraint tightening," *IFAC-PapersOnLine*, vol. 58, no. 15, pp. 79–84, 2024.
[32] G. Pan, R. Ou, and T. Faulwasser, "On a stochastic fundamental lemma and its use for data-driven optimal control," *IEEE Trans. Autom. Control*, vol. 68, no. 10, pp. 5922–5937, 2022.
[33] ——, "Towards data-driven stochastic predictive control," *Int. J. Robust Nonlinear Control*, vol. 35, no. 7, pp. 2588–2610, 2025.
[34] ——, "On data-driven stochastic output-feedback predictive control," *IEEE Trans. Autom. Control*, vol. 70, no. 5, pp. 2948 – 2962, 2024.
[35] ——, "Data-driven stochastic output-feedback predictive control: Recursive feasibility through interpolated initial conditions," in *L4DC*. PMLR, 2023, pp. 980–992.
[36] V. Kučera, "The discrete Riccati equation of optimal control," *Kybernetika*, vol. 8, no. 5, pp. 430–447, 1972.
[37] M. Cannon, Q. Cheng, B. Kouvaritakis, and S. V. Raković, "Stochastic tube MPC with state estimation," *Automatica*, vol. 48, no. 3, pp. 536–541, 2012.
[38] M. Farina, L. Giulioni, L. Magni, and R. Scattolini, "An approach to output-feedback MPC of stochastic linear discrete-time systems," *Automatica*, vol. 55, pp. 140–149, 2015.
[39] E. Joa, M. Bujarbaruah, and F. Borrelli, "Output feedback stochastic MPC with hard input constraints," in *Proc. ACC*. IEEE, 2023, pp. 2034–2039.
[40] J. Ridderhof, K. Okamoto, and P. Tsiotras, "Chance constrained covariance control for linear stochastic systems with output feedback," in *Proc. IEEE CDC*. IEEE, 2020, pp. 1758–1763.
[41] P. Hokayem, E. Cinquemani, D. Chatterjee, F. Ramponi, and J. Lygeros, "Stochastic receding horizon control with output feedback and bounded controls," *Automatica*, vol. 48, no. 1, pp. 77–88, 2012.
[42] J. Köhler and M. N. Zeilinger, "Recursively feasible stochastic predictive control using an interpolating initial state constraint," *IEEE Control Syst. Let.*, vol. 6, pp. 2743–2748, 2022.
[43] H. Schlüter and F. Allgöwer, "Stochastic model predictive control using initial state optimization," *IFAC-PapersOnLine*, vol. 55, no. 30, pp. 454–459, 2022.
[44] M. Ono and B. C. Williams, "Iterative risk allocation: A new approach to robust model predictive control with a joint chance constraint," in *Proc. IEEE CDC*. IEEE, 2008, pp. 3427–3432.
[45] P. J. Goulart, E. C. Kerrigan, and J. M. Maciejowski, "Optimization over state feedback policies for robust control with constraints," *Automatica*, vol. 42, no. 4, pp. 523–533, 2006.
[46] R. Li, J. W. Simpson-Porco, and S. L. Smith, "Stochastic data-driven predictive control with equivalence to stochastic MPC," *arXiv preprint arXiv:2312.15177*, 2023.
[47] M. Cannon, B. Kouvaritakis, and X. Wu, "Model predictive control for systems with stochastic multiplicative uncertainty and probabilistic constraints," *Automatica*, vol. 45, no. 1, pp. 167–172, 2009.
[48] L. Hewing, K. P. Wabersich, and M. N. Zeilinger, "Recursively feasible stochastic model predictive control using indirect feedback," *Automatica*, vol. 119, p. 109095, 2020.
[49] D. Alpago, F. Dörfler, and J. Lygeros, "An extended Kalman filter for data-enabled predictive control," *IEEE Control Syst. Let.*, vol. 4, no. 4, pp. 994–999, 2020.
[50] P. Van Overschee and B. De Moor, "N4SID: Subspace algorithms for the identification of combined deterministic-stochastic systems," *Automatica*, vol. 30, no. 1, pp. 75–93, 1994.
[51] J. C. Willems, P. Rapisarda, I. Markovsky, and B. L. De Moor, "A note on persistency of excitation," *IFAC Syst & Control L*, vol. 54, no. 4, pp. 325–329, 2005.
[52] T. N. E. Greville, "Note on the generalized inverse of a matrix product," *Siam Review*, vol. 8, no. 4, pp. 518–521, 1966.

Ruiqi Li (S'22) received the B.Sc. degree in Honours Physics from the University of Waterloo, ON, Canada in 2019 and the B.Sc. degree in physics from Beijing Institute of Technology, Beijing, China in 2019. He is currently working towards the Ph.D. degree in Electrical and Computer Engineering at the University of Waterloo, ON, Canada. His research interest includes data-driven control, model predictive control and optimization.

John W. Simpson-Porco (S'10–M'15–SM'22) received the B.Sc. degree in engineering physics from Queen's University, Kingston, ON, Canada in 2010, and the Ph.D. degree in mechanical engineering from the University of California at Santa Barbara, Santa Barbara, CA, USA in 2015. He is currently an Associate Professor of Electrical and Computer Engineering at the University of Toronto, Toronto, ON, Canada.

Stephen L. Smith (S'05–M'09–SM'15) received the B.Sc. degree in engineering physics from Queen's University, Canada, in 2003, the M.A.Sc. degree in electrical and computer engineering from the University of Toronto, Canada, in 2005, and the Ph.D. degree in mechanical engineering from the University of California, Santa Barbara, USA, in 2009. He is currently a Professor in the Department of Electrical and Computer Engineering at the University of Waterloo, Canada, where he holds a Canada Research Chair in Autonomous Systems.