AN ONLINE LEARNING APPROACH TO ALGORITHMIC BIDDING FOR VIRTUAL TRADING

Lang Tong

School of Electrical & Computer Engineering

Cornell University, Ithaca, NY

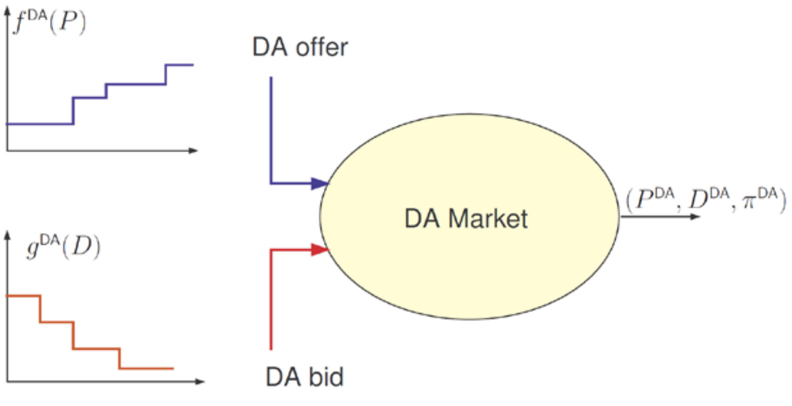
Joint work with Sevi Baltaoglu and Qing Zhao

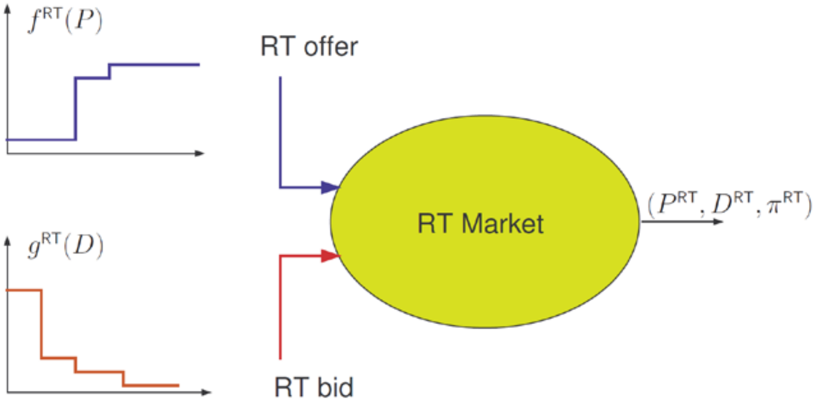
3/27/2018

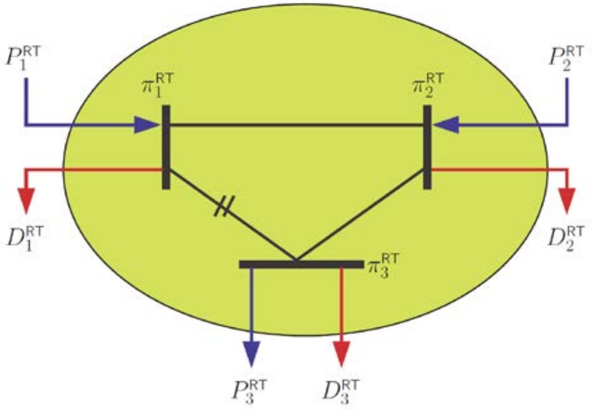
Outline

- Virtual transaction market
  - rational, snap shot statistics, and market mechanism
- Algorithmic bidding
  - bidding, clearing, and settlement models
  - A simple online learning approach
  - Optimal bidding with risk-neutral and risk-averse metrics
- Online learning algorithm
- Tests on real traces
- Related work

Two-settlement market and locational marginal prices





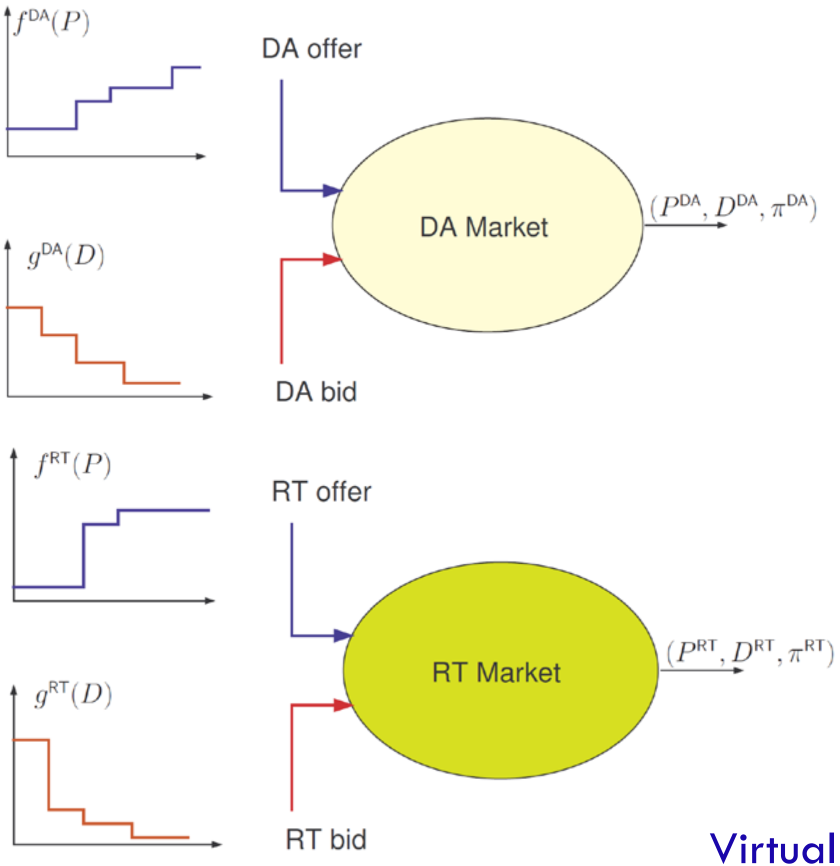


$$\text{DA gen \$} = \pi^{DA} P^{DA}$$

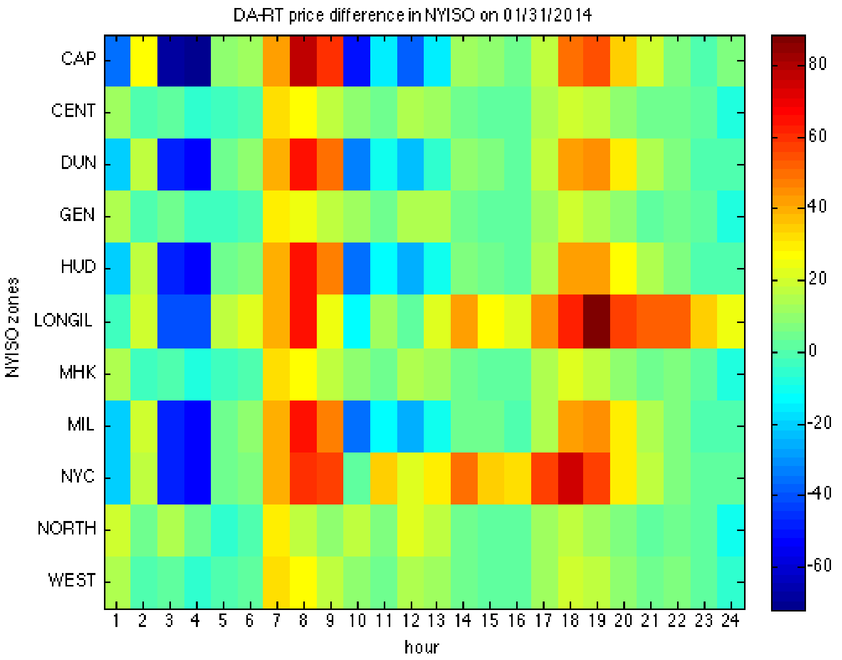
$$\text{RT gen \$} = \pi^{RT}(P^{RT} - P^{DA})$$

$$\text{Total \$} = \pi^{DA} P^{DA} + \pi^{RT}(P^{RT} - P^{DA})$$

LMP spread in time and space



LMP spread $\Delta^{LMP} = \lambda^{DA} - \lambda^{RT}$



Virtual bidding enables arbitraging across time & locations

Virtual bids and virtual transactions

- Virtual bids are submitted and cleared in the DAM and settled in the RTM.
- Virtual transactions are financial transactions. They represent virtual generation, demand, and bilateral scheduling in the DA operation only; they do not exist in the RT operation.
- Three types virtual bids (Q, c):
  - Increment (INC) offer emulates generation offer; it offers to sell Q in the DAM and buy Q in the RTM.
  - Decrement (DEC) bid emulates a demand bid; it bids to buy Q in the DAM and sell Q in the RTM.
  - Up-to-congestion (UTC) bid emulates a bilateral transaction bid; it bids to deliver Q in from source S to destination D.

$$\text{INC profit} = (\lambda^{\text{DA}} - \lambda^{\text{RT}})Q. \\ = \Delta^{\text{LMP}} Q$$

$$\text{DEC profit} = (\lambda^{\text{RT}} - \lambda^{\text{DA}})Q. \\ = -\Delta^{\text{LMP}} Q$$

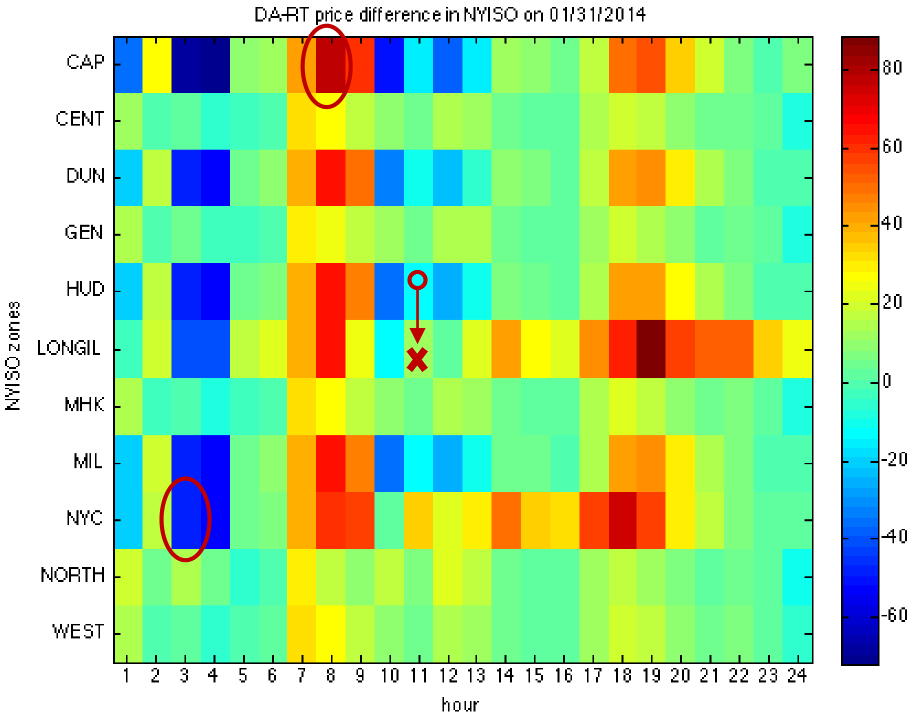
$$\text{UTC profit} = \Big((\lambda_D^{\text{RT}} - \lambda_S^{\text{RT}}) \\ - (\lambda_D^{\text{DA}} - \lambda_S^{\text{DA}})\Big)Q.$$

Virtual bids examples

$$\text{INC profit} = \Delta^{\text{LMP}} Q$$

$$\text{DEC profit} = -\Delta^{\text{LMP}} Q$$

$$\text{UTC profit} = (\Delta_S^{\text{LMP}} - \Delta_D^{\text{LMP}}) Q$$



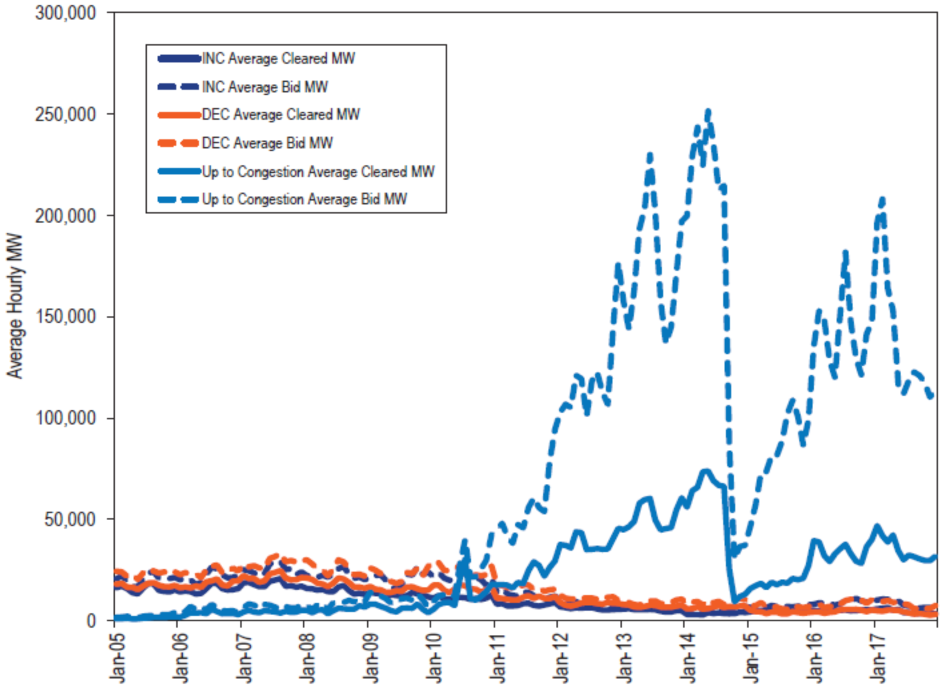
LMP spread $\Delta^{\text{LMP}} = \lambda^{\text{DA}} - \lambda^{\text{RT}}$

- A 10MW INC offer for 8AM CAP, if cleared, generates \$800.
- A 10MW DEC bid for 3AM NYC, if cleared, generates \$600
- A 10MW UTC bid for 1PM HUD to LONGIL, if cleared, generates -\$300.

Only if we know the spread….

Cleared Virtual bids

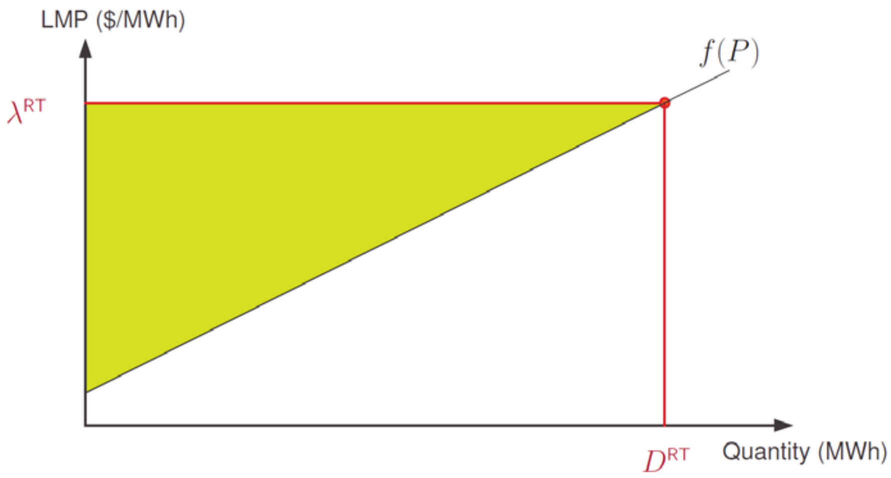
PJM monthly bid and cleared INCs, DECs, and UTCs 2005-2017

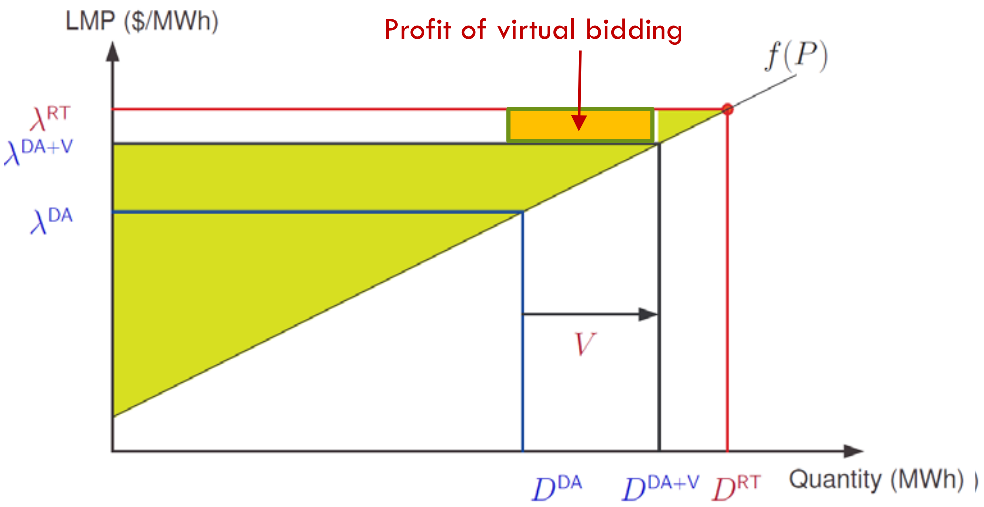


Source: 2018 state of market report for PJM electricity markets

Rationale of virtual transactions

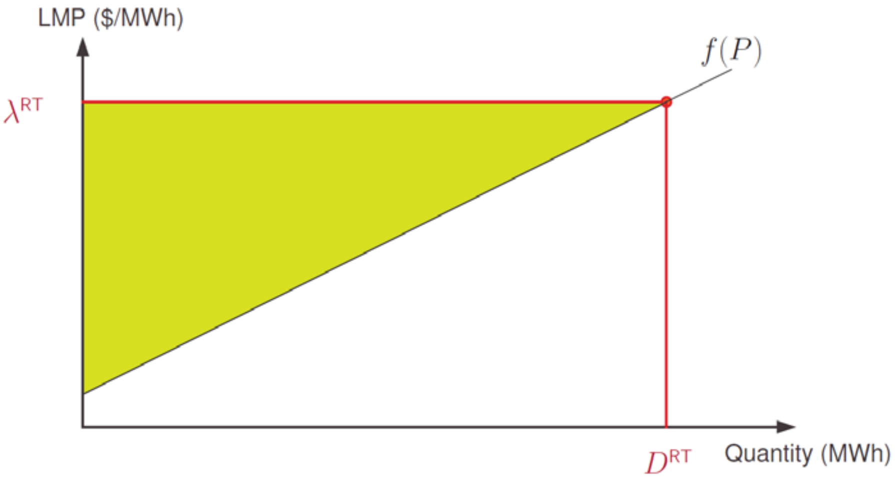
- Enhancing market efficiency by promoting price convergence,
- Mitigating market power through the addition of competitive entities,
- Adding liquidity to the market, and
- Allowing physical participants to hedge against risks.

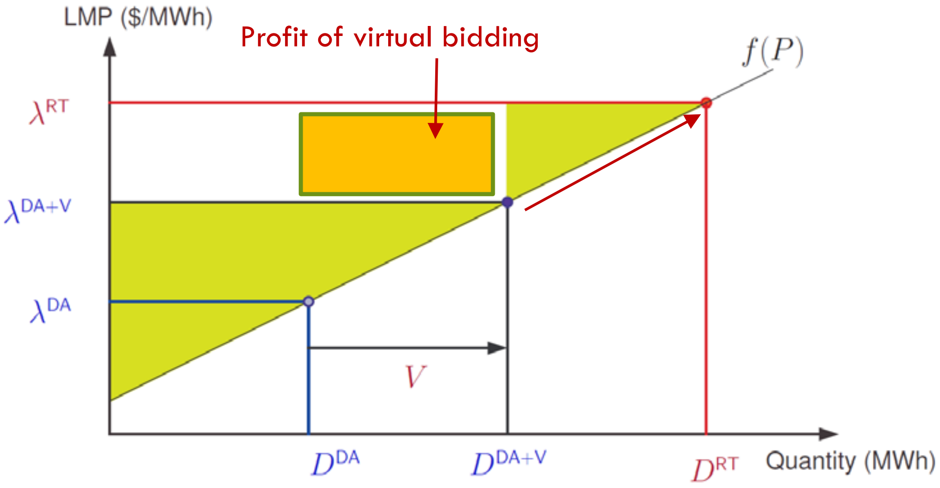




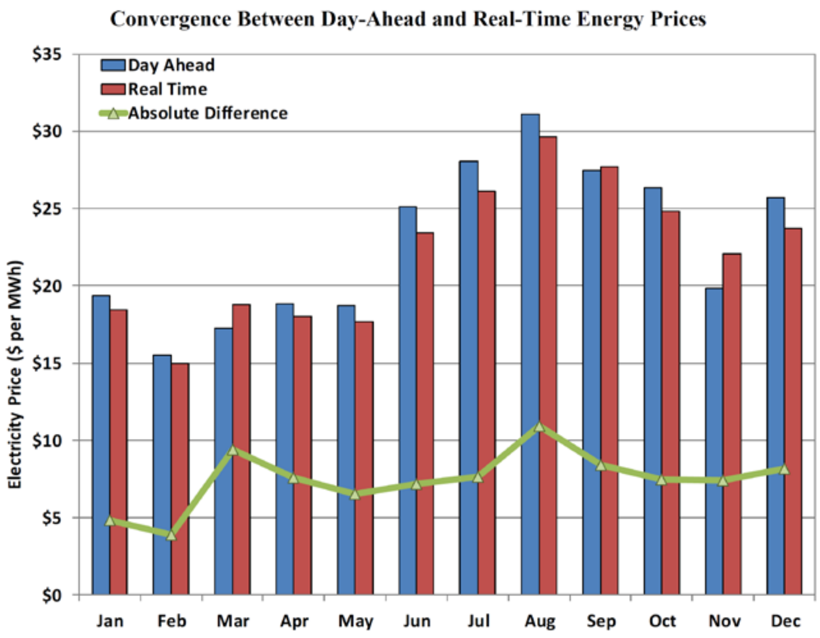
Rationale of virtual transactions

- Enhancing market efficiency by promoting price convergence,
- Mitigating market power through the addition of competitive entities,
- Adding liquidity to the market, and
- Allowing physical participants to hedge against risks.

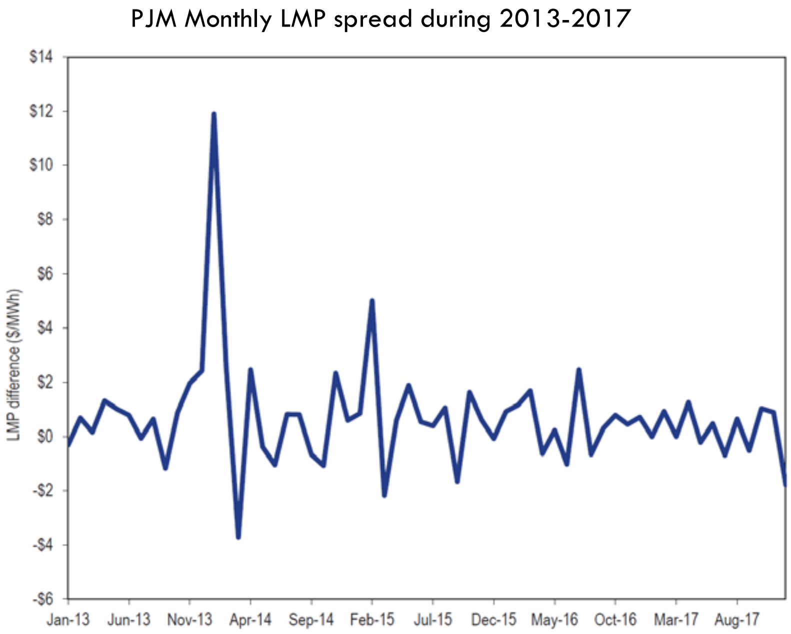




LMP spread traces and statistics



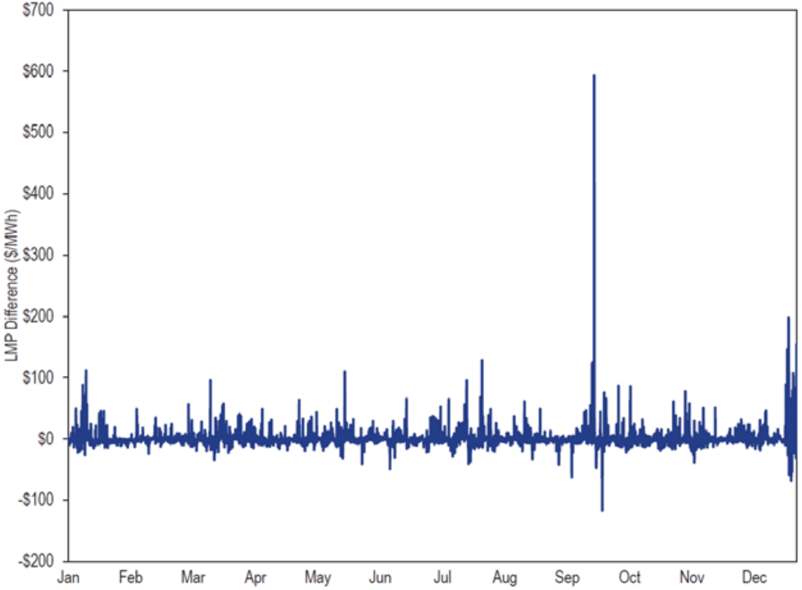
Source: 2016 state of market report for ERCOT electricity markets



Source: 2018 state of market report for PJM electricity markets

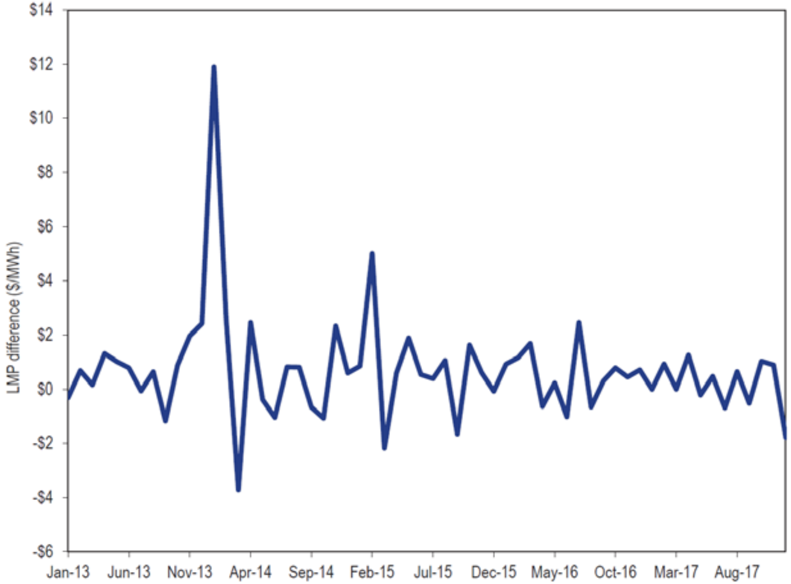
LMP spread traces and statistics

Daily LMP spread in 2017 in PJM



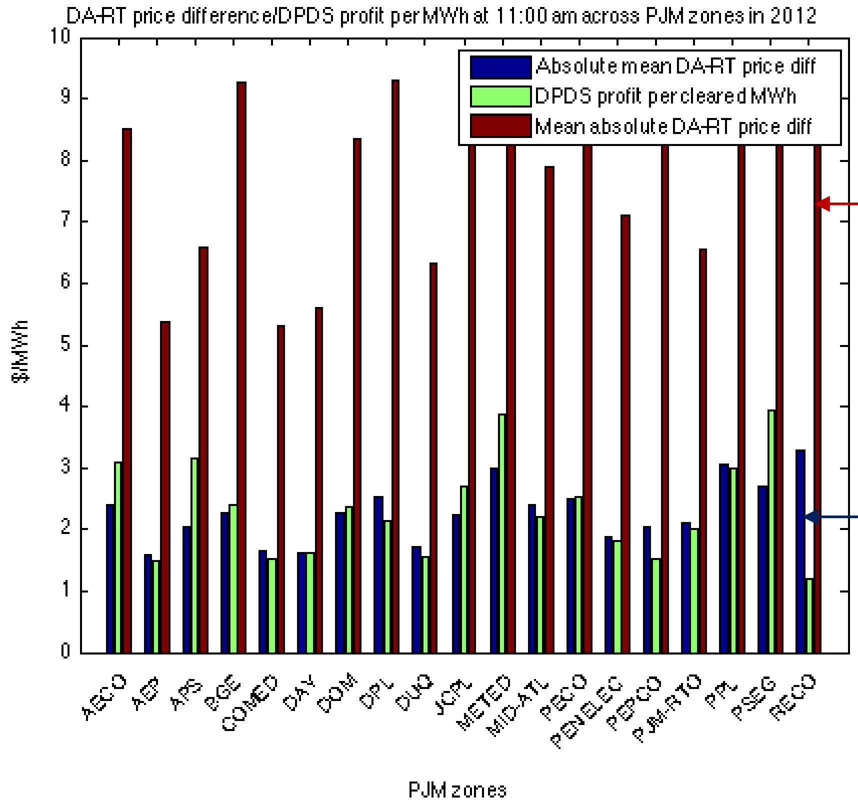
Source: 2018 state of market report for PJM electricity markets

PJM Monthly LMP spread during 2013-2017



Source: 2018 state of market report for PJM electricity markets

Average of the absolute vs. absolute of the average



$\mathbb{E}(|\lambda^{\text{DA}} - \lambda^{\text{RT}}|)$

Upper bound on performance of on-line learning

$|\mathbb{E}(\lambda^{\text{DA}} - \lambda^{\text{RT}})|$

Upper bound on performance of off-line learning

Key challenges in algorithmic virtual bidding

- Large number of trading options with limited budget to explore
  - NYISO allows virtual bids on 11 zones (264 options)
  - PJM allows virtual bids at 1556 different locations (37,344 options)
- Random DA and RT prices with unknown, dependent, and nonstationary distributions.
  - Prefer online learning that dynamically tracks operating conditions.
  - Premium on short-term cumulative reward over asymptotic performance

Bidding model: action space and information structure

- K trading options to bid on
- The vector of bids for day t: $x_t = [x_{t,1}, \ldots, x_{t,K}]^\mathsf{T}$

- The vector of DA prices on day t: $\lambda_t = [\lambda_{t,1}, \ldots, \lambda_{t,K}]^\mathsf{T}$
- The vector of RT prices on day t: $\pi_t = [\pi_{t,1}, \ldots, \pi_{t,K}]^\mathsf{T}$
- Before choosing the bid for day t, $I_{t-1} = \{x_i, \lambda_i, \pi_i\}_{i=1}^{t-1}$.

- **A bidding policy μ:** A sequence of decision rules, *i.e.*, $\mu = (\mu_0, \mu_1 \ldots, \mu_{T-1})$, such that, at time $t-1$,

$$\mu_{t-1} : I_{t-1} \to x_t.$$

Online learning policy and objective

- Payoff from a virtual demand bid:

Bid

Clearing condition

$$(\pi_{t,k} - \lambda_{t,k})\mathbb{1}\{x_{t,k} \geq \lambda_{t,k}\}$$

RTM income

Payment in DAM

- **Objective of the virtual trader:** (Without loss of generality for both INC/DEC)

$$\begin{aligned} \underset{\mu}{\text{maximize}} \quad & \mathbb{E}\left(\sum_{t=1}^{T}(\pi_t - \lambda_t)^{\intercal}\mathbb{1}\{x_t^{\mu} \geq \lambda_t\}\right) \\ \text{subject to} \quad & \|x_t^{\mu}\|_1 \leq B, \qquad \forall t = 1, ..., T, \\ & x_t^{\mu} \geq 0, \qquad \forall t = 1, ..., T, \end{aligned}$$

Empirical "risk" minimization (ERM)

- **Objective of the virtual trader:**

$$\begin{aligned} \underset{\mu}{\text{maximize}} \quad & \mathbb{E}\left(\sum_{t=1}^{T}(\pi_t - \lambda_t)^\intercal \mathbb{1}\{x_t^\mu \geq \lambda_t\}\right) \\ \text{subject to} \quad & \|x_t^\mu\|_1 \leq B, \qquad \forall t = 1, \ldots, T, \\ & x_t^\mu \geq 0, \qquad \forall t = 1, \ldots, T, \end{aligned}$$

Empirical accumulative profit

$$\bar{r}_{t,k}(x_k) = \frac{1}{t}\sum_{i=1}^{t}(\pi_{i,k} - \lambda_{i,k})\mathbb{1}\{x_k \geq \lambda_{i,k}\}.$$

- **ERM objective**:

$$\max_{x \in \mathcal{F}} \bar{r}_t(x) = \max_{x \in \mathcal{F}} \sum_{k=1}^{K} \bar{r}_{t,k}(x_k)$$

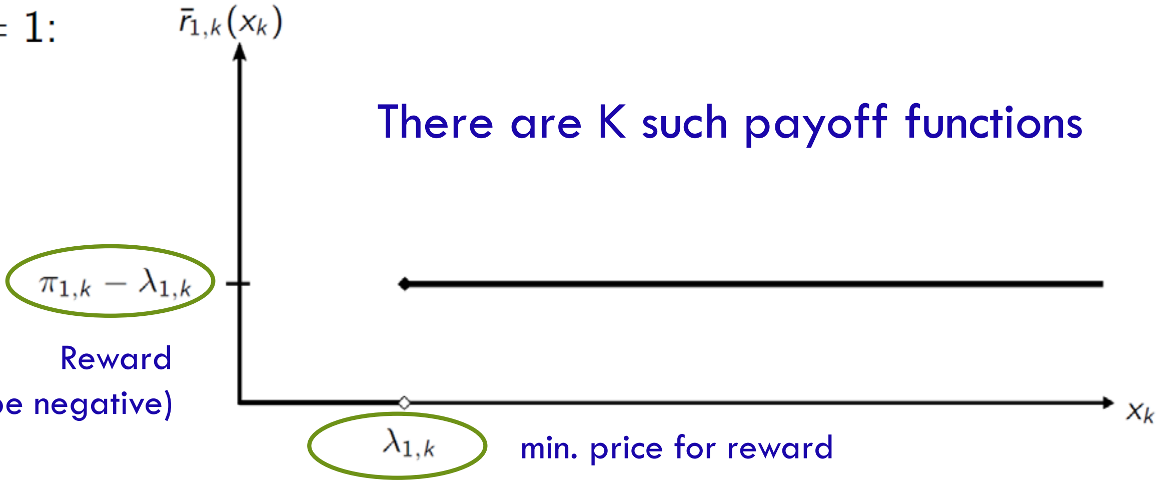
where $\mathcal{F} = \{x \in \Re^K : x \geq 0, \|x\|_1 \leq B\}$.

Empirical risk minimization (ERM)

$$\max_{x \in \mathcal{F}} \bar{r}_t(x) = \max_{x \in \mathcal{F}} \sum_{k=1}^{K} \bar{r}_{t,k}(x_k) \qquad \bar{r}_{t,k}(x_k) = \frac{1}{t} \sum_{i=1}^{t} (\pi_{i,k} - \lambda_{i,k}) \mathbb{1}\{x_k \geq \lambda_{i,k}\}.$$

$$\mathcal{F} = \{x \in \Re^K : x \geq 0, \|x\|_1 \leq B\}.$$

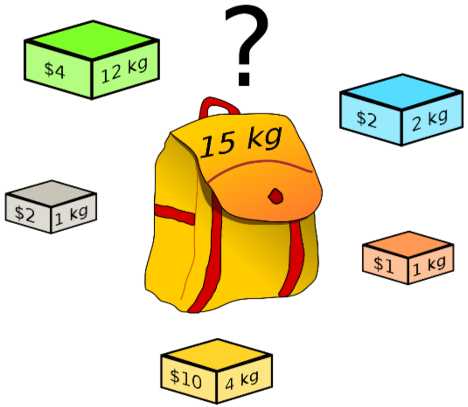
at $t = 1$:



There are K such payoff functions

Reward
(can be negative)

min. price for reward



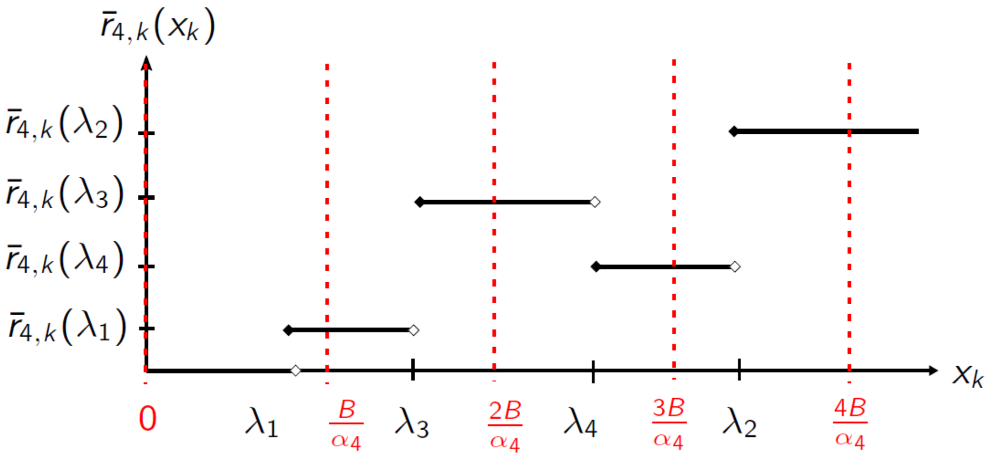
Knapsack problem
(NP hard)

Dynamic programming on discrete set (DPDS)

$$\max_{x_{t+1}\in\mathcal{F}_t} \bar{r}_t(x_{t+1}) = \max_{x_{t+1}\in\mathcal{F}_t} \sum_{k=1}^{K} \bar{r}_{t,k}(x_{t+1,k}). \quad \bar{r}_{t,k}(x_k) = \frac{1}{t}\sum_{i=1}^{t}(\pi_{i,k} - \lambda_{i,k})\mathbb{1}\{x_k \geq \lambda_{i,k}\}.$$

$$\mathcal{F}_t = \{x \in \mathcal{F} : x_k \in \{0, B/\alpha_t, 2B/\alpha_t, ..., B\}, \forall k\}$$

- At $t = 4$:



- Can be solved by DP
- Complexity:
  - Full information $O(KT^2)$
  - Rolling window $O(K)$

From risk-neutral to risk-averse

$$\begin{aligned} \underset{\mu}{\text{maximize}} \quad & \mathbb{E}\left(\sum_{t=1}^{T} (\pi_t - \lambda_t)^\intercal \mathbb{1}\{x_t^\mu \geq \lambda_t\}\right) \\ \text{subject to} \quad & \|x_t^\mu\|_1 \leq B, \qquad \forall t = 1, \dots, T, \\ & x_t^\mu \geq 0, \qquad \forall t = 1, \dots, T, \end{aligned}$$

$$\max_{\mu:\, x_t^\mu \in \mathcal{F} \;\forall\; t} \mathbb{E}\left(\sum_{t=1}^{T} r^{(\rho)}(x_t^\mu)\right)$$

$$r^{(\rho)}(x_t) = \sum_{k=1}^{K} (r_k(x_{t,k}) - \rho v_k(x_{t,k}))$$

Mean return

Variance

Sharpe Ratio:

$$S = \frac{\mathbb{E}(R - R_o)}{\sqrt{\mathrm{Var}(R)}}$$

$$r_k(x_{t,k}) = \mathbb{E}((\pi_{t,k} - \lambda_{t,k})\mathbb{1}\{x_{t,k} \geq \lambda_{t,k}\} | x_{t,k})$$

$$v_k(x_{t,k}) = \mathbb{E}\left(((\pi_{t,k} - \lambda_{t,k})\mathbb{1}\{x_{t,k} \geq \lambda_{t,k}\} - r_k(x_{t,k}))^2 | x_{t,k}\right)$$

Optimality of DPDS (under i.i.d. assumption)

$$\max_{\mu:\, x_t^{\mu} \in \mathcal{F}\ \forall\ t} \mathbb{E}\left(\sum_{t=1}^{T} r^{(\rho)}(x_t^{\mu})\right)$$

- **Optimal bid under known distribution:** $x^* = \arg\max_{x \in \mathcal{F}} r^{(\rho)}(x)$

- **Regret of policy μ:** $\mathcal{R}_T^{\mu}(f) = \sum_{t=1}^{T} \mathbb{E}\left(r^{(\rho)}(x^*) - r^{(\rho)}(x_t^{\mu})\right)$

monotonically increasing for any policy μ.

Optimality of DPDS (under i.i.d. assumption)

Theorem

Let x_{t+1}^{DPDS} denote the bid of DPDS policy for day $t+1$. Under nominal assumptions[2]*, for $t \geq 2$,*

$$\mathbb{E}(r^{(\rho)}(x^*) - r^{(\rho)}(x_{t+1}^{DPDS})) \leq \boxed{C_1\sqrt{\log t/t} + C_2 t^{-1/2}}$$

Regret per day decreases

and for $T > 1$,

$$\mathcal{R}_T^{DPDS}(f) \leq \boxed{C\sqrt{T \log T},}$$

Cumulative regret increases decreases

where C, C_1, and C_2 are positive constants. Furthermore, for any bidding policy μ, there exists a case such that

$$R_T^{\mu}(f) \geq C'\sqrt{T},$$

Matching (almost) lower bound

where C' is a positive constant.

[2] S. Baltaoglu, L. Tong, and Q. Zhao, "Algorithmic Bidding for Virtual Trading in Electricity Markets." Available: https://arxiv.org/abs/1802.03010

Tests on historical data: NYISO & PJM 2006-2016

UCBID-GR (Continuum multi-armed bandits)

- Inspired by UCBID algorithm of Weed et al. [8]
- Starting from the most profitable option, it sets the bid of an option equal to its average RT price until there isn't any sufficient budget left.

SVM-GR (Support vector machine)

- Inspired by the use of SVM by Tang et al.[6] to determine if a demand or a supply bid is profitable for an option by looking at the price spreads occured in previous week.

UCBID-GR (Stochastic approximation)

- A variant of Kiefer-Wolfowitz stochastic approximation method

Cumulative profit

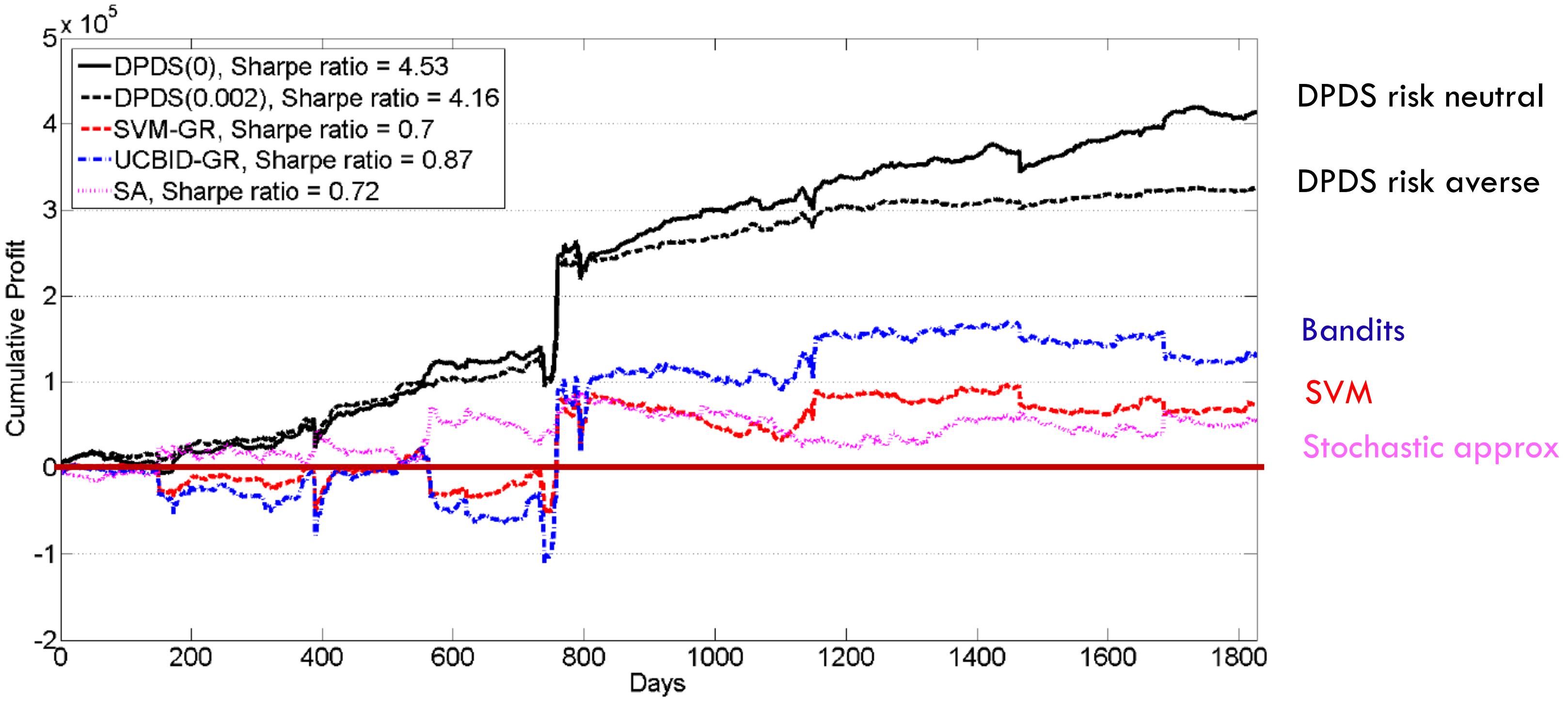


Figure: Cumulative profit trajectory from 2012 to 2016 in NYISO for B=$250,000 after an initial training with 2011 data. (Sharpe Ratio of S&P 500 for the same period is 2.10.)

Annual profit (NYISO)

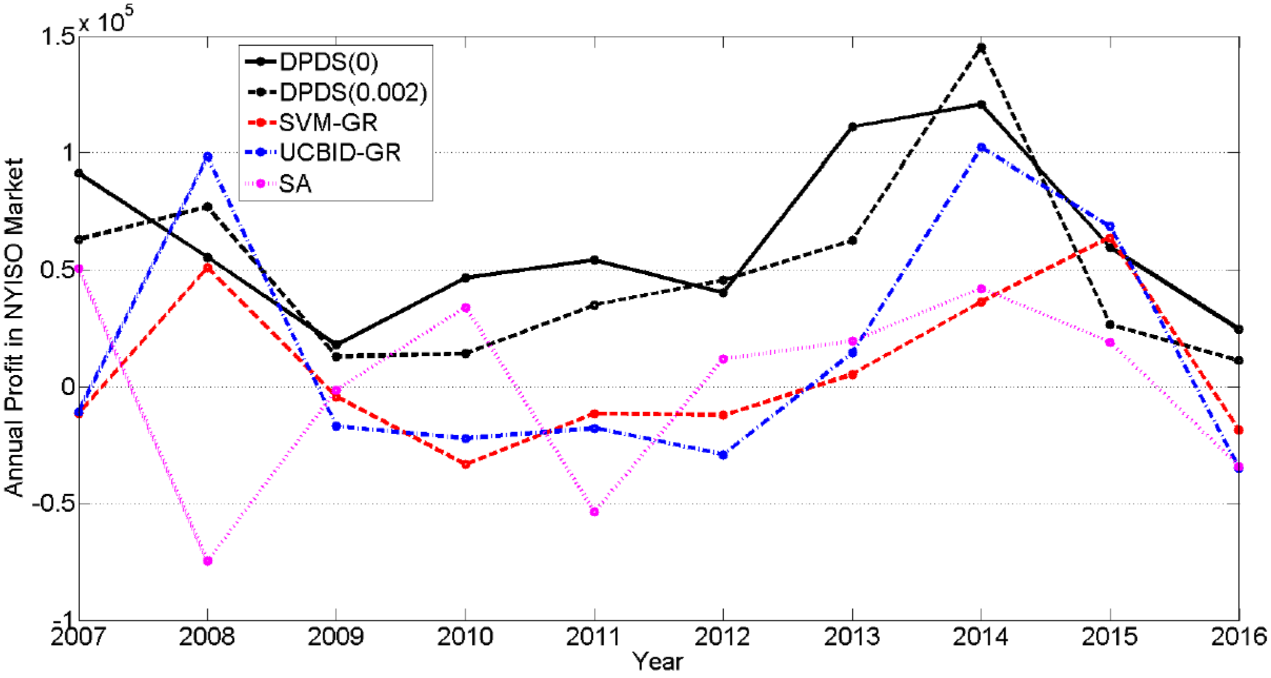


Figure: Annual profit performance in NYISO for $B = \$250,000$ (For each year, an initial training with previous year's data was performed.)

DPDS risk neutral

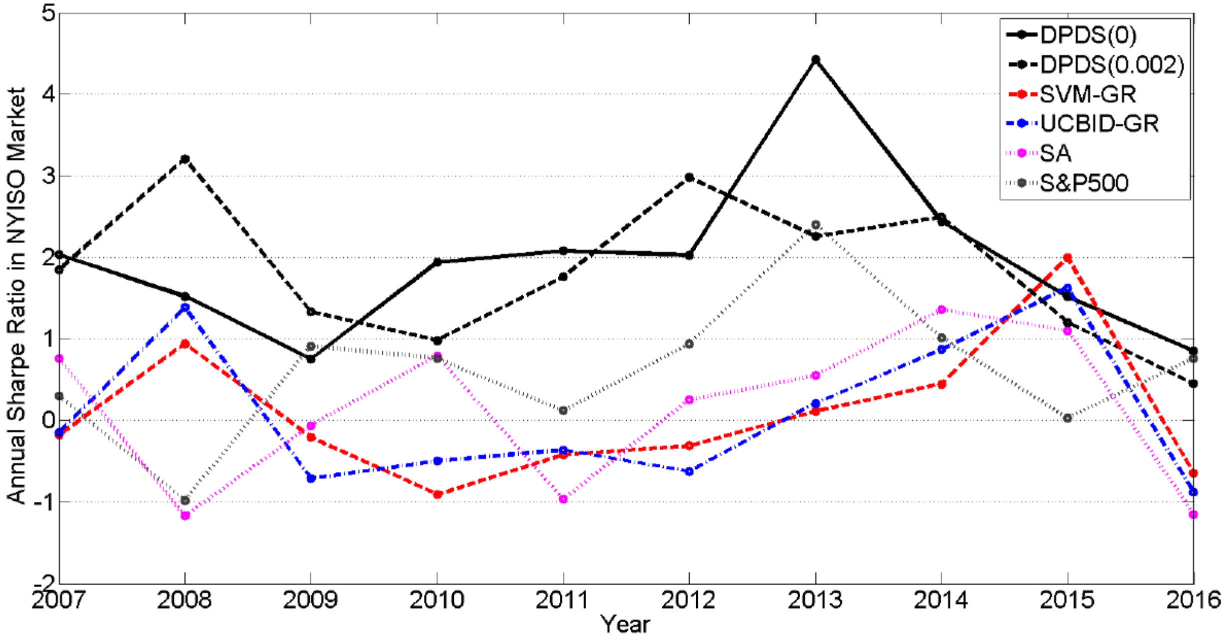
DPDS risk averse

SVM

Bandits

Stochastic approx

Sharpe ratio (NYISO)



DPDS risk neutral
S&P 500

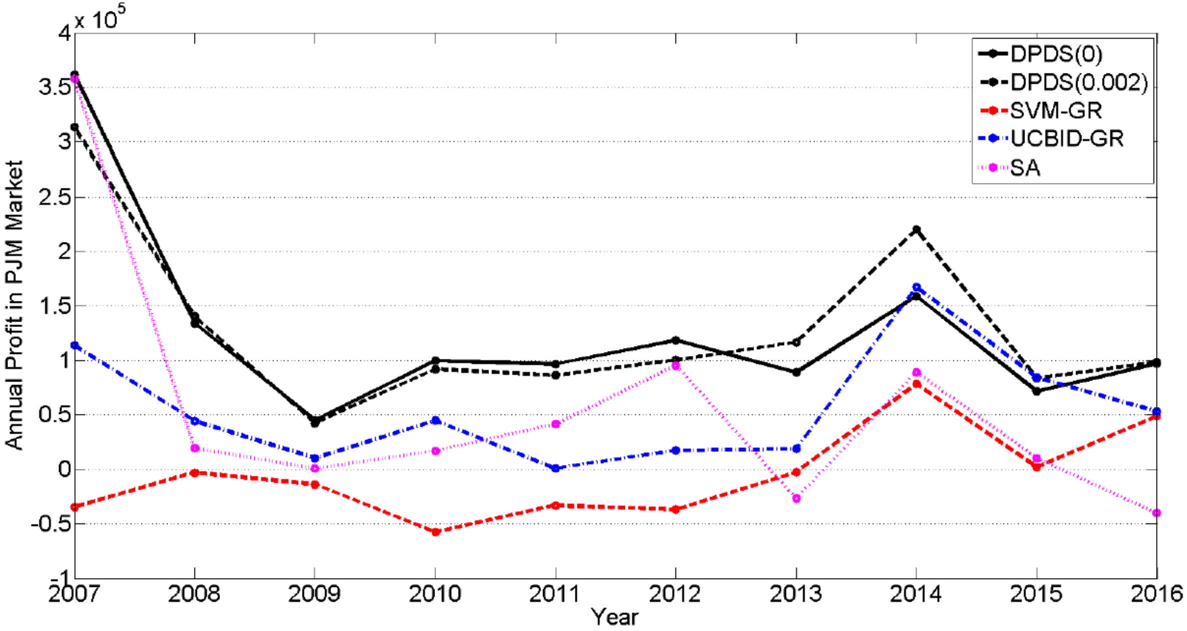
DPDS risk averse

SVM

Bandits
Stochastic approx

Figure: Annual Sharpe ratio performance in NYISO for $B = \$250,000$ (For each year, an initial training with previous year's data was performed.)

Annual profit (PJM)



DPDS risk neutral

DPDS risk averse

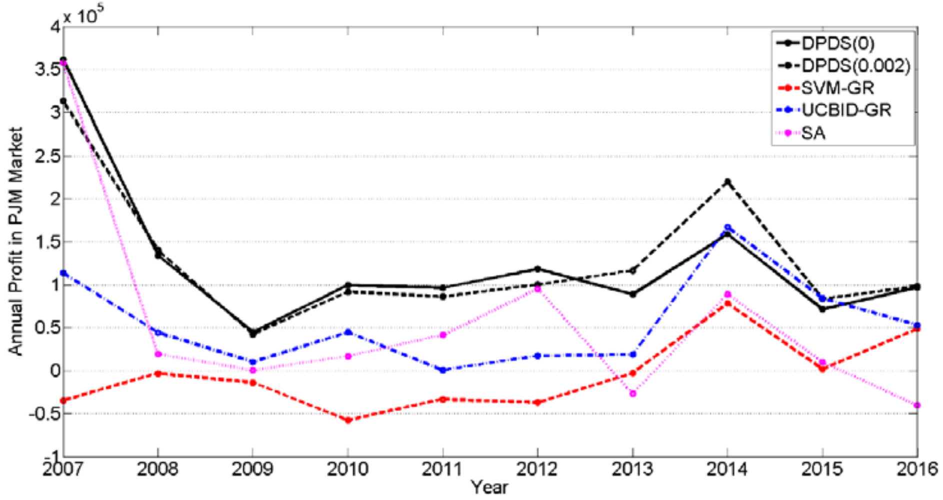
Bandits

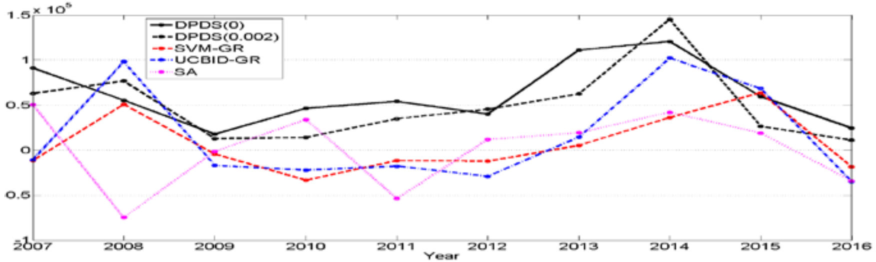
SVM

Stochastic approx

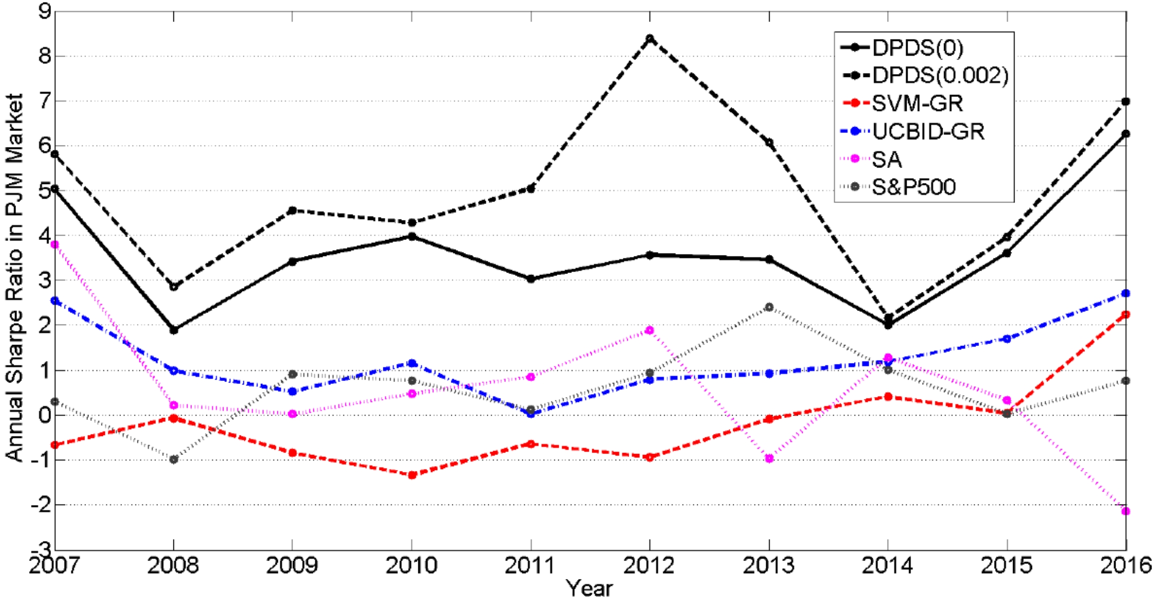
Figure: Annual profit performance in PJM for $B = \$250,000$ (For each year, an initial training with previous year's data was performed.)

Annual profit (PJM vs. NYISO)





Sharpe ratio (PJM)



DPDS risk averse

DPDS risk neutral

Bandits

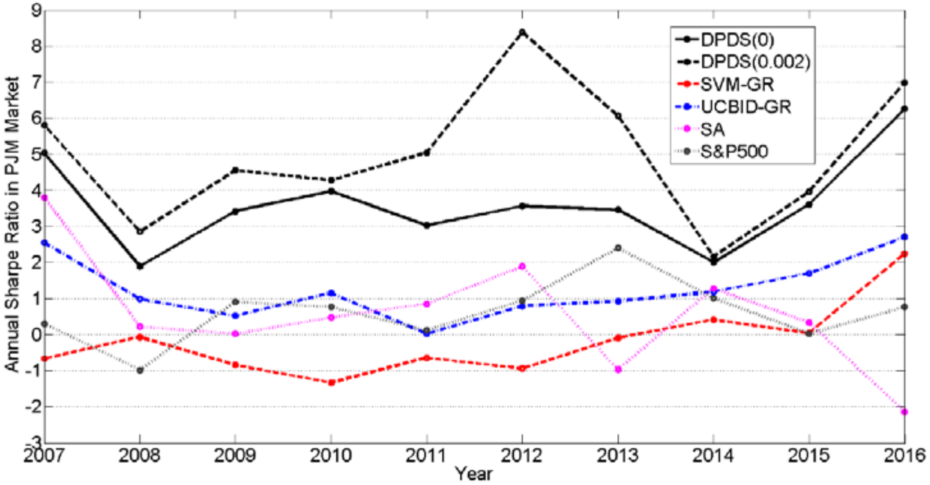
SVM

S&P 500

Stochastic approx

Figure: Annual profit performance in PJM for $B = \$250,000$ (For each year, an initial training with previous year's data was performed.)

Sharpe ratio (PJM vs. NYISO)



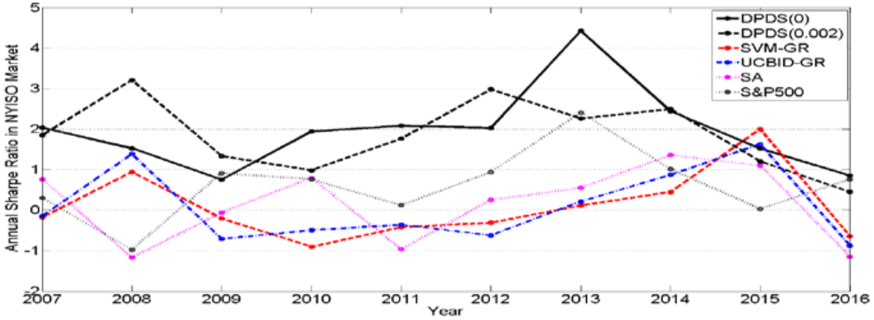
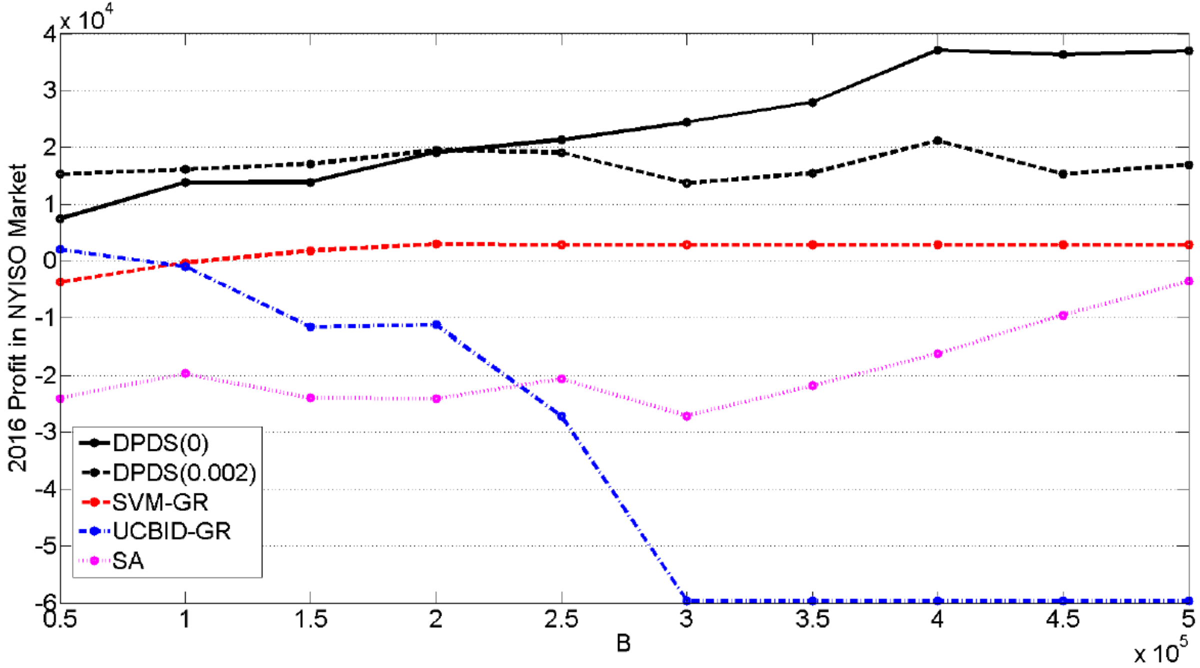


Figure: Annual profit performance in PJM for $B = \$250,000$ (For each year, an initial training with previous year's data was performed.)

Profit vs. budget (NYISO)



DPDS risk neutral

DPDS risk averse

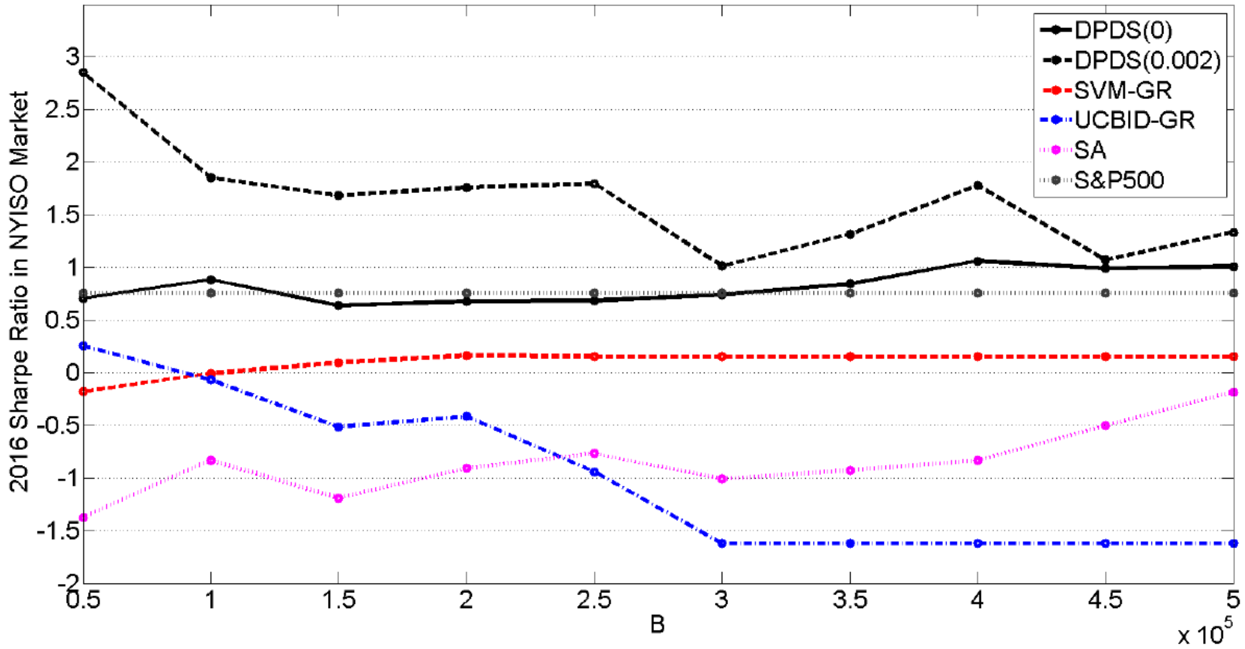
SVM

Stochastic approx

Bandits

Figure: 2016 profit performance in NYISO under different budget levels after an initial training with 2014 and 2015 data

Sharpe ratio vs. budget (NYISO)



DPDS risk averse

DPDS risk neutral

S&P 500

SVM

Stochastic approx

Bandits

Figure: 2016 Sharpe ratio performance in NYISO under different budget levels after an initial training with 2014 and 2015 data

Some relevant literature

1. S. Baltaoglu, L. Tong, Q. Zhao, "Algorithmic bidding for virtual trading in electricity markets," [Online https://arxiv.org/pdf/1802.03010.pdf]
2. J. Weed, V. Perchet, and P. Rigollet, "Online learning in repeated auctions," In 29th Annu. Conf. Learning Theory, 2016.
3. W. Tang, R. Rajagopal, K. Poolla, and P. Varaiya, "Model and data analysis of two-settlement electricity market with virtual, bidding" In IEEE 55th Conf. Decision and Control, 2016.
4. W. Tang, R. Rajagopal, K. Poolla, and P. Varaiya, "Impact of virtual bidding on financial and economic efficiency of wholesale electricity markets" Working paper.
5. J. Mather, E. Bitar, and K. Poolla, "Virtual bidding: Equilibrium, learning, and the wisdom of crowds," IFAC-PapersOnLine, 2017.
6. R. Li, A. J. Svoboda, and S. S. Oren, "Efficiency impact of convergence bidding in the California electricity market," J. Regul. Econ., Dec. 2015.
7. A. Jha and F. A. Wolak, "Testing for market efficiency with transactions costs: An application to convergence bidding in wholesale electricity markets, 2015.
8. J. E. Parsons, C. Colbert, J. Larrieu, T. Martin, and E. Mastrangelo, "Financial arbitrage and ecient dispatch in wholesale electricity markets," MIT Center for Energy and Environmental Policy Research No. 15-002, 2015.

Summary of results

- We consider on algorithmic bidding for virtual trading in electricity markets
- We develop an online learning (polynomial time) technique aimed at dealing with unknown and dependent price distributions for both risk-neutral and risk-averse objectives.
- Dynamic programming on discrete set (DPDS)
  - is order optimal in the growth rate of accumulative payoff, and
  - yields significant profit based on 10yr historical data in NYISO and PJM.
- Empirical results show that both PJM and NYISO are profitable, although PJM market seems to present better opportunities for traders.