Robust-Adaptive Control of Linear Systems: beyond Quadratic Costs

Edouard Leurent, Denis Efimov, Odalric-Ambrym Maillard

Motivation

- **Adaptive**: estimate the dynamics along the way
- **Robust**: avoid failures, maximize worst-case outcomes

Related work

- Robust Dynamic Programming [e.g. Iyengar 2005]
  - ↳ Finite states only
- Quadratic costs (LQ) [e.g. Dean et al. 2017]
  - ↳ Stabilization only

Setting

$$\dot{x}(t) = A(\theta)x(t) + Bu(t) + D\omega(t) \tag{1}$$

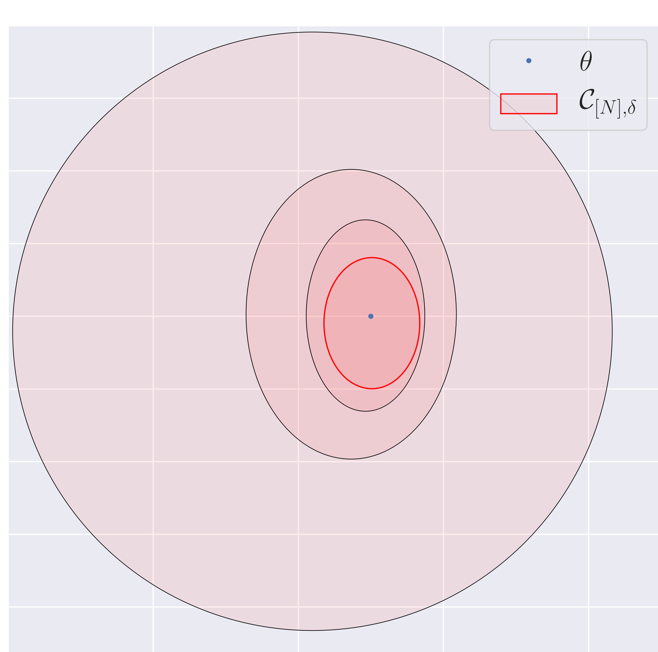
Model Estimation

$$\mathbb{P}\left(\theta \in \mathcal{C}_{N,\delta}\right) \geq 1 - \delta, \tag{2}$$

Robust control

$$\underbrace{\sup_{\mathbf{u}\in(\mathbb{R}^q)^{\mathbb{N}}} \inf_{\substack{\theta\in\mathcal{C}_{N,\delta}\\ \omega\in[\underline{\omega},\overline{\omega}]^{\mathbb{R}}}} \left[\sum_{n=N+1}^{\infty} \gamma^n R(x_n(\mathbf{u},\omega))\right]}_{V^r(\mathbf{u})} \tag{3}$$

Model Estimation



$C_{N,\delta}$ shrinks with the number of samples N.

Assumption 1 (Structure).

$$A(\theta) = A + \sum_{i=1}^{d} \theta_i \phi_i, \tag{4}$$

Assumption 2 (Noise Model). *Assume*

- *sub-Gaussian observation:* $\mathbb{E}\left[\exp\left(u^\intercal \eta\right)\right] \leq \exp\left(\frac{1}{2}u^\intercal \Sigma_p u\right)$
- *bounded disturbance:* $\underline{\omega}(t) \leq \omega(t) \leq \overline{\omega}(t)$

Theorem 1 (Matricial version of Abbasi-Yadkori et al. 2011).

Let

$$\theta_{N,\lambda} = G_{N,\lambda}^{-1} \sum_{n=1}^{N} \Phi_n^\intercal \Sigma_p^{-1} y_n,$$

$$G_{N,\lambda} = \sum_{n=1}^{N} \Phi_n^\intercal \Sigma_p^{-1} \Phi_n + \lambda I_d \in \mathbb{R}^{d\times d}.$$

Then, with probability at least $1-\delta$

$$\|\theta_{N,\lambda} - \theta\|_{G_{N,\lambda}} \leq \beta_N(\delta), \tag{5}$$

with $\beta_N(\delta) \overset{def}{=} \sqrt{2\ln\left(\frac{\det(G_{N,\lambda})^{1/2}}{\delta \det(\lambda I_d)^{1/2}}\right)} + (\lambda d)^{1/2} S.$

Interval Prediction

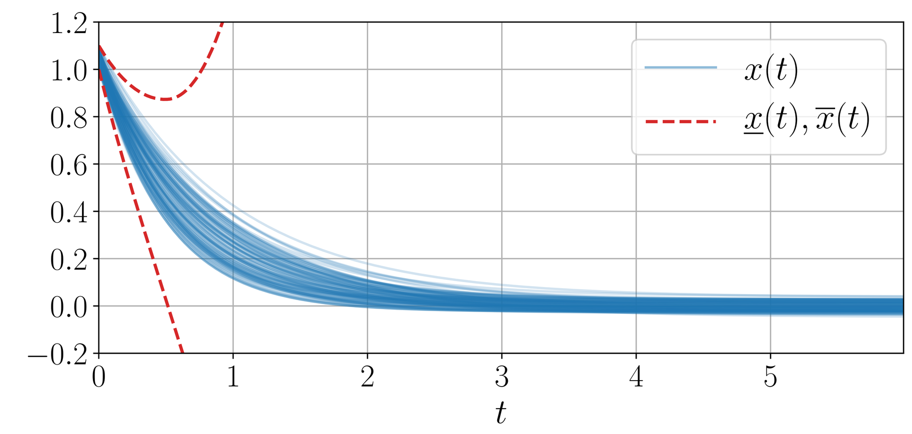
Having observed N samples, given $\theta \in \mathcal{C}_{N,\delta}$, we want

$$\underline{x}(t) \leq x(t) \leq \overline{x}(t), \qquad \forall t \geq t_N. \tag{6}$$

Proposition (Simple predictor of Efimov et al. 2012).

$$\dot{\underline{x}}(t) = \underline{A}^+\underline{x}^+(t) - \overline{A}^+\underline{x}^-(t) - \underline{A}^-\overline{x}^+(t) + \overline{A}^-\overline{x}^-(t) + Bu(t) + D^+\underline{\omega}(t) - D^-\overline{\omega}(t),$$
$$\dot{\overline{x}}(t) = \overline{A}^+\overline{x}^+(t) - \underline{A}^+\overline{x}^-(t) - \overline{A}^-\underline{x}^+(t) + \underline{A}^-\underline{x}^-(t) + Bu(t) + D^+\overline{\omega}(t) - D^-\underline{\omega}(t),$$

ensures the inclusion property (6).

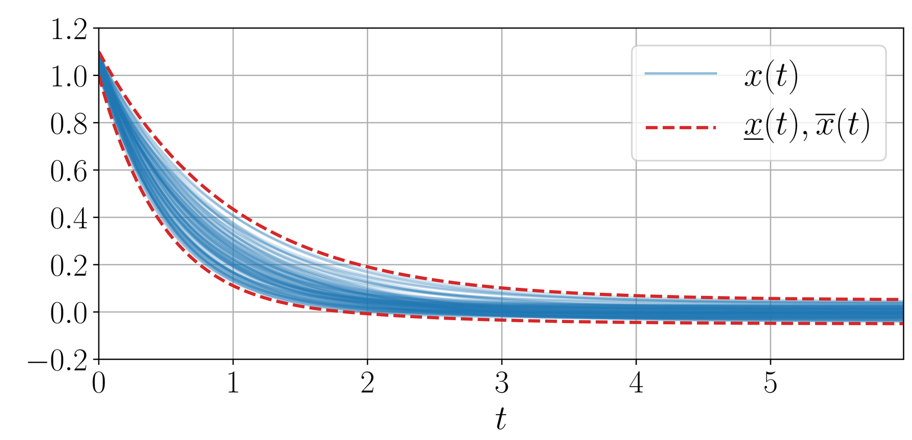


Assumption 3. *There exists orthogonal Z such that $Z^\intercal A_N Z$ is Metzler.*

Proposition (Enhanced predictor of Leurent et al. 2019).

$$\begin{aligned}\dot{\underline{x}}(t) &= A_N\underline{x}(t) - \Delta A_+\underline{x}^-(t) - \Delta A_-\overline{x}^+(t) + Bu(t) + D^+\underline{\omega}(t) - D^-\overline{\omega}(t),\\ \dot{\overline{x}}(t) &= A_N\overline{x}(t) + \Delta A_+\overline{x}^+(t) + \Delta A_-\underline{x}^-(t) + Bu(t) + D^+\overline{\omega}(t) - D^-\underline{\omega}(t),\end{aligned}$$

ensures the inclusion property (6) *under Assumption 3.*



Robust Control

Definition 1 (Surrogate objective). *Let $\underline{x}, \overline{x}$ following* (6) *and*

$$\hat{V}^r(\mathbf{u}) \overset{def}{=} \sum_{n=N+1}^{\infty} \gamma^n \underline{R}_n(\mathbf{u}) \quad \textit{where} \quad \underline{R}_n(\mathbf{u}) \overset{def}{=} \min_{x\in[\underline{x}_n(\mathbf{u}),\overline{x}_n(\mathbf{u})]} R(x).$$

Theorem 2 (Lower bound).

$$\hat{V}^r(\mathbf{u}) \leq V^r(\mathbf{u})$$

Theorem 3 (Suboptimality bound). *Under two conditions:*

1. *a Lipschitz regularity assumption for the reward function R;*
2. *a stability condition: there exist $P > 0, Q_0 \in \mathbb{R}^{p\times p}$, $\rho > 0$, and $N_0 \in \mathbb{N}$ such that*

$$\forall N > N_0, \quad \begin{bmatrix} A_N^\intercal P + PA_N + Q_0 & P|D| \\ |D|^\intercal P & -\rho I_r \end{bmatrix} < 0;$$

we can bound the suboptimality with probability at least $1-\delta$, for a planning budget K, as:

$$V(a_\star) - \hat{V}^r(a_K) \leq \underbrace{\Delta_\omega}_{\textit{robustness to disturbances}} + \underbrace{\mathcal{O}\left(\frac{\beta_N(\delta)^2}{\lambda_{\min}(G_{N,\lambda})}\right)}_{\textit{estimation error}} + \underbrace{\mathcal{O}\left(K^{-\frac{\log 1/\gamma}{\log \kappa}}\right)}_{\textit{planning error}}.$$

Corollary 1 (Asymptotic near-optimality). *Under an additional persistent excitation (PE) assumption*

$$\exists \underline{\phi}, \overline{\phi} > 0 : \forall n \geq n_0, \quad \underline{\phi}^2 \leq \lambda_{\min}(\Phi_n^\intercal \Sigma_p^{-1} \Phi_n) \leq \overline{\phi}^2,$$

the stability condition 2. of Theorem 3 can be relaxed to its limit $A_N \to A(\theta)$ and

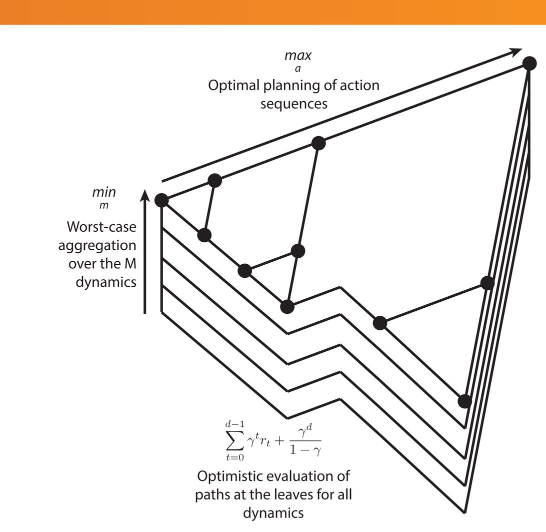
$$\mathcal{O}\left(\frac{\beta_N(\delta)^2}{\lambda_{\min}(G_{N,\lambda})}\right) = \mathcal{O}\left(\frac{\log\left(N^{d/2}/\delta\right)}{N}\right).$$

Multi-Model Extension

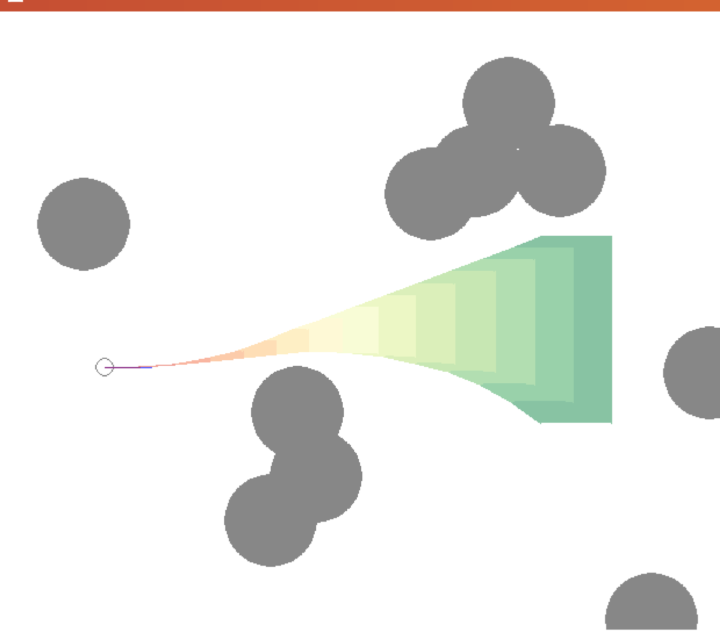
Assumption 4 (Multi-model ambiguity). *(A, ϕ) from* (4) *lies within a finite set of M models.*

Model adequacy If $y \notin \mathcal{P}^m$, the model (A_m, ϕ_m) can be confidently rejected.

Proposition 1 (Robust selection). *With discrete ambiguity, the robust version of* `OPD` *enjoys the same regret bound as* `OPD` *and recovers V^r exactly.*

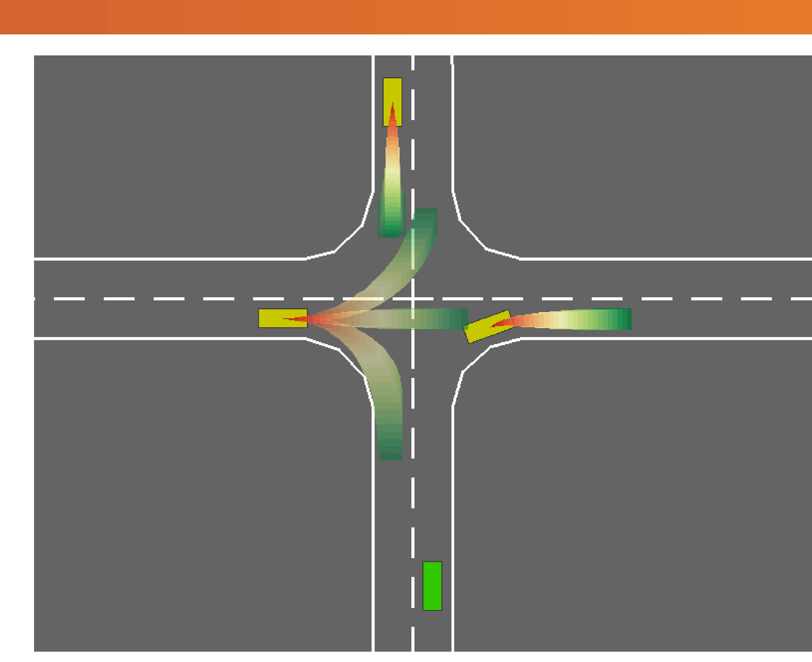


Experiments



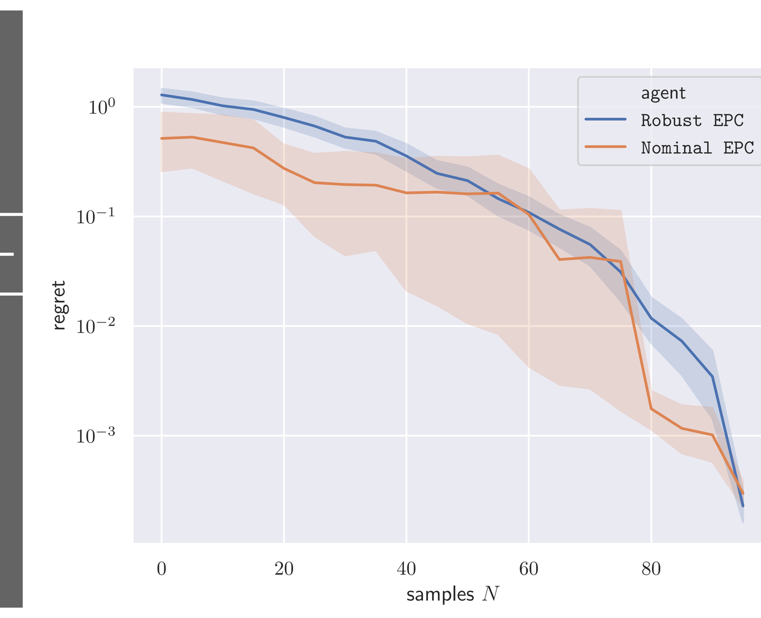
Obstacle avoidance

	failures	min	avg ± std
Oracle	0%	11.6	14.2 ± 1.3
Nominal	4%	2.8	$\mathbf{13.8} \pm 2.0$
Robust	**0%**	**10.4**	13.0 ± 1.5
DQN[a]	6%	1.7	12.3 ± 2.5



Unsignalized intersection

	failures	min	avg ± std
Oracle	0%	6.9	7.4 ± 0.5
Nominal 1	4%	5.2	$\mathbf{7.3} \pm 1.5$
Nominal 2	33%	3.5	6.4 ± 0.3
Robust	**0%**	**6.8**	7.1 ± 0.3
DQN[a]	3%	5.4	6.3 ± 0.6



Mean regret with respect to N

[a] After training on 3000 episodes

References

[1] Yasin Abbasi-Yadkori et al. "Improved Algorithms for Linear Stochastic Bandits". In: *Advances in Neural Information Processing Systems 24*. Ed. by J. Shawe-Taylor et al. Curran Associates, Inc., 2011, pp. 2312–2320.

[2] Sarah Dean et al. "On the Sample Complexity of the Linear Quadratic Regulator". In: *ArXiv* abs/1710.01688 (2017).

[3] D. Efimov et al. "Interval Estimation for LPV Systems Applying High Order Sliding Mode Techniques". In: *Automatica* 48 (2012), pp. 2365–2371.

[4] Garud N. Iyengar. "Robust Dynamic Programming". In: *Mathematics of Operations Research* 30 (2005), pp. 257–280.

[5] E. Leurent et al. "Interval Prediction for Continuous-Time Systems with Parametric Uncertainties". In: *Proc. IEEE Conference on Decision and Control (CDC)*. Nice, 2019.