GCEP Final Technical Report: Machine Learning, Renewable Energy, and Electricity Markets

Stefan Wager (GSB & Statistics)
Kuang Xu (GSB & Electrical Engineering)

Abstract

Motivated by applications in future smart-grid energy trading markets, we study experimental design in large-scale stochastic systems with substantial uncertainty and structured cross-unit interference. We consider the problem of a platform that seeks to optimize supply-side payments p in a centralized marketplace where different suppliers interact via their effects on the overall supply-demand equilibrium, and propose a class of local experimentation schemes that can be used to optimize these payments without perturbing the overall market equilibrium. We show that, as the system size grows, our scheme can estimate the gradient of the platform's utility with respect to p while perturbing the overall market equilibrium by only a vanishingly small amount. We can then use these gradient estimates to optimize p via any stochastic first-order optimization method. These results stem from the insight that, while the system involves a large number of interacting units, any interference can only be channeled through a small number of key statistics, and this structure allows us to accurately predict feedback effects that arise from global system changes using only information collected while remaining in equilibrium.

Investigators

Stefan Wager, Assistant Professor of Operations, Information & Technology, Stanford Graduate School of Business; Assistant Professor of Statistics (by courtesy)
Kuang Xu, Assistant Professor of Operations, Information & Technology, Stanford Graduate School of Business; Assistant Professor of Electrical Engineering (by courtesy)
Xinkun Nie, PhD Student, Stanford Computer Science
Ruiyang Song, PhD Student, Stanford Electrical Engineering

Introduction

The amount of electricity generated from renewable sources has substantially increased in recent years worldwide. One challenging consequence of this development is that many renewable resources—such as solar and wind—are notoriously volatile, and thus make capacity planning in the electricity market more difficult; see Denholm and Margolis [2007] and Smith et al. [2007] for general discussions. There is considerable interest in using predictive modeling and marketplace incentives to mitigate the volatility of renewable energy sources. One lesson from the tech sector is that, in order to successfully integrate such quantitative tools into renewable energy markets, we should expect to have to run experiments to collect empirical evidence on what works and what doesn't. The difficulty, however, is that in an interconnected marketplace, any incentives given to one market participant may alter their behavior in a way that affects other participants also, thus breaking standard approaches to experimentation that rely on the lack of interference. The goal of this research was to develop methods for experimental design that are robust to rich system-wide interference that is bound to occur in energy markets.

Randomized controlled trials are widely used to guide decision making across several areas, ranging from classical industrial and agricultural settings [Fisher, 1935] to the modern tech sector [Athey and Luca, 2019, Kohavi et al., 2009, Tang et al., 2010]. Much of the existing work on experimental design, however, has focused on settings where we can intervene separately on different units, i.e., there is no cross-unit interference, and this lack of interference plays a key role in justifying standard analyses of randomized trials [Imbens and Rubin, 2015], and the question of how to run experiments when simple A/B tests fail has proven to be challenging. Some approaches have been proposed that assume sparse interference patterns. For example, when studying Internet ad auctions, Basse et al. [2016], Kohavi et al. [2009] and Ostrovsky and Schwarz [2011] note that the auction type used for one ad keyword does not meaningfully affect how advertisers bid for other keywords, and then consider experiments that randomly assign keywords, rather than advertisers, to different conditions. Similarly, a social network might try to deploy different versions of a feature in different countries, and hope that the number of cross-border links is small enough to induce only negligible interference. The limitation of these approaches, however, is that the power of any experiment is limited by the number of non-interfering clusters available: For example, if a platform has 200 million customers in 100 countries, but chooses to randomize by country, then the largest effective sample size they can use for any experiment is 100, and not 200 million.

The focus of our research is an alternative approach to experimentation in stochastic systems, where a large number of, if not all, units interfere with one another. For concreteness, we focus on the problem of setting supply side payments in a centralized marketplace, where available demand is randomly allocated to a set of available suppliers. In these systems, different suppliers interact via their effects on the overall supply-demand equilibrium: the more suppliers choose to participate in the marketplace, the less demand on average an individual supplier would be able to serve in equilibrium. The objective of the system designer is to identify the optimal payment that maximizes the platform's utility. Note that conventional randomized experimentation schemes that assume no interference fail in this system: For example, if we double the per-transaction payments made to a random half of suppliers, these suppliers will increase their production levels and reduce the amount of demand available to the remaining suppliers, and thus reduce their incentives to produce.

We consider a simple model of such a centralized marketplace, and design a class of "local" experimentation schemes that—by carefully leveraging the structure of the marketplace—enable us to optimize payments without perturbing the overall market equilibrium. More specifically, we perturb the per-transaction payment p_i available to the i-th supplier by a small mean-zero shock, i.e., $p_i = p + \zeta\varepsilon_i$ where $0 < \zeta \ll 1$ and $\varepsilon_i = \pm 1$ independently and uniformly at random. Then, in the limit where the number of suppliers is large, we show that we can estimate the gradient of the platform's utility with respect to p while perturbing the overall market equilibrium by only a vanishingly small amount. We can then use these gradient estimates to optimize p via any stochastic first-order optimization method, such as stochastic gradient descent and its extensions.

The driving insight behind our result is that, although there is dependence across the behavior of a large number of units in the system, any such interference can only be channeled through a small number of key statistics: in our example, the total supply made available by all suppliers. Then, if we can intervene on individual units without meaningfully affecting the key statistics, we can obtain meaningful information about the system at a cost that scales sub-linearly in the number of units.

Our model can be applied to studying the operations and optimization of a future whole-sale market or smart-grid energy-sharing platform that connects a large network of interconnected generators (conventional or renewable) with buyers. In this system, prior to the settlement period (e.g., a day-ahead market) the market operator announces a per-unit payment for each unit of energy to be distributed. Given this knowledge, each generator makes a decision on whether or not it wishes to participate in the market, which may carry a generator-dependent opportunity cost. During the settlement period, the actual demand and renewable generation at the generators are realized, and energy distributed according to a certain matching mechanism. Crucially, the amount of energy that each generator will be able to sell will depend on the random realization of the demand, as well as on the total number of active generators available. As such, the participation decision of any individual generator has a global impact on the market dynamics in equilibrium that will subsequently influence all other generators. The local randomization method developed in our work allows the operators to efficiently identify the optimal per-unit payment, by sequentially perturbing the payments offered across the generators in a symmetric manner, and using their resulting participation behavior to derive an accurate estimate of the gradient for the overall profit function with respect to the payment.

Background

The problem of experimental design under interference has received considerable attention in the statistics literature. The dominant paradigm has focused on robustness to interference, and on defining estimands in settings where some units may be exposed to spillovers from treating other units [Aronow and Samii, 2017, Athey et al., 2018, Basse et al., 2019, Eckles et al., 2017, Hudgens and Halloran, 2008, Manski, 2013, Sobel, 2006]. Depending on applications, the exposure patterns may be simple (e.g., the units are clustered such that exposure effects are contained within clusters) or more complicated (e.g., the units are connected in a network, and two units far from each other in graph distance are not exposed to each others' treatments). Unlike this line of work that seeks robustness to interference driven by potentially complex and unknown mechanisms, the local randomization scheme proposed here crucially relies on having a stochastic model that lets us explain interference. Then, because all inference acts via a simple statistic, we can move beyond simply seeking robustness to interference and can in fact accurately predict interference effect using information gathered in equilibrium.

The idea that one can distill insights of a structural model down to the relationship between a small number of observable statistics has a long tradition in economics [Chetty, 2009], and can often be used for practical counterfactual analysis without needing to fit complicated structural models. Here, we use such an argument for experimental design rather than to guide methods for observational study analysis.

Our approach to optimizing p using gradients obtained from local experimentation intersects with the literature on noisy zeroth-order optimization [e.g., Spall, 2005], which aims to optimize a function $f(x)$ by sequentially evaluating f at points $x_1, x_2, \ldots$, and obtaining in return noisy versions of the function values $f(x_1), f(x_2), \ldots$ A number of zeroth-order optimization methods first generate noisy gradient estimates of the function by comparing adjacent function values, and subsequently use these estimates in a first-order optimization method [Jamieson et al., 2012, Duchi et al., 2012, Ghadimi and Lan, 2013, Nesterov and Spokoiny, 2017]. In our model, this approach would amount to estimating utility gradients via global experimentation, by comparing the empirical utilities observed at two different payment levels. Compared to this literature, our paper exploits a cross-sectional structure not present in most existing zeroth-order models: We show that our local experimentation approach, which offers slightly different payments across a large number of units, is far more efficient at estimating the gradient than global experimentation, which offers all units the same payment on a given day. Notably, local experimentation is beneficial even if the final objective is to identify a single optimal payment for all units. Such cross-sectional signals would be lost if we abstracted away the multiplicity of units, and only treated the average payment as a decision variable to be optimized.

The limiting regime that we use, one in which the system size tends to infinity, is often known as the mean-field limit. It has a long history in the study of large-scale stochastic systems, such as the many-server regime in queueing networks [Halfin and Whitt, 1981, Vvedenskaya et al., 1996, Bramson et al., 2012, Tsitsiklis and Xu, 2012, Stolyar, 2015] and interacting particle systems [Mézard et al., 1987, Sznitman, 1991, Graham and Méléard, 1994]. Likewise, our proposed method leverages a key property of the mean-field limit: While changes to the behavior of a single unit may have significant impact on other units in a finite system, such interference diminishes as the system size grows and, in the limit, the behaviors among any finite set of units become asymptotically independent from one another, a phenomenon known as the propagation of chaos [Sznitman, 1991, Graham and Méléard, 1994, Bramson et al., 2012]. This asymptotic independence property underpins the effectiveness of our local experimentation scheme, and ensures that small, symmetric payment perturbations do not drastically alter the equilibrium demand-supply dynamics. Our work thus suggests that, just as mean-field models have been successful in the analysis of stochastic systems, they may be a useful paradigm for designing experiments in large stochastic systems.

Results

For conciseness, we present our main results in the context of a simple setting inspired by a centralized marketplace for freelance suppliers that operates over a number of periods. In each period, the high-level objective of the decision maker (i.e., operator of the platform) is to match demand with a pool of potential suppliers in such a manner that maximizes the platform's expected utility. To do so, the decision maker offers payments to each potential supplier individually, who in turn decides whether to become active/available based upon their belief of future revenue. Our main question is how the decision maker can use experimentation efficiently to discover their revenue-maximizing payment, despite not knowing the detailed parameterization of the model, and the presence of substantial stochastic uncertainty.

Here, we discuss an elementary variant of our model that, despite its simplicity, lets us highlight some key properties of our approach; a more flexible stochastic model for which our results hold is described in a working paper, available from the authors. Each day $k = 1, ..., K$ there are $i = 1, ..., n$ potential suppliers, and demand for D_k identical tasks to be accomplished. A central platform offers a payment p_{ik} to each supplier that they can earn by servicing a unit of demand. The suppliers can accurately anticipate demand D_k (e.g., via their knowledge of local weather or events); however, the platform may not be able to. Given their knowledge of p_{ik} and D_k, each supplier independently chooses to become "active"; we write $Z_{ik} = 1$ for active suppliers and $Z_{ik} = 0$ else. Then, demand D_k is randomly allocated to active suppliers (where each active supplier can service at most one unit of demand).

Our key assumption is that each supplier chooses to become active based on their expected revenue conditionally on being active: They first compute $q_{D_k}(p_{\cdot k})$, their allocation rate (rate at which they will be matched with demand) conditionally on being active (which also depends on the payments offered to other supplies), and then take expected revenue to be $p_{ik} q_{D_k}(p_{\cdot k})$. For now, we assume the following simple functional form

$$\mathbb{P}\left[Z_{ik} \,\middle|\, p_{\cdot k},\, D_k\right] = \frac{1}{1 + e^{-\beta\left(p_{ik} q_{D_k}(p_{\cdot k}) - B_i\right)}}, \tag{1}$$

where B_i measures the value of a random supplier-specific outside option. This specific functional form does not matter for our results—the important point is suppliers only interact via $q_D(p)$. Finally, assume that the utility of the platform is as follows (later, we will relax the functional form of R),

$$U = R(T,\, D) - \bar{p} \min\left\{T,\, D\right\}, \quad R(T,\, D) = CD\left(1 - e^{-D/T}\right), \tag{2}$$

where $T = \sum_{i=1}^{n} Z_i$ is the total number of active suppliers, $\bar{p}$ is the average payment made to suppliers who service a unit of demand, and C is a scaling constant.

Figure 1 shows a simple example of an equilibrium resulting from this model in the limit as n gets large in a setting where all suppliers are offered the same price p, for a specific realization of demand D; here $D = 0.3n$, $C = 100$, and $\log(B_i/20)$ has a standard Gaussian distribution. Overall, we see that the expected fraction of active suppliers $\mu_D(p) = \mathbb{E}_p\left[T/n \,\middle|\, D\right]$ increases with p, whereas $q_D(p)$ decreases in p. There is a noticeable kink in the slope of μ_D once q_D falls below 1 due to equilibrium effects.

In order to optimize payments p, the platform needs to understand the expected utility function $\chi(p) = \mathbb{E}_p\left[U/n\right]$ well enough to find its maximum. Before presenting our proposal, we briefly outline two baselines that are often discussed in practice.

- **Global experimentation:** On each day $k = 1, ..., K$, we randomly choose a payment p_k that is made available to all suppliers. We then observe utility U_k, which acts as a noisy estimate of $\chi(p_k)$. Given sufficient time periods, this approach will consistently find the maximizer p^* of $\chi(\cdot)$; however, the cost of experimentation is substantial, as we may end up repeatedly offering poorly chosen payments p_k to all suppliers in the marketplace.

- **Classical A/B experimentation:** On each day $k = 1, ..., K$, we choose a small random fraction of suppliers and offer them a payment p_k, while everyone else gets offered the status quo payment p_0. We could then try to use the behavior of suppliers offered p_k to estimate $\chi(p_k)$. This approach allows for cheap experimentation because most of the suppliers get offered p_0. However, it will not consistently recover the optimal payment p^* because it ignores feedback effects: When we raise payments, more suppliers opt to join the market and so the rate at which any given supplier is matched with demand goes down—and this attenuates the payment-sensitivity of supply relative to what is predicted by A/B testing.

Our goal is to use high-level information about the stochastic system described above to design a new experimental framework that lets us avoid the problems of both approaches described above: We want our experimental scheme to be consistent for p^* like global experimentation, but also to be cost-effective (like classical A/B testing) in that it only requires small perturbations to the status quo.

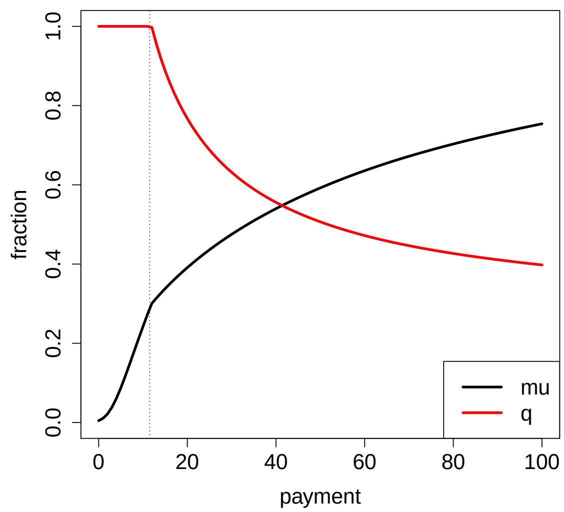

Fig. 1: Example of large-sample behavior of market, conditionally on a realization of demand D. Here $\mu_D(p)$ represents the that a random supplier decides to become active at p, whereas $q_D(p)$ is the expected allocation rate across active suppliers.

Estimating Utility Gradients in Equilibrium The driving insight behind our approach is that it is possible to learn about $\chi(p)$ via unobstrusive randomization by randomly perturbing the prices p_{ik} offered to supplier i in time period k: We propose setting

$$p_{ik} = p_k + \zeta \varepsilon_{ik}, \qquad \varepsilon_{ik} \overset{\text{iid}}{\sim} \{\pm 1\} \tag{3}$$

uniformly at random, where $\zeta > 0$ is a (small) constant that governs the magnitude of the perturbations. Then, by regressing market participation Z_{ik} on the price perturbations ε_{ik}, we can estimate the average local payment sensitivity of suppliers ignoring feedback effects [Imbens and Angrist, 1994]

$$\Delta_D(p) = \left\{ \frac{d}{dp'} \mathbb{E}\left[\frac{1}{1 + \exp\left(-\beta\left(p' q_D(p) - B_i\right)\right)} \,\middle|\, D \right] \right\}_{p'=p}. \tag{4}$$

This quantity Δ is not directly of interest for optimizing p, as it ignores feedback effects. However, it turns out that Δ captures relevant information for optimizing p and, given our generic model structure, we find that as the system size n grows

$$\lim_{n\to\infty} \frac{d\mu_D(p)}{dp} - \left(1 - \frac{d\mu_D(p)}{dp} \frac{p}{\mu_D(p)} 1\left(\left\{\frac{D}{n} < \mu_D(p)\right\}\right)\right) \Delta_D(p) = 0, \tag{5}$$

where all terms in (5) other than Δ and the actual average sensitivity $d\mu_D(p)/dp$ (which accounts for feedback) are readily observable from the data. The upshot is that, once we can estimate Δ via local randomization (3), we can also get estimates of $d\mu_D(p)/dp$ by solving (5), and finally obtain noisy $\widehat{\Gamma}$ gradients for $d\chi(p)/dp$. A formal statement is available in our working paper.

We can thus use unobstrusive randomization as in (3) to estimate the gradient $d\chi(p)/dp$ around any chosen price p. The final step of our proposed approach, **local experimentation**, is to use these gradients to guide any first-order optimization method. For example, one could use gradient descent and, at each time period k estimate $d\chi(p_k)/dp_k$ as $\widehat{\Gamma}_k$, and then set $p_{k+1} = p_k + \eta\widehat{\Gamma}_k$. In our experiments, we use variant of stochastic gradient descent called adagrad [Duchi et al., 2011]; in principle, one could also use more sophisticated methods that rely on acceleration [Cohen et al., 2018].

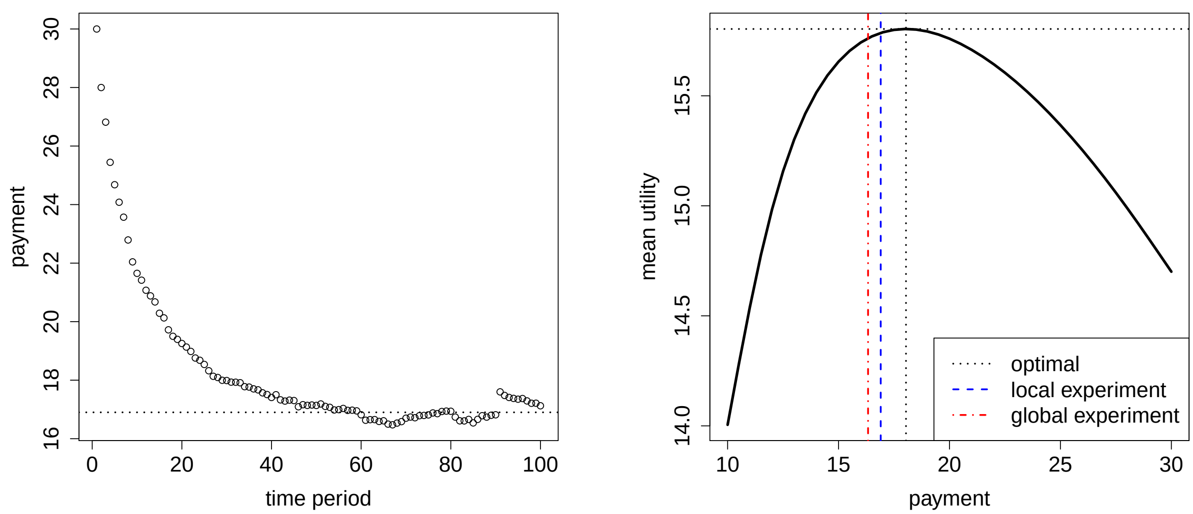

Fig. 2: Results from learning p via local experimentation. The left panel shows the evolution of the p_k over time. The right panel compares the value of p_k averaged over the last 50 steps of our algorithm to the value of p that optimizes mean scaled utility, and a payment $\hat{p}$ learned via global experimentation.

Figure 2 shows results on a simple simulation experiment in the setting of Figure 1, where the scaled demand D/n follows a beta(15, 35) distribution. We initialize the system at $p_1 = 30$, and then each day run price perturbations as in (3) to guide a price update using adagrad. We see that the system quickly converges to a near-optimal price of around 17.

We also compare our results to what one could obtain using the baseline of global experimentation, where we randomize the price $p_k \sim \text{Uniform}(10, 30)$ in each time step and measure resulting utility U_k, and then choose the final payment $\hat{p}$ by maximizing a smooth estimate of the expectation of U_k given p_k. The left panel of Figure 3 shows the resulting (p_k, U_k) pairs, as well as the resulting $\hat{p}$. As seen in the right panel of Figure 2, the final $\hat{p}$ obtained via this method is a reasonable estimate of the optimal p. However, the price of experimentation incurred for finding this $\hat{p}$ is huge: As shown in the right panel of Figure 3, after the first few days, the global experimentation approach systematically achieves lower daily utilities U_k than local experimentation because it often uses very poor choices of p_k.

In principle, it may be possible to design a smarter version of global experimentation than the random search pursued above; for example, it may be possible to leverage recent advances in Bayesian optimization to rule out poor choices of p early on [Letham et al., 2018]. However, designing a variant of global experimentation that is competitive with our approach when n is large appears to be a challenging task. For example, any direct alternative to our method that seeks to estimate the gradient $d\chi(p)/dp$ via global experimentation would need to undertake perturbations that scales linearly in n in order to get any finite error bounds (or, equivalently, the average per-supplier perturbation would need to be of constant order as n gets large). In contrast, local experimentation allows for per-supplier perturbations that scale only a little slower than $1/\sqrt{n}$. Thus, any approach to global experimentation that seeks to estimate $d\chi(p)/dp$ will require unboundedly larger perturbations than our approach as the number of suppliers n gets large.

Conclusions

In order reduce global greenhouse gas emissions, it is important for us to make good use of available resources for renewable energy production. Many sources of renewable energy, however, are controlled by independent parties and face substantial uncertainty both on the supply and demand side, thus making optimal coordi-

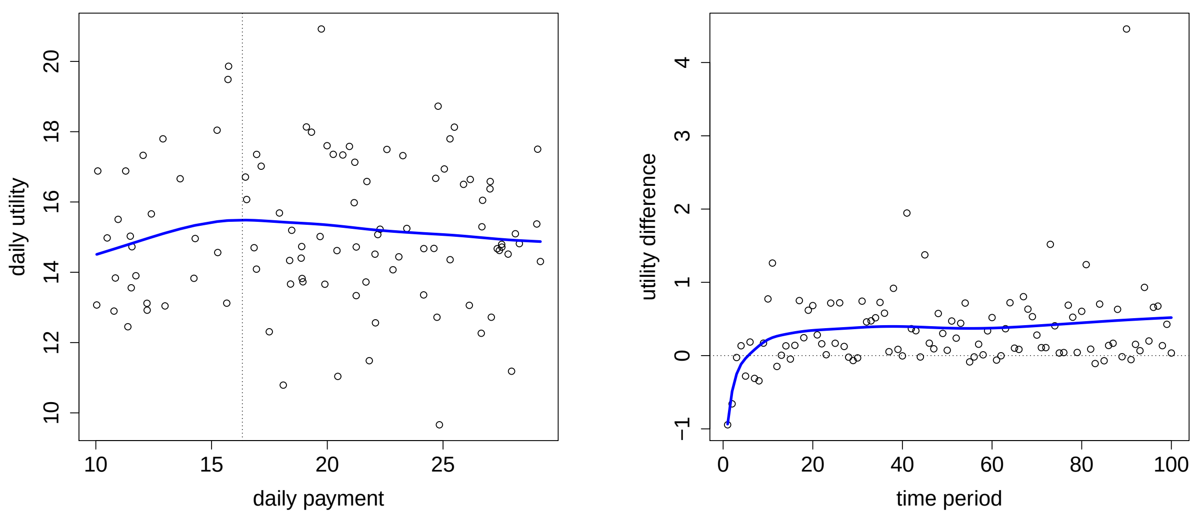

Fig. 3: Results from learning p via global experimentation. The left panel shows pairs (p_k, U_k) resulting from daily experiments, along with an estimate of the optimal p. The right panel shows the (scaled) difference in daily utility between our local experimentation approach and the global experimentation baseline (both approaches worked using the same demand sequence D_k).

nation of these resources challenging. The outcome of our research was a method for experimentation that will enable us to design marketplace incentives that encourage individual energy producers act in a way that is aligned with overall social welfare considerations. As the complexity of renewable energy markets grows, we expect such tools to become ever more important in enabling us to learn how to design efficient energy markets.

We have thus far been focusing on a one-shot model where the matching and settlement occurs in a single time period. An important next step in our program is to consider systems with sequential stochastic dynamics, such as a network of renewable generators that employ large-scale on-site or grid energy storage. The problem faced by the system designer becomes substantially more complex in the sequential setting, shifting from solving a static stochastic optimization problem to that of a state-dependent Markov decision problem (MDP). As a first step, we have been working on using black-box architectures to speedup solving large-scale MDPs motivated by renewable energy management with battery storage. We have devised an algorithm that decomposes a large MDP into smaller sub-problems, and show that it achieves a high competitive ratio for all MDP with a small diameter.

Upcoming Presentations

June 2019: Stanford Workshop on Marketplace Innovation (Stanford, CA)
July 2019: Applied Probability Society Conference (Brisbane, Australia)
September 2019: From Controlled Trials to Big Data and Back (Lausanne, Switzerland)
October 2019: INFORMS (Seattle, WA)

We are preparing a working paper for submission to a leading journal in operations research.

References

Peter M Aronow and Cyrus Samii. Estimating average causal effects under general interference, with application to a social network experiment. *The Annals of Applied Statistics*, 11(4):1912–1947, 2017.

Susan Athey and Michael Luca. Economists (and economics) in tech companies. *Journal of Economic Perspectives*, 33(1):209–30, 2019.

Susan Athey, Dean Eckles, and Guido W Imbens. Exact p-values for network interference. *Journal of the American Statistical Association*, 113(521):230–240, 2018.

Guillaume W Basse, Hossein Azari Soufiani, and Diane Lambert. Randomization and the pernicious effects of limited budgets on auction experiments. In *Artificial Intelligence and Statistics*, pages 1412–1420, 2016.

Guillaume W Basse, Avi Feller, and Panos Toulis. Randomization tests of causal effects under interference. *Biometrika*, 2019.

Maury Bramson, Yi Lu, and Balaji Prabhakar. Asymptotic independence of queues under randomized load balancing. *Queueing Systems*, 71(3):247–292, 2012.

Raj Chetty. Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods. *Annual Review of Economics*, 1(1):451–488, 2009.

Michael B Cohen, Jelena Diakonikolas, and Lorenzo Orecchia. On acceleration with noise-corrupted gradients. *arXiv preprint arXiv:1805.12591*, 2018.

Paul Denholm and Robert M Margolis. Evaluating the limits of solar photovoltaics (pv) in traditional electric power systems. *Energy policy*, 35(5):2852–2861, 2007.

John Duchi, Elad Hazan, and Yoram Singer. Adaptive subgradient methods for online learning and stochastic optimization. *Journal of Machine Learning Research*, 12(Jul):2121–2159, 2011.

John C Duchi, Peter L Bartlett, and Martin J Wainwright. Randomized smoothing for stochastic optimization. *SIAM Journal on Optimization*, 22(2):674–701, 2012.

Dean Eckles, Brian Karrer, and Johan Ugander. Design and analysis of experiments in networks: Reducing bias from interference. *Journal of Causal Inference*, 5(1), 2017.

Ronald A Fisher. *The Design of Experiments.* Oliver and Boyd, Edinburgh, 1935.

Saeed Ghadimi and Guanghui Lan. Stochastic first-and zeroth-order methods for nonconvex stochastic programming. *SIAM Journal on Optimization*, 23(4):2341–2368, 2013.

Carl Graham and Sylvie Méléard. Chaos hypothesis for a system interacting through shared resources. *Probability Theory and Related Fields*, 100(2):157–174, 1994.

Shlomo Halfin and Ward Whitt. Heavy-traffic limits for queues with many exponential servers. *Operations research*, 29(3):567–588, 1981.

Michael G Hudgens and M Elizabeth Halloran. Toward causal inference with interference. *Journal of the American Statistical Association*, 103(482):832–842, 2008.

Guido W. Imbens and Joshua D. Angrist. Identification and estimation of local average treatment effects. *Econometrica*, 62(2):467–475, 1994.

Guido W Imbens and Donald B Rubin. *Causal Inference in Statistics, Social, and Biomedical Sciences.* Cambridge University Press, 2015.

Kevin G Jamieson, Robert Nowak, and Ben Recht. Query complexity of derivative-free optimization. In *Advances in Neural Information Processing Systems*, pages 2672–2680, 2012.

Ramesh Johari, Pete Koomen, Leonid Pekelis, and David Walsh. Peeking at a/b tests: Why it matters, and what to do about it. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1517–1525. ACM, 2017.

Ron Kohavi, Roger Longbotham, Dan Sommerfield, and Randal M Henne. Controlled experiments on the web: survey and practical guide. *Data mining and knowledge discovery*, 18(1):140–181, 2009.

Benjamin Letham, Brian Karrer, Guilherme Ottoni, and Eytan Bakshy. Constrained Bayesian optimization with noisy experiments. *Bayesian Analysis*, 2018.

Charles F Manski. Identification of treatment response with social interactions. *The Econometrics Journal*, 16(1):S1–S23, 2013.

Marc Mézard, Giorgio Parisi, and Miguel Virasoro. *Spin glass theory and beyond: An Introduction to the Replica Method and Its Applications*, volume 9. World Scientific Publishing Company, 1987.

Yurii Nesterov and Vladimir Spokoiny. Random gradient-free minimization of convex functions. *Foundations of Computational Mathematics*, 17(2):527–566, 2017.

Michael Ostrovsky and Michael Schwarz. Reserve prices in internet advertising auctions: a field experiment. *EC*, 11:59–60, 2011.

J Charles Smith, Michael R Milligan, Edgar A DeMeo, and Brian Parsons. Utility wind integration and operating impact state of the art. *IEEE transactions on power systems*, 22(3):900–908, 2007.

Michael E Sobel. What do randomized studies of housing mobility demonstrate? causal inference in the face of interference. *Journal of the American Statistical Association*, 101(476):1398–1407, 2006.

James C Spall. *Introduction to stochastic search and optimization: estimation, simulation, and control*, volume 65. John Wiley & Sons, 2005.

Alexander L Stolyar. Pull-based load distribution in large-scale heterogeneous service systems. *Queueing Systems*, 80(4):341–361, 2015.

Alain-Sol Sznitman. Topics in propagation of chaos. In *Ecole d'été de probabilités de Saint-Flour XIX1989*, pages 165–251. Springer, 1991.

Diane Tang, Ashish Agarwal, Deirdre O'Brien, and Mike Meyer. Overlapping experiment infrastructure: More, better, faster experimentation. In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 17–26. ACM, 2010.

John N Tsitsiklis and Kuang Xu. On the power of (even a little) resource pooling. *Stochastic Systems*, 2(1): 1–66, 2012.

Nikita Dmitrievna Vvedenskaya, Roland L'vovich Dobrushin, and Fridrikh Izrailevich Karpelevich. Queueing system with selection of the shortest of two queues: An asymptotic approach. *Problemy Peredachi Informatsii*, 32(1):20–34, 1996.

Contact

Stefan Wager: `swager@stanford.edu`
Kuang Xu: `kuangxu@stanford.edu`
Xinkun Nie: `xinkun@stanford.edu`
Ruiyang Song: `ruiyangs@stanford.edu`