Analysis of Networks: Privacy in Bayesian Networks and Problems in Lattice Models

A thesis submitted in fulfilment of the requirements

for the degree of Doctor of Philosophy

The University of Melbourne, 2017

Zuhe Zhang

School of Mathematics and Statistics

The University of Melbourne

Australia

March 2017

To My Parents

Abstract

This thesis deals with differential privacy in Bayesian inference, probabilistic graphical models and information-theoretic settings. It also studies the expansion property and enumeration problems of certain subgraphs of networks.

The contributions of this thesis fall into three main categories:

(i) We establish results for Bayesian inference, providing a posterior sampling algorithm preserving differential privacy by placing natural conditions on the priors. We prove bounds on the sensitivity of the posterior to training data, which delivers a measure of robustness, from which differential privacy follows within a decision-theoretic framework. We provide bounds on the mechanism's utility and on the distinguishability of datasets. These bounds are complemented by a novel application of Le Cam's method to obtain lower bounds. We also explore inference on probabilistic graphical models specifically, in terms of graph structure. We show how the posterior sampling mechanism lifts to probabilistic graphical models and bound KL-divergence when releasing an empirical posterior based on a modified prior. We develop an alternate approach that uses the Laplace mechanism to perturb posterior parameterisations, and we apply techniques for released marginal tables that maintain consistency in addition to privacy, by adding Laplace noise in the Fourier domain. We also propose a maximum a posteriori estimator that leverages the exponential mechanism.

(ii) We generalize a prior work that considered differential privacy as a trade-off between information leakage and utility in noisy channels. By assuming certain symmetric properties of the graphs induced by the Hamming-1 adjacency relation on datasets, the authors showed the relation between utility and differential privacy. We prove the utility results still hold without *any assumption* on the structure of induced graphs. Our analysis applies to the graph of datasets induced by any symmetric relation, therefore is applicable to generalized notions of differential privacy.

(iii) In a different direction in graph analysis within statistical mechanics, we discover the relation between graph energy per vertex of a regular lattice and that of its clique-inserted lattice using spectral techniques. We obtain the asymptotic energy per vertex of 3-12-12 and 3-6-24 lattices. We derive the formulae expressing the number of spanning trees and dimer covering of the k-th iterated clique-inserted lattices in terms of those of the original one. We show that new families of expander networks can be constructed from the known ones by clique-insertion. We modify the transfer matrix method and use it to obtain upper and lower bound for the entropy of number independent sets on the 4-8-8 lattice. We show that the boundary conditions have no effect on the entropy constant. We also introduce a random graph model, where we study the annealed entropy of independent set per vertex. We show that the annealed entropy can be computed in terms of the largest eigenvalue (in modulus) of corresponding expected transfer matrix. Experiments suggest that this annealed entropy is highly correlated to the corresponding Shannon entropy.

Declaration

This is to certify that:

(1) Unless otherwise stated, this thesis comprises only my original work towards the PhD;

(2) Due acknowledgement has been made to all other material used; and

(3) The thesis is less than 100,000 words in length.

— Zuhe Zhang

Preface

This thesis reports on the body of work completed throughout the author's PhD research programme. Some of this research has been published with co-authors as follows:

A paper based on Chapter 3 has been accepted by *Journal of Machine Learning Research* (2017) under the title "Differential Privacy for Bayesian Inference through Posterior Sampling" with Christos Dimitrakakis, Blaine Nelson, Aikaterini Mitrokotsa and Benjamin I. P. Rubinstein.

A paper based on Chapter 4 has been published in *Proceedings of the 30th AAAI Conference on Artificial Intelligence* (AAAI'2016) under the title "On the Differential Privacy of Bayesian Inference" with Benjamin I. P. Rubinstein and Christos Dimitrakakis.

A paper based on Chapter 5 is being prepared for publication under the working title "Differential Privacy and Information Leakage" with Benjamin I. P. Rubinstein and Sanming Zhou.

A paper based on Chapter 6 has been published in *Journal of Statistical Mechanics: Theory and Experiment* (2013) under the title "Some Physical and Chemical Indices of Clique-inserted Lattices".

A paper based on Chapter 7 has been published in *Journal of Statistical Physics* (2014) under the title "Merrifield-Simmons Index and Its Entropy of the 4-8-8 Lattice".

A paper based on Chapter 8 is being prepared for publication under the title "The Number of Independent Sets in Randomly Triangulated Grid Graphs" with Yin Chen.

Acknowledgements

I will be forever grateful to my supervisors Benjamin I. P. Rubinstein and Sanming Zhou for their guidance, assistance and support through my PhD study. Without their support this thesis would not have been possible.

I would like to thank my co-authors: Christos Dimitrakakis, Xiaogang Liu, Blaine Nelson, Aikaterini Mitrokotsa and Yin Chen for the collaborations, special thanks go to Christos for all the help with research.

Thanks to Brendan McKay for suggesting the randomly triangulated grid graph. Without that, Chapters 8 would not have been possible.

My gratitude is also given to Zhibo Chen and Nicholas Witte for the suggestions in writing research papers.

I would also like to thank Richard Brak and Peter Forrester for their time and efforts in serving as members of my advisory panel.

My thanks go to Maggie as well for her support, company and proofreading of the thesis.

Lastly, I would dedicate this thesis to my parents for their unconditional love and support.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Background and Related Work

Network structures underlie a broad range of physical phenomena and information concepts, and the global properties of such structures play critical roles in their influence. This thesis explores networks from different viewpoints, using a range of mathematical tools of analysis. Building on the formal foundations of differential privacy, we explore how the independence structure of joint random variable models - encompassed as the graph structure of Bayesian networks - influence privacy and utility of inference mechanisms. The algebraic structure of graph relations enable us to prove links between differential privacy and information leakage in noisy-channel settings. In the area of statistical mechanics we demonstrate how to use mathematical tools in algebra and combinatorics to solve the enumeration problem of certain type of substructures on lattices or random networks which arise in physics and chemistry.

1.1.1 Differential Privacy

In an era of big data analysis and personal computing, collecting individual information is increasingly central to decision making across different domains. Meanwhile, the increase of privacy concerns prevents researchers from making full use of data. Past privacy breach reports by Fung et al. [2010], Narayanan and

Shmatikov [2008] have shown that various ad-hoc approaches failed to anonymize public records “linkage attacks" (to identify personal records by linking different databases). Therefore the concept of differential privacy, which was proposed by Dwork et al. [2006], quickly drew the attention of the theoretical computer science community by providing semantic guarantees of performing computation and releasing information about a sensitive dataset without revealing personal information about any individual.

The concept of differential privacy formalizes the idea that a “private” mechanism should not reveal whether any individual is included in the input or not, much less what their data are. It quantifies the privacy “cost” of an algorithm such that researchers can develop mechanisms which achieve a good trade-off between privacy and utility. Such requirements of privacy are of growing interest in the computer science and statistics communities due to the impact on individual privacy by real-world data analytics.

Dwork et al. [2006] proposed the first differentially-private mechanism, the *Laplace mechanism*, that is based on output perturbation through adding noise. The immediate follow-up work focused on the constructions of differential privacy preserving methods which have good utility by reducing the amount of noise injected [Nissim et al., 2007]. McSherry and Talwar [2007] proposed the *exponential mechanism* that releases a response with probability exponential in a utility function describing the usefulness of each response, with the best response having maximal utility. Other generic privatising mechanisms include Gaussian Dwork and Roth [2014], Bernstein Aldà and Rubinstein [2017] and more. Chaudhuri and Monteleoni [2008], Chaudhuri et al. [2011] proposed an approach that can be employed for privatising regularised empirical-risk minimization by adding a random term to the primal objectives. Rubinstein et al. [2012] proposed a set of privacy preserving classification methods using support vector machines with an output perturbation approach. Other learning algorithms including principal component analysis [Chaudhuri et al., 2012], the functional mechanism [Zhang et al., 2012] and trees [Jagannathan et al., 2009] have also been adapted to maintain differential privacy. Kifer and Machanavajjhala [2011] proved a no-free-lunch theorem, which defines non-privacy as a game, to argue that it is not possible to provide privacy and utility without making assumptions about how the data are generated.

They also proposed the Pufferfish framework that can be used to generate new private definitions that are customized to the requirements of a given application [Kifer and Machanavajjhala, 2012]. Inspired by the Pufferfish framework He et al. [2014] presented a class of privacy definition, called Blowfish privacy, that allows the data publisher to use a policy to specify the information that must be kept secret and the constraints that may be known about the data. Let us also mention that Kasiviswanathan and Smith [2008] defined the semantic privacy that provides the differential privacy guarantees in terms of the inferences drawn by a Bayesian adversary.

In statistics, Dwork and Lei [2009] made the first connection between (frequentist) robust statistics and differential privacy, developing mechanisms for the interquartile, median and $\boldsymbol{B}$-robust regression. Wasserman and Zhou [2010] introduced the concept of privacy as hypothesis testing where an adversary wishes to distinguish two datasets. Hall et al. [2013] studied differential privacy on Functional data. Dwork et al. [2015] studied how to guarantee the validity of statistical inference in adaptive data analysis.

Other areas where researchers have shown interesting relations with differential privacy include mechanism design from algorithmic game theory [Nissim et al., 2012], geometry [Hardt and Talwar, 2010] and information theory [Mir, 2012, Alvim et al., 2011a]. This is far from an exhaustive list. We refer the reader to the monograph on differential privacy [Dwork and Roth, 2014] and the reference therein for more details.

Our vision for differentially private mechanism for Bayesian inference is that they could be incorporated into probabilistic programming framework using systems techniques. Several systems have been developed to ease implementation of differentially-private mechanisms, with Barthe et al. [2016] providing an overview of contributions from Programming Languages. Dynamic approaches track privacy budget expended at runtime, typically through basic operations on data with known privacy loss, with the PINQ [McSherry, 2009, McSherry and Mahajan, 2010] and Airavat [Roy et al., 2010] systems being examples. These create a C# LINQ-like interface and a framework for bringing differential privacy to MapReduce, respectively. To complement PINQ and Airavat, Haeberlen et al. [2011] presented a design that is effective against covert channels. Haeberlen et al. [2011] presented GUPT that is secure against side-channel attacks. Fuzz [Reed and Pierce, 2010, Palamidessi and Stronati, 2012] offers a higher-order functional language whose static type system tracks sensitivity based on linear logic, so that differential privacy is guaranteed by typechecking.

Beyond academic interest, differentially-private mechanisms have also been applied in releasing or collecting aggregation information from data in government and commercial projects, such as the U.S. Census Bureau project called OnTheMap [Machanavajjhala et al., 2008], the RAPPOR project [Erlingsson et al., 2014] from Google and the iOS 10 update from Apple [Russell Brandom, 2016].

Privacy in Bayesian Networks Probabilistic graphical models have been used to preserve privacy. Zhang et al. [2014] learned a graphical model from data, in order to generate *surrogate data* for release. Note that their mechanism PrivBayes does not do Bayesian inference, the Bayesian network approach is to factorize a joint distribution in the frequentist model. Williams and McSherry [2010] fit a model to the response of private mechanisms to clean up output and improve accuracy. Xiao and Xiong [2012] similarly used Bayesian credible intervals to increase the utility of query responses.

Williams and McSherry [2010] improved the utility of differentially-private releases by calculating posteriors in a noisy measurement model. Beside their work, though Bayesian networks are widely used in different applications where privacy is important, there exists little research in private inference under the Bayesian setting until Dimitrakakis et al. [2014] first established conditions for differentially-private Bayesian inference. Dimitrakakis et al. (2014; 2017) introduced a differentially-private mechanism for Bayesian inference based on posterior sampling–a mechanism on which we build. Zheng [2015] considered further refinements. Wang et al. [2015] explored Monte Carlo approaches to Bayesian inference using the same mechanism, while Mir [2012] was the first to establish differential privacy of the Gibbs estimator [McSherry and Talwar, 2007] by minimizing risk bounds. Recently, Foulds et al. [2016] proposed an alternative to posterior sampling mechanism based on the Laplace mechanism and showed it is as asymptotically efficient as non-private posterior inference under general assumptions.

In this thesis, we work towards addressing the following challenges in the study of differential privacy under Bayesian inference:

1. How can differential privacy be accomplished in Bayesian statistical learning?
2. Can the existing Bayesian inference machinery provide a level of privacy?
3. Is there a relationship between a joint model's factorisation (conditional independence assumptions) and the level of privacy?

The first challenge has been studied in [Dimitrakakis et al., 2014], we extend the discussion in Chapter 3. The answers to the second and third are positive as discussed in Chapter 4.

Differential Privacy and Information Leakage From an information-theoretic perspective, any mechanism that releases a statistic leaks some information about the individual participants. Therefore, it is natural to consider the trade-off between information leakage and utility for privacy-preserving algorithms. Mir [2012] first formulated differential privacy in an information-theoretic framework. Duchi et al. [2013] provided information-theoretic bounds on mutual information and Kullback-Leibler divergence that depend on the privacy guarantees. This direction seeks to bridge a large community in communications that has formed around information-theoretic notions of privacy, to differential privacy.

1.1.2 Statistics of Certain Types of Subgraphs on Networks

Networks, lattices and molecule structures in the theory of engineering, computer science, statistical physics and chemistry are considered as graphs realized in the real world. Some substructures and invariants of graphs play an important role in these fields and the enumeration of these structures is a useful way to characterize networks and it is always a great challenge to obtain exact solutions or estimate related invariants. The reader may refer to the monographs by [Baxter, 1982] and [Borwein et al., 2013] that summarize such challenge in lattice models.

One of the oldest problem in this field appeared in Kirchoff's electrical networks theory, where spanning trees can be used to compute the current in networks Kirchhoff [1847]. Since then the enumeration of spanning trees has been widely studied [Greenhill et al., 2013, Lyons, 2005, Shrock and Wu, 2000, Teufl and Wagner, 2010].

Solving Lattice models such as Ising model or dimer model is a classical topic in statistical mechanics. It is sometimes related to the enumeration of certain subgraphs of lattices, which is also a topic of interest in chemistry. For instance, the number of perfect matchings in Pauling's resonant theory [Pauling, 1939] can be used to determine the Pauling bond order which is correlated with experimentally determined bond length of various benzenoid hydrocarbons [Pauling, 1980]. Fowler and Rushbrooke [1937] considered the same enumeration problem which was introduced as the dimer problem in order to describe the absorption of diatomic molecules on crystal surface. The diatomic molecules are modelled as rigid dimers each of which occupies two adjacent sites and no lattice site is covered by more than one dimer. After more than three decades, Fisher [1961], Kasteleyn [1963], Temperley and Fisherpp [1961] solved the dimer problem on plane quadratics lattices independently. Subsequently there have been many further developments dealing with the dimer problem of plane quadratic lattices with different boundary conditions. Cohn et al. [1996] provided a proof for the explicit expression of the number of perfect matchings on Aztec diamond. Sachs and Zernitz [1994] obtained the entropy constant of dimers of another type of finite plane quadratic lattices. In chemistry quite a few results have been published on this topic, especially in the study of Benzenoid hydrocarbons which are usually modelled as planar honeycomb lattices with different boundary condition. For details see the book Cyvin and Gutman [2013] and the references cited therein. Unlike the dimer problem, in the monomer-dimer problem and lattice gas model, the entropy constants are independent of the boundary conditions.

In statistical physics the two-dimensional gas model assumes that all of the gas molecules lie at the grid sites and only interact with their grid-neighbours. The grid is taken to be rigid and square, so the limit of partition function per vertex is called the "hard square constant". Baxter et al. [1980] were first to consider this problem. The model has also been studied on the triangular and hexagonal lattices Baxter [1982, 1999], Domb and Green [1972], Finch [1999], Pearce and Seaton [1988]. This problem has also been studied by mathematicians who are interested in the counting of independent sets [Neil J Calkin, 1998].

The monomer-dimer problem also originates from crystal physics where it has been used to model the behavior of systems of diatomic molecules (dimers) and single atoms absorbed on the surface of a crystal. This surface is represented as a lattice and is exposed to a gas consisting of monomers and dimers. The diatomic molecules are modelled as dimers each of which occupies two adjacent sites and no lattice site is covered by more than one dimer. The other lattice sites that are not covered by the dimers are regarded as occupied by monomers. The number of all possible monomer-dimer arrangements (or monomer-dimer coverings) is equal to the number of matchings of the lattice. In chemistry, Hosoya [1971] introduced number of matchings of a molecular graph as a topological index. Many early works were surveyed in the book by Gutman and Polansky [2012].

Other than enumerating substructures, computing algebraic invariants of the graphs of networks, lattices and molecule structures is another way to capture their behavior. Conversely some new algebraic invariants are inspired by concepts from other fields. One famous example is the introduction of graph energy. As the linear algebraic approach to approximate the solution of Schrödinger equation of conjugated hydrocarbons - a class of organic molecule as studied in Huckel molecular orbital theory [Coulson et al., 1978]. Inspired by Huckel molecular orbital theory, Gutman [2001] introduced the total π- electron energy of a molecular graph as the sum of obsolete values of the spectra of its adjacency matrix. This concept has been extended to general graphs. Today this has become a fruitful topic not only in mathematical chemistry but also in algebraic graph theory. Li et al. [2012] described the development of this field.

Another topic we study in this thesis is the expansion property. A network has a nice expansion property if it is both sparse and highly connected. It is known that the expansion property of a graph depends on its spectra. For details see the survey written by Hoory et al. [2006].

1.2 Thesis Outline and Main Contributions

This thesis consists of three main parts:

- Differential privacy on networks;

- Differential privacy and information leakage;

- Some problems in lattice models.

These three parts cover Chapter 3 through Chapter 8 with this introduction and concluding chapters in addition.

We study differential privacy in the Bayesian paradigm in Chapters 3 and 4, in which we wish to release the results of Bayesian inference on privacy-sensitive data.

In Chapter 3, we propose a posterior sampling algorithm that preserves differential privacy by placing conditions on the priors. We prove bounds on the sensitivity of the posterior to the data, which delivers a measure of robustness, from which differential privacy follows within a decision-theoretic framework. We provide bounds on the mechanisms utility and on the distinguishability of datasets. These bounds are complemented by a novel application of Le Cam's method to obtain lower bounds. We show that with the right choice of prior, Bayesian inference is both private and robust. Our results demonstrate that robustness and privacy appear to be deeply linked: not only can the same sufficient conditions achieve both privacy and robustness, but privacy can also imply robustness, and robustness implies privacy. This chapter is joint work with Christos Dimitrakakis, Aikaterini Mitrokotsa, Blaine Nelson and Benjamin Rubinstein [Dimitrakakis et al., 2017], and a follow up to the preliminary work of [Dimitrakakis et al., 2014]. Co-authors are responsible for the query model and dataset distinguishability part.

In Chapter 4, we explore inference on probabilistic graphical models in terms of graph structure. Our main contributions are four different algorithms for private Bayesian inference on probabilistic graphical models. These include two mechanisms for adding noise to the Bayesian posterior updates, either directly to the posterior parameters, or to their Fourier transform so as to preserve update consistency. We also utilise the posterior sampling mechanism introduced in Chapter 3, for which we prove bounds for the specific but general case of discrete Bayesian networks, and we introduce a maximum-a-posteriori private mechanism. Our analysis includes utility and privacy bounds, with a novel focus on the influence of graph structure on privacy. Worked examples and experiments with Bayesian naïve Bayes and Bayesian linear regression illustrate the application of our mechanisms. This chapter is joint work with Benjamin Rubinstein and Christos Dimitrakakis [Zhang et al., 2016].

Chapter 5 presents a significant generalization on the utility results in Alvim et al. [2011b,a], where the authors modelled a query system in terms of an information-theoretic channel and showed that differential privacy implies a bound on the min-entropy leakage. They also showed that the parameter of ϵ-differential privacy implies a bound on utility. By viewing the possible input databases as a graph whose nodes correspond to databases and whose adjacency is determined by the adjacency of the databases, the authors proved the bounds on a special class of symmetric graphs by manipulating the channel matrices. We show that the assumption on graph symmetry is redundant and these results can be generalized to arbitrary graphs. Given that the original symmetric assumption is quite restricted, our generalization can be considered as a significant improvement. This chapter is joint work with Benjamin Rubinstein and Sanming Zhou.

The second part comprises Chapters 6, 7 and 8, which investigate the statistics of certain combinatorial objects on both deterministic and random networks.

In Chapter 6, we recall the relationship between the spectra of an r-regular lattice and that of its clique-inserted lattice, and investigate the *graph energy* statistics. As an application, the asymptotic energies per vertex of the 3-12-12 and 3-6-24 lattices are computed. We also develop formulae expressing the numbers of spanning trees and dimer coverings of the k-th iterated clique-inserted lattices in terms of that of the original. Moreover, we show that new families of expander graphs can be constructed from known expanders by clique-inserting [Zhang, 2013].

In Chapter 7, we investigate the statistics of vertex independent sets on some (random) networks using the transfer matrix method. In this chapter, we first propose the concept of transfer multiplicity and the multi-step transfer matrices methods to study more complicated lattices where the single step transfer matrix approach as in [Neil J Calkin, 1998] is not compatible. We demonstrate our method on the 4-8-8 lattice by providing numerical results of the number of independent sets and a rigorous bound on its entropy. We also show that this entropy constant of a two dimensional lattice with free boundary condition is the same as the entropy constants of the corresponding cylindrical and toroidal lattices [Zhang, 2014].

In Chapter 8, we investigate the annealed entropy of independent set per site on a random graph model suggested by Brendan Mckay. We show that this annealed entropy is asymptotically equal to the largest eigenvalue (in modulus) of the random graph's expected transfer matrix. We provide extensive numerical results and find a strong correlation between the annealed entropy and the Shannon entropy of it corresponding underlying distribution. This chapter is joint work with Yin Chen.

Chapter 2

Differential Privacy

In this chapter, we recall some basic definitions and preliminary results from the differential privacy literature. Before that, let us first consider an application scenario of differential privacy. Consider a national Census Bureau such as the U.S. Census Bureau or the Australian Bureau of Statistics. For such a bureau to conduct their legislated business, they must collect very detailed datasets from citizens and businesses at a broad scale, and then release findings to government through aggregate statistics so as to facilitate data-driven policy making. However, such a bureau depends on the trust of the public, in order to collect truthful information. An untrusting public may submit false information in order to protect privacy. Historically such bureaus have been quite forward-thinking, therefore, in adopting privacy-enhancing technologies *e.g.*, Machanavajjhala et al. [2008]. For example consider the problem of fitting a regression on demographic features for predicting annual income. A bureau may have a large dataset with all relevant co-variates, and wish to release such a model for 3rd parties to make subsequent predictions on test individuals, while protecting the privacy of the sensitive training data. Differential privacy provides a framework for verifying that the released model provides this privacy protection. We revisit this example in Section 4.4.3.

2.1 Definitions

Let $\mathcal{X}$ denote a set and $\boldsymbol{x} \in \mathcal{S} = \mathcal{X}^n$ denote a dataset over $\mathcal{X}$ of length n. Each row of the dataset[1] represents the data of an individual. $\boldsymbol{x}$, $\boldsymbol{y}$ are defined as neighbouring datasets (denoted as $x \sim y$) if they differ in a single row, that is, $\boldsymbol{x}$ is at *Hamming*-1 *distance* of y. A mechanism takes datasets $\boldsymbol{x} \in \mathcal{S}$ as input and outputs responses. The concept of differential privacy [Dwork et al., 2006] states that for any two neighbouring datasets, the probability of privacy-preserving mechanism producing any given response is almost the same.

Definition 2.1.1. *[Differential Privacy] A randomized mechanism M provides ϵ-differential privacy if for all neighbouring input datasets $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{S}$ and all measurable $R \subseteq Range(M)$,*

$$Pr[M(\boldsymbol{x}) \in R] \leq e^{\epsilon} Pr[M(\boldsymbol{y}) \in R],$$

where $\epsilon \geq 0$.

Note that, equivalently, differential privacy can also be defined in terms of a family of probability distribution [McGregor et al., 2010]. We will adopt this view of a mechanism as conditional probability distribution in the next two chapters. Intuitively, differential privacy formalizes the idea that a "private" mechanism should not reveal whether any individual is included in the input or not. It provides a strong guarantee that makes it infeasible for an adversary with unbounded computational resources and knowledge of the mechanism up to randomness, to distinguish neighbouring datasets based on the output even if the attacker knows all of the dataset except for the one entry. In the next chapter, we will generalize the definition of neighbouring datasets to encode alternative kinds of desired secrecy, the semantics of which can be understood via the Pufferfish privacy work [Kifer and Machanavajjhala, 2012].

Dwork et al. [2006] showed the above definition can be weakened by allowing for a small probability of the privacy protection failing.

Definition 2.1.2. *[Approximate Differential Privacy] A randomized mechanism M provides (ϵ, δ)-differential privacy if for all neighbouring input datasets $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{S}$ and*

[1] Convention dictates that datasets are represented as n by d matrices when the space $\mathcal{X}$ is $\mathbb{R}^d$. In that case the rows correspond to elements of the dataset.

all measurable $R \subseteq Range(M)$,

$$Pr[M(\boldsymbol{x}) \in R] \leq e^{\epsilon} Pr[M(\boldsymbol{y}) \in R] + \delta,$$

where $\epsilon, \delta \geq 0$.

If $\delta = 0$, approximate differential privacy collapses to ϵ-differential privacy. We can think of this generalization as allowing the mechanism to violate the restriction of the probability ratio, with probability less than δ.

2.2 Laplace and Exponential Mechanisms

Now let us recall two popular ϵ differentially-private mechanisms which we will use as building blocks for some of our algorithms in the next two chapters. One way to achieve differential privacy is to add random noise to the true response to "blur" it. Dwork et al. [2006] proposed adding noise from the Laplace distribution as it is symmetric, exponentially concentrated and matches a convenient form of global sensitivity below. This Laplace noise method depends on the *global sensitivity* of a function, a form of Lipschitz condition:

Definition 2.2.1. *[Global sensitivity] The* L_1 *global sensitivity of a function* $f : \mathcal{S} \rightarrow \mathbb{R}^d$ *is*

$$\varDelta f = max_{\boldsymbol{x} \sim \boldsymbol{x}'} \|f(\boldsymbol{x}) - f(\boldsymbol{x}')\|_1$$

Theorem 2.2.2. *[Laplace Mechanism] Given any function* $f : \boldsymbol{x} \rightarrow \mathbb{R}^d$, *the mechanism* $M_f(\boldsymbol{x}) = f(\boldsymbol{x}) + Lap(\varDelta f/\epsilon)$ *provides* ϵ*-differential privacy where* $Lap(\varDelta f/\epsilon)$ *denotes the* $\boldsymbol{d}$*-vector whose elements are* i.i.d. *random variables drawn from the Laplace Distribution with zero mean and scale parameter* $\varDelta f/\epsilon$.

Proof. Consider any pair of neighbouring datasets $\boldsymbol{x}$ and $\boldsymbol{y}$ at arbitrary output $\boldsymbol{z}$, we have:

$$\begin{aligned}
\frac{Pr(f(\boldsymbol{x}) + Lap(\varDelta f/\epsilon) = \boldsymbol{z})}{Pr(f(\boldsymbol{y}) + Lap(\varDelta f/\epsilon) = \boldsymbol{z})} &= \frac{\exp\left(-\frac{|\boldsymbol{z}-f(\boldsymbol{x})|\epsilon}{\varDelta f}\right)}{\exp\left(-\frac{|\boldsymbol{z}-f(\boldsymbol{y})|\epsilon}{\varDelta f}\right)} \\
&= \exp\left(\frac{\epsilon}{\varDelta f}(|\boldsymbol{z} - f(\boldsymbol{y})| - |\boldsymbol{z} - f(\boldsymbol{x})|)\right)
\end{aligned}$$

$$
\begin{aligned}
&\leq \exp\left(\frac{\epsilon|f(\boldsymbol{x}) - f(\boldsymbol{y})|}{\Delta f}\right)\\
&\leq \exp(\epsilon).
\end{aligned}
$$

□

In this thesis, we will also use another method of providing ϵ-differential privacy called the *exponential mechanism*, and proposed by McSherry and Talwar [2007]. This mechanism is defined with respect to some utility function $\boldsymbol{u}$ which scores preference $u(x, r)$ of response $r \in \boldsymbol{R}$ given a query x. It assumes a base measure over the range $\boldsymbol{R}$. Intuitively, the response that maximizes the utility score is preferred and the exponential mechanism releases close responses with high probability. The exponential mechanism is a generalization of the Laplace mechanism.

Theorem 2.2.3 (Exponential Mechanism)**.** *The exponential mechanism $\boldsymbol{M}$ outputs an element r with probability proportional to* $\exp\left(\frac{\epsilon u(x,r)}{2\Delta u}\right)$ *and preserves ϵ-differential privacy, where $\Delta \boldsymbol{u}$ is the global sensitivity of the utility function, that is:*

$$
\Delta u = \max_{r} \max_{x \sim y} |u(x, r) - u(y, r)|,
$$

Proof. For any two neighbouring datasets $\boldsymbol{x}$ and $\boldsymbol{y}$ and $r \in \boldsymbol{R}$, we have:

$$
\begin{aligned}
\frac{Pr[M(\boldsymbol{x}, R, u, \epsilon) = r]}{Pr[M(\boldsymbol{y}, R, u, \epsilon) = r]} &= \frac{\exp(\frac{\epsilon u(x,r)}{2\Delta u}) \sum_{r'} \exp(\frac{\epsilon u(y,r')}{2\Delta u})}{\exp(\frac{\epsilon u(y,r)}{2\Delta u}) \sum_{r'} \exp(\frac{\epsilon u(x,r')}{2\Delta u})}\\
&\leq \exp\left(\frac{\epsilon(u(x, r) - u(y, r))}{2\Delta u}\right) \frac{\sum_{r'} \exp(\frac{\epsilon u(x,r')+\Delta u}{2\Delta u})}{\sum_{r'} \exp(\frac{\epsilon u(x,r')}{2\Delta u})}\\
&\leq \exp(\epsilon/2)\left(\frac{\exp(\epsilon/2) \sum_{r'} \exp(\frac{\epsilon u(x,r')}{2\Delta u})}{\sum_{r'} \exp(\frac{\epsilon u(x,r')}{2\Delta u})}\right)\\
&= \exp(\epsilon)
\end{aligned}
$$

For a continuous $\boldsymbol{R}$, the proof still follows by replacing the sums with the integrals.
□

The exponential mechanism is a very general mechanism as it does not require continuous or even numeric responses. We will discuss its connection to some of the mechanisms we propose in the thesis.

2.3 Composition Property

Differential privacy satisfies useful composition properties that assist in building up complex differentially-private mechanisms from simple privacy-preserving operations. Dwork et al. [2006] showed that given k independent (ϵ_i, δ_i) mechanisms: $M_i(x), i = 1, \cdots, k$,

$$M_{[k]}(x) \triangleq (M_1(x), \cdots, M_k(x))$$

is $(\sum_{i=1}^{k} \epsilon_i, \sum_{i=1}^{k} \delta_i)$ differentially-private. This demonstrates that the privacy level degrades through composition. For tighter bound of k-fold adaptive composition, the reader may refer to Dwork and Roth [2014] and the references therein.

Chapter 3

Differential Privacy and Bayesian Inference

In this chapter[1], we provide two sufficient conditions on the likelihood and prior to guarantee robustness (low sensitivity to data perturbation) of the posterior distribution in Bayesian Inference. As a result, sampling from the posterior can achieve a level of privacy and utility, essentially "for free". To prove this, we consider differential privacy in a framework of generalized neighbouring datasets, sample spaces and probability distributions. We also propose PSAQR: a mechanism that responds to queries that seek to maximise the expected utility under a Bayesian decision-theoretic framework, and analyze its privacy and utility. We also study how many samples from the posterior the Adversary need in order to distinguish two input databases with high probability. A number of examples of simple conjugate-pairs is provided to demonstrate the results.

3.1 Notation and Definitions

This section displays the notation, definitions and setting that will be used in the differential privacy and Bayesian inference themes of this thesis.

[1] Extending the work of [Dimitrakakis et al., 2014], and correcting some of the proofs therein.

Bayesian inference. This and the next chapters focus on the *Bayesian inference* setting, where a posterior distribution is formed by the statistician $\mathscr{B}$ from a prior distribution ξ and a training dataset x. We assume that data $x \in \mathcal{S}$ is drawn from the distribution $P_{\theta^\star}$ on $\mathcal{S}$, parameterised by $\theta^\star$, from a family of distributions $\mathcal{F}_\Theta$. $\mathscr{B}$ defines a parameter set Θ indexing the family of distributions $\mathcal{F}_\Theta$ on $(\mathcal{S}, \mathfrak{S}_\mathcal{S})$, where $\mathfrak{S}_\mathcal{S}$ is a σ-algebra on $\mathcal{S}$

$$\mathcal{F}_\Theta \triangleq \{P_\theta : \theta \in \Theta\}.$$

We use p_θ to denote the corresponding densities[2] when necessary. In order to perform inference in the Bayesian setting, $\mathscr{B}$ selects a prior measure ξ on $(\Theta, \mathfrak{S}_\Theta)$ that reflects $\mathscr{B}$'s subjective beliefs about which θ is more likely to be true, *a priori*. i.e. For any measurable set of $B \in \mathfrak{S}_\Theta$, $\xi(B)$ represents $\mathscr{B}$'s prior belief of that $\theta^\star \in B$. Generally, the posterior distribution after observing $x \in \mathcal{S}$ is

$$\xi(B \mid x) = \frac{\int_B p_\theta(x)\,\mathrm{d}\xi(\theta)}{\phi(x)} \quad , \tag{3.1}$$

where ϕ is the corresponding marginal density given by

$$\phi(x) \triangleq \int_\Theta p_\theta(x)\,\mathrm{d}\xi(\theta) \quad .$$

Privacy. Recall that Definition 2.1.1 is defined on the neighbouring datasets that differ in only one individual record, it guarantees privacy in the sense of secrecy of an individual record even when the attacker may possess knowledge of the remainder of the database. These semantics are described and generalized in the Blowfish framework [He et al., 2014]. We generalize the concept of neighbouring datasets by equipping $\mathcal{S}$ with a pseudo-metric[3] $\rho : \mathcal{S} \times \mathcal{S} \to \mathbb{R}_+$, and define neighbourhood through distance to encode much boarder notions of adversary's knowledge.

Definition 3.1.1 ((ϵ, δ)-differential privacy under ρ.). *A conditional distribution* $P(\cdot \mid x)$ *on* $(\Theta, \mathfrak{S}_\Theta)$ *is* (ϵ, δ) differentially-private under *a pseudo-metric* $\rho : \mathcal{S} \times \mathcal{S} \to$

[2] I.e. the Radon-Nikodym derivative of P_θ relative to some dominating measure ν.

[3] Meaning that $\rho(x, y) = 0$ does not necessarily imply $x = y$.

$\mathbb{R}_+$ *if, for all* $B \in \mathfrak{S}_\Theta$ *and for any* $x \in \mathcal{S}$,

$$P(B \mid x) \le e^{\epsilon\rho(x,y)} P(B \mid y) + \delta\rho(x,y) \quad \forall y \in \mathcal{S}\,.$$

In the definition above, differential privacy is seen as a measure of smoothness, provided that mechanisms are considered as conditional distributions that correspond to posterior distributions in our Bayesian setting. Note that in Definition 2.1.1, two neighbouring datasets are defined as datasets in Hamming distance one.

Remark 3.1.2. *If* $\mathcal{S} = \mathcal{X}^n$ *and* $\rho(x,y) = \sum_{i=1}^n \mathbb{I}\{x_i \neq y_i\}$ *is the Hamming distance, Definition 3.1.1 is analogous to the standard* (ϵ, δ)*-differential privacy. When considering only* $(\epsilon, 0)$*- differential privacy or* $(0, \delta)$*-privacy, it is an equivalent notion.*

Proof. For $(\epsilon, 0)$-DP, let $\rho(x,z) = \rho(z,y) = 1$, i.e. the data differ in one element. Then, from the standard DP, we have $P(B \mid x) \le e^\epsilon P(B \mid z)$ and so $P(B \mid x) \le e^{2\epsilon} P(B \mid y) = e^{\rho(x,y)\epsilon} P(B \mid y)$. By induction, this holds for any pair of x, y. Similarly, for $(0, \delta)$-DP, by induction we obtain $P(B \mid x) \le P(B \mid y) + \delta\rho(x,y)$. □

Let us show that this generalization of differential privacy satisfies the standard composition property. Composition permits building of complex differentially-private mechanisms based on simple differentially-private algorithmic building blocks.

Theorem 3.1.3 (Composition)**.** *Let conditional distributions* $P(\cdot \mid x)$ *on* $(\Theta, \mathfrak{S}_\Theta)$ *be* (ϵ, δ) differentially-private under *a pseudo-metric* $\rho : \mathcal{S} \times \mathcal{S} \to \mathbb{R}_+$ *and* $P'(\cdot \mid x)$ *on* $(\Theta', \mathfrak{S}'_{\Theta'})$ *be* (ϵ', δ')-differentially private under *the same pseudo-metric. Then the conditional distribution on the product space* $(\Theta \times \Theta', \mathfrak{S}_\Theta \otimes \mathfrak{S}'_{\Theta'})$ *given by*

$$Q(B \times B' \mid x) = P(B \mid x) P(B' \mid x), \forall B \times B' \in \mathfrak{S}_\Theta \otimes \mathfrak{S}'_{\Theta'}$$

satisfies $(\epsilon + \epsilon', \delta + \delta')$ differentially-private under *pseudo-metrics* ρ*. Here* $\mathfrak{S}_\Theta \otimes \mathfrak{S}'_{\Theta'}$ *is the product* σ*-algebra on* $\Theta \times \Theta'$*.*

Proof. By definition of Q and the privacy of P, P',

$$Q(B \times B' \mid x) \le [e^{\epsilon\rho(x,y)} P(B \mid y) + \delta\rho(x,y)] P'(B' \mid x)$$

$$\leq e^{\epsilon\rho(x,y)} P(B \mid y)[e^{\epsilon'\rho(x,y)} P'(B' \mid y) + \delta'\rho(x,y)] + \delta\rho(x,y)$$
$$\leq e^{(\epsilon+\epsilon')\rho(x,y)} P(B \mid y) P'(B' \mid y) + (\delta + \delta')\rho(x,y)$$

In some cases, this generalization can be made equivalent to the definition of Pufferfish privacy proposed by Kifer and Machanavajjhala [2012], a privacy concept with Bayesian semantics. The reader can refer to [Bassily et al., 2013, Chatzikokolakis et al., 2013] for more discussion on generalized neighbouring datasets and the use of metrics in differential privacy.

3.2 Background and Setting

In statistical decision theory, uncertainty is taken into account with decision making under the Bayesian framework, which is attractive as it enables the machinery of probability to be applied in making predictions and modelling. To be more specific, based on the Bayesian paradigm, the world can be described by using probabilistic models with some families of likelihood distributions and prior beliefs on missing likelihood parameters. As more data being observed, a so called posterior belief can be formed by adjusting prior belief through the calculus of probability. The posterior belief can then be released to the world for subsequent modelling and decision makings under uncertainty.

Unfortunately, as the data collected by the statistician is sometimes sensitive, there can a concern that the sensitive information in the original data may be divulged when any information, in terms of the posterior distribution itself or any decisions made based on the calculation of the posterior, is released by the statistician. Currently, in order to codify the information leaking, a framework of differential privacy and various extensions has been developed. The purpose of this framework is to measure the amount of input information that can be leaked through its output. The leakage of input information is bounded provided the algorithm is differentially-private.

In this chapter, we consider how to build differentially-private algorithms based on the Bayesian framework. In particular, we aim to determine what choice of prior enables differential privacy for decisions based on the posterior distribution. Under a decision-theoretic framework, a unified understanding of privacy and learning in adversarial environments is obtained. We show that, under suitable assumptions, uniformly good utility with a fixed privacy budget in the differential privacy setting can be achieved through Bayesian inference and posterior sampling. Apart from that, strong connections between robustness and privacy are illustrated as well.

We show that a base level of data privacy through the posterior distribution is guaranteed by the Bayesian statistician's choice of prior distribution that enables them to respond to external queries safely. A trade-off on privacy leakage and accurate response to query needs to be made, based on how many samples should be used in the estimation of Bayesian models from sensitive data. Our proposed approach is particularly useful in situations where Bayesian inference is already in use by providing examples in the exponential family. Our setting is however entirely general and not limited to specific distribution families, or i.i.d. observations. The general framework is summarised below.

Summary of setting We consider the problem faced by a statistician $\mathscr{B}$ who analyzes data and communicates her findings to a third party $\mathscr{A}$. While $\mathscr{B}$ wants to learn as much as possible *from* the data, she does not want $\mathscr{A}$ to learn *about* any individual datum. For example, in a case where $\mathscr{A}$ is an insurance agency and the data are medical records, $\mathscr{B}$ wants to convey the efficacy of drugs to the agency while without revealing the specific illnesses of individuals in the population. There are no assumptions being made on the data x, and the protocol of interaction between $\mathscr{B}$ and $\mathscr{A}$ is shown below for non-private inference.

1. $\mathscr{B}$ selects a model family $(\mathcal{F}_\Theta)$ and a prior (ξ).

2. $\mathscr{B}$ observes data x and forms the posterior $\xi(\theta|x)$ but does not reveal it.

3. $\mathscr{A}$ is allowed to see $\mathcal{F}_\Theta$ and ξ and is computationally unbounded.
   For steps $t = 1, 2, \ldots$

4. $\mathscr{A}$ sends his utility function u and a query q_t to $\mathscr{B}$.

5. $\mathscr{B}$ responds with the r_t maximising u that depends on the posterior.

To elaborate, based on this framework, the problem commands the choice of the model family $\mathcal{F}_\Theta$, and the prior knowledge of $\mathscr{B}$ acts as a determinant on the choice of ξ, which also influences the privacy level achieved. Informally speaking, better privacy can be achieved by more informative priors because the posterior is less dependent on the data. As publicly available information should be reflected in the prior, it can be assumed to be public. The posterior distribution $\xi(\theta \mid x)$ remains private as it summarizes the statistician's conclusion drawn from the observed data x.

The interaction with $\mathscr{A}$ is indicated in the second part of the process, where a decision-theoretic viewpoint is adopted to achieve the characters of the optimal responses to queries. To be more specific, a utility function $u_\theta(q_t, r_t)$ that $\mathscr{A}$ wants to maximise is built based on the assumption of the existence of a "true" parameter $\theta \in \Theta$. For instance, in the case where a normal distribution has parameters of $\theta = (\mu, \Sigma)$ and an example query q_t is *"what is the expected value* $\mathrm{E}_\theta x_i = \mu$ *of the distribution?"*, the optimal response r_t would be a real vector that depends on the utility function. A possible utility function is the negative squared L_2 distance:

$$u_\theta(q_t = \text{“what is the mean?"}, r_t) = -\|\mathrm{E}_\theta x_i - r_t\|_2^2.$$

Even if θ is unknown, $\mathscr{B}$ can obtain the information about it through a posterior distribution, which takes over the expectation of the expected utility $\mathrm{E}_\xi(u \mid q_t, r_t, x)$ that has been maximised by the optimal response of $\mathscr{B}$ under standard decision-theoretic notions. However this deterministic response cannot be differentially-private.

In this chapter, the use of *posterior sampling* to respond to queries is advocated and the posterior sampling mechanism draws a set of $\hat{\Theta}$ of i.i.d. samples from the posterior distribution. As a result, all the responses only depend on the posterior through $\hat{\Theta}$. As only a single sample set of $\hat{\Theta}$ is taken when we define the algorithm, no more information about the data than what we infer from $\hat{\Theta}$ can be leaked when further queries by the opponent arrives. This enables us to respond to any number of queries with a confined privacy budget, and at the same time achieve good utility.

In section 4, we show that differentially-private responses and robustness of the posterior can be achieved, provided that $\mathcal{F}_\Theta$ and ξ are selected appropriately.[4]

[4] To be more specific about robustness, that small changes in the data result in small changes in the posterior in terms of the KL- divergence.

Further, upper and lower bounds are proven on distinguishing ϵ-close datasets. Lastly, we bound the loss in utility incurred due to privacy. The implication of the results we obtain is that robustness and privacy are linked via smoothness. The robustness of Bayesian learning algorithms is attributed to smooth mappings where their output (*e.g.*, a spam filter) changes little with perturbations to input (*e.g.*, similar training corpora): outliers have decreased effect and unknown information about the data can not be detected by adversaries easily. This suggests robustness and privacy can be achieved at the same time and they are linked to each other deeply.

Based on generalized differential privacy to dataset distances, outcome spaces and distribution families, we provide a uniform mathematical treatment on the privacy and robustness attributes of Bayesian inference, with distinct contributions as follows:

- Under certain regularity conditions on the prior distribution ξ or likelihood family $\mathscr{F}_\Theta$, the posterior distribution is shown to be *robust*: small alterations in the dataset result in small posterior changes.

- We promote a novel *posterior sampling mechanism* that is private.[5] Our approach is different from other common mechanisms in differential privacy as it is based on the non-private (Bayesian) learning framework without alteration.

- Necessary and sufficient conditions for differentially-private Bayesian inference are provided.

- The notion of *dataset distinguishability* is introduced based on which we provide finite-sample bounds for our mechanism: the size of $\hat{\Theta}$ needs to be determined for $\mathscr{A}$ to differentiate between two datasets with high probability.

- We also provide examples of conjugate-pair distributions where our assumptions hold, to illustrate the application of our results.

[5] Although previously used *e.g.*, for efficient exploration in reinforcement learning [Thompson, 1974, Osband et al., 2013], posterior sampling has not previously been employed for privacy.

3.3 Our Main Assumptions

In this section, we introduce two assumptions that one could make on the smoothness of the family $\mathcal{F}_\Theta$ with respect to some metric d on $\mathbb{R}_+$ such that close datasets $x, y \in \mathcal{S}$ result in posterior distributions that are close as measured by KL-divergence. The first assumption states that the likelihood is smooth for all parameterizations of the family. Firstly, let us define our notion of smoothness. Let $f(x, \theta) \triangleq \ln p_\theta(x)$ be the log probability of x under θ. The Lipschitz constant for a parameter value θ is

$$\ell(\theta) \triangleq \inf \{ u : |f(x, \theta) - f(y, \theta)| \leq u\rho(x, y) \forall x, y \in \mathcal{S} \} . \tag{3.2}$$

Our first assumption is uniform smoothness for all parameters.

Assumption 1 (Lipschitz continuity)**.** *We assume that there exists some $L < \infty$ such that*

$$\ell(\theta) \leq L, \qquad \forall \theta \in \Theta. \tag{3.3}$$

In other words, this assumption says that the log probability is uniformly Lipschitz with respect to ρ over all parameter values. But it might be difficult for this assumption to hold uniformly over Θ generally. This can be seen by the following counterexample for the Bernoulli family of distributions: when the parameter is 0, then any sequence $x = 0, 0, \ldots$ has probability 1, while any sequence containing a 1 has probability 0. To avoid such problems, we relax the assumption by only requiring that $\mathscr{B}$'s *prior* probability ξ is concentrated in the regions of the family for which the likelihood is smoothest:

Assumption 2 (Stochastic Lipschitz continuity; Norkin, 1986)**.** *Firstly, we define the subset of parameter values*

$$\Theta_L \triangleq \{ \theta \in \Theta : \ell(\theta) \leq L \} \tag{3.4}$$

to be those parameters for which Lipschitz continuity holds with Lipschitz constant

L. Thus there are some constants $c, L_0 > 0$ such that, for all $L \geq L_0$

$$\xi(\Theta_L) \geq 1 - \exp(-c(L - L_0)) \ . \tag{3.5}$$

This weaker assumption is easier to meet while still generates useful guarantees by not requiring uniform smoothness. Note that L_0 is determined by the nature of the likelihood distributions. It reflects that certain levels of smoothness are not achievable for certain likelihood functions. In fact, in Section 3.6, we prove that this assumption is satisfied by many important example distribution families.

In the next section, we show that verifying our assumptions for a distribution of a single random variable lifts to a corresponding property for the product distribution on i.i.d. samples.

Lemma 3.3.1. *If $\mathcal{F}_\Theta$ satisfies Assumption 1 (resp. Assumption 2) with respect to pseudo-metric ρ and constant L (or c and L_0), then, for any fixed $n \in \mathbb{N}$, the product family $\mathcal{F}_\Theta^n$ with densities $p_\Theta^n(\{x_i\}) = \prod_{i=1}^n p_\Theta(x_i)$ satisfies the same assumption with respect to*

$$\rho^n(\{x_i\}, \{y_i\}) = \textstyle\sum_{i=1}^n \rho(x_i, y_i)$$

and constant L (or c and L_0).

Proof. For Assumption 1, the proof follows directly from the definition of the absolute log-ratio distance, namely

$$\begin{aligned} |\ln p_\theta^n(\{x_i\}) - \ln p_\theta^n(\{y_i\})| &\leq \textstyle\sum_{i=1}^n |\ln p_\theta(x_i) - \ln p_\theta(y_i)| \\ &\leq L \textstyle\sum_{i=1}^n \rho(x_i, y_i) \ . \end{aligned}$$

For Assumption 2, consider the sub-family Θ_L from Eq. (3.4) for marginal p_θ and pseudo-metric ρ, and define the corresponding sub-family Θ_L^n in terms of product distribution p_θ^n and pseudo-metric ρ^n. Then the same argument as above shows that $\Theta_L \subseteq \Theta_L^n$. Therefore, the same prior and parameters c and L_0 yield the lower bound of Eq. (3.5), for Θ_L^n. □

3.3.1 Sufficient Statistics as a Necessary Condition

The extent to which our assumptions hold for a particular family of distributions $\mathcal{F}_\Theta$ is mainly determined by ρ. And the choice of metric is also important for achieving differential privacy. Now we specifically consider metrics defined in terms of a difference in statistics

$$\rho(x, y) \triangleq \|\tau(x) - \tau(y)\|, \tag{3.6}$$

where $\tau : \mathcal{S} \to \mathcal{V}$ is a statistic mapping from datasets to a normed vector space.

Necessity for assumptions. In that case, our assumptions imply that f must be a *sufficient* statistic, since if $\tau(x) = \tau(y)$ then $\rho(x, y) = 0$ and it follows that $P_\theta(x) = P_\theta(y)$. More generally, ρ must be such that if the distance between x, y is zero, then their probabilities should be equal. We will provide some examples of such statistics for conjugate distributions in the exponential family in Section 3.6. That means a metric that simply ignores part of the data can not be used, for example.

Necessity for differential privacy. Similarly, the definition of differential privacy (Definition 3.1.1) implies that f must be a *Bayes-sufficient* statistic. This means that for any x, y

$$f(x) = f(y) \quad \Rightarrow \quad \xi(B \mid x) = \xi(B \mid y).$$

Note that this is a slightly weaker condition than a sufficient statistic, which is necessary for our assumptions to hold.

3.3.2 Summary of Results

Given the above assumptions, we show the following results: Firstly, if we choose an informative prior ξ, then the resulting posterior is robust in terms of KL-divergence to small changes in the data; Secondly, the posterior distribution is differentially-private; Thirdly, this implies that sampling from the posterior can be used as part of a differentially-private mechanism. We complement these with results on how easily an adversary can distinguish two similar datasets from posterior samples; Finally, we characterise the trade-off between utility and privacy, stated here informally for ease of exposition:

Claim 1. *If $\mathscr{A}$ prefers to use the prior $\xi^\star$, but $\mathscr{B}$ uses a prior ξ satisfying Assumption 1, and $\mathscr{A}$'s utility is bounded in $[0, 1]$, the following is true for the posterior sampling mechanism with N samples:*

- *The mechanism is $2NL$ differentially-private;*
- *$\mathscr{A}$'s utility loss is $O\Big([1 - \xi^\star(\Theta_L)] + \sqrt{1/N}\Big)$ w.h.p., where Θ_L is the support of ξ.*

The following sections discuss our main results in detail. We begin by proving that our assumptions result in robust posteriors, in the sense that the KL-divergence between posteriors arising from similar datasets is small. Then we show that they also result in differentially-private posterior distributions, and analyze the resulting posterior sampling mechanism. We conclude with some examples and a discussion of related work.

3.4 Robustness of the Posterior Distribution

We now show that the above assumptions provide guarantees on the robustness of the posterior. That is, if the distance between two datasets x, y is small, then so is the distance between the two resulting posteriors, $\xi(\cdot \mid x)$ and $\xi(\cdot \mid y)$. We prove this result for the case where we measure the distance between the posteriors in terms of the well-known KL-divergence

$$D\left(P \parallel Q\right) = \int_S \ln \frac{\mathrm{d}P}{\mathrm{d}Q} \,\mathrm{d}P \ . \tag{3.7}$$

The following theorem shows that any distribution family $\mathcal{F}_\Theta$ and prior ξ satisfying one of our assumptions is robust, in the sense that the posterior does not change significantly with small changes to the dataset. It is notable that our mechanisms are simply tuned through the choice of prior.

Theorem 3.4.1. *When ξ is a prior distribution on Θ and $\xi(\cdot \mid x)$ and $\xi(\cdot \mid y)$ are the respective posterior distributions for datasets $x, y \in \mathcal{S}$, the following results hold*

1. *Under a pseudo-metric ρ and $L > 0$ satisfying Assumption 1,*

$$D\left(\xi(\cdot \mid x) \parallel \xi(\cdot \mid y)\right) \leq 2L\rho(x, y) \ ; \tag{3.8}$$

2. *Under a pseudo-metric ρ and $c > 1$ satisfying Assumption 2*

$$D\left(\xi(\cdot \mid x) \parallel \xi(\cdot \mid y)\right) \leq C_\xi^{\mathcal{F}_\Theta}\left(1 + 2L_0 + c^{-1}\right)\rho(x, y) \ , \tag{3.9}$$

where $C_\xi^{\mathcal{F}_\Theta}$ is the ratio between the maximum and marginal likelihoods (3.11).

Proof. Let us now tackle claim 1. First, we can decompose the KL-divergence into two parts.

$$\begin{aligned} D\left(\xi(\cdot \mid x) \parallel \xi(\cdot \mid y)\right) &= \int_\Theta \ln \frac{\mathrm{d}\xi(\theta \mid x)}{\mathrm{d}\xi(\theta \mid y)} \,\mathrm{d}\xi(\theta \mid x) \\ &= \int_\Theta \ln \frac{p_\theta(x)}{p_\theta(y)} \,\mathrm{d}\xi(\theta \mid x) + \int_\Theta \ln \frac{\phi(y)}{\phi(x)} \,\mathrm{d}\xi(\theta \mid x) \\ &\leq \int_\Theta \left|\ln \frac{p_\theta(x)}{p_\theta(y)}\right| \mathrm{d}\xi(\theta \mid x) + \int_\Theta \ln \frac{\phi(y)}{\phi(x)} \,\mathrm{d}\xi(\theta \mid x) \\ &\leq L\rho(x, y) + \left|\ln \frac{\phi(y)}{\phi(x)}\right| \ . \end{aligned} \tag{3.10}$$

From Assumption 1, $p_\theta(y) \leq \exp(L\rho(x, y))p_\theta(x)$ for all θ so:

$$\begin{aligned} \phi(y) &= \int_\Theta p_\theta(y) \,\mathrm{d}\xi(\theta) \\ &\leq \exp(L\rho(x, y)) \int_\Theta p_\theta(x) \,\mathrm{d}\xi(\theta) = \exp(L\rho(x, y))\phi(x) \ . \end{aligned}$$

Combining this with (3.10) we obtain

$$D\left(\xi(\cdot \mid x) \parallel \xi(\cdot \mid y)\right) \leq 2L\rho(x, y) \ .$$

Claim 2 is dealt with similarly. Once more, we can break down the distance in parts. In more detail, we first write:

$$D\left(\xi(\cdot \mid x) \parallel \xi(\cdot \mid y)\right) \leq \underbrace{\int_{\Theta} \left| \ln \frac{p_\theta(x)}{p_\theta(y)} \right| \mathrm{d}\xi(\theta \mid x)}_{A} + \underbrace{\int_{\Theta} \ln \frac{\phi(y)}{\phi(x)} \mathrm{d}\xi(\theta \mid x)}_{B} \ ,$$

as before. Now, let us re-write the A term as

$$\int_{\Theta} \left| \ln \frac{p_\theta(x)}{p_\theta(y)} \right| \frac{p_\theta(x)}{\phi(x)} \mathrm{d}\xi(\theta) \leq \sup_{\theta'} \frac{p_{\theta'}(x)}{\phi(x)} \int_{\Theta} \left| \ln \frac{p_\theta(x)}{p_\theta(y)} \right| \mathrm{d}\xi(\theta) \ ,$$

so that the left-hand side term is the ratio between the maximal likelihood and marginal likelihood. Using the same steps, we can bound B in the same manner.

Now, let us define a data-dependent and a data-independent bound:

$$C_{\xi}^{\mathcal{F}_\Theta}(x) \triangleq \sup_{\theta} \frac{p_\theta(x)}{\phi(x)} \ , \qquad\qquad C_{\xi}^{\mathcal{F}_\Theta} \triangleq \sup_{x} C_{\xi}^{\mathcal{F}_\Theta}(x) \ . \tag{3.11}$$

Replacing, we obtain:

$$D\left(\xi(\cdot \mid x) \parallel \xi(\cdot \mid y)\right) \leq C_{\xi}^{\mathcal{F}_\Theta} \underbrace{\int_{\Theta} \left| \ln \frac{p_\theta(x)}{p_\theta(y)} \right| \mathrm{d}\xi(\theta)}_{A} + \underbrace{\int_{\Theta} \ln \frac{\phi(y)}{\phi(x)} \mathrm{d}\xi(\theta \mid x)}_{B} \ .$$

Now, to bound the individual terms, we start from A and note that theorem 3 of [Norkin, 1986] on the Lipschitz property of the expectation of stochastic Lipschitz functions applies.

Theorem 3.4.2. *[Norkin, 1986] If ξ is a probability measure on Θ and $f : \mathcal{S} \times \Theta \to \mathbb{R}$ is a ξ-measurable function, such that for any $\theta \in \Theta$, $f(\cdot, \theta)$ is $\ell(\theta)$-Lipschitz, then the function $f_\xi(x) \triangleq \mathrm{E}_\xi f(x, \theta)$ is L_ξ-Lipschitz, where $L_\xi = \mathrm{E}_\xi \ell(\theta)$.*

Recall that the expectation of a non-negative random variable can be written in terms of its CDF F as $\int_0^\infty [1 - F(t)] \, \mathrm{d}t$. In our case, $\ell(\theta)$ is a random variable on Θ, and we can write its cumulative distribution function as

$$F(t) \triangleq \xi\left(\{\, \theta \in \Theta : \ell(\theta) \leq t \,\}\right) = \xi(\Theta_t) \ ,$$

by the definition of Θ_t. It follows that $\ln p_\theta(x)$ is L_ξ-Lipschitz, where through the formula for the expectation of positive variables:

$$L_\xi = \int_0^\infty [1 - \xi(\Theta_t)] \,\mathrm{d}t \le L_0 \xi(\Theta_{L_0}) + [1 - \xi(\Theta_{L_0})] \int_0^\infty e^{-ct} \,\mathrm{d}t \le L_0 + c^{-1} \ . \tag{3.12}$$

So, term A becomes $C_\xi^{\mathcal{F}_\Theta} \left(L_0 + c^{-1}\right) \rho(x, y)$.

Now let us move on to term B. For technical reasons, we start by considering a pair x, y such that $\rho(x, y) \le c - 1$. This also implies that $c > 1$, since the distance cannot be negative.

$$\frac{\phi(x)}{\phi(y)} \overset{(a)}{=} \int_\Theta \frac{p_\theta(x)}{\phi(y)} \,\mathrm{d}\xi(\theta) \overset{(b)}{\le} \int_\Theta \frac{p_\theta(y) e^{\ell(\theta)\rho(x,y)}}{\phi(y)} \,\mathrm{d}\xi(\theta) \overset{(c)}{\le} C_\xi^{\mathcal{F}_\Theta} \int_\Theta e^{\ell(\theta)\rho(x,y)} \,\mathrm{d}\xi(\theta) \ . \tag{3.13}$$

Note that $\left\{\theta \in \Theta : e^{\ell(\theta)\rho(x,y)} \le t\right\} = \left\{\theta \in \Theta : \ell(\theta) \le \rho(x, y)^{-1} \ln t\right\} = \Theta_{\rho(x,y)^{-1} \ln t}$. So the CDF of the random variable $e^{\ell(\theta)}$ is $F(t) = \xi(\Theta_{\rho(x,y)^{-1} \ln t})$. Then

$$\begin{aligned}
\mathrm{E}_\xi e^{\ell(\theta)\rho(x,y)} &= \mathrm{E}_\xi[e^{\ell(\theta)\rho(x,y)} \mid \ell \le L_0] \xi(\Theta_{L_0}) + \mathrm{E}_\xi[e^{\ell(\theta)\rho(x,y)} \mid \ell > L_0][1 - \xi(\Theta_{L_0})] \\
&\le e^{L_0\rho(x,y)} + \rho(x, y) \int_{t_0}^\infty t^{\rho(x,y)-1} [1 - \xi(\Theta_{\ln t})] \,\mathrm{d}t \\
&\le e^{L_0\rho(x,y)} + \rho(x, y) \int_{t_0}^\infty e^{\ln t[\rho(x,y)-1]} e^{-c(\ln t - L_0)} \,\mathrm{d}t \qquad \text{(where } t_0 = e^{L_0}) \\
&= e^{L_0\rho(x,y)} + \rho(x, y) \int_{t_0}^\infty e^{\ln t[\rho(x,y)-c-1]+cL_0} \,\mathrm{d}t \\
&= e^{L_0\rho(x,y)} + \rho(x, y) e^{cL_0} \int_{t_0}^\infty t^{\rho(x,y)-c-1} \,\mathrm{d}t \\
&= e^{L_0\rho(x,y)} + \rho(x, y) e^{cL_0} \frac{t_0^{\rho(x,y)-c}}{c - \rho(x, y)} \\
&= e^{L_0\rho(x,y)} + \rho(x, y) e^{cL_0} \frac{e^{L_0(\rho(x,y)-c)}}{c - \rho(x, y)} \\
&\le e^{L_0\rho(x,y)} + \rho(x, y) e^{cL_0} e^{L_0(\rho(x,y)-c)} \\
&= e^{L_0\rho(x,y)} + \rho(x, y) e^{L_0\rho(x,y)} = (1 + \rho(x, y)) e^{L_0\rho(x,y)} \le e^{(1+L_0)\rho(x,y)}.
\end{aligned}$$

Consequently, $\ln \phi(x)/\phi(y) \le C_\xi^{\mathcal{F}_\Theta}(1 + L_0)\rho(x, y)$.

To handle larger distances ρ, we can simply apply the above result repeatedly between k data sets $z_1, \dots, z_k$, where $z_1 = x$, $z_k = y$ and such that $\rho(z_i, z_{i+1}) < c - 1$. By chaining logarithmic ratios, i.e. using the fact that $\ln \phi(x)/\phi(y) = \ln \phi(x)/\phi(z) + \ln \phi(z)/\phi(y)$ we can now extend our result to general pairs for term B. Replacing those terms, we obtain:

$$D\left(\xi(\cdot \mid x) \parallel \xi(\cdot \mid y)\right) \leq C_{\xi}^{\mathcal{F}_\Theta}\left(1 + 2L_0 + c^{-1}\right)\rho(x, y) \ .$$

If the intermediate points do not exist under ρ, we can scale it properly, thus obtaining the final result. □

Note that the second claim bounds the KL-divergence in terms of $\mathscr{B}$'s prior belief that L is small, which is expressed via the constant c. The larger c is, the less prior mass is placed in large L and so the more robust inference becomes. On the other hand, choosing c to be too large may decrease efficiency.

3.4.1 Alternative Analysis

We show an alternative result to Theorem 3.4.1 below.

Theorem 3.4.3. *When $d : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ is the absolute log-ratio distance, ξ is a prior distribution on Θ and $\xi(\cdot \mid x)$ and $\xi(\cdot \mid y)$ are the respective posterior distributions for datasets $x, y \in \mathcal{S}$, the following results hold:*

1. *Under a metric ρ and $L > 0$ satisfying Assumption 1,*

$$D\left(\xi(\cdot \mid x) \parallel \xi(\cdot \mid y)\right) \leq 2L\rho(x, y) \ .$$

2. *Under a metric ρ and $c > 0$ satisfying Assumption 2 and satisfying $\rho(x, y) < (1 - \epsilon)c$ uniformly for all x, y for some $\epsilon \in (0, 1)$,*

$$D\left(\xi(\cdot \mid x) \parallel \xi(\cdot \mid y)\right) \leq M \cdot \max\{\rho(x, y), 1\} \ .$$

where

$$M = \left(\frac{\kappa}{c} + L_0\left(\frac{1}{1 - e^{-\omega}} + 1\right) + \ln C_{\xi}^{\mathcal{F}_\Theta} + \ln\left(e^{-L_0 \delta c}\left(e^{-\omega(1-\delta)} - e^{-\omega}\right)^{-1} + e^{L_0(1-\delta)c}\right)\right) C_{\xi}^{\mathcal{F}_\Theta};$$

constants $\kappa = 4.91081$ *and* $\omega = 1.25643$*; and* $C_\xi^{\mathcal{F}_\Theta}$ *defined in* (3.11)*. Note that*

$$M = O\left(\left(\frac{1}{c} + \ln C_\xi^{\mathcal{F}_\Theta} + L_0\right) C_\xi^{\mathcal{F}_\Theta}\right).$$

Proof. Using the same steps as in the proof of Theorem 3.4.1, we have

$$D\left(\xi(\cdot \mid x) \parallel \xi(\cdot \mid y)\right) \le C_\xi^{\mathcal{F}_\Theta} \left(\underbrace{\int_\Theta \left| \ln \frac{p_\theta(x)}{p_\theta(y)} \right| \mathrm{d}\xi(\theta)}_{A} + \underbrace{\int_\Theta \ln \frac{\phi(y)}{\phi(x)} \mathrm{d}\xi(\theta)}_{B} \right) . \tag{3.14}$$

Now, to bound the individual terms, we start from A and write it as a sum of integrals that partitions Θ. Let $\Theta_{[a,b]} \triangleq \Theta_b \setminus \Theta_a$. Then $\xi(\Theta_{[a,b]}) = \xi(\Theta_b) - \xi(\Theta_a) \le e^{-ca}$, as $\Theta_b \supset \Theta_a$, while $\xi(\Theta_b) \le 1$ and $\xi(\Theta_a) \ge 1 - e^{-ca}$ from Ass 2. We can thus partition Θ into disjoint sets corresponding to uniformly sized intervals $[L_0 + (L-1)\alpha, L_0 + L\alpha)$ of size $\alpha > 0$ indexed by L. We bound the divergence on each partition and sum over L.

$$\int_\Theta \left| \ln \frac{p_\theta(x)}{p_\theta(y)} \right| \mathrm{d}\xi(\theta) = \sum_{L=1}^{\infty} \int_{\Theta_{[L_0+(L-1)\alpha, L_0+L\alpha]}} \left| \ln \frac{p_\theta(x)}{p_\theta(y)} \right| \mathrm{d}\xi(\theta) + \int_{\Theta_{[0,L_0]}} \left| \ln \frac{p_\theta(x)}{p_\theta(y)} \right| \mathrm{d}\xi(\theta) \tag{3.15}$$

$$\overset{(a)}{\le} \rho(x,y) \sum_{L=1}^{\infty} (L_0 + L\alpha) \int_{\Theta_{[L_0+(L-1)\alpha, L_0+L\alpha]}} \mathrm{d}\xi(\theta) + L_0 \rho(x,y) \tag{3.16}$$

$$\overset{(b)}{\le} \rho(x,y)[\alpha \sum_{L=1}^{\infty} L e^{-c\alpha(L-1)} + L_0 \sum_{L=0}^{\infty} e^{-c\alpha L} + L_0] \tag{3.17}$$

$$\overset{(c)}{=} \rho(x,y) \left[\alpha \left(1 - e^{-c\alpha}\right)^{-2} + \frac{L_0}{1 - e^{-\alpha c}} + L_0 \right] , \tag{3.18}$$

where $(a), (b)$ are from Assumption 1, equation (3.4) and (3.3) respectively, and (c) is via the geometric series. Now let us move on to term B. Since the logarithmic term does not depend on θ, this is simply bounded by $|\ln \frac{\phi(y)}{\phi(x)}|$. We now attempt to bound this as follows:

$$\phi(y) = \int_\Theta p_\theta(y) \,\mathrm{d}\xi(\theta) = \sum_{L=1}^{\infty} \int_{\Theta_{[L_0+(L-1)\alpha, L_0+L\alpha]}} p_\theta(y) \,\mathrm{d}\xi(\theta) \; + \; \int_{\Theta_{[0,L_0]}} p_\theta(y) \,\mathrm{d}\xi(\theta) \tag{3.19}$$

$$\leq \sum_{L=1}^{\infty} e^{(L_0+L\alpha)\rho(x,y)} \int_{\Theta_{[L_0+(L-1)\alpha, L_0+L\alpha]}} p_\theta(x)\,\mathrm{d}\xi(\theta) + p_\theta(x) e^{L_0\rho(x,y)} \tag{3.20}$$

$$\frac{\phi(y)}{\phi(x)} \leq \sum_{L=1}^{\infty} e^{(L_0+\alpha L)\rho(x,y)} \int_{\Theta_{[L_0+(L-1)\alpha, L_0+L\alpha]}} [p_\theta(x)/\phi(x)]\,\mathrm{d}\xi(\theta) + e^{L_0\rho(x,y)}\ [p_\theta(x)/\phi(x)] \tag{3.21}$$

$$\leq C_\xi^{\mathcal{F}_\Theta} \left[\sum_{L=1}^{\infty} e^{(L_0+\alpha L)\rho(x,y)} \int_{\Theta_{[L_0+(L-1)\alpha, L_0+L\alpha]}} \mathrm{d}\xi(\theta) + e^{L_0\rho(x,y)} \right] \tag{3.22}$$

$$\leq C_\xi^{\mathcal{F}_\Theta} \left[e^{\alpha c + L_0(\rho(x,y)-c)} \sum_{L=1}^{\infty} e^{\alpha L(\rho(x,y)-c)} + e^{L_0\rho(x,y)} \right] \tag{3.23}$$

$$\leq C_\xi^{\mathcal{F}_\Theta} \left[e^{L_0(\rho(x,y)-c)} (e^{-\alpha\rho(x,y)} - e^{-c\alpha})^{-1} + e^{L_0\rho(x,y)} \right] \ . \tag{3.24}$$

Note that the series converge only if $\rho(x,y) < c$. Let us assume that there exists $1 > \epsilon > 0$ such that $(1-\epsilon)c \geq \rho(x,y)$. Then we have that

$$D\left(\xi(\cdot \mid x) \parallel \xi(\cdot \mid y)\right) \leq C_\xi^{\mathcal{F}_\Theta} \rho(x,y) \left(\alpha(1-e^{-c\alpha})^{-2} + \frac{L_0}{1-e^{-\alpha c}} + L_0 \right) \tag{3.25}$$

$$+ C_\xi^{\mathcal{F}_\Theta} \left(\ln C_\xi^{\mathcal{F}_\Theta} + \ln(e^{L_0(\rho(x,y)-c)}(e^{-\alpha\rho(x,y)} - e^{-c\alpha})^{-1} + e^{L_0\rho(x,y)}) \right) \tag{3.26}$$

$$\leq C_\xi^{\mathcal{F}_\Theta} \rho(x,y) \left(\underbrace{\alpha(1-e^{-c\alpha})^{-2} + \frac{L_0}{1-e^{-\alpha c}} + L_0}_{D} \right) \tag{3.27}$$

$$+ C_\xi^{\mathcal{F}_\Theta} \left(lnC_\xi^{\mathcal{F}_\Theta} + \underbrace{\ln\left(e^{-L_0\epsilon c}(e^{-\alpha(1-\epsilon)c} - e^{-c\alpha})^{-1} + e^{L_0(1-\epsilon)c} \right)}_{E} \right) \tag{3.28}$$

$$\leq C_\xi^{\mathcal{F}_\Theta} (D + lnC_\xi^{\mathcal{F}_\Theta} + E) \max(\rho(x,y), 1) \tag{3.29}$$

but it remains to tune the constant α.

Tuning the bound. Note that L_0 depends on the likelihoods, for simplicity let us pick α based on the case $L_0 = 0$. In such case, we have

$$D\left(\xi(\cdot \mid x) \parallel \xi(\cdot \mid y)\right) \leq C_\xi^{\mathcal{F}_\Theta} \left(\underbrace{\alpha(1-e^{-c\alpha})^{-2}}_{F} + \ln C_\xi^{\mathcal{F}_\Theta} + \underbrace{\alpha c - \ln(e^{c\alpha\epsilon} - 1)}_{G} \right) \tag{3.30}$$

This bound holds for any size parameter $\alpha > 0$ and is convex for $\alpha > 0$, $c > 0$. Thus, there is an optimal choice for α that minimizes this bound. The optimal choice is given by the solution of the transcendental equation which is obtained by differentiating w.r.t to α and setting the result to zero. Since there is no analytical solution to the transcendental equation, we tune this bound by examining the minimum of F and G separately. Note that for F, we have that the optimal $\alpha_1^* = \frac{\omega}{c}$ where ω is the unique non-zero solution to $e^\omega = 2\omega + 1$. As the $\omega \approx 1.25643$ is the unique positive solution to $e^\omega = 2\omega + 1$, we have $\alpha_A^* = 1.25643/c$. For G, the minimum point is $\alpha_B^* = \frac{1}{c\epsilon} \ln \frac{c}{c-c\epsilon}$. By mean value theorem there exists a $z \in (c - c\epsilon, c)$ such that $\alpha_B^* = 1/z$, that is $\alpha_B^* \in (1/c, 1/(c - c\epsilon))$. Note that α_A^* is in $(1/c, 1/(c - c\epsilon))$ when $\epsilon \geq 0.21$. We can pick $\alpha^* = 1.25643/c$ here. □

3.5 Privacy and Utility

We next examine the differential privacy of the posterior distribution. We show in Section 3.5.1 that this can be achieved under either of our assumptions. The result can also be interpreted as the differential privacy of a *posterior sampling mechanism* for responding to queries (described in Section 3.5.2), for which we prove a bound on the utility depending on the number of samples taken. Section 3.5.3 examines an alternative notion of privacy, *dataset distinguishability*, similar to Wasserman and Zhou [2010]. For this, we prove a bound on privacy, that also depends on the number of samples taken. Together, these exhibit a trade off between utility and privacy controlled by choosing the number of samples appropriately, in a manner described in Section 3.5.4.

3.5.1 Differential Privacy of Posterior Distributions

We consider our generalized notion of differential privacy for posterior distributions (Definition 3.1.1), and show that the type of differential privacy exhibited by the posterior depends on which assumption holds.

Theorem 3.5.1. *1. Under Assumption 1, for all $x, y \in \mathcal{S}$, $B \in \mathfrak{S}_\Theta$:*

$$\xi(B \mid x) \leq \exp\{2L\rho(x, y)\}\xi(B \mid y) ,$$

i.e. the posterior ξ is $(2L, 0)$ differentially-private under pseudo-metric ρ.

2. *Under a pseudo-metric ρ and $c > 1$ satisfying Assumption 2, $C_{\xi}^{\mathcal{F}_\Theta}$ defined in (3.11), for all $x, y \in \mathcal{S}$, $B \in \mathfrak{S}_\Theta$:*

$$|\xi(B \mid x) - \xi(B \mid y)| \leq \sqrt{\frac{C_{\xi}^{\mathcal{F}_\Theta}}{2}(1 + 2L_0 + c^{-1})\rho(x, y)},$$

i.e. the posterior ξ is $\left(0, O(\sqrt{C_{\xi}^{\mathcal{F}_\Theta}(L_0 + 1/c)})\right)$-differentially private[6] under pseudo-metric $\sqrt{\rho}$.

Proof. For part 1, we assumed that there is an $L > 0$ such that $\forall x, y \in \mathcal{S}$, $\left|\log \frac{p_\theta(x)}{p_\theta(y)}\right| \leq L\rho(x, y)$, implying $\frac{p_\theta(x)}{p_\theta(y)} \leq \exp\{L\rho(x, y)\}$. Further, in the proof of Theorem 3.4.1, we showed that $\phi(y) \leq \exp\{L\rho(x, y)\}\phi(x)$ for all $x, y \in \mathcal{S}$. From Eq. (3.1), we can then combine these to bound the posterior of any $B \in \mathfrak{S}_\Theta$ as follows for all $x, y \in \mathcal{S}$:

$$\xi(B \mid x) = \frac{\int_B \frac{p_\theta(x)}{p_\theta(y)} p_\theta(y) \, \mathrm{d}\xi(\theta)}{\phi(y)} \cdot \frac{\phi(y)}{\phi(x)} \leq \exp\{2L\rho(x, y)\}\xi(B \mid y) \ .$$

For part 2, note that the KL-divergence of the posteriors under assumption is bounded by (3.9). Now, recall Pinsker's inequality [cf. Fedotov et al., 2003]:

$$D\left(Q \| P\right) \geq \|Q - P\|_{\mathrm{TV}}^2 \triangleq 2 \sup_B |Q(B) - P(B)|^2 \tag{3.31}$$

This yields: $|\xi(B \mid x) - \xi(B \mid y)| \leq \sqrt{\frac{1}{2} D\left(\xi(\cdot \mid x) \parallel \xi(\cdot \mid y)\right)} \leq \sqrt{\frac{1}{2} C_{\xi}^{\mathcal{F}_\Theta}(1 + 2L_0 + c^{-1})\rho(x, y)}$.

The difference between the two bounds' form is due to the fact that the first claim has a direct proof and the second claim arises from Theorem 3.4.1.

Finally, we show that posterior distributions are also randomly differentially-private.

Corollary 3.5.2. *Under Assumption 2:*

$$\mathbb{P}\left[\forall B \in \mathfrak{S}_\Theta : \xi(B \mid x) \leq \exp\left\{2L\rho(x, y)\right\} \xi\left(B \mid y\right), \forall x, y \in \mathcal{S}\right] \geq 1 - \exp(-c(L - L_0)) \ .$$

[6] This holds, for example, for hamming distance as in the Beta-Binomial example presented in Lemma 3.6.3.

That is, the posterior ξ is $(2L, 0, \exp(-c(L - L_0))$*-randomly differentially-private under pseudo-metric* ρ*.*

This is a conceptually different definition from the original RDP, as the measure over which the randomness is defined is not the data distribution, but the prior measure ξ.

This property of the posterior distribution directly leads to the definition of a posterior sampling mechanism which will be differentially private. This is explained in the following section.

3.5.2 Posterior Sampling Mechanism

Given that we have a full posterior distribution which is differentially-private, we can use it to define a private mechanism. We may allow the adversary to submit an arbitrary set of queries $\{ q_t \}$ with each $q_t \in \mathcal{Q}$. Each query warrants a response r_t in a set of possible responses $\mathcal{R}$. The adversary is allowed to condition the queries on our previous responses.

We extend the approach of Dimitrakakis et al. [2014] to take some utility function u into account, which scores preferences of responses given a query. The first step is to simply draw a number of samples from the posterior, as in the original approach (Algorithm 3.5.2). After the algorithm calculates the posterior distribution $\xi(\cdot \mid x)$, N parameter samples are drawn from it, producing a parameter set $\hat{\Theta}$. Thereafter, responses depend only on the utility function and the sample $\hat{\Theta}$, and we do not draw new samples after every query. This allows us to work with a fixed privacy budget.

Algorithm 3.5.1: BAPS: Bayesian Posterior Sampling

```
input prior $\xi$, data $x \in \mathcal{S}$
Calculate posterior $\xi(\theta \mid x)$.
for $k = 1, \dots, N$ do
    Sample $\theta^{(k)} \sim \xi(\theta \mid D)$.
end for
return $\hat{\Theta} = \left\{ \theta^{(k)} : k = 1, \dots, N \right\}$.
```

Corollary 3.5.3. *Algorithm 3.5.1 is differentially private under the conditions of Theorem 3.5.1, namely:*

1. *Under a pseudo-metric ρ and $L > 0$ satisfying Assumption 1, the algorithm is $(2NL, 0)$-differentially private under pseudo-metric ρ; or*

2. *Under a pseudo-metric ρ and $c > 1$ satisfying Assumption 2, $C_\xi^{\mathcal{F}_\Theta}$ defined in (3.11), the algorithm is $\left(0, O(N\sqrt{C_\xi^{\mathcal{F}_\Theta}(L_0 + 1/c)})\right)$-differentially private under pseudo-metric $\sqrt{\rho}$.*

Proof. This follows directly from Theorems 3.5.1 and the composition property as the algorithm samples from the posterior distribution, which is differentially private. □

Utility and optimal responses. We assume a collection of utility functions $\mathcal{U} = \{u_\theta : \theta \in \Theta\}$, such that the optimal response for a given parameter θ is the one that maximises a utility function $u_\theta : \mathcal{Q} \times \mathcal{R} \to [0, 1]$. If we know the true parameter θ, then we should respond to any query q with $r \in \arg\max_r u_\theta(q, r)$. However, since θ is unknown, we must select a method for conveying the required information. In a Bayesian setting, there are three main approaches we could employ. The standard methodology is to maximise *expected utility* with respect to the posterior. This corresponds to marginalising out θ, and responding with:

$$r_t \in \arg\max_r \int_\Theta u_\theta(q_t, r) \,\mathrm{d}\xi(\theta \mid x) \ .$$

The second is to use the *maximum a posteriori* value of θ. The final, which we employ here, is to use sampling; i.e. to reply to each query using parameters sampled from the posterior. This allows us to reply to arbitrary queries without compromising privacy, since the most information an adversary could obtain is the set of sampled parameters. By adjusting the number of samples used, we can easily trade off between privacy and utility.

After this we respond to a series of queries. For the t-th received query q_t, the algorithm returns the optimal response over the sampled parameter set $\hat{\Theta}$, in the manner shown in Algorithm 3.5.2. Since we allow arbitrary queries, the third party could simply ask for $\hat{\Theta}$ with a suitable choice of the utility function. Then if u is bounded, it is easy to show that the loss due to sampling is bounded.

Algorithm 3.5.2: PSAQR: Posterior Sample Query Response

```
input Parameter sample $\hat{\Theta}$.
for $t = 1, \dots$ do
    Observe query $q_t \in Q$, perhaps depending on $r_1, \dots, r_{t-1}$ and $q_1, \dots, q_{t-1}$.
    return $r_t \in \arg\max_r \sum_{\theta \in \hat{\Theta}} u_\theta(q_t, r)$
end for
```

Lemma 3.5.4. *The returned responses of PSAQR have a utility which is within* $O\left(\sqrt{\ln(1/\delta)/N}\right)$ *of the optimal value with probability at least* $1 - \delta$ *for any* $\delta > 0$.

Proof. Sampling N times from the posterior gives us the estimate of the utility function

$$\hat{u}_\xi(q, r) = \frac{1}{N} \sum_{\theta \in \hat{\Theta}} u_\theta(q, r),$$

which with probability at least $1 - \delta$ satisfies $|\hat{u}_\xi(q, r) - u(q, r)| < \sqrt{\frac{\ln(2/\delta)}{2N}} = \epsilon$, $\forall r, q$, via Hoeffding's inequality and the boundedness of u. Consequently, we can be at most 2ϵ-away from the optimal. □

Now that we have shown bounds on the utility for the algorithm above, we turn to the issue of how utility and privacy can be optimally tuned. First, we try to quantify the amount of samples an adversary needs to distinguish two datasets.

3.5.3 Distinguishability of Datasets

We want to relate the size of the sample $\hat{\Theta}$ and the amount of information about x that can be obtained by the adversary $\mathscr{A}$. Specifically, how well $\mathscr{A}$ can distinguish x from all alternative datasets y should be bounded. $\mathscr{A}$ has to decide whether $\mathscr{B}$'s posterior is $\xi(\cdot \mid x)$ or $\xi(\cdot \mid y)$ within the posterior sampling query model but he can only do so within some neighbourhood ϵ of the original data. Here, we bound the error $\mathscr{A}$ made in determining the posterior in terms of the number of samples used. This is similar to the dataset-size bounds on queries in interactive models of differential privacy [Dwork et al., 2006] and the privacy as hypothesis testing [Kairouz et al., 2015, Wasserman and Zhou, 2010] where an adversary wants to distinguish the dataset from two alternatives.

For this section, a utility function with an optimal response of $\hat{\Theta}$ is considered and this corresponds to the most powerful query possible based on the model shown in Algorithm 3.5.2. Then the adversary can approximate the posterior up to some sample errors by forming the empirical distribution. As a result, his power measured by the number of samples needed to distinguish between x and y is governed by the bounds on the KL-divergence between the empirical and actual distributions.

We first need a finite sample bound on the quality of the empirical distribution due to the sampling model. By constructing the empirical distribution on any sub-algebra $\mathfrak{S}$, the adversary could try to differentiate different posteriors.

Lemma 3.5.5. *For any $\delta \in (0, 1)$, let $\mathscr{M}$ be a finite partition of the sample space $\mathcal{S}$, of size $m \le \log_2 \sqrt{1/\delta}$, generating the σ-algebra $\mathfrak{S} = \sigma(\mathscr{M})$. Let $x_1, \dots, x_n \sim P$ be i.i.d. samples from a probability measure P on $\mathcal{S}$, let $P_{|\mathfrak{S}}$ be the restriction of P on $\mathfrak{S}$ and let $\hat{P}^n_{|\mathfrak{S}}$ be the empirical measure on $\mathfrak{S}$. Then, with probability at least $1 - \delta$,*

$$\left\|\hat{P}^n_{|\mathfrak{S}} - P_{|\mathfrak{S}}\right\|_1 \le \sqrt{\frac{3}{n} \ln \frac{1}{\delta}} \ . \tag{3.32}$$

Proof. (Note that in this proof, ε, δ do not refer to the privacy parameters.) We use the inequality due to Weissman et al. [2003] on the ℓ_1 norm, which states that for any multinomial distribution P with m outcomes, the ℓ_1 deviation of the empirical distribution $\hat{P}_n$ after n draws from the multinomial satisfies

$$\mathbb{P}\left(\left\|\hat{P}_n - P\right\|_1 \ge \varepsilon\right) \le (2^m - 2)e^{-\frac{1}{2}n\varepsilon^2}, \qquad \forall \varepsilon > 0 \ .$$

The right hand side is bounded by $e^{m \ln 2 - \frac{1}{2}n\varepsilon^2}$. Substituting $\varepsilon = \sqrt{\frac{3}{n} \ln \frac{1}{\delta}}$, we obtain

$$\begin{aligned}\mathbb{P}\left(\left\|\hat{P}_n - P\right\|_1 \ge \sqrt{\frac{3}{n} \ln \frac{1}{\delta}}\right) &\le e^{m \ln 2 - \frac{3}{2} \ln \frac{1}{\delta}} \\ &\le e^{\log_2 \sqrt{\frac{1}{\delta}} \ln 2 - \frac{3}{2} \ln \frac{1}{\delta}} = e^{\frac{1}{2} \ln \frac{1}{\delta} - \frac{3}{2} \ln \frac{1}{\delta}} = \delta \ ,\end{aligned}$$

where the second inequality follows from $m \leq \log_2 \sqrt{1/\delta}$. □

We can combine this bound on the adversary's estimation error with the bound in Theorem 3.4.1 on the KL-divergence between posteriors resulting from similar data to obtain a measure of how fine a distinction between datasets the adversary can make after a finite number of draws from the posterior.

Theorem 3.5.6. *Under Assumption 1, the adversary can distinguish between data* x, y *with probability* $1 - \delta$ *if*

$$\rho(x, y) \geq \frac{3}{4Ln} \ln \frac{1}{\delta} \ .$$

Under Assumption 2, this becomes

$$\rho(x, y) \geq \frac{3}{2n\left(C_{\xi}^{\mathcal{F}_\Theta} + 2L_0 + c^{-1}\right)} \ln \frac{1}{\delta} \ .$$

Proof. Recall that the data processing inequality states that, for any sub-algebra $\mathfrak{S}$,

$$\left\| Q_{|\mathfrak{S}} - P_{|\mathfrak{S}} \right\|_1 \leq \| Q - P \|_1 \ .$$

Using this and Pinsker's inequality (3.31) we get

$$\begin{aligned} 2L\rho(x, y) &\geq D\left(\xi(\cdot \mid x) \| \xi(\cdot \mid y)\right) \\ &\geq \frac{1}{2} \| \xi(\cdot \mid x) - \xi(\cdot \mid y) \|_1^2 \\ &\geq \frac{1}{2} \left\| \xi_{|\mathfrak{S}}(\cdot \mid x) - \xi_{|\mathfrak{S}}(\cdot \mid y) \right\|_1^2 \ . \end{aligned}$$

On the other hand, due to (3.32) the adversary's ℓ_1 error in the posterior distribution is bounded by $\sqrt{\frac{3}{n} \ln \frac{1}{\delta}}$ with probability $1 - \delta$. In order for him to be able to distinguish the two different posteriors, it must hold that

$$\left\| \xi_{|\mathfrak{S}}(\cdot \mid x) - \xi_{|\mathfrak{S}}(\cdot \mid y) \right\|_1 \geq \sqrt{\frac{3}{n} \ln \frac{1}{\delta}} \ .$$

Using the above inequalities, we can bound the error in terms of the distinguishability of the real dataset x from an arbitrary set y as

$$4L\rho(x, y) \geq \frac{3}{n} \ln \frac{1}{\delta} \ .$$

Rearranging, we obtain the required result. The second case is treated similarly to obtain

$$\left(C_{\xi}^{\mathcal{F}_{\Theta}} + 2L_0 + c^{-1}\right)\rho(x, y) \geq \frac{1}{2} \left\|\xi_{|\mathfrak{S}}(\cdot \mid x) - \xi_{|\mathfrak{S}}(\cdot \mid y)\right\|_1^2 \geq \frac{3}{2n} \ln \frac{1}{\delta} \ .$$

Consequently, either smoother likelihoods (i.e. decreasing L), or a larger concentration on smoother likelihoods (i.e. increasing c), increases the effort required by the adversary and reduces the sensitivity of the posterior. Note that, unlike the results obtained for differential privacy of the posterior sampling mechanism, these results have the same algebraic form under both assumptions.

3.5.4 Trading off Utility and Privacy

By construction, in our setting there are three ways to tune privacy. The first is the choice of family; the second is the choice of prior; and the third is how many samples N to draw. The choice of family is usually fixed due to other considerations. However, we have the choice of either tuning the prior, so that we can satisfy our assumptions with some suitable constants L or c, or by tuning the number of samples N in the posterior sampling framework.

The following lemma bounds the regret we suffer in terms of utility when the private posterior used is ξ, in the case where the posterior we would like to use (assuming no privacy constraints) was $\xi^\star$.

Lemma 3.5.7. *If our utility is bounded in* $[0, 1]$*, the private posterior we use is* ξ*, while the ideal posterior is* $\xi^\star$*, then the regret suffered is bounded by* $2\|\xi - \xi^\star\|_1$*.*

Proof. Let $r, r^\star$ be the optimal responses under $\xi, \xi^\star$ respectively. For notational convenience, let $u_\xi = \int_\Theta u_\theta \, \mathrm{d}\xi(\theta)$ denote the expected utility under a belief ξ. Then our regret is

$$\begin{aligned} u_\xi(q, r) - u_\xi(q, r^\star) &= u_\xi(q, r) - u_{\xi^\star}(q, r) \\ &\quad + u_{\xi^\star}(q, r) - u_{\xi^\star}(q, r^\star) \end{aligned}$$

$$
\begin{aligned}
&+ u_{\xi^\star}(q, r^\star) - u_\xi(q, r^\star) \\
&\le 2 \left\| \xi - \xi^\star \right\|_1 \ .
\end{aligned}
$$

This follows from the fact that

$$
\begin{aligned}
u_\xi(q, r) - u_{\xi^\star}(q, r) &= \int_\Theta u_\theta(q, r) \,\mathrm{d}[\xi - \xi^\star](\theta) \\
&\le \|u\|_\infty \|\xi - \xi^\star\|_1
\end{aligned}
$$

and the boundedness of u. The third term is dealt with identically. Note that for the second term: $u_{\xi^\star}(q, r) - u_{\xi^\star}(q, r^\star) \le 0$ since $r^\star$ maximises $u_{\xi^\star}$. □

Lastly, consider the case where $\mathscr{B}$, being a true Bayesian, is convinced that $\xi^\star$ is the correct prior distribution to use, but needs to use the prior ξ in order to achieve privacy. The following theorem bounds the expected KL-divergence between the two resulting posteriors.

Lemma 3.5.8. *If $\forall \theta \in \Theta$, $|\ln \xi^\star(\theta)/\xi(\theta)| \le \eta$, then the expected KL-divergence is*

$$
\mathrm{E}_{x \sim \phi^\star} D(\xi^\star(\cdot \mid x) \| \xi(\cdot \mid x)) \le 2\eta \ ,
$$

where $\phi^\star$ is the $\xi^\star$ marginal distribution.

Proof. Let $\phi^\star(x) = \int_\Theta p_\theta(x) \,\mathrm{d}\xi^\star(x)$ be the prior marginal distribution. Then the $\xi^\star$-expected KL-divergence between the two posteriors is

$$
\begin{aligned}
&\sum_x \int_\Theta \ln \frac{\mathrm{d}\xi^\star(\theta \mid x)}{\mathrm{d}\xi(\theta \mid x)} \,\mathrm{d}\xi^\star(\theta \mid x) \phi^\star(x) \\
&\le \sum_x \int_\Theta \left(\left| \ln \frac{\mathrm{d}\xi^\star(\theta)}{\mathrm{d}\xi(\theta)} \right| + \left| \ln \frac{\phi(x)}{\phi^\star(x)} \right| \right) \mathrm{d}\xi^\star(\theta \mid x) \phi^\star(x) \\
&\le 2\eta \ .
\end{aligned}
$$

The first term $\left| \ln \frac{\mathrm{d}\xi^\star(\theta)}{\mathrm{d}\xi(\theta)} \right|$ is bounded by η by assumption. From the same assumption, it follows that $\phi(x) = \int_\Theta p_\theta(x) \,\mathrm{d}\xi(\theta) \le \int_\Theta p_\theta(x) e^\eta \,\mathrm{d}\xi^\star(\theta) = e^\eta \phi^\star(x)$, and so the second term is also bounded by η. □

We can now combine Lemmas 3.5.4 and 3.5.7 with Lemma 3.5.8, to obtain the following result:

Corollary 3.5.9. *If $\mathscr{A}$ has a preferred prior $\xi^\star$, while the private prior used by $\mathscr{B}$ is ξ and it satisfies the conditions of Lemma 3.5.8, then the loss of $\mathscr{A}$ in terms of the $\xi^\star$-expected utility is $O\left(\eta + \sqrt{\ln(1/\delta)/N}\right)$, with probability at least $1 - \delta$.*

Consequently, if $\mathscr{A}$ believes the correct prior should be $\xi^\star$, he can use the private posterior sample to make decisions, incurring a small loss. Finally, we already showed that $\mathscr{A}$ cannot distinguish between data that are closer than $O(1/N)$ with high probability. Hence, in this setting we can tune N to trade off utility and privacy.

The following theorem characterises the link between the choice of prior, the number of samples, privacy and utility directly. This connects several of our results in one place.

Theorem 3.5.10. *If, instead of using a non-private prior $\xi^\star$, we use a prior ξ restricted on Θ_L (such that it satisfies Assumption 1 with constant L) and generate N samples from the posterior, then*

(a) the sample is $2LN$ differentially-private;

(b) the loss of $\mathscr{A}$ in terms of the $\xi^\star$-expected utility is $O\left([1 - \xi^\star(\Theta_L)] + \sqrt{\ln(1/\delta)/N}\right)$, with probability at least $1 - \delta$ for any $\delta > 0$.

Proof. For (a), due to composition, N repetitions give $2LN$ differential privacy. For (b), let Θ_L be the support of ξ. Since ξ is the restriction of $\xi^\star$ on Θ_L, it holds that

$$\begin{aligned}\left\|\xi - \xi^\star\right\|_1 &= \xi(\Theta_L) - \xi^\star(\Theta_L) + \xi^\star(\Theta \setminus \Theta_L) - \xi(\Theta \setminus \Theta_L)\\ &= 2[1 - \xi^\star(\Theta_L)] \ .\end{aligned}$$

We now just need to couple this with Lemmas 3.5.7 and 3.5.4 to directly obtain the stated bound on the utility. □

In practice, our choice of ξ gives us a base amount of privacy that depends only on L. By keeping ξ fixed and increasing N, we can easily trade off privacy and utility.

Finally, we should note that the adversary could choose any arbitrary estimator ψ to guess x. Section 3.5.5 below describes how to apply Le Cam's method to obtain matching lower bounds in this case, by defining *dataset estimators* as a model for the adversary.

3.5.5 Lower Bounds

It is possible to apply standard minimax theory to obtain lower bounds on the rate of convergence of the adversary's estimate to the true data. In order to do so, we can for example apply the method due to Le Cam [1973], which places lower bounds on the expected distance between an estimator and the true parameter. In order to apply it in our case, we simply replace the the role parameter space with the dataset space.

Le Cam's method assumes the existence of a family of probability measures indexed by some parameters, with the parameter space being equipped with a pseudo-metric. In our setting, we use Le Cam's method in a slightly unorthodox but very natural manner. Define the family of probability measures on Θ to be

$$\Xi \triangleq \{ \xi(\cdot \mid x) : x \in \mathcal{S} \},$$

the family of posterior measures in the parameter space, for a specific prior ξ. Consequently, now $\mathcal{S}$ plays the role of the parameter space, while ρ is used as the pseudo-metric. The original family $\mathcal{F}_\Theta$ plays no further role in this construction, other than a way to specify the posterior distributions from the prior.

Now let ψ be an arbitrary estimator of the unknown data x. As in [Le Cam, 1973], we extend ρ to subsets of $\mathcal{S}$ via

$$\rho(A, B) \triangleq \inf \{ \rho(x, y) : x \in A, y \in B \} \ , \quad A, B \subset \mathcal{S} \ .$$

Now we can re-state the following well-known lemma for our specific setting.

Lemma 3.5.11 (Le Cam's method)**.** *Let ψ be an estimator of x on Ξ taking values in the metric space $(\mathcal{S}, \rho)$. Suppose that there are well-separated subsets $\mathcal{S}_1, \mathcal{S}_2$ such that $\rho(\mathcal{S}_1, \mathcal{S}_2) \geq 2\delta$. Suppose also that Ξ_1, Ξ_2 are subsets of Ξ such that $x \in \mathcal{S}_i$ for*

$\xi(\cdot \mid x) \in \Xi_i$. *Then*

$$\sup_{x\in\mathcal{S}} \mathrm{E}_{\xi}(\rho(\psi, x) \mid x) \geq \delta \sup_{\xi_i \in co(\Xi_i)} \|\xi_1 \wedge \xi_2\| \ .$$

This lemma has an interesting interpretation in our case. The quantity

$$\mathrm{E}_{\xi}(\rho(\psi, x) \mid x) = \int_{\Theta} \rho(\psi(\theta), x)\, \mathrm{d}\xi(\theta \mid x)$$

is the expected distance between the real data x and the guessed data $\psi(\theta)$ when θ is drawn from the posterior distribution. Consequently, it is possible to apply this method directly to obtain results for specific families of posteriors. These would be dependent on the family, the prior and the metric. While we shall not engage in this exercise, we point the interested reader to [Yu et al., 1997], which provides two simple examples with minimax rates of $O(n^{-4/9})$ and $O(n^{-4/5})$.

3.6 Examples Satisfying our Assumptions

In what follows we study, for different choices of likelihood and corresponding conjugate prior, what constraints can be placed on the prior's concentration to guarantee a desired level of privacy. These case studies closely follow the pattern in differential privacy research where the main theorem for a new mechanism are sufficient conditions on (*e.g.*, Laplace) noise levels to be introduced to a response in order to guarantee a level ϵ of ϵ-differential privacy.

For exponential families, we have the canonical form

$$p_{\theta}(x) = h(x) \exp\left\{\eta_{\theta}^{\top} \tau(x) - A(\eta_{\theta})\right\},$$

where $h(x)$ is the base measure, η_θ is the distribution's natural parameter corresponding to θ, $\tau(x)$ is the distribution's sufficient statistic, and $A(\eta_\theta)$ is its log-partition function. For distributions in this family, under the absolute log-ratio distance, the family of parameters Θ_L of Assumption 2 must satisfy, for all $x, y \in \mathcal{S}$,

$$\left|\ln \frac{h(x)}{h(y)} + \eta_{\theta}^{\top} \left(\tau(x) - \tau(y)\right)\right| \leq L\rho(x, y).$$

If the left-hand side has an amenable form, then we can quantify the set Θ_L for which this requirement holds. Particularly, for distributions where $h(x)$ is constant and $\tau(x)$ is scalar (*e.g.*, Bernoulli, exponential, and Laplace), this requirement simplifies to $\frac{|\tau(x)-\tau(y)|}{\rho(x,y)} \leq \frac{L}{\eta_\theta}$. One can then find the supremum of the left-hand side independent from θ, yielding a simple formula for the feasible L for any θ. In the following examples, we are making the conventional assumption in machine learning that data are bounded ($\|x\| \leq B$). Also we use $\xi(\theta)\mathbb{1}_{[c_1,c_2]}$ to denote the trimmed density function that densities of $\xi(\theta)$ outside $[c_1, c_2]$ is projected to c_1 or c_2.

Lemma 3.6.1 (Exponential-Exponential conjugate prior). *The exponential distribution $\mathcal{Exp}(x;\theta)$ with a trimmed exponential conjugate prior $\theta \sim \mathcal{Exp}(\theta;\lambda)\mathbb{1}_{[c_1,c_2]}$, $\lambda > 0$, satisfies Assumption 2 with parameter $c = \lambda$, $L_0 = c_1$, $C_\xi^{\mathcal{F}_\Theta} = c_2 / \min\left\{c_1 e^{-c_1 B}, c_2 e^{-c_2 B}\right\}$ and metric $\rho(x,y) = |x-y|$.*

Consequently, the trimmed-exponential prior results in a posterior sampling mechanism that is $(0,\delta)$-DP under ρ, with $\delta = \sqrt{\frac{1}{2} C_\xi^{\mathcal{F}_\Theta}(1 + 2c_1 + 1/\lambda)}$. It is also $(0,\delta)$-DP under the classical definition if $x, y \in [0,1]$.

Proof. Since $Exp(x;\theta)$ is monotonic decreasing in x and concave as a function of θ, we have $\inf_{\{\|x\| \leq B, \theta \in [c_1,c_2]\}} Exp(x;\theta) = \min\left\{c_1 e^{-c_1 B}, c_2 e^{-c_2 B}\right\} \leq \phi(x)$. Then we have

$$C_\xi^{\mathcal{F}_\Theta} = c_2 / \min\left\{c_1 e^{-c_1 B}, c_2 e^{-c_2 B}\right\} \ .$$

Next we compute the absolute log-ratio distance for any x_1 and x_2 according to the exponential likelihood function:

$$|\ln p_\theta(x_1) - \ln p_\theta(x_2)| = \theta |x_1 - x_2| \ .$$

Thus, for $\theta \in [c_1, c_2]$, under Assumption 2, using $\rho(x,y) = |x-y|$, the set of feasible parameters for any $L > c_1$ is $\Theta_L = (c_1, L)$. Note the density of the renormalized exponential prior on $[c_1, c_2]$ is given by $K\lambda e^{-\lambda\theta}$, where $K = (e^{-\lambda c_1} - e^{-\lambda c_2})^{-1}$. Thus the CDF at L of this density is $K\left(e^{-\lambda c_1} - e^{-\lambda L}\right)$ for $L \in [c_1, c_2]$ and 1 for $L \geq c_2$. It is natural to choose L_0 to be c_1. Then we need to find c such that

$$\xi(\Theta_L) = \int_{c_1}^{L} K\lambda e^{-\lambda\theta} d\theta = K(e^{-\lambda c_1} - e^{-\lambda L}) \geq 1 - e^{-c(L-c_1)}$$

for $L \in (c_1, c_2)$. By plugging K into the inequality, we have

$$e^{-c(L-c_1)} \geq \frac{e^{-\lambda(L-c_2)} - 1}{e^{-\lambda(c_1-c_2)} - 1} \ .$$

Since $e^{-\lambda(L-c_2)} \leq e^{-\lambda(c_1-c_2)}$, it is sufficiency to find c such that $e^{-c(L-c_1)} \geq e^{-\lambda(L-c_1)}$. Therefore we can have $c = \lambda$. □

Lemma 3.6.2 (Laplace-Exponential conjugate prior). *The distribution Laplace$(x; s, \mu)$ with a trimmed exponential conjugate prior $1/s = \theta \sim Exp(\theta; \lambda)\mathbb{1}_{[c_1,c_2]}$, $\mu \in \mathbb{R}$, $s \geq 1/L$, $\lambda > 0$ satisfies Assumption 2 with parameters $c = \lambda$, $L_0 = c_1$,*

$$C_\xi^{\mathcal{F}_\Theta} = \begin{cases} \frac{c_2}{2\min\left\{\frac{1}{2c_2}, \frac{1}{2c_1}\exp\left(\frac{-B-\mu}{c_1}\right)\right\}} \ , & x < \mu \\ \frac{c_2}{2\min\left\{\frac{1}{2c_2}, \frac{1}{2c_1}\exp\left(\frac{\mu-B}{c_1}\right)\right\}} \ , & x \geq \mu \end{cases} \ ,$$

and metric $\rho(x, y) = |x - y|$.

Proof. Note that *Laplace*$(x; s, \mu)$ is monotonic decreasing in x if $x < \mu$, and increasing in x if $x \geq \mu$. Since *Laplace*$(x; s, \mu)$ is concave as a function of s, we have $\phi(t) \geq \min\left\{\frac{1}{2c_2}, \frac{1}{2c_1}\exp\left(\frac{-B-\mu}{c_1}\right)\right\}$ if $x < \mu$ and $\phi(t) \geq \min\left\{\frac{1}{2c_2}, \frac{1}{2c_1}\exp\left(\frac{\mu-B}{c_1}\right)\right\}$ if $x \geq \mu$. Thus, we can take

$$C_\xi^{\mathcal{F}_\Theta} = \begin{cases} \frac{c_2}{2\min\left\{\frac{1}{2c_2}, \frac{1}{2c_1}\exp\left(\frac{-B-\mu}{c_1}\right)\right\}} \ , & x < \mu \\ \frac{c_2}{2\min\left\{\frac{1}{2c_2}, \frac{1}{2c_1}\exp\left(\frac{\mu-B}{c_1}\right)\right\}} \ , & x \geq \mu \end{cases} \ .$$

For any x_1 and x_2, the absolute log-ratio distance for this distribution can be bounded as

$$\begin{aligned} &|\ln p_{\mu,s}(x_1) - \ln p_{\mu,s}(x_2)| \\ =&\frac{1}{s}\,\big|\|x_1 - \mu\| - \|x_2 - \mu\|\big| \leq \frac{1}{s}\|x_1 - x_2\| \ , \end{aligned}$$

where the inequality follows from the triangle inequality on $\|\cdot\|$. Thus, if we use $\rho(x, y) = \|x - y\|$, the set of feasible parameters for Assumption 2 is $\mu \in \mathbb{R}$ and $\frac{1}{s} = \theta \leq L$. Again we can use the trimmed exponential prior with rate parameter $\lambda > 0$ for the inverse scale, $\frac{1}{s}$, and similar to the previous example, Assumption 2 is satisfied with $c = \lambda$ and $L_0 = c_1$. □

Lemma 3.6.3 (Beta-Binomial conjugate prior). *The Binomial distribution $\mathcal{B}inom(\theta, n)$, with prior $\theta \sim \mathcal{B}eta(\alpha,\beta)$, $\alpha = \beta > 1$ satisfies Assumption 2 for $L_0 = \ln n$, $c = 2^{-2\alpha+1}/B(\alpha)$, where $B(\alpha)$ denotes the beta function with parameters $\alpha = \beta$,*

$$C_{\xi}^{\mathcal{F}_\Theta} = B(\alpha)/B\left(\frac{n+2\alpha-1}{2}, \frac{n+2\alpha+1}{2}\right)$$

and metric $\rho(x,y) = \|x-y\|_1$, where $x,y \in \{0,1\}^n$.

Proof. Here, we consider data drawn from a Binomial distribution with a beta prior on its proportion parameter, θ. Thus, the likelihood and prior functions are

$$\begin{aligned} p_{\theta,n}(X = k) &= \tbinom{n}{k}\theta^k(1-\theta)^{n-k} \\ \xi_0(\theta) &= \tfrac{1}{B(a,b)}\theta^{a-1}(1-\theta)^{b-1} \end{aligned} ,$$

where $k \in \{0,1,2,\ldots,n\}$, $a,b \in \mathbb{R}_+$ and $B(a,b)$ is the beta function. The resulting posterior is a Beta-Binomial distribution. Again we consider the application of Assumption 2 to this Beta-Binomial distribution. For this purpose, we must quantify the parameter sets Θ_L for a given $L > 0$ according to a distance function. The absolute log-ratio distance between the Binomial likelihood function for any pair of arguments, k_1 and k_2, is

$$|\ln p_{\theta,n}(k_1) - \ln p_{\theta,n}(k_2)| = \left|\Delta_n(k_1,k_2) + (k_1-k_2)\ln\tfrac{\theta}{1-\theta}\right|$$

where $\Delta_n(k_1,k_2) \triangleq \ln\binom{n}{k_1} - \ln\binom{n}{k_2}$. By substituting this distance into the supremum of Eq. (3.4), we seek feasible values of $L > 0$ for which the supremum is non-negative; here, we explore the case where $\rho((n,k_1),(n,k_2)) \triangleq |k_1-k_2|$. Without loss of generality, we assume $k_1 > k_2$, and thus require that

$$\sup_{k_1>k_2}\left|\tfrac{\Delta_n(k_1,k_2)}{k_1-k_2} + \ln\tfrac{\theta}{1-\theta}\right| \leq L \ . \tag{3.33}$$

However, by the definition of $\Delta_n(k_1,k_2)$, the ratio $\frac{\Delta_n(k_1,k_2)}{k_1-k_2}$ is in fact the slope of the chord from k_2 to k_1 on the function $\ln\binom{n}{k}$. Since the function $\ln\binom{n}{k}$ is concave in

k, this slope achieves its maximum and minimum at its boundary values; i.e. it is maximised for $k_1 = 1$ and $k_2 = 0$ and minimised for $k_1 = n$ and $k_2 = n - 1$. Thus, the ratio attains a maximum value of $\ln n$ and a minimum of $-\ln n$ for which the above supremum is simply $\ln n + \left|\ln \frac{\theta}{1-\theta}\right|$. From Eq. (3.33), we therefore have, for all $L \geq \ln n$:

$$\Theta_L = \left[\left(1 + \tfrac{e^L}{n}\right)^{-1}, \left(1 + \tfrac{n}{e^L}\right)^{-1}\right] \ .$$

We want to bound $\xi(\Theta_L)$. We know that: $\xi(\Theta_L) = 1 - \xi\left(\Theta_L^\complement\right)$ where $\Theta_L^\complement$ is the complement of Θ_L. So $\xi(\Theta_L^\complement)$ is composed of two symmetric intervals: $\left[0, \left(1 + \frac{e^L}{n}\right)^{-1}\right)$ and $\left(\left(1 + \frac{n}{e^L}\right)^{-1}, 1\right]$. We selected $\alpha = \beta$, therefore the mass must concentrate at $1/2$, as we have $\alpha > 1$.

Due to symmetry, the mass outside of Θ_L is two times that is the first interval. This is:

$$\frac{2}{B(\alpha, \alpha)} \int_0^{\frac{p}{1+p}} x^{\alpha-1}(1-x)^{\alpha-1}\,\mathrm{d}x \ .$$

where p denotes $ne^{-L} \in [0, 1]$, Therefore c is upper bounded by

$$\ln\left(\frac{2A(p)}{B(\alpha, \alpha)}\right) / (L_0 - L) = \ln\left(\frac{2A(p)}{B(\alpha, \alpha)}\right) / \ln p,$$

where $A(p)$ denotes the incomplete Beta function $\int_0^{\frac{p}{1+p}} x^{\alpha-1}(1-x)^{\alpha-1}dx$. Note that we have

$$A'(p) = \frac{p^{\alpha-1}}{(1+p)^{2\alpha}} \ ,$$

$$A''(p) = \frac{p^{\alpha-2}[(\alpha-1)(1+p) - 2\alpha p]}{(1+p)^{2\alpha+1}} \ .$$

Claim 2. $H(p) = \alpha A(p) - \frac{p^\alpha}{(1-p)(1+p)^{2\alpha-1}} \leq 0$ *for all* $p \in (0, 1)$.

Proof. Calculating derivatives and simplifying

$$\begin{aligned} & H'(p) \\ = \ & \alpha A'(p) - \frac{\alpha p^{\alpha-1}(1-p)(1+p)^{2\alpha-1} - p^\alpha\left[(2\alpha-1)(1-p)(1+p)^{2\alpha-2} - (1+p)^{2\alpha-1}\right]}{[(1-p)(1+p)^{2\alpha-1}]^2} \end{aligned}$$

$$\begin{aligned}
&= \frac{\alpha p^{\alpha-1}}{(1+p)^{2\alpha}} - \frac{\alpha p^{\alpha-1}(1-p)(1+p) - p^{\alpha}[(2\alpha-1)(1-p)-(1+p)]}{(1-p)^2(1+p)^{2\alpha}} \\
&= \frac{p^{\alpha-1}}{(1+p)^{2\alpha}}\left(\alpha - \frac{\alpha(1-p^2)-2p(\alpha-1-p\alpha)}{(1-p)^2}\right) \\
&= \frac{p^{\alpha-1}}{(1+p)^{2\alpha}(1-p)^2}\left(\alpha(1-2p+p^2)-\alpha(1-p^2)+2p(\alpha-1-\alpha p)\right) \\
&= \frac{-2p^{\alpha}}{(1+p)^{2\alpha}(1-p)^2} < 0\ .
\end{aligned}$$

Therefore $H(p)$ is strictly decreasing. Then combined with $H(0) = 0$, we claim follows. □

Claim 3. *$G(p) = p\frac{A'(p)}{A(p)}\ln p - \ln\frac{2A(p)}{B(\alpha,\alpha)} < 0$ for all $p \in (0,1)$.*

Proof. Again taking derivatives

$$\begin{aligned}
G'(p) &= \frac{A'(p)}{A(p)}(1+\ln p) + p\ln p\frac{A''(p)A(p)-A'(p)^2}{A(p)^2} - \frac{A'(p)}{A(p)} \\
&= \frac{\ln p}{A(p)^2}(A(p)A'(p)+pA''(p)A(p)-pA'(p)^2) \\
&= \frac{\ln p}{A(p)^2}\left[\frac{p^{\alpha-1}}{(1+p)^{2\alpha}}A(p)\left(1+\frac{(\alpha-1)(1+p)-2\alpha p}{1+p}\right)-\frac{p^{2\alpha-1}}{(1+p)^{4\alpha}}\right] \\
&= \frac{\ln p}{A(p)^2}\frac{p^{\alpha-1}}{(1+p)^{2\alpha+1}}\left[\alpha(1-p)A(p)-\frac{p^{\alpha}}{(1+p)^{2\alpha-1}}\right] \\
&= \frac{p^{\alpha-1}}{(p+1)^{2\alpha+1}A(p)^2}H(p)\ln p(1-p) > 0\ .
\end{aligned}$$

So $G(p)$ is strictly increasing. Combined with $\lim_{p\to1} G(p) = 0$, the claim follows. □

Claim 4. *$F(p) = \ln\left(2I_{\frac{p}{1+p}}(\alpha)\right)/\ln p$ is decreasing in $p \in (0,1)$, where the incomplete Beta function $I_{\frac{p}{1+p}}(\alpha) = A(p)/B(\alpha,\alpha)$.*

Proof. Taking derivatives

$$\begin{aligned}
F'(p) &= \frac{1}{\ln^2 p}\left(\frac{A'(p)}{A(p)}\ln p - \frac{1}{p}\ln\frac{2A(p)}{B(\alpha,\alpha)}\right) \\
&= \frac{1}{p\ln^2 p}\left(\frac{A'(p)}{A(p)}p\ln p - \ln\frac{2A(p)}{B(\alpha,\alpha)}\right)
\end{aligned}$$

$$= \frac{1}{p\ln^2 p} G(p) < 0 \; .$$

Therefore $\ln\left(2I_{\frac{p}{1+p}}(\alpha)\right)/\ln p$ is monotonic decreasing in p. Thus the minimum value of $F(p)$ is $\frac{1}{B(\alpha)2^{2\alpha-1}}$ as $p \to 1$, which we can take as our c in this example.

Let us consider $C_{\xi}^{\mathcal{F}_\Theta}$ for this example. We have

$$\frac{p_\theta(x)}{\phi(x)} = \frac{B(\alpha,\beta)\theta^x(1-\theta)^{n-x}}{B(\alpha+x, n+\beta-x)} \; ,$$

where $\theta \in [0,1]$ and $x \in [0,1,\ldots,n]$. Note that

$$\frac{B(\alpha+x+1, n+\beta-x-1)}{B(\alpha+x, n+\beta-x)} = \frac{\Gamma(\alpha+x+1)\Gamma(n+\beta-x-1)}{\Gamma(\alpha+x)\Gamma(n+\beta+1)} = \frac{\alpha+x}{n+\beta-x-1} \; .$$

So $B(\alpha+x+1, n+\beta-x-1) \le B(\alpha+x, n+\beta-x)$ if $x \le \frac{n+\beta-\alpha-1}{2}$; $B(\alpha+x+1, n+\beta-x-1) > B(\alpha+x, n+\beta-x)$ otherwise. Thus

$$B(\alpha+x, n+\beta-x) \ge B\left(\frac{n+\alpha+\beta-1}{2}, \frac{n+\alpha+\beta+1}{2}\right) \; .$$

Hence we can take $C_{\xi}^{\mathcal{F}_\Theta} = B(\alpha,\beta)/B\left(\frac{n+\alpha+\beta-1}{2}, \frac{n+\alpha+\beta+1}{2}\right)$. □

We next present two results on normal distributions.

Lemma 3.6.4 (Normal distribution with known mean and unknown variance)**.** *The normal distribution* $N(x;\mu,\sigma^2)$ *with a trimmed exponential prior* $1/\sigma^2 = \theta \sim \mathcal{E}xp(\theta;\lambda)\mathbb{1}_{[c_1,c_2]}$ *satisfies Assumption 2 with parameter* $c = \frac{2\lambda}{\max\{|\mu|,1\}}$, $L_0 = \frac{c_1\max\{|\mu|,1\}}{2}$,

$$C_{\xi}^{\mathcal{F}_\Theta} = \min\left\{\sqrt{c_2/c_1}\exp\left(\frac{c_1c_2^2}{2}\right), \exp\left(\frac{c_2^3}{2}\right)\right\}$$

and metric $\rho(x,y) = \left|x^2-y^2\right| + 2|x-y|$.

Proof. Since $N(x;\mu,\theta)$ is decreasing in x^2 and concave as a function of θ. We have $\phi(t) \ge \inf_{\{x| \|x\| \le B\},\theta\in[c_1,c_2]} N(x;\mu,\theta) = \min\left\{\sqrt{\frac{c_1}{2\pi}}e^{\frac{-c_1c_2^2}{2}}, \sqrt{\frac{c_2}{2\pi}}e^{\frac{-c_2^3}{2}}\right\}$. Then we can take

$$C_{\xi}^{\mathcal{F}_\Theta} = \min\left\{\sqrt{c_2/c_1}e^{\frac{c_1c_2^2}{2}}, e^{\frac{c_2^3}{2}}\right\}$$

For the normal distribution, (3.4) requires: $2L\rho(x,y)\sigma^2 \ge |2\mu - x - y|\,|x-y|$. Taking the absolute log ratio of the Gaussian densities we have

$$\begin{aligned}&\frac{1}{2\sigma^2}\left|\left((x-\mu)^2 - (y-\mu)^2\right)\right| \\ \le&\frac{\max\{\,|\mu|, 1\,\}}{2\sigma^2}\left(\left|x^2 - y^2\right| + 2\,|x-y|\right).\end{aligned}$$

Consequently, we can set $\rho(x,y) = \left|x^2 - y^2\right| + 2\,|x-y|$ and $L(\mu,\sigma) = \frac{\max\{\,|\mu|,1\,\}}{2\sigma^2}$. Again, the trimmed exponential prior is given by $K\lambda e^{-\lambda\theta}$, where $K = (e^{-\lambda c_1} - e^{-\lambda c_2})^{-1}$. Thus the CDF at L of this density is $K\left(e^{-\lambda c_1} - e^{-\lambda L}\right)$ for $L \in [\frac{c_1\max\{\,|\mu|,1\,\}}{2}, \frac{c_2\max\{\,|\mu|,1\,\}}{2}]$ and 1 for $L \ge \frac{c_2\max\{\,|\mu|,1\,\}}{2}$. Thus the CDF at L of this density is $K\left(e^{-\lambda c_1} - e^{\frac{-2\lambda L}{\max\{\,|\mu|,1\,\}}}\right)$. We choose L_0 to be $\frac{c_1\max\{\,|\mu|,1\,\}}{2}$. Then we need to find c such that

$$\xi(\Theta_L) = \int_{c_1}^{L} K\lambda e^{-\lambda\theta}d\theta = K(e^{-\lambda c_1} - e^{-\lambda L}) \ge 1 - e^{-c\left(L - \frac{c_1\max\{\,|\mu|,1\,\}}{2}\right)}.$$

By plugging K into the inequality, we have

$$e^{-c\left(L - \frac{c_1\max\{\,|\mu|,1\,\}}{2}\right)} \ge \frac{e^{\frac{-2\lambda L}{\max\{\,|\mu|,1\,\}}+\lambda c_2} - 1}{e^{-\lambda(c_1-c_2)} - 1}.$$

Since $e^{-\lambda\left(\frac{2\lambda L}{\max\{\,|\mu|,1\,\}}-c_2\right)} \le e^{-\lambda(c_1-c_2)}$, it is sufficiency to find c such that

$$e^{-c\left(L - \frac{c_1\max\{\,|\mu|,1\,\}}{2}\right)} \ge e^{-\lambda\left(\frac{2L}{\max\{\,|\mu|,1\,\}}-c_1\right)}.$$

This is equivalent to have c satisfying

$$c\left(L - \frac{c_1\max\{\,|\mu|,1\,\}}{2}\right) \le \lambda\left(\frac{2L}{\max\{\,|\mu|,1\,\}} - c_1\right).$$

Then we can take $c = \frac{2\lambda}{\max\{\,|\mu|,1\,\}}$ to satisfy the above inequality. □

Unbounded observation spaces are generally a problem for privacy, even for finite parameter spaces, generally because likelihoods become vanishingly small, thus making log likelihood ratios arbitrarily large. However, the following two examples circumvent this problem. In the first example, we consider a general multivariate extension of Lemma 3.6.4.

Lemma 3.6.5 (Multivariate normal distribution)**.** *The multivariate normal distribution $N(x; \mu, A^{-1})$ satisfies our Assumption 1 with $L = \frac{1}{2}(\sum_{i=1}^{n} \lambda_i^2)^{\frac{1}{2}} \max\{1, \|\mu\|_2\}$ under metric $\rho(x, y) = \|xx^\top - yy^\top\|_F + 2\|x - y\|_2$. When $\mu = 0$, Assumption 1 is satisfied with $L = \frac{1}{2}(\sum_{i=1}^{n} \lambda_i^2)^{\frac{1}{2}}$ under metric $\rho(x, y) = \|(xx^\top - yy^\top)\|_F$.*

Proof. Consider the likelihood log-ratio distance of multivariate normal distributions with precision matrix A:

$$\frac{1}{2}|x^\top Ax - y^\top Ay| \ ,$$

where A is positive definite with eigenvalues $\lambda_1 \geq \ldots \geq \lambda_n > 0$). For simplicity, assume the mean to be a zero vector then

$$\begin{aligned}
|x^\top Ax - y^\top Ay| &= \left|\sum_{i,j} x_i x_j A_{i,j} - \sum_{i,j} y_i y_j A_{i,j}\right| \\
&= \left|\sum_{i,j} A_{i,j}(x_i x_j - y_i y_j)\right| \\
&= |Tr(A(xx^\top - yy^\top)')| \\
&\leq [Tr(A^2)Tr((xx^\top - yy^\top)(xx^\top - yy^\top)')]^{\frac{1}{2}} \\
&= \left(\sum_{i=1}^{n} \lambda_i^2\right)^{\frac{1}{2}} \|(xx^\top - yy^\top)\|_F \ .
\end{aligned}$$

For mean equal to μ, we have

$$\frac{1}{2}|(x^\top - \mu)A(x - \mu) - (y^\top - \mu)A(y - \mu)| \ .$$

By the above analysis we have the difference being bounded by

$$\frac{1}{2}\left(\sum_{i=1}^{n} \lambda_i^2\right)^{\frac{1}{2}} \|(x - \mu)(x - \mu)' - (y - \mu)(y - \mu)')\|_F \ .$$

Note that

$$\begin{aligned}\|(x-\mu)(x-\mu)' - (y-\mu)(y-\mu)')\|_F &= \|xx^\top - \mu(x^\top - y^\top) - (x-y)\mu' - yy^\top\|_F \\ &\leq \|xx^\top - yy^\top\|_F + 2\|\mu(x-y)'\|_F \\ &= \|xx^\top - yy^\top\|_F + 2\|\mu\|_2\|(x-y)'\|_2 \\ &\leq \max\{1, \|\mu\|_2\}(\|xx^\top - yy^\top\|_F + 2\|x-y\|_2) \ .\end{aligned}$$

The above examples demonstrate that our assumptions are reasonable. In fact, for several of them we recover standard choices of prior distributions.

3.7 Discussion

We have demonstrated a unified framework for private and secure inference under a Bayesian setting where the inference can be both robust and private under concentration conditions on the prior. Firstly, we prove that posterior distributions with small KL-divergence can be achieved by similar datasets. Secondly, we show that the posterior is differentially-private, which makes it possible for us to use a general posterior sampling mechanism to reply to queries. And at the same time, the desired trade-off between privacy and utility can be achieved easily by adjusting the quantity of samples used.

This framework might serve as a basic building block to enable further sophisticated and private Bayesian inference provided by the fact that there is no extra machinery being required. As an additional step toward this goal, we have shown how to derive analytical expressions for well-known distribution families and discrete Bayesian networks by using our framework. Finally, the amount of effort that an attacker requires to breach privacy when observing samples from posterior is bounded. This provides a principled guide for determining the appropriate level of access granted to query the posterior while ensuring privacy.

Chapter 4

Differential Privacy in Bayesian Networks

In this section, we extend our discussion on differential privacy in our Bayesian setting to *probabilistic graphical models* (PGM), which are popular as a tool for modelling conditional dependence assumptions. We develop the first set of mechanisms for Bayesian inference on the flexible probabilistic graphical model framework (*cf.* Table 4.1). Our mechanisms consider graph structure and include a purely Bayesian approach that only places conditions on the prior. To apply the posterior sampling mechanism on probabilistic graphical models, we show Assumption 1 (Assumption 2) of the previous chapter lift to graphs of random variables, and bound KL-divergence when releasing an empirical posterior based on a modified prior. We develop an alternate approach that uses the Laplace mechanism to perturb posterior parameterisations, and we apply techniques due to Barak et al. [2007], who released marginal tables that maintain consistency in addition to privacy, by adding Laplace noise to posterior updates in the Fourier domain. Our motivation is novel: we wish to guarantee privacy against omniscient attackers and stealth against unsuspecting third parties. We complement our study with a *maximum a posteriori* estimator that leverages the exponential mechanismdue to McSherry and Talwar [2007]. Our utility and privacy bounds connect privacy and graph/dependency structure, and are complemented by illustrative experiments with Bayesian naïve Bayes and Bayesian linear regression. Our mechanisms and theoretical bounds are the first to establish such a link between the graph structure of probabilistic graphical models and privacy.

	DBN only	**Privacy**	**Utility type**	**Utility bound**
Laplace	✓	$(\epsilon, 0)$	closeness of posterior	$O(mn \ln n)\left[1 - \exp\left(-\frac{n\epsilon}{2\lvert \mathrm{I} \rvert}\right)\right] + \sqrt{-O(mn \ln n) \ln \delta}$
Fourier	✓	$(\epsilon, 0)$	close posterior params	$\frac{4\lvert \mathcal{N}_{\mathrm{I}} \rvert}{\epsilon}\left(2^{\lvert \pi_i \rvert} \log \frac{\lvert \mathcal{N}_{\mathrm{I}} \rvert}{\delta} + t\lvert \mathcal{N}_{\mathrm{I}} \rvert\right)$
Sampler	✗	$(2L, 0)$ if Lipschitz; or $(0, \sqrt{M/2})$ stochastic Lipschitz	expected utility functional wrt posterior	$O\left(\eta + \sqrt{\ln(1/\delta)/N}\right)$ Dimitrakakis et al. [2017]
MAP	✗	$(\epsilon, 0)$	closeness of MAP	$\mathbb{P}(S_{2t}^{c}) \leq \exp(-\epsilon t)/\xi(S_t)$

Table 4.1: Summary of the privacy/utility guarantees for this chapter's mechanisms.

4.1 Problem Setting

Let us extend our setting in Chapter 3 to Bayesian networks and repeat the game for convenience. We assume that $\mathscr{B}$ is using *Bayesian inference* to draw conclusions from observations of a system of random variables by updating a prior distribution on parameters (i.e. *latent* variables) to a posterior. Still, our goal is to release an approximation to the posterior that preserves differential privacy.

Remark 4.1.1. *In Chapter 3, we use x to denote a dataset since the material is more abstract on random variables of a conjugate pair. In this chapter, we use* **D** *to denote a dataset as it refers to the observations on a system random variables.*

Consider a Bayesian statistician $\mathscr{B}$ estimating the parameters $\boldsymbol{\theta}$ of some family of distributions $\mathcal{F}_{\Theta} = \{p_{\boldsymbol{\theta}} : \boldsymbol{\theta} \in \Theta\}$ on a system of r.v.'s $\boldsymbol{X} = \{X_i : i \in \mathrm{I}\}$, where I is an index set, with observations denoted $x_i \in \mathcal{X}_i$, where $\mathcal{X}_i$ is the sample space of X_i. $\mathscr{B}$ has a prior distribution[1] ξ on Θ reflecting her prior belief, which she updates on an observation $\boldsymbol{x}$ to obtain posterior

$$\xi(B \mid \boldsymbol{x}) = \frac{\int_B p_{\boldsymbol{\theta}}(\boldsymbol{x}) \, \mathrm{d}\xi(\boldsymbol{\theta})}{\phi(\boldsymbol{x})}, \quad \forall B \in \mathfrak{S}_{\Theta}$$

where $\phi(\boldsymbol{x}) \triangleq \int_{\Theta} p_{\boldsymbol{\theta}}(\boldsymbol{x}) \, d\xi(\boldsymbol{\theta})$. Posterior updates are iterated over an *i.i.d.*dataset $D \in \mathcal{D} = (\prod_i \mathcal{X}_i)^n$ to $\xi(\cdot \mid D)$.

$\mathscr{B}$'s goal is to communicate her posterior distribution to a third party $\mathscr{A}$, while limiting the information revealed about the original data. From the point of view of the data provider, $\mathscr{B}$ is a trusted party.[2] However, she may still inadvertently reveal

[1] Precisely, a probability measure on a σ-algebra $(\Theta_i, \mathfrak{S}_{\Theta_i})$.

[2] Cryptographic tools for untrusted $\mathscr{B}$ do not prevent information leakage to $\mathscr{A}$ *cf. e.g.*, Pagnin et al. [2014].

information. We assume that $\mathscr{A}$ is computationally unbounded, and has knowledge of the prior ξ and the family $\mathcal{F}_\Theta$. To guarantee that $\mathscr{A}$ can gain little additional information about $\boldsymbol{D}$ from their communication, $\mathscr{B}$ uses Bayesian inference to learn from the data, and a differentially-private posterior to ensure disclosure to $\mathscr{A}$ is carefully controlled.

4.1.1 Probabilistic Graphical Models

Our main results focus on PGMs which model conditional independence assumptions with joint factorisation

$$p_{\boldsymbol{\theta}}(\boldsymbol{x}) = \prod_{i \in \mathrm{I}} p_{\boldsymbol{\theta}}\left(x_i \mid x_{\pi_i}\right), \quad x_{\pi_i} = \left\{x_j : j \in \pi_i\right\} ,$$

where $\boldsymbol{\pi_i}$ denote the parents of the $\boldsymbol{i}$-th variable in a Bayesian network—a directed acyclic graph with r.v.'s as nodes.

Example 1. *For concreteness, we illustrate some of our mechanisms on systems of Bernoulli r.v.'s $X_i \in \{0, 1\}$. In that case, we represent the conditional distribution of X_i given its parents as Bernoulli with parameters $\theta_{i,j} \in [0, 1]$:*

$$(X_i \mid X_{\pi_i} = j) \sim \mathcal{B}ernoulli(\theta_{i,j}) \ .$$

The choice of conjugate prior $\xi(\boldsymbol{\theta}) = \prod_{i,j} \xi_{i,j}(\theta_{i,j})$ has Beta marginals with parameters $\alpha_{i,j}, \beta_{i,j}$, so that:

$$(\theta_{i,j} \mid \alpha_{i,j} = \alpha, \beta_{i,j} = \beta) \sim \mathcal{B}eta(\alpha, \beta) \ .$$

Given observation $\boldsymbol{x}$, the updated posterior Beta parameters are $\alpha_{i,j} := \alpha_{i,j} + x_i$ and $\beta_{i,j} := \beta_{i,j} + (1 - x_i)$ if $x_{\pi_i} = j$.

4.2 Privacy by Posterior Perturbation

One approach to differential privacy is to use additive Laplace noise [Dwork et al., 2006]. Previous work has focused on the addition of noise directly to the outputs of a non-private mechanism. We are the first to apply Laplace noise to the posterior parameter updates. Note that the notion of differential privacy in this section is the standard ϵ-differential privacy.

4.2.1 Laplace Mechanism on Posterior Updates

Under the setting of Example 1, we can add Laplace noise to the posterior parameters. Algorithm 4.2.1 releases perturbed parameter updates for the Beta posteriors, calculated simply by *counting*. It then adds zero-mean Laplace-distributed noise to the updates $\Delta\boldsymbol{\omega} = (\cdots, \Delta\alpha_{i,j}, \Delta\beta_{i,j}, \cdots)$. This is the final dependence on D. Finally, the perturbed updates $\Delta\boldsymbol{\omega}'$ are truncated at zero to rule out invalid Beta parameters and are upper truncated at n. This yields an upper bound on the raw updates and facilitates an upper bound on the utility loss of KL-divergence.Note that this truncation only improves utility (relative to the utility pre-truncation), and does not affect privacy.

Algorithm 4.2.1: Laplace Mechanism on Posterior Updates

```
**Input** data $D$; graph $\mathrm{I}, \{\pi_i \mid i \in \mathrm{I}\}$; parameter $\epsilon > 0$;
calculate posterior updates: $\Delta\alpha_{i,j}, \Delta\beta_{i,j}$ for all
    $i \in \mathrm{I}, j \in \{0,1\}^{|\pi_i|}$;
perturb updates: $\Delta\alpha'_{i,j} \triangleq \Delta\alpha_{i,j} + \mathrm{Lap}\left(\frac{2|\mathrm{I}|}{\epsilon}\right)$,
    $\Delta\beta'_{i,j} \triangleq \Delta\beta_{i,j} + \mathrm{Lap}\left(\frac{2|\mathrm{I}|}{\epsilon}\right)$;
truncate: $Z^{(1)}_{i,j} \triangleq \mathbf{1}_{[0,n]}(\Delta\alpha'_{i,j}), Z^{(2)}_{i,j} \triangleq n - Z^{(2)}_{i,j}$, where $\mathbf{1}_{[0,n]}$ denotes the indicator
    function on $[0,n]$;
output $\mathbf{Z}_{i,j} = (Z^{(1)}_{i,j}, Z^{(2)}_{i,j})$.
```

Privacy. To establish differential privacy of our mechanism, we must calculate a Lipschitz condition for the vector $\boldsymbol{\Delta\omega}$, the *global sensitivity* Dwork et al. [2006].

Lemma 4.2.1. *For any* neighbouring *datasets* $D, \tilde{D}$*, the corresponding updates* $\Delta\boldsymbol{\omega}, \Delta\tilde{\boldsymbol{\omega}}$ *satisfy* $\|\Delta\boldsymbol{\omega} - \Delta\tilde{\boldsymbol{\omega}}\|_1 \leq 2|\mathrm{I}|$.

Proof. By changing the observations of one datum, at most two counts associated with each X_i can change by 1.

Corollary 4.2.2. *Algorithm 4.2.1 preserves* ϵ*-differential privacy.*

Proof. Based on Lemma 4.2.1, the intermediate $\Delta\boldsymbol{\omega}'$ preserve ϵ-differential privacy by Dwork et al. [2006]. Since truncation depends only on $\Delta\boldsymbol{\omega}'$, the $\boldsymbol{Z}$ preserves the same privacy.

Utility on Updates. Before bounding the effect on the posterior of the Laplace mechanism, we demonstrate a utility bound on the posterior update counts.

Proposition 4.2.3. *With probability at least* $1 - \delta$*, for* $\delta \in (0, 1)$*, the update counts computed by Algorithm 4.2.1 are close to the non-private counts*

$$\|\Delta\boldsymbol{\omega} - \Delta\boldsymbol{\omega}'\|_\infty \le \frac{2|\mathrm{I}|}{\epsilon} \ln\left(\frac{2m}{\delta}\right) ,$$

where $m = \sum_{i \in I} 2^{|\pi_i|}$.

Proof. Let us denote the event of a Laplace sample exceeding $z > 0$ in absolute value as A_k, $k \in 1, \cdots, 2m$. Consider the probability of an event that none of the $2m$ *i.i.d.*Laplace noise we add to each count exceed $z > 0$ in absolute value:

$$\begin{aligned} 1 - \mathbb{P}[\cup_{k=1}^{2m}\{A_k\}] &\ge 1 - \sum_{k=1}^{2m} \mathbb{P}[A_i] \\ &= 1 - 2m \exp(-z\epsilon/2|\mathrm{I}|)). \end{aligned}$$

To make sure this probability is no smaller than $1 - \delta$, we need z to be at most to $\frac{2|\mathrm{I}|}{\epsilon} \ln(\frac{2m}{\delta})$. This bound states that w.h.p., none of the updates can be perturbed beyond $O(|\mathrm{I}|^2/\epsilon)$. This implies the same bound on the deviation between $\Delta\boldsymbol{\omega}$ and the revealed truncated $\boldsymbol{Z}$.

Utility on Posterior. We derive our main utility bounds for Algorithm 4.2.1 in terms of posteriors. We abuse notation, and use ξ to refer to the prior density; its meaning will be apparent from context. Given priors $\xi_{i,j}(\theta_{i,j}) = \mathrm{Beta}\left(\alpha_{i,j}, \beta_{i,j}\right)$, the posteriors on n observations are

$$\xi_{i,j}(\theta_{i,j}|D) = \mathcal{B}eta(\alpha_{i,j} + \Delta\alpha_{i,j}, \beta_{i,j} + \Delta\beta_{i,j}) \ .$$

The privacy-preserving posterior parametrised by the output of Algorithm 4.2.1 is

$$\xi'_{i,j}(\theta_{i,j}|D) = \mathcal{B}eta\left(\alpha_{i,j} + Z^{(1)}_{i,j}, \beta_{i,j} + Z^{(2)}_{i,j}\right) \ .$$

It is natural to measure utility by the KL-divergence between the joint product posteriors $\xi(\boldsymbol{\theta}|D)$ and $\xi'(\boldsymbol{\theta}|D)$, which is the sum of the component-wise divergences, with each having known closed form. In our analysis, the divergence is a random quantity, expressible as the sum $\sum_{i,j}^{m} f_{i,j}(\mathbf{Z}_{i,j})$, where the randomness is due to the added noise. The following result based on concentration inequality (McDiarmid) is reported in [Zhang et al., 2016] to show that this sum of random variable is not too large with high probability.

Theorem 4.2.4. *Let* $m = \sum_{i\in\mathrm{I}} 2^{|\pi_i|}$. *Assume that* $Z_{i,j}$ *are independent and* f *is a mapping from* $\mathcal{Z}^m$ *to* $\mathbb{R}$*:* $f(\cdots, z_{i,j}, \cdots) \triangleq \sum_{i,j} f_{i,j}(z_{i,j})$*. Given* $\delta > 0$*, we have*

$$\mathbb{P}\left[f(\mathbf{Z}) \geq \mathrm{E}(f(\mathbf{Z})) + \left(-\frac{1}{2}\sum_{i,j} c_{i,j} \ln\delta\right)^{\frac{1}{2}}\right] \leq \delta$$

where $c_{i,j} \leq (2n+1)\ln[(\alpha_{i,j}+n+1) + (\beta_{i,j}+n+1))$ *and* $\mathrm{E}(f_{i,j}(\mathbf{Z}_{i,j})] \leq n\ln((\alpha_{i,j} + \varDelta\alpha_{i,j})(\beta_{i,j} + \varDelta\beta_{i,j})) = \mathcal{U}$*.*
Moreover, when $n \geq b = \frac{2|\mathrm{I}|}{\epsilon}$*, the bound for expectation can be refined as the following*

$$\ln[(\alpha_{i,j}+n+1)(\beta_{i,j}+n+1)]\left(\frac{n}{2}\exp\left(-\frac{n\epsilon}{2|\mathrm{I}|}\right)\right) \ .$$

The loss of utility measured by KL-divergence is no more than

$$O\left(mn\ln n\right)\left[1-\exp\left(-\frac{n\epsilon}{2|\mathrm{I}|}\right)\right] + \sqrt{-O\left(mn\ln n\right)\ln\delta}$$

with probability at least $1-\delta$*.*

However, we will prove a bound of the total utility loss in Theorem 4.2.7 that surpasses the above result. As in [Zhang et al., 2016], we need to assume that $\alpha_{i,j}$ and $\beta_{i,j}$ are larger than the only turning point of the $\varGamma$ function which is between 1 and 2; $\alpha_{i,j}, \beta_{i,j} \geq 2$ is sufficient.[3] For simplicity, we consider a related mechanism

[3] To cover more priors, we could assume that $\alpha_{i,j}$ is bounded away from zero, and that $\varGamma$ at this

that truncates the perturbed $\alpha_{i,j}$ update, then forms the $\beta_{i,j}$ update via $n - \mathbf{Z}_{i,j}^{(1)}$. The resulted bound is the same for the present mechanism in $\mathcal{O}$ notation and the analysis is very similar. Before we bound this random variable, let us prove the following lemmas.

Lemma 4.2.5. *For constants a and $t \geq 0$, $(a+t)\ln(a+t) - a\ln a \leq t\ln(a+t) + t$.*

Proof. This follows from applying the Mean Value Theorem to the function $x\ln(x)$ on the interval $[a, a+t]$.

Lemma 4.2.6. *For positive integers $x, y, z \geq 2$, the log-ratio of Gamma functions: $\ln\left(\frac{\Gamma(x+z)}{\Gamma(y)}\right) \leq \min\{y\ln(x+y), (x+y+1)\ln(x+y+1) - x\ln(x) - y + 1\}$.*

Proof.

$$\begin{aligned}
\ln\left(\frac{\Gamma(x+y)}{\Gamma(x)}\right) &= \ln\left(\frac{\Gamma(x)\prod_{r=1}^{y}(x+r)}{\Gamma(x)}\right)\\
&= \sum_{r=1}^{y}\ln(x+r)\\
&\leq \int_{0}^{y+1}\ln(x+t)dt\\
&=(x+y+1)\ln(x+y+1) - x\ln(x) - y + 1
\end{aligned}$$

Alternatively, we can simply argue that

$$\ln\left(\frac{\Gamma(x+y)}{\Gamma(x)}\right) = \sum_{r=1}^{y}\ln(x+r) \leq y\ln(x+y),$$

but this bound is worse when y is much larger than x. □

Theorem 4.2.7. *Let $m = \sum_{i\in \mathrm{I}} 2^{|\pi_i|}$. Assume that $\mathbf{Z}_{i,j}$ are independent and f is a mapping from $\mathcal{Z}^m$ to $\mathbb{R}$: $f(\cdots, z_{i,j}, \cdots) \triangleq \sum_{i,j} f_{i,j}(z_{i,j})$. Given $\delta > 0$, the loss of utility measured by KL-divergence $\sum_{i,j}^{m} f_{i,j}(\mathbf{Z}_{i,j})$ is no more than $\mathcal{O}\left(\frac{|I|}{\epsilon}\ln\frac{m}{\delta}\ln\left(\frac{|I|}{\epsilon}\ln\frac{m}{\delta}\right)\right)$ with probability $1-\delta$.*

parameter is maximum below 2 and proceed from there for the second case.

Proof. First we show the component-wise divergence $f_{i,j}(\mathbf{Z}_{i,j})$ is bounded. It is known that the absolute value of Laplace noise with zero mean and scale ϵ is bounded by $1/\epsilon \ln(1/\delta)$ with probability $1-\delta$. So, with probability $1-\delta$, the noise injected to $f_{i,j}(\mathbf{Z}_{i,j})$ is bounded by $\frac{2|\mathrm{I}|}{\epsilon}\ln(1/\delta)$.

Since $B(x,y) = \Gamma(x)\Gamma(y)/\Gamma(x+y)$ and $\Delta\alpha_{i,j} + \Delta\beta_{i,j} = Z^{(1)}_{i,j} + Z^{(2)}_{i,j}$, we have

$$\frac{B(\alpha'_{i,j}, \beta'_{i,j})}{B(\alpha_{i,j} + \Delta\alpha_{i,j}, \beta_{i,j} + \Delta\beta_{i,j})} = \frac{\Gamma(\alpha'_{i,j})}{\Gamma(\alpha_{i,j} + \Delta\alpha_{i,j})} + \frac{\Gamma(\beta'_{i,j})}{\Gamma(\beta_{i,j} + \Delta\beta_{i,j})}.$$

Thus we have the following bound by applying Lemma 4.2.6 and using the inequalities of digamma function $\ln(x-1) \leq \phi(x) \leq \ln(x)$,

$$\begin{aligned}
|f_{i,j}(\mathbf{Z}_{i,j})| \leq & \ln\frac{\Gamma(\alpha'_{i,j})}{\Gamma(\alpha_{i,j})} + \ln\frac{\Gamma(\beta'_{i,j})}{\Gamma(\beta_{i,j})} + (\alpha_{i,j} - \alpha'_{i,j})\phi(\alpha_{i,j}) \\
& + (\beta_{i,j} - \beta'_{i,j})\phi(\beta_{i,j}) + (\alpha'_{i,j} + \beta'_{i,j} - \alpha_{i,j} - \beta_{i,j})\phi(\alpha_{i,j} + \beta_{i,j}) \\
& \leq \ln\frac{\Gamma\left(\left\lceil \alpha_{i,j} + \ 2|\mathrm{I}|/\epsilon \ln(1/\delta)\right\rceil\right)}{\Gamma\left(\left\lfloor \alpha_{i,j}\right\rfloor\right)} + \ln\frac{\Gamma\left(\left\lceil \beta_{i,j} + 2|\mathrm{I}|/\epsilon \ln(1/\delta)\right\rceil\right)}{\Gamma\left(\left\lfloor \beta_{i,j}\right\rfloor\right)} \\
& + \frac{4|\mathrm{I}|}{\epsilon}\ln(1/\delta)\ln(\alpha_{i,j} + \beta_{i,j}) \\
& \leq \lceil 2|\mathrm{I}|/\epsilon \ln(1/\delta)\rceil \ln\left(\left\lceil \alpha_{i,j} + 2|\mathrm{I}|/\epsilon \ln[(1/\delta)\right\rceil \left\lceil \beta_{i,j} + 2|\mathrm{I}|/\epsilon \ln(1/\delta)\right\rceil\right) \\
& + 4|\mathrm{I}|/\epsilon \ln(1/\delta)\ln(\alpha_{i,j} + \beta_{i,j}) \\
& \leq \lceil 6|\mathrm{I}|/\epsilon \ln(1/\delta)\rceil \ln\left(\left\lceil \alpha_{i,j} + 2|\mathrm{I}|/\epsilon \ln[(1/\delta)\right\rceil \left\lceil \beta_{i,j} + 2|\mathrm{I}|/\epsilon \ln(1/\delta)\right\rceil\right)
\end{aligned}$$

holding with high probability at least $1-\delta$, starting at the second inequality. This inequality follows from the concentration of Laplace, monotonicity of Γ and that Γ coincides with factorial on the integers. Let us denote $|\mathrm{I}|/\epsilon \ln(1/\delta))$ as q. Then the component-wise divergence $f(Z_{i,j})$ is bounded by $O(q\ln q)$ with probability $1-\delta$.

$$\begin{aligned}
& \mathbb{P}\left(\sum_{i=1}^{m} |f(\mathbf{Z}_{i,j})| \geq mO(q\ln q)\right) \\
\leq & \mathbb{P}\left(\bigcup_{i=1}^{m} \{|f(\mathbf{Z}_{i,j})| \geq O(q\ln q)\right)
\end{aligned}$$

$$\leq \sum_{i=1}^{m} \mathbb{P}\left((|f(\boldsymbol{Z}_{i,j}) \geq O(q \ln q)\right)$$
$$\leq m\delta.$$

Let $\delta' = m\delta$, we have $\delta = \delta'/m = \exp(-\frac{q\epsilon}{|I|})$. Then $q = \frac{|I|}{\epsilon} \ln \frac{m}{\delta'}$, we have

$$\mathbb{P}\left(\sum_{i=1}^{m} |f(\boldsymbol{Z}_{i,j})| \geq O\left(\frac{|I|}{\epsilon} \ln \frac{m}{\delta'} \ln\left(\frac{|I|}{\epsilon} \ln \frac{m}{\delta'}\right)\right)\right) \leq \delta'.$$

□

Note that m depends on the structure of the network: bounds are better for networks with an underlying graph having smaller average in-degree; more conditional independence improves privacy.

4.2.2 Laplace Mechanism in the Fourier Domain

Algorithm 4.2.1 follows *Kerckhoffs's Principle* [Kerckhoffs, 1883] of "no security through obscurity": differential privacy defends against a mechanism-aware attacker. However *additional stealth* may be required in certain circumstances. An oblivious observer will be tipped off to our privacy-preserving activities by our independent perturbations, which are likely inconsistent with one-another (*e.g.*, noisy counts for X_1, X_2 and X_2, X_3 will say different things about X_2). To achieve differential privacy and stealth, we turn to Barak et al. [2007]'s study of consistent marginal contingency table release. This section presents a particularly natural application to Bayesian posterior updates.

Denote by $h \in \mathbb{R}^{\{0,1\}^{|\mathbb{I}|}}$ the *contingency table* over r.v.'s Iinduced by D: i.e. for each combination of variables $j \in \{0,1\}^{|\mathbb{I}|}$, component or *cell* h_j is a non-negative count of the observations in D with characteristic j. Geometrically h is a real-valued function over the $|\mathbb{I}|$-dimensional Boolean hypercube. Then the parameter delta's of our first mechanism correspond to cells of $(|\pi_i| + 1)$-way marginal contingency tables $\mathrm{C}^{\overline{\pi}_i}(h)$ where vector $\overline{\pi}_i \triangleq \pi_i + e_i$ and the projection/marginalisation operator is defined as

$$\left(\mathrm{C}^{j}(h)\right)_{\gamma} \triangleq \sum_{\eta : \langle \eta, j \rangle = \gamma} h_{\eta} \ . \tag{4.1}$$

We wish to release these statistics as before, however we will not represent them under their Euclidean coordinates but instead in the Fourier basis $\{f^j : j \in \{0,1\}^{|\mathrm{I}|}\}$ where

$$f^j_\gamma \triangleq (-1)^{\langle \gamma, j\rangle} 2^{-|\mathrm{I}|/2} .$$

Due to this basis structure and linearity of the projection operator, any marginal contingency table must lie in the span of few projections of Fourier basis vectors [Barak et al., 2007]:

Theorem 4.2.8. *For any table $h \in \mathbb{R}^{\{0,1\}^{|\mathrm{I}|}}$ and set of variables $j \in \{0,1\}^{|\mathrm{I}|}$, the marginal table on j satisfies $\mathrm{C}^j(h) = \sum_{\gamma \preceq j} \langle f^\gamma, h\rangle \, \mathrm{C}^j(f^\gamma)$.*

This states that marginal j lies in the span of only those (projected) basis vectors f^γ with γ contained in j. The number of values needed to update X_i is then $2^{|\pi_i|+1}$, potentially far less than suggested by (4.1). To release updates for two r.v.'s $i, j \in \mathrm{I}$ there may well be significant overlap $\langle \overline{\pi}_i, \overline{\pi}_j\rangle$; we need to release once, coefficients $\langle f^\gamma, h\rangle$ for γ in the downward closure of variable neighbourhoods:

$$\mathcal{N}_{\mathrm{I}} \triangleq \bigcup_{i \in \mathrm{I}} \bigcup_{j \preceq \overline{\pi}_i} j .$$

Privacy. By [Barak et al., 2007, Theorem 6] we can apply Laplace additive noise to release these Fourier coefficients.

Corollary 4.2.9. *For any $\epsilon > 0$, releasing for each $\gamma \in \mathcal{N}_{\mathrm{I}}$ the Fourier coefficient $\langle f^\gamma, h\rangle + \mathrm{Lap}\left(2|\mathcal{N}_{\mathrm{I}}|\epsilon^{-1}2^{-|\mathrm{I}|/2}\right)$ (and Algorithm 4.2.2) preserves ϵ-differential privacy.*

Remark 4.2.10. *Since $|\mathcal{N}_{\mathrm{I}}| \leq |\mathrm{I}|2^{1+\max_{i\in\mathrm{I}} \mathrm{indeg}(i)}$, at worst we have noise scale $|\mathrm{I}|2^{2+\max_i \mathrm{indeg}(i)-|\mathrm{I}|/2}/\epsilon$. This compares favourably with Algorithm 4.2.1's noise scale provided no r.v. is child to more than half the graph. Moreover the denser the graph—the more overlap between nodes' parents and the less conditional independence assumed—the greater the reduction in scale. This is intuitively appealing.*

Consistency. What is gained by passing to the Fourier domain, is that the perturbed marginal tables of Corollary 4.2.9 are consistent: anything in the span of projected Fourier basis vectors correspond to some valid contingency table on I with (possibly negative) real-valued cells [Barak et al., 2007].

Algorithm 4.2.2: Laplace Mechanism in the Fourier Domain

```
Input data $D$; graph $\mathrm{I}, \{\pi_i \mid i \in \mathrm{I}\}$; prior parameters $\boldsymbol{\alpha}, \boldsymbol{\beta} \geq \mathbf{0}$; parameters $t, \epsilon > 0$
define contingency table $h \in \mathbb{R}^{\{0,1\}^{|\mathrm{I}|}}$ on $D$
define downward closure $\mathcal{N}_\mathrm{I} = \bigcup_{i \in \mathrm{I}} \bigcup_{j \preceq \bar{\pi}_i} j$
for $\gamma \in \mathcal{N}_\mathrm{I}$ do
    Fourier coefficient $z_\gamma = \langle f^\gamma, h \rangle + \mathrm{Lap}\left(\frac{2|\mathcal{N}_\mathrm{I}|}{\epsilon 2^{|\mathrm{I}|/2}}\right)$
end for
increment first coefficient $z_\mathbf{0} \leftarrow z_\mathbf{0} + \frac{4t|\mathcal{N}_\mathrm{I}|^2}{\epsilon 2^{|\mathrm{I}|/2}}$
for $i \in \mathrm{I}$ do
    project marginal for $X_i$ as $h^i = \sum_{\gamma \preceq \bar{\pi}_i} z_\gamma C^{\bar{\pi}_i}(f^\gamma)$
    for $j \preceq \pi_i$ do
        output posterior param $\left(\alpha_{ij} + h^i_{e_i+j}, \beta_{ij} + h^i_j\right)$
    end for
end for
```

Non-negativity. So far we have described the first stage of Algorithm 4.2.2. The remainder yields *stealth* by guaranteeing releases that are non-negative w.h.p. We adapt an idea of Barak et al. [2007] to increase the coefficient of Fourier basis vector $f^\mathbf{0}$, affecting a small increment to each cell of the contingency table. While there is an exact minimal amount that would guarantee non-negativity, it is data dependent. Thus our efficient $O(|\mathcal{N}_\mathrm{I}|)$-time approach is randomised.

Corollary 4.2.11. *For $t > 0$, adding $4t|\mathcal{N}_\mathrm{I}|^2\epsilon^{-1}2^{-k/2}$ to $f^\mathbf{0}$'s coefficient induces a non-negative table w.p. $\geq 1 - \exp(-t)$.*

Parameter t trades off between the probability of non-negativity and the resulting (minor) loss to utility. In the rare event of negativity, re-running Algorithm 4.2.2 affords another chance of stealth at the cost of privacy budget ϵ. We could alternatively truncate to achieve validity, sacrificing stealth but not privacy.

Utility. Analogous to Proposition 4.2.3, each perturbed marginal is close to its unperturbed version w.h.p.

Theorem 4.2.12. *For each $i \in \mathrm{I}$ and $\delta \in (0, 1)$, the perturbed tables in Algorithm 4.2.2 satisfy with probability at least $1 - \delta$:*

$$\left\|\mathrm{C}^{\bar{\pi}_i}(h) - h^i\right\|_1 \quad \leq \quad \frac{4|\mathcal{N}_\mathrm{I}|}{\epsilon}\left(2^{|\pi_i|}\log\frac{|\mathcal{N}_\mathrm{I}|}{\delta} + t|\mathcal{N}_\mathrm{I}|\right) \ .$$

Proof. We follow the proof of [Barak et al., 2007, Theorem 7]. If $X \sim \mathsf{Lap}(b)$ then by the CDF of the Laplace $\mathbb{P}(|X| > R) = \exp(-R/b)$ where $R > 0$. By the union bound for $\{X_j\}_{j \in \mathcal{N}_\mathrm{I}} \overset{i.i.d.}{\sim} \mathsf{Lap}(b)$, we have w.h.p. none is large $\mathbb{P}\left(\forall j \in \mathcal{N}_\mathrm{I}, |X_j| \leq b\log(|\mathcal{N}_\mathrm{I}|/\delta)\right) \geq 1 - \delta$ for $\delta \in (0, 1)$. Since $\|f^j\|_1 = 2^{k/2}$ for each $j \subseteq \mathrm{I}$ it follows with probability at least $1 - \delta$, that $\forall j \in \mathcal{N}_\mathrm{I} \backslash \{\emptyset\}, \left\|z_j f^j - \langle f^j, h\rangle f^j\right\|_1 \leq \frac{2|\mathcal{N}_\mathrm{I}|}{\epsilon}\log\frac{|\mathcal{N}_\mathrm{I}|}{\delta}$. For $f^{\mathbf{0}}$ the additional increment comes at an additional cost of $4t|\mathcal{N}_\mathrm{I}|^2/\epsilon$. Putting everything together, we note that $2^{|\pi_i|+1}$ Fourier coefficients represent h_i including $f^{\mathbf{0}}$. □

Note that the scaling of this bound is reasonable since the table h^i involves $2^{|\pi_i|+1}$ cells.

4.3 Privacy by Posterior Sampling

For general Bayesian networks, $\mathscr{B}$ can release samples from the posterior instead of perturbed samples of the posterior's parametrisation. We now develop a calculus of building up (stochastic) Lipschitz properties of systems of r.v.'s that are locally (stochastic) Lipschitz. Given smoothness of the entire network, differential privacy and utility of posterior sampling follow by the results of the previous chapter.

4.3.1 (Stochastic) Lipschitz Smoothness of Networks

The distribution family $\{p_\theta : \theta \in \Theta\}$ on outcome space $\mathcal{S}$, equipped with pseudo metric[4] ρ, is *Lipschitz continuous* if

Assumption 3 (Lipschitz Continuity). *Let $d(\cdot, \cdot)$ be a metric on $\mathbb{R}$. There exists $L > 0$ such that, for any $\theta \in \Theta$:*

$$d(p_\theta(x), p_\theta(y)) \leq L\rho(x, y), \forall x, y \in \mathcal{S}.$$

[4] Meaning that $\rho(x, y) = 0$ does not necessarily imply $x = y$.

We fix the distance function d to be the absolute log-ratio (*cf.* differential privacy). Consider a general Bayesian network. The following lemma shows that the individual Lipschitz continuity of the conditional likelihood at every $i \in \mathrm{I}$ implies the global Lipschitz continuity of the network.

Lemma 4.3.1. *If there exists* $\boldsymbol{L} = (L_1, \cdots, L_{|\mathrm{I}|}) \geq \boldsymbol{0}$ *such that* $\forall i \in \mathrm{I}$, $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathcal{X} = \prod_{i=1}^{|\mathrm{I}|} \mathcal{X}_i$ *we have* $d(p_{\boldsymbol{\theta}}(x_i|x_{\pi_i}), p_{\boldsymbol{\theta}}(y_i|y_{\pi_i})) \leq L_i \rho_i(x_i, y_i)$*, then* $d(p_{\boldsymbol{\theta}}(\boldsymbol{x}), p_{\boldsymbol{\theta}}(\boldsymbol{y})) \leq \|\boldsymbol{L}\|_\infty \boldsymbol{\rho}(\boldsymbol{x}, \boldsymbol{y})$ *where* $\boldsymbol{\rho}(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i=1}^{|I|} \rho_i(x_i, y_i)$.

Proof.

$$\begin{aligned} d(p_{\boldsymbol{\theta}}(x), p_{\boldsymbol{\theta}}(y)) &= \left| \log \prod_{i=1}^{|\mathrm{I}|} \frac{p_{\boldsymbol{\theta}}(x_i|x_{\pi_i})}{p_{\boldsymbol{\theta}}(y_i|y_{\pi_i})} \right| \\ &\leq \sum_{i=1}^{|\mathrm{I}|} \left| \log \frac{p_{\boldsymbol{\theta}}(x_i|x_{\pi_i})}{p_{\boldsymbol{\theta}}(y_i|y_{\pi_i})} \right| \\ &= \sum_{i=1}^{|\mathrm{I}|} d(p_{\boldsymbol{\theta}}(x_i|x_{\pi_i}), p_{\boldsymbol{\theta}}(y_i|y_{\pi_i})) \\ &\leq \sum_{i=1}^{|\mathrm{I}|} L_i \rho_i(x_i, y_i) \\ &\leq \|\boldsymbol{L}\|_\infty \|\boldsymbol{\rho}(\boldsymbol{x}, \boldsymbol{y})\|_1 . \end{aligned}$$

Note that while Lipschitz continuity holds uniformly for some families *e.g.*, the exponential distribution, this is not so for many useful distributions such as the Bernoulli. In such cases a relaxed assumption requires that the prior be concentrated on smooth regions.

Assumption 4 (Stochastic Lipschitz Continuity). *Let the set of* L*-Lipschitz* θ *be*

$$\Theta_L \triangleq \left\{ \theta \in \Theta : \sup_{x,y \in \mathcal{S}} \{ d(p_\theta(x), p_\theta(y)) - L\rho(x,y) \} \leq 0 \right\}$$

Then there exists constants $c, L_0 > 0$ *such that,* $\forall L \geq L_0$*:* $\xi(\Theta_L) \geq 1 - e^{-cL}$.

Lemma 4.3.2. *For the conditional likelihood at each node* $i \in \mathrm{I}$*, define the set* $\Theta_{i,L}$ *of parameters for which Lipschitz continuity holds with Lipschitz constant* L*. If*

$\exists \boldsymbol{c} = (c_1, \cdots, c_{|\mathrm{I}|})$ *such that* $\forall i, L \geq L_0$, $\xi(\Theta_{i,L}) \geq 1 - e^{-c_i L}$, *then* $\xi(\Theta_L) \geq 1 - e^{-c' L}$, *where* $c' = \min_{i \in \mathrm{I}} c_i - \ln |\mathrm{I}|/L_0$ *when* $|\mathrm{I}| \leq e^{L_0 \min_{i \in \mathrm{I}} c_i}$.

Proof. Define

$$\Theta_{i,L} = \left\{ \boldsymbol{\theta} \in \Theta : \sup_{x,y \in \mathcal{X}} \{ d(p_{\boldsymbol{\theta}}(x_i|x_{\pi_i}), p_{\boldsymbol{\theta}}(y_i|y_{\pi_i})) - L\rho_i(x_i, y_i) \} \leq 0 \right\}.$$

By taking $\rho(x, y) = \sum_i \rho_i(x_i, y_i)$, we have

$$\begin{aligned}
\bigcap_{i=1}^{|\mathrm{I}|} \tilde{\Theta}_{i,L} =& \left\{ \boldsymbol{\theta} \in \Theta : \sup_{x_i, y_i \in X_i} \{ d(p_{\boldsymbol{\theta}}(x_i|x_{\pi_i}), p_{\boldsymbol{\theta}}(y_i|y_{\pi_i})) \leq L\rho_i(x_i, y_i) \}, \forall i \in \mathrm{I} \right\} \\
\subseteq& \left\{ \boldsymbol{\theta} \in \Theta : \sup_{x_i, y_i \in X_i} \left\{ \sum_{i=1}^{|\mathrm{I}|} d(p_{\boldsymbol{\theta}}(x_i|x_{\pi_i}), p_{\boldsymbol{\theta}}(y_i|y_{\pi_i})) \leq L \sum_{i=1}^{|\mathrm{I}|} \rho_i(x_i, y_i) \right\} \right\} \\
\subseteq& \{ \boldsymbol{\theta} \in \Theta : \sup \{ d(p_{\boldsymbol{\theta}}(x), p_{\boldsymbol{\theta}}(y)) - L\rho(x, y) \} \leq 0 \} \\
=& \Theta_L
\end{aligned}$$

Therefore, we have that the set of $\boldsymbol{\theta} \in \Theta$ satisfying the Stochastic Lipschitz continuity for conditional likelihood of every $i \in \mathrm{I}$ in the Bayesian network is a subset of the set of $\boldsymbol{\theta}$ satisfying the global Stochastic Lipschitz continuity for same L.

Note that $(\bigcap_{i=1}^{|\mathrm{I}|} \Theta_{i,L})^c = \bigcup_{i=1}^{|\mathrm{I}|} (\Theta_{i,L})^c$ and $\xi((\Theta_{i,L})^c) = 1 - \xi(\Theta_{i,L}) \leq e^{-c_i L}$. Then we have

$$\xi[(\cap_{i=1}^{|\mathrm{I}|} \Theta_{i,L})^c] \leq \sum_{i=1}^{|\mathrm{I}|} \xi(\Theta_{i,L})^c) \leq \sum_{i=1}^{|\mathrm{I}|} e^{-c_i L}.$$

Therefore, we have

$$\xi(\Theta_L) \geq \xi(\cap_{i=1}^{|\mathrm{I}|} \Theta_{i,L}) \geq 1 - \sum_{i=1}^{|\mathrm{I}|} e^{-c_i L} \geq 1 - N e^{-\min_i c_i L}.$$

Take $c' = \min\{c_i\}_{i=1} - \ln |\mathrm{I}|/L_0$, we have $\xi(\Theta_L) \geq 1 - e^{-c' L}$. □

Therefore, 4.3.3 asserts differential privacy of the Bayesian network's posterior.

Theorem 4.3.3. *Differential privacy is satisfied using the log-ratio distance, for all* $B \in \mathfrak{S}_\Theta$ *and* $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{\mathcal{X}}$*:*

1. *Under the conditions in Lemma 4.3.1*

$$\xi(B \mid \boldsymbol{x}) \leq \exp\{2L\rho(\boldsymbol{x},\boldsymbol{y})\}\xi(B \mid \boldsymbol{y}),$$

i.e. the posterior ξ *is* $(2\|\boldsymbol{L}\|_\infty, 0)$*-differentially private under pseudo-metric* $\rho(\boldsymbol{x},\boldsymbol{y})$*.*

2. *Under the conditions in Lemma 4.3.2, if* $\rho(\boldsymbol{x},\boldsymbol{y}) \leq (1-\delta)c$ *uniformly for all* $\boldsymbol{x},\boldsymbol{y}$ *for some* $\delta \in (0,1)$*,*

$$|\xi(B \mid \boldsymbol{x}) - \xi(B \mid \boldsymbol{y})| \leq \sqrt{\frac{M}{2} \cdot \max\{\rho(\boldsymbol{x},\boldsymbol{y}), 1\}},$$

where $M = \left(\frac{\kappa}{c} + L_0(\frac{1}{1-e^{-\omega}} + 1) + \ln C \;\; + \ln\left(e^{-L_0\delta c}(e^{-\omega(1-\delta)} - e^{-\omega})^{-1} + e^{L_0(1-\delta)c}\right)\right) C$*; constants* $\kappa = 4.91081$ *and* $\omega = 1.25643$*;* $C = \prod_i^{|I|} C_i$*; and*

$$C_i \;\; = \;\; \sup_{x \in \mathcal{X}} \frac{p_{\theta^\star_{i,\mathrm{MLE}}}(x_i \mid x_{\pi_i})}{\int_{\Theta_i} p_{\theta_i}(x_i \mid x_{\pi_i}) d\xi(\theta_i)} \quad ,$$

the ratio between the maximum and marginal likelihoods of each likelihood function. Note that $M = O\left(\left(\frac{1}{c} + \ln C + L_0\right) C\right)$ *i.e. the posterior* ξ *is* $\left(0, \sqrt{\frac{M}{2}}\right)$*-differentially private under pseudo-metric* $\sqrt{\rho}$ *for* $\rho(\boldsymbol{x},\boldsymbol{y}) \geq 1$*.*

4.3.2 MAP by the Exponential Mechanism

As an application of the posterior sampler, we now turn to releasing MAP point estimates via the exponential mechanism due to McSherry and Talwar [2007], which samples responses from a likelihood exponential in some score function. By selecting a utility function that is maximised by a target non-private mechanism, the exponential mechanism can be used to privately approximate that target with high utility. It is natural then to select as our utility u the posterior likelihood $\xi(\cdot|D)$. This u is maximised by the MAP estimate.

Formally, Algorithm 4.3.1, under the assumptions of Theorem 4.3.3, outputs response θ with probability proportional to $\exp(\epsilon u(D,\theta)/2\Delta)$ times a base measure $\mu(\theta)$. Here Δ is a Lipschitz coefficient for u with sup-norm on responses and pseudo-

Algorithm 4.3.1: Mechanism for MAP Point Estimates

```
1: Input data $D$; prior $\xi(\cdot)$; appropriate smoothness parameters $c, L, M > 0$;
   parameters distance $r > 0$, privacy $\epsilon > 0$
2: calculate posterior $\xi(\theta|D)$
3: set $\Delta = \begin{cases} \sqrt{Lr}\ , & \text{if Lipschitz continuous} \\ \sqrt{0.5M}\ , & \text{if stochastic Lipschitz} \end{cases}$
4: output $\hat{\theta}$ sampled $\propto \exp\left(\frac{\epsilon\xi(\theta|D)}{2\Delta}\right)\xi(\theta)$
```

metric ρ on datasets as in the previous section. Providing the base measure is non-trivial in general, but for discrete finite outcome spaces can be uniform [McSherry and Talwar, 2007]. For our mechanism to be broadly applicable, we can safely take $\mu(\theta)$ as $\xi(\theta)$.[5]

Corollary 4.3.4. *Algorithm 4.3.1 preserves ϵ-differential privacy wrt pseudo-metric ρ up to distance $r > 0$.*

Proof. The sensitivity of the posterior score function corresponds to the computed Δ under either Lipschitz assumptions. The result then follows from [McSherry and Talwar, 2007, Theorem 6].

Utility for Algorithm 4.3.1 follows from McSherry and Talwar [2007], and states that the posterior likelihood of responses is likely to be close to that of the MAP.

Lemma 4.3.5. *Let $\theta^\star = \max_\theta \xi(\theta|D)$ with maximizer the MAP estimate, and let $S_t = \{\theta \in \Theta : \xi(\theta|D) > \theta^\star - t\}$ for $t > 0$. Then $\mathbb{P}(S_{2t}^c) \leq \exp(-\epsilon t)/\xi(S_t)$.*

4.4 Experiments

Having proposed a number of mechanisms for approximating exact Bayesian inference in the general framework of probabilistic graphical models, we now demonstrate our approaches on two simple, well-known PGMs: the (generative) naïve Bayes classifier, and (discriminative) linear regression. This section, illustrates how our approaches are applied, and supports our extensive theoretical results with

[5] In particular the base measure guarantees we have a proper density function: if $u(D, \theta)$ is bounded by M, then we have normalising constant $\int_\theta \exp(\epsilon u(D, \theta))\mu(\theta)d\theta \leq \exp(M\epsilon) < \infty$.

experimental observation. We focus on the trade-off between privacy and utility (accuracy and MSE respectively), which involves the (private) posterior via a predictive posterior distribution in both case studies.

4.4.1 Bayesian Discrete Naïve Bayes

We review the derivation of the naïve Bayes predictive posterior for two cases applied in our experiments.

Recall that when the random variables in the network are all Bernoulli's with Beta conjugate priors:

$$\mathbb{P}(Y = y|\boldsymbol{X} = \boldsymbol{x}) \propto \int_{\Theta} p_\theta(y) \prod_{i=1}^{d} p_\theta(x_i|y)\, \xi(\theta)\, d\theta.$$

The integral decouples into the product of (where α,β refer to the y posterior)

$$\begin{aligned}
&\int_{\Theta} p_\theta(y)\, \xi(\theta)\, d\theta \\
= &\int_0^1 \frac{\theta^{\alpha+y-1}(1-\theta)^{\beta+(1-y)-1}}{B(\alpha,\beta)} d\theta \\
= &\frac{B(\alpha+y,\beta+1-y)}{B(\alpha,\beta)} \times \int_0^1 \frac{\theta^{\alpha+y-1}(1-\theta)^{\beta+(1-y)-1}}{B(\alpha+y,\beta+1-y)} d\theta \\
= &\frac{B(\alpha+y,\beta+1-y)}{B(\alpha,\beta)} \\
= &\frac{\Gamma(\alpha+y)\Gamma(\beta+1-y)}{\Gamma(\alpha+\beta+1)} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \\
= &\frac{\alpha^y \beta^{1-y}}{\alpha+\beta} \ .
\end{aligned}$$

and terms (where α,β refer to the $x_i \mid y$ posterior)

$$\int_0^1 \frac{\theta^{\alpha+x_i-1}(1-\theta)^{\beta+(1-x_i)-1}}{B(\alpha,\beta)} d\theta = \frac{\alpha^{x_i}\beta^{1-x_i}}{\alpha+\beta} \ ,$$

computed in the same way.

4.4.2 Sampling

Given an empirical CDF sampled from our posterior sampling mechanism, we can approximate by posterior sampling:

- Repeat many times for both $y = 0, y = 1$:
  - Sample $\hat{\theta}_y, \hat{\theta}_{x_1,y}, \ldots, \hat{\theta}_{x_d,y}$
  - Plug-in the sampled parameters and fixed r.v.'s into the product of densities to obtain an unnormalised probability estimate
- Average the obtained estimates, for each $y = 0, y = 1$
- Normalise

We modify the above slightly so that we sample from a truncated posterior. This allows us to assume a minimal probability ω assigned to any sub-event in the naïve Bayes network, so that the joint distribution satisfies Assumption 1. Trivially in particular this yields a differential privacy level given by $\epsilon = 2\log(1/\omega)$. Given a desired privacy budget ϵ we can therefore select $\omega = \exp(-\epsilon/2)$. We then simply rejection sample when sampling above, to obtain samples from the truncated posterior. This is the posterior sampler algorithm used in the naïve Bayes experiments.

An illustrative example for our mechanisms is a Bayesian naïve Bayes model on Bernoulli class Y and attribute variables X_i, with full conjugate Beta priors. This PGM directly specialises the running Example 1. We synthesised data generated from a naïve Bayes model, with 16 features and 1000 examples. Of these we trained our mechanisms on only 50 examples, with uniform Beta priors. We formed predictive posteriors for $Y|\boldsymbol{X}$ from which we thresholded at 0.5 to make classification predictions on the remaining, unseen test data so as to evaluate classification accuracy. The results are reported in Figure 4.1, where average performance is taken over 100 repeats to account for randomness in train/test split, and randomised mechanisms.

The small size of this data represents a challenge in our setting, since privacy is more difficult to preserve under smaller samples Dwork et al. [2006]. As expected, privacy incurs a sacrifice to accuracy for all private mechanisms.

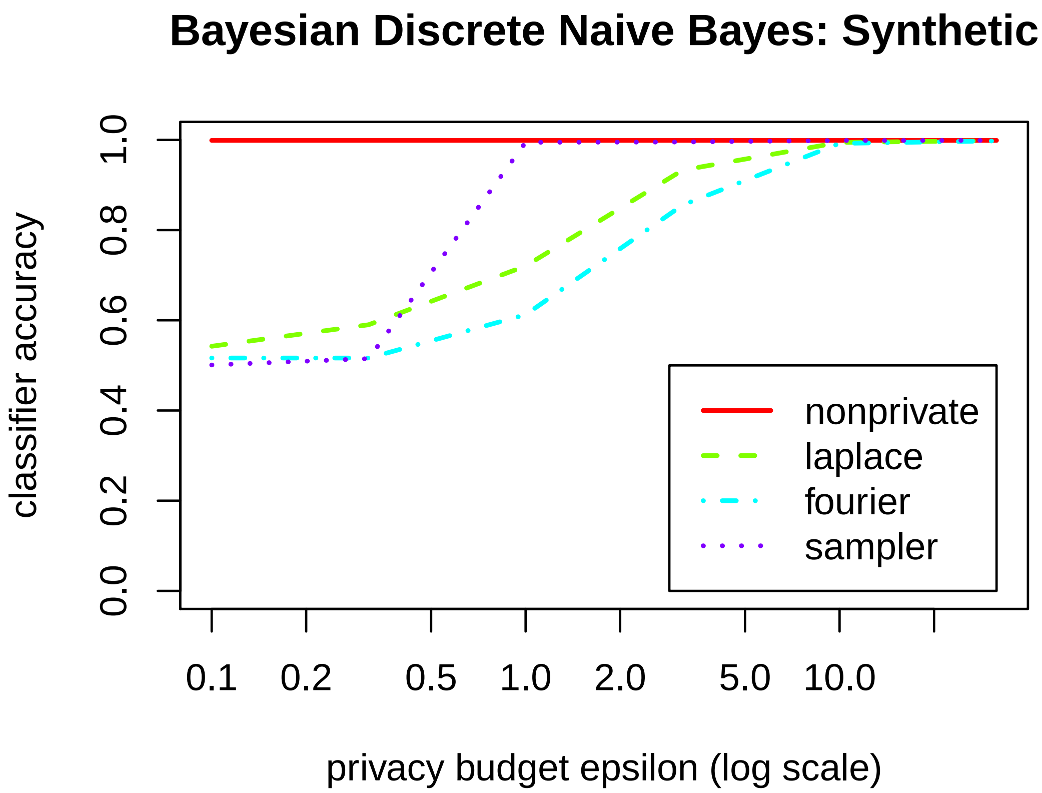

Fig. 4.1: Effect on Bayesian naïve Bayes predictive-posterior accuracy of varying the privacy level.

For both Laplace mechanisms that perturb posterior updates, note that the d Boolean attributes and class label (being sole parent to each) yields nodes $|\mathrm{I}| = d+1$ and downward closure size $|\mathcal{N}_{\mathrm{I}}| = 2d + 2$. Following our generic mechanisms, the noise added to sufficient statistics is independent on training set size, and is similar in scale. The parameter t, that trades off between the probability of non-negativity and the resulting loss to utility, was set for the Fourier approach, so that stealth was achieved 90% of the time—those times that contributed to the plot. Due to the small increments to cell counts for Fourier, necessary to achieve its *additional stealth property*, we expect a *small decrease to utility which is borne out in Figure 4.1.*

For the posterior sampler mechanism, while we can apply Assumption 2 to a

Bernoulli-Beta pair to obtain a generalized form of (ϵ, δ)-differential privacy, we wish to compare with our ϵ-differentially-private mechanisms and so choose a route which satisfies Assumption 1. We trim the posterior before sampling, so that probabilities are lower-bounded. Figure 4.1 demonstrates that for small ϵ, the minimal probability at which to trim is relatively large resulting in a poor approximate posterior. But past a certain threshold, *the posterior sampler eventually outperforms the other private mechanisms.*

4.4.3 Bayesian Linear Regression

We next explore a system of continuous r.v.'s in Bayesian linear regression, for which our posterior sampler is most appropriate. We model label Y as *i.i.d.*Gaussian with known-variance and mean a linear function of features, and the linear weights endowed with multivariate Gaussian prior with zero mean and spherical covariance. To satisfy Assumption 1 we conservatively truncate the Gaussian prior and sample from the resulting truncated posterior; form a predictive posterior; then compute mean squared error. To evaluate our approach we used the U.S. census records dataset from the *Integrated Public Use Microdata Series* Minnesota Population Center [2009] with 370k records and 14 demographic features.

To predict *Annual Income*, we train on 10% data with the remainder for testing. Figure 4.2 displays MSE under varying prior precision b (inverse of covariance) and weights with bounded norm $10/\sqrt{b}$ (chosen conservatively). As expected, more concentrated prior (larger b) leads to worse MSE for both mechanisms, as stronger priors reduce data influence. Compared with linear regression, private regression suffers only slightly worse MSE. At the same time the posterior sampler enjoys increasing privacy (that is proportional to the bounded norm). Let us denote a set of observations $D = \{(x_1, y_1), \dots, (x_n, y_n)\}$ where

$$x_i = (x_i^{(1)}, \dots, x_i^{(d)}) \in \mathbb{R}^d, y_i \in \mathbb{R}.$$

In the model we assume that Y_i are independent given xw. Recall that a normal

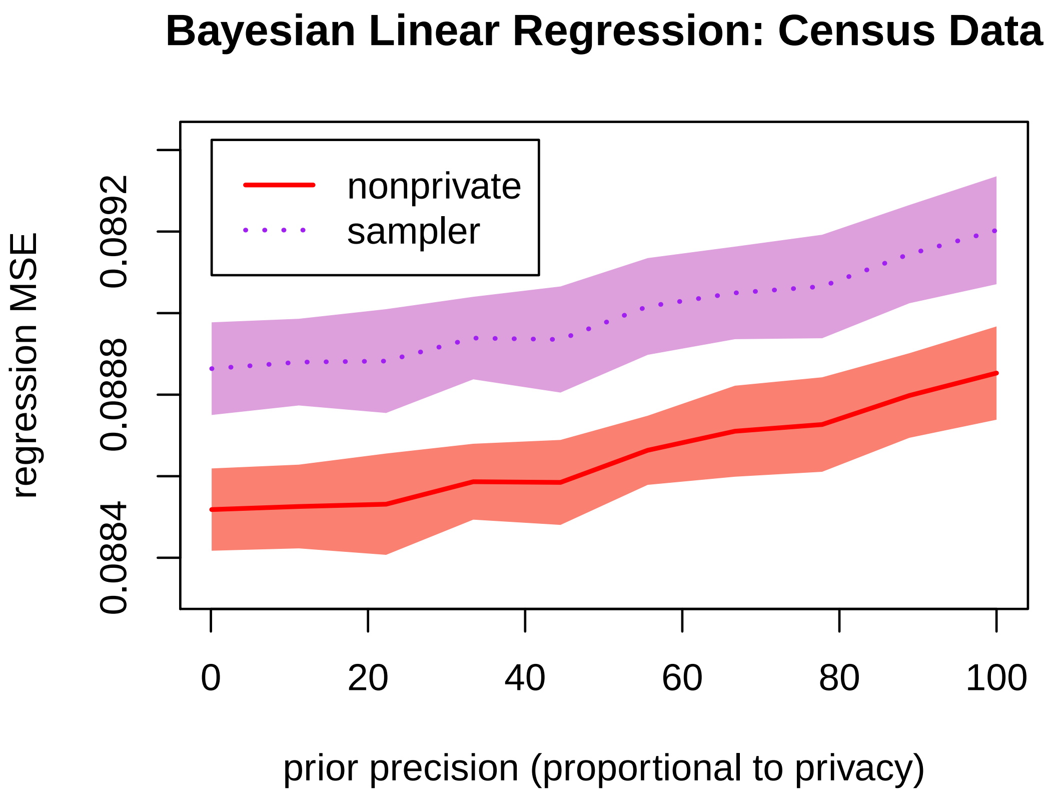

Fig. 4.2: Effect on linear regression of varying prior concentration. Bands indicate standard error over repeats.

linear regression model with i.i.d Gaussian noise is given as follows,

$$y_i = \sum_{j=1}^{d} x_i^{(j)} w^{(j)} + \epsilon_i, \epsilon_i \sim N(0, \sigma^2).$$

The normal likelihood function, as a product of likelihoods for each of the individual components of $y = (y_1, \ldots, y_n)$, is given by

$$p_w(y|x, w; \sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\frac{1}{2\sigma^2}(y-xw)^T(y-xw)}.$$

Given observations $\boldsymbol{D}$, we are interested in computing the sensitivity (in terms of data/observation) of this likelihood, that is $\sup_{w,D,D'} |\ln \frac{p_w(D)}{p_w(D')}|$. Note that $\ln \frac{p_w(D)}{p_w(D')} = \ln \frac{\prod_i p_w(x_i,y_i)}{\prod_i p_w(x'_i,y'_i)} = \sum_i \ln \frac{p_w(x_i,y_i)}{p_w(x'_i,y'_i)}$.

For simplicity, assume that the precision of Y is 1. Let $f_w(D)$ denote the log-likelihood, we have

$$|f_w(D) - f_w(D')| \leq \sum_i |f_w(x_i, y_i) - f_w(x'_i, y'_i).|$$

Note that by Mean Value Theorem, we have

$$\begin{aligned} & f_w(x_i, y_i) - f_w(x'_i, y'_i) \\ = \; & \nabla f_w((1-c)(x_i^{(1)}, \ldots, x_i^{(d)}, y_i) \\ & +c(x_i'^{(1)}, \ldots, x_i'^{(d)}, y'_i)) \cdot (x_i^{(1)} - x_i'^{(1)}, \ldots, x_i^{(d)} - x_i'^{(d)}, y_i - y'_i) \end{aligned}$$

Therefore by the Cauchy-Schwarz inequality we have:

$$\begin{aligned} \Delta f_w(x_i, y_i) &\leq \|\nabla f_w\|_2 \|(\Delta x_i, \Delta y_i)\|_2 \\ \Delta f_w(D, D') &\leq \sum_{i=1}^n \Delta f_w(x_i, y_i) \\ &\leq \|\nabla f_w\|_2 \sum_{i=1}^n \|(\Delta x_i, \Delta y_i)\|_2 \end{aligned}$$

Note that

$$\begin{aligned} df_w(x_i, y_i)/dx_i^{(j)} &= \frac{1}{2\sigma^2}(x_i^{(j)} w^T w - y_i w^{(j)}) \\ df_w(x_i, y_i)/dy_i &= \frac{1}{2\sigma^2}(y_i - w^T x_i^T) \end{aligned}$$

Recall that in linear regression, it is common to assume that every tuple (x_i, y_i) in the database satisfies $\|x_i\|_2 \leq 1$ and $\|y_i\|_2 \leq 1$, we have

$$\begin{aligned} \|\nabla f_w\|_2 \leq & \frac{1}{2\sigma^2} \sum_{i=1}^n \left(y_i - w^T x_i^T + \sum_{j=1}^d \left(x_i^{(j)} \|w\|_2 - y_i w^{(j)} \right) \right) \\ \leq & \frac{n}{2\sigma^2}(1 + 2\|w\|_1 + d\|w\|_2) \\ \leq & \frac{n}{2\sigma^2}(1 + (d+2)\|w\|_1) \end{aligned}$$

Hence the log-likelihood of normal regression satisfies Assumption 1 for $\rho(D, D') = \sum_{i=1}^{n} \|(\varDelta x_i, \varDelta y_i)\|_2$ under the condition that w is bounded.

For normal linear regression with bounded w, it is natural to choose a prior of w with truncated normal density, that is

$$p(w) \propto N(0, \Lambda^{-1})\mathbf{1}\{\|w\|_2 \leq 1\}$$

(In experiments we vary the norm bound for truncation with Λ. Our argument extends immediately.) As we show below, this truncated normal prior is still a conjugate prior for Gaussian likelihood.

Lemma 4.4.1. *The truncated Gaussian prior and the Gaussian likelihood of linear regression is a conjugate pair and the resulted posterior is a truncated Gaussian distribution.*

Proof. By Bayes's rule,

$$\begin{aligned} p(w|D) &\propto p(D|w)p(w) \\ &\propto N(w|\mu_n, \Sigma_n)\mathbf{1}\{\|w\|_2 \leq 1\} \end{aligned}$$

where $\mu_n = (X^T X + \sigma^2 \Lambda)^{-1} X^T y$ and $\Sigma_n = \sigma^2 (X^T X + \sigma^2 \Lambda)^{-1}$.

Therefore the posterior BAPS (Bayesian Posterior Sampling) on $p(w|D)$ is $2L(w)$-differentially private, where $L(w) = \frac{n}{2\sigma^2}(1 + 2\|w\|_1 + d\|w\|_2)$. □

4.5 Discussion

We have presented a suite of mechanisms for differentially-private inference in graphical models, in a Bayesian framework. The first two perturb posterior parameters to achieve privacy. This can be achieved either by performing perturbations in the original parameter domain, or in the frequency domain via a Fourier transform. Our third mechanism relies on the choice of a prior, in combination with posterior sampling. We complement our mechanisms for releasing the posterior, with private MAP point estimators. Throughout we have proved utility and privacy bounds for our mechanisms, which in most cases depend on the *graph structure of the Bayesian* *network: naturally, conditional independence affects privacy.* We support our new mechanisms and analysis with applications to two concrete models, with experiments exploring the privacy-utility trade-off.

Chapter 5

Differential Privacy and Information Leakage

In this chapter, we extend the prior work [Alvim et al., 2011a,b], which study differential privacy in an information-theoretic framework using Rényi min-entropy. By assuming certain symmetric properties of the graphs induced by the adjacency relation (Hamming-1 neighbouring) on datasets, they showed that differential privacy implies a bound on utility and provide an method that builds an optimal differentially-private mechanism. However, as we will show in this chapter, the above results actually hold without *any assumption* on the structure of induced graphs. Our result does not only break the limitation (in Alvim et al. [2011a,b]) of input datasets, it also allows us to consider differential privacy in this information-theoretic framework without any restriction on how we define "neighbouring datasets".

5.1 Information Leakage and Utility Model

As a generalization of the Shannon entropy, the Rényi entropy of order α (where $\alpha > 0, \alpha \neq 1$) of a random variable X is defined as

$$H_\alpha(X) = \frac{1}{1-\alpha} \log_2 \sum_{x \in \mathcal{X}} p(x)^\alpha.$$

In particular, we are interested in the so called *min-entropy* that is the limit of $H_\alpha(X)$ as α approaches infinity. It is known that $H_\infty(X) = -\log_2 \max_{x\in\mathcal{X}} p(x)$. For conditional entropy, we adopt the definition proposed in Dodis et al. [2004]:

$$H_\infty(X|Y) = -\log_2 \sum_{y\in\mathcal{Y}} p(y) \max_{x\in\mathcal{X}} p(x|y)$$

The min-entropy leakage is defined as $I_\infty(X;Y) = H_\infty(X) - H_\infty(X|Y)$. Braun et al. [2009] proved the worse-case leakage is obtained at the uniform input distribution, and it is equal to the sum of maxima of each column in the channel matrix: $\sum_{y\in\mathcal{Y}} \max_{x\in\mathcal{X}} p(y|x)$. Taking X as the input and Y as the output of a channel matrix, this min-entropy leakage is a natural measurement of the information that the attacker can learn about the database by observing the reported answers.

Alvim et al. [2011a] also use the converse of the *Bayes risk* to measure the utility of the channel matrix. Assume that Y is the true answer of the query function and Z is the perturbed output of the channel matrix. The author use the binary utility function to derive that the expected utility is given by

$$U(Y,Z) = \sum_z \max_y p(y,z).$$

This utility is closely relation to the min-entropy and mini-entropy leakage:

$$H_\infty(Y|Z) = -\log_2 U(Y,Z) \qquad I_\infty(Y;Z) = H_\infty(X) + \log_2 U(Y,Z).$$

5.2 Induced Graphs and Their Automorphism Groups

We adopt the notation in Alvim et al. [2011a] and view mechanisms as channel matrices. More precisely, we assume that A and B are random variables with domains $\mathcal{A}$ and $\mathcal{B}$. Let M be a channel matrix with input A and output B. The matrix M represents the conditional probability $p_{B|A}(\cdot|\cdot)$. Since B is determined by A and M, we use $B(M,A)$ to represent the dependency. For simplicity, we use $H^M_\infty(A)$ to denote the conditional min-entropy $H_\infty(A|B(M,A))$ and $I^M_\infty(A)$ to denote $I_\infty(A;B(M,A))$. As in the proofs of Alvim et al. [2011a], we use i,h,l ranging over rows and j,k ranging over columns of M. We use $\max^j M$ to denote the maximum value of column j over all rows i, i.e. $\max^j M = \max_i M_{i,j}$.

In this channel model, we fix a finite set of individuals participating in the input datasets. Two datasets are adjacent if and only if they differ for the value of exactly one individual. Note that the adjacency of two datasets is a symmetric relation. Therefore, a graph structure can be obtained on the datasets. For instance, $(\mathcal{A}, \sim)$ denotes the underlying graph resulted by the adjacency relation on all datasets in $\mathcal{A}$. In other words, $(\mathcal{A}, \sim)$ is the graph with all datasets in $\mathcal{A}$ as the vertices such that two datasets are adjacent if and only if they differ at exactly one coordinate (i.e. with Hamming distance one).

Let us introduce some standard definitions in group theory and symmetric graphs.

Definition 5.2.1. *Let G be a group acting on $\mathcal{A}$. The action of G on $\mathcal{A}$ is called transitive if for any $x, y \in \mathcal{A}$, there exists $g \in G$ such that $g(x) = y$.*

Definition 5.2.2. *Let G be a group acting on a set $\mathcal{A}$ and $a \in \mathcal{A}$. The $G - orbit$ on $\mathcal{A}$ containing a is defined as $Ga = \{g(a) | g \in G\}$.*

Definition 5.2.3. *An automorphism g of a graph $(\mathcal{A}, \sim)$ is a permutation of $\mathcal{A}$ such that $\{h, k\}$ is an edge of $(\mathcal{A}, \sim)$ if and only if $\{g(h), g(k)\}$ is an edge of $(\mathcal{A}, \sim)$.*

The set of automorphisms of $(\mathcal{A}, \sim)$ equipped with the usual composition of permutations is a group, called the automorphism group of $(\mathcal{A}, \sim)$ and denoted by $\mathrm{Aut}(\mathcal{A}, \sim)$. Any subgroup of $\mathrm{Aut}(\mathcal{A}, \sim)$ is called an automorphism group of $(\mathcal{A}, \sim)$.

Definition 5.2.4. *The distance between two vertices of the graph $(\mathcal{A}, \sim)$ is length of a shortest path between the two vertices. Given an integer $d \geq 0$, we define $Border_d(a)$ to denote the set of vertices at distance d from a in the graph.*

5.3 Improvement of the Utility Results

In this section, we show that Lemma 5 in [Alvim et al., 2011a] (as shown below) can be improved by dropping the condition on the induced graph $(\mathcal{A}, \sim)$.

Lemma 5.3.1. *Let M be the matrix of a channel with the same input and output alphabet $\mathcal{A}$ and $\sim$ a symmetric relation on $\mathcal{A}$ such that $(\mathcal{A}, \sim)$ has an automorphism* *with a single orbit. Assume that the maximum value of each column of M is on the diagonal, that is $M_{i,i} = \max^i M$ for all $i \in \mathcal{A}$. If M provides ϵ-differential privacy, then we can construct a new channel matrix M' such that:*

1. *M' provides ϵ-differential privacy;*
2. *$M'_{i,i} = M'_{h,h}$ for all $i, h \in \mathcal{A}$;*
3. *$M'_{i,i} = \max^i M$ for all $i \in \mathcal{A}$;*
4. *$H_\infty^M(A) = H_\infty^{M'}(A)$.*

This lemma plays a central role in proving the utility results in [Alvim et al., 2011a,b]. By saying that $(\mathcal{A}, \sim)$ has an automorphism with a single orbit, the authors mean that the orbits of the cyclic group are generated by the automorphism. Note that Hamming graphs generally have no automorphism such that the cyclic group generated by it is transitive on the vertex set. Hence the symmetric condition of induced graphs here is very restrictive. Our results show that the above lemma and the theorems implied by it can be proven without any assumption on the structure of $(\mathcal{A}, \sim)$.

Lemma 5.3.2. *Let M be the matrix of a channel with the same input and output alphabet $\mathcal{A}$ and $\sim$ a symmetric relation on $\mathcal{A}$. Assume that the maximum value of each column of M is on the diagonal, that is $M_{i,i} = \max^i M$ for all $i \in \mathcal{A}$. If M provides ϵ-differential privacy, then we can construct a new channel matrix M' such that:*

1. *M' provides ϵ-differential privacy;*
2. *$M'_{i,i} = M'_{h,h}$ for all $i, h \in \mathcal{A}$;*
3. *$M'_{i,i} = \max^i M$ for all $i \in \mathcal{A}$;*
4. *$H_\infty^M(A) = H_\infty^{M'}(A)$.*

Proof. Let G denote a subgroup of the automorphism group $Aut(\mathcal{A}, \sim)$ of the graph $(\mathcal{A}, \sim)$. For each pair $h, k \in \mathcal{A}$, define

$$M'_{h,k} = (1/|G|) \sum_{g \in G} M_{g(h),g(k)}. \tag{5.1}$$

First we prove that M' provides ϵ-differential privacy. Since for each $h \sim l$ and $g \in G$, we have $g(h) \sim g(l)$. For each pair $h \sim l$, we have $g(h) \sim g(l)$ and therefore for every k

$$M'_{h,k} = (1/|G|) \sum_{g \in G} M_{g(h),g(k)} \le (1/|G|) \sum_{g \in G} e^{\epsilon} M_{g(l),g(k)} = e^{\epsilon} M'_{l,k}. \tag{5.2}$$

We then prove that for every h, M'_h is a probability distribution

$$\sum_{k=1}^{|\mathcal{A}|} M'_{h,k} = \sum_{k=1}^{|\mathcal{A}|} (1/|G|) \sum_{g \in G} M_{g(h),g(k)} = (1/|G|) \sum_{g \in G} \sum_{k=1}^{|\mathcal{A}|} M_{g(h),g(k)} = 1. \tag{5.3}$$

We claim that the diagonal contains the maximal value of each column, because

$$M'_{k,k} = (1/|G|) \sum_{g \in G} M_{g(k),g(k)} \ge (1/|G|) \sum_{g \in G} M_{g(h),g(k)} = M'_{h,k}. \tag{5.4}$$

Finally, we show that $H_\infty^M(A) = H_\infty^{M'}(A)$. Suppose G has r orbits on $\mathcal{A}$, say, $\mathcal{A}_1, \ldots, \mathcal{A}_r$ with sizes $a_1, \ldots, a_r$, respectively. Then $\{\mathcal{A}_1, \ldots, \mathcal{A}_r\}$ is a partition of $\mathcal{A}$. Fix a vertex $k_i \in \mathcal{A}_i$ and let $G_{k_i} \triangleq \{g \in G : g(k_i) = k_i\}$ denote the stabilizer of k_i in G for each i. Then $g(k_i) = g'(k_i)$ if and only if $g'G_{k_i} = gG_{k_i}$. Since G is transitive on $\mathcal{A}_i$, by the Orbit-Stabilizer Lemma, $a_i = |\mathcal{A}_i| = |G|/|G_{k_i}|$. Let $\{g_{i1}, \ldots, g_{ia_i}\}$ be a set of representatives of $[G : G_{k_i}]$ (the set of left cosets of G_{k_i} in G). Then $\mathcal{A}_i = \{g_{i1}(k_i), \ldots, g_{ia_i}(k_i)\}$.

Since G is transitive on $\mathcal{A}_i$, by the Orbit-Stabilizer Lemma we have $|G| = |G_k||\mathcal{A}| = |G_k|n$ for any $k \in \mathcal{A}$, where $G_k \triangleq \{g \in G : g(k) = k\}$ is the stabilizer of k under the action of G. Note that G_k is a subgroup of G. Moreover, $g(k) = g'(k)$ if and only if $g' \in gG_k$. Let $\{g_0, g_1, \ldots, g_{n-1}\}$ be a set of representatives of $[G : G_k]$ (the set of left cosets of G_k in G). Then $\sum_{g \in G} M_{g(k),g(k)} = |G_k| \sum_{i=0}^{n-1} M_{g_i(k),g_i(k)} = |G_k| H_\infty^M(A)$. Therefore, we have

$$H_\infty^{M'}(A) = (1/|G|) \sum_{k=0}^{n-1} \Big(|G_k| H_\infty^M(A)\Big) = (1/n) \sum_{k=0}^{n-1} H_\infty^M(A) = H_\infty^M(A). \tag{5.5}$$

This completes the proof. □

By applying this new lemma, we immediately improve Theorem 3 in [Alvim et al., 2011a] as follows.

Theorem 5.3.3. *Let $\mathcal{H}$ be a randomization mechanism for the randomized function $\mathcal{K}$ and query function f, and assume that $\mathcal{K}$ provides ϵ-differential privacy. For an element a of $\mathcal{A}$ (the input alphabet of Y), let n denote the maximum distance from a in $\mathcal{A}$ and $c = \min_{1 \le d \le n} |Border_d(a)|$. Then*

$$U(Y, Z) \le \frac{(e^\epsilon)^n(1 - e^\epsilon)}{(e^\epsilon)^n(1 - e^\epsilon) + c(1 - (e^\epsilon)^n)}. \tag{5.6}$$

Since c is ranging from 1 to n, c must be non-zero, so this bound does tell us something. This upper bound would decrease as c and n increase. Given the premise of this chapter, the underlying graph is vertex-transitive and so no matter what vertex we choose the bound is the same. While for a general graph, we need to find the vertex such that c and n are large in order to get a good bound.

Similarly, Theorem 4 in [Alvim et al., 2011a] also follows without assuming that the induced graph has an automorphism with a single orbit.

The utility results in [Alvim et al., 2011a,b] hold for any matrix M' defined by $M'_{h,k} = (1/|G|) \sum_{g \in G} M_{g(h),g(k)}$ for any subgroup G of $\mathrm{Aut}(\mathcal{A}, \sim)$. No symmetric assumption of $\mathrm{Aut}(\mathcal{A}, \sim)$ need to be made at all. Our improvement extends the original results to a much broader class of database structures.

Chapter 6

Some Statistical Properties of Clique-inserted Lattices

We now turn to study other network statistics that are also related to graph structure. In this and the next two chapters, we will investigate the statistics of certain combinatorial objects on certain networks. In statistical mechanics, crystals are modelled by lattices whose atoms and bonds are represented by lattice vertices and edges respectively. Solving the enumeration problem of particular types of subgraphs on special types of lattices is an important area which is related to physical properties of materials in the real world. For instance, the dimer problem is related to the absorption of diatomic molecules on crystal surface. The Ising model is a mathematical model of ferromagnetism and the simplest statistical model of phase transition [Baxter, 1982, Gallavotti, 2013]. The enumeration of spanning trees and independent sets were also discussed in [Shrock and Wu, 2000]. Graph operations such as *subdivision-vertex joins*, *subdivision-edge joins* [Liu and Zhang, 2017] and *clique-insertion* [Zhang et al., 2009], preserve nice spectral properties. In this chapter, we recall the relationship between the spectra of an r-regular lattice and that of its clique-inserted lattice, and investigate the *graph energy* statistics. As an application, the asymptotic energy per vertex of the 3-12-12 and 3-6-24 lattices are computed. We also develop formulae expressing the numbers of spanning trees and dimer coverings of the k-th iterated clique-inserted lattices in terms of that of the original. Moreover, we show that new families of expander graphs can be constructed from known expanders by clique-inserting.

The operation of replacing every vertex of an r-regular lattice H by a complete graph of order r is called *clique-insertion*, and the resulting lattice is called the *clique-inserted lattice* of H. For any given r-regular lattice, applying this operation iteratively, an infinite family of r-regular lattices is generated. Some interesting lattices including the 3-12-12 lattice can be constructed this way.

6.1 Introduction

In the study of lattice statistical mechanics, one family of 2-dimensional lattices that have received much attention are constructed by replacing each vertex of r-regular lattices with a complete graph of order r such that each of the r new vertices corresponds to one of the incident edges. (To avoid triviality, we assume $r \geq 3$ throughout this thesis.) Such lattices include the martini [Scullard, 2006, Teufl and Wagner, 2010, Wu, 2006b], the 3-12-12 [Shrock and Wu, 2000, Teufl and Wagner, 2010, Wu, 2006b], the 3-6-24 [Guo et al., 2004] and the modified bath room lattices [Teufl and Wagner, 2010]. Following [Zhang et al., 2009], this operation of transforming each vertex of an r-regular graph to an r-clique (complete graph of order r) is called *clique-insertion*, and the graph obtained this way is called the clique-inserted graph of the original graph. From a given r-regular lattice H, the operation of clique-insertion can also be performed, and the resulting lattice $C(H)$ is called the *clique-inserted lattice* of the original lattice.

Throughout the rest of the thesis, we always assume that G denotes an undirected simple graph. Note that in the language of graph theory, the clique-insertion operation on a graph G can be described as taking the line graph of the subdivision graph of G. For any given regular lattice H, by iterating this operation, a set of hierarchical regular lattices, namely, iterated clique-inserted lattices can be obtained. Denote by $\{C^k(H)\}_{k\geq 0}$ the sequence of clique-inserted lattices with $C^0(H) \equiv H$ and $C^{k+1}(H) = C(C^k(H))$. Start with the hexagonal lattice, the 3-12-12, 3-6-24 and 3-6-12-48 lattices (refer to [Guo et al., 2004] for definitions of these lattices) can be generated by clique-insertion. In this case, the clique-insertion operation on a lattice is equivalent to the fundamental “Y-Delta" transformation (also known as the star-triangle transformation) on the subdivision graph of the original lattice. By this observation we obtained the relations between some physical and chemical indices of r-regular lattices and their k-th clique-inserted lattices. With such relations, we can compute some indices of certain complex lattices easily based on the results of well-studied lattices such as the square and hexagonal lattices.

In this chapter, we consider the lattices produced by the operation of clique-insertion on regular lattices with free, cylindrical and toroidal boundary conditions. We will discuss the *energy per vertex*, *average resistance* (the *Kirchhoff index* over the number of pairs of vertices) and the entropy of spanning-tree and dimer models of such lattices. We will also use the operation of clique-insertion to construct new families of expander graphs from known ones.

The dimer problem, the study of absorption diatomic molecules on crystal surface and phase transition in the Ising model, have attracted the attention of many physicists as well as mathematicians. For some classical works, we refer to [Fisher, 1961, Teufl and Wagner, 2010, Wu, 2006b]. Cayley [1889] and Kirchhoff [1847] presented the problem of enumeration of spanning trees of graphs, and further work in statistical physics has appeared in both the physics and mathematics literature. For a good survey, the reader is referred to [Shrock and Wu, 2000]. In the 1930s, Hückel proposed a method for finding approximate solutions of the Schrödinger equation of a class of organic molecules, the so-called conjugated hydrocarbons. In the framework of this modelization, the total π-electron energy, can be approximated by the sum of the absolute values of eigenvalues of the molecular graphs under certain chemical-based conditions. Gutman abstracted a mathematical notion from this application-driven analysis on molecular graphs, therefore he defined *graph energy* as a graph invariant [Gutman, 1978, 2001]. Since then, graph energy has been studied extensively by chemists and mathematicians. Yan and Zhang [2009] proposed the energy per vertex problem for lattice systems and showed that the energy per vertex of 2-dimensional lattices is independent from the boundary conditions, in certain settings. For a comprehensive survey of results and common proof methods obtained on graph energy, see the monograph on graph energy [Li et al., 2012] and references cited therein. *Expander graphs* were first defined by Bassalygo and Pinsker in the early 70's. These graphs are regular sparse graphs with strong connectivity properties, measured by vertex, edge or spectral expansion as described in [Hoory et al., 2006]. For a graph, having such a property has significant implications in various disciplines including complexity theory, computer networks, statistical mechanics and so on.

The rest of the chapter is organized as follows. The expression of the energy and Kirchhoff index of k-th iterated clique-inserted lattices of regular lattices are discussed in Sections 2 and 3, respectively. As an application, we compute the energy per vertex of the 3-12-12 and 3-6-24 lattices. In Section 4, we show that, given z_H as the entropy of spanning trees of an r-regular lattice H, the entropy of spanning trees of $C^k(H)$ (the k-th iterated clique-inserted graph of H) is given by $r^{-k}(z_H + s_k(r)\ln r(r+2))$ where $s_k(r) = (r/2-1)(r^k-1)/(r-1)$. We will also show that when H is cubic, the free energy per dimer of $C^k(H)$ is $\frac{1}{3}\ln 2$. In Section 5, inspired by Liu and Zhou's work [Liu and Zhou, 2014], we show that by applying the clique-insertion operation iteratively on an expander family, new families of expander graphs can be obtained. We propose clique-insertion as a modification to extend the size of computer networks, with their expansion properties being preserved to a certain degree.

6.2 Asymptotic Energy

Let $G = (V(G), E(G))$ be a graph with vertex set $V(G) = \{v_1, v_2, \ldots, v_n\}$ and edge set $E(G)$. The adjacency matrix of G, denoted by $A(G)$, is the $n \times n$ symmetric matrix such that $a_{ij} = 1$ if vertices v_i and v_j are adjacent and 0 otherwise. Let $d_G(v_i)$ be the degree of vertex v_i of G. The *Line graph* $L(G)$ of G, is the graph such that each vertex of $L(G)$ represents an edge of G and two vertices of $L(G)$ are adjacent if and only if their corresponding edges of G share a common end vertex in G. The *subdivision graph* $S(G)$ of a graph G is the graph obtained by inserting a new vertex into every edge of G. It is easy to see that $C(G) = L(S(G))$. The *energy* of a graph G with n vertices, denoted by $\mathcal{E}(G)$, is defined by

$$\mathcal{E}(G) = \sum_{i=1}^{n} |\lambda_i|,$$

where λ_i's are the eigenvalues of the adjacency matrix of G. The *asymptotic energy* per vertex of G [Yan and Zhang, 2009] is defined by $\lim_{|V(G)|\to\infty} \frac{\mathcal{E}(G)}{|V(G)|}$.

Lemma 6.2.1. *[Yan and Zhang, 2009] Suppose that* $\{G_n\}$ *is a sequence of finite simple graphs with bounded average degree such that* $\lim\limits_{n\to\infty} |V(G_n)| = \infty$ *and* $\lim\limits_{n\to\infty} \frac{\mathcal{E}(G_n)}{|V(G_n)|} = h \neq 0$*. If* $\{G'_n\}$ *is a sequence of spanning subgraphs of* $\{G_n\}$ *such that* $\lim\limits_{n\to\infty} \frac{|\{v\in V(G'_n): d_{G'_n}(v)=d_{G_n}(v)\}|}{|V(G_n)|} = 1$*, then* $\lim\limits_{n\to\infty} \frac{\mathcal{E}(G'_n)}{|V(G'_n)|} = h$*. That is,* G_n *and* G'_n *have the same asymptotic energy.*

Lemma 6.2.2. *[Zhang et al., 2009] Let* G *be an* r*-regular graph with* n *vertices and* m *edges. Suppose that the eigenvalues of* G *are* $\lambda_1 = r \geq \lambda_2 \geq \ldots \geq \lambda_n$*. Then the eigenvalues of the clique-inserted graph* $C(G)$ *of* G *are* $\frac{r-2\pm\sqrt{r^2+4(\lambda_i+1)}}{2}$, $i = 1, 2, \ldots, n$*, besides* -2 *and* 0 *each with multiplicity* $m - n$*.*

From Lemma 6.2.2, we immediately obtain the following corollary.

Corollary 6.2.3. *Let* G *be an* r*-regular* $(r \geq 3)$ *graph with* n *vertices and eigenvalues* $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$*, the energy of the clique-inserted graph of* G *is*

$$\mathcal{E}(C(G)) = \sum_{i=1}^{n} \sqrt{r^2 + 4(\lambda_i + 1)} + n.$$

We will use this result to calculate the asymptotic energy per vertex of the 3-12-12 lattice and its clique-inserted lattice in the rest of this section.

6.2.1 3-12-12 lattice

Our notation for the hexagonal lattices follows [Yan and Zhang, 2009]. The hexagonal lattices on a $n\times m$ torus, denoted by $H^t(n, m)$, are illustrated in Figure 6.1, where $(a_1, a_1^*), (a_2, a_2^*), \ldots, (a_{m+1}, a_{m+1}^*), (b_1, b_1^*), (b_2, b_2^*), \ldots, (b_{n+1}, b_{n+1}^*)$ are edges in $H^t(n, m)$.

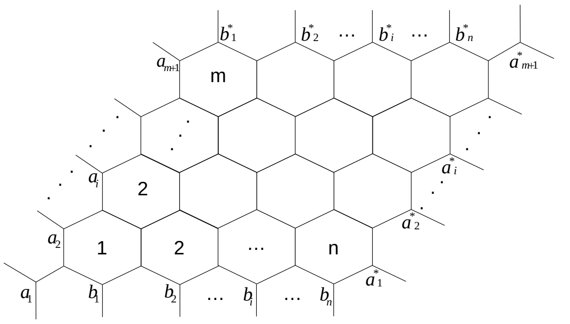

Fig. 6.1: $H^t(n,m)$ with toroidal boundary condition

By the definition of a clique-inserted lattice, it is easy to see that each 3-12-12 lattice on the same geometry is a clique-inserted-graph of $H^t(n,m)$, denoted as $T^t(n,m)$ (see Figure 6.2). Note that $(a_1,a_1^*),(a_2,a_2^*),\ldots,(a_{m+1},a_{m+1}^*)$, $(b_1,b_1^*),(b_2,b_2^*),\ldots,$ (b_{n+1},b_{n+1}^*) are edges in $T^t(m,n)$. If we delete edges $(b_1,b_1^*),(b_2,b_2^*),\ldots,(b_{n+1},b_{n+1}^*)$ from $T^t(n,m)$, then the 3-12-12 lattice with cylindrical boundary condition, denoted by $T^c(n,m)$ (see Figure 6.2) can be obtained. If we delete the edges $(a_1,a_1^*),(a_2,a_2^*),\ldots,$ (a_{m+1},a_{m+1}^*) from $T^c(m,n)$, then the 3-12-12 lattice with free boundary condition, denoted by $T^f(m,n)$ (see Figure 6.2) can be obtained.

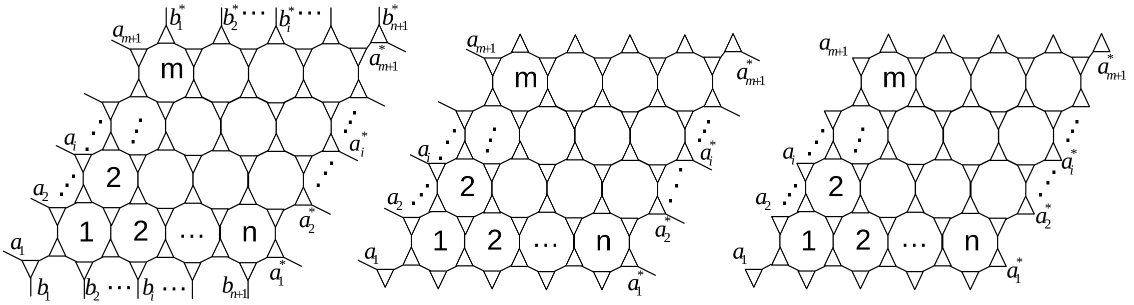

Fig. 6.2: The 3-12-12 lattice $T^t(n,m)$ (left), $T^c(n,m)$ (middle), and $T^f(n,m)$.

Note that almost all vertices of $T^c(m,n)$ and $T^f(m,n)$ are of degree 3. Since $T^f(m,n)$ and $T^c(m,n)$ are spanning subgraphs of $T^t(m,n)$, by Lemma 6.2.1 we have

$$\lim_{n,m\to\infty}\frac{\mathcal{E}(T^t(n,m))}{6mn}=\lim_{n,m\to\infty}\frac{\mathcal{E}(T^c(n,m))}{6mn}=\lim_{n,m\to\infty}\frac{\mathcal{E}(T^f(n,m))}{6mn}$$

It is shown in [Yan and Zhang, 2009] that the eigenvalues of $H^t(n,m)$ are:

$$\pm\sqrt{3+2\cos\frac{2i\pi}{n+1}+2\cos\frac{2j\pi}{m+1}+2\cos\left(\frac{2i\pi}{n+1}+\frac{2j\pi}{m+1}\right)},\quad 0\le i\le n, 0\le j\le m.$$

Since $T^t(n,m)$ is the clique-inserted graph of $H^t(n,m)$, we have

$$\begin{aligned}
&\mathcal{E}(T^t(n,m))\\
&=\sum_{i=0}^{n}\sum_{j=0}^{m}\sqrt{13+4\sqrt{3+2\cos\frac{2i\pi}{n+1}+2\cos\frac{2j\pi}{m+1}+2\cos\left(\frac{2i\pi}{n+1}+\frac{2j\pi}{m+1}\right)}}\\
&+\sum_{i=0}^{n}\sum_{j=0}^{m}\sqrt{13-4\sqrt{3+2\cos\frac{2i\pi}{n+1}+2\cos\frac{2j\pi}{m+1}+2\cos\left(\frac{2i\pi}{n+1}+\frac{2j\pi}{m+1}\right)}}+2mn\\
&=\sum_{i=0}^{n}\sum_{j=0}^{m}\sqrt{26+2\sqrt{121-32\cos\frac{2i\pi}{n+1}-32\cos\frac{2j\pi}{m+1}-32\cos\left(\frac{2i\pi}{n+1}+\frac{2j\pi}{m+1}\right)}}+2mn.
\end{aligned}$$

Thus, the average energy per vertex of the 3-12-12 lattice can be expressed as

$$\begin{aligned}
&\lim_{n,m\to\infty}\frac{\mathcal{E}(T^t(n,m))}{6mn}\\
&=\frac{1}{3}+\frac{1}{24\pi^2}\int_0^{2\pi}\int_0^{2\pi}\sqrt{26+2\sqrt{121-32\cos x-32\cos y-32\cos(x+y)}}dxdy\\
&=1.4825\ldots.
\end{aligned}$$

The last line follows by a numerical integration. Therefore, the 3-12-12 lattices $T^t(n,m), T^c(n,m)$, and $T^f(n,m)$ with toroidal, cylindrical, and free boundary conditions have the same asymptotic energy ($\approx 8.895mn$).

6.2.2 3-6-24 lattice

The clique-inserted lattice of $T^t(m,n)$ is a lattice with toroidal boundary condition, denoted by $S^t(m,n)$, illustrated in Figure 6.3). Note that $(a_1,a_1^*),(a_2,a_2^*),\ldots,$ (a_{m+1},a_{m+1}^*), (b_1,b_1^*), (b_2,b_2^*), $\ldots,(b_{n+1},b_{n+1}^*)$ are edges in $S^t(m,n)$. If we delete edges $(b_1,b_1^*),(b_2,b_2^*),\ldots,(b_{n+1},b_{n+1}^*)$ from $S^t(n,m)$, then the 3-6-24 lattice with cylindrical boundary condition, denoted by $S^c(n,m)$ (see Figure 6.3)) can be obtained. If we delete edges $(a_1,a_1^*),(a_2,a_2^*),\ldots,(a_{m+1},a_{m+1}^*)$ from $S^c(m,n)$, then the 3-6-24 lattice

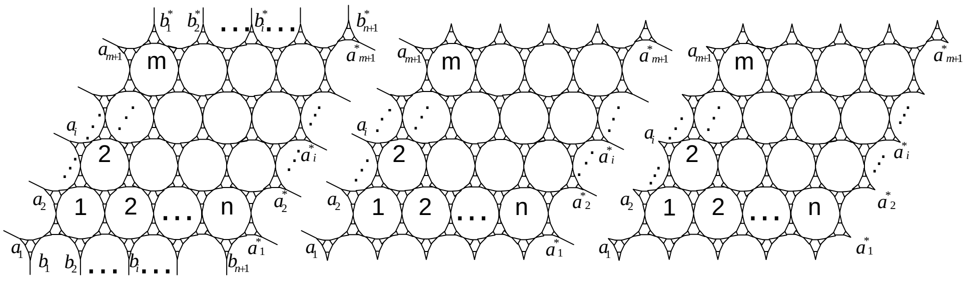
b_1^* b_2^* ... b_i^* ... b_{n+1}^*
a_{m+1}
m
a_{m+1}^*
a_i
2
a_i^*
a_2
1
2
...
n
a_2^*
a_1
a_1^*
b_1 b_2 ... b_i ... b_{n+1}

Fig. 6.3: The 3-6-24 lattice $S^t(n,m)$ (left), $S^c(n,m)$ (middle), and $S^f(n,m)$.

with free boundary condition, denoted by $S^f(m,n)$ (see Figure 6.3) can be obtained.

Note that $S^f(m,n)$ and $S^c(m,n)$ are spanning subgraphs of $S^t(m,n)$, by Lemma 6.2.1 we have

$$\lim_{n,m\to\infty}\frac{\mathcal{E}(S^t(n,m))}{18mn} = \lim_{n,m\to\infty}\frac{\mathcal{E}(S^c(n,m))}{18mn} = \lim_{n,m\to\infty}\frac{\mathcal{E}(S^f(n,m))}{18mn}.$$

The energy of the clique-inserted-graph of 3-12-12 lattice can be obtained by

$$\begin{aligned}
&\mathcal{E}(S^t(n,m))\\
&= \sum_{i=0}^{n}\sum_{j=0}^{m}\sqrt{30+2\sqrt{173-16\sqrt{3+2\cos\frac{2i\pi}{n+1}+2\cos\frac{2j\pi}{m+1}+2\cos\left(\frac{2i\pi}{n+1}+\frac{2j\pi}{m+1}\right)}}}\\
&+ \sum_{i=0}^{n}\sum_{j=0}^{m}\sqrt{30+2\sqrt{173+16\sqrt{3+2\cos\frac{2i\pi}{n+1}+2\cos\frac{2j\pi}{m+1}+2\cos\left(\frac{2i\pi}{n+1}+\frac{2j\pi}{m+1}\right)}}}\\
&+ \sqrt{5}mn+\sqrt{13}mn+6mn.
\end{aligned}$$

Then the average energy per vertex of the clique-inserted lattice of the 3-12-12 lattice is given by

$$\begin{aligned}
&\lim_{n,m\to\infty}\frac{\mathcal{E}(S^t(n,m))}{18mn}\\
&= \frac{1}{72\pi^2}\int_0^{2\pi}\int_0^{2\pi}\Bigg(\sqrt{30+2\sqrt{173-16\sqrt{3+2\cos x+2\cos y+2\cos(x+y)}}}
\end{aligned}$$

$$+ \sqrt{30 + 2\sqrt{173 + 16\sqrt{3 + 2\cos x + 2\cos y + 2\cos(x+y)}}}\Bigg)dxdy$$
$$+ \frac{\sqrt{5} + \sqrt{13} + 6}{18}$$
$$= 1.4908\ldots.$$

Thus, the lattices $S^t(n,m)$, $S^c(n,m)$, and $S^f(n,m)$ with toroidal, cylindrical, and free boundary conditions have the same asymptotic energy ($\approx 26.8344mn$).

6.3 Spanning Trees and Dimer Coverings

6.3.1 Spanning Trees

Let $N_{ST}(G)$ denote the number of spanning trees of G. When G is a periodic lattice in finite dimension $D > 1$, $N_{ST}(G)$ has asymptotic exponential growth. Define the quantity z_G by

$$z_G = \lim_{n\to\infty} \frac{1}{n} \ln N_{ST}(G).$$

This quantity, corresponding to the free energy per site in the thermodynamic limit, is called *bulk free energy*. The following lemma indicates the relation between the number of spanning trees of a regular lattice and of its k-th iterated clique-inserted lattice.

Lemma 6.3.1. *[Yan et al., 2008] Let G be an r-regular graph with n vertices. Then the number of spanning trees of the iterated clique-inserted-graphs $C^k(G)$ of G can be expressed by $N_{ST}(C^k(G)) = r^{ns-k}(r+2)^{ns+k}N_{ST}(G)$, where $s = s_k(r) = (r/2-1)(r^k - 1)/(r-1)$.*

Therefore, we have the following proposition.

Proposition 6.3.2. *Let H be an r-regular lattice. For $C^k(H)(k = 0, 1, 2, \ldots)$, the rate of growth of the number of spanning trees, $z_{C^k(H)}$, is given by $r^{-k}(z_H + s\ln r(r+2))$, where $s = (r/2-1)(r^k-1)/(r-1)$ and z_H denotes the rate of growth of spanning trees of H.*

The next theorem implies that the boundary condition does not affect the bulk limit of a lattice.

Theorem 6.3.3. *[Lyons, 2005] Let $\langle G_n \rangle$ be a tight sequence of finite connected graphs with bounded average degree such that*

$$\lim_{n\to\infty} \mid V(G)_n \mid^{-1} \mid \{x \in V(G'_n); deg_{G'_n}(x) = deg_{G_n}(x)\} \mid = 1,$$

then $\lim\limits_{n\to\infty} \mid V(G)_n \mid^{-1} \log \ N_{ST}(G'_n) = h.$

For the hexagonal lattice, z_{hc} is $0.8076649\ldots$ as shown in [Shrock and Wu, 2000]. Thus, by Proposition 6.3.2 and Theorem 6.3.3, we have that for the 3-12-12 and 3-6-24 lattices with toroidal, cylindrical and free boundary condition,

$$z_{3-12-12} = 0.7205633\ldots$$

$$z_{3-6-24} = 0.6915295\ldots.$$

6.3.2 Dimer Coverings

Let $M(G)$ denote the number of dimer coverings (perfect matchings) of G. The free energy per dimer of G, denoted by Z_G, is defined as $Z_G = \lim\limits_{n\to\infty} \frac{2}{n} \ln M(G)$. Given the number of vertices and edges of a connected graph, the number of dimer coverings of the graph and of its line graph have the following relation.

Lemma 6.3.4. *[Dong et al., 2013] Let G be a 2-connected graph of order n and size m, where m is even and $\Delta(G)$ is the maximum degree of G. Then $M(L(G)) \geq 2^{m-n+1}$, where the equality holds if and only if $\Delta(G) \leq 3$.*

With this general result, we can readily obtain the following.

Proposition 6.3.5. *Let H be a cubic lattice with toroidal boundary conditions. The free energy per dimer of $C^k(H)$ $(k = 1, 2, 3, \ldots)$ is equal to $\frac{1}{3} \ln 2$.*

Proof. Assume that H has n vertices. Since $C^k(H)$ is the line graph of the subdivision of $C^{k-1}(H)$, by Lemma 6.3.4 we have $Z_{C^k}(G) = \lim\limits_{n\to\infty} \frac{2}{3^k n} \ln 2^{3^k n - \frac{5}{6} \cdot 3^k n + 1} = \frac{1}{3} \ln 2$. □

Example 6.3.6. *Let $R^t(m,n)$ be the k-th iterated clique-inserted lattice of the hexagonal lattice $H^t(m,n)$ with toroidal boundary. Note that the corresponding lattice $R^c(m,n)$ ($R^f(m,n)$) with cylindrical (free) boundary condition can be considered as the line graph of a graph which differs from $S(C^{k-1}(H^tm,n))$ by a small number (small in the sense that the number is $o(mn)$ as m,n approach infinity) of edges. Therefore, by applying Lemma 6.3.4, we have $Z_{R^t(m,n)} = Z_{R^c(m,n)} = Z_{R^f(m,n)} = \frac{1}{3}\ln 2$.* □

In general, when a cubic lattice is a line graph, the free energy per dimer of plane lattices are the same as that of the corresponding cylindrical and toroidal lattices. However, this may not be true when a cubic lattice is not a line graph. The hexagonal lattice is such a counterexample as shown in [Yan et al., 2008].

6.4 Expansion Property

Let $D(G) = \mathrm{diag}(d_G(v_1), d_G(v_2), \ldots, d_G(v_n))$ be the diagonal matrix of vertex degree of G. The Laplacian matrix of G is $L(G) = D(G) - A(G)$. The eigenvalues of $L(G)$, denoted by $\mu_1 \le \mu_2 \le \cdots \le \mu_n$ are called the Laplacian spectrum of G. It is well known that μ_2, called the *algebraic connectivity* of G, is greater than 0 if and only if G is a connected graph. The *spectral gap* of G is defined as the difference of the largest and the second largest eigenvalues of $A(G)$. Note that for a regular graph, $\mu_i = r - \lambda_i$ for $i = 1, 2, \ldots, n$, which implies that its spectral gap is equal to its algebraic connectivity. Here we use spectral gap to quantify the expansion property, that is, a family of regular graphs is an expander family if and only if there is a positive lower bound for their spectral gaps, and the larger the bound the better the expansion. This characterization can be formulated as follows:
An infinite family of regular graphs, $G_1, G_2, G_3, \ldots$, is called a family of ε-expander graphs [Hoory et al., 2006], where $\varepsilon > 0$ is a fixed constant, if (i) all these graphs are k-regular for a fixed integer $k \ge 3$; (ii) $\mu_2 \ge \varepsilon$ for $i = 1, 2, 3, \ldots$; and (iii) $n_i = |V(G_i)| \to \infty$ as $i \to \infty$. Note that Lemma 6.2.2 implies that

$$\mu_2(C(G)) = \frac{r + 2 - \sqrt{(r+2)^2 - 4\mu_2(G)}}{2}.$$

Denote the function iteration of

$$f(x) = \frac{r + 2 - \sqrt{(r+2)^2 - 4x}}{2}$$

by $f^1(x) = f(x)$ and $f^{k+1}(x) = f(f^k(x))$ for $k = 1, 2, 3, \ldots$.

One primary application of expander graphs is in designing robust computer networks. In the study of computer networks, it would be helpful to find simple and local graph operations to enlarge networks such that the new networks share similar topological properties with the old ones. For instance, Saad and Schultz [1988] studied the mapping which maps grid to hypercubes and found many topological properties are preserved under such an operation. In our case, applying clique-insertion on networks can be considered as replacing each workstation by a cluster (modelled by a complete graph) and rewiring them properly. By the following result, we will see that this provides a modest modification to enlarge the networks such that their expansion properties are maintained in some sense.

Proposition 6.4.1. *Suppose $G_1, G_2, G_3, \ldots,$ is a family of r-regular ϵ-expander graphs. Then $C^k(G_1), C^k(G_2), C^k(G_3), \ldots,$ is a family of r-regular $f^k(\epsilon)$-expander graphs.*

Proof. Let $x = (\frac{2}{r+2})^2\epsilon$, then

$$\begin{aligned} f(\epsilon) &= \frac{r+2}{2}(1 - \frac{\sqrt{(r+2)^2 - 4\epsilon}}{2}) \\ &= \frac{(r+2)}{2}(1 - \sqrt{1-x}) \\ &= \frac{(r+2)}{2}\left(\frac{1}{2}x + \frac{1}{8}x^2 + \frac{1}{16}x^3 + \cdots\right) \approx \frac{\epsilon}{(r+2)}. \end{aligned}$$

The statement holds by definition. □

This implies that the lower bound of the spectral gaps of the new expander family obtained by clique-insertion is a linear term of that of the original expander family. Note that it is simple and intuitive enough to perform realistic operations on networks according to clique-insertion. So even if the expansion properties of clique-inserted lattices are not exceptional, it is still meaningful to consider clique-insertion as an approach to extend computer networks, because in reality, the trade-off between performance and simplicity needs to be taken into account.

Let us apply clique-insertion to the famous expander family $X^{p,q}$ of Lubostzky, Phillips and Sarnak [Davidoff et al., 2003]. Recall that for a fixed real number $0 < \gamma < 1/6$ and sufficiently large q, the spectral gap of $X^{p,q}$ is bounded from below by $\varepsilon(r) = (p+1) - p^{\frac{5}{6}+\gamma} - p^{\frac{1}{6}-\gamma}$. By Proposition 6.4.1, for a fixed odd prime p, $C(X^{p,q})$ is a $(p + 3 - \sqrt{p^2 + 2p + 4p^{\frac{5}{6}+\gamma} + 4p^{\frac{1}{6}-\gamma} + 5})/2$-expander family with degree $p + 1$. More generally, $C^k(X^{p,q})$ is a $f^k((p + 1) - p^{\frac{5}{6}+\gamma} - p^{\frac{1}{6}-\gamma})$-expander family.

Chapter 7

Enumeration of Independent Sets on the 4-8-8 Lattice

In this and the next chapter, we continue our discussion of network statistics. In particular, we investigate the statistics of vertex independent sets on some (random) networks using the transfer matrix method. In this chapter, we will first propose the concept of transfer multiplicity and the multi-step transfer matrices method to study more complicated lattices where the single step transfer matrix approach as in [Neil J Calkin, 1998] is not compatible. We demonstrate our method on the 4-8-8 lattice by providing numerical results of the number of independent sets and a rigorous bound of its entropy. We will also show that this entropy constant of a two dimensional lattice with free boundary condition is the same as the entropy constants of the corresponding cylindrical and toroidal lattices.

7.1 Background

A typical problem in lattice statistics is to count the number of ways of putting particles on lattice sites such that no two share the same sites or are in adjacent sites. In particular, for the two dimensional lattice gas model it is assumed that all of the gas molecules lie on the vertices of a square lattice and only interact with their four grid neighbours. The grid is taken to be rigid and square, so the limit of partition function per vertex is called "the hard square constant" [Baxter et al., 1980]. The model has also been studied on the triangular and hexagonal lattices [Baxter, 1982, 1999, Domb and Green, 1972, Finch, 1999, Pearce and Seaton, 1988].

In combinatorics, this enumeration problem is equivalent to counting vertex independent sets on graphs. This problem also exists in structural chemistry, where Merrifield and Simons defined a topological space for chemical graphs [Merrifield and Simmons, 1989]. The cardinality of that topological space, Merrified-Simmons index, is the number of independent sets on the chemical graphs. The properties of this index of some types of benzenoids and polyominoes have been studied in [Gutman, 1993, Ren and Zhang, 2007, Zhang and Tian, 2003, Zeng and Zhang, 2007]. The 4-8-8 lattice is an Archimedean tilling which has been used to describe phase transitions in the layered hydrogen-boded $SnCl_2 \cdot 2H_2O$ crystal [Salinas and Nagle, 2000] in physical systems. The triangular Kagomé lattice corresponds to the positions of Cu atoms in the fabricated materials $Cu_9X_2(cpa)_6{\cdot}xH_2O$ (cpa= 2-carboxypentonic acid, an derivative of asorbic acid; X-F,Cl,Br) [Maruti and ter Haar, 1994]. Various physical models such as the spanning trees, dimer covering and bond percolation have been studied on these two lattices [Allen, 1974, Haji-Ankabari and Ziff, 2009, Liu and Yan, 2013, Loh et al., 2008, Salinas and Nagle, 2000, Shrock and Wu, 2000, Wu, 2006b, Yan et al., 2008].

7.2 Transfer Matrix

Let m and n be positive integers. $G_{m,n}$ denote a finite section of the 4-8-8 lattice whose hexagons are arranged in m rows and n columns as shown in Fig.7.1.

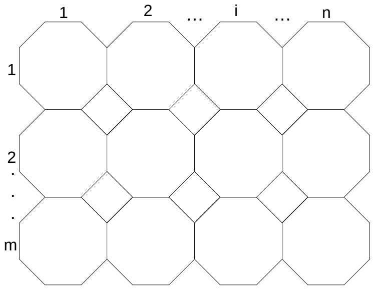

Fig. 7.1: Grid graph $G_{m,n}$

Given an independent set S of $G_{m,n}$, a portion of S that lies in a fixed column of $G_{m,n}$ can be represented by either an $(m+1)$-vector or a $2m$-vector of 0's and 1's, where a 1 indicates that the vertex is in S and a 0 indicates that the vertex is not in S. Thus, any independent set of $G_{m,n}$ can be represented by $3n+1$ column vectors.

Example 2. *Fig.7.2 below shows an independent set S in $G_{3,4}$. The portions of S that lie in each of the columns are represented by the respective vectors:*
$(1,0,0,0,1,0),(0,1,0,1),(1,0,1,0),(0,0,1,0,0,1),(0,0,1,0),(0,1,0,1),(1,0,0,1,1,0),$
$(0,0,0,0),(1,1,1,1),(0,0,0,0,0,0),(0,1,0,0),(0,0,1,1),(0,1,1,0,0,0).$

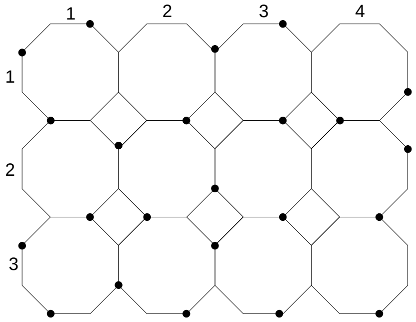

Fig. 7.2: An independent set of $G_{3,4}$.

Let $\mathbf{P}_m$ denote the set of all possible $(m+1)$-vectors which may represent a column in an independent set of $G_{m,n}$. The $(3k+2)$-th and $(3k+3)$-th columns of an independent set of $G_{m,n}$ are chosen from $\mathbf{P}_m$, where $0 \leq k \leq n-1$. Clearly, $\mathbf{P}_m$ is a collection of $(m+1)$-vectors of 0's and 1's, and the cardinality of $\mathbf{P}_m$ is 2^{m+1}. Similarly, let $\mathbf{Q}_m$ denote the set of all possible $2m$-vectors which may represent a $(3k+1)$-th column in an independent set of $G_{m,n}$, where $0 \leq k \leq n$. It is easy to see that $\mathbf{Q}_m$ consists of $2m$-vectors of 0's and 1's in which no consecutive 1's occupy the positions of the $(2k-1)$-th and $2k$-th entries, for $1 \leq k \leq m$. Since there are three possibilities at each pair of consecutive $(2k-1)$-th and $2k$-th positions, the set Q_m has 3^m vectors.

Let $\mathbf{v}(i)$ denote the i-th element of a vector $\mathbf{v}$. Note that two column vectors $\mathbf{v}$, $\mathbf{v}'$ need to fit into one of the following two cases to be a possible consecutive pair of columns in an independent set of $G_{m,n}$,
(i) $\mathbf{v}$ and $\mathbf{v}'$ are chosen from $\mathbf{P}_m$ and they are orthogonal to each other;

(ii) $\mathbf{v}$, $\mathbf{v}'$ are chosen from $\mathbf{P}_m$ and $\mathbf{Q}_m$ respectively and they satisfy the following conditions:

(a) $\mathbf{v}(1) \cdot \mathbf{v}'(1) = 0$,

(b) $\mathbf{v}(i) \cdot \mathbf{v}'(2i-2) = 0$ and $\mathbf{v}(i) \cdot \mathbf{v}'(2i-1) = 0$ for $i = 2, 3, \ldots, m$,

(c) $\mathbf{v}(m+1) \cdot \mathbf{v}'(2m) = 0$.

For fixed m and n, all possible independent sets S of $G_{m,n}$ can be obtained by a gluing process described below. First, we take a vector from $\mathbf{Q}_m$ such that it corresponds to the first column of S, and denote it as $\mathbf{u}$. Second, to the right of $\mathbf{u}$ we glue a vector $\mathbf{v}$ selected from $\mathbf{P}_m$, making sure that $\mathbf{u}$ and $\mathbf{v}$ satisfy (ii). Third, we glue a vector $\mathbf{v}'$ from $\mathbf{P}_m$ to the right of $\mathbf{v}$ under the condition that $\mathbf{v}$ and $\mathbf{v}'$ satisfy (i). Then, we glue $\mathbf{u}'$ to the right of $\mathbf{v}'$ such that $\mathbf{v}'$ and $\mathbf{u}'$ satisfy (ii). Repeat the above procedure until the $(3n+1)$-th column is glued.

Let us call the gluing of the $(i+1)$-th column v_{i+1} to the right of i-th column v_i as step i. Thus the transfer matrix representing step one, denoted $T_{m_1} = (T_{v_1,v_2})$, is a $3^m \times 2^{m+1}$ matrix whose rows are indexed by vectors of Q_m and columns are indexed by vectors of P_m, where $T_{v_1,v_2} = 1$ if v_1 and v_2 represent possible consecutive pair of columns in an independent set of $G_{m,n}$ and $T_{v_1,v_2} = 0$ otherwise. Note that the matrix depends only on n.

Similarly, the transfer matrix for step two is a $2^{m+1} \times 2^{m+1}$ matrix T_{m_2} whose rows and columns are indexed by vectors of P_m. The entry of T_{m_2} in position (α, β) is 1 if the vectors represented by α, β are orthogonal, and is 0 otherwise. The transfer matrix, T_{m_3}, for step three is the transpose of T_{m_1}. Note that T_{m_1} is the transfer matrix for every step i when $i = 3k+1$ $(0 \leq k \leq n-1)$, T_{m_2} is the transfer matrix for every step i when $i = 3k+2$ $(0 \leq k \leq n-1)$ and T_{m_3} is the transfer matrix for every step i when $i = 3k+3$ $(0 \leq k \leq n-1)$. Thus, if we take the transfer matrix of $G_{m,n}$ to be T_m=$T_{m_1}T_{m_2}T_{m_3}$, then it is well known that for the number of independent sets of $G_{m,n}$, we have

$$f(m,n) = \mathbf{1} \cdot T_m^n \mathbf{1},$$

where $\mathbf{1}$ denotes the all-one column vector. We call T_m a triple-step transfer matrix since it is given as the product of three transfer matrices.

When $m = 2$, the transfer matrices T_{2_1}, T_{2_2} and T_{2_3} are:

$$T_{2_1} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad T_{2_2} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad T_{2_3} = T_{2_1}^{\top}.$$

Thus, the triple-step transfer matrix of $G_{m,n}$ is:

$$T_2 = \begin{pmatrix} 27 & 18 & 18 & 18 & 12 & 18 & 18 & 12 & 12 \\ 18 & 9 & 12 & 12 & 6 & 12 & 12 & 6 & 8 \\ 18 & 12 & 9 & 9 & 6 & 9 & 12 & 8 & 6 \\ 18 & 12 & 9 & 9 & 6 & 9 & 12 & 8 & 6 \\ 12 & 6 & 6 & 6 & 3 & 6 & 8 & 4 & 4 \\ 18 & 12 & 9 & 9 & 6 & 9 & 12 & 8 & 6 \\ 18 & 12 & 12 & 12 & 8 & 12 & 9 & 6 & 6 \\ 12 & 6 & 8 & 8 & 4 & 8 & 6 & 3 & 4 \\ 12 & 8 & 6 & 6 & 4 & 6 & 6 & 4 & 3 \end{pmatrix}.$$

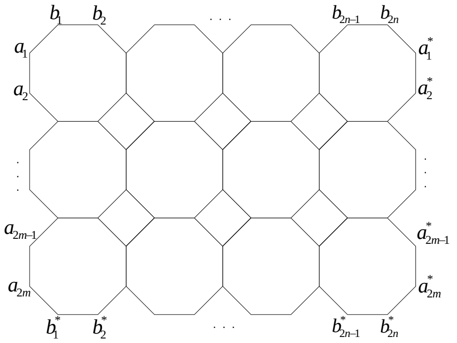

Fig. 7.3: The 4-8-8 lattice $G_{m,n}$ with free boundary condition

By identifying edges (a_i, a_{i+1}) with (a_i^*, a_{i+1}^*) $(1 \leq i \leq 2m-1)$ of $G_{m,n}$ respectively, the 4-8-8 lattice with cylindrical boundary condition, denoted as $H_{m,n}$, can be obtained (see Figure 7.3). Note that here the graph can be seen as drawn on a vertical cylinder.

Similar to the discussion of $G_{m,n}$, any independent set I of $H_{m,n}$ can be generated by gluing vectors from $\mathbf{P}_{n-1}$ and $\mathbf{Q}_n$. Note that if $\mathbf{v}$ and $\mathbf{v}'$ are chosen from $\mathbf{P}_{n-1}$ and $\mathbf{Q}_n$, they are a possible consecutive pair of vectors in I if and only if

$$\mathbf{v}(1) \cdot \mathbf{v}'(1) + \mathbf{v}(1) \cdot \mathbf{v}'(2n) = 0, \mathbf{v}(i) \cdot \mathbf{v}'(2i-2) = 0,$$

and

$$\mathbf{v}(i) \cdot \mathbf{v}'(2i-1) = 0,$$

for $i = 2, 3, \ldots, n$.
I can be obtained by the following procedure . Firstly let us take a vector from $\mathbf{Q}_n$ such that it corresponds to the first row of I, and denote it as $\mathbf{u}$. Secondly, we glue a vector $\mathbf{v}$ from $\mathbf{P}_{n-1}$ to the bottom of $\mathbf{u}$, ensuring that $\mathbf{u}$ and $\mathbf{v}$ are legitimate consecutive vectors in I. Thirdly, we glue a vector $\mathbf{v}'$ from $\mathbf{P}_{n-1}$ to the bottom of $\mathbf{v}$ such that $\mathbf{v}$ and $\mathbf{v}'$ are orthogonal. Then we glue $\mathbf{u}'$ to the bottom of $\mathbf{v}'$, making sure that $\mathbf{v}'$ and $\mathbf{u}'$ are possible consecutive vectors in I. Repeat the above procedure until the $(3m+1)$-th column is glued.

Consider the transfer matrix of $H_{m,n}$. The transfer matrix B_{n_1}, which represents every $(3k+1)$-th $(0 \leq k \leq n-1)$ step, can be defined as a $3^n \times 2^n$ matrix of 0's and 1's as follows. The rows of B_{n_1} are indexed by vectors of $\mathbf{Q}_n$ and columns are indexed by vectors of $\mathbf{P}_{n-1}$, and the entry of B_{n_1} in position (α, β) is 1 if α, β represent a possible consecutive pair of rows in an independent set of $H_{m,n}$, and is 0 otherwise. Let B_{n_3} denote the transfer matrix that represents every $(3k+3)$-th $(0 \leq k \leq n-1)$ step. It is not difficult to see that B_{n_3} is the transpose of B_{n_1}. The transfer matrix B_{n_2} that represents every $(3k+2)$-th $(1 \leq k \leq n-1)$ step is a $2^n \times 2^n$ matrix whose rows and columns are indexed by vectors of $\mathbf{P}_{n-1}$. The entry of B_{n_2} in position (α, β) is 1 if α, β are orthogonal, and is 0 otherwise. Thus if we take the transfer matrix of $H_{m,n}$ to be $B_n = B_{n_1} B_{n_2} B_{n_3}$, for the number of independent sets of $H_{m,n}$, we have

$$g(m,n) = \mathbf{1} \cdot B_n^m \mathbf{1},$$

where **1** denotes the all-one column vector.

When $n = 2$, the transfer matrices B_{2_1}, B_{2_2} and B_{2_3} are

$$B_{2_1} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad B_{2_2} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad B_{2_3} = B_{2_1}^{\top}.$$

Therefore, the triple-step transfer matrix of $H_{m,n}$ is

$$B_2 = \begin{pmatrix} 9 & 6 & 6 & 6 & 4 & 6 & 6 & 6 & 4 \\ 6 & 3 & 4 & 4 & 2 & 4 & 3 & 3 & 2 \\ 6 & 4 & 3 & 3 & 2 & 3 & 4 & 4 & 2 \\ 6 & 4 & 3 & 3 & 2 & 3 & 4 & 4 & 2 \\ 4 & 2 & 2 & 2 & 1 & 2 & 2 & 2 & 1 \\ 6 & 4 & 3 & 3 & 2 & 3 & 4 & 4 & 2 \\ 6 & 3 & 4 & 4 & 2 & 4 & 3 & 3 & 2 \\ 6 & 3 & 4 & 4 & 2 & 4 & 3 & 3 & 2 \\ 4 & 2 & 2 & 2 & 1 & 2 & 2 & 2 & 1 \end{pmatrix}.$$

The 4-8-8 lattice with toroidal boundary condition, denoted as $T_{m,n}$, can be obtained by identifying edges (b_i, b_{i+1}) with (b_i^*, b_{i+1}^*) $(1 \le i \le 2n-1)$ of $H_{m,n}$ respectively (see Fig. 7.3).

In general, for a two dimensional lattice $P_{m,n}$ with free boundary condition and transfer matrix T_m, let us denote its corresponding lattices with cylindrical boundary condition and toroidal boundary condition by $C_{m,n}$ (with transfer matrix B_n) and $T_{m,n}$ respectively.

We say that the square lattice is of transfer multiplicity one since computing the number of independent sets for it only needs to employ single-step transfer matrices as shown in [Neil J Calkin, 1998]. This may not hold for other lattices. For example, in [Zhang, 2006], for a generalized Aztec diamond we need to introduce double-step transfer matrices, therefore the transfer multiplicity of a generalized Aztec diamond is two. Based on the discussion above, the 4-8-8 lattice is of transfer multiplicity three. The triangular Kagomé lattice has multiplicity four since quadruple-step transfer matrices need to be involved in computing its Merrifield-Simmons index. The entropy constant of a two dimensional lattice with free boundary condition is defined by

$$\eta = \lim_{m,n\to\infty} f(m,n)^{1/k(m,n)},$$

where $f(m,n)$ is the number of independent sets of $P_{m,n}$ and $k(m,n)$ denotes the number of vertices of $P_{m,n}$. Before we show that the entropy constants for a lattice with these three different boundary conditions are the same, let us establish a relation between the transfer matrix of $P_{m,n}$ ($C_{n,m}$) and the number of independent sets of $C_{n,m}$ ($T_{n,m}$) by the following lemma.

Theorem 7.2.1. *Let $P_{m,n}$, who has the same real symmetric transfer matrix in both vertical and horizontal directions and denoted by T_m, be a two dimensional lattice with free boundary condition for given positive integers m and n. Let $C_{m,n}$, whose transfer matrix is denoted by B_n, be the corresponding lattice with cylindrical boundary condition and $T_{m,n}$ be the corresponding lattice with toroidal boundary condition. Then the trace of T_m^n is equal to the number of independent sets of $C_{m,n}$ and the trace of B_n^m is equal to the number of independent sets of $T_{m,n}$.*

Proof. Recall that $C_{m,n}$ can be obtained by identifying the leftmost and the rightmost columns of $P_{m,n}$. Thus there is a bijection between the independent sets of $C_{m,n}$ and the independent sets of $P_{m,n}$ whose leftmost and rightmost column vectors are the same. And the latter is the trace of T_m^n.

Similarly, $T_{m,n}$ can be obtained by identifying the top row and the bottom row of $C_{m,n}$. Thus there is a bijection between the independent sets of $T_{m,n}$ and the independent sets of $C_{m,n}$ whose corresponding top and bottom row vectors are the same. And the latter is the trace of B_n^m. □

Neil J Calkin [1998] proved the existence of the entropy constant of the square lattice and established its upper and lower bounds. Their approach is valid for the lattices with the same symmetric transfer matrices in both horizontal and vertical directions. Thus, the bounds can be generalized and stated as follows.

Theorem 7.2.2. *For both T_m and B_n, the entropy constant of the Merrifield-Simmons index is lower bounded by $(\frac{\lambda_{p+2q}}{\lambda_{2q}})^{1/p}$, where λ's are the largest eigenvalues of corresponding T's. And the upper bound of the entropy constant is $(\xi_{2k})^{1/2k}$, where ξ's are the largest eigenvalues of corresponding B's.*

7.3 Numerical Results

Note that the 4-8-8 lattice has the same symmetric transfer matrix in both horizontal and vertical directions. By Theorem7.2.1 and Theorem 7.2.2, we can derive the lower and upper bound of entropy constant for the 4-8-8 lattice.

Let $p = 2$, $q = 3$ and $k = 4$ in Theorem 7.2.2, since the largest eigenvalues of T_6, T_8 and B_8 are given by $105606.367915106937\ldots$, $3510407.307349548675\ldots$ and $1220870.544468111359\ldots$ respectively, we have

$$5.765456527051\ldots \leq \lim_{m,n\to\infty} f(m,n)^{1/mn} \leq 5.765456529341\ldots,$$

where $f(m,n)$ is the Merrifield-Simmons index of the 4-8-8 lattice $G_{m,n}$. Since the number of vertices of $G_{m,n}$ is $4mn + 2m + 2n$, we can see that the entropy constant of the 4-8-8 lattice is

$$\eta = \lim_{m,n\to\infty} f(m,n)^{1/4mn+2m+2n},$$

which is between $1.549560101247\ldots$ and $1.549560101400\ldots$.

Some numerical results of Merrifield-Simmons index of the 4-8-8 lattice with various boundary conditions are presented in the following tables.

Table 7.1: The numerical results of Merrifield-Simmons index of the 4-8-8 lattice $G_{m,n}$ with free boundary condition

n	$f(1,n)$	$f(2,n)$	$f(3,n)$	$f(4,n)$
1	47	779	12887	213203
2	779	74753	7144259	682899449
3	12887	7144259	3935795807	2168825644779
4	213203	682899449	2168825644779	6890559309027361
5	3527231	65275954619	1195119855466343	21891571255589181023
6	58354523	6239500931057	658564832936254227	69550417915510873198592
7	965417447	596412129451379	362898854276036750752	220964515515328798010638336
8	15971869859	57008955129453545	199973597225797700796416	702013283095176099770505625600
9	264238674767	5449287169347507444	110194449810190513907695616	
n	$f(5,n)$	$f(6,n)$	$f(7,n)$	$f(8,n)$
1	3527231	58354523	965417447	15971869859
2	65275954619	6239500931057	596412129451379	57008955129453545
3	1195119855466343	658564832936254227	362898854276036750752	199973597225797700796416
4	21891571255589181059	69550417915510873198592	220964515515328798010638336	702013283095176099770505625600
5	400989410760263959456896	734495814243670259853269 4016		
6	7344958142436702598532694016			
n	$f(9,n)$			
1	264238674767			
2	5449287169347507444			
3	110194449810190513907695616			

Table 7.2: The numerical results of Merrifield-Simmons index of the 4-8-8 lattice $H_{m,n}$ with cylindrical boundary condition

m	$g(m,1)$	$g(m,2)$	$g(m,3)$	$g(m,4)$
1	15	275	4527	74915
2	81	9223	872817	83484359
3	435	308007	167161419	92210259455
4	2337	10288055	32024704905	101890687050895
5	12555	343638751	6135187215867	112585183506906592
6	67449	11478130575	1175359722363369	124402293749613166592
7	362355	383389471367	225171683166443808	137459739741954729050112
8	1946673	12805873388263	43137676077348913152	151887714595812392365457408
9	10458075	427738384795215	8264179008000619446272	167830070726599585943198367744
m	$g(m,5)$	$g(m,6)$	$g(m,7)$	$g(m,8)$
1	1239375	20504243	339222255	5612094275
2	7979610641	762745733191	72908185459505	6969039275744391
3	50808099454115	27997712914113532	15427992474894299136	8501518342886229278720
4	323676087966247552	1028340702944578895872	3267078212691260344369152	1037964136440012759351322214
5	2061947765452025364480	37769283626796604821864448	691822641713709142974429921280	
6	13135457408094420540063744	1387206347469407026513191108608		
7	83678274939057768517536317440			

Table 7.3: The numerical results of Merrifield-Simmons index of the 4-8-8 lattice $S_{m,n}$ with toroidal boundary condition

m	$s(m,1)$	$s(m,2)$	$s(m,3)$	$s(m,4)$
1	5	29	155	833
2	29	1127	37235	1244819
3	155	37235	7023713	1347121127
4	833	1244819	1347121127	1490864952223
5	4475	41575979	258054952385	1647144129284543
6	24041	1388718947	49437673996703	1820049164089801216
7	129155	46385594795	9471107791983884	2011082476723567853568
8	693857	1549359408899	1814444828776562944	2222168789066396131655680
9	3727595	51751291603979	347605577830867337216	2455410851801241684155564032
m	s(m,5)	s(m,6)	s(m,7)	s(m,8)
1	4475	24041	129155	693857
2	41575979	1388718947	46385594795	1549359408899
3	258054952385	49437673996703	9471107791983884	1814444828776548608
4	1647144129284543	1820049164089801216	2011082476723567853568	2222168789066506995499008
5	10491381016588933120	66835240986806694445056	425767589675573485997391872	2712315179113658136571668332544
6	66835240986806694445056	2454789445716724048587653120	90160437291592209800637701947392	
7	425767589675573485997391872	90160437291592209800637701947392		
8	2712315179113658136571668332544			

This transfer matrix method for computing Merrifield-Simmons index is also valid for non-planar lattices. For instance, the square lattice with crossed diagonal bonds (obtained by adding two crossed diagonals to each square inner face) is non-planar; its Merrifield-Simmons index can be computed by using multi-step transfer matrix and it can be showed that the entropy constant is between $1.342542258\ldots$ and $1.342652572\ldots$. Unlike the Merrifield-Simmons index, dimer entropy is affected by the boundary conditions, see [Kasteleyn, 1963, 1987, Yan et al., 2008].

Chapter 8

Independent Sets on the Randomly Triangulated Grid Graphs

Let us continue our discussion on the statistics of independent sets in this chapter, on a random graph model. Let $G_{m,n}$ be a grid graph of square cells in m rows and n columns. Each square cell of the graph can be independently triangulated by adding a diagonal edge in two ways. Let $G_{m,n}(p)$ denote the random graph obtained by admitting identical independent Bernoulli distribution on these two types of triangulation for each square cell. In this chapter we consider the enumeration of independent sets on $G_{m,n}(p)$. By using the transfer matrix approach we obtained the annealed of independent sets per site on $G_{m,n}(p)$, i.e. $\eta_m(p) = \lim_{n\to\infty} \ln f(m,n;p)/mn$, where $f(m,n;p)$ denote the expected number of independent sets on $G_{m,n}(p)$. Subsequently, we found very strong correlation between this annealed entropy per site and the Shannon entropy of the corresponding Bernoulli distribution. This problem was suggested to the author by Brendan McKay during the 36th Australasian Conference on Combinatorial Mathematics and Combinatorial Computing.

8.1 Background and Related Work

In recent years, the enumeration problem of certain types of subgraphs has been discussed on random graphs. Zdeborová and Mézard [2006] obtained analytical results of the annealed entropy and quenched entropy of dimer problems in regular and Erdős-Rényi random graphs by means of the cavity method. Krivelevich et al. [2003] studied the asymptotic value of independence numbers of Erdős-Rényi random graphs. Greenhill et al. [2013] presented an asymptotic formula for the expectation of spanning trees on uniformly random regular graphs and the asymptotic distribution of the number of spanning trees on uniformly cubic graphs were also provided. Based on the structure of Graphene, a random planar hexagonal lattice model whose samples are from real world were considered in [Ren et al., 2012]. Its dimer problem and dimer-monomer problem were studied in [Ren et al., 2012] and [Ren et al., 2014] respectively. Inspired by the prior work, in this Chapter, we extend our discussion of lattice gas model to the random triangulated square lattice.

8.2 The Random Triangulation

Let $G_{m,n}$ (where m and n always denote positive integers) denote a grid graph whose square cells are arranged in m rows and n columns. The number of vertices in $G_{m,n}$ is $(m+1)(n+1)$. Each square cell of the graph $G_{m,n}$ can be triangulated by adding a diagonal edge in one of the two directions. Assume that the probabilities of choosing these two directions are p (north-east direction) and $1-p$ (north-west direction) respectively as shown in Figure 8.1. Let $G_{m,n}(p)$ denote the random graph obtained by admitting identically independent $\mathcal{Bernoulli}(p)$ on these two types of triangulation for each square cell. Let us illustrate this type of random graph by examples, when $m = 2$ or 3, in Figure 8.2 and Figure 8.3 respectively.

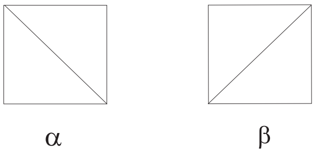

Fig. 8.1: α-type with probability p and β-type with probability $1-p$.

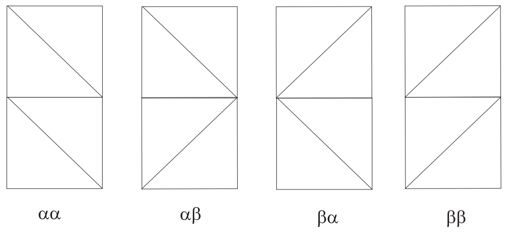

Fig. 8.2: Sample space of random graph $G_{2,1}(p)$

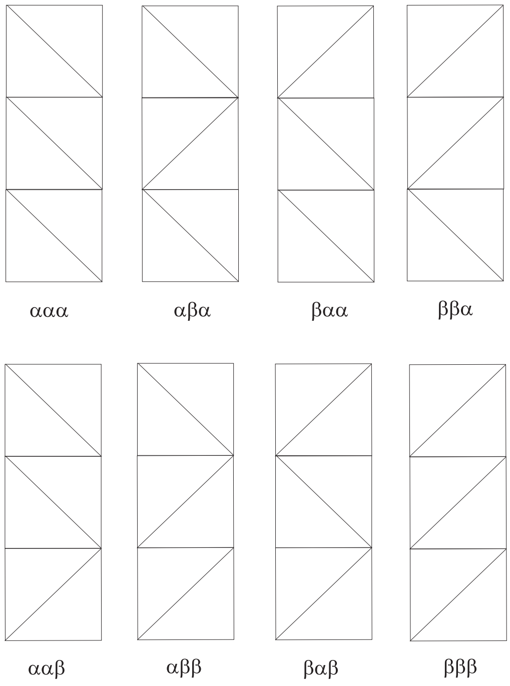

Fig. 8.3: Sample space of random graph $G_{3,1}(p)$

Similar to Chapter 7, we denote the collection of binary $(m+1)$-vectors containing no consecutive 1's as C_m that represents all possible independent sets in a column of $G_{m,n}$. The cardinality of C_m is F_{m+2} (the $(m+2)$-th Fibonacci number). For any one of the 2^m samples in the sample space of $G_{m,1}(p)$, we can represent the independent sets of the sample graph by a transfer matrix $T_m^{(i)}$ of order F_{m+2} $(i = 1, \dots, 2^m)$, whose rows and columns are indexed by vectors in C_m. More precisely, for any $v, v' \in C_m$ the element of $T_m^{(i)}$ in position (v, v') is 1 if the vertex sets corresponding to v, v' consist an independent set of $G_{m,1}$ and is 0 otherwise. Taking the expectation of all transfer matrices of the $G_{m,n}(p)$, we obtain the expected transfer matrix $T_m(p)$ as a weighted sum of transfer matrices for all 2^{mn} samples. The graphs in the sample space of $G_{2,1}(0.3)$ and their corresponding transfer matrices $T_m^{(i)}, i = 1, \ldots, 2^{mn}$ are given below. The expectation of the transfer matrices for all samples is given by $T_2(0.3)$.

$$T_2^{(1)} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}, T_2^{(2)} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \end{pmatrix}$$

$$T_2^{(3)} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}, T_2^{(4)} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$T_2(0.3) = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0.3 & 1 & 0 \\ 1 & 0.7 & 0 & 0.3 & 0.21 \\ 1 & 1 & 0.7 & 0 & 0 \\ 1 & 0 & 0.21 & 0 & 0 \end{pmatrix}.$$

Note that a sample of $G_{m,n}(p)$ can be constructed by gluing n samples (with replacement) of $G_{m,1}(p)$ successively. We show that the expected number of independent sets on $G_{m,n}(p)$ can be computed by the sum of entries of $T_m^n(p)$.

8.3 Main Results

In this section, we show that the expected number of independent sets on $G_{m,n}(p)$ can be computed by the sum of entries of $T_m^n(p)$.

Theorem 8.3.1. *The expectation of the number of independent sets on $G_{m,n}(p)$ is given by*

$$f(m, n; p) = \mathbf{1}T_m^n(p)\mathbf{1}^T.$$

Proof. Let $p(T_m^{(i)})$ denote the probability of choosing $G_{m,1}^{(i)}$, $i = 1, \ldots, 2^m$ from the sample space of $G_{m,1}(p)$ and $f_i(m,n)$ denote the number of independent sets on $G_{m,n}^{(i)}$. Then we have

$$\begin{aligned}
f(m,n;p) &= \sum_{i_1 i_2 \cdots i_n} p(T_m^{(i_1)})p(T_m^{(i_2)}) \cdots p(T_m^{(i_n)}) f_{i_1 i_2 \cdots i_n}(m,n) \\
&= \sum_{i_1 i_2 \cdots i_n} p(T_m^{(i_1)})p(T_m^{(i_2)}) \cdots p(T_m^{(i_n)}) \mathbf{1} T_m^{(i_1)} T_m^{(i_2)} \cdots T_m^{(i_n)} \mathbf{1}^T \\
&= \sum_{i_1 i_2 \cdots i_n} \mathbf{1} \left(p(T_m^{(i_1)})p(T_m^{(i_2)}) \cdots p(T_m^{(i_n)}) T_m^{(i_1)} T_m^{(i_2)} \cdots T_m^{(i_n)} \right) \mathbf{1}^T \\
&= \mathbf{1} \left(\sum_{i_1, i_2 \ldots, i_n} p(T_m^{(i_1)})p(T_m^{(i_2)}) \cdots p(T_m^{(i_n)}) T_m^{(i_1)} T_m^{(i_2)} \cdots T_m^{(i_n)} \right) \mathbf{1}^T \\
&= \mathbf{1} \left(p(T_m^{(1)}) T_m^{(1)} + p(T_m^{(2)}) T_m^{(2)} + \cdots + p(T_m^{(2^m)}) T_m^{(2^m)} \right)^n \mathbf{1}^T \\
&= \mathbf{1} T_m^n(p) \mathbf{1}^T,
\end{aligned}$$

where the choice of $i_1, i_2, \ldots, i_n \in \{1, 2, 3, \ldots, 2^m\}$ are repeatable. So the index set is of cardinality 2^{mn}. □

The *annealed entropy* of the number of independent set of $G_{m,n}(p)$ is defined as $\ln f(m,n;p)/mn$. Consider the limit of the annealed entropy of independent set per square when n approaches infinity, that is

$$\eta_m(p) = \lim_{n \to \infty} \ln f(m,n;p)/mn.$$

Let $\lambda_{m,1}(p)$ denote the largest eigenvalue (in modulus) of $T_m(p)$. In the following, we prove the relation between $\eta_m(p)$ and the largest eigenvalue (in modulus) of $T_m(p)$.

Theorem 8.3.2. *The limit of the annealed entropy of independent set per site of $G_{m,n}(p)$ as $n \to \infty$ is given by*

$$\eta_m(p) = \ln \lambda_{m,1}(p)/m.$$

Proof. It is easy to verify that $T_m(p)$ is an irreducible matrix, which implies that there is a positive (simple) eigenvalue of $T_m(p)$ that exceeds the modulus of all other eigenvalues (the existence of such an eigenvalue is guaranteed by the Perron-

Frobenius Theorem). Let k denote the Fibonacci number F_{m+2}, the characteristic polynomial of $T_m(p)$ can be written as:

$$x^k + c_1x^{k-1} + c_2x^{k-2} + \cdots + c_{k-1}x + c_k = 0$$

By the Hamilton-Caylay Theorem we have

$$T_m^k + c_1T_m^{k-1} + c_2T_m^{k-2} + \cdots + c_{k-1}T_m + c_kI = \mathbf{0},$$

where I and $\mathbf{0}$ denote the F_{m+2} by F_{m+2} identity matrix and matrix of all 0's respectively. Then we obtain the recurrence relation of the cardinality of independent sets in the form of

$$f(m, j; p) + c_1f(m, j-1; p) + \cdots + c_{k-1}f(m, j-k+1; p) + c_kf(m, j-k; p) = 0,$$

where $j = k+1, k+2, \ldots$. By the theory of linear difference equations with constant coefficients, the homogeneous solution to the above equation is a linear combination in the form of $C_i\lambda_{m,i}(p)^m$, where C_i is the constant coefficient determined by the initial conditions, i.e. the initial values of $f(m, 1; p)$, $f(m, 2; p)$, ..., $f(m, k; p)$. Let $\lambda_{m,1}(p), \ldots, \lambda_{m,l}(p)$ denote the distinct eigenvalues of $T_m(p)$ with algebraic multiplicity $1, e_2, \ldots, e_l$ respectively. Thus,

$$f(m, n; p) = c_1\lambda_{m,1}^n(p) + \sum_{i=2}^{l}\sum_{j=1}^{e_i} a_{ij}m^{e_i-j}\lambda_{m,j}^m(p),$$

where a_{ij} are constant coefficients. Then we have

$$\eta_m(p) = \lim_{n\to\infty} \ln f(m, n; p)/mn = \ln \lambda_{m,1}(p)/m.$$

That completes the proof. □

8.4 Experiments

Our experiment also supports Theorem 8.3.2. We report some of the values of $\ln \lambda_{m,1}(p)/m$ and $\ln(f(m,n;p))/mn$ with $n = 100$ in Table 8.1 and Table 8.2 respectively.

m	$p = 0.3$	$p = 0.5$	$p = 0.9$
1	0.77932296726864	0.78213773426906	0.77085787963316
2	0.56617745580448	0.56777543927018	0.56156482815422
3	0.48903371824254	0.49063344553942	0.48436459137243
4	0.45193825107408	0.45340621564412	0.44769945558867
5	0.42929548846124	0.43072758929728	0.42515832764671
6	0.41430479539063	0.41569901252843	0.41028859427341
7	0.40356799600337	0.40493967122235	0.39962012472550
8	0.39552370200930	0.39687701253179	0.39163320947690
9	0.38926462553250	0.39060412280449	0.38541668087957
10	0.38425806847677	0.38558636349262	0.38044486767815
11	0.38016158614475	0.38148076455478	0.37637657279850

Table 8.1: The eigenvalue approximation

m	$p = 0.3$	$p = 0.5$	$p = 0.9$
1	0.78964843564591	0.79245384580572	0.78121022063039
2	0.57368686057915	0.57529460853630	0.56904352987573
3	0.49552251420460	0.49712828026730	0.49083451592644
4	0.45793589169533	0.45941235502240	0.45367054368683
5	0.43499371732890	0.43643409238543	0.43083081447020
6	0.41980460313994	0.42120757999449	0.41576114593824
7	0.40892577348357	0.41030632482455	0.40495032385232
8	0.40077503278234	0.40213738856693	0.39685644356716
9	0.39443314552337	0.39578179095878	0.39055680108048
10	0.38936034439968	0.39069787982204	0.38551846531132
11	0.38520966133618	0.38653815221735	0.38139575476258

Table 8.2: The annealed entropy approximation

Let us denote the Shannon entropy of Bernoulli distribution $\mathcal{B}ernoulli(p)$ as $H(p)$. Intuitively, $H(p)$ of the Bernoulli distribution that $G_{m,n}(p)$ follows should describe the uncertainty of our model. The numerical results suggest that the annealed entropy $\eta_m(p)$ achieve the maximum when $p = 1/2$ as $H(p)$ does. We set $n = 100$ and illustrate the relation between $H(p)$ and $\eta_m(p)$ for different values of m and p in Figure 8.4. For each value of m, we plot $\eta_m(p)$ against $H(p)$ in the same color and calculate the linear regression line with least squares. Our simulation suggests that all these linear regressions have small residual standard errors and each regression is performing better than "random noise" as a predictor, based on the output F-statistic. Therefore, we conclude that there is a strong correlation between $\eta_m(p)$ and $H(p)$.

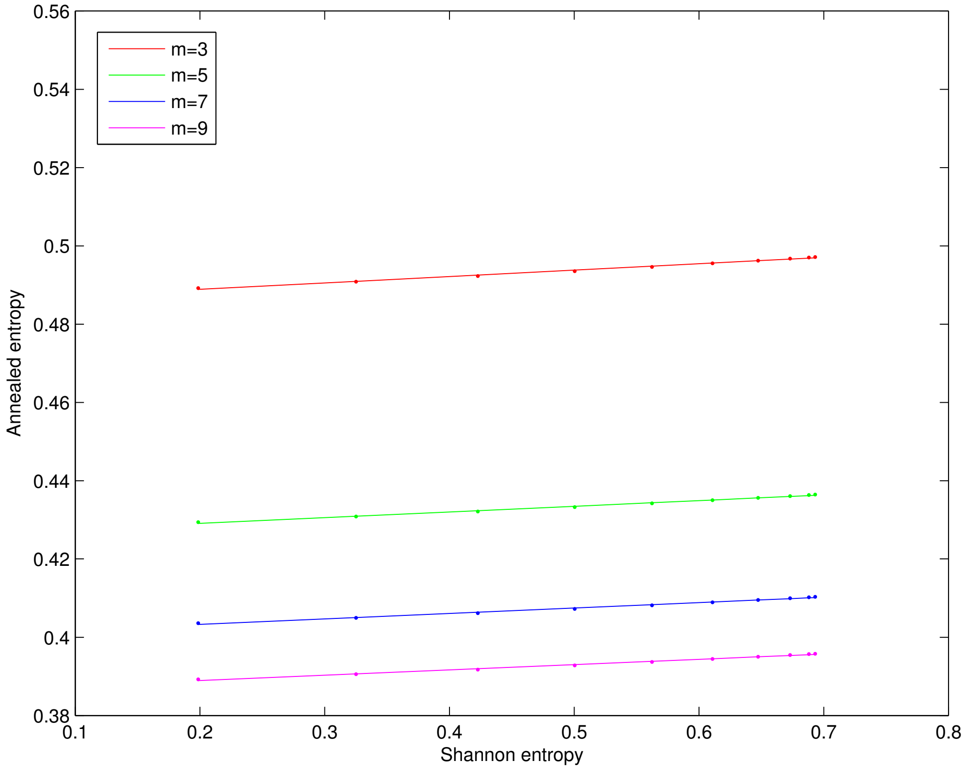

Fig. 8.4: Relation between the annealed entropy and the Shannon entropy

Chapter 9

Future Research Directions

We conclude this thesis by listing a few research directions that deserve further investigation.

In Chapter 3, we indicate that Bayesian inference can be differentially-private provided the prior is set appropriately, but the effects on learning have not been fully examined. Even though larger c can improve privacy, the high concentration of prior would result in the inhibition on learning. Therefore, the trade-off between privacy and utility could be optimized by the choice of c, through which the number of samples can be controlled.

In Chapter 4, we investigate the link between posterior sampling and the exponential mechanism. By setting the utility function as the log-likelihood and base measure as the prior, the density function in the exponential mechanism has similar form to the posterior distribution. We use this to release approximate MAP point estimates. Wang et al. [2015] noted this connection independently. In our framework, privacy is achieved by setting ϵ to a sufficiently small value; it is noteworthy that this is also how robustness results for altered Bayesian inference is obtained by Grünwald [2012]. That connection suggests we can gain both privacy and efficiency in some cases. We have proven that privacy is achievable by modifying the prior which is consistent with the base measure in the exponential mechanism. Therefore, we believe that the examination on settings where both ϵ and the prior measure may be adjusted could be fruitful.

Graph structure is known to be deeply linked with computational complexity of exact Bayesian inference, and mixing times for sampling-based inference. Chapter 4 demonstrates how graph structure (i.e. conditional independence structure) also impacts on differential privacy. Are there further connections between graph structure and differential privacy in other mechanisms developed here and elsewhere?

Beyond sampling-based methods, it is natural to consider differential privacy in exact inference. There are algorithms such as message passing and junction tree which perform exact inference on graphical models. Variational method such as mean field approximation can be easily used to preserve privacy since approximation with a different model might provide privacy guarantees for free. Finally, could existing frameworks for probabilistic programming such as *Stan* [Stan Development Team, 2015] be adapted to provide differential privacy in a usable manner? A solution to this question would be of broad interest.

In Chapter 7, we develop an approach to count independent sets on the 4-8-8 lattice by modifying the classical transfer matrix method. However, Yao-ban Chan reminded the author that there is a more powerful approach: the corner transformation method can be applied to this kind of problem and it usually provides better approximations. Also, the 4-8-8 lattice has larger entropy constant of the number of independent sets comparing to that of the square lattice [Neil J Calkin, 1998] while the square lattice has larger entropy than that of the square lattice with crossed diagonal bonds as shown in Chapter 7. Note that the vertices of the 4-8-8 lattice are of degree three while the vertices of the square lattice and square lattice with crossed diagonal bonds are of degree four and eight respectively. Is the entropy of the number of independent sets negatively correlated to the average degree of a lattice? Intuitively, the answer is likely to be a yes but further investigation needs to be made.

On the random triangulated grid graph $G_{m,n}(p)$, our experiments suggest that there is a strong correlation between the annealed entropy per site of the independent sets and the Shannon entropy of underlying Bernoulli distribution when n approaches infinity. However, to derive an analytical relation between these two entropies would be a challenge. The double limit of this annealed entropy as m and n approach infinite at the same rate represents a more difficult task.

Bibliography

Francesco Aldà and Benjamin I. P. Rubinstein. The Bernstein mechanism: Function release under differential privacy. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI'2017)*, 2017.

GR Allen. Dimer models for the antiferroelectric transition in copper formate tetrahydrate. *The Journal of Chemical Physics*, 60(8):3299–3309, 1974.

Mário S Alvim, Miguel E Andrés, Konstantinos Chatzikokolakis, Pierpaolo Degano, and Catusciam Palamidessi. Differential privacy: on the trade-off between utility and information leakage. In *International Workshop on Formal Aspects in Security and Trust*, pages 39–54. Springer, 2011a.

Mário S Alvim, Miguel E Andrés, Konstantinos Chatzikokolakis, and Catuscia Palamidessi. On the relation between differential privacy and quantitative information flow. In *International Colloquium on Automata, Languages, and Programming*, pages 60–76. Springer, 2011b.

Miguel E Andrés, Nicolás E Bordenabe, Konstantinos Chatzikokolakis, and Catuscia Palamidessi. Geo-indistinguishability: Differential privacy for location-based systems. In *Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security*, pages 901–914. ACM, 2013.

Boaz Barak, Kamalika Chaudhuri, Cynthia Dwork, Satyen Kale, Frank McSherry, and Kuna Talwar. Privacy, accuracy, and consistency too: A holistic solution to contingency table release. In *Proceedings of the Twenty-sixth ACM Symposium on Principles of Database Systems*, PODS '07, pages 273–282. ACM, 2007.

Gilles Barthe, Marco Gaboardi, Justin Hsu, and Benjamin Pierce. Programming language techniques for differential privacy. *ACM SIGLOG News*, 3(1):34–53, 2016.

Raef Bassily, Adam Groce, Jonathan Katz, and Adam Smith. Coupled-worlds privacy: Exploiting adversarial uncertainty in statistical data privacy. In *Foundations of Computer Science (FOCS), 2013 IEEE 54th Annual Symposium on*, pages 439–448. IEEE, 2013.

R J Baxter. *Exactly Solved Models in Statistical Mechanics.* Academic Press, London, 1982.

R J Baxter. Planar lattice gas with nearest-neighbor exclusion. *Annals of Combinatorics*, 3(2-4):191–203, 1999.

R J Baxter, I G Enting, and S K Tsang. Hard-square lattice gas. *Journal of Statistical Physics*, 22(4):465–489, 1980.

James O Berger. *Statistical Decision Theory and Bayesian Analysis.* Springer-Verlag, 1985.

Peter J Bickel and Kjell A Doksum. *Mathematical Statistics: Basic Ideas and Selected Topics*, volume 1. Holden-Day Company, 2001.

Jonathan M Borwein, ML Glasser, and RC McPhedran. *Lattice Sums Then and Now.* Number 150. Cambridge University Press, 2013.

Olivier Bousquet and André Elisseeff. Stability and generalization. *Journal of Machine Learning Research*, 2(Mar):499–526, 2002.

Christelle Braun, Konstantinos Chatzikokolakis, and Catuscia Palamidessi. Quantitative notions of leakage for one-try attacks. *Electronic Notes in Theoretical Computer Science*, 249:75–91, 2009.

Arthur Cayley. A theorem on trees. *Quart. J. Math.*, (8), 1889.

Konstantinos Chatzikokolakis, Miguel E Andres, Nicolas Emilio Bordenabe, and Catuscia Palamidessi. Broadening the scope of differential privacy using metrics. In *Privacy Enhancing Technologies*, pages 82–102, 2013.

Kamalika Chaudhuri and Daniel Hsu. Convergence rates for differentially private statistical estimation. In *Proceedings of the International Conference on Machine Learning. International Conference on Machine Learning*, volume 2012, page 1327. NIH Public Access, 2012.

Kamalika Chaudhuri and Claire Monteleoni. Privacy-preserving logistic regression. In *Advances in Neural Information Processing Systems 21*, NIPS, pages 289–296, 2008.

Kamalika Chaudhuri, Claire Monteleoni, and Anand D Sarwate. Differentially private empirical risk minimization. *Journal of Machine Learning Research*, 12 (Mar):1069–1109, 2011.

Kamalika Chaudhuri, Anand Sarwate, and Kaushik Sinha. Near-optimal differentially private principal components. In F Pereira, C J C Burges, L Bottou, and K Q Weinberger, editors, *Advances in Neural Information Processing Systems 25*, pages 989–997. 2012.

Henry Cohn, Noam Elkies, and James Propp. Local statistics for random domino tilings of the Aztec diamond. *Duke Math. J.*, 85(1):117–166, 10 1996.

Charles Alfred Coulson, Brian O'Leary, and Roger B Mallion. *Hückel Theory for Organic Chemists.* Academic Press, 1978.

Sven J Cyvin and Ivan Gutman. *Kekulé Structures in Benzenoid Hydrocarbons*, volume 46. Springer Science & Business Media, 2013.

Giuliana Davidoff, Peter Sarnak, and Alain Valette. *Elementary Number Theory, Group Theory and Ramanujan Graphs.* Cambridge University Press, 2003.

Morris H DeGroot. *Optimal Statistical Decisions.* John Wiley & Sons, 1970.

Christos Dimitrakakis, Blaine Nelson, Aikaterini Mitrokotsa, and Benjamin I P Rubinstein. Robust and private Bayesian inference. In *International Conference on Algorithmic Learning Theory*, pages 291–305. Springer, 2014.

Christos Dimitrakakis, Blaine Nelson, Zuhe Zhang, Aikaterini Mitrokotsa, and Benjamin I P Rubinstein. Differential privacy for Bayesian inference through posterior sampling. *Journal of Machine Learning Research*, (18), 2017. to appear; arXive:1306.1066.

Yevgeniy Dodis, Leonid Reyzin, and Adam Smith. Fuzzy extractors: How to generate strong keys from biometrics and other noisy data. In *International Conference on the Theory and Applications of Cryptographic Techniques*, pages 523–540. Springer, 2004.

Cyril Domb and Melville S Green, editors. *Phase Transitions and Critical Phenomena*. Academic Press, London, 1972.

Fengming Dong, Weigen Yan, and Fuji Zhang. On the number of perfect matchings of line graphs. *Discrete Applied Mathematics*, 161(6):794–801, 2013.

John C Duchi, Michael I Jordan, and Martin J Wainwright. Local privacy and statistical minimax rates. Technical Report 1302.3203, arXiv, 2013.

Cynthia Dwork. Differential privacy. In *ICALP*, pages 1–12, 2006.

Cynthia Dwork and Jing Lei. Differential privacy and robust statistics. In *Proceedings of the forty-first annual ACM symposium on Theory of computing*, pages 371–380. ACM, 2009.

Cynthia Dwork and Aaron Roth. The algorithmic foundations of differential privacy. *Foundations and Trends® in Theoretical Computer Science*, 9(3-4):211–407, 2014.

Cynthia Dwork and Adam Smith. Differential privacy for statistics: what we know and what we want to learn. *Journal of Privacy and Confidentiality*, 1(2):135–154, 2009.

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

Cynthia Dwork, Vitaly Feldman, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Aaron Roth. Preserving statistical validity in adaptive data analysis. Technical Report 1411.2664, arXiv, 2014.

Cynthia Dwork, Vitaly Feldman, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Aaron Roth. The reusable holdout: Preserving validity in adaptive data analysis. *Science*, 349(6248):636–638, 2015.

Konrad Engel. On the Fibonacci number of an M x N lattice. *Fibonacci Quarterly*, 28(1):72–78, 1990.

Úlfar Erlingsson, Vasyl Pihur, and Aleksandra Korolova. Rappor: Randomized aggregatable privacy-preserving ordinal response. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*, pages 1054–1067. ACM, 2014.

Alexei A Fedotov, Peter Harremoës, and Flemming Topsoe. Refinements of Pinsker's inequality. *IEEE Transactions on Information Theory*, 49(6):1491–1498, 2003.

Steven R Finch. Several constants arising in statistical mechanics. *Annals of Combinatorics*, 3(2-4):323–335, 1999.

Michael E Fisher. Statistical mechenics of dimers on a plane lattice. *Physical Review*, 124(6):1664, 1961.

James Foulds, Joseph Geumlek, Max Welling, and Kamalika Chaudhuri. On the theory and practice of privacy-preserving Bayesian data analysis. *arXiv preprint arXiv:1603.07294*, 2016.

R H Fowler and G S Rushbrooke. An attempt to extend the statistical theory of perfect solutions. *Transactions of the Faraday Society*, 33:1272–1294, 1937.

Benjamin Fung, Ke Wang, Rui Chen, and Philip S Yu. Privacy-preserving data publishing: A survey of recent developments. *ACM Computing Surveys (CSUR)*, 42(4):14, 2010.

Marco Gaboardi, Andreas Haeberlen, Justin Hsu, Arjun Narayan, and Benjamin C Pierce. Linear dependent types for differential privacy. *ACM SIGPLAN Notices*, 48(1):357–370, 2013.

Giovanni Gallavotti. *Statistical Mechanics: A Short Treatise*. Springer Science & Business Media, 2013.

Catherine Greenhill, Matthew Kwan, and David Wind. On the number of spanning trees in random regular graphs. *arXiv preprint arXiv:1309.6710*, 2013.

Peter Grünwald. The safe Bayesian: Learning the learning rate via the mixability gap. In *Proceedings of the 23rd International Conference on Algorithmic Learning Theory*, ALT, pages 169–183, 2012.

Peter D Grünwald and A P Dawid. Game Theory, maximum Entropy, minimum Discrepancy, and robust Bayesian decision theory. *The Annals of Statistics*, 32 (4):1367–1433, 2004.

ZZ Guo, Kwok Y Szeto, and Xiujun Fu. Damage spreading on two-dimensional trivalent structures with Glauber dynamics: Hierarchical and random lattices. *Physical Review E*, 70(1):016105, 2004.

Ivan Gutman. The energy of a graph. *Berichte aus der Mathematic-Statist. Sekt. Forschungszentrum Graz*, pages 1–22, 1978.

Ivan Gutman. Extremal hexagonal chains. *Journal of Mathematical Chemistry*, 12 (1):197–210, 1993.

Ivan Gutman. The energy of a graph: Old and new results. *Algebraic Combinatorics and Applications*, pages 196–211, 2001.

Ivan Gutman and Bojan Mohar. The quasi-Wiener and the Kirchhoff indices coincide. *Journal of Chemical Information and Computer Sciences*, 36(5):982–985, 1996.

Ivan Gutman and Oskar E Polansky. *Mathematical Concepts in Organic Chemistry*. Springer Science & Business Media, 2012.

Andreas Haeberlen, Benjamin C Pierce, and Arjun Narayan. Differential privacy under fire. In *USENIX Security Symposium*, 2011.

A Haji-Ankabari and R M Ziff. Bond percolation on the triangular kagomé lattice. *Physical Review B*, 2009.

Rob Hall, Alessandro Rinaldo, and Larry Wasserman. Random differential privacy. *Journal of Privacy and Confidentiality*, 4(2), 2011.

Rob Hall, Alessandro Rinaldo, and Larry Wasserman. Differential privacy for functions and functional data. *Journal of Machine Learning Research*, 14(Feb):703–727, 2013.

Frank R Hampel, Elvezio M Ronchetti, Peter J Rousseeuw, and Werner A Stahel. *Robust Statistics: The Approach Based on Influence Functions*. John Wiley and Sons, 1986.

Moritz Hardt and Kunal Talwar. On the geometry of differential privacy. In *Proceedings of the forty-second ACM symposium on Theory of computing*, pages 705–714. ACM, 2010.

Xi He, Ashwin Machanavajjhala, and Bolin 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. ACM, 2014.

Shlomo Hoory, Nathan Linial, and Avi Wigderson. Expander graphs and their applications. *Bulletin of the American Mathematical Society*, 43(4):439–561, 2006.

Haruo Hosoya. A newly proposed quantity characterizing the topological nature of structural isomers of saturated hydrocarbons. *Bulletin of the Chemical Society of Japan*, 44(9):2332–2339, 1971.

Peter J Huber. Robust estimation of a location parameter. *The Annals of Mathematical Statistics*, 35(1):73–101, 1964.

Peter J Huber. *Robust Statistics.* John Wiley and Sons, 1981.

G Jagannathan, K Pillaipakkamnatt, and R N Wright. A practical differentially private random decision tree classifier. In *IEEE International Conference on Data Mining Workshops*, pages 114–121, dec 2009.

Anthony D Joseph, Pavel Laskov, Fabio Roli, J Doug Tygar, and Blaine Nelson. Machine learning methods for computer security (Dagstuhl Perspectives Workshop 12371). *Dagstuhl Manifestos*, 3(1):1–30, 2013.

Peter Kairouz, Sewoong Oh, and Pramod Viswanath. The composition theorem for differential privacy. *arXiv preprint arXiv:1311.0776*, 2013.

Peter Kairouz, Sewoong Oh, and Pramod Viswanath. The composition theorem for differential privacy. In *Proceedings of The 32nd International Conference on Machine Learning*, ICML, pages 1376–1385, 2015.

Shiva Prasad Kasiviswanathan and Adam Smith. On theSemantics' of differential privacy: A Bayesian formulation. *arXiv preprint arXiv:0803.3946*, 2008.

P W Kasteleyn. *The Statistics of Dimers on a Lattice*, pages 281–298. Birkhäuser Boston, Boston, MA, 1987.

Pieter W Kasteleyn. Dimer statistics and phase transitions. *Journal of Mathematical Physics*, 4(2):287–293, 1963.

Auguste Kerckhoffs. La cryptographie militaire. *Journal des Sciences Militaires*, IX:5–83 January; 161–191, February, 1883.

Daniel Kifer and Ashwin Machanavajjhala. No free lunch in data privacy. In *Proceedings of the 2011 ACM SIGMOD International Conference on Management of data*, pages 193–204. ACM, 2011.

Daniel Kifer and Ashwin Machanavajjhala. A rigorous and customizable framework for privacy. In *Proceedings of the 31st ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems*, pages 77–88. ACM, 2012.

Gustav Kirchhoff. Über die Auflösung der Gleichungen, auf welche man bei der Untersuchungen der linearen Vertheilung galvanischer Ströme geführt wird. *Ann. Phys. Chem.*, 148:497–508, 1847.

Douglas J Klein and T G Schmalz. Exact enumeration of long-range-ordered dimer coverings on the square-planar lattice. *Physical Review B*, 41(4):2244, 1990.

Douglas J Klein, G E Hite, and T G Schmalz. Transfer-matrix method for subgraph enumeration: Applications to polypyrene fusenes. *Journal of Computational Chemistry*, 7(4):443–456, 1986.

Michael Krivelevich, Benny Sudakov, Van H Vu, and Nicholas C Wormald. On the probability of independent sets in random graphs. *Random Structures and Algorithms*, 22(1):1–14, 2003.

Lucien Le Cam. Convergence of estimates under dimensionality restrictions. *The Annals of Statistics*, pages 38–53, 1973.

Xueliang Li, Yongtang Shi, and Ivan Gutman. *Graph Energy*. Springer, New York, 2012.

M. Lichman. UCI machine learning repository, 2013. URL http://archive.ics.uci.edu/ml.

Xiaogang Liu and Zuhe Zhang. Spectra of subdivision-vertex join and subdivision-edge join of two graphs. *Bulletin of the Malaysian Mathematical Sciences Society*, pages 1–17, 2017.

Xiaogang Liu and Sanming Zhou. Spectra of the neighbourhood corona of two graphs. *Linear and Multilinear Algebra*, 62(9):1205–1219, 2014.

Xiaoyun Liu and Weigen Yan. The triangular kagomé lattices revisited. *Physica A: Statistical Mechanics and its Applications*, 392(22):5615–5621, 2013.

Yen Lee Loh, Dao-Xin Yao, and Erica W Carlson. Dimers on the triangular kagome lattice. *Physical Review B*, 78(22):224410, 2008.

Russell Lyons. Asymptotic enumeration of spanning trees. *Combinatorics, Probability and Computing*, 14(04):491–522, 2005.

Ashwin Machanavajjhala, Daniel Kifer, John Abowd, Johannes Gehrke, and Lars Vilhuber. Privacy: Theory meets practice on the map. In *IEEE 24th International Conference on Data Engineering, 2008*, pages 277–286. IEEE, 2008.

Sanchit Maruti and Leonard W ter Haar. Magnetic properties of the two-dimensional "triangles-in-triangles" kagomé lattice $\mathrm{cu_9x_2(cpa)_6}$ (x=f, ci, br). *Journal of Applied Physics*, 75(10):5949–5951, 1994.

Andrew McGregor, Ilya Mironov, Toniann Pitassi, Omer Reingold, Kunal Talwar, and Salil Vadhan. The limits of two-party differential privacy. In *Foundations of Computer Science (FOCS), 2010 51st Annual IEEE Symposium on*, pages 81–90. IEEE, 2010.

Frank McSherry and Ratul Mahajan. Differentially-private network trace analysis. *ACM SIGCOMM Computer Communication Review*, 40(4):123–134, 2010.

Frank McSherry and Kunal Talwar. Mechanism design via differential privacy. In *FOCS*, pages 94–103, 2007.

Frank D McSherry. Privacy integrated queries: an extensible platform for privacy-preserving data analysis. In *Proceedings of the 2009 ACM SIGMOD International Conference on Management of Data*, pages 19–30. ACM, 2009.

Richard E Merrifield and Howard E Simmons. *Topological Methods in Chemistry*. Wiley, New York, 1989.

Minnesota Population Center. Integrated public use microdata series - international: Version 5.0. 2009, 2009.

Darakhshan Mir. Differentially-private learning and information theory. In *Proceedings of the 2012 Joint EDBT/ICDT Workshops*, pages 206–210. ACM, 2012.

Prashanth Mohan, Abhradeep Thakurta, Elaine Shi, Dawn Song, and David Culler. GUPT: privacy preserving data analysis made easy. In *Proceedings of the 2012*

ACM SIGMOD International Conference on Management of Data, pages 349–360. ACM, 2012.

Arvind Narayanan and Vitaly Shmatikov. Robust de-anonymization of large sparse datasets. In *2008 IEEE Symposium on Security and Privacy (sp 2008)*, pages 111–125. IEEE, 2008.

Herbert S Wilf Neil J Calkin. The number of independent sets in a grid graph. *SIAM Journal on Discrete Mathematics*, 11(1):54–60, 1998.

Kobbi Nissim, Sofya Raskhodnikova, and Adam Smith. Smooth sensitivity and sampling in private data analysis. In *Proceedings of the Thirty-ninth Annual ACM Symposium on Theory of Computing*, pages 75–84. ACM, 2007.

Kobbi Nissim, Rann Smorodinsky, and Moshe Tennenholtz. Approximately optimal mechanism design via differential privacy. In *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference*, pages 203–213. ACM, 2012.

V I Norkin. Stochastic Lipschitz functions. *Cybernetics and Systems Analysis*, 22 (2):226–233, 1986.

Ian Osband, Daniel Russo, and Benjamin Van Roy. (More) efficient reinforcement learning via posterior sampling. In *Advances in Neural Information Processing Systems*, NIPS, pages 3003–3011, 2013.

Elena Pagnin, Christos Dimitrakakis, Aysajan Abidin, and Aikaterini Mitrokotsa. On the leakage of information in biometric authentication. In *Indocrypt 2014*, pages 265–280. Springer, 2014.

Catuscia Palamidessi and Marco Stronati. Differential privacy for relational algebra: improving the sensitivity bounds via constraint systems. In Herbert Wiklicky and Mieke Massink, editors, *QAPL - Tenth Workshop on Quantitative Aspects of Programming Languages*, volume 85, pages 92–105, 2012.

Linus Pauling. *The Nature of Chemical Bonds.* Cornell University Press, 1939.

Linus Pauling. Bond numbers of and bond lengths in tetrabenzo[de,no,st,c1,d1]heptacene and other condensed aromatic hydrocarbons: a valence-bond treatment. *Acta Crystallographica Section B: Structural Science*, 1980.

Paul A Pearce and Katherine A Seaton. A classical theory of hard squares. *Journal of Statistical Physics*, 53(5-6):1061–1072, 1988.

Jason Reed and Benjamin C Pierce. Distance makes the types grow stronger: a calculus for differential privacy. *ACM Sigplan Notices*, 45(9):157–168, 2010.

Haizhen Ren and Fuji Zhang. Double hexagonal chains with maximal Hosoya index and minimal Merrifield-Simmons index. *Journal of Mathematical Chemistry*, 42 (4):679–690, 2007.

Haizhen Ren, Fuji Zhang, and Jianguo Qian. Dimer coverings on random multiple chains of planar honeycomb lattices. *Journal of Statistical Mechanics: Theory and Experiment*, 2012(08):P08002, 2012.

Haizhen Ren, Fuji Zhang, and Jianguo Qian. Monomer-dimer problem on random planar honeycomb lattice. *Journal of Mathematical Physics*, 55(2):23304, 2014.

Indrajit Roy, Srinath TV Setty, Ann Kilzer, Vitaly Shmatikov, and Emmett Witchel. Airavat: Security and privacy for MapReduce. In *NSDI*, volume 10, pages 297–312, 2010.

Benjamin I P Rubinstein, Peter L Bartlett, Ling Huang, and Nina Taft. Learning in a large function space: privacy-preserving mechanisms for SVM learning. *Journal of Privacy and Confidentiality*, 4(1):4, 2012.

Russell Brandom. This is what Apple's differential privacy means for iOS 10, 2016.

Youcef Saad and Martin H Schultz. Topological properties of hypercubes. *IEEE Transactions on computers*, 37(7):867–872, 1988.

Horst Sachs and Holger Zernitz. Remark on the dimer problem. *Discrete Applied Mathematics*, 51(1):171–179, 1994.

S R Salinas and J F Nagle. Theory of the phase transition in the layered hydrogen-bonded $sncl_2 \cdot 2h_2o$ crystal. *Physical Review B*, 2000.

Christian R Scullard. Exact site percolation thresholds using a site-to-bond transformation and the star-triangle transformation. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 73(1):1–6, 2006.

Robert Shrock and Fa Y Wu. Spanning trees on graphs and lattices in d dimensions. *Journal of Physics A: Mathematical and General*, 33(21):3881, 2000.

Stan Development Team. Stan: A c++ library for probability and sampling, version 2.7.0, 2015. URL http://mc-stan.org/.

H. N. V. Temperley and Michael E Fisherpp. Dimer problem in statistical mechanics-an exact result. *Philosophical Magazine*, 6(68):1061–1063, 1961.

E Teufl and S Wagner. On the number of spanning trees on various lattices. *Journal of Physics A: Mathematical and Theoretical*, 43(41):415001, 2010.

William R Thompson. On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika*, 25(3-4):285–294, 1974.

Vladimir N Vapnik. *The Nature of Statistical Learning Theory.* Springer-Verlag, 1995.

Ulrike Von Luxburg, Agnes Radl, and Matthias Hein. Hitting and commute times in large graphs are often misleading. *arXiv preprint arXiv:1003.1266*, 2010.

Yu-Xiang Wang, Stephen Fienberg, and Alex Smola. Privacy for free: posterior sampling and stochastic gradient Monte Carlo. In David Blei and Francis Bach, editors, *Proceedings of the 32nd International Conference on Machine Learning (ICML-15)*, pages 2493–2502, 2015.

Larry Wasserman and Shuheng Zhou. A statistical framework for differential privacy. *Journal of the American Statistical Association*, 105(489):375–389, 2010.

Tsachy Weissman, Erik Ordentlich, Gadiel Seroussi, Sergio Verdu, and Marcelo J Weinberger. Inequalities for the l1 deviation of the empirical distribution. *Hewlett-Packard Labs, Tech. Rep*, 2003.

Oliver Williams and Frank McSherry. Probabilistic inference and differential privacy. In *NIPS'10*, pages 2451–2459, 2010.

Fa Y Wu. New critical frontiers for the Potts and percolation models. *Physical review letters*, 96(9):090602, 2006a.

Fa Y Wu. Dimers on two-dimensional lattices. *International Journal of Modern Physics B: Condensed Matter Physics, Statistical Physics, Applied Physics*, 20 (32):5357–5371, 2006b.

Yonghui Xiao and Li Xiong. Bayesian inference under differential privacy. (1203.0617), 2012.

Weigen Yan and Zuhe Zhang. Asymptotic energy of lattices. *Physica A: Statistical Mechanics and its Applications*, 388(8):1463–1471, 2009.

Weigen Yan, Yeong-Nan Yeh, and Fuji Zhang. Dimer problem on the cylinder and torus. *Physica A: Statistical Mechanics and its Applications*, 387(24):6069–6078, 2008.

Weigen Yan, Yeong Nan Yeh, and Fuji Zhang. The asymptotic behavior of some indices of iterated line graphs of regular graphs. *Discrete Applied Mathematics*, 160(7-8):1232–1239, 2012.

Bin Yu, Fano Assouad, and Lucien Le Cam. In *Festschrift for Lucien Le Cam*, pages 423–435. Springer, 1997.

Lenka Zdeborová and Marc Mézard. The number of matchings in random graphs. *Journal of Statistical Mechanics: Theory and Experiment*, 2006(05):P05003, 2006.

Yanqiu Zeng and Fuji Zhang. Extremal polyomino chains on k-matchings and k-independent sets. *Journal of Mathematical Chemistry*, 42(2):125–140, 2007.

Fuji Zhang, Yi-Chiuan Chen, and Zhibo Chen. Clique-inserted-graphs and spectral dynamics of clique-inserting. *Journal of Mathematical Analysis and Applications*, 349(1):211–225, 2009.

Jun Zhang, Zhenjie Zhang, Xiaokui Xiao, Yin Yang, and Marianne Winslett. Functional mechanism: regression analysis under differential privacy. *Proc. VLDB Endowment*, 5(11):1364–1375, 2012.

Jun Zhang, Graham Cormode, Cecilia M. Procopiuc, Divesh Srivastava, and Xiaokui Xiao. PrivBayes: private data release via Bayesian networks. *Proceedings of the 2014 ACM SIGMOD International Conference on Management of Data - SIGMOD '14*, (1):1423–1434, 2014.

Lian-Zhu Zhang and Feng Tian. Extremal catacondensed benzenoids. *Journal of Mathematical Chemistry*, 34(1-2):111–122, 2003.

Zuhe Zhang. Merrifield-Simmons index of generalized Aztec diamond and related graphs. *MATCH*, 2006.

Zuhe Zhang. Some physical and chemical indices of clique-inserted lattices. *Journal of Statistical Mechanics: Theory and Experiment*, 2013(10):P10004, 2013.

Zuhe Zhang. Merrifield-Simmons index and its entropy of the 4-8-8 lattice. *Journal of Statistical Physics*, 154(4):1113–1123, 2014.

Zuhe Zhang, Benjamin I P Rubinstein, and Christos Dimitrakakis. On the differential privacy of Bayesian inference. Technical Report https://hal.inria.fr/hal-01234215, HAL, 2015.

Zuhe Zhang, Benjamin I P Rubinstein, and Christos Dimitrakakis. On the differential privacy of Bayesian inference. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence*, AAAI, 2016.

Shijie Zheng. The Differential Privacy of Bayesian Inference. Bachelor's thesis, Harvard College, 2015.