A stochastic generalized Nash equilibrium model for platforms competition in the ride-hail market

Filippo Fabiani and Barbara Franci

Abstract—The inherent uncertainties in the ride-hailing market complicate the pricing strategies of on-demand platforms that compete each other to offer a mobility service while striving to maximize their profit. Looking at this problem as a stochastic generalized Nash equilibrium problem (SGNEP), we design a distributed, stochastic equilibrium seeking algorithm with Tikhonov regularization to find an optimal pricing strategy. The proposed iterative scheme does not require a potentially infinite number of samples of the random variable to perform the stochastic approximation, thus making it appealing from a practical perspective. Moreover, we show that the algorithm returns a Nash equilibrium under mere monotonicity assumptions and a careful choice of the step size sequence, obtained by exploiting the specific structure of the SGNEP at hand.

I. INTRODUCTION

In the last few years, we have been experiencing a dizzying growth of the ride-hailing market [1], where on-demand ride-hailing platforms, such as Uber, Lyft and Didi Chuxing, have to settle up suitable pricing strategies to be attractive on two nearly complementary fronts: customers and drivers. Each ride-hailing firm, indeed, not only competes for customers with the other firms, but also strives to secure an as wide as possible fleet of “loyal” drivers so that it can meet possibly growing customers’ demand, which often may not be predicted accurately. In this framework, competition among the platforms can be naturally described through a stochastic generalized Nash equilibrium problem (SGNEP): the firms aim at maximizing their expected valued profit function, trying to satisfy the demand for rides, which is typically uncertain, while sharing the market with the other platforms.

SGNEPs amount to a collection of coupled stochastic optimization problems, which are challenging to address if one aims at finding a stochastic generalized Nash equilibrium (SGNE) in a distributed fashion. The difficulties are mainly due to the constraints coupling the agents’ strategies, and the presence of uncertainty. While standard approaches well accommodate the first issue [2], [3], the second one is more delicate and known techniques typically approximate the expected-value pseudogradient mapping of the game by exploiting available realizations of the uncertainty [4], [5].

As a next step, one should design a distributed algorithm, with provable convergence guarantees to an equilibrium solution of the SGNEP at hand. Among the numerous algorithms for classic stochastic optimization [6], [7], only few of them are amenable to solve SGNEPs, as for instance the forward-backward (SFB) algorithm [3] and related variants, namely the relaxed forward-backward (SRFB) [8] or the projected-reflected-gradient (SPRG) algorithms [9]. These procedures however are affected by common drawbacks: the monotonicity assumption of the operators involved and the number of samples necessary for the approximation. In fact, both SFB and SPRG algorithms converge in case the pseudogradient mapping is strongly monotone or cocoercive, which are rather strong assumptions. Moreover, FB-based methods take as an approximation the average over an increasing, possibly infinite, number of samples of the uncertainty, which is fairly impractical. These reasons motivate us to modify the traditional FB algorithm with a Tikhonov (Tik) regularization method [10, Ch. 12] tailored for SGNEPs. Introducing a regularization sequence is a well-known technique to weaken strong assumptions [5], [11], since it allows one to obtain a strongly monotone operator starting from a merely monotone one [10]. We summarize all these considerations in Table I.

In [5], a distributed version of the Tikhonov regularization algorithm was introduced for SGNEPs without coupling constraints. We show here that the generalization to SGNEPs is possible, albeit non-trivial, as the time-varying nature of the step sizes and regularization step poses technical challenges to be treated carefully when using operator splitting techniques. We hence summarize our contributions as follows:

- We propose a noncooperative model for on-demand ride-hailing platforms under a regulated pricing scenario and uncertainties, resulting into a SGNEP (§II, III);
- We propose a distributed Tikhonov regularization-based algorithm that exploit a finite number uncertainty realizations to perform the stochastic approximation (§IV);
- We show that the algorithm converges to a SGNE i) under mere monotonicity of the pseudogradient mapping, and ii) with a careful choice of the step size sequence that exploits the structure of the problem (see Table I).

| MONOTONICITY | Tik | SpFB [3] | SPRG [9] | SRFB [8] |
|--------------|-----|----------|----------|----------|
| # SAMPLES    | 1   | $N_k$    | $N_k$    | $N_k$    |
| STEP SIZE    | $\alpha_k$ | $\alpha$ | $\alpha$ | $\alpha$ |

TABLE I: FB-based algorithms for SGNEPs which converge with monotonicity (✓) or stronger assumptions (✗). $N_k$ indicates an increasing number of samples, while $\alpha_k$ the time-varying step size sequence, as opposed to $\alpha$. 

At the time of the work, F. Fabiani was with the Department of Engineering Science, University of Oxford, OX1 3PJ, United Kingdom. Currently, he is with the IMT School for Advanced Studies Lucca, Piazza San Francesco 19, 55100 Lucca, Italy (filippo.fabiani@imtlucca.it). B. Franci is with the Department of Data Science and Knowledge Engineering, Maastricht University, P.O. Box 616, NL–6200 MD Maastricht, The Netherlands (b.franci@maastrichtuniversity.nl).
We corroborate our theoretical results on a numerical instance of the proposed ride-hailing competition model (§V).

A. Notation and Preliminaries

a) Notation: \( \mathbb{N} \) indicates the set of natural numbers and \( \mathbb{R} \) is the set of real numbers. \( \langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \) denotes the standard inner product and \( \| \cdot \| \) is the associated Euclidean norm. Given a vector \( x \in \mathbb{R}^n \), \( x_{\min} = \min_{i=1,\ldots,n} x_i \). We indicate that a matrix \( A \) is positive definite with \( A > 0 \). Given a symmetric \( W > 0 \), the \( W \)-induced inner product is \( \langle x, y \rangle_W = (Wx, y) \) and the associated norm is \( \| x \|_W = \sqrt{(Wx, x)} \). \( I_d \) is the identity operator. \( i_X \) is the indicator function of the set \( X \) and the set-valued mapping \( N_x : \mathbb{R}^n \to \mathbb{R}^n \) denotes the associated normal cone operator [10].

b) Operator theory: Let \( \text{gra}(F) = \{(x, u) : u \in F(x)\} \) be the graph of \( F : X \subseteq \mathbb{R}^n \to \mathbb{R}^n \). Then, \( F \) is said to be: monotone on \( X \) if \( \langle F(x) - F(y), x - y \rangle \geq 0 \), for all \( x, y \in X \); \( \mu \)-strongly monotone on \( X \) if there exists a constant \( \mu > 0 \) such that \( \langle F(x) - F(y), x - y \rangle \geq \mu \| x - y \|^2 \), for all \( x, y \in X \); maximal monotone if there exists no monotone operator \( G : X \to \mathbb{R}^n \) such that \( \text{gra} G \) properly contains \( \text{gra} F \); \( \ell \)-Lipschitz continuous with constant \( \ell > 0 \) if for all \( x, y \in X \), \( \| F(x) - F(y) \| \leq \ell \| x - y \| \). The projection operator onto \( C \) is the operator defined as \( \text{proj}_C(x) = \arg \min_{z \in C} \| z - x \|^2 \).

c) Stochastic setting: \( \mathbb{E}_\xi \) represent the mathematical expectation with respect to the distribution of the random variable \( \xi(\omega) \), in the probability space \((\mathbb{E}, \mathcal{F}, \mathbb{P})\). We avoid expressing the dependency on \( \omega \) when clear from the context and often simply write \( \mathbb{E} \). We use i.i.d. to indicate independent identically distributed random variables.

II. On-Demand Competing Ride-Hailing Firms

We take inspiration from [13]–[15] to examine how \( N \in \mathbb{N} \) on-demand competing ride-hailing platforms design their pricing strategies under a regulated pricing scenario and the presence of uncertainties. Specifically, given a continuum of potential riders of mass \( C_h > 0 \) for each area of interest \( h \in \mathcal{H} \subseteq \mathbb{N} \) (e.g., suburbs, city centres, airports), each firm \( i \in \mathcal{I} := \{1, \ldots, N\} \) aims at maximizing its profit by setting:

i) a price \( p_{i,h} \geq 0 \) for the on-demand ride-hailing service to attract as many customers as possible in the \( h \)-th area, with

\[
\frac{1}{N} \sum_{i \in \mathcal{I}} p_{i,h} \leq \bar{p}_h, \quad \forall h \in \mathcal{H}, \tag{1}
\]

capping the averaged maximum price allowed \( \bar{p}_h \geq 0 \), usually imposed by consumers’ associations; ii) a wage \( w_{i,h} \geq \bar{w} > 0 \) for the registered drivers on the \( i \)-th ride-hailing platform, which is also typically regulated by institutions [13], [16], to meet the resulting customers’ demand. As one may expect, since profit maximization is a consideration, \( p_{i,h} \) and \( w_{i,h} \) shall be necessarily interdependent.

Similar to [14], [17], we assume that the fraction of customers who choose the \( i \)-th platform’s service in the \( h \)-th area, i.e., the demand for the \( i \)-th firm, is characterized as:

\[
d_{i,h} = \frac{K_{i,h}}{\sum_{j \in \mathcal{I}_h \setminus \{i\}} K_{j,h}} \left( \bar{p} - p_{i,h} + \frac{\theta_i}{\sum_{j \in \mathcal{I}_h \setminus \{i\}} p_{j,h}} \right), \tag{2}
\]

where \( K_{i,h} \geq 0 \) denotes the number of registered drivers on the \( i \)-th ride-hailing platform who prefer to work in the \( h \)-th area, \( \bar{p} > \max_{h \in \mathcal{H}} p_h \) is a maximum service price, and \( \theta_i \in [0,1] \) models the substitutability of the service provided by each firm. When \( \theta_i \) is close to 0, the service of the \( i \)-th platform is almost independent from the others, while \( \theta_i \) close to 1 means that it is fully substitutable, thus obtaining a perfect competition market. We discuss in details the role played by this parameter in §V with a numerical example.

The demand request in (2) does not account for the willingness of the drivers to actually provide a service, which is key to meet the customers’ demand and hence maximize the profit. In fact, any driver provides service in a prescribed area \( h \in \mathcal{H} \) only if its earn is greater than their opportunity cost, here denoted by \( \delta_{i,h} \), which we let coincide with a random variable. For each firm and area, we assume the wage be given by \( w_{i,h} = \beta p_{i,h} \), where the parameter \( \beta \geq 0 \) denotes the commission ratio that the platform should pay to its driver, typically regulated by a third party (e.g., governments [13]). Thus, for all \( i \in \mathcal{I} \) and \( h \in \mathcal{H} \), we introduce the effective demand as

\[
d_{i,h}^{\text{eff}} = d_{i,h} \mathbb{P}[w_{i,h} \geq \delta_{i,h}], \tag{3}
\]

which coincides with the portion of customers’ demand for which a ride-hailing company can actually claim a payment. Roughly speaking, a driver is willing to provide a service only if the wage she gets matches (at least) her expectations. This directly leads us to define the fraction of drivers that give services on the \( i \)-th platform as:

\[
k_{i,h} = K_{i,h} \mathbb{P}[w_{i,h} \geq \delta_{i,h}], \tag{4}
\]

As a consequence, the effective demand in (3) can be equivalently obtained from (2) by replacing \( K_{i,h} \) with \( k_{i,h} \). However, it is unlikely that the \( i \)-th firm is aware of the total number of potential drivers, \( \sum_{j \in \mathcal{I}} K_{j,h} \), both for privacy reasons and possible multiple registrations. Thus, we define \( C_h(\xi) = \sum_{j \in \mathcal{I}} K_{j,h} \) as the unknown fraction of passengers in area \( h \in \mathcal{H} \), which allows us to explicitly account for the uncertain parameter \( \xi \). In accordance, the demands in (2) and (3) turn into random variables \( d_{i,h}(\xi) \) and \( d_{i,h}^{\text{eff}}(\xi) \).

As commonly adopted in the literature [13], [15], we restrict attention to the case where the drivers’ willingness is uniformly distributed in \( [\bar{w}, \bar{w}_{i,h}] \), so that \( \mathbb{P}[w_{i,h} \geq \delta_{i,h}] = (\bar{w}_{i,h} - \bar{w})/(\bar{w}_{i,h} - \bar{w}) \). This not only allows us to consider one source of uncertainty, but also to make the constraint in (4) convex. Note that a similar argument can be adopted for any distribution concave in the induced demand \( d_{i,h}(\xi) \) (e.g., exponential, Pareto). Then, the stochastic optimization problem associated to each platform amounts to:

\[
\forall i \in \mathcal{I}: \left\{ \begin{array}{l}
\max_{(p_{i,h}, w_{i,h}) \in \mathcal{H}} \mathbb{E}_\xi \left[ \sum_{h \in \mathcal{H}} (p_{i,h} d_{i,h}^{\text{eff}}(\xi) - w_{i,h} k_{i,h}) \right] \\
\text{s.t.} \quad (1), (3), (4), p_{i,h} \geq 0, \forall h \in \mathcal{H}, \\
w_{i,h} \in [\bar{w}, \bar{w}_{i,h}], \forall h \in \mathcal{H}.
\end{array} \right. \tag{5}
\]

The first part of the cost function amounts to the profit of the \( i \)-th firm to provide a service to the customers, while the second one considers the costs for providing a service to the drivers. Unlike [13], [14], [17], the cost
functions in (5) accounts for the number of actual drivers who decide to provide a service rather than the whole fleet of registered ones, $K_{i,h}$. After replacing the equality constraints in (3) and (4), we obtain a collection of mutually coupled stochastic optimization problems, where the cost functions (cubic in $p_{i,h}$, due to the distribution of $\delta_{i,h}$) are affected by the uncertain fraction of potential riders, i.e., $C_h(\xi)$. The proposed model relies on a common assumption in the literature: those riders who do not get assigned to a driver when they seek service from the ride-hailing platforms, e.g., because of excess demand for rides, use a different mean of transportation. The need for a common platform handling competition in ride-hailing mobility to satisfy the customers' demand has been indeed recently explored in, e.g., [18], [19]. Throughout the paper we treat this model as a SGNEP, and we propose an algorithm to compute an equilibrium solution.

III. STOCHASTIC GENERALIZED NASH EQUILIBRIUM PROBLEM

To compact the notation, we rewrite the problem in (5) as

$$\forall i \in \mathcal{I} : \min \{ J_i(x_i, x_{-i}) | x_i \in \Omega_i, g(x_i, x_{-i}) \leq 0 \}. \quad (6)$$

where $x_i = \text{col}((p_{i,h})_{h \in \mathcal{H}}) \in \mathbb{R}^n$, $n := |\mathcal{H}|$, $x = \text{col}((x_i)_{i \in \mathcal{I}}) \in \mathbb{R}^{nN}$, and $x_{-i} = \text{col}((x_j)_{j \neq i})$. Moreover, we indicate the set of local constraints of firm $i$ as $\Omega_i \in \mathbb{R}^n$ and let $\Omega := \prod_{i \in \mathcal{I}} \Omega_i$. On the other hand, the set of coupling constraints arising from (1) in a general form read as

$$\mathcal{X} := \Omega \cap \{ y \in \mathbb{R}^{nN} | g(y) \leq 0_m \}, \quad (7)$$

where $g : \mathbb{R}^{nN} \to \mathbb{R}^m$. We indicate with $\mathcal{X}_i(x_{-i})$ the piece of coupling constraints corresponding to agent $i$, which is affected by the decision variables of the other agents $x_{-i}$. We stress that the formulation of the SGNEP in (6) is standard [2], [8], and hence the theory we develop applies to all the SGNEPs satisfying the assumptions introduced next.

Assumption 1: For each $i \in \mathcal{I}$, the set $\Omega_i$ is nonempty, closed and convex. The set $\mathcal{X}$ satisfies Slater’s constraint qualification.

Assumption 2: The mapping $g$ in (7) has a separable form, i.e., $g(x) := \sum_{i=1}^N g_i(x_i)$, for convex differentiable functions $g_i : \mathbb{R}^n \to \mathbb{R}$, $i \in \mathcal{I}$ and it is $\ell_2$-Lipschitz continuous. Its gradient $\nabla g$ is bounded, i.e., $\sup_{x \in \mathcal{X}} \| \nabla g(x) \| \leq B \nabla g$. □

Given the stochastic nature of our collection of problems, we indicate the cost function of each agent $i \in \mathcal{I}$ as $J_i(x_i, x_{-i}) := \mathbb{E}[J_i(x_i, x_{-i}, \xi)]$, for some measurable function $J_i : \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}$. We assume that $\mathbb{E}[J_i(x, \xi)]$ is well defined for all feasible $x \in \mathcal{X}$ [20].

Assumption 3: For all $i \in \mathcal{I}$ and $x_{-i} \in \mathcal{X}_{-i}$ the function $J_i(\cdot, x_{-i})$ is convex and continuously differentiable. □

The goal of the firms is hence to solve (6) to find a SGNE, i.e., a strategy profile where no agent can decrease its cost function by unilaterally deviating from its decision. Formally, a SGNE is a collective vector $x^* \in \mathcal{X}$ such that for all $i \in \mathcal{I}$,

$$J_i(x^*_i, x^*_{-i}) \leq \inf \{ J_i(g, x^*_{-i}) | y \in \mathcal{X}_i(x^*_{-i}) \}. \quad (8)$$

Existence of a SGNE for the game in (6) is guaranteed under suitable assumptions [20, §3.1], though uniqueness does not hold in general [20, §3.2]. Within all possible Nash equilibria, we focus on those coinciding with the solutions of a suitable stochastic variational inequality (SVI) [5]. Thus, we introduce the pseudogradient mapping of the game as

$$\bar{F}(x) := \text{col} \left( (\mathbb{E}[\nabla x_i J_i(x_i, x_{-i})])_{i \in \mathcal{I}} \right). \quad (8)$$

Flipping the expected value and the gradient follows from differentiability of $J_i(\cdot, x_{-i})$ (Assumption 3) [21, Th. 7.44]. Then, the SVI associated to the SGNEP in (6) reads as

$$\langle \bar{F}(x^*), x - x^* \rangle \geq 0, \quad \text{for all } x \in \mathcal{X}. \quad (9)$$

If Assumptions 1–3 hold, any solution to $\text{SVI}(\mathcal{X}, \bar{F})$ in (9) is a SGNE of the game in (6), while the converse does not necessarily hold. A game may have a Nash equilibrium while the associated (S)VI may have no solution [22, Prop. 12.7].

Assumption 4: The SVI in (9) admits solution. □

We call variational equilibria (v-SGNE) the SGNE that are also solution to $\text{SVI}(\mathcal{X}, \bar{F})$ in (9) with $\bar{F}$ in (8) and $\mathcal{X}$ in (7). These equilibria can be characterized in terms of the Karush–Kuhn–Tucker (KKT) conditions of the coupled optimization problems in (6), i.e., a $x^*$ is a v-SGNE if and only if the following inclusion is satisfied for $\lambda \in \mathbb{R}^m_{\geq 0}$ [12, Th. 4.6]:

$$0 \in \mathcal{T}(x, \lambda) := \left[ \bar{F}(x) + N_{\Omega}(x) + \nabla g(x)^\top \lambda \right] \left[ N_{\mathbb{R}^m_{\geq 0}}(\lambda) - \nabla g(x) \right], \quad (10)$$

where $\mathcal{T} : \mathcal{X} \times \mathbb{R}^m_{\geq 0} \to \mathbb{R}^{nN} \times \mathbb{R}^m$ is a set-valued mapping. According to [23, Th. 3.1], [24, Th 3.1], the v-SGNE are those equilibria such that the shared constraints have the same dual variable for all the agents, i.e., $\lambda_i = \lambda$ for all $i \in \mathcal{I}$, and solve the $\text{SVI}(\mathcal{X}, \bar{F})$ in (9). Then, the v-SGNE of the game in (6) correspond to the zeros of $\mathcal{T}$, which can be split as the sum of two operators, $\mathcal{T} = A + B$, with

$$A : \begin{bmatrix} x \\ \lambda \end{bmatrix} \mapsto \begin{bmatrix} \bar{F}(x) \\ 0 \end{bmatrix} + \begin{bmatrix} \nabla g(x)^\top \lambda \\ -g(x) \end{bmatrix}, \quad (11)$$

B : \begin{bmatrix} x \\ \lambda \end{bmatrix} \mapsto \begin{bmatrix} N_{\Omega}(x) \\ N_{\mathbb{R}^m_{\geq 0}}(\lambda) \end{bmatrix}.$$

IV. DISTRIBUTED STOCHASTIC TIKHONOV RELAXATION

We now discuss in details the sequence of instructions summarized in Algorithm 1. For the local decision variable $x_i$ the projection onto $\Omega_i$ guarantees that the local constraints are always satisfied, while the coupling constraints are enforced asymptotically through the (nonnegative, due to the projection onto $\mathbb{R}^m_{\geq 0}$) dual variable $\lambda_i$. The auxiliary variable $z_i$, instead, forces consensus on the dual variables [2], [23].

The set of agents $j$ whose decision variables explicitly affect the cost function of agent $i$ are denoted by $\mathcal{N}_i^L$. Let us then introduce the graph $G^\lambda = (\mathcal{I}, \mathcal{E}^\lambda)$ through which a local copy of the dual variable is shared, along with of the auxiliary one, $z_i \in \mathbb{R}^m$. The set of edges $\mathcal{E}^\lambda$ of the multiplier graph $G^\lambda$, is given by: $(i,j) \in \mathcal{E}^\lambda$ if player $j$ shares its $\{\lambda_j, z_j\}$ with player $i$. For all $i \in \mathcal{I}$, the neighboring agents in $G^\lambda$ form the set $\mathcal{N}_i^L = \{ j \in \mathcal{I} : (i,j) \in \mathcal{E}^\lambda \}$. Under these premises, Algorithm 1 is distributed in the sense that each agent knows its own problem data and communicates with
Algorithm 1 Distributed stochastic Tikhonov relaxation

Initialization: $x_0^i ∈ Ω_i, λ_0^i ∈ ℜ^m_≥0$, and $z_0^i ∈ ℜ^m$ for $i ∈ N^p_i$, then updates:

$$x_{i+1} = \text{proj}_Ω(x_i - α_i^j γ_i (F_i(x_i, k^i, ξ^i) + \nabla g_i(x_i)\top λ_i^k + ε_i^j x_i))$$

$$z_{i+1} = z_i - α_i^j ν_i (\sum_{j∈N_i^k} w_{ij}(λ_i^k - λ_j^k) + ε_i^j z_i)$$

$$λ_{i+1} = \text{proj}_{N^m_≥0}(λ_i + α_i^j τ_i (g_i(x_i) - ε_i^j λ_i^k) + α_i^j τ_i \sum_{j∈N^k_i} w_{ij}[(z_i^k - z_j^k) - (λ_i^k - λ_j^k)])$$

By exploiting the approximation in (13) of the expected value mapping $F$, we replace the operator $A$ in (12) with $A$, the operator that involves $F_i(x, ξ)$ in place of $F(x)$; the remaining terms of $A$ are not affected by the approximation.

Algorithm 1 can now be rewritten in compact form as $u_{k+1} = (Id + Φ_k^{-1} B)^{-1}(u_k - Φ_k^{-1} (A u_k + ε_k^i u_k))$, where $Φ_k > 0$ contains the inverse of step size sequences $Φ_k = α_k^{-1} Φ = α_k^{-1} \text{diag}(γ^{-1}, ν^{-1}, τ^{-1})$, (14) with $γ^{-1}$, $ν^{-1}$, $τ^{-1}$ being diagonal matrices, and $ε_k = \text{diag}(ε_k^i) ∈ ℜ^TxT$, $T = nN + 2Nm$, contains the regularization steps. This last term is what typically characterizes the Tikhonov regularization scheme [5, 10].

A. Convergence analysis

We now study the convergence properties of Algorithm 1. First, to ensure that $A$ and $B$ have the properties that we use for the analysis, we make the following assumption:

Assumption 6: $F$ in (8) is monotone and $ℓ_f$-Lipschitz continuous for some $ℓ_f > 0$. □

Lemma 1: [8, Lemma 2 and 4] Let Assumptions 5 and 6 hold true, and let $Φ > 0$. Then, the operators $A$ and $B$ in (12) have the following properties.

1) $A$ is monotone and $ℓ_A$-Lipschitz continuous.
2) $B$ is maximally monotone.
3) $Φ^{-1} A$ is monotone and $ℓ_A$-Lipschitz continuous.
4) $Φ^{-1} B$ is maximally monotone. □

The Lipschitz constants in Lemma 1 depends on those of $g_i$, $i ∈ I$, and $F$ (Assumptions 1 and 6, respectively). Their specific expressions, however, are not relevant for our analysis, and therefore we point to [8] for additional details.

Remark 2: Introducing the regularization term makes $A + ε_k$ strongly monotone [10, Th. 12.2.3]. Thus, mere monotonicity of the pseudogradient mapping is enough to show convergence. In addition, note that the regularization term is added to the operator $A$ and not to force strong monotonicity of the SVI in (9). In fact, starting from a strongly monotone pseudogradient, the resulting operator $A$ can be at most cocoercive [2, Lemma 5 and 7], [25, Lemma 2 and 4]. □

Taking few samples as in (13) is realistic and computationally tractable, at the price of requiring an additional assumption on the step sizes. Specifically, we indicate next how to choose the parameters in Algorithm 1 to ensure that they are vanishing to control the approximation error [5].

Assumption 7: The step size sequence $(α^k)_{k∈N}$ and the regularizing sequences $(ε^k)_{k∈N}$, $j = 1, ..., T$, are such that $α^k = (k + ε^k)^{-a}$ and $ε^k = (k + ζ^k)^{-b}$ for $k ≥ 0$, where each $α$ and $ζ$ are selected from a uniform distribution on the intervals $[α, η]$ and $[ζ, ξ]$, respectively, for some $0 < η < η$ and $0 < ξ < ζ$ and $a, b ∈ (0, 1)$, $a + b < 1$, and $a > b$. □

Let us define the filtration $F = \{F_k\}_{k∈N}$, that is, a family of $σ$-algebras such that $F_0 = σ(X_0)$ and $F_k = σ(X_{k}, ξ_1, ξ_2, ..., ξ_k)$ for all $k ≥ 1$, such that $F_k ⊆ F_{k+1}$ for all $k ∈ N$. In words, $F_k$ contains the information up to iteration $k$. Since we consider the approximation in (13), we let $Δ^k = col(Δ^k, 0, 0)$ be the stochastic error, i.e.,
$\Delta^k = \tilde{A}(u^k) - \hat{A}(u^k, \xi^k)$. Then, an additional assumption is finally needed to regulate its asymptotic behaviour.

**Assumption 8:** The step size sequence $(\alpha^k)_{k \in \mathbb{N}}$, the regularization sequence $(\epsilon^k)_{k \in \mathbb{N}}$, and the stochastic error $\Delta^k$ satisfy $\lim_{k \to \infty} (\alpha^k / \epsilon_{k, \min}) \mathbb{E}[\|\Delta^k\|^2_2 \mid F_k] = 0$ and $\sum_{k=0}^\infty \alpha^k \mathbb{E}[\|\Delta^k\|^2_2 \mid F_k] < \infty$ a.s.. □

We are now ready to state our main convergence result.

**Theorem 1:** Let Assumptions 1 - 8 hold true. Then, the sequence $(x^k)_{k \in \mathbb{N}}$ generated by Algorithm 1 with $\tilde{F}$ as in (13) converges a.s. to a v-SGNE in the game in (6). □

**Proof:** Convergence to the primal-dual solution follows similarly to [5] by using the $\Phi$-induced metric. Specifically, with similar steps as in [5, Prop. 1 and 2] and by letting $y^k$ be the sequence generated by the centralized Tikhonov method [5, Lemma 3], we obtain $\mathbb{E}[\|u^{k+1} - y^k\|^2_2 \mid F_k] \leq (1-c(\alpha^k, \epsilon^k))\|u^{k+1} - y^k\|^2_2 + d(\alpha^k, \epsilon^k) + \alpha^2 k \mathbb{E}[\|\Delta^k\|^2_2 \mid F_k]$, where $d(\alpha^k, \epsilon^k)$ and $\alpha^2 k \mathbb{E}[\|\Delta^k\|^2_2 \mid F_k]$ vanish as $k \to \infty$ (Assumptions 7 and 8) and $c(\alpha^k, \epsilon^k) \in [0, 1]$ is such that [26, Lemma 4.7] can be applied to conclude that $\lim_{k \to \infty} \|u^{k+1} - y^k\|^2 = 0$. Then, the statements recalled in Remark 1 guarantee convergence of $(x^k)_{k \in \mathbb{N}}$ to a v-SGNE of the SGNEP in (6).

**Remark 3:** Note that our proof does not follow directly from the one of [5, Prop. 1 and 2], which is similar, albeit tailored for SNPEs, and thus not applicable to a generalized game setting. Our proof therefore applies to a generic zero-finding problem of a monotone inclusion $0 \in \tilde{A}(u) + \tilde{B}(u)$, independently from the fact that it comes from the KKT conditions associated to (10). This also explains the need of a step size matrix $\Phi_k$ with structure as in (14). □

### B. Discussion on the variable step size sequence

We traditionally identify two main approaches to perform stochastic approximations: either we take one sample as in (13), or we take the average over possibly infinite number of realizations. The idea behind taking a larger number is that the variance of the stochastic error disappears with the iterations [6], [7]. By following classic results in convergence analysis [2], [3], [8], however, it turns out that this latter approach is computationally expensive, and hence taking just one or a finite number of samples is preferable. This simplification comes at the price of choosing a vanishing step size to control the approximation error, i.e., the time-varying matrix $\Phi_k$, thus possibly a time-varying metric for the convergence analysis. Although convergence can be guaranteed in such cases [26], [27], some additional assumptions on the metric should be satisfied. Specifically, $\Phi_k$ should be chosen so that

$$\sup_{k \in \mathbb{N}} \Phi_k^{-1} < \infty \land \forall k > 0 \ (1 + \eta_k) \Phi_k^{-1} \geq \Phi_k^{-1}, \quad (15)$$

where $(\eta^k)_{k \in \mathbb{N}}$ is a nonnegative sequence such that $\sum_k \eta^k < \infty$. Unfortunately, this contradicts the fact that the step size sequence should be decreasing. Loosely speaking, the motivation for (15) stems from the fact that, with a variable metric, it is hard to prove whether the algorithm converges to a zero of the mapping or to a zero of the sequence. On the other hand, given the specific structure of our matrix $\Phi_k$, we overcome this issue by using a fixed matrix $\Phi$ as a metric, and then by pre-multiplying $\Phi$ with a vanishing step as in (14). Note that this formulation allows us to preserve the distributed nature of the algorithm.

We also note that a “separable” matrix as $\Phi_k$ in (14) cannot be used for every iterative distributed algorithm. In the standard SFB [3], for instance, the matrix $\Phi_k$ serves as a preconditioning matrix that has non-zero off-diagonal entries. Pre-multiplying such entries for a quantity (even fixed) would compromise the corresponding KKT conditions, thus making impossible to produce distributed iterations.

### V. Numerical simulations

We now validate both the model in §II and Algorithm 1 numerically. Specifically, we consider an instance of the competition among $N = 5$ on-demand ride-hailing firms over $|H| = 10$ areas with main parameters in Table II.

First, we test the effect of the step size sequence on the convergence of Algorithm 1 by recalling that, in view of Assumption 7, $\alpha^k = (k + \eta)^{-\alpha}$. In particular, Fig. 1a shows how the rate of convergence is affected by the exponent $\alpha$ while in Fig. 1b we plot the effect of the base $\eta$. In the first example, we choose $\eta = 10^8$ and $\epsilon^k = (k + 10^8)^{-0.15}$, while for the second one $\eta = 0.7$ and $\epsilon^k = (k + 10^8)^{-0.2}$. The thick lines indicate the average performance while the transparent areas are the variability over 10 runs of the algorithm.

Successively, we examine how the competition parameters $\theta_i, i \in I$, affect the equilibrium solution of the SGNEP in (5) against $K = 1000$ random realizations of $\delta_i, h$ In view of (4), given some drivers’ expectation level $\delta^{(i,j)}_h \sim \mathcal{U}(w_i, \bar{w}_{i,j})$ and an equilibrium strategy profile $x^* = \text{col}((x^*_i)_{i \in I})$ with $x^*_i = \text{col}((\tilde{p}^*_i, h_i)_{h_i \in H})$, if $\delta^{(i,j)}_h \leq \beta \tilde{p}^*_i$, then $k^{(i,j)}_{i,j} = K_{i,j}$. The number of drivers providing a service allows each firm to actually request a payment from the customers, thus earning $J_i^{(i,j)} := \mathbb{E}_\xi \left[ \sum_{h_i \in H} p^*_{i,h} (d^{(i,j)}_h (\xi) - \beta k^{(i,j)}_{i,j}) \right]$. This value, however, can be different (possibly smaller) from the expected profit $J_i := \mathbb{E}_\xi \left[ \sum_{h \in H} p^*_{i,h} (d^{(i)}_h (\xi) - \beta k_{i,h}) \right]$, i.e., the value function associated to the equilibrium strategy profile. Then, for each firm $i \in I$, Fig. 1c shows how the ratio between these two quantities changes with the level of competition in the ride-hailing market. Specifically, we note that when firms operate in an almost oligopoly regime, i.e., $\theta_i \in [0.2, 0.4]$ for all $i \in I$, the averaged percentage is small and grows when the market becomes more competitive. For small values of $\theta_i$, indeed, each company is almost independent and tends to select prices to match the lower bound $w / \beta$ – this also
coincides with the strategy adopted by firms with only few registered drivers, thus explaining the behavior of firms 2 and 5 – whereas for larger values of $\theta$, the firms are entitled to significantly raise their prices, thus allowing to meet drivers’ expectations with a higher probability. Such trend is confirmed by the numerical results shown in Fig. 1d, which reports the customers’ demand satisfaction, measured as $\sum_{i \in I} k_{i,h}^{(j)}/C_h$, for each area $h \in \mathcal{H}$. In almost oligopoly regimes the customers’ request is met few times only, while it grows significantly when also the competition increases.

VI. CONCLUSION

We have proposed a model for the ride-hailing market under a regulated pricing scenario and uncertainties involving several platforms that compete to offer mobility services. To optimize the operations of these interdependent problems, we have recast the model as a stochastic Nash equilibrium problem for which we have designed a distributed, Tikhonov regularization-based algorithm that enjoys convergence guarantees to a Nash equilibrium. The proposed algorithm exploits only a finite number of samples of the uncertainty to perform the stochastic approximation, and it requires mere monotonicity of the game mapping to establish convergence.

REFERENCES

[1] NYC OpenData. (2022) For-hire vehicle base aggregate report. [Online] https://opendata.cityofnewyork.us/.
[2] P. Yi and L. Pavel, “An operator splitting approach for distributed generalized Nash equilibria computation,” Automatica, vol. 102, pp. 111–121, 2019.
[3] B. Franci and S. Grammatico, “A distributed forward-backward algorithm for stochastic generalized Nash equilibrium seeking,” IEEE Transactions on Automatic Control, 2020.
[4] H. Robbins and S. Monro, “A stochastic approximation method,” The Annals of Mathematical Statistics, pp. 400–407, 1951.
[5] J. Koshal, A. Nedic, and U. V. Shanbhag, “Regularized iterative stochastic approximation methods for stochastic variational inequality problems,” IEEE Transactions on Automatic Control, vol. 58, no. 3, pp. 594–609, 2013.
[6] A. Iusem, A. Jofré, R. I. Oliveira, and P. Thompson, “Extragradient method with variance reduction for stochastic variational inequalities,” SIAM Journal on Optimization, vol. 27, no. 2, pp. 686–724, 2017.
[7] R. I. Boj, P. Mertikopoulos, M. Staudigl, and P. T. Vuong, “Minibatch forward-backward-forward methods for solving stochastic variational inequalities,” Stochastic Systems, vol. 11, no. 2, pp. 112–139, 2021.
[8] B. Franci and S. Grammatico, “Stochastic generalized Nash equilibrium seeking in merely monotone games,” IEEE Transactions on Automatic Control, 2021.
[9] ——, “Distributed projected-reflected-gradient algorithms for stochastic generalized Nash equilibrium problems,” in 2021 European Control Conference (ECC). IEEE, 2021, pp. 369–374.
[10] F. Facchinei and J.-S. Pang, Finite-dimensional variational inequalities and complementarity problems. Springer Science & Business Media, 2007.
[11] A. Kannan and U. V. Shanbhag, “Distributed computation of equilibria in monotone Nash games via iterative regularization techniques,” SIAM Journal on Optimization, vol. 22, no. 4, pp. 1177–1205, 2012.
[12] F. Facchinei and C. Kanzow, “Generalized Nash equilibrium problems,” Annals of Operations Research, vol. 175, no. 1, pp. 177–211, 2010.
[13] Y. Zhong, T. Yang, B. Cao, and T. Cheng, “On-demand ride-hailing platforms in competition with the taxi industry: Pricing strategies and government supervision,” International Journal of Production Economics, vol. 243, p. 108301, 2022.
[14] E. J. He, S. Savin, J. Goh, and C.-P. Teo, “Off-platform threats in on-demand services,” Available at SSRN 3550640, 2020.
[15] K. Bimpikis, O. Candogan, and D. Saban, “Spatial pricing in ride-sharing networks,” Operations Research, vol. 67, no. 3, pp. 744–769, 2019.
[16] R. Beer, C. Brakewood, S. Rahman, and J. Visconti, “Qualitative analysis of ride-hailing regulations in major American cities,” Transportation Research Record, vol. 2650, no. 1, pp. 84–91, 2017.
[17] T. W. McGuire and R. Staelin, “An industry equilibrium analysis of downstream vertical integration,” Marketing Science, vol. 2, no. 2, pp. 161–191, 1983.
[18] V. Pandey, J. Montiel, C. Gambella, and A. Simonetto, “On the needs for MaaS platforms to handle competition in ridesharing mobility,” Transportation Research Part C: Emerging Technologies, vol. 108, pp. 269–288, 2019.
[19] F. Fabiani, A. Simonetto, and P. J. Goulart, “Personalized incentives as feedback design in generalized Nash equilibrium problems,” IEEE Transactions on Automatic Control, 2021, (Under review – available at arxiv.org/abs/2203.12948).
[20] U. Ravat and U. V. Shanbhag, “On the characterization of solution sets of smooth and nonsmooth convex stochastic Nash games,” SIAM Journal on Optimization, vol. 21, no. 3, pp. 1168–1199, 2011.
[21] A. Shapiro, D. Dentcheva, and A. Ruszczynski, Lectures on stochastic programming: modeling and theory. SIAM, 2021.
[22] D. P. Palomar and Y. C. Eldar, Convex optimization in signal processing and communications. Cambridge university press, 2010.
[23] F. Facchinei, A. Fischer, and V. Piccialli, “On generalized Nash games and variational inequalities,” Operations Research Letters, vol. 35, no. 2, pp. 159–164, 2007.
[24] A. Auslender and M. Teboulle, “Lagrangian duality and related multiplier methods for variational inequality problems,” SIAM Journal on Optimization, vol. 10, no. 4, pp. 1097–1115, 2000.
[25] B. Franci and S. Grammatico, “A damped forward-backward algorithm for stochastic generalized Nash equilibrium seeking,” in 2020 European Control Conference (ECC). IEEE, 2020, pp. 1117–1122.
[26] ——, “Convergence of sequences: A survey,” Annual Reviews in Control, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1367578822000037
[27] P. L. Combettes and B. C. Vu, “Variable metric forward-backward splitting with applications to monotone inclusions in duality,” Optimization, vol. 63, no. 9, pp. 1289–1318, 2014.