Noise-Robust Tolls in Atomic Congestion Games

Davide Sipione [*] **Leonardo Cianfanelli** [*] **Giacomo Como** [*]
Fabio Fagnani [*] **Shinkyu Park** [**]

[*] *Department of Mathematical Sciences, Politecnico di Torino, Italy (e-mail: {name.surname}@ polito.it)*
[**] *Electrical and Computer Engineering, King Abdullah University of Science and Technology (KAUST), Thuwal 23955, Saudi Arabia (e-mail: shinkyu.park@kaust.edu.sa).*

Abstract: We consider a congestion game with a finite number of agents selecting paths according to a logit dynamics with fixed noise level. We propose an iterative distributed tolling scheme and prove its convergence to a toll vector that minimizes the expected total travel time for every noise level.

Keywords: Transportation Networks, Control of Networks, Multi-agent systems, Game Theories, Cyber-physical urban systems

1. INTRODUCTION

We consider a congestion game with a finite number of agents selecting paths according to a logit dynamics with fixed noise level. When the population is infinitely large, marginal cost tolls can restore the social optimum under the assumption of perfect rationality (Beckmann et al. (1956)), but these assumptions rarely hold. In fact, in presence of noise, the invariant distribution of the logit dynamics concentrates, as the number of agents grows large, on perturbed equilibria, which are in general sub-optimal (Sandholm (2011)). In this work, we establish an iterative distributed algorithm for toll design in congestion games to minimize the expected total travel time on the network for every noise level, when the number of players is large and finite. This work is related to Park and Leonard (2025), where an incentive mechanism that guarantees convergence to desirable configurations was provided for every noise level in contractive games with a continuum of users.

Notation: Let Δ and $\Delta_n = \Delta \cap \frac{1}{n}\mathbb{Z}_+^m$ denote the m-dimensional simplex and its discrete counterpart, whose size may be deduced from the context. Let $\delta^\theta(x)$ denote the Dirac delta centered in θ.

2. PROBLEM SETTING

We model the network by a directed multigraph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. We consider a single origin-destination pair o, d in $\mathcal{V}$, with d reachable from o, and let Γ denote the non-empty set of $o-d$ paths. We define the link-path incidence matrix A in $\{0,1\}^{\mathcal{E}\times\Gamma}$ with entries $A_{e\gamma} = 1$ if and only if e belongs to γ. We consider n indistinguishable players, each choosing a path in Γ, and collect the empirical path frequencies into the state vector x in Δ_n, with x_γ denoting the fraction of players choosing γ. The cost of path γ is

$$c_\gamma^\theta(x) = \sum_{e\in\mathcal{E}} A_{e\gamma}\left(\tau_e\left((Ax)_e\right) + \omega_e^\theta\left((Ax)_e\right)\right), \qquad (1)$$

where, for every link e in $\mathcal{E}$, $\tau_e : \mathbb{R}_+ \to \mathbb{R}_+$ is a continuously differentiable, strictly increasing, positive, and convex delay function that accounts for congestion, and $\omega_e^\theta : \mathbb{R}_+ \to \mathbb{R}_+$ is a monetary *toll* parametrized by θ. Since the cost functions depend on the path distribution, this setting defines a population game.

The population state evolves by the following discrete-time Markov chain $X(t)$ on Δ_n, called *logit dynamics*. At each time step, an agent is selected uniformly at random and chooses a random path according to a logit probability distribution. That is, for every state x in Δ_n and paths i, j in Γ with $x_i > 0$, the transition probability is

$$P_{xy}^{(n)} = \begin{cases} x_i \dfrac{e^{-\eta^{-1}c_j^\theta(y)}}{\sum\limits_{k\in\Gamma} e^{-\eta^{-1}c_k^\theta(x+\frac{1}{n}(u^k-u^i))}} & \text{if } y = x + \dfrac{1}{n}(u^j - u^i) \\ 0 & \text{otherwise}, \end{cases}$$

where $\eta > 0$ is the noise level and u^i is the vector with 1 in the i-th entry and 0 elsewhere. The Markov chain $X(t)$ is irreducible for every noise level. Hence, it admits a unique invariant distribution π_n, which depends on the toll functions.

The goal of the planner is to design a vector of toll functions to minimize the expected total travel time, i.e.,

$$\min_{\{\omega_e:\mathbb{R}_+\to\mathbb{R}_+\}_{e\in\mathcal{E}}} \mathbb{E}_{\pi_n}\left[\phi(X)\right], \qquad (2)$$

where

$$\phi(x) = \sum_{\gamma\in\Gamma} x_\gamma \sum_{e\in\mathcal{E}} A_{e\gamma}\tau_e\left((Ax)_e\right), \qquad (3)$$

is the total travel time.

3. OPTIMAL TOLLS

In this work, we shall assume that $\mathcal{V} = \{o, d\}$ and that the network is composed of E parallel links from o to d, so that $\Gamma = \mathcal{E}$. This implies that $\phi(x)$ is strictly convex and therefore

$$\mathcal{X}^* = \arg\min_{x\in\Delta}\{\phi(x)\}\,.$$

is a singleton, i.e., $\mathcal{X}^* = \{x^*\}$. Furthermore, we shall assume for ease of exposition that x^* lies in $\mathrm{int}(\Delta)$, representing the interior of Δ. We remark that x^* is the flow that minimizes the total travel time when the number of agents is infinitely large.

In the limit of infinitely large population, the marginal cost tolls

$$\tilde{\omega}_e(x_e) = x_e\tau_e'(x_e)\,,$$

are known to align the unique equilibrium of the game with the social optimum. However, in presence of noise, these tolls are no longer optimal (Sandholm (2011)). The main idea of this work is to combine marginal cost tolls with noise-dependent tolls

$$\bar{\omega}_e(\theta) = -\eta\log(\theta_e)\ ,$$

with θ in Δ, to address the distortion introduced by the noise. We then propose an iterative distributed algorithm for updating θ in such a way that the toll vector

$$\omega_e^\theta(x_e) = \tilde{\omega}_e(x_e) + \bar{\omega}_e(\theta_e)\,,\quad \forall e\in E\,, \tag{4}$$

converges to the solution of (2) for every noise level $\eta > 0$ and for every initial condition θ^0 in $\mathrm{int}(\Delta)$, in the limit of large population.

The parameter θ is updated through a discrete-time dynamical system that evolves on a much slower time scale than the underlying Markov chain. Let π_n^θ denote the invariant distribution of the Markov chain $X(t)$ with tolls (4). Then, let $\psi : \Delta \to \Delta$ defined by

$$\psi_n(\theta) = \mathbb{E}_{\pi_n^\theta}[X] = \sum_{x\in\Delta_n} x\pi_n^\theta(x)\,,$$

and consider the discrete-time dynamical system

$$\theta^{k+1} = \psi_n(\theta^k),\quad k\geq 0\,.$$

We also define the map

$$\Psi(\theta) = \arg\min_{x\in\Delta}\left\{\eta^{-1}\phi(x) + D_{KL}(x||\theta)\right\}\ ,$$

where $D_{KL}(x||\theta) = \sum_{e\in\mathcal{E}} x_e\log\left(x_e/\theta_e\right)$.

It is possible to verify (cf. Lemma 4) that $\psi_n(\theta)$ converges uniformly to $\Psi(\theta)$, which prompts the study of the limit dynamical system

$$\theta^{k+1} = \Psi(\theta^k)\,. \tag{5}$$

Proposition 1. The limit dynamical system (5) admits a fixed point $\theta^* = x^*$. Moreover, θ^* is the unique fixed point in $\mathrm{int}(\Delta)$, and

$$\lim_{k\to+\infty}\theta^k = \theta^*\,,\quad \forall\theta^0\in\mathrm{int}(\Delta)\,.$$

Proof. The proof is deferred to Appendix B.

We can now establish our main result.

Theorem 1. Consider the toll $\omega^\theta(x)$ defined in (4). For every initial condition θ^0 in $\mathrm{int}(\Delta)$ and for every $\epsilon > 0$, there exists k_ϵ in $\mathbb{N}$ and, for every $k > k_\epsilon$, an integer $n_{\epsilon,k}$ in $\mathbb{N}$ such that

$$\mathbb{E}_{\pi_n^{\theta^k}}\Big[\phi(X)\Big] - \inf_{\pi\in\Delta_n}\mathbb{E}_\pi\Big[\phi(X)\Big] < \epsilon.$$

for every $n > n_{\epsilon,k}$.

Proof. The proof is deferred to Appendix C.

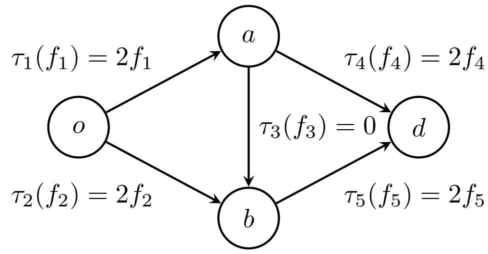

Fig. 1. Wheatstone network.

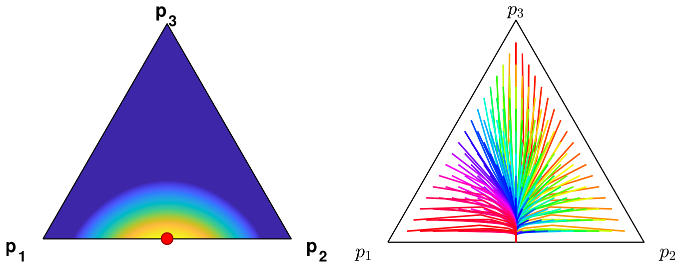

Fig. 2. Left: Invariant distribution of the Markov chain with $n = 500$ players and noise level $\eta = 1$. In red, the social optimum. Right: Trajectories of θ^k under (5) for different initial conditions θ^0, all converging to $\theta^* = (1/2, 1/2, 0)$.

4. EXAMPLE

We now provide a numerical example suggesting that our results may be extended to cases where the social optimum does not lie in the interior of the simplex and to networks with non-parallel links. Consider the Wheatstone network in Figure 1, with path set $\Gamma = \{(o, a, d), (o, b, d), (o, a, b, d)\}$, which we shall denote by p_1, p_2, and p_3, respectively.

We consider a congestion game with $n = 500$ players, and compute the invariant distribution with noise level $\eta = 1$ with tolls (4). As shown in the left plot of Figure 2, the invariant distribution concentrates around the social optimum $(1/2, 1/2, 0)$, which lies on the boundary. The right plot of Figure 2 reports the evolution of θ^k initialized at different values of θ^0, all converging to $\theta^* = (1/2, 1/2, 0)$.

5. CONCLUSION

We proposed an iterative distributed tolling scheme that steers the invariant distribution of a Markov chain toward the social optimum. Future work aim to extend these results to the case when the system-optimum lies on the boundary and to general networks.

REFERENCES

Beckmann, M.J., McGuire, C.B., and Winsten, C.B. (1956). *Studies in the Economics of Transportation*. Yale University Press, New Haven.

Cover, T.M. and Thomas, J.A. (1991). *Elements of Information Theory*. Wiley & Sons, New York.

Mei, W. and Bullo, F. (2017). Lasalle invariance principle for discrete-time dynamical systems: A concise and self-contained tutorial. *arXiv preprint arXiv:1710.03710*.

Monderer, D. and Shapley, L. (1996). Potential games. *Games and Economic Behavior*, 14, 124–143.

Park, S. and Leonard, N. (2025). Learning with delayed payoffs in population games using Kullback–Leibler divergence regularization. *IEEE Transactions on Automatic Control*, PP, 1–16.

Sandholm, W.H. (2011). *Population Games and Evolutionary Dynamics.* MIT Press.

Appendix A. PRELIMINARY LEMMAS

We start by characterizing the invariant distribution of the Markov Chain $X(t)$.

Lemma 1. The Markov Chain $X(t)$ admits the following unique invariant distribution

$$\pi_n^\theta(x) = \frac{\binom{n}{nx} \prod_{a\in\Gamma} \theta_a^{nx_a} e^{-\eta^{-1}\bar{\phi}_n(x)}}{\sum_{y\in\Delta_n} \binom{n}{ny} \prod_{a\in\Gamma} \theta_a^{ny_a} e^{-\eta^{-1}\bar{\phi}_n(y)}}\,, \tag{A.1}$$

where, for every x in Δ_n,

$$\bar{\phi}_n(x) = \sum_{e\in\mathcal{E}} \sum_{j=0}^{n(Ax)_e} \left(\tau_e\left(\frac{j}{n}\right) + \tilde{\omega}_e\left(\frac{j}{n}\right)\right)\,.$$

Proof. First, from (Monderer and Shapley (1996)) it follows that since the game under consideration is a congestion game, it is a potential game, i.e., it exists a function $\phi_n : \Delta_n \to \mathbb{R}$ such that given x, y in Δ_n and $x = y + \frac{1}{n}(u_i - u_j)$ it holds that

$$\phi_n(x) - \phi_n(y) = c_i(x) - c_j(y)\,. \tag{A.2}$$

Moreover, it also defines the potential of the game as

$$\begin{aligned}
\phi_n(x) &= \sum_{e\in\mathcal{E}} \sum_{j=0}^{n(Ax)_e} c_e(j) = \sum_{e\in\mathcal{E}} \sum_{j=0}^{n(Ax)_e} \tau_e\Big(\frac{j}{n}\Big) + \omega_e^\theta\Big(\frac{j}{n}\Big)\\
&= \sum_{e\in\mathcal{E}} \sum_{j=0}^{n(Ax)_e} \Big(\tau_e\Big(\frac{j}{n}\Big) + \tilde{\omega}_e\Big(\frac{j}{n}\Big) + \bar{\omega}_e(\theta_e)\Big)\\
&= \sum_{e\in\mathcal{E}} \sum_{j=0}^{n(Ax)_e} \Big(\tau_e\Big(\frac{j}{n}\Big) + \tilde{\omega}_e\Big(\frac{j}{n}\Big) - \eta\log(\theta_e)\Big)\\
&= \sum_{e\in\mathcal{E}} \sum_{j=0}^{n(Ax)_e} \Big(\tau_e\Big(\frac{j}{n}\Big) + \tilde{\omega}_e\Big(\frac{j}{n}\Big)\Big) - n\eta \sum_{e\in\mathcal{E}} x_e \log(\theta_e)\\
&= \bar{\phi}_n(x) - n\eta \sum_{e\in\mathcal{E}} x_e \log(\theta_e)\,.
\end{aligned} \tag{A.3}$$

Then, consider two states x and y in Δ_n such that $x = y + \frac{1}{n}(u_i - u_j)$ and note that

$$\begin{aligned}
\frac{P_{yx}^{(n)}}{P_{xy}^{(n)}} &= \frac{y_j e^{-\eta^{-1} c_i^\theta(x)}}{x_i e^{-\eta^{-1} c_j^\theta(y)}} \frac{\sum_{a\in\Gamma} e^{-\eta^{-1}\left(c_a^\theta\left(x + \frac{1}{n}(u_a - u_i)\right)\right)}}{\sum_{a\in\Gamma} e^{-\eta^{-1}\left(c_a^\theta\left(y + \frac{1}{n}(u_a - u_j)\right)\right)}}\\
&= \frac{y_j}{x_i} \frac{e^{-\eta^{-1}\phi_n(x)}}{e^{-\eta^{-1}\phi_n(y)}} = \frac{y_j}{x_i} \frac{\prod_{a\in\Gamma} \theta_a^{nx_a} e^{-\eta^{-1}\bar{\phi}_n(x)}}{\prod_{a\in\Gamma} \theta_a^{ny_a} e^{-\eta^{-1}\bar{\phi}_n(y)}}\\
&= \frac{y_j}{x_i} \frac{\theta_i e^{-\eta^{-1}\bar{\phi}_n(x)}}{\theta_j e^{-\eta^{-1}\bar{\phi}_n(y)}}\,,
\end{aligned} \tag{A.4}$$

where the second equality follows from (A.2) and the fact that

$$\begin{aligned}
x + \frac{1}{n}(u_a - u_i) &= y + \frac{1}{n}(u_i - u_j) + \frac{1}{n}(u_a - u_i)\\
&= y + \frac{1}{n}(u_a - u_j)\,,
\end{aligned}$$

so that the sums in the numerator and the denominator simplify, the third one from the properties of logarithms, and the fourth follows from $x = y + \frac{1}{n}(u_i - u_j)$. We now prove that

$$\frac{\pi_n^\theta(x)}{\pi_n^\theta(y)} = \frac{P_{yx}^{(n)}}{P_{xy}^{(n)}}\,,$$

that is, the Markov chain is reversible with invariant distribution π_n^θ. This conditions is verified with

$$\pi_n^\theta(x) = \frac{\binom{n}{nx} \prod_{a\in\Gamma} \theta_a^{nx_a} e^{-\eta^{-1}\bar{\phi}_n(x)}}{\sum_{z\in\Delta_n} \binom{n}{nz} \prod_{a\in\Gamma} \theta_a^{nz_a} e^{-\eta^{-1}\bar{\phi}_n(z)}}\,.$$

This follows from the fact that for every x, y in Δ_n such that $x = y + \frac{1}{n}(u_i - u_j)$ it holds that

$$\begin{aligned}
\frac{\pi_n^\theta(x)}{\pi_n^\theta(y)} &= \frac{\binom{n}{nx} \prod_{a\in\Gamma} \theta_a^{nx_a} e^{-\eta^{-1}\bar{\phi}_n(x)}}{\binom{n}{ny} \prod_{a\in\Gamma} \theta_a^{ny_a} e^{-\eta^{-1}\bar{\phi}_n(y)}}\\
&= \frac{\frac{n!}{(nx_1)!,\ldots,(nx_{|\Gamma|})!} \prod_{a\in\Gamma} \theta_a^{nx_a} e^{-\eta^{-1}\bar{\phi}_n(x)}}{\frac{n!}{(ny_1)!,\ldots,(ny_{|\Gamma|})!} \prod_{a\in\Gamma} \theta_a^{ny_a} e^{-\eta^{-1}\bar{\phi}_n(y)}}\\
&= \frac{(ny_i)!(ny_j)!}{(nx_i)!(nx_j)!} \frac{\theta_i^{nx_i}\theta_j^{nx_j}}{\theta_i^{ny_i}\theta_j^{ny_j}} \frac{e^{-\eta^{-1}\bar{\phi}_n(x)}}{e^{-\eta^{-1}\bar{\phi}_n(y)}}\\
&= \frac{y_j}{x_i} \frac{\theta_i}{\theta_j} \frac{e^{-\eta^{-1}\bar{\phi}_n(x)}}{e^{-\eta^{-1}\bar{\phi}_n(y)}}\,,
\end{aligned}$$

where the third and fourth equalities follows from the fact that $x = y + \frac{1}{n}(u_i - u_j)$. Notice that this is exactly equal to (A.4), which concludes the proof. ∎

The next result shows that, as n grows, the invariant distribution π_n^θ concentrates around $\Psi(\theta)$. To this, let us introduce the ρ-neighborhood of $\Psi(\theta)$ by

$$U_{\rho,n}(\theta) = \Big\{x \in \Delta_n : ||x - \Psi(\theta)|| < \rho\Big\}\,. \tag{A.5}$$

and its complement set, i.e.,

$$U_{\rho,n}^c(\theta) = \Delta_n \setminus U_{\rho,n}(\theta)\,.$$

Moreover, for every set M we define the invariant distribution of M as

$$\pi_n^\theta(M) := \sum_{x\in M} \pi_n^\theta(x)\,.$$

Lemma 2. For every compact $K \subset \mathrm{int}(\Delta)$, for every $\rho > 0$, and for every $\epsilon > 0$, there exists $n_{\epsilon,K}$ such that

$$\sup_{\theta\in K} \pi_n^\theta(U_{\rho,n}^c(\theta)) < \epsilon$$

for every $n > n_{\epsilon,K}$.

Proof. Recall from Lemma 1 that the invariant distribution is in the form (A.1), and notice that it can be equivalently written as

$$\pi_n^\theta(x) = \frac{\exp(B_n^\theta(x))}{Z_n^\theta} ,$$

where

$$\begin{aligned} B_n^\theta(x) &= \sum_{k=1}^n \log k - \sum_{i=1}^m \sum_{k=1}^{nx_i} \log k + n \sum_{a\in\Gamma} x_a \log(\theta_a) + \\ &- \eta^{-1}\bar{\phi}_n(x) \end{aligned}$$

and

$$Z_n^\theta = \sum_{x\in\Delta_n} \exp\big(B_n^\theta(x)\big) .$$

Consider a sequence $(k_n)_n$ such that k_n/n belongs to Δ_n for every n, and $k_n/n \to x$, and note that

$$\begin{aligned} \lim_{n\to+\infty} \frac{1}{n}\bar{\phi}_n\Big(\frac{k_n}{n}\Big) &= \lim_{n\to+\infty} \frac{1}{n} \sum_{e\in\mathcal{E}} \Big(\sum_{j=0}^{(Ak_n)_e} \tau_e\Big(\frac{j}{n}\Big) + \tilde{\omega}_e\Big(\frac{j}{n}\Big)\Big) \\ &= \sum_{e\in\mathcal{E}} \int_0^{(Ax)_e} (\tau_e(s) + \tilde{\omega}_e(s))\, ds \\ &= \sum_{e\in\mathcal{E}} \int_0^{(Ax)_e} \Big(\frac{d}{ds}(s\tau_e(s))\Big)\, ds \\ &= \sum_{e\in\mathcal{E}} \int_0^{(Ax)_e} \Big(\frac{d}{ds}(s\tau_e(s))\Big)\, ds \\ &= \sum_{e\in\mathcal{E}} (Ax)_e \tau_e\left((Ax)_e\right) \\ &= \sum_{e\in\mathcal{E}} \sum_{\gamma\in\Gamma} A_{e\gamma} x_\gamma \tau_e\left((Ax)_e\right) \\ &= \sum_{\gamma\in\Gamma} x_\gamma \sum_{e\in\mathcal{E}} A_{e\gamma} \tau_e\left((Ax)_e\right) \\ &= \phi(x) . \end{aligned} \tag{A.6}$$

where the third equality follows from the fact that

$$\begin{aligned} \frac{d}{ds}(s\tau_e(s)) &= \tau_e(s) + s\tau_e'(s) \\ &= \tau_e(s) + \tilde{\omega}_e(s) . \end{aligned}$$

Moreover, notice that from Stirling formula it holds that

$$\begin{aligned} &\lim_{n\to+\infty} \frac{1}{n}\Big(\sum_{k=1}^n \log k - \sum_{a\in\Gamma}\sum_{k=1}^{nx_a} \log k + n \sum_{a\in\Gamma} x_a \log(\theta_a)\Big) \\ &= \lim_{n\to+\infty} \frac{1}{n}\Big(n\log(n) - n + o(n) - \sum_{a\in\Gamma} \big(nx_a \log(nx_a) - nx_a \\ &\quad + o(n)\big) + n \sum_{a\in\Gamma} x_a \log(\theta_a)\Big) \\ &= \lim_{n\to+\infty} \log(n) - 1 - \sum_{a\in\Gamma} x_a \log(n) - \sum_{a\in\Gamma} x_a \log(x_a) + \\ &+ \sum_{a\in\Gamma} x_a + \sum_{a\in\Gamma} x_a \log(\theta_a) \\ &= \lim_{n\to+\infty} \log(n) - 1 - \log(n) - \sum_{a\in\Gamma} x_a \log(x_a) + 1 + \\ &+ \sum_{a\in\Gamma} x_a \log(\theta_a) \\ &= \sum_{a\in\Gamma} x_a \Big(\log(\theta_a) - \log(x_a)\Big) = \sum_{a\in\Gamma} x_a \log\Big(\frac{\theta_a}{x_a}\Big) \\ &= -D_{KL}(x||\theta) . \end{aligned} \tag{A.7}$$

Hence, as n grows we have that

$$\frac{1}{n} B_n^\theta(x) \xrightarrow{n\to+\infty} -F^\theta(x) , \tag{A.8}$$

where $F^\theta(x) = \eta^{-1}\phi(x) + D_{KL}(x||\theta)$. Note that $F^\theta(x)$ is a strictly convex function, since both ϕ and D_{KL} are strictly convex.

We then aim to prove that this convergence is uniform for every x in Δ_n and for every θ in Δ. In particular, consider

$$\begin{aligned} &\Big|\frac{1}{n}B_n^\theta(x) - (-F^\theta(x))\Big| \le \\ &\Big|\frac{1}{n}\Big(\sum_{k=1}^n \log k - \sum_{a\in\Gamma}\sum_{k=1}^{nx_a} \log k + n\sum_{a\in\Gamma} x_a \log(\theta_a)\Big) + D_{KL}(x||\theta)\Big| \\ &+ \Big|\frac{\eta^{-1}}{n}\bar{\phi}_n(x) - \eta^{-1}\phi(x)\Big| \le \\ &\Big|\frac{1}{n}\Big(\sum_{k=1}^n \log k - \sum_{a\in\Gamma}\sum_{k=1}^{nx_a} \log k\Big) + \sum_{a\in\Gamma} x_a \log(x_a)\Big| \\ &+ \Big|\frac{\eta^{-1}}{n}\bar{\phi}_n(x) - \eta^{-1}\phi(x)\Big| . \end{aligned}$$

Notice that every dependence on θ has vanished, so the convergence is uniform for every θ in Δ. We are then left to ensure uniformness for every x in Δ_n. First, notice that since τ_e and $\tilde{\omega}_e$ are continuous on $[0, 1]$, they are also uniformly continuous and bounded so that

$$\sup_{\theta\in\Delta} \sup_{x\in\Delta_n} \Big|\frac{\eta^{-1}}{n}\bar{\phi}_n(x) - \eta^{-1}\phi(x)\Big| \xrightarrow{n\to+\infty} 0 .$$

Moreover, the convergence in (A.7) is uniform over $x \in \Delta_n$. Indeed, by (Cover and Thomas (1991)), for every $x \in \Delta_n$,

$$\frac{1}{(n+1)^m} e^{nH(x)} \le \frac{n!}{\prod_{a\in\Gamma}(nx_a)!} \le e^{nH(x)},$$

where $m = |\Gamma|$ and

$$H(x) = -\sum_{a\in\Gamma} x_a \log x_a.$$

Taking logarithms and dividing by n, we obtain

$$\begin{aligned} \sup_{x\in\Delta_n} &\left|\frac{1}{n}\left(\sum_{k=1}^n \log k - \sum_{a\in\Gamma}\sum_{k=1}^{nx_a} \log k\right) + \sum_{a\in\Gamma} x_a \log x_a\right| \\ &\le \frac{m\log(n+1)}{n}. \end{aligned}$$

Since $m\log(n+1)/n \to 0$, the Stirling-type convergence is uniform over Δ_n. Hence, (A.8) is uniform both in θ and in x.

Then, given a compact $K \subset \mathrm{int}(\Delta)$Since $\Psi(\theta)$ is the unique minimizer of $F^\theta(x)$, there exists $c_{\rho,K} > 0$ such that

$$F^\theta(x) \ge F^\theta(\Psi(\theta)) + c_{\rho,K} , \quad \forall x \in U_\rho^c(\theta) , \forall\theta \in K . \tag{A.9}$$

Then, Equation (A.8) implies that for large n it holds

$$-\frac{c_{\rho,K}}{4} \le \frac{1}{n}B_n^\theta(x) + F^\theta(x) \le \frac{c_{\rho,K}}{4} . \tag{A.10}$$

By considering the upper bound of (A.10) and applying Equation (A.9) we obtain

$$\begin{aligned} B_n^\theta(x) &\le -n\Big(F^\theta(\Psi(\theta)) + c_{\rho,K} - \frac{c_{\rho,K}}{4}\Big) \\ &= -n\Big(F^\theta(\Psi(\theta)) + \frac{3c_{\rho,K}}{4}\Big) \end{aligned}$$

which implies that

$$\sum_{x\in U_\rho^c(\theta)} e^{B_n^\theta(x)} \le |\Delta_n| e^{-n\left(F^\theta(\Psi(\theta))+\frac{3c_{\rho,K}}{4}\right)} \le (n+1)^m e^{-n\left(F^\theta(\Psi(\theta))+\frac{3c_{\rho,K}}{4}\right)} . \tag{A.11}$$

Then, consider a $x_n(\theta)$ in Δ_n such that

$$\sup_{\theta\in K} ||x_n(\theta)-\Psi(\theta)|| \to 0$$

and consider the lower bound of Equation (A.10), i.e.,

$$B_n^\theta(x_n(\theta)) \ge -n\Big(F^\theta(x_n(\theta)) + \frac{c_{\rho,K}}{4}\Big) .$$

By uniform continuity of F^θ and sufficiently large n it holds that

$$F^\theta(x_n(\theta)) - F^\theta(\Psi(\theta)) \le \frac{c_{\rho,K}}{4} ,$$

which implies that

$$\begin{aligned} B_n^\theta(x_n(\theta)) &\ge -n(F^\theta(\Psi(\theta)) + \frac{c_{\rho,K}}{4} + \frac{c_{\rho,K}}{4}) \\ &= -n(F^\theta(\Psi(\theta)) + \frac{c_{\rho,K}}{2}) . \end{aligned}$$

From this it follows that

$$Z_n^\theta \ge e^{B_n^\theta(x_n(\theta))} \ge e^{-n\left(F^\theta(\Psi(\theta))+\frac{c_{\rho,K}}{2}\right)} . \tag{A.12}$$

By combining together Equations (A.11) and (A.12) we achieve

$$\pi_n^\theta(U_\rho^c(\theta)) = \frac{\sum_{x\in U_\rho^c(\theta)} \exp(B_n^\theta(x))}{Z_n} \le (n+1)^m e^{-n\frac{c_{\rho,K}}{4}}$$

and since the bound has no dependence on θ the claim follows. ■

As a consequence of this result, the expected total travel time computed with respect to the invariant distribution π_n^θ converges uniformly in θ to the expected total travel time computed with respect to a delta function centered in $\Psi(\theta)$, as the next result shows.

Lemma 3. For every compact $K \subset \mathrm{int}(\Delta)$ and for every $\epsilon > 0$, there exists $\tilde{n}_{\epsilon,K}$ such that

$$\sup_{\theta\in K} |\mathbb{E}_{\pi_n^\theta}[\phi(X)] - \mathbb{E}_{\delta^{\Psi(\theta)}}[\phi(X)]| < \epsilon$$

for every $n > \tilde{n}_{\epsilon,K}$.

Proof. Since $\phi(x)$ is continuous on Δ, it holds that, for every $\rho > 0$,

$$\begin{aligned} &\sup_{\theta\in K} |\mathbb{E}_{\pi_n^\theta}[\phi(X)] - \mathbb{E}_{\delta^{\Psi(\theta)}}[\phi(X)]| = \\ &= \sup_{\theta\in K} \Big| \sum_{x\in\Delta_n} \pi_n^\theta(x)\big(\phi(x)-\phi(\Psi(\theta))\big)\Big| \\ &\le \sup_{\theta\in K} \Big(\sup_{x\in U_{\rho,n}(\theta)} |\phi(x)-\phi(\Psi(\theta))| + 2||\phi||_\infty \pi_n^\theta(U_{\rho,n}^c(\theta))\Big) . \end{aligned}$$

We now aim to upper bound the two terms in the external superior. Since ϕ is continuous, for every K and for every $\epsilon > 0$, there exists ρ_ϵ such that

$$\sup_{\theta\in K} \sup_{x\in U_{\rho_\epsilon,n}(\theta)} \Big\{ |\phi(x)-\phi(\Psi(\theta))\Big\} < \frac{\epsilon}{2} .$$

Moreover, Lemma 2 guarantees that

$$2||\phi||_\infty \sup_{\theta\in K} \pi_n^\theta(U_{\rho_\epsilon,n}^c(\theta)) < \frac{\epsilon}{2} ,$$

for every $n > n_{\epsilon,K}$. Summing the two contributions we obtain the statement. ■

We now prove that $\psi_n(\theta)$ converges uniformly to $\Psi(\theta)$.

Lemma 4. For every compact $K \subset \mathrm{int}(\Delta$ and for every $\epsilon > 0$, it exists $\bar{n}_{\epsilon,K}$ such that

$$\sup_{\theta\in K} \|\psi_n(\theta) - \Psi(\theta)\| < \epsilon \tag{A.13}$$

for every $n > \bar{n}_{\epsilon,K}$.

Proof. From Lemma 2, the invariant distributions π_n^θ concentrates around $\Psi(\theta)$ as n grows large. Namely, for every compact $K \subset \mathrm{int}(\Delta)$ and for every $\rho > 0$, we have

$$\lim_{n\to+\infty} \sup_{\theta\in K} \pi_n^\theta(U_{\rho,n}^c(\theta)) = 0 .$$

Let $D = \sup_{x,y\in\Delta} \|x-y\|_\infty \le 1$. Then,

$$\begin{aligned} \sup_{\theta\in K} \|\psi_n(\theta)-\Psi(\theta)\|_\infty &= \sup_{\theta\in K} \Big\| \sum_{x\in\Delta_n} (x-\Psi(\theta))\pi_n^\theta(x)\Big\|_\infty \\ &\le \sup_{\theta\in K} \sum_{x\in U_{\rho,n}(\theta)} \|x-\Psi(\theta)\|_\infty \pi_n^\theta(x) \\ &\quad + \sup_{\theta\in K} \sum_{x\in U_{\rho,n}^c(\theta)} \|x-\Psi(\theta)\|_\infty \pi_n^\theta(x) \\ &< \rho + D \sup_{\theta\in K} \pi_n^\theta(U_\rho^c(\theta)) , \end{aligned}$$

where the first step follows from the definition of $\psi_n(\theta)$ and the fact that $\sum_{x\in\Delta_n} \pi_n^\theta(x) = 1$. The second step follows by splitting the sum into a sum of the elements inside and outside the set $U_{\rho,n}(\theta)$, respectively. Lastly, the third steps follows from the definition of $U_{\rho,n}(\theta)$ (A.5) and the worst-case bound D. From Lemma 2, and since $\rho > 0$ is arbitrary, the claim follows. ■

The next result proves that, if the k-th iteration of the map Ψ for an initial condition θ^0 is arbitrarily close to θ^*, then also the k-th iteration of the map ψ_n is arbitrarily close to θ^* for sufficiently large n.

Lemma 5. Consider $k \in \mathbb{N}$ and $\theta^0 \in \mathrm{int}(\Delta)$ such that

$$\|\Psi^k(\theta^0) - \theta^*\| < \varepsilon ,$$

for some $\varepsilon > 0$. Then there exists $n_{\varepsilon,k} \in \mathbb{N}$ such that, for every $n > n_{\varepsilon,k}$,

$$\|\psi_n^k(\theta^0) - \theta^*\| < 2\varepsilon.$$

Proof. Let

$$\theta^r := \Psi^r(\theta^0), \qquad r = 0,\dots,k.$$

Since the finite set $\{\theta^0,\dots,\theta^k\}$ is contained in $\mathrm{int}(\Delta)$, there exists $\delta > 0$ such that

$$K = \bigcup_{r=0}^k \overline{B_\delta(\theta^r)}$$

is compact and contained in $\mathrm{int}(\Delta)$.

By Lemma 4, we have

$$\sup_{\theta\in K} \|\psi_n(\theta) - \Psi(\theta)\| \xrightarrow{n\to+\infty} 0.$$

We now prove by induction that, for every fixed $r = 0,\dots,k$,

$$\psi_n^r(\theta^0) \xrightarrow{n\to+\infty} \Psi^r(\theta^0).$$

The claim is trivial for $r = 0$. Suppose it holds for some $r < k$. Then, for large enough n, $\psi_n^r(\theta^0) \in K$, since $\theta^k = \Psi^k(\theta^0)$ belongs to K. Hence, by the triangle inequality

$$\begin{aligned}\|\psi_n^{r+1}(\theta^0) - \Psi^{r+1}(\theta^0)\| &= \|\psi_n(\psi_n^r(\theta^0)) - \Psi(\Psi^r(\theta^0))\| \\ &\le \|\psi_n(\psi_n^r(\theta^0)) - \Psi(\psi_n^r(\theta^0))\| + \\ &+ \|\Psi(\psi_n^r(\theta^0)) - \Psi(\Psi^r(\theta^0))\| .\end{aligned}$$

The first term converges to zero by the uniform convergence of ψ_n to Ψ on K, while the second term converges to zero by the induction hypothesis and the continuity of Ψ. Therefore,

$$\psi_n^{r+1}(\theta^0) \xrightarrow{n\to+\infty} \Psi^{r+1}(\theta^0).$$

By induction, this holds in particular for $r = k-1$, namely

$$\psi_n^{k}(\theta^0) \xrightarrow{n\to+\infty} \Psi^{k}(\theta^0).$$

Therefore, there exists $n_{\varepsilon,k} \in \mathbb{N}$ such that, for every $n > n_{\varepsilon,k}$,

$$\|\psi_n^k(\theta^0) - \Psi^k(\theta^0)\| < \varepsilon.$$

Using the triangle inequality and combining this inequality with $\|\Psi^k(\theta^0) - \theta^*\| < \varepsilon$, we obtain

$$\begin{aligned}\|\psi_n^k(\theta^0) - \theta^*\| &\le \|\psi_n^k(\theta^0) - \Psi^k(\theta^0)\| + \|\Psi^k(\theta^0) - \theta^*\| \\ &< 2\varepsilon.\end{aligned}$$

This concludes the proof. ■

Appendix B. PROOF OF PROPOSITION 1

An internal fixed point θ^* of $\Psi(\theta)$ must inevitably satisfy

$$\theta^* = \arg\min_{x\in\Delta} \left\{\eta^{-1}\phi(x) + D_{KL}(x||\theta^*)\right\} .$$

Hence, the first-order optimality conditions imply that

$$\eta^{-1}\nabla\phi(\theta^*) + \nabla D_{KL}(\theta^*||\theta^*) = c\mathbb{1} ,$$

where the right-hand side $c\mathbb{1}$ follows from the constraint that x lives in Δ. Since $\nabla D_{KL}(\theta^*||\theta^*) = \mathbb{1}$, we are left with

$$\eta^{-1}\nabla\phi(\theta^*) = (c-1)\mathbb{1} ,$$

which implies that the fixed point θ^* of $\Psi(\theta)$ is an internal stationary point of ϕ, and is therefore its unique minimum since ϕ is strictly convex. The existence and uniqueness of an internal fixed point follows from the fact that x^* is stationary for $\phi(x)$, from the assumption that x^* is internal, and from strict convexity of ϕ.

$$\theta^{k+1} = \arg\min_{x\in\Delta}\{\eta^{-1}\phi(x) + D_{KL}(x||\theta^k)\} ,$$

and since the derivative of $D_{KL}(x\|\theta)$ diverges as x approaches the boundary of Δ, the minimum θ^{k+1} is internal. Hence, the first-order optimality conditions imply that, for every z in Δ,

$$\langle \eta^{-1}\nabla\phi(\theta^{k+1}) + \nabla D_{KL}(\theta^{k+1}||\theta^k), z - \theta^{k+1}\rangle = 0 . \quad \text{(B.1)}$$

Since ϕ is convex, it holds that

$$\phi(\theta^{k+1}) - \phi(z) \le \langle \nabla\phi(\theta^{k+1}), \theta^{k+1} - z\rangle . \quad \text{(B.2)}$$

By combining (B.1) and (B.2) we achieve

$$\eta^{-1}(\phi(\theta^{k+1}) - \phi(z)) \le \langle \nabla D_{KL}(\theta^{k+1}||\theta^k), z - \theta^{k+1}\rangle . \quad \text{(B.3)}$$

We can expand the right-hand side as

$$\begin{aligned}\langle \nabla D_{KL}(\theta^{k+1}||\theta^k), z - \theta^{k+1}\rangle &= \sum_{\gamma\in\Gamma}(z_\gamma - \theta_\gamma^{k+1})\log\big(\frac{\theta_\gamma^{k+1}}{\theta_\gamma^k}\big) + \\ &+ \sum_{\gamma\in\Gamma}(z_\gamma - \theta_\gamma^{k+1}) \\ &= \sum_{\gamma\in\Gamma}(z_\gamma - \theta_\gamma^{k+1})\big(\log\big(\frac{\theta_\gamma^{k+1}}{\theta_\gamma^k}\big)\big)\end{aligned}$$

Notice that

$$\begin{aligned}&D_{KL}(z||\theta^k) - D_{KL}(z||\theta^{k+1}) - D_{KL}(\theta^{k+1}||\theta^k) \\ &= \sum_{\gamma\in\Gamma} z_\gamma \log\left(\frac{z_\gamma}{\theta_\gamma^k}\right) - \sum_{\gamma\in\Gamma} z_\gamma \log\left(\frac{z_\gamma}{\theta_\gamma^{k+1}}\right) + \\ &- \sum_{\gamma\in\Gamma}\theta_\gamma^{k+1}\log\left(\frac{\theta_\gamma^{k+1}}{\theta_\gamma^k}\right) \\ &= \sum_{\gamma\in\Gamma} z_\gamma \log\left(\frac{\theta_\gamma^{k+1}}{\theta_\gamma^k}\right) - \sum_{\gamma\in\Gamma}\theta_\gamma^{k+1}\log\left(\frac{\theta_\gamma^{k+1}}{\theta_\gamma^k}\right) \\ &= \sum_{\gamma\in\Gamma}(z_\gamma - \theta_\gamma^{k+1})\log\left(\frac{\theta_\gamma^{k+1}}{\theta_\gamma^k}\right) .\end{aligned}$$

Therefore, we can rewrite Equation (B.3) as

$$\begin{aligned}\eta^{-1}(\phi(\theta^{k+1}) - \phi(z)) \le\ & D_{KL}(z||\theta^k) - D_{KL}(z||\theta^{k+1}) \\ & - D_{KL}(\theta^{k+1}||\theta^k) .\end{aligned} \quad \text{(B.4)}$$

Now let us fix $z = x^*$. This implies by the strict convexity of ϕ that $\phi(\theta^{k+1}) - \phi(x^*) \ge 0$ and inequality is strict if and only if $\theta^{k+1} = x^*$. Hence, we can further rewrite (B.4) as

$$\eta^{-1}(\phi(\theta^{k+1}) - \phi(x^*)) + D_{KL}(\theta^{k+1}||\theta^k) \le V^k - V^{k+1} ,$$

where $V^k = D_{KL}(x^*\|\theta^k)$. Since the left-hand side of such equation is non-negative, it follows that

$$V^k - V^{k+1} \ge 0 \Rightarrow V^{k+1} \le V^k ,$$

and the inequality is strict if and only if $\theta^{k+1} = \theta^k = x^*$. Hence, $D_{KL}(x^*\|\theta^k)$ is a Lyapunov function and LaSalle principle for discrete-time dynamical systems (Mei and Bullo (2017)) ensures convergence to the largest invariant set in Δ. We now aim to exclude convergence to fixed points on the boundary of Δ. If we sum for consecutive step we achieve

$$\begin{aligned}&\sum_{k=0}^{K}\Big(\eta^{-1}(\phi(\theta^{k+1}) - \phi(x^*)) + D_{KL}(\theta^{k+1}||\theta^k)\Big) \\ \le\ &\sum_{k=0}^{K}\Big(V^k - V^{k+1}\Big) ,\end{aligned}$$

or equivalently

$$\sum_{k=0}^{K}\Big(\eta^{-1}(\phi(\theta^{k+1}) - \phi(x^*)) + D_{KL}(\theta^{k+1}||\theta^k)\Big) \le V^0 - V^{K+1} . \quad \text{(B.5)}$$

Clearly, $V^0 - V^{K+1} \le V^0$ and moreover given x^* in $\mathrm{int}(\Delta)$ and θ^0 in $\mathrm{int}(\Delta)$ we also have that $V^0 < +\infty$. Hence, we can rewrite Equation (B.5) as

$$\sum_{k=0}^{K}\Big(\eta^{-1}(\phi(\theta^{k+1}) - \phi(x^*))\Big) \le V^0 ,$$

and if we consider the infinite sum we get

$$\sum_{k=0}^{+\infty}\Big(\eta^{-1}(\phi(\theta^{k+1}) - \phi(x^*))\Big) \le V^0 . \quad \text{(B.6)}$$

Since the left-hand side of Equation (B.6) is the infinite sum of non-negative terms and $V^0 < +\infty$, we conclude

$$\phi(\theta^{k+1}) - \phi(x^*)) \to 0 , \quad k \to +\infty ,$$

which implies that

$$\phi(\theta^k) \to \phi(x^*) , \quad k \to +\infty .$$

Finally, since $\phi(x^*)$ is the unique minimum of ϕ it holds that

$$\theta^k \to x^* \,, \quad k \to +\infty \,.$$

$\blacksquare$.

Appendix C. PROOF OF THEOREM 1

First, note that, for parallel links,

$$C(X) = \sum_{\gamma\in\Gamma} X_\gamma A_{e\gamma}\tau_e\left((AX)_e\right) = \sum_{e\in\mathcal{E}} X_e\tau_e(X_e)\,.$$

Let

$$\theta^r := \Psi^r(\theta^0), \qquad r = 0,\ldots,k.$$

Since the finite set $\{\theta^0,\ldots,\theta^k\}$ is contained in $\mathrm{int}(\Delta)$, there exists $\delta > 0$ such that

$$K(\theta^0) = \bigcup_{r=0}^{k} \overline{B_\delta(\theta^r)}$$

is compact and contained in $\mathrm{int}(\Delta)$. By the proof of Lemma 5 we know that for large enough n, the set $\{\psi_n^0(\theta^0),\ldots,\psi_n^k(\theta^0)\}$ is in $K(\theta^0)$.

Using the fact that $\psi_n^k(\theta^0)$ belongs to $K(\theta^0)$, it follows from Lemma 3 that, for every $\epsilon > 0$, there exists a finite $\tilde{n}_{\epsilon,\theta^0} := \tilde{n}_{\epsilon,K(\theta^0)}$ such that

$$\left|\mathbb{E}_{\pi_n^{\psi_n^k(\theta^0)}}[\phi(X)] - \mathbb{E}_{\delta^{\Psi(\psi_n^k(\theta^0))}}[\phi(X)]\right| < \frac{\epsilon}{2}\,, \tag{C.1}$$

for every $n > \tilde{n}_{\epsilon,\theta^0}$ and θ in $K(\theta^0)$. From continuity of $\phi(x)$, we have that for every $\epsilon > 0$, there exists σ_ϵ such that

$$||x-y|| < \sigma_\epsilon \implies |\phi(x)-\phi(y)| < \epsilon/2\,. \tag{C.2}$$

Since $\Psi(\theta)$ is the minimizer of a continuous and strictly convex function, it is continuous. Therefore, for every $\sigma_\epsilon > 0$, there exists $\bar{\sigma}_\epsilon > 0$ such that

$$||\theta-\theta^*|| < \bar{\sigma}_\epsilon \implies ||\Psi(\theta)-\Psi(\theta^*)|| < \sigma_\epsilon\,. \tag{C.3}$$

From Proposition 1, it follows that there exists k_ϵ such that

$$||\Psi^k(\theta^0)-\theta^*|| < \bar{\sigma}_\epsilon/2\,, \quad \forall k > k_\epsilon\,. \tag{C.4}$$

Using (C.4) and Lemma 5, for every $k > k_\epsilon$ there exists $n_{\epsilon,k}$ in $\mathbb{N}$ such that

$$||\psi_n^k(\theta^0)-\theta^*|| < \bar{\sigma}_\epsilon\,, \quad \forall n > n_{\epsilon,k}\,. \tag{C.5}$$

We can then use (C.3) with $\theta = \psi_n^k(\theta^0)$ and get that

$$||\Psi(\psi_n^k(\theta^0))-\theta^*|| = ||\Psi(\psi_n^k(\theta^0))-\Psi(\theta^*)|| < \sigma_\epsilon\,. \tag{C.6}$$

By (C.2), we then get that

$$\begin{aligned} &|\mathbb{E}_{\delta^{\Psi(\psi_n^k(\theta^0))}}[\phi(X)] - \mathbb{E}_{\delta^{\theta^*}}[\phi(X)]| \\ &= \phi(\Psi(\psi_n^k(\theta^0))) - \phi(\theta^*) < \frac{\epsilon}{2}\,, \end{aligned} \tag{C.7}$$

for every $n > n_{\epsilon,k}$ Using (C.1), (C.7) and the triangle inequality, we get that, for every $k > k_\epsilon$,

$$\left|\mathbb{E}_{\pi_n^{\psi_n^k(\theta^0)}}[\phi(X)] - \mathbb{E}_{\delta^{\theta^*}}[\phi(X)]\right| < \epsilon, \tag{C.8}$$

for every $n > \max(n_{\epsilon,k},\tilde{n}_{\epsilon,\theta^0})$. Recall now that $\theta^* = x^*$, where

$$x^* = \arg\min_{x\in\Delta} \sum_{e\in\mathcal{E}} \phi(x)\,.$$

and that $\theta^k = \psi_n^k(\theta^0)$. Hence, from (C.8)

$$\begin{aligned} &\mathbb{E}_{\pi_n^{\theta^k}}[\phi(X)] - \inf_{\pi\in\Delta_n} \mathbb{E}_\pi[\phi(X)] \\ &< \mathbb{E}_{\pi_n^{\theta^k}}[\phi(X)] - \mathbb{E}_{\delta^{\theta^*}}[\phi(X)] < \epsilon\,. \end{aligned}$$

This concludes the proof. $\blacksquare$