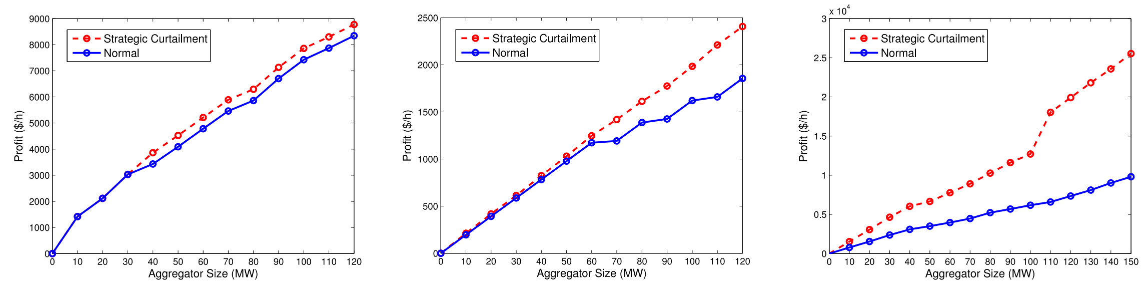


Fig. 3. The profit under the normal (no-curtailment) condition and under (optimal) strategic curtailment, as a function of size of the aggregator in IEEE test case networks: a) IEEE 14-Bus Case, b) IEEE 30-Bus Case, and c) IEEE 57-Bus Case. The difference between the two curves is the curtailment profit.

flow, and the real-time LMPs are 20.0, 25.0, 25.0, 35.0, 28.7, 24.0 $\$/MWh$, respectively.

Assume that the aggregator owns node 1 and aims to increase its profit by curtailing the generation at this node. It can be seen that by curtailing just 0.15 MW generation at node 1 (i.e., from 375.20 MW to 375.05 MW), the binding/non-binding constraints in problem (1) change, and as a result the ISO will determine the new LMPs as 25.8, 25.0, 25.0, 35.0, 30.6, 24.0 $\$/MWh$. Fig. 2 shows the LMPs, before and after the curtailment. In this case, the curtailment profit is $\gamma = 25.8 \times 375.05 - 20 \times 375.20 = 2172\ {}^{\$}/_{h}$, which means that the aggregator has been able to increase its profit by 2172 $\$/h$ during that dispatch interval.

B. Case Studies

We simulate the behavior of aggregators with different sizes, i.e., different number of buses, in a number of different networks. We use the IEEE 14-, 30-, and 57-bus test cases. Since studying market manipulation makes sense only when there is congestion in the network, we scale the demand (or equivalently the line flow limits) until there is some congestion in the network. In order to examine the profit and market power of aggregator as a function of its size, we assume that the way aggregator grows is by sequentially adding random buses to its set (more or less like the way, e.g., a solar firm grows). Then at any fixed set of buses, it can choose different curtailment strategies to maximize its profit. In other words, for each of its nodes it should decide whether to curtail or not (assuming that the amount of curtailment has been fixed to a small portion). We assume that the total generation of the aggregator in each bus is 10 MW and it is able to curtail 1% of it (0.1 MW).

For each of the three networks, Fig. 3 shows the profit for a random sequence of nodes. Comparing the no-curtailment profit with the strategic-curtailment profit reveals an interesting phenomenon. As the size of the aggregator (number of its buses) grows, not only does the profit increase (which is expected), but also the difference between the two curves increase, which is the "curtailment profit." More specifically, the latter does not need to happen in theory. However in practice, it is observed most of the time, and it highlights that larger aggregators have higher incentive to behave strategically, and they can indeed gain more from curtailment.

The other important question is what is the impact of strategic curtailment on the price of each bus of the network (not necessarily just the aggregator's buses). This is important in many scenarios like the effect of such coordinated manipulations on consumers or the effect of competing firms on each other. Fig. 4 shows a heat map of an aggregator's impact on the prices in the IEEE 14-bus network. As one can see, the price of other buses can often be highly impacted as well.

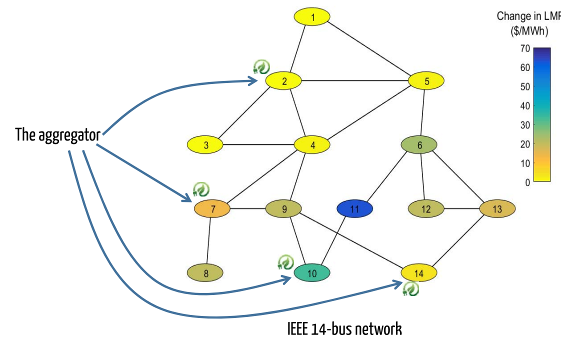


Fig. 4. A heat map of the impact of coordinated curtailment on the prices in the IEEE 14-bus network. Aggregator nodes are 2, 7, 10, and 14.

V. Optimizing Curtailment Profit

The aggregator's profit maximization problem is challenging to analyze, as one would expect given its bilevel form. In fact, bilevel linear programming is NP-hard to approximate up to any constant multiplicative factor in general [47]. Furthermore, the objective of the program (5) is quadratic (bilinear) in the variables, rather than linear. This combination of difficulties means that we cannot hope to provide a complete analytic characterization of the behavior of a profit maximizing aggregator.

In this section, we begin with the case of a single-bus aggregator and build to the case of general multi-bus aggregators in acyclic networks. For the single-bus aggregator, the optimal curtailment can be found exactly, in polynomial time. For the general case, we cannot provide an exact algorithm, but we do provide a practical approximation algorithm for general multi-bus aggregators in acyclic networks (e.g., distribution networks).

A. An Exact Algorithm for Single-Node Aggregators in Arbitrary Networks

Even in the simplest case, when the aggregator has only a single node, i.e., its entire generation is located in a

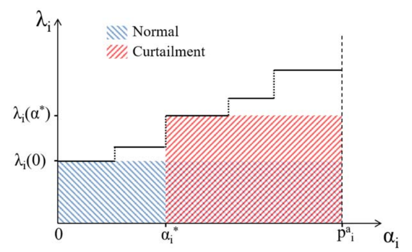


Fig. 5. The LMP at bus i as a function of curtailed generation at that bus. Shaded areas indicate the aggregator's revenue at the normal condition and at the curtailment.

single bus, it is not trivial how to solve the aggregator's profit maximization problem.

The first step toward solving the problem is already difficult. In particular, in order to understand the effect of curtailment on the profit, we first need to understand how does curtailment impact the prices – an impact which is not monotonic in general. Although LMPs are not monotonic in general, it turns out that in single-bus curtailment, the LMP is indeed monotonic with respect to the curtailment. The proof of the following lemma is in Appendix B.

Lemma 1: The LMP of any bus i is monotonically increasing with respect to the curtailment at that bus. That is

$$\lambda_i(\alpha') \geq \lambda_i(\alpha)$$

if $\alpha'_i > \alpha_i$, and $\alpha'_j = \alpha_j$ for all $j = [n]\backslash\{i\}$.

A consequence of the above lemma is that the price λ_i is a monotonically increasing staircase function of α_i, for any bus i, as depicted in Fig. 5. As α_i increases, if the binding constraints of (1) do not change, the dual variables remain the same, and thus the LMPs remain the same (constant intervals). Once a constraint becomes binding/non-binding, the LMP jumps to the next level.

In Fig. 5, the two shaded areas show profit at the normal condition and at the curtailment. The difference between the two areas is the curtailment profit. In particular, if the red area is larger than the blue one, the aggregator is able to earn a positive curtailment profit on bus i. The optimal curtailment α_i^* also happens where the red area is maximized. It should be clear that the optimal curtailment always happens at the verge of a price change, not in the middle of a constant interval (otherwise it can be increased by curtailing less).

Given the knowledge of the network and state estimates, it is possible to find the jump points (i.e., where the binding constraints change) and evaluate them for profitability. Therefore, if there are not too many jumps, an exhaustive search over the jump points can yield the optimal curtailment. Based on this observation, we have the following theorem, which is proven in Appendix C.

Theorem 1: The exact optimal curtailment for an aggregator with a single bus, in an arbitrary network with t lines, can be found by an algorithm with running time $O(t^{3.373})$.

Clearly, this approach does not extend to large multi-bus aggregators. The following section uses a different and more sophisticated algorithmic approach for that setting.

B. An Approximation Algorithm for Multi-Bus Aggregators in Radial Networks

In this section, we show that the aggregator profit maximization problem, while hard in general, can be solved in an approximate sense to determine an approximately-feasible approximately-optimal curtailment strategy in polynomial time using an approach based on dynamic programming. In particular, we show that an ϵ-approximation of the optimal curtailment profit can be obtained using an algorithm with running time that is linear in the size of the network and polynomial in $\frac{1}{\epsilon}$.

Before we state the main result of this section, we introduce the notion of an approximate solution to (5) in the following definition.

Definition 3: A solution $(\alpha, f, \lambda^-, \lambda^+, \mu^-, \mu^+, \nu)$ to (5) is an ϵ-**accurate solution** if the constraints are violated by at most ϵ and $\gamma(\alpha) \geq \gamma^* - \epsilon$.

Note that, if one is simply interested in approximating γ^* (as a market regulator would be), the ϵ-constraint violation is of no consequence, and an ϵ-accurate solution of (5) suffices to compute an ϵ-approximation to γ^*.

Given the above notion of approximation, our main theorem is as follows (proof in Appendix D):

Theorem 2: An ϵ-accurate solution to the optimal aggregator curtailment problem (5) for an n-bus radial network can be found by an algorithm with running time $cn\left(\frac{1}{\epsilon}\right)^9$ where c is a constant that depends on the parameters $p_i^a, B, d, p, \underline{f}, \overline{f}$. On a linear (feeder line) network, the running time reduces to $cn\left(\frac{1}{\epsilon}\right)^6$.

We now give an informal description of the approximation algorithm. Consider a radial distribution network with nodes labeled $i \in [n]$, (where 1 denotes the substation bus, where the radial network connects to the transmission grid). Radial distribution networks have a *tree* topology (they do not have cycles). We denote bus 1 as the *root* of the tree, and buses with only one neighbor as *leaves*. Every node (except the root) has a unique *parent*, defined as the first node on the unique path connecting it to the root node. The set of nodes k that have a give node i as its parent are said to be its *children*. It can be shown that the strategic curtailment problem on any radial distribution network can be expressed as an equivalent problem on a network where each node has maximum degree 3 (known as a *binary tree*, see Appendix D). Thus, we can limit our attention to networks of this type, where every node has a unique parent and at most 2 children.

For a node i, let $c_1(i), c_2(i)$ denote its children, as in Fig. 6 (where $c_1 = \emptyset$, $c_2 = \emptyset$ is allowed since a node can have fewer than two children). We use the shorthand

$$p^{net}(i) = f_{c_1(i)} + f_{c_2(i)} - f_i - (p_i - \alpha_i - d_i).$$

Constraint (4a) reduces to $\underline{\Delta p_i} \leq p^{net}(i) \leq \overline{\Delta p_i}$, where $f_1 = 0$ and $f_\emptyset = 0$. The matrix H in (4c) is an empty matrix (the nullspace of the matrix B is of dimension 0), so this constraint can be dropped. Using this additional structure, the problem (5) can be rewritten (after some algebra) as:

$$\underset{\lambda,f,\alpha}{\text{maximize}} \quad \sum_{i=1}^{n} \lambda_i\big(p_i^a - \alpha_i\big) \tag{6a}$$

$$\text{subject to} \quad 0 \le \alpha_i \le p_i^a, \quad i \in [n] \tag{6b}$$

$$\underline{\Delta p}_i \le p^{net}(i) \le \overline{\Delta p}_i, \quad i \in [n] \tag{6c}$$

$$\underline{f}_i \le f_i \le \bar{f}_i, \quad i \in [n] \setminus \{1\} \tag{6d}$$

$$\lambda_i \begin{cases} \le c_i, & \text{if } p^{net}(i) = \underline{\Delta p}_i \\ = c_i, & \text{if } \underline{\Delta p}_i < p^{net}(i) < \overline{\Delta p}_i, \\ \ge c_i, & \text{if } p^{net}(i) = \overline{\Delta p}_i \end{cases} \quad i \in [n] \tag{6e}$$

$$\lambda_{c_j(i)} - \lambda_i \begin{cases} \ge 0, & \text{if } f_i = \underline{f}_i \\ = 0, & \text{if } \underline{f}_i < f_i < \bar{f}_i, \\ \le 0, & \text{if } f_i = \bar{f}_i \end{cases} \quad i \in [n], j = 1, 2 \tag{6f}$$

where λ_i is the LMP at bus i. Note that we assumed that there is some aggregator generation and potential curtailment at every bus (however this is not restrictive, since we can simply set $p_i^a = 0$ at buses where the aggregator owns no assets).

Define $x_i = (\lambda_i, f_i, \alpha_i)$, it is easy to see that (6) is of the form

$$\max_x \quad \sum_{i=1}^{n} g_i(x_i)$$
$$\text{s.t.} \quad h_i\big(x_i, x_{c_1(i)}, x_{c_2(i)}\big) \le 0, \quad i \in [n]$$

for some functions $g_i(.)$ and $h_i(.)$. This form is amenable to dynamic programming, since if we fix the value of x_i, the optimization problem for the subtree under i is decoupled from the rest of the network. Set $\kappa_n(x) = 0$, define κ_i for $i < n$ recursively as

$$\kappa_i(x) = \max_{\substack{x_{c_1(i)}, x_{c_2(i)} \\ h_i(x, x_{c_1(i)}, x_{c_2(i)}) \le 0}} \sum_{j=1}^{2} g_{c_j(i)}\big(x_{c_j(i)}\big) + \kappa_{c_j(i)}\big(x_{c_j(i)}\big).$$

Then, the optimal value can be computed as $\gamma^* = \max_x \kappa_1(x) + g_1(x)$. However, the above recursion requires an infinite-dimensional computation at every step, since the value of κ_i needs to be calculated for *every* value of x. To get around this, we note that the variables λ_i, f_i, α_i are bounded, and hence x_i can be discretized to lie in a certain set $\mathcal{X}_i$ such that every feasible x_i is at most $\delta(\epsilon_i)$ away (in infinity-norm sense) from some point in $\mathcal{X}_i$ (Lemma 2). The discretization error can be quantified, and this error bound can be used to relax the constraint to $h_i(x_i, x_{i+1}) \le \epsilon$ guaranteeing that any solution to (5) is feasible for the relaxed constraint. This allows us to define a dynamic program (Algorithm 1).

The algorithm essentially starts at the leaves of the tree and proceeds towards the root, at each stage updating κ for nodes whose children have already been updated (stopping at root). Along with the discretization error analysis in Appendix D, this essentially concludes Theorem 3.

It is worth noting that previous work on distribution level markets have used AC power flow models (at least in some approximate form) due to the importance of voltage constraints and reactive power in a distribution system [48]. Our approach extends in a straightforward way to this setting as well, as the dynamic programming structure remains preserved (the KKT conditions will simply be replaced by the corresponding conditions for the AC based market clearing mechanism).

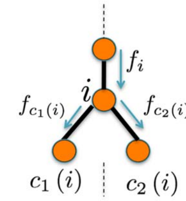


Fig. 6. The representation of a binary tree. For any node i, the children are denoted by $c_1(i), c_2(i)$.

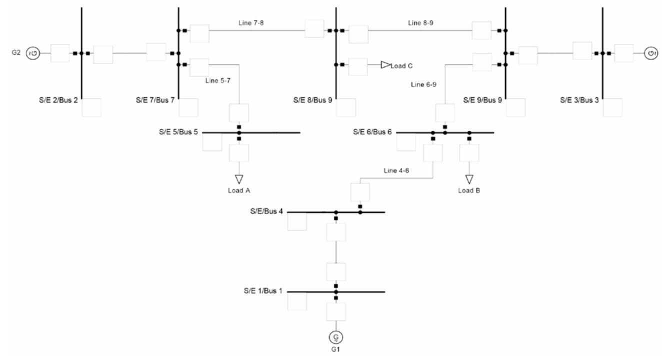


Fig. 7. The 9-bus acyclic network from [49], used for the evaluation of the proposed approximation algorithm.

Algorithm 1 Dynamic Programming on Binary Tree

$S \leftarrow \{i : c_1(i) = \emptyset, c_2(i) = \emptyset\}$
$\kappa_i(x) \leftarrow 0 \ \forall x \in \mathcal{X}_i, i \in S$
while $|S| \le n$ **do**
  $S' \leftarrow \{i \notin S : c_1(i), c_2(i) \in S\}$
  $\forall i \in S', \forall x \in \mathcal{X}_i$:

$$\kappa_i(x) \leftarrow \max_{\substack{x_1' \in \mathcal{X}_{c_1(i)}, x_2' \in \mathcal{X}_{c_2(i)} \\ h_i(x, x_1', x_2') \le \epsilon}} \sum_{j=1,2} g_{c_j(i)}\big(x_j'\big) + \kappa_{c_j(i)}\big(x'\big)$$

  $S \leftarrow S \cup S'$
end while
$\gamma \leftarrow \max_{x \in \mathcal{X}_1} \kappa_1(x) + g_1(x)$

C. Evaluation of the Approximation Algorithm

To evaluate the performance of our approximation algorithm on acyclic networks, we run it on a number of small test networks and compare the results with the brute-force optimal values. The algorithm indeed finds solutions within the prespecified error range (and often exact) in reasonable time.

As an example, for an acyclic version of the IEEE 9-bus network (taken from [49], see Fig. 7), we demonstrate the suboptimality gap of the solution versus the running time in Fig. 8. At each point of the graph, the error percentage (y-axis) is bounded by a constant factor of ϵ. Clearly the smaller ϵ we choose, the longer is the running time, but the smaller is the error. As one can see the error drops pretty quickly.

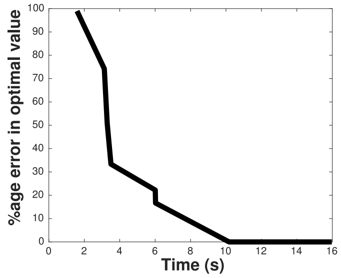


Fig. 8. The difference from the optimal solution as a function of the running time of the algorithm, in the 9-bus network with 1% curtailment allowance.

We should remark that the network chosen here was small in order to allow for comparison with the optimal value. However, the main advantage of our algorithm is that it is scalable, while the brute-force becomes intractable quickly.

VI. Concluding Remarks

Understanding the potential for market manipulation by aggregators is crucial for electricity market efficiency in the new era of renewable energy. In this paper, we characterized the profit an aggregator can make by strategically curtailing generation in the ex-post market as the outcome of a bi-level optimization problem. This model captures the realistic price clearing mechanism in the electricity market. We showed through simulations on realistic test cases that there is potentially large profit for aggregators by manipulating the LMPs in the electricity market. When the aggregator is located in a single bus, we have shown that the locational marginal price is monotonically increasing with the curtailment, and we have an exact polynomial-time algorithm to solve the aggregators profit maximization problem.

The aggregator's strategic curtailment problem in a general setting is a difficult bi-level optimization problem, which is intractable. However, we showed that for radial distribution networks (where aggregators are likely located) there is an efficient algorithm to approximate the solution up to arbitrary precision. We also demonstrated via simulation on a distribution test case that our algorithm can efficiently find the approximately optimal curtailment strategy.

We view this paper as a first step in understanding market power of aggregators, and more generally, towards market design for integrating renewable energy and demand response from geographically distributed sources. With the result of our paper, it is interesting to ask what can the operator do to address this problem. In particular, how to design market rules for aggregators to maximize the contribution of renewable energy yet mitigate the exercise of market power. Also, extending the analysis to the case of multiple aggregators in the market is another interesting direction for future research.

Appendix A
Connections Between Curtailment Profit and Market Power

As mentioned earlier, there has been significant work on market power in electricity markets, but work is only beginning to emerge on the market power of renewable generation producers. One important work from [10], and the following is the proposed notion of market power from that work.

Definition 4: For $\alpha_i^* \geq 0$, the **market power** (ability) of the aggregator is defined as

$$\eta_i = \left(\frac{\lambda_i(\alpha^*) - \lambda_i(0)}{\lambda_i(0)}\right) / \left(\frac{\alpha_i^*}{p_i^a}\right) \tag{7}$$

In this definition the value of η_i captures the ability of the generator/aggregator to exercise market power. Intuitively, in a market with high value of η_i, the aggregator can significantly increase the price by curtailing a small amount of generation.

Interestingly, the optimal curtailment profit is closely related to this notion of market power. We summarize the relationship in the following proposition.

Proposition 1: If the curtailment profit γ is positive then the market power $\eta_i > 1$. Furthermore, the larger the curtailment profit is, the higher is the market power.

Proof: From the definition of $\gamma(\alpha^*) = \lambda_i(\alpha^*)(p_i^a - \alpha_i^*) - \lambda_i(0)p_i^a$ it follows that

$$\begin{aligned}\frac{\gamma(\alpha^*)}{\lambda_i(0)\big(p_i^a - \alpha_i^*\big)} &= \frac{\lambda_i(\alpha^*)}{\lambda_i(0)} - \frac{p_i^a}{p_i^a - \alpha_i^*} \\ &= 1 + \frac{\lambda_i(\alpha^*) - \lambda_i(0)}{\lambda_i(0)} - \left(1 - \frac{\alpha_i^*}{p_i^a}\right)^{-1} \\ &\simeq 1 + \frac{\lambda_i(\alpha^*) - \lambda_i(0)}{\lambda_i(0)} - \left(1 + \frac{\alpha_i^*}{p_i^a}\right) \\ &= \frac{\lambda_i(\alpha^*) - \lambda_i(0)}{\lambda_i(0)} - \frac{\alpha_i^*}{p_i^a}. \end{aligned} \tag{8}$$

Therefore we have

$$\begin{aligned}\frac{p_i^a}{\lambda_i(0)\big(p_i^a - \alpha_i^*\big)\alpha_i^*}\gamma\big(\alpha^*\big) &= \left(\frac{\lambda_i(\alpha^*) - \lambda_i(0)}{\lambda_i(0)}\right) / \left(\frac{\alpha_i^*}{p_i^a}\right) - 1 \\ &= \eta_i - 1.\end{aligned}$$

Since the left-hand side parameters are all positive, if $\gamma(\alpha^*) > 0$, we can conclude that $\eta_i > 1$. Moreover, it is clear that the larger the value of $\gamma(\alpha^*)$ is, the higher the value of η_i is. Note that we used the approximation $(1 - \frac{\alpha_i^*}{p_i^a})^{-1} \simeq 1 + \frac{\alpha_i^*}{p_i^a}$, since the curtailment is small with respect to the generation; however, the right-hand side expression (8) is an upper bound on the left-hand side anyway, and the result holds exactly. ■

This proposition highlights that the notion of market power in [10] is consistent with an aggregator seeking to maximize their curtailment profit, and higher curtailment profit corresponds to more market power.

Appendix B
Proof of Lemma 1

Let us take a look at the ISO's optimization problem (1), which is a linear program. It is not hard to see that the dual of this problem is as follows.

$$\begin{aligned}\underset{\lambda^-, \lambda^+, \mu^-, \mu^+, \nu}{\text{maximize}} \quad & \big(\underline{\Delta p} + p - \alpha - d\big)^T \lambda^- \\ & + \big(-p + \alpha + d - \overline{\Delta p}\big)^T \lambda^+ + \underline{f}^T \mu^- - \overline{f}^T \mu^+ \end{aligned} \tag{9a}$$

$$\text{subject to} \quad B^T\left(c+\lambda^+-\lambda^-\right)-\mu^-+\mu^++H^T\nu=0 \tag{9b}$$

$$\lambda^-, \lambda^+, \mu^-, \mu^+ \geq 0 \tag{9c}$$

If one focuses on the terms involving α_i for a certain i, the objective of the above optimization problem is in the form: $(\underline{\Delta p}_i+p_i-\alpha_i-d_i)\lambda_i^-+(-p_i+\alpha_i+d_i-\overline{\Delta p_i})\lambda_i^+$ plus a linear function of the rest of the variables (i.e., the rest of λ^-, λ^+, as well as μ^-, μ^+, ν). There is no α in the constraints, and the first two terms of this objective are the only parts where α_i appears (and with opposite signs).

We need to show that if α_i is changed to $\alpha_i+\delta$ for some $\delta > 0$, then $c_i+\lambda_i^{+new}-\lambda_i^{-new} \geq c_i+\lambda_i^+-\lambda_i^-$, where $\lambda_i^{+new}, \lambda_i^{-new}$ are the optimal solutions of the new problem.

We prove this in a general setting. Consider the following two optimization problems.

$$f^* = \sup_{\substack{x_1,x_2\in\mathbb{R}\\ x_3\in\mathbb{R}^m}} \quad a_1x_1+a_2x_2+a_3^Tx_3 \tag{10a}$$

$$\text{s.t.} \quad (x_1,x_2,x_3)\in S \tag{10b}$$

$$f^{*new} = \sup_{\substack{x_1,x_2\in\mathbb{R}\\ x_3\in\mathbb{R}^m}} \quad (a_1-\delta)x_1+(a_2+\delta)x_2+a_3^Tx_3 \tag{11a}$$

$$\text{s.t.} \quad (x_1,x_2,x_3)\in S \tag{11b}$$

Assume that the optimal values of the problems are attained at (x_1^*, x_2^*, x_3^*) and $(x_1^{*new}, x_2^{*new}, x_3^{*new})$, respectively.

We claim that $x_2^{*new}-x_1^{*new} \geq x_2^*-x_1^*$ (This precisely implies the LMP condition in our case, i.e., $\lambda_i^{+new}-\lambda_i^{-new} \geq \lambda_i^+-\lambda_i^-$).

Suppose by way of contradiction that $x_2^{*new}-x_1^{*new} < x_2^*-x_1^*$.

We know that $a_1x_1^*+a_2x_2^*+a_3^Tx_3^* \geq a_1x_1+a_2x_2+a_3^Tx_3, \ \forall(x_1,x_2,x_3)\in S$.

Therefore we have

$$\begin{aligned}
&(a_1-\delta)x_1^{*new}+(a_2+\delta)x_2^{*new}+a_3^Tx_3^{*new}\\
&\quad = a_1x_1^{*new}+a_2x_2^{*new}+a_3^Tx_3^{*new}-\delta x_1^{*new}+\delta x_2^{*new}\\
&\quad \leq a_1x_1^*+a_2x_2^*+a_3^Tx_3^*+\delta\left(x_2^{*new}-x_1^{*new}\right)\\
&\quad < a_1x_1^*+a_2x_2^*+a_3^Tx_3^*+\delta\left(x_2^*-x_1^*\right)\\
&\quad = (a_1-\delta)x_1^*+(a_2+\delta)x_2^*+a_3^Tx_3^*.
\end{aligned}$$

The first inequality above follows from the fact that $(x_1^{*new}, x_2^{*new}, x_3^{*new}) \in S$. Now the above implies that $(x_1^{*new}, x_2^{*new}, x_3^{*new})$ is not the optimal solution of (11), and it is a contradiction.

As a result, $x_2^{*new}-x_1^{*new} \geq x_2^*-x_1^*$. ∎

APPENDIX C
PROOF OF THEOREM 1

Since we are in the single-bus curtailment regime, α has only one nonzero component. For the sake of convenience, we denote that element itself by a scalar α throughout this proof (no α is vector in this proof). The proof consists of the following two pieces: 1) From each jump point, the point where the next jump happens can be computed in polynomial time, 2) There are at most polynomially (in this case even linearly) many jumps.

Assuming that the solution to the program (1) is unique, for any fixed value of α, exactly t of the constraints (1b), (1c), (1d) are binding (active). We can express these binding constraints as

$$Af = b(\alpha),$$

where $A \in \mathbb{R}^{t\times t}$ is an invertible matrix, and $b(\alpha) \in \mathbb{R}^t$ is a vector that depends on α. As long as the binding constraints do not change, the matrix A is fixed and the optimal solution is linear in α (i.e., $f = A^{-1}b(\alpha)$). Then, for simplicity, we can express the solution as $f(\alpha) = f_0+\alpha a$, for some t-vectors f_0 and a.

Now if we look at the non-binding (inactive) constraints of (1), they can also be expressed as

$$\tilde{A}f < \tilde{b},$$

for some matrix $\tilde{A}$ and vector $\tilde{b}$ of appropriate dimensions. Inserting f into this set of inequalities yields $\tilde{A}f_0+\alpha\tilde{A}a < \tilde{b}$, or equivalently

$$\alpha\left(\tilde{A}a\right)_i < \tilde{b}_i-\left(\tilde{A}f_0\right)_i,$$

for all $i = 1, 2, \ldots, (2n+2t-rank(G))$.

Now we need to figure out that, with increasing α, which of the non-binding constraints becomes binding first and with exactly how much increase in α. If for some i we have $(\tilde{A}a)_i \leq 0$, then it is clear that increasing α cannot make constraint i binding. If $(\tilde{A}a)_i > 0$ then the constraint can be written as

$$\alpha < \frac{\tilde{b}_i-\left(\tilde{A}f_0\right)_i}{\left(\tilde{A}a\right)_i}.$$

Computing the right-hand side for all i, and taking their minimum, tells us exactly which constraint will become binding next and how much change in the current value of α results in that.

The complexity of this procedure is $O(t^{2.373})$ for computing f_0 and a, plus $O(t(2n+2t)) = O(nt+t^2)$ for computing the lowest bound among all the constraints. Hence the complexity is $O(t^{2.373})$.

The above procedure describes how the next jump point can be computed efficiently from the current point. The exact same procedure can be repeated for reaching the subsequent jump points. All remains to show is the second piece of the proof, which is that the number of jump points are bounded polynomially. To show the last part note that by increasing α, if a binding constraint becomes non-binding, it will not become binding again. As a result, each constraint can change at most twice, and therefore the number of jumps is at most twice the number of constraints. Thus, the number of jumps is $O(n+t)$, and the overall complexity of the algorithm is $O((n+t)t^{2.373}) = O(t^{3.373})$. ∎

APPENDIX D
PROOF OF THEOREM 2

Lemma 2 (δ-Discretization): Given a set $C \subset [\underline{L_1}, \overline{L_1}] \times \cdots \times [\underline{L_k}, \overline{L_k}]$, there exists a finite set $\mathcal{X}$ such that

$$\forall z \in C \quad \exists z' \in \mathcal{X}, \max_{1\leq i\leq k} |z_i - z_i'| \leq \delta$$

and $\mathcal{X}$ contains at most V/δ^k points, where $V = \prod_{i=1}^{k}(\overline{L_i} - \underline{L_i})$ is a constant (the volume of the box). $\mathcal{X}$ is said to be an δ-discretization of C and written as $\mathcal{X}(\delta)$.

Lemma 3 (Reduction to Binary Tree): Any tree with arbitrary degrees can be reduced to a binary tree by introducing additional dummy nodes to the network.

Proof: Take any node b in the tree with some parent a and $k > 2$ children $c_1, \ldots, c_k$. There exists $m > 0$ such that $2^m < k \leq 2^{m+1}$ for some m. We will show that this subgraph can be made a binary tree by introducing $O(k)$ dummy nodes (in m levels) between b and its children. The additional nodes and edges are defined as follows:

$$\begin{aligned}
&b \to b_0,\ b \to b_1,\\
&b_0 \to b_{00},\ b_0 \to b_{01},\ b_1 \to b_{10},\ b_1 \to b_{11},\\
&b_{00} \to b_{000},\ b_{00} \to b_{001},\ \ldots,\ b_{11} \to b_{111},
\end{aligned}$$

up to m levels:

$$b_{0\ldots 00} \to c_1,\ b_{0\cdots 00} \to c_2,\ b_{0\cdots 01} \to c_3\ \cdots$$

This is transparently a binary tree with $O(k)$ nodes. Each of the new nodes has zero injection, and effectively the incoming flow from its parent is just split in some way between its children. This in fact enforces the flow conservation constraint at b. Similar construction can be applied to any node of the tree with more than two children, until no such node exists. It can be seen that the number of nodes in the new graph is still linear in n. ■

So any tree can be transformed to a binary one by the above procedure. For the rest of the analysis we focus on the ϵ-approximation of the dynamic program on the resulting binary tree. The optimization problem (5) on a binary tree, can be written after some algebra as the following.

$$\max_{\lambda,f,\alpha} \quad \sum_{i=1}^{n} \lambda_i\big(p_i^a - \alpha_i\big) \tag{12a}$$

$$\text{subject to} \quad 0 \leq \alpha_i \leq p_i^a, \quad i = 1, \ldots, n \tag{12b}$$

$$\underline{\Delta p}_i \leq f_{c1(i)} + f_{c2(i)} - f_i - p_i + \alpha_i + d_i \leq \overline{\Delta p}_i, \quad i = 1, \ldots, n \tag{12c}$$

$$\underline{f}_i \leq f_i \leq \overline{f}_i, \quad i = 2, \ldots, n \tag{12d}$$

$$\begin{cases} (\lambda_i - c_i)\Big(f_{c1(i)} + f_{c2(i)} - f_i - p_i + \alpha_i + d_i - \underline{\Delta p}_i\Big) \geq 0 \\ (\lambda_i - c_i)\big(f_{c1(i)} + f_{c2(i)} - f_i - p_i + \alpha_i + d_i - \overline{\Delta p}_i\big) \geq 0, \end{cases} \quad i = 1, \ldots, n \tag{12e}$$

$$\begin{cases} \big(\lambda_i - \lambda_{c_j(i)}\big)\Big(\underline{f}_{c_j(i)} - f_{c_j(i)}\Big) \geq 0 \\ \big(\lambda_i - \lambda_{c_j(i)}\big)\Big(\overline{f}_{c_j(i)} - f_{c_j(i)}\Big) \geq 0, \end{cases} \quad i = 1, \ldots, n, \quad j = 1, 2 \tag{12f}$$

The constraints $0 \leq \alpha_i \leq p_i^a$ and $\underline{f}_i \leq f_i \leq \overline{f}_i$, along with a prior bound on lambda $\underline{\lambda} \leq \lambda \leq \overline{\lambda}$ can be used to define the box where $x_i = (\lambda_i, f_i, \alpha_i)$ lives. Then an ϵ-accurate solution is a solution to the following problem.

$$\max_{\lambda,f,\alpha} \quad \sum_{i=1}^{n} \lambda_i(p_i - \alpha_i) \tag{13a}$$

subject to

$$\underline{\Delta p}_i - \epsilon \leq f_{c1(i)} + f_{c2(i)} - f_i - p_i + \alpha_i + d_i \leq \overline{\Delta p}_i + \epsilon, \quad i = 1, \ldots, n \tag{13b}$$

$$\begin{cases} (\lambda_i - c_i)\Big(f_{c1(i)} + f_{c2(i)} - f_i - p_i + \alpha_i + d_i - \underline{\Delta p}_i\Big) \geq -\epsilon \\ (\lambda_i - c_i)\big(f_{c1(i)} + f_{c2(i)} - f_i - p_i + \alpha_i + d_i - \overline{\Delta p}_i\big) \geq -\epsilon, \end{cases} \quad i = 1, \ldots, n \tag{13c}$$

$$\begin{cases} \big(\lambda_i - \lambda_{c_j(i)}\big)\Big(\underline{f}_{c_j(i)} - f_{c_j(i)}\Big) \geq -\epsilon \\ \big(\lambda_i - \lambda_{c_j(i)}\big)\Big(\overline{f}_{c_j(i)} - f_{c_j(i)}\Big) \geq -\epsilon, \end{cases} \quad i = 1, \ldots, n, \quad j = 1, 2 \tag{13d}$$

Assuming a δ-discretization of the constraint set, each of the constraints (as well as ϵ-accuracy of the objective) imposes a bound on the value of δ. For example constraint (13c) requires $4\delta \leq \epsilon$ (Note that we could have defined different deltas $\delta^\lambda, \delta^f, \delta^\alpha$ for different variables and in that case we had $3\delta^f + \delta^a \leq \epsilon$, but for simplicity we take all the deltas to be the same). Similar bounds on δ can be obtained from the other constraints, and taking the lowest upper-bound, implies the existence of a constant c' (that depends on the parameters) such that $\delta \leq \epsilon/c'$.

As a result we have a δ-discretization with $|\mathcal{X}| = V/\delta^3 = c'^3 V/\epsilon^3$ number of points, for any node. Therefore, the computational complexity over any node will be $|\mathcal{X}|^3$, because we have $|\mathcal{X}|$ many values for the node itself and $|\mathcal{X}|$ many values for any of its two children. Since there are n nodes, the overall complexity of the algorithm will simply be $n|\mathcal{X}|^3 = nc'^9 V^3/\epsilon^9 = cn/\epsilon^9$. ■

REFERENCES

[1] K. Bullis, *Why Solarcity Is Succeeding in a Difficult Solar Industry*, MIT Technology Review, Cambridge, MA, USA, 2012. [Online]. Available: https://www.technologyreview.com/s/427830/why-solarcity-is-succeeding-in-a-difficult-solar-industry/

[2] J. John. (2014). *SolarCity and Tesla: A Utility's Worst Nightmare*. [Online]. Available: http://www.greentechmedia.com/articles/read/SolarCitys-Networked-Grid-Ready-Energy-Storage-Fleet

[3] E. Wesoff. (2016). *Earnings From SunPower, Tesla, Enphase, Plus New Funding for Sunnova, SolarCity, 1366, Siva, Nexeon*. [Online]. Available: http://www.greentechmedia.com/articles/read/Earnings-from-SunPower-Tesla-Enphase-and-New-Funding-for-Sunnova-SolarCi

[4] *Solar Energy Industries Association (SEIA)*. Accessed on May 26, 2016. [Online]. Available: http://www.seia.org/state-solar-policy

[5] (Jan. 2016). *2015 Top 500 North American Solar Contractors*. [Online]. Available: http://www.solarpowerworldonline.com

[6] N. Litvak, "U.S. commercial solar landscape 2016-2020," GTM Res., Boston, MA, USA, Tech. Rep., 2016. [Online]. Available: http://www.greentechmedia.com/research/report/us-commercial-solar-landscape-2016-2020

[7] C. Honeyman, "U.S. residential solar economic outlook 2016-2020," GTM Res., Boston, MA, USA, Tech. Rep. [Online]. Available: http://www.greentechmedia.com/research/report/us-residential-solar-economic-outlook-2016-2020

[8] *Reforming the Energy Vision (REV) Proceeding—NYS SmartGrid Consortium*. [Online]. Available: http://nyssmartgrid.com/innovation-highlights/rev-proceeding

[9] *FERC Grants NYISO Request to Give Behind-the-Meter Resources Market Access*. [Online]. Available: http://www.utilitydive.com/news/ferc-grants-nyiso-request-to-give-behind-the-meter-resources-market-access/419791/

[10] Y. Yu, B. Zhang, and R. Rajagopal, "Do wind power producers have market power and exercise it?" in *Proc. IEEE PES Gen. Meeting Conf. Expo.*, 2014, pp. 1–5.

[11] P. Twomey, R. Green, K. Neuhoff, and D. Newbery, "A review of the monitoring of market power: The possible roles of TSOs in monitoring for market power issues in congested transmission systems," Tech. Rep., 2015. [Online]. Available: https://doi.org/10.17863/CAM.5068

[12] J. Bushnell *et al.*, *An International Comparison of Models for Measuring Market Power in Electricity*, Energy Model. Forum Stanford Univ., Stanford, CA, USA, 1999.

[13] A. Sheffrin and J. Chen, "Predicting market power in wholesale electricity markets," in *Proc. Western Conf. Adv. Regul. Competition*, South Lake Tahoe, CA, USA, 2002.

[14] "*Market Power and Competitiveness*, California Independent Syst. Operator, Folsom, CA, USA, 1998. [Online]. Available: http://www.caiso.com/1c5f/1c5fbe6a1a720.html

[15] *ERCOT Protocols*, Elect. Rel. Council Texas, Austin, TX, USA, 2001. [Online]. Available: http://www.ercot.com

[16] R. Schmalensee and B. W. Golub, "Estimating effective concentration in deregulated wholesale electricity markets," *RAND J. Econ.*, vol. 15, no. 1, pp. 12–26, 1984.

[17] D. T. Scheffman and P. T. Spiller, "Geographic market definition under the U.S. department of justice merger guidelines," *J. Law Econ.*, vol. 30, no. 1, pp. 123–147, 1987.

[18] S. Borenstein, J. Bushnell, E. Kahn, and S. Stoft, "Market power in California electricity markets," *Utilities Policy*, vol. 5, nos. 3–4, pp. 219–236, 1995.

[19] S. S. Oren, "Economic inefficiency of passive transmission rights in congested electricity systems with competitive generation," *Energy J.*, vol. 18, no. 1, pp. 63–83, 1997.

[20] J. B. Cardell, C. C. Hitt, and W. W. Hogan, "Market power and strategic interaction in electricity networks," *Resource Energy Econ.*, vol. 19, nos. 1–2, pp. 109–137, 1997.

[21] L. Xu and R. Baldick, "Transmission-constrained residual demand derivative in electricity markets," *IEEE Trans. Power Syst.*, vol. 22, no. 4, pp. 1563–1573, Nov. 2007.

[22] L. Xu and Y. Yu, "Transmission constrained linear supply function equilibrium in power markets: Method and example," in *Proc. PowerCon Int. Conf. Power Syst. Technol.*, vol. 3. Kunming, China, 2002, pp. 1349–1354.

[23] S. Bose, C. Wu, Y. Xu, A. Wierman, and H. Mohsenian-Rad, "A unifying market power measure for deregulated transmission-constrained electricity markets," *IEEE Trans. Power Syst.*, vol. 30, no. 5, pp. 2338–2348, Sep. 2015.

[24] P. Twomey and K. Neuhoff, "Wind power and market power in competitive markets," *Energy Policy*, vol. 38, no. 7, pp. 3198–3210, 2010.

[25] A. L. Ott, "Experience with PJM market operation, system design, and implementation," *IEEE Trans. Power Syst.*, vol. 18, no. 2, pp. 528–534, May 2003.

[26] T. Zheng and E. Litvinov, "Ex post pricing in the co-optimized energy and reserve market," *IEEE Trans. Power Syst.*, vol. 21, no. 4, pp. 1528–1538, Nov. 2006.

[27] S. Bi and Y. J. Zhang, "False-data injection attack to control real-time price in electricity market," in *Proc. IEEE Glob. Commun. Conf. (GLOBECOM)*, 2013, pp. 772–777.

[28] S. Bi and Y. J. Zhang, "Using covert topological information for defense against malicious attacks on dc state estimation," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 7, pp. 1471–1485, Jul. 2014.

[29] L. Xie, Y. Mo, and B. Sinopoli, "False data injection attacks in electricity markets," in *Proc. 1st IEEE Int. Conf. Smart Grid Commun. (SmartGridComm)*, 2010, pp. 226–231.

[30] L. Xie, Y. Mo, and B. Sinopoli, "Integrity data attacks in power market operations," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 659–666, Dec. 2011.

[31] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Trans. Inf. Syst. Security*, vol. 14, no. 1, p. 13, 2011.

[32] Y. Lin, P. Barooah, S. Meyn, and T. Middelkoop, "Experimental evaluation of frequency regulation from commercial building HVAC systems," *IEEE Trans. Smart Grid*, vol. 6, no. 2, pp. 776–783, Mar. 2015.

[33] H. Hao, A. Kowli, Y. Lin, P. Barooah, and S. Meyn, "Ancillary service for the grid via control of commercial building HVAC systems," in *Proc. Amer. Control Conf. (ACC)*, 2013, pp. 467–472.

[34] B. Zhang, A. Y. S. Lam, A. D. Domínguez-García, and D. Tse, "An optimal and distributed method for voltage regulation in power distribution systems," *IEEE Trans. Power Syst.*, vol. 30, no. 4, pp. 1714–1726, Jul. 2015.

[35] D. B. Arnold, M. Negrete-Pincetic, M. D. Sankur, D. M. Auslander, and D. S. Callaway, "Model-free optimal control of VAR resources in distribution systems: An extremum seeking approach," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 3583–3593, Sep. 2015.

[36] M. Negrete-Pincetic and S. Meyn, "Markets for differentiated electric power products in a smart grid environment," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Vancouver, BC, Canada, 2012, pp. 1–7.

[37] Y. Chen, A. Bušić, and S. Meyn, "Individual risk in mean field control with application to automated demand response," in *Proc. IEEE 53rd Annu. Conf. Decis. Control (CDC)*, Los Angeles, CA, USA, 2014, pp. 6425–6432.

[38] M. C. Ferris, S. P. Dirkse, and A. Meeraus, "Mathematical programs with equilibrium constraints: Automatic reformulation and solution via constrained optimization," Dept. Comput. Sci., Univ. Wisconsin at Madison, Madison, WI, USA, Tech. Rep. 02/11, 2002.

[39] O. Ben-Ayed and C. E. Blair, "Computational difficulties of bilevel linear programming," *Oper. Res.*, vol. 38, no. 3, pp. 556–560, 1990.

[40] Z. H. Gümüş and C. A. Floudas, "Global optimization of nonlinear bilevel programming problems," *J. Glob. Optim.*, vol. 20, no. 1, pp. 1–31, 2001.

[41] S. P. Dirkse and M. C. Ferris, "The path solver: A nommonotone stabilization scheme for mixed complementarity problems," *Optim. Methods Softw.*, vol. 5, no. 2, pp. 123–156, 1995.

[42] Y. Makarov *et al.*, "Incorporation of wind power resources into the California energy market," in *Proc. WINDPOWER Conf. Exhibit.*, 2005.

[43] (2004). *Shift Factor: Methodology and Example*. [Online]. Available: http://www.caiso.com/docs/2004/02/13/200402131609438684.pdf

[44] D. B. Patton, P. LeeVanSchaick, and J. Chen, "2013 assessment of the electricity markets in new england," Potomac Econ., Fairfax, VA, USA, Tech. Rep., 2014.

[45] *Pricing Enhancements: Revised Straw Proposal*, California Independent Syst. Operator, Folsom, CA, USA, 2014. [Online]. Available: https://www.caiso.com/Documents/RevisedStrawProposal_PricingEnhancements.pdf

[46] C. Coffrin, D. Gordon, and P. Scott, "NESTA, the NICTA energy system test case archive," *arXiv preprint arXiv:1411.0359*, 2014.

[47] S. Dempe, V. Kalashnikov, G. A. Pérez-Valdés, and N. Kalashnikova, *Bilevel Programming Problems: Theory, Algorithms and Applications to Energy Networks*. Heidelberg, Germany: Springer, 2015.

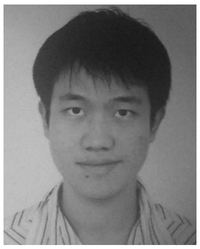
[48] E. Ntakou and M. Caramanis, "Distribution network electricity market clearing: Parallelized PMP algorithms with minimal coordination," in *Proc. IEEE 53rd Annu. Conf. Decis. Control (CDC)*, Los Angeles, CA, USA, 2014, pp. 1687–1694.

[49] B. Kocuk, S. S. Dey, and X. A. Sun, "Inexactness of SDP relaxation and valid inequalities for optimal power flow," *IEEE Trans. Power Syst.*, vol. 31, no. 1, pp. 642–651, Jan. 2016.

Navid Azizan Ruhi received the B.S. degree form the Sharif University of Technology, Tehran, Iran, and the M.S. degree from the University of Southern California, Los Angeles, CA, USA, in 2013 and 2015, respectively, both in electrical engineering. He is currently pursuing the Ph.D. degree in computing and mathematical sciences with the California Institute of Technology. His research interests span optimization, learning, and control, in networked systems.

He was the first place winner and a gold medalist at the 21st National Physics Olympiad in Iran. He is also a recipient of the Chancellor's Award from Sharif University of Technology, the Annenberg Fellowship from the University of Southern California, the CMS Fellowship from the California Institute of Technology, and the Best Student Paper Award at the 2016 ACM GREENMETRICS.

Krishnamurthy Dvijotham received the Ph.D. degree in computer science and engineering with the University of Washington, Seattle, in 2014, and the bachelor's degree from IIT Bombay in 2008. He is a Researcher with the Optimization and Control Group, Pacific Northwest National Laboratory and a Research Assistant Professor with Washington State University, Pullman, WA, USA. His research interests center on algorithmic problems in large scale infrastructure networks and societal systems. He was a recipient of Awards at leading conferences in optimization, AI, and machine learning.



Niangjun Chen received the B.A. degree in computer science from the University of Cambridge in 2011. He is currently pursuing the Ph.D. degree with the Department of Computing and Mathematical Sciences, California Institute of Technology, where he is a member of the Rigorous Systems Research Group and Netlab. He was a Research Engineer with the Institute of Infocomm Technology in Singapore in 2012. His research interests include online algorithms, optimization, and game theory and their applications to distributed control and mechanism design for power systems, data centers, and electricity markets.

Adam Wierman is a Professor with the Department of Computing and Mathematical Sciences, California Institute of Technology, where he currently serves as Executive Officer. He is also the Director of the Information Science and Technology Initiative, Caltech. He is the Founding Director of the Rigorous Systems Research Group and the Co-Director of the Social and Information Sciences Laboratory. His research interests center around resource allocation and scheduling decisions in computer systems and services. He was a recipient of the 2011 ACM SIGMETRICS Rising Star Award, the 2014 IEEE Communications Society William R. Bennett Prize, and has been co-author on papers that received the Best Paper Awards at ACM SIGMETRICS, IEEE INFOCOM, IFIP Performance (twice), IEEE Green Computing Conference, IEEE Power & Energy Society General Meeting, and ACM GREENMETRICS.