Privacy-cost trade-offs in smart electricity metering systems

Giulio Giaconi[1], Deniz Gündüz[1] ✉, H. Vincent Poor[2]

[1]Department of Electrical and Electronic Engineering, Imperial College London, London, SW7 2AZ, UK
[2]Department of Electrical Engineering, Princeton University, Princeton, NJ 08544, USA
✉ E-mail: d.gunduz@imperial.ac.uk

Abstract: Trade-offs between privacy and cost are studied for a smart grid consumer, whose electricity consumption is monitored in almost real time by the utility provider (UP) through smart meter (SM) readings. It is assumed that an electrical battery is available to the consumer, which can be utilised both to achieve privacy and to reduce the energy cost by demand shaping. Privacy is measured via the mean squared distance between the SM readings and a target load profile, while time-of-use pricing is considered to compute the cost incurred. The consumer can also sell electricity back to the UP to further improve the privacy-cost trade-off. Two privacy-preserving energy management policies (EMPs) are proposed, which differ in the way the target load profile is characterised. A more practical EMP, which optimises the energy management less frequently, is also considered. Numerical results are presented to compare the privacy-cost trade-off of these EMPs, considering various privacy indicators.
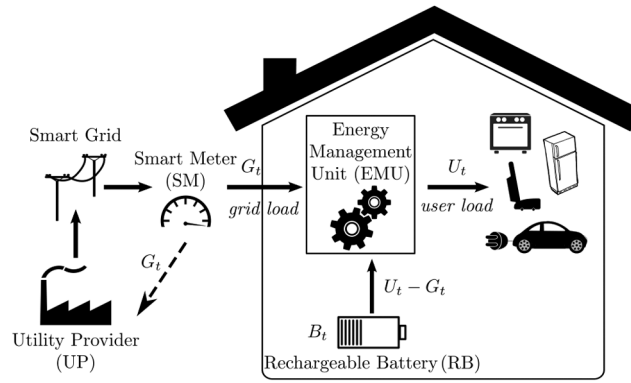
1 Introduction

Smart meters (SMs) are pivotal components of the smart grid, because they enable two-way communication between each household and the utility provider (UP), the entity that sells energy to consumers. Benefits include having more accurate electricity bills, detecting energy theft and outages faster, introducing time-of-use (ToU) tariffs to match demand with available resources, integrating microgeneration systems, e.g. photovoltaic panels and wind farms, and residential energy storage solutions, and the possibility for the consumers to sell energy to the grid. For these reasons, the SM roll-out is proceeding rapidly and is attracting considerable investments. However, an SM's ability to monitor a user's electricity consumption in almost real-time entails serious implications about consumer privacy. In fact, non-intrusive appliance load monitoring techniques can distinguish the power signatures of specific appliances from aggregated household SM measurements, revealing sensitive information about a consumer's life, such as her presence at home, religious beliefs and disabilities [1, 2]. SM privacy is also critical for businesses, e.g. factories and data centres, as power consumption data may reveal information about the state of their businesses.

1.1 Privacy-aware SM techniques

Privacy-preserving methods for SMs can be classified into two families. First, the *SM data manipulation* (SMDM) family [3], encompasses methods that modify SM measurements before reporting them to the UP, and includes *data obfuscation* [4], *aggregation* [5], *anonymisation* [6], *down-sampling* [7], and *data-sharing prevention* [8] approaches. However, these techniques suffer from several shortcomings [3]. First, obfuscation approaches add noise to the SM readings, causing a mismatch between the reported values and the real energy consumption, which prevents distribution systems operators (DSOs), i.e. the entities that operate and manage the grid, and UPs from accurately monitoring the grid state. Second, anonymisation and aggregation techniques that include the presence of a trusted third party (TTP) only shift the problem of trust from one entity (UP/DSO) to another (TTP). Third, DSOs, UPs, or more generally, any eavesdropper can embed additional sensors right outside a household or a business to monitor the energy consumption, without fully relying on SM

readings. The second family of privacy-preserving approaches, called the *user demand shaping* (UDS) family [3], overcomes these issues by modifying the consumer's actual electricity consumption, called the *user load*, rather than the data sent to the UP. This is achieved by exploiting physical resources, e.g. rechargeable batteries (RBs) or renewable energy sources (RESs), making the user load as different as possible from the SM measurements, called the *grid load* [9–11]. However, UDS techniques may require an initial investment by the user as physical resources need to be installed at the user's premises. Using RBs for privacy preservation may also lead to a quicker physical degradation of RBs [12]. Moreover, the cost of energy may increase when providing privacy via some UDS approaches. Hence, it is important to consider these aspects when considering UDS techniques for privacy preservation.

In this paper, we adopt UDS techniques because they report the energy taken from the grid accurately, and employ physical resources, e.g. RBs and RESs, which are becoming increasingly available to the consumers. Our aim is to jointly minimise the information leaked about a user and the cost of electricity. While a widely accepted definition of privacy is elusive, privacy is achieved when it is not possible to distinguish a specific appliance load from the aggregated household energy consumption [9]. Statistical techniques measure privacy loss by the mutual information between the user and the grid loads [10, 11, 13–17], or by computing approximations of it [18]; however, this requires the knowledge of the underlying statistics, and the results are typically valid under various simplifications, e.g. assuming independent and identically distributed user load, and over sufficiently long time horizons. An alternative approach is based on the idea that a high degree of privacy can be achieved by flattening the power consumption around a *target load profile*, e.g. minimising the distance from a completely private profile [10, 19–21]. The target load profile can be set to be a constant value over time, typically equal to the average consumption [10, 19, 20]. In this model, it is assumed that the energy management unit (EMU), i.e. the system that implements the privacy-preserving energy management policy (EMP) at the user's premises, knows, or, accurately predicts, the load profile for the time period of interest, and obtains the optimal EMP by solving an optimisation problem. On the other hand, a completely constant consumption may not be practically viable or desirable, since the energy cost may vary greatly during the system

Fig. 1 *System model. U_t, G_t and $U_t - G_t$ are the user load, the grid load, and the energy drawn from the RB at the time t, respectively. The dashed line represents the meter readings being accurately reported to the UP*

operation due to ToU tariffs. Hence, in [21], the EMU is allowed to target a different fixed power value for each price period. The flexibility of the latter approach leads to a better overall privacy-cost trade-off; however, such a piecewise constant target profile implies also an inherent information leakage compared to a constant target profile. We follow up on [10, 19–21], and measure the privacy leakage as the squared distance between the grid and target load profiles; however, differently from those works, we consider a more general target load profile, and assume that the consumer has only a partial knowledge of her future energy consumption and energy cost. We note that privacy leakage may be measured as the distance between the grid and user loads. However, if the aim is to increase such distance, this might lead to a potentially deterministic strategy for the EMU, e.g. produce a low grid load when the user load is high and vice-versa. Since we also assume that the UP knows the optimal strategy implemented by the EMU, such a strategy would result in a better estimate of the user load by the UP. On the contrary, trying to match the grid load to a specific target profile would make it harder to estimate the user load, e.g. flattening the grid load independent of the user load reveals only the average energy consumption.

The main contributions of this paper are as follows:

i.   While full information about the future electricity consumption is assumed to be available at the EMU in [10, 21], which we call the *long horizon* model (LHM), here we consider a more realistic scenario whereby the consumer's future consumption profile is only partially known to the EMU in a *receding horizon* manner, which we call the *short horizon* model (SHM). The optimal solution at any time is computed only based on the currently available information within the *prediction horizon* by adopting a model predictive controller, recently implemented in an SM setting in [14]. We present a detailed comparison of the results for SHM and LHM.
ii. We introduce a target load profile computed as a low-pass filtered version of the user load, since higher-frequency components of a user's consumption profile leak more information about her behaviour, compared to lower-frequency components. To the authors' knowledge, this is the first time that such a target profile has been studied in the SM privacy-preservation literature.
iii. We propose a more practical EMP that updates the optimal strategy less frequently. The optimal solution is computed in batch, reducing the computational load at the expense of the privacy-cost trade-off. Finally, we compare the privacy-cost trade-offs for all the schemes using real consumption and pricing data.

The remaining of this paper is organised as follows. In Section 2, we present the system model, while in Sections 3 and 4, we consider the SHM for a constant and a filtered target load profile, respectively. A more practical EMP with less regular policy updates is analysed in Section 5, while conclusions are drawn in Section 6.

1.2 Notation

For integers $0 < a < b$, U_a^b denotes the sequence $[U_a, U_{a+1}, \ldots, U_b]$, while $U^b \triangleq U_1^b$. The positive part $[x]^+$ is equal to x if $x > 0$, and 0 otherwise. When solving optimisation problems, we denote the optimal value of a variable with a star, e.g. G^* denotes the optimal value of the parameter G.

2 System model

We consider the discrete-time system depicted in Fig. 1, where t represents one-time slot (TS) of duration D seconds, for $1 \le t \le N$, where N is the time horizon of interest. For TS t, the user load, i.e. the total power requested by all the household appliances within TS t, is denoted by $U_t \in \mathcal{U}$, while the grid load is $G_t \in \mathcal{G}$. We remark that the TSs in our model correspond to time instants when the electricity is actually requested by the user and drawn from the grid, rather than the typically longer sampling interval used for sending SM measurements to the UP. We assume that the SM measures and records the output power values at each TS because our aim is to protect consumers' privacy not only from the UP, but also from the DSO or any other attacker that may deploy a sensor on the consumer's power line recording the electricity consumption in almost real time. An RB of capacity B_{max} is installed at the user's premises, which is used to both filter the user load to provide privacy, and to shift energy intake from the grid to minimise electricity costs. The EMU computes the amount of energy to draw from the grid, G_t, and to exchange with the RB, $U_t - G_t$. Let $B_t \in [0, B_{max}]$ denote the amount of energy in the RB at the end of TS t, and we set $B_0 = 0$. The RB is charging if $G_t - U_t > 0$, and discharging if $G_t - U_t < 0$. The user's electricity consumption and the electricity price are assumed to be known for H_F TSs beyond the current TS, naming H_F as the *prediction horizon*. Additionally, we assume that the EMU keeps memory about the past H_P TSs, which we call the *past horizon*. At each TS t, the EMU computes an EMP for the following H_F TSs, using its knowledge of U_t and the electricity cost within the prediction horizon, and its knowledge of the user load, grid load, and the RB level of energy within the past horizon.

Instead of the LHM scenario, where full information about a consumer's future energy consumption is assumed to be known over the whole time period of interest, in this paper, we consider a more realistic scenario and assume that only partial knowledge of the consumer's future energy consumption is available to the EMU. The SHM assumption is motivated by the difficulty in obtaining reliable longer-term predictions of a consumer's energy consumption.

2.1 System constraints

Let $\overline{t + H_F} \triangleq \min\{t + H_F, N\}$. We do not allow wasting grid energy; that is, there are no battery overflows, i.e. we impose $B_{t-1} + (G_t - U_t)D \le B_{max}, \forall t$. This means that, at any time t and

considering a prediction horizon of H_F TSs, the EMU has to satisfy the following constraint:

$$B_{t-1} + \sum_{s=t}^{\tau}(G_s - U_s)D \leq B_{\max}, \qquad (1)$$

where $t \leq \tau \leq \overline{t + H_F}$.

While additional energy can be stored in the RB for future use, we do not allow demand rescheduling, so that the user's energy demands are always satisfied at the time of request, i.e. we impose: $G_t D \geq U_t D - B_{t-1}$, $\forall t$. This leads to the following constraint for the EMU at the time t:

$$\sum_{s=t}^{\tau}(U_s - G_s)D \leq B_{t-1}, \qquad (2)$$

where $t \leq \tau \leq \overline{t + H_F}$. This constraint is also implicitly verified by the equation expressing the evolution of the energy level in the battery

$$0 \leq B_{t+1} = B_t + G_{t+1}D - U_{t+1}D. \qquad (3)$$

The power the RB can be charged or discharged at is constrained by $\hat{P}_c$ and $\hat{P}_d$, respectively. Thus, $\forall t$ we have

$$G_t - U_t \leq \hat{P}_c, \qquad (4)$$

$$U_t - G_t \leq \hat{P}_d. \qquad (5)$$

The model could be made more accurate by introducing further constraints, e.g. battery charging and discharging efficiency parameters, which we leave for future research focusing on the practical implications of the proposed UDS techniques.

We study and compare the two scenarios in which energy can or cannot be sold to the UP. The price of energy sold to the grid is set equal to the price of energy bought from it, i.e. we adopt the *net metering* approach, in which the SM can measure bi-directional energy flows [22]. If energy cannot be sold, then

$$G_t \geq 0, \quad \forall t, \qquad (6)$$

whereas if energy can be sold, we have

$$G_t : \begin{cases} \geq 0, & \text{if energy is purchased from the UP,} \\ < 0, & \text{if energy is sold to the UP.} \end{cases} \qquad (7)$$

Given $(U_t, B_t) = (u_t, b_t)$, $B_{\max}$ and the constraints (4)–(6), the set of feasible energy requests at a time t is given by

$$\bar{\mathcal{G}}_t(u_t, b_t) \triangleq \left\{ g_t \in \mathcal{G} : \left[u_t - \min\left\{ \frac{b_t}{D}, \hat{P}_d \right\} \right]^+ \leq g_t \right.$$
$$\left. \leq u_t + \min\left\{ \hat{P}_c, \frac{B_{\max} - b_t}{D} \right\} \right\}. \qquad (8)$$

If selling energy to the UP is allowed, then the feasible set is as in (8), but without the $[\cdot]^+$ operator.

The EMP computes the grid load at each TS while satisfying the above constraints. We consider a model predictive control approach, whereby the user load and the cost of energy are known beforehand within the prediction horizon $[t+1, \ldots, t+H_F]$, and the goal is to jointly minimise the information leaked about a user's energy consumption as well as the cost the user incurs to purchase energy from the UP. While non-causal knowledge of the electricity price for the typical range of interest is a realistic assumption in today's energy networks, non-causal knowledge of power consumption is appropriate for appliances whose activity can be accurately predicted, e.g. refrigerators, boilers, heaters and electric vehicles. We note that the setting studied in [21], which assumes all

future energy consumption and cost information to be known beforehand, is a lower bound on the setting studied in this paper, as more information leads to a better privacy-cost trade-off.

Let the target load at time t be denoted by W_t. We measure the privacy leakage as the average variance of the grid load G^N from the target load profile W^N:

$$\mathcal{P} \triangleq \frac{1}{N}\sum_{t=1}^{N}(G_t - W_t)^2, \qquad (9)$$

according to which, perfect privacy is achieved when $G_t = W_t$, $\forall t$. We adopt squared distance in (9) not to discriminate between negative and positive deviations of G_t from W_t. The average cost incurred by the user is given by

$$\mathcal{C} \triangleq \frac{1}{N}\sum_{t=1}^{N}G_t C_t, \qquad (10)$$

where C_t is the cost of energy at the time t, which is determined by the specific ToU tariff employed by the UP.

2.2 Simulation settings

We use real SM consumption traces from the UK Dale dataset [23]. We convert the original resolution of 6 seconds to 10 minutes to reduce the computational complexity. We consider a Tesla Powerwall 2 [24] as RB, for which $B_{\max} = 13.5$ kWh, and $\hat{P}_c = \hat{P}_d = 5$ kW. We consider a ToU tariff that was offered in the UK [25], in which the off-peak price is 4.99 p/kWh during 23:00 to 6:00, the medium price is 11.99 p/kWh during 6:00 to 16:00 and during 19:00 to 23:00, and the peak price is 24.99 p/kWh during 16:00 to 19:00. All the simulation results are obtained for a time interval spanning 14 consecutive days, to average over a considerably large amount of data. For simplicity, we will mostly present numerical results when selling energy to the grid is not allowed, unless energy selling leads to significantly different results.
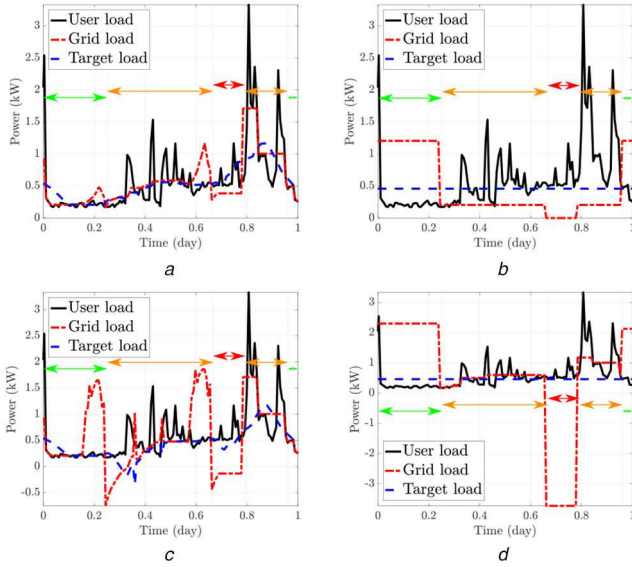
3 Target load as a constant value

In this section, following up on [9, 10, 20], we assume that the goal of the EMU is to keep the grid load as constant as possible. In [11], it is assumed that all the future user load and energy cost values are known, and the EMU can fix a target value for the whole duration, e.g. one whole day. In our model the information available to the EMU on U_t and C_t is limited to the prediction horizon, and changes over time; thus, the target load cannot be constant, and its variability depends on the length of the past and prediction horizons. In this section, given the knowledge of the cost of energy and the user's power consumption, the aim is to characterise both the optimal target load W^* and the optimal grid load G^* so as to optimise the overall privacy-cost trade-off.

Given the nature of the objective functions and the constraints, pairs of $(\mathcal{P}, \mathcal{C})$ form a convex region and the optimal points can be characterised by the Pareto boundary of this region. Hence, the objective can be cast as a weighted sum of privacy leakage (9) and cost (10):

$$\min_{G_t^{\overline{t+H_F}}, W_t} \alpha \sum_{\tau=\overline{t-H_P}}^{\overline{t+H_F}}(G_\tau - W_t)^2 + (1-\alpha)\sum_{\tau=t}^{\overline{t+H_F}}G_\tau C_\tau, \qquad (11)$$

where $0 \leq \alpha \leq 1$ is the weighting parameter, i.e. if $\alpha = 0$ only cost of energy is minimised, whereas if $\alpha = 1$ only information leakage is minimised; and $\overline{t - H_P} \triangleq \max\{t - H_P, 0\}$. We remark that setting the value of α is up to the consumer, who is in charge of deciding whether to focus more on protecting her privacy or on saving costs. The result of the minimisation in (11) is the grid load for the current TS and the entire duration of the prediction horizon $G_t^{\overline{t+H_F}}$, and the target load W_t. Equation (11) characterises the target

Fig. 2 *Power profiles for $\alpha = 0.5$ and $H_F = H_P = 2$ h. In the figures, the arrows of green, orange, and red colours denote time intervals characterised by off-peak, medium, and peak price for the electricity cost, respectively*
(a) SHM, no energy selling, *(b)* LHM, no energy selling, *(c)* SHM, energy selling, *(d)* LHM, energy selling

load value W_t for the finite prediction horizon, which leads to the optimal privacy-cost trade-off over this horizon, based on the available information. At TS $t + 1$, the minimisation (11) is carried out again based on the additional information that becomes available, i.e. G_{t+H_F+1} and C_{t+H_F+1}, and G_{t+1} and W_{t+1} are determined. The past horizon $\sum_{\tau = t - H_P}^{t-1} (G_\tau - W_\tau)^2$ is considered when optimising for the privacy objective, since it ensures smoother variations of the overall target load profile. We note that, since privacy and cost in (11) may have significantly different magnitudes, they need to be further normalised to get the Pareto optimal solution consistent with α [26].

Remark 1: Differently from [21], here we do not impose the RB to be emptied at the end of each time window $[t - H_P, t + H_F]$, since here the end of the prediction horizon does not typically coincide with the end of the time horizon of interest, and the energy remaining in the RB can be utilised in the following TSs. Since the algorithm jointly minimises privacy leakage and cost, the RB is normally emptied at the end of the time horizon of interest N, unless α is high. If $\alpha \to 1$ and the RB is large, a sustained demand of energy may take place in the short term, which is ultimately constrained by $B_{\max}$ and $\hat{P}_c$.

Fig. 2 shows the load profiles of SHM and LHM over one day. As expected, the LHM provides better performance as the resultant profile is much flatter and hides most of the consumption spikes, which typically reveal more information about the user's behaviour. However, also the SHM leads to a reasonable suppression of consumption peaks, despite relying on much less data. Also, the SHM reveals more information about the low-frequency variation of user's energy consumption, which, however, is common across households, and thus provides only a limited amount of personal information. Moreover, Figs. 2*a* and 2*c* show that the peaks of the grid load generated by SHM are not necessarily aligned with those of the user load.

3.1 Solution to the optimisation problem (11)

In the following, we consider the optimisation problem (11) for $\alpha \neq \{0, 1\}$. We analyze first the optimal solution when selling of energy is allowed, i.e. (6) does not hold. Based on (11) and the constraints (1), (2) and (4), (5), we define the Lagrangian function in (12) at the bottom of this page, where $\lambda_\tau^{(j)} \geq 0$, for $1 \leq j \leq 4$, are the Lagrange multipliers, and $t \leq \tau \leq \overline{t + H_F}$. Denoting the vectors in bold, we have $\boldsymbol{\lambda} = [\boldsymbol{\lambda}^{(1)}, \boldsymbol{\lambda}^{(2)}, \boldsymbol{\lambda}^{(3)}, \boldsymbol{\lambda}^{(4)}]$. The slackness conditions are imposed on the inequality constraints, for $\tau = t, t + 1, \ldots, \overline{t + H_F}$:
(see (12))

$$\lambda_\tau^{(1)} \left[D \sum_{s=t}^{\tau} (G_s - u_s) - B_{\max} + B_{t-1} \right] = 0, \qquad (13)$$

$$\lambda_\tau^{(2)} \left[D \sum_{s=t}^{\tau} (U_s - G_s) - B_{t-1} \right] = 0, \qquad (14)$$

$$\lambda_\tau^{(3)} (G_\tau - U_\tau - \hat{P}_c) = 0, \qquad (15)$$

$$\lambda_\tau^{(4)} (U_\tau - G_\tau - \hat{P}_d) = 0. \qquad (16)$$

Let $a_\tau \triangleq D \sum_{s=\tau}^{\overline{t+H_F}} (\lambda_s^{(2)} - \lambda_s^{(1)}) - \lambda_\tau^{(3)} + \lambda_\tau^{(4)}$, and $\tilde{C}_\tau \triangleq ((1 - \alpha)/2\alpha) C_\tau$. Applying the Karush–Kuhn–Tucker (KKT) conditions and setting the gradient of the Lagrangian to zero, we obtain the following expressions:

$$G_\tau^* = \frac{a_\tau}{2\alpha} - \tilde{C}_\tau + W_t^*, \quad \text{for } \tau = t, \ldots, \overline{t + H_F}, \qquad (17)$$

$$W_t^* = \frac{\sum_{\tau = t - H_P}^{\overline{t + H_F}} G_\tau^*}{1 + \min\{H_P, t\} + \min\{H_F, N - t\}}. \qquad (18)$$

The optimal solution for the grid load given in (17) resembles the classical water-filling algorithm [27]. However, differently from the classical water-filling formulation, here the water level, $G_\tau^* + \tilde{C}_\tau = (a_\tau/2\alpha) + W_t^*$, is not constant, but varies over time due to the instantaneous power constraints. The optimal solutions given by (17) and (18) depend on the values of the Lagrangian multipliers and can be determined numerically.

When $\alpha = 0$, the only objective is to minimise the cost, and (11) reduces to a linear program, which can be solved using standard linear programming solvers. On the other hand, when the user is not concerned about the cost, i.e. $\alpha = 1$, (11) leads to a quadratic program analogous to the general case.

When energy selling is not allowed, the constraint (6) holds, and the Lagrangian in (12) is modified accordingly. The slackness conditions are given in (13)–(16), as well as, for $\tau = t, \ldots, \overline{t + H_F}$: $\lambda_\tau^{(5)} G_\tau = 0$, and $\lambda^{(6)} W_t = 0$. Let $\tilde{a}_\tau \triangleq D \sum_{s=\tau}^{\overline{t+H_F}} (\lambda_s^{(2)} - \lambda_s^{(1)}) - \lambda_\tau^{(3)} + \lambda_\tau^{(4)} + \lambda_\tau^{(5)}$. Then, we obtain the following expressions, counterparts of (17) and (18):

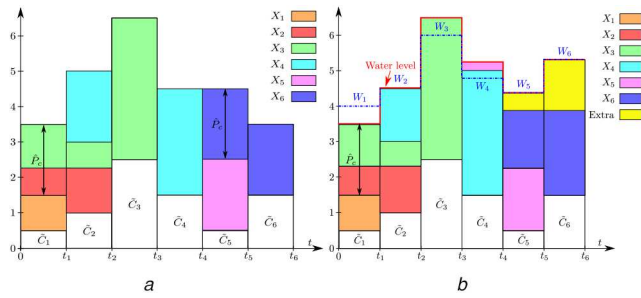$$G_\tau^* = \left[\frac{\tilde{a}_\tau}{2\alpha} - \tilde{C}_\tau + W_t^* \right]^+, \quad \text{for } \tau = t, \ldots, \overline{t + H_F}, \qquad (19)$$

$$W_t^* = \left[\frac{\sum_{\tau = t - H_P}^{\overline{t + H_F}} G_\tau^* + \lambda^{(6)}}{1 + \min\{H_P, t\} + \min\{H_F, N - t\}} \right]^+. \qquad (20)$$

$$\mathscr{L}(G_t^{\overline{t + H_F}}, W_t, \boldsymbol{\lambda}) = \alpha \sum_{\tau = t - H_P}^{\overline{t + H_F}} (G_\tau - W_t)^2 + (1 - \alpha) \sum_{\tau = t}^{\overline{t + H_F}} G_\tau C_\tau + \sum_{\tau = t}^{\overline{t + H_F}} \lambda_\tau^{(1)} \left[D \sum_{s=t}^{\tau} (G_s - U_s) - B_{\max} + B_{t-1} \right]$$
$$+ \sum_{\tau = t}^{\overline{t + H_F}} \lambda_\tau^{(2)} \left[D \sum_{s=t}^{\tau} (U_s - G_s) - B_{t-1} \right] + \sum_{\tau = t}^{\overline{t + H_F}} \lambda_\tau^{(3)} (G_\tau - U_\tau - \hat{P}_c) + \sum_{\tau = t}^{\overline{t + H_F}} \lambda_\tau^{(4)} (U_\tau - G_\tau - \hat{P}_d).$$
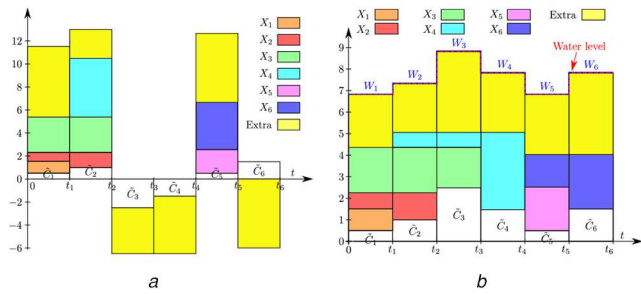$$(12)$$

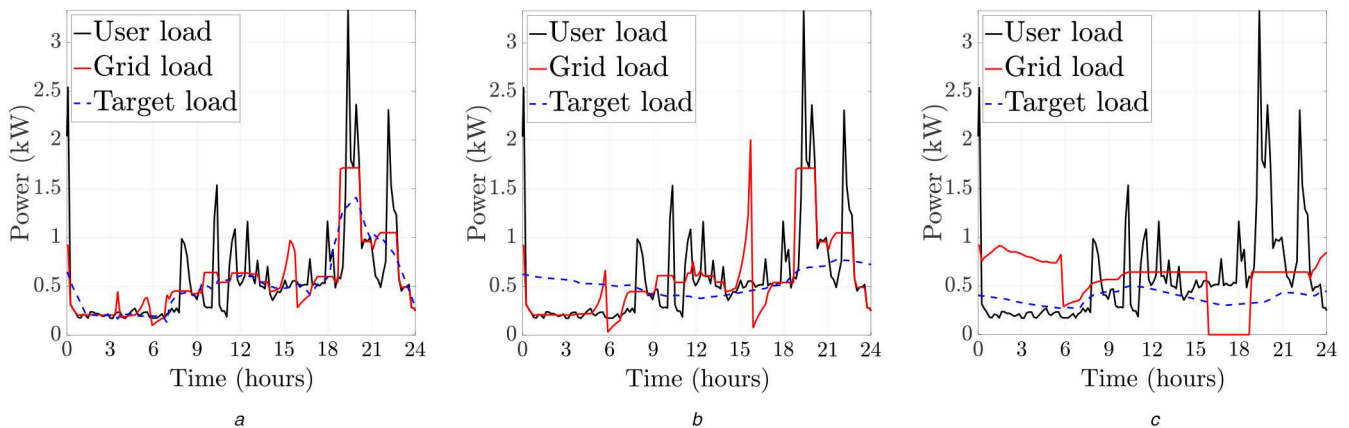## 3.2 Illustration of the water-filling solution

Here we present the solution for some simple scenarios to acquire an intuition on the solution of the optimisation problem (11), and on its water-filling interpretation. Assume energy selling is not allowed, $N = 6$, and $D = 1$, i.e. power and energy can be used interchangeably. Consider $u^6 = [1, 2, 6, 5, 2, 4]$ and $c^6 = [1, 2, 5, 3, 1, 3]$, $H_P = H_F = 2$ TSs, and an RB with $B_{max} = 4$ and $\hat{P}_c = \hat{P}_d = 2$. Fig. 3a shows the optimal solution for $\alpha = 0$, $G^{*,6} = [3, 4, 4, 3, 4, 2]$. Since the electricity is cheaper for $t = \{1, 2\}$, more energy is requested from the grid at these TSs and stored in the RB to satisfy the demand at later TSs. However, such energy is limited by B_{max} and $\hat{P}_c$. At $t = 1$, $\hat{P}_c$ limits the grid load, since $G_1^* = u_1 + \hat{P}_c = 3$, and the level of energy in the RB at the end of the first TS is $B_1 = \hat{P}_c = 2$. At $t = 2$, the grid load is limited by $\hat{P}_c$ and B_{max} simultaneously, and $G_2^* = u_2 + \hat{P}_c = 4$, and $B_2 = B_{max} = 4$. Note that, although the third TS is the most expensive, $G_3^* = 4$ because the RB cannot be discharged by more than 2 units of energy ($\hat{P}_d = 2$). For the same reason, $G_4^* = 3$, while the remaining energy was stored in the second TS. Similar

considerations hold for the last two TSs. Fig. 3b illustrates the optimal solution for $\alpha = 1$, $G^{*,6} = [3, 3.5, 4, 3.75, 3.88, 3.81]$. In this scenario, the EMP tries to match G_t to the target load W_t, even at the cost of asking more energy than needed. The energy demand at $t = 1$ is the same as the case $\alpha = 0$, whereas at $t = 2$ less energy is stored in the battery to be used at $t = 4$, so that the water level matches the target load in this TS (at the expense of a higher cost at $t = 4$). It is noteworthy that more energy than needed is requested at $t = 4$, and used to satisfy the demand at $t = 5$, which is cheaper. Finally, more energy than needed is requested at $t = \{5, 6\}$ (depicted in yellow), allowing the EMP to match the target and the grid load. Such energy is stored in the RB for future use.

Fig. 4 shows the same scenario with $B_{max} = 20$, $\hat{P}_c = \hat{P}_d = 10$, and when energy can be sold. When $\alpha = 0$, Fig. 4a shows that energy is bought when it is cheaper and sold back to the grid when it is more expensive, maximising the user's profit. $\tilde{C}_3$ and $\tilde{C}_4$ are plotted as negative since energy is sold in these TSs. In fact, the RB is emptied of the energy stored during the first two TSs at the end of the fourth TS, and it is emptied again of the energy stored during the fifth TS at the end of the sixth TS. When $\alpha = 1$, Fig. 4b shows that a larger RB permits greater flexibility but also boosts the amount of energy requested. This is not necessarily a disadvantage, as such energy can be used at a later TS.
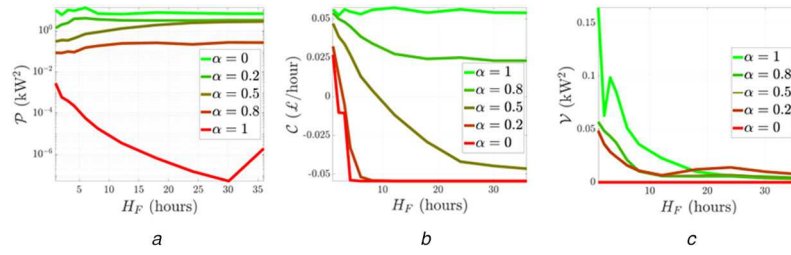
3.3 Impact of the duration of prediction and past horizons

Fig. 5 shows the load profiles for various combinations of H_F and H_P. As expected, a larger H_F produces flatter target and grid loads (see Fig. 5c), compared to a smaller H_F (see Fig. 5a). However, when H_F is larger the resulting grid load is more distant from the target load, thus resulting in a higher information leakage according to our definition of privacy leakage in (9). In fact, when H_F is small, the grid load values that are compared to the target load are few, and the EMP can determine a target load that is close to the grid load within the analysed time window. On the contrary, when H_F is large, the EMP needs to find a single target load that matches a longer interval of grid load values; as a result, the target load may be less representative for some periods. A larger H_P leads to a flatter target load (see Fig. 5b), however, the grid load is spikier compared to considering a larger H_F.

Fig. 6 shows the average information leakage $\mathcal{P}$, average cost $\mathcal{C}$, and average target load variance with respect to H_F. The average target load variance, which can be considered as another privacy indicator, is defined as
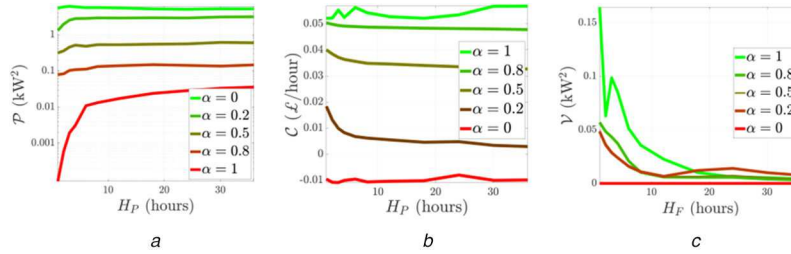
$$\mathcal{V} \triangleq \frac{1}{N} \sum_{t=1}^{N} (W_t - \mu_W)^2, \qquad (21)$$

where μ_W is the mean of W overtime. When α is small, i.e. the main focus is to minimise the cost, a larger H_F reduces the average cost up to a certain extent, beyond which it cannot be further
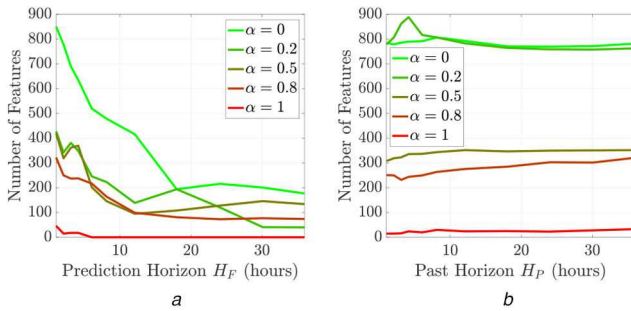

Fig. 3 *Optimal grid and target load profiles, for $B_{max} = 4$, $\hat{P}_c = \hat{P}_d = 2$, when energy selling is not allowed*
(a) $\alpha = 0$, (b) $\alpha = 1$

Fig. 4 *Optimal grid and target load profiles, for $B_{max} = 20$, $\hat{P}_c = \hat{P}_d = 10$, when energy selling is allowed*
(a) $\alpha = 0$, (b) $\alpha = 1$

Fig. 5 *Comparison between various past and prediction horizons for $\alpha = 0.5$, when energy selling is not allowed*
(a) $H_P = H_F = 1$ h, (b) $H_P = 12$ h, $H_F = 1$ h, (c) $H_P = 1$ h, $H_F = 12$ h

Fig. 6 *Impact of the prediction horizon H_F on leakage, cost, and target load variance ($H_P = 2$ h)*
(a) Energy selling, *(b)* Energy selling, *(c)* No energy selling

Fig. 7 *Impact of the past horizon H_P on leakage, cost and target load variance ($H_F = 2$ h)*
(a) Energy selling, *(b)* Energy selling, *(c)* No energy selling

Fig. 8 *Number of features versus H_F and H_P, no energy selling*
(a) $H_P = 2$ h, *(b)* $H_F = 2$ h

reduced (see Fig. 6*b*); whereas for $\alpha = 1$ the cost does not change considerably with H_F. Opposite considerations hold for the information leakage, which even slightly increases when $\alpha \neq 1$ (see Fig. 6*a*). This is due to the fact that a longer prediction horizon generates a grid load that is more distant from the target load, except for $\alpha = 1$, when the focus is on privacy only. Fig. 6*c* shows that increasing H_F induces a smaller variance on the target load.

Fig. 7, which shows $\mathscr{P}$, $\mathscr{C}$, and $\mathscr{V}$ with respect to H_P, exhibits similar behaviours to those in Fig. 6, with some notable differences. The *y*-axis ranges are more limited here, confirming that the knowledge of past consumption is less critical for the EMP compared to the knowledge of future consumption and costs. This explains the far smaller reduction in cost in Fig. 7*b*, as compared to Fig. 6*b*, and the increase in the information leakage when $\alpha = 1$ in Fig. 7*a*. Fig. 7*c* shows that the target variance is higher and more variable when $\alpha = 1$; however, this corresponds to the case in which the grid load is closer to the target load, i.e. the most private scenario according to our original privacy measure. This contradiction of the two privacy indicators shows that evaluating the variance of the target load does not fully reflect the level of privacy achieved, as defined in (9).

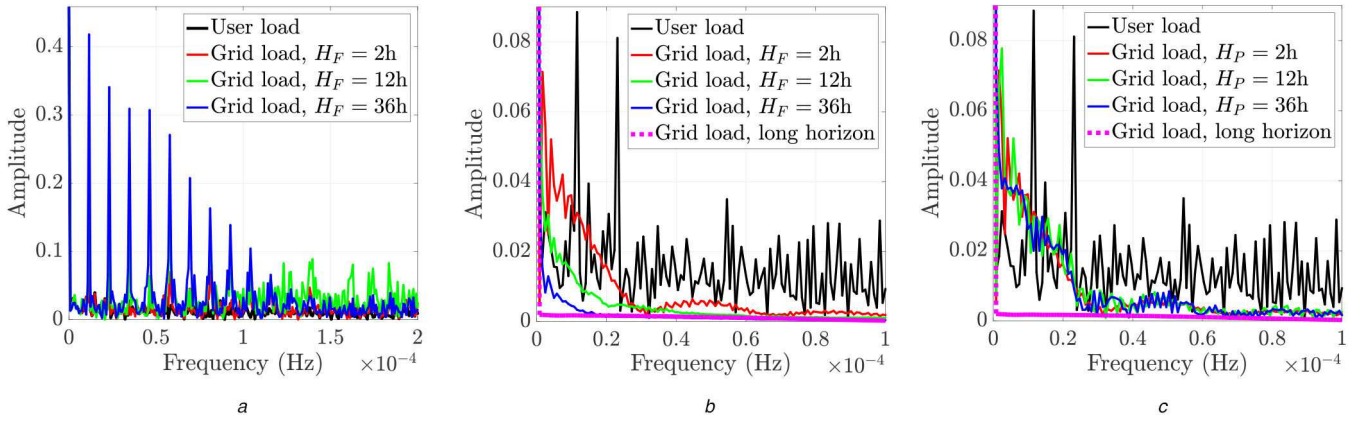### 3.4 Alternative privacy measures

As opposed to the LHM studied in [10], where W_t is fixed throughout the operation time, here W_t is allowed to vary over time. Therefore, the squared distance between G_t and W_t may not be sufficient as a privacy measure on its own. Accordingly, we consider alternative measures of privacy to see the impact of the proposed model predictive control framework on those measures. One of the objectives of privacy-preserving algorithms for SM data

is to mask the difference between successive power measurements, called *features*, which non-intrusive appliance load monitoring algorithms exploit to identify appliances' switch-on/off events [28]. Thus, it is possible to evaluate an EMP's performance against such algorithms by computing the number of features present in the grid load [20]. We classify as features those differences that are larger or equal to 50 W, which represent a typical household electricity consumption of lights. Fig. 8 shows the number of features with respect to H_F and H_P. A larger H_F leads to a reduction in the number of features in the grid load (Fig. 8*a*); however, H_P does not seem to have any influence on this (Fig. 8*b*).

Another way of assessing the performance of privacy-preserving algorithms is by analysing the power spectrum of the resulting grid load. In fact, the higher-frequency components of the grid load spectrum correspond typically to more sensitive information about a user's energy consumption [29]. Fig. 9 shows the grid load spectra corresponding to using different values of H_F and H_P. Larger values of H_F lead to better suppression of higher-frequency components when $\alpha = 1$ (Fig. 9*b*), whereas for $\alpha = 0$ even additional high-frequency components are introduced (Fig. 9*a*). When $\alpha = 1$, increasing H_P also attenuates the higher-frequency components (Fig. 9*c*), but less markedly compared to increasing H_F. As the spectral analysis of the grid load better captures the information leaked, in the following section, we consider a privacy-preserving approach whose aim is to filter out directly the higher-frequency components of the user load.
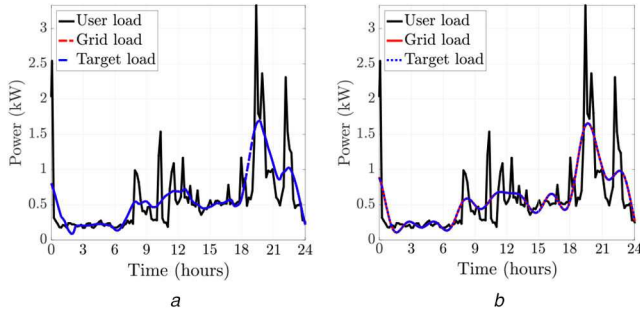
4 Target load as filtered user load

When the target load is set to a constant value, one can consider this as the DC component of the Fourier transform of the user load profile. If the grid load can be maintained at the average value of the user load at all times, this is equivalent to filtering out all the positive frequency components of the user load profile. However, as shown in Section 3, this is not always possible due to the RB capacity and power constraints, and the information leakage is measured as average squared error distance from this constant DC component. In this section, we consider a more general target load profile, obtained by low-pass filtering the user load, which is equivalent to removing only the high-frequency variations. The motivation for this is two-fold: First, the EMU is able to better approximate the target load profile by keeping the low-frequency components; and secondly, the high-frequency components are the ones that leak more information about user behaviour. Low-frequency devices are those that typically have a continuous periodic operation, e.g. the fridge, and are not particularly privacy sensitive. We would like to remark that, differently from the

Fig. 9 *Power spectra versus H_F and H_P, and no energy selling*
(a) $\alpha = 0$, $H_P = 2\,h$, *(b)* $\alpha = 1$, $H_P = 2\,h$, *(c)* $\alpha = 1$, $H_F = 2\,h$

Fig. 10 *Filtered target load scenario with a cut-off frequency of 0.1 mHz,*
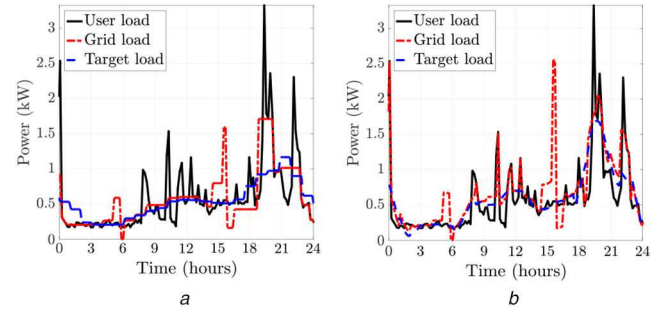$\alpha = 1$, *and* $H_F = H_P = 2\,h$
(a) SHM, *(b)* LHM

Fig. 11 *Practical EMP for* $\alpha = 0.5$, $H_F = H_P = 2\,h$, $T_S = 1\,h$, *and cut-off frequency set to 0.1 mHz*
(a) Constant target, *(b)* Filtered target

previous section, here W is not an optimisation variable but it is determined based only on the user load. The optimisation problem is expressed as

$$\min_{G_t^{\overline{t+H_F}}} \alpha \sum_{\tau = t}^{\overline{t+H_F}} (G_\tau - W_\tau)^2 + (1-\alpha) \sum_{\tau = t}^{\overline{t+H_F}} G_\tau C_\tau, \qquad (22)$$

where $W_t, W_{t+1}, \ldots, W_{\overline{t+H_F}}$ are obtained as low-pass filtered versions of the user load, subject to the same constraints of the constant target scenario, i.e. (1), (2) and (4), (5). The target load at time t, W_t, is computed as follows. The EMU selects the only available user load, i.e. that one within $[\overline{t-H_P}, \overline{t+H_F}]$, and computes its spectral representation by means of the discrete Fourier transform. Then, a low-pass filter with a predefined cut-off frequency is applied. Finally, the inverse transform provides the target load profile W_t. We note that, although H_P does not appear explicitly in (22), the target load computed at time t is determined by low-pass filtering the user load within the time window $[\overline{t-H_P}, \overline{t+H_F}]$ to prevent the target load from varying dramatically over different TSs. When $\alpha = 0$, (22) reduces to the linear program of the previous section. The optimal solutions to (22) can be characterised by following the same steps of Section 3, apart from W_t, which here is not an optimisation variable. The optimal solutions are given in (17) and (19) for the scenarios where selling energy is allowed and not allowed, respectively. Fig. 10 compares the SHM and the LHM, showing that SHM generates profiles that are smooth and similar to that of the LHM, despite relying only on the knowledge of 2 h of future electricity consumption.

5 More practical EMP

In the previous sections, it is assumed that the EMU solves the optimisation problem at each TS. However, in practice it may not be feasible to obtain the future predictions at each TS, and may be impractical to compute the target profile so often. Thus, in this section, we consider a more practical EMP where the optimisation problem is solved once every T_S TSs. The optimisation problems at time t are still given by (11) and (22) for the constant and filtered target load scenarios, respectively, such that the sequences $G_t^{t+T_S}$ and $W_t^{t+T_S}$ are obtained at time t on the basis of the available information for TSs $[t - H_P, t + H_F]$, where $H_F \geq T_S$.
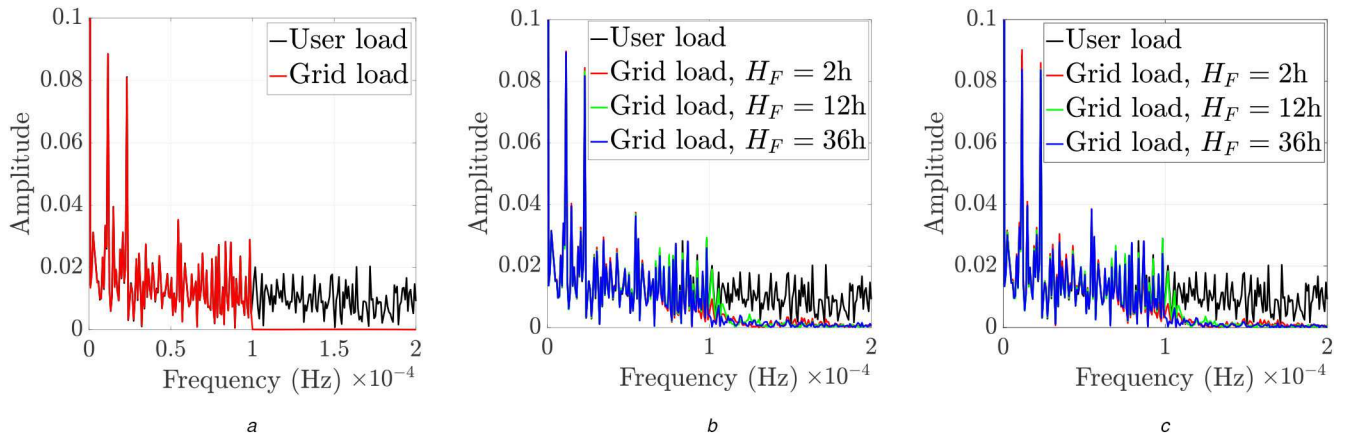
Fig. 11 shows that the practical EMP for a constant and a filtered target load creates a piecewise target and grid loads, similar to the piecewise target load profile approach [21]. Due to the discontinuities introduced in the grid load profile, spikes at high frequencies may appear in the spectrum of the grid load produced by this strategy, leading to a higher privacy loss. When $\alpha = 1$, Fig. 12 shows that the practical EMP reaches virtually the same performance of the SHM, despite computing the grid load six times less often than the LHM.

Finally, in Fig. 13, we present the privacy-cost trade-offs for the various scenarios we have discussed when energy selling is allowed. This figure clearly highlights the increasing loss in performance due to the decreasing amount of information available to the EMU when moving from the LHM to the SHM and to the more practical EMP.
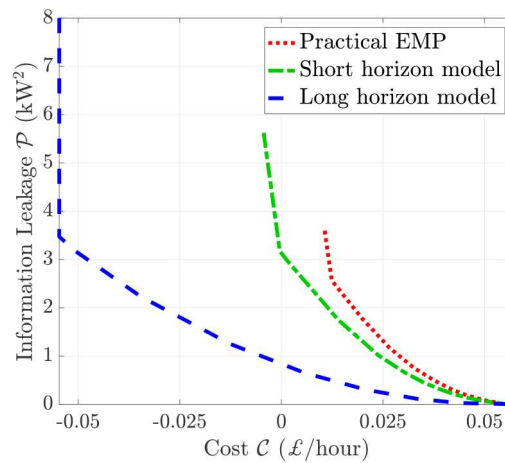
6 Conclusions

We have studied the joint optimisation of privacy and cost for an SM system equipped with an RB. Privacy is measured via the mean squared-error between the SM measurements and a target load profile, which is set to be either a constant function or a low-pass filtered version of the user load. We assume that only partial information about the user's future electricity consumption and electricity cost is known to the EMU, and we cast the joint privacy and cost optimisation as a model predictive control problem. The scenario in which the user is allowed to sell excess energy to the UP is studied, which is shown to achieve a better privacy-cost trade-off. The optimal solutions for the constant and filtered target load profiles have been characterised, highlighting their water-filling interpretation. The privacy-cost trade-off has been

Fig. 12 *Filtered target load scenario with a cut-off frequency of 0.1 mHz, $\alpha = 1$, and no energy selling. For the SHM and the practical EMP, we set $H_P = 2$ h and $T_S = 1$ h*
(a) LHM, *(b)* SHM, *(c)* Practical EMP

Fig. 13 *Privacy-cost trade-off, for $H_P = H_F = 2$ h, $T_S = 1$ h, energy selling allowed, and constant target load for the SHM, LHM, and the practical EMP*

characterised for the various scenarios, and detailed numerical simulations and alternative privacy measures have been presented. As a future extension of this study, one can consider a generalisation of the SHM and LHM models by introducing errors in the predictions of the future energy consumption profile at the EMU. It is reasonable to assume that the prediction error will increase gradually for more distant time instants in the future.

7 Acknowledgments

8 References

[1] Quinn, E.: 'Privacy and the new energy infrastructure', *Soc. Sci. Res. Netw.*, 2009, http://dx.doi.org/10.2139/ssrn.1370731
[2] Rouf, I., Mustafa, H., Xu, M., *et al.*: 'Neighborhood watch: security and privacy analysis of automatic meter reading systems'. Proc. ACM Conf. on Computer and Communications Security, Raleigh, NC, USA, 2012, pp. 462–473
[3] Giaconi, G., Gündüz, D., Poor, H.V.: 'Privacy-aware smart metering: progress and challenges', *IEEE Signal Process. Mag.*, 2018, **35**, (6), pp. 59–78
[4] Kim, Y., Ngai, E.C., Srivastava, M.B.: 'Cooperative state estimation for preserving privacy of user behaviors in smart grid'. Proc. IEEE Int. Conf. Smart Grid Communications, Brussels, Belgium, 2011, pp. 178–183
[5] Bohli, J.M., Sorge, C., Ugus, O.: 'A privacy model for smart metering'. Proc. IEEE Int. Conf. Communications, Cape Town, South Africa, 2010, pp. 1–5
[6] Petrlic, R.: 'A privacy-preserving concept for smart grids'. Sicherheit in vernetzten Systemen: 18. DFN Workshop, Hamburg, Germany, 2010, pp. B1–B14
[7] Cárdenas, A., Amin, S., Schwartz, G.A.: 'Privacy-aware sampling for residential demand response programs'. Proc. ACM Int. Conf. High Confidence Networked Syst., Beijing, China, 2012
[8] Molina-Markham, A., Shenoy, P., Fu, K., *et al.*: 'Private memoirs of a smart meter'. Proc. ACM Workshop Embedded Sensing Syst. for Energy-Efficiency in Building, Zurich, Switzerland, 2010, pp. 61–66
[9] Kalogridis, G., Efthymiou, C., Denic, S.Z., *et al.*: 'Privacy for smart meters: towards undetectable appliance load signatures'. Proc. IEEE Int. Conf. Smart Grid Communications, Gaithersburg, MD, USA, 2010, pp. 232–237
[10] Tan, O., Gómez-Vilardebó, J., Gündüz, D.: 'Privacy-cost trade-offs in demand-side management with storage', *IEEE Trans. Inf. Foren. Sec.*, 2017, **12**, (6), pp. 1458–1469
[11] Giaconi, G., Gündüz, D., Poor, H.V.: 'Smart meter privacy with renewable energy and an energy storage device', *IEEE Trans. Inf. Foren. Sec.*, 2018, **13**, (1), pp. 129–142
[12] Avula, R.R., Oechtering, T.J., Månsson, D.: 'Privacy-preserving smart meter control strategy including energy storage losses'. Proc. IEEE PES Innovative Smart Grid Technologies Conf. Europe., Sarajevo, Bosnia-Herzegovina, 2018, pp. 1–6
[13] Giaconi, G., Gündüz, D.: 'Smart meter privacy with renewable energy and a finite capacity battery'. Proc. IEEE Int. Workshop Signal Processing Advances in Wireless Communications, Edinburgh, UK, 2016, pp. 1–5
[14] Chin, J.X., Rubira, T.T.D., Hug, G.: 'Privacy-protecting energy management unit through model-distribution predictive control', *IEEE Trans. Smart Grid*, 2017, **8**, (6), pp. 3084–3093
[15] Yao, J., Venkitasubramaniam, P.: 'The privacy analysis of battery control mechanisms in demand response: revealing state approach and rate distortion bounds', *IEEE Trans. Smart Grid*, 2015, **6**, (5), pp. 2417–2425
[16] Gómez-Vilardebó, J., Gündüz, D.: 'Smart meter privacy for multiple users in the presence of an alternative energy source', *IEEE Trans. Inf. Foren. Sec.*, 2015, **10**, (1), pp. 132–141
[17] Li, S., Khisti, A., Mahajan, A.: 'Information-theoretic privacy for smart metering systems with a rechargeable battery', *IEEE Trans. Inf. Theory*, 2018, **64**, (5), pp. 3679–3695

[18]   Chin, J., Giaconi, G., Tinoco De Rubira, T., *et al.*: 'Considering time correlation in the estimation of privacy loss for consumers with smart meters'. Proc. Power Systems Computation Conf., Dublin, Ireland, 2018, pp. 1–7

[19]   Yang, L., Chen, X., Zhang, J., *et al.*: 'Optimal privacy-preserving energy management for smart meters'. Proc. IEEE Int. Conf. on Computer Communications, Toronto, ON, Canada, 2014, pp. 513–521

[20]   Yang, L., Chen, X., Zhang, J., *et al.*: 'Cost-effective and privacy-preserving energy management for smart meters', *IEEE Trans. Smart Grid*, 2015, **6**, (1), pp. 486–495

[21]   Giaconi, G., Gündüz, D., Poor, H.V.: 'Optimal demand-side management for joint privacy-cost optimization with energy storage'. Proc. IEEE Int. Conf. Smart Grid Communications, Dresden, Germany, 2017, pp. 265–270

[22]   Payne, A.M., Duke, R.D., Williams, R.H.: 'The impact of net metering on the residential rooftop PV market'. Proc. IEEE Photovoltaic Specialists Conf., Anchorage, AK, USA, 2000, pp. 1391–1394

[23]   Kelly, J., Knottenbelt, W.: 'The UK-DALE dataset, domestic appliance-level electricity demand and whole-house demand from five UK homes', *Sci. Data*, 2015, **2**, p.150007

[24]   'Tesla Powerwall'. Available at https://www.tesla.com/en_GB/powerwall

[25]   'Green Energy UK offers first electricity tariff based on time of day'. Available at https://www.theguardian.com/money/2017/jan/03/green-energy-uk-launches-first-time-of-dayelectricity-tariff

[26]   Narzisi, G.: 'Multi-objective optimization, a quick introduction, New York University Lectures', 2008

[27]   Cover, T.M., Thomas, J.A.: '*Elements of information theory*' (Wiley-Interscience, New York, NY, USA, 1991)

[28]   Prudenzi, A.: 'A neuron nets based procedure for identifying domestic appliances pattern-of-use from energy recordings at meter panel'. Proc. IEEE Power Engineering Society Winter Meeting, New York, NY, USA, 2002, vol. 2, pp. 941–946

[29]   Engel, D., Eibl, G.: 'Wavelet-based multiresolution smart meter privacy', *IEEE Trans. Smart Grid*, 2017, **8**, (4), pp. 1710–1721