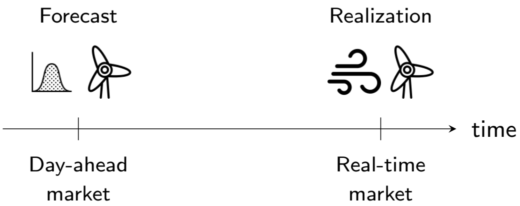
On the Existence of AI Equilibrium in Electricity Markets

Vladimir Dvorkin

Department of Electrical Engineering and Computer Science
University of Michigan

5th Workshop on Foundation Models of the Electric Grid
Harvard John A. Paulson School of Engineering and Applied Sciences
March, 2026

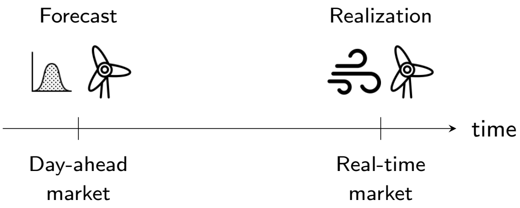
UNIVERSITY OF MICHIGAN

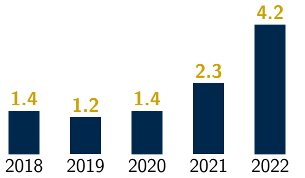


- Two-stage electricity markets to manage uncertainty of renewables:
  - Day-ahead market: minimize the cost of power supply w.r.t. forecast
  - Real-time market: least-cost re-dispatch to accommodate forecast errors

- As renewable penetration increases, the cost of real-time re-dispatch also increases

Two-stage electricity markets

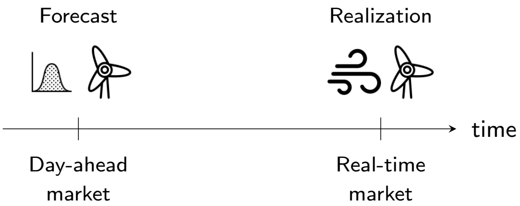


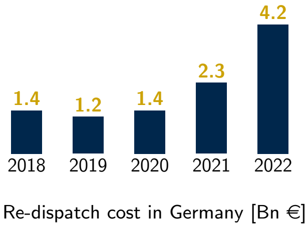


Re-dispatch cost in Germany [Bn €]

- Two-stage electricity markets to manage uncertainty of renewables:
  - Day-ahead market: minimize the cost of power supply w.r.t. forecast
  - Real-time market: least-cost re-dispatch to accommodate forecast errors

- As renewable penetration increases, the cost of real-time re-dispatch also increases

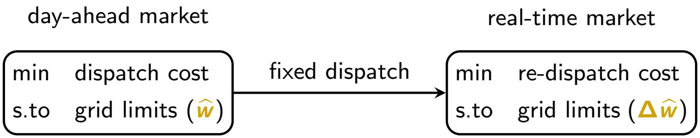




- Two-stage electricity markets to manage uncertainty of renewables:
  - Day-ahead market: minimize the cost of power supply w.r.t. forecast
  - Real-time market: least-cost re-dispatch to accommodate forecast errors

- As renewable penetration increases, the cost of real-time re-dispatch also increases

How to make renewable power generation less expensive for the system?


day-ahead market
real-time market
min dispatch cost
s.to grid limits ($\widehat{w}$)
fixed dispatch
min re-dispatch cost
s.to grid limits ($\Delta\widehat{w}$)

day-ahead market		real-time market
min dispatch cost s.to grid limits $(\widehat{w})$	$\xrightarrow{\text{fixed dispatch}}$	min re-dispatch cost s.to grid limits $(\boldsymbol{\Delta}\widehat{w})$

Improving cost efficiency across day-ahead and real-time markets:

- Stochastic electricity market design [PZP10, M^{+}12, Dvo19]:
  - + Co-optimization of dispatch and re-dispatch decisions
  - + Least-cost solution in *expectation*
  - – Market properties only hold in *expectation*

$$\begin{aligned} \min \quad & \text{dispatch cost} + \mathbb{E}_{\mathbb{P}_{\boldsymbol{\Delta}\widehat{w}}}[\text{re-dispatch cost}] \\ \text{s.to} \quad & \text{grid limits } (\widehat{w}, \boldsymbol{\Delta}\widehat{w}) \text{ for all } \boldsymbol{\Delta}\widehat{w} \sim \mathbb{P}_{\boldsymbol{\Delta}\widehat{w}} \end{aligned}$$

day-ahead market → fixed dispatch → real-time market

day-ahead market			real-time market	
min	dispatch cost	fixed dispatch →	min	re-dispatch cost
s.to	grid limits $(\widehat{w})$		s.to	grid limits $(\Delta\widehat{w})$

Improving cost efficiency across day-ahead and real-time markets:

- Stochastic electricity market design [PZP10, M+12, Dvo19]:
  - \+ Co-optimization of dispatch and re-dispatch decisions
  - \+ Least-cost solution in *expectation*
  - – Market properties only hold in *expectation*

$$\begin{aligned} \text{min} \quad & \text{dispatch cost} + \mathbb{E}_{\mathbb{P}_{\Delta\widehat{w}}}[\text{re-dispatch cost}] \\ \text{s.to} \quad & \text{grid limits } (\widehat{w}, \Delta\widehat{w}) \text{ for all } \Delta\widehat{w} \sim \mathbb{P}_{\Delta\widehat{w}} \end{aligned}$$

- Approximating stochastic market efficiency within deterministic markets:
  - Improved scheduling of renewbales [M+14]
  - Cost-aware reserve requirements [DDM18]
  - Cost-aware transmission allocation [JKP17, DP19]

Two-stage electricity markets (continued)

day-ahead market

min dispatch cost
s.to grid limits ($\widehat{w}$)

fixed dispatch →

real-time market

min re-dispatch cost
s.to grid limits ($\Delta\widehat{w}$)

Improving cost efficiency across day-ahead and real-time markets:

- Stochastic electricity market design [PZP10, M+12, Dvo19]:
  - \+ Co-optimization of dispatch and re-dispatch decisions
  - \+ Least-cost solution in *expectation*
  - – Market properties only hold in *expectation*

$$\begin{aligned} \text{min} \quad & \text{dispatch cost} + \mathbb{E}_{\mathbb{P}_{\Delta\widehat{w}}}[\text{re-dispatch cost}] \\ \text{s.to} \quad & \text{grid limits } (\widehat{w}, \Delta\widehat{w}) \text{ for all } \Delta\widehat{w} \sim \mathbb{P}_{\Delta\widehat{w}} \end{aligned}$$

- Approximating stochastic market efficiency within deterministic markets:
  - Improved scheduling of renewbales [M+14]
  - Cost-aware reserve requirements [DDM18]
  - Cost-aware transmission allocation [JKP17, DP19]

Can AI bridge the gap between ideal stochastic and current deterministic market designs?

Typical grid optimization problem:

$$
\begin{aligned}
\min_{\mathbf{p},\mathbf{q},\mathbf{u}}.\quad & c(\mathbf{p}_\theta) + s(\mathbf{u}_\theta) && \textit{generation and UC cost}\\
\text{s.to}\quad & f(\mathbf{p}_\theta, \mathbf{q}_\theta, \mathbf{w}_\theta) = \mathbf{0} \quad : \boldsymbol{\lambda}_\theta && \textit{power flow equations}\\
& g(\mathbf{p}_\theta, \mathbf{q}_\theta, \mathbf{w}_\theta) \leqslant b(\mathbf{u}_\theta) && \textit{gen, flow, voltage limits}
\end{aligned}
$$

Power systems optimization meets machine learning (ML)

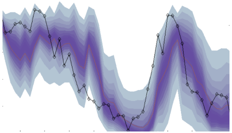
Typical grid optimization problem:

$$\begin{aligned} \min_{\mathbf{p},\mathbf{q},\mathbf{u}} \quad & c(\mathbf{p}_\theta) + s(\mathbf{u}_\theta) && \textit{generation and UC cost} \\ \text{s.to} \quad & f(\mathbf{p}_\theta, \mathbf{q}_\theta, \mathbf{w}_\theta) = \mathbf{0} \quad : \boldsymbol{\lambda}_\theta && \textit{power flow equations} \\ & g(\mathbf{p}_\theta, \mathbf{q}_\theta, \mathbf{w}_\theta) \leqslant b(\mathbf{u}_\theta) && \textit{gen, flow, voltage limits} \end{aligned}$$

Typical grid optimization problem:

$$\begin{aligned} \min_{\mathbf{p},\mathbf{q},\mathbf{u}}.\quad & c(\mathbf{p}_\theta) + s(\mathbf{u}_\theta) && \textit{generation and UC cost} \\ \text{s.to}\quad & f(\mathbf{p}_\theta, \mathbf{q}_\theta, \mathbf{w}_\theta) = \mathbf{0} \quad : \boldsymbol{\lambda}_\theta && \textit{power flow equations} \\ & g(\mathbf{p}_\theta, \mathbf{q}_\theta, \mathbf{w}_\theta) \leqslant b(\mathbf{u}_\theta) && \textit{gen, flow, voltage limits} \end{aligned}$$

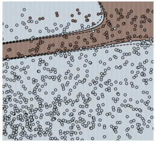
Net load forecasting [PCK07, ZHS24]



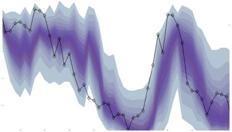
Typical grid optimization problem:

$$\begin{aligned} \min_{\mathbf{p},\mathbf{q},\mathbf{u}}.\quad & c(\mathbf{p}_\theta) + s(\mathbf{u}_\theta) && \textit{generation and UC cost} \\ \text{s.to}\quad & f(\mathbf{p}_\theta, \mathbf{q}_\theta, \mathbf{w}_\theta) = \mathbf{0} \quad : \boldsymbol{\lambda}_\theta && \textit{power flow equations} \\ & g(\mathbf{p}_\theta, \mathbf{q}_\theta, \mathbf{w}_\theta) \leqslant b(\mathbf{u}_\theta) && \textit{gen, flow, voltage limits} \end{aligned}$$

SVM-based prediction of on/off gen. status [PK24]



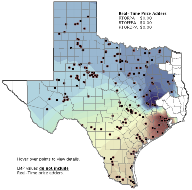
Net load forecasting [PCK07, ZHS24]



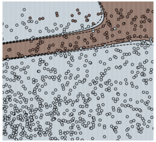
Typical grid optimization problem:

$$
\begin{aligned}
\min_{\mathbf{p},\mathbf{q},\mathbf{u}} \quad & c(\mathbf{p}_\theta) + s(\mathbf{u}_\theta) && \textit{generation and UC cost} \\
\text{s.to} \quad & f(\mathbf{p}_\theta, \mathbf{q}_\theta, \mathbf{w}_\theta) = \mathbf{0} \quad : \boldsymbol{\lambda}_\theta && \textit{power flow equations} \\
& g(\mathbf{p}_\theta, \mathbf{q}_\theta, \mathbf{w}_\theta) \leqslant b(\mathbf{u}_\theta) && \textit{gen, flow, voltage limits}
\end{aligned}
$$

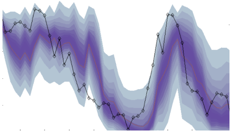
Real-time electricity pricing via GNNs [LWZ21]



SVM-based prediction of on/off gen. status [PK24]



Net load forecasting [PCK07, ZHS24]



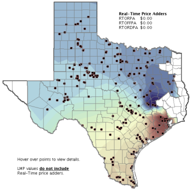
Typical grid optimization problem:

$$\min_{\mathbf{p},\mathbf{q},\mathbf{u}} \quad c(\mathbf{p}_\theta) + s(\mathbf{u}_\theta) \qquad \textit{generation and UC cost}$$

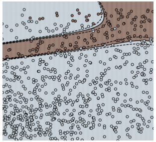
$$\text{s.to} \quad f(\mathbf{p}_\theta, \mathbf{q}_\theta, \mathbf{w}_\theta) = \mathbf{0} \quad : \boldsymbol{\lambda}_\theta \qquad \textit{power flow equations}$$

$$g(\mathbf{p}_\theta, \mathbf{q}_\theta, \mathbf{w}_\theta) \leqslant b(\mathbf{u}_\theta) \qquad \textit{gen, flow, voltage limits}$$

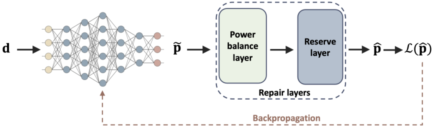
Real-time electricity pricing via GNNs [LWZ21]



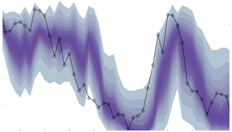
SVM-based prediction of on/off gen. status [PK24]



Optimal generation dispatch prediction using optimization proxies [CTVH23]



Net load forecasting [PCK07, ZHS24]



Typical grid optimization problem:

$$\min_{\mathbf{p},\mathbf{q},\mathbf{u}} \quad c(\mathbf{p}_\theta) + s(\mathbf{u}_\theta) \qquad \textit{generation and UC cost}$$
$$\text{s.to} \quad f(\mathbf{p}_\theta, \mathbf{q}_\theta, \mathbf{w}_\theta) = \mathbf{0} \quad : \boldsymbol{\lambda}_\theta \qquad \textit{power flow equations}$$
$$g(\mathbf{p}_\theta, \mathbf{q}_\theta, \mathbf{w}_\theta) \leqslant b(\mathbf{u}_\theta) \qquad \textit{gen, flow, voltage limits}$$

Does a market equilibrium exist when decision-making is delegated to AI?

Agenda

Agenda

ML errors = decision errors

DC optimal power flow:

$$\min_{\underline{\mathbf{p}}\leqslant\mathbf{p}\leqslant\overline{\mathbf{p}}} \quad \mathbf{p}^\top\mathbf{C}\mathbf{p}+\mathbf{c}^\top\mathbf{p}$$

$$\text{s.to} \quad \mathbf{1}^\top(\mathbf{p}+\widehat{\mathbf{w}}-\mathbf{d})=\mathbf{0}\;:\;\lambda_b$$

$$|\mathbf{F}(\mathbf{p}+\widehat{\mathbf{w}}-\mathbf{d})|\leqslant\overline{\mathbf{f}}\;:\;\boldsymbol{\lambda}_{\overline{f}},\boldsymbol{\lambda}_{\underline{f}}$$

- Electricity market clearing based on DC-OPF
- Relies on the *forecast* $\widehat{\mathbf{w}}$ of wind power generation
- Forecast errors $\rightarrow$ pricing errors via market optimization
- May not be a dominant generation resource, yet still exposes the entire electricity trading to errors

Locational marginal prices:

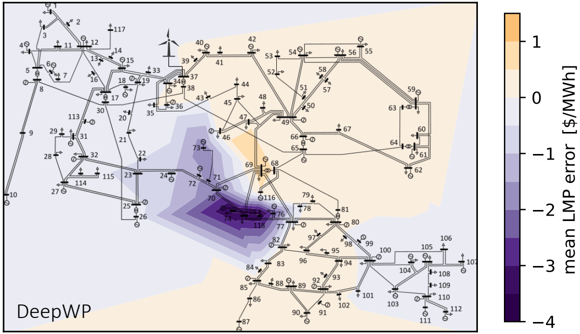
$$\boldsymbol{\lambda}(\widehat{\mathbf{w}})=\underbrace{\mathbf{1}\cdot\lambda_b(\widehat{\mathbf{w}})}_{\text{uniform part}}-\underbrace{\mathbf{F}^\top(\boldsymbol{\lambda}_{\overline{f}}(\widehat{\mathbf{w}})-\boldsymbol{\lambda}_{\underline{f}}(\widehat{\mathbf{w}}))}_{\text{congestion part}}$$

DC optimal power flow:

$$\min_{\underline{\mathbf{p}}\leqslant\mathbf{p}\leqslant\overline{\mathbf{p}}} \quad \mathbf{p}^\top \mathbf{C}\mathbf{p} + \mathbf{c}^\top \mathbf{p}$$

$$\text{s.to} \quad \mathbf{1}^\top(\mathbf{p} + \widehat{\mathbf{w}} - \mathbf{d}) = \mathbf{0} \;:\; \lambda_b$$

$$|\mathbf{F}(\mathbf{p} + \widehat{\mathbf{w}} - \mathbf{d})| \leqslant \overline{\mathbf{f}} \;:\; \boldsymbol{\lambda}_{\overline{f}}, \boldsymbol{\lambda}_{\underline{f}}$$



Thanks to Robert Mieth of Rutgers for the help with this visualization

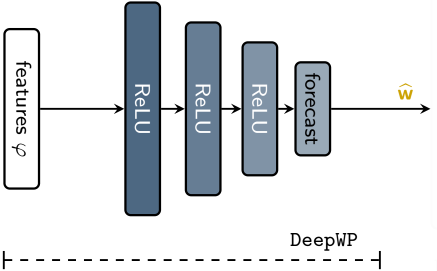
Locational marginal prices:

$$\boldsymbol{\lambda}(\widehat{\mathbf{w}}) = \underbrace{\mathbf{1}\cdot\lambda_b(\widehat{\mathbf{w}})}_{\text{uniform part}} - \underbrace{\mathbf{F}^\top(\boldsymbol{\lambda}_{\overline{f}}(\widehat{\mathbf{w}}) - \boldsymbol{\lambda}_{\underline{f}}(\widehat{\mathbf{w}}))}_{\text{congestion part}}$$

Forecast errors from a single wind power plant propagate into locational marginal price (LMP) errors across the IEEE 118-Bus RTS. Electricity at certain buses is systematically over- or under-priced [DF23].

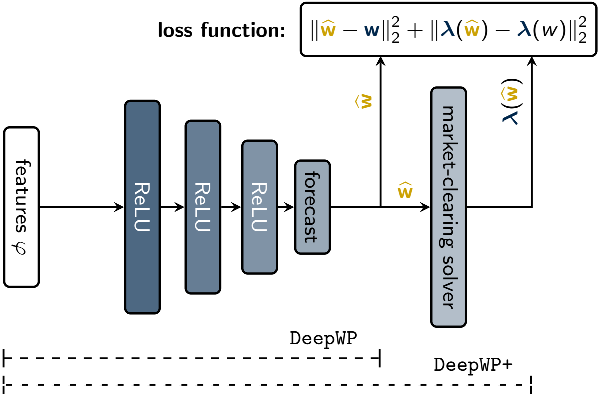
- Dataset $\{(\varphi_1, \mathbf{w}_1), \dots, (\varphi_m, \mathbf{w}_m)\}$ of wind power records, with features φ and measurements $\mathbf{w}$
- Two deep learning architectures `DeepWP` and `DeepWP+` for wind power forecasting:

loss function: $\|\widehat{\mathbf{w}} - \mathbf{w}\|_2^2$



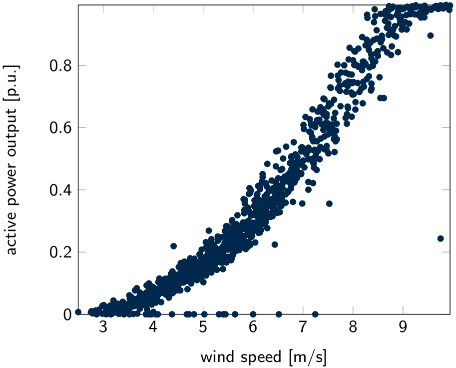
Market clearing as a deep learning layer

- Dataset $\{(\varphi_1, \mathbf{w}_1), \dots, (\varphi_m, \mathbf{w}_m)\}$ of wind power records, with features φ and measurements $\mathbf{w}$
- Two deep learning architectures `DeepWP` and `DeepWP+` for wind power forecasting:

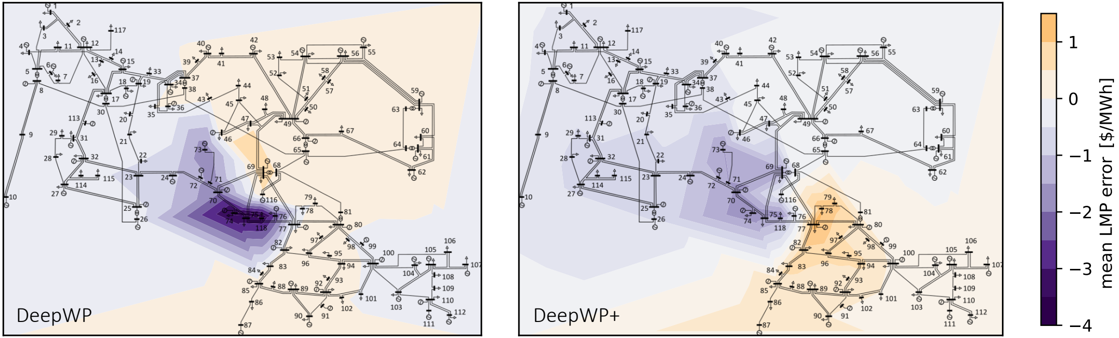


- DeepWP+ incorporates market clearing as an optimization layer [DAK17], which informs on pricing errors

- 1,000 wind power records from a real wind power turbine:
  - Active power output
  - Wind speed and direction
  - Blade pitch angle
- `DeepWP` has 4 hidden layers with 30 neurons each. `DeepWP+` additionally includes an opt. layer
- ADAM optimizer with varying learning rate



IEEE 118-bus system



`DeepWP`: Forecast error objective – LMP errors $[-4, 1]$ \$/MWh

`DeepWP+`: LMP error objective – LMP errors $[-1, 1]$ \$/MWh

Underlying trade-offs between forecast and price errors

case	DeepWP			DeepWP+					
	RMSE($\widehat{w}$)	RMSE($\widehat{\lambda}$)	CVaR($\widehat{\lambda}$)	RMSE($\widehat{w}$)		RMSE($\widehat{\lambda}$)		CVaR($\widehat{\lambda}$)	
	MWh	\$/MWh	\$/MWh	MWh	gain	\$/MWh	gain	\$/MWh	gain
14_ieee	0.35	0.62	1.52	0.35	+0.6%	0.61	−0.6%	1.50	−0.8%
57_ieee	2.31	11.03	34.64	2.60	+11.2%	10.72	−2.9%	33.59	−3.1%
24_ieee	4.08	8.62	37.70	4.51	+9.6%	8.33	−3.5%	36.35	−3.7%
39_epri	5.94	11.15	31.21	6.43	+7.6%	10.19	−9.4%	28.02	−11.4%
73_ieee	4.02	5.12	16.21	5.51	+26.9%	4.24	−20.8%	13.41	−20.9%
118_ieee	2.29	3.59	11.32	2.60	+12.1%	2.88	−24.7%	9.06	−25.0%

- Price errors reduction comes at the expense of forecast error
- Price error reduction is more significant in larger networks

For more results, including **price fairness**: https://arxiv.org/pdf/2308.01436

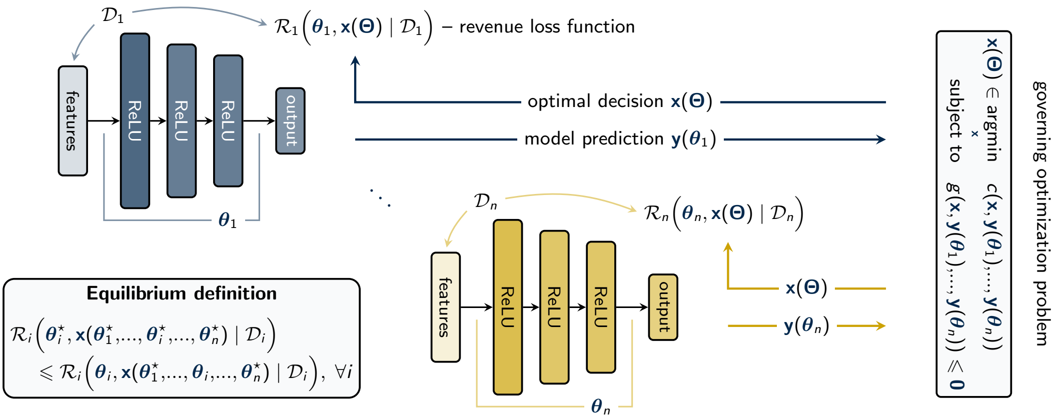
Agenda

Introduction

Controlling the impact of ML errors on electricity pricing

Nash equilibrium of ML models in electricity markets

Insights for foundational models in power systems



- Many systems governed by optimization (pricing, scheduling, social planning, optimal control,...)
- Arbitrary AI models (regression, deep learning, reinforcement learning, trees)
- From Nash to Generalized Nash Equilibrium: more insights, algorithms, etc.
- Does the design of the governing optimization steer AI models to operational and economic equilibrium?

Baseline approach to wind power forecasting:

- Collect a training dataset $\mathcal{D} = \{(\varphi_1, \mathbf{w}_1), \dots, \varphi_n, \mathbf{w}_n)\}$
- Machine learning model $\mathbb{W}_{\boldsymbol{\theta}} : \mathcal{F} \mapsto \mathcal{W}$ with parameter $\boldsymbol{\theta}$
- Learn optimal parameter $\boldsymbol{\theta}^\star$ by minimizing a prediction loss

$$\underset{\|\boldsymbol{\theta}\|_1 \leqslant \tau}{\min.} \quad \mathcal{L}(\boldsymbol{\theta} \,|\, \mathcal{D}) = \frac{1}{n} \sum_{i=1}^{n} \|\mathbb{W}_{\boldsymbol{\theta}}(\varphi_i) - \mathbf{w}_i\|_2^2$$

Baseline approach to wind power forecasting:

- Collect a training dataset $\mathcal{D} = \{(\varphi_1, \mathbf{w}_1), \dots, \varphi_n, \mathbf{w}_n)\}$
- Machine learning model $\mathbb{W}_{\boldsymbol{\theta}} : \mathcal{F} \mapsto \mathcal{W}$ with parameter $\boldsymbol{\theta}$
- Learn optimal parameter $\boldsymbol{\theta}^\star$ by minimizing a prediction loss

$$\underset{\|\boldsymbol{\theta}\|_1 \leqslant \tau}{\text{min.}} \quad \mathcal{L}(\boldsymbol{\theta} \,|\, \mathcal{D}) = \frac{1}{n} \sum_{i=1}^{n} \|\mathbb{W}_{\boldsymbol{\theta}}(\varphi_i) - \mathbf{w}_i\|_2^2$$

Revenue-optimal forecasting [PCK07, CK19, WSC23]:

1. Day-ahead stage: LMP $\boldsymbol{\lambda}_1$ pricing the forecast of wind power
2. Real-time stage: LMP $\boldsymbol{\lambda}_2$ pricing any forecast deviation

$$\underset{\|\boldsymbol{\theta}\|_1 \leqslant \tau}{\text{max.}} \quad \mathcal{R}^W(\boldsymbol{\theta} \,|\, \mathcal{D}, \boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = \frac{1}{n} \sum_{i=1}^{n} \Big(\underbrace{\boldsymbol{\lambda}_{1i} \mathbb{W}_{\boldsymbol{\theta}}(\varphi_i)}_{\text{day-ahead revenue}} + \underbrace{\boldsymbol{\lambda}_{2i} (\mathbf{w}_i - \mathbb{W}_{\boldsymbol{\theta}}(\varphi_i))}_{\text{real-time revenue}} \Big)$$

Optimization of wind power producers

$$\underset{\|\boldsymbol{\theta}_j\|\leqslant\tau_j}{\text{max.}} \quad \underbrace{\mathcal{R}^W(\boldsymbol{\theta}_j|\mathcal{D}_j,\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2)}_{\text{expected revenue}} - \gamma\cdot\underbrace{\mathcal{L}(\boldsymbol{\theta}_j|\mathcal{D}_j)}_{\text{regularization by prediction loss}}$$

for all producers $j \in 1,\dots,b$

feature selection

regularization
by prediction loss

expected revenue

Regression equilibrium in two-stage electricity markets [Dvo24]

Optimization of wind power producers

$$\max_{\|\boldsymbol{\theta}_j\|\leqslant\tau_j} \quad \mathcal{R}^W(\boldsymbol{\theta}_j|\mathcal{D}_j,\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2) - \gamma\cdot\mathcal{L}(\boldsymbol{\theta}_j|\mathcal{D}_j)$$

for all producers $j \in 1,\dots,b$

Optimization of controllable generators

$$\max_{\mathbf{p}_i,\mathbf{r}_i\in\mathcal{G}} \quad \underbrace{\mathcal{R}^G(\mathbf{p}_i,\mathbf{r}_i \mid \boldsymbol{\lambda}_{1i},\boldsymbol{\lambda}_{2i})}_{\text{revenue from generation and regulation}} - \underbrace{c(\mathbf{p}_i,\mathbf{r}_i)}_{\text{generation and regulation cost}}$$

for all training samples $i \in \mathcal{D}_{1:b}$

revenue from generation and regulation

generation and regulation cost

Regression equilibrium in two-stage electricity markets [Dvo24]

Optimization of wind power producers

$$\max_{\|\boldsymbol{\theta}_j\| \leqslant \tau_j} \quad \mathcal{R}^W(\boldsymbol{\theta}_j | \mathcal{D}_j, \boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) - \gamma \cdot \mathcal{L}(\boldsymbol{\theta}_j | \mathcal{D}_j)$$

for all producers $j \in 1, \ldots, b$

Optimization of controllable generators

$$\max_{\mathbf{p}_i, \mathbf{r}_i \in \mathcal{G}} \quad \mathcal{R}^G(\mathbf{p}_i, \mathbf{r}_i \mid \boldsymbol{\lambda}_{1i}, \boldsymbol{\lambda}_{2i}) - c(\mathbf{p}_i, \mathbf{r}_i)$$

for all training samples $i \in \mathcal{D}_{1:b}$

Others

Optimization of wind power producers

$$\max_{\|\theta_j\|\leqslant\tau_j} \quad \mathcal{R}^W(\theta_j|\mathcal{D}_j, \boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) - \gamma\cdot\mathcal{L}(\theta_j|\mathcal{D}_j)$$

for all producers $j \in 1, \dots, b$

Optimization of controllable generators

$$\max_{\mathbf{p}_i,\mathbf{r}_i\in\mathcal{G}} \quad \mathcal{R}^G(\mathbf{p}_i, \mathbf{r}_i \mid \boldsymbol{\lambda}_{1i}, \boldsymbol{\lambda}_{2i}) - c(\mathbf{p}_i, \mathbf{r}_i)$$

for all training samples $i \in \mathcal{D}_{1:b}$

Others

Market-clearing conditions at the 1st and 2nd stages

$$\mathbf{0} \leqslant \boldsymbol{\lambda}_{1i} \perp \mathbf{p}_i + \sum_{j=1}^{b} \mathbb{W}_{\theta_j}(\varphi_i) - \mathbf{d} \geqslant \mathbf{0}, \quad \mathbf{0} \leqslant \boldsymbol{\lambda}_{2i} \perp \mathbf{r}_i + \sum_{j=1}^{b} \mathbf{w}_{ji} - \sum_{j=1}^{b} \mathbb{W}_{\theta_j}(\varphi_i) \geqslant \mathbf{0}$$

for all training samples $i \in \mathcal{D}_{1:b}$

Optimization of wind power producers

$$\max_{\|\theta_j\|\leqslant\tau_j} \quad \mathcal{R}^W(\theta_j|\mathcal{D}_j, \boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) - \gamma \cdot \mathcal{L}(\theta_j|\mathcal{D}_j)$$

for all producers $j \in 1, \dots, b$

Optimization of controllable generators

$$\max_{\mathbf{p}_i, \mathbf{r}_i \in \mathcal{G}} \quad \mathcal{R}^G(\mathbf{p}_i, \mathbf{r}_i \mid \boldsymbol{\lambda}_{1i}, \boldsymbol{\lambda}_{2i}) - c(\mathbf{p}_i, \mathbf{r}_i)$$

for all training samples $i \in \mathcal{D}_{1:b}$

Others

Market-clearing conditions at the 1st and 2nd stages

$$\mathbf{0} \leqslant \boldsymbol{\lambda}_{1i} \perp \mathbf{p}_i + \sum_{j=1}^{b} \mathbb{W}_{\theta_j}(\varphi_i) - \mathbf{d} \geqslant \mathbf{0}, \quad \mathbf{0} \leqslant \boldsymbol{\lambda}_{2i} \perp \mathbf{r}_i + \sum_{j=1}^{b} \mathbf{w}_{ji} - \sum_{j=1}^{b} \mathbb{W}_{\theta_j}(\varphi_i) \geqslant \mathbf{0}$$

for all training samples $i \in \mathcal{D}_{1:b}$

Assumptions:

- Class of ML models $\mathbb{W}_\theta$ is **convex** in θ, e.g., kernel regression
- Training datasets are such that $n \gg \text{card}[\varphi]$ (unique regression solution)
- The intersection of private feasible regions is compact (at least one feasible dispatch $\forall i \in \mathcal{D}_{1:b}$)

Main result: regression equilibrium exists and is unique!

Optimization of wind power producers

$$\max_{\|\theta_j\|\leqslant\tau_j}\quad \mathcal{R}^W(\theta_j|\mathcal{D}_j,\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2)-\gamma\cdot\mathcal{L}(\theta_j|\mathcal{D}_j)$$

for all producers $j\in 1,\dots,b$

Optimization of controllable generators

$$\max_{\mathbf{p}_i,\mathbf{r}_i\in\mathcal{G}}\quad \mathcal{R}^G(\mathbf{p}_i,\mathbf{r}_i\mid\boldsymbol{\lambda}_{1i},\boldsymbol{\lambda}_{2i})-c(\mathbf{p}_i,\mathbf{r}_i)$$

for all training samples $i\in\mathcal{D}_{1:b}$

Others

Market-clearing conditions at the 1st and 2nd stages

$$\mathbf{0}\leqslant\boldsymbol{\lambda}_{1i}\perp\mathbf{p}_i+\sum_{j=1}^{b}\mathbb{W}_{\theta_j}(\varphi_i)-\mathbf{d}\geqslant\mathbf{0},\quad \mathbf{0}\leqslant\boldsymbol{\lambda}_{2i}\perp\mathbf{r}_i+\sum_{j=1}^{b}\mathbf{w}_{ji}-\sum_{j=1}^{b}\mathbb{W}_{\theta_j}(\varphi_i)\geqslant\mathbf{0}$$

for all training samples $i\in\mathcal{D}_{1:b}$

Equilibrium regression profile $\Theta^\star=(\theta_1^\star,\dots,\theta_b^\star)$, such that:

- Feasible operation of the power gird and markets
- Maximized wind power profits, with no incentives to deviate
- Minimized expected dispatch costs across the two markets

Optimization of wind power producers

$$\max_{\|\theta_j\|\leqslant\tau_j} \quad \mathcal{R}^W(\theta_j|\mathcal{D}_j, \boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) - \gamma \cdot \mathcal{L}(\theta_j|\mathcal{D}_j)$$

for all producers $j \in 1, \ldots, b$

Optimization of controllable generators

$$\max_{\mathbf{p}_i, \mathbf{r}_i \in \mathcal{G}} \quad \mathcal{R}^G(\mathbf{p}_i, \mathbf{r}_i \mid \boldsymbol{\lambda}_{1i}, \boldsymbol{\lambda}_{2i}) - c(\mathbf{p}_i, \mathbf{r}_i)$$

for all training samples $i \in \mathcal{D}_{1:b}$

Others

Market-clearing conditions at the 1st and 2nd stages

$$\mathbf{0} \leqslant \boldsymbol{\lambda}_{1i} \perp \mathbf{p}_i + \sum_{j=1}^{b} \mathbb{W}_{\theta_j}(\varphi_i) - \mathbf{d} \geqslant \mathbf{0}, \quad \mathbf{0} \leqslant \boldsymbol{\lambda}_{2i} \perp \mathbf{r}_i + \sum_{j=1}^{b} \mathbf{w}_{ji} - \sum_{j=1}^{b} \mathbb{W}_{\theta_j}(\varphi_i) \geqslant \mathbf{0}$$

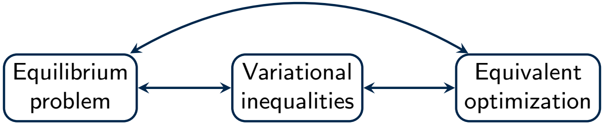
for all training samples $i \in \mathcal{D}_{1:b}$

Equilibrium regression profile $\Theta^\star = (\theta_1^\star, \ldots, \theta_b^\star)$, such that:

- Feasible operation of the power gird and markets
- Maximized wind power profits, with no incentives to deviate
- Minimized expected dispatch costs across the two markets

How to compute equilibrium regression $\Theta^\star$?

Connection to variational inequalities theory [S+10]



- Equilibrium problem: stacks many private optimization problems
- Variational inequalities (VI): analyzes the interaction between private optimization problems
- In some special cases (like ours), VI connects equilibrium to a centralized optimization

For more details visit Appenix A in: https://arxiv.org/pdf/2405.17753

Computing regression equilibrium: centralized optimization

- By the Symmetry Principle Theorem [FP03], there exists an equivalent opt solving the equilibrium
- ... which happens to minimizes the expected generation and regulation costs
- ... thus enhancing the temporal coordination of day-ahead and real-time markets

$$
\begin{aligned}
\min_{\Theta,\mathbf{p},\mathbf{r}} \quad & \frac{1}{n}\sum_{i=1}^{n} c(\mathbf{p}_i,\mathbf{r}_i) + \gamma \left\|\Theta\varphi_i - \mathbf{w}_i\right\|_2^2 && \text{regularized expected cost} \\
\text{s.to} \quad & \mathbf{1}^\top(\mathbf{p}_i + \Theta\varphi_i - \mathbf{d}) = 0, && \text{day-ahead power balance} \\
& \mathbf{1}^\top(\mathbf{r}_i - \Theta\varphi_i + \mathbf{w}_i) = 0, && \text{real-time power balance} \\
& |\mathbf{F}(\mathbf{p}_i + \Theta\varphi_i - \mathbf{d})| \leqslant \overline{\mathbf{f}}, && \text{day-ahead power flow limit} \\
& |\mathbf{F}(\mathbf{p}_i + \Theta\varphi_i - \mathbf{d}) \\
& \quad + \mathbf{F}(\mathbf{r}_i - \Theta\varphi_i + \mathbf{w}_i)| \leqslant \overline{\mathbf{f}}, && \text{real-time power flow limit} \\
& \underline{\mathbf{p}} \leqslant \mathbf{p}_i + \mathbf{r}_i \leqslant \overline{\mathbf{p}}, && \text{generation limit} \\
& |\mathbf{r}_i| \leqslant \overline{\mathbf{r}}, \quad \forall i = 1,\dots,n, && \text{regulation limit} \\
& |\Theta| \leqslant \tau && \text{equilibrium feature selection}
\end{aligned}
$$

- For more details visit Appenix A in: https://arxiv.org/pdf/2405.17753

Computing regression equilibrium: Walrasian auction and ADMM

Step 1 Primal update of each wind power producer:

$$\boldsymbol{\theta}^{k} \leftarrow \underset{\|\boldsymbol{\theta}\|\leqslant\tau}{\max.} \quad \mathcal{R}^{W}(\boldsymbol{\theta}|\mathcal{D},\boldsymbol{\lambda}_1^k,\boldsymbol{\lambda}_2^k) - \gamma\cdot\mathcal{L}(\boldsymbol{\theta}|\mathcal{D})$$

Step 2 Primal update of each conventional generator:

$$\mathbf{p}^{k},\mathbf{r}^{k} \leftarrow \underset{\mathbf{p},\mathbf{r}\in\mathcal{G}}{\max.} \quad \mathcal{R}^{G}(\mathbf{p},\mathbf{r}\mid\boldsymbol{\lambda}_1^k,\boldsymbol{\lambda}_2^k) - c\,(\mathbf{p},\mathbf{r})$$

Step 3 Electricity price updates:

$$\boldsymbol{\lambda}_1^{k+1} \leftarrow \left[\boldsymbol{\lambda}_1^{k} - \varrho\left(\mathbf{p}^{k} + \mathbb{W}_{\boldsymbol{\theta}^k}(\varphi) - \mathbf{d}\right)\right]_{+}$$

$$\boldsymbol{\lambda}_2^{k+1} \leftarrow \left[\boldsymbol{\lambda}_2^{k} - \varrho\left(\mathbf{r}^{k} + \mathbf{w} - \mathbb{W}_{\boldsymbol{\theta}^k}(\varphi)\right)\right]_{+}$$

- Resembles Walrasian auction: Equilibrium is computed via price exchange
- Proprietary training datasets are localized and not exchanged

Step 1 Primal update of each wind power producer:

$$\theta^k \leftarrow \max_{\|\theta\|\leqslant\tau} \quad \mathcal{R}^W(\theta|\mathcal{D},\boldsymbol{\lambda}_1^k,\boldsymbol{\lambda}_2^k) - \gamma\cdot\mathcal{L}(\theta|\mathcal{D}) \underbrace{+\frac{\varrho}{2}\left\|\mathbf{p}^{k-1}+\mathbb{W}_{\theta}(\varphi)-\mathbf{d}\right\|_2^2+\frac{\varrho}{2}\left\|\mathbf{r}^{k-1}+\mathbf{w}-\mathbb{W}_{\theta}(\varphi)\right\|_2^2}_{\text{ADMM feasibility terms}}$$

Step 2 Primal update of each conventional generator:

$$\mathbf{p}^k,\mathbf{r}^k \leftarrow \max_{\mathbf{p},\mathbf{r}\in\mathcal{G}} \quad \mathcal{R}^G(\mathbf{p},\mathbf{r}\mid\boldsymbol{\lambda}_1^k,\boldsymbol{\lambda}_2^k) - c(\mathbf{p},\mathbf{r}) \underbrace{+\frac{\varrho}{2}\|\mathbf{p}+\mathbb{W}_{\theta^{k-1}}(\varphi)-\mathbf{d}\|_2^2+\frac{\varrho}{2}\|\mathbf{r}+\mathbf{w}-\mathbb{W}_{\theta^{k-1}}(\varphi)\|_2^2}_{\text{ADMM feasibility terms}}$$

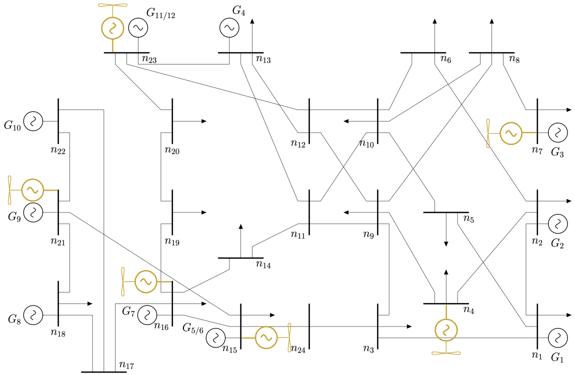
Step 3 Electricity price updates:

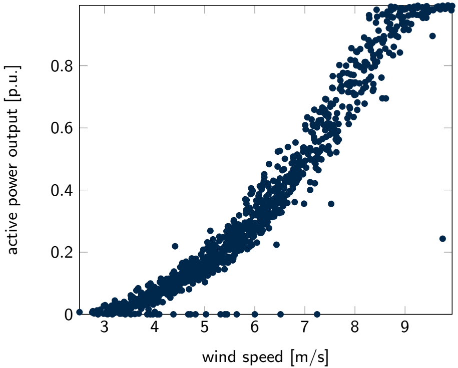
$$\boldsymbol{\lambda}_1^{k+1} \leftarrow \left[\boldsymbol{\lambda}_1^k - \varrho\left(\mathbf{p}^k+\mathbb{W}_{\theta^k}(\varphi)-\mathbf{d}\right)\right]_+$$

$$\boldsymbol{\lambda}_2^{k+1} \leftarrow \left[\boldsymbol{\lambda}_2^k - \varrho\left(\mathbf{r}^k+\mathbf{w}-\mathbb{W}_{\theta^k}(\varphi)\right)\right]_+$$

- Resembles Walrasian auction: Equilibrium is computed via price exchange
- Proprietary training datasets are localized and not exchanged

Experiments on a modified IEEE 24-Bus RTS

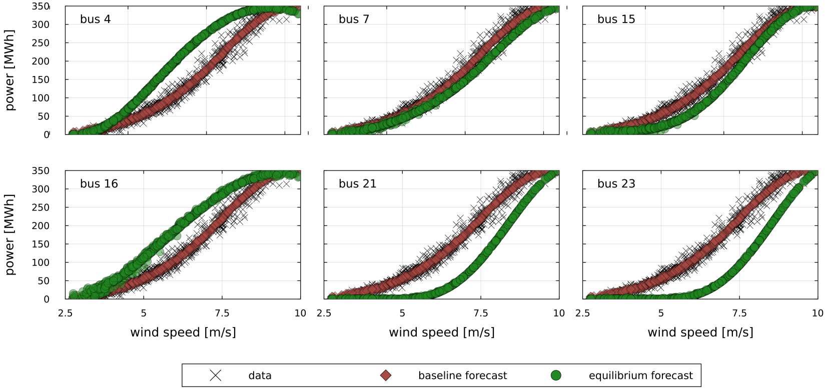




- 6 with farms with identical data and features
- Cover 38.4% of load at peak generation
- Kernel regression with 30 transformed features
- 5, 000 training and 10, 000 testing samples

- Although data is the same, how do equilibrium forecasts depend on the wind farm location in the grid?
- What are the equilibrium benefits in terms profits (any incentives to deviate?) and cost of electricity?

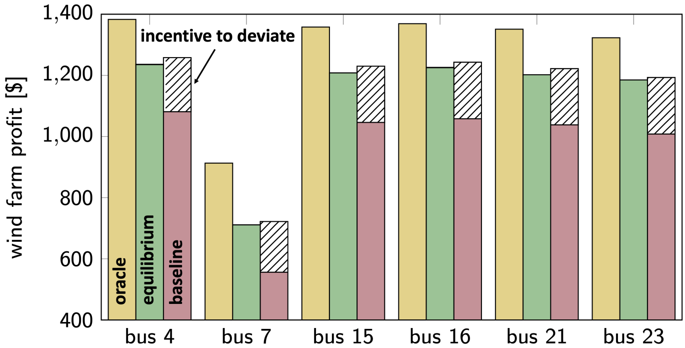
Baseline versus Equilibrium forecasts



- **Baseline:** minimizes a prediction error
- **Equilibrium:** maximizes wind farm profits

Systematic over- or under-prediction depending on the wind farm's location in the grid

Wind farm profits and incentives to deviate



- Equilibrium regression yields larger profits for all wind farms
- There are large profit incentives to unilaterally deviate from the baseline regression
- And (almost) no incentives to deviate from the equilibrium regression

Regression	RMSE, MWh	Average dispatch cost, $			Total dispatch cost error, $	
		total	day-ahead	real-time	average	$\text{CVaR}_{10\%}$
Oracle	———	37,246	37,246	———	———	———
Baseline	88	39,223	37,459	1,764	1,977	8,626
Equilibrium	395	38,326	38,154	172	1,080	3,555

- Baseline regression: minimal forecast error, yet results in large real-time cost
- Equilibrium regression: large forecast errors, withholds cheap generation from the day-ahead market; yet, results in very cheap real-time re-dispatch
- Saving of 2.4% on average, and 13.6% on average across 10% of the worst-case scenarios

Agenda

Introduction

Controlling the impact of ML errors on electricity pricing

Nash equilibrium of ML models in electricity markets

Insights for foundational models in power systems

Supply equilibrium structure to foundation models:

- Current practice: foundational models endowed with power flow physics
- Proposal: integrate market design and equilibrium conditions for market-focused fine-tuning

What foundation models can do:

- Learn from historical market data: is AI adoption converging to equilibrium?
- Track equilibrium in real time: are deviations transient or structural?
- Diagnose non-convergence: which market design features or AI application scopes prevent equilibrium?

ML impacts on pricing

Thank you for your attention!

Regression equilibrium

- Thomas Carriere and George Kariniotakis, *An integrated approach for value-oriented energy forecasting and data-driven decision-making application to renewable energy trading*, IEEE Trans. Smart Grid **10** (2019), no. 6, 6933–6944.
- Wenbo Chen, Mathieu Tanneau, and Pascal Van Hentenryck, *End-to-end feasible optimization proxies for large-scale economic dispatch*, IEEE Transactions on Power Systems (2023).
- Priya Donti, Brandon Amos, and J Zico Kolter, *Task-based end-to-end model learning in stochastic optimization*, Advances in neural information processing systems **30** (2017).
- Vladimir Dvorkin, Stefanos Delikaraoglou, and Juan M Morales, *Setting reserve requirements to approximate the efficiency of the stochastic dispatch*, IEEE Trans. Power Syst. **34** (2018), no. 2, 1524–1536.
- Vladimir Dvorkin and Ferdinando Fioretto, *Price-aware deep learning for electricity markets*, arXiv preprint arXiv:2308.01436 (2023).
- Stefanos Delikaraoglou and Pierre Pinson, *Optimal allocation of hvdc interconnections for exchange of energy and reserve capacity services*, Energy Systems **10** (2019), no. 3, 635–675.

- Yury Dvorkin, *A chance-constrained stochastic electricity market*, IEEE Trans. Power Syst. **35** (2019), no. 4, 2993–3003.
- Vladimir Dvorkin, *Regression equilibrium in electricity markets*, arXiv preprint arXiv:2405.17753 (2024).
- Francisco Facchinei and Jong-Shi Pang, *Finite-dimensional variational inequalities and complementarity problems*, Springer-Verlag New York, Inc., 2003.
- Tue Vissing Jensen, Jalal Kazempour, and Pierre Pinson, *Cost-optimal atcs in zonal electricity markets*, IEEE Trans. Power Syst. **33** (2017), no. 4, 3624–3633.
- Shaohui Liu, Chengyang Wu, and Hao Zhu, *Graph neural networks for learning real-time prices in electricity market*, arXiv preprint arXiv:2106.10529 (2021).
- Juan M Morales et al., *Pricing electricity in pools with wind producers*, IEEE Trans. Power Syst. **27** (2012), no. 3, 1366–1376.
- ______, *Electricity market clearing with improved scheduling of stochastic production*, Eur. J. Oper. Res. **235** (2014), no. 3, 765–774.

- Pierre Pinson, Christophe Chevallier, and George N Kariniotakis, *Trading wind generation from short-term probabilistic forecasts of wind power*, IEEE Trans. Power Syst. **22** (2007), no. 3, 1148–1156.
- Farzaneh Pourahmadi and Jalal Kazempour, *Unit commitment predictor with a performance guarantee: A support vector machine classifier*, IEEE Transactions on Power Systems (2024).
- Geoffrey Pritchard, Golbon Zakeri, and Andrew Philpott, *A single-settlement, energy-only electric power market for unpredictable and intermittent participants*, Oper. Res. **58** (2010), no. 4-part-2, 1210–1219.
- Gesualdo Scutari et al., *Convex optimization, game theory, and variational inequality theory*, IEEE Signal Process. Mag. **27** (2010), no. 3, 35–49.
- Dariush Wahdany, Carlo Schmitt, and Jochen L Cremer, *More than accuracy: end-to-end wind power forecasting that optimises the energy system*, Electric Power Syst. Res. **221** (2023), 109384.
- Yufan Zhang, Jiajun Han, and Yuanyuan Shi, *Risk-aware value-oriented net demand forecasting for virtual power plants*, arXiv preprint arXiv:2406.10434 (2024).