Optimal Local Bayesian Differential Privacy over Markov Chains

Darshan Chakrabarti, Jie Gao, Aditya Saraf, Grant Schoenebeck, Fang-Yi Yu

November 27, 2020

Abstract

In the literature of data privacy, differential privacy is the most popular model. An algorithm is differentially private if its outputs with and without any individual's data are indistinguishable. In this paper, we focus on data generated from a Markov chain and argue that Bayesian differential privacy (BDP) offers more meaningful guarantees in this context. Our main theoretical contribution is providing a mechanism for achieving BDP when data is drawn from a binary Markov chain. We improve on the state-of-the-art BDP mechanism and show that our mechanism provides the *optimal* noise-privacy tradeoffs for any local mechanism up to negligible factors. We also briefly discuss a non-local mechanism which adds *correlated* noise. Lastly, we perform experiments on synthetic data that detail when DP is insufficient, and experiments on real data to show that our privacy guarantees are robust to underlying distributions that are not simple Markov chains.

1 Introduction

In the literature of privacy, differential privacy (DP) [2] is the most popular model. An algorithm is differentially private if its outputs with and without any individual's data are (nearly) indistinguishable. It guards against what an adversary learns about an agent *due directly* to that agent's participation. But this does not offer guarantee on how much an adversary can infer about personal data from the perturbed output. Specifically, with correlated data, DP's guarantee is insufficient, as understanding the correlation can bolster inference significantly [6]. Real world data exhibits several natural correlation structures. Social networks mediate interactions and influence which often lead to strongly correlated personal attributes. Similarly, spatial and temporal proximity lead to strong correlations in data from sensors recorded as discretized time series. Some examples include data from human mobility traces, power grids, health data from personal wearable devices, and US census data. In many applications the correlation structure can be learned from historical data and so should be assumed to be public knowledge. Let the *correlated advantage* be how much an adversary can infer about private data from the perturbed output, given explicit knowledge of the correlation structure. We want a privacy definition and mechanism that bounds the correlated advantage.

To incorporate correlated data, the Pufferfish framework, proposed by Kifer and Machanavajjhala [7], allows for robust privacy guarantees based on a specified data generation model and secret pairs. Bayesian differential privacy (BDP), initially proposed by Yang et al. [14], is an instantiation of the Pufferfish privacy framework and a strict generalization of differential privacy. The distinction between DP and BDP is most salient when comparing which adversaries they protect against. DP protects only against adversaries who know all but one tuple, while BDP simply quantifies over adversaries who know arbitrarily many tuples. As an example, suppose we are trying to protect Alice's time series data. Differential privacy guarantees *only* that someone who knows Alice's location at all but one time step will not learn very much by analyzing the sanitized data set (their posterior after seeing the sanitized data will not change much) [14]. BDP guarantees that, regardless of what the adversary knows about Alice's mobility data, they will not learn much by analyzing the sanitized data set.[1] Presumably, Alice is not interested in protecting her data only against an adversary who knows almost all of it. Analogously, our privacy guarantee should not depend on how many data points the adversary knows.

1.1 Our Contribution

Our main theoretical contribution is providing a mechanism for achieving BDP when the data is drawn from a binary Markov chain. We focus on *non-interactive* mechanisms, which return "sanitized" (i.e. noisy) estimates of the real database. These are called non-interactive since the sanitized databases can be queried offline without further privacy loss; in contrast, query-based mechanisms must be rerun with additional privacy losses for each query. In order to achieve *local* privacy guarantees, our primary mechanism adds independent noise to each tuple. Local privacy means that no centralized curator needs to have access to the agents' true data. Instead, the owners of each tuple can privately sanitize their entry before submission, which would be ideal for IoT settings. When privacy guarantees are framed in terms of trust or persuasive power, this property is extremely attractive.[2]

Our model is simple yet fundamental; we represent data correlations via Markov chains, and these preliminary results consider only binary state spaces. We assume that the data is *positively* correlated, as in the case of location data, where an individual is more likely to stay in the same location than leave (given a fine-enough time scale). We improve on the state-of-the-art BDP mechanism and show that our mechanism provides the *optimal* noise-privacy tradeoffs (among local/independent noise mechanisms) up to negligible factors. This is significant because the previous general results only provide a sufficient bound on the noise [8, 15]. The main challenge with finding an exact bound is describing how the privacy loss evolves through the Markov chain. We also consider a mechanism which adds *correlated* noise to the data, but find no additional improvement to noise-privacy tradeoffs. Lastly, we perform experiments on real and synthetic data. We first demonstrate that DP does not bound the correlated advantage, by providing a concrete, correlation aware attack that more than doubles the DP bound. Further, even on real data that is not entirely Markovian, a neural network adversary, using Long Short Term Memory (LSTM) models, cannot surpass our mechanism's privacy bounds, suggesting that our mechanism is *robust* to varying correlation structures in practice.

1.2 Related Work

There exists a large body of work regarding the limitation of differential privacy with correlated data.[3] As a result, there have been many new correlation-aware privacy definitions in recent years. Zhao et al. [15] consider a definition equivalent to BDP that they term *dependent differential privacy*. Their proofs show that many attractive properties of DP, like post-processing and composition guarantees, also hold for BDP. Liu et al. [8] consider a different privacy definition that they also term dependent differential privacy; however, their definition is quite distinct, as it does not imply DP [15]. Naim et al. [9] consider a privacy definition rooted in information theory, that is more applicable to Internet privacy.

As far as privacy preserving mechanisms, the original BDP paper considers a mechanism for the sum query, with data drawn from a Gaussian query model [14]. We model data generated from a Markov chain, so their mechanism does not apply. Song et al. [13] provide a very general mechanism

[1] Notice that when the data at each time step is independent, BDP reduces to DP; the adversary's knowledge of Alice's location at time x_i is irrelevant to their knowledge of Alice's location at time x_j.

[2] Though we mostly focus on the setting where every tuple belongs to one individual, our mechanism is still applicable to the conventional setting.

[3] See [1, 4, 6, 13–16] for examples in various contexts.

along with guarantees for any privacy definition in the Pufferfish framework. However, their model may be computationally intractable, and must be re-run (with additional privacy loss) for multiple queries. Zhao et al. [15] provides a very general reduction theorem which explains how ε-BDP is implied by a lower ε'-DP. This result, while very powerful, does not produce optimal noise-privacy tradeoffs, as we explore in section 3.

To the best of the authors' knowledge, this is the first paper with an explicit optimal mechanism for local, non-interactive privacy in correlated settings.

Roadmap In Section 2, we review some basic definitions, and clearly differentiate the semantic guarantees of DP and BDP. In Section 3, we describe the noise-privacy tradeoffs of our mechanism, provide high level proof sketches, and briefly discuss a correlated noise variant of our mechanism. The detailed proofs are put in the Appendix after conclusion. In Section 4, we run several experiments and compare with prior results. We show that ε-DP requires less noise than ε-BDP, and also experimentally evaluate how much privacy degrades when an adversary knows the data correlation structures. We then compare our noise-privacy tradeoffs with previous results, showing that we improve on the state-of-the-art. To round out our experiments, we evaluate our mechanism on real world heart rate data which may not be purely Markovian, showing that (1) the Markov chain modelling assumption cannot be exploited by a neural network adversary; (2) our privacy protection mechanism is robust to real world correlation patterns.

2 Preliminaries

Let $X = X_1, \ldots, X_n$ represent a time series, where $X_i \in \{0, 1\}$ corresponds to the value at the i^{th} timestep. The data curator would like to publish the time series but does not want anyone to be able to know, with high confidence, the exact value at any given moment. We now define the data generation model and provide formal privacy definitions for DP and BDP. Then, we explain the semantic difference between the two privacy definitions, and provide arguments in favor of BDP for our context. First, we define *negligible* functions, which we use to explain how our mechanism is optimal.

Definition 1. *A negligible function is a function $\mu : \mathbb{N} \to \mathbb{R}$ that is asymptotically bounded by every inverse polynomial.*

A common example is 2^{-n}, an inverse exponential. When we use negligible, we mean negligible in the length of the Markov chain.

2.1 Markov Chains

A sequence of random variables $(X_1, \ldots, X_n)$ is a *Markov chain* with state space Ω, initial distribution π_0, and transition matrix P if 1) the initial state X_1 is sampled from π_0, and 2) for all $x, x' \in \Omega$, all $t \geq 1$, and all events $H_t = \{X_1 = x_1, \ldots, X_{t-1} = x_{t-1}, X_t = x\}$, we have

$$\Pr[X_{t+1} = x' \mid H_t] = \Pr[X_{t+1} = x' \mid X_t = x] = P_{x,x'}$$

In this paper, we consider the special case where $\Omega = \{0, 1\}$. We use (Ω, π_0, P) to denote a Markov chain. We call a Markov chain *lazy* if $P(x, x) > 1/2$ for all $x \in \Omega$, meaning that the chain is more likely to stay in the current state than switch states.

A distribution π on Ω is a *stationary distribution* of the Markov chain (Ω, π_0, P) if $\pi = \pi P$. Thus, if X_1 is drawn from π, the marginal distribution of any state X_i is also given by π.

2.2 Differential Privacy

A randomized mechanism $\mathcal{M}$ is a function with domain $\mathcal{X}$ consisting of all possible input databases, and range S denoting all possible outputs.

Definition 2 (Dwork et al. [2]). *A randomized mechanism $\mathcal{M}$ is said to be* ε-differentially private *(*ε-DP*) if, for any databases x and y that differ in exactly one tuple (i.e. one data point),*

$$\sup_{s \in S} \frac{\Pr[\mathcal{M}(x) = s]}{\Pr[\mathcal{M}(y) = s]} \leq e^{\varepsilon}.$$

Databases x and y are said to be *neighbors* and ε is referred to as the privacy budget. Note that lower ε entails more privacy.

2.3 Bayesian Differential Privacy

For BDP, we now assume that the domain $\mathcal{X}$ of $\mathcal{M}$ is generated according to some probabilistic model; thus, it makes sense to consider, for $X \in \mathcal{X}$, $\Pr[X]$ or $\Pr[M(X) = s|X_i = x_i]$. We also explicitly introduce adversaries $A = A(i, K)$, where $i \in [n]$ denotes the tuple A is trying to infer and $K \subseteq [n] \setminus \{i\}$ denotes the tuples A already knows.

Definition 3. *The Bayesian Differential Privacy Loss (BDPL) of $\mathcal{M}$ with respect to A is defined as:*

$$\mathrm{BDPL}(A; \mathcal{M}) \stackrel{\text{def}}{=} \sup_{x_i, x_i', \mathbf{x}_K, s} \frac{\Pr[\mathcal{M}(X) = s|x_i, \mathbf{x}_K]}{\Pr[\mathcal{M}(X) = s|x_i', \mathbf{x}_K]}$$

In the above definition, x_i and x_i' correspond to two different values for tuple X_i. Also, note that the probability is taken over the mechanism and data generation process. It should be interpreted as the probability of the database being X, given x_i and $\mathbf{x}_K$, and then observing $\mathcal{M}(X) = s$.

Definition 4. *A randomized mechanism $\mathcal{M}$ is said to be ε-Bayesian differential private (ε-BDP) [14], if*

$$\sup_{A} \mathrm{BDPL}(A; \mathcal{M}) \leq e^{\varepsilon}$$

where the adversaries range over all possible i and K.

ε-DP is equivalent to requiring $\mathrm{BDPL}(A; \mathcal{M}) \leq e^{\varepsilon}$, for all adversaries $A = A(i, [n] \setminus \{i\})$, i.e all adversaries that know all but one tuple. It follows that ε-BDP is at least as strong as ε-DP.

2.4 Semantic Differences between DP and BDP

Differential privacy forms privacy guarantees *without* a model for how the data is generated. However, if an adversary has reasonable background knowledge regarding the data distributions, it may be the case that an ε-differentially private mechanism will produce an output that is disproportionately more likely (from the perspective of the adversary) given one of two neighboring datasets. This limitation of DP is well understood; Dwork et al. [2] reframes this limitation in terms of the *semantic* guarantee that DP provides. Suppose you are a medical researcher tasked with convincing users to divulge sensitive health data, which will then be published online. Differential privacy, according to the semantic interpretation, can be used as a *tool* to encourage participation. Individuals, who can only control their own participation in the study, know that they will receive minimal (privacy) harms *directly tied* to their participation in the study. Viewed this way, DP is an entirely end-user focused persuasive tool. However, an ethical researcher may not only like to persuade users to participate, but also to understand and limit the harms caused by the study itself. DP answers the question of whether a single user ought to participate; BDP answers the question of whether the study ought to be performed (in terms of the privacy "cost" of the study). Put another way, BDP persuades the researcher to publish a sanitized copy of their data, by more comprehensively limiting the harms to any study participant.

We stress that in this medical example, since the database consists of records that each belong to different individuals, DP can still be a reasonable choice (e.g. if the study could significantly help the participants, and the additional noise hinders utility). However, when considering data from a single individual, like mobility or heartbeat time series data, (standard) DP does not provide sufficient guarantees.[4] BDP provides much more meaningful guarantees than DP in these settings.

3 Optimal Mechanisms for Bayesian Differential Privacy

In this section, we introduce our main theoretical results and some high-level proof sketches. We consider a variant of the canonical randomized response mechanism, applied to data with correlations modeled by a Markov chain.

Suppose we have a dataset $X = (X_1, \dots, X_n)$ generated from a Markov chain on $\{0, 1\}$. For example, this data could be discretized heart rate data, indicating whether a person's heart rate is currently elevated or not. We assume the Markov chain is stationary and has transition matrix:

$$P = \begin{bmatrix} 1-q & q \\ r & 1-r \end{bmatrix}, \text{ where } q, r \in (0, 0.5)$$

The initial state is generated from the stationary distribution, $\pi = \left(\frac{r}{q+r}, \frac{q}{q+r}\right)$, so the marginals are given by π. Such a Markov chain is called *lazy*, since we are more likely to stay in the current state than transition. Our proposed mechanism, Algorithm 1, takes in database X, adds independent noise to each data point, and outputs sanitized database Z. We perturb a 0 state to a 1 state with

---

Algorithm 1: Independent Noise Mechanism

Input: A sequence of data $(x_1, \dots, x_n)$ sampled from Markov chain (Ω, π, P), and noise levels $\rho_0, \rho_1 \in [0, 1/2)$

Result: Sanitized data $(z_1, \dots, z_n)$.

for $i = 1, \dots, n$ **do**

Generate z_i such that $\Pr[Z_i = z \mid X_i = x] = B_{x,z}$ independently, where

$$B = \begin{bmatrix} 1-\rho_0 & \rho_0 \\ \rho_1 & 1-\rho_1 \end{bmatrix}$$

end

---

probability ρ_0 and a 1 state to 0 with probability ρ_1. We use two separate noise levels to account for cases where one state is much more likely than the other; we can add less noise to the more popular state to reduce the total expected noise significantly. Notice also that this mechanism can be implemented in a distributed fashion, to guarantee local privacy. We want this mechanism to satisfy ε-BDP, and our goal is to determine ε in terms of parameters ρ_0, ρ_1, q and r.

BDP requires us to protect against attackers who might know any number of tuples. To simplify our work, we will show later that the most ignorant adversary requires the most noise to protect against.[5] This result, while somewhat counterintuitive, matches other cases with positive data

[4] There are definitions such as group differential privacy [2] that apply to this setting, but they require significantly more noise. See [13] for a more thorough analysis.

[5] This in turn implies that DP, which assumed a fully informed adversary, will not provide a strong privacy guarantee.

correlations [14]. So, by Theorem 2, ε-BDP can be reduced to the following conditions:

$$\max_{i\in[n],\mathbf{z}\in\{0,1\}^n} \frac{\Pr(\mathbf{Z}=\mathbf{z}|X_i=0)}{\Pr(\mathbf{Z}=\mathbf{z}|X_i=1)} \le e^{\varepsilon} \tag{1}$$

$$\max_{i\in[n],\mathbf{z}\in\{0,1\}^n} \frac{\Pr(\mathbf{Z}=\mathbf{z}|X_i=1)}{\Pr(\mathbf{Z}=\mathbf{z}|X_i=0)} \le e^{\varepsilon} \tag{2}$$

Where $\mathbf{Z}=\mathbf{z}$ denotes the event that $Z_1=z_1,\ldots,Z_n=z_n$. When it is clear from context, we will sometimes use $\Pr[z_{i:j}|x_k]$ as a short form of $\Pr[Z_{i:j}=z_{i:j}|X_k=x_k]$. We will often refer to Equation 1 and Equation 2 as *likelihood ratios*, as they measure the probability of the data given competing hypotheses $X_i=0$ and $X_i=1$.

We can view the original and sanitized data as a Hidden Markov Model (HMM), with hidden states X, observed states Z, transition matrix P and emission matrix B. We are interested in computing $\Pr[\mathbf{z}|x_j]$. Given concrete q,r,ρ_0,ρ_1, this can be solved via the Forwards-Backwards algorithm, but we are interested in providing a closed form with these as variables. The theorem below is our main technical result, providing a tight, closed form bound.

Theorem 1. *Let $0<q,r,\rho_0,\rho_1<0.5$. For all $n\ge 1$, all $\mathbf{z}\in\{0,1\}^n$, and all $i\in[n]$,*

$$\begin{aligned}
&\frac{\Pr[\mathbf{Z}=\mathbf{z}\mid X_i=0]}{\Pr[\mathbf{Z}=\mathbf{z}\mid X_i=1]} \le \frac{a^2}{cd}, \text{ where}\\
&a=\big((1-q)^2(1-\rho_0)^2-2(1-q-r-qr)(1-\rho_0)\rho_1\\
&\qquad+(1-r)^2\rho_1^2\big)^{\frac{1}{2}}+(1-\rho_0)(1-q)-\rho_1(1-r)\\
&c=2r\rho_1,\quad d=2r(1-\rho_0).
\end{aligned}$$

Moreover, the maximum happens at $\mathbf{z}=\mathbf{0}$ and $i\approx n/2$. With these parameters, the bound is tight up to negligible factors (in the length of the Markov chain).

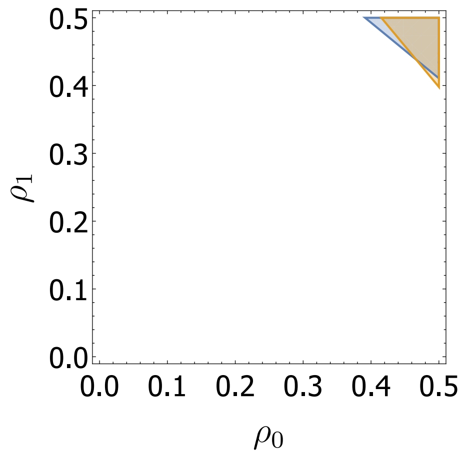
The theorem says that we achieve (nearly) optimal noise-privacy tradeoffs, for our data model and independent noise algorithm. By symmetry, the likelihood ratio for Equation 2 simply swaps q with r and ρ_0 with ρ_1 in the theorem. Notice that there are infinitely many solutions for ρ_0 and ρ_1 for fixed ε,q,r. In Figure 1 we plot the regions of feasible noise parameters for a fixed Markov chain (each region corresponds to one of the BDP constraints). The figure suggests a linear relationship between ρ_0 and ρ_1. Indeed, this can be rigorously shown, and implies that any linear objective function (like the expected noise) can be minimized by looking only at a constant number of points. We omit the exact forms of these linear constraints in this version.

Corollary 1. *For a fixed privacy budget ε and known Markov chain parameters q,r, to minimize the expected noise (i.e. $\rho_0\cdot\Pr[X_i=0]+\rho_1\cdot\Pr[X_i=1]$), one can determine the optimal ρ_0 and ρ_1 in constant time, up to negligible error in n.*

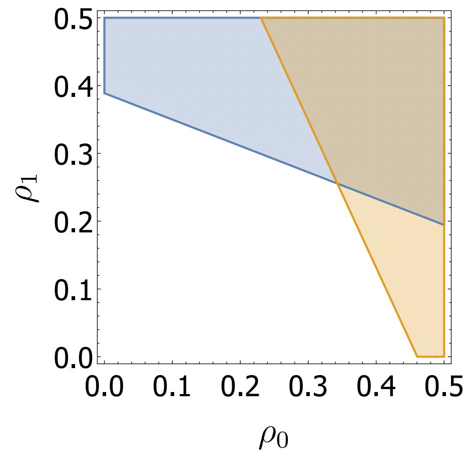
For many of our comparative discussions, we focus on the special case of a symmetric Markov chain, where the states have identical transition probabilities and noise levels. We get the following simplified results:

Corollary 2. *Assume the Markov chain is symmetric. Let $\theta=q=r$ and $\rho=\rho_0=\rho_1$. For all $n\ge 1$, all $z\in\{0,1\}^n$, and all $i\in[n]$:*

$$\begin{aligned}
&\frac{\Pr[z\mid X_i=0]}{\Pr[z\mid X_i=1]} \le \frac{1-\rho}{\rho}\left(\frac{a}{c}\right)^2, \text{ where}\\
&a=\sqrt{\theta^2+(1-2\theta)(1-2\rho)^2}+(1-\theta)(1-2\rho)\\
&c=2\theta(1-\rho)
\end{aligned}$$



(a)



(b)

Figure 1: Noise Constraints for BDP, on a Markov chain with $q = 0.2$, $r = 0.35$, and $\varepsilon = 0.5$ (a) or $\varepsilon = 2$ (b). Each plane represents one of the constraints; the intersection represents the feasible set of noise parameters.

Moreover, given θ, ρ and n, the maximum happens at $\mathbf{z} = \mathbf{0}$ and $i = \lfloor n/2 \rfloor$. Further, given θ, for a desired privacy level ε, it suffices to set:

$$\rho \geq \frac{4 + \theta(\theta e^{\varepsilon} - 2) - \sqrt{\theta^2 e^{\varepsilon}(4 + \theta(\theta e^{\varepsilon} - 4))}}{8 + 2\theta(\theta e^{\varepsilon} + \theta - 4)}$$

This closed form is not exact; to see how much this approximation loses, see Figure 3.

3.1 Proof Sketch for Theorem 1

We prove Theorem 1 in four parts. First, we write the likelihood ratio (LR) in terms of modified α, β recurrences for HMMs. Next, we show that for any fixed index i, the value of $\mathbf{z}$ that maximizes the LR is $\mathbf{z} = \mathbf{0}$, i.e. the all-zero database. Then, we find a closed form for the α and β recurrences when $\mathbf{z} = \mathbf{0}$ via matrix diagonalization. This closed form allows us to prove the bound in the theorem. Lastly, we show that for $\mathbf{z} = \mathbf{0}$, the LR is maximized with $i \approx n/2$, and that our bound is tight up to inverse exponential factors. We adopt this approach because a closed form of the LR in terms of the parameters is hard to compute for arbitrary $\mathbf{z}$.

Because the algorithm can be viewed as a hidden Markov model, the likelihood ratio can be computed by a slight variant of the standard Forwards-Backwards decomposition [12]. Given an outcome $\mathbf{z} \in \{0, 1\}^n$ we can define the forward probability from state $1 \leq t \leq n$ as:

$$\alpha_t(x; \mathbf{z}) := \Pr[\mathbf{Z}_{1:t} = \mathbf{z}_{1:t} \mid X_t = x] \tag{3}$$

Similarly we define the backward probability as:

$$\beta_t(x; \mathbf{z}) := \Pr[\mathbf{Z}_{t+1:n} = \mathbf{z}_{t+1:n} \mid X_t = x] \tag{4}$$

To simplify notation, we sometimes omit the outcome $\mathbf{z}$ when there is no ambiguity. With the forward and backward probability, we can compute the likelihood ratio easily:

$$\frac{\Pr[\mathbf{Z} = \mathbf{z} \mid X_i = 0]}{\Pr[\mathbf{Z} = \mathbf{z} \mid X_i = 1]} = \frac{\alpha_i(0; \mathbf{z})}{\alpha_i(1; \mathbf{z})} \cdot \frac{\beta_i(0; \mathbf{z})}{\beta_i(1; \mathbf{z})}$$

Therefore, to maximize the likelihood ratio, it suffices to maximize the above α and β ratios independently. The α and β terms can be written recursively as[6]:

$$\begin{aligned}\alpha_{i+1}(x;\mathbf{z}) &= \alpha_i(0)\,B_{x,z_{i+1}}\,P_{x,0} + \alpha_i(1)\,B_{x,z_{i+1}}\,P_{x,1}\\ \beta_i(x;\mathbf{z}) &= \beta_{i+1}(0)B_{0,z_{i+1}}P_{x,0} + \beta_{i+1}(1)B_{1,z_{i+1}}P_{x,1}\end{aligned}$$

where P and B are the transition and noise matrices. Thus, to compute a closed form of the ratios, we must work with a concrete $\mathbf{z}$.

Lemma 1. *For all $i \in [n]$, the likelihood ratio is maximized when $\mathbf{z} = \mathbf{0}$. That is, for all $\mathbf{z} \in \{0,1\}^n$:*

$$\frac{\Pr[\mathbf{Z}=\mathbf{z}|X_i=0]}{\Pr[\mathbf{Z}=\mathbf{z}|X_i=1]} \leq \frac{\Pr[\mathbf{Z}=\mathbf{0}|X_i=0]}{\Pr[\mathbf{Z}=\mathbf{0}|X_i=1]}.$$

The proof can be found in Appendix A.1. Since we are interested in computing the maximum likelihood ratio, over all $\mathbf{z}$ and i, the above lemma shows that we can restrict attention to $\mathbf{z} = 0$. To compute the theorem upper bound, we use the following lemma:

Lemma 2. *With $\mathbf{z} = \mathbf{0}$, for all $i \in [n]$:*

$$\frac{\alpha_i(0;\mathbf{z})}{\alpha_i(1;\mathbf{z})} \leq \frac{a}{c} \text{ and } \frac{\beta_i(0;\mathbf{z})}{\beta_i(1;\mathbf{z})} \leq \frac{a}{d}$$

The proof can be found in Appendix A.2 and involves writing the α and β recurrences as matrices and diagonalizing for a closed form. However, the above lemma does not indicate how tight this bound is. To conclude the proof of the theorem, we show that a tuple that's roughly in the middle of the Markov chain is the easiest to attack, and that at this tuple, the bound we provide on the likelihood ratio is very tight.

Lemma 3.

$$\operatorname*{argmax}_{i\in[n]} \frac{\Pr[\mathbf{Z}=\mathbf{0}|X_i=0]}{\Pr[\mathbf{Z}=\mathbf{0}|X_i=1]} = \frac{n}{2} \pm \mathcal{O}(1)$$

Where $\mathcal{O}(1)$ refers to some constant that depends on the Markov chain parameters and noise parameters, but not on n, the length of the Markov chain. Further, with $i = n/2$, the bound in Theorem 1 is tight up to inverse exponential factors.

The proof can be found in Appendix A.3. With this, we've shown that our mechanism provides the optimal noise-privacy tradeoffs for BDP in our setting, up to negligible factors.

3.2 The Most Ignorant Adversary is Hardest to Protect Against

In Theorem 1, we consider the likelihood ratio with respect to the adversary who knows none of the tuples. In order to show that our mechanism satisfies BDP, we show that this most ignorant adversary is the hardest to protect against. Although this may seem counter-intuitive, we will show later than the noise required for DP (which protects against the most informed adversary) is significantly lower than the noise in Theorem 1. Here, we will show that the BDPL increases monotonically between these two extremes. We do so by showing that for any adversary A, an adversary A' who knows one fewer tuple cannot have lower privacy loss:

Theorem 2. *Let $A = A(i, K)$ be an arbitrary adversary and let $U = [n] \setminus (K \cup \{i\})$. K represents the known tuples and U the unknown tuples for A. Consider adversary $A' = A(i, K')$, where $K' = K \setminus \{j\}$, for arbitrary $j \in K$. Then:*

[6] The base cases are $\alpha_0(x;\mathbf{z}) = 1$ and $\beta_n(x;\mathbf{z}) = 1$ for all $x, \mathbf{z}$.

(a) *if j is not adjacent to U or i,* $\mathrm{BDPL}(A') = \mathrm{BDPL}(A)$*.*

(b) *if j is adjacent to U or i,* $\mathrm{BDPL}(A') > \mathrm{BDPL}(A)$*.*

Further, the maximum $\mathrm{BDPL}(A')$ *occurs with* $\mathbf{z}, \mathbf{x}_K, x_i = \mathbf{0}$.

In particular, this theorem implies that as the set U of unknown tuples grows, in order for each unknown tuple to impact (and increase) the likelihood ratio, the unknown tuples must form a contiguous set around tuple x_i. The adversary's uncertainty about x_i "spreads" and is amplified by U; the amount that $\mathbf{z}$ reveals is thus maximized when $U = [n] \setminus \{i\}$. The proof is fairly involved, but follows the same basic structure as the proof of Theorem 1. It's obvious that the BDPL will always be maximized with $x_i, \mathbf{x}_K, \mathbf{z} = \mathbf{0}$, since the Markov chain is lazy and the mechanism is more likely to keep a bit than flip it (this can be shown formally via induction, similar to Lemma 1). Note that via conditional independence we can write:

$$\begin{aligned}\mathrm{BDPL}(A) &= \frac{\Pr[\mathbf{z}|x_i, \mathbf{x}_K]}{\Pr[\mathbf{z}|\bar{x}_i, \mathbf{x}_K]} \\ &= \frac{\Pr[z_i|x_i]}{\Pr[z_i|\bar{x}_i]} \cdot \frac{\Pr[\mathbf{z}_U|x_i, x_K]}{\Pr[\mathbf{z}_U|\bar{x}_i, x_K]}\end{aligned}$$

We then come up with a recurrence for the second term and find a closed form. With this, we show that BDPL of an adversary with k unknown tuples contiguous with x_i is proportional to a function $h(k)$. Lastly, we numerically verify that $h(k)$ is increasing in k to complete the proof. The full proof can be found in Appendix A.4. For simplicity, we present the proof for the symmetric case, but the result holds in the asymmetric case as well.

3.3 Can Correlated Noise Help?

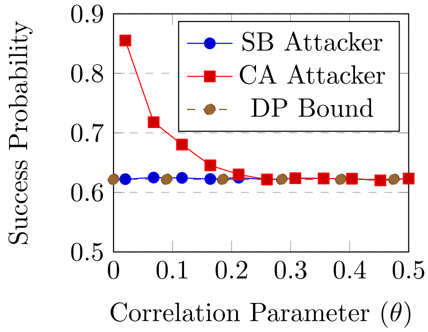
We briefly report results on a mechanism which uses *correlated* noise rather than independent noise to hide the data. We restrict attention to data $\mathbf{x} \in \{0,1\}^n$ generated from a *symmetric*, stationary, lazy Markov chain. The noise $\mathbf{y} \in \{0,1\}^n$ is also generated from a stationary Markov chain with parameters:

$$\begin{aligned}B &= \begin{bmatrix} 1-\rho_0 & \rho_0 \\ \rho_1 & 1-\rho_1 \end{bmatrix}, \text{ where } \rho_0 < \rho_1 \in (0,1) \\ &\qquad \text{and } \pi_B = \left(\frac{\rho_1}{\rho_0+\rho_1}, \frac{\rho_0}{\rho_0+\rho_1}\right)\end{aligned}$$

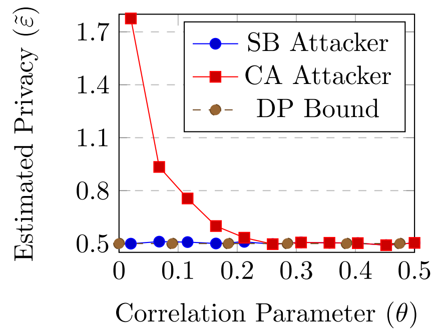
The mechanism constructs the sanitized database by XORing the noise chain and data chain elementwise. This mechanism includes independent noise as a special case (when $\rho_0 + \rho_1 = 1$). Notice that the expected noise for any bit is $\frac{\rho_0}{\rho_0+\rho_1}$, so if $\rho_0 \geq \rho_1$, then our mechanism is more likely to flip a bit than not. So, without loss of generality, we assume that $\rho_0 < \rho_1$ – each bit is more likely to be preserved than flipped.

With the further assumptions that $\rho_0 + \rho_1 \leq 1$ and $\theta < \rho_0$ (made for mathematical convenience), we computed the optimal noise for this mechanism (up to negligible factors). However, we found that the average noise was the same as our independent noise mechanism.[7] This shows that our independent noise mechanism has the optimal noise-privacy tradeoffs over not just all independent noise mechanisms, but also over this subset of correlated noise mechanisms.

[7] We found a closed form for ρ_0 and ρ_1 and computationally verified the expected noise equivalence. Proofs are omitted in this version.



(a)



(b)

Figure 2: Experimental results confirm that the Single Bit (SB) attacker is bounded by DP. However, the Correlation Aware (CA) attacker can dramatically violate the DP bounds on highly correlated data. The probability of successfully recovering a hidden bit is plotted in (a) and the estimated privacy costs of the reconstruction probabilities are plotted in (b). Here, the true $\varepsilon = 0.5$, and we generate data from a symmetric Markov chain, with transition parameter θ. Lower θ means more correlation.

4 Comparative and Experimental Evaluations

We first establish the necessity of a correlation-aware privacy definition, by demonstrating the efficacy of a concrete reconstruction attack on synthetic data that is protected by a differentially privacy mechanism. We then illustrate how our mechanism outperforms the state-of-the-art on BDP. Lastly, we perform experiments on Markov-like real data to demonstrate that a neural network (with a Long Short Term Memory architecture) cannot exploit any additional correlation structures to breach our BDP guarantees.

4.1 DP is Insufficient Against Correlated Advantage

We first analyze the relationship between ε and ρ for ε-DP. Our mechanism $\mathcal{M}$ achieves ε-DP if for all $\mathbf{z} \in \{0,1\}^n$ and all neighboring databases $\mathbf{x}, \mathbf{y} \in \{0,1\}^n$:

$$\frac{\Pr[\mathcal{M}(\mathbf{x}) = \mathbf{z}]}{\Pr[\mathcal{M}(\mathbf{y}) = \mathbf{z}]} \leq e^{\varepsilon}$$

Let $\mathbf{x}$ and $\mathbf{z}$ be arbitrary. Let k be the number of indices that $\mathbf{x}$ and $\mathbf{z}$ agree in. Note that $\Pr[\mathcal{M}(\mathbf{x}) = \mathbf{z}] = (1-\rho)^k \rho^{n-k}$. Since $\mathbf{y}$ can differ in exactly 1 bit from $\mathbf{x}$:

$$\frac{\Pr[\mathcal{M}(\mathbf{x}) = \mathbf{z}]}{\Pr[\mathcal{M}(\mathbf{y}) = \mathbf{z}]} \leq \max\left(\frac{\rho}{1-\rho}, \frac{1-\rho}{\rho}\right) \leq \frac{1-\rho}{\rho}$$

Thus, we get ε-DP when:

$$\rho \geq \frac{1}{e^{\varepsilon}+1} \tag{5}$$

However, ε-DP only considers the most knowledgeable attacker [8], and thus this noise level is insufficient to achieve ε-BDP. This is somewhat counter intuitive – one would expect that the most

[8] In DP, the adversary need not know the correlation structure. However, when the adversary knows all but one data point, BDP and DP are equivalent.[14]

knowledgeable adversary is also the most difficult to protect against. However, one can also think about the most knowledgeable adversary as having the least uncertainty about X_i. If the adversary knows the values of $\mathbf{x}_{-i}$, then they can already use the Markov chain to compute an approximate distribution of X_i. In fact, the *only* information they get from observing $\mathbf{z}$ is encoded in $\Pr[X_i|Z_i = z_i]$, since all $\mathbf{z}_{-i}$ are conditionally independent of X_i given X_{i-1}, X_{i+1}. On the other hand, the adversary who knows none of the data points gets a lot more information from $\mathbf{z}$. Their prior over X_i is simply π, but they could use the standard Forward-Backward inference algorithm to compute a much better informed posterior.

We also present some experimental results to highlight the practical ramifications of BDP over DP. First, consider an attacker with no information about the data generation. To attack variable X_i, the best the attacker can do is guess z_i (since $\rho < 1/2$). Call this the Single Bit (SB) Attacker. On the other hand, a Correlation Aware (CA) attacker can run the Forward-Backward algorithm and output the most likely x_i, given $\mathbf{z}$. We ran a comparison of these two attackers on synthetic data. The data consists of length 30 Markov chains. For a range of $\theta \in [0, 0.5]$, we generated 100 different databases from a symmetric Markov chain with transition parameter θ, and for each model, we generated 1000 different sanitized databases. We then calculated the frequency of correctly guessing hidden state X_{15} based on the sanitized database, for each type of attacker.[9]

To calibrate our expectations, let the success probability be p_s. For ε differential privacy, distinguishing the middle bit is the same as distinguishing $\mathcal{M}(x) = z$ from $\mathcal{M}(y) = z$, where x and y differ in the middle bit. So

$$\frac{p_s}{1-p_s} \leq \frac{\Pr[\mathcal{M}(x) = z]}{\Pr[\mathcal{M}(y) = z]} \leq e^\varepsilon \iff p_s \leq \frac{e^\varepsilon}{1+e^\varepsilon}$$

Which, with $\varepsilon = 0.5$, works out to roughly $p_s \leq 0.622$. Our experimental results confirm that this bounds the SB attacker. However, the CA attacker beats this bound significantly with highly correlated data. In Figure 2, we plot this and the privacy budget "charged" by each attacker, i.e. the ε that satisfies the previous equation for the observed q. Our results suggest that as the data are more correlated, we should add more noise to achieve the same privacy guarantees; this is exactly what we see with BDP, since ρ is inversely proportional to θ in Theorem 1.

4.2 Comparison to Previous BDP Results

Here, we briefly compare our results to that in [15] which uses a privacy definition equivalent to BDP. They provide very general theorems that reduce the problem of computing the noise required for ε-BDP to computing the noise for ε'-DP, where $\varepsilon > \varepsilon'$. In other words, they determine the "price" (in terms of the privacy budget) of data correlations. Theorem 3 in [15] states that ε-BDP $\Longleftarrow$ ε_3-DP with:

$$\varepsilon_3 = \varepsilon - 6 \ln \max_{x_{j+1}, x_j, x'_j \in \{0,1\}} \frac{\Pr[X_j = x_{j+1}|X_j = x_j]}{\Pr[X_j = x_{j+1}|X_j = x'_j]}$$

So, in our simple case, $\varepsilon_3 = \varepsilon - 6 \ln \frac{1-\theta}{\theta}$. Using Equation 5 to solve for ρ in terms of ε:

$$\rho \geq \frac{2(1-\theta)^6}{\theta^6 e^\varepsilon + (1-\theta)^6} \tag{6}$$

However, with this, they cannot provide ρ for ε below $6 \ln \frac{1-\theta}{\theta}$ (since ε_3 can't be negative). To fix this, they provide a stronger theorem which constructs a piecewise linear function.

[9] Our theoretical results show that the middle state maximizes the likelihood ratio, and so should be the easiest to distinguish.

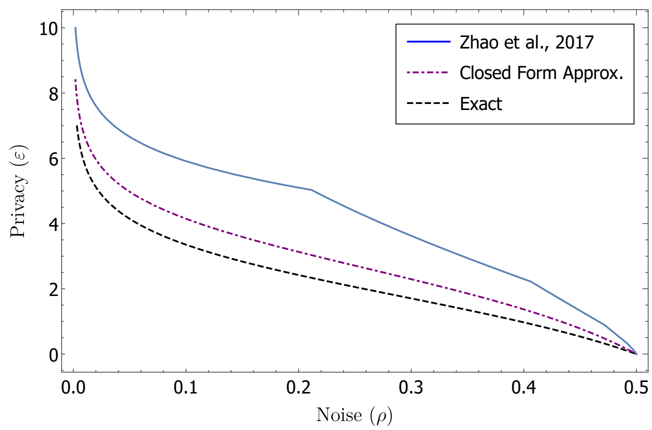


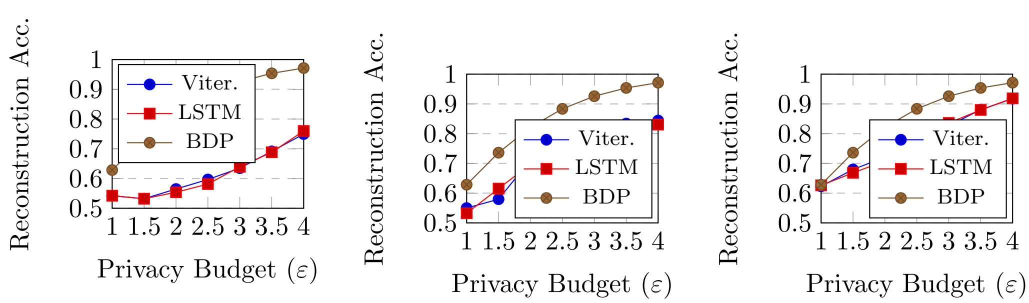
Figure 3: We plot the noise-privacy curves for a symmetric Markov chain with $\theta = 0.35$ and $n = 30$. The top line is the mechanism from [15], the middle line our closed form approximation and the bottom line our exact bound in Corollary 2. Lower noise-privacy curves are better.

Proposition 1 (Adapted from [15])**.** *Let* $0 < \theta < 0.5$*. Then,* ε*-BDP* $\Longleftarrow$ ε_6*-DP where:*

$$\varepsilon_6 = \max_{1 \le t \le n/2} \frac{\varepsilon - 6 \ln \max_{i,j,k \in \{0,1\}} \frac{\Pr[X_t = k | X_0 = j]}{\Pr[X_t = k | X_0 = i]}}{2t - 1}$$
$$= \max_{1 \le t \le n/2} \frac{\varepsilon - 6 \ln \frac{1 + (1-2\theta)^t}{1 - (1-2\theta)^t}}{2t - 1}$$

However, there's no obvious closed form for ρ in terms of ε and θ in this expression. Our work provides an approximate closed form and we have significantly better noise-privacy tradeoffs, as shown in Figure 3.

4.3 Experimental Evaluation on Heartbeat Data



(a) High correlation, synthetic data

(b) High correlation, real data

(c) Low correlation, real data

Figure 4: The reconstruction accuracy for both the Viterbi and the LSTM attackers is lower than our BDP bounds, on both synthetic (a) and real (b,c) data. There two real datasets correspond to a high correlation setting ($q = 0.0893, r = 0.1092, n = 26923$) and a low correlation setting ($q = 0.2384, r = 0.3831, n = 16859$). The synthetic data is chosen from the high correlation parameters.

We also experiment with real world data. Real world data is not necessarily drawn from a true

Markov chain – there are often more complex correlation models at play, which may be unknown even to the data curator. Thus, a natural question is whether our guarantees extended meaningfully to data that is only approximately Markovian. In other words, can an adversary exploit any (hidden) underlying correlation in approximately Markovian data to violate the expected privacy guarantees of our mechanism? We provide some evidence to the contrary.

We used a database containing heart rate data from several subjects while sleeping [3, 10]. We choose this setting because research suggests that heart rate data while sleeping exhibits many forms of correlation. During deep sleep, heart rates are relatively Markovian while during REM[10] sleep, they have long-time correlations (e.g. X_{10} influences X_{100}). And light sleep is a middle ground between these two extremes [11]. This is therefore an ideal setting to test whether (erroneously) modelling the data generation via Markov chains is problematic. The heart rate time series we use are around $15,000-25,000$ samples long, sampled at around 2Hz over 6 hours. We convert the heart rate data to binary data streams by clustering samples below and above the mean. To apply our algorithm, we first compute the empirical q and r from the binary data by observing the frequency of state transitions. Then, we sanitize the data via Algorithm 1.

To see if a more sophisticated adversary can defeat the bounds of our mechanism, we construct an neural network adversary with a Long Short-Term Memory (LSTM) [5] architecture that tries to reconstruct the private database given the sanitized database. We collate the data with the input being a window of sanitized bit and the output being the hidden bit in the middle of this window (the window size is a hyperparameter). Our neural architecture is simple. We embed the inputs in a four dimensional space, feed it through the LSTM, and then apply a linear layer to the last LSTM state, using a sigmoid for our output. Despite being simple, choosing an appropriate window size enables our architecture to detect more complex correlation structures than a simple HMM. We contrast the LSTM attacker with a Viterbi attacker, which simply runs the Viterbi algorithm on the perturbed sequence to predict the private database.

To calibrate our expectations, let the success probability be p_s. For ε-BDP, we can provide a very similar bound as with ε-DP:

$$p_s \leq \frac{e^{\varepsilon}}{\max(\frac{q}{r}, \frac{r}{q}) + e^{\varepsilon}}$$

The only difference from our earlier bound is due to the asymmetry of the Markov chain. Unlike DP however, we show that BDP does bound the correlation advantage.

As a sanity check, we first show that the LSTM attacker matches the Viterbi attacker on synthetic data constructed from a heavily correlated Markov chain. Then, we show that the LSTM attacker approximately matches the Viterbi attacker on two datasets – the first is heavily correlated, with $q = 0.0893, r = 0.1092$, while the second has very asymmetric correlation, with $q = 0.2384, r = 0.3831$. We report the reconstruction accuracy of the Viterbi attacker on the entire database, and the validation accuracy of the LSTM attacker using 10-fold cross validation. In Figure 4, we plot the reconstruction accuracy for both attackers as the privacy budget ε varies, in each setting. We also plot the theoretical bound that BDP guarantees, showing that both attackers are significantly below the BDP bound. Note that, unlike our DP experiments, the reconstruction accuracy *decreases* for a fixed ε with more correlated data. This is because BDP takes the worst-case probability over all possible sanitizations, and the difference between the worst-case and the average-case broadens significantly as the data gets more correlated (this is why the SB attacker was extremely close to the DP bound in the previous section).

Our results suggest that our mechanism is somewhat robust to data that's not strictly drawn from a Markov chain. In other words, despite sanitizing the data while assuming it is drawn from a Markov chain, the more sophisticated LSTM attacker was not able to get a significant advantage over the Viterbi attacker. This is despite the fact that the data exhibits more complex correlations [11], which an LSTM could learn.

[10] Rapid eye movement

5 Conclusion

We investigated privacy definitions for correlated data, and found that Bayesian differential privacy (BDP) provides meaningful guarantees in the presence of data correlations. We described an optimal mechanism for achieving BDP over binary Markov chains. Our mechanism is both local, and can thus be implemented in a distributed manner, and non-interactive, outputting a sanitized database. Finally, in a series of experiments, we showed the harms of using DP in correlated settings, the improvement of our mechanism over the previous state-of-the-art, and of the robustness of our mechanism to approximately Markovian heart rate data.

We briefly suggest three possible directions to build upon this work. First, a basic goal would be to extend our model from binary Markov chains to any Markov chain with a finite state space. Similarly, our results can be extended from Markov chains to Markov random fields, which would better model data correlations found in social networks. Second, at the end of Section 3, we briefly discussed a correlated noise mechanism (which would lose local privacy), but found that it didn't lower the expected noise necessary. However, further investigation could reveal that correlated noise wins on other metrics, like the mutual information between the input and sanitized databases. Also, it may be possible to get better noise-privacy tradeoffs by dropping the some of the assumptions made for mathematical convenience.

Lastly, at the end of Section 4, we discuss a more preliminary idea to reduce the amount of noise necessary with highly correlated data. BDP protects against the privacy loss when the mechanism produces the worst-case sanitization (i.e. when the all zero database is output), but the difference between this worst-case and the average-case can be quite large with sanitized data, as our experiments show. So one potential idea is to prevent the mechanism from outputting sanitizations close to the worst case; this would not provide local privacy, but may achieve better noise-privacy tradeoffs.

A Detailed Proofs

A.1 Proof of Lemma 1

Proof. The following two claims yield the lemma:

1. The ratio of forward probability, $\alpha_i(0;\mathbf{z})/\alpha_i(1;\mathbf{z})$ is maximized when $z_{1:t} = \mathbf{0}$.
2. The ratio of backward probability, $\beta_i(0;\mathbf{z})/\beta_i(1;\mathbf{z})$ is maximized when $z_{t+1:n} = \mathbf{0}$.

For the first claim, we go by induction on i. For the base case, $i = 1$, because ρ_0 and ρ_1 are smaller than 1/2:

$$\begin{aligned}\frac{\alpha_1(0;1)}{\alpha_1(1;1)} =& \frac{\Pr[Z_1 = 1 \mid X_1 = 0]}{\Pr[Z_1 = 1 \mid X_1 = 1]} = \frac{\rho_0}{1-\rho_1} \leq 1 \\ \leq& \frac{1-\rho_0}{\rho_1} = \frac{\Pr[Z_1 = 0 \mid X_1 = 0]}{\Pr[Z_1 = 0 \mid X_1 = 1]} = \frac{\alpha_1(0;0)}{\alpha_1(1;0)}.\end{aligned}$$

Now, suppose the claim is true for $i \geq 1$. Using basic probability theory, $\alpha_{i+1}(x)$ can be recursively computed from α_i:

$$\alpha_{i+1}(x;\mathbf{z}) = \alpha_i(0)\, B_{x,z_{i+1}}\, P_{x,0} \cdot + \alpha_i(1)\, B_{x,z_{i+1}}\, P_{x,1}$$

Where P and B are the transition and noise matrices, respectively. Therefore the α ratio is:

$$\begin{aligned}\frac{\alpha_{i+1}(0)}{\alpha_{i+1}(1)} =& \frac{\alpha_i(0)\, B_{0,z_{i+1}}\, P_{0,0} + \alpha_i(1)\, B_{0,z_{i+1}}\, P_{0,1}}{\alpha_i(0)\, B_{1,z_{i+1}}\, P_{1,0} + \alpha_i(1)\, B_{1,z_{i+1}}\, P_{1,1}} \\ =& \frac{B_{0,z_{i+1}}}{B_{1,z_{i+1}}} \cdot \frac{\frac{\alpha_i(0)}{\alpha_i(1)}(1-q) + q}{\frac{\alpha_i(0)}{\alpha_i(1)} r + (1-r)}\end{aligned}$$

To maximize this, note that the first term only depends on z_{i+1} and the second term only depends on the value of $\mathbf{z}_{1:i}$. Hence, we can maximize these two terms separately. The first term is identical to the base case, and is thus maximized by $z_{i+1} = 0$. For the second term, one can simply take the derivative with respect to the α ratio, and since $(1-q)(1-r) > qr$, this term is is increasing in $\frac{\alpha_i(0)}{\alpha_i(1)}$. Therefore, by our inductive hypothesis, the second term is maximized when $\mathbf{z}_{1:t} = \mathbf{0}$. Combining both terms, we've shown that the ratio $\frac{\alpha_{i+1}(0)}{\alpha_{i+1}(1)}$ is maximized when $\mathbf{z}_{1:t+1} = \mathbf{0}$. This completes the proof of our first claim.

The second claim is almost exactly symmetrical, except we must go by induction from $i = n$ downwards, since the β recurrence is:

$$\beta_i(x; \mathbf{z}) = \beta_{i+1}(0) B_{0,z_{i+1}} P_{x,0} + \beta_{i+1}(1) B_{1,z_{i+1}} P_{x,1}$$

The remaining steps are symmetrical. □

A.2 Proof of Lemma 2

Proof. Throughout this proof, let $\mathbf{z} = \mathbf{0}$ and let $i \in [n]$ be arbitrary. From our earlier derivations, the recurrence for α is:

$$\begin{bmatrix} \alpha_i(0) \\ \alpha_i(1) \end{bmatrix} = \begin{bmatrix} (1-\rho_0)(1-q) & (1-\rho_0)q \\ \rho_1 r & \rho_1(1-r) \end{bmatrix} \begin{bmatrix} \alpha_{t-1}(0) \\ \alpha_{t-1}(1) \end{bmatrix}$$

With base case $\alpha_0(x) = 1$. Now, solving this recurrence for arbitrary i just involves taking powers of the 2×2 matrix:

$$\begin{bmatrix} \alpha_i(0) \\ \alpha_i(1) \end{bmatrix} = \begin{bmatrix} (1-\rho_0)(1-q) & (1-\rho_0)q \\ \rho_1 r & \rho_1(1-r) \end{bmatrix}^t \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

To solve this, we diagonalize the matrix by finding the eigenvalues and eigenvectors. Let V's columns be the eigenvectors and Λ be the diagonal matrix of eigenvalues. Then:

$$\begin{bmatrix} \alpha_i(0) \\ \alpha_i(1) \end{bmatrix} = V \Lambda^i V^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

First, we compute the eigenvalues and eigenvectors. Let $x = (1-\rho_0)(1-q)$ and $y = \rho_1(1-r)$. Then:

$$\begin{aligned}
\lambda_1 &= 0.5\Big(x + y + \sqrt{(x+y)^2 - 4\rho_1(1-q-r)(1-\rho_0)}\Big) \\
\lambda_2 &= 0.5\Big(x + y - \sqrt{(x+y)^2 - 4\rho_1(1-q-r)(1-\rho_0)}\Big) \\
v_1 &= \begin{bmatrix} x - y + \sqrt{(x-y)^2 + 4qr(1-\rho_0)\rho_1} \\ 2r\rho_1 \end{bmatrix} \overset{\text{def}}{=} \begin{bmatrix} a \\ c \end{bmatrix} \\
v_2 &= \begin{bmatrix} x - y - \sqrt{(x-y)^2 + 4qr(1-\rho_0)\rho_1} \\ 2r\rho_1 \end{bmatrix} \overset{\text{def}}{=} \begin{bmatrix} b \\ c \end{bmatrix}
\end{aligned}$$

Putting this back into the recurrence:

$$\begin{aligned}
\begin{bmatrix} \alpha_i(0) \\ \alpha_i(1) \end{bmatrix} &= \begin{bmatrix} ac\lambda_1^i - bc\lambda_2^i & ab(\lambda_2^i - \lambda_1^i) \\ c^2(\lambda_1^i - \lambda_2^i) & ac\lambda_2^i - bc\lambda_1^i \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\
&= \begin{bmatrix} a(c-b)\lambda_1^i + b(a-c)\lambda_2^i \\ c(c-b)\lambda_1^i + c(a-c)\lambda_2^i \end{bmatrix}
\end{aligned}$$

Now we can compute the desired quantity:

$$\frac{\alpha_i(0)}{\alpha_i(1)} = \frac{a(c-b)\lambda_1^i + b(a-c)\lambda_2^i}{c(c-b)\lambda_1^i + c(a-c)\lambda_2^i} \to \frac{a}{c} \text{ as } i \to \infty \tag{7}$$

We now turn our attention to the β recurrence and perform very similar steps:

$$\begin{bmatrix}\beta_i(0)\\ \beta_i(1)\end{bmatrix} = \begin{bmatrix}(1-\rho_0)(1-q) & \rho_1 q\\ (1-\rho_0)r & \rho_1(1-r)\end{bmatrix}^{n-i}\begin{bmatrix}1\\ 1\end{bmatrix}$$

The eigenvalues of this matrix are the same, and for the eigenvectors, the only change is that c changes to $d \stackrel{\text{def}}{=} 2r(1-\rho_0)$. Thus, our final ratio is:

$$\frac{\beta_i(0)}{\beta_i(1)} = \frac{a(d-b)\lambda_1^{n-i} + b(a-d)\lambda_2^{n-i}}{d(d-b)\lambda_1^{n-i} + d(a-d)\lambda_2^{n-i}} \to \frac{a}{d} \text{ as } n-i \to \infty \tag{8}$$

We now claim that the α and β ratios from Equation 7 and Equation 8 are bounded above by their limit, i.e. that:

$$\frac{\alpha_i(0)}{\alpha_i(1)} \leq \frac{a}{c} \text{ and } \frac{\beta_i(0)}{\beta_i(1)} \leq \frac{a}{d}$$

To do so, we first claim that $a \geq c, d \geq 0 \geq b$. We omit the proof of this claim in this version. Now, we can rewrite Equation 7 as:

$$\frac{\alpha_i(0)}{\alpha_i(1)} = \frac{a + b\frac{a-c}{c-b}\left(\frac{\lambda_2}{\lambda_1}\right)^i}{c + c\frac{a-c}{c-b}\left(\frac{\lambda_2}{\lambda_1}\right)^i} \stackrel{\text{def}}{=} \frac{a+b'}{c+c'}$$

Now, notice that $a \geq a + b'$, since $b' \leq 0$ and $c \leq c + c'$ since $c' \geq 0$. Thus, a/c is an upper bound for the α ratio. We can do the exact same thing for the β ratio, since we also know that $a \geq d$, concluding the proof. □

A.3 Proof of Lemma 3

Proof. Recall from earlier that:

$$\frac{\Pr[\mathbf{Z} = \mathbf{z} \mid X_i = 0]}{\Pr[\mathbf{Z} = \mathbf{z} \mid X_i = 1]} = \frac{\alpha_i(0)}{\alpha_i(1)} \cdot \frac{\beta_i(0)}{\beta_i(1)}$$

Let $\sigma = \lambda_2/\lambda_1$, $A = a(c-b)$, $A' = a(d-b)$, $B = -b(a-c)$, $B' = -b(a-d)$, $C = c(c-b)$, $C' = d(d-b)$, $D = c(a-c)$, and $D' = d(a-d)$. Now, from Equation 7 and Equation 8, we know that:

$$\begin{aligned}
\frac{\alpha_j(0)}{\alpha_j(1)} &= \frac{a(c-b)\lambda_1^i + b(a-c)\lambda_2^i}{c(c-b)\lambda_1^i + c(a-c)\lambda_2^i} = \frac{A - B\sigma^i}{C + D\sigma^i}\\
\frac{\beta_i(0)}{\beta_i(1)} &= \frac{a(d-b)\lambda_1^{n-i} + b(a-d)\lambda_2^{n-i}}{d(d-b)\lambda_1^{n-i} + d(a-d)\lambda_2^{n-i}} = \frac{A' - B'\sigma^{n-i}}{C' + D'\sigma^{n-i}}\\
\frac{\Pr[\mathbf{z} \mid X_i = 0]}{\Pr[\mathbf{z} \mid X_i = 1]} &= \frac{AA' - AB'\sigma^{n-i} - A'B\sigma^i + BB'\sigma^n}{CC' + CD'\sigma^{n-i} + C'D\sigma^i + DD'\sigma^n}
\end{aligned}$$

Since only the two middle terms (in the numerator and denominator) involve i, we focus our attention there. We want to combine these terms, so notice that:

$$\begin{aligned}
AB' &= \frac{(a-d)(c-b)}{(a-c)(d-b)} \cdot A'B\\
CD' &= \frac{(a-d)(c-b)}{(a-c)(d-b)} \cdot C'D
\end{aligned}$$

Let $\gamma = \frac{(a-d)(c-b)}{(a-c)(d-b)}$. Then, we can rewrite the likelihood ratio as:

$$\frac{\Pr[\mathbf{Z} = \mathbf{z} \mid X_i = 0]}{\Pr[\mathbf{Z} = \mathbf{z} \mid X_i = 1]} = \frac{AA' - A'B(\gamma\sigma^{n-i} + \sigma^i) + BB'\sigma^n}{CC' + C'D(\gamma\sigma^{n-i} + \sigma^i) + DD'\sigma^n}$$

Now, since $A', B, C',$ and D are all positive (by our previous claims), by minimizing $\gamma\sigma^{n-i} + \sigma^i$, we maximize the ratio (all other terms are independent of i). Since γ is also positive, we can apply the AM-GM inequality, which says that $0.5(\gamma\sigma^{n-i} + \sigma^i) \geq \sqrt{\gamma\sigma^n}$. Note that the right hand side is a constant with respect to i. The AM-GM inequality also tells us that we get equality iff $\sigma^i = \gamma\sigma^{n-i}$, so doing so minimizes $\gamma\sigma^{n-i} + \sigma^i$. Taking logs:

$$\begin{aligned} i \log \sigma &= \log \gamma + (n-i) \log \sigma \\ i &= \frac{\log \gamma + n \log \sigma}{2 \log \sigma} = \frac{n}{2} + \frac{\log \gamma}{2 \log \sigma} = \frac{n}{2} \pm \mathcal{O}(1) \end{aligned}$$

We now argue that this shows that our theorem bound is tight up to negligible factors. Recall that the exact α ratio is:

$$\frac{\alpha_i(0)}{\alpha_i(1)} = \frac{a + b\frac{a-c}{c-b}\left(\frac{\lambda_2}{\lambda_1}\right)^i}{c + c\frac{a-c}{c-b}\left(\frac{\lambda_2}{\lambda_1}\right)^i}$$

Notice that because $\sigma = \lambda_2/\lambda_1 < 1$, the α ratio approaches a/c exponentially quickly as i increases; similarly, the β ratio approaches a/d exponentially quickly as $n-i$ increases. Since we know that the maximum index is on the order of $n/2$, we know that both ratios approach their limit up to some inverse exponential factors. (This also means that we can ignore the constant with again some inverse exponential loss.) In other words, with $z = \mathbf{0}$ and $i = n/2$:

$$\frac{\Pr[\mathbf{Z} = \mathbf{0} \mid X_{n/2} = 0]}{\Pr[\mathbf{Z} = \mathbf{0} \mid X_{n/2} = 1]} = \frac{a^2 + o(\sigma^{n/2})}{cd + o(\sigma^{n/2})}$$

□

A.4 Proof of Theorem 2

As mentioned in the body, we present the proof for a symmetric Markov chain here for simplicity, but the result holds for the asymmetric case as well. Recall that we assume $0 < \theta < 0.5$, where θ is the transition parameter (the probability that the state changes).

Proof. It's obvious that the BDPL will always be maximized with $x_i, \mathbf{x}_K, \mathbf{z} = \mathbf{0}$, since the Markov chain is lazy and the mechanism is more likely to keep a bit than flip it (this can be shown formally via induction, similar to Lemma 1). So, let $\mathbf{z}, x_i, \mathbf{x}_K$ all be set to zero here. The privacy losses can be written as:

$$\begin{aligned} BDPL(A) &= \frac{\Pr[z_i|x_i]}{\Pr[z_i|\bar{x}_i]} \cdot \frac{\Pr[\mathbf{z}_U|x_i, x_j, \mathbf{x}_{K'}]}{\Pr[\mathbf{z}_U|\bar{x}_i, x_j, \mathbf{x}_{K'}]} \\ BDPL(A') &= \frac{\Pr[z_i|x_i]}{\Pr[z_i|\bar{x}_i]} \cdot \frac{\Pr[\mathbf{z}_U, z_j|x_i, \mathbf{x}_{K'}]}{\Pr[\mathbf{z}_U, z_j|\bar{x}_i, \mathbf{x}_{K'}]} \end{aligned}$$

Notice that if j is not adjacent to U:

$$BDPL(A') = \frac{\Pr[z_i|x_i]}{\Pr[z_i|\bar{x}_i]} \cdot \frac{\Pr[\mathbf{z}_U|x_i, \mathbf{x}_{K'}]}{\Pr[\mathbf{z}_U|\bar{x}_i, \mathbf{x}_{K'}]} \cdot \frac{\Pr[z_j|x_i, \mathbf{x}_{K'}]}{\Pr[z_j|\bar{x}_i, \mathbf{x}_{K'}]}$$

The first two fractions are simply $BDPL(A)$ (since by conditional independence $\Pr[\mathbf{x}_U|x_i, x_j, \mathbf{x}_{K'}] = \Pr[\mathbf{x}_U|x_i, \mathbf{x}_{K'}]$). For the last fraction, notice first that if j is not adjacent to i, this last fraction is simply one, since we can ignore x_i via conditional independence. Otherwise, we use the following lemma.

Lemma 4. *Suppose j is $i-1$ or $i+1$, and let j' be $i-2$ or $i+2$ respectively. Then:*

$$\frac{\Pr[z_j|x_i, x_{j'}]}{\Pr[z_j|\bar{x}_i, x_{j'}]} = 2\frac{(1-\rho)(1-\theta)^2 + \rho\theta^2}{(1-\theta)^2 + \theta^2}$$

The lemma involves just a basic calculation, and so we omit the proof. Since j is not adjacent to U, we know that j is between i and a known tuple; thus the lemma applies and the BDPL increases.

Now, consider the case where j is adjacent to U. Without loss of generality, assume j is to the left of U, and write U as $U = U_L = \{i-k, \ldots, i-1\} \cup U_R\{i+1, \ldots, r\}$, the union of the portion to the left and right of x_i (so $j = i-k-1$). Note that U_L or U_R could be empty. By conditional independence:

$$\begin{aligned}
BDPL(A) &= \frac{\Pr[z_i|x_i]}{\Pr[z_i|\bar{x}_i]} \cdot \frac{\Pr[\mathbf{z}_{U_L}|x_i, x_j]}{\Pr[\mathbf{z}_{U_L}|\bar{x}_i, x_j]} \cdot \frac{\Pr[\mathbf{z}_{U_R}|x_i, x_r]}{\Pr[\mathbf{z}_{U_R}|\bar{x}_i, x_r]} \\
BDPL(A') &= \frac{\Pr[z_i|x_i]}{\Pr[z_i|\bar{x}_i]} \cdot \frac{\Pr[\mathbf{z}_{U_L}, z_j|x_i, x_{j-1}]}{\Pr[\mathbf{z}_{U_L}, z_j|\bar{x}_i, x_{j-1}]} \cdot \frac{\Pr[\mathbf{z}_{U_R}|x_i, x_r]}{\Pr[\mathbf{z}_{U_R}|\bar{x}_i, x_r]}
\end{aligned}$$

For comparative purposes, we ignore the first and last terms in the loss expressions. We also substitute $i-k-1$ for j:

$$\begin{aligned}
BDPL(A) &\propto \frac{\Pr[\mathbf{z}_{i-k:i-1}|x_i, x_{i-k-1}]}{\Pr[\mathbf{z}_{i-k:i-1}|\bar{x}_i, x_{i-k-1}]} \\
BDPL(A') &\propto \frac{\Pr[\mathbf{z}_{i-k-1:i-1}|x_i, x_{i-k-2}]}{\Pr[\mathbf{z}_{i-k-1:i-1}|\bar{x}_i, x_{i-k-2}]}
\end{aligned}$$

Now, let $\gamma_k(x, y) = \Pr[z_{i-k:i-1}, X_i = x, X_{i-k-1} = y]$. Then:

$$\begin{aligned}
BDPL(A) &\propto \frac{\gamma_k(x_i, x_{i-k-1}) / \Pr[x_i, x_{i-k-1}]}{\gamma_k(x_i, x_{i-k-1}) / \Pr[x_i, x_{i-k-1}]} \\
&= \frac{\gamma_k(x_i, x_{i-k-1})}{\gamma_k(\bar{x}_i, x_{i-k-1})} \cdot \frac{\Pr[\bar{x}_i, x_{i-k-1}]}{\Pr[x_i, x_{i-k-1}]} \\
BDPL(A') &\propto \frac{\gamma_{k+1}(x_i, x_{i-k-2})}{\gamma_{k+1}(\bar{x}_i, x_{i-k-2})} \cdot \frac{\Pr[\bar{x}_i, x_{i-k-2}]}{\Pr[x_i, x_{i-k-2}]}
\end{aligned}$$

Since $\mathbf{z} = \mathbf{0}$, we can write γ recursively as:

$$\begin{aligned}
\gamma_k(x, 0) &= (1-\theta)(1-\rho)\gamma_{k-1}(x, 0) + \theta\rho\gamma_{k-1}(x, 1) \\
\gamma_k(x, 1) &= \theta(1-\rho)\gamma_{k-1}(x, 0) + (1-\theta)\rho\gamma_{k-1}(x, 1)
\end{aligned}$$

For $x \in \{0, 1\}$. This exactly matches the β recurrence from our main proof, except for the base case. When $k = 0$, $\gamma_0(0, 0), \gamma_0(1, 1) = 0.5(1-\theta)$ and $\gamma_0(0, 1), \gamma_0(1, 0) = 0.5\theta$. The eigenvectors and eigenvalues are:

$$\lambda_1 = 0.5(\sqrt{(1-\theta)^2 - (1-2\theta)(1-2\rho)^2} + (1-\theta))$$
$$\lambda_2 = 0.5(-\sqrt{(1-\theta)^2 - (1-2\theta)(1-2\rho)^2} + (1-\theta))$$
$$v_1 = \begin{bmatrix} \sqrt{\theta^2 + (1-2\theta)(1-2\rho)^2} + (1-\theta)(1-2\rho) \\ 2\theta(1-\rho) \end{bmatrix} \overset{\text{def}}{=} \begin{bmatrix} a \\ c \end{bmatrix}$$
$$v_2 = \begin{bmatrix} -\sqrt{\theta^2 + (1-2\theta)(1-2\rho)^2} + (1-\theta)(1-2\rho) \\ 2\theta(1-\rho) \end{bmatrix} \overset{\text{def}}{=} \begin{bmatrix} b \\ c \end{bmatrix}$$
$$V = \begin{bmatrix} v_1 & v_2 \end{bmatrix} = \begin{bmatrix} a & b \\ c & c \end{bmatrix}, V^{-1} = \begin{bmatrix} c & -b \\ -c & a \end{bmatrix}$$

Thus, we get the following closed forms:

$$\begin{aligned} \gamma_k(0,0) &= (1-\theta)(ac\lambda_1^k - bc\lambda_2^k) + \theta ab(\lambda_2^k - \lambda_1^k) \\ &= \lambda_1^k a\big((1-\theta)c - \theta b\big) + \lambda_2^k b\big(\theta a - (1-\theta)c\big) \\ \gamma_k(1,0) &= \theta(ac\lambda_1^k - bc\lambda_2^k) + (1-\theta)ab(\lambda_2^k - \lambda_1^k) \\ &= \lambda_1^k a\big(\theta c - (1-\theta)b\big) + \lambda_2^k b\big((1-\theta)a - \theta c\big) \end{aligned}$$

Turning attention back to the BDPL, let $l = \lambda_2/\lambda_1$; note $l < 1$. We define the following functions:

$$f(k) \overset{\text{def}}{=} \frac{\Pr[X_i = 1, X_{i-k-1} = 0]}{\Pr[X_i = 0, X_{i-k-1} = 0]} = \frac{1-(1-2\theta)^{k+1}}{1+(1-2\theta)^{k+1}}$$
$$g(k) \overset{\text{def}}{=} \frac{\gamma_k(0,0)}{\gamma_k(0,1)} = \frac{a(c(1-\theta) - b\theta) + bl^k(a\theta - c(1-\theta))}{a(c\theta - b(1-\theta)) + bl^k(a(1-\theta) - c\theta)}$$

Thus, $BDPL(A) \propto f(k)g(k)$ and $BDPL(A') \propto f(k+1)g(k+1)$. From this closed form, we confirmed numerically that $h(k) \overset{\text{def}}{=} f(k)g(k)$ is increasing in k. This completes the proof. $\square$

References

[1] R. Chen, B. C. Fung, S. Y. Philip, and B. C. Desai. Correlated network data publication via differential privacy. *The VLDB Journal*, 23(4):653–676, 2014.

[2] C. Dwork, A. Roth, et al. *The algorithmic foundations of differential privacy*, volume 9. Now Publishers, Inc., 2014.

[3] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley. Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals. *circulation*, 101(23): e215–e220, 2000.

[4] X. He, A. Machanavajjhala, and B. Ding. Blowfish privacy: Tuning privacy-utility trade-offs using policies. In *Proceedings of the 2014 ACM SIGMOD international conference on Management of data*, pages 1447–1458, 2014.

[5] S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural Comput.*, 9(8):1735–1780, Nov. 1997. ISSN 0899-7667. doi: 10.1162/neco.1997.9.8.1735. URL https://doi.org/10.1162/neco.1997.9.8.1735.

[6] D. Kifer and A. Machanavajjhala. No free lunch in data privacy. In *Proceedings of the 2011 ACM SIGMOD International Conference on Management of data*, pages 193–204, 2011.

[7] D. Kifer and A. Machanavajjhala. Pufferfish: A framework for mathematical privacy definitions. *ACM Transactions on Database Systems (TODS)*, 39(1):1–36, 2014.

[8] C. Liu, S. Chakraborty, and P. Mittal. Dependence makes you vulnberable: Differential privacy under dependent tuples. In *NDSS*, volume 16, pages 21–24, 2016.

[9] C. Naim, F. Ye, and S. El Rouayheb. On-off privacy with correlated requests. In *2019 IEEE International Symposium on Information Theory (ISIT)*, pages 817–821. IEEE, 2019.

[10] C.-K. Peng, J. E. Mietus, Y. Liu, G. Khalsa, P. S. Douglas, H. Benson, and A. L. Goldberger. Exaggerated heart rate oscillations during two meditation techniques. *International Journal of Cardiology*, 70(2):101 – 107, 1999. ISSN 0167-5273. doi: https://doi.org/10.1016/S0167-5273(99)00066-2. URL http://www.sciencedirect.com/science/article/pii/S0167527399000662.

[11] T. Penzel, J. W. Kantelhardt, C.-C. Lo, K. Voigt, and C. Vogelmeier. Dynamics of heart rate and sleep stages in normals and patients with sleep apnea. *Neuropsychopharmacology*, 28(1): S48–S53, 2003.

[12] S. J. Russell and P. Norvig. *Artificial intelligence: a modern approach*. Pearson, 2020.

[13] S. Song, Y. Wang, and K. Chaudhuri. Pufferfish privacy mechanisms for correlated data. In *Proceedings of the 2017 ACM International Conference on Management of Data*, pages 1291–1306, 2017.

[14] B. Yang, I. Sato, and H. Nakagawa. Bayesian differential privacy on correlated data. In *Proceedings of the 2015 ACM SIGMOD international conference on Management of Data*, pages 747–762, 2015.

[15] J. Zhao, J. Zhang, and H. V. Poor. Dependent differential privacy for correlated data. In *2017 IEEE Globecom Workshops (GC Wkshps)*, pages 1–7. IEEE, 2017.

[16] T. Zhu, P. Xiong, G. Li, and W. Zhou. Correlated differential privacy: Hiding information in non-iid data set. *IEEE Transactions on Information Forensics and Security*, 10(2):229–242, 2014.