Avoiding Unintended Consequences:
How Incentives Aid Information Provisioning in Bayesian Congestion Games

Bryce L. Ferguson, Philip N. Brown, and Jason R. Marden

Abstract— **When users lack specific knowledge of various system parameters, their uncertainty may lead them to make undesirable deviations in their decision making. To alleviate this, an informed system operator may elect to *signal* information to uninformed users with the hope of persuading them to take more preferable actions. In this work, we study public and truthful signalling mechanisms in the context of Bayesian congestion games on parallel networks. We provide bounds on the possible benefit a signalling policy can provide with and without the concurrent use of monetary incentives. We find that though revealing information can reduce system cost in some settings, it can also be detrimental and cause worse performance than not signalling at all. However, by utilizing both signalling and incentive mechanisms, the system operator can guarantee that revealing information does not worsen performance while offering similar opportunities for improvement. These findings emerge from the closed form bounds we derive on the benefit a signalling policy can provide. We provide a numerical example which illustrates the phenomenon that revealing more information can degrade performance when incentives are not used and improves performance when incentives are used.**

I. INTRODUCTION

In many of today's large-scale cyber-physcial or socio-technical systems, operating conditions are greatly affected by the decisions of the systems users, e.g., power grid demand [1], transit traffic [2], and supply-chain commerce or marketing [3], [4]. Though users often make decisions in their own self interest, the system behavior that emerges need not be optimal [5]. This inefficiency can be further exacerbated by users' uncertainty over the state of the system [6], e.g., uncertainty on current electricity prices, traffic rates, or product quality. By increasing their sensing and communication technologies, the operators of these systems gain the opportunity to learn the unknown system parameters and may choose to fully or partially share this information with their uninformed users.

The emergence of these new sensing and communication technologies opens the door to new methods for improving system performance. One such method is that of information signalling by a well informed central authority [7]–[10]. By partially revealing their information about system parameters to uninformed users, the signaller allows the system users the opportunity to form new beliefs about their environment. If the signaller reveals this information strategically, they may alter user behavior in such a way that the overall system performance is improved; for example, Google/Apple maps can reveal the travel times of certain routes to guide driver decision making in a way that can alter aggregate driving patterns and improve performance for the user population [11]. One may initially think that all information should be shared with the users; however, this need not be optimal and could further degrade system performance. The main focus of this work is understanding when and how a system operator should inform users about the system state and how much they can improve performance by doing so.

We study the principles of information signalling in the context of *Bayesian congestion games*, where a group of users must route themselves through a congestable network while the exact congestion characteristics of each path are unknown. The users possess a prior belief over system parameters that determine each paths congestion rate, for example, the belief there was an accident on a road or the chance weather has affected driving conditions. The system operator can determine the values of these parameters and decide how to strategically reveal this information to their system users. This model has been used to study the how signalling policies should be designed [12]–[14], what behavior is likely to emerge [15], and the associated performance of specific signalling structures [11], [16]. The results are typically limited to computational methods for finding signalling policies or a binary classification of whether information can or cannot help and frequently study simple, two path settings. Additionally, for ease of analysis, much of the work in this area often assumes the signals are private (sent to individual users) and can be deceitful. To better understand the true efficacy of revealing information, we consider public and fully truthful signals. The main results of this work provide closed form bounds on the capabilities information signalling can provide in a parallel Bayesian congestion game with arbitrary polynomial latency functions.

Signalling mechanisms are becoming a topic of increasing research in their ability to influence user behavior, however, this is not the only influencing mechanism at a system operator's disposal. Inventive mechanisms, where users are assessed monetary penalties or rewards based on their actions, have long been studied as an effective means of coordinating system behavior [17]–[21]. The interplay between incentives and signalling is an emerging area of study, and has up until now been limited to studying mechanisms where users must pay to acquire information [22]–[24]. To the best of our knowledge, no current work has studied how monetary incentives and information signalling can be used

B. L. Ferguson (corresponding author) and J. R. Marden are with the Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA, {blferguson, jrmarden}@ece.ucsb.edu.

P.N. Brown is with the Department of Computer Science, University of Colorado at Colorado Springs, {philip.brown}@uccs.edu.
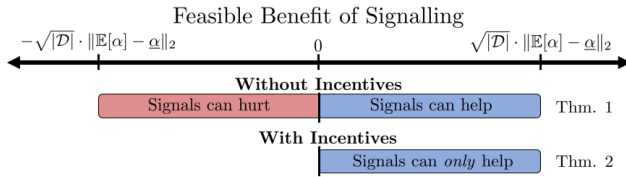
Fig. 1: Illustration of the results in this paper: without incentives, revealing information via signalling can feasibly improve or worsen system cost, however, with the concurrent use of incentives, revealing information can only help and offers similar improvement capabilities.

concurrently to improve system performance.

In this work, we provide bounds on the possible benefit a signalling policy can provide with and without the concurrent use of monetary incentives. When the system designer uses only information signalling, Theorem 1 shows that though signalling has the capability to improve system performance, there exist scenarios where revealing information can create worse system behavior. Conversely, when the system operator can concurrently use signalling and incentive mechanisms, we show that they can guarantee that adding information does not worsen system performance while still offering similar capabilities for improvement. Interestingly, we find that these bounds on the benefit of signalling do not depend on the exact form of the latency functions, in contrast to the many works on selfish routing where the 'steepness' of the latency function causes greater inefficiency [25]; rather, the possible benefit of information signalling increases with the complexity of the model of the unknown parameters and the amount of randomness.

We bolster these conclusions with a numerical example in Section V, in which we find that without incentives, providing more information to the users makes system performance worse; however, when utilizing incentives, providing more information to the users improves system performance.

II. MODEL

A. Congestion Games

In this work, we consider the framework of network congestion games. Consider a parallel, directed graph as a set E of n edges that each connect a single source s to terminal t and can be traversed by a divisible mass of traffic. Let $f_e \geq 0$ denote the mass of traffic utilizing edge $e \in E$, and let $f = \{f_e\}_{e \in E}$ denote a *flow* in the network where $r = \sum_{e \in E} f_e$ is the rate of traffic or demand in the network[1].

When a larger number of users traverse the same path, the congestion on that path increases. To characterize this, each edge is endowed with a *latency function* $\ell_e(f_e) = \sum_{d \in \mathcal{D}} \alpha_{e,d} \cdot (f_e)^d$, where $\alpha_{e,d} \geq 0$ for all $e \in E$ and $d \in \mathcal{D}$. The latency function is a positive, increasing, convex polynomial where $\mathcal{D} = (d_1, \ldots, d_k)$ and $d_i \in \mathbb{Z}_{\geq 0}$ expresses the possible degrees[2], e.g., $\mathcal{D} = \{0, 1\}$ represents affine congestion rates [18], $\mathcal{D} = \{0, 4\}$ can represent the well-known Bureau of Public Roads (BPR) latency functions, commonly used to model the congestion characteristics of

---

[1] Without loss of generality, we can assume $r = 1$, the proof of this is in Remark 1 in the appendix.

[2] We assume that $0 \in \mathcal{D}$ is always satisfied.

physical roads [26], [27], and $\mathcal{D} = \{0, \ldots, D\}$ can represent any positive, convex, increasing polynomial up to degree D [21]. A congestion game of this form can thus be described by the tuple $G = (E, r, \{\alpha_{e,d}\}_{e \in E, d \in \mathcal{D}})$. Let $\alpha \in \mathbb{R}_{\geq 0}^{|E| \cdot |\mathcal{D}|}$ be a vector containing the coefficients of each polynomial term in each latency function. Notice that the size and values in α encode the size of the network and each edge's latency function; as such, we will use α as a variable that encodes all relevant information about the network structure throughout.

The overall congestion in the system can be measured by the *total latency*

$$\mathcal{L}(f; \alpha) = \sum_{e \in E} f_e \cdot \ell_e(f_e), \tag{1}$$

where each $\ell_e(\cdot)$ is inherently determined by α. We denote an optimal flow by $f^{\mathrm{opt}} \in \arg\min_f \mathcal{L}(f; \alpha)$. However, the emergent traffic flow need not be determined by a central planner. When users can choose their own path through the network, let $x \in [0, r]$ denote the index of an infinitesimal agent who uses an edge e_x. When agents are left to choose routes that minimize their own observed travel time, i.e., each agent possesses a cost function $J_x(e; f) = \ell_e(f_e)$, then a plausible behavior that can emerge in the system is that of a *Nash flow* f^{Nf}, which satisfies

$$J_x(e_x; f) \leq J_x(e'; f), \quad \forall e' \in E, \ x \in [0, r]. \tag{2}$$

These system states are those where no user has an incentive to change their action and need not be optimal [28]; additionally, the total latency in any Nash flow in a game of this form is the same [29].

B. Information Signalling

To see what opportunities a system designer has in reducing the system's total latency in a Nash flow, we consider the role of information signalling. Consider that the vector α is a random variable with distribution $\mu_0(x) = \mathbb{P}[\alpha = x]$ and support A, where the system operator has knowledge of this parameter's realization, while the system users do not. If the users reach a flow f, we extend (1) to be the *expected total latency* over a distribution μ,

$$\mathcal{L}(f; \mu) = \mathbb{E}_{\alpha \sim \mu} \left[\sum_{e \in E} f_e \cdot \ell_e(f_e) \right].$$

Because $\ell_e(\cdot)$ is determined by α, it is a random variable.

As a method to coordinate system behavior and induce more desirable system states, the system designer may utilize a *signalling policy* π, in which they reveal only a space in which the realization lies. The signalling policy can be parameterized as $\pi = \{\pi_1, \ldots, \pi_m\}$, where $\pi_i \subset A$, $\pi_i \neq \emptyset$, $\cup_{i=1}^m \pi_i = A$, and $\pi_i \cap \pi_j = \emptyset$ for all $i, j \in \{1, \ldots, m\}$. The signalling policy π thus forms a partition of the support of α, where agents are told which subset π_i the realization α is restricted to. These signals are assumed to be *public and truthful*, and can be used to alter the users' belief over the system state. Let $\mu_i(x) = \mathbb{P}[\alpha = x | \alpha \in \pi_i]$ denote the posterior belief users form when they receive π_i, and let $p_i = \int_{x \in \pi_i} \mu_0(x) dx$ denote the probability signal π_i is sent.

Under a signalling policy π, agents may change their chosen path based on which signal they receive. Let $\mathbf{f} = \{f(\pi_i)\}_{i=1}^{m}$ denote the tuple containing a flow that occurs at the reception of each signal, and let $\sigma_x = \{\sigma_x(\pi_i) \in E\}_{i=1}^{m}$ denote the edge user $x \in [0, r]$ selects after receiving each signal. When each agent adopts a strategy based on the information system's signals, the system designer now cares about the expected total latency, expressed as

$$\mathcal{L}(\mathbf{f}; \mu_0, \pi) = \sum_{i=1}^{m} p_i \cdot \mathcal{L}(f(\pi_i); \mu_i). \qquad (3)$$

An agent's cost will now be their expected travel time,

$$J_x(\sigma_x; \mathbf{f}, \mu_0, \pi) = \sum_{i=1}^{m} p_i \cdot \mathop{\mathbb{E}}_{\alpha \sim \mu_i} \left[\ell_{\sigma_x(\pi_i)} \left(f_{\sigma_x(\pi_i)}(\pi_i) \right) \right].$$

We can now define a *Bayes-Nash Equilibrium* as a tuple $(\mathbf{f}^{\mathrm{BNf}}, \sigma^{\mathrm{BNf}})$ as a set of strategies where no agent elects to unilaterally change, i.e.,

$$J_x(\sigma_x^{\mathrm{BNf}}; \mathbf{f}^{\mathrm{BNf}}, \mu_0, \pi) \leq J_x(\sigma'; \mathbf{f}^{\mathrm{BNf}}, \pi, \mu_0),$$
$$\forall \sigma' \in E^m, \ x \in [0, r]. \quad (4)$$

Our main focus in this work is understanding what opportunities a system designer has in lowering the expected total latency by way of information provisioning, i.e., comparing $\mathcal{L}(\mathbf{f}^{\mathrm{BNf}}; \mu, \pi)$ with $\mathcal{L}(\mathbf{f}^{\mathrm{BNf}}; \mu, \emptyset)$ (where the use of $\emptyset$ where a π was expected denotes the case where no information is shared with users). To quantify this improvement in system performance, we define the *benefit to system cost* as our performance metric

$$\mathbf{B}(\pi; \mu) = \mathcal{L}(\mathbf{f}^{\mathrm{BNf}}; \mu, \emptyset) - \mathcal{L}(\mathbf{f}^{\mathrm{BNf}}; \mu, \pi), \qquad (5)$$

which measures reduction in system cost from utilizing a signal policy π. The system operator's objective is to institute a signalling structure that reduces the system cost or, equivalently, has a positive benefit. The system operator need also avoid cases where the signalling policy increases the system cost or has a negative benefit to performance.

C. Illustrative Example

To illustrate the model, consider the Bayesian congestion game shown in Fig. 2. In this problem instance, the network has two edges $E = \{1, 2\}$, with latency functions $\ell_1(f_1) = \alpha_{1,0}$ and $\ell_2(f_2) = \alpha_{2,1} f_2$. The latency coefficients $\alpha_{1,0}, \alpha_{2,1}$ are unknown to the users and instead they possess a prior belief μ_0 that is distributed over their support $A = [0, 1]^2$.

Though the users do not observe the realizations of $\alpha_{1,0}, \alpha_{2,1}$, they can receive some information from the system operator. Under the signalling policy $\pi = \{\pi_1, \pi_2, \pi_3\}$, if the realization $\alpha \in \pi_i$ then the system operator sends a signal and the users update their belief to the Bayesian posterior μ_i with this knowledge. For example if $\alpha_{1,0} = 0.2$ and $\alpha_{2,1} = 0.6$, then the system designer will inform users that $\alpha \in \pi_2$ and the users will update their beliefs to $\mu_2(x) = \mathbb{P}[x = \alpha | \alpha \in \pi_2]$. Given a signalling policy π, the expected total latency of the Bayes-Nash flow is computed over $\alpha \in A$, where the flow changes based on which region α is realized.
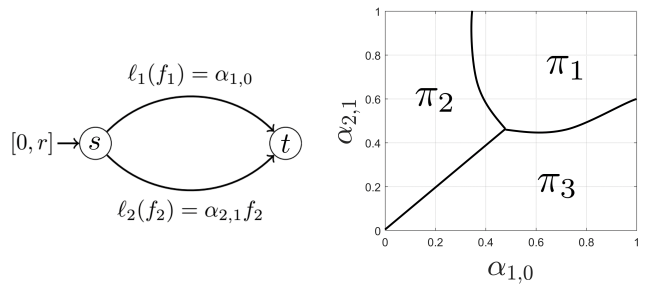

Fig. 2: Two-link, parallel, Bayesian congestion game. One edge possesses a linear latency function, the other a constant latency function. The coefficients of each of these latency functions $\alpha_{1,0}, \alpha_{2,1}$ are unknown but distributed with prior μ_0 over $A = [0, 1]^2$. A signalling policy $\pi = \{\pi_1, \pi_2, \pi_3\}$ partitions A to map realizations to signals.

D. Summary of Contributions

The main contributions of this work come in characterizing the capabilities a system designer has in improving the system cost by revealing information about the system state to its users. We do this by deriving bounds on the benefit a signalling policy can provide. In Section III, we derive these bounds in the setting where an information provider can implement a public and truthful signalling policy; In Theorem 1, we find that there exists significant opportunities for the system designer to provide a positive benefit and improve performance via signalling, however, there is also the possibility that providing information has a negative benefit to system performance and can actually increase system cost, making it difficult for a system designer to guarantee improvement.

In Section IV, we add to the model by considering the case where a system designer need not only use an information signal, but may also design incentives they can levy on the users. In this setting, we find in Theorem 2 that when utilizing well designed tolls, the system designer has similar capabilities in reducing congestion but can now guarantee the signalling policy will never increase system cost, i.e., the benefit is always positive. These results show that implementing incentive and signalling mechanisms concurrently can help a system designer avoid worsening system performance while maintaining their capabilities to improve it.

Finally, in Section V, we provide a numerical simulation of the benefit of signalling. When using a specific from of signalling policy that forms a uniform grid over the support, Fig. 3 shows the change in benefit when more information is revealed (i.e., higher granularity of the grid). We see in this example that without incentives, there is an increasingly negative benefit to revealing information, however with incentives, the benefit is positive.

III. SIGNALLING ALONE

Consider a realization of a congestion game G with latency function coefficients α. Let $\mathcal{L}^{\mathrm{Nf}}(\alpha)$ denote the total latency in a Nash flow in this network (recall that this value is unique for each α, [29]). When a system designer seeks to improve system performance via a public and truthful information system, Theorem 1 provides bounds on the benefit a signalling policy can provide.

Theorem 1. *Consider the class of parallel Bayesian congestion games with polynomial latency functions whose degrees come from the set $\mathcal{D}$. For any distribution over the latency coefficients μ_0 and any signalling policy π, the benefit in the expected total latency of a Bayes-Nash flow from signalling satisfies*

$$-\sqrt{|\mathcal{D}|}\cdot\|\mathbb{E}[\alpha]-\underline{\alpha}\|_2 \leq \mathbf{B}(\pi;\mu_0) \leq \sqrt{|\mathcal{D}|}\cdot\|\mathbb{E}[\alpha]-\underline{\alpha}\|_2, \quad (6)$$

where $\mathbb{E}[\alpha] = \int_{x\in A} x \cdot \mu_0(x)dx$, and $\underline{\alpha} \in \mathbb{R}_{\geq 0}^{|E|\cdot|\mathcal{D}|}$ such that $\underline{\alpha}_{e,d} = \inf\{\mathrm{supp}(\alpha_{e,d})\}$ for each $e \in \bar{E}$, $d \in \mathcal{D}$. Additionally, there exists a μ_0 such that for any $\pi \neq \emptyset$,

$$\mathbf{B}(\pi;\mu_0) = \sqrt{|\mathcal{D}|} \cdot \|\mathbb{E}[\alpha] - \underline{\alpha}\|_2. \quad (7)$$

Further, if $d \in \mathcal{D}$ where $d > 0$, then there exists a μ_0 such that for any $\pi \neq \emptyset$,

$$\mathbf{B}(\pi;\mu_0) = -\|\mathbb{E}[\alpha] - \underline{\alpha}\|_2. \quad (8)$$

The proof of Theorem 1 appears at the end of this section.

Theorem 1 reveals the capabilities a signalling policy has in improving system performance. It also reveals the reality that revealing information can make system performance worse. The bounds for the benefit of a signalling policy depends on the number of terms considered in each latency function $|\mathcal{D}|$ as well as the distance between the average system state and the edge of its support $\|\mathbb{E}[\alpha] - \underline{\alpha}\|_2$. One can think that the number of latency terms $|\mathcal{D}|$ characterizes the complexity of the model of network congestion while $\|\mathbb{E}[\alpha] - \underline{\alpha}\|_2$ measures the amount of uncertainty about the system parameters. Additionally, (7) and (8) shows that there exist situations where regardless of what signalling policy is chosen, revealing information can greatly benefit or hinder system performance.

To prove Theorem 1, we first propose three lemmas. The proof of each appears in the appendix. Lemma 1 shows how we can characterize Bayes-Nash flows by looking at the expected latency coefficients after receiving a signal.

Lemma 1. *With a prior μ_0 and a signalling policy π, the Bayes-Nash flow $\mathbf{f}^{\mathrm{BNf}}$ can be characterized by $\{\bar{f}^{\mathrm{Nf}}(\pi_i)\}_{i=1}^m$, where $\bar{f}^{\mathrm{Nf}}(\pi_i)$ is the Nash flow in the network G with coefficients $\overline{\alpha}_i = \mathbb{E}[\alpha|\alpha \in \pi_i]$.*

Next, as a step in proving the bounds of Theorem 1, we show how we can calculate the expected total latency.

Lemma 2. *For a given flow f, the expected total latency with distribution μ over the coefficients α is equal to the total latency in the network with the expected coefficients, i.e., $\mathcal{L}(f;\mu) = \mathcal{L}(f;\mathbb{E}_{\alpha\sim\mu}[\alpha])$.*

Utilizing Lemma 1 and Lemma 2 together, we can see that the expected total latency in a Bayes-Nash flow is equal to the weighted average of the total latency in the expected network after receiving each signal. This fact will be used in proving Theorem 1.

Finally, in Lemma 3, we provide a fact about the function $\mathcal{L}^{\mathrm{Nf}}(\alpha)$ which bounds the difference in total latency between any two realizations of edge latency coefficients.

Lemma 3. *Consider the class of parallel congestion games with polynomial latency functions with degrees coming from the set $\mathcal{D}$. Let $a, b \in \mathbb{R}_{\geq 0}^{|E|\cdot|\mathcal{D}|}$ be two possible sets of coefficients for a congestion game with edge set E, then*

$$\sqrt{|\mathcal{D}|} \geq \frac{\mathcal{L}^{\mathrm{Nf}}(a) - \mathcal{L}^{\mathrm{Nf}}(b)}{\|a - b\|_2}. \quad (9)$$

Proof of Theorem 1: We start by proving the lower bound on $\mathbf{B}(\pi;\mu_0)$, which quantifies how much the use of a signalling policy can worsen the system performance. Consider the prior μ_0 and signalling policy π that sends a signal to inform users that the state realization $\alpha \in \pi_i$. If μ_i is the posterior formed from receiving signal π_i, let $\overline{\alpha}_i = \mathbb{E}_{\alpha\sim\mu_i}[\alpha]$.

To prove the lower bound, first, define the set $\hat{A} = \prod_{e\in E, d\in\mathcal{D}}[\inf\{\mathrm{supp}(\alpha_{e,d}), \sup\{\mathrm{supp}(\alpha_{e,d})]$ that is the smallest box in $\mathbb{R}_{\geq 0}^{|E|\cdot|\mathcal{D}|}$ that contains $\mathrm{supp}(\alpha)$. Note that $\underline{\alpha}$ is the corner of this box that is closest to the origin. Let, $\hat{\mathcal{L}}^{\mathrm{Nf}}$ be the concave closure of the function $\mathcal{L}^{\mathrm{Nf}}$ over $\hat{A}$, i.e.,

$$\hat{\mathcal{L}}^{\mathrm{Nf}} = \sup\left\{z|(\alpha, z) \in \mathrm{Conv}_{\hat{A}}(\mathcal{L}^{\mathrm{Nf}})\right\},$$

where $\mathrm{Conv}_{\hat{A}}(\mathcal{L}^{\mathrm{Nf}})$ denotes the convex hull of the graph of $\mathcal{L}^{\mathrm{Nf}}$ over the domain $\hat{A}$. With this, we show

$$\mathbf{B}(\pi;\mu_0) = \mathcal{L}(\mathbf{f}^{\mathrm{BNf}};\mu,\emptyset) - \mathcal{L}(\mathbf{f}^{\mathrm{BNf}};\mu,\pi) \quad (10a)$$

$$= \mathcal{L}^{\mathrm{Nf}}(\overline{\alpha}_0) - \sum_{i=1}^m p_i \cdot \mathcal{L}^{\mathrm{Nf}}(\overline{\alpha}_i) \quad (10b)$$

$$\geq \mathcal{L}^{\mathrm{Nf}}(\underline{\alpha}) - \hat{\mathcal{L}}^{\mathrm{Nf}}(\overline{\alpha}_0) \quad (10c)$$

$$= \hat{\mathcal{L}}^{\mathrm{Nf}}(\underline{\alpha}) - \hat{\mathcal{L}}^{\mathrm{Nf}}(\overline{\alpha}_0) \quad (10d)$$

$$\geq -\sqrt{|\mathcal{D}|} \cdot \|\mathbb{E}[\alpha] - \underline{\alpha}\|_2, \quad (10e)$$

where (10b) holds from Lemma 1 and Lemma 2, (10c) holds from $\mathcal{L}^{\mathrm{Nf}}$ monotonically increasing in α and $\underline{\alpha}_{e,d} \leq \alpha_{e,d}$ for all $e \in E$, $d \in \mathcal{D}$, and $\alpha \in A$, as well as the concave closure $\hat{\mathcal{L}}^{\mathrm{Nf}}$ being greater than any concave combination of points in A and $\sum_{i=1}^m p_i\overline{\alpha}_i = \overline{\alpha}_0$, (10d) holds from $\mathcal{L}^{\mathrm{Nf}}(\underline{\alpha}) = \hat{\mathcal{L}}^{\mathrm{Nf}}(\underline{\alpha})$ due to $\underline{\alpha}$ being a corner of $\hat{A}$, and (10e) holds from Lemma 3 along with the observation that the maximum gradient in the concave closure $\hat{\mathcal{L}}^{\mathrm{Nf}}$ must also occur in the original function $\mathcal{L}^{\mathrm{Nf}}$ by the intermediate value theorem.

Now, we prove the upper bound using similar methods:

$$\mathbf{B}(\pi;\mu_0) = \mathcal{L}(\mathbf{f}^{\mathrm{BNf}};\mu,\emptyset) - \mathcal{L}(\mathbf{f}^{\mathrm{BNf}};\mu,\pi) \quad (11a)$$

$$= \mathcal{L}^{\mathrm{Nf}}(\overline{\alpha}_0) - \sum_{i=1}^m p_i \cdot \mathcal{L}^{\mathrm{Nf}}(\overline{\alpha}_i) \quad (11b)$$

$$\leq \mathcal{L}^{\mathrm{Nf}}(\overline{\alpha}_0) - \mathcal{L}^{\mathrm{Nf}}(\underline{\alpha}) \quad (11c)$$

$$\leq \sqrt{|\mathcal{D}|} \cdot \|\mathbb{E}[\alpha] - \underline{\alpha}\|_2. \quad (11d)$$

Together, these two bounds show the range of attainable performance improvements by utilizing a signalling policy π. The fact that $\mathbf{B}(\pi;\mu_0)$ can be negative shows that providing information to users need not always help. In fact, for any signalling policy π, there exist scenarios where adding information can be detrimental to system performance.

To see (7), consider a two link parallel network where $\ell_1(f_1) = 1$, and $\ell_2(f_2) = \sum_{d\in\mathcal{D}}\beta(f_2)^d$, i.e., $\alpha_{2,d} = \beta$

for all $d \in \mathcal{D}$. When $\beta \leq 1/|\mathcal{D}|$, $f_2 = 1$ and $\mathcal{L}^{\mathrm{Nf}}(\beta) = \beta \cdot |\mathcal{D}|$. When $\beta > 1$, $f_1 > 0$ and $\mathcal{L}^{\mathrm{Nf}}(\beta) = 1$. Consider a distribution μ_0 where $\beta = 0$ with probability $1 - \epsilon$ and $\frac{1}{\epsilon \cdot |\mathcal{D}|}$ with probability ϵ. The expected value of β is thus $1/|\mathcal{D}|$ The expected total latency without signalling is $\mathcal{L}^{\mathrm{Nf}}(\beta = 1/|\mathcal{D}|) = 1$. Because any signal π will reveal the two possible realizations of β, the expected total latency with π is

$$\mathcal{L}^{\mathrm{Nf}}(0)(1 - \epsilon) + \mathcal{L}^{\mathrm{Nf}}\left(\frac{1}{\epsilon \cdot |\mathcal{D}|}\right)\epsilon \rightarrow \mathcal{L}^{\mathrm{Nf}}(0) = 0$$

as $\epsilon \rightarrow 0$. The benefit of π is thus

$$\mathcal{L}(\mathbf{f}^{\mathrm{BNf}}; \mu_0, \emptyset) - \mathcal{L}(\mathbf{f}^{\mathrm{BNf}}; \mu_0, \pi) = 1 = -\sqrt{|\mathcal{D}|} \cdot \|\mathbb{E}[\alpha] - \underline{\alpha}\|_2,$$

where α is a vector with $|\mathcal{D}|$ entries of value $\frac{1}{|\mathcal{D}|}$.

To see (8), consider the two link parallel network where $\ell_1(f_1) = (f_1)^d + \beta$ and $\ell_2(f_2) = \beta(f_2)^d + 1$ and $f_1 + f_2 = 1$; in this congestion game, $\beta = \alpha_{1,0} = \alpha_{2,d}$ is a single parameter that represents two, correlated coefficients. It is difficult to characterize the Nash flow in closed form, however, we can utilize the following two facts $(1) \frac{\partial}{\partial \beta} \mathcal{L}^{\mathrm{Nf}}(\beta)|_{\beta=0} = 0$ and (2) $\frac{\partial}{\partial \beta} \mathcal{L}^{\mathrm{Nf}}(\beta)|_{\beta \rightarrow \infty} = 1$. Let μ_0 be a distribution on β. From the first fact, $\mathcal{L}(\mathbf{f}^{\mathrm{BNf}}; \mu_0, \emptyset) = \mathcal{L}^{\mathrm{Nf}}(\overline{\beta}) \rightarrow \mathcal{L}^{\mathrm{Nf}}(0)$ as $\overline{\beta} \rightarrow 0$. Now, consider the prior distribution $\mu_0(x) = \{0, \text{ w.p. } 1 - \epsilon, \overline{\beta}/\epsilon, \text{ w.p. } \epsilon\}$. Any signalling policy π will reveal which β as 0 or $\overline{\beta}/\epsilon$, as such,

$$\mathcal{L}(\mathbf{f}^{\mathrm{BNf}}; \mu_0, \pi) = \mathcal{L}^{\mathrm{Nf}}(0)(1 - \epsilon) + \mathcal{L}^{\mathrm{Nf}}(\overline{\beta}/\epsilon)\epsilon.$$

From fact (2) above, as $\epsilon \rightarrow 0$, $\mathcal{L}(\mathbf{f}^{\mathrm{BNf}}; \mu_0, \pi) \rightarrow \overline{\beta}$. From these two facts, with sufficiently small $\overline{\beta}$,

$$\mathcal{L}(\mathbf{f}^{\mathrm{BNf}}; \mu_0, \emptyset) - \mathcal{L}(\mathbf{f}^{\mathrm{BNf}}; \mu_0, \pi) \rightarrow -\overline{\beta} = -\|\mathbb{E}[\beta] - \underline{\beta}\|_2,$$

where $\underline{\beta} = 0$ to match (8). $\qquad \square$

IV. SIGNALLING & INCENTIVES

Theorem 1 showed that revealing information has the possibility of increasing or decreasing system cost. This section deals with this fact by studying the advantages of concurrently utilizing signals and monetary incentives to influence user behavior. Monetary incentives are a well studied method of influencing user behavior [17]–[21], but, to the authors' best knowledge, the use of information signalling and monetary incentives in tandem has yet to be studied in the context of traffic networks.

Consider a congestion game G; an incentive designer can apply an incentive $\tau_e \in \mathbb{R}$ to each edge $e \in E$ to change the cost experienced by users utilizing that edge, i.e.,

$$J_x(e; f) = \ell_e(f_e) + \tau_e.$$

To understand how incentives and signals interact, consider a *signal-aware incentive mechanism* $T(\pi_i; \pi, \mu_0)$ that assigns tolls $\{\tau_e(\pi_i)\}_{i=1}^m$ dependent on the signal broadcast by the information provider. A player $x \in [0, r]$ with the strategy

σ_x now observes an expected cost of

$$J_x(\sigma_x; \mathbf{f}, \mu_0, \pi, T)$$
$$= \sum_{i=1}^m p_i \cdot \mathbb{E}_{\alpha \sim \mu_i} \left[\ell_{\sigma_x(\pi_i)} \left(f_{\sigma_x(\pi_i)}(\pi_i) \right) + \tau_{\sigma_x(\pi_i)}(\pi_i) \right].$$

The Bayes-Nash flow definition remains as shown in (4), but now with users' tolled cost. We now seek to understand the effectiveness of jointly implementing a signalling policy π and an incentive mechanism T. As such, we extend the definition of (5) which quantifies the gain in system performance to include the effect of an incentive mechanism T, i.e.,

$$\mathbf{B}(\pi; \mu, T) = \mathcal{L}(\mathbf{f}^{\mathrm{BNf}}; \mu, \emptyset, T) - \mathcal{L}(\mathbf{f}^{\mathrm{BNf}}; \mu, \pi, T). \quad (12)$$

In Proposition 1, we first identify a signal-aware incentive mechanism that is optimal for a given signalling policy π.

Proposition 1. *Let μ_0 be a prior on the coefficients in a parallel congestion game with polynomial latency functions, and let π be a signalling policy. An optimal signal aware incentive mechanism $T^\star$ (i.e., maximizes $\mathbf{B}(\pi; \mu, T)$) assigns tolls as*

$$\tau_e^\star(\pi_i) = \sum_{d \in \mathcal{D}} d \cdot \mathbb{E}_{\alpha_{e,d} \sim \mu_i} [\alpha_{e,d}](x_e)^d, \quad (13)$$

where $x \in \arg\min_f \mathcal{L}(f; \mathbb{E}_{\alpha \sim \mu_i}[\alpha])$.

This incentive mechanism is reminiscent of a marginal cost toll in the deterministic setting, and possesses many of the same properties in terms of influencing Nash flow behavior to be optimal.

Consider a realization of a congestion game G with latency function coefficients α. Let $\mathcal{L}^\star(\alpha)$ denote the total latency in the Nash flow in this network under the incentive mechanism $T^\star$.

Theorem 2 provides bounds on the benefit a signalling policy can provide while also utilizing the signal-aware incentive mechanism $T^\star$. We see that by concurrently utilizing incentives and signalling, the system designer can guarantee the benefit of signalling is non-negative.

Theorem 2. *Consider the class of parallel Bayesian congestion games with polynomial latency functions whose degrees come from the set $\mathcal{D}$. For any distribution over the latency coefficients μ_0 and any signalling policy π, the decrease in the expected total latency of a Bayes-Nash flow from signalling satisfies*

$$0 \leq \mathbf{B}(\pi; \mu_0, T^\star) \leq \sqrt{|\mathcal{D}|} \cdot \|\mathbb{E}[\alpha] - \underline{\alpha}\|_2, \quad (14)$$

where $\mathbb{E}[\alpha] = \int_{x \in A} x \cdot \mu_0(x) dx$, and $\underline{\alpha} \in \mathbb{R}_{\geq 0}^{|E| \cdot |\mathcal{D}|}$ such that $\underline{\alpha}_{e,d} = \inf\{\mathrm{supp}(\alpha_{e,d})\}$ for each $e \in E$, $d \in \mathcal{D}$.

Comparing the bounds on the benefit of a signalling policy with and without the use of incentives (i.e., (6) and (14)), we see that incentives can make the use of signals more robust and guarantee they improve performance while allowing for similar opportunities to improve performance. We further support this conclusion in Section V by providing
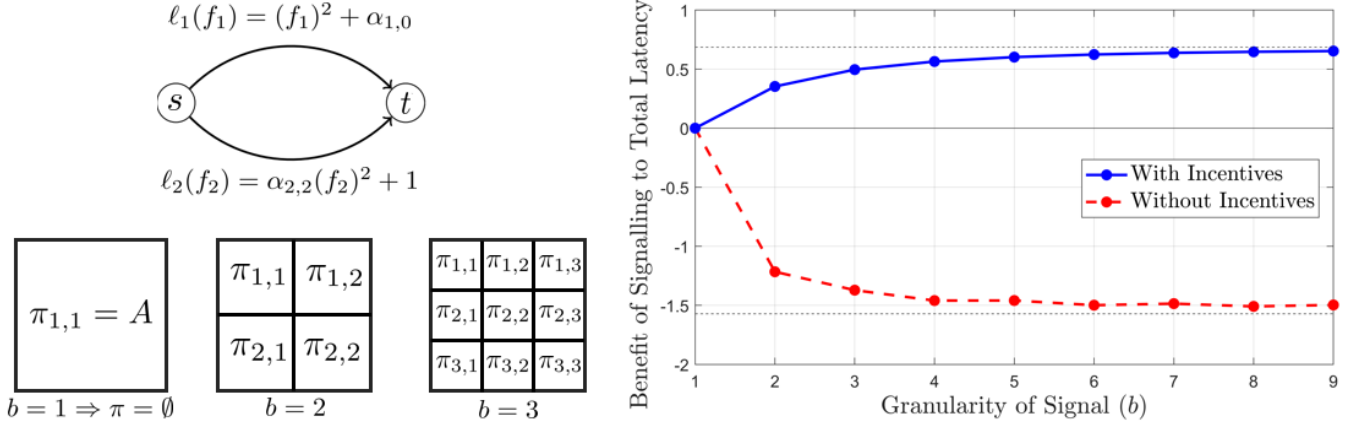
Fig. 3: The benefit of revealing information with and without the concurrent use of incentives. π is the uniform-grid signal structure where the support A is partitioned into a grid with granularity b; as b increases, more information is revealed to the users. At left, the benefit of using the uniform-grid signalling policy π^b is shown for the setting described in Section V with and without the concurrent use of the incentive mechanism $T^\star$. When incentives are used, revealing information provides a positive benefit and improves performance. With no incentives, the benefit becomes negative and revealing information worsens system cost.

a numerical example and comparing the benefit of revealing information with and without incentives.

Before we prove Theorem 2, we state the following lemma that is similar to Lemma 3 but applies to $\mathcal{L}^\star$.

Lemma 4. *Consider the class of parallel congestion games with polynomial latency functions with degrees coming from the set $\mathcal{D}$. Let $a, b \in \mathbb{R}_{\geq 0}^{|E| \cdot |\mathcal{D}|}$ be two possible sets of coefficients for a congestion game with edge set E, then*

$$\sqrt{|\mathcal{D}|} \geq \frac{\mathcal{L}^\star(a) - \mathcal{L}^\star(b)}{\|a - b\|_2}. \tag{15}$$

Proof of Theorem 2: We start with the proof of the lower bound. Consider the prior μ_0 and signalling policy π that sends a signal to inform users that the state realization $\alpha \in \pi_i$. If μ_i is the posterior formed from receiving signal π_i, then let $\overline{\alpha}_i = \mathbb{E}_{\alpha \sim \mu_i}[\alpha]$. When utilizing the incentive mechanism $T^\star$ which assigns incentives as stated in (13), then Proposition 1 states that the equilibrium flow that emerges when signal π_i is received will be $f^\star(\pi_i) \in \arg\min \mathcal{L}(f; \overline{\alpha}_i)$; as such $\mathcal{L}^\star(\overline{\alpha}_i)$ is the total latency that occurs.

First, we observe that $\mathcal{L}^\star(\alpha)$ is concave. $\mathcal{L}^\star(\alpha)$ can be expressed as the infimum over $f \in \Delta(E)$ for a given α,

$$\mathcal{L}^\star(\alpha) = \inf_{f \in \Delta(E)} \mathcal{L}(f; \alpha) = \inf_{f \in \Delta(E)} \sum_{e \in E} \sum_{d \in \mathcal{D}} \alpha_{e,d}(f_e)^{d+1}.$$

Observing that $\sum_{e \in E} \sum_{d \in \mathcal{D}} \alpha_{e,d}(f_e)^{d+1}$ is affine in α (and thus concave in α) for each f, we can invoke that the pointwise infimum over a class of functions that are each concave is itself, concave [30]. Thus $\mathcal{L}^{\mathrm{opt}}(\alpha)$ is concave though need not be affine. Now, consider the total latency in a Bayes-Nash flow with signal policy π and incentive $T^\star$,

$$\mathcal{L}(\mathbf{f}^{\mathrm{BNf}}; \mu_0, \pi, T^\star) = \sum_{i=1}^{m} p_i \cdot \mathcal{L}^\star(\overline{\alpha}_i) \tag{16a}$$

$$\leq \mathcal{L}^\star(\sum_{i=1}^{m} p_i \cdot \overline{\alpha}_i) \tag{16b}$$

$$= \mathcal{L}^\star(\overline{\alpha}_0) = \mathcal{L}(\mathbf{f}^{\mathrm{BNf}}; \mu_0, \emptyset, T^\star), \tag{16c}$$

where the (16a) holds from Lemma 1 and Lemma 2 and (16b) holds from the concavity of $\mathcal{L}^\star$. From this, we can see that $\mathbf{B}(\pi; \mu_0, T^\star) \geq 0$, i.e., while utilizing incentive scheme $T^\star$, signalling cannot increase system cost.

For the upper bound, we show that

$$\mathbf{B}(\pi; \mu_0, T^\star) = \mathcal{L}(\mathbf{f}^{\mathrm{BNf}}; \mu, \emptyset, T^\star) - \mathcal{L}(\mathbf{f}^{\mathrm{BNf}}; \mu, \pi, T^\star)$$

$$= \mathcal{L}^\star(\overline{\alpha}_0) - \sum_{i=1}^{m} p_i \cdot \mathcal{L}^\star(\overline{\alpha}_i) \tag{17a}$$

$$\leq \mathcal{L}^\star(\overline{\alpha}_0) - \mathcal{L}^\star(\underline{\alpha}) \tag{17b}$$

$$\leq \sqrt{|\mathcal{D}|} \cdot \|\mathbb{E}[\alpha] - \underline{\alpha}\|_2, \tag{17c}$$

where the (17a) holds from Lemma 1 and Lemma 2, (17b) holds from $\mathcal{L}^\star$ monotonically increasing with α, and (17c) holds from Lemma 4. $\qquad\square$

V. NUMERICAL EXAMPLE

To understand how the benefit of signalling changes as more information is revealed, we offer the following numerical example. Consider a Bayesian congestion game with two edges, $\ell_1(f_1) = f_1^2 + \alpha_{1,0}$ and $\ell_2(f_2) = \alpha_{2,2} f_2^2 + 1$, where $\alpha_{1,0}$ and $\alpha_{2,2}$ are parameters unknown to the user. We consider that these two parameters are drawn from a truncated normal distribution, i.e., let

$$z \sim \mathcal{N}\left(\begin{bmatrix} 15 \\ 15 \end{bmatrix}, 180 \cdot \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\right),$$

and define the prior over $\alpha_{1,0}, \alpha_{2,2}$ as $\mu_0(x) = \mathbb{P}[x = z | z \in A]$ where $A = [0, 60]^2$ is their support.

Now, we analyze the benefit of the *uniform-grid signalling policy* with and without the concurrent use of the signal-aware incentive mechanism $T^\star$ as defined in Proposition 1. Let b be an integer representing the granularity of the signalling mechanism, i.e., the number of times A is partitioned along each dimension as shown in Fig. 3, i.e.,

$$\pi_{i,j} = \left[\frac{60}{b}(i-1), \frac{60}{b}i\right] \times \left[\frac{60}{b}(j-1), \frac{60}{b}j\right],$$

essentially forming a uniform grid over A.

In Fig. 3, we plot the benefit for using the uniform-grid signalling policy with and without the concurrent use of the incentive mechanism $T^\star$, i.e., $\mathbf{B}(\pi; \mu_0)$ and $\mathbf{B}(\pi; \mu_0, T^\star)$. Observe that when no incentives are used, increasing the amount of information revealed to users (i.e., larger b) causes the benefit to become increasingly negative; meaning as more information is revealed, the signalling policy makes the system performance worse. Conversely, while using incentive mechanism $T^\star$, as more information is added, the benefit becomes increasingly positive and revealing information now improves performance.

VI. Conclusion

This work studies the capabilities of an information signalling system with and without the concurrent use of incentives. We show that when signalling alone, revealing information has the possibility to increase system cost, however, with the concurrent use of incentives and signals, revealing information can only improve the system performance. Future work will look into how to design optimal signalling policies while using incentives and will consider how to co-design information and signals when each mechanism has different amounts of available information (e.g., non-signal-dependent incentives or semi informed incentives).

References

[1] P. Palensky and D. Dietrich, "Demand side management: Demand response, intelligent energy systems, and smart loads," *IEEE Transactions on Industrial Informatics*, vol. 7, no. 3, pp. 381–388, aug 2011.

[2] A. C. Pigou, *The Economics of Welfare*. Macmillan, New York, 1920.

[3] M. Esmaeili, M. B. Aryanezhad, and P. Zeephongsekul, "A game theory approach in seller-buyer supply chain," *European Journal of Operational Research*, vol. 195, no. 2, pp. 442–448, jun 2009.

[4] O. T. Oreagba, O. O. Ogunnaike, and O. J. Kehinde, "Capitalizing on Game Theory for Optimal Marketing Decision in Service Industry: Evidence From Telecommunication Industry in Nigeria," *SAGE Open*, vol. 11, no. 2, p. 21582440211023199, 2021.

[5] G. Hardin, "The Tragedy of the Commons," *Science*, vol. 162, no. 3859, pp. 1243–1248, 1968.

[6] J. Liu, S. Amin, and G. Schwartz, "Effects of information heterogeneity in Bayesian routing games," *arXiv preprint:1603.08853*, 2016.

[7] E. Kamenica and M. Gentzkow, "Bayesian Persuasion," *American Economic Review*, vol. 101, no. 6, pp. 2590–2615, oct 2011.

[8] I. Goldstein and C. Huang, "Bayesian Persuasion in Coordination Games," *American Economic Review*, vol. 106, no. 5, pp. 592–596, may 2016.

[9] D. Bergemann and S. Morris, "Information design: A unified perspective," *Journal of Economic Literature*, vol. 57, no. 1, pp. 44–95, 2019.

[10] F. Sezer, H. Khazaei, and C. Eksin, "Social Welfare Maximization and Conformism via Information Design in Linear-Quadratic-Gaussian Games," *arXiv preprint:2102.13047*, 2021.

[11] R. Lindsey, T. Daniel, E. Gisches, and A. Rapoport, "Pre-trip information and route-choice decisions with stochastic travel conditions: Theory," *Transportation Research Part B: Methodological*, vol. 67, no. C, pp. 187–207, 2014.

[12] M. Castiglioni, A. Celli, A. Marchesi, and N. Gatti, "Signaling in Bayesian Network Congestion Games: the Subtle Power of Symmetry," feb 2020. *arXiv preprint:2002.05190*

[13] Y. Zhu and K. Savla, "A Semidefinite Approach to Information Design in Non-atomic Routing Games," *CoRR*, vol. abs/2005.0, 2020.

[14] M. Wu and S. Amin, "Information design for regulating traffic flows under uncertain network state," in *Allerton Conference on Communication, Control, and Computing*. IEEE, 2019, pp. 671–678.

[15] M. Wu, S. Amin, and A. E. Ozdaglar, "Value of Information in Bayesian Routing Games," aug 2018. *arXiv preprint:1808.10590*

[16] H. Tavafoghi and D. Teneketzis, "Informational incentives for congestion games," in *55th Annual Allerton Conference on Communication, Control, and Computing, Allerton 2017*, vol. 2018-Jan, pp. 1285–1292.

[17] L. Fleischer, K. Jain, and M. Mahdian, "Tolls for Heterogeneous Selfish Users in Multicommodity Networks and Generalized Congestion Games," in *Proc. 45th IEEE Symp. on Foundations of Computer Science*, Rome, Italy, 2004, pp. 277–285.

[18] B. L. Ferguson, P. N. Brown, and J. R. Marden, "Utilizing Information Optimally to Influence Distributed Network Routing," in *IEEE Conference on Decision and Control*, vol. 2019-Decem. , pp. 5008–5013.

[19] ——, "The Effectiveness of Subsidies and Tolls in Congestion Games," *IEEE Transactions on Automatic Control*, pp. 1–1, 2021.

[20] R. Cole, Y. Dodis, and T. Roughgarden, "Pricing network edges for heterogeneous selfish users," in *Proc. of the 35th ACM symp. on Theory of computing*, New York, USA, 2003, pp. 521–530.

[21] D. Paccagnan, R. Chandan, B. L. Ferguson, and J. R. Marden, "Optimal Taxes in Atomic Congestion Games," *ACM Trans. Econ. Comput.*, vol. 9, no. 3, aug 2021.

[22] A. de Palma, R. Lindsey, and N. Picard, "Risk Aversion, the Value of Information, and Traffic Equilibrium," *Transportation Science*, vol. 46, 2012.

[23] G. Egorov and K. Sonin, "Persuasion on Networks," National Bureau of Economic Research, Working Paper 27631, jul 2020.

[24] N. Heydaribeni and A. Anastasopoulos, "Joint Information and Mechanism Design for Queues with Heterogeneous Users," in *IEEE Conference on Decision and Control (CDC)*, 2021, pp. 3806–3811.

[25] J. R. Correa, A. S. Schulz, and N. E. Stier-Moses, "A geometric approach to the price of anarchy in nonatomic congestion games," *Games and Economic Behavior*, vol. 64, no. 2, pp. 457–469, 2008.

[26] F. Farokhi and K. H. Johansson, "A piecewise-constant congestion taxing policy for repeated routing games," *Transportation Research Part B: Methodological*, vol. 78, pp. 123–143, 2015.

[27] R. Singh and R. Dowling, "Improved Speed-Flow Relationships: Application To Transportation Planning Models," *Seventh TRB Conference on the Application of Transportation Planning Methods*, no. 4, pp. 340–349, 2002.

[28] T. Roughgarden, "How Bad Is Selfish Routing?" *J. ACM*, vol. 49, no. 2, pp. 236–259, 2002.

[29] A. Mas-Colell, "On a Theorem of Schmeidler," *Mathematical Economics*, vol. 13, pp. 201–206, 1984.

[30] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[31] R. W. Rosenthal, "A class of games possessing pure-strategy Nash equilibria," *International Journal of Game Theory*, vol. 2, no. 1, pp. 65–67, dec 1973.

Appendix

Remark 1. *Without loss of generality, we can assume a unit traffic rate, $r = 1$, even when $r \sim \nu$ is a random variable.*

proof: Consider a congestion game G with demand r. Define a mapping $Q(G, \gamma)$ that outputs a new congestion game $\hat{G}$ with latency functions $\hat{\ell}_e(x) = \sum_{d \in \mathcal{D}} \frac{\alpha_{e,d}}{d+1}(f_e)^d$. Let f be a flow in G with total traffic r. Now, consider the flow $\gamma f = \{\gamma \cdot f_e\}_{e \in E}$ in $\hat{G}$. Each edge $e \in E$ will have latency

$$\hat{\ell}_e(\gamma f_e) = \sum_{d \in \mathcal{D}} \frac{\alpha_{e,d}}{\gamma^{d+1}}(\gamma f_e)^d = \frac{1}{\gamma}\ell_e(f_e).$$

Notice that latency on each edge is scaled by $1/\gamma$ and the preference structure is preserved; therefore, if f is a Nash flow in G then γf is a Nash flow in $\hat{G}$. Further,

$$\mathcal{L}(\gamma f; \hat{G}) = \sum_{e \in E} \gamma f_e \hat{\ell}_e(\gamma f_e) = \sum_{e \in E} f_e \ell_e(f_e) = \mathcal{L}(f; G),$$

and the two networks will have the same total latency.

If $(\alpha, r) \sim \mu_0$, e.g., $\mu_0(x, y) = \mathbb{P}[\alpha = x, r = y]$, then consider that $\hat{\alpha} \sim \hat{\mu}_0$, where $\hat{\mu}_0(z) = \sum_{x,y | Q(x, 1/y) = z} \mu_0(x, y)$. Now $\hat{\alpha}$ has the same distribution over total latency. $\square$

Proof of Lemma 1: Consider the prior μ_0 on α and the signalling policy π. If the signal π_i is sent to users, they update their belief via Bayesian inference to $\mu_i(x) = \frac{\mu_0(x)}{p_i}\mathbb{1}_{[x \in \pi_i]}$. In a flow f, user $x \in [0, r]$ experiences an expected cost of

$$J_x(e_x; f, \mu_i) = \underset{\alpha \sim \mu_i}{\mathbb{E}}\left[\sum_{d \in \mathcal{D}} \alpha_{e_x,d}(f_{e_x})^d\right]$$
$$= \sum_{d \in \mathcal{D}} \mathbb{E}[\alpha_{e_x,d}|\alpha \in \pi_i](f_e)^d.$$

Note that if f were not a Nash flow in the parallel congestion game with coefficients $\mathbb{E}[\alpha|\alpha \in \pi_i]$, then (2), at least one user $x \in [0, r]$ would be able to deviate their strategy $\sigma_x(\pi_i)$ and experience lower cost. Therefore, the only Bayes-Nash flows occur when $f(\pi_i)$ is a Nash flow with respect to $\mathbb{E}[\alpha|\alpha \in \pi_i]$ for all $i \in \{1, \dots, m\}$. Further, because the total latency in a Nash flow is unique, so too is the expected total latency in a Bayes-Nash flow. $\square$

Proof of Lemma 2: Consider the distribution μ_0 on α, and let f be a flow in the network. The expected total latency

$$\mathcal{L}(f; \mu) = \underset{\alpha \sim \mu}{\mathbb{E}}\left[\sum_{e \in E} f_e \cdot \sum_{d \in \mathcal{D}} \alpha_{e,d}(f_e)^d\right]$$
$$= \sum_{e \in E} f_e \cdot \sum_{d \in \mathcal{D}} \underset{\alpha \sim \mu}{\mathbb{E}}[\alpha_{e,d}](f_e)^d$$
$$= \mathcal{L}(f; \underset{\alpha \sim \mu}{\mathbb{E}}[\alpha]),$$

which follows from linearity of expected value. $\square$

Proof of Lemma 3 & Lemma 4: We will provide the proof for $\mathcal{L}^{\text{Nf}}$ and show that the same approach holds for $\mathcal{L}^\star$. We assume $r = 1$, which is without loss of generality from Remark 1. Let f be the Nash flow in the parallel congestion game with polynomial coefficients α, i.e., $\mathcal{L}^{\text{Nf}}(\alpha) = \sum_{e \in E}\sum_{d \in \mathcal{D}} \alpha_{e,d}f_e^{d+1}$. Note that the derivative with respect to one coefficient satisfies $0 \leq \frac{\partial}{\partial \alpha_{e,d}}\mathcal{L}^{\text{Nf}}(\alpha) \leq f_e^{d+1}$, as no new users will choose edge e if $\alpha_{e,d}$ unilaterally increases, i.e., $\frac{\partial}{\partial \alpha_{e,d}}f_e \leq 0$ where f is the Nash flow with α, and the total latency will not decrease by increasing an edges latency function[3]. Now, the gradient of the $\mathcal{L}^{\text{Nf}}(\alpha)$ must satisfy

$$\|\nabla\mathcal{L}^{\text{Nf}}(\alpha)\|_2 \leq \sqrt{\sum_{e \in E}\sum_{d \in \mathcal{D}}(f_e^{d+1})^2}$$
$$= \sqrt{\sum_{d \in \mathcal{D}}\left((f_1)^{2(d+1)} + \cdots + (f_n)^{2(d+1)}\right)}$$
$$\leq \sqrt{\sum_{d \in \mathcal{D}}(f_1 + \cdots + f_n)^{2(d+1)}}$$
$$= \sqrt{|\mathcal{D}|}.$$

Where the first inequality holds from the observation above, and the second holds from the super-additivity of positive polynomials in the positive orthant.

---

[3] $\mathcal{L}^{\text{Nf}}$ is continuous but need not be continuously-differentiable. However, if at each point the largest sub-gradient is used, the proof still follows.

Finally, consider two sets of coefficients $a, b \in A$. Observe

$$\frac{\mathcal{L}^{\text{Nf}}(a) - \mathcal{L}^{\text{Nf}}(b)}{\|a - b\|_2}$$
$$= \frac{1}{\|a - b\|_2}\int_{\lambda=0}^1 (a - b)^T\nabla\mathcal{L}^{\text{Nf}}(\lambda a + (1 - \lambda)b)d\lambda$$
$$\leq \frac{1}{\|a - b\|_2}\int_{\lambda=0}^1 \|a - b\|_2 \cdot \|\nabla\mathcal{L}^{\text{Nf}}(\lambda a + (1 - \lambda)b)\|_2 d\lambda$$
$$\leq \int_{\lambda=0}^1 \sqrt{|\mathcal{D}|}d\lambda = \sqrt{|\mathcal{D}|}.$$

Where the first inequality holds from Cauchy-Schwarz, and the second holds from our observation above on the norm of the gradient of $\mathcal{L}^{\text{Nf}}(\alpha)$.

We utilized the fact in a Nash flow f, no more users will utilize an edge when unilaterally increasing one of the latency coefficients on that edge. Because this is also true of the flow $f^\star$ that emerges from incentive mechanism $T^\star$ (which is in fact the optimal flow), the same bound holds as presented in (15). $\square$

Proof of Proposition 1: Consider that the users receive signal π_i from signalling policy π and prior μ_0 (forming posterior μ_i), and realize a flow of f. From 2, the expected total latency in this flow is equal to the total latency of this flow in the expected network, i.e., $\mathcal{L}(f; \mu_i) = \mathcal{L}(f; \overline{\alpha}_i)$, where $\overline{\alpha} = \mathbb{E}_{\alpha \sim \mu_i}[\alpha]$. Thus, an optimal flow at the reception of signal π_i is one that satisfies $f^{\text{opt}}(\pi_i) \in \arg\min_f \mathcal{L}(f; \overline{\alpha}_i)$.

We now look for a incentives that will influence users such that $f^{\text{opt}}(\pi_i)$ becomes a Nash flow in a congestion game with latency coefficients $\overline{\alpha}_i$. To do so, we note that G with flow-varying incentive functions $\tau_e(f_e)$ is a potential game [31] with potential function

$$\Phi(f; \alpha) = \sum_{e \in E}\int_0^{f_e} \ell_e(x) + \tau_e(x)dx.$$

As such, the flow in $\arg\min_f \Phi(f; \alpha)$ is a Nash equilibrium. For the polynomial latency functions considered in this work, let $\tau_e(x) = \sum_{d \in \mathcal{D}} d \cdot \alpha_{e,d}(x)^d$. Now, the potential function becomes

$$\Phi(f; \alpha) = \sum_{e \in E}\int_0^{f_e}\sum_{d \in \mathcal{D}} \alpha_{e,d}x^d + d\alpha_{e,d}x^d dx$$
$$= \sum_{e \in E}\sum_{d \in \mathcal{D}} f_e\alpha_{e,d}(f_e)^d = \mathcal{L}(f; \alpha),$$

and the Nash flow that minimizes Φ also minimizes $\mathcal{L}$; as such, $f^{\text{opt}}(\pi_i)$ becomes a Nash equilibrium in the game with the coefficients $\overline{\alpha}_i$.

Finally, notice that by selecting the fixed incentive $\tau_e^\star(\pi_i) = \tau_e(f_e^{\text{opt}}(\pi_i))$, the equilibrium conditions do not change and $f^{\text{opt}}(\pi_i)$ remains a Nash flow. Nash flows retain the same uniqueness properties under fixed incentives, and thus assigning $\tau^\star(\pi_i)$ minimizes the expected total latency when π_i is sent. If this is done for each signal, the total latency with each signal will me minimal and so too will the overall expected total latency, making $T^\star$ an optimal incentive mechanism. $\square$