Distributed Differentially Private Model Predictive Control for Energy Storage

M. Zellner [*], **T. Tinoco De Rubira** [*], **G. Hug** [*],
M. N. Zeilinger [**],

[*] *Power Systems Laboratory, ETH Zurich, Switzerland*
[**] *Institute for Dynamic Systems and Control, ETH Zurich*

Abstract: Smart meters enable a variety of new and useful applications for achieving a smart grid. Unfortunately, they also pose privacy risks by revealing sensitive information about consumers. In this work, techniques that exploit the benefits of smart meters and at the same time mitigate privacy risks are explored. More specifically, we consider the control of local energy storage devices of a group of consumers with the goal of minimizing energy cost and providing aggregate load smoothing to the grid. A differentially private distributed model predictive controller (MPC) is proposed by extending a framework recently proposed in the literature for this application. By using a distributed proximal gradient algorithm, the energy cost of consumers is minimized while keeping the load profile of each consumer private. Aggregate load smoothing is achieved by exchanging net consumption information between consumers. Since information about the consumer load profiles could be inferred from theses exchanges, the proposed controller is designed to ensure Differential Privacy (DP) and to protect this information. Numerical experiments are used to study a case with 30 consumers. The resulting trade-off between privacy and performance is studied as well as the effects of having a trustworthy or untrustworthy mediator handling the information exchanges between consumers.

Keywords: MPC, Differential Privacy, Power systems, Smart grids, Energy Storage Operation

1. INTRODUCTION

Smart Meters are advanced energy meters that provide more precise measurements for the consumption of electrical energy as well as increased temporal resolution. The detailed data available from a smart meter enables a variety of new and useful applications for the grid. A promising new opportunity that is considered in this paper is the use of local storage devices for minimizing the energy cost of consumers while providing certain services to the grid, *e.g.*, demand or load smoothing. Smart meters are gradually being deployed in more and more countries. According to an estimate by the European Commission (2014), there will be an expected installed base of 200 million smart meters by 2020 in the European Union. Despite their benefits, a big hurdle for the deployment of these measurement devices are security and privacy concerns, which have triggered an intense public debate and propelled significant research efforts. For example, the possibility to detect which movie is watched based on smart meter data is shown by Greveler et al. (2012). Examples of the type of private information that can be extracted or inferred from smart meters are described by Molina-Markham et al. (2010). Other works have explored ways to quantify and mitigate privacy issues with smart meters while still exploiting their benefits and have been largely based on information theory. For example, an approach for calculating the information leakage with the help of a finite state model is described by Varodayan and Khisti (2011). In Rajagopalan et al. (2011), smart meter data is protected by obfuscation and the privacy-utility trade-off is analyzed. Gómez-Vilardebó and Gündüz (2015) quantify privacy by mutual information and achieve privacy protection by distorting the net load profile using an alternative energy source. In Giaconi et al. (2015), the information leakage rate is calculated for the case of batteries with zero and infinite capacity. In addition to information-theoretic methods, techniques based on Differential Privacy (DP), a concept introduced by Dwork and Roth (2014), have received significant attention. DP provides a rigorous measure of privacy and ensures important properties such as resilience to post-processing. Sandberg et al. (2015) use DP to protect privacy when using smart meter data for state estimation in distribution networks. In the work by Zhao et al. (2014), a battery is used to ensure DP of the load profile of a single consumer. For the case of charging a group of electric vehicles, Han et al. (2014) propose a differentially private algorithm that uses distributed optimization.

In this work, we follow the line of research of Han et al. (2014) and use distributed optimization and DP to optimally operate local storage devices of a group of consumers while protecting their privacy. In contrast to the work of Han et al. (2014), a composite privacy sensitive objective is considered in the context of MPC with trustworthy and untrustworthy mediators exchanging information with the consumers. The objective consists of two parts: the minimization of energy cost of each consumer and the smoothing of the load profile, which is desirable to the grid as it results in lower operation costs. For solving the resulting optimization problems, an algorithm by Duchi

and Singer (2009) designed for problems with a composite objective is proposed. This algorithm allows handling the separable part of the objective locally, while handling the coupling part through periodic information exchanges. Consumer privacy loss during the information exchanges is mitigated by adding noise to the data in such a way that a desired level of DP is ensured. Naturally, increasing the level of privacy degrades the performance of the controller.

This paper is structured as follows: In Section 2, the problem of controlling local energy storage is described. In Section 3 an MPC that protects the load profile of the consumers using distributed optimization and DP is proposed. In Section 4, experimental results that characterize the performance of the controller under different conditions as well as the trade-off between protecting privacy and performance are shown. Lastly, in Section 5 the key contributions of the work are summarized.

2. PROBLEM DEFINITION

This work considers the control of local energy storage of a group of consumers with the goal of minimizing energy cost while providing load smoothing to the grid. Smooth load profiles are desirable to the grid operator because they reduce the use of primary and secondary control.

In the following, the problem setup is described and the optimal control problem is formulated. We consider a discrete-time system composed of N consumers. Each consumer $i \in \{1, \dots, N\}$ has a local storage device, *e.g.*, battery, and a renewable energy source, *e.g.*, solar. The consumers are connected to the distribution grid and can buy energy from the grid as well as sell excess energy to the grid. We assume the group has market access such that it can benefit from varying energy prices.
The battery charging power of consumer i in time interval t is denoted by $q_i(t)$, and the battery state of charge (SOC) at the end of interval t is denoted by $e_i(t)$. The battery dynamics are given by

$$q_i(t) = \frac{1}{\Delta t}\left(e_i(t) - e_i(t-1)\right), \tag{1}$$

where Δt represents the duration of one time interval. Conversion losses are neglected. The charging power and SOC are limited as

$$q_i^{\min} \le q_i(t) \le q_i^{\max} \tag{2a}$$
$$e_i^{\min} \le e_i(t) \le e_i^{\max}. \tag{2b}$$

It is assumed that each consumer has to pay for the energy consumed and receives money for the energy produced. The price for all consumers is equal. The price $c^b(t)$ for buying energy is assumed to be higher than the price $c^s(t)$ for selling energy. The net consumption of consumer i is denoted by $p_i(t) := l_i(t) - r_i(t) + q_i(t)$, where $l_i(t)$ is the load consumption and $r_i(t)$ is the renewable generation in time interval t.

The cost or revenue associated with the net consumption of each consumer i is given by the following piecewise linear function:

$$g(p_i(t), t) := \begin{cases} c^b(t)p_i(t) & \text{if } p_i(t) \ge 0 \\ c^s(t)p_i(t) & \text{if } p_i(t) < 0. \end{cases} \tag{3}$$

The vector containing the battery state and input of consumer i at time t is denoted by $x_i(t) := (q_i(t), e_i(t))$. The vector containing the battery state and input of all consumers at time t is denoted by $x(t) := (x_1(t), \dots, x_N(t))$.

In addition to the individual cost, the smoothness of the aggregate load profile of the group of consumers is assessed by the following measure:

$$f(x(t), x(t-1)) := \left(p(t) - p(t-1)\right)^2, \tag{4}$$

where $p(t) := \sum_{i=1}^{N} p_i(t)$.

The goal is to find the optimal policy π^* from the set of all causal and implementable policies $\mathcal{P}$ for operating the local storage device of each consumer. This policy minimizes energy cost and provides demand smoothing services to the grid. Using the definitions, this problem can be formulated as follows:

$$\pi^* = \underset{\pi \in \mathcal{P}}{\operatorname{argmin}} \lim_{T \to \infty} \mathsf{E}\left[\frac{1}{T}\sum_{t=0}^{T}\sum_{i=1}^{N} g(p_i(t), t) + \gamma \sum_{t=0}^{T} f(x(t), x(t-1))\right], \tag{5}$$

where $\gamma \ge 0$ is a parameter that controls the level of aggregate load smoothing, and $\mathsf{E}[\cdot]$ denotes expectation, since in the general case consumer load, prices and renewable generation can be stochastic. The constraints (1) and (2) are captured in $\mathcal{P}$.

The output of the control policy π are the charging powers of all batteries at every time step $t > 0$, *i.e.*,

$$q(t) = \pi(x(0), \dots, x(t-1), l(0), \dots, l(t), r(0), \dots, r(t)),$$

where $l(t)$ is the vector containing the load, and $r(t)$ is the vector containing the renewable generation for all consumers.

3. DISTRIBUTED DIFFERENTIALLY PRIVATE MPC

3.1 Model Predictive Control

In order to approximately solve the energy storage control problem in (5), a model predictive control approach is applied. It consists of solving an optimization subproblem at each time t to obtain a control sequence for a finite prediction horizon $\{t, \dots, t+T\}$. Then, only the current action is applied to the system and the process is repeated in a receding horizon fashion.

At time t, the MPC subproblem for the energy storage problem in the non-stochastic case is

$$\underset{x}{\text{minimize}} \quad h(x) \tag{6a}$$
$$\text{subject to} \quad A_i x_i = b_i, \quad i = 1, \dots, N, \tag{6b}$$
$$x_i^{\min} \le x_i \le x_i^{\max}, \quad i = 1, \dots, N, \tag{6c}$$

where x is the vector containing the battery power and SOC for all consumers and times in the prediction horizon, and $x_i := (x_i(t), \dots, x_i(t+T))$ for each consumer i. The objective function h is given by the sum of a smoothing component $\bar{f}$, weighted by a parameter γ, and energy cost components $\bar{g}$ for each consumer:

$$h(x) := \sum\nolimits_{i=1}^{N} \bar{g}(x_i) + \bar{f}(x) \tag{7}$$

where $\bar{g}$ and $\bar{f}$ are defined as

$$\bar{g}(x_i) := \sum_{\tau=t}^{t+T} g(p_i(\tau), \tau), \; \bar{f}(x) := \gamma \sum_{\tau=t}^{t+T} f(x(\tau), x(\tau-1)).$$

Constraint (6b) captures the battery dynamics as defined in (1), and constraint (6c) captures the battery limits as defined in (2).

3.2 Distributed Solution of the MPC Problem

To keep the load profile of each consumer local, the concept of distributed optimization is applied. In the following, two distributed algorithms for solving the MPC subproblem (6) are described. Specifically, the Distributed Projected Gradient algorithm (DProjG) and the Distributed Proximal Gradient algorithm (DProxG) are introduced. The former has been used by Han et al. (2014) for a problem related to electric vehicle charging that has a similar structure as that of (6), but does not consider a composite objective. The latter is applied in this paper for solving subproblem (6) since it can exploit the composite structure of the objective function (7). Handling the coupling part of the objective requires communicating net consumption estimates to a mediator. The proposed optimization algorithm for addressing this problem is related to the Forward-Backward Splitting algorithm proposed by Duchi and Singer (2009).

DProjG consists of a gradient step with respect to the complete objective, followed by distributed projections onto the feasible set of the consumers, namely,

$\mathcal{C}_i := \left\{ x_i \,\middle|\, A_i x_i = b_i,\, x_i^{\min} \le x_i \le x_i^{\max} \right\},\, i \in \{1, \dots, N\}.$

This algorithm is formally stated below:

Distributed Projected Gradient Algorithm:

$$x_i^{(k+1)} = \Pi_{\mathcal{C}_i}\left(x_i^{(k)} - \alpha_k \xi^{(k)} \right) \tag{8}$$

for each i, where $\xi^{(k)}$ is in the subdifferential $\delta h(x^{(k)})$ of h at point $x^{(k)}$, and $\Pi_{\mathcal{C}_i}$ denotes the projection operator defined by $\Pi_{\mathcal{C}_i}(y) := \operatorname{argmin}_{z \in \mathcal{C}_i} \left(\frac{1}{2} \|z - y\|_2^2 \right)$, and $\|\cdot\|_2$ is the Euclidean norm. The step sizes α_k have to fulfill $\sum_k \alpha_k = \infty$, $\lim_{k\to\infty} \alpha_k = 0$.

By including the energy cost part of the objective into the local projection (which then becomes a proximal step), one can obtain DProxG as proposed by Duchi and Singer (2009) and Combettes and Wajs (2005). For a nonlinear cost function such as the one defined in (3), using DProxG can improve the performance with respect to DProjG, as it handles the cost function inside the proximal operator rather than by a (less effective) gradient step. Formally, DProxG is given as follows:

Distributed Proximal Gradient Algorithm:

$$\bar{x}_i^{(k+1)} = \quad \operatorname{prox}_{\alpha_k \bar{g}}\left(x_i^{(k)} - \alpha_k \nabla_{x_i} \bar{f}\left(x^{(k)} \right) \right) \tag{9a}$$

for each i, where $\operatorname{prox}_{\alpha_k \bar{g}}$ is the proximal operator as defined by Parikh and Boyd (2014):

$$\operatorname{prox}_{\alpha_k \bar{g}}(y) = \operatorname*{argmin}_{z \in \mathcal{C}_i}\left(\alpha_k \bar{g}(z) + \frac{1}{2}\|z - y\|_2^2 \right). \tag{9b}$$

The step sizes α_k have to satisfy $\sum_k \alpha_k = \infty$, $\sum_k \alpha_k^2 < \infty$.

If the cost function is linear, *e.g.*, the prices for buying and selling energy are equal, it can be shown that DProjG and DProxG are equivalent. The proof is omitted for brevity and is presented in Zellner (2016).

In DProxG, the coordination between the consumers works as follows: In each iteration of the algorithm, the individual consumers share their net consumption estimates with a mediator. The mediator uses these inputs to compute the gradient $\nabla_{x_i} \bar{f}(x^{(k)})$ of the coupling part of the objective, which it then broadcasts to all consumers as it is equal for all of them. A detailed derivation of the gradient is shown in Appendix A. Using the gradient, the individual consumers can perform a proximal step to update their local variables for the next iteration. In contrast to solving the problem in a centralized manner, which requires consumers to release all information, DProxG allows using load profiles of consumers locally to minimize energy cost, and only requires sharing net consumption estimates for achieving load smoothing, which is critical to preserve the privacy of the load profile. The privacy aspects also apply for DProjG. However, DProxG offers superior performance for the considered application problem, which will be illustrated in Section 4.

3.3 Differential Privacy

In order to achieve smoothing, DProxG requires information exchanges between the different consumers via a mediator in the form of the net consumption estimates, which can reveal information about the load profile. In order to protect privacy during theses exchanges, the concept of *Differential Privacy* proposed by Dwork and Roth (2014) is applied. Privacy is achieved by adding noise to queries executed on a database containing the privacy-sensitive data.

3.3.1. Definitions To apply the concept of DP to our problem, the load profiles of the consumers are treated as the contents of a privacy-sensitive database, while the information exchanges of DProxG are treated as queries on this database.

Definition 1. Database
A database $D \in \mathbb{R}^{N(T+1)}$ is defined as the set of privacy-sensitive load consumptions $l_i := (l_i(t), \dots, l_i(t+T))$ of each consumer $i \in \{1, \dots, N\}$ during an MPC prediction horizon: $D = \{l_i\}_{i=1}^N$.

DP provides the guarantee that adjacent databases are ϵ-indistinguishable in the sense of Definition 3 below, requiring the definition of an adjacency relation.

Definition 2. Adjacency Relation
Databases D and D' are adjacent if and only if there exists an $i \in \{1, \dots, N\}$ such that $\|l_i - l_i'\|_1 \le \delta$ and $l_j = l_j'$ for all $j \neq i$, where δ is a preselected positive scalar.

Using Definitions 1 and 2, DP can be defined as follows:

Definition 3. Differential Privacy [Dwork and Roth (2014)]
A randomized algorithm $\mathcal{M}$ is ϵ-differentially private (d.p.) if for all subsets S of possible outputs of $\mathcal{M}$, and for all (D, D') that are adjacent, it holds that

$$P[\mathcal{M}(D') \in S] \le \exp(\epsilon) P[\mathcal{M}(D) \in S],$$

where $P[\cdot]$ denotes probability.

An important property is that for algorithms involving multiple queries, the epsilons "add up" as stated by the General Composition Theorem (Dwork and Roth (2014)):

Theorem 4. General Composition Theorem
Let $T_1 : D \to T_1(D) \in \mathcal{C}_1$ be ϵ-differentially private, and for $k \ge 2$, $T_k : (D, s_1, \dots, s_{k-1}) \to T_k(D, s_1, \dots, s_{k-1}) \in \mathcal{C}_k$ be ϵ-d.p. private, for all given tuples of d.p. signals $(s_1, \dots, s_{k-1}) \in \Pi_{j=1}^{k-1} \mathcal{C}_j$. Then, for all adjacent

D, D' and all $S \subseteq \Pi_{j=1}^{k} \mathcal{C}_j$, $P((T_1, \ldots, T_k) \in S) \leq \exp(k\epsilon) P'((T_1, \ldots, T_k) \in S)$, where P' denotes the probability when D' is used.

3.3.2. Differentially Private MPC Using the definitions from the previous section, DP can be applied to obtain an ϵ-d.p. version of the MPC described in Section 3.1. This is achieved by adding noise to the information exchanges using the Laplace mechanism. We consider two different cases: One where the mediator can be trusted with the net consumption estimates, and one where the mediator cannot be trusted. In the former case, the net consumption estimates are shared with the mediator, who then creates a d.p. version of the gradient of $\bar{f}$ and broadcasts it to the consumers. In the latter case, the net consumption estimates of the individual consumers need to be d.p. signals. In contrast to the trusted mediator case, the consumers add noise to their net consumption estimates before sharing them with the mediator. Using these d.p. signals, the mediator then computes the gradient of the coupling objective. The two cases are illustrated in Figure 1, where the bold lines denote d.p. signals and the dashed lines denote unprotected privacy-sensitive signals.

Definition 5. Laplace Mechanism [Dwork and Roth (2014)]
Given any function $\varphi : \mathbb{R}^n \to \mathbb{R}^m$, the Laplace mechanism is defined as: $\mathcal{M}_m(D, \varphi, \epsilon) = \varphi(D) + (Y_1, \ldots, Y_m)$, where Y_i are i.i.d. random variables drawn from the Laplace distribution with scale Δ/ϵ, denoted as $\text{Lap}(x|\Delta/\epsilon)$, where Δ is the sensitivity of φ with respect to changes between adjacent databases.

Theorem 6. DP of the Laplace mechanism [Dwork and Roth (2014)]
The Laplace mechanism ensures ϵ-differential privacy.

Using the Laplace mechanism from Definition 5, a mechanism $\mathcal{M}_m$ that creates a differentially private version of a specific signal broadcast at iteration k of DProxG can be obtained. The mechanism takes the gradient $\nabla_{x_i} \bar{f}(x^{(k)})$ (for the trustworthy mediator case), or the net consumption estimates of the different consumers (for the untrustworthy mediator case) and creates an ϵ_k-differentially private version of these signals. The privacy level ϵ_k depends on the scale of the noise added and the sensitivity $\Delta^{(k)}$ of these signals, which will be derived in Section 3.3.3. In the untrustworthy case, the gradient is also differentially-private (d.p.) due to DP's resilience to post processing. Due to Theorem 4, the privacy levels ϵ_k "add up" for multiple queries of the protected signals. This captures the fact that we "loose" an ϵ_k level of privacy in every iteration as more and more information gets released. Therefore, for K iterations, a privacy level of $\sum_{k=1}^{K} \epsilon_k$ is achieved. It follows from this and Theorem 6 that in order to make the algorithm ϵ-differentially private when executed for K iterations, the scale $b^{(k)}$ of the noise added by the Laplace mechanism at iteration k can be set to satisfy

$$b^{(k)} = \frac{\Delta^{(k)}}{\epsilon_k} = \frac{K\Delta^{(k)}}{\epsilon} \,. \tag{10}$$

Since determining the optimal K a priori is hard, in practice K is set to a fixed number of iterations. Using the differentially private versions of the signals, we can now formulate a differentially private version of DProxG which we call DP-DProxG:
DP-DProxG

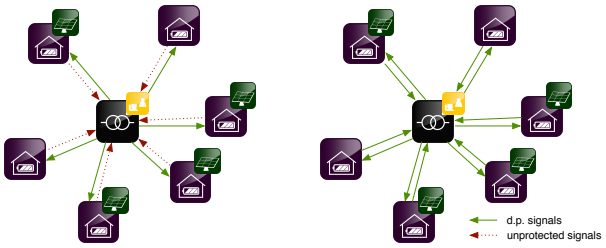

Fig. 1. Information Exchanges in the untrustworthy (left) and trustworthy mediator case (right).

- Initialize $\{x_i^{(1)}\}_{i=1}^{N}$
- Trustworthy Mediator Case
  - Consumers send $p_i^{(k)} = l_i - r_i + q_i^{(k)}$.
  - Mediator forms and broadcasts $\hat{g}^{(k)} := \nabla_{x_i} \bar{f}(x^{(k)}) + w^{(k)}$, where $w^{(k)} \sim \text{Lap}(x|b^{(k)})$.
- Untrustworthy Mediator Case
  - Consumers send $\hat{p}_i^{(k)} := p_i^{(k)} + w_i^{(k)}$, where $w_i^{(k)} \sim \text{Lap}(x|b^{(k)})$.
  - Mediator forms and broadcasts $\hat{g}^{(k)} := \nabla_{x_i} \bar{f}(\hat{p}^{(k)})$.
- Each consumer updates

$$\hat{x}_i^{(k+1)} = \text{prox}_{\alpha_k \bar{g}} \left(x_i^{(k)} - \alpha_k \hat{g}^{(k)} \right) \tag{11}$$

$$x_i^{(k+1)} = (1-\theta) x_i^{(k)} + \theta \hat{x}_i^{(k+1)} \tag{12}$$

  for each i, where $\theta \in [0, 1]$.

The acceleration step (12) proposed by Nesterov (2003) is added to improve the convergence of the algorithm in the given stochastic setting. It follows from the General Composition Theorem (Theorem 4) and Theorem 6 that DP-DProxG provides ϵ-DP.

3.3.3. Sensitivity Analysis In order to determine the correct amount of noise to be added to the query outputs, the sensitivity Δ of the queries is required. For the case of a trustworthy mediator, this is the sensitivity of the gradient with respect to changes between adjacent databases of load profiles. If the mediator is not trusted, the sensitivity of the net consumption estimates needs to be computed.

As shown in Appendix A, the gradient of $\bar{f}$ with respect to x_i can be written as

$$\nabla_{x_i} \bar{f}(x) = 2\gamma P_q^\top \tilde{D} \Big(\sum\nolimits_{j=1}^{N} (u_j + P_q x_j) \Big) \,, \tag{13}$$

where $u_j := (l_j(t) - r_j(t), \ldots, l_j(t+T) - r_j(t+T))$.

Proofs of the following Lemmas and Theorems can be found in Appendix B.

Lemma 7. Sensitivity of the iterate
Given $\hat{g}^{(1)}, \hat{g}^{(2)}, \ldots, \hat{g}^{(k-1)}$, the ℓ_1-sensitivity of the iterate $x_i^{(k)}$ with respect to changes in the database is zero, *i.e.*,

$$\| x_i^{(k)}(D') - x_i^{(k)}(D) \|_1 = 0 \,, \tag{14}$$

where D and D' are adjacent databases according to Definition 2.

Theorem 8. Sensitivity of the gradient
In the trustworthy mediator case, given $\hat{g}^{(1)}, \hat{g}^{(2)}, \ldots, \hat{g}^{(k-1)}$, the ℓ_1 sensitivity of the gradient $\nabla_{x_i} \bar{f}$ at $x^{(k)}$ is $\Delta^{(k)} = 2\gamma \| P_q^\top \tilde{D} \|_1 \delta$, where δ is defined in Definition 2.

Theorem 9. Sensitivity of the net consumption estimates
In the untrustworthy mediator case, when the query out-

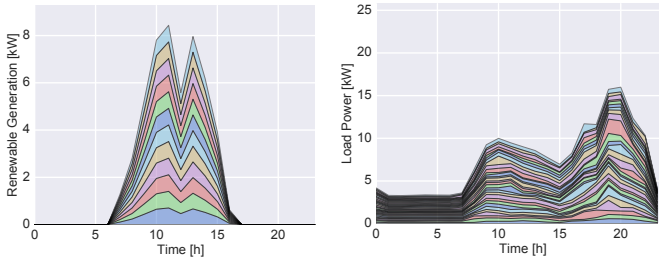

Fig. 2. Individual and aggregate load and solar profiles.

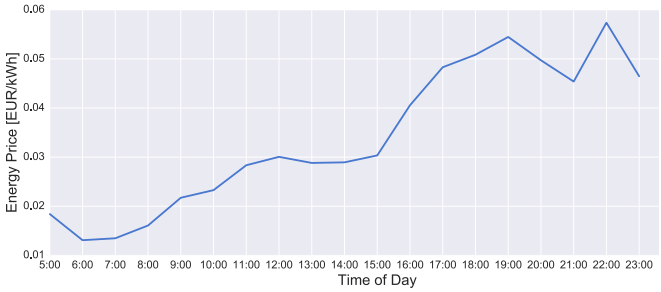

Fig. 3. Price for buying energy

puts $\hat{p}_i^{(1)}, \hat{p}_i^{(2)}, \ldots, \hat{p}_i^{(k-1)}$ are given, the ℓ_1-sensitivity of $p_i^{(k)}$ is $\Delta^{(k)} = \delta$.

4. NUMERICAL EXPERIMENTS

To evaluate the controller proposed in the previous section, a number of experiments are conducted. As a test case for the experiments, a group of 30 consumers is considered. All consumers are equipped with a battery with a capacity of 2 kWh and a maximum power of 50 kW. The consumers are all assumed to be residential. To generate realistic load profiles, the load profile generator `ALPG` described by Hoogsteen et al. (2016) is used with different consumer types (singles, families, retired, etc). Figure 2 shows the generated load profiles of the individual consumers and the aggregate load profile for the simulation horizon of one day. A total of 12 consumers are equipped with renewable generation in the form of solar. To model a realistic solar profile, the `PVWatts` solar profile generator [1] is used with a DC system size of 1 kW. The resulting solar generation profiles are shown in Figure 2. For the prices, values from the European Energy Exchange for the year 2010 are used. The energy price realization during the simulation period is shown in Figure 3.

The proposed MPC is implemented in Python using different tools and frameworks. The modeling of the group of consumers is done with the software package `PFNET` [2], which was extended by the package `MP-PFNET` [3] in this work to support multiple time intervals. For evaluating the proximal and projection operators, the `CVXPY` modeling framework of Diamond and Boyd (2016) is used together with the solver `ECOS` by Domahidi et al. (2013).

[1] http://pvwatts.nrel.gov
[2] http://ttinoco.github.io/PFNET
[3] https://github.com/martinzellner/mp-pfnet

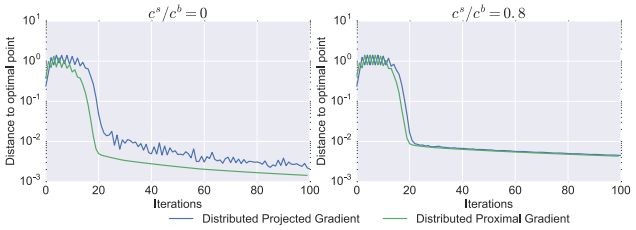

Fig. 4. Algorithm performance with different energy cost functions.

4.1 Distributed Optimization Algorithms

To compare the performance of the two algorithms described in Section 3.2, they are applied to problem (6). Both algorithms (without the acceleration step) are run for 100 iterations and c^s/c^b ratios of 0 (energy provision not compensated) and 0.8. The load smoothing parameter γ is set to 0.1. The step sizes α_k are set to $1/k$. The performance of the algorithms is shown in Figure 4. From the plots, it can be observed that the performance improvement that can be achieved by using DProxG is dependent on the properties of the separable part of the objective, which in our case is $\sum_{i=1}^{N} \bar{g}(x_i)$. If the separable objective is nonlinear, *e.g.*, piecewise linear, DProxG outperforms DProjG. However, when the separable objective is close to linear, the benefits of DProxG are only marginal. This is consistent with the fact that both algorithms (without the acceleration step) are equivalent when the separable part of the objective is linear.

4.2 Differential Privacy

To evaluate the effect of different privacy levels, several experiments are performed. Since DP adds randomness, each experiment is executed 10 times, averaging the results over multiple realizations. The load smoothing parameter γ is set to 100, the step sizes α_k are set to $1/k$ and the acceleration parameter θ is set arbitrarily to 0.5. More information on how to tune θ can be found in Duchi and Singer (2009). Six different values for ϵ are considered, ranging from $\log(3)$ (high privacy) to $\log(10^9)$ (very low privacy) for a number of iterations K ranging from 0 to 30. Figure 5 shows the expected sub-optimality achieved for different privacy levels as a function of the total number of iterations K for the cases of trustworthy and untrustworthy mediators. The optimal objective value without any consideration of privacy is indicated by $h(x^*)$. Not trusting the mediator leads to poorer performance, i.e., a more suboptimal solution. This is consistent with the fact that the variance of the gradient components for the trustworthy case can be shown to be smaller compared to the untrustworthy case. Also, it is apparent that there exists a "sweet-spot" for the number of total iterations K between iteration 1 and 15, depending on the privacy level. When running the algorithm for more iterations, the variance of the noise needs to be higher to compensate for the increased information leakage due to more information exchanges. Similar to what was found in Han et al. (2014), there exists a trade-off between the achieved privacy and the algorithm performance. For high privacy levels, only a few iterations can be carried out before the noise required to guarantee the desired privacy level keeps the algorithm

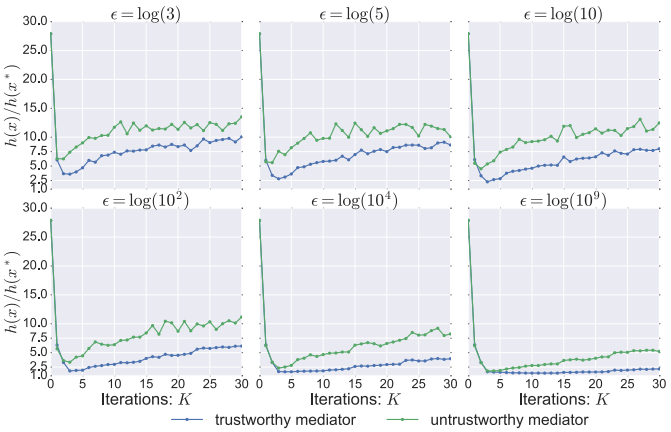

Fig. 5. Performance-privacy trade-off.

from converging to the optimal solution. For very low privacy levels, the noise is smaller, resulting in better performance.

4.3 Model Predictive Control

The proposed MPC is simulated for a time period from 5:00 to 23:00 with a prediction horizon of 8 hours and a discretization interval of 1 hour. The number of iterations K is set to 2 for the untrustworthy mediator case and 4 for the trustworthy case according to the results shown in Figure 5. The price ratios, step sizes and acceleration parameter used are the same as those used for the experiment of Section 4.2. Four cases are considered: two cases with smoothing for a medium privacy level of $\epsilon = \log(10)$ with trustworthy and untrustworthy mediators, one case with smoothing and no privacy $\epsilon = \infty$, and one case without smoothing (for which privacy is not an issue because then the problem becomes completely decoupled). The simulation results obtained are shown in Figure 6. From the plot of the net consumption, it is obvious that the MPC achieves considerable smoothing whenever the smoothing term is considered in the objective. For the two cases with medium privacy, the smoothing performance is only slightly worse compared to the case without privacy. Considering the accumulated energy cost, it is visible that over time, the energy cost is higher when smoothing and privacy are considered, which is expected.

5. CONCLUSION

In this work, a differentially private distributed MPC for energy storage management was proposed to achieve three goals: to minimize the energy cost of a group of consumers, to smoothen their aggregate load profile, and to protect their individual privacy. Smoothing the aggregate load profile is a task that requires communication between the different consumers via a mediator and can therefore reveal privacy-sensitive information. Distributed optimization was considered for solving the MPC subproblems. The distributed proximal gradient algorithm exploits the composite structure of the objective to achieve superior performance compared to the distributed projected gradient algorithm and uses the load profile of each consumer locally for minimizing the energy cost. In order to achieve aggregate load profile smoothing while maintaining privacy, a differentially private algorithm was proposed that

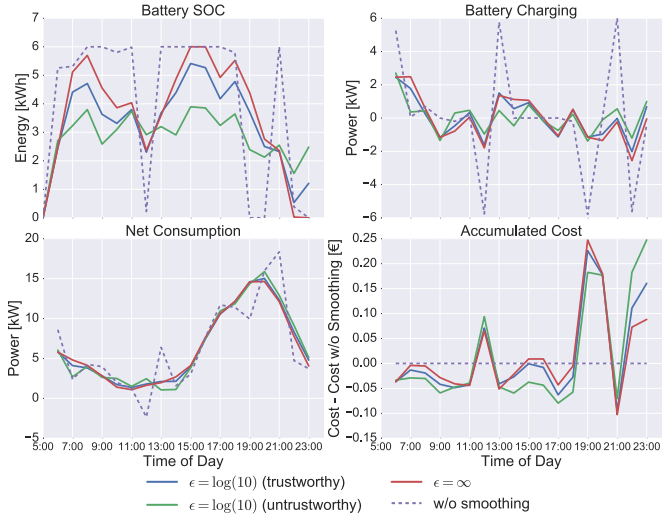

Fig. 6. MPC performance for the different privacy levels.

protects the net consumption estimates of the individual consumers, as they can potentially be used to infer details about the consumer load profiles. Using the differentially private MPC, the trade-off between privacy and performance was characterized for the cases of trustworthy and untrustworthy mediators. The results confirm the intuition that if the mediator handling the information exchanges can be trusted, the variance of the noise required to achieve a certain privacy level is lower, resulting in better performance of the algorithm.

REFERENCES

Combettes, P.L. and Wajs, V.R. (2005). Signal Recovery by Proximal Forward-Backward Splitting. *Multiscale Modeling & Simulation*, 4(4).

Diamond, S. and Boyd, S. (2016). CVXPY: A Python-Embedded Modeling Language for Convex Optimization. *Journal of Machine Learning Research*, 17(83).

Domahidi, A., Chu, E., and Boyd, S. (2013). ECOS: An SOCP solver for embedded systems. In *European Control Conference (ECC)*, 3071–3076.

Duchi, J. and Singer, Y. (2009). Efficient online and batch learning using forward backward splitting. *Journal of Machine Learning Research*, 10, 2899–2934.

Dwork, C. and Roth, A. (2014). The Algorithmic Foundations of Differential Privacy. *Foundations and Trends in Theoretical Computer Science*, 9(2013).

European Commission (2014). Benchmarking smart metering deployment in the EU-27 with a focus on electricity. *Report for the European Comission*.

Giaconi, G., Gunduz, D., and Poor, H.V. (2015). Smart meter privacy with an energy harvesting device and instantaneous power constraints. In *IEEE International Conference on Communications*.

Gómez-Vilardebó, J. and Gündüz, D. (2015). Smart meter privacy for multiple users in the presence of an alternative energy source. *IEEE Transactions on Information Forensics and Security*.

Greveler, U., Justus, B., and Loehr, D. (2012). Multimedia content identification through smart meter power usage profiles. *Computers, Privacy and Data Protection*.

Han, S., Topcu, U., and Pappas, G.J. (2014). Differentially private distributed protocol for electric vehicle charging.

In *52nd Annual Allerton Conference on Communication, Control, and Computing*.

Hoogsteen, G., Molderink, A., Hurink, J.L., and Smit, G.J.M. (2016). Generation of Flexible Domestic Load Profiles to Evaluate Demand Side Management Approaches. *2016 IEEE International Energy Conference (ENERGYCON)*.

Molina-Markham, A., Shenoy, P., Fu, K., Cecchet, E., and Irwin, D. (2010). Private memoirs of a smart meter. *Proc. of the 2nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Building - BuildSys '10*, 61–66.

Nesterov, Y.E. (2003). *Introductory Lectures on Convex Optimization: A Basic Course*. Kluwer Academic Publishers.

Parikh, N. and Boyd, S. (2014). Proximal Algorithms. *Foundations and Trends in Optimization*, 1(3), 123–231.

Rajagopalan, S.R., Sankar, L., Mohajer, S., and Poor, H.V. (2011). Smart meter privacy: A utility-privacy framework. In *2011 IEEE International Conference on Smart Grid Communications*, 190–195.

Sandberg, H., Dn, G., and Thobaben, R. (2015). Differentially private state estimation in distribution networks with smart meters. In *2015 54th IEEE Conference on Decision and Control (CDC)*, 4492–4498.

Varodayan, D. and Khisti, A. (2011). Smart meter privacy using a rechargeable battery: Minimizing the rate of information leakage. In *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing*.

Zellner, M. (2016). *Protecting Privacy in Smart Grids via Distributed Optimization and Differential Privacy*. Master thesis, ETH Zurich.

Zhao, J., Jung, T., Wang, Y., and Li, X. (2014). Achieving differential privacy of data disclosure in the smart grid. *Proceedings - IEEE INFOCOM*, (Ilm), 504–512.

Appendix A. GRADIENT DERIVATION

In the following, the gradient $\nabla_{x_i}\bar{f}$ of the smoothing objective of the MPC subproblem (7) is derived. From (4), this is derived starting with

$$f(x(\tau), x(\tau-1)) = \Big(\sum\nolimits_{i=1}^{N} \left(u_i(\tau) - u_i(\tau-1) + q_i(\tau) - q_i(\tau-1)\right)\Big)^2,$$

where $u_i(\tau) := l_i(\tau) - r_i(\tau)$ denotes the uncontrolled part of the net consumption. Let u_i and q_i be the vectors containing the uncontrolled and battery power of consumer i for all $\tau \in \{t, \dots, t+T\}$, *i.e.*, $u_i = (u_i(t), \dots, u_i(t+T))$, $q_i = (q_i(t), \dots, q_i(t+T))$. Using this, the smoothing objective $\bar{f}$ can be expressed as

$$\bar{f}(x) = \gamma \big\| \sum\nolimits_{i=1}^{N} (Du_i + Dq_i) \big\|_2^2, \tag{A.1}$$

where $D^{T\times(T+1)}$ denotes the difference matrix . Using the projection matrix $P_q^{(T+1)\times 2(T+1)}$ that extracts the battery powers from x_i, and defining $u := \sum_{i=1}^{N} u_i$, $\tilde{D} := D^\top D$, (A.1) can be rewritten as

$$\begin{aligned}\bar{f}(x) &= \gamma \left\| Du + DP_q \sum_{k=1}^{N} x_k \right\|_2^2 \\ &= \gamma u^\top \tilde{D} u + 2\gamma u^\top \tilde{D} \sum_{k=1}^{N} P_q x_k + \gamma \sum_{k\neq i} \sum_{m\neq i} x_k^\top P_q^\top \tilde{D} P_q x_m \\ &\quad + 2\gamma \sum_{k\neq i} x_k^\top P_q^\top \tilde{D} P_q x_i + \gamma x_i^\top P_q^\top \tilde{D} P_q x_i \,.\end{aligned}$$

Using the above, the gradient of $\bar{f}$ with respect to x_i can be derived to be

$$\begin{aligned}\nabla_{x_i}\bar{f}(x_i) &= 2\gamma P_q^\top \tilde{D} u + 2\gamma P_q^\top \tilde{D} P_q \sum_{k\neq i} x_k + 2\gamma P_q^\top \tilde{D} P_q x_i \\ &= 2\gamma P_q^\top \tilde{D} u + 2\gamma P_q^\top \tilde{D} \textstyle\sum_{k=1}^{N} P_q x_k \\ &= 2\gamma P_q^\top \tilde{D} \Big(\textstyle\sum_{k=1}^{N} \left(u_k + P_q x_k\right)\Big).\end{aligned}$$

Appendix B. SENSITIVITY DERIVATION

Proof of Lemma 7 For $k = 1$ it holds that

$$\|x_i^{(k)}(D') - x_i^{(k)}(D)\| = 0\,, \tag{B.1}$$

because the initial point is arbitrary and therefore independent of the database. Now consider the case $k > 1$ and assume that (B.1) holds for all $j \in \{1, \dots, k-1\}$. For notational convenience, let

$$v_i^{(k)}(D) := x_i^{(k)}(D) - \alpha_k \nabla_{x_i}\bar{f}\left(x^{(k)}(D)\right).$$

From the non-expansiveness of the proximal operator stated by Parikh and Boyd (2014) it follows that

$$\begin{aligned}&\|x_i^{(k)}(D') - x_i^{(k)}(D)\| \\ &= \left\|\mathrm{prox}_{\alpha_{k-1}\bar{g}_i}\left(v_i^{(k-1)}(D')\right) - \mathrm{prox}_{\alpha_{k-1}\bar{g}_i}\left(v_i^{(k-1)}(D)\right)\right\| \\ &\leq \|v_i^{(k-1)}(D') - v_i^{(k-1)}(D)\|\,.\end{aligned}$$

The fact that $\nabla_{x_i}\bar{f}(x^{(k-1)})$ is given implies that

$$\|v_i^{(k-1)}(D') - v_i^{(k-1)}(D)\| = \|x_i^{(k-1)}(D') - x_i^{(k-1)}(D)\| \tag{B.2}$$

holds. The inductive hypothesis then gives that the right hand side of (B.2) is zero.

Proof of Theorem 8 The sensitivity $\Delta^{(k)}$ of the gradient can be bounded using the sensitivity of the load and the sensitivity of the iterate as derived in Lemma 7:

$$\begin{aligned}&\|\nabla_{x_i}\bar{f}(x^{(k)}(D)) - \nabla_{x_i}\bar{f}(x^{(k)}(D'))\|_1 \\ =&\Big\|2\gamma P_q^\top \tilde{D} \textstyle\sum_{j=1}^{N} \left(l_j(D) - r_j + P_q x_j^{(k)}(D)\right) \\ &- 2\gamma P_q^\top \tilde{D} \textstyle\sum_{j=1}^{N} \left(l_j(D') - r_j + P_q x_j^{(k)}(D')\right)\Big\| \\ \leq& 2\gamma\Big(\|P_q^\top \tilde{D}\|_1 \|l_i(D) - l_i(D')\|_1 \\ &+ \|P_q^\top \tilde{D} P_q\|_1 \textstyle\sum_{j=1}^{N} \big\|x_j^{(k)}(D) - x_j^{(k)}(D')\big\|_1\Big) \\ =& 2\gamma \|P_q^\top \tilde{D}\|_1 \|l_i(D) - l_i(D')\| \leq 2\gamma \|P_q^\top \tilde{D}\|_1 \delta\,.\end{aligned}$$

Note that the sensitivity with respect to the gradient is equal for all k and depends on the load smoothing factor γ and the database adjacency relation.

Proof of Theorem 9 The sensitivity $\Delta_i^{(k)}$ of $p_i^{(k)}$ is dependent on the sensitivity of the iterate and the load profile:

$$\begin{aligned}\|p_i^{(k)}(D) - p_i^{(k)}(D')\|_1 \leq& \big\|x_i^{(k)}(D) - x_i^{(k)}(D')\big\|_1 \\ &+ \|l_i^{(k)}(D) - l_i^{(k)}(D')\|_1\,.\end{aligned}$$

The sensitivity of the load profile $l_i^{(k)}$ is given by the adjacency definition, *i.e.*, $\|l_i^{(k)}(D) - l_i^{(k)}(D')\| \leq \delta$. Using this and Lemma 7, the overall sensitivity of the public signal $p_i^{(k)}$ is $\|p_i^{(k)}(D) - p_i^{(k)}(D')\| \leq \delta$.