ETH zürich

Automatic Control Laboratory

Stochastic Data-Driven Predictive Control:
Regularization, Estimation, and Constraint Tightening

Mingzhou Yin, Andrea Iannelli, Roy S. Smith

July 17, 2024, SYSID 2024

The LTI stochastic data-driven predictive control problem

- **Data-driven:** Input-disturbance-output trajectory in place of model parameters

$$Z = \begin{bmatrix} z_0^d & \cdots & z_{M-1}^d \end{bmatrix}, \; z_i^d := \text{col}\left(u_{t_i}^d, \ldots, u_{t_i+L-1}^d \middle| w_{t_i}^d, \ldots, w_{t_i+L-1}^d \middle| y_{t_i}^d, \ldots, y_{t_i+L-1}^d\right)$$

- **Predictive:** $L = L_0$ (init. cond. length) $+ L'$ (prediction horizon)

$$\hat{\mathbf{y}}^t \sim \mathcal{D}(\underbrace{\hat{\mathbf{u}}^t := (\hat{u}_k^t)_{k=0}^{L'-1}}_{\text{control inputs}}, \underbrace{\mathbf{w}^t := (w_k)_{k=t-L_0}^{t+L'-1}}_{\text{disturbance sequence}}, \underbrace{\mathbf{u}_{\text{ini}}^t := (u_k)_{k=t-L_0}^{t-1}, \mathbf{y}_{\text{ini}}^t := (y_k)_{k=t-L_0}^{t-1}}_{\text{input \& output init. cond.}})$$

- **Control:** Receding horizon control

$$\min_{(\hat{u}_k^t)_{k=0}^{L'-1}} \; J_t := \sum_{k=0}^{L'-1} \left\| \hat{u}_k^t \right\|_R^2 + \mathbb{E}\left[\sum_{k=0}^{L'-1} \left\| \hat{y}_k^t - r_{t+k} \right\|_Q^2 \right]$$

$$\text{s.t.} \quad \hat{u}_k^t \in \mathcal{U}_{t+k}, \; H^{t+k} \hat{y}_k^t \leq q^{t+k} \text{ with high probability}, \; \forall\, k = 0, \ldots, L'-1$$

The LTI stochastic data-driven predictive control problem

Stochastic: Two sources of uncertainties

- Zero-mean i.i.d. noise cov $(v_t) = \sigma^2 \mathbb{I}_{n_y} \rightarrow$ uncertainties in signal matrix Z & output initial condition $\mathbf{y}_{\text{ini}}^t : \mathbb{E}\left[\mathbf{y}_{\text{ini}}^t\right] = \bar{\mathbf{y}}_{\text{ini}}^t$, cov $\left(\mathbf{y}_{\text{ini}}^t\right) = P_t$
- Uncertainties in online disturbance sequence $\mathbf{w}^t : \mathbb{E}\left[\mathbf{w}^t\right] = \bar{\mathbf{w}}^t$, cov $\left(\mathbf{w}^t\right) = \Sigma_w$

Agenda:

- Accurate predictor under multiple uncertainties
- Tractable formulation of the expected cost
- Tractable formulation of output contraints

The stochastic predictor[1]

Stochastic data-driven prediction

Consider augmented inputs $\psi_t := \text{col}\,(u_t, w_t)$. The distribution of stochastic output prediction can be characterized by $\mathbb{E}\left[\hat{\mathbf{y}}^t \middle| \, g^t\right] = \bar{\mathbf{y}}^t$, cov $\left(\hat{\mathbf{y}}^t \middle| \, g^t\right) = \Sigma^t$, where

$$\begin{aligned}
\bar{\mathbf{y}}^t &:= Y_f g^t - \Gamma(Y_p g^t - \bar{\mathbf{y}}^t_{\text{ini}}), \\
\Sigma^t &:= \Gamma P_t \Gamma^\top + \Gamma_w \Sigma_w \Gamma_w^\top + \left\| g^t \right\|_2^2 T, \\
T &:= \sigma^2 \left(\Gamma\Gamma^\top + \mathbb{I}_{n_y L'} \right), \quad \Gamma_w := (Y_f - \Gamma Y_p) R_3, \\
\Gamma &:= \text{col}\left(CA^{L_0}, \ldots, CA^{L-1} \right) \text{col}\left(C, \ldots, CA^{L_0 - 1} \right)^\dagger .
\end{aligned}$$

[1] Yin, M., Iannelli, A., and Smith, R.S. (2022). Data-driven prediction with stochastic data: Confidence regions and minimum mean-squared error estimates. In European Control Conference (ECC), 853–858.

The stochastic predictor

- g^t can be been as a hyperparameter, typically selected by solving

$$g^t := \underset{g}{\text{argmin}} \ \left\| Y_p g - \bar{\mathbf{y}}^t_{\text{ini}} \right\|_S^2 + \lambda \|g\|_2^2 = \begin{bmatrix} R_1 & R_2 & R_3 & R_4 \end{bmatrix} \text{col}\left(\mathbf{u}^t_{\text{ini}}, \mathbf{u}^t, \bar{\mathbf{w}}^t, \bar{\mathbf{y}}^t_{\text{ini}}\right)$$
$$\text{s.t.} \quad \Psi g = \text{col}\left(\mathbf{u}^t_{\text{ini}}, \hat{\mathbf{u}}^t, \bar{\mathbf{w}}^t\right)$$

- The model parameter Γ can be estimated consistently by $\hat{\Gamma}_Z = Y_f R_4 \left(Y_p R_4\right)^{-1}$ under mild conditions

Stochastic control cost

Expected control cost

The expected control cost is quadratic with

$$J_t = \left\|\hat{\mathbf{u}}^t\right\|_{\bar{R}}^2 + \left\|\bar{\mathbf{y}}^t - \mathbf{r}^t\right\|_{\bar{Q}}^2 + \text{tr}\left(\bar{Q}T\right)\left\|g^t\right\|_2^2 + \text{constant},$$

where $\bar{R} := \mathbb{I}_{L'} \otimes R$, $\bar{Q} := \mathbb{I}_{L'} \otimes Q$, $\mathbf{r}^t := (r_{t+k})_{k=0}^{L'-1}$, $T := \sigma^2\left(\Gamma\Gamma^\top + \mathbb{I}_{n_y L'}\right)$.

- The $\|g^t\|_2^2$-regularization term is commonly seen in DDPC literature, but this result provides a practical approach to selecting the weighting factor

Initial condition estimation

- In standard DDPC, the output initial condition $\mathbf{y}_{\text{ini}}^t$ is directly measured
  $\Rightarrow$ constant covariance = measurement error
- In MPC, the initial condition x_t is estimated from both measurement y_t and previous prediction $x_{t|t-1}$
  $\Rightarrow$ diminishing error covariance

- **Idea:** Update $\mathbf{y}_{\text{ini}}^t$ with Kalman-filtered measurement from previous prediction
- **Method:** Consider the stochastic predictor as a non-minimal state-space "model" with "state"

$$\bar{x}_t := \text{col}\left(u_{t-L_0}, \ldots, u_{t-1}, y^0_{t-L_0}, \ldots, y^0_{t-1}\right)$$

Initial condition estimation

$$\begin{cases} \bar{x}_{t+1} &= \begin{bmatrix} \Lambda^{n_u} & \mathbf{0} \\ \mathbf{0} & \Lambda^{n_y} \end{bmatrix} \bar{x}_t + \begin{bmatrix} \mathbf{0} \\ \hat{u}_0^t \\ \mathbf{0} \\ \bar{y}_0^t \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ e_0^t \end{bmatrix}, \\ \zeta_{t+1} &= \begin{bmatrix} \mathbf{0} & \mathbb{I}_{n_y} \end{bmatrix} \bar{x}_{t+1} + v_t = y_t^0 + v_t = y_t \end{cases} \qquad \bar{x}_t = \begin{bmatrix} u_{t-L_0} \\ \vdots \\ u_{t-1} \\ y_{t-L_0}^0 \\ \vdots \\ y_{t-1}^0 \end{bmatrix}$$

- Λ: upper shift operator
- e_0^t: one-step-ahead prediction error with covariance Σ_0^t
- v_t: measurement noise with variance $\sigma^2 \mathbb{I}_{n_y}$
- Standard Kalman filter design can be done

Chance constraint satisfaction

Both element-wise chance constraints

$$\text{Pr}\,({h_i^{t+k}}^\top \hat{y}_k^t \leq q_i^{t+k}) \geq p,\ \forall\, i = 1, \ldots, n_c, k = 0, \ldots, L' - 1$$

and set-wise chance constraints

$$\text{Pr}\left(H^{t+k} \hat{y}_k^t \leq q^{t+k}\right) \geq p,\ \forall\, k = 0, \ldots, L' - 1$$

can be guaranteed by constraint tightening.

- Unlike usual uncertainty assumptions, error depends on inputs via g^t
- Online calculation of constraint tightening is required

Chance constraint satisfaction

Constraint tightening

Define augmented linear constraints $\bar{H}^t \mathbf{y} \leq \bar{q}^t$. The constraint

$$\bar{q}^t - \bar{H}^t \bar{\mathbf{y}}^t \geq \mu \sqrt{\text{diag}\left(\bar{H}^t \Sigma^t \bar{H}^{t\top}\right)} \qquad (\triangle)$$

guarantees the satisfaction of 1) element-wise chance constraints if $\mu \geq \sqrt{\frac{1}{1-p} - 1}$ and 2) set-wise chance constraints if $\mu \geq \sqrt{\frac{n_y}{1-p}}$.

Proof sketch: 1) uses one-sided Chebyshev's inequality

$$\Pr\left(\bar{h}_i^t \hat{\mathbf{y}}^t - \bar{h}_i^t \bar{\mathbf{y}}^t \leq \sqrt{\frac{1}{1-p} - 1} \cdot \text{std}\left(\bar{h}_i^t \hat{\mathbf{y}}^t\right)\right) \geq p, \ \forall i;$$

2) uses multi-dimensional Chebyshev's inequality

$$\Pr\left(e_k^{t\top} \left(\Sigma_k^t\right)^{-1} e_k^t \leq \frac{n_y}{1-p}\right) \geq p$$

Chance constraint satisfaction

- Unfortunately, the tightened constraints are non-convex
- Convex surrogate obtained by using $\sqrt{\sum_i a_i} \leq \sum_i \sqrt{a_i}$

Convex surrogate of chance constraints

($\triangle$) is guaranteed by second-order cone constraint

$$\bar{q}^t - \bar{H}^t \bar{\mathbf{y}}^t \geq \mu \left(\mathbf{c}_1 + \mathbf{c}_2 \left\| g^t \right\|_2 \right),$$

where

$$\mathbf{c}_1 := \sqrt{\text{diag}\left(\bar{H}^t \left(\Gamma P_t \Gamma^\top + \Gamma_w \Sigma_w \Gamma_w^\top \right) \bar{H}^{t\top} \right)}, \quad \mathbf{c}_2 := \sqrt{\text{diag}\left(\bar{H}^t T \bar{H}^{t\top} \right)}.$$

Stochastic indirect DDPC

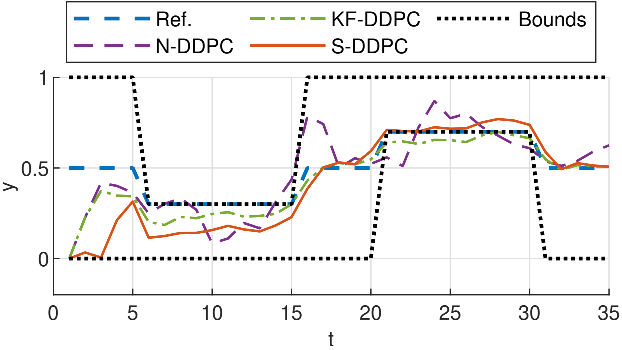
1: Select a data-driven predictor and calculate predictor parameters $R_1, R_2, R_3, R_4, \hat{\Gamma}_Z, T$.
2: Initialize the Kalman filter.
3: **for** $t \leftarrow 0, 1, \ldots$ **do**
4: Update Kalman filtered initial conditions $\text{col}\left(\mathbf{u}^t_{\text{ini}}, \mathbf{y}^t_{\text{ini}}\right) \leftarrow \bar{x}_{t,t}$, $P_t \leftarrow P_{t,t}$.
5: Solve

$$
\begin{aligned}
\hat{\mathbf{u}}^t \leftarrow \underset{\hat{\mathbf{u}}^t}{\text{argmin}} \quad & \left\|\hat{\mathbf{u}}^t\right\|^2_{\bar{R}} + \left\|\bar{\mathbf{y}}^t - \mathbf{r}^t\right\|^2_{\bar{Q}} + \text{tr}\left(\bar{Q}T\right)\left\|g^t\right\|^2_2 \\
\text{s.t.} \quad & g^t = \begin{bmatrix} R_1 & R_2 & R_3 & R_4 \end{bmatrix} \text{col}\left(\mathbf{u}^t_{\text{ini}}, \mathbf{u}^t, \bar{\mathbf{w}}^t, \bar{\mathbf{y}}^t_{\text{ini}}\right), \\
& \bar{\mathbf{y}}^t = Y_f g^t - \Gamma(Y_p g^t - \bar{\mathbf{y}}^t_{\text{ini}}), \\
& \bar{q}^t - \bar{H}^t \bar{\mathbf{y}}^t \geq \mu\left(\mathbf{c}_1 + \mathbf{c}_2\left\|g^t\right\|_2\right), \\
& \hat{u}^t_k \in \mathcal{U}_{t+k},\ \forall\, k = 0, \ldots, L' - 1.
\end{aligned}
$$

6: Apply $u_t = \hat{u}^t_0$ to the system and measure y_t.
7: **end for**

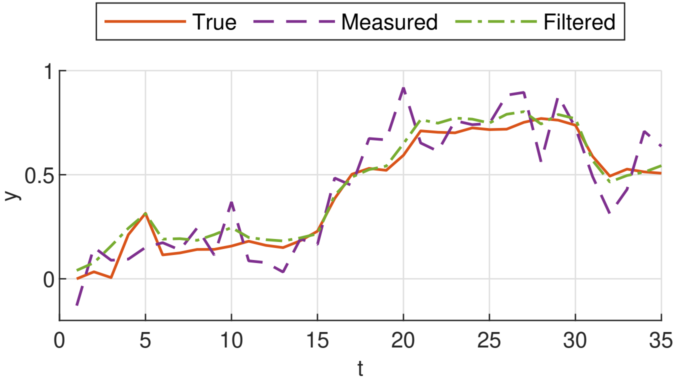
Numerical example

- Fourth-order dynamics
- $L_0 = 4$, $L' = 10$, $Q = 20$, $R = 1$, $\sigma^2 = 0.01$, $p = 0.95$, $\bar{\mathbf{w}}^t = \mathbf{0}$, $\Sigma_w = 0.001 \cdot \mathbb{I}$
- Gaussian noise and disturbance
- 500 offline data points
- Lower and upper output bounds
- No input constraints



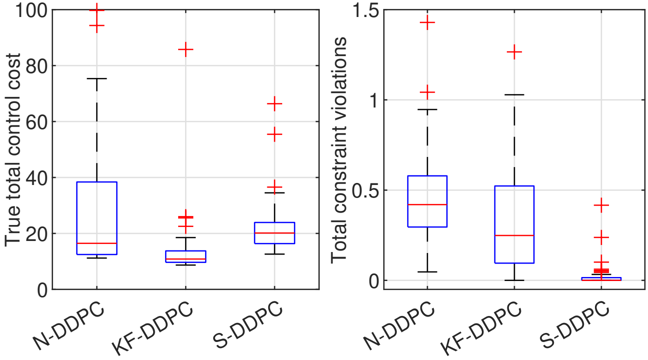
N-DDPC: nominal DDPC
KF-DDPC: DDPC with Kalman filtering
S-DDPC: DDPC with KF & constraint tightening

Numerical example



N-DDPC: nominal DDPC
KF-DDPC: DDPC with Kalman filtering
S-DDPC: DDPC with KF & constraint tightening

Numerical example



- 50 Monte Carlo simulations
- Constraint violation $:= \sum_t \max\left(H^t y_t - q^t, 0\right)$

N-DDPC: nominal DDPC
KF-DDPC: DDPC with Kalman filtering
S-DDPC: DDPC with KF & constraint tightening

ETH zürich

Stochastic Data-Driven Predictive Control: Regularization, Estimation, and Constraint Tightening

- A tuning-free regularizer design in the control cost
- Improved initial condition estimation by Kalman filtering
- Reliable constraint satisfaction by constraint tightening