Online Stochastic Optimization of Unknown Linear Dynamical Systems

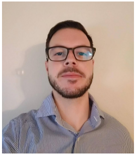
Emiliano Dall'Anese
University of Colorado Boulder

NREL RAES Workshop | Virtual | September 8, 2021

Acknowledgments

Co-authors:

G. Bianchin
CU Boulder



M. Vaquero
IE University

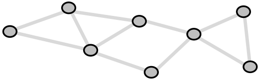
J. Cortés
UC San Diego

References: [Online] arXiv 2108.13040
arXiv 2103.16067 (to appear at IEEE CDC'21)

Acknowledgments: NSF Awards CMMI 2044946 and 2044900
NREL UGA-0-41026-1480

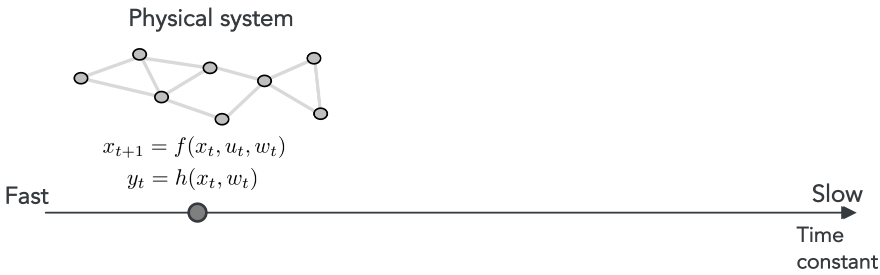
10,000-feet view

Physical system

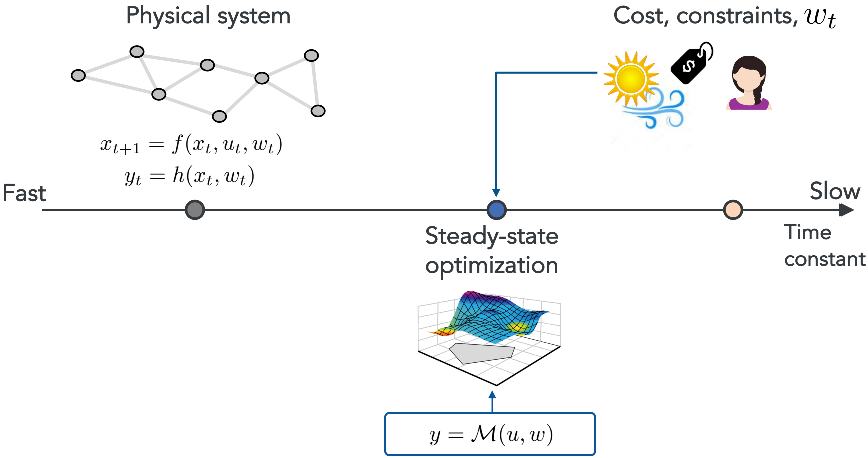


$$x_{t+1} = f(x_t, u_t, w_t)$$
$$y_t = h(x_t, w_t)$$

10,000-feet view

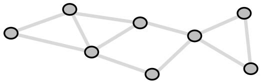


10,000-feet view



10,000-feet view

Physical system



$$x_{t+1} = f(x_t, u_t, w_t)$$
$$y_t = h(x_t, w_t)$$

Cost, constraints, w_t

Fast — Slow

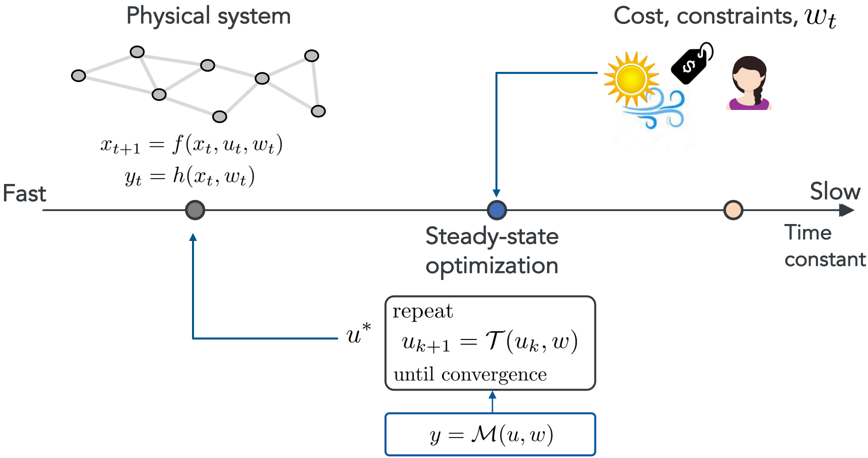
Time constant

Steady-state optimization

repeat
$$u_{k+1} = \mathcal{T}(u_k, w)$$
until convergence

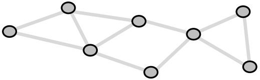
$$y = \mathcal{M}(u, w)$$

10,000-feet view



10,000-feet view

Physical system



$$x_{t+1} = f(x_t, u_t, w_t)$$
$$y_t = h(x_t, w_t)$$

Cost, constraints, w_t

Fast — ∞-fast — Steady-state optimization — Computation + sensing — Slow

Time constant

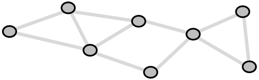
repeat
$$u_{k+1} = \mathcal{T}(u_k, w)$$
until convergence

$$y = \mathcal{M}(u, w)$$

w_t known

10,000-feet view

Physical system



$$x_{t+1} = f(x_t, u_t, w_t)$$
$$y_t = h(x_t, w_t)$$

Cost, constraints, w_t

Fast ——— Slow

Time constant

Steady-state optimization

repeat
$u_{k+1} = \mathcal{T}(u_k, w)$
until convergence

$y = \mathcal{M}(u, w)$

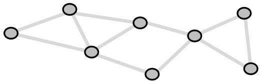
Plant ∞-fast

Computation

w_t known

10,000-feet view

Physical system



$$x_{t+1} = f(x_t, u_t, w_t)$$
$$y_t = h(x_t, w_t)$$

Cost, constraints, w_t

Fast —— Slow

Time constant

Steady-state optimization

repeat
$u_{k+1} = \mathcal{T}(u_k, w)$
until convergence

$y = \mathcal{M}(u, w)$

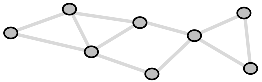
Plant ∞-fast

~~Computation~~

w_t ~~known~~

Feedback-based Optimization

Physical system



$$x_{t+1} = f(x_t, u_t, w_t)$$
$$y_t = h(x_t, w_t)$$

Cost, constraints, w_t

Fast — Slow

Time constant

Steady-state optimization

online algorithm

$$u_{t+1} = \mathcal{C}(u_t, y_t)$$

$$y = \mathcal{M}(u, w)$$

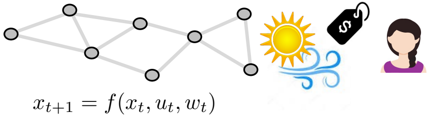
Plant ∞-fast

~~Computation~~

w_t ~~known~~

Feedback-based Optimization

Physical system   Cost, constraints, w_t



$$x_{t+1} = f(x_t, u_t, w_t)$$
$$y_t = h(x_t, w_t)$$

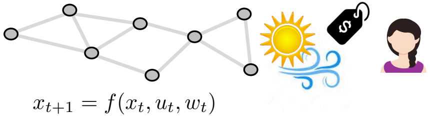
~~Plant ∞-fast~~

~~Computation~~

w_t ~~known~~

Online Optimization as Feedback Control

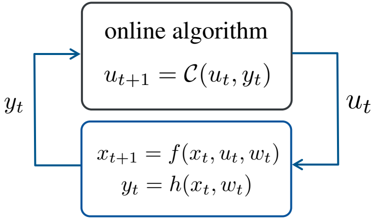
Physical system   Cost, constraints, w_t



$$x_{t+1} = f(x_t, u_t, w_t)$$
$$y_t = h(x_t, w_t)$$

Fast —— Optimization —— Slow

Time constant



~~Plant ∞-fast~~

~~Computation~~

w_t ~~known~~

A Sample of Works

System as **algebraic map**:

- [Bolognani et.al.'13] Static, applications to voltage control
- [Bernstein et al'14] Online gradient, applications to voltage control
- [Dall'Anese-Simonetto'16] Online primal-dual, application to voltage control

A Sample of Works

System as **algebraic map**:

- [Bolognani et.al.'13] Static, applications to voltage control
- [Bernstein et al'14] Online gradient, applications to voltage control
- [Dall'Anese-Simonetto'16] Online primal-dual, application to voltage control

Dynamical system model:

- [Jokic et al'09]: Static cost, KKT controller
- [Brunner et.al.'12]: Input affine systems, static cost, outp constraints
- [Colombino et al'18]: LTI systems, time-varying cost, exponential ISS
- [Lawrence et al'18]: LTI systems, static cost, joint stabilization + optimizat.
- [Menta et al'19]: LTI systems, static cost, power system application
- [Hauswirth et al'20]: Nonlinear sys., static cost, asymptotic stability
- [Bianchin et al'20]: Switched LTI, time-varying, E-ISS, hybrid Nesterov
- [Bianchin et al'21]: LTI systems, time-varying, projected primal-dual, E-ISS

A Sample of Works

System as **algebraic map**:

- [Bolognani et.al.'13] Static, applications to voltage control
- [Bernstein et al'14] Online gradient, applications to voltage control
- [Dall'Anese-Simonetto'16] Online primal-dual, application to voltage control

Dynamical system model:

- [Jokic et al'09]: Static cost, KKT controller
- [Brunner et.al.'12]: Input affine systems, static cost, outp constraints
- [Colombino et al'18]: LTI systems, time-varying cost, exponential ISS
- [Lawrence et al'18]: LTI systems, static cost, joint stabilization + optimizat.
- [Menta et al'19]: LTI systems, static cost, power system application
- [Hauswirth et al'20]: Nonlinear sys., static cost, asymptotic stability
- [Bianchin et al'20]: Switched LTI, time-varying, E-ISS, hybrid Nesterov
- [Bianchin et al'21]: LTI systems, time-varying, projected primal-dual, E-ISS

This work: Unknown LTI systems, unknown disturbances, data-driven control synthesis, exponential ISS

System Model

Linear plant with unknown (A, B, C, D, E)

$$x_{k+1} = Ax_k + Bu_k + Ew_k,$$
$$y_k = Cx_k + Dw_k,$$

$u_k \in \mathbb{R}^m \rightarrow$ Control input

$w_k \in \mathbb{R}^r \rightarrow$ Unknown, with distrib. $\mathcal{W}_k$

$y_k \in \mathbb{R}^p \rightarrow$ Output

System Model

Linear plant with unknown (A, B, C, D, E)

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + Ew_k, \\ y_k &= Cx_k + Dw_k, \end{aligned}$$

$u_k \in \mathbb{R}^m \rightarrow$ Control input

$w_k \in \mathbb{R}^r \rightarrow$ Unknown, with distrib. $\mathcal{W}_k$

$y_k \in \mathbb{R}^p \rightarrow$ Output

(As.) System: Controllable + Observable + A is Schur stable

System Model

Linear plant with unknown (A, B, C, D, E)

$$x_{k+1} = Ax_k + Bu_k + Ew_k,$$
$$y_k = Cx_k + Dw_k,$$

$u_k \in \mathbb{R}^m \to$ Control input
$w_k \in \mathbb{R}^r \to$ Unknown, with distrib. $\mathcal{W}_k$
$y_k \in \mathbb{R}^p \to$ Output

(As.) System: Controllable + Observable + A is Schur stable

$\Rightarrow$ Steady-state output: $y = \underbrace{C(I-A)^{-1}B}_{:=G}\, u + \underbrace{(D + C(I-A)^{-1}E)}_{:=H}\, w.$

Problem Formulation

Equilibrium-selection problem:

$$\begin{aligned} u_k^* \in \arg\min_{\bar{u},\bar{y}} \quad & \mathbb{E}_{w_k \sim \mathcal{W}_k} \left[\phi(\bar{u}, \bar{y})\right] \\ \text{s.t.} \quad & \bar{y} = G\bar{u} + H w_k \end{aligned}$$

Problem Formulation

Equilibrium-selection problem:

$$u_k^* \in \arg\min_{\bar{u}} \quad \mathbb{E}_{w_k \sim \mathcal{W}_k} \left[\phi(\bar{u}, G\bar{u} + Hw_k) \right]$$

Problem Formulation

Equilibrium-selection problem:

$$u_k^* \in \arg\min_{\bar{u}} \quad \mathbb{E}_{w_k \sim \mathcal{W}_k} \left[\phi(\bar{u}, G\bar{u} + Hw_k) \right]$$

Goal: Design a feedback controller $u_{k+1} = \mathcal{C}(u_k, y_k)$ to drive the input and output of the unknown system

$$x_{k+1} = Ax_k + Bu_k + Ew_k$$
$$y_k = Cx_k + Dw_k$$

to $\xi_k^* := (u_k^*, y_k^*)$, $y_k^* = Gu_k^* + Hw_k$.

Problem Formulation

Equilibrium-selection problem:

$$u_k^* \in \arg\min_{\bar{u}} \quad \mathbb{E}_{w_k \sim \mathcal{W}_k} \left[\phi(\bar{u}, G\bar{u} + Hw_k)\right]$$

Goal: Design a feedback controller $u_{k+1} = \mathcal{C}(u_k, y_k)$ to drive the input and output of the unknown system

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + Ew_k \\ y_k &= Cx_k + Dw_k \end{aligned}$$

to $\xi_k^* := (u_k^*, y_k^*)$, $y_k^* = Gu_k^* + Hw_k$.

Stochastic, steady-state regulation problem

Goal $\Rightarrow$ $\mathbb{E}[\|\xi_k - \xi_k^*\|] \leq \beta(k, \|\xi_0 - \xi_0^*\|) + \gamma(\sup\{\|\xi_k^* - \xi_{k-1}^*\|\})$

Assumptions

Problem

$$u_k^* \in \arg\min_{\bar{u}} \quad \mathbb{E}_{w_k \sim \mathcal{W}_k} \left[\phi(\bar{u}, G\bar{u} + Hw_k)\right]$$

Assumptions

Problem

$$u_k^* \in \arg\min_{\bar{u}} \quad \mathbb{E}_{w_k \sim \mathcal{W}_k} \left[\phi(\bar{u}, G\bar{u} + Hw_k) \right]$$

(As.)

(i) $u \mapsto \phi(u, y)$ is μ-strongly convex

(ii) $y \mapsto \phi(u, y)$ is ℓ-Lipschitz

(iii) $u \mapsto \nabla\phi(u, y)$ is ℓ_u^∇-Lipschitz, $y \mapsto \nabla\phi(u, y)$ is ℓ_y^∇-Lipschitz

Unknown System

Challenge: G is unknown.

$$u_k^* \in \arg\min_{\bar{u}} \quad \mathbb{E}_{w_k \sim \mathcal{W}_k} \left[\phi(\bar{u}, G\bar{u} + Hw_k) \right]$$

Unknown System

Challenge: G is unknown. Suppose $\hat{G}$ is an estimate of G; then

$$u_k^* \in \arg\min_{\bar{u}} \quad \mathbb{E}_{z_k \sim \mathcal{Z}_k(\bar{u})} \left[\phi(\bar{u}, \hat{G}\bar{u} + z_k) \right]$$

$z_k := (G - \hat{G})\bar{u} + H w_k$

Unknown System

Challenge: G is unknown. Suppose $\hat{G}$ is an estimate of G; then

$$u_k^* \in \arg\min_{\bar{u}} \quad \mathbb{E}_{z_k \sim \mathcal{Z}_k(\bar{u})} \left[\phi(\bar{u}, \hat{G}\bar{u} + z_k) \right]$$

$z_k := (G - \hat{G})\bar{u} + Hw_k$

Challenges:

(C1) Distribution of $\bar{z}$ is parametrized by the $\bar{u}$ [Drusvyatskiy-Xiao'20]

(C2) $\mathcal{Z}_k(\bar{u})$ is unknown; how to design the controller?

(C3) How to obtain $\hat{G}$ (without estimating A, B, and C)?

Optimization with Decision-Dependent Distributions

(C1) Distribution of $\bar{z}$ is parametrized by $\bar{u}$

Optimization with Decision-Dependent Distributions

(C1) Distribution of $\bar{z}$ is parametrized by $\bar{u}$

Definition [Perdomo et al'20]. u_k^{so} is a stable optimizer at time k if:

$$u_k^{\text{so}} = \arg\min_{\bar{u}} \quad \mathbb{E}_{\mathcal{Z}_k(u_k^{\text{so}})} \left[\phi(\bar{u}, \hat{G}\bar{u} + \bar{z})\right]$$

Moreover, let $x_k^{\text{so}} := (I - A)^{-1}(Bu_k^{\text{so}} + Ew_k)$

Optimization with Decision-Dependent Distributions

(C1) Distribution of $\bar{z}$ is parametrized by $\bar{u}$

Definition [Perdomo et al'20]. u_k^{so} is a stable optimizer at time k if:

$$u_k^{\text{so}} = \arg\min_{\bar{u}} \quad \mathbb{E}_{\mathcal{Z}_k(u_k^{\text{so}})}\left[\phi(\bar{u}, \hat{G}\bar{u} + \bar{z})\right]$$

Moreover, let $x_k^{\text{so}} := (I - A)^{-1}(Bu_k^{\text{so}} + Ew_k)$

Stable optimizers are optimal with respect to
the distribution they induce

Optimization with Decision-Dependent Distributions

(C1) Distribution of $\bar{z}$ is parametrized by $\bar{u}$

Definition [Perdomo et al'20]. u_k^{so} is a stable optimizer at time k if:

$$u_k^{\text{so}} = \arg\min_{\bar{u}} \quad \mathbb{E}_{\mathcal{Z}_k(u_k^{\text{so}})}\left[\phi(\bar{u}, \hat{G}\bar{u} + \bar{z})\right]$$

Moreover, let $x_k^{\text{so}} := (I - A)^{-1}(Bu_k^{\text{so}} + Ew_k)$

Stable optimizers are optimal with respect to the distribution they induce

Prop. If $y \mapsto \phi(u, y)$ in Lipschitz and $u \mapsto \phi(u, y)$ is μ-strongly convex, then

$$\|u_k^* - u_k^{\text{so}}\| \leq \frac{2\ell\|G - \hat{G}\|}{\mu\sigma^2_{\text{min}}(\hat{G})}$$

where $\sigma^2_{\text{min}}(\hat{G})$ is smallest singular value of $\hat{G}$

Synthesizing Controllers with Unknown Disturbance

(C2) $\mathcal{Z}_k(\bar{u})$ is unknown; how to design the controller?

If $\mathcal{Z}_k(\bar{u})$ is known: open-loop gradient-based controller:

$$u_{k+1} = u_k - \eta \mathbb{E}_{\bar{z} \sim \mathcal{Z}_k(\bar{u})} [\nabla_u \phi(u_k, \hat{G} u_k + \bar{z}) + \hat{G}^\top \nabla_y \phi(u_k, \hat{G} u_k + \bar{z})]$$

Synthesizing Controllers with Unknown Disturbance

(C2) $\mathcal{Z}_k(\bar{u})$ is unknown; how to design the controller?

If $\mathcal{Z}_k(\bar{u})$ is known: open-loop gradient-based controller:

$$u_{k+1} = u_k - \eta \mathbb{E}_{\bar{z} \sim \mathcal{Z}_k(\bar{u})}[\nabla_u \phi(u_k, \hat{G}u_k + \bar{z}) + \hat{G}^\top \nabla_y \phi(u_k, \hat{G}u_k + \bar{z})]$$

Idea: just close the loop!

$$u_{k+1} = u_k - \eta(\nabla_u \phi(u_k, y_k) + \hat{G}^\top \nabla_y \phi(u_k, y_k))$$
$$x_{k+1} = Ax_k + Bu_k + Ew_k, \quad y_k = Cx_k + Dw_k$$

Synthesizing Controllers with Unknown Disturbance

(C2) $\mathcal{Z}_k(\bar{u})$ is unknown; how to design the controller?

If $\mathcal{Z}_k(\bar{u})$ is known: open-loop gradient-based controller:

$$u_{k+1} = u_k - \eta \mathbb{E}_{\bar{z} \sim \mathcal{Z}_k(\bar{u})} [\nabla_u \phi(u_k, \hat{G} u_k + \bar{z}) + \hat{G}^\top \nabla_y \phi(u_k, \hat{G} u_k + \bar{z})]$$

Idea: just close the loop!

$$u_{k+1} = u_k - \eta(\nabla_u \phi(u_k, y_k) + \hat{G}^\top \nabla_y \phi(u_k, y_k))$$
$$x_{k+1} = A x_k + B u_k + E w_k, \quad y_k = C x_k + D w_k$$

Key features:

- Map $G u_k + H w_k$ replaced by y_k → **Online + feedback**
- Expectation replaced by one sample → **Stochastic gradient**

Data-Driven Transfer Function

(C3) How to obtain $\hat{G}$ (without estimating A, B, and C)?

Data-Driven Transfer Function

(C3) How to obtain $\hat{G}$ (without estimating A, B, and C)?

- For signal $z_{[k,k+T]} = (z_k, \dots, z_{k+T})$, $z_k \in \mathbb{R}^\sigma$, define the Hankel matrix:

$$Z_{t,q} = \begin{bmatrix} z_0 & z_1 & \cdots & z_{q-1} \\ z_1 & z_2 & \cdots & z_q \\ \vdots & \vdots & \ddots & \vdots \\ z_{t-1} & z_t & \cdots & z_{q+t-2} \end{bmatrix} \in \mathbb{R}^{\sigma t \times q}$$

Data-Driven Transfer Function

(C3) How to obtain $\hat{G}$ (without estimating A, B, and C)?

- For signal $z_{[k,k+T]} = (z_k, \ldots, z_{k+T})$, $z_k \in \mathbb{R}^\sigma$, define the Hankel matrix:

$$Z_{t,q} = \begin{bmatrix} z_0 & z_1 & \cdots & z_{q-1} \\ z_1 & z_2 & \cdots & z_q \\ \vdots & \vdots & \ddots & \vdots \\ z_{t-1} & z_t & \cdots & z_{q+t-2} \end{bmatrix} \in \mathbb{R}^{\sigma t \times q}$$

Def. [Willems et al'05] The signal $z_{[0,T-1]}$, $z_k \in \mathbb{R}^\sigma, \forall k \in \{0, \ldots, T-1\}$ is persistently exciting (or sufficiently rich) of order t if $Z_{t,q}$ has full row rank σt.

Data-Driven Transfer Function

(C3) How to obtain $\hat{G}$ (without estimating A, B, and C)?

- Assume availability of set of sample data $y_{[0,T]}$, $u_{[0,T]}$, $w_{[0,T]}$

Data-Driven Transfer Function

(C3) How to obtain $\hat{G}$ (without estimating A, B, and C)?

- Assume availability of set of sample data $y_{[0,T]}$, $u_{[0,T]}$, $w_{[0,T]}$

- Define
$$y^{\text{diff}} := (y_1 - y_0,\ y_2 - y_1, \ldots,\ y_T - y_{T-1}),$$
$$w^{\text{diff}} := (w_1 - w_0,\ w_2 - w_1, \ldots,\ w_T - w_{T-1}),$$

Data-Driven Transfer Function

(C3) How to obtain $\hat{G}$ (without estimating A, B, and C)?

- Assume availability of set of sample data $y_{[0,T]}$, $u_{[0,T]}$, $w_{[0,T]}$

- Define
$$y^{\text{diff}} := (y_1 - y_0,\ y_2 - y_1, \ldots,\ y_T - y_{T-1}),$$
$$w^{\text{diff}} := (w_1 - w_0,\ w_2 - w_1, \ldots,\ w_T - w_{T-1}),$$

Theorem. Suppose the inputs are PE of order $n + \nu$ ($\nu =$ obs. index). Let $q := T - \nu + 1$. Then:

(i) $\exists\ M \in \mathbb{R}^{q \times m\nu}$:
$$Y^{\text{diff}}_{\nu,q} M = 0, \qquad W^{\text{diff}}_{\nu,q} M = 0,$$
$$U_{\nu,q} M = \mathbb{1}_\nu \otimes I_m, \quad W_{\nu,q} M = 0,$$

(ii) $G = [Y_{\nu,q}]_i M$

Data-Driven Transfer Function (2)

(CH3) How to approximate $\hat{G}$, since it depends on (A, B, C)

If disturbance is unknown, $W_{\nu,q}^{\text{diff}} M = 0$ and $W_{\nu,q} M = 0$ may not hold

Data-Driven Transfer Function (2)

(CH3) How to approximate $\hat{G}$, since it depends on (A, B, C)

If disturbance is unknown, $W_{\nu,q}^{\text{diff}} M = 0$ and $W_{\nu,q} M = 0$ may not hold

Theorem Assume $u_{[0,T-1]}$ and $w_{[0,T-1]}$ are $(n+\nu)$-PE.

Let $\hat{M} \in \mathbb{R}^{q \times m\nu}$ be such that $Y_{\nu,q}^{\text{diff}} \hat{M} = 0$ and $U_{\nu,q} \hat{M} = \mathbb{1}_\nu \otimes I_m$.

Let $\hat{G} = [Y_{\nu,q}]_i \hat{M}$. Then

$$\hat{G} - G = \left(C(I-A)^{-1} E [W_{\nu,q}]_i + D[W_{\nu,q}]_i + C(I-A)^{-1} C^\dagger D [W_{\nu,q}^{\text{diff}}]_i \right) \hat{M}$$

Synthesizing Controllers with Unknown Disturbance

Interconnected system:

$$u_{k+1} = u_k - \eta(\nabla_u \phi(u_k, y_k) + \hat{G}^\top \nabla_y \phi(u_k, y_k))$$
$$x_{k+1} = Ax_k + Bu_k + Ew_k, \quad y_k = Cx_k + Dw_k$$

Synthesizing Controllers with Unknown Disturbance

Theorem. Let $\{\xi_k\}_{k\in\mathbb{Z}_+}$ be a sequence of the interconnected system. Under the current (As), the following holds for any $k \geq 1$:

$$\mathbb{E}[\|\xi_{k+1} - \xi_{k+1}^{\mathsf{SO}}\|] \leq \beta_1 \mathbb{E}\left[\|u_k - u_k^{\mathsf{SO}}\|\right] + \beta_2 \mathbb{E}\left[\|x_k - x_k^{\mathsf{SO}}\|\right] + \gamma_1 \mathbb{E}[\|e_k\|] + \gamma_2 \|u_{k+1}^{\mathsf{SO}} - u_k^{\mathsf{SO}}\| + \gamma_3 \mathbb{E}[\sup_{t\geq 0} \|x_{t+1}^{\mathsf{SO}} - x_t^{\mathsf{SO}}\|]$$

$$\beta_1 = \sqrt{1-\eta\mu} + \eta\hat{\ell}^{\nabla}\|G - \hat{G}\| \qquad \beta_2 = \sqrt{\frac{\bar{\lambda}(P)}{\underline{\lambda}(P)}\left(1 - (1-\kappa)\frac{\underline{\lambda}(Q)}{\bar{\lambda}(P)}\right)} + \eta\hat{\ell}^{\nabla}\|C\|$$

$$\gamma_1 = \eta, \quad \gamma_2 = 1, \quad \gamma_3 = \max\{\sqrt{\frac{2\bar{\lambda}(P)}{\kappa\underline{\lambda}(Q)}}, \frac{4\|A^\top P\|}{\kappa\underline{\lambda}(Q)}\}$$

Synthesizing Controllers with Unknown Disturbance

Theorem. Let $\{\xi_k\}_{k\in\mathbb{Z}_+}$ be a sequence of the interconnected system. Under the current (As), the following holds for any $k \geq 1$:

$$\mathbb{E}[\|\xi_{k+1} - \xi_{k+1}^{\text{SO}}\|] \leq \beta_1 \mathbb{E}\left[\|u_k - u_k^{\text{SO}}\|\right] + \beta_2 \mathbb{E}\left[\|x_k - x_k^{\text{SO}}\|\right] + \gamma_1 \mathbb{E}[\|e_k\|] + \gamma_2 \|u_{k+1}^{\text{SO}} - u_k^{\text{SO}}\| + \gamma_3 \mathbb{E}[\sup_{t\geq 0} \|x_{t+1}^{\text{SO}} - x_t^{\text{SO}}\|]$$

$$\beta_1 = \sqrt{1-\eta\mu} + \eta\hat{\ell}^{\nabla}\|G - \hat{G}\| \qquad \beta_2 = \sqrt{\frac{\bar{\lambda}(P)}{\underline{\lambda}(P)}\left(1 - (1-\kappa)\frac{\underline{\lambda}(Q)}{\bar{\lambda}(P)}\right)} + \eta\hat{\ell}^{\nabla}\|C\|$$

$$\gamma_1 = \eta, \quad \gamma_2 = 1, \quad \gamma_3 = \max\{\sqrt{\frac{2\bar{\lambda}(P)}{\kappa\underline{\lambda}(Q)}}, \frac{4\|A^\top P\|}{\kappa\underline{\lambda}(Q)}\}$$

Corollary: $\exists\ \eta$ such that $\beta_1, \beta_2 \in (0,1)$ and $\exists\ \sigma_e > 0, \sigma_x > 0$:

$$\mathbb{E}[\|\xi_k - \xi_k^{\text{SO}}\|] \leq c^k \|\xi_0 - \xi_0^{\text{SO}}\| + \sigma_e \sup_{t\geq 0} \mathbb{E}\|e_t\|] + \sigma_x \mathbb{E}[\sup_{t\geq 0} \|\xi_{t+1}^{\text{SO}} - \xi_t^{\text{SO}}\|]$$

Building Blocks of the Proof

Bounding controller error:

- Let $\mathcal{C}_\theta(u) := u - \eta \mathbb{E}_{z \sim \mathcal{Z}(\theta)} \left[\nabla \Phi(u, z) \right]$ (true-gradient update)
- Let $\hat{\mathcal{C}}(u, y) := u - \eta \nabla \Phi(u, y)$ (sample-gradient update)

Building Blocks of the Proof

Bounding controller error:

- Let $\mathcal{C}_\theta(u) := u - \eta \mathbb{E}_{z \sim \mathcal{Z}(\theta)} \left[\nabla \Phi(u, z)\right]$ (true-gradient update)
- Let $\hat{\mathcal{C}}(u, y) := u - \eta \nabla \Phi(u, y)$ (sample-gradient update)

$$\mathbb{E}[\|u_{k+1} - u_k^{\mathsf{so}}\|] \leq \underbrace{\mathbb{E}[\|\hat{\mathcal{C}}(u_k, \tilde{x}) - \mathcal{C}_{u_k}(u_k, \tilde{x})\|]}_{\text{Gradient Error}} + \underbrace{\|\mathcal{C}_{u_k}(u_k, \tilde{x}) - \mathcal{C}_{u_k^{\mathsf{so}}}(u_k, \tilde{x}_k)\|}_{\text{Distributional Shift Error}}$$

$$+ \underbrace{\|\mathcal{C}_{u_k^{\mathsf{so}}}(u_k, \tilde{x}_k) - \mathcal{C}_{u_k^{\mathsf{so}}}(u_k, 0)\|}_{\text{System Dynamics Error}} + \underbrace{\|\mathcal{C}_{u_k^{\mathsf{so}}}(u_k, 0) - \mathcal{C}_{u_k^{\mathsf{so}}}(u_k^{\mathsf{so}}, 0)\|}_{\text{Contractivity}}$$

Building Blocks of the Proof

Bounding controller error:

- Let $\mathcal{C}_\theta(u) := u - \eta \mathbb{E}_{z \sim \mathcal{Z}(\theta)} [\nabla \Phi(u, z)]$ (true-gradient update)
- Let $\hat{\mathcal{C}}(u, y) := u - \eta \nabla \Phi(u, y)$ (sample-gradient update)

$$\mathbb{E}[\|u_{k+1} - u_k^{\text{so}}\|] \leq \underbrace{\mathbb{E}[\|\hat{\mathcal{C}}(u_k, \tilde{x}) - \mathcal{C}_{u_k}(u_k, \tilde{x})\|]}_{\text{Gradient Error}} + \underbrace{\|\mathcal{C}_{u_k}(u_k, \tilde{x}) - \mathcal{C}_{u_k^{\text{so}}}(u_k, \tilde{x}_k)\|}_{\leq c_1 \|u_k - u_k^{\text{so}}\| \text{ ``Calmness''}}$$

$$+ \underbrace{\|\mathcal{C}_{u_k^{\text{so}}}(u_k, \tilde{x}_k) - \mathcal{C}_{u_k^{\text{so}}}(u_k, 0)\|}_{\leq c_2 \|\tilde{x}_k\| \text{ ``Ease''}} + \underbrace{\|\mathcal{C}_{u_k^{\text{so}}}(u_k, 0) - \mathcal{C}_{u_k^{\text{so}}}(u_k^{\text{so}}, 0)\|}_{\leq c_3 \|u_k - u_k^{\text{so}}\| \text{ ``Contraction''}}$$

Building Blocks of the Proof

Bounding controller error:

- Let $\mathcal{C}_\theta(u) := u - \eta \mathbb{E}_{z \sim \mathcal{Z}(\theta)} [\nabla \Phi(u, z)]$ (true-gradient update)
- Let $\hat{\mathcal{C}}(u, y) := u - \eta \nabla \Phi(u, y)$ (sample-gradient update)

$$\mathbb{E}[\|u_{k+1} - u_k^{\mathsf{SO}}\|] \leq \underbrace{\mathbb{E}[\|\hat{\mathcal{C}}(u_k, \tilde{x}) - \mathcal{C}_{u_k}(u_k, \tilde{x})\|]}_{\text{Gradient Error}} + \underbrace{\|\mathcal{C}_{u_k}(u_k, \tilde{x}) - \mathcal{C}_{u_k^{\mathsf{SO}}}(u_k, \tilde{x}_k)\|}_{\leq c_1 \|u_k - u_k^{\mathsf{SO}}\| \text{ "Calmness"}}$$

$$+ \underbrace{\|\mathcal{C}_{u_k^{\mathsf{SO}}}(u_k, \tilde{x}_k) - \mathcal{C}_{u_k^{\mathsf{SO}}}(u_k, 0)\|}_{\leq c_2 \|\tilde{x}_k\| \text{ "Ease"}} + \underbrace{\|\mathcal{C}_{u_k^{\mathsf{SO}}}(u_k, 0) - \mathcal{C}_{u_k^{\mathsf{SO}}}(u_k^{\mathsf{SO}}, 0)\|}_{\leq c_3 \|u_k - u_k^{\mathsf{SO}}\| \text{ "Contraction"}}$$

Bounding plant error (a.s.):

$$\|\tilde{x}_{k+1}\| \leq \underbrace{c_4 \|\tilde{x}_k\|}_{\text{"Transient Effect"}} + \underbrace{c_5 \sup\nolimits_{t \geq 0} \|x_{k+1}^{\mathsf{SO}} - x_k^{\mathsf{SO}}\|}_{\text{"Equilibrium Shift"}}$$

Fleet of (electric) vehicles of a ride service provider

Application to Electric Ride Service

Fleet of (electric) vehicles of a ride service provider

- Elasticity of the demand: $d_k^{ij} = \delta_k^{ij}\left(1 - \theta^{ij}\frac{p_k^{ij}}{p_{\max}^{ij}}\right)$
- Idle-vehicle occupancy in region i at time k: $x_k^i \in \mathbb{R}_{\geq 0}$
- Travel time from i to j: $\sigma_k^{ij,\tau}$

Application to Electric Ride Service

Fleet of (electric) vehicles of a ride service provider

- Elasticity of the demand: $d_k^{ij} = \delta_k^{ij}\left(1 - \theta^{ij}\frac{p_k^{ij}}{p_{\max}^{ij}}\right)$
- Idle-vehicle occupancy in region i at time k: $x_k^i \in \mathbb{R}_{\geq 0}$
- Travel time from i to j: $\sigma_k^{ij,\tau}$

Dynamics: $x_{k+1}^i = x_k^i - \sum_{j\in\mathcal{V}} a_{ij}x_k^i + \sum_{j\in\mathcal{V}} a_{ji}x_k^j - \sum_{j\in\mathcal{V}} d_k^{ij}$

$$\underbrace{+\sum_{j\in\mathcal{V}}\sum_{\tau=k-T}^{k-1}\sigma_\tau^{ji,k-\tau}d_\tau^{ji} + m_k^i}_{=w_k^i}$$

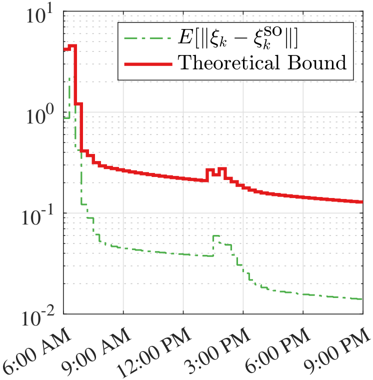
Application to Electric Ride Service

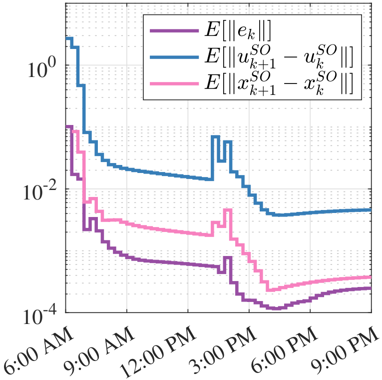
Fleet of (electric) vehicles of a ride service provider

Problem

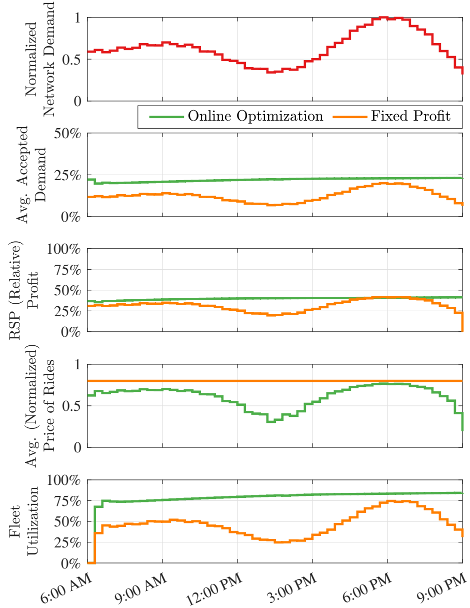
$$
\begin{aligned}
\min_{p,x,d} \quad & \sum_{i\in\mathcal{V}}\sum_{j\in\mathcal{V}} c^{ij}d^{ij} - p^{ij}d^{ij} + \varrho\|x - x_k^{\mathrm{ch}}\|^2, \\
\text{s.t.} \quad & 0 = -\sum_{j\in\mathcal{V}} a_{ij}x^i + \sum_{j\in\mathcal{V}} a_{ji}x^j - \sum_{j\in\mathcal{V}} d^{ij} + w_k^i, \\
& d^{ij} = \delta_k^{ij}\left(1 - \theta^{ij}p^{ij}/p^{ij}_{\text{max}}\right), \\
& d^{ij} \geq 0,\ x^i \geq 0, \quad \forall i, j, \in \mathcal{V},
\end{aligned}
$$

Boundedness of the Error





Application to Electric Ride Service



Conclusions

- Stochastic gradient as feedback controller for LTI systems
- Estimate of the steady-state transfer function using recorded data
- Estimates for the error evolution and E-ISS in expectation

Conclusions

- Stochastic gradient as feedback controller for LTI systems
- Estimate of the steady-state transfer function using recorded data
- Estimates for the error evolution and E-ISS in expectation
- Application to EV fleets for ride sharing
- Future app: electrified transportation and autonomous energy systems

Conclusions

- Stochastic gradient as feedback controller for LTI systems
- Estimate of the steady-state transfer function using recorded data
- Estimates for the error evolution and E-ISS in expectation
- Application to EV fleets for ride sharing
- Future app: electrified transportation and autonomous energy systems

Thanks!