Privacy Odometers and Filters: Pay-as-you-Go Composition

Ryan Rogers[*]   **Aaron Roth**[†]   **Jonathan Ullman**[‡]   **Salil Vadhan**[§]

October 27, 2016

Abstract

In this paper we initiate the study of adaptive composition in differential privacy when the length of the composition, and the privacy parameters themselves can be chosen adaptively, as a function of the outcome of previously run analyses. This case is much more delicate than the setting covered by existing composition theorems, in which the algorithms themselves can be chosen adaptively, but the privacy parameters must be fixed up front. Indeed, it isn't even clear how to define differential privacy in the adaptive parameter setting. We proceed by defining two objects which cover the two main use cases of composition theorems. A *privacy filter* is a stopping time rule that allows an analyst to halt a computation before his pre-specified privacy budget is exceeded. A *privacy odometer* allows the analyst to track realized privacy loss as he goes, without needing to pre-specify a privacy budget. We show that unlike the case in which privacy parameters are fixed, in the adaptive parameter setting, these two use cases are distinct. We show that there exist privacy filters with bounds comparable (up to constants) with existing privacy composition theorems. We also give a privacy odometer that nearly matches non-adaptive private composition theorems, but is sometimes worse by a small asymptotic factor. Moreover, we show that this is inherent, and that *any* valid privacy odometer in the adaptive parameter setting must lose this factor, which shows a formal separation between the filter and odometer use-cases.

---

[*]Department of Applied Mathematics and Computational Science, University of Pennsylvania. `ryrogers@sas.upenn.edu`.

[†]Department of Computer and Information Sciences, University of Pennsylvania. `aaroth@cis.upenn.edu`. Supported in part by an NSF CAREER award, NSF grant CNS-1513694, and a grant from the Sloan Foundation.

[‡]College of Computer and Information Science, Northeastern University. `jullman@ccs.neu.edu`

[§]Center for Research on Computation & Society and John A. Paulson School of Engineering & Applied Sciences, Harvard University. `salil@seas.harvard.edu`. Work done while visiting the Department of Applied Mathematics and the Shing-Tung Yau Center at National Chiao-Tung University in Taiwan. Also supported by NSF grant CNS-1237235, a grant from the Sloan Foundation, and a Simons Investigator Award.

Contents

1 Introduction

Differential privacy [DMNS06] is a stability condition on a randomized algorithm, designed to guarantee individual-level privacy during data analysis. Informally, an algorithm is differentially private if any pair of close inputs map to similar probability distributions over outputs, where similarity is measured by two parameters ε and δ. Informally, ε measures the *amount of privacy* and δ measures the *failure probability* that the privacy loss is much worse than ε. A signature property of differential privacy is that it is preserved under *composition*—combining many differentially private subroutines into a single algorithm preserves differential privacy and the privacy parameters degrade gracefully. Composability is essential for both privacy and for algorithm design. Since differential privacy is composable, we can design a sophisticated algorithm and prove it is private without having to reason directly about its output distribution. Instead, we can rely on the differential privacy of the basic building blocks and derive a privacy bound on the whole algorithm using the composition rules.

The composition theorem for differential privacy is very strong, and holds even if the choice of which differentially private subroutine to run is adaptive—that is, the choice of the next algorithm may depend on the output of previous algorithms. This property is essential in algorithm design, but also more generally in modeling unstructured sequences of data analyses that might be run by a human data analyst, or even by many data analysts on the same data set, while only loosely coordinating with one another. Even setting aside privacy, it can be very challenging to analyze the statistical properties of general adaptive procedures for analyzing a dataset, and the fact that adaptively chosen differentially private algorithms compose has recently been used to give strong guarantees of statistical validity for adaptive data analysis [DFH+15, BNS+16].

However, all the known composition theorems for differential privacy [DMNS06, DKM+06, DRV10, KOV15, MV16] have an important and generally overlooked caveat. Although the choice of the next subroutine in the composition may be adaptive, the number of subroutines called and choice of the privacy parameters ε and δ for each subroutine must be fixed in advance. Indeed, it is not even clear how to define differential privacy if the privacy parameters are not fixed in advance. This is generally acceptable when designing a single algorithm (that has a worst-case analysis), since worst-case eventualities need to be anticipated and budgeted for in order to prove a theorem. However, it is *not* acceptable when modeling the unstructured adaptivity of a data analyst, who may not know ahead of time (before seeing the results of intermediate analyses) what he wants to do with the data. When controlling privacy loss across multiple data analysts, the problem is even worse.

As a simple stylized example, suppose that $\mathcal{A}$ is some algorithm (possibly modeling a human data analyst) for selecting *statistical queries*[1] as a function of the answers to previously selected queries. It is known that for any one statistical query q and any data set $\mathbf{x}$, releasing the perturbed answer $\hat{a} = q(\mathbf{x}) + Z$ where $Z \sim \mathrm{Lap}(1/\varepsilon)$ is a Laplace random variable, ensures $(\varepsilon, 0)$-differential privacy. Composition theorems allow us to reason about the composition of k such operations, where the queries can be chosen adaptively by $\mathcal{A}$, as in the following simple program.

Example1(x):

> For $i = 1$ to k:

---

[1]A *statistical query* is parameterized by a predicate ϕ, and asks "how many elements of the dataset satisfy ϕ?" Changing a single element of the dataset can change the answer to the statistical query by at most 1.

Let $q_i = \mathcal{A}(\hat{a}_1, \ldots, \hat{a}_{i-1})$ and let $\hat{a}_i = q_i(\mathbf{x}) + \mathrm{Lap}(1/\varepsilon)$.
Output $(\hat{a}_1, \ldots, \hat{a}_k)$.

The "basic" composition theorem [DMNS06] asserts that Example1 is $(\varepsilon k, 0)$-differentially private. The "advanced" composition theorem [DRV10] gives a more sophisticated bound and asserts that (provided that ε is sufficiently small), the algorithm satisfies $(\varepsilon \sqrt{8k \ln(1/\delta)}, \delta)$-differential privacy for any $\delta > 0$. There is even an "optimal" composition theorem [KOV15] too complicated to describe here. These analyses crucially assume that both the number of iterations k and the parameter ε are fixed up front, even though it allows for the queries q_i to be adaptively chosen.[2]

Now consider a similar example where the number of iterations is not fixed up front, but actually depends on the answers to previous queries. This is a special case of a more general setting where the privacy parameter ε_i in every round may be chosen adaptively—halting in our example is equivalent to setting $\varepsilon_i = 0$ in all future rounds.

Example2$(\mathbf{x}, \tau)$:
Let $i \leftarrow 1$, $\hat{a}_1 \leftarrow q_1(\mathbf{x}) + \mathrm{Lap}(1/\varepsilon)$.
While $\hat{a}_i \leq \tau$:
    Let $i \leftarrow i + 1$, $q_i = \mathcal{A}(\hat{a}_1, \ldots, \hat{a}_{i-1})$, and let $\hat{a}_i = q_i(\mathbf{x}) + \mathrm{Lap}(1/\varepsilon)$.
Output $(\hat{a}_1, \ldots, \hat{a}_i)$.

Example2 cannot be said to be differentially private *ex ante* for any non-trivial fixed values of ε and δ, because the computation might run for an arbitrarily long time and privacy may degrade indefinitely. What can we say about privacy after we run the algorithm? If the algorithm/data-analyst happens to stop after k rounds, can we apply the composition theorem *ex post* to conclude that it is $(\varepsilon k, 0)$- and $(\varepsilon \sqrt{8k \log(1/\delta)}, \delta)$-differentially private, as we could if the algorithm were constrained to always run for at most k rounds?

In this paper, we study the composition properties of differential privacy when *everything*—the choice of algorithms, the number of rounds, and the privacy parameters in each round—may be adaptively chosen. We show that this setting is much more delicate than the settings covered by previously known composition theorems, but that these sorts of *ex post* privacy bounds do hold with only a small (but in some cases unavoidable) loss over the standard setting. We note that the conceptual discussion of differential privacy focuses a lot on the idea of arbitrary composition and our results give more support for this conceptual interpretation.

1.1 Our Results

We give a formal framework for reasoning about the adaptive composition of differentially private algorithms when the privacy parameters themselves can be chosen adaptively. When the parameters are chosen non-adaptively, a *composition theorem* gives a high probability bound on the worst case *privacy loss* that results from the output of an algorithm. In the adaptive parameter setting, it no longer makes sense to have fixed bounds on the privacy loss. Instead, we propose two kinds of primitives capturing two natural use cases for composition theorems:

---

[2]The same analysis holds for hetereogeneous parameters $(\varepsilon_1, \ldots, \varepsilon_k)$ are used in each round as long as they are all fixed in advance. For basic composition εk is replaced with $\sum_{i=1}^{k} \varepsilon_i$ and for advanced composition $\varepsilon \sqrt{k}$ is replaced with $\sqrt{\sum_{i=1}^{k} \varepsilon_i^2}$.

1. A *privacy odometer* takes as input a global failure parameter δ_g. After every round i in the composition of differentially private algorithms, the odometer outputs a number τ_i that may depend on the *realized* privacy parameters ε_i, δ_i in the previous rounds. The privacy odometer guarantees that with probability $1 - \delta_g$, for every round i, τ_i is an upper bound on the privacy loss in round i.

2. A *privacy filter* is a way to cut off access to the dataset when the privacy loss is too large. It takes as input a global privacy "budget" $(\varepsilon_g, \delta_g)$. After every round, it either outputs CONT ("continue") or HALT depending on the privacy parameters from the previous rounds. The privacy filter guarantees that with probability $1 - \delta_g$, it will output HALT before the privacy loss exceeds ε_g. When used, it guarantees that the resulting interaction is $(\varepsilon_g, \delta_g)$-DP.

A tempting heuristic is to take the *realized* privacy parameters $\varepsilon_1, \delta_1, \ldots, \varepsilon_i, \delta_i$ and apply one of the existing composition theorems to those parameters, using that value as a privacy odometer or implementing a privacy filter by halting when getting a value that exceeds the global budget. However this heuristic *does not* necessarily give valid bounds.

We first prove that the heuristic *does* work for the basic composition theorem [DMNS06] in which the parameters ε_i and δ_i add up. We prove that summing the realized privacy parameters yields both a valid privacy odometer and filter. The idea of a privacy filter was also considered in [ES15], who show that basic composition works in the privacy filter application.

We then show that the heuristic breaks for the advanced composition theorem [DRV10]. However, we give a valid privacy filter that gives the same asymptotic bound as the advanced composition theorem, albeit with worse constants. On the other hand, we show that, in some parameter regimes, the asymptotic bounds given by our privacy filter *cannot* be achieved by a privacy odometer. This result gives a formal separation between the two models when the parameters may be chosen adaptively, which does not exist when the privacy parameters are fixed. Finally, we give a valid privacy odometer with a bound that is only slightly worse asymptotically than the bound that the advanced composition theorem would give if it were used (improperly) as a heuristic. Our bound is worse by a factor that is never larger than $\sqrt{\log\log(n)}$ (here, n is the size of the dataset) and for some parameter regimes is only a constant.

2 Privacy Preliminaries

Differential privacy is defined based on the following notion of similarity between two distributions.

Definition 2.1 (Indistinguishable)**.** Two random variables X and Y taking values from domain $\mathcal{D}$ are (ε, δ)-indistinguishable, denoted as $X \approx_{\varepsilon, \delta} Y$, if $\forall S \subseteq \mathcal{D}$,

$$\mathbb{P}[X \in S] \le e^{\varepsilon} \mathbb{P}[Y \in S] + \delta \quad \text{and} \quad \mathbb{P}[Y \in S] \le e^{\varepsilon} \mathbb{P}[X \in S] + \delta.$$

There is a slight variant of indistinguishability, called *point-wise indistinguishability*, which is nearly equivalent, but will be the more convenient notion for the generalizations we give in this paper.

Definition 2.2 (Point-wise Indistinguishable)**.** Two random variables X and Y taking values from $\mathcal{D}$ are (ε, δ)-point-wise indistinguishable if with probability at least $1 - \delta$ over either $a \sim X$ or $a \sim Y$, we have

$$\left| \log\left(\frac{\mathbb{P}[X = a]}{\mathbb{P}[Y = a]} \right) \right| \le \varepsilon.$$

Lemma 2.3 ([KS14]). *Let X and Y be two random variables taking values from $\mathcal{D}$. If X and Y are (ε, δ)-point-wise indistinguishable, then $X \approx_{\varepsilon,\delta} Y$. Also, if $X \approx_{\varepsilon,\delta} Y$ then X and Y are $\left(2\varepsilon, \frac{2\delta}{e^\varepsilon \varepsilon}\right)$-point-wise indistinguishable.*

We say two databases $\mathbf{x}, \mathbf{x}' \in \mathcal{X}^n$ are *neighboring* if they differ in at most one entry, i.e. if there exists an index $i \in [n]$ such that $\mathbf{x}_{-i} = \mathbf{x}'_{-i}$. We can now state differential privacy in terms of indistinguishability.

Definition 2.4 (Differential Privacy [DMNS06]). A randomized algorithm $\mathcal{M} : \mathcal{X}^n \to \mathcal{Y}$ with arbitrary output range $\mathcal{Y}$ is (ε, δ)-differentially private (DP) if for every pair of neighboring databases $\mathbf{x}, \mathbf{x}'$:

$$\mathcal{M}(\mathbf{x}) \approx_{\varepsilon,\delta} \mathcal{M}(\mathbf{x}').$$

We then define the privacy loss $\text{Loss}_\mathcal{M}(a; \mathbf{x}, \mathbf{x}')$ for outcome $a \in \mathcal{Y}$ and neighboring datasets $\mathbf{x}, \mathbf{x}' \in \mathcal{X}^n$ as

$$\text{Loss}_\mathcal{M}(a; \mathbf{x}, \mathbf{x}') = \log\left(\frac{\mathbb{P}[\mathcal{M}(\mathbf{x}) = a]}{\mathbb{P}[\mathcal{M}(\mathbf{x}') = a]}\right).$$

We note that if we can bound $\text{Loss}_\mathcal{M}(a; \mathbf{x}, \mathbf{x}')$ for any neighboring datasets $\mathbf{x}, \mathbf{x}'$ with high probability over $a \sim \mathcal{M}(\mathbf{x})$, then Theorem 2.3 tells us that $\mathcal{M}$ is differentially private. Moreover, Theorem 2.3 also implies that this approach is without loss of generality (up to a small difference in the parameters). Thus, our composition theorems will focus on bounding the privacy loss with high probability.

A useful property of differential privacy is that it is preserved under post-processing without degrading the parameters:

Theorem 2.5 (Post-Processing [DMNS06]). *Let $\mathcal{M} : \mathcal{X}^n \to \mathcal{Y}$ be (ε, δ)-DP and $f : \mathcal{Y} \to \mathcal{Y}'$ be any randomized algorithm. Then $f \circ \mathcal{M} : \mathcal{X}^n \to \mathcal{Y}'$ is (ε, δ)-DP.*

We next recall a useful characterization from [KOV15]: any DP algorithm can be written as the post-processing of a simple, canonical algorithm which is a generalization of *randomized response*.

Definition 2.6. For any $\varepsilon, \delta \geq 0$, we define the *randomized response* algorithm $\text{RR}_{\varepsilon,\delta} : \{0, 1\} \to \{0, \top, \bot, 1\}$ as the following (Note that if $\delta = 0$, we will simply write the algorithm $\text{RR}_{\varepsilon,\delta}$ as RR_ε.)

$$\begin{aligned}
\mathbb{P}[\text{RR}_{\varepsilon,\delta}(0) = 0] &= \delta & \mathbb{P}[\text{RR}_{\varepsilon,\delta}(1) = 0] &= 0 \\
\mathbb{P}[\text{RR}_{\varepsilon,\delta}(0) = \top] &= (1-\delta)\frac{e^\varepsilon}{1+e^\varepsilon} & \mathbb{P}[\text{RR}_{\varepsilon,\delta}(1) = \top] &= (1-\delta)\frac{1}{1+e^\varepsilon} \\
\mathbb{P}[\text{RR}_{\varepsilon,\delta}(0) = \bot] &= (1-\delta)\frac{1}{1+e^\varepsilon} & \mathbb{P}[\text{RR}_{\varepsilon,\delta}(1) = \bot] &= (1-\delta)\frac{e^\varepsilon}{1+e^\varepsilon} \\
\mathbb{P}[\text{RR}_{\varepsilon,\delta}(0) = 1] &= 0 & \mathbb{P}[\text{RR}_{\varepsilon,\delta}(1) = 1] &= \delta
\end{aligned}$$

Kairouz, Oh, and Viswanath [KOV15] show that any (ε, δ)–DP algorithm can be viewed as a post-processing of the output of $\text{RR}_{\varepsilon,\delta}$ for an appropriately chosen input.

Theorem 2.7 ([KOV15], see also [MV16]). *For every (ε, δ)-DP algorithm $\mathcal{M}$ and for all neighboring databases $\mathbf{x}^0$ and $\mathbf{x}^1$, there exists a randomized algorithm T where $T(\text{RR}_{\varepsilon,\delta}(b))$ is identically distributed to $\mathcal{M}(\mathbf{x}^b)$ for $b \in \{0, 1\}$.*

This theorem will be useful in our analyses, because it allows us to without loss of generality analyze compositions of these simple algorithms $\text{RR}_{\varepsilon,\delta}$ with varying privacy parameters.

We now define the adaptive composition of differentially private algorithms in the setting introduced by [DRV10] and then extended to *heterogenous* privacy parameters in [MV16], in which all of the privacy parameters are fixed prior to the start of the computation. The following "composition game" is an abstract model of composition in which an adversary can adaptively select between neighboring datasets at each round, as well as a differentially private algorithm to run at each round – both choices can be a function of the realized outcomes of all previous rounds. However, crucially, the adversary must select at each round an algorithm that satisfies the privacy parameters which have been fixed ahead of time – the choice of parameters cannot itself be a function of the realized outcomes of previous rounds. We define this model of interaction formally in Algorithm 1 where the output is the *view* of the adversary $\mathcal{A}$ which includes any random coins she uses $R_{\mathcal{A}}$ and the outcomes $A_1, \cdots, A_k$ of every round.

Algorithm 1 $\text{FixedParamComp}(\mathcal{A}, \mathcal{E} = (\mathcal{E}_1, \cdots, \mathcal{E}_k), b)$, where $\mathcal{A}$ is a randomized algorithm, $\mathcal{E}_1, \cdots, \mathcal{E}_k$ are classes of randomized algorithms, and $b \in \{0, 1\}$.

Select coin tosses $R_{\mathcal{A}}^b$ for $\mathcal{A}$ uniformly at random.
for $i = 1, \cdots, k$ **do**
$\quad \mathcal{A} = \mathcal{A}(R_{\mathcal{A}}^b, A_1^b, \cdots, A_{i-1}^b)$ gives neighboring datasets $\mathbf{x}^{i,0}, \mathbf{x}^{i,1}$, and $\mathcal{M}_i \in \mathcal{E}_i$
$\quad \mathcal{A}$ receives $A_i^b = \mathcal{M}_i(\mathbf{x}^{i,b})$
return view $V^b = (R_{\mathcal{A}}^b, A_1^b, \cdots, A_k^b)$

Definition 2.8 (Adaptive Composition [DRV10], [MV16]). We say that the sequence of parameters $\varepsilon_1, \cdots, \varepsilon_k \geq 0$, $\delta_1, \cdots, \delta_k \in [0, 1)$ satisfies $(\varepsilon_g, \delta_g)$-differential privacy under adaptive composition if for every adversary $\mathcal{A}$, and $\mathcal{E} = (\mathcal{E}_1, \cdots, \mathcal{E}_k)$ where $\mathcal{E}_i$ is the class of $(\varepsilon_i, \delta_i)$-DP algorithms, we have $\text{FixedParamComp}(\mathcal{A}, \mathcal{E}, \cdot)$ is $(\varepsilon_g, \delta_g)$-DP in its last argument, i.e. $V^0 \approx_{\varepsilon_g, \delta_g} V^1$.

We first state a basic composition theorem which shows that the adaptive composition satisfies differential privacy where "the parameters just add up."

Theorem 2.9 (Basic Composition [DMNS06], [DKM+06]). *The sequence $\varepsilon_1, \cdots, \varepsilon_k$ and $\delta_1, \cdots \delta_k$ satisfies $(\varepsilon_g, \delta_g)$-differential privacy under adaptive composition where*

$$\varepsilon_g = \sum_{i=1}^{k} \varepsilon_i, \quad \text{and} \quad \delta_g = \sum_{i=1}^{k} \delta_i.$$

We now state the advanced composition bound from [DRV10] which gives a quadratic improvement to the basic composition bound.

Theorem 2.10 (Advanced Composition). *For any $\hat{\delta} > 0$, the sequence $\varepsilon_1, \cdots, \varepsilon_k$ and $\delta_1, \cdots \delta_k$ where $\varepsilon = \varepsilon_i$ and $\delta = \delta_i$ for all $i \in [k]$ satisfies $(\varepsilon_g, \delta_g)$-differential privacy under adaptive composition where*

$$\varepsilon_g = \varepsilon(e^{\varepsilon} - 1)k + \varepsilon\sqrt{2k \log(1/\hat{\delta})}, \quad \text{and} \quad \delta_g = k\delta + \hat{\delta}.$$

This theorem can be easily generalized to hold for values of ε_i that are not all equal (as done in [KOV15]). However, this is not as all-encompassing as it would appear at first blush, because this straightforward generalization would not allow for the *values* of ε_i and δ_i to be chosen adaptively by the data analyst. Indeed, the definition of differential privacy itself (Definition 2.4) does not straightforwardly extend to this case. The remainder of this paper is devoted to laying out a framework for sensibly talking about the privacy parameters ε_i and δ_i being chosen adaptively by the data analyst, and to prove composition theorems (including an analogue of Theorem 2.10) in this model.

3   Composition with Adaptively Chosen Parameters

We now introduce the model of composition with adaptive parameter selection, and define privacy in this setting.

3.1   Definition

We want to model composition as in the previous section, but allow the adversary the ability to also choose the privacy parameters $(\varepsilon_i, \delta_i)$ as a function of previous rounds of interaction. We will define the view of the interaction, similar to the view in `FixedParamComp`, to be the tuple that includes $\mathcal{A}$'s random coin tosses $R_{\mathcal{A}}$ and the outcomes $A = (A_1, \cdots, A_k)$ of the algorithms she chose. Formally, we define an *adaptively chosen privacy parameter composition game* in Algorithm 2 which takes as input an adversary $\mathcal{A}$, a number of rounds of interaction k,[3] and an experiment parameter $b \in \{0, 1\}$.

Algorithm 2 `AdaptParamComp`$(\mathcal{A}, k, b)$

Select coin tosses $R_{\mathcal{A}}^b$ for $\mathcal{A}$ uniformly at random.
for $i = 1, \cdots, k$ **do**
$\quad \mathcal{A} = \mathcal{A}(R_{\mathcal{A}}^b, A_1^b, \cdots, A_{i-1}^b)$ gives neighboring $\mathbf{x}^{i,0}, \mathbf{x}^{i,1}$, parameters $(\varepsilon_i, \delta_i)$, $\mathcal{M}_i$ that is $(\varepsilon_i, \delta_i)$-DP
$\quad \mathcal{A}$ receives $A_i^b = \mathcal{M}_i(\mathbf{x}^{i,b})$
return view $V^b = (R_{\mathcal{A}}^b, A_1^b, \cdots, A_k^b)$

We then define the privacy loss with respect to `AdaptParamComp`$(\mathcal{A}, k, b)$ in the following way for a fixed view $v = (r, a)$ where r represents the random coin tosses of $\mathcal{A}$ and we write $v_{<i} = (r, a_1, \cdots, a_{i-1})$:

$$\texttt{Loss}(v) = \log\left(\frac{\mathbb{P}\left[V^0 = v\right]}{\mathbb{P}\left[V^1 = v\right]} \right) = \sum_{i=1}^{k} \log\left(\frac{\mathbb{P}\left[\mathcal{M}_i(\mathbf{x}^{i,0}) = v_i | v_{<i}\right]}{\mathbb{P}\left[\mathcal{M}_i(\mathbf{x}^{i,1}) = v_i | v_{<i}\right]} \right) \stackrel{\text{def}}{=} \sum_{i=1}^{k} \texttt{Loss}_i(v_{\leq i}). \tag{1}$$

Note that the privacy parameters $(\varepsilon_i, \delta_i)$ depend on the previous outcomes that $\mathcal{A}$ receives. We will frequently shorten our notation $\varepsilon_t = \varepsilon_t(v_{<t})$ and $\delta_t = \delta_t(v_{<t})$ when the outcome is understood.

---

[3]Note that in the adaptive parameter composition game, the adversary has the option of effectively stopping the composition early at some round $k' < k$ by simply setting $\varepsilon_i = \delta_i = 0$ for all rounds $i > k'$. Hence, the parameter k will not appear in our composition theorems the way it does when privacy parameters are fixed. This means that we can effectively take k to be infinite. For technical reasons, it is simpler to have a finite parameter k, but the reader should imagine it as being an enormous number(say the number of atoms in the universe) so as not to put any constraint at all on the number of rounds of interaction with the adversary.

It no longer makes sense to claim that the privacy loss of the adaptive parameter composition experiment is bounded by any fixed constant, because the privacy parameters (with which we would presumably want to use to bound the privacy loss) are themselves random variables. Instead, we define two objects which can be used by a data analyst to control the privacy loss of an adaptive composition of algorithms.

The first object, which we call a *privacy odometer* will be parameterized by one global parameter δ_g and will provide a running real valued output that will, with probability $1 - \delta_g$, upper bound the privacy loss at each round of any adaptive composition in terms of the *realized* values of ε_i and δ_i selected at each round.

Definition 3.1 (Privacy Odometer). A function $\text{COMP}_{\delta_g} : \mathbb{R}^{2k}_{\geq 0} \to \mathbb{R} \cup \{\infty\}$ is a *valid privacy odometer* if for all adversaries in $\text{AdaptParamComp}(\mathcal{A}, k, b)$, with probability at most δ_g over $v \sim V^0$:

$$|\text{Loss}(v)| > \text{COMP}_{\delta_g}(\varepsilon_1, \delta_1, \cdots, \varepsilon_k, \delta_k).$$

The second object, which we call a *privacy filter*, is a stopping time rule. It takes two global parameters $(\varepsilon_g, \delta_g)$ and will at each round either output CONT or HALT. Its guarantee is that with probability $1 - \delta_g$, it will output HALT if the privacy loss has exceeded ε_g.

Definition 3.2 (Privacy Filter). A function $\text{COMP}_{\varepsilon_g, \delta_g} : \mathbb{R}^{2k}_{\geq 0} \to \{\text{HALT}, \text{CONT}\}$ is a *valid privacy filter* for $\varepsilon_g, \delta_g \geq 0$ if for all adversaries $\mathcal{A}$ in $\text{AdaptParamComp}(\mathcal{A}, k, b)$, the following "bad event" occurs with probability at most δ_g when $v \sim V^0$:

$$|\text{Loss}(v)| > \varepsilon_g \quad \text{and} \quad \text{COMP}_{\varepsilon_g, \delta_g}(\varepsilon_1, \delta_1, \cdots, \varepsilon_k, \delta_k) = \text{CONT}.$$

We note two things about the usage of these objects. First, a valid privacy odometer can be used to provide a running upper bound on the privacy loss at each intermediate round: the privacy loss at round $k' < k$ must with high probability be upper bounded by $\text{COMP}_{\delta_g}(\varepsilon_1, \delta_1, \ldots, \varepsilon_{k'}, \delta_{k'}, 0, 0, \ldots, 0, 0)$ – i.e. the bound that results by setting all future privacy parameters to 0. This is because setting all future privacy parameters to zero is equivalent to stopping the computation at round k', and is a feasible choice for the adaptive adversary $\mathcal{A}$. Second, a privacy filter can be used to guarantee that with high probability, the stated privacy budget ε_g is never exceeded – the data analyst at each round k' simply queries $\text{COMP}_{\varepsilon_g, \delta_g}(\varepsilon_1, \delta_1, \ldots, \varepsilon_{k'}, \delta_{k'}, 0, 0, \ldots, 0, 0)$ *before* she runs algorithm k', and runs it only if the filter returns CONT. Again, this is guaranteed because the continuation is a feasible choice of the adversary, and the guarantees of both a filter and an odometer are quantified over all adversaries. We give the formal description of this interaction where $\mathcal{A}$ uses the privacy filter in Algorithm 3.

From the way we have defined a valid privacy filter, we have the following proposition:

Proposition 3.3. *If* $\text{COMP}_{\varepsilon_g, \delta_g}$ *is a valid privacy filter then the views* V_{F}^0 *and* V_{F}^1 *of the adversary from* $\text{PrivacyFilterComp}\left(\mathcal{A}, k, b; \text{COMP}_{\varepsilon_g, \delta_g}\right)$ *with* $b = 0$ *and* $b = 1$ *respectively, are* $(\varepsilon_g, \delta_g)$*-point-wise indistinguishable and hence* $V_{\text{F}}^0 \approx_{\varepsilon_g, \delta_g} V_{\text{F}}^1$.

3.2 Focusing on Randomized Response

In [KOV15, MV16] Theorem 2.7 was used to show that for ordinary composition (Theorem 2.8), it suffices to analyze the composition of randomized response. In this section, we show something similar for privacy odometers and filters. Specifically, we show that we can simulate

---

Algorithm 3 `PrivacyFilterComp`$(\mathcal{A}, k, b; \texttt{COMP}_{\varepsilon_g, \delta_g})$

Select coin tosses $R_{\mathcal{A}}^b$ for $\mathcal{A}$ uniformly at random.
for $i = 1, \cdots, k$ **do**
    $\mathcal{A} = \mathcal{A}(R_{\mathcal{A}}^b, A_1^b, \cdots, A_{i-1}^b)$ gives neighboring $\mathbf{x}^{i,0}, \mathbf{x}^{i,1}$, $(\varepsilon_i, \delta_i)$, and $\mathcal{M}_i$ that is $(\varepsilon_i, \delta_i)$-DP
    if $\texttt{COMP}_{\varepsilon_g, \delta_g}(\varepsilon_1, \delta_1, \cdots, \varepsilon_i, \delta_i, 0, 0, \cdots, 0, 0) = \texttt{HALT}$ **then**
        $A_i, \cdots, A_k = \perp$
        BREAK
    else
        $\mathcal{A}$ receives $A_i^b = \mathcal{M}_i(\mathbf{x}^{i,b})$
return   view $V_{\texttt{F}}^b = (R_{\mathcal{A}}^b, A_1^b, \cdots, A_k^b)$

`AdaptParamComp`$(\mathcal{A}, k, b)$ by defining a new adversary that chooses the differentially private algorithm $\mathcal{M}_i$ of adversary $\mathcal{A}$, but uses the randomized response algorithm from Theorem 2.6 each round along with a post-processing function, which together determine the distribution for $\mathcal{M}_i$.

In Algorithm 4, we define the new composition game `SimulatedComp`$(\mathcal{A}, k, b)$ with adversary $\mathcal{A}$ that outputs the view W^b, which includes the internal randomness $R_{\mathcal{A}}^b$ of $\mathcal{A}$ with the randomized response outcomes $Z^b = (Z_1^b, \cdots, Z_k^b)$. From Theorem 2.7, we know that we can simulate any (ε, δ)-DP algorithm as a randomized post-processing function T on top of $\texttt{RR}_{\varepsilon, \delta}$. Thus given the outcomes prior to round i, $\mathcal{A}$ selects $\mathcal{M}_i$, which is equivalent to selecting a post-processing function T_i. Note that we can simulate T_i as a deterministic function P_i with access to random coins $R_{\texttt{SIM}_i}^b$, i.e. $P_i\left(\texttt{RR}_{\varepsilon_i, \delta_i}(b); R_{\texttt{SIM}_i}^b\right) \sim T_i\left(\texttt{RR}_{\varepsilon_i, \delta_i}(b)\right)$. We then include the random coins $R_{\texttt{SIM}}^b = \left(R_{\texttt{SIM}_1}^b, \cdots, R_{\texttt{SIM}_k}^b\right)$ in the view of adversary $\mathcal{A}$ in `SimulatedComp`$(\mathcal{A}, k, b)$. From the view $W^b = \left(R_{\mathcal{A}}^b, R_{\texttt{SIM}}, Z_1^b, \cdots, Z_k^b\right)$, $\mathcal{A}$ would be able to reconstruct the privacy parameters selected each round along with algorithms $\mathcal{M}_1, \cdots, \mathcal{M}_k$ used, which would also determine the post-processing functions $P_1, \cdots, P_k$.

Algorithm 4 `SimulatedComp`$(\mathcal{A}, k, b)$

Select coin tosses $R_{\mathcal{A}}^b$ for $\mathcal{A}$ uniformly at random.
for $i = 1, \cdots, k$ **do**
    $\mathcal{A} = \mathcal{A}\left(R_{\mathcal{A}}^b, Y_1^b, \cdots, Y_{i-1}^b\right)$ gives neighboring $\mathbf{x}^{i,0}, \mathbf{x}^{i,1}$, parameters $(\varepsilon_i, \delta_i)$, $\mathcal{M}_i$ that is $(\varepsilon_i, \delta_i)$-DP.

    Let P_i be a deterministic post-processing function, such that

$$P_i\left(\texttt{RR}_{\varepsilon_i, \delta_i}(b); R_{\texttt{SIM}_i}^b\right) \sim \mathcal{M}_i\left(\mathbf{x}^{i,b}\right) \tag{2}$$

    for uniformly random $R_{\texttt{SIM}_i}^b$.
    Compute $Z_i^b = \texttt{RR}_{\varepsilon_i, \delta_i}(b)$ and $Y_i^b = P_i(Z_i^b; R_{\texttt{SIM}_i}^b)$.
    $\mathcal{A}$ receives Y_i^b.
return   view $W^b = (R_{\mathcal{A}}^b, R_{\texttt{SIM}}^b, Z_1^b, \cdots, Z_k^b)$, where $R_{\texttt{SIM}}^b = \left(R_{\texttt{SIM}_1}^b, \cdots, R_{\texttt{SIM}_k}^b\right)$.

From the way that we have defined P_i in (2), for each fixed value of the internal randomness of $\mathcal{A}$, the view of $\mathtt{AdaptParamComp}(\mathcal{A}, k, b)$ is distributed identically to a post-processing of the view W^b from $\mathtt{SimulatedComp}(\mathcal{A}, k, b)$.

Lemma 3.4. *For every adversary $\mathcal{A}$, the deterministic function P defined as*

$$P\left(R_{\mathcal{A}}^b, R_{\mathrm{SIM}}^b, Z_1^b, \cdots, Z_k^b\right) = \left(R_{\mathcal{A}}^b, P_1(Z_1^b; R_{\mathrm{SIM}_1}^b), \cdots, P_k(Z_k^b; R_{\mathrm{SIM}_k}^b)\right) \tag{3}$$

ensures $P\left(\mathtt{SimulatedComp}(\mathcal{A}, k, b)\right)$ and $\mathtt{AdaptParamComp}(\mathcal{A}, k, b)$ are identically distributed.

Since $R_{\mathcal{A}}^b$ is the first argument of both random variables, they are also identically distributed conditioned on any fixed value of $R_{\mathcal{A}}^b$. This point-wise equivalence for every value of the internal randomness allows us to without loss of generality analyze *deterministic* adversaries and post-processing functions of $\mathtt{SimulatedComp}(\mathcal{A}, k, b)$ in order to reason about the view of $\mathtt{AdaptParamComp}(\mathcal{A}, k, b)$. Because the randomness is fixed, for clarity, we will omit the random coins $R_{\mathcal{A}}^b$ from the view of both composition games for the rest of the analysis.

We will now show that it is sufficient to prove bounds in which ε_i may be adaptively chosen at each round, and in which $\{\delta_i\} \equiv 0$ uniformly. We do this by giving a generic way to extend a bound in the $\delta_i = 0$ case to a bound that holds when the δ_i may be non-zero. Define a slight modification of Algorithm 4 called $\widetilde{\mathtt{SimulatedComp}}(\mathcal{A}, k, b)$ which is the same as $\mathtt{SimulatedComp}(\mathcal{A}, k, b)$ except that it computes $\widetilde{Z}_i^b = \mathrm{RR}_{\varepsilon_i}(b)$ (where $\delta_i = 0$) and sets $\widetilde{Y}_i^b = P_i\left(\widetilde{Z}_i^b; R_{\mathrm{SIM}_i}^b\right)$. We then define the final view of the adversary $\mathcal{A}$ in $\widetilde{\mathtt{SimulatedComp}}(\mathcal{A}, k, b)$ as $\widetilde{W}^b$ where

$$\widetilde{W}^b = \left(R_{\mathrm{SIM}}^b, \widetilde{Z}_1^b, \cdots, \widetilde{Z}_k^b\right) \qquad \text{and} \qquad \widetilde{V}^b = \left(\widetilde{Y}_1^b, \cdots, \widetilde{Y}_k^b\right) = P\left(\widetilde{W}^b\right) \tag{4}$$

for $P(\cdot)$ given in (3). We then say that $\widetilde{\mathrm{COMP}}_{\delta_g}$ (also $\widetilde{\mathrm{COMP}}_{\varepsilon_g, \delta_g}$) is a valid privacy odometer (filter) when $\{\delta_i\} \equiv 0$ if over all deterministic adversaries $\mathcal{A}$ in $\widetilde{\mathtt{SimulatedComp}}(\mathcal{A}, k, b)$ the condition in Theorem 3.1 (Theorem 3.2) holds with probability at most δ_g over $\widetilde{v} \sim P\left(\widetilde{W}^b\right)$ except now the privacy loss is given as

$$\widetilde{\mathrm{Loss}}(\widetilde{v}) = \log\left(\frac{\mathbb{P}\left[\widetilde{V}^0 = \widetilde{v}\right]}{\mathbb{P}\left[\widetilde{V}^1 = \widetilde{v}\right]}\right) = \sum_{i=1}^{k} \log\left(\frac{\mathbb{P}\left[P_i\left(\mathrm{RR}_{\varepsilon_i}(0); R_{\mathrm{SIM}_i}^0\right) = \widetilde{v}_i | \widetilde{v}_{<i}\right]}{\mathbb{P}\left[P_i\left(\mathrm{RR}_{\varepsilon_i}(1); R_{\mathrm{SIM}_i}^1\right) = \widetilde{v}_i | \widetilde{v}_{<i}\right]}\right) \overset{\text{def}}{=} \sum_{i=1}^{k} \widetilde{\mathrm{Loss}}_i(\widetilde{v}_{\leq i}). \tag{5}$$

The following result gives the connection between valid privacy odometers and filters in the modified game $\widetilde{\mathtt{SimulatedComp}}(\mathcal{A}, k, b)$ with the original definitions given in Theorems 3.1 and 3.2.

Lemma 3.5. *If $\widetilde{\mathrm{COMP}}_{\delta_g}$ is a valid privacy odometer when $\{\delta_i\} \equiv 0$, then for every $\delta_g' \geq 0$, $\mathrm{COMP}_{\delta_g + \delta_g'}$ is a valid privacy odometer where*

$$\mathrm{COMP}_{\delta_g + \delta_g'}(\varepsilon_1, \delta_1, \cdots, \varepsilon_k, \delta_k) = \begin{cases} \infty & \text{if } \sum_{i=1}^{k} \delta_i > \delta_g' \\ \widetilde{\mathrm{COMP}}_{\delta_g}(\varepsilon_1, 0, \cdots, \varepsilon_k, 0) & \text{otherwise} \end{cases}.$$

If $\widetilde{\mathrm{COMP}}_{\varepsilon_g, \delta_g}$ is a valid privacy filter when $\{\delta_i\} \equiv 0$, then for every $\delta_g' \geq 0$, $\mathrm{COMP}_{\varepsilon_g, \delta_g + \delta_g'}$ is a valid privacy filter where

$$\mathrm{COMP}_{\varepsilon_g, \delta_g + \delta_g'}(\varepsilon_1, \delta_1, \cdots, \varepsilon_k, \delta_k) = \begin{cases} \mathtt{HALT} & \text{if } \sum_{i=1}^{k} \delta_i > \delta_g' \\ \widetilde{\mathrm{COMP}}_{\varepsilon_g, \delta_g}(\varepsilon_1, 0, \cdots, \varepsilon_k, 0) & \text{otherwise} \end{cases}.$$

11

Proof. Let $W = (R_{\text{SIM}}, Z_1, \cdots, Z_k)$ be the view of $\mathcal{A}$ in $\texttt{SimulatedComp}(\mathcal{A}, k, 0)$ and $\widetilde{W} = (R_{\text{SIM}}, \widetilde{Z}_1, \cdots, \widetilde{Z}_k)$ be her view in $\widetilde{\texttt{SimulatedComp}}(A, k, 0)$ (where $\{\delta_i\} \equiv 0$). We will also write the view of $\texttt{AdaptParamComp}(\mathcal{A}, k, 0)$ as $V = (A_1, \cdots, A_k)$ and the post-processing functions of $\mathcal{A}$ as P_i from (2). As in (3), we will use the notation $P(W) = \big(P_1(Z_1; R_{\text{SIM}_i}), \cdots, P_k(Z_k; R_{\text{SIM}_i})\big)$ and similarly for $\widetilde{V} = P(\widetilde{W})$. Recall that from Theorem 3.4 that we know $V \sim P(W)$, even if $\mathcal{A}$ were randomized.

Consider the following method of sampling from $\text{RR}_{\varepsilon, \delta}$: first select outcome $\widetilde{z}$ from $\text{RR}_{\varepsilon}(0)$, then with probability $1 - \delta$ set $z = \widetilde{z}$ – otherwise set $z = 0$. Note that this samples from the correct distribution for $\text{RR}_{\varepsilon, \delta}(0)$. We can thus couple draws from $\text{RR}_{\varepsilon}(0)$ and $\text{RR}_{\varepsilon, \delta}(0)$, so for our setting we write the coupled random variable as: $\mathbf{V} = (V, \widetilde{V})$.

We then define the following sets:

$$\mathcal{F} = \{(w = (r, z), \widetilde{w} = (r, \widetilde{z})) : \exists t \in [k] \text{ s.t. } z_t \neq \widetilde{z}_t\}, \quad \mathcal{G}_t = \left\{v : \sum_{i=1}^{t} \delta_i(v_{<i}) \leq \delta'_g\right\},$$

$$\mathcal{F}_t = \{(w = (r, z), \widetilde{w} = (r, \widetilde{z})) : z_t \neq \widetilde{z}_t \text{ and } z_i = \widetilde{z}_i \quad \forall i < t\}, \quad \mathcal{H} = \left\{v : |\widehat{\text{Loss}}(v)| \geq \widetilde{\text{COMP}}_{\delta_g}(\varepsilon_1, 0, \cdots, \varepsilon_k, 0)\right\}.$$

We then want to show that we can bound the privacy loss with high probability. Specifically,

$$\mathbb{P}_{\mathbf{V}}\left[|\text{Loss}(V)| \geq \widetilde{\text{COMP}}_{\delta_g}(\varepsilon_1, 0, \cdots, \varepsilon_k, 0) \quad \wedge \quad \sum_{t=1}^{k} \delta_t \leq \delta'_g\right] \leq \delta_g + \delta'_g.$$

where each ε_i is a function of the outputs of the prefix $\widetilde{V}_{<i}$ of the full view $\widetilde{V}$ from $\widetilde{\texttt{SimulatedComp}}(\mathcal{A}, k, 0)$. We now show that the quantity that we want to bound can be written as the probability of the coupled random variables $\mathbf{V}$ and $\mathbf{W} = \big(W, \widetilde{W}\big)$ being contained in the sets that we defined above.

$$\mathbb{P}_{\mathbf{V} \sim \big(P(W), P(\widetilde{W})\big)}\left[|\text{Loss}(V)| \geq \widetilde{\text{COMP}}_{\delta_g}(\varepsilon_1, 0, \cdots, \varepsilon_k, 0) \quad \wedge \quad \sum_{t=1}^{k} \delta_t \leq \delta'_g\right]$$

$$\leq \mathbb{P}\left[(\mathbf{W} \in \mathcal{F} \quad \wedge \quad V \in \mathcal{G}_k) \quad \vee \quad (V \in \mathcal{H} \quad \wedge \quad \mathbf{W} \notin \mathcal{F})\right]$$

$$\leq \mathbb{P}\left[\mathbf{W} \in \mathcal{F} \quad \wedge \quad V \in \mathcal{G}_k\right] + \mathbb{P}\left[\widetilde{V} \in \mathcal{H}\right]$$

$$\leq \mathbb{P}\left[\mathbf{W} \in \mathcal{F} \quad \wedge \quad V \in \mathcal{G}_k\right] + \delta_g \tag{6}$$

Note, that if $\sum_{i=1}^{k} \delta_i(v_{<i}) \leq \delta_g$ then we must have $\sum_{i=1}^{t} \delta_i(v_{<i}) \leq \delta_g$ for each $t < k$, so that $\mathcal{G}_k \subseteq \mathcal{G}_t$. We then use the fact that $\{\mathcal{F}_t : t \in [k]\}$ forms a partition of $\mathcal{F}$, i.e. $\mathcal{F} = \bigcup_{t=1}^{k} \mathcal{F}_t$ and $\mathcal{F}_i \cap \mathcal{F}_j = \emptyset$ for $i \neq j$, to obtain the following:

$$\mathbb{P}[\mathbf{W} \in \mathcal{F} \quad \wedge \quad V \in \mathcal{G}_k] = \sum_{t=1}^{k} \mathbb{P}[\mathbf{W} \in \mathcal{F}_t \quad \wedge \quad V \in \mathcal{G}_k] \leq \sum_{t=1}^{k} \mathbb{P}[\mathbf{W} \in \mathcal{F}_t \quad \wedge \quad V \in \mathcal{G}_t].$$

$$\sum_{t=1}^{k} \mathbb{P}\big[\mathbf{W} \in \mathcal{F}_t \quad \wedge \quad V \in \mathcal{G}_t\big]$$

$$\leq \sum_{t=1}^{k} \mathbb{P}\Big[\mathbf{W} \in \mathcal{F}_t \quad \wedge \quad \widetilde{V} \in \mathcal{G}_t\Big]$$

$$= \sum_{t=1}^{k} \sum_{\widetilde{v} \in \mathcal{G}_t} \mathbb{P}\Big[\widetilde{V} = \widetilde{v}\Big] \mathbb{P}[\mathbf{W} \in \mathcal{F}_t | \widetilde{v}]$$

$$\leq \sum_{t=1}^{k} \sum_{\widetilde{v} \in \mathcal{G}_t} \mathbb{P}\Big[\widetilde{V} = \widetilde{v}\Big] \delta_t(\widetilde{v}_{<t}).$$

We now switch the order of summation to obtain our result

$$\sum_{t=1}^{k} \sum_{\widetilde{v} \in \mathcal{G}_t} \mathbb{P}\Big[\widetilde{V} = \widetilde{v}\Big] \delta_t(\widetilde{v}_{<t}) = \sum_{\widetilde{v}} \mathbb{P}\Big[\widetilde{V} = \widetilde{v}\Big] \sum_{t:\sum_{i=1}^{t} \delta_i(\widetilde{v}_{<i}) \leq \delta_g'} \delta_t(\widetilde{v}_{<t}) \leq \sum_{\widetilde{v}} \mathbb{P}\Big[\widetilde{V} = \widetilde{v}\Big] \delta_g' = \delta_g'. \qquad (7)$$

We then combine this with (6) to prove our first statement for the privacy odometer.

Using the same notation as above, we now move to proving the statement for the privacy filter. It suffices to prove the following:

$$\mathbb{P}_{\mathbf{V} \sim (P(W), P(\widetilde{W}))}\Big[|\mathrm{Loss}(V)| \geq \varepsilon_g \quad \wedge \quad V \in \mathcal{G}_k \quad \wedge \quad \widetilde{\mathrm{COMP}}_{\varepsilon_g, \delta_g}(\varepsilon_1, 0, \cdots, \varepsilon_k, 0) = \mathrm{CONT}\Big] \leq \delta_g + \delta_g'.$$

We now define a slight variant of $\mathcal{H}$ from above:

$$\mathcal{H}_{\varepsilon_g} = \Big\{v : \big|\widetilde{\mathrm{Loss}}(v)\big| \geq \varepsilon_g\Big\}.$$

Similar to what we showed in (6) for the privacy odometer, we have

$$\mathbb{P}_{\mathbf{V}}\Big[|\mathrm{Loss}(V)| \geq \varepsilon_g \quad \wedge \quad V \in \mathcal{G}_k \quad \wedge \quad \widetilde{\mathrm{COMP}}_{\varepsilon_g, \delta_g}(\varepsilon_1, 0, \cdots, \varepsilon_k, 0) = \mathrm{CONT}\Big]$$

$$\leq \mathbb{P}_{\mathbf{V} \sim (P(W), P(\widetilde{W}))}\Big[\big((\mathbf{W} \in \mathcal{F} \wedge V \in \mathcal{G}_k) \vee (V \in \mathcal{H}_{\varepsilon_g} \wedge \mathbf{W} \notin \mathcal{F})\big) \quad \wedge \quad \widetilde{\mathrm{COMP}}_{\varepsilon_g, \delta_g}(\varepsilon_1, 0, \cdots, \varepsilon_k, 0) = \mathrm{CONT}\Big]$$

$$\leq \mathbb{P}[\mathbf{W} \in \mathcal{F} \quad \wedge \quad V \in \mathcal{G}_k] + \mathbb{P}\Big[\widetilde{V} \in \mathcal{H}_{\varepsilon_g} \quad \wedge \quad \widetilde{\mathrm{COMP}}_{\varepsilon_g, \delta_g}(\varepsilon_1, 0, \cdots, \varepsilon_k, 0) = \mathrm{CONT}\Big]$$

$$\leq \mathbb{P}[\mathbf{W} \in \mathcal{F} \quad \wedge \quad V \in \mathcal{G}_k] + \delta_g$$

$$\leq \delta_g' + \delta_g$$

where the last inequality follows from (6) and (7). $\qquad \square$

3.3 Basic Composition

We first give an adaptive parameter version of the basic composition in Theorem 2.9.

Theorem 3.6. *For each nonnegative δ_g, COMP_{δ_g} is a valid privacy odometer where*

$$\mathrm{COMP}_{\delta_g + \delta_g'}(\varepsilon_1, \delta_1, \cdots, \varepsilon_k, \delta_k) = \begin{cases} \infty & \text{if } \sum_{i=1}^{k} \delta_i > \delta_g' \\ \sum_{i=1}^{k} \varepsilon_i & \text{otherwise} \end{cases}.$$

13

Additionally, for any $\varepsilon_g, \delta_g \geq 0$, $\mathrm{COMP}_{\varepsilon_g,\delta_g}$ is a valid privacy filter where

$$\mathrm{COMP}_{\varepsilon_g,\delta_g}(\varepsilon_1, \delta_1, \cdots, \varepsilon_k, \delta_k) = \begin{cases} \mathrm{HALT} & \text{if } \sum_{i=1}^{k} \delta_i > \delta_g' \text{ or } \sum_{i=1}^{k} \varepsilon_i > \varepsilon_g \\ \mathrm{CONT} & \text{otherwise} \end{cases}.$$

Proof. We use Theorems 3.4 and 3.5 so that we need to only reason about any deterministic adversary in $\widetilde{\mathrm{SimulatedComp}}(\mathcal{A}, k, b)$. We know that $(\varepsilon, 0)$-DP is closed under post-processing from Theorem 2.5, so that for any (randomized) post-processing function T, we have $T(\mathrm{RR}_\varepsilon(0)) \approx_{\varepsilon,0} T(\mathrm{RR}_\varepsilon(1))$ and by Theorem 2.3 we know that $T(\mathrm{RR}_\varepsilon(0))$ and $T(\mathrm{RR}_\varepsilon(1))$ are $(\varepsilon, 0)$-point-wise indistinguishable for any post-processing function T. The proof then follows simply from the definition of (pure) differential privacy, so for all possible views $\widetilde{v}$ of the adversary in $P\left(\widetilde{\mathrm{SimulatedComp}}(\mathcal{A}, k, b)\right)$:

$$\left|\widetilde{\mathrm{Loss}}(\widetilde{v})\right| \leq \sum_{i=1}^{k} \left|\log\left(\frac{\mathbb{P}\left[P_i\left(\mathrm{RR}_{\varepsilon_i}(0); R^0_{\mathrm{SIM}_i}\right) = \widetilde{v}_i | \widetilde{v}_{<i}\right]}{\mathbb{P}\left[P_i\left(\mathrm{RR}_{\varepsilon_i}(1); R^1_{\mathrm{SIM}_i}\right) = \widetilde{v}_i | \widetilde{v}_{<i}\right]}\right)\right| \leq \sum_{i=1}^{k} \varepsilon_i(\widetilde{v}_{<i})$$

where we explicitly write the dependence of the choice of ε_i by $\mathcal{A}$ at round i on the view from the previous rounds as $\varepsilon_i(\widetilde{v}_{<i})$ □

4 Concentration Preliminaries

We give a useful concentration bound that will be pivotal in proving an improved valid privacy odometer and filter from that given in Theorem 3.6. We first present a concentration bound for *self normalized processes*.

Lemma 4.1 (Corollary 2.2 in [dlPKLL04])**.** *If A and $B > 0$ are two random variables such that*

$$\mathbb{E}\left[\exp\left(\lambda A - \frac{\lambda^2}{2}B^2\right)\right] \leq 1 \tag{8}$$

for all $\lambda \in \mathbb{R}$, then for all $\delta \leq 1/e$, $\beta > 0$ we have

$$\mathbb{P}\left[|A| \geq \sqrt{(B^2 + \beta)\left(2 + \log\left(\frac{B^2}{\beta} + 1\right)\right)\log(1/\delta)}\right] \leq \delta.$$

To put this bound into context, suppose that B is a constant and we apply the bound with $\beta = B^2$. Then the bound simplifies to

$$\mathbb{P}\left[|A| \geq O\left(B\sqrt{\log(1/\delta)}\right)\right] \leq \delta,$$

which is just a standard concentration inequality for any subgaussian random variable A with standard deviation B.

We will apply Theorem 4.1 to random variables coming from martingales defined from the privacy loss functions. To set this up, we present some notation: let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability triple where $\emptyset = \mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots \subseteq \mathcal{F}$ is an increasing sequence of σ-algebras. Let X_i be a real-valued

$\mathcal{F}_i$-measurable random variable, such that $\mathbb{E}[X_i|\mathcal{F}_{i-1}] = 0$ a.s. for each i. We then consider the martingale where

$$M_0 = 0 \qquad M_k = \sum_{i=1}^{k} X_i, \qquad \forall k \geq 1. \tag{9}$$

We then use the following result which gives us a pair of random variables to which we can apply Theorem 4.1.

Lemma 4.2 (Lemma 2.4 in [vdG02]). *For M_k defined in (9), if there exists two random variables $C_i < D_i$ that are $\mathcal{F}_{i-1}$-measurable for $i \geq 1$*

$$C_i \leq X_i \leq D_i \qquad a.s. \quad \forall i \geq 1.$$

and we define U_k as

$$U_0^2 = 0, \qquad U_k^2 = \sum_{i=1}^{k} (D_i - C_i)^2, \quad \forall k \geq 1 \tag{10}$$

then

$$\exp\left[\lambda M_k - \frac{\lambda^2}{8} U_k^2\right]$$

is a supermartingale for all $\lambda \in \mathbb{R}$.

We then obtain the following result from combining Theorem 4.1 with Theorem 4.2.

Theorem 4.3. *Let M_k be defined as in (9) and satisfy the hypotheses of Theorem 4.2. Then for every fixed $k \geq 1$, $\beta > 0$ and $\delta \leq 1/e$, we have*

$$\mathbb{P}\left[|M_k| \geq \sqrt{\left(\frac{U_k^2}{4} + \beta\right)\left(2 + \log\left(\frac{U_k^2}{4\beta} + 1\right)\right)\log(1/\delta)}\right] \leq \delta$$

Given Theorem 3.5, we will focus on finding a valid privacy odometer and filter when $\{\delta_i\} \equiv 0$. Our analysis will then depend on the privacy loss $\widetilde{\text{Loss}}(\widetilde{V})$ from (5) where $\widetilde{V}$ is the view of the adversary in $\texttt{Simu}\widetilde{\texttt{lated}}\texttt{Comp}(\mathcal{A}, k, 0)$. We then focus on the following martingale in our analysis:

$$\widetilde{M}_k = \sum_{i=1}^{k} \left(\widetilde{\text{Loss}}_i(\widetilde{V}_{\leq i}) - \widetilde{\mu}_i\right) \qquad \text{where} \qquad \widetilde{\mu}_i = \mathbb{E}\left[\widetilde{\text{Loss}}_i(\widetilde{V}_{\leq i})\big|\widetilde{V}_{<i}\right]. \tag{11}$$

We can then bound the conditional expectation $\widetilde{\mu}_i$ with the following result from [DR16] that improves on an earlier result from [DRV10] by a factor of 2.

Lemma 4.4 ([DR16]). *For $\widetilde{\mu}_i$ defined in (11), we have $\widetilde{\mu}_i \leq \varepsilon_i (e^{\varepsilon_i} - 1)/2$.*

5 Advanced Composition for Privacy Filters

We next show that we can essentially get the same asymptotic bound as Theorem 2.10 for the privacy filter setting using the bound in Theorem 4.3 for the martingale given in (11).

Theorem 5.1. *We define $\mathcal{K}$ as the following*

$$\mathcal{K} \overset{def}{=} \sum_{j=1}^{k} \varepsilon_j \left(\frac{e^{\varepsilon_j} - 1}{2} \right) + \sqrt{ 2 \left(\sum_{i=1}^{k} \varepsilon_i^2 + \frac{\varepsilon_g^2}{28.04 \cdot \log(1/\delta_g)} \right) \left(1 + \frac{1}{2} \log \left(\frac{28.04 \cdot \log(1/\delta_g) \sum_{i=1}^{k} \varepsilon_i^2}{\varepsilon_g^2} + 1 \right) \right) \log(2/\delta_g) }.$$

(12)

$\texttt{COMP}_{\varepsilon_g, \delta_g}$ *is a valid privacy filter for $\delta_g \in (0, 1/e)$ and $\varepsilon_g > 0$ where*

$$\texttt{COMP}_{\varepsilon_g, \delta_g} (\varepsilon_1, \delta_1, \cdots, \varepsilon_k, \delta_k) = \left\{ \begin{array}{ll} \texttt{HALT} & \text{if } \sum_{i=1}^{k} \delta_i > \delta_g/2 \text{ or } \mathcal{K} > \varepsilon_g \\ \texttt{CONT} & \text{otherwise} \end{array} \right. .$$

Note that if we have $\sum_{i=1}^{k} \varepsilon_i^2 = O\left(1/\log(1/\delta_g) \right)$ and set $\varepsilon_g = \Theta\left(\sqrt{ \sum_{i=1}^{k} \varepsilon_i^2 \log(1/\delta_g) } \right)$ in (12), we are then getting the same asymptotic bound on the privacy loss as in [KOV15] and in Theorem 2.10 for the case when $\varepsilon_i = \varepsilon$ for $i \in [k]$. If $k\varepsilon^2 \le \frac{1}{8\log(1/\delta_g)}$, then Theorem 2.10 gives a bound on the privacy loss of $\varepsilon \sqrt{8k \log(1/\delta_g)}$. Note that there may be better choices for the constant 28.04 that we divide ε_g^2 by in (12), but for the case when $\varepsilon_g = \varepsilon \sqrt{8k \log(1/\delta_g)}$ and $\varepsilon_i = \varepsilon$ for every $i \in [n]$, it is nearly optimal.

Proof of Theorem 5.1. Note that Theorem 3.5 allows us to concentrate on showing that we can find an optimal privacy filter when $\{\delta_i\} \equiv 0$. We then focus on the martingale $\widetilde{M}_k$ given in (11). In order to apply Theorem 4.3 we set the lower bound for $\widetilde{M}_i$ to be $C_i = (-\varepsilon_i - \widetilde{\mu}_i)$ and upper bound to be $D_i = (\varepsilon_i - \widetilde{\mu}_i)$ in order to compute U_k^2 from (10). We then have for the martingale in (11) that

$$U_k^2 = 4 \sum_{i=1}^{k} \varepsilon_i^2.$$

We can then directly apply Theorem 4.3 to get the following for $\beta = \left(\frac{\varepsilon_g}{\sqrt{28.04 \cdot \log(1/\delta_g)}} \right)^2 > 0$ with probability at least $1 - \delta_g/2$

$$|\widetilde{M}_k| \le \sqrt{ 2 \left(\sum_{i=1}^{k} \varepsilon_i^2 + \frac{\varepsilon_g^2}{28.04 \cdot \log(1/\delta_g)} \right) \left(1 + \frac{1}{2} \log \left(\frac{28.04 \cdot \log(1/\delta_g) \sum_{i=1}^{k} \varepsilon_i^2}{\varepsilon_g^2} + 1 \right) \right) \log(2/\delta_g) }.$$

We can then obtain a bound on the privacy loss with probability at least $1 - \delta_g/2$ over $\widetilde{v} \sim \widetilde{V}^0$

$$\left| \widetilde{\text{Loss}}(\widetilde{v}) \right| \le \sum_{i=1}^{k} \widetilde{\mu}_i + \sqrt{ 2 \left(\sum_{i=1}^{k} \varepsilon_i^2 + \frac{\varepsilon_g^2}{28.04 \cdot \log(1/\delta_g)} \right) \left(1 + \frac{1}{2} \log \left(\frac{28.04 \cdot \log(1/\delta_g) \sum_{i=1}^{k} \varepsilon_i^2}{\varepsilon_g^2} + 1 \right) \right) \log(2/\delta_g) }$$

$$\le \sum_{i=1}^{k} \varepsilon_i (e^{\varepsilon_i} - 1) + \sqrt{ 2 \left(\sum_{i=1}^{k} \varepsilon_i^2 + \frac{\varepsilon_g^2}{28.04 \cdot \log(1/\delta_g)} \right) \left(1 + \frac{1}{2} \log \left(\frac{28.04 \cdot \log(1/\delta_g) \sum_{i=1}^{k} \varepsilon_i^2}{\varepsilon_g^2} + 1 \right) \right) \log(2/\delta_g) }.$$

$\square$

16

6 Advanced Composition for Privacy Odometers

One might hope to achieve the same sort of bound on the privacy loss from Theorem 2.10 when the privacy parameters may be chosen adversarially. However we show that this cannot be the case for any valid privacy odometer. In particular, even if an adversary selects the same privacy parameter $\varepsilon = o(\sqrt{\log(\log(n)/\delta_g)/k})$ each round but can adaptively select a time to stop interacting with `AdaptParamComp` (which is a restricted special case of the power of the general adversary – stopping is equivalent to setting all future $\varepsilon_i, \delta_i = 0$), then we show that there can be no valid privacy odometer achieving a bound of $o(\varepsilon\sqrt{k\log(\log(n)/\delta_g)})$. This gives a separation between the achievable bounds for a valid privacy odometers and filters. But for privacy applications, it is worth noting that δ_g is typically set to be (much) smaller than $1/n$, in which case this gap disappears (since $\log(\log(n)/\delta_g) = (1 + o(1))\log(1/\delta_g)$).

Theorem 6.1. *For any $\delta_g \in (0, O(1))$ there is no valid $\texttt{COMP}_{\delta_g}$ privacy odometer where*

$$\texttt{COMP}_{\delta_g}(\varepsilon_1, 0, \cdots, \varepsilon_k, 0) = \sum_{i=1}^{k} \varepsilon_i \left(\frac{e^{\varepsilon_i} - 1}{e^{\varepsilon_i} + 1} \right) + o\left(\sqrt{\sum_{i=1}^{k} \varepsilon_i^2 \log(\log(n)/\delta_g)} \right) \tag{13}$$

In order to prove Theorem 6.1, we use the following anti-concentration bound for a sum of random variables.

Lemma 6.2 (Lemma 8.1 in [LT91])**.** *Let $X_1, \cdots, X_k$ be a sequence of mean zero i.i.d. random variables such that $|X_1| < a$ and $\sigma^2 = \mathbb{E}\left[X_1^2\right]$. For every $\alpha > 0$ there exists two positive constants C_α and c_α such that for every x satisfying $\sqrt{k}\sigma C_\alpha \leq x \leq c_\alpha \frac{k\sigma^2}{a}$ we have*

$$\mathbb{P}\left[\sum_{i=1}^{k} X_i \geq x \right] \geq \exp\left[-(1 + \alpha)\frac{x^2}{2k\sigma^2} \right]$$

For $\gamma \in [1/2, 1)$, we define the random variables $\xi_i \in \{-1, 1\}$ where

$$\mathbb{P}[\xi_i = 1] = \gamma \quad \mathbb{P}[\xi_i = -1] = 1 - \gamma. \tag{14}$$

Note that $\mathbb{E}[\xi_i] \overset{\text{def}}{=} \mu = 2\gamma - 1$ and $\mathbb{V}[\xi_i] \overset{\text{def}}{=} \sigma^2 = 1 - \mu^2$. We then consider the sequence of i.i.d. random variables $X_1, \cdots, X_n$ where $X_i = (\xi_i - \mathbb{E}[\xi_i])$. We denote the sum of X_i as

$$M_n = \sum_{i=1}^{n} X_i. \tag{15}$$

We then apply Theorem 6.2 to prove an anti-concentration bound for the martingale given above.

Lemma 6.3 (Anti-Concentration)**.** *Consider the partial sums M_t defined in (15) for $t \in [n]$. There exists a constant C such that for all $\delta \in (0, O(1))$ and $n > \Omega\left(\log(1/\delta) \cdot \left(\frac{1+\mu}{\sigma} \right)^2 \right)$ we have*

$$\mathbb{P}\left[\exists t \in [n] \text{ s.t. } M_t \geq C\sigma\sqrt{t\log(\log(n)/\delta)} \right] \geq \delta.$$

Proof. By Theorem 6.2, we know that there exists constants C_1, C_2, C_3 and large N such that for all $m > N \cdot \left(\frac{1+\mu}{\sigma}\right)^2$ and $x \in \left[1, C_3 \sqrt{m} \frac{\sigma}{1+\mu}\right]$, we have

$$\mathbb{P}\left[\sum_{i=1}^{m} X_i \geq C_1 \sqrt{m} \sigma x\right] \geq e^{-C_2 x^2}.$$

Rather than consider every possible $t \in [n]$, we consider $j \in \left\{m_\delta, m_\delta^2, \cdots, m_\delta^{\lfloor \log_{m_\delta}(n) \rfloor}\right\}$ where $m_\delta \in \mathbb{N}$ and $m_\delta > m \log(1/\delta)$. We then have for a constant C that

$$\mathbb{P}\left[\exists t \in [n] \text{ s.t. } M_t \geq C\sigma \sqrt{t \log(1/\delta)}\right] \geq \mathbb{P}\left[\exists j \in \left[\lfloor \log_{m_\delta}(n) \rfloor\right] \text{ s.t. } M_{m_\delta^j} \geq C\sigma \sqrt{m_\delta^j \log(1/\delta)}\right]$$

$$= \sum_{j=1}^{\lfloor \log_{m_\delta}(n) \rfloor} \mathbb{P}\left[M_{m_\delta^j} \geq C\sigma \sqrt{m_\delta^j \log(1/\delta)} \,\middle|\, M_{m_\delta^\ell} \leq C\sigma \sqrt{m_\delta^\ell \log(1/\delta)} \quad \forall \ell < j\right]$$

$$\geq \sum_{j=1}^{\lfloor \log_{m_\delta}(n) \rfloor} \mathbb{P}\left[M_{m_\delta^j} \geq C\sigma \left(\sqrt{m_\delta^j \log(1/\delta)} + \sqrt{m_\delta^{j-1} \log(1/\delta)}\right)\right]$$

$$= \sum_{j=1}^{\lfloor \log_{m_\delta}(n) \rfloor} \mathbb{P}\left[M_{m_\delta^j} \geq C\sigma \left(1 + 1/\sqrt{m_\delta}\right) \sqrt{m_\delta^j \log(1/\delta)}\right]$$

$$\geq \sum_{j=1}^{\lfloor \log_{m_\delta}(n) \rfloor} \mathbb{P}\left[M_{m_\delta^j} \geq 2C\sigma \sqrt{m_\delta^j \log(1/\delta)}\right]$$

Thus, we set $C = \frac{C_1}{2\sqrt{C_2}}$ and then for any δ such that $\sqrt{C_2} < \sqrt{\log(1/\delta)} < C_3 \sqrt{C_2} \sqrt{m_\delta} \frac{\sigma}{1+\mu}$, we have

$$\mathbb{P}\left[\exists t \in [n] \text{ s.t. } M_t \geq C\sigma \sqrt{t \log(1/\delta)}\right] \geq \lfloor \log_{m_\delta}(n) \rfloor \delta.$$

$\square$

Algorithm 5 Stopping Time Adversary $\mathcal{A}_{\varepsilon,\delta}$ with constant C

---

    for $i = 1, \cdots, k$ **do**

        $\mathcal{A}_{\varepsilon,\delta} = \mathcal{A}_{\varepsilon,\delta(C, Y_1, \cdots, Y_{i-1})}$ gives datasets $\{0,1\}$, parameter $(\varepsilon, 0)$ and $\mathtt{RR}_\varepsilon$ to $\mathtt{AdaptParamComp}$.

        $\mathcal{A}_{\varepsilon,\delta}$ receives $Y_i \in \{\top, \bot\}$.

        if $Y_i = \top$ **then**

            $X_i = \varepsilon$

        else

            $X_i = -\varepsilon$

        if $\sum_{j=1}^{i}\left(X_j - \varepsilon\frac{e^\varepsilon - 1}{e^\varepsilon + 1}\right) \geq C\left(\varepsilon\sqrt{t \log(\log(n)/\delta)}\right)$, **then**

            $\varepsilon_{i+1}, \cdots \varepsilon_k = 0$

        BREAK

We next use Theorem 6.3 to prove that we cannot have a bound like Theorem 2.10 in the adaptive privacy parameter setting, which uses the *stopping time adversary* given in Algorithm 5.

Proof of Theorem 6.1. Consider the stopping time adversary $\mathcal{A}_{\varepsilon,\delta_g}$ from Algorithm 5 for a constant C that we will determine in the proof. Let the number of rounds $k = n$ and $\varepsilon = 1/n$. In order to use Theorem 6.3 we define $\gamma = \frac{e^\varepsilon}{1+e^\varepsilon}$ from (14). Because we let ε depend on n, we have $\mu \equiv \mu_n = \frac{e^{1/n}-1}{e^{1/n}+1} = O(1/n)$ and $\sigma \equiv \sigma_n = 1 - \mu_n^2 = 1 - O(1/n^2)$ which gives $\frac{1+\mu_n}{\sigma_n} = \Theta(1)$. We then relate the martingale in (15) with the privacy loss for this particular adversary in AdaptParamComp($\mathcal{A}_{\varepsilon,\delta_g}, n, 0$) with view V who sets $X_t = \pm\varepsilon$ each round,

$$\sum_{j=1}^t \left(X_j - \frac{\mu_n}{n}\right) = \frac{1}{n}M_t \qquad \forall t \in [n].$$

Hence, at any round t if $\mathcal{A}_{\varepsilon,\delta_g}$ finds that

$$\frac{1}{n}M_t \geq C\left(\frac{1}{n}\sqrt{t\log(\log(n)/\delta_g)}\right) \tag{16}$$

then she will set all future $\varepsilon_i = 0$ for $i > t$. To find the probability that (16) holds in any round $t \in [n]$ we use Theorem 6.3 with the constant C from the lemma statement to say that (16) occurs with probability at least δ_g.

Assume that $\text{COMP}_{\varepsilon_g}$ is a valid privacy odometer and (13) holds. We then know that with probability at least $1 - \delta_g$ over $v \sim V^b$ where V^b is the view for AdaptParamComp($\mathcal{A}_{1/n,\delta_g}, n, b$)

$$|\text{Loss}(v)| \leq \text{COMP}_{\delta_g}(\varepsilon_1, 0, \cdots, \varepsilon_k, 0)$$

$$\implies \left|\sum_{i=1}^t \text{Loss}_i(v_{\leq i})\right| = t \cdot \frac{\mu_n}{n} + o\left(\frac{1}{n}\sqrt{t\log\left(\frac{\log(n)}{\delta_g}\right)}\right) \qquad \forall t \in [n]$$

But this is a contradiction given that the bound in (16) at any round $t \in [n]$ occurs with probability at least δ_g. $\qquad\square$

We now utilize the bound from Theorem 4.3 to obtain a concentration bound on the privacy loss.

Lemma 6.4. COMP_{δ_g} *is a valid privacy odometer for* $\delta_g \in (0, 1/e)$ *where* $\text{COMP}_{\delta_g}(\varepsilon_1, \delta_1, \cdots, \varepsilon_k, \delta_k) = \infty$ *if* $\sum_{i=1}^k \delta_i > \delta_g/2$ *and otherwise for any* $\beta > 0$,

$$\text{COMP}_{\delta_g}(\varepsilon_1, \delta_1, \cdots, \varepsilon_k, \delta_k) = \sum_{j=1}^k \varepsilon_j(e^{\varepsilon_j}-1)/2 + \sqrt{2\left(\sum_{i=1}^k \varepsilon_i^2 + \beta\right)\left(1 + \frac{1}{2}\log\left(\frac{\sum_{i=1}^k \varepsilon_i^2}{\beta} + 1\right)\right)\log(2/\delta_g)}.$$

Proof. We will follow a similar argument as in Theorem 5.1 where we use the same martingale $\widetilde{M}_k$ from (11). We can then directly apply Theorem 4.3 to get the following for any $\beta > 0$ with probability at least $1 - \delta_g/2$

$$\left|\widetilde{M}_k\right| \leq \sqrt{2\left(\sum_{i=1}^k \varepsilon_i^2 + \beta\right)\left(1 + \frac{1}{2}\log\left(\frac{\sum_{i=1}^k \varepsilon_i^2}{\beta} + 1\right)\right)\log(2/\delta_g)}$$

$\qquad\square$

We next need to determine what value of $\beta > 0$ to choose in the above result. If we were to set $\beta = \sum_{i=1}^k \varepsilon_i^2$, then we would get asymptotically close to the same bound as in Theorem 2.10, however, the ε_i are random variables, and their realizations cannot be used in setting β; further, we know from Theorem 6.1 that such a bound cannot hold in this setting.

We now give our main positive result for privacy odometers, which is similar to our privacy filter in Theorem 5.1 except that δ_g is replaced by $\delta_g/\log(n)$, as is necessary from Theorem 6.1. Note that the bound incurs an additive $1/n^2$ loss to the $\sum_i \varepsilon_i^2$ term that is present without privacy. In any reasonable setting of parameters, this translates to at most a constant-factor multiplicative loss, because there is no utility running any differentially private algorithm with $\varepsilon_i < \frac{1}{10n}$ (we know that if A is $(\varepsilon_i, 0)$-DP then $A(\mathbf{x})$ and $A(\mathbf{x}')$ for any pair of inputs have statistical distance at most $e^{\varepsilon_i n} - 1 < 0.1$, and hence the output is essentially independent of the input - note that a similar statement holds for $(\varepsilon_i, \delta_i)$-DP.)

Theorem 6.5 (Advanced Privacy Odometer). COMP_{δ_g} *is a valid privacy odometer for* $\delta_g \in (0, 1/e)$ *where* $\text{COMP}_{\delta_g}(\varepsilon_1, \delta_1, \cdots, \varepsilon_k, \delta_k) = \infty$ *if* $\sum_{i=1}^k \delta_i > \delta_g/2$, *otherwise if* $\sum_{i=1}^k \varepsilon_i^2 \in [1/n^2, 1]$ *then*

$$\text{COMP}_{\delta_g}(\varepsilon_1, \delta_1, \cdots, \varepsilon_k, \delta_k) = \sum_{i=1}^k \varepsilon_i\left(\frac{e^{\varepsilon_i} - 1}{2}\right) + 2\sqrt{\sum_{i=1}^k \varepsilon_i^2\left(1 + \log\left(\sqrt{3}\right)\right)\log(4\log_2(n)/\delta_g)}. \qquad (17)$$

and if $\sum_{i=1}^k \varepsilon_i^2 \notin [1/n^2, 1]$ *then* $\text{COMP}_{\delta_g}(\varepsilon_1, \delta_1, \cdots, \varepsilon_k, \delta_k)$ *is equal to*

$$\sum_{i=1}^k \varepsilon_i\left(\frac{e^{\varepsilon_i} - 1}{2}\right) + \sqrt{2\left(1/n^2 + \sum_{i=1}^k \varepsilon_i^2\right)\left(1 + \frac{1}{2}\log\left(1 + n^2\sum_{i=1}^k \varepsilon_i^2\right)\right)\log(4\log_2(n))/\delta_g)}. \qquad (18)$$

Proof. We again focus on a valid privacy odometer for $\{\delta_i\} \equiv 0$ and the martingale $\widetilde{M}_k$ from Equation (11). We then discretize the choices of $\beta \in \mathcal{D}_n \overset{\text{def}}{=} \{1/n^2, 2/n^2, 4/n^2, \cdots, 1/2, 1\}$ in Theorem 6.4, and then take a union bound over all $\beta \in \mathcal{D}_n$ to say that for the martingale $\widetilde{M}_k$ in (11) the following holds with probability at least $1 - \delta_g$ simultaneously over all $\beta \in \mathcal{D}_n$

$$|\widetilde{M}_k| \leq \sqrt{2\left(\sum_{i=1}^k \varepsilon_i^2 + \beta\right)\left(1 + \frac{1}{2}\log\left(\frac{\sum_{i=1}^k \varepsilon_i^2}{\beta} + 1\right)\right)\log(2\log_2(n^2)/\delta_g)}.$$

Thus, for each realization $\sum_{i=1}^k \varepsilon_i^2 \in [1/n^2, 1]$, we can select β to be the largest value in $\mathcal{D}_n$ that is just below $\sum_{i=1}^k \varepsilon_i^2$, i.e. $1 \leq \frac{\sum_{i=1}^k \varepsilon_i^2}{\beta} \leq 2$. This then gives the following bound with probability at least $1 - \delta_g/2$ when $\sum_{i=1}^k \varepsilon_i^2 \in [1/n^2, 1]$,

$$|\widetilde{M}_k| \leq \sum_{j=1}^k \sqrt{2\left(2\sum_{i=1}^k \varepsilon_i^2\right)\left(1 + \frac{1}{2}\log(2 + 1)\right)\log(4\log_2(n)/\delta_g)}.$$

For the bound given in (18), we set $\beta = 1/n^2$. Hence, we would have with probability at least $1 - \delta_g/2$ when $\sum_{i=1}^k \varepsilon_i^2 \notin [1/n^2, 1]$,

$$|\widetilde{M}_k| \leq \sum_{j=1}^k \sqrt{2\left(1/n^2 + \sum_{i=1}^k \varepsilon_i^2\right)\left(1 + \frac{1}{2}\log\left(1 + n^2\sum_{i=1}^k \varepsilon_i^2\right)\right)\log(4\log_2(n)/\delta_g)}.$$

20

□

In the above theorem, we only allow privacy parameters such that $\sum_{i=1}^{k} \varepsilon_i^2 \in [1/n^2, 1]$. This assumption is not too restrictive, since the output of a single ($\ll 1/n$)-differentially private algorithm is nearly independent of its input. More generally, we can replace $1/n^2$ with an arbitrary "granularity parameter" γ and require that $\sum_{i=1}^{k} \varepsilon_i^2 \in [\gamma, 1]$. When doing so, $\log_2(n^2)/\delta_g$ in (17) will be replaced with $\log_2(1/\gamma)/\delta_g$. For example, we could require that $\varepsilon_1 \geq \delta_g$, in which case we can choose $\gamma = \delta_g^2$, which would not affect our bound substantially.

Acknowledgements

References

[BNS+16] Raef Bassily, Kobbi Nissim, Adam D. Smith, Thomas Steinke, Uri Stemmer, and Jonathan Ullman. Algorithmic stability for adaptive data analysis. In *Proceedings of the 48th Annual ACM on Symposium on Theory of Computing*, STOC, 2016.

[DFH+15] Cynthia Dwork, Vitaly Feldman, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Aaron Leon Roth. Preserving statistical validity in adaptive data analysis. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing*, pages 117–126. ACM, 2015.

[DKM+06] Cynthia Dwork, Krishnaram Kenthapadi, Frank McSherry, Ilya Mironov, and Moni Naor. Our data, ourselves: Privacy via distributed noise generation. In *Advances in Cryptology-EUROCRYPT 2006*, pages 486–503. Springer, 2006.

[dlPKLL04] Victor H. de la Pea, Michael J. Klass, and Tze Leung Lai. Self-normalized processes: exponential inequalities, moment bounds and iterated logarithm laws. *Ann. Probab.*, 32(3):1902–1933, 07 2004.

[DMNS06] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating noise to sensitivity in private data analysis. In *TCC '06*, pages 265–284, 2006.

[DR16] Cynthia Dwork and Guy N. Rothblum. Concentrated differential privacy. *CoRR*, abs/1603.01887, 2016.

[DRV10] Cynthia Dwork, Guy N. Rothblum, and Salil P. Vadhan. Boosting and differential privacy. In *51th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2010, October 23-26, 2010, Las Vegas, Nevada, USA*, pages 51–60, 2010.

[ES15] Hamid Ebadi and David Sands. Featherweight PINQ. *CoRR*, abs/1505.02642, 2015.

[KOV15] Peter Kairouz, Sewoong Oh, and Pramod Viswanath. The composition theorem for differential privacy. In *Proceedings of the 32nd International Conference on Machine Learning, ICML 2015, Lille, France, 6-11 July 2015*, pages 1376–1385, 2015.

[KS14] S.P. Kasiviswanathan and A. Smith. On the 'Semantics' of Differential Privacy: A Bayesian Formulation. *Journal of Privacy and Confidentiality*, Vol. 6: Iss. 1, Article 1, 2014.

[LT91] M. Ledoux and M. Talagrand. *Probability in Banach Spaces: Isoperimetry and Processes*. A Series of Modern Surveys in Mathematics Series. Springer, 1991.

[MV16] Jack Murtagh and Salil P. Vadhan. The complexity of computing the optimal composition of differential privacy. In *Theory of Cryptography - 13th International Conference, TCC 2016-A, Tel Aviv, Israel, January 10-13, 2016, Proceedings, Part I*, pages 157–175, 2016.

[vdG02] Sara A van de Geer. *On Hoeffding's inequality for dependent random variables*. Springer, 2002.