2015 IEEE 54th Annual Conference on Decision and Control (CDC)
December 15-18, 2015. Osaka, Japan

Differential Privacy of Populations in Routing Games

Roy Dong, Walid Krichene, Alexandre M. Bayen, and S. Shankar Sastry

Abstract— **As our ground transportation infrastructure modernizes, the large amount of data being measured, transmitted, and stored motivates an analysis of the privacy aspect of these emerging cyber-physical technologies. In this paper, we consider privacy in the routing game, where the origins and destinations of drivers are considered private. This is motivated by the fact that this spatiotemporal information can easily be used as the basis for inferences for a person's activities. More specifically, we consider the differential privacy of the mapping from the amount of flow for each origin-destination pair to the traffic flow measurements on each link of a traffic network. We use a stochastic online learning framework for the population dynamics, which is known to converge to the Nash equilibrium of the routing game. We analyze the sensitivity of this process and provide theoretical guarantees on the convergence rates as well as differential privacy values for these models. We confirm these with simulations on a small example.**

I. Introduction

With the decreasing cost and size of technologies, our ground transportation infrastructure is increasingly modernizing with new sensor systems, control algorithms, and actuation modalities. Although these technologies promise great gains in traffic performance, such as level of service or equity [1], an unprecedented amount of data is being measured, transmitted, and stored, and an analysis of the privacy aspect of this emerging cyber-physical technology is needed.

There have been a multitude of privacy conceptions in the philosophical and legal literatures. From an engineering perspective, the most commonly used paradigms are *control over information* and *secrecy* [2].

In the abstract, control over information generally requires transparency to the person about what data is being collected and stored, consent to the transmission of this data to any parties, and an ability to correct mistakes in the data. As an example of how this conception works in practice, control over information forms the foundation of the Federal Trade Commission's Fair Information Practices.

On the other hand, secrecy focuses on which new inferences can be made about a person due to the information contained in the data; in this paradigm, a privacy breach occurs when there is a revelation of information that was previously not known, and the person felt that the information was private.

Throughout this paper, our conception of privacy will focus on the secrecy paradigm. In other words, we will focus on what new inferences can be made from the data collected by sensors in ground traffic infrastructures.

In the context of traffic systems, we consider the case where the origin and destination are considered private. This is motivated by the fact that this spatiotemporal information can easily used as the basis for inferences for a person's activities. For example, an executive at the carsharing company Uber claimed he could tell when its users were having an affair in a blog post [3].

More specifically, we consider the differential privacy of the mapping from population sizes, i.e. the amount of flow for each origin-destination pair, to the traffic flow measurements on each link of a traffic network.

A popular modeling assumption is that the traffic flow is *atomless*, i.e. a single vehicle cannot unilaterally affect the flows on links [4], [5], [6]. This is designed to match our intuition that, under normal conditions, one vehicle does not contribute significantly to traffic.

However, this implies that, through our models, one vehicle has no effect on the traffic flow measurements. Thus, in this paper, we consider differential privacy with respect to population sizes: how much does traffic flow change when a non-negligible mass of vehicles switch origin-destination pairs?

This framework is applicable for when some aggregator wants to protect the privacy of several drivers. For example, Google can analyze how much it reveals about its users when it provides routes through Google Maps. Alternatively, companies can consider how much is revealed through their shipping patterns, since this detailed data can allow inferences about important business information, such as which consumer markets are being targeted, which companies are in the supply chain, and which locations have potential for future expansion.

To model the dynamics of the driver populations, we use an online learning model in which, at iteration t, each population chooses a distribution over its paths. The joint decision of all populations determines the flows over the edges of the network, which, in turn, determines the costs over paths. These costs are then revealed to the populations, and given this information, they can update their distributions. This online learning model has been applied to routing games in [7], where the authors show that any no-regret strategy is guaranteed to converge to an equilibrium. The same model is also used in [8], where the authors show that if each population applies a mirror descent algorithm, the joint distribution converges to a Nash equilibrium.

Our contribution is an analysis of the differential privacy of the dynamics of the driver populations. In this article, we

R. Dong, W. Krichene, A. M. Bayen, and S. S. Sastry are with the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, Berkeley, CA, 94707, USA {roydong,walid,bayen,sastry}@eecs.berkeley.edu

978-1-4799-7886-1/15/$31.00 ©2015 IEEE

consider a stochastic version of the model in [8], in which the populations only have access to a noisy measurement of the path costs. The presence of noise is essential in providing differential privacy, while still guaranteeing convergence to the equilibrium, using results from stochastic optimization [9], [10].

The rest of the paper is organized as follows. In Section II, we review the engineering literature on privacy and develop some of the theory of differential privacy. In Section III, we introduce the routing game in the context of privacy. In Sections IV and V, we provide a learning model based on stochastic mirror descent. Here, we present theory on convergence rates and analyze the differential privacy of the routing game. In Section VI, we present a numerical example and we conclude in Section VII.

II. Differential privacy

A. Previous work

Motivated by changing technologies, there has been a lot of recent research considering the issue of privacy. In this section, we will try to summarize the mathematical results in this line of research most relevant to this paper, noting both that the field is too rich for a comprehensive literature review and that privacy is a complicated social phenomenon of which a mathematical model is only one facet.

From a mathematical perspective, there have been several definitions of privacy. We seek to quickly survey a few definitions.

There has been work in inferential privacy, which seeks to bound the probability an adversary with a fixed set of information can correctly infer a hidden parameter, and uses a hypothesis testing model [11].

Additionally, there has been work in information-theoretic based definitions of privacy, which uses the mutual information between a private parameter and the publicly observable data [12], [13] or the conditional entropy of a private parameter given the observables [14].

Throughout this paper, we will focus on a definition of privacy first introduced in [15], called *differential privacy*. This definition was originally designed for databases taking values in a finite alphabet, but has since been extended to consider the output of optimization algorithms [16], [17], [18], [19] and dynamical systems [20]. For a more detailed analysis of the interpretation of differential privacy, we refer the reader to [21].

Our work is closest to that in [19], where the authors considered the differential privacy of constraint sets in the context of gradient descent. Additionally, the work in [16] is of relevance, as it provides several minimax bounds for stochastic mirror descent, considered in this paper.

B. Theory

In this section, we will formally define differential privacy, as well as present results needed in future sections.

First, let $(\Omega, \mathcal{A}, P)$ denote our underlying probability space. Also, let Θ be a set equipped with a symmetric binary relation Adj, called the *adjacency* relation. The set Θ contains the possible values for a private parameter. Intuitively, the adjacency relation indicates which values should be roughly indistinguishable from the observable data. Although we never consider distributions or measures on Θ, for brevity we will often treat Θ as a measurable space, where any subset of Θ is measurable.

Furthermore, let $(S, \mathcal{S})$ denote a measurable space and let $Y : \Theta \times \Omega \to S$ be a mapping such that $Y(\theta, \cdot)$ is measurable for every $\theta \in \Theta$. In other words, given θ, $Y(\theta, \cdot)$ is a random element in S. For shorthand, we will write Y_θ to represent $Y(\theta, \cdot)$.

We can now present the definition of differential privacy.

Definition 1. Differential privacy*: We say a measurable mapping* $Y : \Theta \times \Omega \to S$ *is* (ϵ, δ)*-differentially-private if for all measurable sets* $B \in \mathcal{S}$ *and any* $\theta, \theta' \in \Theta$ *such that* $\mathrm{Adj}(\theta, \theta')$*:* $P(Y_\theta \in B) \leq \exp(\epsilon) P(Y_{\theta'} \in B) + \delta$.

If $\delta = 0$*, we will say this mapping is* ϵ*-differentially-private.*

We note two consequences of this definition. The first lemma appears in [20]. The second lemma allows us to use tail bounds when analyzing differential privacy in certain contexts as we will see in Section IV.

Lemma 1. [20]*: If a mapping* Y *is* (ϵ, δ)*-differentially private then* $Eg(Y_\theta) \leq \exp(\epsilon) Eg(Y_{\theta'}) + \delta$ *holds for all bounded measurable real-valued functions* g *and all* $\theta, \theta' \in \Theta$ *such that* $\mathrm{Adj}(\theta, \theta')$.

Lemma 2. *Fix some event* E*. Suppose* $P(E) \geq 1 - \delta'$ *and that, for all measurable sets* $B \in \mathcal{S}$ *and all* $\theta, \theta' \in \Theta$ *such that* $\mathrm{Adj}(\theta, \theta')$*:* $P(\{Y_\theta \in B\} \cap E) \leq \exp(\epsilon) P(\{Y_{\theta'} \in B\} \cap E) + \delta$*. Then,* Y *is* $(\epsilon, \delta + \delta')$*-differentially-private.*

A result we will use in future sections is how differentially private mappings can be composed.

Proposition 1. Adaptive composition*: Suppose* $Y_1 : \Theta \times \Omega \to S_1$ *is* (ϵ_1, δ_1)*-differentially-private and* $Y_2 : \Theta \times S_1 \times \Omega \to S_2$ *is a measurable mapping such that* $Y_2(\cdot, s, \cdot)$ *is* (ϵ_2, δ_2)*-differentially-private for each fixed* $s \in S_1$*. Then the mapping* $(\theta, \omega) \mapsto (Y_1(\theta, \omega), Y_2(\theta, Y_1(\theta, \omega), \omega))$ *is* $(\epsilon_1 + \epsilon_2, \exp(\epsilon_2)\delta_1 + \delta_2)$*-differentially-private.*

Additionally, we can induct on Proposition 1. For brevity, we will sometimes write $Y_t(\theta, Y_1(\theta), \ldots, Y_{t-1}(\theta), \cdot)$ simply as $Y_t(\theta)$.

Corollary 1. Repeated adaptive composition*: Suppose* $Y_1 : \Theta \times \Omega \to S_1$ *is* (ϵ_1, δ_1)*-differentially-private and* $Y_t : \Theta \times S_1 \times \ldots S_{t-1} \times \Omega \to S_t$ *is a measurable mapping such that* $Y_t(\cdot, s_1, \ldots, s_{t-1}, \cdot)$ *is* (ϵ_t, δ_t)*-differentially-private for each fixed* $(s_1, \ldots, s_{t-1}) \in S_1 \times \cdots \times S_{t-1}$ *and* $1 < t \leq T$*.*

Then, the mapping $(\theta, \omega) \mapsto (Y_1(\theta), Y_2(\theta), \ldots, Y_T(\theta))$ *is* $\left(\sum_{t=1}^T \epsilon_t, \sum_{t=1}^T \exp\left[\sum_{t'=t+1}^T \epsilon_{t'}\right] \delta_t\right)$*-differentially-private.*

Finally, we note that the Gaussian distribution guarantees differential privacy.

Definition 2. Sensitivity*: The ℓ_2 sensitivity of a function $f : \Theta \to \mathbb{R}$ is given by:*

$$\Delta_2 f = \sup_{\theta,\theta' \in \Theta : \mathrm{Adj}(\theta,\theta')} \|f(\theta) - f(\theta')\|_2 \tag{1}$$

Definition 3. *The zero-mean Gaussian distribution on $\mathbb{R}$ with variance parameter σ^2, denoted* Gauss(σ^2)*, has the density $y \mapsto \frac{1}{\sqrt{(2\pi\sigma^2)}} \exp\left(\frac{-|y|^2}{2\sigma^2}\right)$ with respect to the Lebesgue measure.*

Proposition 2. Gaussian mechanism [21]*: For $\epsilon \in (0, 1)$, and $b^2 > 2\ln(1.25/\delta)$, the mapping $Y_\theta = f(\theta) + Z$, where $Z_i \overset{iid}{\sim} \mathrm{Gauss}(\sigma^2)$ for some $\sigma \geq b\Delta_2 f/\epsilon$, is (ϵ, δ)-differentially-private.*

III. The Routing Game

The routing game is given by: a directed graph $G = (V, E)$, a set of non-decreasing, Lipschitz continuous edge cost functions $c_e : \mathbb{R}_+ \to \mathbb{R}_+$, $e \in E$, a finite set of origin-destination pairs $(o_i, d_i) \in V \times V$, indexed by $i \in \{1, \ldots, I\}$, and a finite set of populations P_k, indexed by $k \in \{1, \ldots, K\}$.

For a given origin-destination pair (o_i, d_i), let $\mathcal{P}_i$ be the set of simple paths connecting o_i to d_i, and let $M_i \in \mathbb{R}^{|E| \times |\mathcal{P}_i|}$ be the edge-path incidence matrix, where, for all $(e, p) \in E \times \mathcal{P}_i$, $(M_i)_{e,p} = 1$ if $e \in p$ and $(M_i)_{e,p} = 0$ otherwise.

A population P_k is given by a private vector $\theta_k \in \mathbb{R}^I_+$, which specifies, for each origin-destination pair (o_i, d_i), the total mass of traffic $(\theta_k)_i$ that belongs to this population, and that travels from o_i to d_i. We assume there is some upper bound on the total size of the populations. Furthermore, we will define an adjacency relationship between private vectors.

Assumption 1. *It is common knowledge that θ is bounded. That is, there exists an $A_\theta < \infty$ such that, for every population k, $\|\theta_k\|_\infty \leq A_\theta$, and each population and outside observers know this bound.*

Definition 4. *Two private parameters of populations $(\theta_k)_{k\in[K]}$ and $(\theta'_k)_{k\in[K]}$ are* adjacent *if there exists a k^* such that $\theta_k = \theta'_k$ for $k \neq k^*$ and $\|\theta_{k^*} - \theta'_{k^*}\|_\infty \leq c$.*

Recall that the adjacency relationship provides defines which pairs of private parameters should be roughly indistinguishable. Here, c is a constant that will be determined by the populations, modeling the maximum amount that a single population can increase or decrease the flow in one origin-destination pair without having a significant effect on observable data.

The action set of population P_k is a distribution vector $x_k \in \Delta^{\mathcal{P}_1} \times \cdots \times \Delta^{\mathcal{P}_I}$, where $\Delta^{\mathcal{P}_i} = \left\{m \in \mathbb{R}^{|\mathcal{P}_i|}_+ : \sum_{p\in\mathcal{P}_i} m_p = 1\right\}$ is the set of probability distributions over $\mathcal{P}_i$. In other words, every population chooses, for each origin-destination pair (o_i, d_i), how to distribute its mass across the available paths $\mathcal{P}_i$. For notational convenience, we will write $(x_k)_{\mathcal{P}_i}$ to denote the sub-vector $((x_k)_p)_{p\in\mathcal{P}_i} \in \Delta^{\mathcal{P}_i}$, so that $x_k = ((x_k)_{\mathcal{P}_1}, \ldots, (x_k)_{\mathcal{P}_I})$.

The flow allocations of all populations $(x_k)_{k\in[K]}$ determine the edge flows, defined as follows: the flow on edge e is $\phi_e(x_1, \ldots, x_K) = \sum_{k=1}^K \sum_{i=1}^I (\theta_k)_i \sum_{p\in\mathcal{P}_i} (x_k)_p 1_{(e\in p)}$. The vector of edge flows can be written simply in terms of the incidence matrices: $\phi(x_1, \ldots, x_K) = \sum_{k=1}^K \sum_{i=1}^I (\theta_k)_i M_i (x_k)_{\mathcal{P}_i}$. The edge flows and edge costs determine the path costs. That is, the cost on path $p \in \mathcal{P}_i$ is given by $\ell_p(x_1, \ldots, x_K) = \sum_{e\in p} c_e(\phi_e(x_1, \ldots, x_K))$. We will denote by $\ell_{\mathcal{P}_i}(x_1, \ldots, x_K)$ the vector $(\ell_p(x_1, \ldots, x_K))_{p\in\mathcal{P}_i}$, and $\ell = (\ell_{\mathcal{P}_1}, \ldots, \ell_{\mathcal{P}_I}) \in \mathbb{R}^{\mathcal{P}_1}_+ \times \cdots \times \mathbb{R}^{\mathcal{P}_I}_+$.

Finally, the cost for population P_k under distributions $x_1, \ldots, x_K$ is $\sum_{i=1}^I (\theta_k)_i \sum_{p\in\mathcal{P}_i} ((x_k)_{\mathcal{P}_i})_p \ell_p(x_1, \ldots, x_K)$, which we will denote, more concisely, as $\langle x_k, \ell(x_1, \ldots, x_K)\rangle_{\theta_k}$. Here we define the inner product as follows: for all $x, y \in \mathbb{R}^{\mathcal{P}_1} \times \cdots \times \mathbb{R}^{\mathcal{P}_I}$, $\langle x, y\rangle_\theta = \sum_{i=1}^I \theta_i \sum_{p\in\mathcal{P}_i} x_p y_p$.

A. Nash equilibria and the Rosenthal potential function

Definition 5. *A collection of population distributions $(x_k)_{k\in[K]}$ is a* Nash equilibrium *(also called* Wardrop equilibrium *in the traffic literature), if for every $k \in [K]$ and every $y \in \Delta^{\mathcal{P}_1} \times \cdots \times \Delta^{\mathcal{P}_I}$: $\langle x_k, \ell(x_1, \ldots x_K)\rangle_{\theta_k} \leq \langle y, \ell(x_1, \ldots x_K)\rangle_{\theta_k}$. That is, no driver can improve their cost by unilaterally changing their path.*

Next, we show that the set of Nash equilibria of the game are exactly the set of minimizers of the Rosenthal potential, defined as: $f(x_1, \ldots, x_K) = \sum_{e\in E} \int_0^{\phi_e(x_1,\ldots,x_K)} c_e(u)du$.

Proposition 3. *The Rosenthal potential is convex, and its gradient with respect to x_k is: $\nabla_{x_k} f(x_1, \ldots, x_K) = \sum_{i=1}^I (\theta_k)_i \ell_{\mathcal{P}_i}(x_1, \ldots, x_K)$.*

Corollary 2. *The set of Nash equilibria of the game is exactly the set of solutions of the following convex problem:*

$$\begin{aligned} &\textit{minimize} && f(x_1, \ldots, x_K) \\ &\textit{subject to} && x_k \in \Delta^{\mathcal{P}_1} \times \cdots \times \Delta^{\mathcal{P}_I} \textit{ for all } k \in [K] \end{aligned} \tag{2}$$

IV. Stochastic mirror descent dynamics and convergence to Nash equilibria

A. Online learning model

We consider the following online learning model of the game: at each iteration $t \in \{1, 2, \ldots, T\}$, every population P_k chooses a distribution vector $x_k^{(t)} \in \Delta^{\mathcal{P}_1} \times \cdots \times \Delta^{\mathcal{P}_I}$. The combined choice of all populations determines the path loss vector $\ell(x_1^{(t)}, \ldots, x_K^{(t)})$, which we will denote simply by $\ell^{(t)}$. The loss of population k is then given by the inner product $\left\langle \ell^{(t)}, x_k^{(t)} \right\rangle_{\theta_k}$.

At the end of iteration t, a stochastic loss vector $\hat{\ell}^{(t)}$, is revealed to all populations. Intuitively, one can think of $\hat{\ell}^{(t)}$ as a noisy version of $\ell^{(t)}$. The precise assumptions on the process $(\hat{\ell}^{(t)})$ will be given in Assumption 5.

B. Population dynamics

Our population dynamics take the following form.

Assumption 2. *We assume that for each population P_k, the stochastic process $(x_k^{(t)})$ follows stochastic mirror descent dynamics, given in Algorithm 1.*

Algorithm 1 Stochastic mirror descent dynamics for population k, with initial distribution $x_k^{(0)}$, learning rates $(\eta_k^{(t)})$, and distance generating function ψ_k.

```
for $t \in \{0, \dots, T-1\}$ do
    Observe $\hat{\ell}^{(t)}$
    Update
    $x_k^{(t+1)} = \underset{x_k \in \Delta^{\mathcal{P}_1} \times \cdots \times \Delta^{\mathcal{P}_I}}{\arg\min} \left\langle \hat{\ell}^{(t)}, x_k \right\rangle_{\theta_k} + \frac{1}{\eta_k^{(t)}} D_{\psi_k}(x_k, x_k^{(t)})$
end for
```

These dynamics correspond to a stochastic version of the dynamics used in [8].

Here, ψ_k is a distance generating function defined and C^1 on $\Delta^{\mathcal{P}_1} \times \cdots \times \Delta^{\mathcal{P}_I}$, and D_{ψ_k} is the Bregman divergence induced by ψ_k, defined as follows: $D_{\psi_k}(x_k, y_k) = \psi(x_k) - \psi(y_k) - \langle \nabla\psi(y_k), x_k - y_k \rangle$.

Assumption 3. *For all k, ψ_k is strongly convex with respect to a reference norm $\|\cdot\|$. That is, there exists $\ell_{\psi_k} > 0$ such that for all $x_k, y_k \in \Delta^{\mathcal{P}_1} \times \cdots \times \Delta^{\mathcal{P}_I}$: $D_{\psi_k}(x_k, y_k) \geq \frac{\ell_{\psi_k}}{2}\|x_k - y_k\|^2$.*

See Chapter 11 in [22] for an account of the properties of Bregman divergences. We will further assume that the norm $\|\cdot\|$ decomposes into a sum of norms defined on each of the simplexes.

Assumption 4. *The norm $\|\cdot\|$ on $\mathbb{R}^{\mathcal{P}_1} \times \cdots \times \mathbb{R}^{\mathcal{P}_I}$ can be decomposed as follows: $\|x_k\| = \sum_{i \in I} \|(x_k)_{\mathcal{P}_i}\|$.*

Mirror descent is a general class of first-order optimization methods, used extensively both in convex optimization [23] and online learning [22], [24]. In particular, projected gradient descent and entropic descent (a.k.a. the Hedge algorithm) are instances of the mirror descent method, for the appropriate choices of the distance generating function (see, for example, [25]).

In our model, since each population is updating its distribution vector using mirror descent dynamics, we can write the joint update: $(x^{(t+1)}, \dots, x_K^{(t+1)}) = \arg\min_x f(x^{(t)}) + \left\langle \nabla f(x^{(t)}), x - x^{(t)} \right\rangle + D^{(t)}(x, x^{(t)})$. Here the minimization is taken across x in $(\Delta^{\mathcal{P}_1} \times \cdots \times \Delta^{\mathcal{P}_I})^K$ and we used the expression of the gradient $\nabla f(x^{(t)})$, given in Proposition 3, and defined: $D^{(t)}(x, x^{(t)}) = \sum_k \frac{1}{\eta_k^{(t)}} D_{\psi_k}(x_k, x_k^{(t)})$.

The joint update expression can be interpreted as a local approximation of the potential function f: the first term $f(x^{(t)}) + \left\langle \nabla f(x^{(t)}), x - x^{(t)} \right\rangle$ is simply the linear approximation of f around $x^{(t)}$, and the second term $D^{(t)}(x, x^{(t)})$ is a strongly convex function which penalizes deviation from the previous iterate $x^{(t)}$. By this observation, one can think of the joint dynamics of all populations as implementing a stochastic mirror descent on the Rosenthal potential f.

C. Suboptimality bounds on stochastic mirror descent

We now review some guarantees of the stochastic mirror descent method. First, we need to make assumptions on the stochastic process $(\hat{\ell}^{(t)})$ and the distance generating functions ψ_k.

Assumption 5. *Throughout the paper, we will assume that:*

1) *For all t, $\hat{\ell}^{(t)}$ is unbiased, that is, $\mathbb{E}\left[\hat{\ell}^{(t)} | \mathcal{F}_{t-1}\right] = \ell^{(t)}$, where $(\mathcal{F}_t)$ is the natural filtration of the process $(\hat{\ell}^{(t)})$.*
2) *$\hat{\ell}^{(t)}$ is uniformly bounded in the squared dual norm, that is, there exists L such that for all t: $\mathbb{E}\left[\|\ell^{(t)}\|_*^2\right] \leq L$, where $\|\cdot\|_*$ is the dual norm defined as follows: $\|\ell\|_* = \sup_{\|x\| \leq 1} \langle x, \ell \rangle$.*
3) *For all k, there exists D_k such that D_{ψ_k} is bounded on $\Delta^{\mathcal{P}_1} \times \cdots \times \Delta^{\mathcal{P}_I}$ by D_k.*

Proposition 4 (Theorem 4 in [10]). *Suppose that each population P_k follows a stochastic mirror descent dynamics as in Algorithm 1, and suppose that the learning rates are given by $\eta_k^{(t)} = c_k t^{-\alpha_k}$ with $c_k > 0$ and $\alpha_k \in (0, 1)$. Then for all $t \geq 1$, it holds that: $\mathbb{E}\left[f(x^{(t)})\right] - f^\star \leq \left(1 + \sum_{\tau=1}^{t} \frac{1}{\tau}\right) \sum_{k=1}^{K} \left(\frac{1}{t^{1-\alpha_k}} \frac{D_k}{c_k} + \frac{c_k L}{2\ell_{\psi_k}(1-\alpha_k)} \frac{1}{t^{\alpha_k}}\right)$.*

In particular, the system converges to the set of Nash equilibria in expectation, in the sense that $\mathbb{E}\left[f(x^{(t)})\right] \to f^\star$ at the rate $\mathcal{O}(t^{-\bar{\alpha}} \log t)$ where $\bar{\alpha} = \min_k \min(\alpha_k, 1 - \alpha_k)$.

D. Sensitivity analysis of the stochastic mirror descent update

In this Section, we study the sensitivity of the stochastic process $\hat{\ell}^{(t)}(x^{(t)})$ to changes in the private parameter θ.

First, we consider how the flow allocations change due to a change in mass on some origin-destination pairs. In this case, we hold the observed loss vector $\hat{\ell}^{(t)}$ fixed and will invoke Corollary 1 afterward.

Proposition 5. *Fix a loss vector $\hat{\ell}^{(t)}$ and consider the stochastic mirror descent update for population P_k: $x_k^{(t+1)}(\theta_k) = \arg\min_{x_k} \left\langle \hat{\ell}^{(t)}, x_k \right\rangle_{\theta_k} + \frac{1}{\eta_k^{(t)}} D_{\psi_k}(x_k, x_k^{(t)})$. Here the minimization is taken across $\Delta^{\mathcal{P}_1} \times \cdots \times \Delta^{\mathcal{P}_I}$ and x^k is viewed as a function of the mass vector θ_k. Then for all $\theta_k, \theta'_k \in \mathbb{R}_+^I$: $\|x_k^{(t+1)}(\theta_k) - x_k^{(t+1)}(\theta'_k)\| \leq \frac{\eta_k^{(t)} \|\hat{\ell}^{(t)}\|_*}{\ell_{\psi_k}} \|\theta_k - \theta'_k\|_\infty$.*

We have bounded how much a change in the private parameter affects the distribution on paths. Now, we analyze how the flows are affected by changes in distribution.

We will use the notation $\phi(x; \theta)$, which makes the dependence of edge flows on the parameter θ explicit. Also, $x^{(t+1)}(\theta)$ will be shorthand for $(x_1^{(t+1)}(\theta_1), \dots, x_K^{(t+1)}(\theta_K))$. Also, let $\|\cdot\|_a$ denote an arbitrary norm on the space of edge flows.

Lemma 3. *For any* $\mathrm{Adj}(\theta, \theta')$, *we have:* $\|\phi(x^{(t+1)}(\theta); \theta) - \phi(x^{(t+1)}(\theta'); \theta')\|_a \leq cA_x \left[A_\Delta + A_\theta \frac{\eta_k^{(t)} \|\hat{\ell}^{(t)}\|_*}{\ell_{\psi_k}}\right]$ *Here, A_θ is as given in Assumption 1 and $A_x = \sup_{\|x_k\| \leq 1} \left\|\sum_{i=1}^{I} M_i(x_k)_{\mathcal{P}_i}\right\|_a$, $A_\Delta = \sup_{x_k \in \Delta^{\mathcal{P}_1} \times \cdots \times \Delta^{\mathcal{P}_I}} \|x_k\|$.*

We have bounded the effect of a change in the private parameter on the flows. Thus, we can state the sensitivity of the loss vector at time $t+1$ due to a small differential in the private parameter θ, when the observed loss vector at time t is held fixed.

Theorem 1. Sensitivity of the loss function: *For any* $\mathrm{Adj}(\theta, \theta')$*:* $\|\ell(x^{(t+1)}(\theta); \theta, x^{(t)}, \hat{\ell}^{(t)}) - \ell(x^{(t+1)}(\theta'); \theta', x^{(t)}, \hat{\ell}^{(t)})\| \leq cA_\ell A_x \left[A_\Delta + A_\theta \frac{\max_{k\in[K]}(\eta_k^{(t)})\|\hat{\ell}^{(t)}\|_*}{\min_{k\in[K]}(\ell_{\psi_k})}\right]$.

Here, A_x, A_Δ*, and* A_θ *are as defined in Assumption 1 and Lemma 3, and* A_ℓ *denotes the Lipschitz constant of the function* $\ell : \phi \mapsto \ell(\phi)$ *with respect to the norm* $\|\cdot\|_a$ *on the domain and* $\|\cdot\|$ *on the codomain.*

Note that the sensitivity of $\ell^{(t+1)}$ depends on t through the learning rate $\eta_k^{(t)}$.

V. Differential privacy of the routing game

In this Section, we use results from the previous sections to give privacy guarantees on the routing game when the loss vectors are observed with Gaussian noise.

Also, recall that the Gaussian mechanism preserves (ϵ, δ) differential privacy, and the privacy values depend on the variance of the mechanism and the sensitivity of the function. At each iteration t, we suppose that the populations observe $\hat{\ell}^{(t)} = \ell(x^{(t)}) + Z_t$ where $(Z_t)_p \overset{iid}{\sim} \mathrm{Gauss}(\sigma^2)$.

First, we observe that for each path p, since the loss function ℓ_p is continuous on the compact set $\left(\Delta^{\mathcal{P}_1} \times \cdots \times \Delta^{\mathcal{P}_I}\right)^K$, it is bounded. Therefore, there exists $M > 0$ such that for all $x \in \left(\Delta^{\mathcal{P}_1} \times \cdots \times \Delta^{\mathcal{P}_I}\right)^K$, $\|\ell(x)\|_\infty \leq M$.

Theorem 2. *After* T *iterations, the mapping* $\theta \mapsto (\hat{\ell}^{(1)}, \dots, \hat{\ell}^{(T)})$ *is* (ϵ, δ) *differentially private, where:* $\epsilon = \sum_{t=1}^T \epsilon_t$ *and* $\delta = \sum_{t=1}^T \exp\left[\sum_{t'=t+1}^T \epsilon_{t'}\right] \delta_t + \delta'$*. Here,* a *is any positive constant and* $\delta', \epsilon_t, \delta_t$ *are any constants that satisfy the following constraints:*

$$1 - \delta' = (1 - 2\exp(-a^2/2\sigma^2))^{T\sum_{i=1}^I |\mathcal{P}_i|}$$

$$\epsilon_t > \frac{cA_\ell A_x (2\ln(1.25/\delta_t))^{1/2}}{\sigma^2} \times \left[A_\Delta + A_\theta \frac{\max_{k\in[K]}(\eta_k^{(t)})(\sum_{i=1}^I |\mathcal{P}_i|)^{1/2}(M+a)}{\min_{k\in[K]}(\ell_{\psi_k})}\right]$$

A_x, A_Δ, A_θ*, and* A_ℓ *are as defined in Assumption 1, Lemma 3, and Theorem 1.*

Note that a can be chosen to be any positive constant, and, in effect, provides a trade-off between the ϵ and the δ parameters.

VI. Numerical Example

Consider the routing game played on the network in Figure 1, with the following populations:

1) Population P_1 has mass vector $\theta_1 = (1, 0)$, and follows stochastic mirror descent dynamics with learning rates $\mathcal{O}(t^{-.5})$.
2) Population P_2 has mass vector $\theta_2 = (.2, 1.2)$, and follows stochastic mirror descent dynamics with learning rates $\mathcal{O}(t^{-.2})$.

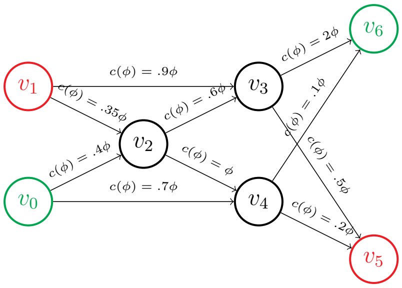

Fig. 1. Example network with two origin-destination pairs: (v_0, v_6) and (v_1, v_5).

The losses are taken to be linear. The resulting path loss functions are bounded by $M = 2$. We simulate the game for $T = 200$ iterations, with Gaussian noise with standard deviation $\sigma \in \{.01, .1, .4\}$.

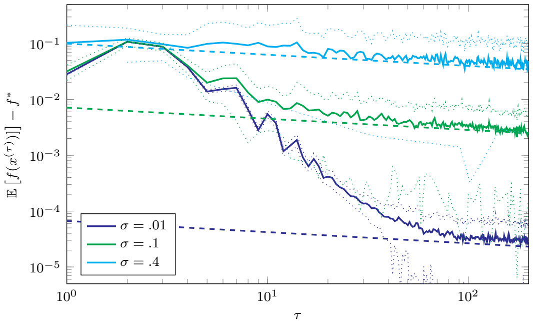

Fig. 2. Potential function values $f(x^{(\tau)})$ as a function of the iteration τ, for different values of σ. The solid and dotted lines show, respectively, the average and the standard deviation over 150 runs of the simulation. The dashed lines show the $\tilde{\mathcal{O}}(t^{-.2})$ asymptotic rate predicted by Proposition 4.

Figure 2 shows the values of the potential function for the different values of σ. The asymptotic rate is consistent with $\tilde{\mathcal{O}}(t^{-\min(\alpha_1,\alpha_2)}) = \tilde{\mathcal{O}}(t^{-.2})$ rate predicted by Proposition 4. The variance of the noise σ^2 significantly affects the value of the expected potential. The effect of σ can also be observed in Figure 3, which shows the path flows for both populations, for $\sigma \in \{.01, .4\}$. Besides the effect of the noise level, we also observe that because the learning rates of population P_2 have a slower decay rate, its updates are more aggressive, which is reflected in the trajectories of its path flows.

Additionally, we consider the differential privacy of these observable traffic flows. Applying Theorem 2, we plot the differential privacy values as a function of the number of iterations in Figure 4. Generally, we are able to mask a small amount of population flow, but should c grow too large, the bounds quickly become trivial, i.e. $\delta = 1$. Furthermore, this value at which we can no longer meaningfully guarantee privacy can be thought of as the rate at which populations must shift origin-destination pairs to retain some level of privacy guarantee.

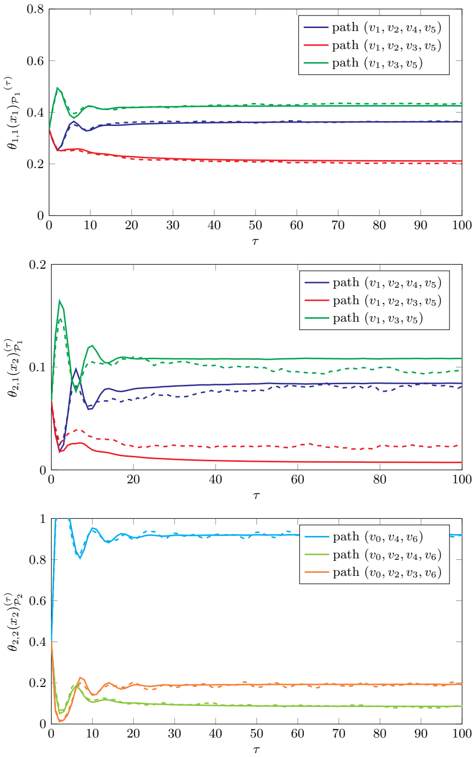

Fig. 3. Path flows for each population, averaged over 150 runs, for $\sigma = .01$ (solid lines) and $\sigma = .4$ (dashed lines)

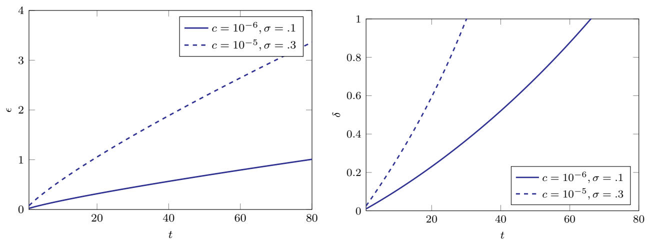

Fig. 4. A plot of the values of ϵ, δ for which differential privacy holds, as a function of t, the number of iterations. Here, (c, σ) are taken to be $(10^{-6}, .1)$ then $(10^{-5}, .3)$, and a is taken to be 2. For larger values of c, the privacy guarantees are only meaningful for shorter periods of time.

VII. Conclusion

In this paper, we considered the privacy of the origins and destinations of drivers when the nominal traffic losses are observable with Gaussian noise. Considering a general online learning model based on stochastic mirror descent, and noting that the routing game is a potential game, we can think of the dynamics of drivers as optimizing the Rosenthal potential.

We analyzed the sensitivity of each update step as a function of the masses for each origin-destination pair, which allowed us to bound the influence of this private information on the observable traffic losses. Additionally, we provided bounds on the convergence rates for different levels of noise, which provides insight into the relationship between how long it takes traffic flows to settle at equilibrium and how much is revealed by these observable traffic costs.

References

[1] Transportation Research Board, “Transportation Research Board 2011 annual report,” The National Academies, Tech. Rep., 2011.

[2] D. J. Solove, “Conceptualizing privacy,” *California Law Review*, vol. 90, p. 1087, 2002.

[3] Z. Tufekci and B. King. (2014, Dec.) We can’t trust Uber. New York Times.

[4] W. H. Sandholm, “Potential games with continuous player sets,” *Journal of Economic Theory*, vol. 97, no. 1, pp. 81–108, 2001.

[5] T. Roughgarden, “Routing games,” in *Algorithmic game theory*. Cambridge University Press, 2007, ch. 18, pp. 461–486.

[6] W. Krichene, B. Drighès, and A. Bayen, “Learning nash equilibria in congestion games,” *SIAM Journal on Control and Optimization (SICON), to appear*, 2015.

[7] A. Blum, E. Even-Dar, and K. Ligett, “Routing without regret: on convergence to nash equilibria of regret-minimizing algorithms in routing games,” in *Proceedings of the twenty-fifth annual ACM symposium on Principles of distributed computing*, ser. PODC ’06. New York, NY, USA: ACM, 2006, pp. 45–52.

[8] W. Krichene, S. Krichene, and A. Bayen, “Convergence of mirror descent dynamics in the routing game,” in *European Control Conference (ECC), accepted*, 2015.

[9] A. Juditsky, A. Nemirovski, and C. Tauvel, “Solving variational inequalities with stochastic mirror-prox algorithm,” *Stoch. Syst.*, vol. 1, no. 1, pp. 17–58, 2011. [Online]. Available: http://dx.doi.org/10.1214/10-SSY011

[10] S. Krichene, W. Krichene, R. Dong, and A. Bayen, “Convergence of stochastic mirror descent and applications to distributed optimization,” in *Internation Conference on Machine Learning (ICML), in review*, 2015.

[11] R. Dong, L. Ratliff, H. Ohlsson, and S. S. Sastry, “Fundamental limits of nonintrusive load monitoring,” in *Proc. of the 3rd Int. Conf. on High Confidence Networked Systems*, ser. HiCoNS ’14. New York, NY, USA: ACM, 2014, pp. 11–18. [Online]. Available: http://doi.acm.org/10.1145/2566468.2566471

[12] L. Sankar, S. Kar, R. Tandon, and H. Poor, “Competitive privacy in the smart grid: An information-theoretic approach,” in *2011 IEEE Int. Conf. on Smart Grid Communications (SmartGridComm)*, Oct 2011, pp. 220–225.

[13] S. Salamatian, A. Zhang, F. du Pin Calmon, S. Bhamidipati, N. Fawaz, B. Kveton, P. Oliveira, and N. Taft, “How to hide the elephant-or the donkey-in the room: Practical privacy against statistical inference for large data,” *IEEE GlobalSIP*, 2013.

[14] P. Venkitasubramaniam, “Privacy in stochastic control: A markov decision process perspective,” in *2013 51st Annu. Allerton Conf. on Communication, Control, and Computing (Allerton)*, Oct 2013, pp. 381–388.

[15] C. Dwork, “Differential privacy,” in *Proc. of the Int. Colloq. on Automata, Languages and Programming*. Springer, 2006, pp. 1–12.

[16] J. C. Duchi, M. I. Jordan, and M. J. Wainwright, “Privacy aware learning,” *arXiv*, 2012.

[17] J. Hsu, Z. Huang, A. Roth, and Z. S. Wu, “Jointly private convex programming,” *arXiv*, 2014.

[18] Z. Huang, S. Mitra, and N. Vaidya, “Differentially private distributed optimization,” *arXiv*, 2014.

[19] S. Han, U. Topcu, and G. J. Pappas, “Differentially private distributed constrained optimization,” *arXiv*, 2014.

[20] J. Le Ny and G. Pappas, “Differentially private filtering,” *IEEE Trans. Autom. Control*, vol. 59, pp. 341–354, Feb 2014.

[21] C. Dwork and A. Roth, *The Algorithmic Foundations of Differential Privacy*. Foundations and Trends in Theoretical Computer Science, 2014.

[22] N. Cesa-Bianchi and G. Lugosi, *Prediction, learning, and games*. Cambridge University Press, 2006.

[23] A. S. Nemirovsky and D. B. Yudin, *Problem complexity and method efficiency in optimization*, ser. Wiley-Interscience series in discrete mathematics. Wiley, 1983.

[24] S. Bubeck and N. Cesa-Bianchi, “Regret analysis of stochastic and nonstochastic multi-armed bandit problems,” *Foundations and Trends in Machine Learning*, vol. 5, no. 1, pp. 1–122, 2012.

[25] A. Beck and M. Teboulle, “Mirror descent and nonlinear projected subgradient methods for convex optimization,” *Oper. Res. Lett.*, vol. 31, no. 3, pp. 167–175, May 2003.