PRIVIC: A privacy-preserving method for incremental collection of location data

Sayan Biswas, Catuscia Palamidessi

PRIVIC: A privacy-preserving method for incremental collection of location data

Sayan Biswas
sayan.biswas@inria.fr
INRIA and LIX, École Polytechnique
Palaiseau, France

Catuscia Palamidessi
catuscia@lix.polytechnique.fr
INRIA and LIX, École Polytechnique
Palaiseau, France

ABSTRACT

With recent advancements in technology, the threats of privacy violations of individuals' sensitive data are surging. Location data, in particular, have been shown to carry a substantial amount of sensitive information. A standard method to mitigate the privacy risks for location data consists in adding noise to the true values to achieve geo-indistinguishability. However, we argue that geo-indistinguishability alone is not sufficient to cover all privacy concerns. In particular, isolated locations are not sufficiently protected by the state-of-the-art Laplace mechanism (LAP) for geo-indistinguishability. In this paper, we focus on a mechanism that can be generated by using the Blahut-Arimoto algorithm (BA) from rate-distortion theory. We show that the BA mechanism, in addition to providing geo-indistinguishability, enforces an elastic metric that mitigates the problem of isolation. We then proceed to study the utility of BA in terms of the precision of the statistics that can be derived from the reported data, focusing on the inference of the original distribution. To this purpose, we apply the iterative Bayesian update (IBU), an instance of the famous expectation-maximization method from statistics, that produces the most likely distribution for any obfuscation mechanism. We show that BA harbours a better statistical utility than LAP for high levels of privacy, and becomes comparable as the level of privacy decreases. Remarkably, we point out a surprising connection, namely that BA and IBU, two apparently unrelated methods that were developed for completely different purposes, are dual to each other. Exploiting this duality and the privacy-preserving properties of BA, we propose an iterative method, PRIVIC, for a privacy-friendly incremental collection of location data from users by service providers. In addition to extending the privacy guarantees of geo-indistinguishability and retaining a better statistical utility than LAP, PRIVIC also provides an optimal trade-off between information leakage and quality of service. We illustrate the soundness and functionality of our method both analytically and with experiments.

CCS CONCEPTS

• **Security and privacy**; • **Mathematics of computing → Information theory**; **Probability and statistics**;

KEYWORDS

location-privacy, geo-indistinguishability, rate-distortion theory, privacy-utility trade-off

ACM Reference Format:
Sayan Biswas and Catuscia Palamidessi. 2023. PRIVIC: A privacy-preserving method for incremental collection of location data. .

, ,
2023.

1 INTRODUCTION

As the need and development of various kinds of research and analysis using personal data are becoming more and more significant, the risk of privacy violations of sensitive information of the data owners is also increasing manifold. One of the most successful proposals to address the issue of privacy protection is *differential privacy (DP)* [22, 23], a mathematical property that makes it difficult for an attacker to detect the presence of a record in a dataset. This is typically achieved by answering queries performed on the dataset in a (controlled) noisy fashion. Lately, the *local variant of differential privacy (LDP)* [21] has gained popularity due to the fact that the noise is applied at the data owner's end without needing a trusted curator. LDP is particularly suitable for situations where a data owner is a user who communicates her personal data in exchange for some service. One such scenario is the use of location-based services (LBS), where a user typically sends her location in order to obtain information like the shortest path to a destination, nearby points of interest, traffic information, etc. The security and the convenience of implementing the local model directly on a user's device (tablets, smartphones, etc.) make LDP very appealing.

Typically, in exchange of their service, providers incrementally collect their users' data, and then make them available to other parties which process them to provide useful statistics to companies and institutions. Obviously, the statistical precision of the collected data is essential for the quality of the analytics performed (*statistical utility*). However, injecting noise locally into the data to protect the privacy of the users usually has a negative effect on the statistical utility. Additionally, the noise degrades the *quality of service* (QoS) as well, since, obviously, the service results from the elaboration of the information received.

Substantial research has been done to address the privacy-utility trade-off in the context of DP. In LDP, the primary focus has been to optimize the utility from the data collector's perspective, i.e., devising mechanisms and post-processing methods that would allow deriving the most accurate statistics from the collection of the noisy data [21, 53]. In contrast, in domains such as location-privacy, the focus usually has been on optimizing the QoS, i.e., the utility from the point of view of the users. In particular, this is the case for the framework proposed by Shokri et al. [45, 47].

We argue that it is important to meet the interest of all parties involved, and hence to consider both kinds of utility at the same time. Hence, a first goal of this paper is to develop a *location-privacy preserving mechanism (LPPM)* that, in addition to providing formal location-privacy guarantees, preserves as much as possible *both* the statistical utility and the QoS.

Now, one may think that statistical utility and QoS are aligned, since they both benefit from preserving as much original information as possible under the privacy constraint. However, this is not true in general: the optimization of statistical utility does not necessarily imply a significant improvement in the QoS, nor vice-versa. A counterexample is provided by Example 1.1 later in this section. Hence, the preservation of both statistical utility and QoS is more tricky than it may appear at first sight.

One of the methods which has been proposed to protect location privacy is the so-called *geo-indistinguishability* [4], which essentially obfuscates locations based on the distance between them. This idea works particularly well for protecting the precision of the location as it ensures that an attacker would not be able to differentiate between points which are close on the map by observing the reported noisy location. At the same time, it does not inject an enormous amount of noise that would be necessary to make far-away locations indistinguishable. Although this approach of distance-based obfuscation seems enticing at a first glance, one of the issues it poses is that it may leave the geo-spatially isolated locations vulnerable, i.e., identifiable despite being formally geo-indistinguishable [16]. To improve the situation, [16] introduced the notion of *elastic distinguishability metrics*, which essentially leads to inject more noise when the location to protect is isolated.

The *Blahut-Arimoto algorithm (BA)* [6, 8] from rate-distortion theory (a branch of information theory) Pareto-optimizes the trade-off between mutual information (MI) and average distortion. This property is appealing in the context of privacy because MI is often considered a measure of information leakage and average distortion is a commonly used metric for quantifying QoS. Moreover, BA was proven to satisfy geo-indistinguishability in [39] opening a door to study it as a potential LPPM. In this paper, we start off by exploring the privacy-preserving properties of BA and comparing them with those of the *Laplace mechanism (LAP)* [4] which is considered as the state-of-the-art mechanism for geo-indistinguishability. In the process, we show that, aside from being formally geo-indistinguishabile, BA offers an elastic distinguishability metric and, hence, protects even the most isolated points in the map, unlike LAP. We then examine the statistical utility, focusing on the estimation of the most general statistical information, namely the distribution of the original data, and show that the statistical utility of BA outperforms that of LAP for high levels of privacy, eventually becoming comparable as the level of privacy decreases. For both mechanisms, we consider the "best" estimation, i.e., the most likely distribution, that could produce the observed result which is computed using the *iterative Bayesian update (IBU)* [2], an instance of the famous *expectation maximization (EM)* method from statistics. Moreover, we prove an intriguing duality between BA and IBU. This is surprising because both algorithms were developed in different contexts, using different concepts and metrics, and for completely different purposes.

One important point to note is that the BA requires knowledge of the original distribution to provide the optimal mechanism. When it is fed with only an approximation of the distribution, it only provides an approximated result. We acknowledge that the distribution of the original data is usually off-limits and, even when available, it typically gets outdated over time. In any case, we can soundly assume that it is not available because it is essentially the reason for collecting the data. Hence we have a vicious circle: we want to collect data in a privacy-friendly fashion to estimate the original distribution while wanting to use a privacy mechanism that requires knowing a good approximation of the original distribution. Motivated by this dilemma, we propose PRIVIC, an incremental data collection method providing extensive privacy protection for the users of LBS while retaining a high utility for both the service providers and the users and ensuring that both parties, acting in their best interest, would benefit from the end mechanism catering to their corresponding privacy and utility requirements.

Finally, we empirically illustrated the convergence of our method and its privacy-utility trade-off. The experiments also demonstrate the efficacy of combining BA and IBU, in that the estimation of the original distribution is very accurate, especially when measured using a notion of distance between distributions compatible with the ground distance used to measure the QoS (e.g., the Earth Mover's distance). All the experiments were performed using real location data from the Gowalla dataset for Paris and San Francisco.

In summary, the key contributions of this paper are:

(1) We show, analytically and with experiments on real datasets, that the Blahut-Arimoto mechanism, in addition to being geo-indistinguishable, fosters an elastic distinguishability metric. As such, it protects the privacy of the isolated locations in the space, which the standard Laplace mechanism for geo-indistinguishability fails at.
(2) We empirically show that the Blahut-Arimoto mechanism provides a better statistical utility than the Laplace one for high levels of privacy and eventually they become comparable as the level of privacy decreases.
(3) We establish a surprising duality between the Blahut-Arimoto algorithm and the iterative Bayesian update, demonstrating a connection between the fields of rate-distortion theory and the expectation-maximization method from statistics.
(4) We propose an iterative method (PRIVIC) that, based on the approximate knowledge of the original distribution, which gets more precise as more (noisy) data get collected, produces a geo-indistinguishable LPPM with an elastic distinguishability metric, which incrementally tends to optimize the trade-off with the QoS (as more data get collected) and provides high statistical utility.
(5) We characterize the long-term behaviour of PRIVIC by translating it to the framework of Markov chains and illustrate, with experiments on real location datasets, that its estimation improves iteratively and, eventually, converges to the most likely true distribution.

Related Work. The clash between privacy and utility has been widely studied in the literature [10, 35]. Optimization techniques for DP and utility for statistical databases have been analyzed by the community from various perspectives [30, 32, 37]. There have been works focusing on devising privacy mechanisms that are optimal to limit the privacy risk against Bayesian inference attacks while maximizing the utility [45, 47]. In [39], Oya et al. examine an optimal LPPM w.r.t. various privacy and utility metrics for the user.

In [40], Oya et al. consider the optimal LPPM proposed by Shokri et al. in [47] which maximizes a notion of privacy (the *adversarial error*) under some bound on the QoS. The construction of the optimal LPPM requires the knowledge of the original distribution,

and [40] uses the EM method to estimate it and design *blank-slate models* that they empirically show to outperform the traditional hardwired models. However, a problem with their approach is that there may exist LPPMs which are optimal in the sense of [47], but with no statistical utility, see Example 1.1 below. Furthermore, for the mechanisms considered in [40] the EM method may fail to converge to the true distribution. Indeed, there are counterexamples shown in [24], which also points out several mistakes in the results of [2], on which [40] intrinsically relies to prove the convergence of their method.

Example 1.1. Consider three collinear locations, a, b and c, where b lies in between a and c at a unit distance from each of them. Assume that the prior distribution on these three locations is uniform and that the constraint on the utility is that it should not exceed $2/3$. Then a mechanism that optimizes the QoS in the sense of [47] is the one that maps all locations to b. However, this mechanism has no statistical utility, as the b's do not provide any information about the original distribution. Indeed, given n obfuscated locations (i.e., n b's) all distributions on a, b and c of the form $k_a/n, k_b/n, k_c/n$ with $k_a + k_b + k_c = n$, have the same likelihood to be the original one.

[42] proposed a method for generating privacy mechanisms that tend to minimize mutual information using an ML-based approach. However, this work assumes the knowledge of the exact prior from the beginning, unlike ours. Moreover, [42] does not provide formal guarantees for location privacy (e.g., geo-indistinguishability) which is one of the main aspects captured by our work. In [51], Zhang et al. consider the Blahut-Arimoto algorithm in the context of location privacy. However, their proposed method also requires the knowledge of the prior distribution to construct the LPPM. Additionally, [51] focuses on measuring privacy for the trace of a single user. On the contrary, our notion of privacy is in the spirit of "group privacy", i.e., for a community of users.

The Laplace mechanism has been rigorously studied in the literature in various scenarios as the cutting-edge standard to achieve geo-indistinguishability [4, 7, 29] and has been proven to be optimal for one-dimensional data w.r.t. Bayesian utility [26]. Despite the wide popularity, it has been recently criticized due to its limitation to protect geo-spatially isolated points from being identified by adversaries [16]. The authors of [16] addressed this concern by proposing the idea of *elastic distinguishability metrics*.

Our paper also considers mutual information (MI) as an additional privacy guarantee. MI and its closely related variants (e.g. conditional entropy) have been shown to nurture a compatible relationship with DP [20]. MI measures the correlation between observations and secrets, and its use as a privacy metric is widespread in the literature. Some key examples are: gauging anonymity [15, 52], estimating privacy in training ML models with a typical cross entropy loss function [1, 34, 42, 48], and assessing location-privacy [39].

A popular choice of utility metric for the users is the *average distortion*, which quantifies the expected quality loss of the service due to the noise induced by the mechanism. Such a metric has gained the spotlight in the community [4, 9, 14, 17, 47] due to its intuitive and simple nature. On the other hand, a standard notion of statistical utility for the data consumer is the precision of the estimation of the distribution on the original data from that of the noisy data. Iterative Bayesian update [2, 3] provides one of the most flexible and powerful estimation techniques and has recently become in the focus of the community [24, 25].

Incremental and privacy-friendly data collection has been explored both in the context of k-anonymity [5, 11, 12] and DP [33, 49]. However, to the best of our knowledge, the problem of providing a rather robust privacy guarantee while preserving utility for both data owners and data consumers has has not been addressed by the community so far.

Plan of the paper. Section 2 introduces preliminary ideas from the literature relevant to our paper. Section 3 highlights BA as an LPPM because of its extensive privacy-preserving properties. Section 4 establishes the duality between BA and IBU. Sections 5 and 6, respectively, explain our proposed method (PRIVIC) and characterize its long-term behaviour by modelling it as a Markov chain. Section 7 exhibits the working of PRIVIC with experiments using real locations from the Gowalla dataset illustrating the convergence of our method. Section 8 concludes. Appendices A, B, and C contain the proofs of the theorems derived in the paper, relevant tables supporting the experimental analysis of PRIVIC, and further mathematical study dissecting the functioning of PRIVIC, respectively.

The code used for implementing our mechanism for experiments is available at `https://anonymous.4open.science/r/PRIVIC`.

2 PRELIMINARIES

2.1 Standards of privacy

Definition 2.1 (*d-privacy*, a.k.a. *metric privacy* [13])**.** For any space $\mathcal{X}$ equipped with a metric $d : \mathcal{X}^2 \mapsto \mathbb{R}_{\geq 0}$ and an output space $\mathcal{Y}$, a mechanism $\mathcal{R} : \mathcal{X} \mapsto \mathcal{Y}$ is *d-private* if $\mathbb{P}[\mathcal{R}(x) = y] \leq e^{\epsilon d(x,x')} \mathbb{P}[\mathcal{R}(x') = y]$ for every $x, x' \in \mathcal{X}$ and $y \in \mathcal{Y}$.

Note that:

- Setting d as the *discrete metric* on any $\mathcal{X}$, we obtain the definition of *local differential privacy (LDP)* [21].
- Setting $\mathcal{X} = \mathcal{Y} = \mathbb{R}^2$ and d as the *Euclidean metric*, we get the definition of *geo-indistinguishability* [4].

Definition 2.2 (Mutual information[44])**.** Let (X, Y) be a pair of random variables defined over the discrete space $\mathcal{X} \times \mathcal{Y}$ such that μ is the joint PMF of X and Y, and p_X and p_Y are the marginal PMFs of X and Y, respectively, and $p_{X|Y}$ is the conditional probability of X given Y. Then the (Shannon) *entropy* of X, $H(X)$, is defined as $H(X) = -\sum_{x \in \mathcal{X}} p_X(x) \log p_X(x)$. The *residual entropy* of X given Y is defined as $H(X|Y) = \sum_{y \in \mathcal{Y}} p_Y(y) H(X|Y = y) = -\sum_{y \in \mathcal{Y}} p_Y(y) \sum_{x \in \mathcal{X}} p_{X|Y} p(x|y) \log p_{X|Y}(x|y)$, and, finally, the *mutual information (MI)* is given by:

$$I(X|Y) = H(X) - H(X|Y) = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} \mu(x, y) \log \frac{\mu(x, y)}{p_X(x) p_Y(y)}$$

2.2 Notions of utility

Definition 2.3 (Quality of service)**.** For discrete spaces $\mathcal{X}$ and $\mathcal{Y}$, let $d: \mathcal{X} \times \mathcal{Y} \to \mathbb{R}_{\geq 0}$ be any distortion metric (a generalization of the notion of distance). Let X be a random variable on $\mathcal{X}$ with PMF $p_{\mathcal{X}}$ and C be any randomizing mechanism where C_{xy} is the

probability of x being mapped by C into y. We define the *quality of service (QoS)* of X for C as the *average distortion w.r.t. d*, given as:

$$AvgD(X, C, d) = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p_{\mathcal{X}}(x) C_{xy} d(x, y)$$

Definition 2.4 (Full-support probability distribution)**.** Let θ be a probability distribution defined on the space $\mathcal{X}$. θ is a *full-support* distribution on $\mathcal{X}$ if $\theta(x) > 0$ for every $x \in \mathcal{X}$.

Definition 2.5 (Iterative Bayesian update [2])**.** Let C be a privacy mechanism that locally obfuscates points from the discrete space $\mathcal{X}$ to $\mathcal{Y}$ such that $C_{xy} = \mathbb{P}(y|x)$ for all $x, y \in \mathcal{X}, \mathcal{Y}$. Let $X_1, \ldots, X_n$ be i.i.d. random variables on $\mathcal{X}$ following some PMF $\pi_{\mathcal{X}}$. Let Y_i denote the random variable of the output when X_i is obfuscated with C.

Let $\overline{y} = \{y_1, \ldots, y_n\}$ be a realisation of $\{Y_1, \ldots, Y_n\}$ and q be the empirical distribution obtained by counting the frequencies of each y in $\overline{y}$. The *iterative Bayesian update (IBU)* estimates $\pi_{\mathcal{X}}$ by converging to the maximum likelihood estimate (MLE) of $\pi_{\mathcal{X}}$ with the knowledge of $\overline{y}$ and C. IBU works as follows:

(1) Start with any full-support PMF θ_0 on $\mathcal{X}$.
(2) Iterate $\theta_{t+1}(x) = \sum_{y \in \mathcal{Y}} q(y) \frac{\theta_t(x) C_{xy}}{\sum_{z \in \mathcal{X}} \theta_t(z) C_{zy}}$ for all $x \in \mathcal{X}$.

The convergence of IBU has been studied in [2, 24]. For a given set of observed locations, the limiting estimate $\hat{\pi}_{\mathcal{X}} = \lim_{t \to \infty} \theta_t$ is well-defined by the privacy mechanism in use, C, and the starting PMF θ_0. We will functionally denote $\hat{\pi}$ as $\Gamma_{\text{IBU}}(\theta_0, C)$.

Definition 2.6 (Earth mover's distance [36])**.** Let π_1 and π_2 be PMFs defined over a discrete space of locations $\mathcal{X}$. For a metric $d: \mathcal{X}^2 \mapsto \mathbb{R}_{\geq 0}$, the *earth mover's distance (EMD)* (aka the *Kantorovich–Rubinstein metric*) is defined as

$$EMD(\pi_1, \pi_2) = \min_{\mu \in \Pi(\pi_1, \pi_2)} \mu(x, y) d(x, y)$$

where $\Pi(\pi_1, \pi_2)$ is the set of all joint distributions over $\mathcal{X}^2$ such that for any $\eta \in \Pi(\pi_1, \pi_2)$, $\sum_{x \in \mathcal{X}} \eta(x_0, x) = \pi_1(x_0)$ and $\sum_{x \in \mathcal{X}} \eta(x, x_0) = \pi_2(x)$ for every $x_0 \in \mathcal{X}$.

EMD is considered a canonical way to lift a distance on a certain domain to a distance between distributions on the same domain.

Definition 2.7 (Statistical utility)**.** Let C be a privacy mechanism that obfuscates data on the discrete space $\mathcal{X}$. Let $\pi_{\mathcal{X}}$ be the PMF of the original locations and let $\hat{\pi}_{\mathcal{X}}$ be its estimate by IBU. Then we define the *statistical utility* of the mechanism C as $EMD(\hat{\pi}_{\mathcal{X}}, \pi_{\mathcal{X}})$.

2.3 Optimization of MI and QoS

Definition 2.8 (Blahut-Arimoto algorithm [6, 8])**.** Let X be a random variable on the discrete space $\mathcal{X}$ with PMF $\pi_{\mathcal{X}}$ and $\boldsymbol{C}(\mathcal{X}, \mathcal{Y})$ be the space of all channels encoding $\mathcal{X}$ to $\mathcal{Y}$. For a distortion $d: \mathcal{X} \times \mathcal{Y} \mapsto \mathbb{R}_{\geq 0}$ and fixed $d^* \in \mathbb{R}^+$, we wish to find the channel $\hat{C} \in \boldsymbol{C}(\mathcal{X}, \mathcal{Y})$ that minimizes MI given the bound d^* on distortion:

$$\hat{C} = \underset{\substack{C \in \boldsymbol{C}(\mathcal{X}, \mathcal{Y}) \\ AvgD(X, C, d) \leq d^*}}{\arg\min} I(X|Y_{X,C})$$

where, for any $C \in \boldsymbol{C}(\mathcal{X})$, $Y_{X,C}$ is the random variable on $\mathcal{Y}$ denoting the output of the encoding of X. The *Blahut-Arimoto algorithm (BA)* provides an iterative method to construct $\hat{C}$ as follows:

Table 1: Key notations

Notation	Meaning
$\mathcal{X} = \{x_1, \ldots, x_m\}$	Finite space of source locations
$\boldsymbol{C}$	Space of all stochastic channels on $\mathcal{X}$
$\Pi_{\mathcal{X}}$	Simplex of all full_support PMFs on $\mathcal{X}$
n	Number of samples
$X = \{\hat{x}_1, \ldots, \hat{x}_n\}$	Sample of original locations
$Y = \{y_1, \ldots, y_n\}$	Noisy locations
d^*	Maximum average distortion
β	Loss parameter of RD function
$\pi_{\mathcal{X}}$	PMF of the original locations (true PMF)
$C^{(0)}$	Uniform channel over $\mathcal{X}$, i.e., $C^{(0)}_{xy} = \frac{1}{\lvert\mathcal{X}\rvert}\ \forall x, y \in \mathcal{X}$
$\lambda_{\text{BA}}(\pi_{\mathcal{X}}, C^{(0)})$	Limiting channel by BA starting with $\pi_{\mathcal{X}} \in \Pi_{\mathcal{X}}$, $C^{(0)} \in \boldsymbol{C}$
$\Gamma_{\text{IBU}}(\theta, C)$	MLE of $\pi_{\mathcal{X}}$ by IBU starting with $\theta \in \Pi_{\mathcal{X}}$ under $C \in \boldsymbol{C}$
$\Gamma^t_{\text{IBU}}(\theta, C)$	Estimate by t^{th} iteration of IBU starting with $\theta \in \Pi_{\mathcal{X}}$ under $C \in \boldsymbol{C}$
N_{BA}	Number of iterations needed for BA to converge
N_{IBU}	Number of iterations needed for IBU to converge
$\Lambda(\theta, N)$	Estimate of $\pi_{\mathcal{X}}$ by PRIVIC after N iterations starting with $\theta \in \Pi_{\mathcal{X}}$
$\mathcal{P}_{\mathcal{X},k}$	Discretized $\Pi_{\mathcal{X}}$ with each component of the PMFs divided in k parts
$\phi(\theta\|\theta')$	Prob. of PRIVIC estimating $\theta' \in \mathcal{P}_{\mathcal{X},K}$ starting from $\theta \in \mathcal{P}_{\mathcal{X},k}$
Φ	Transition matrix of PRIVIC as a Markov chain over $\mathcal{P}_{\mathcal{X},k}$
ψ	Stationary distribution of the Markov chain of PRIVIC
$\mathcal{S}^t$	PMF in $\mathcal{P}_{\mathcal{X},k}$ estimated by PRIVIC after t iterations

(1) Start with any channel $C^{(0)}$ s.t. $C^{(0)}_{xy} > 0$ for all $x, y \in \mathcal{X}$.
(2) Iterate:

$$c_t(y) = \sum_{x \in \mathcal{X}} \pi_{\mathcal{X}}(x) C^{(t)}_{xy} \tag{1}$$

$$C^{(t+1)}_{xy} = \frac{c_t(y) \exp\{-\beta d(x, y)\}}{\sum_{z \in \mathcal{X}} c_t(z) \exp\{-\beta d(x, z)\}} \tag{2}$$

where $\beta > 0$ is the negative of the slope of the *rate-distortion* function $RD(X, d^*) = \min_{C \in \boldsymbol{C}(X)} I(X|Y_{X,C})$ under $AvgD(X, C, d) \leq d^*$. We call β the *loss parameter*, capturing the role of d^* in BA.

Remark 1. The equations (1) and (2) above define two transformations $\mathcal{F} : \boldsymbol{C}(\mathcal{X}, \mathcal{Y}) \to \boldsymbol{D}(\mathcal{X})$ and $\mathcal{G} : \boldsymbol{D}(\mathcal{X}) \to \boldsymbol{C}(\mathcal{X}, \mathcal{Y})$, where $\boldsymbol{D}(\mathcal{X})$ is the space of distributions on $\mathcal{X}$, so that $c_t = \mathcal{F}(C^{(t)})$ and $C^{(t+1)} = \mathcal{G}(c_t)$.

Remark 2. In [19], Csizar proved the convergence of BA when $\mathcal{X}$ is finite. The limit $\lim_{n \to \infty} (\mathcal{G} \circ \mathcal{F})^n (C^{(0)})$ is the optimal channel $\hat{C}$ (parametrized by β), and it is uniquely determined by the prior $\pi_{\mathcal{X}}$ and by the initial channel $C^{(0)}$. We will also denote $\hat{C}$ by $\lambda_{\text{BA}}(\pi_{\mathcal{X}}, C^{(0)})$ to underline this dependency. Note that $\hat{C}$ is a fixpoint of $\mathcal{G} \circ \mathcal{F}$, i.e. $\hat{C} = (\mathcal{G} \circ \mathcal{F})(\hat{C})$, and that $\hat{c} = \mathcal{F}(\hat{C})$ is a fixpoint of $\mathcal{F} \circ \mathcal{G}$.

Remark 3. In [39], Oya et al. proved that, when d is the Euclidean metric, the channel $\hat{C}$ obtained from BA with loss parameter β satisfies 2β-geo-indistinguishability.

In the context of the location-privacy, as addressed in this work, we obfuscate the original locations to points in the same space and, hence, for the rest of the paper we consider the spaces of the secrets and the noisy locations to be the same, i.e., $\mathcal{X} = \mathcal{Y}$.

3 LOCATION-PRIVACY WITH THE BLAHUT-ARIMOTO ALGORITHM

3.1 Motivation

MI is often used as an information theoretical notion of privacy, and average distortion is a standard choice for measuring QoS, as discussed in the introduction. Moreover, Remark 3 formally links the BA channel with geo-indistinguishability, thus providing DP-like guarantees, and Definition 2.8 shows that the BA channel optimizes between MI and average distortion, thus advocating for it being an optimal LPPM w.r.t. QoS.

In this section, we investigate the privacy protection offered by BA beyond geo-indistinguishability, study the statistical utility it renders, and compare it with LAP, the canonical mechanism for geo-indistinguishability.

3.2 Mutual information as a measure of privacy: an operational interpretation

Even though MI has often been used as the notion of privacy in the literature, it has also been critized for being too abstract and for missing a clear connection with natural privacy guarantees. Here we provide an operational interpretation in terms of an attacker and an oracle, which should clarify this point.

We assume that the attacker's goal is to find out the true location, and her means is to query the oracle. The oracle can only answer "yes" or "no". We also assume that the attacker wants to minimize the expected number of queries necessary to discover the true location and that she is rational, i.e., she chooses the best strategy to minimize the expected number of queries.

A strategy corresponds to a binary search tree, where the intermediate nodes represent the queries, and the leaves represent the true locations, labelled with their prior probability. The attacker starts from the root and tries to reach the correct leaf. On each node, the attacker asks the corresponding query and then she goes left or right depending on the answer of the oracle.

It can be proved that the best strategy corresponds to a perfectly balanced tree (or, if not possible, an "almost perfectly balanced" tree) from the point of view of the probability mass. This can be obtained by setting each query to split the probability mass into equal (or as similar as possible) parts between its left and right subtrees. (The query could be of the type "does the true location belong to S", for some S.) It can also be proved that for such a tree, the length of the path from the root to a leaf x is equal (or approximately equal) to $-\log p(x)$. Hence, the expected length of a path (i.e., the number of queries to get to discover the right x is $-\sum_x p(x)\log p(x)$, i.e., the entropy $H(X)$. Hence, we could consider $H(X)$ as a measure of robustness against the attack.

If the attacker sees an obfuscated location y, the attack is similar, except that now y gives some information about the true location x, i.e., the search tree must be constructed using the posterior $p(y|x)$. Furthermore, the expectation must be calculated by considering all the possible y's. This results in measuring the robustness to the attack, after the disclosure of the noisy location, as the residual entropy $H(X|Y)$. Note that it always holds that $H(X|Y) \leq H(X)$, which corresponds to the fact that the secret is expected to be more vulnerable when some information related to it is revealed.

Finally, the mutual information, defined as $I(X|Y) = H(X) - H(X|Y)$, represents how much the revelation of the obfuscated location decreases the robustness of the attack. The smaller is MI, the more private is the mechanism.

3.3 Elastic location-privacy with BA

One of the concerns harboured by geo-indistinguishability is that it treats the space in a uniform way, thus making isolated locations vulnerable to an attacker that knows the prior distribution. This issue has been raised and addressed by Chatzikokolakis et al. in [16] where the authors introduce a variant of LAP based on an *elastic distinguishability metrics*, which they refer to as *elastic mechanisms*. Such mechanism obfuscates locations not only by considering the Euclidean distance between them but also by taking into account an abstract attribute of the reported location, called *mass*, which is a parameter of the definition.

Formally, is $\mathcal{R}_{\text{elas}}$ is an elastic mechanism with privacy parameter ϵ defined on $\mathcal{X}$, then, for all $x, y \in \mathcal{X}$, $\mathcal{R}_{\text{elas}}$ must satisfy:

$$\mathbb{P}[\mathcal{R}_{\text{elas}}(x) = y] \propto \exp\{-\epsilon d_{\text{E}}(x, y)\} \tag{3}$$

$$\mathbb{P}[\mathcal{R}_{\text{elas}}(x) = y] \propto q(y) \tag{4}$$

where $q(y)$ is the mass of the reported location.

$\mathcal{R}_{\text{elas}}$, unlike LAP, protects a point in a densely populated area (e.g. city) and a geo-spatially isolated point (e.g. island) differently by considering not only the ground distance between the true and the reported locations but also the mass of the reported location. The exact mechanism depends of course on how we define the notion of mass. A natural way, and the most meaningful from the privacy point of view, is to set the mass of y to be the probability to be reported (from any true location x). Under this definition, the interpretation of (4) is in the spirit of group privacy: given a true location x, we tend to report with higher probability those locations y that are reported with high probability from other locations as well, so that it becomes harder to re-identify x as the original one. Note that this property is not incompatible with the geo-indistinguishability guarantee. However, LAP does not capture it.

Obviously, the definition of mass as probability to be reported would be circular, because it would depend on the mechanism, which in turn is defined in terms of the mass. The authors of [16] do not explain how this mechanism could be constructed. Fortunately, the following theorem shows that an elastic mechanism of this kind can be constructed using BA. The proof is in Appendix A.

Theorem 3.1. The privacy channel generated by BA produces an elastic location-privacy mechanism.

Note also that there can be many mechanisms satisfying (3) and (4) (also with the mass interpreted as probability). The one produced by the BA is the mechanism that offers the best QoS among these. Finally, a consequence of the connection with BA is that it provides an understanding of the elastic mechanism in terms of information theory and of the attacker illustrated in previous section.

Experimental validation. Having furnished the theoretical foundation, we now enable ourselves to empirically validate that BA, indeed, satisfies the properties of the elastic mechanism unlike LAP, its state-of-the-art geo-indistinguishable counterpart. We perform experiments using real location data from the Gowalla

dataset [18, 31]. We consider 10,078 Gowalla check-ins from a central part of Paris bounded by latitudes (48.8286, 48.8798) and longitudes (2.2855, 2.3909) covering an area of 8Km×6Km discretized with a 16×12 grid.

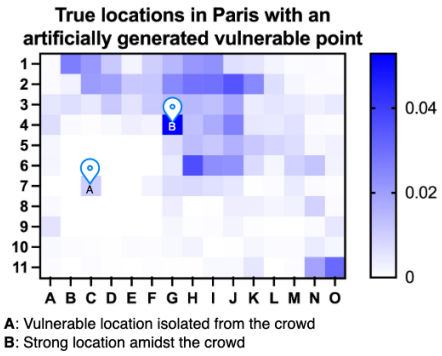


Figure 1: Gowalla check-in locations in Paris with an artificially planted vulnerable point, A, in isolation, and a strong point, B, in a crowd.

In order to demonstrate the property of an elastic mechanism shown by BA, we artificially introduced an "island" amidst the locations in Paris by choosing a grid A in a low-density area of the dataset (in the south-west region), assigning the probability mass of the grids around A to 0, and dumping this cumulative mass from the surrounding region to A, ensuring that the sum of the probability masses of all the grids remain to be 1. We call A as a *vulnerable location* in the map as it's isolated from the crowd. To visualize the elastic behaviour of the mechanisms for locations in crowded regions, we consider another grid B in the central part of the map which has a high probability mass and has a highly populated surrounding – we refer to such a grid B as a *strong location* in the map. Figure 1 illustrates the selection of vulnerable and strong locations in the Paris dataset.

For the mechanism derived from BA with a loss parameter β, we know, by Remark 3, that the privacy parameter ϵ is 2β, which we use to tune the privacy level of LAP in order to demonstrate the elastic property of the two mechanisms under the same level of privacy. Figure 2 illustrates the probability distribution of reporting a privatized point on the map by obfuscating the vulnerable and the strong locations with different levels of geo-indistinguishability – we vary the value of ϵ to be 0.4, 1.2, 1.6, 2 to capture the property of the elastic mechanism shown by BA unlike that by LAP.

In particular, Figure 2a shows the obfuscation distribution of the vulnerable location on the map. By comparing with the distribution of the true locations in Paris given by Figure 1, we observe that when the value of ϵ is low (privacy is high), the reported location with BA is likely to be mapped to a nearby densely populated place. For example, with $\epsilon = 0.2$, the highest level of privacy considered in the experiments, the location reported by BA will most probably be around the most crowded region of Paris. As ϵ increases, the location most likely to be reported by BA systematically moves to a densely populated region closer and closer to the true vulnerable location. LAP, on the other hand, always obfuscates the every location around its true position in the map – varying the value of ϵ changes the spread of the distribution around the true location. This might be problematic as the vulnerable location is known to be isolated and, hence, even being reported somewhere nearby would potentially result in it being identified.

For example, we would like to highlight the setting of $\epsilon = 1.6$ for the vulnerable location to show that the distribution of the location reported by LAP is almost completely around the true vulnerable point covering an area which is deserted, i.e., there is no realistic chance of someone being located in that region. Thus, despite providing formal 1.6-geo-indistinguishability, LAP fails to protect such a vulnerable location from being potentially identified. BA, on the other hand, does the job quite well, adhering to the principles of the elastic mechanism – it distributes the reported location in the crowded areas nearby providing a sense of camouflage amidst the crowd in addition to 1.6-geo-indistinguishability.

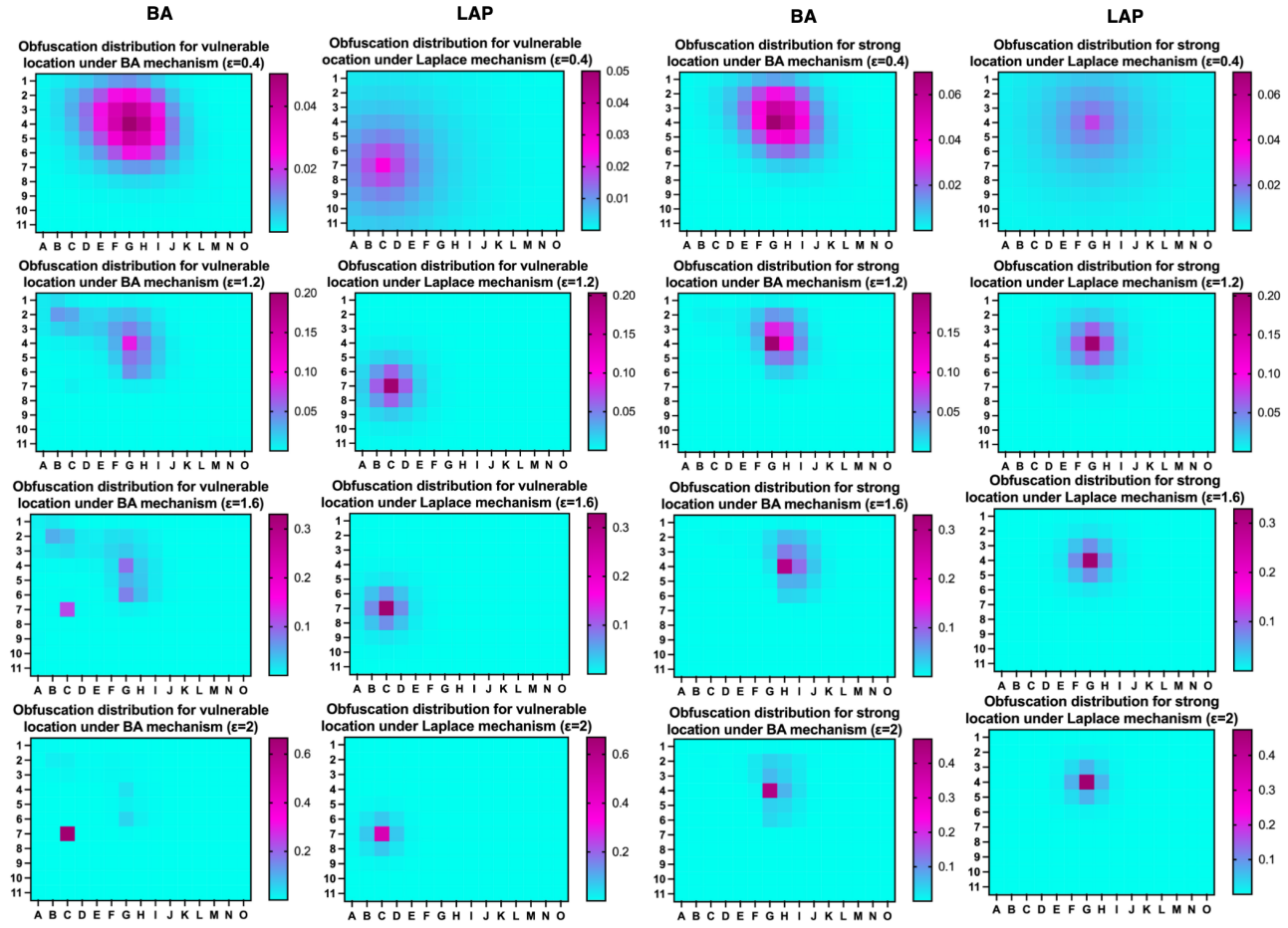
In the case of privatizing the strong location, Figure 2b shows that both BA and LAP behave similarly by concealing the point around its true position. This would not give rise to a similar issue as for the vulnerable location because, by definition, the strong location B is already positioned in a highly dense region of the map and, hence, being privatized, it will still remain among the crowd with a high probability.

Focusing on the utility of individual users, we note that due to theories from Nash equilibrium [38] and Hotelling's spatial competition [28], a huge fraction of the typical points of interests (POIs) like cinemas, theatres, restaurants, retails, etc. lie in crowded areas syncing with the distribution of population. Therefore, for an isolated point in the map who is located in some extremely unpopulated area (e.g. some forest or island far from the city), the closest POI is usually going to be in the nearest urban region, i.e., region on the map with a high density of population. Suppose A is one such isolated location and let A_{BA} and A_{LAP} be the reported locations for A obfuscated with BA and LAP, respectively. Due to the elastic property of BA, A_{BA} will is likely to be at a nearby crowded location to A, while A_{LAP} is likely to be around the true location A. Let P_{BA} and P_{LAP} be the nearest POIs from the reported locations A_{BA} and A_{LAP}, respectively. The most likely scenario is that P_{BA} and P_{LAP} are almost at a similar place under the assumption that typical POIs follow the distribution of the crowd and, and, therefore, a vulnerable user has to travel a similar distance from their true position in both the cases, except that under LAP, the privacy of A will be compromised much more than that under BA.

3.4 Statistical utility: BA vs LAP

Now we proceed to empirically compare the statistical utility of BA and LAP by performing experiments on the locations obtained from the Gowalla dataset for two different cities: Paris and San Francisco. In addition to the same setting for the Gowalla check-ins in Paris as considered in the experiments of Section 3.3, here we also test for 123,025 check-in locations from the Gowalla dataset in a northern part of San Francisco bounded by latitudes (37.7228, 37.7946) and longitudes (-122.5153, -122.3789) covering an area of 12Km×8Km discretized with a 24×17 grid. The locations were privatized with BA and LAP under varying levels of privacy – the loss parameter, β, for BA ranged from 0.2 to 5.0, which implies that the value of the geo-indistinguishability parameter, ϵ, ranged from 0.4 (very high level of privacy) to 10.0 (almost no privacy). To account for the randomness in the process of generating the sanitized locations, 5 simulations were run for each setting of the privacy parameter for obfuscating every location in both datasets.

Figure 3 reveals that BA possesses a significantly better statistical utility than LAP for a high level of privacy (for $\beta \in (0.4, 1.4]$



(a) Reporting distribution the vulnerable location A **(b) Reporting distribution of the strong location B**

Figure 2: Distribution of privatizing the vulnerable and the strong locations for different levels of privacy. Top down, the rows illustrates the results for $\epsilon = 0.4, 1.2, 1.6, 2$, respectively.

and $\beta \in (0, 1)$, i.e., ϵ up to 2.8 and 2, in Paris and San Francisco datasets, respectively). As the level of privacy decreases, the EMD between the true and the estimated PMFs converge to 0 in both the mechanisms, as we would expect, fostering the maximum possible statistical utility with, practically, no privacy guarantee.

Summarizing the results from Sections 3.3 and 3.4, we can establish that:

- in addition to providing a formal geo-indistinguishability guarantee, BA also gives rise to an LPPM with an elastic distinguishability metric to enhance the privacy of the vulnerable and isolated population.
- BA optimizes the trade-off between QoS and privacy.
- the statistical utility for high levels of privacy is significantly better for BA than LAP.

Therefore, we conclude that BA is a key contender for providing a comprehensive notion of location-privacy while preserving the utility of the data for both the users and the service providers.

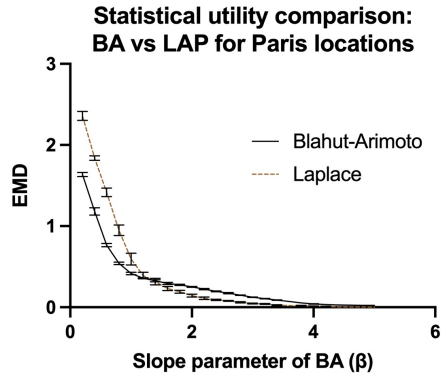
4 DUALITY BETWEEN IBU AND BA

We now explore an intriguing relationship between BA and IBU. Letting $(\mathcal{X}, d)$ to be a finite metric space, let X be a random variable on $\mathcal{X}$ with PMF $\pi_{\mathcal{X}}$. Recalling the iteration of BA from (1) and (2):

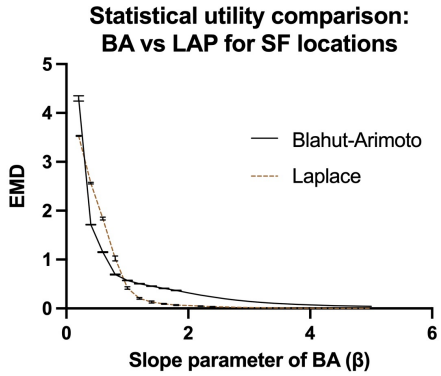
$$c_t(y) = \sum_{x \in \mathcal{X}} \pi_{\mathcal{X}}(x) C_{xy}^{(t)} \text{ and } C_{xy}^{t+1} = \frac{c_t(y) \exp\{-\beta d(x,y)\}}{\sum_{z \in \mathcal{X}} c_t(z) \exp\{-\beta d(x,z)\}}$$

$$\therefore c_{t+1}(y) = \sum_{x \in \mathcal{X}} \pi_{\mathcal{X}}(x) C_{xy}^{(t+1)} = \sum_{x \in \mathcal{X}} \pi_{\mathcal{X}}(x) \frac{c_t(y) \exp\{-\beta d(x,y)\}}{\sum_{z \in \mathcal{X}} c_t(z) \exp\{-\beta d(x,z)\}} \tag{5}$$

Comparing it with the iteration of IBU as in Definition 2.5, we observe that (5) BA is dual to IBU. Indeed, consider an exponential mechanism of the form $C = c \cdot \exp\{-\beta d(x,y)\}$ Flipping the roles of x and y in (5), and replacing the input distribution $\pi_{\mathcal{X}}$ with the empirical distribution in output to C, we obtain the iterative step of IBU.



(a) Statistical utility for BA and Laplace on locations in Paris



]]

(b) Statistical utility for BA and Laplace on locations in SF

Figure 3: EMD between the true and the estimated distributions for locations in Paris and San Francisco under BA and Laplace mechanisms denoting the statistical utility for the two mechanisms in two different datasets.

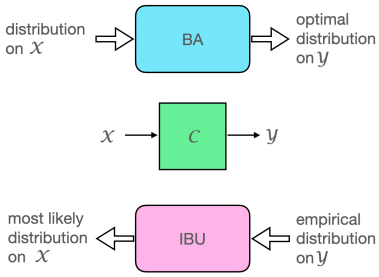


Figure 4: Illustration of the duality between BA and IBU.

Figure 4 illustrates the duality. Due to this duality and taking advantage of the fact that BA converges, i.e., $\lim_{t\to\infty} c_t$ exists, we obtain a new proof of the following theorem, alternative to the proof given by ElSalamouny et al. in [24].

Theorem 4.1. IBU always converges to the MLE of the true prior for a set of observed locations on a finite space privatized with a planar geometric mechanism with Euclidean metric.

Proof. Immediate from the duality between IBU and BA by instantiating $\mathcal{X}$ as a set of locations and d as the Euclidean metric. □

5 PRIVIC: A PRIVACY-PRESERVING METHOD FOR INCREMENTAL DATA COLLECTION

To ensure that the produced mechanism is truly optimal, the BA needs a good approximation of the prior distribution. At the beginning we cannot assume to have such knowledge, but as the service providers incrementally collect data from their users, we can use these data to refine the estimation of the prior and get a better mechanism. These data, however, are obfuscated by the privacy mechanism, hence it is not obvious that the estimation of the prior really improves in the process. We show that this is the case, and, summarizing all results obtained for BA so far, we propose a method that facilitates the service providers to incrementally collect data and gradually achieve a highly with respect to the QoS. We shall refer to our proposed method for **PRIV**acy-preserving **I**ncremental **C**ollection of location data as *PRIVIC*.

In the scope of our proposed method, we shall consider locations sampled from a finite space $\mathcal{X} = \{x_1, \ldots, x_m\}$. Let the *true distribution* or *true PMF* on $\mathcal{X}$ (from which the users' locations are sampled) be $\pi_{\mathcal{X}}$. Note that we do not assume the knowledge of $\pi_{\mathcal{X}}$ in our method. In this work, to be able to achieve geo-indistinguishability, we shall adhere to the Euclidean metric d_{E} to measure ground distance between locations. PRIVIC proceeds as follows:

1. Start with any stochastic channel $C^{(0)}: \mathcal{X}^2 \mapsto \mathbb{R}^+$ and a full-support PMF on θ_0 defined on $\mathcal{X}$. Set $\hat{C}^{(0)} = C^{(0)}$.
2. In step $t \geq 1$:
   i) For a fixed the maximum average distortion, set $\hat{C}^{(t)} = \lambda_{\mathrm{BA}}\left(\theta_{t-1}, C^{(0)}\right)$.
   ii) The users locally obfuscate their true locations with $\hat{C}^{(t)}$ to report the noisy samples to the service provider who obtains the empirical distribution of the observed locations $\boldsymbol{q} = \{\boldsymbol{q}(x): x \in \mathcal{X}\}$.
   iii) $\theta_t = \Gamma_{\mathrm{IBU}}\left(\theta_{t-1}, \hat{C}^{(t)}\right)$.

Algorithm 1: PRIVIC

Input: Full-support PMF: θ_0, Positive stochastic channel: $C^{(0)}$, Loss parameter: β, No. of iterations: N, No. of iterations of BA: N_{BA}, No. of iterations of IBU: N_{IBU};
Output: Estimation of true PMF: $\hat{\pi}_{\mathcal{X}}$;
Assign $\hat{C}^{(0)} \leftarrow C^{(0)}$;
Set $t \leftarrow 0$;
while $t \leq N$ **do**
  $\hat{C}^{(t+1)} = \mathrm{BA}(\theta_t, \hat{C}^{(0)}, \beta, N_{\mathrm{BA}})$;
  $\mathcal{L} \leftarrow \{y_1, \ldots, y_n\}$: Noisy locations reported by users after obfuscating their true locations with $\hat{C}^{(t)}$;
  $\boldsymbol{q} \leftarrow \{q(x): x \in \mathcal{X}\}$: Empirical PMF obtained from $\mathcal{L}$ by the service provider;
  $\theta_{t+1} \leftarrow \mathrm{IBU}(\hat{C}^{(t+1)}, \theta_t, \boldsymbol{q}, N_{\mathrm{IBU}})$;
  $t \leftarrow t + 1$;
$\hat{\pi}_{\mathcal{X}} \leftarrow \theta_N$;
Return: $\hat{\pi}_{\mathcal{X}}$

Algorithm 2: Blahut-Arimoto algorithm (BA)

Input: PMF: π, Initial channel $C^{(0)}$, Loss parameter: $\beta > 0$, No. of iterations: N_{BA};
Output: Channel giving minimum mutual information for maximum avg. distortion encapsulated by β: $\hat{C}$;
Function $\texttt{BA}(\pi, C^{(0)}, \beta, N_{BA})$**:**
  Set $t \leftarrow 0$;
  while $t \leq N_{BA}$ **do**
    $c_t(y) \leftarrow \sum_{x \in \mathcal{X}} \pi(x) C_{xy}^{(t)}$;
    $C_{xy}^{(t+1)} \leftarrow \frac{c_t(y) \exp\{-\beta d_{\text{E}}(x,y)\}}{\sum_{z \in \mathcal{X}} c_t(z) \exp\{-\beta d_{\text{E}}(z,y)\}}$;
    $t \leftarrow t + 1$
  $\hat{C} \leftarrow C^{(N_{\text{BA}})}$;
  Return: $\hat{C}$

Algorithm 3: iterative Bayesian update (IBU)

Input: Privacy channel: C, Full-support PMF: θ_0, Empirical PMF from observed data: $\boldsymbol{q}$, No. of iterations: N_{IBU};
Output: MLE of true PMF: $\hat{\pi}$;
Function $\texttt{IBU}(C, \theta_0, \boldsymbol{q}, N_{IBU})$**:**
  Set $t \leftarrow 0$;
  while $t \leq N_{IBU}$ **do**
    $\theta_{t+1}(x) \leftarrow \sum_{y \in \mathcal{X}} \boldsymbol{q}(y) \frac{C_{xy}\theta_t(x)}{\sum_{z \in \mathcal{X}} C_{zy}\theta_t(z)}$;
    $t \leftarrow t + 1$
  $\hat{\pi} \leftarrow \theta_{N_{\text{IBU}}}$;
  Return: $\hat{\pi}$;

The goal of PRIVIC is to construct an obfuscation channel that guarantees formal geo-indistinguishability, acts as an elastic mechanism, and optimizes between MI and QoS. To circumvent any bias for BA, we initiate it with a uniform channel in each iteration of PRIVIC, i.e., $C_{xy}^{(0)} = 1/|\mathcal{X}|$ for all $x, y \in \mathcal{X}$. Let the privacy channel generated this way after N iterations, for a fixed of maximum average distortion and starting from a uniform initial channel, be functionally represented as $\Lambda(\theta_0, N)$, as PRIVIC is entirely defined and its functioning is determined as a function of the starting full-support PMF θ_0 by fixing the rest of the parameters.

It is of utmost importance to preserve the statistical utility of the privatized data under PRIVIC. To evaluate the statistical utility of $\Lambda(\theta_0, N)$, we measure the EMD between the true and the estimated PMFs at the end of N iterations of PRIVIC. Thus, the quantity $EMD(\hat{\pi}_{\mathcal{X}}, \pi_{\mathcal{X}})$ parameterizes the utility of $\Lambda(\theta_0, N)$ for the service providers. Here we use the same Euclidean distance as the underlying metric for computing, both, the EMD and the average distortion – this consistency threads together and complements the notion of *utility* of the service providers and that from the sense of the QoS of the users. The opportunity to capture the essence of the same metric that quantifies QoS and statistical utility for a privacy channel was one of the motivations to use EMD to compare the true and the estimated distributions. On this note, we observe one of the most crucial properties of BA being aligned with IBU in the context of PRIVIC given by Theorem 5.1.

Theorem 5.1. Implementing BA with a starting PMF θ_0 on $\mathcal{X}$ and a uniform initial channel $C^{(0)}$, there is a unique MLE for the prior for a set of observed locations on $\mathcal{X}$ obfuscated with $\lambda_{\text{BA}}\left(\theta_0, C^{(0)}\right)$.

Theorem 5.1 (the proof is provided in Appendix A.) essentially shows that in each iteration of PRIVIC, the unique MLE given by the noisy locations under the channel generated by BA will be estimated by IBU. This is one of the major aspects where PRIVIC triumphs over the method proposed by Oya et al. in [40] which relies on the flawed theoretical results by [2] and adheres to the idea of optimal LPPMs presented by Shokri et al. in [46]. As discussed in Section 1, ElSalamouny et al. in [24] illustrate how various LPPMs, which would be optimal by Shokri et al.'s standards in [46], may not have a unique MLE for a given set of observed data obfuscated by them and, hence, the method proposed by Oya et al. in [40] may not converge. This is one of the principal reasons why the EM method highlighted in [40] is not reliable to estimate the true PMF to a desirable degree of accuracy.

6 PRIVIC AS A MARKOV CHAIN

Remark 2 and Theorem 5.1, along with the privacy-preserving properties shown by BA, attest that in a single iteration, PRIVIC would spawn a unique channel guaranteeing geo-indistinguishability, having elastic distinguishability metric, and optimizing the information theoretical notion of privacy and QoS of the users, and eventually it would converge to the unique MLE of the original distribution observing the noisy locations sanitized with the privacy channel it engenders. Now we proceed to to examine the convergence of PRIVIC in depth.

First of all, we acknowledge that discretizing the probability simplex is reasonable under the realistic computational boundaries. Let $\Pi_{\mathcal{X}} = \{\theta : \sum_{x \in \mathcal{X}} \theta(x) = 1,\ 0 < \theta(x) \leq 1\}$ be the probability simplex of full-support PMFs on the finite space of locations $\mathcal{X}$. We discretize the interval $(0, 1]$ in k equal intervals, allowing $\theta(x)$ to take the values from $(0, 1]_k = \{\frac{1}{k}, \ldots, \frac{k-1}{k}, 1\}$ for every $x \in \mathcal{X}$ and for every $\theta \in \Pi_{\mathcal{X}}$. Note that complying to our need or the available computational capacity , k can be made as large as desired – the only requirement is to have a finite k. For example, in the case of using Python as a computational resource, k could be assigned something as large as $\approx 1e308$.

An important consequence of such a discretization of $\Pi_{\mathcal{X}}$ is that it guarantees the finiteness of the probability simplex. Let $\mathcal{P}_{\mathcal{X},k} = \{\theta : \sum_{x \in \mathcal{X}} \theta(x) = 1,\ \theta(x) \in (0, 1]_k\}$ be the discretized probability simplex on $\mathcal{X}$. Note that $|\mathcal{P}_{\mathcal{X},k}| < k^{|\mathcal{X}|-1}$, i.e., the size of the discretized probability simplex is finite. An alternative perspective to this is that $\mathcal{P}_{\mathcal{X},k}$ introduces a discretized mesh within the continuous $\Pi_{\mathcal{X}}$ on $\mathcal{X}$, and therefore, every full-support PMF θ on $\mathcal{X}$ lies on the discrete mesh. The coarseness of this mesh is tuned by the value of k is. In particular, $\lim_{k \to \infty} \mathcal{P}_{\mathcal{X},k} = \Pi_{\mathcal{X}}$. For a given k, let K denote the size of the discretized probability simplex on $\mathcal{X}$, i.e., $|\mathcal{P}_{\mathcal{X},k}| = K < \infty$. Figure 5 illustrates an example of such a discretization of a probability simplex on a location space of size 3. In this example, K is the number of points of intersection in the

grids inside the area of the triangle denoting the three-dimensional probability simplex for full-support PMFs.

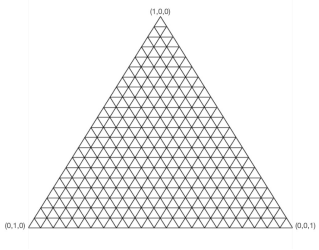


Figure 5: A visualization of a discretized simplex on $\mathcal{X}$ of size 3 showing that the continuous simplex (the entire area inside the triangle) has been made finite with the mesh.

Now we align this idea of making the probability simplex finite with the aim to model the long-term behaviour of PRIVIC. We note that starting with any PMF $\theta_0 \in \mathcal{P}_{\mathcal{X},k}$, the chance of PRIVIC ending up at some $\theta' \in \mathcal{P}_{\mathcal{X},k}$ as an approximation of the true prior at the end of one iteration is determined by the noise injected by the channel generated by $\lambda_{\text{BA}}\left(\theta_0, C^{(0)}\right)$. In other words, in t^{th} iteration of PRIVIC, feeding θ_{t-1} to BA and implementing it with the uniform channel $C^{(0)}$, the limiting channel of BA, $\hat{C}^{(t)}$, is uniquely determined (Remark 2), and once the true locations are obfuscated with $\hat{C}^{(t)}$, IBU gives the unique MLE of the true prior for the observed set of locations under $\hat{C}^{(t)}$ (Theorem 5.1). Thus, the only door of randomness in the t^{th} cycle of PRIVIC is the addition of noise done by $\hat{C}^{(t)}$. In particular, the probability of ending up in certain $\theta' \in \mathcal{P}_{\mathcal{X},k}$ after one cycle of PRIVIC starting from some $\theta \in \mathcal{P}_{\mathcal{X},k}$ is some function F of the optimal channel that BA converged to within this cycle, i.e., $\mathbb{P}\left[\Lambda(\theta, 1) = \theta'\right] = F\left(\lambda_{\text{BA}}\left(\theta, C^{(0)}\right)\right)$.

We take a step further towards abstraction by looking at the probability of an iteration of PRIVIC estimating a certain PMF as the true prior given that it starts from a PMF. Let $\phi: \mathcal{P}^2_{\mathcal{X},k} \mapsto \mathbb{R}^+$ be such that $\phi(\theta'|\theta) = \mathbb{P}\left[\Lambda(\theta, 1) = \theta'\right]$ for every $\theta, \theta' \in \mathcal{P}_{\mathcal{X},k}$. $\phi(\theta'|\theta)$. This essentially denotes the probability of a cycle of PRIVIC landing up in a certain $\theta' \in \mathcal{P}_{\mathcal{X},k}$ staring from some $\theta \in \mathcal{P}_{\mathcal{X},k}$. This creates an environment to model PRIVIC as a *Markov chain* over the finite *state space* $\mathcal{P}_{\mathcal{X},k}$ with a transition matrix Φ such that $\Phi_{\theta,\theta'} = \phi(\theta'|\theta)$ for every distribution $\theta, \theta' \in \mathcal{P}_{\mathcal{X},k}$. With this interpretation, let $\mathcal{S}^t$ be the random variable denoting the state in $\mathcal{P}_{\mathcal{X},k}$ PRIVIC is at in its t^{th} iteration.

Remark 4. Due to the discretization of the probability simplex $\Pi_{\mathcal{X}}$ into $\mathcal{P}_{\mathcal{X},k}$, PRIVIC can be modelled as discrete-time a Markov chain on a finite state space.

Definition 6.1 (Markov chain preliminaries)**.**

(1) A discrete-time Markov chain with a transition matrix P over finite state space is *irreducible* if for all states x, y, there is $0 \leq t < \infty$ such that $P^t_{xy} > 0$, where P^t is the transition matrix for the Markov chain at time t.
(2) For a discrete-time Markov chain with a transition matrix P over finite state space starting from any state x, let $T(x) = \{t \geq 1 : P^t(x, x) > 0\}$ be the set of all time-steps at which it returns to x with a non-zero probability. Then the *period* of state x is $\gcd T(x)$.
(3) A discrete-time Markov chain over finite state space is called *aperiodic* if the period of all of its states is 1.
(4) For a discrete-time Markov chain with a transition matrix P over finite state space, a distribution ψ on the state space is called a *stationary distribution* iff $\psi P = \psi$.
(5) For a discrete-time Markov chain with a transition matrix P over finite state space, let the *first hitting time* for a state x be defined as $\tau(x) = \min\{t \geq 1 : \mathcal{S}^t = x\}$.

Lemma 6.1. PRIVIC is an irreducible and aperiodic Markov chain.

Now with Lemma 6.1 (the proof in provided in Appendix A) we have laid the foundations to analyze the long-term behaviour of our proposed mechanism seen as a Markov chain. Aiming to investigate the limiting behaviour of PRIVIC if it is run for long enough, we concede to a well-known result in probability theory: an irreducible discrete-time Markov chain over a finite state space has a unique stationary distribution (Theorem 3.3 in [27]). As a consequence, we get the following theorem.

Theorem 6.2. PRIVIC, seen as a discrete-time Markov chain with transition matrix Φ over the finite state space $\mathcal{P}_{\mathcal{X},k}$, has a unique stationary distribution ψ over $\mathcal{P}_{\mathcal{X},k}$ and it is given by $\psi(\theta) = \frac{1}{\mathbb{E}(\tau(\theta))}$ for every $\theta \in \mathcal{P}_{\mathcal{X},k}$.

Proof. Immediate from Corollary 39 and Theorem 54 by Serfozo in [43]. Explicitly, Theorem 3.3 in [27]. □

Exploiting the elevation of PRIVIC to a Markov chain over a finite state space and as a direct derivative of Theorem 6.2, we can conclude the following interesting result.

Theorem 6.3. Let ψ be the unique stationary distribution for the discrete-time Markov chain of PRIVIC over the finite state space $\mathcal{P}_{\mathcal{X},k}$. Then, over time, the estimation of the true PMF given by PRIVIC follows the distribution ψ, i.e., $\lim_{t\to\infty} \mathcal{S}^t \sim \psi$.

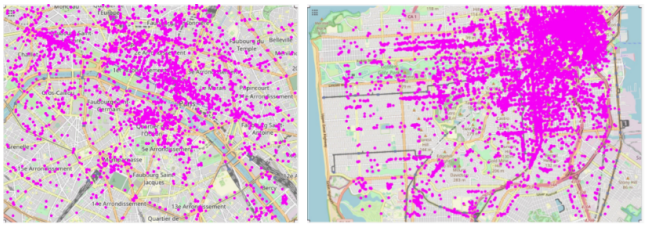
Proof. An immediate corollary of the *Perron–Frobenius theorem* [50]. An explicit proof has been given by Freedman in Theorem 4.9 of [27]. □

One very interesting interpretation of Theorem 6.3 is that irrespective of the starting distribution fed into PRIVIC, after enough number of iterations, it will estimate the true PMF following a fixed distribution that can be computed independently and beforehand using ψ as in Theorem 6.2.

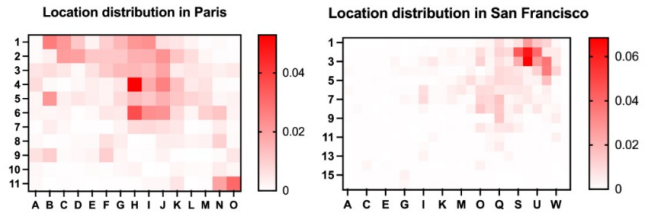
7 EXPERIMENTAL ANALYSIS OF PRIVIC

In this section, we describe the empirical results obtained by carrying out experiments to illustrate and validate the working of our proposed method. Like in the previous experiments to compare the statistical utilities of BA and LAP, as elaborated in Section 3.4, we use real locations from the same regions in Paris and San Francisco from the Gowalla dataset [18, 31]. In particular, we consider Gowalla check-ins from (i) a northern part of San Francisco bounded by latitudes (37.7228, 37.7946) and longitudes (-122.5153, -122.3789) covering an area of 12Km×8Km discretized with a 24×17 grid; (ii) a central part of Paris bounded by latitudes (48.8286, 48.8798) and longitudes (2.2855, 2.3909) covering an area of 8Km×6Km discretized with a 16×12 grid. In this setting, we work with 123,108 check-in locations in San Francisco and 10,260 check-in locations in Paris.

Figure 6a shows the particular points of check-in from Paris and San Francisco and Figure 6b highlights their distribution.



(a) Check-in locations in Paris and San Francisco

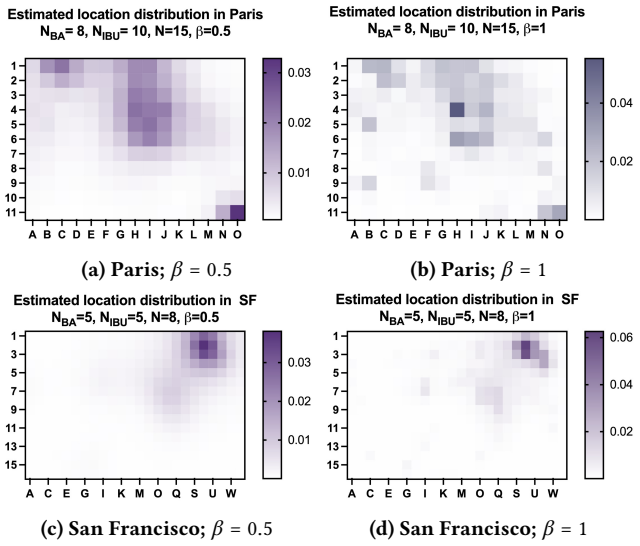


(b) Density of the original locations from Paris and San Francisco

Figure 6: (a) visualizes the original locations from Gowalla dataset from Paris and San Francisco. (b) illustrates a heatmap representation of the locations in the two cities as to capture the distribution of the data.

We implemented PRIVIC on the locations from Paris and San Francisco separately to judge its performance on real data with very different priors. In both the cases, we ran our mechanism until it empirically converged to the estimated PMF – 15 cycles of PRIVIC were required for the Paris dataset where each cycle comprised of 8 iterations of BA and 10 iterations of IBU, while in case of San Francisco, PRIVIC needed 8 cycles to converge with 5 iterations of BA and IBU each in every cycle. In both cases, we assigned the value of the loss parameter signifying the QoS of the users, β, to be 0.5 and 1. This was done to test the performance of PRIVIC in estimating the true PMF under two different levels of privacy. Each experiment was run for 5 rounds of simulation to calibrate the randomness of the sampling and obfuscation. In each cycle of PRIVIC, across all the settings, BA was initiated with the uniform channel $C^{(0)}$ and a uniform distribution over the space of locations as the "starting guess" of the true distribution.

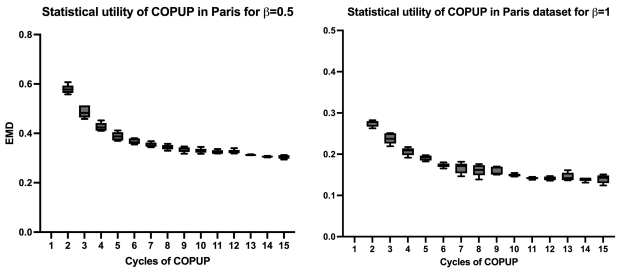
With $\beta = 1$, BA produces a geo-indistinguishable mechanism that injects less local noise than that obtained with $\beta = 0.5$. As a result, PRIVIC obtains a more accurate estimate of the true PMF for the former as we would expect. However, in both cases, the EMD between the true and the estimated PMFs is very low, indicating a good performance of the PRIVIC mechanism in preserving the statistical utility. Moreover, for both Paris and San Francisco, PRIVIC seems to significantly improve its estimation of the true PMF with every iteration until it converges to the MLE. Comparing Figures 7 and 6b, we see that the estimations of the true distributions of the locations in Paris and San Francisco by IBU under PRIVIC for both the settings of the loss parameter are fairly accurate. However, as we would anticipate, the statistical utility for $\beta = 1$ is better than that for $\beta = 0.5$.



(a) Paris; $\beta = 0.5$ **(b) Paris; $\beta = 1$**

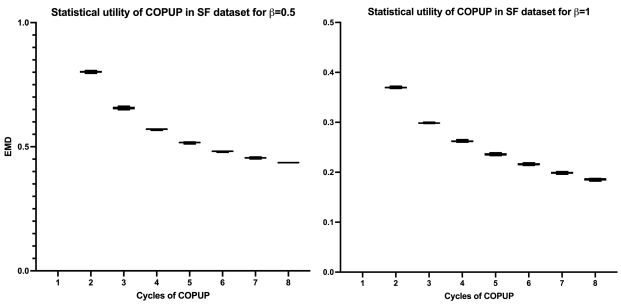
(c) San Francisco; $\beta = 0.5$ **(d) San Francisco; $\beta = 1$**

Figure 7: Visualization of the estimated true distribution of the locations in Paris ((a) and (b)) and San Francisco ((c) and (d)) by PRIVIC after its convergence; the first column is for $\beta = 0.5$ and the second column is for $\beta = 1$.



(a) $\beta = 0.5$ **(b) $\beta = 1$**

Figure 8: (a) and (b) show the EMD between the true PMF of the Paris locations and its estimation by PRIVIC in each of its cycle for $\beta = 0.5$ and $\beta = 1$, respectively.

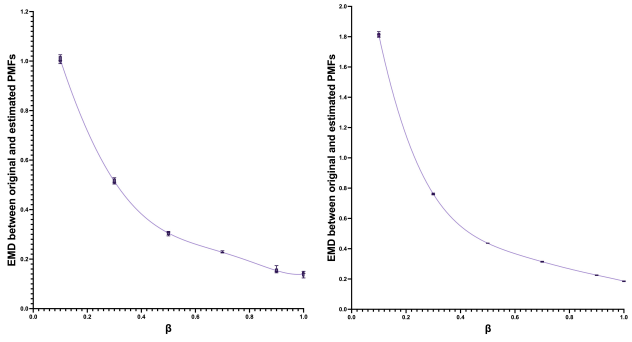


(a) $\beta = 0.5$ **(b) $\beta = 1$**

Figure 9: (a) and (b) show the EMD between the true PMF of the San Francisco locations and its estimation by PRIVIC in each of its cycle for $\beta = 0.5$ and $\beta = 1$, respectively.

Now we shift our attention to analyze the performance of PRIVIC in preserving the statistical utility of the two datasets, and its convergence behaviour. Figure 8 shows us the EMD between the true distribution of the locations in Paris and its estimate by IBU under PRIVIC in each of its 15 cycles under the two settings of privacy (β = 0.5, 1). One of the most crucial observations here is that the EMD between the true and the estimated PMFs seems to decrease with the number of iterations and it finally converges, implying that the estimation of PMFs given by PRIVIC seems to improve at the end of each cycle and, eventually, it converges to the MLE of the prior of the noisy locations, giving the estimate of the true PMF. This, empirically, suggests the convergence of the entire method. This is a major difference from the work of [40] which, as we pointed out before, has the potential of encountering with a LPPM which is optimal according to the standards set by Shokri et al. in [46] but the EM method used to estimate the true distribution would fail to converge for that mechanism as illustrated in Example 1.1. We observe a very similar trend for the San Francisco dataset. Figure 9 shows the statistical utility of the channel generated by PRIVIC under each of its 8 cycles for β = 0.5 and β = 1. The explicit values of the EMD between the true and the estimated PMFs on the location data from Paris and San Francisco for both the settings of the loss parameter can be found in Tables 2 and 3 in Appendix B.

In the next part of the experiments, we set ourselves to dissect the trend of the statistical utility harboured by PRIVIC w.r.t. the level of geo-indistinguishability it guarantees. We recall that the higher the value of β, the lesser the local noise that is injected into the data, and, hence, the worse will be the statistical utility, staying consistent with our observations in Figure 7. We continue working with the location data from Paris and San Francisco obtained from the Gowalla dataset in the same framework as described before. We consider β taking the values 0.1, 0.3, 0.5, 0.7, 0.9, 1, and for each value of the loss parameter, we run PRIVIC on both the datasets, i.e., using the same number of iterations as in the previous experiments. We adhere to 5 rounds of simulation for each β to account for the randomness generated in the obfuscation process.



(a) Paris **(b) San Francisco**

Figure 10: (a) and (b) illustrate that EMD between the true and the estimated distributions of the locations in Paris and San Francisco, respectively, after the empirical convergence of PRIVIC for the different values of the loss parameters β.

Figure 10 shows us that the difference between the true and the estimated PMFs under PRIVIC starts by sharply decreasing and then eventually stabilizes with an increase in the value of the loss parameter. In other words, as the intensity of the local noise decreases, we will end up estimating the unique MLE of the original distribution while optimizing MI and the users' QoS. Both the location datasets result in a Pareto curve showing a similar trend. This depicts an improvement of the estimated PMF until it converges to the true distribution. This observation is complements the Pareto-optimality of MI with the maximum average distortion as studied in *rate-distortion theory* [44], and thus, we empirically weave together the two ends of utility with the information theoretical notion of privacy under PRIVIC.

Discussion. We recall Theorem 6.3 suggesting that the limiting behaviour of the estimated PMF by PRIVIC will follow a unique distribution ψ over $\mathcal{P}_{\mathcal{X},k}$ and ψ can be obtained with the help of Theorem 6.2. Figures 8 and 9 highlight this point by empirically demonstrating the convergence of PRIVIC to a unique PMF across all the rounds of experiments, illustrating that, in fact, the MLE has the highest chance of being estimated by PRIVIC when it converges.

As a justification to the applicability and the working of our method, in a setting where the service providers periodically collect location data from clients, it is reasonable to assume that, over time, they would like to maximize their utility by accurately approximating the true distribution of the population for improving their service in various aspects (crowd management, security enhancement, WLAN hotspot positioning, etc.). BA guarantees geo-indistinguishability, acts as an elastic location-privacy mechanism, and optimizes between MI and the data owners' QoS when it initiates with the true prior. Therefore, as every iteration of PRIVIC improves the estimation of the original distribution, as seen in Figures 3a and 3b, which is used as the starting distribution in its next cycle, the overall privacy protection and its trade-off with QoS of the users will also improve, motivating the users and the service providers comply with PRIVIC to act in their best interests and, in turn, engaging them in a positive feedback loop to maximize the corresponding privacy and utility goals.

8 CONCLUSION

We have bridged some ideas from information theory and statistics to develop a method allowing an incremental collection of location data while protecting the privacy of the data owners, upholding their quality of service, and preserving the statistical utility for the data consumers. Specifically, we have proposed the Blahut-Arimoto algorithm as a location-privacy mechanism, showing its extensive privacy-preserving properties and its other advantages over the state-of-the art Laplace mechanism for geo-indistinguishability. Further, we have exhibited its duality with the iterative Bayesian update and explored this connection to present an iterative method (PRIVIC) for incremental collection of location data with formal guarantees of geo-indistinguishability and an elastic distinguishability metric, while optimizing the QoS of the users and their privacy from an information theoretical perspective. Moreover, PRIVIC allows to efficiently estimate the MLE of the distribution of the original data and, thus, yields a high statistical utility for the service

providers. We have laid out a theoretical framework to characterize the long-term behaviour of our method by modelling it as a Markov chain. Finally, we have illustrated the convergence and the general functioning of PRIVIC with experiments on real location datasets. We believe that our results can be extended easily to other kinds of data, including those with high dimensions, and to other notions of distortion measures since the analysis carried out in this paper does not depend on the notion of distance used.

REFERENCES

[1] Martín Abadi and David G. Andersen. 2016. Learning to Protect Communications with Adversarial Neural Cryptography. *CoRR* abs/1610.06918 (2016). arXiv:1610.06918 http://arxiv.org/abs/1610.06918

[2] Dakshi Agrawal and Charu C. Aggarwal. 2001. On the Design and Quantification of Privacy Preserving Data Mining Algorithms. In *Proceedings of the Twentieth ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems* (Santa Barbara, California, USA) *(PODS '01)*. Association for Computing Machinery, New York, NY, USA, 247–255. https://doi.org/10.1145/375551.375602

[3] Rakesh Agrawal, Ramakrishnan Srikant, and Dilys Thomas. 2005. Privacy preserving OLAP. In *Proceedings of the 2005 ACM SIGMOD international conference on Management of data*. 251–262.

[4] Miguel E. Andrés, Nicolás E. Bordenabe, Konstantinos Chatzikokolakis, and Catuscia Palamidessi. 2013. Geo-Indistinguishability: Differential Privacy for Location-Based Systems. In *Proceedings of the 2013 ACM SIGSAC Conference on Computer &; Communications Security* (Berlin, Germany) *(CCS '13)*. Association for Computing Machinery, New York, NY, USA, 901–914. https://doi.org/10.1145/2508859.2516735

[5] Adeel Anjum, Guillaume Raschia, Marc Gelgon, Abid Khan, Naveed Ahmad, Mansoor Ahmed, Sabah Suhail, M Masoom Alam, et al. 2017. τ-safety: A privacy model for sequential publication with arbitrary updates. *computers & security* 66 (2017), 20–39.

[6] Suguru Arimoto. 1972. An algorithm for computing the capacity of arbitrary discrete memoryless channels. *IEEE Trans. Inf. Theory* 18 (1972), 14–20.

[7] Ugur Ilker Atmaca, Sayan Biswas, Carsten Maple, and Catuscia Palamidessi. 2022. A privacy preserving querying mechanism with high utility for electric vehicles. *arXiv preprint arXiv:2206.02060* (2022).

[8] Richard E. Blahut. 1972. Computation of channel capacity and rate-distortion functions. *IEEE Trans. Inform. Theory* 18 (1972), 460–473.

[9] Nicolás E. Bordenabe, Konstantinos Chatzikokolakis, and Catuscia Palamidessi. 2014. Optimal Geo-Indistinguishable Mechanisms for Location Privacy. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security* (Scottsdale, Arizona, USA) *(CCS '14)*. Association for Computing Machinery, New York, NY, USA, 251–262. https://doi.org/10.1145/2660267.2660345

[10] Justin Brickell and Vitaly Shmatikov. 2008. The Cost of Privacy: Destruction of Data-Mining Utility in Anonymized Data Publishing. In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (Las Vegas, Nevada, USA) *(KDD '08)*. Association for Computing Machinery, New York, NY, USA, 70–78. https://doi.org/10.1145/1401890.1401904

[11] Ji-Won Byun, Tiancheng Li, Elisa Bertino, Ninghui Li, and Yonglak Sohn. 2009. Privacy-Preserving Incremental Data Dissemination. *J. Comput. Secur.* 17, 1 (jan 2009), 43–68.

[12] Ji-Won Byun, Yonglak Sohn, Elisa Bertino, and Ninghui Li. 2006. Secure Anonymization for Incremental Datasets. In *Secure Data Management*, Willem Jonker and Milan Petković (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 48–63.

[13] Konstantinos Chatzikokolakis, Miguel E. Andrés, Nicolás E. Bordenabe, and Catuscia Palamidessi. 2013. Broadening the Scope of Differential Privacy Using Metrics. In *The 13th Privacy Enhancing Technologies Symposium (Lecture Notes in Computer Science, Vol. 7981)*, De Cristofaro, Emiliano, Wright, and Matthew (Eds.). Springer, Bloomington, Indiana, United States, 82–102. https://doi.org/10.1007/978-3-642-39077-7

[14] Konstantinos Chatzikokolakis, Ehab ElSalamouny, and Catuscia Palamidessi. 2017. Efficient Utility Improvement for Location Privacy. *Proceedings on Privacy Enhancing Technologies* 2017, 4 (2017), 308–328. https://doi.org/doi:10.1515/popets-2017-0051

[15] Konstantinos Chatzikokolakis, Catuscia Palamidessi, and Prakash Panangaden. 2008. Anonymity Protocols as Noisy Channels. *Information and Computation* 206, 2–4 (2008), 378–401. https://doi.org/10.1016/j.ic.2007.07.003

[16] Konstantinos Chatzikokolakis, Catuscia Palamidessi, and Marco Stronati. 2015. Constructing elastic distinguishability metrics for location privacy. *Proceedings on Privacy Enhancing Technologies* 2015, 2 (2015), 156–170. https://doi.org/doi:10.1515/popets-2015-0023

[17] Konstantinos Chatzikokolakis, Catuscia Palamidessi, and Marco Stronati. 2015. Constructing elastic distinguishability metrics for location privacy. *Proceedings on Privacy Enhancing Technologies* 2015, 2 (2015), 156–170. https://doi.org/doi:10.1515/popets-2015-0023

[18] Eunjoon Cho, Seth A Myers, and Jure Leskovec. 2011. Friendship and mobility: user movement in location-based social networks. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*. 1082–1090.

[19] Imre Csiszar. 1974. On the computation of rate-distortion functions (Corresp.). *Information Theory, IEEE Transactions on* 20 (02 1974), 122 – 124. https://doi.org/10.1109/TIT.1974.1055146

[20] Paul Cuff and Lanqing Yu. 2016. Differential Privacy As a Mutual Information Constraint. In *Proceedings of the 23rd ACM SIGSAC Conference on Computer and Communications Security (CCS)* (Vienna, Austria) *(CCS '16)*. ACM, New York, NY, USA, 43–54. https://doi.org/10.1145/2976749.2978308

[21] John C. Duchi, Michael I. Jordan, and Martin J. Wainwright. 2013. Local Privacy and Statistical Minimax Rates. In *2013 IEEE 54th Annual Symposium on Foundations of Computer Science*. 429–438. https://doi.org/10.1109/FOCS.2013.53

[22] Cynthia Dwork, Krishnaram Kenthapadi, Frank McSherry, Ilya Mironov, and Moni Naor. 2006. Our Data, Ourselves: Privacy Via Distributed Noise Generation. In *Advances in Cryptology - EUROCRYPT 2006*, Serge Vaudenay (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 486–503.

[23] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. 2006. Calibrating Noise to Sensitivity in Private Data Analysis. In *Theory of Cryptography*, Shai Halevi and Tal Rabin (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 265–284.

[24] Ehab ElSalamouny and Catuscia Palamidessi. 2019. Full Convergence of the Iterative Bayesian Update and Applications to Mechanisms for Privacy Protection. *CoRR* abs/1909.02961 (2019). arXiv:1909.02961 http://arxiv.org/abs/1909.02961

[25] Ehab ElSalamouny and Catuscia Palamidessi. 2020. Generalized Iterative Bayesian Update and Applications to Mechanisms for Privacy Protection. In *2020 IEEE European Symposium on Security and Privacy (EuroS P)*. 490–507. https://doi.org/10.1109/EuroSP48549.2020.00038

[26] Natasha Fernandes, Annabelle McIver, and Carroll Morgan. 2021. The Laplace Mechanism has optimal utility for differential privacy over continuous queries. In *36th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2021*. IEEE, 1–12. https://doi.org/10.1109/LICS52264.2021.9470718

[27] Ari Freedman. 2017. CONVERGENCE THEOREM FOR FINITE MARKOV CHAINS. https://math.uchicago.edu/~may/REU2017/REUPapers/Freedman.pdf

[28] Esther Gal-or. 1982. Hotelling's spatial competition as a model of sales. *Economics Letters* 9, 1 (1982), 1–6. https://doi.org/10.1016/0165-1765(82)90089-1

[29] Filippo Galli, Sayan Biswas, Kangsoo Jung, Catuscia Palamidessi, and Tommaso Cucinotta. 2022. Group privacy for personalized federated learning. *arXiv preprint arXiv:2206.03396* (2022).

[30] Arpita Ghosh, Tim Roughgarden, and Mukund Sundararajan. 2012. Universally Utility-maximizing Privacy Mechanisms. *SIAM J. Comput.* 41, 6 (2012), 1673–1693. https://doi.org/10.1137/09076828X arXiv:https://doi.org/10.1137/09076828X

[31] Gowalla [n. d.]. The Gowalla dataset. [Online]. https://snap.stanford.edu/data/loc-gowalla.html. (Accessed on 10/08/2021).

[32] Mangesh Gupte and Mukund Sundararajan. 2010. Universally Optimal Privacy Mechanisms for Minimax Agents. In *Proceedings of the Twenty-Ninth ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems* (Indianapolis, Indiana, USA) *(PODS '10)*. Association for Computing Machinery, New York, NY, USA, 135–146. https://doi.org/10.1145/1807085.1807105

[33] Mehmet Emre Gursoy, Acar Tamersoy, Stacey Truex, Wenqi Wei, and Ling Liu. 2021. Secure and Utility-Aware Data Collection with Condensed Local Differential Privacy. *IEEE Transactions on Dependable and Secure Computing* 18, 5 (2021), 2365–2378. https://doi.org/10.1109/TDSC.2019.2949041

[34] Chong Huang, Peter Kairouz, Xiao Chen, Lalitha Sankar, and Ram Rajagopal. 2017. Context-aware generative adversarial privacy. *Entropy* 19, 12 (1 12 2017). https://doi.org/10.3390/e19120656

[35] Stratis Ioannidis, Andrea Montanari, Udi Weinsberg, Smriti Bhagat, Nadia Fawaz, and Nina Taft. 2014. Privacy Tradeoffs in Predictive Analytics. In *The 2014 ACM International Conference on Measurement and Modeling of Computer Systems* (Austin, Texas, USA) *(SIGMETRICS '14)*. Association for Computing Machinery, New York, NY, USA, 57–69. https://doi.org/10.1145/2591971.2592011

[36] L. V. Kantorovich. 1960. *Mathematical Methods of Organizing and Planning Production*. Vol. 6. INFORMS. Issue 4. https://doi.org/10.1287/mnsc.6.4.366

[37] Chao Li, Michael Hay, Vibhor Rastogi, Gerome Miklau, and Andrew McGregor. 2010. Optimizing Linear Counting Queries under Differential Privacy. In *Proceedings of the Twenty-Ninth ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems* (Indianapolis, Indiana, USA) *(PODS '10)*. Association for Computing Machinery, New York, NY, USA, 123–134. https://doi.org/10.1145/1807085.1807104

[38] John Nash. 1951. Non-Cooperative Games. *Annals of Mathematics* 54, 2 (1951), 286–295. http://www.jstor.org/stable/1969529

[39] Simon Oya, Carmela Troncoso, and Fernando Pérez-González. 2017. Back to the Drawing Board: Revisiting the Design of Optimal Location Privacy-preserving Mechanisms. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security* (Dallas, Texas, USA). ACM, 1959–1972. https:

//doi.org/10.1145/3133956.3134004
[40] Simon Oya, Carmela Troncoso, and Fernando Pérez-González. 2019. Rethinking Location Privacy for Unknown Mobility Behaviors. In *2019 IEEE European Symposium on Security and Privacy (EuroS P)*. 416–431. https://doi.org/10.1109/EuroSP.2019.00038
[41] Hari Palaiyanur and Anant Sahai. 2008. On the uniform continuity of the rate-distortion function. In *2008 IEEE International Symposium on Information Theory*. 857–861. https://doi.org/10.1109/ISIT.2008.4595108
[42] Marco Romanelli, Kostantinos Chatzikokolakis, and Catuscia Palamidessi. 2020. Optimal Obfuscation Mechanisms via Machine Learning. In *2020 IEEE 33rd Computer Security Foundations Symposium (CSF)*. 153–168. https://doi.org/10.1109/CSF49147.2020.00019
[43] Richard Serfozo. 2009. *Basics of applied stochastic processes*. Springer Science & Business Media.
[44] C. E. Shannon. 1948. A mathematical theory of communication. *The Bell System Technical Journal* 27, 3 (1948), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x
[45] Reza Shokri. 2015. Privacy Games: Optimal User-Centric Data Obfuscation. *Proceedings on Privacy Enhancing Technologies* 2015, 2 (2015), 299–315. https://doi.org/doi:10.1515/popets-2015-0024
[46] Reza Shokri, George Theodorakopoulos, George Danezis, Jean-Pierre Hubaux, and Jean-Yves Le Boudec. 2011. Quantifying Location Privacy: The Case of Sporadic Location Exposure. In *Privacy Enhancing Technologies*, Simone Fischer-Hübner and Nicholas Hopper (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 57–76.
[47] Reza Shokri, George Theodorakopoulos, Carmela Troncoso, Jean-Pierre Hubaux, and Jean-Yves Le Boudec. 2012. Protecting Location Privacy: Optimal Strategy against Localization Attacks. In *Proceedings of the 2012 ACM Conference on Computer and Communications Security* (Raleigh, North Carolina, USA) *(CCS '12)*. Association for Computing Machinery, New York, NY, USA, 617–627. https://doi.org/10.1145/2382196.2382261
[48] Ardhendu Tripathy, Ye Wang, and Prakash Ishwar. 2019. Privacy-Preserving Adversarial Networks. In *57th Annual Allerton Conference on Communication, Control, and Computing, Allerton 2019, Monticello, IL, USA, September 24-27, 2019*. IEEE, 495–505. https://doi.org/10.1109/ALLERTON.2019.8919758
[49] Yue Wang, Xintao Wu, and Donghui Hu. 2016. Using Randomized Response for Differential Privacy Preserving Data Collection.. In *EDBT/ICDT Workshops*, Vol. 1558. 0090–6778.
[50] Wikipedia contributors. 2022. Perron–Frobenius theorem — Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/w/index.php?title=Perron%E2%80%93Frobenius_theorem&oldid=1126682156. [Online; accessed 2-January-2023].
[51] Wenjing Zhang, Ming Li, Ravi Tandon, and Hui Li. 2019. Online Location Trace Privacy: An Information Theoretic Approach. *IEEE Transactions on Information Forensics and Security* 14, 1 (2019), 235–250. https://doi.org/10.1109/TIFS.2018.2848659
[52] Ye Zhu and Riccardo Bettati. 2005. Anonymity vs. Information Leakage in Anonymity Systems. In *Proc. of ICDCS*. IEEE Computer Society, 514–524.
[53] Úlfar Erlingsson, Vasyl Pihur, and Aleksandra Korolova. 2014. RAPPOR: Randomized Aggregatable Privacy-Preserving Ordinal Response. In *Proceedings of the 21st ACM Conference on Computer and Communications Security*. Scottsdale, Arizona. https://arxiv.org/abs/1407.6981

A PROOFS

Theorem 3.1. The privacy channel generated by BA produces an elastic location-privacy mechanism.

Proof. Let $x, y \in \mathcal{X}$ be any true and reported location, respectively. Letting λ to be the limiting channel generated by BA, to show that λ is possesses an elastic distinguishability metric, we need to ensure that:

(1) The probability of reporting y to obfuscate x given by λ should be exponentially reducing w.r.t. the Euclidean distance between x and y, staying consistent with the essence of geo-indistinguishability (the property captured by (3)).
(2) Under λ, the probability of reporting y to obfuscate x should be taking into account the mass around y, i.e., the more geospatially isolated y is in the space, the less likely it should be to report it, as, ideally, we would like to have x being reported amidst a crowd of other locations (the property captured by (4)).

Let's simplify the notation and denote $\mathbb{P}[\lambda(x) = y]$ as $\mathbb{P}_\lambda[y|x]$ and let $q(y)$ be the probability mass of the observed location y. Hence, for being an elastic location-privacy mechanism, λ should satisfy (3) and (4), i.e., we must have:

$$\mathbb{P}_\lambda[y|x] \propto \exp\{-\beta d_{\mathrm{E}}(x, y)\} \tag{6}$$

$$\mathbb{P}_\lambda[y|x] \propto q(y) \tag{7}$$

where $q(y)$ denotes the mass of y.

Therefore, in order to satisfy (6) and (7), it is sufficient to have:

$$\begin{aligned} &\mathbb{P}_\lambda[y|x] \propto \exp\{-\beta d_{\mathrm{E}}(x, y) + \ln q(y)\} \\ \implies &\mathbb{P}_\lambda[y|x] \propto q(y) \exp\{-\beta d_{\mathrm{E}}(x, y)\} \\ &= \frac{q(y) \exp\{\beta d_{\mathrm{E}}(x, y)\}}{\sum_{z \in \mathcal{X}} q(z) \exp\{-\beta d_{\mathrm{E}}(x, z)\}} \end{aligned} \tag{8}$$

Now, it's sufficient to note that that, if we interpret the mass of y as probability of beeing reported by the mechanism, (8) is exactly the fixpoint of $\mathcal{G} \circ \mathcal{F}$, cf. Remarks 1 and 2. □

Theorem 5.1. Implementing BA with a starting PMF θ_0 on $\mathcal{X}$ and a uniform initial channel $C^{(0)}$, there is a unique MLE for the prior for a set of observed locations on $\mathcal{X}$ obfuscated with $\lambda_{\mathrm{BA}}\left(\theta_0, C^{(0)}\right)$.

Proof. Note that the limiting channel generated by BA, $\lambda\left(\theta_0, C^{(0)}\right)$, is of the form $\hat{C}_{xy} = \frac{\beta(y)}{\alpha(x)} \exp\{-\beta d_{\mathrm{E}}(x, y)\}$ for some constants $\alpha > 0$ and $\beta \geq 0$ which just depend on x and y, respectively.

Let the empirical PMF of the observed locations be $\boldsymbol{q} = \{\boldsymbol{q}(x) : x \in \mathcal{X}\}$. Therefore, the t^{th} iterative step of IBU would look like

$$\begin{aligned} \theta_{t+1}(x) &= \sum_{y \in \mathcal{X}} q(y) \frac{\theta_t(x) \hat{C}_{xy}}{\sum_{z \in \mathcal{X}} \theta_t(z) \hat{C}_{zy}} \\ &= \sum_{y \in \mathcal{X}} q(y) \frac{\theta_t(x) \alpha_x^{-1} \exp\{-\beta d_{\mathrm{E}}(x, y)\}}{\sum_{z \in \mathcal{X}} \theta_t(z) \alpha_z^{-1} \exp\{-\beta d_{\mathrm{E}}(z, y)\}} \end{aligned} \tag{9}$$

(9) suggests that this iterative step of IBU is essentially the same as having a mechanism C such that $C_{xy} = \alpha'(x) \exp\{-\beta d_{\mathrm{E}}(x, y)\}$ for some normalising positive constant $\alpha'(x)$ depending on x. The rest of the proof proceeds identically as the proof of Corollary 3 in [24] by ElSalamouny et al.

In particular, in the proof of Corollary 3, ElSalamouny et al. exploited the fact that linear independence of the columns of $G(\mathcal{I}')$, defined in the proof of Corollary 2, is preserved even when each column of $G(\mathcal{I}')$ is scaled by a positive constant. We note that the $G(\mathcal{I}')$ that has been defined is symmetric, i.e., $G(\mathcal{I}')^T = G(\mathcal{I}')$, i.e., the as columns of $G(\mathcal{I})$ are linearly independent, so are the rows.

We observe from (9) that, in our case, each row of $G(\mathcal{I})$ gets scaled by a positive factor, proceeding with the same idea of $G(\mathcal{I})$ as used in the proof of Corollary 2 of [24]. Or in other words, we could think of having each column scaled by a positive factor by considering $G(\mathcal{I})^T$. As the scaling of the columns preserves the linear independence of the columns of $G(\mathcal{I})^T$ (or the rows of $G(\mathcal{I})$), we can conclude that the rows of $G(\mathcal{I})$ scaled by a positive factor

preserve the linear independence of the rows of the unscaled $G(\mathcal{I})$ devised for the simple geometric mechanism by ElSalamouny et al.

Thus, we have $\boldsymbol{v}G'(\mathcal{I}) = \boldsymbol{v}G'(\mathcal{I})^T = 0$ iff $\boldsymbol{v} = 0$, where $G'(\mathcal{I})$ is the matrix where the x^{th} row of $G(\mathcal{I})$ is scaled by $\alpha^{-1}(x)$. This reduces us to satisfy the sufficient condition for having a unique MLE given by Theorem 7 of [24] by the exact same line of idea for the proofs of Corollary 2 and Corollary 3 by ElSalamouny et al. Thus, by Theorem 7 of [24] we can conclude that there is a unique MLE for a set of observed locations which are obfuscated by $\hat{C}$. □

Lemma 6.1. PRIVIC is an irreducible and aperiodic Markov chain.

Proof. **Irreducible:** Let Φ be the transition matrix of PRIVIC seen as a Markov chain over $\mathcal{P}_{\mathcal{X},k}$. In a PRIVIC cycle, we implement BA with a full-support PMF $\theta \in \mathcal{P}_{\mathcal{X},k}$ and the uniform channel $C^{(0)}$ and we converge to an optimal channel $\hat{C}$ given by $\lambda_{\text{BA}}\left(\theta, C^{(0)}\right)$.

We observe from (2) that $\hat{C}_{xy} > 0$ for every $x, y \in \mathcal{X}$. This essentially implies that any given location in $\mathcal{X}$ can be obfuscated to any other location $y \in \mathcal{X}$. In other words, from a given set of input (original) locations $X = \{\hat{x}_1, \ldots, \hat{x}_n\}$, any set of output (noisy) locations $Y = \{y_1, \ldots, y_n\}$ could be reached via the location-privacy channel $\hat{C}$. In particular, with $\hat{C}$, for any input of location samples X, we could give rise to an empirical distribution, $\boldsymbol{q}$, of the observed locations, which, in turn, can harbour any MLE $\theta' \in \mathcal{P}_{\mathcal{X},k}$.

Looking at it alternatively, for a given set of original locations, X, for any target PMF $\theta' \in \mathcal{P}_{\mathcal{X},k}$, we have a non-zero probability of fostering a set of noisy locations, Y, produced by locally obfuscating the locations in X with the channel $\hat{C}$ obtained with $\lambda_{\text{BA}}\left(\theta, C^{(0)}\right)$, which would induce θ' as its MLE under $\hat{C}$ where IBY will converge to at the end of this cycle of PRIVIC. Therefore, $\Phi_{\theta,\theta'} = \phi(\theta|\theta') > 0$ for every $\theta, \theta' \in \mathcal{P}_{\mathcal{X},k}$, implying that, in this setting, PRIVIC induces an irreducible Markov chain.

Aperiodic: The follows very similarly as the previous part. Let Φ be the transition matrix of PRIVIC seen as a Markov chain over $\mathcal{P}_{\mathcal{X},k}$. We showed in the proof of irreducibility that $\Phi_{\theta,\theta'} > 0$ for all $\theta, \theta' \in \mathcal{P}_{\mathcal{X},k}$.

In particular, for every $\theta \in \mathcal{P}_{\mathcal{X},k}$ and for a channel $\hat{C}$ produced by $\lambda_{\text{BA}}\left(\theta, C^{(0)}\right)$, where $C^{(0)}$ is the uniform channel, we can obfuscate any set of true locations X to produce the noisy locations Y which would give rise to the output distribution $\boldsymbol{q}$ over $\mathcal{X}$, empirically computed on the observed set locations Y, such that $\boldsymbol{q}(y) = \sum_{z\in\mathcal{X}} \theta(x)\hat{C}_{xy}$ with a non-zero probability.

A very interesting and beautiful consequence of this is that it would ensure that there is some positive probability that the MLE of the observed locations privatized with $\hat{C}$ is the "starting guess" itself, which was fed into IBU, i.e, $\mathbb{P}(\Gamma_{\text{IBU}}(\theta, \hat{C}) = \theta) > 0$ for every $\theta \in \mathcal{P}_{\mathcal{X},k}$. This shows that the period for every state $\theta \in \mathcal{P}_{\mathcal{X},k}$ in the Markov chain is 1. Therefore, PRIVIC gives rise to an aperiodic Markov chain. □

B TABLES

Table 2: EMD between the true and the estimated PMFs by PRIVIC on the Paris locations.

N	$\beta = 1$					$\beta = 0.5$				
	Round 1	Round 2	Round 3	Round 4	Round 5	Round 1	Round 2	Round 3	Round 4	Round 5
1	2.02262	2.02262	2.02262	2.02262	2.02262	2.02262	2.02262	2.02262	2.02262	2.02262
2	0.27104	0.27796	0.28247	0.27758	0.26276	0.57738	0.57994	0.55791	0.60717	0.56880
3	0.21916	0.23750	0.25035	0.25116	0.23241	0.51324	0.48295	0.47043	0.51285	0.45826
4	0.19156	0.21115	0.21726	0.20913	0.20408	0.43184	0.42398	0.41040	0.45230	0.41119
5	0.18241	0.19264	0.19570	0.19747	0.18728	0.39741	0.38771	0.37284	0.41176	0.36953
6	0.16526	0.18020	0.174578	0.17310	0.17268	0.37818	0.35482	0.36039	0.36375	0.38045
7	0.14643	0.18159	0.16092	0.17222	0.17139	0.35044	0.34383	0.34818	0.35760	0.36804
8	0.13860	0.17605	0.15938	0.17078	0.16192	0.34086	0.32983	0.35769	0.34430	0.34780
9	0.15047	0.16926	0.15153	0.17005	0.15266	0.33137	0.31749	0.34690	0.33840	0.34028
10	0.14734	0.14825	0.14585	0.15001	0.15459	0.3170	0.32975	0.32772	0.34529	0.32670
11	0.14227	0.14135	0.14326	0.13797	0.14507	0.31917	0.31851	0.32689	0.33667	0.32043
12	0.13818	0.14448	0.14703	0.13589	0.14142	0.32137	0.32451	0.31843	0.32556	0.34014
13	0.16111	0.13641	0.14893	0.14208	0.13808	0.31224	0.31219	0.31380	0.31515	0.31159
14	0.13894	0.13111	0.14094	0.14199	0.14192	0.30883	0.30496	0.30578	0.30726	0.30282
15	0.15106	0.14271	0.14601	0.13584	0.12413	0.29405	0.31198	0.30786	0.31167	0.30100

Table 3: EMD between the true and the estimated PMFs by PRIVIC on the San Francisco locations.

N	$\beta = 1$					$\beta = 0.5$				
	Round 1	Round 2	Round 3	Round 4	Round 5	Round 1	Round 2	Round 3	Round 4	Round 5
1	7.37595	7.37595	7.37595	7.37595	7.37595	7.37595	7.37595	7.37595	7.37595	7.37595
2	0.37229	0.37038	0.36784	0.36949	0.36816	0.79621	0.80474	0.80219	0.80670	0.79940
3	0.29828	0.298370	0.30017	0.29784	0.29859	0.64931	0.66292	0.65362	0.65950	0.65285
4	0.26091	0.26029	0.26231	0.26180	0.26518	0.56896	0.57378	0.57338	0.57125	0.56618
5	0.23472	0.23419	0.23337	0.23740	0.23897	0.51672	0.51710	0.51735	0.51880	0.51138
6	0.21367	0.21432	0.21537	0.21777	0.21881	0.48194	0.48299	0.47992	0.48267	0.47791
7	0.19612	0.19761	0.19904	0.20120	0.20067	0.45531	0.45861	0.45122	0.45732	0.45450
8	0.18244	0.18412	0.18741	0.18724	0.18674	0.43588	0.43745	0.43558	0.43704	0.43584

C FURTHER MATHEMATICAL ANALYSIS

Here we delve deeper into the mathematical intricacies of our proposed method.

C.1 Properties of BA under PRIVIC

In [41], Palaiyanur et al. studied the uniform continuity of the rate-distortion function. In order to elaborate on the analytical behaviour BA under PRIVIC, we shall aim to take advantage of the results in [41]. In particular, we note that using any distortion metric d on $\mathcal{X}$, for every location $x \in \mathcal{X}$, we have $d(x, \hat{x}) = 0$ for any $\hat{x} = x$. Therefore, we satisfy the *Condition (Z)* proposed in Section II of [41]. With this, we immediately place ourselves in the environment to profit from Lemma 2 of [41]. In order to compare the estimated distributions at the successive and/or intermediate steps of BA and IBU within PRIVIC, we shall use the *total variation* distance, d_{TV}, and the L^1 norm, d_{L^1}. Note that for any pair of $\theta, \theta' \in \Pi_{\mathcal{X}}$, we have $d_{\text{TV}}(\theta, \theta') = \frac{1}{2} d_{L^1}(\theta, \theta')$.

As $\mathcal{X}$ is finite, we can set $\Delta < \infty$ to be the maximum distortion on $\mathcal{X}$, i.e., $\Delta = \max_{x,y\in\mathcal{X}} d(x, y)$. For the subsequent analysis, let us fix a maximum allowance for the average distortion $d^* \in [0, \Delta]$ under a chosen distortion metric d. Let us, moreover, define the minimum possible non-zero distortion on $\mathcal{X}$ as $\tilde{d} = \min\{d(x, x') : x, x' \in \mathcal{X}, d(x, x') > 0\}$. Then we have the following result.

Lemma C.1. Let $\theta_1, \theta_2 \in \Pi_{\mathcal{X}}$ such that $d_{L^1}(\theta, \theta') < \frac{\tilde{d}}{4\Delta}$. Let X_θ and $X_{\theta'}$ be random variables on $\mathcal{X}$ with PMFs θ and θ', respectively. Then we have:

$$\left|RD(X_\theta, d^*) - RD(X_{\theta'}, d^*)\right| \leq \frac{7\Delta}{\tilde{d}} d_{L^1}(\theta, \theta') \ln \frac{|\mathcal{X}|^2}{d_{L^1}(\theta, \theta')} \tag{10}$$

Proof. Immediate from Lemma 2 of [41]. □

Lemma C.2. Let $A,\ B \in \mathbb{R}^+$ be positive constants. Let $f : \mathbb{R}^+ \mapsto \mathbb{R}^+$ be such that $f(x) = Ax \ln \frac{B}{x}$. Then

$$f^{-1}(x) = \frac{B}{\exp\left\{-W_r\left(-\frac{x}{AB}\right)\right\}}$$

where W_r is *Lambert W function* of some integer order r.

Proof.

$$\begin{aligned}
y = Ax \ln \frac{B}{x} &\implies \frac{y}{A} = x \ln \frac{B}{x} \\
&\implies y' = \frac{\ln x'}{x'} \text{ [where } x' = \frac{B}{x},\ y' = \frac{y}{AB}\text{]} \\
&\implies y'x' = \ln x' \implies y' \exp\{\ln x'\} = \ln x' \\
&\implies y' = \ln x' \exp\{-\ln x'\} \\
&\implies -y' = z \exp\{z\} \text{ [where } z = -\ln x' = -\ln \frac{B}{x}\text{]} \\
&\implies z = W_r(-y') = W_r\left(-\frac{y}{AB}\right) \\
&\implies -\ln \frac{B}{x} = W_r\left(-\frac{y}{AB}\right) \\
&\implies \frac{B}{x} = \exp\left\{-W_r\left(-\frac{y}{AB}\right)\right\} \\
&f^{-1}(x) = \frac{B}{\exp\left\{-W_r\left(-\frac{x}{AB}\right)\right\}}
\end{aligned}$$

□

Definition C.1 (Rate-distortion function [44]). Let $\mathcal{X}$ and $\mathcal{Y}$ be a pair of discrete spaces and $\mathcal{C}(\mathcal{X}, \mathcal{Y})$ be the space of all channels encoding from $\mathcal{X}$ to $\mathcal{Y}$. Suppose X is a random variable (r.v.) in $\mathcal{X}$. Then the *rate-distortion function (RD)* for an r.v. $X \in \mathcal{X}$ and d^* under the distortion metric d is defined as:

$$RD(X, d^*) = \min_{\substack{C \in \mathcal{C}(\mathcal{X},\mathcal{Y}) \\ AvgD(X,C,d) \le d^*}} I(X|Y_{X,C})$$

where $Y_{X,C}$ is the r.v. on $\mathcal{Y}$ that denotes the output of the encoding of X with any $C \in \mathcal{C}(\mathcal{X}, \mathcal{Y})$.

Corollary C.2.1. Setting $A = \frac{7\Delta}{\tilde{d}}$ and $B = |\mathcal{X}|^2$, for $\epsilon > 0$ satisfying $\frac{B}{\exp\{-W_r(-\frac{\epsilon}{BA})\}} < \frac{\tilde{d}}{4\Delta}$, there exists $\delta > 0$ such that for all $\theta, \theta' \in \Pi_{\mathcal{X}}$,

$$d_{L^1}(\theta, \theta') < \delta \implies \left|RD(X_\theta, d^*) - RD(X_{\theta'}, d^*)\right| \le \epsilon$$

where X_θ and $X_{\theta'}$ are random variables on $\mathcal{X}$ with PMFs θ and θ', respectively.

Proof. Setting $\delta = \frac{B}{\exp\{-W_r(-\frac{\epsilon}{AB})\}}$, the result follows immediately from Lemma C.1 and Lemma C.2. □

Essentially, Corollary C.2.1 asserts that the rate-distortion function is uniformly continuous if we only focus on the "small jumps" of the rate-distortion function.

Remark 5. For a starting PMF and an initial channel, we know BA converges to give the channel that estimates the unique RD function [19]. We can, therefore, comment that the uniform continuity of RD function implies the uniform continuity of BA under the same condition of ϵ (jump) as in Corollary C.2.1.

C.2 Properties of IBU under PRIVIC

For the convenience of notation, let us extend the functional representation of IBU Γ_{IBU}, as introduced in Definition 2.5, as follows. For any privacy channel C over $\mathcal{X}$ and starting with any full-support $\theta \in \Pi_{\mathcal{X}}$, let $\Gamma^t_{\text{IBU}}(\theta, C)$ denote the t^{th} iteration of IBU for all $t \in \mathbb{N}$. Therefore, in one cycle of PRIVIC, $\Gamma_{\text{IBU}}(\theta, C) = \Gamma^{N_{\text{IBU}}}_{\text{IBU}}(\theta, C)$.

Lemma C.3. For any privacy channel C over $\mathcal{X}$, having a sufficiently large number of samples implies that every step of IBU is probabilistically uniformly continuous w.r.t. d_{TV}, i.e., as $n \to \infty$, for all $\epsilon > 0$, there exists $\delta > 0$ w.p. 1 such that for all $\theta, \theta' \in \Pi_{\mathcal{X}}$:

$$d_{\text{TV}}(\theta, \theta') < \delta \implies d_{\text{TV}}\left(\Gamma^1_{\text{IBU}}(\theta, C), \Gamma^1_{\text{IBU}}(\theta', C)\right) < \epsilon$$

Proof. Let $\epsilon > 0$.

$$\begin{aligned}
&d_{\text{TV}}\left(\Gamma^1_{\text{IBU}}(\theta, C), \Gamma^1_{\text{IBU}}(\theta', C)\right) \\
&= \frac{1}{2} \sum_{x \in \mathcal{X}} \left|\Gamma^1_{\text{IBU}}(\theta, C)(x) - \Gamma^1_{\text{IBU}}(\theta', C)(x)\right| \qquad (11)
\end{aligned}$$

To prove the result, it is enough to show the uniform continuity of the steps of IBU w.r.t L^1 norm, as $\frac{1}{2} d_{L^1} = d_{\text{TV}}$, as we can always scale δ by half and reduce to the total variation distance. Therefore, (11), under the L^1 norm, reduces to:

$$\sum_{x \in \mathcal{X}} \left| \sum_{y \in \mathcal{X}} q^{(1)}_y \frac{C_{xy}\theta(x)}{\sum_{z \in \mathcal{X}} C_{zy}\theta(z)} - \sum_{y \in \mathcal{X}} q^{(2)}_y \frac{C_{xy}\theta'(x)}{\sum_{z \in \mathcal{X}} C_{zy}\theta'(z)} \right| \qquad (12)$$

Here $q^{(1)}$ and $q^{(2)}$ are the corresponding empirical PMFs generated by the observed locations running IBU starting with θ and θ', respectively. Note that the black-box sampling and obfuscation of the locations that, in turn, give rise to $q^{(1)}$ and $q^{(2)}$ do not depend on θ and θ' respectively. In particular, since both cases of obfuscating the real locations are done by the same channel, as $n \to \infty$, we can assume $\mathbb{P}\left[q^{(1)} = q^{(2)}\right] \to 1$. Let $q = \lim_{n\to\infty} q^{(1)} = \lim_{n\to\infty} q^{(2)}$. Thus, we get an upper bound for (12) w.p. 1 in the form of:

$$\begin{aligned}
&\le \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{X}} \left| q_y C_{xy} \left(\frac{\theta(x)}{\sum_{z \in \mathcal{X}} C_{zy}\theta(z)} - \frac{\theta'(x)}{\sum_{z \in \mathcal{X}} C_{zy}\theta'(z)} \right) \right| \\
&\qquad [\because \text{Triangle Inequality}] \\
&= \sum_{y \in \mathcal{X}} q_y \sum_{x \in \mathcal{X}} C_{xy} \left| \frac{\theta(x)}{\sum_{z \in \mathcal{X}} C_{zy}\theta(z)} - \frac{\theta'(x)}{\sum_{z \in \mathcal{X}} C_{zy}\theta'(z)} \right| \\
&\qquad [\because q_y \ge 0, C_{xy} \ge 0 \,\forall\, x, y] \\
&= \sum_{y \in \mathcal{X}} q_y \sum_{x \in \mathcal{X}} C_{xy} \left| \frac{\theta(x) \sum_{z \in \mathcal{X}} C_{zy}\theta'(z) - \theta'(x) \sum_{z \in \mathcal{X}} C_{zy}\theta(z)}{(\sum_{z \in \mathcal{X}} C_{zy}\theta(z))(\sum_{z \in \mathcal{X}} C_{zy}\theta'(z))} \right| \\
&= \sum_{y \in \mathcal{X}} K_y \sum_{x \in \mathcal{X}} C_{xy} \left| \sum_{z \in \mathcal{X}} C_{zy}(\theta(x)\theta'(z) - \theta'(x)\theta(z)) \right| \\
&\qquad [\because K_y = \frac{q_y}{(\sum_{z \in \mathcal{X}} C_{zy}\theta(z))(\sum_{z \in \mathcal{X}} C_{zy}\theta'(z))}] \\
&\le \sum_{y \in \mathcal{X}} K_y \sum_{x \in \mathcal{X}} C_{xy} \sum_{z \in \mathcal{X}} C_{zy} \left|\theta(x)\theta'(z) - \theta'(x)\theta(z)\right|
\end{aligned}$$

$$
\begin{aligned}
&[\because \text{Triangle Inequality; } C_{zy} \geq 0 \,\forall\, z, y] \\
&= \sum_{y\in\mathcal{X}} K_y \sum_{x\in\mathcal{X}} C_{xy} \sum_{z\in\mathcal{X}} C_{zy} \big|\theta(x)\theta'(z) \\
&\qquad -\theta(x)\theta(z) + \theta(x)\theta(z) - \theta'(x)\theta(z)\big| \\
&= \sum_{y\in\mathcal{X}} K_y \sum_{x\in\mathcal{X}} C_{xy} \sum_{z\in\mathcal{X}} C_{zy} \big|\theta(x)(\theta'(z) \\
&\qquad -\theta(z)) + \theta(z)(\theta(x) - \theta'(x))\big| \\
&\leq \sum_{y\in\mathcal{X}} K_y \sum_{x\in\mathcal{X}} C_{xy} \sum_{z\in\mathcal{X}} C_{zy} \big|\theta(x)(\theta'(z) - \theta(z))\big| \\
&+ \sum_{y\in\mathcal{X}} K_y \sum_{x\in\mathcal{X}} C_{xy} \sum_{z\in\mathcal{X}} C_{zy} \big|\theta(z)(\theta(x) - \theta'(x))\big| \\
&< M^2 \sum_{y\in\mathcal{X}} K_y \sum_{x\in\mathcal{X}} \sum_{z\in\mathcal{X}} \big|\theta(x)(\theta'(z) - \theta(z))\big| \\
&\quad + M^2 \sum_{y\in\mathcal{X}} K_y \sum_{x\in\mathcal{X}} \sum_{z\in\mathcal{X}} \big|\theta(z)(\theta(x) - \theta'(x))\big| \\
&\qquad [\text{Letting } M = \max_{x,y\in\mathcal{X}} C_{xy}] \\
&= M^2 \sum_{y\in\mathcal{X}} K_y \sum_{x\in\mathcal{X}} \theta(x) \sum_{z\in\mathcal{X}} \big|(\theta'(z) - \theta(z))\big| \\
&\quad + M^2 \sum_{y\in\mathcal{X}} K_y \sum_{x\in\mathcal{X}} \big|(\theta(x) - \theta'(x))\big| \sum_{z\in\mathcal{X}} \theta(z) \qquad (13)
\end{aligned}
$$

Letting $d_{L^1}(\theta, \theta') < \delta$, and exploiting the fact that θ is a probability distribution on $\mathcal{X}$, (13) is bounded above by:

$$
M^2 \sum_{y\in\mathcal{X}} K_y \sum_{x\in\mathcal{X}} \theta(x)\delta + M^2 \sum_{y\in\mathcal{X}} K_y \delta \sum_{z\in\mathcal{X}} \theta(z) = 2M^2\delta \sum_{y\in\mathcal{X}} K_y \qquad (14)
$$

If the r^{th} column of C is 0, it would result in a probability of 0 for observing $x_r \in \mathcal{X}$ in the output, which effectively will have no influence on determining the MLE and, therefore, on IBU. Hence, we can ignore the 0-columns. Therefore, we can safely assume $\min_{y\in\mathcal{X}} \sum_{z\in\mathcal{X}} C_{zy}\theta(z) > 0$ for any $\theta \in \Pi_{\mathcal{X}}$. In fact, with n observations, n being large enough, we can reasonably assume $\min_{y\in\mathcal{X}} \sum_{z\in\mathcal{X}} C_{zy}\theta(z) \geq \frac{1}{n}$ for any $\theta \in \Pi_{\mathcal{X}}$. Therefore, $\sum_{y\in\mathcal{X}} K_y \leq \sum_{y\in\mathcal{X}} q_y\, n^2 = n^2$. Hence, using (14), $d_{L^1}(\theta, \theta') < \delta$, or, $d_{\text{TV}}(\theta, \theta') < \delta/2$ implies:

$$
\begin{aligned}
& d_{L^1}\left(\Gamma^1_{\text{IBU}}(\theta, C), \Gamma^1_{\text{IBU}}(\theta', C)\right) < 2M^2\delta n^2 \\
\implies & d_{\text{TV}}\left(\Gamma^1_{\text{IBU}}(\theta, C), \Gamma^1_{\text{IBU}}(\theta', C)\right) < M^2\delta n^2 \qquad (15)
\end{aligned}
$$

So for the given ϵ, if we choose $\delta = \frac{\epsilon}{M^2 n^2}$, we are done. □

As an immediate corollary of Lemma C.3, we can elevate the uniform continuity of a single step of IBU to the entire method and formally assert the following.

Theorem C.4. For a privacy channel $\hat{C}$ derived with BA, defined over $\mathcal{X}$, if we have a sufficiently large number of samples, IBU is a probabistically uniformly continuous transformation w.r.t. d_{TV}, i.e., as $n \to \infty$, for all $\epsilon > 0$, there exists a $\delta > 0$ w.p. 1 such that for every $\theta, \theta' \in \Pi_{\mathcal{X}}$:

$$
d_{\text{TV}}(\theta, \theta') < \delta \implies d_{\text{TV}}\left(\Gamma_{\text{IBU}}\left(\theta, \hat{C}\right), \Gamma_{\text{IBU}}\left(\theta', \hat{C}\right)\right) < \epsilon
$$

Proof. Let $\epsilon > 0$. As $\hat{C}$ is derived from BA, by Theorem 5.1, starting with any PMF in $\Pi_{\mathcal{X}}$ and an observed set of noisy locations, IBU converges uniquely to the most likely prior under $\hat{C}$ which is denoted by $\Gamma_{\text{IBU}}\left(\theta, \hat{C}\right)$. Running two IBUs starting with $\theta \in \Pi_{\mathcal{X}}$ and $\theta' \in \Pi_{\mathcal{X}}$, respectively, let $N_{\text{IBU}} \in \mathbb{N}$ be the number of iterations after which they numerically converge to, or are within a desirable threshold of $\Gamma_{\text{IBU}}\left(\theta, \hat{C}\right)$ and $\Gamma_{\text{IBU}}\left(\theta', \hat{C}\right)$, respectively. Assuming that we observe a sufficiently large number of locations, by Theorem C.3, for any $\epsilon_k > 0$, for every $k \in \mathbb{Z}_{\geq 0}$, there is $\epsilon_{k-1} > 0$ w.p. 1 such that:

$$
\begin{aligned}
& d_{\text{TV}}\left(\Gamma^{k-1}_{\text{IBU}}\left(\theta, \hat{C}\right), \Gamma^{k-1}_{\text{IBU}}\left(\theta', \hat{C}\right)\right) < \epsilon_{k-1} \\
\implies & d_{\text{TV}}\left(\Gamma^{k}_{\text{IBU}}\left(\theta, \hat{C}\right), \Gamma^{k}_{\text{IBU}}\left(\theta', \hat{C}\right)\right) < \epsilon_{k}
\end{aligned}
$$

As $n \to \infty$, setting $\epsilon_{N_{\text{IBU}}} = \epsilon > 0$ and $\epsilon_0 = \delta$, by induction, we will get $\{\epsilon_1, \ldots, \epsilon_{N_{\text{IBU}}-1}\}$ w.p. 1 such that:

$$
\begin{aligned}
d_{\text{TV}}(\theta, \theta') < \delta \iff & d_{\text{TV}}\left(\Gamma^0_{\text{IBU}}\left(\theta, \hat{C}\right), \Gamma^0_{\text{IBU}}\left(\theta', \hat{C}\right)\right) < \epsilon_0 \\
\implies & d_{\text{TV}}\left(\Gamma^1_{\text{IBU}}\left(\theta, \hat{C}\right), \Gamma^1_{\text{IBU}}\left(\theta', \hat{C}\right)\right) < \epsilon_1 \\
& \vdots \\
\implies & d_{\text{TV}}\left(\Gamma^{N_{\text{IBU}}-1}_{\text{IBU}}\left(\theta, \hat{C}\right), \Gamma^{N_{\text{IBU}}-1}_{\text{IBU}}\left(\theta', \hat{C}\right)\right) < \epsilon_{N_{\text{IBU}}-1} \\
\implies & d_{\text{TV}}\left(\Gamma^{N_{\text{IBU}}}_{\text{IBU}}\left(\theta, \hat{C}\right), \Gamma^{N_{\text{IBU}}}_{\text{IBU}}\left(\theta', \hat{C}\right)\right) < \epsilon_{N_{\text{IBU}}}
\end{aligned}
$$

□

Lemma C.5. For a privacy channel $\hat{C}$ derived with BA, defined over $\mathcal{X}$, with a sufficiently large number of samples, IBU is a *Lipschitz continuous* transformation w.r.t. d_{TV}, i.e., as $n \to \infty$, there exists constant $\mathcal{K} \in \mathbb{R}^+$ w.p. 1 such that for any pair of PMFs $\theta, \theta' \in \Pi_{\mathcal{X}}$:

$$
d_{\text{TV}}\left(\Gamma_{\text{IBU}}\left(\theta, \hat{C}\right), \Gamma_{\text{IBU}}\left(\theta', \hat{C}\right)\right) \leq \mathcal{K}\, d_{\text{TV}}(\theta, \theta')
$$

Proof. Let θ, θ' be any pair of PMF in $\Pi_{\mathcal{X}}$. If some $K \in \mathbb{R}^+$ satisfies:

$$
d_{\text{TV}}\left(\Gamma_{\text{IBU}}\left(\theta, \hat{C}\right), \Gamma_{\text{IBU}}\left(\theta', \hat{C}\right)\right) \leq \mathcal{K} d_{\text{TV}}(\theta, \theta')
$$

Let d_{L^1} denote the L^1 norm. Then, as $d_{\text{TV}} = \frac{1}{2} d_{L^1}$, the same $\mathcal{K}$ also satisfies:

$$
d_{\text{TV}}\left(\Gamma_{\text{IBU}}\left(\theta, \hat{C}\right), \Gamma_{\text{IBU}}\left(\theta', \hat{C}\right)\right) \leq \mathcal{K}\, d_{\text{TV}}(\theta, \theta')
$$

To prove the result, it is enough to show the Lipschitz continuity of IBU w.r.t L^1 norm. To do this, first, we aim to show the Lipschitz of each step of IBU. We proceed like the proof of Lemma C.3 and, thus, we have:

$$
\begin{aligned}
& d_{\text{TV}}\left(\Gamma^1_{\text{IBU}}\left(\theta, \hat{C}\right), \Gamma^1_{\text{IBU}}\left(\theta', \hat{C}\right)\right) \\
&= \frac{1}{2} \sum_{x\in\mathcal{X}} \left|\Gamma^1_{\text{IBU}}\left(\theta, \hat{C}\right)(x) - \Gamma^1_{\text{IBU}}\left(\theta', \hat{C}\right)(x)\right| \\
&= \sum_{x\in\mathcal{X}} \left| \sum_{y\in\mathcal{X}} q^{(1)}_y \frac{\hat{C}_{xy}\theta(x)}{\sum_{z\in\mathcal{X}} \hat{C}_{zy}\theta(z)} - \sum_{y\in\mathcal{X}} q^{(2)}_y \frac{\hat{C}_{xy}\theta'(x)}{\sum_{z\in\mathcal{X}} \hat{C}_{zy}\theta'(z)} \right| \qquad (16)
\end{aligned}
$$

Here $q^{(1)}$ and $q^{(2)}$ are the corresponding empirical PMFs generated by the observed locations running IBU starting with θ and θ', respectively. Note that the black-box sampling and obfuscation of

the locations that, in turn, give rise to $q^{(1)}$ and $q^{(2)}$ do not depend on θ and θ' respectively. In particular, since both cases of obfuscating the real locations are done by the same channel, as $n \to \infty$, we can assume $\mathbb{P}\left[q^{(1)} = q^{(2)}\right] \to 1$. Let $q = \lim_{n\to\infty} q^{(1)} = \lim_{n\to\infty} q^{(1)}$. Thus, we get an upper bound for (16) w.p. 1 in the form of:

$$\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{X}}\left|q_y\hat{C}_{xy}\left(\frac{\theta(x)}{\sum_{z\in\mathcal{X}}\hat{C}_{zy}\theta(z)}-\frac{\theta'(x)}{\sum_{z\in\mathcal{X}}\hat{C}_{zy}\theta'(z)}\right)\right|$$

[$\because$ Triangle Inequality]

$$=\sum_{y\in\mathcal{X}}q_y\sum_{x\in\mathcal{X}}\hat{C}_{xy}\left|\frac{\theta(x)}{\sum_{z\in\mathcal{X}}\hat{C}_{zy}\theta(z)}-\frac{\theta'(x)}{\sum_{z\in\mathcal{X}}\hat{C}_{zy}\theta'(z)}\right|$$

[$\because q_y \geq 0, \hat{C}_{xy} \geq 0 \,\forall\, x, y$]

$$=\sum_{y\in\mathcal{X}}q_y\sum_{x\in\mathcal{X}}\hat{C}_{xy}\left|\frac{\theta(x)\sum_{z\in\mathcal{X}}\hat{C}_{zy}\theta'(z)-\theta'(x)\sum_{z\in\mathcal{X}}\hat{C}_{zy}\theta(z)}{(\sum_{z\in\mathcal{X}}\hat{C}_{zy}\theta(z))(\sum_{z\in\mathcal{X}}\hat{C}_{zy}\theta'(z))}\right|$$

$$=\sum_{y\in\mathcal{X}}K_y\sum_{x\in\mathcal{X}}\hat{C}_{xy}\left|\sum_{z\in\mathcal{X}}\hat{C}_{zy}(\theta(x)\theta'(z)-\theta'(x)\theta(z))\right|$$

$$[\because K_y = \frac{q_y}{(\sum_{z\in\mathcal{X}}\hat{C}_{zy}\theta(z))(\sum_{z\in\mathcal{X}}\hat{C}_{zy}\theta'(z))}]$$

$$\leq\sum_{y\in\mathcal{X}}K_y\sum_{x\in\mathcal{X}}\hat{C}_{xy}\sum_{z\in\mathcal{X}}\hat{C}_{zy}\left|\theta(x)\theta'(z)-\theta'(x)\theta(z)\right|$$

[$\because$ Triangle Inequality; $\hat{C}_{zy} \geq 0 \,\forall\, z, y$]

$$=\sum_{y\in\mathcal{X}}K_y\sum_{x\in\mathcal{X}}\hat{C}_{xy}\sum_{z\in\mathcal{X}}\hat{C}_{zy}\left|\theta(x)\theta'(z)-\theta(x)\theta(z)+\theta(x)\theta(z)-\theta'(x)\theta(z)\right|$$

$$=\sum_{y\in\mathcal{X}}K_y\sum_{x\in\mathcal{X}}\hat{C}_{xy}\sum_{z\in\mathcal{X}}\hat{C}_{zy}\left|\theta(x)(\theta'(z)-\theta(z))+\theta(z)(\theta(x)-\theta'(x))\right|$$

$$\leq\sum_{y\in\mathcal{X}}K_y\sum_{x\in\mathcal{X}}\hat{C}_{xy}\sum_{z\in\mathcal{X}}\hat{C}_{zy}\left|\theta(x)(\theta'(z)-\theta(z))\right|+\sum_{y\in\mathcal{X}}K_y\sum_{x\in\mathcal{X}}\hat{C}_{xy}\sum_{z\in\mathcal{X}}\hat{C}_{zy}\left|\theta(z)(\theta(x)-\theta'(x))\right|$$

$$<M^2\sum_{y\in\mathcal{X}}K_y\sum_{x\in\mathcal{X}}\sum_{z\in\mathcal{X}}\left|\theta(x)(\theta'(z)-\theta(z))\right|+M^2\sum_{y\in\mathcal{X}}K_y\sum_{x\in\mathcal{X}}\sum_{z\in\mathcal{X}}\left|\theta(z)(\theta(x)-\theta'(x))\right|$$

[Letting $M = \max_{x,y\in\mathcal{X}}\hat{C}_{xy}$]

$$=M^2\sum_{y\in\mathcal{X}}K_y\sum_{x\in\mathcal{X}}\theta(x)\sum_{z\in\mathcal{X}}\left|(\theta'(z)-\theta(z))\right|+M^2\sum_{y\in\mathcal{X}}K_y\sum_{x\in\mathcal{X}}\left|(\theta(x)-\theta'(x))\right|\sum_{z\in\mathcal{X}}\theta(z)$$

$$=2M^2d_{L^1}(\theta,\theta')\sum_{y\in\mathcal{X}}K_y\,[\because \theta \text{ is a PMF}] \quad (17)$$

Similar to the proof of Lemma C.3, we can ignore the columns of $\hat{C}$ which are all 0 as they would result in a probability of 0 for observing $x_r \in \mathcal{X}$ in the output, which effectively will have no influence on the limit of IBU. Therefore, we can assume w.l.o.g. $\min_{y\in\mathcal{X}}\sum_{z\in\mathcal{X}}\hat{C}_{zy}\theta(z) > 0$ for any $\theta \in \Pi_{\mathcal{X}}$. With sufficiently large n observations, we can reasonably assume $\min_{y\in\mathcal{X}}\sum_{z\in\mathcal{X}}\hat{C}_{zy}\theta(z) > \frac{1}{n}$ for any $\theta \in \Pi_{\mathcal{X}}$. Therefore, $\sum_{y\in\mathcal{X}}K_y \leq \sum_{y\in\mathcal{X}}q_y\,n^2 = n^2$. Hence, by setting $\mathcal{K}_1 = 2M^2n^2 > 0$, we satisfy

$$d_{L^1}\left(\Gamma^1_{\text{IBU}}\left(\theta,\hat{C}\right),\Gamma^1_{\text{IBU}}\left(\theta',\hat{C}\right)\right) < \mathcal{K}_1 d_{L^1}(\theta,\theta')$$
$$\implies d_{\text{TV}}\left(\Gamma^1_{\text{IBU}}\left(\theta,\hat{C}\right),\Gamma^1_{\text{IBU}}\left(\theta',\hat{C}\right)\right) < \mathcal{K}_1 d_{\text{TV}}(\theta,\theta')$$

As $\hat{C}$ is generated via BA, by Theorem 5.1, we have a unique MLE that IBU converges to. Assuming IBU converges in N_{IBU} iterations, for all $t \in \{0,\ldots,N_{\text{IBU}-1}\}$, we will have positive $\{\mathcal{K}_2,\ldots,\mathcal{K}_{N_{\text{IBU}}}\}$ satisfying:

$$d_{\text{TV}}\left(\Gamma^{t+1}_{\text{IBU}}\left(\theta,\hat{C}\right),\Gamma^{t+1}_{\text{IBU}}\left(\theta',\hat{C}\right)\right) < \mathcal{K}_{t+1}d_{\text{TV}}\left(\Gamma^{t}_{\text{IBU}}\left(\theta,\hat{C}\right),\Gamma^{t}_{\text{IBU}}\left(\theta',\hat{C}\right)\right)$$

where $d_{\text{TV}}\left(\Gamma^0_{\text{IBU}}\left(\theta,\hat{C}\right),\Gamma^0_{\text{IBU}}\left(\theta',\hat{C}\right)\right) = d_{\text{TV}}(\theta,\theta')$. Therefore, setting $\mathcal{K} = \prod_{i=1}^{N_{\text{IBU}}}\mathcal{K}_i$ gives us the desired result. □

C.3 Compiling BA and IBU under PRIVIC

Having discussed several very interesting and quite desirable analytical properties of the RD function, BA, and IBU, we are now in a position to comment about the uniform continuity of PRIVIC.

Theorem C.6. PRIVIC is a uniformly continuous transformation w.r.t. d_{TV}, i.e., supposing PRIVIC runs for N iterations, for all $\epsilon > 0$ there exists $\delta > 0$ such that for every $\theta_0, \theta' \in \Pi_{\mathcal{X}}$:

$$d_{\text{TV}}(\theta_0,\theta'_0) < \delta \implies d_{\text{TV}}\left(\Lambda\left(\theta_0,N\right),\Lambda(\theta_0,N)\right) < \epsilon$$

Proof. As $d_{\text{TV}} = \frac{1}{2}d_{L^1}$, note that it is sufficient to prove the uniform continuity of PRIVIC w.r.t. the L^1 norm. It is desirable for us to have the privacy channels obtained from BA starting with any two PMFs, $\theta_0, \theta'_0 \in \Pi_{\mathcal{X}}$, to be as close as possible, in order for us to be able to reduce to the setting of Theorem C.4. We start by choosing a very small ϵ' that we need to be able to have the channels produced by BA with the two starting PMFs to be approximately the same, i.e., $\lambda_{\text{BA}}(\theta_0,C_0) \approx \lambda_{\text{BA}}(\theta'_0,C_0)$. More explicitly, let E be defined as:

$$E = \{\epsilon > 0 : d_{\mathcal{C}}\left(\lambda_{\text{BA}}(\theta_0,C_0),\lambda_{\text{BA}}(\theta'_0)\right) < \epsilon \implies \lambda_{\text{BA}}(\theta_0,C_0) \approx \lambda_{\text{BA}}(\theta'_0)\} \quad (18)$$

where $d_{\mathcal{C}}$ is some metric to measure the distance between channels over the space $\mathcal{C}$ containing all encodings from $\mathcal{X}$ to $\mathcal{X}$. Hence, our focus remains on choosing some $\epsilon' \in E$.

By Remark 5, if we have ϵ' satisfying (18), there is $\hat{\epsilon}'(\epsilon') > 0$ that satisfies $\left|RD(X_{\theta_0},d^*) - RD(X_{\theta'_0},d^*)\right| \leq \hat{\epsilon}'$ where X_θ and $X_{\theta'}$ are random variables on $\mathcal{X}$ following the distributions θ and θ'_0, respectively. Since we want the channels to be almost equal, we gauge the behaviour of the RD function taking very small steps. Therefore, it is reasonable for us to choose ϵ' small enough for

$\hat{\epsilon}'(\epsilon') > 0$ to satisfy $\frac{B}{\exp\left\{-W_r\left(-\frac{\hat{\epsilon}'(\epsilon')}{AB}\right)\right\}} < \frac{\tilde{d}}{4\Delta}$, where A, B are as in Lemma C.2.1.

Therefore, by Corollary C.2.1 and Remark 5, we have the uniform continuity of BA. In particular, for $\epsilon' \in E$, we get a $\delta_1 > 0$ such that

$$d_{L^1}(\theta, \theta') < \delta \implies d_{\mathcal{C}}\left(\lambda_{\mathrm{BA}}(\theta_0, C_0), \lambda_{\mathrm{BA}}(\theta_0')\right) \leq \epsilon'$$

Let $\epsilon > 0$. As $\epsilon' \in E$, we can say $\lambda_{\mathrm{BA}}(\theta_0, C_0) \approx \lambda_{\mathrm{BA}}(\theta_0')$, reducing to the condition needed to deploy Theorem C.4. Indeed, by Theorem C.4, we have $\delta_2 > 0$ such that

$$d_{L^1}(\theta, \theta') < \delta_2 \implies d_{L^1}\left(\Gamma_{\mathrm{IBU}}\left(\theta, \hat{C}\right), \Gamma_{\mathrm{IBU}}\left(\theta', \hat{C}\right)\right) < \epsilon$$

Choosing $\delta = \min\{\delta_1, \delta_2\}$ implies that each step of PRIVIC is a uniformly continuous transformation. An inductive argument, following the same line of extending the proof of Lemma C.3 to that of Theorem C.4, gives us the required result. □