CrossMark

Incentive-compatible demand-side management for smart grids based on review strategies

Jie Xu[*] and Mihaela van der Schaar

Abstract

Demand-side load management is able to significantly improve the energy efficiency of smart grids. Since the electricity production cost depends on the aggregate energy usage of multiple consumers, an important incentive problem emerges: self-interested consumers want to increase their own utilities by consuming more than the socially optimal amount of energy during peak hours since the increased cost is shared among the entire set of consumers. To incentivize self-interested consumers to take the socially optimal scheduling actions, we design a new class of protocols based on review strategies. These strategies work as follows: first, a review stage takes place in which a statistical test is performed based on the daily prices of the previous billing cycle to determine whether or not the other consumers schedule their electricity loads in a socially optimal way. If the test fails, the consumers trigger a punishment phase in which, for a certain time, they adjust their energy scheduling in such a way that everybody in the consumer set is punished due to an increased price. Using a carefully designed protocol based on such review strategies, consumers then have incentives to take the socially optimal load scheduling to avoid entering this punishment phase. We rigorously characterize the impact of deploying protocols based on review strategies on the system's as well as the users' performance and determine the optimal design (optimal billing cycle, punishment length, etc.) for various smart grid deployment scenarios. Even though this paper considers a simplified smart grid model, our analysis provides important and useful insights for designing incentive-compatible demand-side management schemes based on aggregate energy usage information in a variety of practical scenarios.

1 Introduction

Governments and relevant industries are making a significant effort to develop next-generation energy grids ("smart grid") which meet new environmental requirements as well as increased usage demands [1]. To address these challenges, demand-side management (DSM) techniques for smart grids (see e.g., [1, 2]) were proposed as a way to significantly save energy. In a typical configuration, the energy producer (e.g., a utility company) periodically receives usage information from the smart meter affiliated with consumers via a communication network. The energy producer then manages the energy generation/purchase/transmission and bills the consumer based on this usage information. To save energy, reduce cost, and increase reliability, the energy producer can use for

instance smart pricing to encourage consumers to transfer peak-hour consumptions to off-peak hours [3].

In this paper, we focus on DSM which operates based on the knowledge of aggregate information of real-time energy usage of a set of consumers instead of each individual consumer (see Fig. 1). The use of aggregate information instead of individual consumers' information may be due to various reasons, e.g., high cost for large-scale deployment of smart meters for individual houses. Specifically, the utility company generates/purchases/transmits electricity according to the real-time *aggregate* usage information of a set of consumers and charges the consumers for their electricity usage at a price proportional to the cost due to energy generation/purchase/transmission. Since the cost incurred in peak hours is higher than that in off-peak hours, the price is higher if consumers schedule more loads in peak hours. If consumers were cooperative,

*Correspondence: jiexu@ucla.edu
Electrical Engineering Department, University of California, Los Angeles, USA
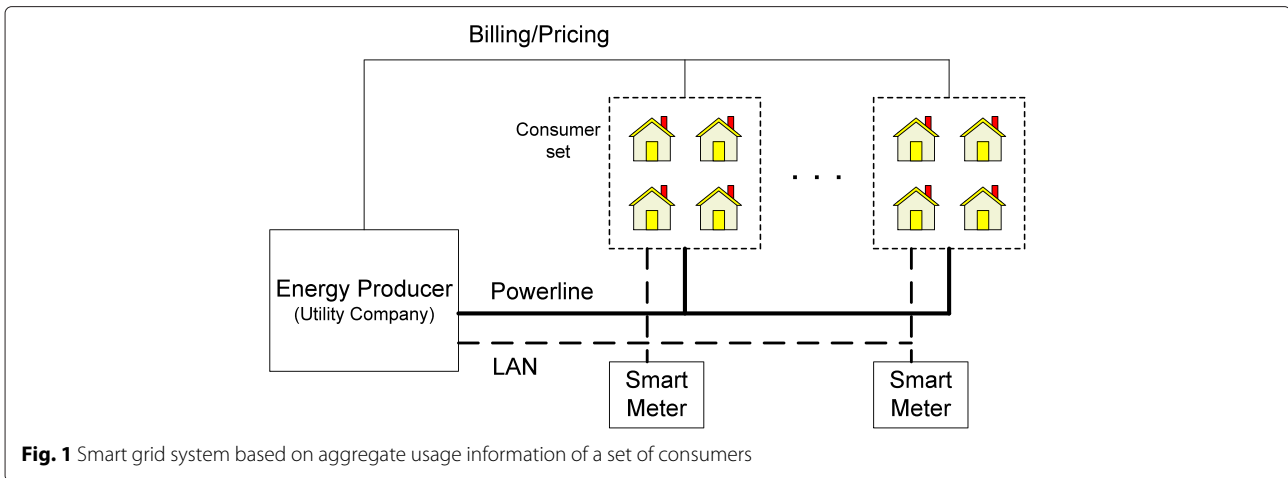
Fig. 1 Smart grid system based on aggregate usage information of a set of consumers

then such a pricing scheme would result in a socially optimal load scheduling by the consumers, meaning that the sum utility of consumers is maximized. However, the self-interested nature of individual consumers imposes some important challenges for the deployment of such solutions. Because the consumers' utilities (i.e., the benefit from energy consumption minus the payment) depend on other consumers' scheduling actions, an energy consumption game emerges: self-interested consumers may want to consume more energy in peak hours while paying a relatively lower cost since this cost is shared among all the participating consumers. Therefore, a crucial problem for such smart grid systems becomes how to incentivize individual consumers to perform the socially optimal load scheduling in order to maximize the overall grid performance. In this paper, we augment existing DSM schemes by proposing a novel technology which incentivizes consumers to reduce their consumption in peak hours. This technology is general and can be deployed in conjunction with many DSM schemes proposed in smart grids based on utility functions so that selfish consumers find it in their self-interest to follow the DSM scheme.

We model participating consumers' interactions as a repeated game [4] in which the energy consumption game is played repeatedly (e.g., every day). We assume that the energy delivery system is deploying a protocol which is designed by the utility company or a third party aiming to maximize the system efficiency (i.e., the sum utility of the consumers in this paper). Besides designing a smart pricing scheme that is based on the aggregate usage pattern of the consumers, this protocol designer also constructs a protocol which recommends a set of scheduling actions to consumers based on the past prices, which depend on the history of all the consumers' energy consumption and scheduling. The protocol designer is only active during the design stage of the protocol and it is passive at run time. In implementing the protocol, the utility company observes

every day the (aggregate) energy consumption pattern of the consumer set and based on this performs billing every billing cycle. The consumers use the past prices to determine their future energy consumption scheduling actions in a self-interested and completely decentralized manner. Note that the protocol designer is passive at run time, and hence, it cannot oblige the consumers to follow the recommended energy scheduling. The consumers will only adopt the recommended consumption scheduling if it is in their self-interest to do so, i.e., if they are better off following the recommended protocol rather than deviating from it. Such a protocol is called incentive-compatible (IC).

To ensure that the self-interested consumers will take the recommended scheduling actions, a punishment phase is incorporated in the recommendation of the protocol designer upon observing deviations. The protocol designer designs review strategy-based protocol which consist of two phases: the review phase and the punishment phase. (Its operation is depicted in Fig. 2.) In the review phases, the consumers are recommended to take the socially optimal scheduling actions. At the end of each review phase, a statistical test is performed (either by the consumers or the utility company) on the prices announced by the utility company. If the test fails (i.e., there is an evidence that some consumers did not follow the recommendation), the system goes into a punishment phase in which consumers perform for a certain time the strongest punishment which can be imposed on a self-interested consumer. The punishment is implemented by the participating consumers using a different energy consumption scheduling. In particular, upon a deviation, all consumers shift more of their energy consumption to peak hours and, hence, the prices for the next billing cycles will be raised. Therefore, the consumer who deviated will be punished by receiving a low utility.

There are two important points that are worth noting. First, even though the consumers who did not

Fig. 2 Operation of protocols based on review strategies
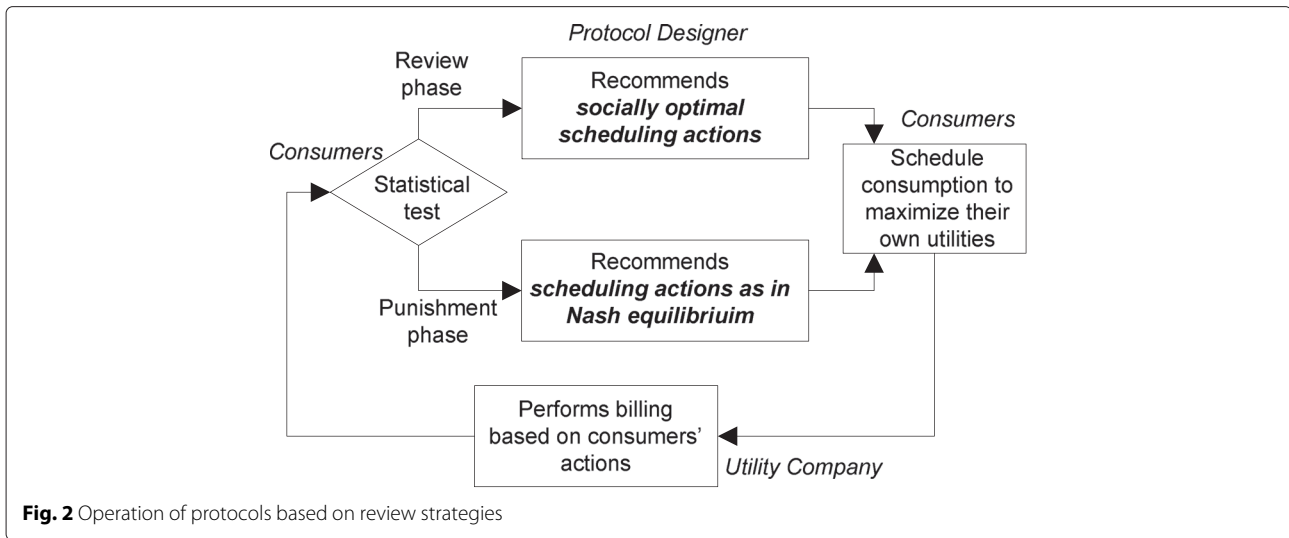
deviate will also receive low utilities due to the raised prices, the protocol is designed in such a way that consumers still want to carry out such a punishment. Second, the utilities obtained by the consumers in the punishment phase are lower than those in the review phase. However, they are the same as in the scenario where there was no review strategy protocol being deployed. Therefore, the review strategy-based protocol is guaranteed to achieve better system performance than that which can be obtained without deploying the protocol. Importantly, the protocol designer needs to carefully design the protocol such that this punishment does not take place too frequently since it reduces all consumers' utilities. We rigorously characterize the performance of deploying review strategies and determine the optimal design (optimal billing cycle, punishment phase length, etc.) for various smart grid deployment scenarios. It is also important to note that the proposed review strategy-based protocol can be deployed in conjunction with any DSM scheme besides the smart pricing scheme which is used for illustration in this paper, which may take into account find-grained load scheduling (e.g., hourly scheduling). The scheduling results of any DSM scheduling schemes can serve as the input of our design framework, and the output is the review strategy-based protocol.

The rest of this paper is organized as follows. In Section 2, related works are discussed. Section 3 models the repeated energy consumption game and formulates the design problem of the DSM with usage data aggregation. Section 4 formally introduces the proposed incentive protocol based on review strategies. Section 5 determines the optimal protocol parameters and evaluates its performance. Section 6 provides simulation results. Section 7 concludes the paper.

2 Related works

A main issue for the efficient deployment of smart grids is the design of DSM [5]. A large body of literature assumes the deployment of smart meters and designs smart pricing schemes to encourage individual consumers to manage their own loads (e.g., by shifting their energy consumption from peak hours to off-peak hours). Among them, real-time pricing (RTP) [6], time-of-use pricing (TOUP) [7], and critical-peak pricing (CPP) [8] represent popular options.

Recent works [3, 9–14] considered consumers' discomfort costs and aimed to jointly minimize the consumers' billing and discomfort costs, by assuming some utility functions. These works can be classified into two categories. In the first category, consumers are assumed to be price-taking, meaning that they do not consider how their consumption will affect the prices. In this case, the decision-making of a *single* foresighted consumer is formulated as a stochastic control problem aiming to maximize its long-term utility [11–13]. Alternatively, in [15, 16], multiple *myopic* consumers aim to maximize their utility, and their decisions are formulated as static optimization problems among *cooperative* users.

The second category assumed that consumers are myopic and price-anticipating, meaning that they take into account how their consumption will affect the prices. In this case, each consumer's electricity usage affects the other consumers' billing costs. These works [3, 17, 18] model the interaction emerging among myopic consumers as one-shot games and studied the Nash equilibrium (NE) of the emerging game. In this paper, we also model the consumers as price-anticipating. However, consumers interact with each other *repeatedly* and are *foresighted*, thereby engaging in a repeated game. It is well-known that the Nash equilibrium in one-shot games

with myopic players is often inefficient. In this paper, we design a novel class of incentive protocols based on review strategies in order to achieve the socially optimal load scheduling in smart grid systems. Prior work [19] also studied DSM in a repeated game setting. However, that work assumes that each individual consumer's action can be perfectly observed, while in this work, only the aggregate scheduling of a set of participating consumers can be observed with noise.

This paper adopts a similar pricing method as in [3] where pricing is performed based on the *aggregate* usage pattern of a set of consumers and consumers are charged proportional to their total daily energy consumption. It is argued in [3] that this proportional charging model is consistent with the existing residential metering models. Nevertheless, our work can be used in conjunction with a variety of existing DSM scheduling methods [6–8]. Augmenting these methods with our proposed incentive protocols is especially important when consumers have incentives to deviate from the optimal scheduling given by a DSM scheme (e.g., when it is performed on the aggregate energy usage information).

The approach proposed in this paper contributes to both the smart grid and the game-theoretic literature dedicated to engineering applications. Review strategies have been adopted in the principal-agent games with discounting in [20] in economics, and such games differ significantly from the smart grid deployment scenario considered in this paper. In [20], the game is played between only two players (i.e., a principal and an agent). However, in the considered smart grid scenario, there are multiple players (i.e., a set of consumers) and their utilities exhibit negative externalities.

Table 1 provides a comparison with existing works.

3 System model
3.1 Power system
We consider a smart grid system with multiple consumers and one energy producer, e.g., a utility company. These consumers receive electricity from the same aggregator which distributes electricity to the consumers. Each consumer is equipped with an energy consumption scheduler (ECS) for scheduling the household energy consumption. A smart meter is connected to the set of consumers from

which it collects and analyzes the energy consumption. This smart meter gathers (almost) accurate readings automatically, at requested time intervals, and relays them to the utility company. Using this information, the aggregator (utility company) can adjust its energy generation, purchase, and transmission accordingly. The communication between the utility company and the consumers' smart meters is done through the local area network (LAN) by using appropriate communication protocols. Let $\mathcal{N}$ denote the set of consumers that share the same aggregator, where the number of consumers is N. Figure 1 illustrates the system model.

Time is discrete and each time slot represents 1 day. Everyday, there are peak hours when consumers' energy consumption demand is high and off-peak hours when consumers' energy consumption demand is low. Nevertheless, our analysis can be easily extended to include finer-grained demand intensities, e.g., hourly scheduling. For consumer $n \in \mathcal{N}$, we assume that its daily schedulable demand does not change and is denoted by d_n. We also denote the schedulable demand vector of all consumers in the set $\mathcal{N}$ by $\boldsymbol{d}$. Consumers can schedule these portions of loads in different hours within a day. Since we consider only peak hours and off-peak hours, we write a_n as the fraction of load that is scheduled in peak hours for consumer n and $1 - a_n$ as the fraction of its load that is scheduled in off-peak hours [1]. Consumers prefer consuming energy in peak hours to off-peak hours. For consumer n, its energy consumption benefit per day is

$$b_n\left(a_n\right) = B\left(b_{\text{peak}} a_n d_n + b_{\text{off-peak}}\left(1 - a_n\right) d_n\right) \qquad (1)$$

where $b_{\text{peak}} > b_{\text{off-peak}} > 0$ and $B(\cdot)$ is a benefit function.

The utility company charges the consumers for their electricity consumption at a price $p\left(D_{\text{peak}}, D_{\text{off-peak}}\right)$ which is set depending on the electricity usage in peak hours and off-peak hours, where $D_{\text{peak}} = \sum_n a_n d_n$ is the total consumption in peak hours and $D_{\text{off-peak}} = \sum_n (1 - a_n) d_n$ is the total consumption in off-peak hours. Clearly, $D_{\text{peak}} + D_{\text{off-peak}} = \sum_n d_n$ is the total usage of all consumers. Let $C_{\text{peak}}(x)$ and $C_{\text{off-peak}}(x)$ represent the costs of generating and distributing x units of electricity

Table 1 Comparison with existing works

	Price-taking or price-anticipating	Myopic or foresighted	Model	Observe individual or aggregate usage
[15, 16]	Price-taking	Myopic	Optimization	Individual
[3, 17, 18]	Price-anticipating	Myopic	One-shot game	Individual
[19]	Price-anticipating	Foresighted	Repeated game	Individual
This work	Price-anticipating	Foresighted	Repeated game	Aggregate

by the energy source in peak hours and off-peak hours, respectively. The price $p\left(D_{\text{peak}}, D_{\text{off}-\text{peak}}\right)$ is set to be proportional to the average unit cost of electricity generation and distribution, i.e.

$$p\left(D_{\text{peak}}, D_{\text{off}-\text{peak}}\right) = \kappa \frac{C_{\text{peak}}\left(D_{\text{peak}}\right) + C_{\text{off}-\text{peak}}\left(D_{\text{off}-\text{peak}}\right)}{D_{\text{peak}} + D_{\text{off}-\text{peak}}} \tag{2}$$

where κ is the revenue/cost ratio. If $\kappa = 1$, then the billing system is budget-balanced and the utility company charges the consumers only the generating/providing energy costs for the utility (i.e., the utility company does not make money and it serves simply as a benevolent energy provider). If $\kappa > 1$, then the difference between the total charges to the consumers and the total energy cost represent the profit made by the utility company. In this paper, the protocol designer is considered to be benevolent and represent the consumer's interests rather than maximizing the utility company's profit. Thus, $\kappa = 1$ in the subsequent analysis.

We make the following standard assumptions on the benefit function and cost functions throughout this paper.

Assumption 1. *(1) The benefit function $B(x)$ is increasing and concave in x. $B(0) = 0$. (2) The cost functions $C_{\text{peak}}(x)$ and $C_{\text{off}-\text{peak}}(x)$ are increasing and strictly convex in x. $C_{\text{peak}}(0) = C_{\text{off}-\text{peak}}(0)$ and $C'_{\text{peak}}(x) \geq C'_{\text{off}-\text{peak}}(x), \forall x \geq 0$.*

Because the smart meter only periodically sends the usage information to the utility company, monitoring the usage pattern $(D_{\text{peak}}, D_{\text{off-peak}})$ is imperfect due to the monitoring noise. Hence, the price is also a noisy function of $(D_{\text{peak}}, D_{\text{off-peak}})$. In particular, we model this by adding a noise term to the price function

$$\hat{p}\left(D_{\text{peak}}, D_{\text{off-peak}}\right) = p\left(D_{\text{peak}}, D_{\text{off}-\text{peak}}\right) + \epsilon \tag{3}$$

where ϵ is induced by the monitoring noise of the usage pattern. Because we assume that the schedulable demand vector $\boldsymbol{d}$ is fixed, the price $p\left(D_{\text{peak}}, D_{\text{off-peak}}\right)$ only depends on the consumption scheduling actions $\boldsymbol{a}$ of the consumers. Therefore, we alternatively write the price $p(\boldsymbol{a})$ as a function of the scheduling action profile $p(\boldsymbol{a})$. Note that $p(\boldsymbol{a})$ is the expected price if $\boldsymbol{a}$ is taken and $\hat{p}(\boldsymbol{a})$ is the actual realized price if $\boldsymbol{a}$ is taken in the noisy environment. When consumers are making scheduling decisions, only $p(\boldsymbol{a})$ is important since consumers can only compute the expected price but not the actual price which has not been realized yet.

By taking both energy consumption benefit and payment into consideration, consumer n's (expected) utility can be written as [2]

$$u_n\left(\boldsymbol{a}\right) = b_n(a_n) - p\left(\boldsymbol{a}\right) d_n. \tag{4}$$

A billing cycle consists of L days. At the beginning of each billing cycle, the consumers determine the consumption scheduling actions for the next L days. At the end of each billing cycle, the utility company posts bills as well as the electricity price of the previous cycle. Using the pricing information, the consumers are able to infer the aggregate daily usage pattern in the last L days. However, since the price does not perfectly reflect the aggregate usage pattern due to the noise term, the consumers' knowledge of the aggregate usage pattern is imperfect.

3.2 Energy consumption game: stage game

The consumers' energy consumption scheduling action profile $\boldsymbol{a}$ determines the price at which the consumers will be charged, thereby leading to the following non-cooperative game $\mathcal{G} = \langle \mathcal{N}, \mathcal{A}, \{u_n(\cdot)\}_{n \in \mathcal{N}} \rangle$ among consumers. The elements of this game are elaborated below:

- Players: consumers in the set $\mathcal{N}$.
- Actions: each consumer $n \in \mathcal{N}$ selects its energy consumption scheduling action $a_n \in \mathcal{A} = [0, 1]$, i.e., the fraction of its energy consumption in peak hours [3].
- Payoffs: the utility for consumer n is the benefit obtained by the energy consumption minus the payment to the utility company as in (4). By separating consumer n's action from other consumers' actions, the utility can also be written as

$$u_n\left(a_n; \boldsymbol{a}_{-n}\right) = b_n\left(a_n\right) - p\left(a_n, \boldsymbol{a}_{-n}\right) d_n \tag{5}$$

where $\boldsymbol{a}_{-n}$, by convention, is the action profile of consumers except consumer n.

Since consumers are self-interested, they will want to selfishly maximize their own utilities. We use Nash equilibrium as the solution concept of this energy consumption game.

Definition 1 (NE). *A Nash equilibrium action profile $\boldsymbol{a}^{NE}$ is such that, $\forall n \in \mathcal{N}, \forall \tilde{a}_n, u_n\left(a_n^{NE}, \boldsymbol{a}_{-n}^{NE}\right) \geq u_n\left(\tilde{u}_n, \boldsymbol{a}_{-n}^{NE}\right)$.*

In NE, no agent can improve its own utility by unilaterally changing its own action. Theorem 1 proves the existence of NE of the considered energy consumption game.

Theorem 1. *There exists at least one NE in the energy consumption game.*

Proof. We will show that energy consumption game is a strictly concave N-person game. The existence of NE for this type of games then directly results from [21].

In order to show that the game is concave, consider any consumer n. We need to show that the utility function $u_n(\boldsymbol{a})$ is strictly concave in a_n. We investigate the two parts consisting the utility function separately as follows.

(1) By Assumption 1, we know that $B(\cdot)$ is a concave function. It is then straightforward to see that $b_n(a_n)$ is concave in a_n.
(2) Now consider the price term. We perform second-order derivative with respect to a_n,

$$\frac{\partial p(\boldsymbol{a})}{\partial a} = \frac{d_n^2 \left(C_{\text{peak}}'' (D_{\text{peak}}) + C_{\text{off-peak}}'' (D_{\text{off-peak}}) \right)}{\sum_n d_n} > 0 \tag{6}$$

Hence, the payment term is strictly convex in a_n.

In sum, $u_n(\boldsymbol{a})$ is strictly concave in a_n. Thus, there exists at least one NE in the energy consumption game. □

It is well known that NE is often not efficient in terms of Pareto-optimality. If an action profile is Pareto-optimal, then no consumer can gain a higher utility without decreasing at least one other consumer's utility by using a different action profile. We provide a formal definition as follows.

Definition 2 (Pareto-optimal (PO)). *An action profile $\boldsymbol{a}$ is Pareto-optimal if there does not exist any other action profile $\tilde{\boldsymbol{a}}$ such that $u_n(\tilde{\boldsymbol{a}}) \geq u_n(\boldsymbol{a}), \forall n$.*

Even though a PO action profile $\boldsymbol{a}^{\text{PO}}$ is superior to $\boldsymbol{a}^{\text{NE}}$, consumers do not have incentives to automatically adopt this action profile in the energy consumption game since they are self-interested. This is because for any action profile that is not a NE, there will always be some consumers who want to schedule a different amount of energy consumption in peak hours to increase their own utilities. In order to provide incentives for the consumers to play the PO action profile, in this paper, we will design a protocol that exploits the ongoing nature of consumers' energy consumption interactions. For notational simplicity, we denote $u_n(\boldsymbol{a}^{\text{NE}})$ as u_n^{NE} and $u_n(\boldsymbol{a}^{\text{PO}})$ as u_n^{PO}. Before we proceed to that, we provide a simple example to illustrate the difference between $\boldsymbol{a}^{\text{NE}}, \boldsymbol{a}^{\text{PO}}$, and the inefficiency of $\boldsymbol{a}^{\text{NE}}$.

Example (Two consumers). Consider two consumers: $N = 2$. Let $\kappa = 1, b_{\text{peak}} = 2, b_{\text{off-peak}} = 1, d_1 = d_2 = 1$ and $B(x) = x$. The cost functions are simple quadratic functions as in [22]: $C_{\text{peak}}(x) = x^2, C_{\text{off-peak}} = 0.5x^2$. The utility of consumer n is therefore $u_n(a_1, a_2) = 1 + a_n - 0.5 \left[(a_1 + a_2)^2 + 0.5 (2 - a_1 - a_2)^2 \right]$.

In this two-consumer game example, the unique symmetric NE action profile is $a_1^{\text{NE}} = a_2^{\text{NE}} = 2/3$. That is, both

consumers schedule 2/3 of their energy consumption in peak hours and no one wants to unilaterally schedule a different amount since that will only reduce its own utility. Thus, the corresponding utilities are $u_1^{\text{NE}} = u_2^{\text{NE}} = 2/3$. However, the sum utility is maximized when both users choose the action profile $a_1^{\text{PO}} = a_2^{\text{PO}} = 0.5$, which is the unique symmetric PO action profile. The corresponding utilities are $u_1^{\text{PO}} = u_2^{\text{PO}} = 5/4$. The sum utility obtained when taking this action profile is 87.5 % higher than that by taking the NE action profile (i.e., 4/3), thereby indicating the inefficiency of NE.

3.3 Repeated energy consumption game

In the repeated energy consumption game, consumers play the energy consumption stage game repeatedly. At the beginning of each billing cycle, the consumers determine the scheduling actions for this billing cycle based on the previous energy consumption history. Specifically, consumer n determines a_n^t for $t = mL+1, mL+2, \ldots, (m+1)L$ at the beginning of billing cycle L based on the prices $p_n^t, t = (m-1)L+1, (m-1)L+2, \ldots, mL$ of the previous billing cycle. Alternatively, consumer n can also determine $a_n^{mL+\tau}$ when day $mL + \tau$ comes. However, because consumers do not obtain additional information about the past energy consumption histories during one billing cycle (i.e., prices are announced only at the end of each billing cycle), these two methods are equivalent for analysis. In this paper, we consider that consumers determine the scheduling actions at the beginning of each billing cycle. Consumers do not directly know the energy consumption history of other consumers. They can only infer this history according to a public signal $z \in \mathcal{Z}$. In this paper, the public signal is binary, representing whether the statistical test is passed or not. Based on this information, the consumers take actions to maximize their *long-term utilities*, which is defined as follows

$$U_n = (1 - \delta) \sum_{t=0}^{\infty} \delta^t E \left\{ u_n(\boldsymbol{a}^t) \right\} \tag{7}$$

That is, the long-term utility is the normalized sum of the discounted expected stage utilities which are induced by the strategy where the expectation is taken over the probability over different sequences of action profiles $\{\boldsymbol{a}^t\}_{t=1}^{\infty}$.

Figure 3 illustrates the timeline of the system and the decision flow of consumers.

3.4 Problem formulation

In this paper, we assume a benevolent designer (i.e., $\kappa = 1$) who aims to design a protocol that maximizes the expected sum utility of all consumers, i.e., $\sum_n E \{u_n(\boldsymbol{a})\}$.

Nevertheless, other performance criteria can also be used in our framework. Formally, the designer wants to achieve an action profile $\boldsymbol{a}^{\text{PO}}$ that solves the optimization below:
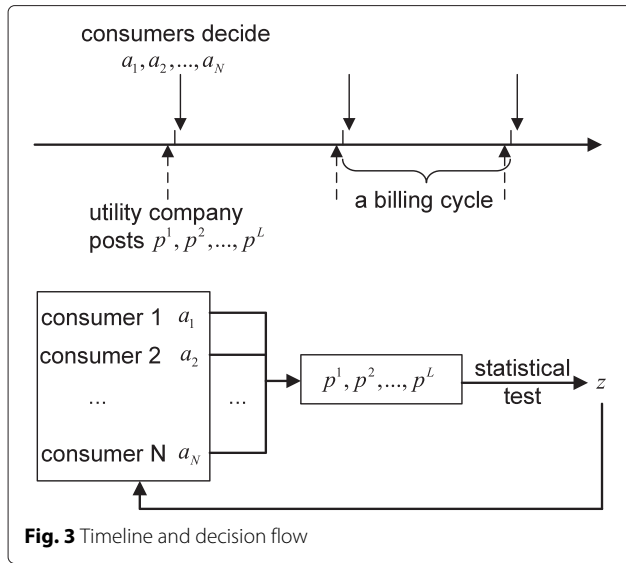
Fig. 3 Timeline and decision flow

$$\text{maximize} \quad \sum_n u_n(\boldsymbol{a})$$
$$\text{subject to} \quad u_n(\boldsymbol{a}) \geq u_n\left(\boldsymbol{a}^{\text{NE}}\right), \forall n \qquad (8)$$
$$\boldsymbol{a} \text{ is Pareto} - \text{optimal}$$

The first constraint requires that no consumer's utility is decreased by applying such a protocol than NE. The second constraint requires the solution be Pareto-optimal. To provide consumers with incentives to take this action profile, the designer designs a protocol that exploits the ongoing interactions of consumers. Let Ψ denote a protocol and $V(\Psi) = \sum_n E\{u_n(\Psi)\}$ as the protocol efficiency where $E\{u_n(\Psi)\}$ is the expected utility of consumer n if all consumers follow the protocol Ψ. The protocol design problem is then formally presented as follows

$$\text{maximize} \quad V(\Psi)$$
$$\text{subject to} \quad \Psi \text{ is incentive compatible} \qquad (9)$$

The constraint requires that all consumers have incentives to follow the protocol Ψ.

4 Protocol based on review strategies

A possible incentive compatible protocol is the grim-trigger strategy [4] which uses the strongest punishment that can be imposed upon deviation. In the trigger strategy, following any point in time at which there is any evidence that any consumer had deviated from any previous recommendation, the protocol designer recommends that each consumer consumes afterwards a large amount of energy in peak hours (i.e., revert to NE action profile forever afterwards). In the two-consumer example, the protocol designer recommends both consumers to take the PO actions $a_1^{\text{PO}} = a_2^{\text{PO}} = 0.5$ in each billing cycle. If there is a deviation, the protocol designer recommends both consumers to take the NE actions $a_1^{\text{NE}} = a_2^{\text{NE}} = 2/3$ forever (i.e., all the way until the consumers unsubscribe

from this service). Therefore, the long-term utility that consumer 1 can receive by following the recommendation is $u_1^{\text{PO}} + \sum_{t=1}^{\infty} \delta^t u_t^{\text{PO}}$, and the long-term utility by deviation is $u_1^d + \sum_{t=1}^{\infty} \delta^t u_t^{\text{NE}}$ where $u_1^d > u_1^{\text{PO}}$ is the one-shot utility by deviation. Because $u_1^{\text{PO}} > u_1^{\text{NE}}$, the grim-trigger strategy may provide sufficient incentives for the consumers to play the PO actions. Since no users want to deviate, the system never enters the NE equilibrium phase, thereby leading to the highest social welfare. However, if there is noise in the prices, the designer will determine that some users have deviated by *accident* almost surely, and hence, the system will eventually enter the punishment phase with probability one and stay there indefinitely. This leads to the lowest social welfare. Therefore, a protocol which allows stopping the punishment after a certain time is needed in such imperfect monitoring scenarios.

4.1 Protocols based on review strategies

In a protocol based on review strategies, there are two types of phases: review phases and punishment phases. Each review phase consists of one billing cycle. Since the length of the billing cycle is designed by the protocol designer, the length of a review phase is a design parameter. Each punishment phase consists of a certain integer number K of billing cycles, and hence, it has KL time slots (i.e., days). The protocol designer recommends a review strategy σ^R for all consumers as follows: the recommended scheduling action profile is $\boldsymbol{a}^{\text{PO}}$ for each day in the review phase; the recommended scheduling action profile is $\boldsymbol{a}^{\text{NE}}$ for each day in the punishment phase. At the end of each billing cycle in the review phase, a statistical test is performed (either by the consumers themselves or the utility company) on the prices of the previous cycle to determine whether other consumers followed or not the recommended strategy. A signal $z \in \mathcal{Z} = \{0, 1\}$ is generated based on the test results with $z = 1$ representing that all consumers followed the recommended strategy and $z = 0$ representing that some consumers deviated from the recommended strategy. If $z = 1$, the system moves to another review phase; if $z = 0$, the system moves to the punishment phase. When a punishment phase ends, the system automatically moves to a new review phase. This is the protocol based on review strategies depicted in Fig. 4.

Importantly, note that the utilities obtained by the consumers in the punishment phase are lower than in the review phase and they are the same as in the scenario where there was no review strategy protocol being deployed. Therefore, the review strategy-based protocol is guaranteed to achieve better system performance than that which can be obtained without deploying the protocol. However, the designer still needs to carefully design the protocol such that the system enters the punishment phase as rarely as possible.
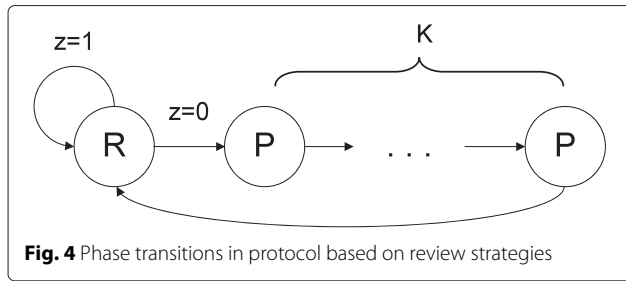
Fig. 4 Phase transitions in protocol based on review strategies

Given this review strategy structure, the protocol can be fully characterized by the length of the review phase (i.e., the billing cycle length) L, the length of the punishment phase KL and the statistical test G. Therefore, we write a protocol based on review strategies as $\Psi(L, K, G)$.

In Table 2, we summarize what information the designer and the consumers have and what computations that they need to perform. Note that when performing appliance level load scheduling, consumer takes into account the specific electricity demands. However, since the protocol will be designed to be incentive-compatible, consumers will perform scheduling according to $\boldsymbol{a}^{\text{PO}}$ in the review phases and according to $\boldsymbol{a}^{\text{NE}}$ in the punishment phases.

4.2 Performance metrics

To evaluate the protocols based on review strategies, we need to define several performance metrics. First, the protocol should be IC since the consumers are self-interested and will only follow recommendations when this is in their self-interest. Particularly, consumers should have incentives to take the PO scheduling actions in the review phase.

Proposition 1. *In a PO action profile, a self-interested consumer will not want to schedule less than the recommended energy consumption in peak hours in the stage game.*

Proof. If consumer n can increase its stage utility by choosing $a_n^d < a_n^{PO}$, then all other consumers' utilities are also increased because the price is reduced as the aggregate usage in peak hours decreases. This causes a contradiction to the definition of "Pareto-optimality". $\square$

Proposition 1 states that, given the daily desired energy demand of a consumer, if the designer recommends to it the PO scheduling action, then the self-interested consumer will schedule no less than the PO energy consumption in peak hours in order to maximize its own utility assuming other consumers are complying with the recommended PO scheduling. Note that this applies even in the case when the consumer's daily demand is low since the consumer is simply re-scheduling its desired amount of energy consumption to different hours. In the two-consumer example, given the recommended PO action profile $a_1 = a_2 = 0.5$, a consumer can only increase its utility by scheduling more than half of its energy consumption in peak hours to increase its own utility. Therefore, we only need to focus on the case $a_n^d > a_n^{\text{PO}}$ when considering consumers' incentive problems and denote the corresponding utility by u_n^d for consumer n. Note that consuming more energy in peak hours increases the price and, hence, consumer n's own payment is increased. However, because the costs are shared among the entire consumer set, consumer n is still able to receive a higher utility $u_n^d > u_n^{\text{PO}}$ by unilaterally increasing its own consumption in peak hours when $\nabla_n u_n(\boldsymbol{a}^{\text{PO}}) > 0$.

At the beginning of each billing cycle, the consumers determine the scheduling actions for this billing cycle based on the previous energy consumption history. We will focus on *constant strategies* during a billing cycle, i.e., consumer n chooses a constant scheduling action every day during a billing cycle. However, our analysis can also be extended to more sophisticated strategies (i.e., the consumer may use different scheduling actions during a billing cycle) by taking the equivalent average scheduling action. Let $U_n(\sigma^R | s = 1)$ and $U_n(\sigma^R | s = 0)$ denote the

Table 2 Information and computation of protocol based on review strategies

The design stage	
1.	Each consumer reports d_n and $B(\cdot)$ to designer.
2.	Designer computes $\boldsymbol{a}^{\text{PO}}$ and $\boldsymbol{a}^{\text{NE}}$.
3.	Designer determines the protocol $\phi(L, K, G)$.
4.	Designer informs each consumer about $\boldsymbol{a}^{\text{PO}}, \boldsymbol{a}^{\text{NE}}, p(\boldsymbol{a}^{\text{PO}})$, and $\Psi(L, K, G)$.
The operation stage (each billing cycle m):	
1.	Each consumer performs its load schedule $a_n^t, \forall t = mL + 1, \ldots, (m+1)L$.
2.	Designer determines $\hat{p}^t, \forall t = mL + 1, \ldots, (m+1)L$ based on the actual load schedule.
3.	Each consumer is charged according to its load d_n and the price $\hat{p}^t, \forall t = mL + 1, \ldots, (m+1)L$ at the end of the current cycle.
4.	The statistical test G is performed using $\hat{p}^t, \forall t = mL + 1, \ldots, (m+1)L$ to determine which phase (review or punishment) the next billing cycle is in.

long-term utilities of consumer n at the beginning of a review phase and a punishment phase, respectively, if all consumers follow the strategy σ. Let $\tilde{\sigma}_n^R$ denote a strategy where consumer n deviates to some $a_n^d > a_n^{\text{PO}}$ in the current review phase and follow the review strategy afterwards and all consumers follow σ^R all the time. The following proposition characterizes the IC condition of a protocol.

Proposition 2. *A protocol based on a review strategy σ^R is IC against a strategy $\tilde{\sigma}_n$ with the deviation action $a_n^d > a_n^{PO}$ for consumer n if and only if*

$$U_n\left(\sigma^R | s = 1\right) \geq U_n\left(\tilde{\sigma}_n | s = 1\right) \qquad (10)$$

Proof. This directly results from one-shot deviation principle in repeated games [4]. $\qquad \square$

Proposition 2 uses the one-shot deviation principle in repeated games and shows that if a consumer cannot gain by unilaterally deviating from the recommended strategy σ^R in the current billing cycle and following afterwards, it cannot gain by switching to any other strategies either and vice versa. In the punishment phase, the recommended strategy is the NE scheduling action profile $\boldsymbol{a}^{\text{NE}}$, and hence, the consumers will not have incentives to deviate from the recommended action. Therefore, we only need to check whether the consumers have incentives to take the recommended PO profile $\boldsymbol{a}^{\text{PO}}$, i.e., whether the one-shot deviation principle holds in the review phase. The left-hand side of (10) represents the long-term utility by following σ^R, and the right-hand side represents the long-term utility for consumer n by deviating to a_n^d only in the current review phase while all other consumers follow σ^R. In Section 4.2, we will use Proposition 2 to construct IC protocols based on review strategies.

The goal of the protocol designer is to maximize the sum utility of all the users. The maximum sum utility is achieved when all consumers take the PO scheduling actions in each time slot, denoted by $V^* = \sum_n u_n^{\text{PO}}$. We call V^* the "first-best" utility which yields the maximum sum utility for the consumers. The efficiency loss of a protocol is defined by $C(\Psi) = (V^* - V(\Psi))/V^*$.

Definition 3. *A protocol Ψ is said to be Δ-Pareto optimal (Δ-PO) if $C(\Psi) < \Delta$.*

A Δ-PO protocol yields a sum utility no less than $1 - \Delta$ of the first-best sum utility. If a protocol Ψ prescribes the scheduling action profile $\boldsymbol{a}^{\text{PO}}$ every day regardless of the history, then $V(\Psi) = V^*$ and thus Ψ achieves the first-best sum utility (0-PO). However, such a protocol is not IC, and hence, the first-best is not achievable.

4.3 Statistical test

In this subsection, we discuss how the public signal based on the prices is constructed. The consumers cannot know the actual scheduling actions made by the other consumers but can only observe the prices announced by the utility company at the end of each billing cycle. According to (3), if all consumers take the recommended optimal scheduling actions, the corresponding price on that day is $\hat{p}\left(\boldsymbol{a}^{\text{PO}}\right) = p\left(\boldsymbol{a}^{\text{PO}}\right) + \epsilon$, where

$$p\left(\boldsymbol{a}^{\text{PO}}\right) = \frac{C_{\text{peak}}\left(\sum_n a_n^{\text{PO}} d_n\right) + C_{\text{off-peak}}\left(\sum_n \left(1 - a_n^{\text{PO}}\right) d_n\right)}{\sum_n d_n}$$

$$(11)$$

If there are some consumers who deviated by consuming more than the recommended amount of energy during the peak hours, the price of that day is increased[4]. Due to the monitoring noise, the actual price is $\hat{p}\left(a_n^d, \boldsymbol{a}_{-n}^{\text{PO}}\right) = p\left(a_n^d, \boldsymbol{a}_{-n}^{\text{PO}}\right) + \epsilon$.

The statistical test determines whether the consumer set was following the recommended strategy in the previous billing cycle by comparing the average value of the prices to a threshold value,

$$z = \begin{cases} 1, \text{if } \frac{1}{L}\sum_{t=1}^{L} \hat{p}^t < \frac{1}{L}\sum_{t=1}^{L} p\left(\boldsymbol{a}^{\text{PO}}\right) + p_{th} \\[2mm] 0, \text{if } \frac{1}{L}\sum_{t=1}^{L} \hat{p}^t \geq \frac{1}{L}\sum_{t=1}^{L} p(\boldsymbol{a}^{\text{PO}}) + p_{th} \end{cases} \qquad (12)$$

We are interested in the following two kinds of error probabilities by performing the statistical test.

- False alarm probability $q_F(L, p_{\text{th}})$: the probability that $z = 0$ when all consumers follow the recommended strategy σ^R, i.e., they all scheduling the optimal amount of energy consumption in peak hours.
- Miss detection probability $q_{M,n}(L, p_{th})$: the probability that $z = 1$ when consumer n deviates to the action a_n^d in the previous billing cycle, i.e., it schedules more than the optimal amount of energy consumption in peak hours while all other consumers take the optimal scheduling.

Note that our design can also be extended to analyze the collusion problem where a subset of consumers colludes in order to gain higher utilities. In that case, we can simply take the consumers who collude as a single consumer with its demand being the aggregate demand.

The following proposition characterizes the impact of the threshold and the billing cycle length on the error probabilities.

Proposition 3. $\forall p_{th} \in \left[0, p\left(a_n^d, \boldsymbol{a}_{-n}^{\text{PO}}\right) - p\left(\boldsymbol{a}^{\text{PO}}\right)\right], \lim_{L \to \infty} q_F\left(L, p_{th}\right) = 0$ *and* $\lim_{L \to \infty} q_{M,n}\left(L, p_{th}\right) = 0$.

Proof. By the law of large numbers, the sample averages converge in probability and almost surely to the expected value as the sample number tends to infinity. Therefore, when the threshold is large than 0, the false alarm probability goes to 0 and when the threshold is smaller than $p\left(a_n^d, \boldsymbol{a}_{-n}^{\mathrm{PO}}\right) - p\left(\boldsymbol{a}^{\mathrm{PO}}\right)$, the miss detection probability goes to 0. □

Proposition 3 informs the protocol designer's selection of the statistical test (i.e., the test threshold). It also proves that, in order to accurately detect self-interested consumers' excessive usage in peak hours, a longer billing cycle should be chosen such that the monitoring mitigated. However, a very long billing cycle also reduces the consumers' valuation of utilities in the future billing cycles, and hence, the punishment may not be strong enough to induce consumers' compliance. Therefore, the optimal billing cycle as well as the punishment length should be carefully designed to induce the optimal performance of the smart grid system.

5 Design and performance analysis

In this section, we design the optimal incentive-compatible protocols based on review strategies and analyze their performance. The outline of this section is as follows.

- We first establish the existence of IC protocols based on review strategies and provide the IC conditions such that the consumers have incentives to perform the recommended scheduling in the review phase.
- Next, we propose a greedy algorithm to determine the optimal design of review strategies.
- We then evaluate the performance of the optimal protocol based on review strategies. Specifically, the proposed protocol is able to achieve Δ-PO of the first-best efficiency.

5.1 Incentive-compatibility

Recall that Proposition 2 provides us with a simple method to determine whether a protocol based on review strategies can be IC. In this subsection, we study how to design protocol parameters according to Proposition 2. To do this, we need to compute $U_n(\sigma^R|s=1)$, $U_n(\sigma^R|s=0)$, and $U_n(\tilde{\sigma}_n|s=1)$ where $s=1$ denotes that the system is in a review phase and $s=0$ denotes that the system is in a punishment phase. These utilities are dependent on each other as follows,

$$
\begin{aligned}
&U_n\left(\sigma^R|s=1\right) \\
&= (1-\delta)\left(\sum_{t=0}^{L-1}\delta^t u_n^{PO} + \delta^L \left[q_F\left(L, p_{\mathrm{th}}\right) U_n\left(\sigma^R|s=0\right)\right.\right. \\
&\left.\left. \quad + \left(1-q_F\left(p_{\mathrm{th}}, L\right)\right) U_n\left(\sigma^R|s=1\right)\right]\right)
\end{aligned}
\tag{13}
$$

and

$$
\begin{aligned}
&U_n\left(\sigma^R|s=0\right) \\
&= (1-\delta)\left(\sum_{t=0}^{KL-1}\delta^t u_n^{NE} + \delta^{KL} U_n\left(\sigma^R|s=1\right)\right)
\end{aligned}
\tag{14}
$$

The first term in (13) is the utility in the current review phase. The second term is the continuation utility after the review phase. With probability $1 - q_F(L, p_{\mathrm{th}})$ the system remains in the review phase; with probability $q_F(L, p_{\mathrm{th}})$, the system moves to a punishment phase due to the monitoring error. The utility of consumer n by choosing a deviation scheduling action a_n^d at the beginning of a review phase is given by,

$$
\begin{aligned}
&U_n\left(\tilde{\sigma}_n|s=1\right) \\
&= (1-\delta)\left(\sum_{t=0}^{L-1}\delta^t u_n^d + \delta^L \left[q_{M,n}\left(L, p_{\mathrm{th}}\right) U_n\left(\sigma^R|s=1\right)\right.\right. \\
&\left.\left. \quad + \left(1-q_{M,n}\left(L, p_{\mathrm{th}}\right)\right) U_n\left(\sigma^R|s=0\right)\right]\right)
\end{aligned}
\tag{15}
$$

The first term in (15) is the utility gain by the deviation in the current review phase which is larger than u_n^{PO}. The second term is the continuation utility after the review phase. With probability $1 - q_{M,n}(L, p_{\mathrm{th}})$ the deviation is detected, so the system moves to a punishment phase; with probability $q_{M,n}(L, p_{\mathrm{th}})$ the system remains in the review phase.

We define the following "incentive ratio" of a protocol based on review strategies for consumer n:

$$
g_n(L, p_{\mathrm{th}}) = \frac{1}{1 - q_F - q_{M,n}} \frac{1 - \delta^L}{\delta^L} \frac{u_n^d - u_n^{\mathrm{PO}}}{u_n^{\mathrm{PO}} - u_n^{\mathrm{NE}}}
\tag{16}
$$

Let us examine the physical meaning of this incentive ratio. Essentially, the numerator represents the long-term utility gain due to deviation, and the denominator represents the maximal long-term utility loss due to the punishment. To enforce consumers to cooperatively optimize their energy consumption, this loss should be positive and larger than the gain obtained when deviating. Therefore, the incentive ratio should be in the range $[0, 1]$. It is worth noting that $g_n(L, p_{th})$ should be strictly less than 1 since the denominator is only an upper bound on the loss induced by the punishment but not the actual loss (which depends on L, K, p_{th}). Theorem 2 provides a condition such that a protocol based on review strategy is IC against a deviation action a_n^d. This condition serves as a guideline for the choices of the proper protocol parameters L, K, and p_{th}.

Theorem 2. *The protocol $\Psi(L, K, p_{th})$ is IC against a_n^d for consumer n if and only if the billing cycle length satisfies*

$0 \leq g_n(p_{th}, L)$ *and the punishment phase length is large enough, i.e.*

$$K \geq \frac{1}{L} \log_\delta (1 - g_n(p_{th}, L)) \quad (17)$$

Proof. According to Proposition 2, we only need to check the utility difference below,

$$
\begin{aligned}
U_n &\left(\sigma^R | s = 1\right) - U_n \left(\tilde{\sigma} | s = 1\right) \\
&= (1 - \delta) \left(\sum_{t=0}^{L-1} \delta^t \left(u_n^{\mathrm{PO}} - u_n^d \right) \right. \\
&\quad \left. + \delta^L (1 - q_F - q_{M,n}) \left(U_n \left(\sigma^R | s = 1\right) - U_n \left(\sigma^R | s = 0\right) \right) \right) \\
&= -(1 - \delta^L) \left(u_n^{\mathrm{PO}} - u_n^d \right) \\
&\quad + \delta^L \left(1 - \delta^{LK} \right) \left(1 - q_F - q_{M,n} \right) \left(u_n^{\mathrm{PO}} - u_n^{\mathrm{NE}} \right)
\end{aligned}
$$
$$(18)$$

In the last equation of $(1 - \delta^L)(u_n^{\mathrm{PO}} - u_n^d)$ is the utility gain in the current review phase billing cycle by deviation, and the remaining term is the utility loss for the future due to the punishment. For Ψ to be IC, the utility loss due to the punishment should exceed the utility gain due to deviation, i.e.,

$$
\begin{aligned}
\left(1 - \delta^L\right) &\left(u_n^{\mathrm{PO}} - u_n^d \right) \leq \\
&\delta^L \left(1 - \delta^{LK} \right) \left(1 - q_F - q_{M,n} \right) \left(u_n^{\mathrm{PO}} - u_n^{\mathrm{NE}} \right)
\end{aligned}
$$
$$(19)$$

The sufficient and necessary condition for the above to hold is $0 \leq g_n(L, p_{th}) < 1$ and $KL \geq \log_\delta \left(1 - g_n(L, p_{th})\right)$. □

5.2 Optimal protocol parameters

We first determine the efficiency of a given protocol. If a protocol is IC, then all consumers follow the recommended strategy and play the PO action profile. Therefore, the efficiency depends on the probability that the system is in review phases and punishment phases due to monitoring errors. Denote these two probabilities by $\eta_R(\Psi)$ and $\eta_P(\Psi) = 1 - \eta_R(\Psi)$, respectively. The efficiency of an IC protocol is thus $V(\Psi) = \sum_n (u_n^{\mathrm{PO}} \eta_R(\Psi) + u_n^{\mathrm{NE}} \eta_P(\Psi))$ where $\eta_R(\Psi)$ is determined in the following Lemma.

Lemma 1. $\eta_R(\Psi) = \frac{1}{1 + K q_F(p_{th}, L)}$

Proof. Solving the stationary distribution of the Markov chain in Fig. 6 yields the result. The transition probabilities of the chain are R-R with probability $1 - q_F$, R-P with q_F, P-P and P-R with probability 1. □

Lemma implies that in order to maximize the system efficiency, the protocol designer should choose L, K, p_{th}

such that $K q_F(L, p_{th})$ is minimized subject to IC conditions in Theorem 2. These design parameters are coupled in a complex manner, and thus, to find the optimal design parameters, it is better to work backwards.

Step 1. We first determine the optimal $K^*(L, p_{th})$ given L, p_{th}. Therefore, the false alarm probability $q_F(L, p_{th})$ and the miss detection probability $q_{M,n}(L, p_{th})$ are also determined. If $0 \leq g_n(L, p_{th})$, the optimal K is chosen as

$$K^*(L, p_{th}) = \left\lceil \max_n \left(\frac{1}{L} \log_\delta \left(1 - g_n(p_{th}, L) \right) \right) \right\rceil \quad (20)$$

Step 2. Given L, the statistical test determines $q_F(L, p_{th})$ and $q_{M,n}(L, p_{th})$, $\forall n \in \mathcal{N}$. Note that $K^*(L, p_{th})$ depends on the statistical test through the term $q_F(L, p_{th}) + q_{M,n}(L, p_{th})$. Therefore, the optimal statistical threshold is chosen as

$$p_{th}^* = \arg\min_{p_{th}} K^*(L, p_{th}) q_F(L, p_{th}) \quad (21)$$

In general, the statistical test threshold p_{th} has two opposite effects on $K^*(L, p_{th}) q_F(L, p_{th})$. Minimizing $q_F(L, p_{th})$ often leads to a large $q_{M,n}(L, p_{th})$, and hence, $q_F(L, p_{th}) + q_{M,n}(L, p_{th})$ may also be large. This by (16) induces a large $K^*(L, p_{th})$. Therefore, the protocol designer has to trade-off when selecting p_{th} between minimizing the false alarm probability $q_F(L, p_{th})$ and the punishment phase length $K^*(L, p_{th})$.

Step 3. The previous two steps provide the optimal p_{th}^* and $K^*(L, p_{th}^*)$ given the review phase length L. However, the space of L includes all positive integer numbers and is infinite. In the following, we determine the upper bound on L such that an IC protocol can be designed.

Proposition 4. *If $\Psi(L, K, p_{th})$ is IC, then*

$$L \leq \min_n \log_\delta \frac{u_n^d - u_n^{PO}}{u_n^d - u_n^{NE}} \quad (22)$$

Proof. Note that

$$g_n(L, p_{th}) > \frac{1 - \delta^L}{\delta^L} \frac{u_n^d - u_n^{PO}}{u_n^{PO} - u_n^{NE}} \quad (23)$$

If (22) does not hold, then there must exist $n \in \mathcal{N}$ such that $g_n(L, p_{th}) > 1$ which violates the IC condition in Theorem 2. Therefore, for a protocol to be IC, (22) must be satisfied. □

Proposition 4 leads to a crucial trade-off of the review phase length. On one hand, the protocol designer wants to choose a longer review phase period L because it improves the monitoring accuracy, and hence, there is a smaller probability that the system goes into a punishment phase due to monitoring errors. On the other hand, a longer review increases the current gain of a consumer in the review phase upon deviation while it reduces the future

loss due to the punishment because of the discounting of the future utility. This requires a stronger punishment (longer punishment phase), and hence, it induces more energy consumption in peak hours as the system stays longer in the punishment phase. More importantly, if the review phase is too long, then even the strongest punishment (i.e., a trigger strategy that prescribes to stay in the punishment phase forever upon deviations) is not able to provide the sufficient incentives for the consumers to schedule the PO energy consumption in the current review phase. Therefore, given consumers' valuation of the future utility (i.e., the discount factor δ) and the structure of the stage game (i.e., $u_n^d, u_n^{\mathrm{PO}}, u_n^{\mathrm{NE}}, \forall n \in \mathcal{N}$), there is a maximum length of the review phase, and hence, the billing cycle should not be too long. To make the protocol IC, the protocol designer must choose a review phase no longer than the upper bound determined in Proposition 4. The optimal review phase length L (i.e., billing cycle) is thus such that it minimizes the product of $K^*\left(L, p_{\mathrm{th}}^*\right)$ and $q_F\left(p_{\mathrm{th}}^*, L\right)$ which have been determined in the previous two steps for a fixed L. Based on the above three design steps, we propose a greedy algorithm (presented in Table 2) to determine the optimal protocols based on review strategies which requires only finite iterations on L. If the candidate threshold p_{th} belongs to a continuous interval, then we need to quantize p_{th} to solve (19) in finite iterations, and hence, the algorithm leads to a sub-optimal protocol.

5.3 Performance evaluation

In the previous subsection, we determine the optimal protocol design. In this subsection, we characterize the performance of these optimal protocols. Specifically, we are interested in determining whether a protocol based on review strategies Ψ can be Δ-PO.

Recall that the first-best efficiency is $V^* = \sum_n u_n^{\mathrm{PO}}$, and the sum utility of a given IC protocol can be determined using the result in Proposition 3. Therefore, if $\eta_R(\Psi)$ is close enough to 1, then $(V(\Psi) - V^*)/V^*$ can be made less than a given Δ. According to Proposition 3, this can be done if q_F is close enough to 0 and if K is finite. That is, an accurate enough statistical test is required. By Proposition 4, if L is long enough, then the monitoring can be accurate enough. However, a long L can be chosen only when the consumers valuate the future utilities sufficiently high enough. The following theorem characterizes the condition when a protocol is Δ-PO.

Theorem 3. $\forall p_{th} \in \left[0, p_n^d - p_n^{PO}\right]$, *for a given* $\Delta \in [0, 1]$, *there exists* $\delta_{min} \in (0, 1)$, *such that* $\forall \delta \geq \delta_{min}$ *there exists a protocol* $\Psi(L, K, p_{th})$ *such that it is* Δ-PO.

Proof. Let us write Δ-PO condition in terms of η_R and q_F. $V(\Psi) \geq (1 - \Delta)V^*$ implies:

$$
\begin{aligned}
& \eta_R \sum_{n \in \mathcal{N}} u_n^{\mathrm{PO}} + (1 - \eta_R) \sum_{n \in \mathcal{N}} u_n^{\mathrm{NE}} \\
& \geq (1 - \Delta) \sum_{n \in \mathcal{N}} u_n^{\mathrm{PO}} \\
& \Rightarrow \eta_R \geq \frac{(1-\Delta)u_n^{\mathrm{PO}} - u_n^{\mathrm{NE}}}{u_n^{\mathrm{PO}} - u_n^{\mathrm{NE}}} \triangleq A \\
& \Rightarrow q_F(L, p_{\mathrm{th}}) \leq \frac{1-A}{KA}
\end{aligned}
\tag{24}
$$

Note that $\Delta \in (0, 1) \to A \in (0, 1)$, $\Delta = 0 \to A = 1$, and $\Delta = 1 \to A = 0$.

Fix $p_{\mathrm{th}} = \bar{p}_{\mathrm{th}} \in [0, p_n^d - p^{PO}]$ and

$$
K = \bar{K} > \max_n \in \mathcal{N} \frac{u_n^d - u_n^{\mathrm{PO}}}{u_n^{\mathrm{PO}} - u_n^{\mathrm{NE}}}
\tag{25}
$$

Now, we select $L = \bar{L}$ such that

$$
q_F(\bar{L}, \bar{p}_{\mathrm{th}}) \leq \frac{1-A}{KA}
\tag{26}
$$

Such $\bar{L}$ exists due to the law of large numbers (see Proposition 3).

$\Psi(\bar{L}, \bar{K}, \bar{p}_{\mathrm{th}})$ is a Δ-PO protocol by the above construction. Now, we show that it is also IC for δ close enough to 1. The condition under which $\Psi(\bar{L}, \bar{K}, \bar{p}_{\mathrm{th}})$ is an IC protocol is:

$$
\frac{1 - \delta^{\bar{L}}}{1 - \delta^{\bar{L}\bar{K}}} \leq \delta^{\bar{L}}\left(1 - q_F - q_{M,n}\right) \frac{u_n^{\mathrm{PO}} - u_n^{\mathrm{NE}}}{u_n^d - u_n^{\mathrm{PO}}}, \forall n \in \mathcal{N}
\tag{27}
$$

For $\delta \to 1$, the above inequality becomes

$$
\bar{K} \geq \max_n \frac{u_n^d - u_n^{\mathrm{NE}}}{u_n^{\mathrm{PO}} - u_n^{\mathrm{NE}}}
\tag{28}
$$

This implies that there exists a $\delta_{\min}$ close to 1 such that for every $\delta \geq \delta_{\min}$ the above inequality defines an IC protocol (the inequality must be strict). Since (28) is exactly that same as (25), the constructed protocol $\Psi(\bar{L}, \bar{K}, \bar{p}_{\mathrm{th}})$ is IC for $\delta \geq \delta_{\min}$. □

Theorem 3 characterizes the asymptotical performance of the protocols based on review strategies. Specifically, when the consumers highly value their future utilities, then the protocol designer can design a protocol based on review strategies whose sum utility is close enough to the first-best. That is, consumers will comply with the recommended scheduling most of the time.

Corollary 1. *If* $\delta \to 1$, *then the efficiency loss of the optimal protocol goes to 0.*

The corollary states that if the consumers do not discount their future utilities, then the protocol designer can design a review strategy-based protocol which is IC and also asymptotically achieves the full efficiency.

Remark: Our analysis depends on the deviation strategy $\boldsymbol{a}^d$. To compute a_n^d, we consider the unilateral deviation by consumer n at the optimal action profile $\boldsymbol{a}^{\mathrm{OPT}}$, i.e.,

$a_n^d = a_n^* = \arg\max_{a_n} \mu_n\left(a_n; \boldsymbol{a}_{-n}^{\text{OPT}}\right)$. In this way, a consumer chooses to deviate to the action that maximizes its own utility. However, if the consumer is "smarter", it can deviate to a slightly lower action than a_n^* to avoid being detected but still gains some increased utility (of course, since there is noise, the probability of being detected is higher if its selected action is closer to a_n^d). Hence, a more practical way to set a_n^d is by setting a maximal tolerable deviation action $a_n^d = (1 + \gamma)a_n^{\text{OPT}}$ where $\gamma < 1$ depends on the maximal tolerable social welfare loss. Since, by doing this the (one-shot) social welfare loss is at most $u_n\left(\boldsymbol{a}^{\text{OPT}}\right) - u_n\left((1 + \gamma)\boldsymbol{a}^{\text{OPT}}\right)$, the designer can determine γ according to the maximal tolerable social welfare loss and set a_n^d accordingly.

6 Simulations

In this section, we present numerical results and assess the performance of our proposed framework. For illustration purposes, we assume that the benefit function takes the linear form $B_n(a_n) = b_{\text{peak}}a_n d_n + b_{\text{off-peak}}(1 - a_n)d_n$, the cost functions have the quadratic forms [3] $C_{\text{peak}}(x) = c_p x^2$, $C_{\text{off-peak}}(x) = 0.5c_p x^2$. We regard c_p as parameter determined by the power generation technology and the power market and will investigate their impact of the protocol design and performance.

Figure 5 shows how the optimal design of the billing cycle varies depending on consumers' valuation of future utilities (reflected by the discount factor δ). On one hand, a longer billing cycle improves the monitoring accuracy, and therefore, the system has a smaller probability to go into the punishment phase. Hence, the efficiency is increasing in the billing cycle. On the other hand, a too long billing cycle reduces the consumers' valuation of future utilities and, beyond some point, it violates the consumers' incentives to comply. Moreover, the optimal billing cycle depends on how much the consumers value the future utilities which is parameterized by the discount factor δ. When δ is large, consumers value more their future utilities, and hence, a longer billing cycle can be used.

Figures 6 and 7 show the impact of environment parameters (the consumer set size and cost function) on the performance of the proposed protocols. In Fig. 8, we investigate the impact of the consumer set size on the performance of our proposed protocols. As the number of consumer increases, the protocol is able to provide sufficient incentives for consumers to follow the recommended optimal energy scheduling actions, and therefore, an improved performance is attained. Fig. 9 varies the cost function parameter c_1 in peak hours and illustrates the corresponding efficiency for the optimal protocol and two protocols with fixed design parameters. As the cost in peak hours increases, the efficiency decreases. The system efficiency of the optimal protocol is 80 to 250 % higher than that when no incentive protocols are deployed. The optimal protocol also outperforms the fixed protocols by 50 % on average. This highlights that designing the optimal protocol could bring significant improvement on the system efficiency. Varying the cost functions represents different utility functions for the consumers. Note that different optimal scheduling schemes are required for these different utilities, and hence, these simulations
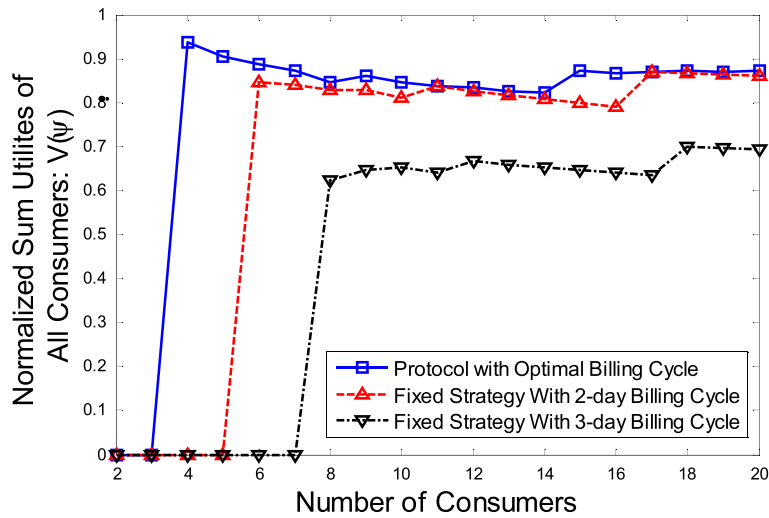

Fig. 5 Optimal billing cycle length (review phase length). ($N = 8, \delta = 0.95, c_p = 0.2, d_n = 1, p_{\text{th}} = 0.5, b_{\text{peak}} = 2, b_{\text{off-peak}} = 1$)
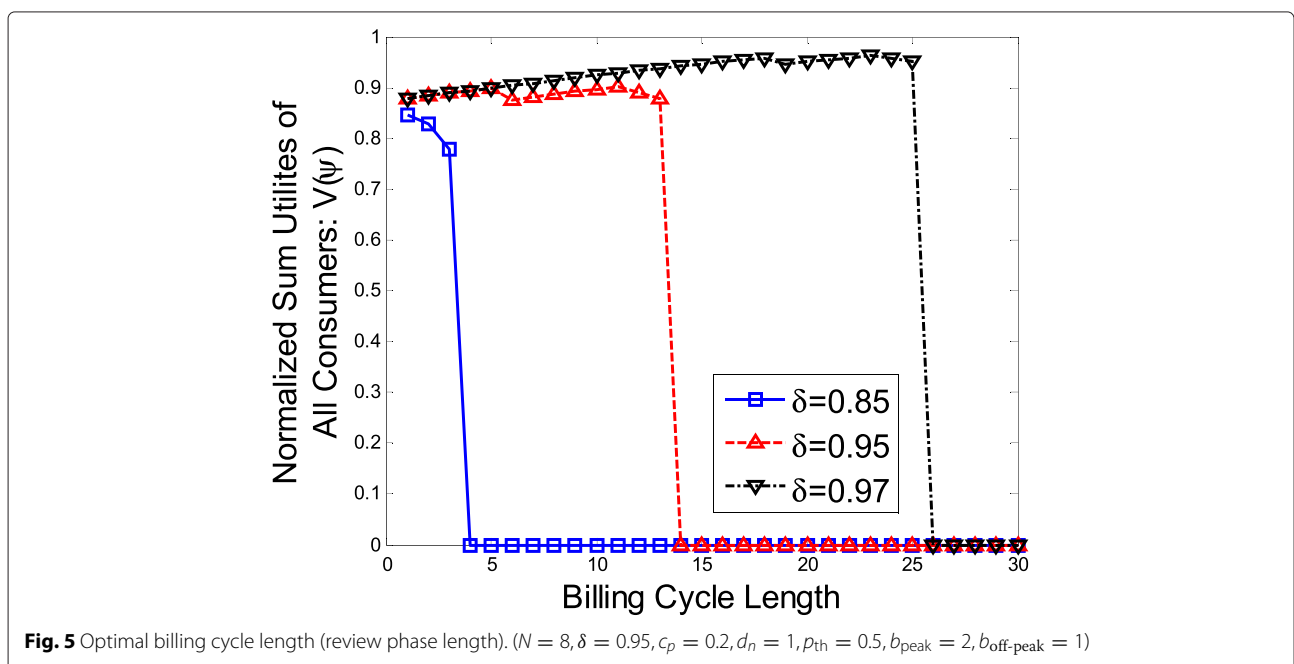
Fig. 6 Impact of the consumer set size on the protocol performance. ($\delta = 0.85, cp = 0.2, d_n = 1, p_{th} = 0.5, b_{peak} = 2, b_{off\text{-}peak} = 1$)

also highlight the capability of the proposed framework to design review strategy-based protocols which are applicable to a variety of DSM schemes.

In Fig. 8, we consider a set of eight consumers with various load demands. The first subplot illustrates the consumers' energy consumption scheduling actions with and without the proposed protocol deployment. With the deployment of the protocol, consumers tend to consume less energy in peak hours. At the same time, their individual utilities are increased as illustrated in the second subplot. This is because the energy cost is smaller, and hence, the bills charged to consumers are smaller. The

third subplot further shows the prices with respect to various discount factors. Note that the price is proportional to the energy cost. Therefore, the lower the price is, the more energy-efficient the system is. When the consumers value more the future utilities (i.e., the discount factor is larger), consumers tend to schedule less energy consumption in peak hours, and therefore, the price is also lower.

Figure 9 illustrates a real-time curve of the prices for the considered set of consumers. Note that the price also reflects the energy generation and transmission cost of this consumer set. The energy price is low in the review phase since consumers schedule less power in peak hours.
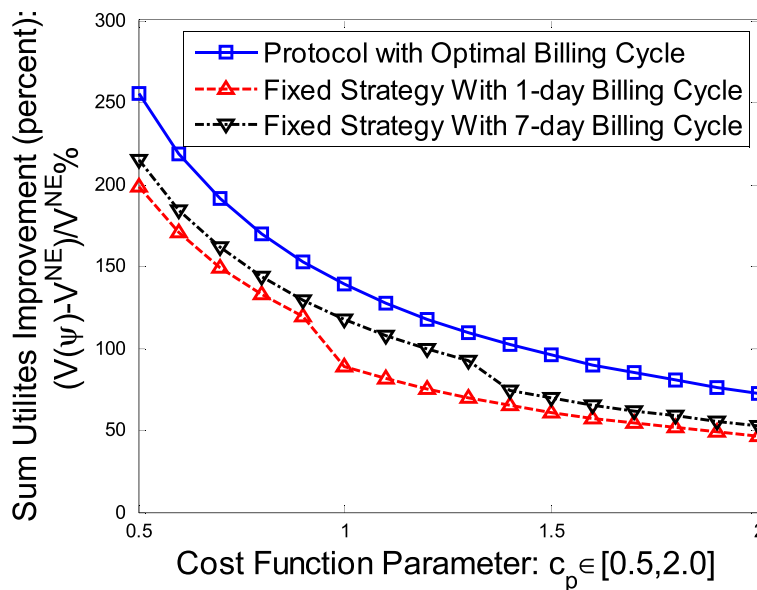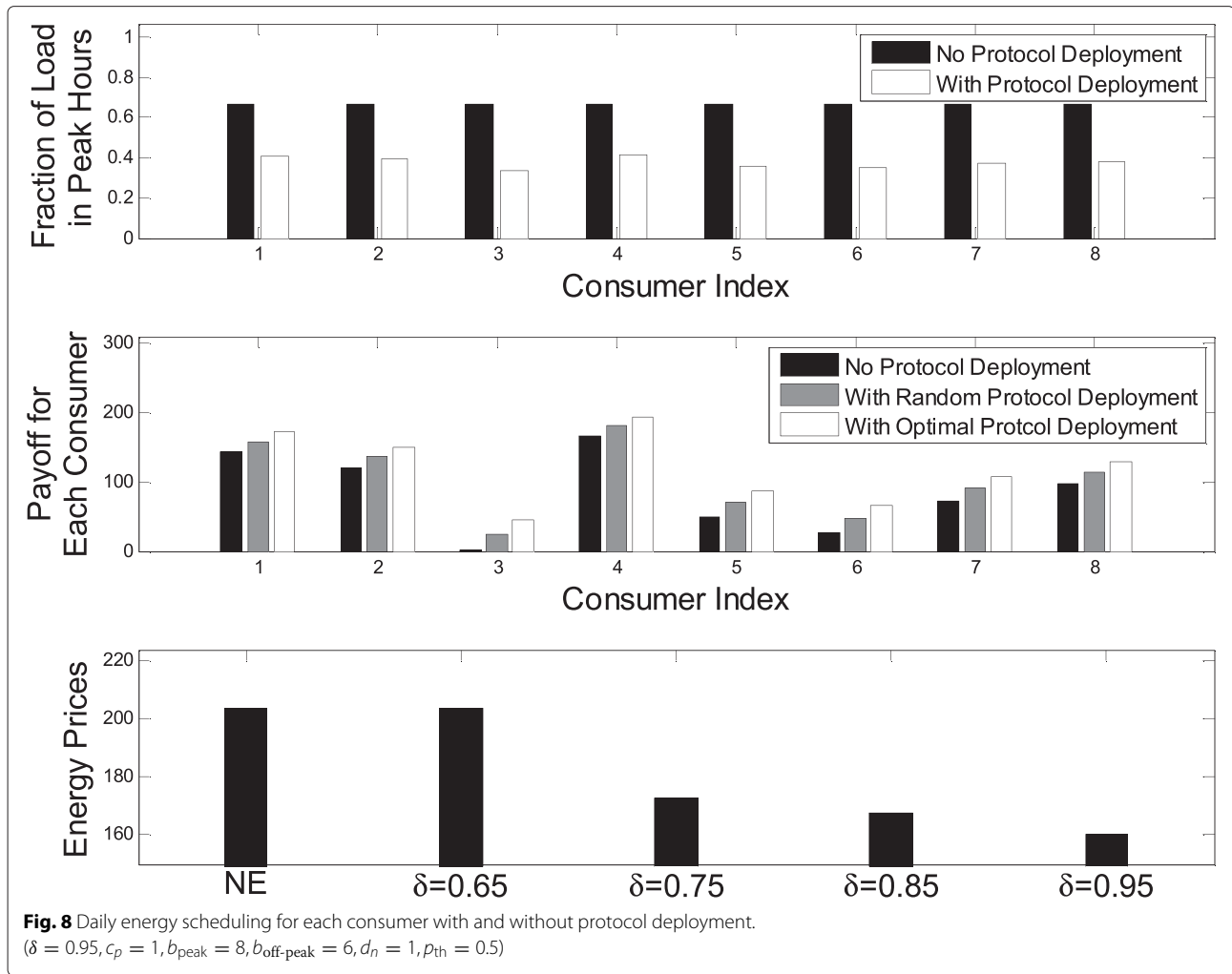

Fig. 7 Impact of the cost functions on the protocol performance. ($N = 8, \delta = 0.99, c_p = 0.2, d = 1, p_{th} = 0.5, b_{peak} = 2, b_{off\text{-}peak} = 1$)

Fig. 8 Daily energy scheduling for each consumer with and without protocol deployment. ($\delta = 0.95, c_p = 1, b_{\text{peak}} = 8, b_{\text{off-peak}} = 6, d_n = 1, p_{\text{th}} = 0.5$)
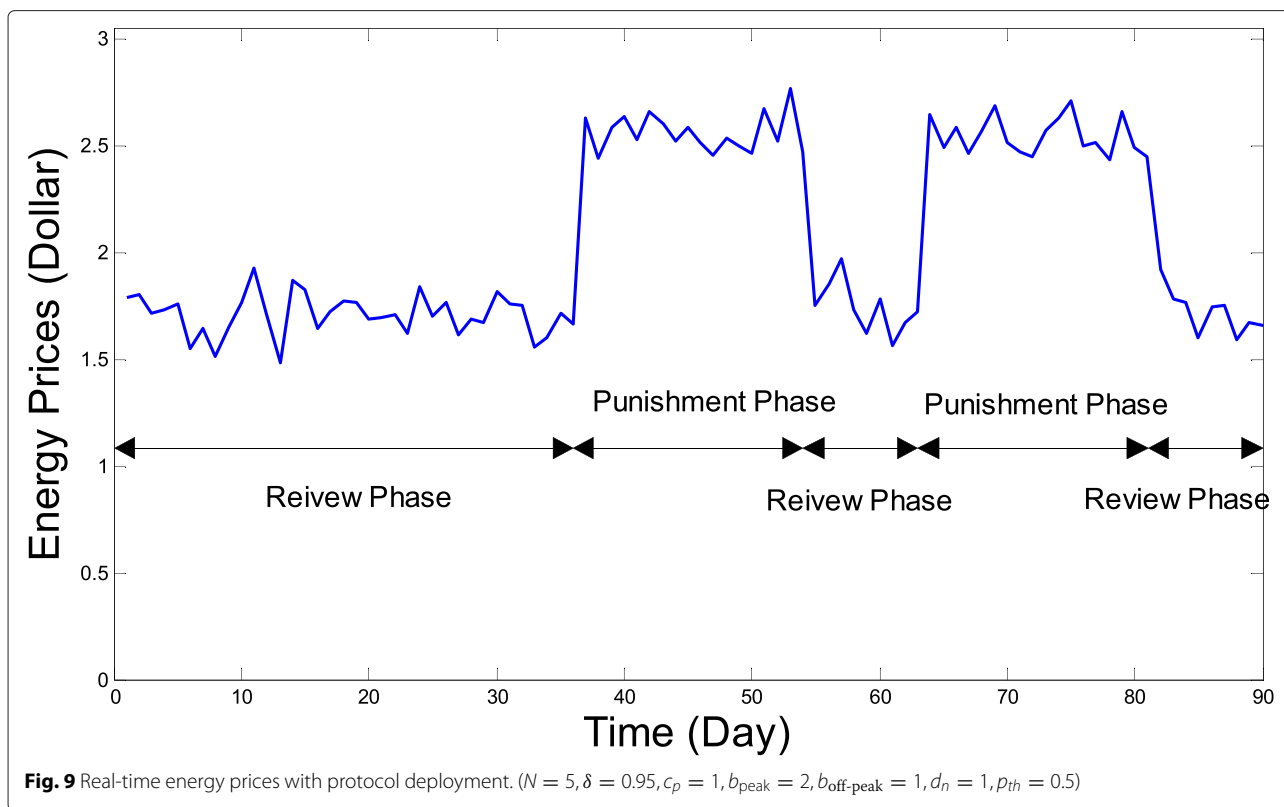
Due to the imperfect monitoring, if the average price of the previous billing cycle exceeds the predetermined threshold, the system goes into the punishment phase. In the punishment phase, the energy price is high because consumers schedule more consumption in peak hours. When the punishment phase ends, the system automatically goes back to a new review phase and the price drops again.

7 Conclusions

In this paper, we augment existing DSM schemes using aggregate usage information of a set of participating consumers by proposing a novel framework which incentivizes consumers to reduce their consumption in peak hours. The proposed technology is review strategy-based protocols. It is general and can be deployed in conjunction with any DSM schemes proposed in smart grids to make it incentive-compatible (i.e., selfish consumers find it in their self-interest to follow it). The proposed protocols are simple, and thus, they are suitable

for practical implementation in smart grids where the energy scheduling actions cannot be perfectly monitored. Even though this paper considers a simplified smart grid model, our analysis provides important and useful insights for designing incentive-compatible demand-side management schemes based on aggregate energy usage information in a variety of practical scenarios.

The success of DSM relies heavily on the energy producer's knowledge of individual consumers' energy usage information. However, a huge concern is that smart meters threaten consumers' privacy as data mining techniques are applied to energy consumption traces in order to infer consumers' habits and behaviors [2, 23, 24]. This information may be used for other purposes besides improving the efficiency of smart grids, thereby giving rise to privacy concerns [25]. To respond to these concerns, in 2010, California's new smart meter privacy law [26] was adopted, which mandates privacy protection for the consumers' energy consumption data. The proposed scheme does not require detailed knowledge of household

Fig. 9 Real-time energy prices with protocol deployment. ($N = 5, \delta = 0.95, c_p = 1, b_{peak} = 2, b_{off\text{-}peak} = 1, d_n = 1, p_{th} = 0.5$)

electricity demand profiles at the appliance level or for a fine time granularity. Instead, it only needs the aggregate usage pattern of individual consumers at large time scales, e.g., the aggregate usage in peak hours of a day and the aggregate usage in off-peak hours of a day. This information is very limited for the detection of the actual consumption traces of individual consumers, and hence, we believe that the proposed scheme can be applied in a wide range of practical deployment scenarios.

Endnotes

[1] The proposed protocol can also be extended to directly recommend per-appliance scheduling actions. However, such an extension would only complicate the notation and presentation of the proposed methods while the proposed methods will remain unchanged.

[2] In practice, the utility function may dynamically change over time. In this paper, we make the common assumption that the utility function is fixed. One way to handle the dynamically changing utility functions is by periodically updating the design of the proposed protocol to adapt to the changes.

[3] The proposed method allows both continuous and discrete value space of actions. For illustration purpose, we consider only continuous value space in this paper.

[4] If consumers consume less than the recommended amount of energy in peak hours, the price is reduced.

Since self-interested consumers only have incentives to deviate to a higher consumption in peak hours (by Proposition 1), we do not regard consuming less in peak hours as a deviation.

References
1. SM Amin, BF Wollenberg, Toward a smart grid: power delivery for the 21st century. IEEE Power Energy Mag. **3**(5), 34–41 (2005)
2. Z Fan, P Kulkarni, S Gormus, C Efthymiou, G Kalogridis, M Sooriyabandara, Z Zhu, S Lambotharan, WH Chin, Smart grid communications: overview of research challenges, solutions, and standardization activities. IEEE Commun. Surv. Tutorials. **15**(1), 21–38 (2013)
3. A-H Mohsenian-Rad, VW Wong, J Jatskevich, R Schober, A Leon-Garcia, Autonomous demand-side management based on game-theoretic energy consumption scheduling for the future smart grid. IEEE Trans. Smart Grid. **1**(3), 320–331 (2010)
4. GJ Mailath, L Samuelson, *Repeated Games and Reputations, vol. 2*. (Oxford University Press, Oxford, UK, 2006)
5. CW Gellings, JH Chamberlin, *Demand-side management: concepts and methods*. (The Fairmont Press Inc., Lilburn, GA, 1987)
6. H Allcott, *Real time pricing and electricity markets*. (Harvard University, 2009)
7. C Triki, A Violi, Dynamic pricing of electricity in retail markets. 4OR. **7**(1), 21–36 (2009)
8. K Herter, Residential implementation of critical-peak pricing of electricity. Energy Policy. **35**(4), 2121–2130 (2007)

9. B Ramanathan, V Vittal, A framework for evaluation of advanced direct load control with minimum disruption. IEEE Trans. Power Syst. **23**(4), 1681–1688 (2008)
10. M Alizadeh, A Scaglione, RJ Thomas, From packet to power switching: digital direct load scheduling. Selected IEEE J. Areas Commun. **30**(6), 1027–1036 (2012)
11. L Jia, L Tong, in *Communication, Control, and Computing (Allerton), 2012 50th Annual Allerton Conference on*. Optimal pricing for residential demand response: a stochastic optimization approach (IEEE Allerton Park, IL, USA, 2012), pp. 1879–1884
12. H-I Su, AEl Gamal, in *Communication, Control, and Computing (Allerton), 2011 49th Annual Allerton Conference on*. Modeling and analysis of the role of fast-response energy storage in the smart grid (IEEE Allerton Park, IL, USA, 2011), pp. 719–726
13. L Huang, J Walrand, K Ramchandran, in *Smart Grid Communications (SmartGridComm), 2012 IEEE Third International Conference on*. Optimal demand response with energy storage management (IEEE Tainan, ROC, 2012), pp. 61–66
14. Z Yu, L Jia, MC Murphy-Hoye, A Pratt, L Tong, Modeling and stochastic control for home energy management. IEEE Trans. Smart Grid. **4**(4), 2244–2255 (2013)
15. N Li, L Chen, SH Low, in *Power and Energy Society General Meeting, 2011 IEEE*. Optimal demand response based on utility maximization in power networks (IEEE San Diego, CA, USA, 2011), pp. 1–8
16. C Joe-Wong, S Sen, S Ha, M Chiang, Optimized day-ahead pricing for smart grids with device-specific scheduling flexibility. IEEE J. Selected Areas Commun. **30**(6), 1075–1085 (2012)
17. C Ibars, M Navarro, L Giupponi, in *Smart grid communications (SmartGridComm), 2010 first IEEE international conference on*. Distributed demand management in smart grid with a congestion game (IEEE Gaithersburg, MD, USA, 2010), pp. 495–500
18. HK Nguyen, JB Song, Z Han, in *Computer Communications Workshops (INFOCOM WKSHPS), 2012 IEEE Conference on*. Demand side management to reduce peak-to-average ratio using game theory in smart grid (IEEE, 2012), pp. 91–96
19. L Song, Y Xiao, M van der Schaar, Demand side management in smart grids using a repeated game framework. IEEE J. Selected Areas Commun. **32**, 1412–1424 (2014)
20. R Radner, Repeated principal-agent games with discounting. Econometrica: J. Econ. Soc. **53**(5), 1173–1198 (1985)
21. JB Rosen, Existence and uniqueness of equilibrium points for concave n-person games. Econometrica: J. Econ. Soc. **33**(3), 520–534 (1965)
22. AJ Wood, BF Wollenberg, *Power Generation, Operation, and Control*. (Wiley, USA, 2012)
23. P McDaniel, S McLaughlin, Security and privacy challenges in the smart grid. IEEE Secur. Privacy. **7**(3), 75–77 (2009)
24. MK El Mahrsi, S Vignes, G Hébrail, M-L Picard, in *Research Challenges in Information Science, 2009. RCIS 2009. Third International Conference on*. A data stream model for home device description (IEEE Fez, Morocco, 2009), pp. 395–402
25. G Kalogridis, C Efthymiou, SZ Denic, TA Lewis, R Cepeda, in *Smart Grid Communications (SmartGridComm) 2010. First IEEE International Conference on*. Privacy for smart meters: towards undetectable appliance load signatures (IEEE Gaithersburg, MD, USA, 2010), pp. 232–237
26. C King, California's new landmark smart meter privacy law (2010). http://www.emeter.com/2010/californias-new-landmark-smart-meter-privacy-law/