# Differentially Private Smart Metering with Battery Recharging

Michael Backes[1,2] and Sebastian Meiser[2]

[1]MPI-SWS, [2]Saarland University

**Abstract.** The energy industry has recently begun using smart meters to take fine-grained readings of energy usage. These smart meters enable flexible time-of-use billing, forecasting, and demand response, but they also raise serious user privacy concerns. We propose a novel technique for provably hiding sensitive power consumption information in the overall power consumption stream. Our technique relies on a rechargeable battery that is connected to the household's power supply. This battery is used to modify the household's power consumption by adding or subtracting noise (i.e., increasing or decreasing power consumption), in order to establish strong privacy guarantees in the sense of differential privacy. To achieve these privacy guarantees in realistic settings, we first investigate the influence of, and the interplay between, capacity and throughput bounds that batteries face in reality. We then propose an integrated method based on noise cascading that allows for recharging the battery on-the-fly so that differential privacy is retained, while adhering to capacity and throughput constraints, and while keeping the additional consumption of energy induced by our technique to a minimum.

## 1 Introduction

The energy industry has recently begun using smart meters to take fine-grained readings of energy usage, enabling flexible time-of-use billing, forecasting, and demand response [8]. The underlying incentive for energy providers is the ability to accurately match energy consumption with its generation in a fine-grained manner, thereby saving electricity and enabling dynamic tariffs with higher rates during peak consumption times. Moreover, the fine-grained metering of energy consumption enables more accurate forecasts, which is expected to lead to an overall saving of energy. Smart metering is currently being widely promoted in the United States, European Union, and Asia as part of the modernization of the electronic grid [1,2]; to this end, $4.3 billion dollars has been allocated by the U.S government for the smart grids [21], with similar programs in progress in the EU and Asia.

In addition to all these undisputed advantages, smart meters also raise serious user privacy concerns [5]: Smart meters provide highly accurate consumption data to the corresponding electricity provider. These data naturally include personal, privacy-sensitive data, e.g., information about when certain devices were active.

If metering is performed sufficiently long in small time intervals, personal information can be disaggregated from the overall consumption stream. For instance, non-intrusive appliance load monitoring techniques [15,18,19] already allow for identifying

common electronic devices such as personal computers, laser printers, or light bulbs in the overall consumption stream [7], and even to tell apart different TV programs [14].

To address these privacy concerns, privacy-aware solutions for smart metering are currently receiving increasing attention both in the research community and in ongoing standardization processes, e.g., [23]. In fact, the current absence of accepted solutions to tackle these privacy concerns caused a deadlock in the mandatory deployment of smart meters in the Netherlands [9], because of the common belief that smart metering is necessarily privacy-invasive. In this paper, we join the line of research that is working on changing this belief: we present a privacy-aware technique for smart metering that achieves strong privacy guarantees while simultaneously preserving the promises of smart metering.

### 1.1 Our Contributions

We propose a novel technique for provably hiding sensitive power consumption information in the overall power consumption stream. Our technique relies on a *rechargeable* battery that is connected to the household's power supply, and that appropriately modifies the overall consumption stream by suitably adding or subtracting noise, in order to establish strong privacy guarantees in the sense of differential privacy.

In addition to economic considerations, any solution must respect the fact that a battery adheres to hard resource constraints, such as its capacity (bounding the overall amount of energy that can be stored) and its throughput (bounding the amount of energy that can be charged/retrieved within a given time interval). Moreover, a battery will naturally get depleted over time if it constantly provides energy that is used as noise; a depleted battery will eventually put all privacy guarantees at stake. These limitations in particular render existing general-purpose approaches infeasible, because they typically require higher capacity and throughput than what a real-life battery can offer; moreover privacy-aware battery recharging is not considered in these approaches.

To achieve strong privacy guarantees in such realistic settings, we propose a novel technique for provably hiding sensitive power consumption information in the overall power consumption stream, using a rechargeable battery as a buffer and applying Laplacian noise to the consumption itself by either providing (discharging) or consuming (charging) energy by the battery. We first investigate the influence of, and the interplay between, capacity and throughput bounds of the battery to the overall approach (while still ignoring battery recharging issues), and develop a technique that achieves privacy guarantees in such resource-bounded settings. Since battery depletion is not prevented, the privacy guarantees are naturally strongest if only metering over a short time interval is considered, and they become weaker for longer time intervals.

We subsequently explore the more involved case of recharging the battery. The complication which arises here is that recharging corresponds to additional energy consumption, which is observable to the adversary by assumption. Thus simply fully recharging the battery enables an observer to determine the amount by which we are recharging the battery, causing our differential privacy guarantees to degenerate over time, similar to the case without battery recharging. To counter this effect, we propose an integrated method that allows for recharging the battery on-the-fly so that differential privacy is retained, while adhering to capacity and throughput constraints, and while

keeping the additional consumption of energy induced by our technique to a minimum. The central idea is to follow a novel cascading approach for generating differentially private noise: we consider the added noise for recharging the battery as a function that one makes differentially private by appropriately adding (a much smaller amount of) noise. To avoid that this small amount of noise is observable, we impose the assumption that this small additional energy consumption can be hidden in the overall consumption stream. Among other options, this can be achieved by continuously drawing a small, constant amount of energy that is sufficient for the recharging process, and by discarding all energy that exceeds the actual noise demand for recharging the battery in a differentially private manner.[1]

We show that meaningful differential privacy guarantees in such resource-bounded settings can be achieved, in particular using privacy-aware battery recharging. More precisely we focus on a simplistic model that captures all aspects necessary for analyzing the benefits of privacy-aware battery-recharging in smart metering. The privacy guarantee is based on hiding individual device activations in a stream of smart meter data. A more comprehensive model that additionally captures activation patterns of devices over several timeslots or the privacy of consumer behavior patterns is considered future work. Moreover, we provide a correspondence between the parameters of the battery such as capacity and throughput with the obtained privacy guarantees, and we evaluate the applicability of our techniques by means of examples.

### 1.2 Further Related Work

Privacy concerns in smart metering have been studied in several existing works in the recent past. Anderson and Fuloria [5, 6] analyze the security economics of electricity metering, in particular the conflicting interests among stakeholders. Quinn [24] and Cavoukian et al. [8] investigate legal aspects of smart meters. The privacy of billing is investigated by Danezis et al. [17,26] and Molina-Markham et al. [22]. They in particular identify the private information that current meters might leak, and propose protocol adaptations for anonymizing individual measurements. In contrast to our work, these works require a trusted third party for anonymization, as well as changes in the existing communication protocols; moreover, in contrast to differential privacy guarantees, the resulting privacy assurances and the overall consequences are less clear. Similarly, Garcia and Jacobs [13] propose to use homomorphic encryption to achieve privacy for individual measurements, but the lack of a proper perturbation of the aggregate does not make the result differentially private, and the resulting privacy interpretations are again unclear.

Prior work on differential privacy in smart metering or on the smart use of batteries to achieve privacy guarantees comprises [3, 4, 10, 12, 17, 20, 25, 27, 28].

The paper that we consider most closely related to ours is the promising contribution of Acs et al. [4]. They were first to propose the smart use of a battery in order to achieve and rigorously show differential privacy guarantees. In contrast to our work, they do not consider battery recharging, and hence only obtain meaningful privacy guarantees

[1] We stress that we wish to avoid wasting any energy in general. Our solution discards only the small amount of energy that arises for generating the noise of the battery recharging process.

if battery exhaustion is not an issue, and hence if metering is performed over a short period of time. Moreover, the magnitude of noise that they apply in their Laplacian technique depends on which appliances will be activated in the stream in the future, which only works in settings in which future activations can be accurately predicted, or at least reasonably estimated.

Papers that strive for differential privacy guarantees, yet without considering a battery (and hence in particular without the corresponding benefits gained from privacy-friendly recharging) include [3, 10, 25, 27]. Acs and Castelluccia [3] use aggregation over a large number of smart meters, add noise to the smart meter output, and encrypt the result before delivery to the energy provider. Danezis et al. [10] propose to add noise to customer bills to hide the user consumption behavior. Rastogi and Nath [25] pursue a similar approach but add noise in a distributed manner to improve performance. These approaches require the currently deployed smart meters to be replaced by new, provably trustworthy ones. Shi et al. [27] investigate untrusted aggregators of data. Their approach induces a separation between billing and the actual consumption of electricity; this allows for cheating behaviors, e.g., by applying noise with a slightly positive attitude, corresponding to seemingly increased energy consumption.

The use of a battery for privacy-preserving smart metering is discussed in [20, 28]. Varodayan and Khisti [28] consider a simplistic model where both the battery and the load of the appliances have Boolean state; differential privacy is not considered there. McLaughlin et al. [20] propose to radically smooth the consumption level to counter some common techniques for non-intrusive appliance load monitoring techniques. We consider this a promising approach; however, it currently still lacks any formalized privacy guarantees.

### 1.3 Outline of the Paper

In Section 2, we review the concept and the definition of differential privacy. Section 3 presents our model of privacy-aware smart metering in the presence of a resource-bounded battery. Section 4 investigates differential privacy guarantees in such resource-bounded settings, yet without taking battery recharging into account. Section 5 proposes our technique for privacy-aware battery recharging, and establishes corresponding differential privacy guarantees. Section 6 highlights the relationship between the individual parameters (such as the battery's resource constraints and measurement times) and the obtained privacy guarantees, and explores two concrete use cases. Section 7 discusses our guarantees and the practical feasibility of our approach. Section 8 concludes.

## 2 Preliminaries

In this paper we use a variant of *differential privacy*, as introduced in [11], as a measurement for the amount of private information leaked by a smart meter. Differential privacy was originally invented as a measurement for the amount of information leaked by answering a statistical query to a database. The notion of differential privacy that we use is *approximate differential privacy*, as introduced in [12]. In contrast to differential privacy, approximate differential privacy allows for an additional error $\delta$.

| **Definitions:** | | | |
|---|---|---|---|
| $\Delta t$ | Time interval between measurements. | $F(\mathcal{D}_i)$ | Noisy version of $f$; no resource bounds. |
| $t_i$ | Point in time defined by $t_i = t_0 + i \cdot \Delta t$. | $\underline{F}_b(\mathcal{D}_i)$ | $F$ with throughput bounds, and $\geq 0$. |
| $\mathcal{D}$ | Set of all possible devices. | $\mathbb{F}(\mathcal{D}_i)$ | $F$ with capacity/throughput bounds. |
| $\mathcal{D}_i$ | Set of all active devices in $i$'th timeslot. | | = load measured by the smart meter. |
| $\Phi$ | Stream of active devices $\mathcal{D}_1, \mathcal{D}_2, \ldots$ | $bl(i)$ | Battery level at time $t_i$. |
| $f(\mathcal{D}_i)$ | Consumption of all devices in $\mathcal{D}_i$. | $\Delta bl(i)$ | Battery charging/discharging in step $i$. |
| | | $\Delta f$ | Sensitivity of the function $f$. |

**Fig. 1.** Notation overview, not including notation for privacy-aware battery recharging (Section 5).

In the original setting of statistical databases, (approximate) differential privacy intuitively ensures that adding a single entry to the database (or deleting one from it) does not significantly change the answer given to differentially private statistical queries. Usually this is achieved by adding noise to the output. From observing the (noisy) answer to the query, a passive observer cannot determine whether a specific entry is included in the data set or not, no matter which additional information an observer might possess about other entries.

The main difference between the data base setting and the smart meter setting is that we are not interested in single readings of a smart meter, or, more formally, single applications of a function to a specific data set. Instead, we wish to apply a function to a *stream* of data. We hence extend the basic definition of (approximate) differential privacy to streams in a standard way, similar to [16].

**Definition 1** **($(\epsilon, \delta)$-Differential Privacy on Streams).** *A probabilistic algorithm $F : \mathcal{P}(\mathcal{D}) \to \mathbb{R}$ for a set $\mathcal{D}$ provides* $(\epsilon, \delta)$-differential privacy on streams *if for all (possibly countably infinite) streams $\Phi, \Phi'$ of sets $\mathcal{D}_k, \mathcal{D}'_k \subseteq \mathcal{D}$, differing in at most one element $d \in \mathcal{D}$ at one point $i$ and all sets $S$ of finite and countably infinite streams over $\mathbb{R}$,*

$$Pr[F(\Phi) \in S] \leq e^{\epsilon} \cdot Pr[F(\Phi') \in S] + \delta,$$

*where with $F(\Phi)$ we denote the stream we get when applying $F$ to each element of the stream $\Phi$ individually. The probability is taken over the randomness of $F$.*

The smart meter measures the energy load sum in every time interval, so $\mathcal{D}$ corresponds directly to the set of all devices, while $\mathcal{D}_k$ and $\mathcal{D}'_k$ correspond to the devices active in a particular time slot.

## 3 Privacy-Aware Smart Metering

In this section we present our model of privacy-aware smart metering by means of a battery. We introduce further notation used in the paper, specify the notion of a household, and define the information gained by the smart meter. We finally define two constraints that we focus on in this paper: the battery's resources *throughput* and *capacity*.

### 3.1 Notation

A household, together with its appliances, is represented by a set of possibly active devices $\mathcal{D}$. We assume this set to be finite, fixed and known to adversaries, i.e., we are able to provide strong privacy guarantees even if $\mathcal{D}$ is known to the adversary. A smart meter measures the energy load on a regular basis. We denote the time interval between two measurements of the smart meter with $\Delta t$. Thus, for our model it suffices to consider a starting time $t_0$ and times $t_i = t_0 + i \cdot \Delta t$ for all natural numbers $i \in \mathbb{N}$.

We assume for simplicity that devices can only be activated/deactivated at times $t_i$. Thus, a device can be either active (consuming energy) or inactive (not consuming energy) throughout the whole interval. We denote the devices that are active in between $t_{i-1}$ and $t_i$ as $\mathcal{D}_i \subseteq \mathcal{D}$. We write $\Phi = [\mathcal{D}_1, \mathcal{D}_2, \ldots]$ for the list/stream of active devices over time. This assumption does not weaken our guarantees: if a device is only partially active in between two time slots, its consumption will be lower (and deviate from the expected consumption), which makes it harder to link the information to the device.

The consumption function $f : \mathcal{D} \to \mathbb{R}$ assigns to each device $d \in \mathcal{D}$ the amount of energy load it consumes during one time slot (of length $\Delta t$). We assume that the consumption of devices $d$ does not vary over time, so $f(d)$ is independent of the time slot $i$ in which the device is active. Although this simplification is in contrast to tome attacks that rely on specific patterns of devices, we can model devices with varying consumption for different time slots by adding one device for each consumption level. The net consumption of all devices in a set $X \subseteq \mathcal{D}$ is expressed by leveraging the function $f$ to the powerset of $\mathcal{D}$, i.e., $f : \mathcal{P}(\mathcal{D}) \to \mathbb{R}$, with $f(X) = \sum_{d \in X} f(d)$.

This quantity is the final output the smart meter can read if no noise is added. To achieve differential privacy, we add noise to the output of $f$. Without considering the limitations of our battery at this stage, we define a probabilistic function $F : \mathcal{P}(\mathcal{D}) \to \mathbb{R}$ with $F(X) = f(X) + r$ with $r \leftarrow Lap\left(\frac{\Delta f}{\epsilon}\right)$, i.e., where $r$ is the noise we add to $f(X)$.

In our model this noise is drawn from a battery. We denote the battery level at the end of a time slot $i$ (i.e., at time $t_i$) with $bl(i)$. Thus, the change during a time slot is denoted $\Delta bl(i) = bl(i) - bl(i-1)$.

### 3.2 Modeling Throughput Restrictions

A battery's throughput denotes the amount of energy we can draw out of the battery or recharge into it during one time slot. Since we use the battery only for generating the Laplacian noise that we add to the net consumption, this means that the throughput constitutes an inherent limit for the amount of noise that can be added in one step. For simplicity reasons the battery behavior is considered linear, i.e., the throughput is independent of its current energy level. In practice this can be achieved, e.g., by using a slightly larger battery and ensuring that it is does not reach the non-linear zones.

The Laplacian noise added by $F$ can, although with small probability, reach values of arbitrary magnitude, which cannot be achieved in deployed solutions. We thus define a *throughput-respecting* function $F_b$ based on $F$ that takes into account the throughput bound $b$ of our battery. Moreover, we extend $F_b$ to its *0-bounded variant* $\underline{F}_b$ by capping the load function for the smart meter at 0; this models that we do not permit to

sell, discard or waste energy for economical reasons, which in particular excludes trivial approaches that consume enormous amounts of energy to boost the application of noise.[2]

**Definition 2 (Throughput-respecting and 0-bounded variant of $F$).** *Given a function $F$ with $F(x) = f(x) + R$ for a deterministic function $f$ and a random variable $R$. Given a bound for the throughput $b$, we define the* throughput-respecting variant $F_b$ of $F$ *as follows:*

$$F_b(x) = \begin{cases} F(x) & \text{if } |R| \leq b \\ f(x) + b & \text{if } R > b \\ f(x) - b & \text{if } -R > b. \end{cases}$$

*We define the* 0-bounded variant $\underline{F}_b$ of $F_b$ *as* $\underline{F}_b(x) = \max(0, F_b(x))$.

### 3.3 Adding Capacity Restrictions

A battery not only limits the energy output during a specific time interval $\Delta t$, but also the total amount of stored energy: its capacity. For the sake of simplicity we consider the capacity to be a fixed value $c$ that does not change over time and that also does not depend on the load drained out of the battery.[3]

The actual output we provide and that is being transmitted by the smart meter depends on the battery's capacity: If the battery is exhausted or fully charged, we naturally cannot add noise in the respective direction to the net load of our devices anymore. Building upon $\underline{F}_b$ as in Definition 2, we define an overall, *bounded mechanism* $\mathbb{F}$ that, starting with an initial battery level $bl(0)$, adds noise only as long as the capacity is not exceeded in either direction. As soon as the capacity is exceeded, $\mathbb{F}$ stops adding noise and output the net demand $f$ of our devices instead. The output of $\mathbb{F}$ constitutes the output that is transmitted to the energy provider by the smart meter.

**Definition 3 (Bounded Mechanism).** *Given a function $F$ with $F(x) = f(x) + R$ for a deterministic function $f$ and a random variable $R$, a capacity bound $c$ and a throughput bound $b$, we define the corresponding* bounded mechanism $\mathbb{F}$ *as follows, where $bl(i-1)$ is the battery level before step $i$, $R_i$ the noise added by $\underline{F}_b$ during step $i$ and $s_k = \sum_{j=1}^{k} R_j$ the sum of all noise added until step $k$:*

$$\mathbb{F}(\mathcal{D}_i) = \begin{cases} f(\mathcal{D}_i) & \text{if } \exists k \leq i.\ s_k > c - bl(0) \vee -s_k > bl(0) \\ \underline{F}_b(\mathcal{D}_i) & \text{otherwise.} \end{cases}$$

*The new battery level is $bl(i) := bl(i-1) + (\mathbb{F}(\mathcal{D}_i) - f(\mathcal{D}_i))$.*

[2] Selling electricity would be an alternative. However, an accurate treatment would additionally require a detailed cost model; moreover selling electricity after drawing it from the provider is typically not economical. We thus do not further consider this case.

[3] In practice, the amount of energy that a battery can provide usually is slightly smaller when under heavy load; we ignore this here.

As soon as the capacity is exceeded, we are facing a situation where our privacy guarantees are at stake. We can, however, give an upper bound for the probability that this happens and integrate it into the overall privacy result that we derive in the upcoming section.

# 4 Privacy-Aware Smart Metering (Without Battery Recharging)

In this section we investigate the privacy guarantees of our bounded mechanism $\mathbb{F}$, i.e., the privacy guarantees that we obtain in a resource-bounded scenario. To this end, we investigate which probabilities influence the statistical distances between $F$ and $\underline{F}_b$ (the influence of throughput constraints) as well as between $\underline{F}_b$ and $\mathbb{F}$ (the influence of capacity constraints), and develop concrete bounds for these probabilities, depending only on the throughput and capacity values of the battery as well as the magnitude of the noise (specified by $\Delta f$ and $\epsilon_1$). Finally, we combine these results in order to show that $\mathbb{F}$ is $(\epsilon_1, \delta_1)$-differentially private for an arbitrary $\epsilon_1$ and for concrete bounds for $\delta_1$, which depend on the constraints of our battery and the chosen value for $\epsilon_1$. We stress that aside from the fact that the battery can be charged when positive noise is added, battery "recharging", i.e., restoring the battery status to a secure value, is not considered in this section. Thus, we can reach situations in which the battery gets depleted (then yielding trivial privacy guarantees with $\epsilon_1$ or $\delta_1$ greater than 1). Battery recharging, and the benefits that can be drawn from it, are addressed in Section 5.

## 4.1 Differential Privacy and Statistical Distance

We start by exploring the relation between the statistical distance of two functions and differential privacy. First, recall that if our battery was unbounded, we could simply realize the function $F$ by computing $F(\mathcal{D}_i) = f(\mathcal{D}_i) + Lap\left(\frac{\Delta f}{\epsilon_1}\right)$ for sets of devices $\mathcal{D}_i \subseteq \mathcal{D}$, where the Laplacian noise is drawn from the (unbounded) battery and where $\Delta f = \max_{d \in \mathcal{D}} f(d)$ is the sensitivity of the function $f$ to which we add the noise. Adding noise in this manner corresponds to the common approach[4] to guarantee $(\epsilon, \delta)$-differential privacy with $\delta = 0$. For $\lambda = \frac{\epsilon}{\Delta f}$, the noise added by the standard technique is $Lap(\frac{1}{\lambda})$, the scaled symmetric exponential distribution with standard deviation of $\sqrt{2}\frac{1}{\lambda}$ with a variance of $2\left(\frac{\Delta f}{\epsilon}\right)^2$. The probability density function is $p(x) = \frac{\lambda}{2} \cdot e^{-|x| \cdot \lambda}$.

We now relate this case to our setting with a resource-bounded battery. To this end, we first show that differential privacy can be transferred between two functions (for increasing values of $\delta$), provided that their statistical distance is sufficiently small.

**Definition 4 (Statistical Distance).** *The* statistical distance *between two distributions X and Y over a set U is defined as*

$$d(X, Y) = \max_{S \subseteq U}(|Pr[X \in S] - Pr[Y \in S]|).$$

[4] For this work we only consider Laplacian noise. Applying other, e.g., already bounded noise distributions or other masking techniques is considered future work.

The following lemma relates differential privacy and the statistical distance.

**Lemma 1.** *Given two probabilistic functions $F$ and $G$ with the same input domain, where $F$ is $(\epsilon, \delta_1)$-differentially private. If for all possible inputs $x$ we have that the statistical distance on the output distributions of $F$ and $G$ is: $d(F(x), G(x)) \leq \delta_2$, then $G$ is $(\epsilon,\ \delta_1 + (e^{\epsilon} + 1)\delta_2)$-differentially private.*

The proofs of all lemmas and theorems are postponed to Appendix D. We note that this lemma is not tailored to our setting of streams, but applies to arbitrary types of inputs.

### 4.2 Privacy Guarantees for Throughput Restrictions

For relating the case with unbounded throughput and the throughput-bounded case, we first determine the statistical distance between $F$ and $F_b$, and subsequently exploit Lemma 1 in a suitable manner. We first observe that if one does not consider streams but only individual timeslots, $F_b$ differs from $F$ if and only if the randomness added by $F$ is of a larger magnitude than the throughput bound $b$. Consequently, the statistical distance between $F$ and $F_b$ can be bounded as follows:

**Lemma 2.** *Given an $(\epsilon, \delta)$-differentially private function $F$ with $F(x) = f(x) + R$ for a deterministic function $f$ and a random variable $R$. Then for all $x$, the statistical distance between $F$ and $F_b$ is at most $d(F(x), F_b(x)) \leq Pr\left[|R| > b\right]$.*

This lemma reasons about single elements, or more precisely, about streams of length 1. However, the probability to exceed the throughput (and thus leak information about the current input set) at one step is independent from all previous and future steps in time. For our results on differential privacy, it is thus sufficient to concentrate on the probability that the throughput is exceeded in exactly that point in time in which the streams might differ. Exceeding the throughput in any other step does not reveal additional information that helps to identify the input string from the perspective of differential privacy. We now derive a concrete bound for this probability, depending on $\epsilon$, the sensitivity $\Delta f$ of $f$, and the throughput bound $b$.

**Lemma 3.** *Given a function $F$ with $F(x) = f(x) + Lap\left(\Delta f/\epsilon\right)$ for a deterministic function $f$, and a throughput bound $b \in \mathbb{R}^+$, the probability that the Laplacian noise $Lap\left(\Delta f/\epsilon\right)$ applied to $f$ is larger than $b$ is bounded by $Pr\left[\left|Lap\left(\frac{\Delta f}{\epsilon}\right)\right| > b\right] = e^{-\frac{b\cdot\epsilon}{\Delta f}}$.*

Moreover, if $F_b$ is $(\epsilon, \delta)$-differentially private, then also its 0-bounded variant $\underline{F_b}$ is $(\epsilon, \delta)$-differentially private, because one can, without further knowledge, compute $\underline{F_b}(x)$ from $F_b(x)$ for every $x$.

### 4.3 Privacy Guarantees for Capacity Restrictions

Including bounds for the capacity requires an approach beyond considering single steps only, since the probability to exceed the capacity in step $i$ also depends on the noise added in previous steps. In fact, if one considered an arbitrarily long time interval during which random Laplacian noise is added, any finite capacity would naturally be exceeded (if there is no recharging). We exclude this case, similar to existing prior works,

by restricting us to consumption streams of a certain length $n$. We exploit how to overcome this restriction by tackling the problem of privacy-aware battery recharging during runtime in Section 5.

Similar to how we deal with throughput restrictions, we exploit the statistical distance (now on streams of length $n$) and subsequently apply Lemma 1. To combine this result with our result on throughput, we immediately bound the distance between $\underline{F}_b$ and $\mathbb{F}$: These functions differ on consumption streams of length $n$ if and only if the capacity is exceeded at least once. Recall that the battery is only used to generate noise added to the net consumption $f$. We first assume that the battery level is optimally placed at $bl(0) = \frac{c}{2}$ at the beginning of our time interval. Consequently, the probability to exceed the capacity is bounded by the probability that the sum of the noise added in all steps exceeds $\frac{c}{2}$.

**Lemma 4.** *Given an $(\epsilon_1, 0)$-differentially private function $F$ with $F(x) = f(x) + Lap\left(\frac{\Delta f}{\epsilon_1}\right)$. If the corresponding bounded mechanism $\mathbb{F}$ has capacity bound $c$ and throughput bound $b$, then for all consumption streams $\Phi$ of length $n$, the statistical distance between $\mathbb{F}$ and $\underline{F}_b$ when starting with battery level $bl(0) = \frac{c}{2}$ is at most*

$$d(\mathbb{F}(\Phi), \underline{F}_b(\Phi)) \leq Pr\left[\exists k \leq n \left| \sum_{j=1}^{k} F_b(\mathcal{D}_j) - f(\mathcal{D}_j) \right| > \frac{c}{2}\right].$$

Before we can derive an estimate for exceeding the capacity, we have to deal with the following additional complication. By definition of $\mathbb{F}$, no additional noise is added as soon as the capacity is exceeded. If this happens before the point where the two streams might differ, all privacy guarantees are lost by definition. If it happens at or after the point where the two streams might differ, the guarantees also break down because exceeding the capacity means leaking the total amount of noise generated by the battery; this information is enough for an adversary to determine which stream he has observed. Consequently, it does not suffice to give a bound for the probability to exceed the capacity in one of the steps, but we have to consider all steps at once. Further, recall that we might cap the noise not only at the throughput bound $b$, but also if the load measured by the smart meter would be negative. In this case, the expected value of the noise would be different than zero. We hence derive an estimate for the probability to exceed the capacity *at least once*, which we can successfully bound in the following lemma:

**Lemma 5.** *Given an $(\epsilon_1, 0)$-differentially private function $F$ with $F(x) = f(x) + Lap\left(\frac{\Delta f}{\epsilon_1}\right)$. For all $t > 0$, the probability that the Laplacian noise exceeds the capacity for $c \geq 2(n + t) \cdot \frac{\Delta f}{\epsilon_1}$ in at least one of the $n$ steps is bounded by*

$$Pr\left[\exists k \leq n \left| \sum_{j=1}^{k} F_b(\mathcal{D}_j) - f(\mathcal{D}_j) \right| > \frac{c}{2}\right] \leq \frac{2n}{t^2}.$$

This estimate constitutes a bound for the statistical distance between $\underline{F}_b$ and $\mathbb{F}$.

### 4.4 Obtaining an Overall Privacy Guarantee

We now combine our results on throughput and capacity constraints to obtain an overall result on differential privacy for $\mathbb{F}$. We consider streams of length $n$ and also impose the assumption that the battery level is set to $bl(0) = \frac{c}{2}$ at the beginning. The following theorem follows directly from the results we have shown in this section.

**Theorem 1.** *Given an $(\epsilon_1, 0)$-differentially private function $F$. If the corresponding bounded mechanism $\mathbb{F}$ has capacity bound $c$ and throughput bound $b$, and $bl(0)$ set to $\frac{c}{2}$, then $\mathbb{F}$ is $(\epsilon_1, \delta_1)$-differentially private on all consumption streams of length $n$ with $\delta_1 = (e^{\epsilon_1} + 1) \cdot (P_b + P_c)$ where $P_b$ is the statistical distance between $F$ and $F_b$ and $P_c$ is the statistical distance between $\underline{F}_b$ and $\mathbb{F}$.*

Obtaining concrete bounds for differential privacy can be achieved by plugging in values for $P_b$ (Lemmas 2 and 3) and $P_c$ (Lemma 4 and 5).

## 5 Privacy-Aware Smart Metering with Battery Recharging

In the last section, we have established privacy guarantees for settings in which battery recharging is not considered. In this section, we propose an integrated method that allows for recharging the battery on-the-fly, so that meaningful privacy guarantees for more comprehensive use cases can be achieved.

We start with a general explanation what makes privacy-aware battery recharging in the context of smart metering a sophisticated task. After that, we describe our solution to overcome the underlying problems, and which additional assumptions we have to impose.

**The General Problem of Privacy-aware Battery Recharging** We develop a privacy-preserving technique for recharging the battery at runtime, i.e., while using the battery for generating noise. Ideally we would simply recharge the battery level to the target level of $\frac{c}{2}$ again every $n$ steps. This would mean to increase or decrease the overall energy consumption accordingly, i.e., by the difference of the current and the target battery level. However, recall that this additional energy consumption is part of the overall energy consumption, which is measured by the smart meter, and hence observable by the adversary. Consequently, an adversary is able to determine the amount by which we are recharging the battery (provided that he knows the sets of activated devices for the recharging step). Thus, an adversary that has sufficient knowledge about the observed consumption stream can exploit this information to compute the noise added in step $i$ as follows: In every step, it computes the difference between the observed load and the expected load. Except for step $i$, where the streams differ, this is exactly equal to the added noise, and hence allows for keeping track of the battery level. As soon as the difference between this forecast of the battery level and the actual battery level is leaked, it is possible to compute the amount of noise added in step $i$ with probability 1. This information is sufficient to distinguish the streams, and hence to break differential privacy.

**Our Solution: Differentially-private Noise Generation via Cascading** We pursue the following idea for countering this effect, which constitutes a novel cascading approach for generating differentially private noise: we consider the amount of recharged energy as a function, and make this function differentially private by appropriately adding noise. We show that the additional noise is much smaller than the noise we add directly to the consumption, essentially since the new noise is only used every $n$ steps instead of every step. If desired, this process can be continued, by making this smaller noise differentially private again, and so on. In this paper, we do not formalize this further, i.e., we work with a cascade of depth one.

In a nutshell, this cascading approach transforms the problem of generating a large amount of noise that must be unobservable for an adversary into generating a much smaller, unobservable amount of noise. However, this smaller amount of noise still corresponds to energy consumption that is measured by the smart meter and thus observable by the adversary; hence if we use the battery itself to generate this additional noise, we still leak the amount of noise added by $\mathbb{F}$ in the long run: Assume we restore the battery level to a state $\frac{c}{2} + r$ for a noisy value $r$. The randomness $r$ hides all but a small part of the information about noise added to the net load in the critical time step $i$. When we recharge the battery again after $n$ additional steps, information about $r$ is leaked. After recharging the battery sufficiently often, the value of $r$ can be estimated precisely with a high probability, and differential privacy breaks down.

In order to circumvent this inherent problem, we impose the assumption that the amount of additional noise can be hidden in the overall consumption using appropriate techniques. We outline two possible techniques for achieving this in practice. First, one can assume the existence of a distinct, small secondary energy source, e.g., home-owned solar panels, that is unobservable by the adversary and solely used for the recharging process. Second, if we drop the assumption that we do not discard any energy at all, we can simply continuously draw a small, constant amount of energy from the primary source that is sufficient for the recharging process, and discard all energy that exceeds the actual battery recharging demand. For simplicity of notation in the following, we assume that this additional energy is stored in a distinct, small second battery, and then used to recharge the primary battery as described below. (In practice, both batteries would typically coincide.) We stress that the amount of energy that is wasted for the recharging process only depends on the amount of secondary noise, but not on the amount by which the (primary) battery is recharged.

### 5.1 The Battery Recharging Mechanism

We define the battery recharging mechanism $\mathbb{F}_c$ as follows: it builds on the definition of $\mathbb{F}$, but instead restores its energy every $n$ steps. We additionally reserve an amount $b_{inc} = b$ of throughput. The total amount of throughput for the battery is thus increased to $b_{total} = b + b_{inc} = 2b$, i.e., the total amount of throughput is twice as high as in the restricted setting for $n$ steps without battery recharging. When $n$ steps have passed, we compare the current battery level $bl(i)$ with the target level $\frac{c}{2}$. We do not try to hide the approximate amount of energy that we need in order to restore the battery. The precise value, however, is hidden by Laplacian noise. We postpone the precise definition of $\mathbb{F}_c$ to the Appendix.

### 5.2 Differential Privacy of the Battery Recharging Mechanism

To obtain a privacy guarantee for $\mathbb{F}_c$, we employ a conservative approach: We first show that when ignoring the leakage due to recharging, $\mathbb{F}_c$ does not leak more information than $\mathbb{F}$, for which we already gave a privacy guarantee. Then, we calculate the leakage due to recharging and combine both results. An outline of the proof, together with the Lemmas that lead to our final result, can be found in the Appendix

Finally we present the main theorem of this paper. It states that the battery-recharging mechanism $\mathbb{F}_c$ constructed in Definition 5 is indeed $(\epsilon_1 + \epsilon_2, \delta_1 + \delta_2)$-differentially private on infinite consumption streams for arbitrary values $\epsilon_1$ and $\epsilon_2$, and we give upper bounds for the values of $\delta_1$ and $\delta_2$, depending on the sensitivity of $f$ ($\Delta f$), the privacy guarantee itself $(\epsilon_1, \epsilon_2)$ and the resource limits of our primary $(b, c)$ and secondary battery ($c_{2nd}$).

**Theorem 2.** *Given an $(\epsilon_1, 0)$-differentially private function $F$ with $F(x) = f(x) + Lap\left(\frac{\Delta f}{\epsilon_1}\right)$ for a deterministic function $f$ and functions $G$ and $\mathbb{G}$ as in Definition 6. If the corresponding capacity-regulating mechanism $\mathbb{F}_c$, when using recharging noise with distribution $Lap\left(\frac{\Delta f}{\epsilon_2}\right)$ has throughput bound $b_{total} = 2 \cdot b = 2 \cdot b_{inc}$ and capacity bound $c_{total} = c + c_{2nd}$, and given a secondary battery that provides at least an amount of $c_{2nd}$ energy every $n$ steps, then for every initial battery level $bl(0)$, $\mathbb{F}_c$ is $(\epsilon_1 + \epsilon_2, \delta)$-differentially private on (possibly infinite) consumption streams with*

$$\delta = (e^{\epsilon_1} + 1) \cdot (P_b + P_c) + (e^{\epsilon_2} + 1) \cdot P_{c_{2nd}}, \text{ where}$$

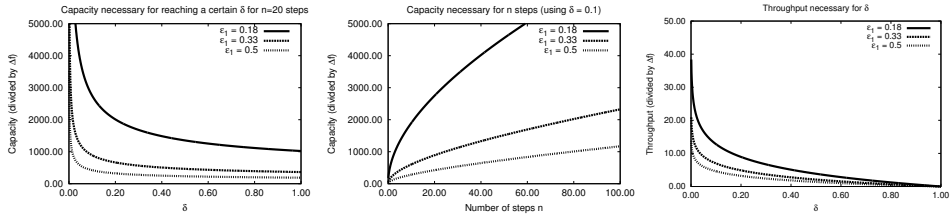
- *$P_b$ is the statistical distance between $F$ and $\underline{F}_b$.*
- *$P_c$ is the statistical distance between $\underline{F}_b$ and $\mathbb{F}$.*
- *$P_{c_{2nd}}$ is the statistical distance between $G$ and $\mathbb{G}$.*

We can formulate several instantiations of this theorem, e.g., by combining the theorem with the concrete bounds for the statistical distances proven in this paper. A corollary for Theorem 2 can be found in the Appendix.

### 5.3 Interpretation

We stress that the bounds derived in these results are not necessarily tight, but they allow for a flexible adjustment to different situations. For instance, we can freely decide the amount of noise to be added to the consumption, or to exclude certain devices from the set of devices we wish to hide, e.g., devices with a very high consumption (in this case we just compute the sensitivity $\Delta f$ over the subset $\mathcal{D}^*$ as $\Delta f = \max_{d \in \mathcal{D}^*} f(d)$). This enables us to derive strong privacy guarantees for those devices that one considers particularly privacy-critical, such as TV, Laptop or other electronic media. Concentrating on particular devices does not require any changes to the physical installation of the battery, but solely a different treatment of the required noise.

If one increases the secondary battery's capacity $c_{2nd}$, we can further reduce the amount of energy that needs to be drawn unobservably, e.g., by means of a secondary energy source: We can compute the probability that the secondary battery is exceeded over $m$ iterations and get the same privacy guarantee for a smaller share of capacity




(a) Capacity required for obtaining a privacy guarantee $\delta$ during $n = 20$ steps.

(b) Capacity required for obtaining $n$ steps for a privacy guarantee $\delta = 0.1$.

(c) Throughput required for obtaining a privacy guarantee $\delta$.

**Fig. 2.** Amount of capacity and throughput required, depending on the parameters $\epsilon$, $\delta$ and $n$.

per iteration. Using this technique and the bounds presented in this paper, the costs for restoring the battery status can (asymptotically) be reduced to $2\frac{\Delta f}{\epsilon_2}$ for each restoring process.

## 6 Evaluation and Concrete Use Cases

In this section, we further highlight the relationship between the individual parameters (such as the battery's resource constraints and measurement time) and the obtained privacy guarantees. For the sake of illustration, we moreover explore a concrete, realistic use case and analyze which privacy guarantees can be achieved under which resource assumptions.

### 6.1 Evaluation

Figure 2(a) displays the relationship between the required battery capacity and the privacy parameter $\delta$ that can be guaranteed by applying Lemma 5. Similarly, Figure 2(b) shows the relationship between this capacity and the number of steps $n$ for which the capacity has to be provided. In Figure 2(c) we depict the relationship between battery throughput and the obtained privacy guarantees. The values $\delta$ in the graphs denote the amount of privacy loss we face for the considered parameters (see Theorem 2). For the graphs, we divided the values for capacity and throughput by the sensitivity $\Delta f$ of our consumption sum function. This allows to reason about the relation of the different parameters independent from the appliances themselves. If, e.g., a TV with a consumption of 130W is to be hidden and we aim at $\varepsilon_1 = 0.33$ and $\delta = 0.1$, the battery has to have a throughput of at least about $9 \cdot 130W$ per step (time in between two readings).

### 6.2 A Concrete Use Case: Hiding TV Activation and Content

For the sake of illustration, we finally investigate the concrete use case of hiding a TV device in the overall consumption stream. We consider three different TV devices with different power consumptions, and we strive for two different security guarantees for

each of these three devices: a) hiding TV activation, and b) hiding which TV program is being watched (while potentially still disclosing that the TV was turned on).

We assume a standard American household with an average consumption of about 30 kWh per day, according to the U.S. Energy Information Administration. Within this household, we consider the following three TV devices: (1) a 42" plasma TV with 335W, (2) a 29" CRT TV with 130W, and (3) a 19" LCD TV with 36W. In the following we write $\Delta_1 f$ to denote the sensitivity we have when to hide the plasma TV, and similarly $\Delta_2 f$ for the CRT TV and $\Delta_2 f$ for the LCD TV. We will work with the following parameters: The smart meter sends the current load sum every $\Delta t = 5$ minutes, which corresponds to one of the most commonly used time intervals in smart meterings [20]. We consider an off-the-shelf rechargeable battery, and we assume that the throughput of the battery is sufficiently high so that the battery can be fully discharged within one hour. We consider an additional resource consumption of 3 kWh per day to recharge the secondary battery.

**Hiding TV Activation** In this setting, we wish to hide if the TV was activated or not. This means that we calibrate the sensitivity $\Delta f$ to the net load of the TV. Naturally, this also hides which TV program is being watched. We show results for the case where only the TV program is hidden in the following example. As a by-product, this approach also hides all appliances that use at most as much energy as the TV.

For our computation we hence obtain the following parameters: $\Delta_1 f = 335W \cdot \Delta t \approx 28Wh$, $\Delta_2 f = 130W \cdot \Delta t \approx 11$Wh, and $\Delta_3 f = 36W \cdot \Delta t = 3$Wh. The values for $\delta$ heavily depend on the selection of $\epsilon$. Note that the optimal choice of $n$ and the optimal relation of $\epsilon_1$ to $\epsilon_2$ also depend on $\epsilon$; additionally the choice of $n$ can influence the guarantees. We aim to achieve a privacy guarantee of $(0.33, 0.1)$-differential privacy in this example; hence we can choose $\epsilon_1 \geq 0$ and $\epsilon_2 \geq 0$ freely as long as $\epsilon_1 + \epsilon_2 \leq 0.33$. We can even choose $n$ freely, which denotes the number of steps between consecutive rechargings.

We exemplarily show several sample calculations (the parameters $\epsilon_1$, $\epsilon_2$, and $n$ have been determined experimentally to obtain improved results for the individual scenarios):

(1) For the 42" plasma TV with 335W, we set $\epsilon_1 \approx 0.13$ and $\epsilon_2 \approx 0.20$ and $n$ to 60 (i.e., we restore the battery status every 5 hours). We then obtain $(0.33, 0.1)$-differential privacy if one uses a battery with $11kWh$ or more.
(2) For the 29" CRT TV with 130W, we set $\epsilon_1 \approx 0.15$ and $\epsilon_2 \approx 0.18$ and $n$ to 50 (i.e., we restore the battery status every 4.17 hours). We then obtain $(0.33, 0.1)$-differential privacy if one uses a battery with $3.7kWh$ or more.
(3) For the 19" LCD TV with 36W, we set $\epsilon_1 \approx 0.21$ and $\epsilon_2 \approx 0.12$ and $n$ to 10 (i.e., we restore the battery status every 50 minutes). We then obtain $(0.33, 0.1)$-differential privacy if one uses a battery with $0.82kWh$ or more.

**Hiding which TV Program is Watched** In this setting, we wish to hide the actual TV program that is being watched (but we do not intend to hide the activation of the TV per se). The program displayed on a TV influences the energy consumption because brighter scenes require a larger consumption of energy. We thus calibrate the sensitivity

to the maximum difference between two TV programs, which is the difference between displaying a white and a dark screen. Note that for our LCD screen we assume that there is at least one non-black pixel, as otherwise the light bulb is turned off completely, resulting in a significantly larger difference in terms of consumption. (If we wanted to cover this case, we would have to use a larger value for the sensitivity.) Consequently, the sensitivity now has to only account for the maximal difference in power consumption of the individual TVs:

(1) For the 42" plasma TV with 335W, the maximal difference in power consumption based on the program is at most 130W (obtained from a power consumption fact sheet for the respective TV); thus $\Delta_1 f \approx 11Wh$. We set $\epsilon_1 \approx 0.15$ and $\epsilon_2 \approx 0.18$ and $n$ to 50 (i.e., we restore the battery status every 4.17 hours). We then obtain $(0.33, 0.1)$-differential privacy if one uses a battery with $3.7kWh$ or more.
(2) For the 29" CRT TV with 130W, the maximal difference in power consumption based on the program is at most 46W (fact sheet); thus $\Delta_2 f \approx 2.3Wh$. We set $\epsilon_1 \approx 0.19$ and $\epsilon_2 \approx 0.14$ and $n$ to 10 (i.e., we restore the battery status every 50 minutes). We then obtain $(0.33, 0.1)$-differential privacy if one uses a battery with $1.2kWh$ or more.
(3) For the 19" LCD TV with 36W, the maximal difference in power consumption based on the program is at most $2W$ (fact sheet); thus $\Delta_3 f = 0.167Wh$. We set $\epsilon_1 \approx 0.26$ and $\epsilon_2 \approx 0.07$ and $n$ to 10 (i.e., we restore the battery status every 50 minutes). We then obtain $(0.33, 0.1)$-differential privacy if one uses a battery with $0.04kWh$ or more.

## 7 Discussion

**Adversary model** In our model, in contrast to other solutions, the smart meter is not trusted. We consider a smart-meter adversary that has access to the power consumption of a household and that can make regular readings of this consumption. This adversary is, in a sense, honest-but-curious. The battery can be bought and installed by the consumer itself, without the need of any cooperation from the smart meter or the electricity company.

**Privacy guarantees** With our solution we can give (mathematically) strong privacy guarantees. However, the interpretation of these results is not trivial.

Formally we can only guarantee to hide a single activation of a single device. In practice, a realistic adversary can not keep track of all other device activations, which means that the uncertainty of an adversary covers more than one activation. However, we can only expect to hide which device (from a set of not-too-greedy devices) was activated and when. If the consumer in question follows a daily routine with almost no variation, our adversary can find out this routine. Moreover our solution does not hide the large bulk of device activations. The adversary might still be able to infer whether or not the consumer is at home (large consumption) or not (small consumption). Our solution does, however, counter many practical attacks, as the addition of random noise makes it hard to analyze the data. The parameters $(\Delta_f, \epsilon)$ should be modified whenever

the privacy policy of a consumer changes, which might be, e.g., after buying a new TV, if this TV consumes more energy.

Hiding the total consumption sum for several (say $k$) points in time, e.g., by setting $\Delta_f \approx k\cdot$ total consumption, naturally is much more expensive. We consider this out of scope for our solution.

**Usefulness** Since our solution adds the noise not to a numerical value, but to the actual consumption of the consumer, the readings of the smart meter are not influenced. In contrast to other works on smart meters, in our case there is no difference between the actual consumption and the smart meter readings (and outputs, if the smart meter is honest).

Thus, in our case the common measure of "usefulness" that is often used when analyzing the practical value of differential privacy should be defined differently. We suggest discussing the practical feasibility of our solution.

**Practical feasibility** The Laplacian noise generation can be done efficiently by applying a relatively simple function to a (normal, uniform) random or pseudorandom variable. The generation of large quantities of noise is stressful for a battery and will in practical use most likely result in a reduced life-time of the battery. We could envision a solution that uses capacitors instead of (or additionally to) a battery to improve the life-cycle of the battery.

In contrast to works that rely on modifying the smart meter, our solution does not come without cost, as installing a sufficiently large battery might be expensive. However, since our solution does not require cooperation, the decision about applying this solution can be made by each individual consumer.

## 8 Conclusions

We have proposed a novel technique for provably hiding sensitive power consumption information in the overall power consumption stream. Our technique relies on a rechargeable battery that is used to modify the household's power consumption by adding or subtracting noise (i.e., increasing or decreasing power consumption), in order to establish strong privacy guarantees in the sense of differential privacy. To achieve these privacy guarantees in realistic settings, we have investigated the influence of, and the interplay between, capacity and throughput bounds that batteries face in reality. Based on these observations, we have proposed an integrated method based on noise cascading that allows for recharging the battery on-the-fly so that differential privacy is retained, while adhering to capacity and throughput constraints, and while keeping the additional consumption of energy induced by our technique to a minimum.

## References

1. Energy independence and security act of 2007. *One Hundred Tenth Congress of the United States of America*, 2007.

2. Directive 2009/72/ec of the european parliament and of the council. *Official Journal of the European Union*, 2009.
3. G. Acs and C. Castelluccia. I have a dream! (differentially private smartmetering). In *Proc. 13th Information Hiding Conference (IH)*, volume 5806 of *LNCS*, pages 118–132. Springer, 2011.
4. G. Acs, C. Castelluccia, and W. Lecat. Protecting against physical resource monitoring. In *Proc. 10th annual ACM workshop on Privacy in the electronic society (WPES)*, pages 23–32. ACM, 2011.
5. R. Anderson and S. Fuloria. On the security economics of electricity metering. In *Workshop on the Economics of Information Security (WEIS)*, 2010.
6. R. Anderson and S. Fuloria. Who controls the off switch? In *Proc. 1st IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 96–101. IEEE Press, 2010.
7. M. Baranski and J. Voss. Detecting patterns of appliances from total load data using a dynamic programming approach. In *Proc. 4th IEEE International Conference on Data Mining (ICDM)*, pages 327–330. IEEE Press, 2004.
8. A. Cavoukian, J. Polonetsky, and C. Wolf. Smartprivacy for the smart grid: embedding privacy into the design of electricity conservation. *Identity in the Information Society*, 3:275–294, 2010.
9. C. Cuijpers. No to mandatory smart metering: does not equal privacy. [Online] available at: http://vortex.uvt.nl/TILTblog/?p=54.
10. G. Danezis, M. Kohlweiss, and A. Rial. Differentially private billing with rebates. *Proc. 13th Information Hiding Conference (IH)*, 2011.
11. C. Dwork. Differential privacy. In *Automata, Languages and Programming*, volume 4052 of *LNCS*, pages 1–12. Springer, 2006.
12. C. Dwork, K. Kenthapadi, F. McSherry, I. Mironov, and M. Naor. Our data, ourselves: Privacy via distributed noise generation. In *Proc. 25th International Cryptology Conference (EUROCRYPT)*, volume 4004 of *LNCS*, pages 486–503. Springer, 2006.
13. F. Garcia and B. Jacobs. Privacy-friendly energy-metering via homomorphic encryption. In *Security and Trust Management*, volume 6710 of *LNCS*, pages 226–238. Springer, 2011.
14. U. Greveler, B. Justus, and D. Loehr. Hintergrund und experimentelle Ergebnisse zum Thema Smart Meter und Datenschutz. Technical report, Fachhochschule Münster, 2011.
15. G. Hart. Nonintrusive appliance load monitoring. *Proc. of the IEEE*, 80(12):1870–1891, 1992.
16. T.-H. Hubert Chan, E. Shi, and D. Song. Private and continual release of statistics. In *Automata, Languages and Programming*, volume 6199 of *LNCS*, pages 405–417. Springer, 2010.
17. K. Kursawe, G. Danezis, and M. Kohlweiss. Privacy-friendly aggregation for the smart-grid. In *Privacy Enhancing Technologies*, volume 6794 of *LNCS*, pages 175–191. Springer, 2011.
18. H. Lam, G. Fung, and W. Lee. A novel method to construct taxonomy electrical appliances based on load signatures. *IEEE Transactions on Consumer Electronics*, 53(2):653–660, 2007.
19. C. Laughman, K. Lee, R. Cox, S. Shaw, S. Leeb, L. Norford, and P. Armstrong. Power signature analysis. *IEEE Power and Energy Magazine*, 1(2):56–63, 2003.
20. S. McLaughlin, P. McDaniel, and W. Aiello. Protecting consumer privacy from electric load monitoring. In *Proc. 18th ACM conference on Computer and communications security (CCS)*, pages 87–98. ACM, 2011.
21. R. Merritt. Stimulus: DoE readies $4.3 billion for smart grid. *EE Times*, 2009.
22. A. Molina-Markham, P. Shenoy, K. Fu, E. Cecchet, and D. Irwin. Private memoirs of a smart meter. In *Proc. 2nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Building (BuildSys)*, pages 61–66. ACM, 2010.

23. T. S. G. I. Panel. Cyber security strategy and requirements. technical report 7628. *National Institute of Standards and Technology*.
24. E. L. Quinn. Privacy and the new energy infrastructure. *Social Science Research Network*, (09):1995–2008, 2009.
25. V. Rastogi and S. Nath. Differentially private aggregation of distributed time-series with transformation and encryption. In *Proc. 2010 international conference on Management of Data (SIGMOD)*, pages 735–746. ACM, 2010.
26. A. Rial and G. Danezis. Privacy-preserving smart metering. In *Proc. 10th annual ACM workshop on Privacy in the electronic society (WPES)*, pages 49–60. ACM, 2011.
27. E. Shi, T.-H. H. Chan, E. Rieffel, R. Chow, and D. Song. Privacy-preserving aggregation of time-series data. In *Proc. 18th Annual Network & Distributed System Security Symposium (NDSS)*, 2011.
28. D. Varodayan and A. Khisti. Smart meter privacy using a rechargeable battery: minimizing the rate of information leakage. In *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2011.

# A Postponed details

We formalize the overall mechanism $\mathbb{F}_c$ as follows:

**Definition 5 (Battery-recharging Mechanism).** *The* battery-recharging mechanism $\mathbb{F}_c$ *on stream* $\Phi = [\mathcal{D}_1, \mathcal{D}_2, \ldots]$ *behaves as follows:*

1. *Determine the target level for recharging:* $x_{goal} = \frac{c}{2} - bl(i) + R^*$*, where* $bl(i)$ *is the current battery level and* $R^*$ *is the additional noise for battery recharging. The energy for* $R^*$ *is either thrown away (if positive) or taken from the secondary battery (if negative). If* $x_{goal} > 0$ *we are charging, if* $x_{goal} < 0$ *we are discharging the battery.*
2. *Initialize* $x_a = 0$*, a counter evaluating the progress we made for reaching our target level.*
3. *Now simulate* $\mathbb{F}$ *on* $\Phi$ *for n steps, internally using a simulated battery level* $bl'$ *starting with* $bl(i)' = \frac{c}{2}$*, but with the following additional computations:*
   - *The amount* $x_{inc}$ *we want to restore in each step is* $b_{inc}$*, but at most as much as we need to reach* $x_{goal}$ *with* $x_a + x_{inc} = x_{goal}$*. If the target level has already been reached,* $x_{inc} = 0$*.*
   - *We output* $\mathbb{F}_c(\mathcal{D}_i) = max(0, \mathbb{F}(\mathcal{D}_i) - x_{inc})$*. We cap the output at* $0$ *to avoid discarding energy.*
   - *The amount of energy by which the battery is restored is added to* $x_a$*.*[5]
   - *As soon as either the real battery level bl or the simulated battery level* $bl'$ *would exceed the capacity in either direction, we stop adding noise.*
4. *Go to 1), reinitialize the variables and repeat the process.*

# B Towards Differentially Private Smart Metering with Battery Recharging

To bound the information leakage in the recharging process, the overall proof proceeds as follows. We first define the so-called *summing* function $g$ that sums up the consumption of all devices in a given stream, without taking noise into account. We then show

[5] This can be different from $x_{inc}$, when, e.g., $\mathbb{F}(\mathcal{D}_i) - x_{inc}$ is negative.

that $f$ and its summing function have the same sensitivity. After that, we define a differentially private version $G$ of this summing function by adding Laplacian noise; finally, we derive a capacity-bounded version $\mathbb{G}$ from $G$ that leaks as much information as our recharging process and that respects the capacity constraints of the second battery.

To establish the differential privacy for $\mathbb{F}_c$, we show that the output of $\mathbb{F}_c$ can be computed from the outputs of $\mathbb{F}$ and $\mathbb{G}$. The overall proof structure to show $\mathbb{F}_c$ differentially private is depicted in Figure 3.

To bound the leakage in the recharging process, our proof proceeds as follows. We define a recharging leakage function $g$ on streams $\Phi$, that sums over the consumption of all devices in the steps in the $n$ preceeding the recharging process.[6] For the $k$'th recharging process, we thus have

$$g_k(\Phi) = \sum_{j=(k-1)n+1}^{k \cdot n} f(\mathcal{D}_j).$$

Adding noise to $g$ yields a differentially private, so-called *$n$-sumning function with noise $G$*; we draw the required noise $r_k$ from the secondary battery. This battery, analogously to our primary battery, has a finite capacity $c_{2nd}$. As for the main consumption, we define the *bounded variant* $\mathbb{G}$ of $G$ that respects the capacity bound. We show that the information leakage in the recharging process is equal to the information leaked by $\mathbb{G}$.

In contrast to $\mathbb{F}$, we do not impose any explicit throughput bound, but only require that the throughput is large enough such that the secondary battery can provide the full amount of $c_{2nd}$ over a period of $n$ steps (the time in between two recharging processes). Another difference is that in contrast to the noise added at every step in time, it suffices here to only add noise once every $n$ steps. We assume the battery to be recharged in at most $n$ steps.

**Definition 6 (Recharging Leakage Functions).** *For a deterministic function $f$ with sensitivity $\Delta f$, a privacy parameter $\epsilon_2$ and a capacity limit $c_{2nd}$, we define the corresponding $n$-*summing function with noise *$G$ and its* bounded variant *$\mathbb{G}$ on streams $\Phi$ with infinite length, or a length that is a multiple of n, as follows: We have $G(\Phi) = [G_1, G_2, G_3, \ldots]$ and $\mathbb{G}(\Phi) = [\mathbb{G}_1, \mathbb{G}_2, \mathbb{G}_3, \ldots]$, respectively, where*

$$G_k = \left( \sum_{j=(k-1)\cdot n+1}^{k \cdot n} f(\mathcal{D}_j) \right) + r_k \text{ with } r_k \leftarrow Lap\left(\frac{\Delta f}{\epsilon_2}\right),$$

$$\mathbb{G}_k = \left( \sum_{j=(k-1)\cdot n+1}^{k \cdot n} f(\mathcal{D}_j) \right) + \begin{cases} c_{2nd} & \text{if } r_k > c_{2nd} \\ -c_{2nd} & \text{if } r_k < -c_{2nd} \\ r_k & \text{otherwise.} \end{cases}$$

Since the noise is drawn according to the Laplacian distribution and of the necessary magnitude, $G$ is $(\epsilon_2, 0)$-differentially private. For $\mathbb{G}$ we use the statistical distance again.

[6] We have $\Delta g = \Delta f$ since two neighboring streams may differ in only one point in time by only one device; thus only one summand may differ by at most $\max_{d \in \mathcal{D}}(f(d)) = \Delta f$

Similar to $F_b$, the statistical distance between $G$ and $\mathbb{G}$ can be estimated with the probability that the Laplacian noise exceeds $c_{2nd}$. We do not repeat this result (Lemma 2) but instead directly give a concrete bound for the statistical distance:

**Lemma 6.** *For all streams $\Phi$ of length n, the statistical distance between $G$ and $\mathbb{G}$ as defined in Definition 6 is bounded by*

$$d(G(\Phi), \mathbb{G}(\Phi)) = e^{-\frac{c_{2nd} \cdot \epsilon_2}{\Delta f}}.$$

Note that by using Lemma 1 we can directly prove differential privacy for $\mathbb{G}$. Since one can easily compute $\mathbb{F}_c$ from $\mathbb{F}$ and $\mathbb{G}$, we can combine the individual results for $\mathbb{F}$ and $\mathbb{G}$ to achieve an overall result for $\mathbb{F}_c$:

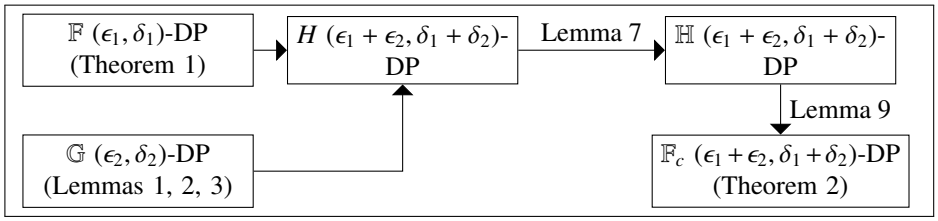



**Fig. 3.** Simplified overview of our proof. Here $(\epsilon, \delta)$-DP stands for $(\epsilon, \delta)$-differentially private.

### B.1 Obtaining Differential Privacy for the Overall Process

For arguing that $\mathbb{F}_c$ is differentially private, we have investigated the information leaked by both the measured consumption $\mathbb{F}$ (without considering battery recharging) and by the battery recharging mechanism $\mathbb{G}$ separately in Section 5. We now combine them in the standard manner by the *combining mechanism $H$* as follows:

**Definition 7 (Combining Mechanism).** *Given a mechanism $\mathbb{F}$ as in Definition 3 and a mechanism $\mathbb{G}$ as in Definition 6, we define the corresponding* combining mechanism *$H$ as*

$$H(\Phi) = (\mathbb{F}(\Phi), \mathbb{G}(\Phi)),$$

*where the battery level for $\mathbb{F}$ is assumed to be reset to $\frac{c}{2}$ every n steps.*

Following the reasoning of the combination result in [12] (Theorem 1), we obtain that $H$ satisfies $(\epsilon_1 + \epsilon_2, \delta_1 + \delta_2)$-differential privacy.

Until now we have only stated that $H$ is useful for computing $\mathbb{F}_c$. We now show how to generate a closely related function $\mathbb{H}$ out of $H$ to which our privacy guarantees can be transferred without loss. To define $\mathbb{H}$, we subtract the sum of the corresponding $n$ outputs $\mathbb{F}(\mathcal{D}_i)$ from $\mathbb{G}$, which is easily computable out of $H$. This yields the difference between the sum of all outputs and the sum of the net consumption (plus one additional share of noise). This exactly corresponds to the amount $x_{goal}$ in Definition 5 by which we want to restore the battery level.

**Definition 8.** *Let $\mathbb{H}$ be defined as*

$$\mathbb{H}(\Phi) = (\mathbb{F}(\Phi), \mathbb{G}^*(\Phi))$$

*where in $\mathbb{G}^*$ each $\mathbb{G}^*_k$ is defined as $\mathbb{G}_k - \sum_{j=(k-1)\cdot n+1}^{k\cdot n} \mathbb{F}(\mathcal{D}_j)$ and where for $\mathbb{F}$ the battery level is assumed to be reset to $\frac{c}{2}$ every n steps.*

Since this computation can be done by every observer as well, we obtain the following lemma.

**Lemma 7.** *Given a function $H = (\mathbb{F}, \mathbb{G})$ with $\mathbb{F}$ as in Definition 3 and $\mathbb{G}$ as in Definition 6. If H satisfies $(\epsilon_1 + \epsilon_2, \delta_1 + \delta_2)$-differential privacy, then $\mathbb{H}$, as defined in Definition 8, satisfies $(\epsilon_1 + \epsilon_2, \delta_1 + \delta_2)$-differential privacy.*

We now combine the results established in this section and show that the output $\mathbb{F}_c$ can indeed be computed from the output of $\mathbb{H}$ without further knowledge. This allows us to imply an $(\epsilon_1 + \epsilon_2, \delta_1 + \delta_2)$-differential privacy guarantee for $\mathbb{F}_c$.

In order to compute $\mathbb{F}_c$ from $\mathbb{H}$, we first have to make sure that the mechanism $\mathbb{F}$, when used internally by $\mathbb{F}_c$ (see Definition 5) entails the same output distribution as $\mathbb{F}$ itself, when using the same consumption stream as input and using an initial battery level of $bl(0) = \frac{c}{2}$, as in Theorem 1.

**Lemma 8.** *When starting with an arbitrary battery level $r_0$, the part $\mathbb{F}$ of $\mathbb{F}_c$ behaves as $\mathbb{F}$ would behave when initialized with $bl(0) = \frac{c}{2}$.*

This also implies that the output distribution of the part $\mathbb{F}$ of $\mathbb{F}_c$ is independent of the initial battery level. Among other results, this allows us to generalize the privacy guarantees for $\mathbb{F}_c$ to infinite streams.

**Lemma 9.** *If $\mathbb{H}$ is $(\epsilon_1 + \epsilon_2, \delta_1 + \delta_2)$-differentially private on (potentially infinitely long) consumption streams, then $\mathbb{F}_c$ is also $(\epsilon_1 + \epsilon_2, \delta_1 + \delta_2)$-differentially private on these streams.*

## C Corollary (real numbers)

**Corollary 1.** *Given an $(\epsilon_1, 0)$-differentially private function $F$ with $F(x) = f(x) + Lap\left(\frac{\Delta f}{\epsilon_1}\right)$ for a deterministic function $f$. If the corresponding capacity regulating mechanism $\mathbb{F}_c$, when using recharging noise with distribution $Lap\left(\frac{\Delta f}{\epsilon_2}\right)$ has throughput bound $b_{total} = 2 \cdot b = 2 \cdot b_{inc}$ and capacity bound $c_{total} = c + c_{2nd}$ with $c \geq 2(n + t) \cdot \frac{\Delta f}{\epsilon}$ for an arbitrary number t and given a secondary battery that provides at least an amount of $c_{2nd}$ energy every n steps, then for every initial battery level $bl(0)$, $\mathbb{F}_c$ is $(\epsilon_1 + \epsilon_2, \delta)$-differentially private on (possibly infinite) consumption streams with*

$$\delta = (e^{\epsilon_1} + 1) \cdot \left( e^{-\frac{b \cdot \epsilon_1}{\Delta f}} + \frac{2n}{t^2} \right) + (e^{\epsilon_2} + 1) \cdot e^{-\frac{c_{2nd} \cdot \epsilon_1}{\Delta f}}$$

## D Postponed Proofs

*Proof (Proof for Lemma 1).* Given a set $S$ and a two neighboring elements $x$ and $x'$.

$$
\begin{aligned}
&Pr[G(x) \in S] \\
&\leq Pr[F(x) \in S] + |Pr[F(x) \in S] - Pr[G(x) \in S]| \\
&\leq e^{\epsilon} \cdot Pr[F(x') \in S] + \delta_1 + |Pr[F(x) \in S] - Pr[G(x) \in S]| \\
&\leq e^{\epsilon} \cdot Pr[F(x') \in S] + \delta_1 + \delta_2 \\
&\leq e^{\epsilon} \cdot (Pr[G(x') \in S] + |Pr[F(x) \in S] - Pr[G(x) \in S]|) + \delta_1 + \delta_2 \\
&\leq e^{\epsilon} \cdot Pr[G(x') \in S]) + \delta_1 + (e^{\epsilon} + 1) \cdot \delta_2
\end{aligned}
$$

The second inequality follows from the fact that $F$ is $(\epsilon, \delta)$-differentially private. The calculation shows that $G$ is $(\epsilon, \delta_1 + (e^{\epsilon} + 1)\delta_2)$-differentially private.

*Proof (Proof for Lemma 2).* We have to prove the following two statements:

i) $\forall x, S.\ Pr[F_b(x) \in S] \leq Pr[F(x) \in S] + Pr[|R| > b]$,
ii) $\forall x, S.\ Pr[F(x) \in S] \leq Pr[F_b(x) \in S] + Pr[|R| > b]$,

For this proof we use the following notation:

- $P_F$ denotes $Pr[F(x) \in S]$; $P_{F_b}$ denotes $Pr[F_b(x) \in S]$.
- $P_=$ denotes $Pr[F(x) = F_b(x)]$, analogously for $P_{\neq}$.
- $P_{A|=}$ denotes $Pr[A|F(x) = F_b(x)]$, analogously for $P_{A|\neq}$.
- We assume for the proof and for these probabilities in particular, that $F_b(x)$ is computed out of $f(x)$ and $F(x)$. Thus $F$ and $F_b$ use the same randomness.

First note that $P_{\neq}$ is exactly equal to the probability $Pr[|R| > b]$. If $R$ is larger than $b$, the noise is cut down by $F_b$, but not by $F$. This is the only possibility for $F$ and $F_b$ to differ. We first show i), given $x$ and $S$:

$$
\begin{aligned}
P_{F_b} &= P_{F_b|=} \cdot P_= + P_{F_b|\neq} \cdot P_{\neq} \\
&= P_{F|=} \cdot P_= + P_{F_b|\neq} \cdot P_{\neq} \\
&= P_{F|=} \cdot P_= + P_{F_b|\neq} \cdot P_{\neq} + P_{F|\neq} \cdot P_{\neq} - P_{F|\neq} \cdot P_{\neq} \\
&= P_F + P_{F_b|\neq} \cdot P_{\neq} - P_{F|\neq} \cdot P_{\neq} \\
&= P_F + P_{\neq} \cdot (P_{F_b|\neq} - P_{F|\neq}) \\
&\leq P_F + P_{\neq}
\end{aligned}
$$

Analogously we show ii):

$$
\begin{aligned}
P_F &= P_{F|=} \cdot P_= + P_{F|\neq} \cdot P_{\neq} \\
&= P_{F_b|=} \cdot P_= + P_{F|\neq} \cdot P_{\neq} \\
&= P_{F_b|=} \cdot P_= + P_{F|\neq} \cdot P_{\neq} + P_{F_b|\neq} \cdot P_{\neq} - P_{F_b|\neq} \cdot P_{\neq} \\
&= P_{F_b} + P_{F|\neq} \cdot P_{\neq} - P_{F_b|\neq} \cdot P_{\neq} \\
&= P_{F_b} + P_{\neq} \cdot (P_{F|\neq} - P_{F_b|\neq}) \\
&\leq P_{F_b} + P_{\neq}
\end{aligned}
$$

Thus, the statistical distance as defined by Definition 4 is at most

$$d(F(x), F_b(x)) \leq Pr\left[|R| > b\right].$$

*Proof (Proof for Lemma 3).* When regarding the noise as a random variable $Lap\left(\frac{\Delta f}{\epsilon}\right)$, which has variance $\frac{2\cdot(\Delta f)^2}{\epsilon^2}$ we can apply the following variant of the Chebyshev-inequality to directly yield this result: *Given a random variable X with variance* $\sigma^2$. *For any* $k \in \mathbb{R}^+/\{0\}$ *we have that:*

$$Pr\left[|X - E[X]| \geq k^2\right] \leq \frac{\sigma^2}{k^2},$$

*where* $E[X]$ *denotes the expected value of X.*

*Proof (Proof for Lemma 4).* The difference between $\mathbb{F}$ and $\underline{F}_b$ is solely in the fact that $\mathbb{F}$ keeps track of the battery level *bl* and reduces the noise whenever it would exceed the capacity bound. Given a consumption stream $\phi$ of length $n$ and a set $S$,

$$\begin{aligned}
&\left|Pr[\mathbb{F}(\Phi) \in S] - Pr[\underline{F}_b(\Phi) \in S]\right| \\
&\leq Pr[\exists k \in \{1, \ldots, n\}. \mathbb{F}(\mathcal{D}_k) \neq \underline{F}_b(\mathcal{D}_k)] \\
&\leq Pr\left[\exists k \in \{1, \ldots, n\}. \left|\sum_{j=1}^{k} \left(\underline{F}_b(\mathcal{D}_j) - f(\mathcal{D}_j)\right)\right| > \frac{c}{2}\right]
\end{aligned}$$

The claim of the lemma follows.

*Proof (Proof for Lemma 5).* Recall that $Lap(\frac{1}{\lambda})$ has the following probability density function:

$$Lap(s, x) = \frac{\lambda}{2} \begin{cases} exp\,(x\lambda) & if\, x < 0 \\ exp\,(-x\lambda)\; if\, x \geq 0 \end{cases}$$

Taking the absolute value leads to the exponentially distributed probability density function:

$$|Lap(s, x)| = \lambda \cdot e^{-x\lambda}$$

which for $\lambda = \frac{\epsilon_1}{\Delta f}$ has the expected value $\frac{\Delta f}{\epsilon_1}$. Thus, the expected value of the sum is

$$E\left[\sum^{n} \left|Lap\left(\frac{\Delta f}{\epsilon_1}\right)\right|\right] = n \cdot \frac{\Delta f}{\epsilon_1}.$$

The variance of $\left|Lap\left(\frac{\Delta f}{\epsilon_1}\right)\right|$ is $2 \cdot \left(\frac{\Delta f}{\epsilon_1}\right)^2$, but since the individual random variables are uncorrelated we have for the sum:

$$var\left(\sum^{n} \left|Lap\left(\frac{\Delta f}{\epsilon_1}\right)\right|\right) = 2n \cdot \left(\frac{\Delta f}{\epsilon_1}\right)^2.$$

For the following computation, we abbreviate $\sum^{n} \left| Lap\left(\frac{\Delta f}{\epsilon_1}\right)\right|$ by $X_n$ for reasons of readability.

$$
\begin{aligned}
& Pr\left[X_n \geq 2 \cdot n \cdot \frac{\Delta f}{\epsilon_1}\right] \\
\leq\ & Pr\left[\left|X_n - n \cdot \frac{\Delta f}{\epsilon_1}\right| \geq n \cdot \frac{\Delta f}{\epsilon_1}\right] \\
=\ & Pr\left[|X_n - E[X_n]| \geq n \cdot \frac{\Delta f}{\epsilon_1}\right] \\
\leq\ & \frac{var(X_n)}{\left(n \cdot \frac{\Delta f}{\epsilon_1}\right)^2} \\
\leq\ & \frac{2n \cdot \left(\frac{\Delta f}{\epsilon_1}\right)^2}{\left(n \cdot \frac{\Delta f}{\epsilon_1}\right)^2} = \frac{2}{n}
\end{aligned}
$$

*Proof (Proof for Theorem 1).* By assumption, $F$ is $(\epsilon_1, 0)$-differentially private. By Lemma 1, $\underline{F_b}$ is $(\epsilon_1, (e^{\epsilon_1} + 1) \cdot P_b)$-differentially private, where $P_b$ is the statistical distance between $F$ and $F_b$. By applying Lemma 1 again we obtain that $\mathbb{F}$ is $(\epsilon_1, (e^{\epsilon_1} + 1) \cdot (P_b + P_c))$-differentially private, where $P_c$ is the statistical distance between $\underline{F_b}$ and $\mathbb{F}$.

*Proof (Proof for Lemma 6).*

The only difference between $G$ and $\mathbb{G}$ is that $\mathbb{G}$ cuts the noise if it exceeds an amount of $\pm c_{cap}$. We can apply Lemma 2 and directly obtain that for all streams $\Phi$ of length $n$, the statistical distance is bounded by

$$
d(G(\Phi), \mathbb{G}(\Phi)) \leq Pr\left[\left|Lap\left(\frac{\Delta f}{\epsilon_2}\right)\right| > c_{cap}\right].
$$

By Lemma 3 we know that the probability for the Laplacian noise to exceed a bound $c_{cap}$ is bounded by

$$
Pr\left[\left|Lap\left(\frac{\Delta f}{\epsilon_2}\right)\right| > c_{cap}\right] \leq \frac{2 \cdot (\Delta f)^2}{\left(c_{cap}\right)^2 \cdot {\epsilon_2}^2}.
$$

This concludes the proof.

*Proof (Proof for Lemma 7).*

Let $\epsilon = \epsilon_1 + \epsilon_2$ and $\delta = \delta_1 + \delta_2$. Now assume for contradiction that: $\exists \Phi, \Phi', \exists S \subseteq \mathbb{R}^{n+1}, \exists i \in \mathbb{N}$ s.t.

$$
Pr[\mathbb{H}(\Phi) \in S\,] > e^{\epsilon} Pr[\mathbb{H}(\Phi') \in S\,] + \delta.
$$

We show that this leads to a contradiction.

$$
\begin{aligned}
& Pr[\mathbb{H}(\Phi) \in S\,] > e^{\epsilon} Pr[\mathbb{H}(\Phi') \in S\,] + \delta \\
\overset{Def}{\Leftrightarrow} & Pr\left[\left(\mathbb{F}(\Phi), \mathbb{G}(\Phi) - \sum_{j=1}^{n} \mathbb{F}(\mathcal{D}_j)\right) \in S\right] > \\
& e^{\epsilon} Pr\left[\left(\mathbb{F}(\Phi'), \mathbb{G}(\Phi') - \sum_{j=1}^{n} \mathbb{F}(\mathcal{D}'_j)\right) \in S^*\right] + \delta
\end{aligned}
$$

Note that $\mathbb{F}(\Phi) \in \mathbb{R}^n$ actually denotes a n-tuple. We define a new set $S^* \subseteq \mathbb{R}^{n+1}$ as follows:

$$
S^* := \left\{ \left( (a_1, \ldots, a_n), b - \sum_{j=1}^{n} a_j \right) \,\middle|\, ((a_1, \ldots, a_n), b) \in S \right\}
$$

so that by definition we have:

$$
(\mathbb{F}(\Phi), \mathbb{G}(\Phi)) \in S^* \Leftrightarrow \left(\mathbb{F}(\Phi), \mathbb{G}(\Phi) - \sum_{j=1}^{n} \mathbb{F}_c(\mathcal{D}_j)\right) \in S. \quad (I)
$$

We combine this with the above and obtain:

$$
\begin{aligned}
& Pr[\mathbb{H}(\Phi) \in S\,] > e^{\epsilon} Pr[\mathbb{H}(\Phi') \in S\,] + \delta \\
\overset{Def}{\Leftrightarrow} & Pr\left[\left(\mathbb{F}(\Phi), \mathbb{G}(\Phi) - \sum_{j=1}^{n} \mathbb{F}(\mathcal{D}_j)\right) \in S\right] > \\
& e^{\epsilon} Pr\left[\left(\mathbb{F}(\Phi'), \mathbb{G}(\Phi') - \sum_{j=1}^{n} \mathbb{F}(\mathcal{D}'_j)\right) \in S\right] + \delta \\
\overset{(I)}{\Leftrightarrow} & Pr\left[(\mathbb{F}(\Phi), \mathbb{G}(\Phi)) \in S^*\right] > e^{\epsilon} Pr\left[(\mathbb{F}(\Phi'), \mathbb{G}(\Phi')) \in S^*\right] + \delta \\
\overset{Def}{\Leftrightarrow} & Pr[H(\Phi) \in S^*] > e^{\epsilon} Pr[H(\Phi') \in S^*] + \delta
\end{aligned}
$$

which contradicts the fact that $H$ satisfies $(\epsilon, \delta)$-differential privacy.

*Proof (Proof for Lemma 8).*

Assume for contradiction that this is not the case. The only possibility for $\mathbb{F}$ in $\mathbb{F}_c$ to behave differently than $\mathbb{F}$ is by exceeding the capacity at a different point in time. There are two possibilities:

- $\mathbb{F}$ in $\mathbb{F}_c$ exceeds the capacity earlier than $\mathbb{F}$:
  Assume that $\mathbb{F}$ in $\mathbb{F}_c$ exceeds the capacity (in either direction) in a step $i$, but internally $|s_i| < \frac{c}{2}$. We distinguish the following cases, where $r_i$ is the noise added in the current step $i$:
  - The recharging process was finished in step $i$ or before.
    This means that $bl(0) + x_a = \frac{c}{2}$. But since $bl(i) = bl(0) + x_a + s_i = \frac{c}{2} + s_i$, the capacity cannot be exceeded as long as $|s_i| < \frac{c}{2}$.

- The recharging process is ongoing in step $i$ with $|x_{inc}| = b_{inc}$ and $sign(s_i) \neq sign(x_{inc})$.
  If $sign(R) \neq sign(s_i)$, then the magnitude of the noise added has been reduced, i.e., a problem occurred in the recharging mechanism: we restore the battery in the wrong direction, i.e., we started at the opposite side of $\frac{c}{2}$.
  Then one of the following cases must hold:
  * $bl(i-1) \geq 0$ and $bl(i) < 0$: Since $|r_i| \leq b = b_{inc} = |x_{inc}|$ we cannot have exceeded the capacity in this step:
  $$bl(i) = bl(i-1) + r_i + x_{inc} \geq bl(i-1) \geq 0$$
  * $bl(i-1) \leq c$ and $bl(i) > c$: Since $|r_i| \leq b = b_{inc} = |x_{inc}|$ we cannot have exceeded the capacity in this step:
  $$bl(i) = bl(i-1) + r_i + x_{inc} \leq bl(i-1) \leq c$$
  * We restore the battery in the wrong direction, i.e., we started at the opposite side of $\frac{c}{2}$. But then $|s_i| > \frac{c}{2}$.
- The recharging process is ongoing in step $i$ with $|x_{inc}| = b_{inc}$ and $sign(s_i) = sign(x_{inc})$.
  We know that $bl(i) = bl(0) + x_a + s_i$ and that $bl(i) = bl(i-1) + r_i + x_{inc}$. We distinguish the following cases:
  * $x_{inc} < 0$. Then $bl(0) + x_a > \frac{c}{2}$. Still we assume that $bl(i) < 0$. Thus, $bl(0) + x_a + s_i < 0 \Rightarrow \frac{c}{2} + s_i < 0 \Rightarrow |s_i| > \frac{c}{2}$
  * $x_{inc} > 0$. Then $bl(0) + x_a < \frac{c}{2}$. Still we assume that $bl(i) > c$. Thus, $bl(0) + x_a + s_i > c \Rightarrow \frac{c}{2} + s_i > c \Rightarrow |s_i| > \frac{c}{2}$.

– $\mathbb{F}$ in $\mathbb{F}_c$ exceeds the capacity later than $\mathbb{F}$:
  This contradicts our definition of $\mathbb{F}_c$, since it simulates a capacity of $\frac{c}{2}$ for $\mathbb{F}$.

*Proof (Proof for Lemma 9).*
$\mathbb{F}_c$ can be computed from $\mathbb{H}$ as follows, where we call this new mechanism $\mathbb{F}_{\mathbb{H}}H$:

1. Set $k = 0$, set $x_{goal} = r^*$, where $r^*$ is drawn as the additional noise for recharging.
2. Initialize $x_a = 0$
3. Now for the next $n$ steps proceed as follows:
   – The amount $x_{inc}$ restored by $\mathbb{F}_c$ in each step is $b_{inc}$, but at most as much as is needed to reach $x_{goal}$ with $x_a + x_{inc} = x_{goal}$. If the the goal has already been reached, $x_{inc} = 0$.
   – Output $max(0, \mathbb{F}(\mathcal{D}_i) - x_{inc})$
   – The difference between the output and $\mathbb{F}(\mathcal{D}_i)$ is added to $x_a$.
4. Increase $k$ by 1 and set $x_{goal} = \mathbb{G}^*(\Phi_k) + x_{goal} - x_a$.
5. Go to 2. to reinitialize $x_a$ and repeat the process.

We show that the computations of $\mathbb{F}_c$ and $\mathbb{F}_{\mathbb{H}}$ are equal, via induction over $k$.

For $k = 0$, both $\mathbb{F}_c$ and $\mathbb{F}_{\mathbb{H}}$ start by setting $x_{goal}$. In $\mathbb{F}_c$ we have $x_{goal} = \frac{c}{2} - \frac{c}{2} + r^*$, where $\mathbb{F}_{\mathbb{H}}$ sets $x_{goal} = r^*$. In both computations $r^*$ is a random number, drawn from the same distribution. $x_a$ is set to 0.

Now $n$ steps follow, where in $\mathbb{F}_c$ the underlying mechanism $\mathbb{F}$ is simulated, while for $\mathbb{F}_{\mathbb{H}}$ the real mechanism $\mathbb{F}$ is used. By Lemma 8 we know that $\mathbb{F}$ behaves the same, independent whether it is simulated by $\mathbb{F}_c$ or used directly by $\mathbb{F}_{\mathbb{H}}$.

Given all computations were equal for all values $k$ we have seen so far. For $k+1$ we again compute $x_{goal}$ in both $\mathbb{F}_c$ and $\mathbb{F}_{\mathbb{H}}$. In $\mathbb{F}_c$ we have

$$
\begin{aligned}
x_{goal} &:= \frac{c}{2} - bl((k+1)\cdot n) + r^* \\
&= \frac{c}{2} - bl(0) - \sum_{j=1}^{(k+1)\cdot n} \Delta bl(j) + r^* \\
&= \frac{c}{2} - bl(k\cdot n) - \sum_{j=k\cdot n+1}^{((k+1)\cdot n)} \Delta bl(j) + r^* \\
&= \frac{c}{2} - bl(k\cdot n) - \sum_{j=k\cdot n+1}^{((k+1)\cdot n)} \Big(f(\mathcal{D}_j) - \mathbb{F}(\mathcal{D}_j)\Big) + r^* \\
&= \frac{c}{2} - bl(k\cdot n) - \sum_{j=k\cdot n+1}^{((k+1)\cdot n)} f(\mathcal{D}_j) - \sum_{j=k\cdot n+1}^{((k+1)\cdot n)} \mathbb{F}(\mathcal{D}_j) + r^* \\
&= \frac{c}{2} - bl(k\cdot n) + \mathbb{G}^*(\Phi_{k+1}) \\
&\overset{IH}{=} x'_{goal} - x_a + \mathbb{G}^*(\Phi_{k+1}),
\end{aligned}
$$

which exactly corresponds to how $x_{goal}$ is computed by $\mathbb{F}_{\mathbb{H}}$. Here $x'_{goal}$ is the previous value of $x_{goal}$. The last equation holds by induction hypothesis: Since the computations have been equal up to iteration $k$, the battery level $bl(k\cdot n)$ is $\frac{c}{2} - x'_{goal} + x_a$.

For the next $n$ steps again for $\mathbb{F}_c$ the underlying mechanism $\mathbb{F}$ is simulated, while for $\mathbb{F}_{\mathbb{H}}$ the real mechanism $\mathbb{F}$ is used with a reset battery level of $\frac{c}{2}$. By Lemma 8 we know that $\mathbb{F}$ behaves in the same way, if it is simulated by $\mathbb{F}_c$ or used directly by $\mathbb{F}_{\mathbb{H}}$.

Since $\mathbb{H}$ is $(\epsilon,\delta)$-differentially private and this transformation from $\mathbb{H}$ to $\mathbb{F}_c$ constitutes a deterministic function, $\mathbb{F}_c$ is $(\epsilon,\delta)$-differentially private.

*Proof (Proof for Theorem 2).* By Theorem 1 we know that $\mathbb{F}$ is $(\epsilon_1,\delta_1)$-differentially private. By Lemma 1 $\mathbb{G}$ is $(\epsilon_2,\delta_2)$-differentially private. Thus, $H$ is $(\epsilon_1+\epsilon_2,\delta_1+\delta_2)$-differentially private. By Lemma 7, $\mathbb{H}$ is $(\epsilon_1+\epsilon_2,\delta_1+\delta_2)$-differentially private. Applying Lemma 9 concludes the proof.