2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

Is Stochastic Mirror Descent Vulnerable to Adversarial Delay Attacks? A Traffic Assignment Resilience Study

Yunian Pan, Tao Li, and Quanyan Zhu*

***Abstract*— *Intelligent Navigation Systems* (INS) are exposed to an increasing number of informational attack vectors, which often intercept through the communication channels between the INS and the transportation network during the data collecting process. To measure the resilience of INS, we use the concept of a Wardrop Non-Equilibrium Solution (WANES), which is characterized by the probabilistic outcome of learning within a bounded number of interactions. By using concentration arguments, we have discovered that any bounded feedback delaying attack only degrades the systematic performance up to order $\tilde{\mathcal{O}}(\sqrt{d^3T^{-1}})$ along the traffic flow trajectory within the Delayed Mirror Descent (DMD) online-learning framework. This degradation in performance can occur with only mild assumptions imposed. Our result implies that learning-based INS infrastructures can achieve Wardrop Non-equilibrium even when experiencing a certain period of disruption in the information structure. These findings provide valuable insights for designing defense mechanisms against possible jamming attacks across different layers of the transportation ecosystem.**

I. Introduction

The real-time routing demand has been significantly growing with the rapid development of the modern *Intelligent Navigation Systems* (INS), in which typical *Online Navigation Platforms* (ONP), such as Google Maps and Waze, receive billions of routing requests per second. It is, therefore, crucial to provide reliable and efficient navigation services for active users, such that the ex-post routing regret is small. The regret-free routing for individuals gives rise to special macroscopic traffic conditions, commonly known as the *Wardrop equilibrium* (WE) [1] in *congestion games*. The seeking of WE is referred to as *traffic assignment* problem. However, the increasing connectivity of transportation networks exposes the INS to a wide variety of *informational attacks* [2].

This paper focuses on a class of *information-delaying attacks* against the INS that aim to intercept the communication channel between the data source/individual users and the navigation center, **delaying** the delivery of traffic condition information for adversarial purposes. A quintessential attack surface is the data transmission process, including the network jamming attacks [3], which are often implemented by sending high-frequency wireless interference, or a sheer volume of network packets to the target communication channels/servers. With the *information delays* of critical traffic conditions, the INS infrastructures are at risk of making improper routing recommendations and misguiding users.

*The authors are with the Department of Electrical and Computer Engineering, Tandon School of Engineering, New York University, Brooklyn, NY, 11201 USA; E-mail: {yp1170,tl2636,qz494}@nyu.edu. This work is partially supported by grants ECCS-1847056 and BCS-2122060 from National Science Foundation (NSF) and grant W911NF-19-1-0041 from Army Research Office (ARO).

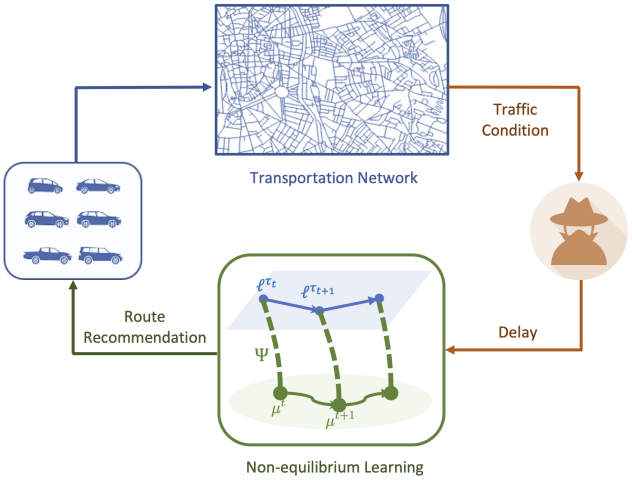

Fig. 1. The traffic conditions are withheld at each timestep due to the presence of the attacker, creating a timing disparity between the traffic flow and traffic latency, disrupting the route recommendation.

Prior studies have shown that relying solely on attack detection and prevention measures is inadequate in the face of pervasive malicious factors [4], [5]. These findings highlight the need to develop resilient mechanisms to endow traffic systems with self-healing capabilities. We adopt the notion of *Wardrop Non-Equilibrium Solution* (WANES) [6], which extends the regret analysis in games [7], [8] to probabilistic setting, to investigate the provable resilience of *Mirror Descent* (MD) based INS under adversarial delays. The schematic view of this framework is illustrated in fig. 1.

To the best of our knowledge, this work is among the first endeavors to analyze stochastic mirror descent under adversarial delays, given existing works often focus on the deterministic or bandit setting. Our initial step involves analyzing the *Delayed Mirror Descent* (DMD) dynamics. DMD is nearly identical to ordinary MD, but with the added feature of delayed latency feedback. Our analysis rests on the telescoping technique and a Martingale approach, which encounters two significant challenges. The first challenge is establishing the per-iterate telescoping inequality due to the potentially large cardinality (up to T) of the delayed latency "bundle" in the general setting. The second challenge is deriving the concentration argument due to the need for special analysis of the maximum of the empirical process. We tackle the problems by making additional assumptions about the attack capacity and the uniform bound of the expected latency function. These assumptions are often practical in

traffic assignment problems.

Without further assumptions imposed, we show that the INS under DMD dynamics with carefully chosen learning parameters endure performance loss up to order $\mathcal{O}(\sqrt{d^3T^{-1}})$ with high probability, which matches the order of delay-free case where $d = 1$. When $d = \mathcal{O}(T^\alpha)$ with $\alpha < 1/3$, the performance loss order indicates a high-probability sub-linear regret bound. This result can be generalized to resilience analysis of online traffic assignments when the Beckman potential is time-varying. It further indicates the probability of a certain class of online learning problems with delayed stochastic feedback.

II. Related Work

The role of mirror descent in the congestion game frameworks has been discussed in the literature [9], [10], where the Hannan consistency in the Cesáro sense was established in the deterministic latency setting. It was later demonstrated in [11] that the convergence rate, although remaining $\mathcal{O}(T^{-1/2})$ in the stochastic domain, can be lifted to $\mathcal{O}(T^{-2})$ leveraging Nesterov's acceleration scheme in the deterministic domain. Our result matches the $T^{-1/2}$ bound in the trivial setting without feedback delays (where $d = 1$).

The gap between online learning and resilience in congestion games was filled in [6], where the setting was extended by considering the adversary and studying the probabilistic non-equilibrium outcome of learning. The adversary is assumed capable of informational manipulation, involving threats both on the physical layers (such as sensors, GPS spoofing [12],) and cyber layers (such as routing attacks and DoS attacks [13].) Our study concerns a general class of attacks that cause delayed feedback to the traffic assignment systems.

There have been versatile delay-handling strategies in the literature. One is pooling multiple independent learners together to process the buffered feedback vectors in a sequential manner, see the analysis of Joulani et al. in [14], or giving the learners the ability to learn from local loss permutations, see Shamir et al. [15]. Another is treating delays as introduced by distributed asynchronous processors, which is explored by Agarwal et al. [16]. The strategies mentioned above can be resource-intensive or make impractical assumptions. Therefore, this work builds upon simplistic models [17], and explores the stochastic feedback setting due to its absence in current literature.

III. Problem Formulation

Repeated Congestion Game A transportation network is abstracted by a directed, finite, and connected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, with nodes $\mathcal{V}$ depicting road junctions, transportation hubs, etc., and edges $\mathcal{E}$ indexing road segments, transportation segments, etc., between different node pairs, we assume that $(v, v) \notin \mathcal{E}$, for all $v \in \mathcal{V}$. The set of Origin-Destination (OD) pairs is $\mathcal{W} \subseteq \mathcal{V} \times \mathcal{V}$. Between each OD pair $w \in \mathcal{W}$ is a set of directed paths $\mathcal{P}_w$, let $\mathcal{P} := \bigcup_{w\in\mathcal{W}} \mathcal{P}_w$.

The vehicles over $\mathcal{G}$ constitute a set of infinitesimal players $\mathcal{X}$, split into distinct populations indexed by different OD pairs, i.e., $\mathcal{X} = \bigcup_{w\in\mathcal{W}} \mathcal{X}_w$ and $\mathcal{X}_w \bigcap \mathcal{X}_{w'} = \emptyset,\ \forall w, w' \in \mathcal{W}$. For each $w \in \mathcal{W}$, let m_w be the traffic demand, i.e., the number of vehicles traveling between w.

For all $w \in \mathcal{W}$, each player $x \in \mathcal{X}_w$ is equipped with an action set $\mathcal{P}_w$ and makes decisions repeatedly, at each round $t = 1, \dots, T \in \mathbb{N}_+$, each player is committed to a single path $p \in \mathcal{P}$. The action profile can be captured a flow vector $\mu \in \Delta := \{\mu \in \mathbb{R}^{|\mathcal{P}|}_{\geq 0} \,|\, \sum_{p\in\mathcal{P}_w} \mu_p = m_w,\ \forall w \in \mathcal{W}\}$. A path flow vector determines an edge flow vector $q \in \mathbb{R}^{|\mathcal{E}|}_{\geq 0}$, through the edge-path incident matrix $\Lambda = [\Lambda^1|, \dots, |\Lambda^{|\mathcal{W}|}] \in \mathbb{R}^{|\mathcal{E}|\times|\mathcal{P}|}$ such that $\Lambda^w_{e,p} = \mathbb{1}_{\{e\in p\}}, \forall e \in \mathcal{E}, w \in \mathcal{W}, p \in \mathcal{P}_w$. In a compact form, $q = \Lambda\mu$.

The total travel time of a road segment is jointly determined by the traffic flow on that road and some stochastic factors, such as weather condition and road incidents, which affects the congestion level. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the underlying probability space; $\omega \in \Omega$ encapsulates the universal latent conditions for $\mathcal{G}$. Let $l_e : \mathbb{R}_{\geq 0} \times \Omega \mapsto \mathbb{R}_+$ be the edge latency function for $e \in \mathcal{E}$, $l : \mathbb{R}^{|\mathcal{E}|}_{\geq 0} \times \Omega \mapsto \mathbb{R}^{|\mathcal{E}|}_+$ be its vector-valued extension, $\ell : \Delta \times \Omega \mapsto \mathbb{R}^{|\mathcal{P}|}_+$ be the path latency function. Fixing $\omega \in \Omega$, one can verify that $\ell = \Lambda^\top l(\Lambda\mu, \omega)$.

Standing Assumption 1 ensures that the latency functions realistically capture the relation between traffic flow and travel time.

Standing Assumption 1: The latency functions l_e are $\mathcal{F}$-measurable, differentiable w.r.t. q_e, for all $e \in \mathcal{E}$, and $\dfrac{\partial l_e(q_e, \omega)}{\partial q_e} > 0$ for all $q_e \geq 0$.

Each path flow profile $\mu \in \Delta$ induces a pushforward probability measure $\mathbb{P}_{\ell,\mu} : \mathcal{B}(\mathbb{R}^{|\mathcal{P}|}_+) \to [0, 1]$ associated with the positive random vector $\ell(\mu, \cdot) : \Omega \mapsto \mathbb{R}^{|\mathcal{P}|}_+$.

The Standing Assumption 2 quantitatively ensures that the latency for a given path flow is relatively stable in the sense that it has a subgaussian tail. In other words, its variance magnitude is controlled by the parameter σ. It also implies that the expected travel time of the paths is bounded by some upper estimate that is linearly related to σ by the oracle-given constant κ. Such a setting is practical in most realistic scenarios.

Standing Assumption 2:

- There exists $\sigma > 0$, such that for all $\mu \in \Delta$, the Euclidean norm $\|z\| = \|\ell(\mu) - \mathbb{E}[\ell(\mu)]\|$ is σ-subgaussian, i.e.,
$$\mathbb{E}[\exp(\frac{\|z\|^2}{\sigma^2})] \leq \exp(1).$$
- There exists a positive constant L, such that,
$$\|\,\mathbb{E}\,\ell(\mu)\| \leq L, \quad \forall \mu \in \Delta.$$

Further, there exists a constant $\kappa > 0$ such that $L \leq \kappa\sigma$.

Resilience under Adversarial Information Delay Wardrop Equilibrium (WE) has been a conventional solution concept in transportation literature that describes the conditions under which the individual users have the least ex-post regret. Since the latent variable is oftentimes unobservable, we consider a meta-version of the congestion game, $\mathcal{G}_c = (\mathcal{G}, \mathcal{W}, \mathcal{X}, \mathcal{P}, \mathbb{E}[\ell(\cdot)])$, with the utility functions replaced by

the expected latency function. This "meta" game gives rise to a solution concept corresponding to Definition 1.

Definition 1 (Mean Wardrop Equilibrium [1]): A path flow $\mu \in \Delta$ is said to be a *Mean Wardrop Equilibrium* (MWE) if $\forall w \in \mathcal{W}$, $\mu_p > 0$ indicates $\mathbb{E}[\ell_p] \leq \mathbb{E}[\ell_{p'}]$ for all $p, p' \in \mathcal{P}_w$. The set of all MWE is denoted by $\boldsymbol{\mu}^*$.

It is well known that at MWE, the Beckman Potential of $\mathcal{G}_c$ is minimized; by Fubini's theorem, it is equivalent to minimizing the *Mean Beckman Potential* (MBP), as in (1),

$$\min_{\mu \in \Delta} \Phi(\mu) := \mathbb{E}\left[\sum_{e \in \mathcal{E}} \int_0^{(\Lambda\mu)_e} l_e(z, \omega) dz\right]. \tag{1}$$

Given Standing Assumption 1, Φ is in general non-strictly convex, $\nabla_\mu \Phi(\mu) = \mathbb{E}_\omega[\Lambda^\top l(\Lambda\mu, \omega)] = \mathbb{E}[\ell(\mu)]$.

By convention, at each time t, within each OD population $w \in \mathcal{W}$, if the infinitesimal players randomize independently according to some mixed strategy $\pi_w^t \in \Delta(\mathcal{P}_w)$ identically, the individual-level and population-level decision makings are equivalent [9], in the sense that $\pi_w^t = \frac{1}{m_w}(\mu_p^t)_{p \in \mathcal{P}_w}$. Let $\mathcal{H}_t$ denote the history of information about μ^t and ℓ^t realizations, a learning algorithm $\mathcal{A}$ maps a history $\mathcal{H}_t$ to μ^{t+1}.

Definition 2 ([6]): For the congestion game $\mathcal{G}_c$, let $(\Delta^T, \mathcal{B}^T)$ be the product space, with $\mathcal{B}^T$ be the product Borel algebra of Δ^T. For any $\epsilon > 0$, define the target set as $\mathcal{C}_\epsilon := \{\mu \in \Delta | \Phi(\mu) - \Phi^* < \epsilon\}$. A probability measure $\mathbb{P}_T$ over $(\Delta^T, \mathcal{B}^T)$ is an (ϵ, δ)-Wardrop Non-Equilibrium solution (WANES) if $\mathbb{P}_T\{(\mu^t)_{t=1}^T \in \Delta^T | \bar{\mu}^T \in \mathcal{C}_\epsilon\} \geq 1 - \delta$, with $\bar{\mu}^T = \frac{1}{T}\sum_{t=1}^T \mu^t$. Furthermore, any learning algorithm $\mathcal{A}$ producing such $\mathbb{P}_T$ is said to be (ϵ, δ)-resilient.

Definition 2 provides additional freedom to analyze the transient behavior.

Attack Model We consider the online traffic navigation scenario, where a feedback-delaying attacker (typically network jamming attacker) can delay the crowdsourcing of traffic latency (typically by launching DoS attacks), withholding the traffic latency of each round for a finite period of time. Consequently, the ONP navigation center is not able to retrieve the exact travel time estimates from the planned paths at each round t. Instead, at every t, a historical "bundle" of latency feedback is revealed to the navigation center, while the real-time traffic latency is to be revealed in the future.

Mathematically, let the attacker's action be a vector $\mathbf{d} = (\min\{d_t, T - t + 1\})_{1 \leq t \leq T}$, where $d_t \in \mathbb{N}_+$ represents the delayed time that the latency information ℓ^t is delivered to the INS. The total budget delay is $D = \sum_{t=1}^T \min\{d_t, T - t + 1\}$, which is of order $\mathcal{O}(T^2)$, as the maximum D is $T(T+1)/2$. We assume that the attacker's maximum per-iterate attack budget is $d := \|\mathbf{d}\|_\infty \ll T$. The INS receiver's information structure [18] at time t includes the latency vector "bundle" indexed by $\mathcal{D}_t = \{k | k + \min\{d_k, T - k + 1\} - 1 = t\}$, i.e., they receive $\mathcal{L}^t := \{\ell^k | k \in \mathcal{D}_t\}$, without access to the "time stamps" of ℓ^k.

IV. Traffic Assignment with Delayed Feedback

The Delayed Mirror Descent Let $\bar{\ell}^t$ be the estimate of the sum of arrived latency "bundle", i.e., $\bar{\ell}^t = \sum_{\tau \in \mathcal{D}_t} \ell^\tau$. The Delayed Mirror Descent (DMD), as shown in Algorithm 1, replaces the latency vector in the ordinary Mirror Descent algorithm with $\bar{\ell}^t$. We use the Bregman divergence that measures the dissimilarity between two iterates.

Definition 3: Given a mirror map $\Psi : \nabla \to \bar{\mathbb{R}}$ and two points $\mu_1, \mu_2 \in \Delta$, the Bregman divergence is $\mathrm{D}_\Psi|_{\mu_2}^{\mu_1} = \Psi(\mu_1) - \Psi(\mu_2) - \langle \nabla\Psi(\mu_1), \mu_1 - \mu_2 \rangle$.

Suppose that the mirror map Ψ is σ_Ψ-strongly convex, we have $\mathrm{D}_\Psi|_{\mu_2}^{\mu_1} \geq \frac{\sigma_\Psi}{2}\|\mu_1 - \mu_2\|^2$. We refer readers to [19] for a more in-depth discussion regarding this notion.

Algorithm 1: Delayed Mirror Descent (DMD)

Input : initialize $\mu^1 \in \Delta$, learning rate η.
for $t \in 1, \ldots, T$ **do**
  for $w \in \mathcal{W}$, $x \in \mathcal{X}_w$, **do**
    INS assigns mixed strategy $\pi^t(\cdot, x) \leftarrow \frac{1}{m_w}(\mu_p^t)_{p \in \mathcal{P}_w}$ to player x;
    player x samples path $A(x) \sim \pi^t(\cdot, x)$;
  end
  Players $\mathcal{X}$ suffer latency $\ell^t \sim \mathbb{P}_{\ell, \mu^t}(\cdot)$;
  INS reveals latency vector $(\ell^k)_{k \in \mathcal{D}_t}$ to $\mathcal{X}$;
  INS updates:

$$\mu^{t+1} \leftarrow \arg\min_{\mu \in \Delta} \langle \mu, \eta \bar{\ell}^t \rangle + D_\Psi(\mu, \mu^t) \tag{2}$$

end

Telescoping Setup Central to the analysis, Lemma 1 quantifies the summation of MBP functional gains along the learning trajectory, setting up conditions for telescoping the sequence.

Let $\mathcal{F}_t = \sigma((z_\tau)_{\tau \in \mathcal{D}_1}, \ldots, (z_\tau)_{\tau \in \mathcal{D}_{t-1}})$ be the filtration of the delayed data generating process for $t = 1, \ldots, T$, note that μ^t is $\mathcal{F}_t$-measurable. We fix an equilibrium flow $\mu^* \in \boldsymbol{\mu}^*$ for simplicity of analysis, and define the process $\xi_t = \eta \langle z^t, \mu^* - \mu^t \rangle$ for $t = 1, \ldots, T$. We also let $\|z_m^t\| := \max_{r \in \cup_{s=\tau_t}^t \mathcal{D}_s} \|z^r\|$ be the maximum of the empirical process generated through $\mathcal{D}_{\tau_t}, \ldots, \mathcal{D}_t$, with $\tau_t = \min \mathcal{D}_t$.

Lemma 1: Under the mirror descent dynamics with latency feedback, the following holds for $t = 1, \ldots, T$,

$$\begin{aligned} &\sum_{\tau \in \mathcal{D}_t} \eta\left(\Phi^\tau - \Phi^*\right) - \frac{2\eta^2 d}{\sigma_\Psi} L^2 + \mathrm{D}_\Psi|_{\mu^{t+1}}^{\mu^*} - \mathrm{D}_\Psi|_{\mu^t}^{\mu^*} \\ &\leq \sum_{\tau \in \mathcal{D}_t} \xi_\tau + \frac{2\eta^2 d}{\sigma_\Psi} \|z_m^t\|^2. \end{aligned} \tag{3}$$

As illustrated in Fig. 2, due to bounded attack capacity d, the cardinality of the delayed latency "bundle" $|\mathcal{D}_t| \leq d$ for every $t \leq 1, \ldots, T$. One can further see that the cardinality of $\cup_{s=\tau_t}^t \mathcal{D}_t$ should be bounded by $2d$.

Proof: Recall that the mirror step can be equivalently written as

$$\mu^{t+1} =: \nabla\Psi^*(\underbrace{\nabla\Psi(\mu^t) - \sum_{\tau \in \mathcal{D}_t} \eta \ell^\tau}_{\text{dual step}}),$$

where $\Psi^* : \mathbb{R}^{|\mathcal{P}|} \to \mathbb{R}$ is the Fenchel conjugate of the mirror map Ψ. Under proper conditions, for $\mu \in \partial\Psi^*(\nu)$, $\nu \in$

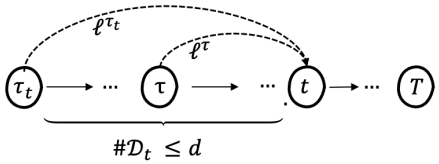

Fig. 2. The per-iterate latency "bundle" is bounded, $t - \tau_t + 1 \le d$.

$\partial\Psi(\mu)$, it holds that $\Psi(\mu) + \Psi^*(\nu) = \langle \mu, \nu \rangle$. Therefore, given a primal flow point μ^t, we define $\nu^{t+1} := \nabla\Psi(\mu^t) - \sum_{\tau\in\mathcal{D}_t} \eta\ell^\tau = \nabla\Psi(\mu^{t+1})$ as the dual latency point. One can verify that $\mu^{t+1} = \nabla\Psi^*(\nu^{t+1})$ and the dual step can be written as $\nu^{t+1} = \nu^t - \sum_{\tau\in\mathcal{D}_t} \eta\ell^\tau$.

Due to the convexity of Φ,

$$\begin{aligned}
&\sum_{\tau\in\mathcal{D}_t} \eta(\Phi(\mu^\tau) - \Phi^*) \le \sum_{\tau\in\mathcal{D}_t} \eta\langle \mathbb{E}\,\ell^\tau, \mu^\tau - \mu^*\rangle \\
&= \sum_{\tau\in\mathcal{D}_t} \eta\langle z^\tau, \mu^* - \mu^\tau\rangle + \eta\langle \ell^\tau, \mu^\tau - \mu^*\rangle \\
&= \sum_{\tau\in\mathcal{D}_t} \xi_\tau + \sum_{\tau\in\mathcal{D}_t} \eta\langle \ell^\tau, \mu^{\tau,-} - \mu^*\rangle + \eta\langle \ell^\tau, \mu^\tau - \mu^{\tau,-}\rangle,
\end{aligned}$$

where we let $\mu^{\tau,-}$ be intermediate primal flow before applying latency ℓ^τ, i.e., let $\mathcal{D}_{t,\tau} := \{r \in \mathcal{D}_t, r < \tau\}$, $\mu^{\tau,-} := \nabla\Psi^*(\nu^{\tau,-}) = \nabla\Psi^*(\nabla\Psi(\mu^t) - \sum_{r\in\mathcal{D}_{t,\tau}} \eta\ell^r)$. Similarly we define the immediate primal flow point after applying latency ℓ^τ as $\mu^{\tau,+} := \nabla\Psi^*(\nu^{\tau,+}) := \nabla\Psi^*(\nabla\Psi(\mu^{\tau,-}) - \eta\ell^\tau)$. One can then verify that $\mu^{\tau_t,-} = \mu^t$ and $\mu^{\max\mathcal{D}_t,+} = \mu^{t+1}$.

Consider the following decomposition for arbitrary $\mu^* \in \Delta$, $\langle \mu^{\tau,-} - \mu^*, \ell^\tau\rangle = \langle \mu^{\tau,-} - \mu^{\tau,+}, \ell^\tau\rangle + \langle \mu^{\tau,+} - \mu^*, \ell^\tau\rangle$. By the first-order optimality condition of (2),

$$\langle \eta\ell^\tau + \nabla\Psi(\mu^{\tau,+}) - \nabla\Psi(\mu^{\tau,-}), \mu^{\tau,+} - \mu^*\rangle \le 0.$$

Apply Pythagorean identity of Bregman divergence: $\mathrm{D}_\Psi|_{\mu^{\tau,-}}^{\mu^{\tau,+}} + \mathrm{D}_\Psi|_{\mu^{\tau,+}}^{\mu^*} - \mathrm{D}_\Psi|_{\mu^{\tau,-}}^{\mu^*} = \langle \nabla\Psi(\mu^{\tau,+}) - \nabla\Psi(\mu^{\tau,-}), \mu^{\tau,+} - \mu^*\rangle$, and we arrive at

$$\eta\langle \ell^\tau, \mu^{\tau,+} - \mu^*\rangle + \mathrm{D}_\Psi|_{\mu^{\tau,-}}^{\mu^{\tau,+}} + \mathrm{D}_\Psi|_{\mu^{\tau,+}}^{\mu^*} - \mathrm{D}_\Psi|_{\mu^{\tau,-}}^{\mu^*} \le 0.$$

Summing over $\tau \in \mathcal{D}_t$, since $\mathrm{D}_\Psi|_{\mu^{\tau,-}}^{\mu^{\tau,+}} \ge \frac{\sigma_\Psi}{2}\|\mu^{\tau,+} - \mu^{\tau,-}\|^2$, by Cauchy-Schwarz inequality and $-b^2 + 2ab \le a^2$,

$$\begin{aligned}
&\sum_{\tau\in\mathcal{D}_t} \eta\langle \ell^\tau, \mu^{\tau,-} - \mu^*\rangle + \frac{\sigma_\Psi}{2}\|\mu^{\tau,+} - \mu^{\tau,-}\|^2 \\
&\le \sum_{\tau\in\mathcal{D}_t} \mathrm{D}_\Psi|_{\mu^{\tau,-}}^{\mu^*} - \mathrm{D}_\Psi|_{\mu^{\tau,+}}^{\mu^*} + \langle \mu^{\tau,-} - \mu^{\tau,+}, \eta\ell^\tau\rangle \\
&\le \sum_{\tau\in\mathcal{D}_t} \mathrm{D}_\Psi|_{\mu^{\tau,-}}^{\mu^*} - \mathrm{D}_\Psi|_{\mu^{\tau,+}}^{\mu^*}. + \eta\|\ell^\tau\|\|\mu^{\tau,-} - \mu^{\tau,+}\|,
\end{aligned}$$

which then gives

$$\begin{aligned}
&\sum_{\tau\in\mathcal{D}_t} \eta\langle \ell^\tau, \mu^{\tau,-} - \mu^*\rangle \\
&\le \sum_{\tau\in\mathcal{D}_t} \mathrm{D}_\Psi|_{\mu^{\tau,-}}^{\mu^*} - \mathrm{D}_\Psi|_{\mu^{\tau,+}}^{\mu^*} + \frac{\eta^2}{2\sigma_\Psi}\|\ell^\tau\|^2 \\
&\le \sum_{\tau\in\mathcal{D}_t} \mathrm{D}_\Psi|_{\mu^{\tau,-}}^{\mu^*} - \mathrm{D}_\Psi|_{\mu^{\tau,+}}^{\mu^*} + \frac{\eta^2}{2\sigma_\Psi}(L^2 + \|z^\tau\|^2).
\end{aligned}$$

Now we analyze $\eta\langle \ell^\tau, \mu^\tau - \mu^{\tau,-}\rangle$: applying Cauchy-Schwarz inequality and triangular inequality several times,

$$\begin{aligned}
&\sum_{\tau\in\mathcal{D}_t} \eta\langle \ell^\tau, \mu^\tau - \mu^{\tau,-}\rangle \le \sum_{\tau\in\mathcal{D}_t} \eta\|\ell^\tau\|\|\mu^\tau - \mu^{\tau,-}\| \\
&\le \sum_{\tau\in\mathcal{D}_t} \eta\|\ell^\tau\| \left(\sum_{s=\tau}^{t-1} \|\mu^s - \mu^{s+1}\| + \|\mu^t - \mu^{\tau,-}\|\right) \\
&= \sum_{\tau\in\mathcal{D}_t} \eta\|\ell^\tau\| \Bigg(\sum_{s=r}^{t-1} \left\|\nabla\Psi^*(\nu^s) - \nabla\Psi^*(\nu^{s+1})\right\| \\
&\qquad + \left\|\nabla\Psi^*(\nu^t) - \nabla\Psi^*(\nu^{\tau,-})\right\|\Bigg).
\end{aligned}$$

Since Ψ is σ_Ψ-strongly convex, its Fenchel conjugate Ψ^* is $\frac{1}{\sigma_\Psi}$-smooth, the R.H.S. becomes

$$\begin{aligned}
&\le \sum_{\tau\in\mathcal{D}_t} \eta\|\ell^\tau\| \left(\frac{1}{\sigma_\Psi}\Big(\sum_{s=\tau}^{t-1} \left\|\nu^s - \nu^{s+1}\right\| + \left\|\nu^t - \nu^{\tau,-}\right\|\Big)\right) \\
&\le \sum_{\tau\in\mathcal{D}_t} \frac{\eta^2}{\sigma_\Psi}\|\ell^\tau\| \left(\sum_{s=\tau}^{t-1} \sum_{r\in\mathcal{D}_s} \|\ell^r\| + \sum_{p\in\mathcal{D}_{t,\tau}} \|\ell^p\|\right).
\end{aligned}$$

To associate the upper estimate with the cardinality of $|\mathcal{D}_t|$, by triangular inequality, $\|\ell^t\| = \|\mathbb{E}\,\ell^t + z^t\| \le L + \|z^t\|$ for all t. Breaking the brackets, we get the R.H.S. becomes

$$\begin{aligned}
&\le \underbrace{\frac{\eta^2 L^2}{\sigma_\Psi} \sum_{\tau\in\mathcal{D}_t} Q_\tau}_{\text{I}} + \underbrace{\frac{\eta^2 L}{\sigma_\Psi} \sum_{\tau\in\mathcal{D}_t} \|z^\tau\| Q_\tau}_{\text{II}} \\
&\quad + \underbrace{\frac{\eta^2}{\sigma_\Psi} \sum_{\tau\in\mathcal{D}_t} \|z^\tau\| \left(\sum_{s=\tau}^{t-1} \sum_{r\in\mathcal{D}_s} \|z^r\| + \sum_{p\in\mathcal{D}_{t,\tau}} \|z^p\|\right)}_{\text{III}},
\end{aligned}$$

where $Q_\tau = |\mathcal{D}_{t,\tau}| + \sum_{s=\tau}^{t-1} |\mathcal{D}_s|$, which essentially counts for the number of all latency vectors other than ℓ^τ, that have been delivered between round τ_t and t.

We fix τ and t and look into round $s \in \{\tau, \dots, t\}$, when $s = t$, consider $q \in \mathcal{D}_{t,\tau}$; when $s < t$, consider $q \in \mathcal{D}_s$. There are two cases, $q < \tau$ and $q \ge \tau$. Consider both cases quantitatively, $Q_\tau \le Q_{\tau,1} + Q_{\tau,2}$:

$$\begin{aligned}
Q_\tau \le &\left(\sum_{q\in\mathcal{D}_{t,\tau}} \mathbb{1}_{\{q\ge\tau\}} + \sum_{s=\tau}^{t-1} \sum_{q\in\mathcal{D}_s} \mathbb{1}_{\{q\ge\tau\}}\right) \\
&+ \left(\sum_{q\in\mathcal{D}_{t,\tau}} \mathbb{1}_{\{q<\tau\}} + \sum_{s=\tau}^{t-1} \sum_{q\in\mathcal{D}_s} \mathbb{1}_{\{q<\tau\}}\right).
\end{aligned}$$

When $q \ge \tau$, by a pigeonhole argument, there are at most d_τ instances, as fixing a q, $q + d_q - 1$ is at most in only one of $\{q, \dots, t-1\}$. Analytically, we have $Q_{\tau,1} = \sum_{q\in\mathcal{D}_{t,\tau}} \mathbb{1}_{\{q\ge\tau\}} + \sum_{s=\tau}^{t-1} \sum_{q\in\mathcal{D}_s} \mathbb{1}_{\{q\ge\tau\}}$, which is $\sum_{s=\tau}^{t-1} \sum_{q=\tau}^{s} \mathbb{1}_{\{q+d_q-1=s\}}$. We rearrange the sum and rewrite it as $\sum_{q=\tau}^{t-1} \sum_{s=q}^{t-1} \mathbb{1}_{\{q+d_q-1=s\}}$ which by observation is bounded by $d_\tau - 1 \le d_\tau$.

When $q < \tau$, fixing q, we have that $Q_{\tau,2}$ can be written as $\sum_{q\in\mathcal{D}_{t,\tau}} \mathbb{1}_{\{q<\tau\}} + \sum_{s=\tau}^{t-1}\sum_{q\in\mathcal{D}_s} \mathbb{1}_{\{q<\tau\}}$ which is essentially $\sum_{s=\tau}^{t}|\mathcal{D}_{s,\tau}|$ and can be further written as $\sum_{q=1}^{\tau-1}\mathbb{1}_{\{q+d_q-1=t\}} + \sum_{s=\tau}^{t-1}\sum_{q=1}^{\tau-1}\mathbb{1}_{\{q+d_q-1=s\}}$. This quantity essentially counts the q's that get delayed into the range $\{\tau, t\}$, i.e., $\sum_{q=1}^{\tau-1}\mathbb{1}_{\{q+d_q-1\in\{\tau,\dots,t\}\}} \le d$.

Hence, we arrive at:

$$\mathrm{I} \le \frac{\eta_1^2 L^2}{\sigma_\Psi}\sum_{\tau\in\mathcal{D}_t} d_\tau + d \le \frac{2\eta_1^2 L^2}{\sigma_\Psi}\sum_{\tau\in\mathcal{D}_t} d,$$

$$\mathrm{II} \le \frac{2\eta^2 L}{\sigma_\Psi}\max_{\tau\in\mathcal{D}_t}\|z^\tau\|\sum_{\tau\in\mathcal{D}_t} d,$$

$$\mathrm{III} \le \frac{2\eta^2}{\sigma_\Psi}\max_{\tau\in\cup_{s=\tau_{\min}}^{t}\mathcal{D}_s}\|z^\tau\|^2\sum_{\tau\in\mathcal{D}_t} d.$$

Let $\|z_m^t\| := \max_{\tau\in\cup_{s=\tau_t}^{t}\mathcal{D}_s}\|z^\tau\|$ be the maximum of the empirical process generated by $\mathcal{D}_{\tau_t},\dots,\mathcal{D}_t$, clearly, $\max_{s\in\mathcal{D}_t}\|z^s\| \le \|z_m^t\|$. Combining I, II, and III, we obtain

$$\begin{aligned}
&\sum_{\tau\in\mathcal{D}_t}\eta\left(\Phi^\tau - \Phi^*\right) + \mathrm{D}_\Psi|_{\mu^{t+1}}^{\mu^*} - \mathrm{D}_\Psi|_{\mu^t}^{\mu^*} \le \sum_{\tau\in\mathcal{D}_t}\xi_\tau \\
&+ \frac{\eta^2}{2\sigma_\Psi}(L^2 + \|z^\tau\|^2 + 2dL^2 + \max_{\tau\in\mathcal{D}_t} 2dL\|z^\tau\| + 2d\|z_m^t\|^2) \\
&\le \sum_{\tau\in\mathcal{D}_t}\xi_\tau + \frac{\eta^2}{2\sigma_\Psi}((1+2d)(L^2+\|z_m^t\|^2) + 2dL\|z_m^t\|) \\
&\le \sum_{\tau\in\mathcal{D}_t}\xi_\tau + \frac{2\eta^2 d}{\sigma_\Psi}(L^2 + \|z_m^t\|^2).
\end{aligned}$$

Rearrange the terms and we arrive at (3). ∎

V. Resilience Analysis

Bounding the Moment Generating Function To verify that iterates (2) lead to a Wardrop Non-equilibrium solution, we need to derive a high probability bound for the functional gaps $\Phi(\mu^t) - \Phi^*$ of the flow trajectory $(\mu^t)_{t=1}^{T+1}$ along the learning process. To this end, we define a set of weights $\{w_t\}_{t=1}^{T+1}$ that serves as the set of variable coefficients inside the moment generating functions of the functional gaps, which allows for the flexibility of compensating the learning rate η. The target of our analysis is the two auxiliary quantities as we have defined in (4).

$$\begin{aligned}
Z_t =& w_{t+1}\sum_{\tau\in\mathcal{D}_t}(\eta(\Phi^\tau - \Phi^*) - \frac{2\eta^2 dL^2}{\sigma_\Psi}) \\
&+ w_{T+1}(\mathrm{D}_\Psi|_{\mu^{t+1}}^{\mu^*} - \mathrm{D}_\Psi|_{\mu^t}^{\mu^*}) \quad \text{for } t = 1,\dots,T. \quad (4)\\
S_t =& \sum_{i=t}^{T} Z_i, \qquad\qquad \text{for } t = 1,\dots,T+1.
\end{aligned}$$

We analyze the moment generating function of S_t conditioned on $\mathcal{F}_{\tau_t}$. By convention, we let $\tau_t = t$ if $\mathcal{D}_t = \emptyset$, thus $\tau_1 = 1$. The core result, as shown in Theorem 1 is a Chernoff-type of bound that gives rise to the main concentration argument of our interest.

Theorem 1: Suppose that $\{w_t\}$ satisfies that, $w_{t+1} + w_{t+1}^2\frac{648d^3\eta^2\sigma^2}{\sigma_\Psi} \le w_t$, and $w_{t+1}\eta^2 d^2 \le \frac{\sigma_\Psi}{432d\sigma^2}$. Then, it holds that for every $1 \le t \le T$ with probability 1,

$$\mathbb{E}[\exp(S_t)|\mathcal{F}_{\tau_t}] \le \exp((w_t - w_{T+1})\mathrm{D}_\Psi|_{\mu^t}^{\mu^*} + C\sum_{i=t}^{T} w_{i+1}\eta^2), \quad (5)$$

where the constant $C := 324\sigma^2 d^3\sigma_\Psi^{-1}(8+\kappa^2)$.

The proof, which is deferred to the Appendix, relies on an induction approach. The intuition behind this approach is that propagation of subgaussian behavior scales with d^2, while the subgaussianity of the maxima of empirical process scales with d. Thus, the stochastic error of the per-iterate upper estimate scales with d^3.

DMD attains Non-equilibrium With all the preparations above, we now turn to the non-equilibrium analysis. Corollary 1 comes from the fact that the stochastic fluctuation will be absorbed by the sequence $\{w_t\}$ under certain conditions.

Corollary 1: Let $w_{T+1} = \frac{\sigma_\Psi}{1296d^3\sigma^2\eta^2(T+1)}$ and $w_t = w_{t+1} + \frac{648w_{t+1}^2 d^3\eta^2}{\sigma_\Psi}$ for all $1 \le t \le T$. The sequence $\{w_t\}$ satisfies the condition required by Theorem 1, and for $\delta \in (0,1)$, the following events hold with probability at least $1-\delta$:

$$\frac{1}{T}\sum_{t=1}^{T}(\Phi(\mu^t) - \Phi^*) \le \mathcal{O}\left(\frac{\mathrm{D}_\Psi|_{\mu^1}^{\mu^*}}{\eta T} + \frac{\sigma^2}{\sigma_\Psi}d^3(1+\ln(\frac{1}{\delta}))\eta\right), \quad (6)$$

and

$$\mathrm{D}_\Psi|_{\mu^{T+1}}^{\mu^*} \le \mathcal{O}\left(\mathrm{D}_\Psi|_{\mu^1}^{\mu^*} + \frac{\sigma^2}{\sigma_\Psi}d^3(1+\ln(\frac{1}{\delta}))\eta^2\right). \quad (7)$$

Setting $\eta = \sqrt{\frac{\mathrm{D}_\Psi|_{\mu^1}^{\mu^*}}{\frac{\sigma^2}{\sigma_\Psi}d^3(1+\ln(\frac{1}{\delta}))T}}$, we have

$$\frac{1}{T}\sum_{t=1}^{T}(\Phi(\mu^t) - \Phi^*) \le \tilde{\mathcal{O}}\left(d^{\frac{3}{2}}\sqrt{\frac{\frac{\sigma^2}{\sigma_\Psi}\mathrm{D}_\Psi|_{\mu^1}^{\mu^*}(1+\ln(\frac{1}{\delta}))}{T}}\right). \quad (8)$$

Proof: For simplicity, let $c_d := 108d^3$. We first show that the sequence $\{w_t\}$ satisfies the conditions required by Theorem 1:

$$w_{t+1} + w_{t+1}^2\frac{6c_d\eta^2\sigma^2}{\sigma_\Psi} \le w_t, \frac{w_{t+1}\eta^2}{\sigma_\Psi} \le \frac{1}{4c_d\sigma^2}.$$

Let $A = 6c_d\sigma^2\sigma_\Psi^{-1}\eta^2(T+1)$. Set $w_{T+1} = \frac{1}{2A}$. For $1 \le t \le T$, set w_t such that $w_{t+1} + w_{t+1}^2\frac{6c_d\eta^2\sigma^2}{\sigma_\Psi} = w_t$, the first condition is automatically satisfied. To verify the second condition, notice that in this setup, $w_t\eta^2 \le \frac{\eta^2}{A} = \frac{\sigma_\Psi}{6c_d\sigma^2(T+1)} \le \frac{\sigma_\Psi}{4c_d\sigma^2}$. Now let $K = (w_1 - w_{T+1})\mathrm{D}_\Psi|_{\mu^1}^{\mu^*} + C\sum_{t=1}^{T} w_{t+1}\eta^2 + \ln(\frac{1}{\delta})$. By Markov inequality and Theorem 1,

$$\begin{aligned}
&\mathbb{P}[S_1 \ge K] \le \exp(-K)\,\mathbb{E}[\exp(S_1)] \\
&\le \exp(-K)\exp((w_1 - w_{T+1})\mathrm{D}_\Psi|_{\mu^1}^{\mu^*} + C\sum_{t=2}^{T+1} w_t\eta^2) \\
&= \delta.
\end{aligned}$$

Since $S_1 = \sum_{t=1}^{T} w_{t+1}\eta \sum_{\tau\in\mathcal{D}_t}(\Phi(\mu^\tau) - \Phi^*) - \frac{2dL^2}{\sigma_\Psi}\sum_{t=1}^{T} w_{t+1}\eta^2 + w_{T+1}(\mathrm{D}_\Psi|_{\mu^T}^{\mu^*} - \mathrm{D}_\Psi|_{\mu^1}^{\mu^*})$, we have that with probability at least $1-\delta$,

$$\sum_{t=1}^{T} w_{t+1}\eta(\sum_{\tau\in\mathcal{D}_t}\Phi(\mu^\tau) - \Phi^*) + w_{T+1}\mathrm{D}_\Psi|_{\mu^{T+1}}^{\mu^*}$$
$$\leq w_1\mathrm{D}_\Psi|_{\mu^1}^{\mu^*} + (\frac{2dL^2}{\sigma_\Psi} + C)\sum_{t=1}^{T} w_{t+1}\eta^2 + \ln(\frac{1}{\delta}).$$

Since $w_{T+1} = \frac{1}{2A}$ and $\frac{1}{2A} \leq w_t \leq \frac{1}{A}$ for $1 \leq t \leq T+1$, we plug them into above and obtain

$$\eta\sum_{t=1}^{T}\sum_{\tau\in\mathcal{D}_t}(\Phi(\mu^\tau) - \Phi^*) + \mathrm{D}_\Psi|_{\mu^{T+1}}^{\mu^*}$$
$$\leq 2\mathrm{D}_\Psi|_{\mu^1}^{\mu^*} + 2(\frac{2dL^2}{\sigma_\Psi} + C)\eta^2 T + 2A\ln(\frac{1}{\delta})$$
$$\leq 2\mathrm{D}_\Psi|_{\mu^1}^{\mu^*} + 2(\frac{\sigma^2}{\sigma_\Psi}(B + A\ln(\frac{1}{\delta})))\eta^2 T,$$

where $B = \mathcal{O}(d^3)$. Dividing both sides by η yields

$$\frac{1}{T}\sum_{t=1}^{T}(\Phi(\mu^t) - \Phi^*) \leq \frac{\mathrm{D}_\Psi|_{\mu^1}^{\mu^*}}{\eta T} + 2(\frac{\sigma^2}{\sigma_\Psi}B + A\ln(\frac{1}{\delta}))\eta$$
$$\leq \mathcal{O}\left(\frac{\mathrm{D}_\Psi|_{\mu^1}^{\mu^*}}{\eta T} + \frac{\sigma^2}{\sigma_\Psi}d^3(1+\ln(\frac{1}{\delta}))\eta\right).$$

and $\mathrm{D}_\Psi|_{\mu^{T+1}}^{\mu^*} \leq 2\mathrm{D}_\Psi|_{\mu^1}^{\mu^*} + 2(\frac{\sigma^2}{\sigma_\Psi}(B + A\ln(\frac{1}{\delta})))\eta^2 T$. Setting $\eta = \sqrt{\frac{\mathrm{D}_\Psi|_{\mu^1}^{\mu^*}}{\frac{\sigma^2}{\sigma_\Psi}d^3(1+\ln(\frac{1}{\delta}))T}}$ gives the results.

■

Corollary 1 immediately implies resilience in the non-equilibrium sense, which is summarized in Proposition 1. A case study for simulated delay attack over the Sioux Fall network is omitted due to page limit[1].

Proposition 1: For $\delta \in (0,1)$, the DMD Algorithm 1 with $\eta = \mathcal{O}(\sqrt{d^{-3}T^{-1}})$ is (ϵ,δ)-resilient, which gives a (ϵ,δ)-WANES, with $\epsilon = \tilde{\mathcal{O}}(\sqrt{\frac{d^3}{T}})$.

Proof: For $\mu^1,\ldots,\mu^T$ produced by the DMD algorithm, by the convexity of Φ^*, $\Phi(\bar{\mu}^T) - \Phi^* \leq \frac{1}{T}\sum_{t=1}^{T}\Phi(\mu^t) - \Phi^*$ which satisfies Corollary 1 with $1-\delta$, hence the statement follows.

■

VI. Conclusion

In this paper, we have investigated the resilience of DMD-based INS under adversarial delay attacks. We made some mild assumptions to handle the challenges that arose in finite-time analysis, obtaining a high probability bound for the performance loss. With the aid of the non-equilibrium notion, we have demonstrated the self-restoring capability of INS to recover from information-delaying attacks.

Future research would focus on developing scalable and distributed strategies to handle adversarial delays to improve the defense mechanism in the face of cyber-physical threats. We would also refine the analysis of concentration arguments, improving the order of existing results to match the lower bound in the deterministic setting.

[1]Interested readers can refer to `https://github.com/UnionPan/Wardrop_attack.git`.

References

[1] John Glen Wardrop. Road paper. some theoretical aspects of road traffic research. *Proceedings of the institution of civil engineers*, 1(3):325–362, 1952.

[2] Yunian Pan and Quanyan Zhu. On poisoned wardrop equilibrium in congestion games. In *International Conference on Decision and Game Theory for Security*, pages 191–211. Springer, 2022.

[3] Zhiheng Xu and Quanyan Zhu. A game-theoretic approach to secure control of communication-based train control systems under jamming attacks. In *Proceedings of the 1st International Workshop on Safe Control of Connected and Autonomous Vehicles*, pages 27–34, 2017.

[4] Quanyan Zhu and Zhiheng Xu. *Cross-Layer Design for Secure and Resilient Cyber-Physical Systems*. Springer, 2020.

[5] Hideaki Ishii and Quanyan Zhu. *Security and Resilience of Control Systems*. Springer, 2022.

[6] Yunian Pan, Tao Li, and Quanyan Zhu. On the resilience of traffic networks under non-equilibrium learning. In *2023 American Control Conference (ACC)*, pages 3484–3489. IEEE, 2023.

[7] Yunian Pan and Quanyan Zhu. Efficient episodic learning of nonstationary and unknown zero-sum games using expert game ensembles. In *2021 60th IEEE Conference on Decision and Control (CDC)*, pages 1669–1676. IEEE, 2021.

[8] Tao Li, Guanze Peng, Quanyan Zhu, and Tamer Başar. The Confluence of Networks, Games, and Learning a Game-Theoretic Framework for Multiagent Decision Making Over Networks. *IEEE Control Systems*, 42(4):35–67, 2022.

[9] Walid Krichene, Benjamin Drighès, and Alexandre Bayen. On the convergence of no-regret learning in selfish routing. In *International Conference on Machine Learning*, pages 163–171. PMLR, 2014.

[10] Walid Krichene, Syrine Krichene, and Alexandre Bayen. Convergence of mirror descent dynamics in the routing game. In *2015 European Control Conference (ECC)*, pages 569–574. IEEE, 2015.

[11] Dong Quan Vu, Kimon Antonakopoulos, and Panayotis Mertikopoulos. Fast Routing under Uncertainty: Adaptive Learning in Congestion Games with Exponential Weights. *Advances in Neural Information Processing Systems*, 18(NeurIPS):14708–14720, 2021.

[12] Jian Lou and Yevgeniy Vorobeychik. Decentralization and security in dynamic traffic light control. In *Proceedings of the Symposium and Bootcamp on the Science of Security*, pages 90–92, 2016.

[13] Teodora Mecheva and Nikolay Kakanakov. Cybersecurity in intelligent transportation systems. *Computers*, 9(4):83, 2020.

[14] Pooria Joulani, Andras Gyorgy, and Csaba Szepesvári. Online learning under delayed feedback. In *International Conference on Machine Learning*, pages 1453–1461. PMLR, 2013.

[15] Ohad Shamir and Liran Szlak. Online learning with local permutations and delayed feedback. In *International Conference on Machine Learning*, pages 3086–3094. PMLR, 2017.

[16] Alekh Agarwal and John C Duchi. Distributed delayed stochastic optimization. In J. Shawe-Taylor, R. Zemel, P. Bartlett, F. Pereira, and K.Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 24. Curran Associates, Inc., 2011.

[17] Kent Quanrud and Daniel Khashabi. Online learning with adversarial delays. In C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 28. Curran Associates, Inc., 2015.

[18] Tao Li, Yuhan Zhao, and Quanyan Zhu. The role of information structures in game-theoretic multi-agent learning. *Annual Reviews in Control*, 53:296–314, 2022.

[19] Amir Beck and Marc Teboulle. Mirror descent and nonlinear projected subgradient methods for convex optimization. *Operations Research Letters*, 31(3):167–175, 2003.