Distributionally Robust Stochastic Data-Driven Predictive Control with Optimized Feedback Gain

Ruiqi Li, John W. Simpson-Porco, and Stephen L. Smith

***Abstract*— We consider the problem of direct data-driven predictive control for unknown stochastic linear time-invariant (LTI) systems with partial state observation. Building upon our previous research on data-driven stochastic control, this paper (i) relaxes the assumption of Gaussian process and measurement noise, and (ii) enables optimization of the gain matrix within the affine feedback policy. Output safety constraints are modelled using conditional value-at-risk, and enforced in a distributionally robust sense. Under idealized assumptions, we prove that our proposed data-driven control method yields control inputs identical to those produced by an equivalent model-based stochastic predictive controller. A simulation study illustrates the enhanced performance of our approach over previous designs.**

I. Introduction

Model predictive control (MPC) is a widely used technique for multivariate control [1], adept at handling constraints on inputs, states and outputs while optimizing complex performance objectives. Constraints typically model actuator limits, or encode safety constraints in safety-critical applications, and MPC employs a system model to predict how inputs influence state evolution. Both deterministic and stochastic frameworks have been developed to account for plant uncertainty in MPC. While *Robust MPC* [2] approaches model uncertainty in a worst-case deterministic sense, work on *Stochastic MPC (SMPC)* [3] has focused on describing model uncertainty probabilistically. SMPC methods optimize over feedback control policies rather than control actions, resulting in performance benefits when compared to the naïve use of deterministic MPC [4], and accommodate probabilistic and risk-aware constraints.

The system model required by MPC (and SMPC) is sometimes obtained from identification, making MPC an *indirect* design method, since one goes from data to a controller through an intermediate modelling step [5]. In contrast, data-driven or *direct* methods seek to compute controllers directly from input-output data, showing promise for complex or difficult-to-model systems [6]. Accounting for constraints in control, *Data-Driven Predictive Control (DDPC)* methods were developed, including Data-Enabled Predictive Control (DeePC) [7]–[9] and Subspace Predictive Control (SPC) [10], both of which have been applied in multiple experiments [11]–[13]. For *deterministic* LTI systems in theory, both DeePC and SPC produce equivalent control actions as from MPC.

Real-world systems often deviate from idealized deterministic LTI models, exhibiting stochastic and non-linear behavior, with noise-corrupted data. To address these challenges, data-driven methods must account for noisy data and measurements. For instance, in SPC applications, required predictor matrices are often computed using denoising techniques [10]. Regularized and distributionally robust DeePC were also developed for stochastic systems [7]–[9]. Unlike in the deterministic case, however, these stochastic adaptations of DeePC and SPC lack theoretical equivalence to model-based MPC.

Recognizing this gap, some recent advancements in DDPC have aimed to establish equivalence with MPC methods for stochastic systems. The work in [14], [15] proposed a DDPC framework for stochastic systems, and their method performs equivalently to SMPC if stochastic signals can be exactly expressed by their Polynomial Chaos Expansion. This paper builds in particular on our previous work [16], where we proposed a data-driven control method for stochastic systems, without estimation of disturbance as required in [14], [15], and established that the method has equivalent control performance to SMPC when offline data is noise-free.

Contribution: This paper contributes towards the continued development of high-performance DDPC methods for stochastic systems. Specifically, in this paper we develop a stochastic DDPC strategy utilizing distributionally robust conditional value-at-risk constraints, providing an improved safety constraint description when compared to our prior work in [16], and providing robustness against non-Gaussian (i.e., possibly heavy-tailed) process and measurement noise. Additionally, in contrast with the fixed feedback gain in [16], we consider control policies where feedback gains are decision variables in the optimization, giving a more flexible parameterization of control policies. As theoretical support for the approach, under technical conditions, we establish equivalence between our proposed design and a corresponding SMPC. Finally, a simulation case study compares and contrasts our design with other recent stochastic and data-driven control strategies.

Notation: Let $M^\dagger$ be the pseudo-inverse of a matrix M. Let $\otimes$ denote the Kronecker product. Let $\mathbb{S}_{+}^{q}$ (resp. $\mathbb{S}_{++}^{q}$) be the set of $q \times q$ positive semi-definite (resp. definite) matrices. Let $\mathrm{col}(M_1, \ldots, M_k)$ (resp. $\mathrm{Diag}(M_1, \ldots, M_k)$) denote the vertical (resp. diagonal) concatenation of matrices / vectors $M_1, \ldots, M_k$. Let $\mathbb{Z}_{[a,b]} := [a,b] \cap \mathbb{Z}$ denote a set of consecutive integers from a to b, and 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

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

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 Setup

Consider a stochastic linear time-invariant (LTI) system

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

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 $v_t \in \mathbb{R}^p$. The system (A, B, C, D) is assumed as a minimal realization, but the matrices themselves 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 probability distributions of w_t and v_t are *unknown*, but we assume that w_t and v_t have zero mean and zero auto-correlation (white noise), are uncorrelated, and their variances $\Sigma^{\mathsf{w}} \in \mathbb{S}^n_+$ and $\Sigma^{\mathsf{v}} \in \mathbb{S}^p_{++}$ are known. The initial state x_0 has given mean $\mu^{\mathsf{x}}_{\mathsf{ini}}$ and variance Σ^{x} and is uncorrelated with the noise. We record these conditions as

$$\mathbb{E}\Big[\begin{bmatrix} w_t \\ v_t \end{bmatrix}\Big] = 0, \quad \mathbb{E}\Big[\begin{bmatrix} w_t \\ v_t \end{bmatrix}\begin{bmatrix} w_s \\ v_s \end{bmatrix}^{\mathsf{T}}\Big] = \begin{bmatrix} \delta_{ts}\Sigma^{\mathsf{w}} & 0 \\ 0 & \delta_{ts}\Sigma^{\mathsf{v}} \end{bmatrix}, \tag{2}$$

$$\mathbb{E}[x_0] = \mu^{\mathsf{x}}_{\mathsf{ini}}, \quad \mathrm{Var}[x_0] = \Sigma^{\mathsf{x}}, \quad \mathbb{E}\Big[x_0 \begin{bmatrix} w_t \\ v_t \end{bmatrix}^{\mathsf{T}}\Big] = 0, \tag{3}$$

with δ_{ts} the Kronecker delta. Assume (A, Σ^{w}) is stabilizable.

In a reference tracking control problem for (1), 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 may be encoded in an instantaneous cost

$$J_t(u_t, y_t) := \|y_t - r_t\|^2_Q + \|u_t\|^2_R \tag{4}$$

to be minimized over a time horizon, with user-selected parameters $Q \in \mathbb{S}^p_+$ and $R \in \mathbb{S}^m_{++}$. This tracking should be achieved subject to constraints on the inputs and outputs. We consider here polytopic constraints, which in a deterministic setting would take the form $E\begin{bmatrix} u_t \\ y_t \end{bmatrix} \leq f$ for all $t \in \mathbb{N}_{\geq 0}$, and for some fixed matrix $E \in \mathbb{R}^{q \times (m+p)}$ and vector $f \in \mathbb{R}^q$. We can equivalently express these constraints as the single constraint $h(u_t, y_t) \leq 0$, where

$$h(u_t, y_t) := \max\nolimits_{i \in \{1,\dots,q\}} e_i^{\mathsf{T}} \begin{bmatrix} u_t \\ y_t \end{bmatrix} - f_i, \tag{5}$$

with $e_i \in \mathbb{R}^{m+p}$ the transposed i-th row of E and $f_i \in \mathbb{R}$ the i-th entry of f. For the system (1) which is subject to (possibly unbounded) stochastic disturbances, the deterministic constraint $h(u_t, y_t) \leq 0$ must be relaxed. Beyond a traditional chance constraint $\mathbb{P}[h(u_t, y_t) \leq 0] \geq 1 - \alpha$ with a violation probability $\alpha \in (0, 1)$, a *conditional value-at-risk (CVaR)* constraint is more conservative; the CVaR at level α of $h(u_t, y_t)$ is defined as the expected value of $h(u_t, y_t)$ in the $\alpha\cdot$ 100% worst cases, and takes extreme violations into account. With the noise distributions unknown, we must further guarantee satisfaction of the CVaR constraint for all possible distributions under consideration. Let $\mathbb{D}$ denote a joint distribution of all random variables in (1) satisfying (2) and (3), and let the *ambiguity set* $\mathcal{D}$ be the set of all such distributions. The *distributionally robust CVaR (DR-CVaR)* constraint [17], [18] is then

$$\sup\nolimits_{\mathbb{D} \in \mathcal{D}} \mathbb{D}\text{-}\mathrm{CVaR}_\alpha[h(u_t, y_t)] \leq 0, \tag{6}$$

where $\mathbb{D}\text{-}\mathrm{CVaR}_\alpha[z]$ is the CVaR value of a random variable $z \in \mathbb{R}$ at level α given distribution $\mathbb{D}$.

If the system matrices A, B, C, D were known, this constrained tracking control problem subject to (6) can be approached using SMPC, as described in Section III-A. Our objective is to develop a data-driven control method that produces equivalent control inputs as produced by SMPC.

III. Stochastic Model-Based and Data-Driven Predictive Control

We introduce a model-based SMPC framework in Section III-A and propose a data-driven control method in Section III-B, with their theoretical equivalence in Section III-C.

A. A framework of Stochastic Model Predictive Control

We focus here on output-feedback SMPC [19]–[21] which typically combines state estimation and feedback control. The formulation here broadly follows our prior work [16], but we now consider a DR-CVaR constraint in place of chance constraints, and we will allow optimization over the feedback gain. This SMPC scheme merges the established works on DR constrained control [17], [18] and output-error feedback [22], while the combined framework is part of our contribution.

1) State Estimation: SMPC follows a receding-horizon strategy and makes decisions for N upcoming steps at each *control step*. At control step $t = k$, we begin with prior information of the mean and variance of state x_k, namely

$$\mathbb{E}[x_k] = \mu^{\mathsf{x}}_k, \qquad \mathrm{Var}[x_k] = \Sigma^{\mathsf{x}}, \tag{7}$$

where the mean μ^{x}_k is computed from a state estimator to be described next; at the initial step $k = 0$, $\mu^{\mathsf{x}}_0 = \mu^{\mathsf{x}}_{\mathsf{ini}}$ is a given parameter as in (3). For simplicity of computation, we let Σ^{x} in (3) and (7) be the steady-state variance through the Kalman filter, as the unique positive semi-definite solution to the associated discrete-time algebraic Riccati equation (DARE) (8a), with observer gain $L_{\mathsf{L}} \in \mathbb{R}^{n \times p}$ in (8b).

$$\Sigma^{\mathsf{x}} = (A - L_{\mathsf{L}}C)\Sigma^{\mathsf{x}}A^{\mathsf{T}} + \Sigma^{\mathsf{w}} \tag{8a}$$

$$L_{\mathsf{L}} := A\Sigma^{\mathsf{x}}C^{\mathsf{T}}(C\Sigma^{\mathsf{x}}C^{\mathsf{T}} + \Sigma^{\mathsf{v}})^{-1} \tag{8b}$$

Estimates $\hat{x}_t$ of future states over the desired horizon are computed through the observer, with *innovation* $\nu_t \in \mathbb{R}^p$,

$$\nu_t := y_t - C\hat{x}_t - Du_t, \qquad t \in \mathbb{Z}_{[k,k+N)} \tag{9a}$$

$$\hat{x}_{t+1} := A\hat{x}_t + Bu_t + L_{\mathsf{L}}\nu_t, \qquad t \in \mathbb{Z}_{[k,k+N)} \tag{9b}$$

$$\hat{x}_k := \mu^{\mathsf{x}}_k \tag{9c}$$

where we utilize in (9b) the observer gain L_{L} in (8b) so that (9) is equivalent to the steady-state Kalman filter. While the noise here is potentially non-Gaussian, the Kalman filter is the best affine state estimator in the mean-squared-error sense, regardless of the distributions of x_k, w_t, v_t once their means and variances are specified as in (2) and (7) [23, Sec. 3.1].

At the control step with condition (7), we can predict future states and outputs by simulating the noise-free model,

$$\bar{x}_{t+1} := A\bar{x}_t + B\bar{u}_t, \qquad t \in \mathbb{Z}_{[k,k+N)} \tag{10a}$$

$$\bar{y}_t := C\bar{x}_t + D\bar{u}_t, \qquad t \in \mathbb{Z}_{[k,k+N)} \tag{10b}$$

$$\bar{x}_k := \mu_k^{\mathsf{x}} \tag{10c}$$

with *nominal inputs* $\bar{u}_t$ as decision variables to be optimized, and with resulting *nominal states* $\bar{x}_t$ and *nominal outputs* $\bar{y}_t$.

2) Feedback Control Policies: Our prior work [16] was based on an affine feedback policy $u_t = \bar{u}_t - K(\hat{x}_t - \bar{x}_t)$ with a fixed feedback gain K. Here we investigate control policies where the feedback gain is a time-varying decision variable. However, the naive parameterization

$$u_t \leftarrow \bar{u}_t - K_t(\hat{x}_t - \bar{x}_t) \tag{11}$$

leads to non-convex bilinear terms of the decision variables $\bar{u}$ and K_t, as $\hat{x}_t, \bar{x}_t$ depend on $\bar{u}_{[k,t)}$ via (9), (10). Thus, we instead apply an *output error feedback* control policy [22]

$$u_t \leftarrow \pi_t(\nu_{[k,t)}) := \bar{u}_t + \textstyle\sum_{s=k}^{t-1} M_t^s \, \nu_s \tag{12}$$

where the nominal input $\bar{u}_t$ and feedback gains $M_t^s \in \mathbb{R}^{m\times p}$ are both decision variables, with innovation ν in (9a). The policy parameterization (12) contains within it the policy (11) as a special case: indeed, for a sequence of gains $K_{[k,k+N)}$, the selection for all $s, t \in \mathbb{Z}_{[k,k+N)}, s \le t$

$$M_t^s \leftarrow (A - BK_{t-1})(A - BK_{t-2}) \cdots (A - BK_s) L_{\mathsf{L}}$$

reduces (12) to (11). Crucially, (12) leads to jointly convex optimization in decision variables $\bar{u}, M_t^s$, as we will see next.

With the estimator (9) and policy (12), both input u_t and output y_t of (1) can be written as affine functions of the decision variables, through direct calculation, with $\bar{y}$ in (10),

$$\begin{bmatrix} u_t \\ y_t \end{bmatrix} = \begin{bmatrix} \bar{u}_t \\ \bar{y}_t \end{bmatrix} + \Lambda_t \, \eta_k, \quad t \in \mathbb{Z}_{[k,k+N)}, \tag{13}$$

where $\eta_k := \mathrm{col}(x_k - \mu_k^{\mathsf{x}}, w_{[k,k+N)}, v_{[k,k+N)}) \in \mathbb{R}^{n_\eta}$ is a vector of uncorrelated zero-mean random variables of dimension $n_\eta := n + nN + pN$, and matrix $\Lambda_t \in \mathbb{R}^{(m+p)\times n_\eta}$ is linearly dependent on the gain matrices M_t^s as

$$\Lambda_t := \begin{bmatrix} \Delta_{t-k}^{\mathsf{U}} \\ \Delta_{t-k}^{\mathsf{Y}} \end{bmatrix} \mathcal{M} \, \Delta^{\mathsf{M}} + \begin{bmatrix} 0_{m\times n_\eta} \\ \Delta_{t-k}^{\mathsf{A}} \end{bmatrix}, \quad t \in \mathbb{Z}_{[k,k+N)}, \tag{14}$$

where $\mathcal{M} \in \mathbb{R}^{mN\times pN}$ is a concatenation of M_t^s

$$\mathcal{M} := \begin{bmatrix} M_k^k & & & \\ M_{k+1}^k & M_{k+1}^{k+1} & & \\ \vdots & \vdots & \ddots & \\ M_{k+N-1}^k & M_{k+N-1}^{k+1} & \cdots & M_{k+N-1}^{k+N-1} \end{bmatrix} \tag{15}$$

and where $\Delta_i^{\mathsf{U}} \in \mathbb{R}^{m\times mN}, \Delta_i^{\mathsf{Y}} \in \mathbb{R}^{p\times mN}, \Delta_i^{\mathsf{A}} \in \mathbb{R}^{p\times n_\eta}$ and $\Delta^{\mathsf{M}} \in \mathbb{R}^{pN\times n_\eta}$ are independent of both decision variables $\bar{u}$ and M_t^s, with expressions available in Appendix A.

3) Deterministic Formulation of Cost and Constraint: Given (13), $\mathrm{col}(u_t, y_t)$ has mean $\mathrm{col}(\bar{u}_t, \bar{y}_t)$ and variance $\Lambda_t \Sigma^\eta \Lambda_t^{\mathsf{T}}$, since η_k has zero mean and the variance $\Sigma^\eta := \mathrm{Diag}(\Sigma^{\mathsf{x}}, I_N \otimes \Sigma^{\mathsf{w}}, I_N \otimes \Sigma^{\mathsf{v}}) \in \mathbb{S}_+^{n_\eta}$ via (2) and (7). Then, the constraint (6) can be equivalently written as a second-order cone (SOC) constraint of the decision variables $\bar{u}$ and M_t^s.

Lemma 1 (SOC Expression of DR-CVaR Constraint)**.** *With $h(u_t, y_t)$ as in* (5)*, for $t \in \mathbb{Z}_{[k,k+N)}$,* (6) *holds iff*

$$2\big(\tfrac{1-\alpha}{\alpha}\big)^{\frac{1}{2}} \big\|(\Sigma^\eta)^{\frac{1}{2}} \Lambda_t^{\mathsf{T}} e_i\big\|_2 \le -e_i^{\mathsf{T}} \begin{bmatrix} \bar{u}_t \\ \bar{y}_t \end{bmatrix} + f_i, \ \ i \in \mathbb{Z}_{[1,q]}. \tag{16}$$

Proof. Substituting (13) into (5), $h(u_t, y_t)$ can be written as

$$h(u_t, y_t) = \max_{i\in\{1,\dots,q\}} e_i^{\mathsf{T}} \Lambda_t \eta_k + e_i^{\mathsf{T}} \mathrm{col}(\bar{u}_t, \bar{y}_t) - f_i,$$

where the random variable η_k has zero mean and variance Σ^η. According to [18, Thm. 3.3], (6) holds if and only if there exist $\theta_t \in \mathbb{R}$ and $\Theta_t \in \mathbb{S}_+^{n_\eta+1}$ satisfying the LMIs

$$0 \ge \alpha\theta_t + \mathrm{Trace}\big[\Theta_t \, \mathrm{Diag}(\Sigma^\eta, 1)\big]$$

$$\Theta_t \succeq \begin{bmatrix} 0_{n_\eta\times n_\eta} & \Lambda_t^{\mathsf{T}} e_i \\ e_i^{\mathsf{T}} \Lambda_t & e_i^{\mathsf{T}} \mathrm{col}(\bar{u}_t, \bar{y}_t) - f_i - \theta_t \end{bmatrix}, \ i \in \mathbb{Z}_{[1,q]}.$$

From [24, Thm. 1], these LMIs are feasible in (θ_t, Θ_t) if and only if (16) holds, which completes the proof. ∎

SMPC problems typically consider the expected cost $\sum_{t=k}^{k+N-1} \mathbb{E}[J_t(u_t, y_t)]$ summing (4) over the horizon, which is equal to a deterministic quadratic function of $\bar{u}$ and M_t^s,

$$\textstyle\sum_{t=k}^{k+N-1} \big[J_t(\bar{u}_t, \bar{y}_t) + \|\mathrm{Diag}(R, Q)^{\frac{1}{2}} \Lambda_t (\Sigma^\eta)^{\frac{1}{2}}\|_{\mathsf{F}}^2\big], \tag{17}$$

given the mean and variance of $\mathrm{col}(u_t, y_t)$ and given that $\mathbb{E}[\|z\|_S^2] = \|\mathbb{E}[z]\|_S^2 + \|S^{\frac{1}{2}} \mathrm{Var}[z]^{\frac{1}{2}}\|_{\mathsf{F}}^2$ for any random vector z and fixed matrix S; $\|\cdot\|_{\mathsf{F}}$ denotes the Frobenius norm.

4) SMPC Optimization Problem and Algorithm: Using the cost (17) and reformulation (16) of constraint (6), we have the SMPC problem as a second-order cone problem (SOCP)

$$\underset{\bar{u}, M_t^s}{\mathrm{minimize}} \ (17) \text{ s.t. } (16) \text{ for } t \in \mathbb{Z}_{[k,k+N)}, (10), (14), \tag{18}$$

which problem has a unique optimal solution when feasible, since (17) is jointly strongly convex[1] in $\bar{u}$ and M_t^s.

The nominal inputs $\bar{u}$ and gains M_t^s determined from (18) complete the parameterization of control policies $\pi_{[k,k+N)}$ in (12), and 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 parameter $N_{\mathrm{c}} \in \mathbb{Z}_{[1,N]}$. The next control step will be set as $t = k + N_{\mathrm{c}}$, and the state mean $\mu_{k+N_{\mathrm{c}}}^{\mathsf{x}}$ in (7) will be iterated as the estimate $\hat{x}_{k+N_{\mathrm{c}}}$ via (9); we let the nominal state $\bar{x}_{k+N_{\mathrm{c}}}$ via (10) be a backup value $\mu_{k+N_{\mathrm{c}}}^{\bar{\mathsf{x}}}$ of $\mu_{k+N_{\mathrm{c}}}^{\mathsf{x}}$ that ensures feasibility of (18) at the new control step [19]. The entire SMPC control process is shown in Algorithm 1.

B. Stochastic Data-Driven Predictive Control (SDDPC)

We develop in this section a data-driven control method, which consists of an offline process for data collection and an online process that makes real-time control decisions.

1) Use of Offline Data: In data-driven control, sufficient offline data is required to capture the system's behavior. Here we explain how we collect data and use it to calculate some quantities required in our control method. We first consider noise-free data and then address the case of noisy data.

Consider a deterministic version of the system (1)

$$x_{t+1} = Ax_t + Bu_t, \qquad y_t = Cx_t + Du_t. \tag{19}$$

[1] Strong convexity in $\bar{u}$ is clear from the first term; strong convexity in M_t^s can be shown by noting that a sub-matrix of $\mathrm{col}(\mathcal{J}_k, \dots, \mathcal{J}_{k+N-1})$ with $\mathcal{J}_t := \mathrm{Diag}(R, Q)^{1/2} \Lambda_t (\Sigma^\eta)^{1/2}$ is $\bar{\mathcal{J}}_{\mathsf{L}} \mathcal{M} \bar{\mathcal{J}}_{\mathsf{R}}$, where $\bar{\mathcal{J}}_{\mathsf{L}} := I_N \otimes R^{1/2}$ and $\bar{\mathcal{J}}_{\mathsf{R}} := (I_{pN} - \Xi(A_{\mathsf{L}})(I_N \otimes L_{\mathsf{L}}))(I_N \otimes (\Sigma^{\mathsf{v}})^{1/2})$ are non-singular.

Algorithm 1 Distributionally Robust Optimized-Gain Stochastic MPC (DR/O-SMPC)

Input: horizon lengths N, N_c, system matrices A, B, C, noise variances $\Sigma^\mathsf{w}, \Sigma^\mathsf{v}$, initial-state mean $\mu^\mathsf{x}_\mathsf{ini}$, cost matrices Q, R, constraint coefficients E, f, and CVaR level α.

1: Compute $\Sigma^\mathsf{x}, L_\mathsf{L}$ via (8) and $\Delta^\mathsf{U}_{[0,N)}, \Delta^\mathsf{Y}_{[0,N)}, \Delta^\mathsf{A}_{[0,N)}, \Delta^\mathsf{M}$ through Appendix A.
2: Initialize the control step $k \leftarrow 0$ and set $\mu^\mathsf{x}_0 \leftarrow \mu^\mathsf{x}_\mathsf{ini}$.
3: Solve $\bar{u}_{[k,k+N)}$ and M^s_t from problem (18).
4: **If** (18) is infeasible **then** Set $\mu^\mathsf{x}_k \leftarrow \mu^{\bar{\mathsf{x}}}_k$, and redo line 3.
5: **for** t **from** k **to** $k + N_\mathrm{c} - 1$ **do**
6: Input $u_t \leftarrow \pi_t(\nu_{[k,t)})$ in (12) to the system (1).
7: Measure y_t from the system (1).
8: Compute ν_t via (9).
9: Set $(\mu^\mathsf{x}_{k+N_\mathrm{c}}, \mu^{\bar{\mathsf{x}}}_{k+N_\mathrm{c}})$ as $(\hat{x}_{k+N_\mathrm{c}}, \bar{x}_{k+N_\mathrm{c}})$ in (9), (10).
10: Set $k \leftarrow k + N_\mathrm{c}$. Go back to line 3.

By assumption, (19) is minimal; 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. 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 (19). The input sequence u^d is assumed to be *persistently exciting* of order $K_\mathrm{d} := L + 1 + n$, i.e., its associated K_d-depth block-Hankel matrix $\mathcal{H}_{K_\mathrm{d}}(u^\mathrm{d}_{[1,T_\mathrm{d}]}) \in \mathbb{R}^{mK_\mathrm{d} \times (T_\mathrm{d} - K_\mathrm{d} + 1)}$, 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. 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 width $h := T_\mathrm{d} - L$ by partitioning associated Hankel matrices as

$$\begin{bmatrix} U_1 \\ U_2 \end{bmatrix} := \mathcal{H}_{L+1}(u^\mathrm{d}_{[1,T_\mathrm{d}]}), \quad \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} := \mathcal{H}_{L+1}(y^\mathrm{d}_{[1,T_\mathrm{d}]}). \tag{20}$$

The data matrices in (20) will now be used to represent a quantity $\mathbf{\Gamma} \in \mathbb{R}^{p \times (mL + pL)}$ related to the system (19),

$$\mathbf{\Gamma} = \begin{bmatrix} \mathbf{\Gamma}_\mathsf{U} & \mathbf{\Gamma}_\mathsf{Y} \end{bmatrix} := \begin{bmatrix} C\mathcal{C} & CA^L \end{bmatrix} \begin{bmatrix} I_{mL} & \\ \mathcal{G} & \mathcal{O} \end{bmatrix}^\dagger, \tag{21}$$

with $\mathcal{C} := [A^{L-1}B, \dots, AB, B]$ the extended controllability matrix and $\mathcal{G} := \mathrm{Toep}(D, CB, \dots, CA^{L-2}B)$ the impulse-response matrix; Toep denotes the block-Toeplitz matrix

$$\mathrm{Toep}(M_1, \dots, M_k) := \begin{bmatrix} M_1 & & & \\ M_2 & M_1 & & \\ \vdots & \ddots & \ddots & \\ M_k & \cdots & M_2 & M_1 \end{bmatrix}.$$

Lemma 2 (Data Representation of $\mathbf{\Gamma}$ and D [16])**.** *If system* (19) *is controllable and the input data $u^\mathrm{d}_{[1,T_\mathrm{d}]}$ is persistently exciting of order $L + 1 + n$, then, given the data matrices in* (20)*, the matrix $\mathbf{\Gamma}$ defined in* (21) *and matrix D in system* (19) *can be expressed as* $[\mathbf{\Gamma}_\mathsf{U}, \mathbf{\Gamma}_\mathsf{Y}, D] = Y_2\, \mathrm{col}(U_1, Y_1, U_2)^\dagger$.

With Lemma 2, the matrices $\mathbf{\Gamma}, D$ are represented using offline data collected from system (19), and will be used as part of the construction for our data-driven control method.

In the case where the measured data is corrupted by noise, as will usually be the case, the pseudo-inverse computation in Lemma 2 is numerically fragile and does not recover the desired matrices $\mathbf{\Gamma}, D$. A standard technique to robustify this computation is to replace the pseudo-inverse $W^\dagger$ of $W := \mathrm{col}(U_1, Y_1, U_2)$ in Lemma 2 with its Tikhonov regularization $(W^\mathsf{T} W + \lambda I_h)^{-1} W^\mathsf{T}$ with a regularization parameter $\lambda > 0$.

2) Auxiliary State-Space Model: The SMPC approach of Section III-A uses as sub-components a state estimator, an affine feedback law and a DR-CVaR constraint. We now leverage the offline data as described in Section III-B-1 to directly design analogs of these components based on data, without knowledge of the system matrices.

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. 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

$$\mathbf{x}_t := \begin{bmatrix} u_{[t-L,t)} \\ \hline y^\circ_{[t-L,t)} \\ \hline \rho_{[t-L,t)} \end{bmatrix}, \quad \mathbf{w}_t := \begin{bmatrix} 0_{mL \times 1} \\ \hline 0_{pL \times 1} \\ \hline 0_{pL(L-1) \times 1} \\ \rho_t \end{bmatrix} \tag{22}$$

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 auxiliary signals $\mathbf{x}_t, \mathbf{w}_t$ together with u_t, y_t, v_t then satisfy the relations given by Lemma 3.

Lemma 3 (Auxiliary Model [16])**.** *For system* (1)*, signals u_t, y_t, v_t and the auxiliary signals $\mathbf{x}_t, \mathbf{w}_t$ in* (22) *satisfy*

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{B}u_t + \mathbf{w}_t, \quad y_t = \mathbf{C}\mathbf{x}_t + Du_t + v_t \tag{23}$$

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

$$\begin{aligned} \mathbf{A} &:= \mathrm{col}(0_{mL \times n_\mathsf{aux}}, 0_{p(L-1) \times n_\mathsf{aux}}, \mathbf{C}, 0_{pL^2 \times n_\mathsf{aux}}) \\ &\qquad + \mathrm{Diag}(\mathcal{D}_m, \mathcal{D}_p, \mathcal{D}_{pL}), \text{ with } \mathcal{D}_q := \begin{bmatrix} & I_{q(L-1)} \\ 0_{q \times q} & \end{bmatrix}, \\ \mathbf{B} &:= \mathrm{col}\big(0_{m(L-1) \times m}, I_m, 0_{p(L-1) \times m}, D, 0_{pL^2 \times m}\big), \\ \mathbf{C} &:= [\mathbf{\Gamma}_\mathsf{U}, \mathbf{\Gamma}_\mathsf{Y}, \mathbf{F} - \mathbf{\Gamma}_\mathsf{Y}\mathbf{E}], \end{aligned}$$

with matrices $\mathbf{\Gamma}_\mathsf{U}, \mathbf{\Gamma}_\mathsf{Y}$ in (21)*, and zero-one matrices* $\mathbf{E} := \mathrm{Toep}(0_{p \times pL}, S_1, \dots, S_{L-1})$ *and* $\mathbf{F} := [S_L, S_{L-1}, \dots, S_1]$ *composed by* $S_j := [0_{p \times (j-1)p}, I_p, 0_{p \times (L-j)p}]$ *for* $j \in \mathbb{Z}_{[1,L]}$.

The output noise v_t in (23) is precisely the same as in (1); $\mathbf{w}_t$ appears now as a new disturbance of zero mean and the variance $\mathbf{\Sigma}^\mathsf{w} := \mathrm{Diag}(0_{(n_\mathsf{aux} - pL) \times (n_\mathsf{aux} - pL)}, \Sigma^\rho)$, where $\Sigma^\rho := \mathcal{O}\Sigma^\mathsf{w}\mathcal{O}^\mathsf{T} \in \mathbb{S}^{pL}_+$ is the variance of ρ_t. The matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}, D$ in (23) are known given offline data described in Section III-B-1, since they only depend on $\mathbf{\Gamma}_\mathsf{U}, \mathbf{\Gamma}_\mathsf{Y}, D$ which are data-representable via Lemma 2. Hence, the auxiliary model (23) can be interpreted as a data-representable realization of the system (1).

3) Data-Driven State Estimation, Feedback and Constraint: The auxiliary model (23) will now be used for both state estimation and constrained feedback control purposes. Suppose we are at a control step $t = k$ in a receding-horizon process. Similar to (7), auxiliary state $\mathbf{x}_k$ has condition $\mathbb{E}[\mathbf{x}_k] = \boldsymbol{\mu}^\mathsf{x}_k$ and $\mathrm{Var}[\mathbf{x}_k] = \mathbf{\Sigma}^\mathsf{x}$, where $\boldsymbol{\mu}^\mathsf{x}_k$ is known from the state estimator to be introduced next; at the initial time $k = 0$, the initial mean $\boldsymbol{\mu}^\mathsf{x}_\mathsf{ini}$ is a parameter; the variance $\mathbf{\Sigma}^\mathsf{x}$ is the unique positive semi-definite solution to DARE (24a),

$$\mathbf{\Sigma}^{\mathsf{x}} = (\mathbf{A} - \mathbf{L}_{\mathsf{L}}\mathbf{C})\mathbf{\Sigma}^{\mathsf{x}}\mathbf{A}^{\mathsf{T}} + \mathbf{\Sigma}^{\mathsf{w}} \tag{24a}$$

$$\mathbf{L}_{\mathsf{L}} := \mathbf{A}\mathbf{\Sigma}^{\mathsf{x}}\mathbf{C}^{\mathsf{T}}(\mathbf{C}\mathbf{\Sigma}^{\mathsf{x}}\mathbf{C}^{\mathsf{T}} + \Sigma^{\mathsf{v}})^{-1} \tag{24b}$$

given $(\mathbf{A}, \mathbf{C})$ detectable and $(\mathbf{A}, \mathbf{\Sigma}^{\mathsf{w}})$ stabilizable [16, Lemma 5]. The state estimator for the auxiliary model (23) is analogous to (9), with observer gain $\mathbf{L}_{\mathsf{L}} \in \mathbb{R}^{n_{\mathsf{aux}} \times p}$ in (24b),

$$\boldsymbol{\nu}_t := y_t - \mathbf{C}\hat{\mathbf{x}}_t - Du_t, \qquad t \in \mathbb{Z}_{[k,k+N)} \tag{25a}$$

$$\hat{\mathbf{x}}_{t+1} := \mathbf{A}\hat{\mathbf{x}}_t + \mathbf{B}u_t + \mathbf{L}_{\mathsf{L}}\boldsymbol{\nu}_t, \qquad t \in \mathbb{Z}_{[k,k+N)} \tag{25b}$$

$$\hat{\mathbf{x}}_k := \boldsymbol{\mu}_k^{\mathsf{x}} \tag{25c}$$

where $\hat{\mathbf{x}}_t$ is the estimate and $\boldsymbol{\nu}_t$ is the innovation. The output-error-feedback policy (12) in SMPC is now extended as $\boldsymbol{\pi}_t(\cdot)$,

$$u_t \leftarrow \boldsymbol{\pi}_t(\boldsymbol{\nu}_{[k,t)}) := \bar{u}_t + \textstyle\sum_{s=k}^{t-1} M_t^s\, \boldsymbol{\nu}_s \tag{26}$$

where the nominal input $\bar{u}_t \in \mathbb{R}^m$ and gain matrices $M_t^s \in \mathbb{R}^{m\times p}$ are decision variables. Let $\bar{\mathbf{x}}_t \in \mathbb{R}^{n_{\mathsf{aux}}}$ and $\bar{\mathbf{y}}_t \in \mathbb{R}^p$ be the resulting nominal state and nominal output as

$$\bar{\mathbf{x}}_{t+1} := \mathbf{A}\bar{\mathbf{x}}_t + \mathbf{B}\bar{u}_t, \qquad t \in \mathbb{Z}_{[k,k+N)}, \tag{27a}$$

$$\bar{\mathbf{y}}_t := \mathbf{C}\bar{\mathbf{x}}_t + D\bar{u}_t, \qquad t \in \mathbb{Z}_{[k,k+N)}, \tag{27b}$$

$$\bar{\mathbf{x}}_k := \boldsymbol{\mu}_k^{\mathsf{x}}. \tag{27c}$$

The SOC formulation of constraint (6) is similar to (16),

$$2\big(\tfrac{1-\alpha}{\alpha}\big)^{\frac{1}{2}} \big\|(\mathbf{\Sigma}^{\eta})^{\frac{1}{2}}\mathbf{\Lambda}_t^{\mathsf{T}} e_i\big\|_2 \le -e_i^{\mathsf{T}}\begin{bmatrix}\bar{u}_t\\ \bar{\mathbf{y}}_t\end{bmatrix} + f_i, \; i \in \mathbb{Z}_{[1,q]} \tag{28}$$

with matrices $\mathbf{\Sigma}^{\eta} := \mathrm{Diag}(\mathbf{\Sigma}^{\mathsf{x}}, I_N \otimes \mathbf{\Sigma}^{\mathsf{w}}, I_N \otimes \Sigma^{\mathsf{v}}) \in \mathbb{S}_+^{n_{\eta\text{-aux}}}$ and $\mathbf{\Lambda}_t \in \mathbb{R}^{(m+p)\times n_{\eta\text{-aux}}}$ with $n_{\eta\text{-aux}} := n_{\mathsf{aux}} + n_{\mathsf{aux}}N + pN$,

$$\mathbf{\Lambda}_t := \begin{bmatrix}\mathbf{\Delta}_{t-k}^{\mathsf{U}}\\ \mathbf{\Delta}_{t-k}^{\mathsf{Y}}\end{bmatrix}\mathcal{M}\mathbf{\Delta}^{\mathsf{M}} + \begin{bmatrix}0_{m\times n_{\eta\text{-aux}}}\\ \mathbf{\Delta}_{t-k}^{\mathsf{A}}\end{bmatrix} \tag{29}$$

where $\mathbf{\Delta}_i^{\mathsf{U}} \in \mathbb{R}^{m\times mN}$, $\mathbf{\Delta}_i^{\mathsf{Y}} \in \mathbb{R}^{p\times mN}$, $\mathbf{\Delta}_i^{\mathsf{A}} \in \mathbb{R}^{p\times n_{\eta\text{-aux}}}$ and $\mathbf{\Delta}^{\mathsf{M}} \in \mathbb{R}^{pN\times n_{\eta\text{-aux}}}$ can be found in Appendix A, and where $\mathcal{M} \in \mathbb{R}^{mN\times pN}$ is a concatenation of M_t^s as in (15).

4) SDDPC Optimization Problem and Algorithm: With the results above, we are now ready to mirror the steps of getting (18) and formulate a distributionally robust optimized-gain Stochastic Data-Driven Predictive Control (SDDPC) problem,

$$\underset{\bar{u}, M_t^s}{\mathrm{minimize}} \; (31) \text{ s.t. } (28) \text{ for } t \in \mathbb{Z}_{[k,k+N)}, (27), (29) \tag{30}$$

where the quadratic cost function is analogous to (17) as

$$\textstyle\sum_{t=k}^{k+N-1}\big[J_t(\bar{u}_t, \bar{\mathbf{y}}_t) + \|\mathrm{Diag}(R,Q)^{\frac{1}{2}}\mathbf{\Lambda}_t(\mathbf{\Sigma}^{\eta})^{\frac{1}{2}}\|_{\mathsf{F}}^2\big]. \tag{31}$$

Problem (30) has a unique optimal solution if feasible, similar as problem (18). The solution $(\bar{u}, M_t^s)$ finishes parameterization of the control policies $\boldsymbol{\pi}_{[k,k+N)}$ via (26), where the first N_{c} policies are applied to the system. At the next control step $t = k + N_{\mathrm{c}}$, the state mean $\boldsymbol{\mu}_{k+N_{\mathrm{c}}}^{\mathsf{x}}$ is iterated as the estimate $\hat{\mathbf{x}}_{k+N_{\mathrm{c}}}$ via (9), with a backup value $\boldsymbol{\mu}_{k+N_{\mathrm{c}}}^{\bar{\mathsf{x}}}$ of $\boldsymbol{\mu}_{k+N_{\mathrm{c}}}^{\mathsf{x}}$ equal to the nominal state $\bar{\mathbf{x}}_{k+N_{\mathrm{c}}}$ via (10). The method is formally summarized in Algorithm 2.

C. Theoretical Equivalence of SMPC and SDDPC

We establish theoretical results in this section, starting by an underlying relation between the means of x_k and $\mathbf{x}_k$.

Lemma 4 (Related Means of x_k and $\mathbf{x}_k$ [16])**.** *If μ_k^{x} is the mean of x_k and $\boldsymbol{\mu}_k^{\mathsf{x}}$ is the mean of $\mathbf{x}_k$, then they satisfy*

$$\mu_k^{\mathsf{x}} = \Phi_{\mathsf{orig}}\,\widetilde{\mu}_k^{\mathsf{x}}, \qquad \boldsymbol{\mu}_k^{\mathsf{x}} = \Phi_{\mathsf{aux}}\,\widetilde{\mu}_k^{\mathsf{x}} \tag{32}$$

for some $\widetilde{\mu}_k^{\mathsf{x}} \in \mathbb{R}^{mL+n(L+1)}$, where matrices $\Phi_{\mathsf{orig}}, \Phi_{\mathsf{aux}}$ are

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

with the matrices $\mathcal{C}$, $\mathcal{O}$, $\mathcal{G}$ defined in Section III-B-1 and $\mathcal{C}_{\mathsf{w}} := [A^{L-1}, \ldots, A, I_n]$, $\mathcal{G}_{\mathsf{w}} := \mathrm{Toep}(0_{p\times n}, C, CA, \ldots, CA^{L-2})$.

As we assume (32) holds, the SMPC and SDDPC problems will have equal feasible and optimal sets.

Proposition 5 (Equivalence of Optimization Problems)**.** *If the parameters $\mu_k^{\mathsf{x}}, \boldsymbol{\mu}_k^{\mathsf{x}}$ satisfy* (32)*, then the optimal (resp. feasible) solution set of SDDPC problem* (30) *is equal to the optimal (resp. feasible) solution set of SMPC problem* (18).

Proof. We first claim that, for all $\bar{u}$ and M_t^s, we have

$$\bar{y}_t = \bar{\mathbf{y}}_t, \qquad \Lambda_t \Sigma^{\eta} \Lambda_t^{\mathsf{T}} = \mathbf{\Lambda}_t \mathbf{\Sigma}^{\eta} \mathbf{\Lambda}_t^{\mathsf{T}} \tag{33}$$

for $t \in \mathbb{Z}_{[k,k+N)}$, which is explained in Appendix B. Given (33), the objective function (17) of problem (18) and objective function (31) of problem (30) are equal, and the constraint (16) in problem (18) and constraint (28) in problem (30) are equivalent. Thus the problems (18) and (30) have the same objective function and constraints, and the result follows. ■

We present in Theorem 7 our main theoretical result, saying that our proposed SDDPC control method and the benchmark SMPC method will result in identical control actions, under idealized conditions in Assumption 6.

Assumption 6 (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 (19); the input data u^{d} is persistently exciting of order $L + 1 + n$.

Algorithm 2 Distributionally Robust Optimized-Gain Stochastic Data-Driven Predictive Control (DR/O-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, and CVaR level α.

1: Compute $\mathbf{\Gamma}$ and D as in Section III-B-1 using data $u^{\mathrm{d}}, y^{\mathrm{d}}$, and formulate matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}$ as in Section III-B-2.
2: Compute $\mathbf{\Sigma}^{\mathsf{x}}, \mathbf{L}_{\mathsf{L}}$ via (24) and $\mathbf{\Delta}_{[0,N)}^{\mathsf{U}}, \mathbf{\Delta}_{[0,N)}^{\mathsf{Y}}, \mathbf{\Delta}_{[0,N)}^{\mathsf{A}}$, $\mathbf{\Delta}^{\mathsf{M}}$ through Appendix A.
3: Initialize the control step $k \leftarrow 0$ and set $\boldsymbol{\mu}_0^{\mathsf{x}} \leftarrow \boldsymbol{\mu}_{\mathsf{ini}}^{\mathsf{x}}$.
4: Solve $\bar{u}_{[k,k+N)}$ and M_t^s from problem (30).
5: **If** (30) is infeasible **then** Set $\boldsymbol{\mu}_k^{\mathsf{x}} \leftarrow \boldsymbol{\mu}_k^{\bar{\mathsf{x}}}$, and redo line 4.
6: **for** t **from** k **to** $k + N_{\mathrm{c}} - 1$ **do**
7:     Input $u_t \leftarrow \boldsymbol{\pi}_t(\boldsymbol{\nu}_{[k,t)})$ in (26) to the system (1).
8:     Measure y_t from the system (1).
9:     Compute $\boldsymbol{\nu}_t$ via (25).
10: Set $(\boldsymbol{\mu}_{k+N_{\mathrm{c}}}^{\mathsf{x}}, \boldsymbol{\mu}_{k+N_{\mathrm{c}}}^{\bar{\mathsf{x}}})$ as $(\hat{\mathbf{x}}_{k+N_{\mathrm{c}}}, \bar{\mathbf{x}}_{k+N_{\mathrm{c}}})$ in (25), (27)
11: Set $k \leftarrow k + N_{\mathrm{c}}$. Go back to line 4

(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 $\widetilde{\mu}^{\mathsf{x}}_{\mathsf{ini}} \in \mathbb{R}^{mL+n+nL}$ satisfying $\mu^{\mathsf{x}}_{\mathsf{ini}} = \Phi_{\mathsf{orig}}\widetilde{\mu}^{\mathsf{x}}_{\mathsf{ini}}$. (Such $\widetilde{\mu}^{\mathsf{x}}_{\mathsf{ini}}$ always exists because Φ_{orig} has full row rank.)

Theorem 7 (Equivalence of SMPC and SDDPC). *Consider system* (1) *with initial state* x_0 *and a specific noise realization* $\{w_t, v_t\}_{t=0}^{\infty}$, *and consider the following two 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 6.*

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. The proof is similar to the proof of [16, Thm. 9], requiring Proposition 5 and the fact that both problems (18) and (30) have unique optimal solutions if feasible. ■

While in practice Assumption 6 may not hold, noisy offline data can be accommodated as discussed in Section III-B-1, and Σ^{ρ} becomes a tuning parameter of our SDDPC method.

IV. Numerical Case Study

In this section, we numerically test our proposed method on a batch reactor system introduced in [25] and applied in [14], [26]. The system has $n = 4$ states, $m = 2$ inputs and $p = 2$ outputs, and the discrete-time system matrices with sampling period 0.1s are

$$\left[\begin{array}{c|c} A & B \\ \hline C & \end{array}\right] = \left[\begin{array}{cccc|cc} 1.178 & .001 & .511 & -.403 & .004 & -.087 \\ -.051 & .661 & -.011 & .061 & .467 & .001 \\ .076 & .335 & .560 & .382 & .213 & -.235 \\ 0 & .335 & .089 & .849 & .213 & -.016 \\ \hline 1 & 0 & 1 & -1 & & \\ 0 & 1 & 0 & 0 & & \end{array}\right].$$

The process/sensor noise on each state/output follows the t-distribution of 2 DOFs scaled by 10^{-4}, which is a heavy-tailed distribution. Control parameters are reported in TABLE I. We collected offline data of length $T_{\mathrm{d}} = 600$ from the noisy system, where the input data was the outcome of a PI controller $U(s) = \begin{bmatrix} 0 & -1/s \\ 2+1/s & 0 \end{bmatrix} Y(s)$ plus a white-noise signal of noise power 10^{-2}. In the online control process, the reference signal is $r_t = [0, 0]^{\mathsf{T}}$ from time 0s to time 30s, alternates between $[0, 0]^{\mathsf{T}}$ and $[0.3, 0]^{\mathsf{T}}$ from 30s to 60s, and is $r_t = [0.5, 0]^{\mathsf{T}}$ from 60s to 90s. With our proposed SDDPC method, the first output signal is in Fig. 1; the signal remains around 0.4 from 60s to 90s because of the safety constraint specified in TABLE I.

For comparison purposes, we implemented the simulation with different controllers. In addition to distributionally robust optimized-gain (DR/O) SMPC and SDDPC in this paper, we applied the SMPC and SDDPC frameworks from [16], which use chance constraints and a fixed feedback gain (CC/F). To observe separate impacts of using the DR constraint and optimized gains, we also implement SMPC and SDDPC with DR constraints and a fixed feedback gain (DR/F). We also compare to DeePC, SPC and deterministic MPC as

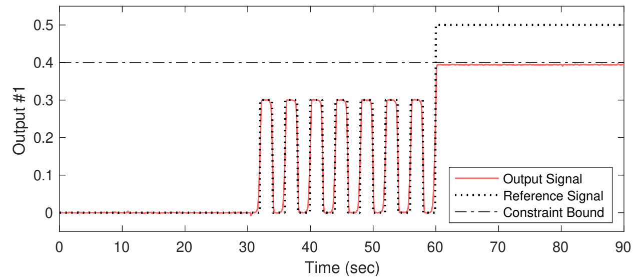

Fig. 1. The system's first output signal with DR/O-SDDPC.

benchmarks. The model used in MPC methods is identified from the same offline data in the data-driven controllers.

The simulation results are summarized in TABLE II. We evaluate (i) the controllers' tracking performance through the tracking cost from 0s to 60s and (ii) the controllers' ability to satisfy constraints according to the cumulative amount of constraint violation between 60s and 90s, when the first output signal hits the constraint margin. When the reference signal is constant (0s–30s), SMPC and SDDPC tracked better than other methods, aligning with the observation in [16]. Comparing DR/F and CC/F methods, the controllers with DR constraints achieved lower amounts of constraint violation (60s–90s), while the tracking performance is slightly worse during 30s–60s when the reference signal has frequent step changes. Comparing DR/O and DR/F methods, we observe that the methods with optimized gain achieved lower tracking costs when the reference signal changes frequently (30s–60s).

V. Conclusions

We proposed a Stochastic Data-Driven Predictive Control (SDDPC) method that accommodates distributionally robust (DR) probability constraints and produces closed-loop control policies with feedback gains determined from optimization. In theory, our SDDPC method can produce equivalent control inputs with associated Stochastic MPC, under specific conditions. Simulation results indicated separate benefits of using DR constraints and optimized feedback gains.

Appendix A. Definition of Δ^{U}_i, Δ^{Y}_i, Δ^{A}_i, Δ^{M}

The matrices $\Delta^{\mathsf{U}}_i \in \mathbb{R}^{m \times mN}$, $\Delta^{\mathsf{Y}}_i \in \mathbb{R}^{p \times mN}$, $\Delta^{\mathsf{A}}_i \in \mathbb{R}^{p \times n_\eta}$ for $i \in \mathbb{Z}_{[0,N)}$ and $\Delta^{\mathsf{M}} \in \mathbb{R}^{pN \times n_\eta}$ in (14) are what follows,

$$\begin{aligned} \mathrm{col}\big(\Delta^{\mathsf{U}}_0, \dots, \Delta^{\mathsf{U}}_{N-1}\big) &:= I_{mN} \\ \mathrm{col}\big(\Delta^{\mathsf{Y}}_0, \dots, \Delta^{\mathsf{Y}}_{N-1}\big) &:= \Xi(A)\,(I_N \otimes B) \\ \mathrm{col}\big(\Delta^{\mathsf{A}}_0, \dots, \Delta^{\mathsf{A}}_{N-1}\big) &:= [\Theta(A), \Xi(A), I_{pN}] \\ \Delta^{\mathsf{M}} &:= [\Theta(A_{\mathsf{L}}), \Xi(A_{\mathsf{L}}), I_{pN} - \Xi(A_{\mathsf{L}})\,(I_N \otimes L_{\mathsf{L}})] \end{aligned}$$

where we let $\Theta(A) := \mathrm{col}(C, CA, \dots, CA^{N-1}) \in \mathbb{R}^{pN \times n}$, $\Xi(A) := \mathrm{Toep}(0_{p\times n}, C, CA, \dots, CA^{N-2}) \in \mathbb{R}^{pN \times nN}$, and similarly define $\Theta(A_{\mathsf{L}}), \Xi(A_{\mathsf{L}})$ with $A_{\mathsf{L}} := A - L_{\mathsf{L}}C$.

The matrices $\boldsymbol{\Delta}^{\mathsf{U}}_i, \boldsymbol{\Delta}^{\mathsf{Y}}_i, \boldsymbol{\Delta}^{\mathsf{A}}_i, \boldsymbol{\Delta}^{\mathsf{M}}$ in (29) are computed (with underlying $\boldsymbol{\Theta}(\mathbf{A}), \boldsymbol{\Xi}(\mathbf{A}), \mathbf{A}_{\mathsf{L}}$) in the same way as above, with $A, B, C, L_{\mathsf{L}}, n$ replaced by $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{L}_{\mathsf{L}}, n_{\mathsf{aux}}$, respectively.

TABLE I

CONTROL PARAMETERS

Time horizon lengths	$L = 5,\ N = 15,\ N_{\mathrm{c}} = 5$
Cost matrices	$Q = 10^3 I_p,\ R = I_m$
Safety constraint coefficients	$E = I_{m+p} \otimes \left[\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right]$
	$f = [.1\ .1\ .5\ .1\ .4\ .4\ .4\ .4]^{\mathsf{T}}$
CVaR level[a]	$\alpha = 0.3$
Variance of v_t for SMPC/SDDPC	$\Sigma^{\mathsf{v}} = 5 \times 10^{-7} I_p$
Variance of ρ_t for SDDPC	$\Sigma^{\rho} = 10^{-7} I_{pL}$
Variance of w_t for SMPC[b]	$\Sigma^{\mathsf{w}} = \mathcal{O}^{\dagger} \Sigma^{\rho} \mathcal{O}^{\dagger\mathsf{T}}$

[a] α is used as the risk bound for chance constrained controllers.
[b] $\mathcal{O}$ is obtained given the identified model (A, B, C, D) in SMPC.

TABLE II

SIMULATION RESULT STATISTICS

Controller	Total Tracking Cost 0s to 30s	Total Tracking Cost 30s to 60s	Cumulative Violation from 60s to 90s
DR/O-SDDPC[a]	0.02	64.2	0
DR/F-SDDPC	0.02	68.9	0
CC/F-SDDPC	0.02	64.9	0.03
DR/O-SMPC	0.02	64.2	0
DR/F-SMPC	0.02	68.0	0
CC/F-SMPC	0.02	64.9	0.01
deterministic MPC	0.09	64.6	0.20
SPC	0.18	65.5	2.23
DeePC	0.18	64.7	0.19

[a] DR – distributionally robust constrained, CC – chance constrained, O – with optimized feedback gain, F – with fixed feedback gain.

APPENDIX B. PROOF OF (33)

Proof. The relation $\bar{y}_t = \bar{\mathbf{y}}_t$ in (33) was established in [16, Claim 7.7]. The other relation in (33) is equivalent to

$$\Delta^{\mathsf{U}}_{t-k} = \boldsymbol{\Delta}^{\mathsf{U}}_{t-k}, \qquad \Delta^{\mathsf{M}} \Sigma^{\eta} (\Delta^{\mathsf{M}})^{\mathsf{T}} = \boldsymbol{\Delta}^{\mathsf{M}} \boldsymbol{\Sigma}^{\eta} (\boldsymbol{\Delta}^{\mathsf{M}})^{\mathsf{T}},$$
$$\Delta^{\mathsf{Y}}_{t-k} = \boldsymbol{\Delta}^{\mathsf{Y}}_{t-k}, \quad \Delta^{\mathsf{A}}_{t-k} \Sigma^{\eta} (\Delta^{\mathsf{A}}_{t-k})^{\mathsf{T}} = \boldsymbol{\Delta}^{\mathsf{A}}_{t-k} \boldsymbol{\Sigma}^{\eta} (\boldsymbol{\Delta}^{\mathsf{A}}_{t-k})^{\mathsf{T}},$$

via the definitions of Λ_t and $\boldsymbol{\Lambda}_t$ in (14) and (29). Given the definitions in Appendix A, the above relations are implied by

1) $CA^q B = \mathbf{C}\mathbf{A}^q \mathbf{B}$ for $q \in \mathbb{Z}_{[0,N)}$,
2) $CA^q \Sigma^{\mathsf{x}} (CA^r)^{\mathsf{T}} = \mathbf{C}\mathbf{A}^q \boldsymbol{\Sigma}^{\mathsf{x}} (\mathbf{C}\mathbf{A}^r)^{\mathsf{T}}$ for $q, r \in \mathbb{Z}_{[0,N)}$,
3) $CA^q \Sigma^{\mathsf{w}} (CA^r)^{\mathsf{T}} = \mathbf{C}\mathbf{A}^q \boldsymbol{\Sigma}^{\mathsf{w}} (\mathbf{C}\mathbf{A}^r)^{\mathsf{T}}$ for $q, r \in \mathbb{Z}_{[0,N-2]}$,
4) $CA_{\mathsf{L}}^q \Sigma^{\mathsf{x}} (CA_{\mathsf{L}}^q)^{\mathsf{T}} = \mathbf{C}\mathbf{A}_{\mathsf{L}}^q \boldsymbol{\Sigma}^{\mathsf{x}} (\mathbf{C}\mathbf{A}_{\mathsf{L}}^r)^{\mathsf{T}}$ for $q, r \in \mathbb{Z}_{[0,N)}$,
5) $CA_{\mathsf{L}}^q \Sigma^{\mathsf{w}} (CA_{\mathsf{L}}^q)^{\mathsf{T}} = \mathbf{C}\mathbf{A}_{\mathsf{L}}^q \boldsymbol{\Sigma}^{\mathsf{w}} (\mathbf{C}\mathbf{A}_{\mathsf{L}}^r)^{\mathsf{T}}$ for $q, r \in \mathbb{Z}_{[0,N-2]}$,
6) $CA_{\mathsf{L}}^q L_{\mathsf{L}} = \mathbf{C}\mathbf{A}_{\mathsf{L}}^q \mathbf{L}_{\mathsf{L}}$ for $q \in \mathbb{Z}_{[0,N-2]}$,

where the relations 1)–6) can be shown given the equalities

$$A\Phi\Phi_{\mathsf{aux}} = \Phi\mathbf{A}\Phi_{\mathsf{aux}}, \quad B = \Phi\mathbf{B}, \quad L_{\mathsf{L}} = \Phi\mathbf{L}_{\mathsf{L}},$$
$$\mathbf{A}\Phi_{\mathsf{aux}} = \Phi_{\mathsf{aux}}\widetilde{A}, \quad \mathbf{B} = \Phi_{\mathsf{aux}}\widetilde{B}, \quad \mathbf{L}_{\mathsf{L}} = \Phi_{\mathsf{aux}}\widetilde{L}_{\mathsf{L}},$$
$$C\Phi\Phi_{\mathsf{aux}} = \mathbf{C}\Phi_{\mathsf{aux}}, \quad \Sigma^{\mathsf{w}} = \Phi\boldsymbol{\Sigma}^{\mathsf{w}}\Phi^{\mathsf{T}}, \quad \Sigma^{\mathsf{x}} = \Phi\boldsymbol{\Sigma}^{\mathsf{x}}\Phi^{\mathsf{T}},$$

established with some matrices $\Phi, \widetilde{A}, \widetilde{B}, \widetilde{L}_{\mathsf{L}}$ according to [16, Claim 5.1, Claim 7.1, Claim 7.4]. ■

REFERENCES

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

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

[3] 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.

[4] 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.

[5] 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.

[6] 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.

[7] J. Coulson, J. Lygeros, and F. Dörfler, "Data-enabled predictive control: In the shallows of the DeePC," in *Proc. ECC*, 2019, pp. 307–312.

[8] ——, "Regularized and distributionally robust data-enabled predictive control," in *Proc. IEEE CDC*, 2019, pp. 2696–2701.

[9] ——, "Distributionally robust chance constrained data-enabled predictive control," *IEEE Trans. Autom. Control*, vol. 67, no. 7, pp. 3289–3304, 2021.

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

[11] 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.

[12] P. G. Carlet, A. Favato, S. Bolognani, and F. Dörfler, "Data-driven predictive current control for synchronous motor drives," in *ECCE*, 2020, pp. 5148–5154.

[13] L. Huang, J. Coulson, J. Lygeros, and F. Dörfler, "Decentralized data-enabled predictive control for power system oscillation damping," *IEEE Trans. Control Syst. Tech.*, vol. 30, no. 3, pp. 1065–1077, 2021.

[14] G. Pan, R. Ou, and T. Faulwasser, "Towards data-driven stochastic predictive control," *Int. J. Robust Nonlinear Control*, 2022.

[15] ——, "On a stochastic fundamental lemma and its use for data-driven optimal control," *IEEE Trans. Autom. Control*, 2022.

[16] 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.

[17] B. P. Van Parys, D. Kuhn, P. J. Goulart, and M. Morari, "Distributionally robust control of constrained stochastic systems," *IEEE Trans. Autom. Control*, vol. 61, no. 2, pp. 430–442, 2015.

[18] S. Zymler, D. Kuhn, and B. Rustem, "Distributionally robust joint chance constraints with second-order moment information," *Math. Program.*, vol. 137, pp. 167–198, 2013.

[19] 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.

[20] E. Joa, M. Bujarbaruah, and F. Borrelli, "Output feedback stochastic mpc with hard input constraints," in *Proc. ACC*, 2023, pp. 2034–2039.

[21] J. Ridderhof, K. Okamoto, and P. Tsiotras, "Chance constrained covariance control for linear stochastic systems with output feedback," in *Proc. IEEE CDC*, 2020, pp. 1758–1763.

[22] P. J. Goulart and E. C. Kerrigan, "Output feedback receding horizon control of constrained systems," *Int J Control*, vol. 80, no. 1, pp. 8–20, 2007.

[23] J. Humpherys, P. Redd, and J. West, "A fresh look at the kalman filter," *SIAM review*, vol. 54, no. 4, pp. 801–823, 2012.

[24] L. E. Ghaoui, M. Oks, and F. Oustry, "Worst-case value-at-risk and robust portfolio optimization: A conic programming approach," *Oper. Res.*, vol. 51, no. 4, pp. 543–556, 2003.

[25] H. Ye, "Scheduling of networked control systems," *IEEE Control Syst.*, vol. 21, no. 1, pp. 57–65, 2001.

[26] C. De Persis and P. Tesi, "Formulas for data-driven control: Stabilization, optimality, and robustness," *IEEE Trans. Autom. Control*, vol. 65, no. 3, pp. 909–924, 2019.