Privacy-Protecting Energy Management Unit through Model-Distribution Predictive Control

Jun-Xing Chin, Student Member, IEEE, Tomas Tinoco De Rubira, and Gabriela Hug, Senior Member, IEEE

Abstract—The roll-out of smart meters in electricity networks introduces risks for consumer privacy due to increased measurement frequency and granularity. Through various Non-Intrusive Load Monitoring techniques, consumer behavior may be inferred from their metering data. In this paper, we propose an energy management method that protects privacy through the minimization of information leakage. The method is based on a Model Predictive Controller that utilizes energy storage and local generation, and that predicts the effects of its actions on the statistics of the actual energy consumption of a consumer and that seen by the grid. Computationally, the method requires solving a Mixed-Integer Quadratic Program of manageable size whenever new meter readings are available. We simulate the controller on generated residential load profiles with different privacy costs in a two-tier time-of-use energy pricing environment. Results show that information leakage is effectively reduced at the expense of increased energy cost. The results also show that, using the proposed controller, the consumer load profile seen by the grid resembles a mixture between that obtained with Non-Intrusive Load Leveling and Lazy Stepping.

Index Terms—Consumer Privacy, Energy Management Unit, Model Predictive Control, Mutual Information, Optimization Methods, Smart Meter

I. INTRODUCTION

Globally, traditional electromechanical electricity meters are being replaced by Smart Meters (SMs) as part of efforts to modernize the grid in order to better manage electricity generation and distribution. In Europe, this is mandated by the European Commission under its Third Energy Package, requiring member states to roll-out SMs where cost-benefit analyses are positive [1]. However, with their ability to provide real-time information on consumer demand, SMs also raise concerns regarding consumer privacy [2], [3]. By measuring energy demand at much higher frequencies than traditional meters, detailed consumer load profiles can be extracted using SM data. Through Non-Intrusive Load Monitoring (NILM) techniques, which were first studied by Hart in 1989 [4] and further developed through the years, e.g., [5], [6], individual appliance usage, and ultimately consumer lifestyle patterns, preferences, and occupancy profiles can be inferred [7]–[9]. Such data may not only be accessible by Utility Providers (UPs), as metering infrastructures are vulnerable to attacks that lead to information leakage [10]. Moreover, if one also considers UPs being untrustworthy parties, privacy concerns stemming from SM roll-outs become much more severe, and can even lead to backlashes against their installation [11].

To counter this, various works have been done in order to protect consumer privacy. While some focus on cyber measures, such as encryption, adding noise to achieve differential privacy [12], and consumer aggregation, others address the problem through physical means, e.g., load control [13] and battery load hiding (BLH) [14], [15]. BLH techniques mask consumer load by charging and discharging a battery to alter the load profile captured by SMs. For example, in [14], the authors use a best-effort algorithm to hide consumer load change, whilst [15] details a load-leveling method that switches its energy consumption based on the battery’s state-of-charge. The authors in [16], on the other hand, introduce a stepping algorithm that improves upon the performance of best-effort and load-leveling algorithms. While privacy loss metrics, e.g., Mutual Information and Fischer Information, have been used to assess BLH schemes such as in [16], the schemes themselves have been primarily heuristics-driven. Cyber-physical methods proposed for minimizing directly a specific privacy loss measure such as mutual information have so far been mainly theoretical [17], [18].

In general, BLH methods proposed in the literature for privacy protection have not directly considered energy cost in the control policies. Any cost savings obtained have been a by-product of the privacy protection scheme. In these methods, privacy protection is closely linked to storage capacity [16], [19], but this alone is unlikely to justify the cost of investment on high-capacity storage devices, which according to estimates by [20] remains high. Utilizing energy storage for minimizing energy cost in addition to providing privacy protection is critical for justifying this investment. Methods that focus on minimizing energy cost alone have been the subject of various recent works. For example, the authors in [21], [22] utilize batteries in a Model Predictive Control (MPC) scheme to minimize energy costs in a varying-price environment by charging during low-price periods to compensate consumer load during high-price periods.

In this paper, we propose a new scheme based on MPC that combines energy cost minimization with privacy protection that directly minimizes mutual information. The proposed scheme introduces binary variables in the MPC optimization subproblems for counting predicted observations and estimating the joint statistics of consumer load and the net load seen by the grid (grid load). This allows the controller to predict the effects of its actions on the mutual information between these two load profiles. Computationally, the proposed scheme requires solving Mixed-Integer Quadratic Programs (MIQPs) of manageable size whenever new meter readings are available.

The rest of this paper is structured as follows: Section II describes the problem considered. Section III presents our proposed solution approach. Section IV details implementation and numerical experiments. Lastly, Section V summarizes this work and provides an outlook on future research.
II. Problem Description

We consider the problem of minimizing energy cost and reducing information leakage about the load profile of an energy consumer. The system considered consists of an Energy Management Unit (EMU) that uses energy storage, local generation, and energy available from the grid to supply the consumer load in a cost-effective and privacy-concerned manner. This is illustrated in Fig. 1. Mathematically, the system is represented by the discrete-time random process

\[ \{(S_t, G_t, C_t, X_t, Y_t) \mid t \in \mathbb{Z}_+\}, \]

where \( S_t \) is the energy supplied to the energy storage, \( G_t \) is the energy supplied by the local generation, \( C_t \) is the price of energy, \( X_t \) is the consumer energy demand, and \( Y_t \) is the energy demand seen by the grid during time interval \( t \), i.e.,

\[ Y_t = X_t + S_t - G_t. \tag{1} \]

The realizations of these random variables are denoted by \( s_t, g_t, c_t, x_t, \) and \( y_t \), respectively, for each time \( t \).

The battery charging energy \( S_t \) is constrained by

\[ S_{\min} \leq S_t \leq S_{\max} \tag{2} \]
during each time interval \( t \), where \( -S_{\min} \geq 0 \) is the maximum discharging energy and \( S_{\max} \geq 0 \) is the maximum charging energy over a single time interval. These constraints are given by the power rating of the device. It is assumed that the charge and discharge efficiencies are equal, i.e., \( \eta_c = \eta_d \). The state of charge of the energy storage device at the beginning of time interval \( t \) is denoted by \( E_t \) and satisfies

\[ E_{t+1} = E_t + \alpha S_t \tag{3} \]

for each \( t \), where

\[ \alpha := \begin{cases} \eta_c, & S_t \geq 0 \\ 1/\eta_d, & S_t < 0. \end{cases} \]

It is also assumed for simplicity that the energy storage device is not used to provide ancillary services to the grid, and as such is constrained by \( 0 \leq E_t \leq E_{\max} \), where \( E_{\max} \) is the maximum energy storage capacity of the device. The local generation energy \( G_t \) is constrained by

\[ 0 \leq G_t \leq G_{\max}, \]

where \( G_{\max} \) is the maximum energy output of the device over a time interval. It is assumed in our system that the local generation is deterministic in nature and that the EMU has perfect prior knowledge regarding its output. Additionally, there is no generation curtailment. The consumer loads \( X_t \) as well as the energy prices \( C_t \) are also assumed to be known to the EMU. In practice, only noisy predictions would be available. With regards to the grid, we assume that no energy feed-in to the grid and no energy wastage are allowed.

The energy cost at each time interval \( t \) is given by \( C_t Y_t \).

On the other hand, loss of consumer privacy is measured by the information leakage rate of consumer load given grid load, which following [23] is given by

\[ I(X^{[t]}; Y^{[t]}) := \int_Y \int_X p_{XY}(x_0, y_0, \ldots, x_t, y_t) \times \log \frac{p_{XY}(x_0, y_0, \ldots, x_t, y_t)}{p_X(x_0, \ldots, x_t) p_Y(y_0, \ldots, y_t)}. \]

Here, \( X^{[t]} := (X_0, \ldots, X_t) \), \( Y^{[t]} := (Y_0, \ldots, Y_t) \), \( p_{XY} \) and \( p_Y \) denote the probability density functions of \( (X^{[t]}, Y^{[t]}) \), \( X^{[t]} \) and \( Y^{[t]} \), respectively, and \( \log \) denotes the base-2 logarithm. Quantifying the cost of privacy loss through a non-negative constant \( \mu \) with units of Rp-per-bit (100 Rp = 1 CHF), we seek to find a causal and implementable control policy \( U \in \mathcal{U} \) that minimizes the total time-average expected cost. Such a policy determines the charge and discharge of the energy storage device during each time interval \( t \) based on the observation history up to that time, i.e.,

\[ S_t = U(X^{[t]}, C^{[t]}, Y^{[t-1]}, G^{[t-1]}), \]

where \( C^{[t]} \), \( G^{[t]} \) and \( S^{[t]} \) are defined in analogous ways to \( X^{[t]} \) and \( Y^{[t]} \). Mathematically, the problem can be posed as finding a policy \( U^* \) such that

\[ U^* = \arg \min_{U \in \mathcal{U}} f(U), \tag{4} \]

where

\[ f(U) := \lim_{T \to \infty} \frac{1}{T+1} \left\{ \sum_{t=0}^{T} E[C_t Y_t] + \mu I(X^{[T]}; Y^{[T]}) \right\} \]

and \( E[\cdot] \) denotes expectation.

III. Model-Distribution Predictive Control

We propose an approach based on model predictive control (MPC) to find a control policy that is close to \( U^* \). At time \( t \), the controller (EMU) observes the realizations \( x_t, g_t, c_t \), and \( e_t \) (energy storage state of charge), determines actions \( s_t, \ldots, s_{t+T} \) for a prediction horizon of length \( T \), executes action \( s_t \) to achieve a desirable grid load \( y_t \), and repeats the process in a receding horizon manner.

Since the exact evaluation of the information leakage rate \( I(X^{[t]}; Y^{[t]}) \) is not possible without knowing the probability density functions \( p_{XY} \) and \( p_Y \), which depend on the control actions, we propose using an approximation. This is done by first assuming that \( X_t \) are independent identically distributed (i.i.d) samples of a random variable \( X_t \) near time \( t \), and similarly that \( Y_t \) are i.i.d samples of a random variable \( Y_t \) near time \( t \). More specifically, this is assumed during the time window \( \{t + T - N + 1, \ldots, t + T\} \), which has
length \( N \gg T \) and covers both the recent past and the entire prediction horizon with respect to time \( t \). While time independence does not hold in reality, we make this strong assumption as a first step in making the problem tractable. Time correlation and other properties will be studied in future work. Letting \( X^{[i]} := (X_{t-M+1}^{[i]}, \ldots, X_{t+T}^{[i]}) \) and \( Y^{[i]} := (Y_{t-M+1}^{[i]}, \ldots, Y_{t+T}^{[i]}) \), where \( M := N - T \), this assumption gives

\[
I(X^{[t+T]}, Y^{[t+T]}) = \frac{t + T + 1}{N} I(X^{[t]}, Y^{[t]}) = (t + T + 1) I(\hat{X}_t; \hat{Y}_t).
\]

(5)

Furthermore, for the purpose of approximating mutual information only, we assume that the random variables \( \hat{X}_t \) and \( \hat{Y}_t \) can take on only a finite number of energy levels, i.e., \( \hat{X}_t \in \{x^1, x^2, \ldots, x^m\} \) and \( \hat{Y}_t \in \{y^1, y^2, \ldots, y^n\} \). The minimum values are zero, while the maximum values can be computed from historical records and the connection capacity available on the distribution feeder cables (or circuit breakers). From this assumption, it follows that

\[
I(\hat{X}_t; \hat{Y}_t) = \sum_{i=1}^{m} \sum_{j=1}^{n} p_{\hat{X}, \hat{Y}}(x^i, y^j) \log \frac{p_{\hat{X}, \hat{Y}}(x^i, y^j)}{p_{\hat{X}}(x^i)p_{\hat{Y}}(y^j)},
\]

(6)

where \( p_{\hat{X}, \hat{Y}}, p_{\hat{X}}, p_{\hat{Y}} \) denote the probability mass functions of \( (\hat{X}_t, \hat{Y}_t), \hat{X}_t, \hat{Y}_t \) and \( \hat{Y}_t \) given \( \hat{X}_t \), respectively. Since \( \hat{X}_t \) and \( \hat{Y}_t \) are assumed to be i.i.d samples of \( \hat{X}_t \) and \( \hat{Y}_t \), respectively, for \( t \in \{t+T - N+1, \ldots, t+T\} \), the statistics of these random variables are approximated by the relative frequency of events during this time window, as illustrated in Fig. 2. More specifically,

\[
p_{\hat{X}, \hat{Y}}(x^i, y^j) \approx \sum_{\tau=t-M+1}^{t+T} \mathbb{I}\{(x_\tau, y_\tau) = (x^i, y^j)\} + \varepsilon
\]

\[
p_{\hat{X}}(x^i) \approx \frac{\sum_{\tau=t-M+1}^{t+T} \mathbb{I}\{x_\tau = x^i\} + n \varepsilon}{N + m \varepsilon},
\]

\[
p_{\hat{Y}}(y^j) \approx \frac{\sum_{\tau=t-M+1}^{t+T} \mathbb{I}\{y_\tau = y^j\} + m \varepsilon}{N + m \varepsilon},
\]

where \( \mathbb{I}\{A\} = 1 \) when \( A \) is true and 0 otherwise, and \( \varepsilon > 0 \). The addition of the positive scalar \( \varepsilon \) in the probability estimates corresponds to additive smoothing [24]. This is done to avoid probability estimates of zero for events for which there are no observations during the time window. This probability estimation strategy, albeit simplistic, constitutes an adequate choice for a first step towards exploring the proposed methodology of including in MPC the ability to estimate the statistical effects of the control actions in a tractable manner. Including more sophisticated probability estimation techniques into this methodology is an interesting subject for future work.

Since at time \( t \) the counting window covers the recent past and the prediction horizon, the probability estimates can be separated into parts that are constants and parts that depend on the controller’s actions during the prediction horizon. More specifically, they can be expressed as

\[
p_{\hat{X}, \hat{Y}}(x^i, y^j) \approx a^{ij}_t + \frac{1}{N \varepsilon} \sum_{\tau=t}^{t+T} z^{ij}_\tau
\]

(7)

\[
p_{\hat{Y}}(y^j) \approx b^j_t + \frac{1}{N \varepsilon} \sum_{\tau=t}^{t+T} \sum_{k=1}^{m} z^{kj}_\tau,
\]

(8)

where \( N \varepsilon := N + m \varepsilon \), \( a^{ij}_t \) and \( b^j_t \) are constants (at time \( t \)) defined by

\[
a^{ij}_t := \sum_{\tau=t-M+1}^{t+T} \mathbb{I}\{(x_\tau, y_\tau) = (x^i, y^j)\} + \varepsilon
\]

\[
b^j_t := \sum_{\tau=t-M+1}^{t+T} \mathbb{I}\{y_\tau = y^j\} + m \varepsilon,
\]

and \( z^{ij}_\tau := \mathbb{I}\{(x_\tau, y_\tau) = (x^i, y^j)\} \) for \( \tau \in \{t, \ldots, t+T\} \), \( i \in \{1, \ldots, m\} \) and \( j \in \{1, \ldots, n\} \) are binary variables that depend on the controller’s actions and forecasts during the prediction horizon. Letting \( c^i_t := p_{\hat{X}}(x^i) \), which is a constant at time \( t \) due to the perfect prediction assumption on consumer load, it follows from (7) and (8) that

\[
p_{\hat{X}}(x^i | t) \approx c^i_t \left( a^{ij}_t + \frac{1}{N \varepsilon} \sum_{\tau=t}^{t+T} z^{ij}_\tau \right)
\]

From these expressions for probability estimates and (6), we then have the approximation

\[
I(\hat{X}_t; \hat{Y}_t) \approx \sum_{i=1}^{m} \sum_{j=1}^{n} \left( a^{ij}_t + \frac{1}{N \varepsilon} \sum_{\tau=t}^{t+T} z^{ij}_\tau \right) \times \left\{ \log \left( a^{ij}_t + \frac{1}{N \varepsilon} \sum_{\tau=t}^{t+T} z^{ij}_\tau \right) - \log \left( b^j_t + \frac{1}{N \varepsilon} \sum_{\tau=t}^{t+T} \sum_{k=1}^{m} z^{kj}_\tau \right) - \log c^i_t \right\},
\]

(9)

the right-hand side of which is a fairly complicated function of the binary variables \( z^{ij}_\tau \). Since \( N \gg T \), this approximation can be simplified by replacing the logarithms with their first-order Taylor expansions around the constant part of their arguments, i.e.,

\[
\log \left( a^{ij}_t + \frac{1}{N \varepsilon} \sum_{\tau=t}^{t+T} z^{ij}_\tau \right) \approx \log a^{ij}_t + \log' \left( a^{ij}_t \right) \left( \frac{1}{N \varepsilon} \sum_{\tau=t}^{t+T} z^{ij}_\tau \right)
\]

\[
\log \left( b^j_t + \frac{1}{N \varepsilon} \sum_{\tau=t}^{t+T} \sum_{k=1}^{m} z^{kj}_\tau \right) \approx \log b^j_t + \log' \left( b^j_t \right) \left( \frac{1}{N \varepsilon} \sum_{\tau=t}^{t+T} \sum_{k=1}^{m} z^{kj}_\tau \right).
\]
Using these Taylor expansions and letting \( \nu := 1/\log_e 2 \), we obtain the approximation
\[
I(\tilde{X}_t; \tilde{Y}_t) \approx \sum_{i=1}^{m} \sum_{j=1}^{n} \left( a_{ij} + \frac{1}{N} \sum_{t=\tau}^{t+T} z_{ij}^t \right) \times \left\{ \log \frac{a_{ij}^{t+T}}{b_i^{t+T} c_i^{t+T}} + \nu \frac{t+T}{a_i N_c} \sum_{t=\tau}^{t+T} z_{ij}^t - \nu \frac{t+T}{b_i^c N_c} \sum_{t=\tau}^{t+T} z_{ij}^t \right\},
\]
the right-hand side of which is now a quadratic function of the binary variables \( z_{ij}^t \), and is denoted by \( I(\tilde{X}_t, \tilde{Y}_t) \).

An important observation is that the curvature of the logarithm functions grows rapidly as zero is approached. Since the constants \( a_{ij}^{t+T} \) and \( b_i^{t+T} \) can be potentially very small for some \( i \) and \( j \) due to the absence of observations of the associated events in the time window \( \{t-M+1, \ldots, t-1\} \) and a poor choice of \( \varepsilon \) for additive smoothing, the Taylor expansions of the logarithms can be very poor. Hence, we propose choosing a bound \( \gamma \) for \( \log' (a_{ij}^{t+T}) \) and \( \log' (b_i^{t+T}) \) that ensures that the logarithms are always expanded sufficiently away from zero. Since
\[
\log' (a_{ij}^{t+T}) = \frac{\nu}{a_i^{t+T}} \leq \frac{N + mn \varepsilon}{\varepsilon} \nu
\]
and
\[
\log' (b_i^{t+T}) = \frac{\nu}{b_i^{t+T}} \leq \frac{\nu}{a_i^{t+T}} = \log' (a_{ij}^{t+T}),
\]
it follows that setting \( \varepsilon \geq N/(\gamma \nu^{-1} - mn) \) ensures that both \( \log'' (a_{ij}^{t+T}) \) and \( \log' (b_i^{t+T}) \) are always bounded by \( \gamma \).

Using (5) and the definition of \( I(\tilde{X}_t, \tilde{Y}_t) \), the ideal objective function for the MPC optimization problem solved at time \( t \) is approximated as follows:
\[
\frac{1}{T+1} \sum_{\tau=t}^{t+T} c_r y_r + \frac{1}{T+1} \mu I(X^{t+T}; Y^{t+T}) \approx \frac{1}{T+1} \sum_{\tau=t}^{t+T} c_r y_r + \mu I(\tilde{X}_t; \tilde{Y}_t).
\]

Hence, the proposed MPC scheme consists of solving at each time \( t \) the optimization problem
\[
\text{minimize} \quad \frac{1}{T+1} \sum_{\tau=t}^{t+T} c_r y_r + \mu \Phi(z) \quad \text{subject to} \quad (s, c, y, z) \in \mathcal{F}_t,
\]
where \( \Phi(z) := I(\tilde{X}_t; \tilde{Y}_t) \) measures privacy loss.
\[
s := \{s_r|_{t+\tau}^{t+T}\}, \quad c := \{c_r|_{t+\tau}^{t+T}\}, \quad y := \{y_r|_{t+\tau}^{t+T}\}
\]
are continuous optimization variables, and \( z \) is the collection of binary optimization variables defined by
\[
z := \left\{ z_{ij}^t \mid \tau \in \{t, \ldots, t+T\}, \quad j \in \{1, \ldots, n\}, \quad i = \arg \min_{k \in \{1, \ldots, m\}} \|x_r - x_k\|_2 \right\}.
\]
Note that not all binary variables used in (9) are treated as optimization variables in (10). This is because many of them are known by the controller to be zero at time \( t \) since \( x_r \) is assumed to be known, and for each \( \tau \in \{t, \ldots, t+T\} \), there is exactly one \( i \) such that
\[
i = \arg \min_{k \in \{1, \ldots, m\}} \|x_r - x_k\|_2.
\]
The constraint \( (s, c, y, z) \in \mathcal{F}_t \) enforces the binary restrictions on \( z \), the system constraints (1), (2) and (3), the constraint \( \sum_{j=1}^{n} z_{ij}^t = 1 \) for each \( \tau \in \{t, \ldots, t+T\} \), and the constraint
\[
z_{ij}^t = 1 \iff j = \arg \min_{k \in \{1, \ldots, n\}} \|y_r - y^k\|_2
\]
for each \( \tau \in \{t, \ldots, t+T\} \) and \( j \in \{1, \ldots, n\} \), which are representable as linear constraints. Problem (10) is therefore an MIQP with \( n(T+1) \) binary variables, roughly \( 3(T+1) \) continuous variables, and roughly \( 7(T+1) \) constraints.

Since this proposed MPC scheme predicts the effects of the controller’s actions on the probability distribution of random variables, we refer to it as Model-Distribution Predictive Control, or MDPC.

### IV. NUMERICAL EXPERIMENTS

The proposed scheme was implemented in YALMIP [25] and MATLAB R2015a using the IBM CPLEX 12.6.3 solver. Simulations were done using consumer load profiles generated with tools from [26] over a period of approximately one month (30 days) with hourly resolution in a two-tier time-of-use energy pricing environment. The computer used for the simulations was an Intel Core i7-4710MQ CPU with 2.50 GHz of clock speed, 16.0 GB of RAM, and running 64-bit Windows 7 Enterprise.

In order to solve the MIQP problems (10) reliably and reduce the number of solution candidates, a convex penalty term \( r(y) \) based on the \( l_1 \)-norm was added to the objective function with a regularization coefficient or weight given by the non-negative scalar \( \sigma \). Without this term, the MIQP solver would spend an excessive amount of time performing branch-and-bound searches and utilize excessive memory resources trying to solve certain instances of (10). More details regarding the design of \( r(y) \) and its effects on the controller’s performance are discussed below.

Unless stated otherwise, the parameters shown in Table I were used for the simulations. Local generation was set to zero throughout the entire simulation period in order to clearly observe the effects of the controller actions. A battery was used as the energy storage device, and the parameters chosen were those of a Tesla Powerwall.

| Parameter | Value |
|-----------|-------|
| \( T \) | 12 | \( \text{Reg. Coefficient: } 10^{-3} \) |
| \( M \) | 120 | Battery Capacity: 6.4 kWh |
| \( N \) | 132 | Battery Power: 3.3 kW |
| \( m \) | 15 | Battery Efficiencies: 96 % |
| \( n \) | 15 | Energy Price (high): 24.6 Rp/kWh |
| \( z \) | 0.1 | Energy Price (low): 13.15 Rp/kWh |
A. Visualization of Load Profiles and Control Actions

Fig. 3 illustrates the actual consumer load and the load seen by the grid over seven days for two different prices of privacy loss, $\mu = 0$ and $\mu = 25$. For $\mu = 0$, the controller optimizes for energy costs only, and thus charges the battery during low-price periods and compensates consumer load during the high-price periods. This can be seen in Fig. 4 and 5, which show the charge and discharge of the battery, the energy prices, and the battery state of charge. With the chosen horizon length of 12, the controller is unable to fully utilize the battery’s capacity as the consumer load “seen” by it is generally less than the battery’s capacity. With a price of privacy loss of $\mu = 25$, the controller now charges over high-price periods as well, as seen in the top plot of Fig. 4, in order to mask low consumer load periods. This results in a grid load curve that shows a stepping behavior similar to that obtained in [16], with periods of load leveling similar to those obtained in [15]. Note that the correlation and thus also the mutual information of consumer and grid loads is greatly reduced for $\mu = 25$. This can be seen more clearly in Fig. 6, which illustrates the estimated distribution of $\tilde{X}_t$ and $\tilde{Y}_t$. For $\mu = 25$ (right plot), the grid load concentrates in the first four levels despite consumer load levels being spread out. For $\mu = 0$, while the correlation between $\tilde{X}_t$ and $\tilde{Y}_t$ is still reduced by the battery charging during low-price hours to compensate consumption during high-price hours, it is still possible for an adversary to reconstruct the consumer load based on observations of the grid load with knowledge of energy prices and the battery system specifications, as discussed in [16].

B. Performance of the MDPC Scheme

The proposed scheme was evaluated using the right-hand side of (9) with a static window of length $N = 718$ that encompasses the entire simulation period of approximately 30 days. This “cumulative” approximation of mutual information is denoted by $I_c$. For reference, a system without a battery using the same simulation setup had a total energy consumption of 535 kWh, energy costs of 125 CHF, and privacy loss of $I_c = 2.58$ bits.
In Fig. 7, the total energy costs and mutual information $I_c$ for different battery capacities are shown as a function of the price $\mu$ of privacy loss. As can be seen from the figure, energy costs increase with the drop in mutual information, as the battery charges during high-price periods to mask low consumer loads, as previously shown in Fig. 4. While it would be possible for consumer behavior to be inferred with $\mu = 0$, the mutual information value with a 1.32 kWh battery for energy cost optimization only is already approximately half that of the no-battery case.

Similar to the findings in [16], the size of the battery relative to the total consumer load over a high-low consumption cycle determines the level of privacy protection that can be achieved. Privacy loss is minimal with a battery capacity that sufficiently covers the consumer load during a typical day. In the case of the simulated consumer load, this battery capacity is approximately 6.4 kWh, with a 12.8 kWh battery only achieving slightly better performance in terms of privacy loss reduction. However, energy cost reduction improves with a larger battery, since it also allows covering the consumption during periods of high price and low load.

As seen in Fig. 7, minimum privacy loss is achieved with a price of privacy loss of about 35 Rp/bit. Beyond this price of privacy loss, the system limitations impede the controller from doing better, and the mutual information is seen to fluctuate around the minimum value. These fluctuations are caused by the inherent randomness that is present in the scheme, as seen in the top plot of Fig. 8, which shows how $I_c$ varies as the price of privacy loss increases even when the energy cost is ignored. It can be seen that the “width” of these fluctuations is roughly 0.2 bits. They arise from the MIQP solver finding multiple solution candidates and choosing different ones, and hence affecting the system trajectory, purely due to the scaling of the objective function, which is completely controlled by $\mu$ when the energy cost is ignored. The regularization term $\sigma r(y)$ added to the objective, which is designed to encourage the controller to choose a set of actions at time $t$ that is similar to those predicted at time $t - 1$ when the forecasts are similar, reduces the amplitude of these fluctuations. This is seen in the bottom plot of Fig. 8, where $I_c$ curves using different levels of regularization when considering energy cost are shown. While lower privacy loss is achievable with less or no regularization, the fluctuation amplitudes are larger. Increasing $\sigma$ reduces the fluctuations in $I_c$, but at the expense of privacy protection.

The effect of the number of quantization levels of $\hat{X}_t$ and $\hat{Y}_t$, which are $m$ and $n$, respectively, and of the charge-discharge rate limit of the battery on the performance of the proposed controller were also studied. Fig. 9 shows $I_c$ as a function of $\mu$ for different number of quantization levels and battery power ratings. For the evaluation of the quantization levels, $(m, n) = (20, 20)$ was used. Increasing the number of quantization levels for the same maximum load allows the controller to potentially achieve lower levels of privacy loss because it is now able to take advantage of the higher resolution to fine-tune its control actions. However, this also generally increases the fluctuation amplitude or variability of $I_c$ due to the MIQP solver finding a larger set of candidate solutions. Nonetheless, smaller quantization levels could also result in large fluctuations when evaluated at higher resolutions due to the coarse level of controller actions, such as in the case of $(m, n) = (10, 10)$. The battery power ratings on the other hand impose a limit on the controller’s ability to mask changes in consumer load. For the considered simulation setup with a maximum consumer load of 3.6 kW, a battery rating of 3.3 kW is almost sufficient. This can be seen in the bottom plot of Fig. 9, where a battery with a 6.6 kW power rating marginally improves performance. A battery with a 1.65 kW rating is sufficient for lower prices of privacy loss, as energy cost reduction is the main driver of control actions in such cases. However, at higher prices of privacy loss the controller is unable to fully mask the consumer load peaks that occur mainly during high-price periods, as shown in Fig. 4.

Additionally, as discussed in Section IV-A, a horizon length of $T = 12$ is insufficient for the controller to predict the load and fully charge the battery to minimize energy costs and mask consumer loads during the later stage of high energy price periods. Hence, the controller was simulated with different horizon lengths to study its performance and computational requirements. Table II summarizes the results obtained. As can be seen from the table, prediction length greatly impacts the controller’s performance because it allows it to anticipate changes that occur farther in the future. Increasing $T$ to 18 reduces energy costs, and generally also reduces privacy-loss, due to the controller being able to fully utilize the capacity of the 6.4 kWh battery. The case of $\mu = 35$ is atypical due to the randomness in the scheme, which is exacerbated by longer horizons since they enlarge the solution space and increase the number of candidate MIQP solutions. It is important to note...
that computational tractability is still maintained with $T = 18$ despite having a median solver time seven times that of $T = 12$. However, $T = 24$ is computationally impractical for a realistic controller as it requires almost nine hours using the current implementation to solve one iteration of the problem that needs to be completed in under an hour to be practical. Note that for the energy pricing structure and consumer load profile in the simulation setting, a horizon length of 18 appears to be sufficient. However, this should not be generalized to other simulation settings and load profiles. In general, using the scheme proposed, a controller that is able to predict longer time periods while maintaining tractability and keeping the size of the solution space manageable should perform better.

C. Time Series Variability

In order to study the change in mutual information with time, we evaluated the controller’s performance with $\mu = 20$ for the three different horizon lengths discussed in Section IV-B using the right-hand side of (9) with a moving window of length $N = 132$. This “time-dependent” approximation of mutual information is denoted by $I_t$ for each time $t$. For each time $t$, the window considered the observation history from time $t − N + 1$ to time $t$, as shown in Fig. 2, and hence it was based on the actual implemented controller actions. The graphs obtained of $I_t$ are shown in the top plot of Fig. 10. As seen from the plot, the variability of $I_t$ decreases with increasing horizon length as the battery is able to better anticipate load and price changes, which allows a more gradual adjustment of mutual information in order to balance energy costs and privacy loss. For the current implementation and load profiles, improvements beyond $T = 18$ are marginal as previously mentioned in Section IV-B, and may even lead to more variability. The variability of $I_t$ of the different prediction horizon lengths was quantified using total variation. The results, which are consistent with the variability of the curves in the top plot of Fig. 10, are shown in Table III. Although longer horizon lengths reduce the variation of $I_t$ as a function of time, they do not eliminate it. The reason for this is that the variability is also caused by the numerous non-unique optimal points of the MIQPs that increase with the horizon length $T$, and by the logarithm linearizations done by the controller, which make it overestimate the mutual information to varying degrees depending on the statistics of $\tilde{X}_t$ and $\tilde{Y}_t$. This can be seen in the bottom plot of Fig. 10, which shows $I_t$ now for the time window $\{t − M + 1, \ldots, t + T\}$ at each time $t$ for $\mu = 20$ and $T = 12$ and the approximation $\tilde{I}(\tilde{X}, \tilde{Y})$ with linearized logarithms used by the controller. As seen on the plot, at certain times, the discrepancy introduced by the linearization leads to the controller taking actions that are expected to reduce mutual information, but in actual fact increases it. For example, this can be seen around the time corresponding to 11 PM on the 23rd of January. Around 1 AM on the 30th of January, on the other hand, this is reversed: mutual information is reduced when the controller predicted actions that supposedly increased mutual information.

\begin{table}[h]
\centering
\caption{Comparison Between Different Horizon Lengths.}
\begin{tabular}{|c|c|c|c|c|c|c|c|c|}
\hline
 & $T = 12$ & $T = 18$ & $T = 24$ & $T = 12$ & $T = 18$ & $T = 24$ & $T = 12$ & $T = 18$ & $T = 24$ \\
\hline
Max Solver Time (s) & 0.196 & 0.881 & 0.450 & 1.94 & 15.2 & 5110 & 3.56 & 61.2 & 6610 \\
Min Solver Time (s) & 0.021 & 0.040 & 0.043 & 0.075 & 0.182 & 1.19 & 0.082 & 0.609 & 1.56 \\
Mean Solver Time (s) & 0.044 & 0.089 & 0.073 & 0.345 & 2.96 & 81.9 & 0.456 & 4.17 & 105 \\
Median Solver Time (s) & 0.031 & 0.073 & 0.062 & 0.308 & 2.11 & 27.8 & 0.382 & 2.56 & 42.3 \\
Mutual Information $I_t$ (bits) & 0.821 & 0.749 & 0.903 & 0.387 & 0.242 & 0.263 & 0.276 & 0.464 & 0.305 \\
Total Energy Costs (CHF) & 108 & 106 & 106 & 113 & 108 & 109 & 115 & 110 & 109 \\
Total Energy Used (kWh) & 547 & 549 & 549 & 552 & 554 & 553 & 552 & 553 & 554 \\
\hline
\end{tabular}
\end{table}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig10.png}
\caption{Effects of prediction horizon lengths on time variability of mutual information (top) and effects of logarithm linearization (bottom)}
\end{figure}

\begin{table}[h]
\centering
\caption{Total variation of $I_t$}
\begin{tabular}{|c|c|c|}
\hline
 & $T = 12$ & $T = 18$ & $T = 24$ \\
\hline
Total Variation & 2.30 & 1.72 & 2.08 \\
Reduction (%) & 0% & −25.2% & −9.57% \\
\hline
\end{tabular}
\end{table}

V. CONCLUSION

In this paper, we proposed an MPC-based controller that balances energy costs and privacy protection based on mutual information. This is done by predicting the effects of the controller’s actions on the statistics of the consumer load and that seen by the grid using counting, and solving MIQP problems of manageable size whenever new meter readings are available. While not computationally scalable with prediction horizon and discretization levels of load seen by the grid, the results obtained showed that the controller is able to reduce information leakage and hence protect consumer privacy. Extensive experiments were carried out in this work to thoroughly characterize the performance of the proposed controller. In particular, the effects of battery size, power ratings, prediction horizon lengths, discretization and regularization were studied. An important result of this work is that it shows that practical schemes that specifically target the minimization of a rigorous
measure of privacy loss, such as the one proposed here, are possible.

Future work will focus on enlarging the reach of the prediction horizon, accounting for time correlation of the load consumptions, including more sophisticated probability estimation techniques, comparing the approach to other BLH schemes, and exploring MIQP relaxation methods to make the scheme more computationally scalable.

REFERENCES

[1] European Union, “Directive of 2009/72/EC of the European Parliament and of the Council of 13 July 2009 Concerning Common Rules for the Internal Market in Electricity and Repealing Directive 2003/54/EC,” Official Journal of the European Union, vol. L211, no. August, pp. 55 – 93, 2009.

[2] E. L. Quinn, “Privacy and the New Energy Infrastructure;” Tech. Rep. 09, 2009. [Online]. Available: http://papers.ssrn.com/abstract=1370731

[3] P. McDaniel and S. McLaughlin, “Security and privacy challenges in the smart grid,” IEEE Security and Privacy, vol. 7, no. 3, pp. 75–77, 2009.

[4] G. W. Hart, “Residential energy monitoring and computerized surveillance via utility ‘power flows’,” IEEE Technology and Society Magazine, vol. 8, no. 2, pp. 12–16, 1989.

[5] H. Altrabalsi, J. Liao, L. Stankovic, and V. Stankovic, “A low-complexity energy disaggregation method: Performance and robustness,” in IEEE Symposium on Computational Intelligence Applications in Smart Grid (CIASG), 2014.

[6] Z. Guanchen, G. Wang, H. Farhangi, and A. Palizban, “Residential electric load disaggregation for low-frequency utility applications,” in IEEE Power & Energy Society General Meeting, 2015.

[7] A. Molina-Markham, P. Shenoy, K. Fu, E. Cecchet, and D. Irwin, “Private memoirs of a smart meter,” in Proceedings of the 2nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Building. Zurich, Switzerland: ACM, 2010, pp. 61–66.

[8] M. A. Lisovich, D. K. Mulligan, and S. B. Wicker, “Inferring personal information from demand-response systems,” IEEE Security and Privacy, vol. 8, no. 1, pp. 11–20, 2010.

[9] U. Grevenler, B. Justus, and D. Loehr, “Multimedia content identification through smart meter power usage profiles,” Computers, Privacy and Data Protection, vol. 1, 2012.

[10] S. McLaughlin, D. Podkuiko, S. Miasdzvezhanka, A. Delozier, and P. McDaniel, “Multi-vendor penetration testing in the advanced metering infrastructure,” in Proceedings of the 26th Annual Computer Security Applications Conference, Austin, Texas, USA, 2010, pp. 107–116.

[11] R. Hoenkamp, G. B. Huijema, and A. J. C. de Moor-van Vugt, “The neglected consumer: the case of the smart meter rollout in the Netherlands,” Renewable Energy Law and Policy Review, no. 4, pp. 269–282, 2011.

[12] G. Acs and C. Castelluccia, “I have a DREAM! (DiffeRentially PrivatE smart Metering),” The 13th Information Hiding Conference (IH), pp. 118–132, 2011.

[13] C. Dong, S. Kaia, D. Irwin, P. Shenoy, and J. Albrecht, “Preventing occupancy detection from smart meters,” IEEE Transactions on Smart Grid, vol. 6, no. 5, pp. 2426–2434, 2015.

[14] G. Calogridis, C. Efthymiou, S. Z. Denic, T. A. Lewis, and R. Cepeda, “Privacy for Smart Meters: Towards Undetectable Appliance Load Signatures,” in First IEEE International Conference on Smart Grid Communications (SmartGridComm), 2010, pp. 232–237.

[15] S. McLaughlin, “Protecting consumer privacy from electric load monitoring,” in Proceedings of the 18th ACM Conference on Computer and Communications Security (CCS ’11), Chicago, Illinois, USA, 2011, pp. 87–98.

[16] W. Yang, N. Li, Y. Qi, W. Qardaji, S. McLaughlin, and P. McDaniel, “Minimizing private data disclosures in the smart grid,” in Proceedings of the 2012 ACM Conference on Computer and Communications Security (CCS ’12). Raleigh, North Carolina, USA: ACM, 2012, p. 415.

[17] G. Giaconi and D. Gunduz, “Smart meter privacy with renewable energy and a finite capacity battery,” in IEEE Workshop on Signal Processing Advances in Wireless Communications, SPAWC, vol. 2016-August, Edinburgh, UK, 2016.

[18] S. Li, A. Khisti, and A. Mahajan, “Structure of optimal privacy-preserving policies in smart-metered systems with a rechargeable battery,” in IEEE Workshop on Signal Processing Advances in Wireless Communications, SPAWC, 2015, pp. 375–379.

[19] O. Tan, D. Gunduz, and H. V. Poor, “Increasing smart meter privacy through energy harvesting and storage devices,” IEEE Journal on Selected Areas in Communications, vol. 31, no. 7, pp. 1331–1341, 2013.

[20] IRENA, “Battery Storage for Renewables: Market Status and Technology Outlook,” Tech. Rep., 2015. [Online]. Available: http://www.irena.org/DocumentDownloads/Publications/IRENA_BatteryStorage_report_2015.pdf

[21] A. Khakimova, A. Shamshimova, D. Sharipova, A. Kusatayeva, V. Ten, A. Bemporad, Y. Familiant, A. Shintemirov, and M. Rubagotti, “Hybrid model predictive control for optimal energy management of a smart house,” in IEEE 15th International Conference on Environment and Electrical Engineering (EEEIC), 2015, pp. 513–518.

[22] W. Haining, M. Ke, D. Zhao Yang, X. Zhao, L. Fengji, and W. Kit Po, “Efficient real-time residential energy management through MILP based rolling horizon optimization,” in IEEE Power & Energy Society General Meeting, 2015.

[23] A. Galka, T. Ozaki, and O. Yamashita, “A new approach to mutual information between pairs of times series;,” in International Symposium on Nonlinear Theory and its Applications (NOLTA2005), Bruges, Belgium, 2005, pp. 626–629.

[24] C. D. Manning, P. Raghavan, and H. Schutze, An Introduction to Information Retrieval. Cambridge University Press, 2009.

[25] J. Lotfberg, “YALMIP : a toolbox for modeling and optimization in MATLAB,” in IEEE International Conference on Computer Aided Control Systems Design, 2004, pp. 284–289.

[26] N. Pfugradt and W. Schafft, “Analysing low-voltage grids using a behaviour based load profile generator,” Renewable Energy and Power Quality Journal, vol. 1, no. 11, pp. 1–5, 2013.

Jun-Xing Chin (S’15) received the B. Eng. degree in electronic engineering and the B.BA degree in general business management from the Hong Kong University of Science and Technology in 2008, and the M.S. degree in electrical technology for sustainable and renewable energy systems from the University of Nottingham, the United Kingdom in 2011. From 2011 to 2015, he was an assistant engineer at the engineering consultancy firm Ove Arup and Partners H.K. Ltd. in Hong Kong. Currently, he is working towards his Ph.D. degree at the Power Systems Laboratory at ETH Zurich, with research emphasis on consumer privacy for smart metering systems, and short-term load forecasting.

Tomas Tinoco De Rubira received the B.S. degree in Electrical and Computer Engineering from the University of California Berkeley in 2008, and the M.S. and Ph.D. degrees in Electrical Engineering from Stanford University in 2011 and 2015, respectively. He is currently a Postdoctoral Research Fellow at the Power Systems Laboratory of ETH Zurich. His research consists on developing and applying mathematical optimization techniques for solving challenging problems, such as enabling large-scale integration of renewable energy in electric power grids. His employment experience includes Aurora Solar Inc. and the Electric Power Research Institute.

Gabriela Hug (S’05, M’08, SM’14) was born in Baden, Switzerland. She received the M.Sc. degree in electrical engineering from the Swiss Federal Institute of Technology in Zurich (ETH Zurich) in 2004 and the Ph.D. degree from the same institution in 2008. After her Ph.D., she worked in the Special Studies Group at Hydro One in Toronto, Canada. From 2009 to 2015, she was an Assistant Professor at Carnegie Mellon University in Pittsburgh, Pennsylvania. Currently, she is an Associate Professor in the Power Systems Laboratory at ETH Zurich. Her research is dedicated to the control and optimization of electric power systems.