E-companion to the paper titled "A Day-to-Day Dynamical Approach to the Most Likely User Equilibrium Problem"

Appendix A Explanation of the Entropy Function

In Section 2.1, we have defined the negative entropy of a route choice strategy $\boldsymbol{p} \in \mathscr{P}$ as

$$\phi(\boldsymbol{p}) = \langle \operatorname{diag}(\boldsymbol{q})\boldsymbol{p}, \log(\boldsymbol{p})\rangle.$$

To explain how this measures "the number of states" (i.e., the different ways travelers can be arranged to produce the route flow corresponding to $\boldsymbol{p}$), suppose that the flow carried by a single traveler is $\tau > 0$ (τ is a small constant). Hence, the number of travelers traveling between each OD pair $w \in \mathscr{W}$ and selecting each route $k \in \mathscr{K}_w$ would be $n_w = q_w/\tau$ and $m_k = f_k/\tau$, respectively. Applying the basic counting principle, the total number of states, after taking the logarithm, reads

$$\begin{aligned}
\log\Big(\prod_{w\in\mathscr{W}} \frac{n_w!}{\prod_{k\in\mathscr{K}_w} m_k!}\Big) &= \sum_{w\in\mathscr{W}} \big(\log(n_w!) - \sum_{k\in\mathscr{K}_w} \log(m_k!)\big) \\
&= \sum_{w\in\mathscr{W}} \big(n_w \log(n_w) - n_w + \mathscr{O}(\log(n_w)) - \sum_{k\in\mathscr{K}_w} m_k \log(m_k) - m_k + \mathscr{O}(\log(m_k))\big), \\
&= \sum_{w\in\mathscr{W}} \big(n_w \log(n_w) + \mathscr{O}(\log(n_w)) - \sum_{k\in\mathscr{K}_w} m_k \log(m_k) + \mathscr{O}(\log(m_k))\big),
\end{aligned}$$

where Stirling's formula gives the second equality, while the relation $n_w = \sum_{k\in\mathscr{K}_w} m_k$ gives the third one. When τ is sufficiently small (as close to the nonatomic setting), we would have $n_w \to \infty$ and $m_k \to \infty$. Hence, the term $\mathscr{O}(\log(n_w))$ and $\mathscr{O}(\log(m_k))$ would become eligible relative to $n_w \log(n_w)$ and $m_k \log(m_k)$. Further noting that

$$\sum_{w\in\mathscr{W}} \big(n_w \log(n_w) - \sum_{k\in\mathscr{K}_w} m_k \log(m_k)\big) = \sum_{w\in\mathscr{W}} \sum_{k\in\mathscr{K}_w} m_k \big(\log(n_w) - \log(m_k)\big) = -\frac{1}{\tau} \cdot \sum_{w\in\mathscr{W}} q_w \sum_{k\in\mathscr{K}_w} p_k \log(p_k),$$

we have

$$\log\Big(\prod_{w\in\mathscr{W}} \frac{n_w!}{\prod_{k\in\mathscr{K}_w} m_k!}\Big) \approx -\frac{1}{\tau} \cdot \sum_{w\in\mathscr{W}} q_w \sum_{k\in\mathscr{K}_w} p_k \log(p_k) = -\frac{1}{\tau} \cdot \langle \operatorname{diag}(\boldsymbol{q})\boldsymbol{p}, \log(\boldsymbol{p})\rangle,$$

when τ is sufficiently small. Dropping τ (which is a constant) then gives rise to the entropy function for evaluating the likelihood of $\boldsymbol{p}$ in our setting.

Appendix B Proofs in Section 3

B.1 Proof of Lemma 1

Proof. We first note that for all $\boldsymbol{s}^t \in \mathbb{R}^{|\mathscr{K}|}$, the corresponding logit choice can be written in the vector form as $q_r(\boldsymbol{s}^t) = \boldsymbol{y}^t / \boldsymbol{\Sigma}^\top \boldsymbol{\Sigma} \boldsymbol{y}^t$, where $\boldsymbol{y}^t = \exp(-r \cdot \boldsymbol{s}^t)$. As each column of $\boldsymbol{\Sigma}$ is a standard unit vector, we have $\log(\boldsymbol{\Sigma}^\top \boldsymbol{\Sigma} \boldsymbol{y}^t) = \boldsymbol{\Sigma}^\top \log(\boldsymbol{\Sigma} \boldsymbol{y}^t)$. Denoting $\boldsymbol{x}^t = \bar{\boldsymbol{\Lambda}} \boldsymbol{p}^t$ for all $t \geq 0$, we can show

$$\begin{aligned}
\langle \boldsymbol{e}, \log(\boldsymbol{p}^t)\rangle &= \langle \boldsymbol{e}, \log(\boldsymbol{y}^t) - \log(\boldsymbol{\Sigma}^\top \boldsymbol{\Sigma} \boldsymbol{y}^t)\rangle = \langle \boldsymbol{e}, -r \cdot \boldsymbol{s}^t - \boldsymbol{\Sigma}^\top \log(\boldsymbol{\Sigma} \boldsymbol{y}^t)\rangle \\
&= \langle \boldsymbol{e}, -r \cdot \boldsymbol{s}^t\rangle = -r \cdot \sum_{i=0}^{t-1} \eta^i \cdot \langle \boldsymbol{e}, \boldsymbol{\Lambda}^\top \boldsymbol{x}^i\rangle + \langle \boldsymbol{e}, -r \cdot \boldsymbol{s}^0\rangle = -r \cdot \langle \boldsymbol{e}, \boldsymbol{s}^0\rangle.
\end{aligned} \tag{31}$$

The first and the second equalities hold due to the earlier discussions, the third and fifth equalities hold because $\boldsymbol{e} \in \ker(\boldsymbol{\Sigma})$ and $\boldsymbol{e} \in \ker(\boldsymbol{\Lambda})$, respectively, and the fourth equality is obtained by applying Equation (7). □

B.2 Proof of Lemma 2

Proof. Per Proposition 2, the KL projection problem (10) can be written as

$$\begin{aligned} \min_{\boldsymbol{p}^* \geq 0} \;& \langle \operatorname{diag}(\boldsymbol{q})\boldsymbol{p}^*, \log(\boldsymbol{p}^*) - \log(\boldsymbol{p}^0) \rangle. \\ \text{s.t. } \;& \boldsymbol{\Sigma}\operatorname{diag}(\boldsymbol{q})\boldsymbol{p}^* = \boldsymbol{d}, \quad \boldsymbol{\Lambda}\operatorname{diag}(\boldsymbol{q})\boldsymbol{p}^* = \boldsymbol{x}^*, \end{aligned} \tag{32}$$

which is evidently a convex program. Hence, any $\bar{\boldsymbol{p}}^* \in \mathscr{P}^*$ solves (32) if and only if there exist $\boldsymbol{\alpha} \in \mathbb{R}^{|\mathscr{K}|}$ and $\boldsymbol{\beta} \in \mathbb{R}^{|\mathscr{A}|}$ such that

$$\begin{cases} \bar{\boldsymbol{p}}^* \geq 0, \quad \log(\bar{\boldsymbol{p}}^*) - \log(\boldsymbol{p}^0) - \boldsymbol{\Sigma}^\mathsf{T}\boldsymbol{\alpha} - \boldsymbol{\Lambda}^\mathsf{T}\boldsymbol{\beta} \geq 0, \\ \langle \bar{\boldsymbol{p}}^*, \log(\bar{\boldsymbol{p}}^*) - \log(\boldsymbol{p}^0) - \boldsymbol{\Sigma}^\mathsf{T}\boldsymbol{\alpha} - \boldsymbol{\Lambda}^\mathsf{T}\boldsymbol{\beta} \rangle = 0. \end{cases} \tag{33}$$

If $\langle \boldsymbol{e}, \log(\bar{\boldsymbol{p}}^*) - \log(\boldsymbol{p}^0) \rangle = 0$ for all $\boldsymbol{e} \in \ker(\boldsymbol{\Sigma}) \cap \ker(\boldsymbol{\Lambda})$, then $\log(\bar{\boldsymbol{p}}^*) - \log(\boldsymbol{p}^0) \in \operatorname{im}(\boldsymbol{\Sigma}^\mathsf{T}) \cup \operatorname{im}(\boldsymbol{\Lambda}^\mathsf{T})$[1], which means one can always find $\boldsymbol{\alpha} \in \mathbb{R}^{|\mathscr{K}|}$ and $\boldsymbol{\beta} \in \mathbb{R}^{|\mathscr{A}|}$ such that

$$\log(\bar{\boldsymbol{p}}^*) - \log(\boldsymbol{p}^0) = \boldsymbol{\Sigma}^\mathsf{T}\boldsymbol{\alpha} + \boldsymbol{\Lambda}^\mathsf{T}\boldsymbol{\beta}.$$

Thus, Condition (33) must be satisfied. □

B.3 Proof of Theorem 1

Proof. Per Lemma 1, for all $\boldsymbol{e} \in \ker(\boldsymbol{\Sigma}) \cap \ker(\boldsymbol{\Lambda})$, the value of $\langle \boldsymbol{e}, \log(\boldsymbol{p}^t) \rangle$ is the same for all $t \geq 0$ given $\boldsymbol{p}^0$. This observation leads to

$$\langle \boldsymbol{e}, \log(\boldsymbol{p}^t) - \log(\boldsymbol{p}^0) \rangle = 0. \tag{34}$$

Noting that the function $\langle \boldsymbol{e}, \log(\boldsymbol{p}) - \log(\boldsymbol{p}^0) \rangle$ is continuous in $\boldsymbol{p}$, we obtain $\langle \boldsymbol{e}, \log(\bar{\boldsymbol{p}}^*) - \log(\boldsymbol{p}^0) \rangle = 0$ by letting $t \to \infty$ in Equation (34). This implies $\bar{\boldsymbol{p}}^*$ is the KL projection of $\boldsymbol{p}^0$ on $\mathscr{P}^*$ according to Lemma 2. □

B.4 Proof of Corollary 3

Proof. We first note that $\operatorname{supp}(\bar{\boldsymbol{p}}^*) \subseteq \mathscr{K}^*$ because $\bar{\boldsymbol{p}}^* \in \mathscr{P}^*$. Suppose $\bar{\boldsymbol{p}}^*$ is the limiting point of the CumLog model but there exists $k \in \mathscr{K}^*$ such that $k \notin \operatorname{supp}(\bar{\boldsymbol{p}}^*)$, then we must have $\bar{p}_k^* = 0$. Construct a function $l(\varepsilon) = D((1-\varepsilon)\cdot\bar{\boldsymbol{p}}^* + \varepsilon\cdot\boldsymbol{p}^0, \boldsymbol{p}^0)$, where $D(\cdot,\cdot)$ is the KL divergence defined in (9). The reader can verify that the derivative of $l(\cdot)$ at $\varepsilon = 0$ is $-\infty$. This means moving from $\bar{\boldsymbol{p}}^*$ toward $\boldsymbol{p}^0$ can reduce the KL divergence. Hence, $\bar{\boldsymbol{p}}^*$ cannot be the solution to the KL projection problem (10), or the limiting point of the CumLog model, a contradiction. □

[1]For any matrix $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, we have $\ker(\boldsymbol{A})^\perp = \operatorname{im}(\boldsymbol{A}^\mathsf{T})$, i.e., the perpendicular complement of $\ker(\boldsymbol{A})$ is $\operatorname{im}(\boldsymbol{A}^\mathsf{T})$.

B.5 Proof of Proposition 7

Proof. If $\boldsymbol{s}^0 = \boldsymbol{\Lambda}^\mathsf{T}\boldsymbol{v}^0$, then for any $\boldsymbol{e} \in \ker(\boldsymbol{\Sigma}) \cap \ker(\boldsymbol{\Lambda})$, we have $\langle \boldsymbol{e}, \boldsymbol{s}^0 \rangle = \langle \boldsymbol{e}, \boldsymbol{\Lambda}^\mathsf{T}\boldsymbol{v}^0 \rangle = \langle \boldsymbol{\Lambda}\boldsymbol{e}, \boldsymbol{v}^0 \rangle = 0$. Thus, Lemma 1 guarantees $\langle \boldsymbol{e}, \log(\boldsymbol{p}^t) \rangle = -r \cdot \langle \boldsymbol{e}, \boldsymbol{s}^0 \rangle = 0$ for any $\boldsymbol{e}$ and $t \geq 0$. As the function $\langle \boldsymbol{e}, \log(\boldsymbol{p}) \rangle$ is continuous in $\boldsymbol{p}$, letting $t \to \infty$ leads to $\langle \boldsymbol{e}, \log(\bar{\boldsymbol{p}}^*) \rangle = 0$ for any $\boldsymbol{e}$, which implies $\bar{\boldsymbol{p}}^*$ is the MEUE strategy according to Proposition 4. □

Appendix C Proofs in Section 4

C.1 Proof of Proposition 8

Proof. First, as there are a finite number of acyclic paths in a network, the discovery process must stop adding new routes after finite days (note that cyclic paths can never be a shortest route as long as the link cost is strictly positive). That is, there must exist $T_1 < \infty$ and $\overline{\mathscr{K}}_+ \subseteq \mathscr{K}$ such that $\mathscr{K}_+^t = \overline{\mathscr{K}}_+$ for all $t \geq T_1$. Starting from $t = T_1$, Algorithm 1 reduces to the original CumLog model without route exploration, applied to solving a "reduced" routing game in which only routes in $\overline{\mathscr{K}}_+$ are available. Denote the route-link incidence matrix corresponding to $\overline{\mathscr{K}}_+$ as $\overline{\boldsymbol{\Lambda}}_+^t$ and define $c_+ : \overline{\mathscr{P}}_+ \to \mathbb{R}^{|\overline{\mathscr{K}}_+|}$ as a map that satisfies

$$c_+(\boldsymbol{p}_+) = \overline{\boldsymbol{\Lambda}}_+^\mathsf{T} u(\boldsymbol{x}), \quad \text{where } \boldsymbol{x} = \overline{\boldsymbol{\Lambda}}_+ \text{diag}(\overline{\boldsymbol{\Sigma}}_+^\mathsf{T}\boldsymbol{d})\boldsymbol{p}_+. \tag{35}$$

By Proposition 5, as long as $r < 1/2\overline{L}$ for some $L \geq \max_{\boldsymbol{p}_+ \in \overline{\mathscr{P}}_+} \|\nabla c_+(\boldsymbol{p}_+)\|_2$, the route choice strategy $\boldsymbol{p}_+^t$ must converge to a fixed point $\overline{\boldsymbol{p}}_+ \in \overline{\mathscr{P}}_+$, which is a UE of the reduced routing game.

We then claim $\overline{\boldsymbol{p}} = [\overline{\boldsymbol{p}}_+; \boldsymbol{0}] \in \mathscr{P}^*$. To simplify the proof, let us assume $|\mathscr{W}| = 1$ without loss of generality; otherwise, we can simply pick one $w \in \mathscr{W}$ to raise the following conflict. Suppose that $\overline{\boldsymbol{p}} \notin \mathscr{P}^*$, then given any $k_0 \in \arg\min_{k \in \mathscr{K}} c_k(\overline{\boldsymbol{p}})$, we must have $c_0 := c_k(\overline{\boldsymbol{p}}) < c_{\min} := \min_{k \in \overline{\mathscr{K}}_+} c_k(\overline{\boldsymbol{p}})$. Hence, $k_0 \in \mathscr{K} \setminus \overline{\mathscr{K}}_+$, i.e., there exist some $k_0 \in \mathscr{K} \setminus \overline{\mathscr{K}}_+$ strictly better than any routes in $\overline{\mathscr{K}}_+$. For notational simplicity, let us define $\boldsymbol{p}^t = [\boldsymbol{p}_+^t; \boldsymbol{0}]$ for all $t \geq T_1$. Since $\boldsymbol{p}_+^t \to \overline{\boldsymbol{p}}_+ \Rightarrow \boldsymbol{p}^t \to \overline{\boldsymbol{p}} \Rightarrow c(\boldsymbol{p}^t) \to c(\overline{\boldsymbol{p}})$ and $\min_{k \in \overline{\mathscr{K}}_+} c(\boldsymbol{p}^t) \to c_{\min}$, there must exist $T_2 > T_1$ such that whenever $t \geq T_2$, we have

$$c_{k_0}(\boldsymbol{p}^t) < c_0 + \delta/3 \quad \text{and} \quad \min_{k \in \overline{\mathscr{K}}_+} c_k(\boldsymbol{p}^t) > c_{\min} - \delta/3, \tag{36}$$

where $\delta = c_{\min} - c_0$. This means on day T_2, route k_0 is strictly better than all routes in $\overline{\mathscr{K}}_+$. Hence, the route discovery process has not stabilized at T_2, which contradicts with the assumption that $\mathscr{K}_+^t$ remains unchanged after $t \geq T_1$. □

C.2 Proof of Proposition 9

Proof. Let $\mathscr{K}_+^t = \overline{\mathscr{K}}_+$ for all $t \geq T_1$ for some $T_1 > 0$. Starting from T_1, Algorithm 2 is reduced to the original CumLog model applied to solving a routing game defined on $\overline{\mathscr{K}}_+$. By viewing $t = T_1$ as an initial point of Algorithm 2, Condition (B) given in Proposition 7 is satisfied. Hence, $\overline{\boldsymbol{p}}_+ = \lim_{t \to \infty} \boldsymbol{p}_+^t$ must be the MEUE strategy of the reduced problem.

We proceed to prove $\overline{\boldsymbol{p}} = [\overline{\boldsymbol{p}}_+; \boldsymbol{0}]$ is the MEUE strategy of the original routing game as long as $\overline{\mathscr{K}}_+ \supseteq \cup_{\boldsymbol{p}^*} \text{supp}(\boldsymbol{p}^*)$. We first define $\phi_+ : \overline{\mathscr{P}}_+ \to \mathbb{R}$ as the negative entropy function of the reduced routing game, which reads

$$\phi_+(\boldsymbol{p}_+) = \langle \text{diag}(\overline{\boldsymbol{q}}_+)\boldsymbol{p}_+, \log(\boldsymbol{p}_+) \rangle, \quad \text{where } \overline{\boldsymbol{q}}_+ = (\overline{\boldsymbol{\Sigma}}_+)^\mathsf{T}\boldsymbol{d}. \tag{37}$$

Denoting $\overline{\mathscr{P}}^*_+ \subseteq \overline{\mathscr{P}}_+$ as the set of UE strategies for the reduced routing game, we claim

$$\min_{\boldsymbol{p}^*_+ \in \overline{\mathscr{P}}^*_+} \phi_+(\boldsymbol{p}^*_+) \leq \min_{\boldsymbol{p}^* \in \mathscr{P}^*} \phi(\boldsymbol{p}^*). \tag{38}$$

To see this, consider a map $h : \mathscr{P}^* \to \overline{\mathscr{P}}_+$ such that $h(\boldsymbol{p}^*) = (p^*_k)_{k \in \overline{\mathscr{K}}_+}$, i.e., it "cuts off" all elements in $\mathscr{K} \setminus \overline{\mathscr{K}}_+$. As $\cup_{\boldsymbol{p}^*} \operatorname{supp}(\boldsymbol{p}^*) \subseteq \overline{\mathscr{K}}_+$, the elements dropped by the map h must all be zero. Hence, we conclude that for all $\boldsymbol{p}^* \in \mathscr{P}^*$, (1) $h(\boldsymbol{p}^*) \in \overline{\mathscr{P}}^*_+$; (2) $\phi_+(h(\boldsymbol{p}^*)) = \phi(\boldsymbol{p}^*)$. Combining both, Equation (38) must hold. Recalling that $\phi_+(\overline{\boldsymbol{p}}_+) = \min_{\boldsymbol{p}^*_+ \in \overline{\mathscr{P}}^*_+} \phi_+(\boldsymbol{p}^*_+)$, we derive $\phi_+(\overline{\boldsymbol{p}}_+) \leq \min_{\boldsymbol{p}^* \in \mathscr{P}^*} \phi(\boldsymbol{p}^*)$. Finally, as $\overline{\boldsymbol{p}} = [\overline{\boldsymbol{p}}_+; \mathbf{0}] \in \mathscr{P}^*$ and $\phi_+(\overline{\boldsymbol{p}}_+) = \phi(\overline{\boldsymbol{p}})$, we have $\phi(\overline{\boldsymbol{p}}) = \min_{\boldsymbol{p}^* \in \mathscr{P}^*} \phi(\boldsymbol{p}^*)$, which means $\overline{\boldsymbol{p}}$ is the MEUE of the original routing game. □