ETH zürich

IEEE CDC2021
60th Conference on Decision and Control
December 13-15, 2021 | Austin, Texas, USA

Data-driven Control in Autonomous Energy Systems

Linbin Huang
Postdoctoral researcher, ETH Zürich

AUTOMATIC CONTROL LABORATORY

Acknowledgements

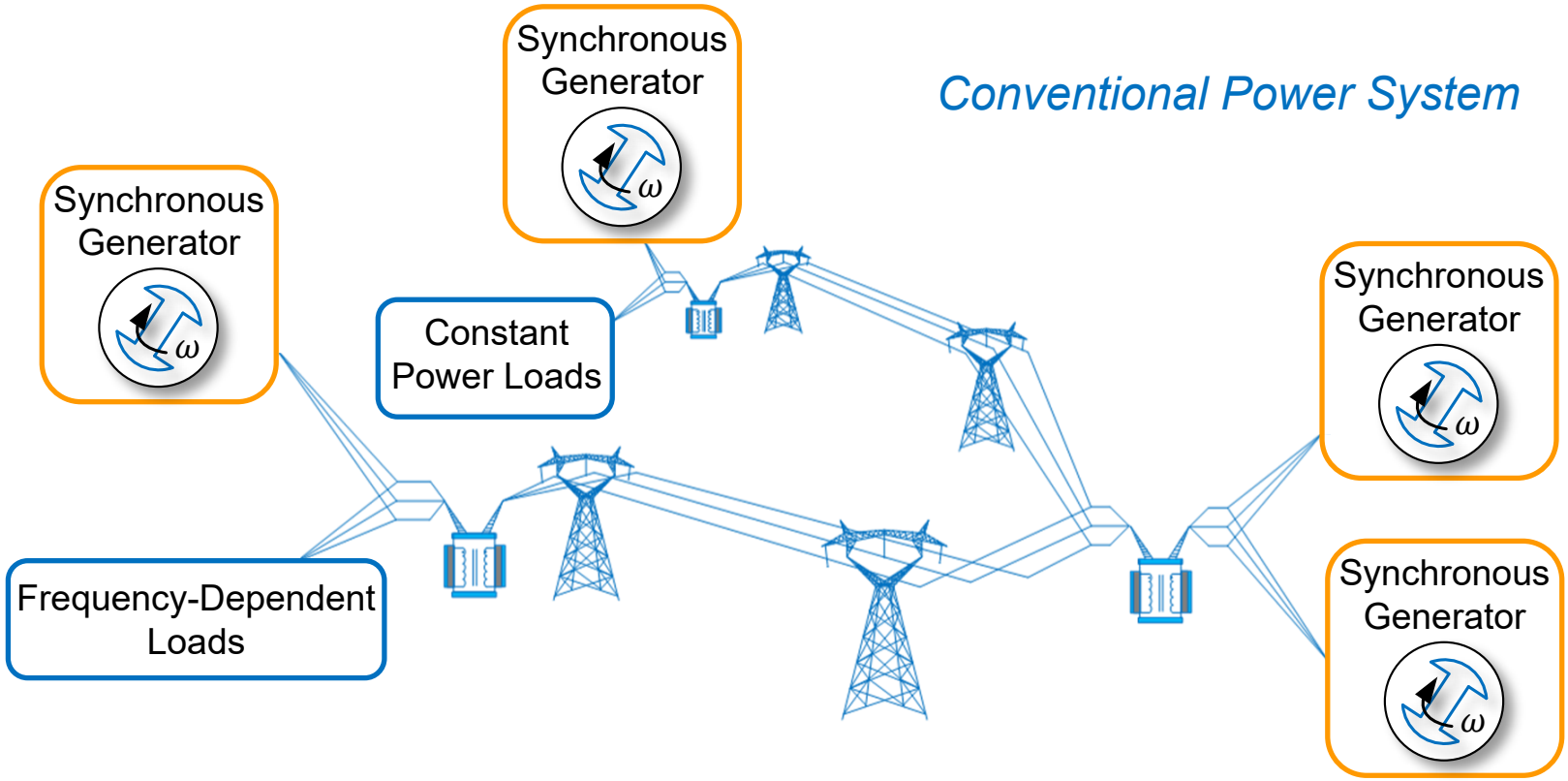
Trevor Zhen

Jeremy Coulson

John Lygeros

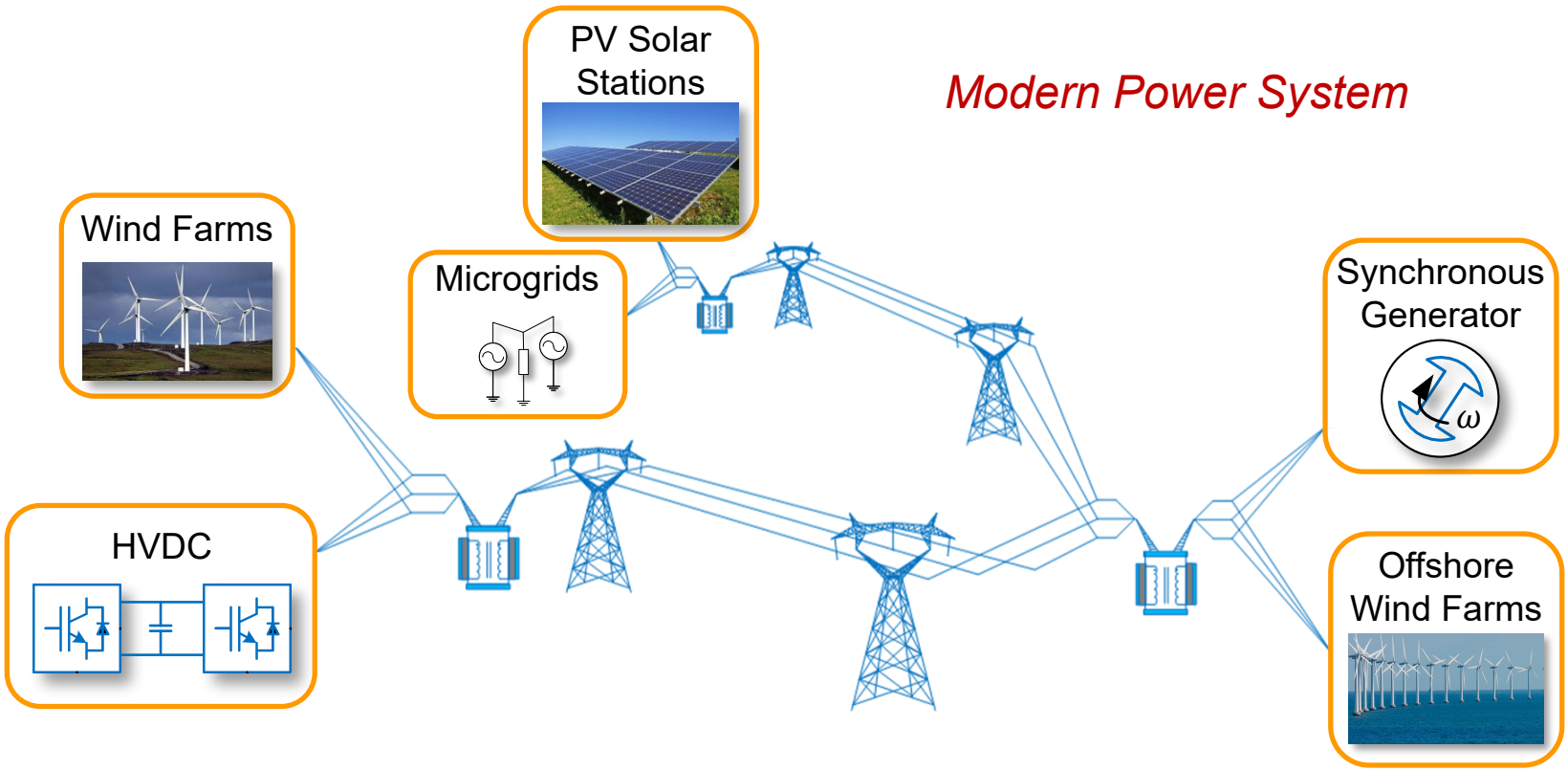
Florian Dörfler

Motivations



The main power source in conventional power system is ***synchronous generator***.

Motivations



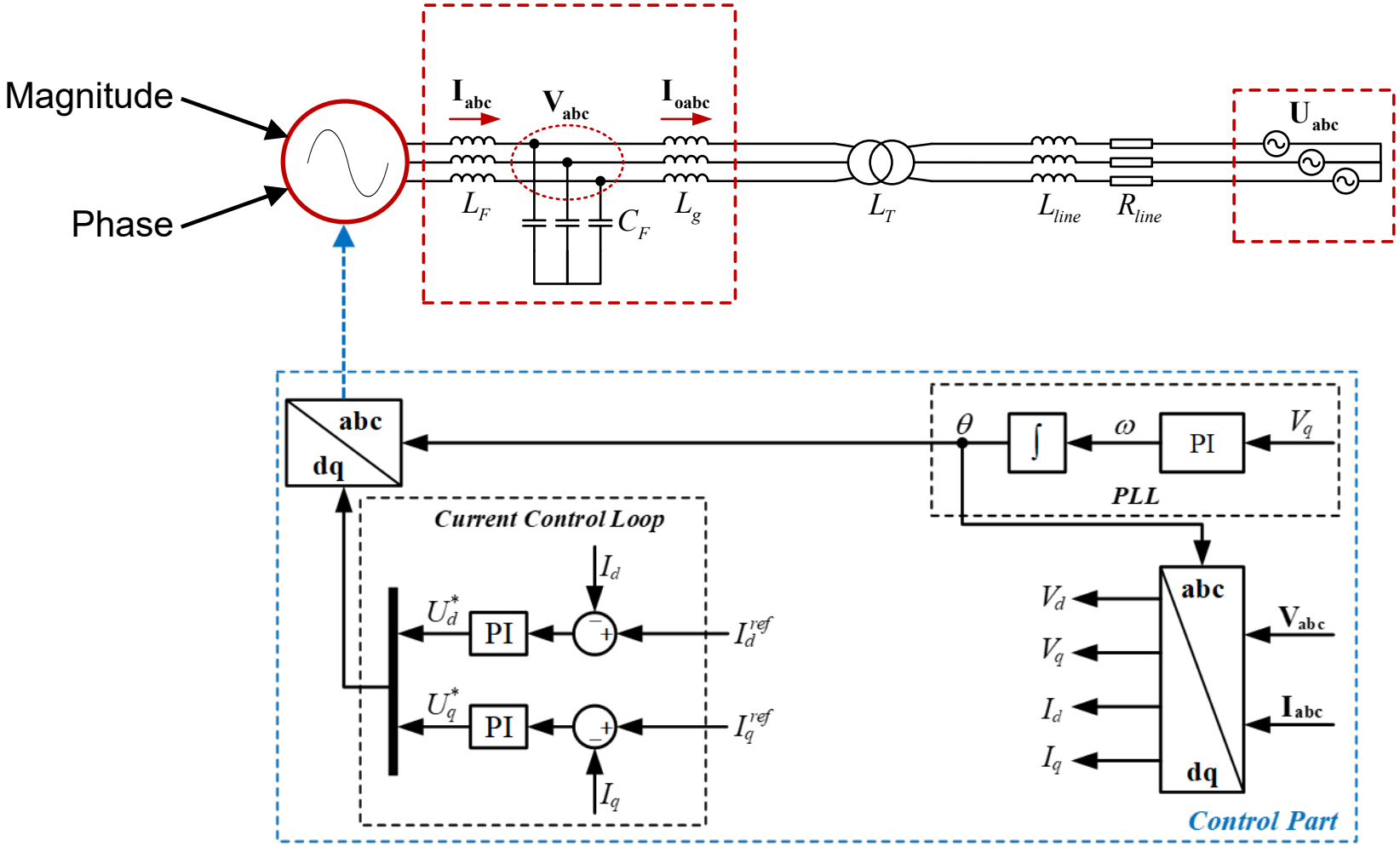
Modern power system contains different ***power-electronic-interfaced devices***, e.g., photovoltaics, wind farms, high-voltage direct current (HVDC) systems.

Motivations

The power system is becoming more and more complex,

also more controllable!

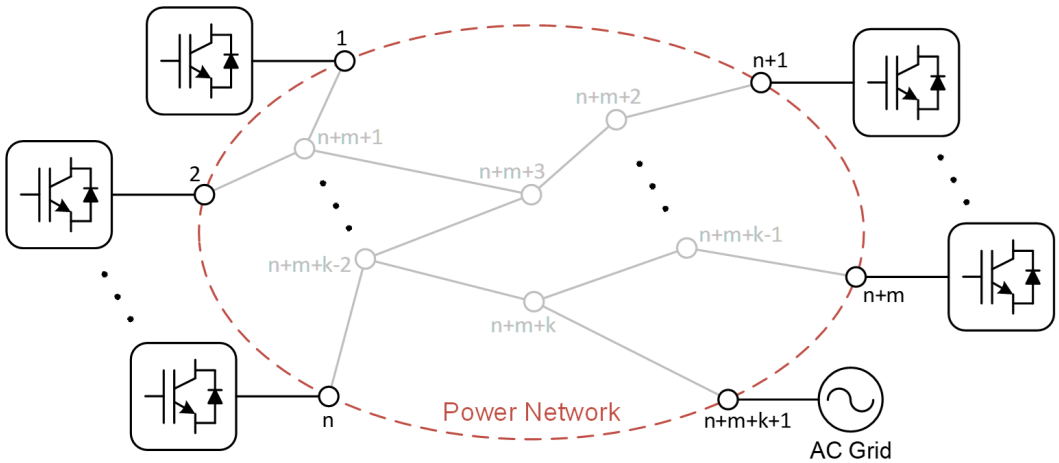
Motivations



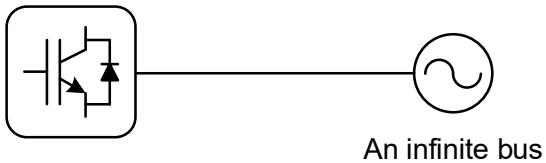
Motivations

- In synchronous generators, we cannot directly control the voltage angle.
- By comparison, we can directly control the voltage magnitude and angle of converters. ➜ We need to design the control scheme more carefully

Motivations



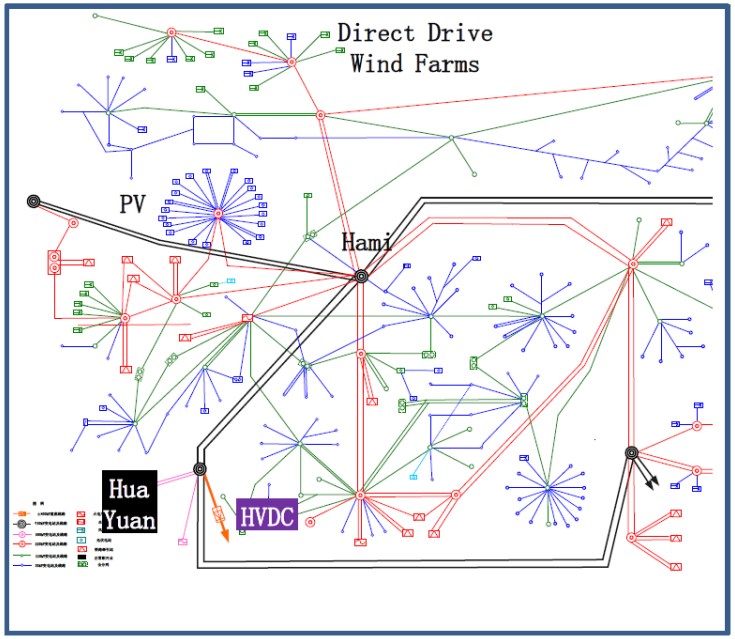
In practice, a power grid is ever-changing and converters can strongly interact with each other.



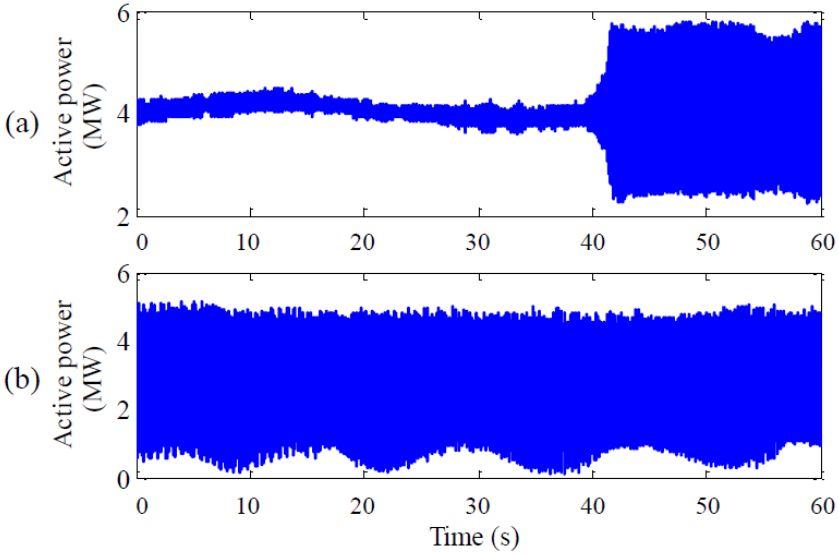
However, we generally assume that the converter is connected to an infinite bus in control design, which could be problematic.

Motivations

Sub-synchronous oscillation events in China



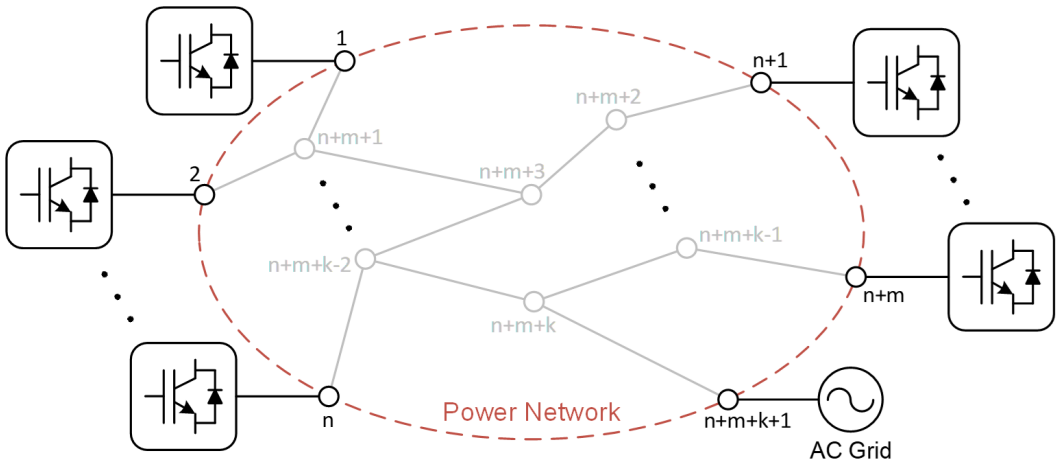
*Data provided by Prof. Xiaorong Xie



Measurements in a Wind Farm
(7/1/2015)

The wind turbines were well tested by the manufacturers before connected to the grid, but instabilities can still arise during online operation.

Motivations

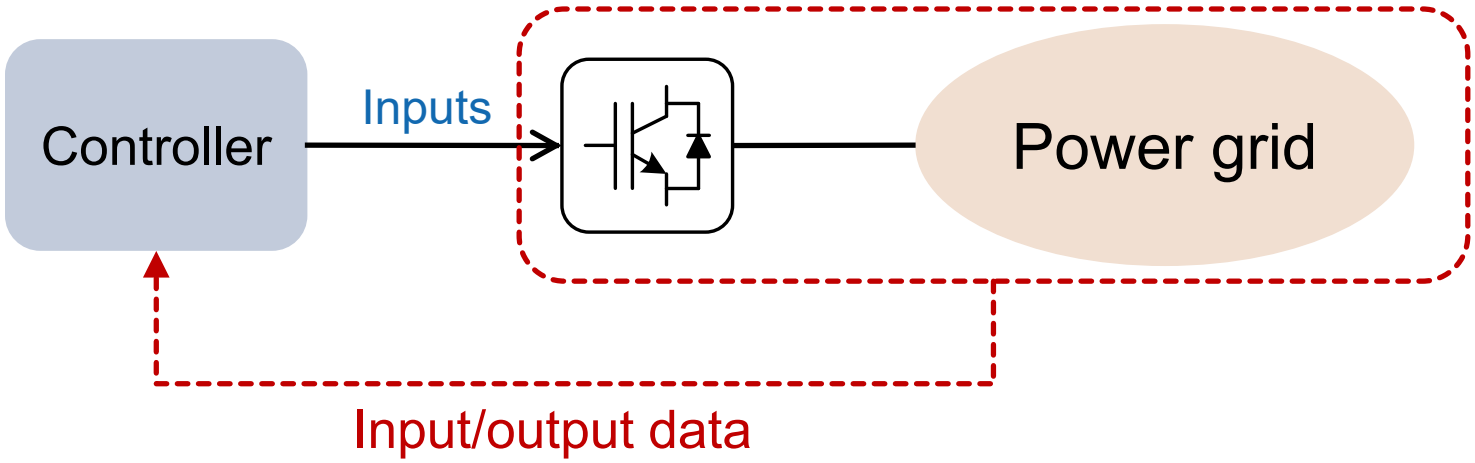


In practice, a power grid is ever-changing and converters can strongly interact with each other.

How to design a controller that can adapt to a complex grid?

Maybe it is a good idea to use input/output data based on the converter's "observations"!

Motivations



- Can we use data to enhance the controller design or to obtain an optimal controller?
- Can we provide robustness against noise in data?

The role of data

- Can we use input and output data to represent the system model?

The role of data

An n-order discrete-time LTI system (m inputs and p outputs):

$$\begin{cases} x_{t+1} = Ax_t + Bu_t \\ y_t = Cx_t + Du_t \end{cases} \qquad (1)$$

Input/output data: $u^{\mathrm{d}}, y^{\mathrm{d}}$ $\begin{cases} u^{\mathrm{d}} = \mathrm{col}(u_1, ..., u_i, ..., u_T) \\ y^{\mathrm{d}} = \mathrm{col}(y_1, ..., y_i, ..., y_T) \end{cases}$ $T \geq (m+1)(T_{\mathrm{ini}} + N + n) - 1$

Construct the Hankel matrix:

$$H = T - T_{\mathrm{ini}} - N + 1$$

$$\begin{array}{r} T_{\mathrm{ini}} \\ N \end{array} \left\{ \begin{bmatrix} \overbrace{U_{\mathrm{P}}} \\ U_{\mathrm{F}} \end{bmatrix} \right. := \mathscr{H}_{T_{\mathrm{ini}}+N}(u^{\mathrm{d}}), \quad \begin{bmatrix} Y_{\mathrm{P}} \\ Y_{\mathrm{F}} \end{bmatrix} := \mathscr{H}_{T_{\mathrm{ini}}+N}(y^{\mathrm{d}})$$

$$\mathscr{H}_L(x) := \begin{bmatrix} x_1 & x_2 & \cdots & x_{T-L+1} \\ x_2 & x_3 & \cdots & x_{T-L+2} \\ \vdots & \vdots & \ddots & \vdots \\ x_L & x_{L+1} & \cdots & x_T \end{bmatrix}$$

x is persistently exciting of order L if $\mathscr{H}_L(x)$ is of full row rank.

Fundamental Lemma

An n-order discrete-time LTI system (m inputs and p outputs):

$$\begin{cases} x_{t+1} = Ax_t + Bu_t \\ y_t = Cx_t + Du_t \end{cases}$$

The *lag* of this system is defined by the smallest integer ℓ so that the observability matrix $\mathscr{O}_\ell(A, C) := \mathrm{col}(C, CA, ..., CA^{\ell-1})$ has rank n.

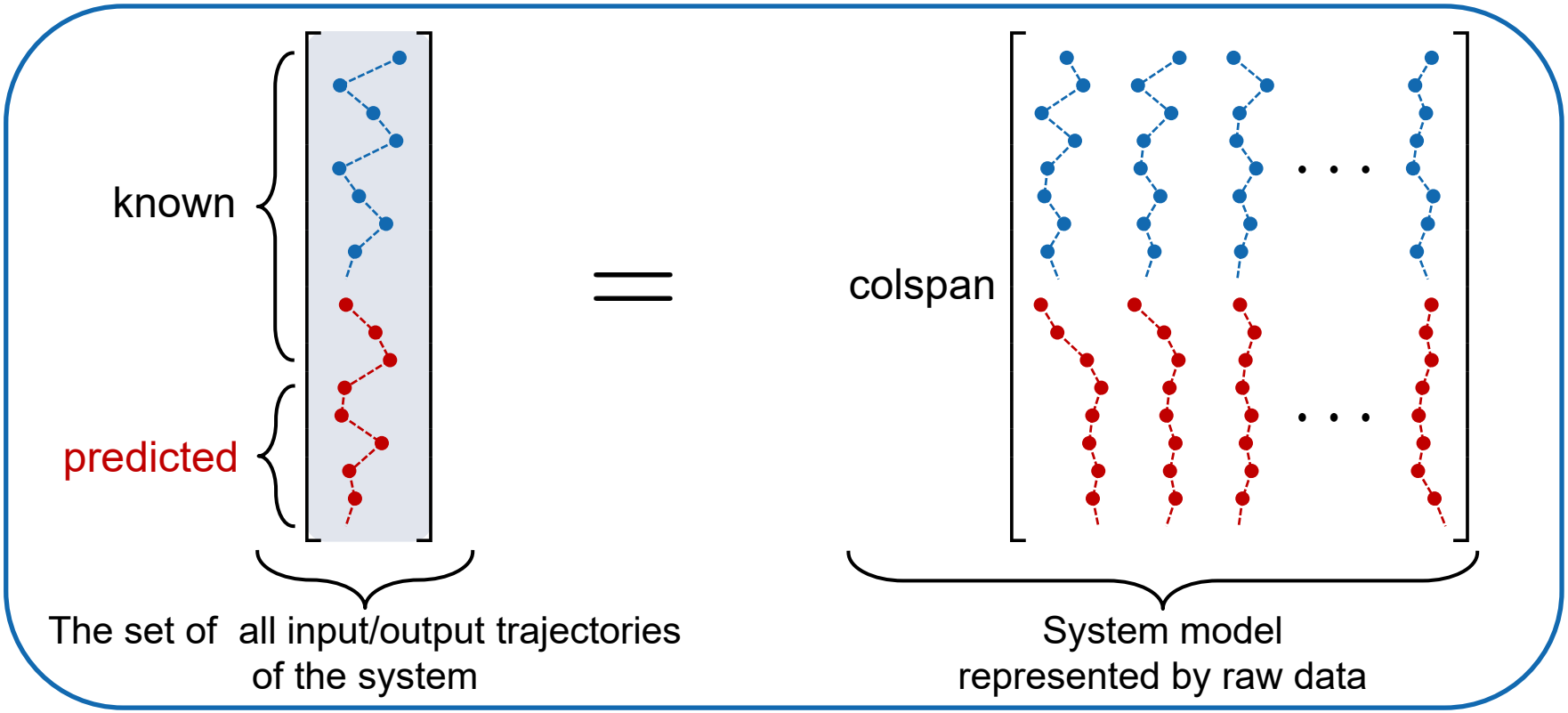
If u^{d} is persistently exciting of order $T_{\mathrm{ini}} + N + n$, then

a trajectory $\begin{bmatrix} u_{\mathrm{ini}} \\ y_{\mathrm{ini}} \\ u \\ y \end{bmatrix}$ belongs to this system if and only if there exists

$g \in \mathbb{R}^H$ such that $\begin{bmatrix} U_{\mathrm{P}} \\ Y_{\mathrm{P}} \\ U_{\mathrm{F}} \\ Y_{\mathrm{F}} \end{bmatrix} g = \begin{bmatrix} u_{\mathrm{ini}} \\ y_{\mathrm{ini}} \\ u \\ y \end{bmatrix}$. [Willems *et al*. 2005]

If $T_{\mathrm{ini}} \geq \ell$, the output trajectory y is uniquely determined for every given u.

The role of data



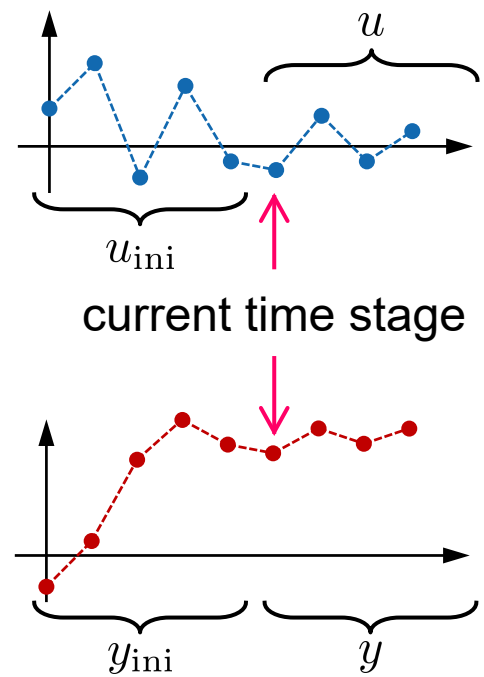
From data to control

$$\begin{bmatrix} U_{\mathrm{P}} \\ Y_{\mathrm{P}} \\ U_{\mathrm{F}} \\ Y_{\mathrm{F}} \end{bmatrix} g = \begin{bmatrix} u_{\mathrm{ini}} \\ y_{\mathrm{ini}} \\ u \\ y \end{bmatrix}$$

The most recent trajectory

Historical data

The future trajectory of length *N*



From data to control

$$\min_{g,\, u\in\mathcal{U},\, y\in\mathcal{Y}} \quad \|u\|_R^2 + \|y - r\|_Q^2$$

$$\text{s.t.} \quad \begin{bmatrix} U_{\mathrm{P}} \\ Y_{\mathrm{P}} \\ U_{\mathrm{F}} \\ Y_{\mathrm{F}} \end{bmatrix} g = \begin{bmatrix} u_{\mathrm{ini}} \\ y_{\mathrm{ini}} \\ u \\ y \end{bmatrix}$$

Data-Enabled Predictive Control (DeePC)

[Coulson et al. 2019]

We solves the above optimization problem in receding horizon manner to get the optimal control inputs.

Apply $(u_t, ..., u_{t+k}) = (u_0^\star, ..., u_k^\star)$ to the system for some time steps $k \leq N - 1$;

Update $\mathrm{col}(u_{\mathrm{ini}}, y_{\mathrm{ini}})$ to the most recent input/output measurements;

Set t to $t + k + 1$ for the DeePC algorithm.

From data to robust control

$$\min_{g,\, u\in\mathcal{U},\, y\in\mathcal{Y}} \quad \|u\|_R^2 + \|y-r\|_Q^2$$

$$\text{s.t.} \quad \begin{bmatrix} U_\mathrm{P} \\ Y_\mathrm{P} \\ U_\mathrm{F} \\ Y_\mathrm{F} \end{bmatrix} g = \begin{bmatrix} u_\mathrm{ini} \\ y_\mathrm{ini} \\ u \\ y \end{bmatrix}$$

How to deal with noise in data?

The data is transparent in this formulation!

Robustification

$$\min_{g} \max_{\xi\in\mathcal{D}} \min_{\substack{(u,y)\in\mathcal{C} \\ \sigma_u,\sigma_y}} \|u\|_R^2 + \|y-r\|_Q^2 + \lambda_u\|\sigma_u\|^2 + \lambda_y\|\sigma_y\|^2$$

$$\text{s.t.} \quad \begin{bmatrix} U_\mathrm{P}(\xi) \\ Y_\mathrm{P}(\xi) \\ U_\mathrm{F}(\xi) \\ Y_\mathrm{F}(\xi) \end{bmatrix} g = \begin{bmatrix} u_\mathrm{ini}(\xi) \\ y_\mathrm{ini}(\xi) \\ u \\ y \end{bmatrix} + \begin{bmatrix} \sigma_u \\ \sigma_y \\ 0 \\ 0 \end{bmatrix}$$

(Robust DeePC)

[Huang et al. 2021]

$$\boxed{Y_\mathrm{P}(\xi) = \hat{Y}_\mathrm{P} + \sum_{j=1}^{n_\xi} Y_\mathrm{P}^{(j)}\xi_j \quad \text{and} \quad y_\mathrm{ini}(\xi) = \hat{y}_\mathrm{ini} + \sum_{j=1}^{n_\xi} y_\mathrm{ini}^{(j)}\xi_j}$$

e.g., $Y_\mathrm{P}(\xi) = \hat{Y}_\mathrm{P} + \begin{bmatrix} \xi_1 & \cdots & \xi_i \\ \vdots & \ddots & \vdots \\ \xi_j & \cdots & \xi_n \end{bmatrix}$

Robust DeePC

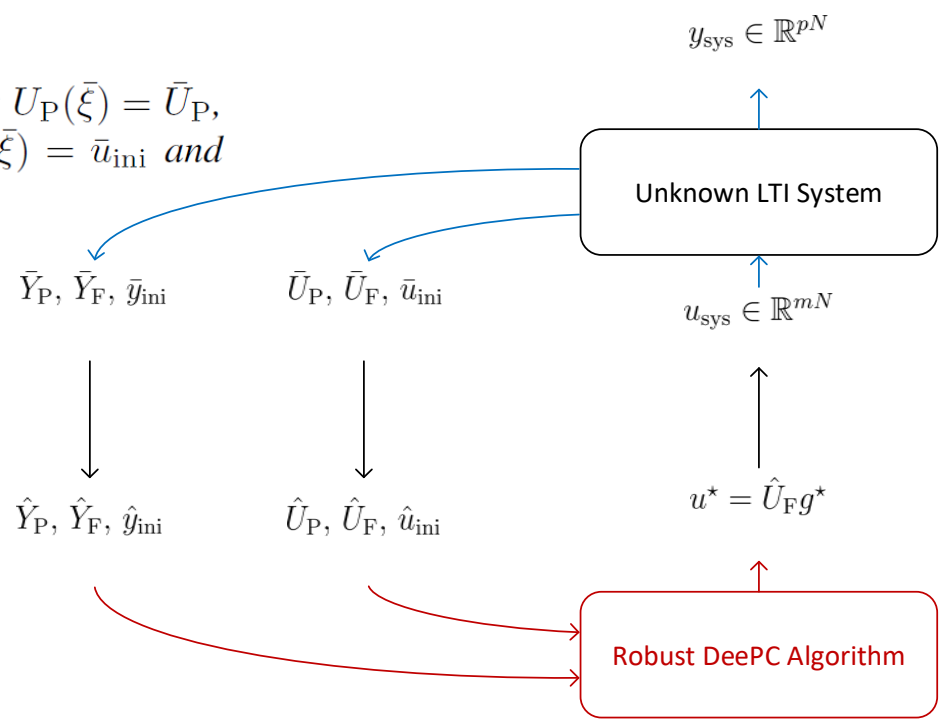
$$\min_{g} \max_{\xi \in \mathcal{D}} \min_{\substack{(u,y)\in\mathcal{C} \\ \sigma_u, \sigma_y}} \|u\|_R^2 + \|y - r\|_Q^2 + \lambda_u \|\sigma_u\|^2 + \lambda_y \|\sigma_y\|^2$$

$$\text{s.t.} \quad \begin{bmatrix} U_{\mathrm{P}}(\xi) \\ Y_{\mathrm{P}}(\xi) \\ U_{\mathrm{F}}(\xi) \\ Y_{\mathrm{F}}(\xi) \end{bmatrix} g = \begin{bmatrix} u_{\mathrm{ini}}(\xi) \\ y_{\mathrm{ini}}(\xi) \\ u \\ y \end{bmatrix} + \begin{bmatrix} \sigma_u \\ \sigma_y \\ 0 \\ 0 \end{bmatrix}$$

(Robust DeePC)

Assumption 1. *There exists a* $\bar{\xi} \in \mathcal{D}$ *such that* $U_{\mathrm{P}}(\bar{\xi}) = \bar{U}_{\mathrm{P}}$, $Y_{\mathrm{P}}(\bar{\xi}) = \bar{Y}_{\mathrm{P}}$, $U_{\mathrm{F}}(\bar{\xi}) = \bar{U}_{\mathrm{F}}$, $Y_{\mathrm{F}}(\bar{\xi}) = \bar{Y}_{\mathrm{F}}$, $u_{\mathrm{ini}}(\bar{\xi}) = \bar{u}_{\mathrm{ini}}$ *and* $y_{\mathrm{ini}}(\bar{\xi}) = \bar{y}_{\mathrm{ini}}$.

The noiseless (perfect) data, though unknown, can be captured by a large enough disturbance set.



Robust DeePC

$$\min_{g} \max_{\xi \in \mathcal{D}} \min_{\substack{(u,y)\in\mathcal{C} \\ \sigma_u,\sigma_y}} \|u\|_R^2 + \|y-r\|_Q^2 + \lambda_u\|\sigma_u\|^2 + \lambda_y\|\sigma_y\|^2 \boxed{= c_{\text{opt}}}$$

$$\text{s.t.} \quad \begin{bmatrix} U_{\text{P}}(\xi) \\ Y_{\text{P}}(\xi) \\ U_{\text{F}}(\xi) \\ Y_{\text{F}}(\xi) \end{bmatrix} g = \begin{bmatrix} u_{\text{ini}}(\xi) \\ y_{\text{ini}}(\xi) \\ u \\ y \end{bmatrix} + \begin{bmatrix} \sigma_u \\ \sigma_y \\ 0 \\ 0 \end{bmatrix}$$

(Robust DeePC)

Assumption 1. *There exists a* $\bar{\xi} \in \mathcal{D}$ *such that* $U_{\text{P}}(\bar{\xi}) = \bar{U}_{\text{P}}$, $Y_{\text{P}}(\bar{\xi}) = \bar{Y}_{\text{P}}$, $U_{\text{F}}(\bar{\xi}) = \bar{U}_{\text{F}}$, $Y_{\text{F}}(\bar{\xi}) = \bar{Y}_{\text{F}}$, $u_{\text{ini}}(\bar{\xi}) = \bar{u}_{\text{ini}}$ *and* $y_{\text{ini}}(\bar{\xi}) = \bar{y}_{\text{ini}}$.

Performance guarantees

Theorem 1. *If Assumption 1 holds and the unknown system (1) does not have input disturbance, then there exist sufficiently large* $(\lambda_u, \lambda_y) \in \mathbb{R}^2_{\geq 0}$ *such that*

[redacted]

[Huang et al. 2021]

Optimization cost

$g^\star$

Control sequence

$u_{\text{sys}} = \hat{U}_{\text{F}} g^\star$

[redacted]

Unknown System (1)

Output sequence

y_{sys}

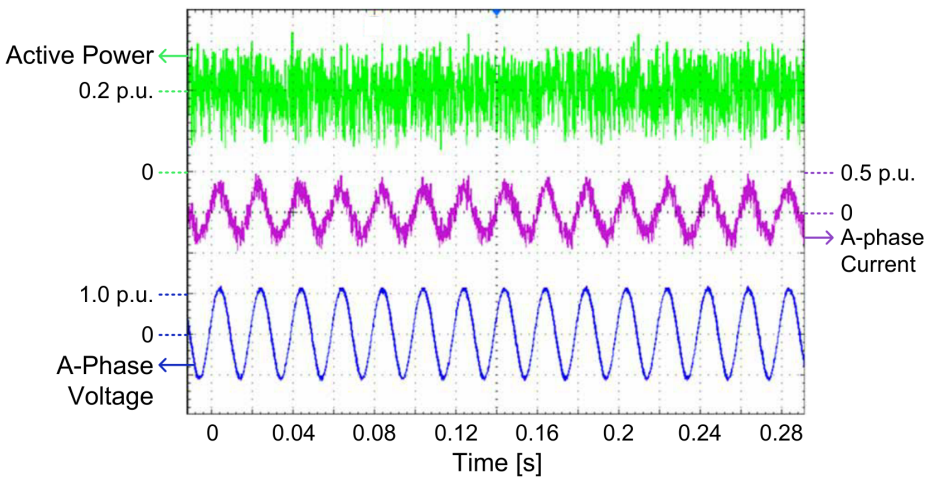
Realized cost

$c_{\text{realized}} = \|u_{\text{sys}}\|_R^2 + \|y_{\text{sys}} - r\|_Q^2$

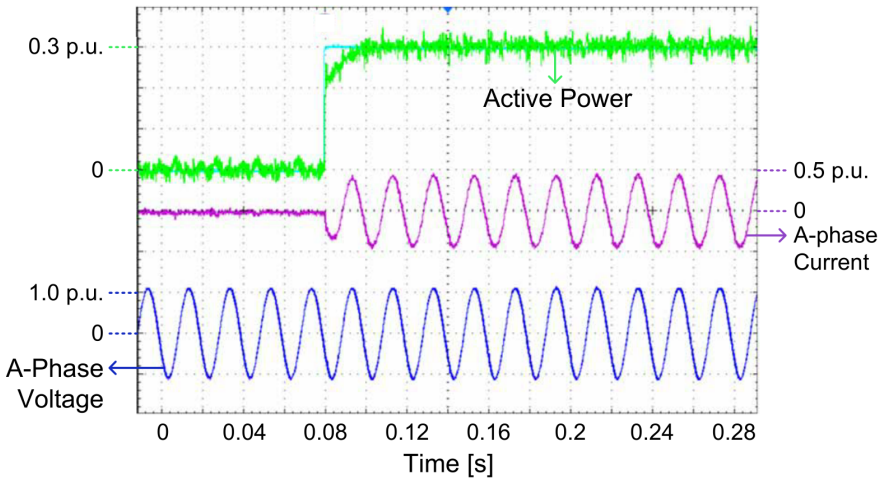
Applications in power converters – Synchronization

Experiment results:

(Closed-form solution of Robust DeePC is used)



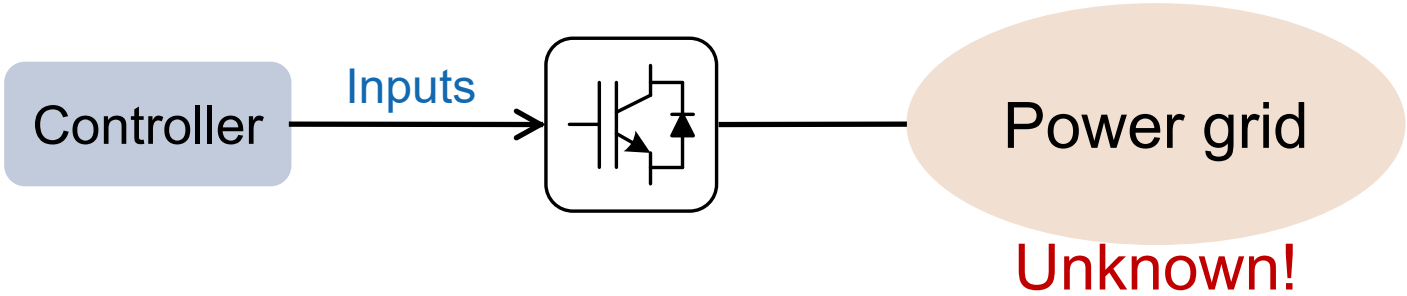
Data collection



DeePC activated

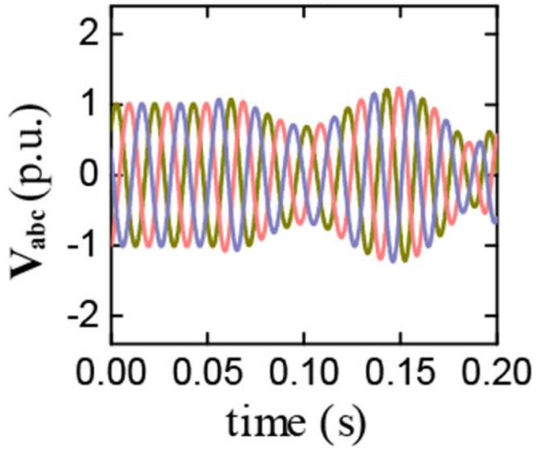
0.08s: active power reference steps from 0 to 0.3(p.u.)

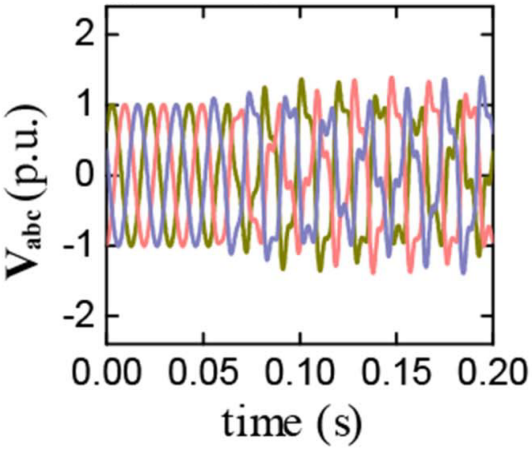
Compared with conventional control design



Conventional control design for power converters is complicated.

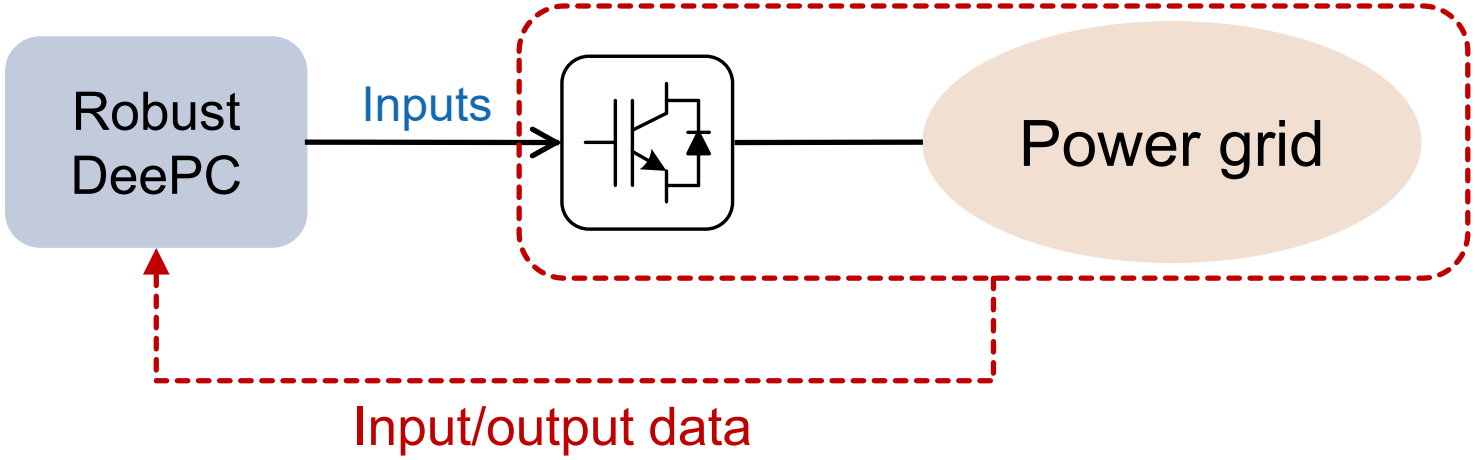
We may need to tune 5 (coupled) PI controllers, and the performance highly depends on the power grid condition.

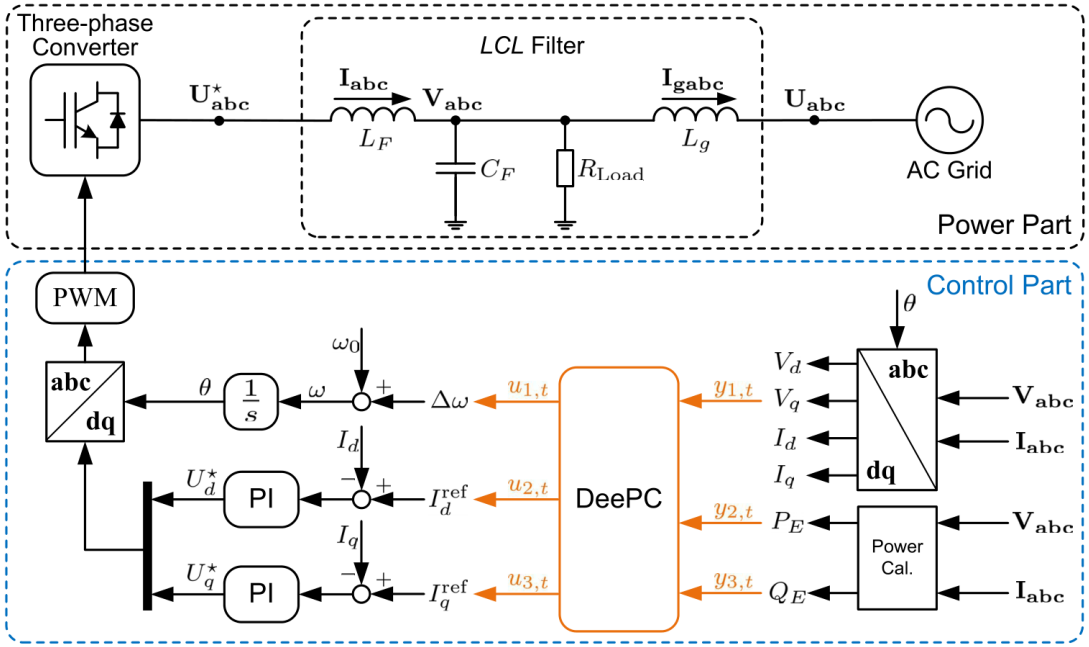




Instabilities!

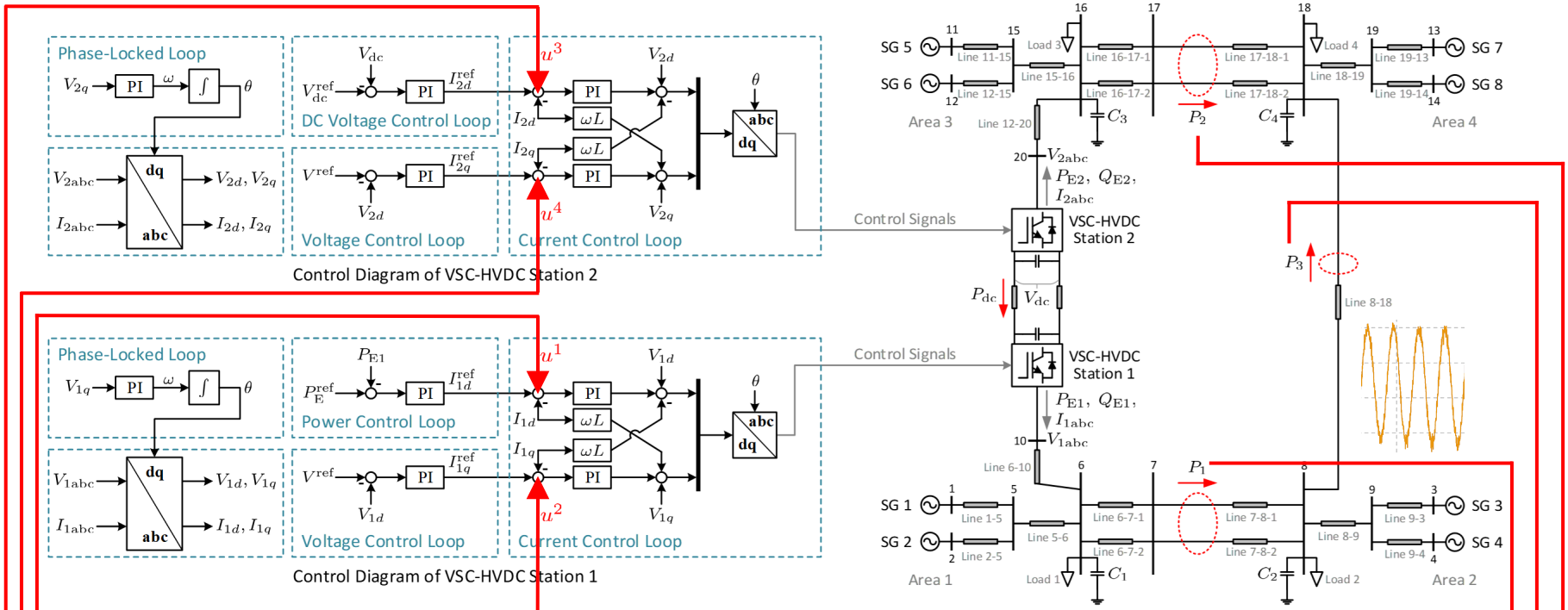
Compared with conventional control design





(Nonlinear)

Applications in power systems – Oscillation Damping



Unknown system ($n = 208$), nonlinear, noisy, delays

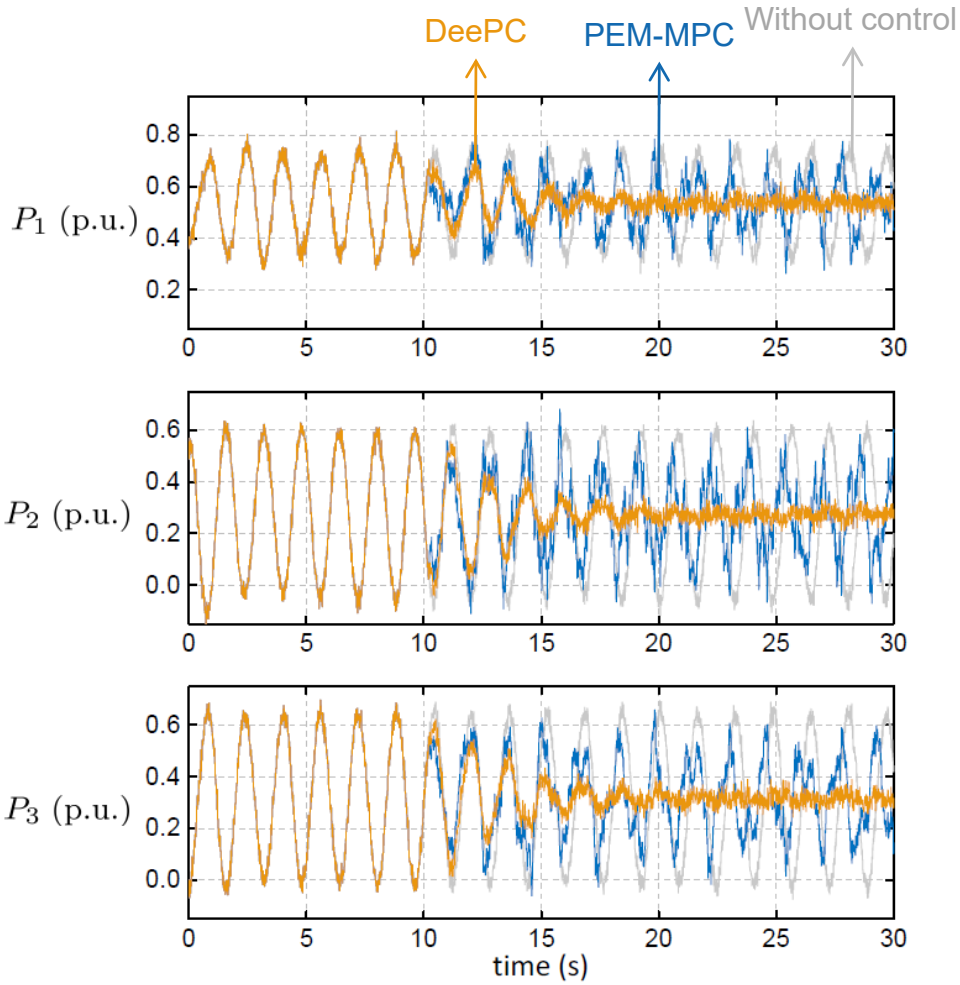
Robust DeePC

System inputs u

System outputs y

Applications in power systems – Oscillation Damping

Oscillations are well eliminated by DeePC !



Sampling time for DeePC is 0.02s.

DeePC is activated at $t = 10$s.

$$T_{\text{ini}} = 60$$
$$N = 120$$
$$T = 1500$$

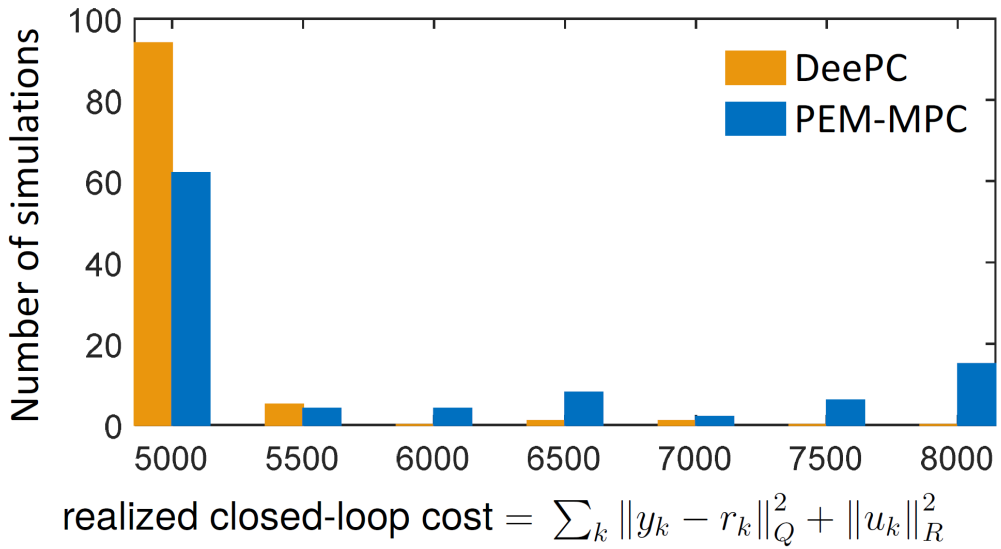
By using OSQP, it takes about 1s to solve the optimization problem. (on Intel Core i5 7200U)

Then, we apply the first 60 elements of the optimal control sequence to the system.

implementable

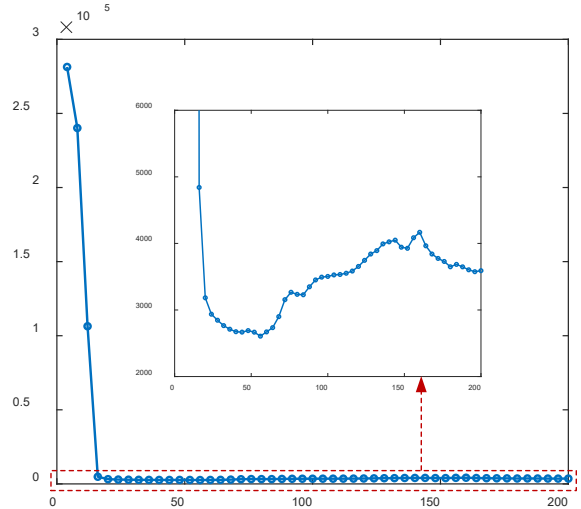
Applications in power systems – Oscillation Damping

We ran a lot of simulations with different data sets…

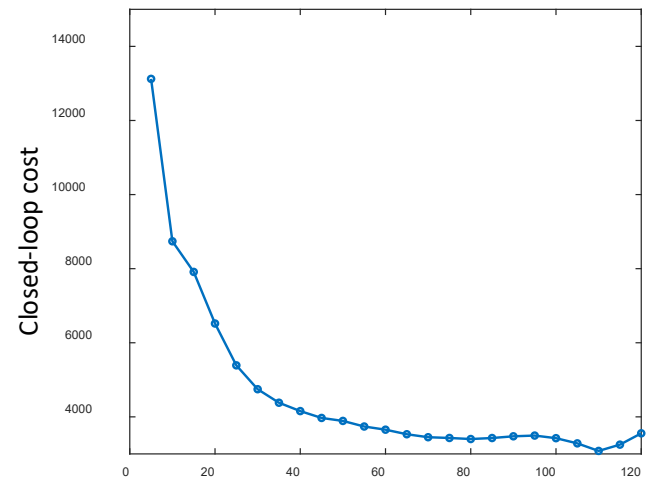


and here we can see the power of *direct robustification*.

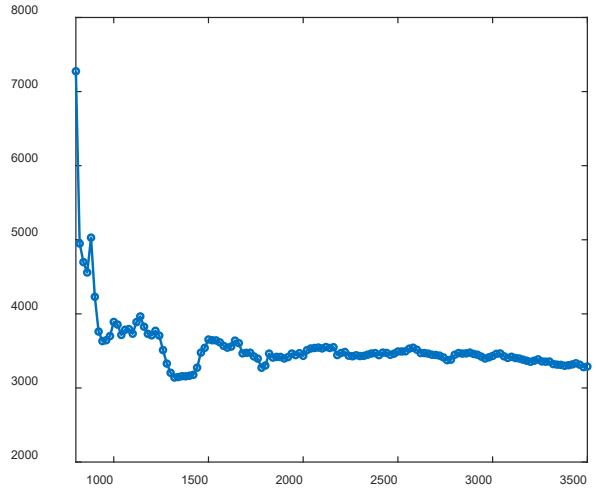
Applications in power systems – Oscillation Damping



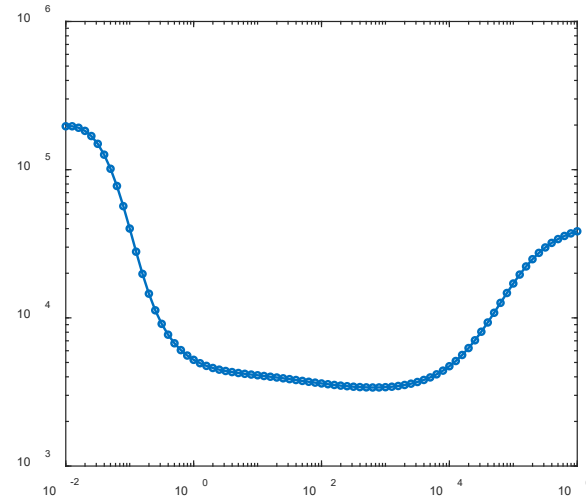
Prediction horizon N



T_{ini} Estimation horizon

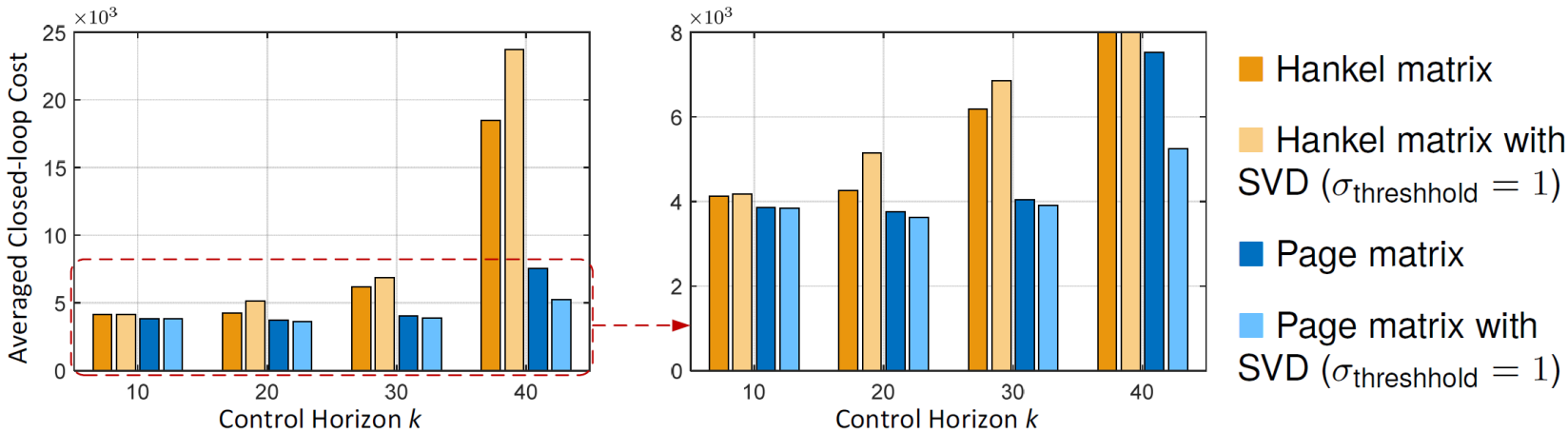


Data length T



λ_g Regularization

Applications in power systems – Oscillation Damping



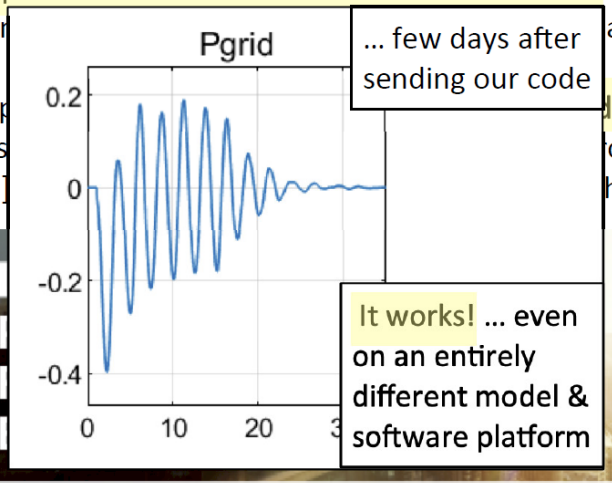
- comparison baseline: Hankel and Page matrices of ***same size***
- ***perfomance***: Page consistency beats Hankel matrix predictors
- offline ***denoising via SVD threshholding*** works wonderfully for Page though obviously not for Hankel (entries are constrained)
- effects very pronounced for ***longer horizon*** (= open-loop time)
- ***price-to-be-paid***: Page matrix predictor requires more data

Applications in power systems – Oscillation Damping

Dear Linbin and Florian,

[...] which I believe could be of importance to our work at Hitachi Power Grids. We do have [...] require off-line tuning that [...] co an adaptive approach would be very interesting.

If p d DeePC approach with our more detailed HVDC sys oblem. Could so some code be made available [...] her to do such a demonstration ? [...]

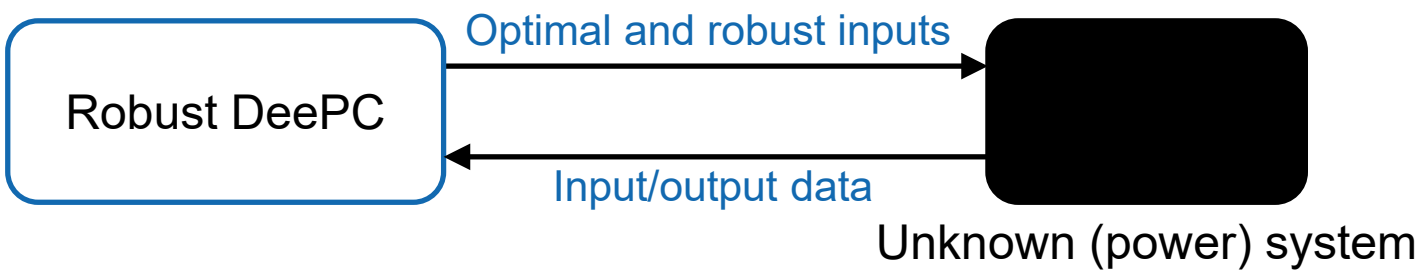


... few days after sending our code

It works! ... even on an entirely different model & software platform

Conclusions

- ✓ Direct data-driven control that provides *robust and optimal* inputs; (End-to-End Robustification)
- ✓ Tractable formulations that can be solved efficiently;
- ✓ Data-driven synchronization and stabilization in power-electronic devices;
- ✓ Data-driven oscillation damping in power systems;



[1] L. Huang, J. Zhen, J. Lygeros, and F. Dörfler. “Robust Data-Enabled Predictive Control: Tractable Formulations and Performance Guarantees.” submitted to IEEE trans. Automatic Control, 2021.

[2] L. Huang, J. Zhen, J. Lygeros, and F. Dörfler. “Quadratic regularization of data-enabled predictive control: Theory and application to power converter experiments.” in 19th IFAC Symposium on System Identification, Padova, 2021.

[3] L. Huang, J. Coulson, J. Lygeros, and F. Dörfler. “Decentralized data-enabled predictive control for power system oscillation damping.” IEEE trans. Control Systems Technology, 2021.

[4] L. Huang, J. Coulson, J. Lygeros, and F. Dörfler. “Data-enabled predictive control for grid-connected power converters.” CDC, 2019.

[5] J. Coulson, J. Lygeros, and F. Dörfler. “Data-enabled predictive control: In the shallows of the DeePC.” ECC, 2019.