Convergence in a Repeated Non-atomic Routing Game
with Partial Signaling

Yixian Zhu Ketan Savla

Abstract

We study the following repeated non-atomic routing game. In every round, nature chooses a state in an i.i.d.
manner according to a publicly known distribution, which influences link latency functions. The system planner makes
private route recommendations to participating agents, which constitute a fixed fraction, according to a publicly known
signaling strategy. The participating agents choose between obeying or not obeying the recommendation according
to cumulative regret of the participating agent population in the previous round. The non-participating agents choose
route according to myopic best response to a calibrated forecast of the routing decisions of the participating agents.
We show that, for parallel networks, if the planner’s signal strategy satisfies the obedience condition, then, almost
surely, the link flows are asymptotically consistent with the Bayes correlated equilibrium induced by the signaling
strategy.

I. INTRODUCTION

Optimal decision-making for agents in modern complex socio-technical systems, such as transportation, is
challenging because of dynamical environments and uncertainties. Personal mobile devices allow the agents to
stay updated on the information, and even make decisions for them. This architecture also allows the possibility for
information design, whereby the source of information, say a system planner, can strategically pick what information
to reveal or make direct route recommendations to steer the traffic flow towards optimal social state.

We consider a non-atomic routing game where the link latency functions are conditional on the state of the
nature, e.g., whether there is a traffic incident or not. This state is generated in an i.i.d. manner from an exogenous
distribution. This distribution is known publicly to all the agents, e.g., from historical data or personal experience, but
the actual realization is known only to the system planner. It is therefore natural to consider that the system planner
could use this informational advantage to recommend routes to the agents in order to optimize a social objective.
Such a mapping from the state realization to route recommendation is referred to as signaling strategy in line with
the information design literature [1]. The recommendations would be obeyed by the agents only if, by following it,
they are better off than by making decision based only on the prior distribution of the states. Such a requirement on
the signaling strategy is formally referred to as obedience constraint. Extension of this constraint to a multi-round
setting is not clear. Specifically, how does an agent determine obedience based on the series of recommendations
and payoffs that she receives, and can a static notion guarantee persistent obedience in a multi-round setting? One naive justification is that the agents follow the recommendations in every round until they have sufficient samples to empirically evaluate the obedience condition. If the strategy indeed satisfies the obedience constraint, then the agents’ evaluation would be so and therefore they will continue following the recommendation subsequently. We rather consider a setting where the agents choose between obeying or not obeying recommendation in every round from the beginning. This choice is determined by the agents’ regret associated with the decision in the previous round and aggregated over all other agents. Moreover, we also allow for partial signaling, i.e., when only a fraction of agents participate in signaling. The rest are assumed to choose routes as best response to their forecast of the link flows induced by the participating agents.

The motivation for the setting comes from the practical consideration that agents have the discretion to use or not to use a navigation system, and if they use one, then whether to obey the recommendation or not on any given trip. It is therefore of interest to understand the long term effectiveness of a route recommendation strategy in such a setting. The feature of our model to let the decision of a participating agent depend on the collective experience of all the participating agents in the previous rounds is inspired by platforms such as Yelp which aggregate feedback from all users to provide a single review rating which is publicly accessible. The decision models in this paper are reminiscent of well-known models studied in the context of convergence to correlated equilibrium, e.g., [2]–[4]. This is not completely surprising given that the obedience constraint has been shown to be equivalent to Bayes correlated equilibrium [5]. Specifically, a participating agent computes regret associated with the decision in the previous round in the same spirit as [4]; however, its decision in the current round depends on the cumulative regret of all the participating agents, and also, the regret determines the propensity to follow the recommendation in the current round as opposed to deviating from the decision in the previous round as in [4]. The model for non-participating agents to choose routes as best response to a forecast of the flow induced by the participating agents is reminiscent of [2]. Our main result is that, if the signaling strategy is obedient, then the non-participating agents’ forecast of the link flows induced by the participating agents converges to actual induced link flows, which in turn converges to the link flows corresponding to the Bayes correlated equilibrium associated with the signaling strategy.

The rest of the paper is organized as follows. Section II formulates the non-atomic routing game with partial signaling, defines the notion of partial information correlated equilibrium, as well as describes the repeated game setting. Section III provides convergence analysis, and Section IV provides illustrative simulation results. Concluding remarks are provided in Section V. The proofs of all the technical results are collected in the Appendix.

We end this section by defining key notations to be used throughout this paper. Let \( \triangle(X) \) denote the set of all probability distributions on \( X \). For a vector \( x \in \mathbb{R}^n \), \( \text{diag}(x) \) will denote the \( n \times n \) diagonal matrix with elements of \( x \) on the main diagonal. For an integer \( n \), we let \( [n] := \{1, 2, \ldots, n\} \). For \( \lambda \geq 0 \), let \( \mathcal{P}_n(\lambda) := \left\{ x \in \mathbb{R}_{\geq 0}^n \mid \sum_{i \in [n]} x_i = \lambda \right\} \) be the \((n-1)\)-dimensional probability simplex of size \( \lambda \). When \( \lambda = 1 \), we shall simply denote the simplex as \( \mathcal{P}_n \) for brevity in notation. For a real number \( r \in \mathbb{R} \), we let \( [r]^+ = \max(r, 0) \), and \( \lfloor r \rfloor \) to be the largest integer no greater than \( r \).
II. Model

A. Non-atomic Routing Game with Partial Signaling

Consider a network consisting of \( n \) parallel links between a single source-destination pair. Without loss of generality, let the agent population generate a unit volume of traffic demand. The link latency functions \( \ell_{\omega,i}(f_i), \ i \in [n] \), give latency on link \( i \) as a function of flow \( f_i \) through them, conditional on the state of the network \( \omega \in \Omega = \{\omega_1, \ldots, \omega_s\} \). Throughout the paper, we shall make the following basic assumption on these functions.

**Assumption 1:** For every \( i \in [n], \omega \in \Omega \), \( \ell_{\omega,i} \) is a non-negative, continuously differentiable and non-decreasing function.

At times, we shall strengthen the assumption to \( \ell_{\omega,i} \) being strictly increasing. For Proposition 1, Lemma 1, Lemma 2 and Theorem 1, we restrict ourselves to general polynomial latency function in the following form:

\[
\ell_{\omega,i}(f_i) = \sum_{d=0}^{D} \alpha_{d,\omega,i} f_i^d, \ i \in [n], \ \omega \in \Omega
\]

(1)

with \( \alpha_{0,\omega,i} \geq 0 \) and \( \alpha_{1,\omega,i} \geq 0 \). We shall also let \( \alpha_{d,\omega} \) refer to the \( n \times 1 \) column vector whose entries are \( \alpha_{d,\omega,i} \), and \( \alpha_{d} \) refer to the \( s \times n \) matrix whose entries are \( \alpha_{d,\omega,i} \).

Let \( \omega \sim \mu_0 \in \text{int}(\triangle(\Omega)) \), for some prior \( \mu_0 \) which is known to all the agents. The agents do not have access to the realization of \( \omega \), but a fixed fraction \( \nu \in [0,1] \) of the agents receives private route recommendations conditional on the realized state. These conditional recommendations are generated by a signal \( \pi = \{\pi_{\omega} \in \mathcal{P}_n(\nu) : \omega \in \Omega\} \).

**Remark 1:** The signal defined above is diagonal atomic [6]. A generalization is \( \pi = \{\pi_{\omega} \in \triangle(\mathcal{P}_n(\nu)) : \omega \in \Omega\} \). We postpone consideration of such a general setting to future work.

B. Repeated Game Setting

Consider the following repeated game setting. The strategy space of a participating agent (P-agent) is \{obey, do not obey\}, whereas the strategy space of a non-participating Bayesian agent (B-agent) is \{1,\ldots,n\}. The do not obey behavior of the P-agents is modeled by the row-stochastic matrix \( P \) with zeros on the diagonal, where \( P_{ij} \) is the fraction of do not obey P-agents who are recommended \( i \) but choose \( j \). The choice of a P-agent is driven by a notion of regret, and the decision of a B-agent is myopic best response to a forecast of P-agent decisions. The details are as follows.

1) Route choice model for P-agents: The outline of the decision making process of P-agents is as follows. At the end of stage \( k \), every P-agent computes difference between the payoffs associated with the recommended choice and the alternative choice(s). These payoff differences are then aggregated over the entire population of P-agents to give \( u(k) \). The average of the initial condition \( m(1) \) and \( u(1), \ldots, u(k) \), denoted as \( m(k+1) \), is then mapped to regret as \( \theta(m(k+1)) \in [0,1] \), which equals the fraction of P-agents who do not follow recommendation in stage \( k+1 \). The details for each step in this process is provided next.

We assume that upon completion of the trips, all agents (P- and B-agents) have access to traffic report from the \( k \)-th stage. This report consists of \( \omega(k) \) and \( \{\ell_{\omega(k),i}\}_{i \in [n]} \). For simplicity, first consider the two-link case.
where $\ell_{\omega(1)} > \ell_{\omega(2)}$. For a $P$-agent who was recommended route 1, irrespective of whether she obeys the recommendation or not, her recommendation is sub-optimal by $\ell_{\omega(1)} - \ell_{\omega(2)}$. In the general case of $n \geq 2$ links, for a $P$-agent who obeys recommendation to take route $i$, the sub-optimality of the recommendation is $\ell_{\omega(i)} - \sum_{j \in [n]} P_{ij} \ell_{\omega(j)}$. On the other hand, for a $P$-agent who does not obey the recommendation to take route $i$ but rather takes route $j$, the sub-optimality of her recommendation is $\ell_{\omega(i)} - \ell_{\omega(j)}$. Taking into account the number of $P$-agents who are recommended different routes and the fraction of them who obey or do not obey the recommendation, the aggregation of payoff difference over the entire $P$-agent population is given by:

$$u(k) := \sum_{i \in [n]} \left( \ell_{\omega(i)} - \sum_{j \in [n]} P_{ij} \ell_{\omega(j)} \right) \pi_{\omega(i)} (1 - \theta(m(k))) + \sum_{i,j \in [n]} \left( \ell_{\omega(i)} - \ell_{\omega(j)} \right) P_{ij} \pi_{\omega(i)} \theta(m(k))$$

$$= \sum_{i,j \in [n]} \left( \ell_{\omega(i)} - \ell_{\omega(j)} \right) P_{ij} \pi_{\omega(i)}$$

$$= \pi_{\omega(k)}^T (I - P) \ell_{\omega(k)}$$

(2)

where $\ell_{\omega(k)}$ is the $n \times 1$ column vectors whose entries are $\ell_{\omega(i), i} \in [n]$. The average of these instantaneous payoff differences (as well as the initial condition $m(1)$) is:

$$m(k + 1) = \frac{1}{k + 1} (m(1) + u(1) + \ldots + u(k))$$

$$= \frac{k}{k + 1} m(k) + \frac{1}{k + 1} u(k)$$

(3)

We adopt the following notion of regret:

$$\theta(m(k)) = \frac{\lfloor m(k) \rfloor^+}{m_{\text{max}}}$$

(4)

where $m_{\text{max}}$ is chosen to be sufficiently large so that $\theta(m(k)) \in [0, 1]$ for all possible values of $m(k)$. Since $\lfloor m(k) \rfloor^+$ can be upper bounded by $\sum_{i \in [n]} \sum_{d=0}^{D_{\omega,i}} \alpha_{\omega,i}$, it can then serve as a lower bound for $m_{\text{max}}$.

$\theta(k) = \theta(m(k))$ is interpreted as the fraction of agents who do not follow recommendation. Therefore, the link flows induced by $P$-agents be given by:

$$x_i(k) = x_i(m(k), \omega(k)) = \nu \pi_{\omega(i)} (1 - \theta(m(k))) + \nu \sum_{j \in [n]} P_{ij} \pi_{\omega(j)} \theta(m(k))$$

which in matrix form becomes:

$$x(k) = x(m(k), \omega(k)) = \nu \pi_{\omega(k)} (1 - \theta(m(k))) + \nu \sum_{j \in [n]} P_{ij} \pi_{\omega(j)} \theta(m(k))$$

$$= \nu \left( \pi_{\omega(k)} - \theta(m(k)) \pi_{\omega(k)} + \theta(m(k)) P^T \pi_{\omega(k)} \right)$$

$$= \nu \pi_{\omega(k)} + \nu \theta(m(k)) (P^T - I) \pi_{\omega(k)}$$

(5)

where $x(m(k), \omega(k)) \in \mathcal{P}_n(\nu)$.

Remark 2: The framework in (2)-(5) is reminiscent of the regret matching framework of [4]. Details on comparison are provided in Section B. A few points are worth special emphasis:

1) The aggregation of payoff difference over all $P$-agents in (2) is inspired by platforms such as Yelp.
2) In the framework of [4], $[m]^+$ is referred to as regret and $m_{\text{max}}$ is the inertia parameter which captures the propensity of a $P$-agent to stick to the action choice in the previous round. On the other hand, in our setup, $\theta = [m]^+/m_{\text{max}}$ is interpreted as the degree of obedience of an individual agent, tuned by the parameter $m_{\text{max}}$.

2) Route choice model for $B$-agents: The B-agents forecast the flow induced by $P$-agents for different realizations of $\omega$ using a model structurally similar to (5), as follows:

$$\hat{x}_i(k) = \hat{x}_i(\hat{\theta}(k), \omega(k)) = \nu \pi_{\omega(k),i} \left( 1 - \hat{\theta}(k) \right) + \nu \sum_{j \in [n]} P_{ji} \pi_{\omega(k),j} \hat{\theta}(k)$$

or in matrix form as:

$$\hat{x}(k) = \hat{x}(\hat{\theta}(k), \omega(k)) = \nu \pi_{\omega(k)} + \nu \hat{\theta}(k) (P^T - I) \pi_{\omega(k)}$$

(6)

where $\hat{\theta}$ is B-agents’ forecast of $\theta$. Let us denote the myopic best response to the forecast as $\bar{y}(\hat{\theta}(k))$. For a given $\theta \in [0, 1]$, $y(\theta)$ is the unique (assuming strictly increasing latency functions $\{\ell_\omega, i\}_{\omega, i}$, more details in Lemma 2) $y \in P_n(1-\nu)$ satisfying:

$$y_i > 0 \implies E_{\omega \sim \mu_0}[\ell_{\omega, i}(\bar{x}_i(\theta, \omega) + y_i)] \leq E_{\omega \sim \mu_0}[\ell_{\omega, j}(\bar{x}_j(\theta, \omega) + y_j)], \quad i, j \in [n]$$

(8)

**Remark 3**: Implementation of (7) requires B-agents to have access to $\theta(k)$. As we assumed in Section II-B.1, B-agents have access to state $\omega(k)$, delays $\{\ell_{\omega(k), i}\}_{i \in [n]}$. B-agents also know the fixed signal $\pi$, as well as explicit form of latency functions. For a given $\hat{\theta}(k)$, (6) gives $\hat{x}_i(k)$. (8) then gives $y(\hat{\theta}(k))$, i.e., the flow induced by $B$-agents is known to them. Assuming the latency functions are strictly increasing, the total link-wise flows can be inferred from other information known to $B$-agents. They can then infer the actual inflows induced by $P$-agents. Using the inverse of (5), B-agents can then get $\theta(k)$.

An equivalent variational inequality characterization of $y(\theta)$ is that it is the unique $y \in P_n(1-\nu)$ satisfying

$$\sum_{i \in [n]} E_{\omega \sim \mu_0}[\ell_{\omega, i}(\bar{x}_i(\theta, \omega) + y_i)](z_i - y_i) \geq 0 \quad \forall z \in P_n(1-\nu)$$

(9)

In matrix form, $y(\theta)$ satisfies:

$$E_{\omega \sim \mu_0} \left[ (z - y(\theta))^T \ell_{\omega}(\hat{x}(\theta, \omega) + y) \right] \geq 0 \quad \forall z \in P(1-\nu)$$

(10)

Note that the total link flow forecasted by $B$-agents is $\hat{x}(\hat{\theta}(k), \omega(k)) + y(\hat{\theta}(k))$, but the actual flow is $x(m(k), \omega(k)) + y(\hat{\theta}(k))$. 

July 26, 2022 DRAFT
Remark 4: 1) One can show that the map defined in (8) is Lipschitz continuous. Detailed proof in Lemma 2.
2) The forecast and best response combination of B agent decisions is reminiscent of [2], where convergence is established to correlated equilibrium set in two player games if the forecast is calibrated to the plays of the opponent. We argue in Section A that, with appropriate adaptation to the non-atomic setting of this paper, the forecast strategy of B-agents could also interpreted as being calibrated to P-agents’ actions.

III. Convergence Analysis and Discussion

We shall establish convergence of (2)-(7) for \( \pi \) which are obedient, i.e., for which there exists \( y \in \mathcal{P}_n(1 - \nu) \) such that:

\[
\sum_{\omega \in \Omega} \ell_{\omega,i}(\pi_{\omega,i} + y_i) \pi_{\omega,i} \mu_0(\omega) \leq \sum_{\omega \in \Omega} \ell_{\omega,j}(\pi_{\omega,j} + y_j) \pi_{\omega,i} \mu_0(\omega), \quad i, j \in [n] \tag{11a}
\]

\[
\sum_{\omega \in \Omega} \ell_{\omega,j}(\pi_{\omega,j} + y_j) y_i \mu_0(\omega) \leq \sum_{\omega \in \Omega} \ell_{\omega,j}(\pi_{\omega,j} + y_j) y_i \mu_0(\omega), \quad i, j \in [n] \tag{11b}
\]

Following [5], \((\pi, y)\) is therefore the Bayes Correlated Equilibrium induced by \(\pi\). We start the convergence analysis by considering the extreme case when there are no B-agents, i.e., when \(\nu = 1\).

Proposition 1: Consider the dynamics in (2)-(5) with \(\nu = 1\), and general polynomial latency functions in (1). For any \(\mu_0\), obedient \(\pi\), and every \(m(1) \in [-m_{\max}, +m_{\max}]\), we have that, almost surely

\[
\lim_{k \to \infty} x_i(k) - \pi_{\omega(k),i} = 0, \quad i \in [n]
\]

Towards consideration of the general case of heterogeneous population, we first provide an analysis of the case in the absence of the aggregator, i.e., for some fixed \(m(k) \equiv m\).

Proposition 2: Consider the dynamics in (2)-(7) for some fixed \(m(k) \equiv m \in [-m_{\max}, +m_{\max}]\). For any \(\mu_0\), obedient \(\pi\), \(\nu \in (0, 1)\), \(\hat{\theta}(1) \in [0, 1]\), we have that

\[
\lim_{k \to \infty} \hat{\theta}(k) = \theta(m)
\]

Next, we extend the analysis to the case when \(m(k)\) evolves according to (2)-(3).

Proposition 3: Consider the dynamics in (2)-(7). For any \(\mu_0\), obedient \(\pi\), \(\nu \in (0, 1)\), \(\hat{\theta}(1) \in [0, 1]\), \(m(1) \in [-m_{\max}, +m_{\max}]\), we have that

\[
\lim_{k \to \infty} \hat{\theta}(k) - \theta(m(k)) = 0
\]

The following results will be useful in proving convergence of link-wise flows.

Lemma 1: Consider the dynamics in (2)-(7) for general polynomial latency functions in (1). For any \(\mu_0\), obedient \(\pi\), \(\nu \in (0, 1)\), and initial conditions \(\hat{\theta}(1) \in [0, 1]\) and \(m(1) \in [-m_{\max}, +m_{\max}]\), there exists a subsequence \(\{m(k_s)\}_s\), whose limit is negative almost surely.

Lemma 2: The mapping \(g(\theta) : [0, 1] \to \mathcal{P}_n(1 - \nu)\) defined in (8) for general polynomial latency functions in (1) is Lipschitz continuous.

We are now ready to state the main result.
Theorem 1: Consider the dynamics in (2)-(7) for general polynomial latency functions in (1). For any $\mu_0$, obedient $\pi, \nu \in (0, 1)$, and initial conditions $\hat{\theta}(1) \in [0, 1]$ and $m(1) \in [-m_{\text{max}}, +m_{\text{max}}]$, we have that, almost surely

$$\lim_{k \to \infty} x_i(k) - \nu \pi \omega(k), i = 0, i \in [n]$$

IV. SIMULATIONS

We report simulation results which suggest that the main convergence result is robust to natural variations of the dynamics in (2)-(7). We consider a network with two parallel links. Unless noted otherwise, in all the scenarios, $\mu_0(\omega) = 0.6 = 1 - \mu_0(\omega)$, $\nu = 0.5$, $\theta(1) = 0.5$, $\hat{\theta}(1) = 0.25$, and the total demand is set to be 1. Also, the link latency functions are affine with coefficients:

$$\alpha_0 = \frac{\omega_1}{\omega_2} \begin{bmatrix} 5 & 25 \\ 20 & 15 \end{bmatrix}, \quad \alpha_1 = \frac{\omega_1}{\omega_2} \begin{bmatrix} 4 & 2 \\ 1 & 2 \end{bmatrix}$$

![Fig. 1](image_url)

Fig. 1: Convergence of regret and its forecast over two parallel links for (a) review aggregation with discounting factor $\lambda = 0.9$, and (b) dynamic participation rate.

In Scenario (a), we replace (3) with discounted review aggregation: $m(k+1) = \lambda m(k) + (1 - \lambda) u(k)$. In Scenario (b), we let the participation rate also be dynamic, and in particular, we let $\nu(k) = \theta(k)$. Figure [1] which shows results for these scenarios, suggests that convergence of regret forecast as well as actual regret to zero is achieved even under these variations.

V. CONCLUSION

In this paper, we provided convergence analysis for a non-atomic routing game with partial signaling. The decision models are reminiscent of adaptive models studied previously in the context of dynamic procedures leading to the
set of correlated equilibria. Our analysis lays the foundation to extend such analysis to repeated scenarios with explicit presence of a coordinator, albeit in the context of non-atomic routing games.

There are several natural directions for future work. We plan to extend the analysis to general increasing latency functions, and to settings involving multiple destinations. The notion of conditional universal consistency in [3] potentially offers a framework to unify and generalize the specific decision models of the participating and non-participating agents considered in this paper. Finally, it would be interesting to consider a multiscale routing decision framework to enable integration of traffic flow dynamics into the decision-theoretic analysis, e.g., as in [7].

REFERENCES

[1] D. Bergemann and S. Morris, “Information design: A unified perspective,” Journal of Economic Literature, vol. 57, no. 1, pp. 44–95, 2019.
[2] D. P. Foster and R. V. Vohra, “Calibrated learning and correlated equilibrium,” Games and Economic Behavior, vol. 21, no. 1-2, p. 40, 1997.
[3] D. Fudenberg and D. K. Levine, “Conditional universal consistency,” Games and Economic Behavior, vol. 29, no. 1-2, pp. 104–130, 1999.
[4] S. Hart and A. Mas-Colell, “A simple adaptive procedure leading to correlated equilibrium,” Econometrica, vol. 68, no. 5, pp. 1127–1150, 2000.
[5] D. Bergemann and S. Morris, “Bayes correlated equilibrium and the comparison of information structures in games,” Theoretical Economics, vol. 11, no. 2, pp. 487–522, 2016.
[6] Y. Zhu and K. Savla, “Information design for non-atomic routing games: Computation and structural properties,” IEEE Transactions on Control of Network Systems, 2020. Under revision. Available at https://viterbi-web.usc.edu/~ksavla/papers/infodesign-sdp.pdf.
[7] G. Como, K. Savla, D. Acemoglu, M. A. Dahleh, and E. Frazzoli, “Stability analysis of transportation networks with multiscale driver decisions,” SIAM Journal on Control and Optimization, vol. 51, no. 1, pp. 230–252, 2013.
[8] M. Abramowitz and I. A. Stegun, Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, Chapter 5, vol. 55. US Government printing office, 1964.
[9] S. Dempe and S. Vogel, “The subdifferential of the optimal solution in parametric optimization,” Technical report, Fachbereich Mathematik, Technische Universität Chemnitz, Chemnitz, Germany, 1997.

APPENDIX

A. Connection to Calibrated Forecast

[2] studies a two player repeated game, in which, in each round, each player forecasts the probability \( p = (p_1, \ldots, p_n) \) of decisions for the other player. Convergence to the set of correlated equilibria is established if the forecasting rule is calibrated. Let \( N(p, k) \) be the number of rounds up to the \( k \)-th round in which the forecast is \( p \), and let \( \rho(p, j, k) \) be the fraction of these rounds for which the other player plays \( j \). Then the forecasting rule is said to be calibrated if:

\[
\lim_{k \to \infty} \sum_p \left| \rho(p, j, k) - p_j \right| \frac{N(p, k)}{k} = 0 \quad \forall j
\]  

(12)

In contrast to [2], our setting considers non-atomic agents, and the utility of every player depends also on the state \( \omega \), which changes from one round to the next. Non-atomicity gives \( B \) agents sufficient samples to compare forecast with outcome on a per round basis. Specifically, the flow forecast \( \hat{x}_j \) (after normalization by \( \nu \)) could be interpreted as the forecast of the probability with which an individual \( P \)-agent will pick route \( j \), and similarly \( x_j \)
is the probability with which a $P$-agent actually picks $j$. Therefore, a reasonable adaptation of (12) to our setting is that the $B$ agents’ forecast is calibrated if:

$$\lim_{k \to \infty} \frac{1}{k} \sum_{t=1}^{k} \left| x_j(t) - \hat{x}_j(t) \right| = 0 \quad \forall j$$

(13)

for almost every i.i.d. realization of $\omega(t) \sim \mu_0$.

Proposition 3 shows that $\lim_{k \to \infty} \hat{\theta}(k) - \theta(m(k)) = 0$. Combine with (5) and (6), one can easily get $\lim_{k \to \infty} x_j(k) - \hat{x}_j(k) = 0$ for $\forall j$. Thus the forecasting rule in (6)-(7) is calibrated in the sense of (13).

B. Connection to Regret-Matching

[4] considers the following repeated game between finite players. In round $k$, a player who chose action $i$ in round $k-1$, switches to $j \neq i$ with probability $\frac{m_{i \to j}(k)^+}{m_{\text{max}}}$, and sticks to $i$ with probability $1 - \sum_{j \neq i} \frac{m_{i \to j}(k)^+}{m_{\text{max}}}$, where $m_{i \to j}(k)$ is a measure of regret for not having played $j$ every time that $i$ was played in the past, and the inertia parameter $m_{\text{max}}$ is sufficiently large to ensure that the probabilities are well defined. Formally, the regret is computed as: $m_{i \to j}(k) := \frac{1}{k-1} \sum_{t=1}^{k-1} \text{player chooses } i(U(j, s(t)) - U(s(t)))$, where $U$ is the utility function of the player under consideration, $s(t)$ is the set of actions chosen by all the players in round $t$, and $s^-(t)$ is the set of actions of all the players except the one under consideration.

The similarities and differences with the decision model of the $P$ agents in our setup is now apparent. An individual $P$ agent computes regret for its action to follow or to not to follow the recommendation in the previous round in the same spirit as [4]. However, this regret is not categorized in terms of the specific action taken in the previous round. Such an absence of conditioning facilitates efficient aggregation of regrets of all the $P$ agents as a simple summation. The time-averaging of regrets is achieved through (3), where $m_{\text{max}}$ plays the role as the inertia parameter similar to [4].
C. Proof of Proposition \[7\]

Plugging \[4\] into \[3\] and substituting \[5\] gives

\[
m(k) = \frac{m(1)}{k} + \frac{1}{k} \sum_{t=1}^{k-1} u(t)
\]

\[
= \frac{m(1)}{k} + \frac{1}{k} \sum_{t=1}^{k-1} \pi_{\omega(t)}^T (I - P) \ell_{\omega(t)}(x(t))
\]

\[
= \frac{m(1)}{k} + \frac{1}{k} \sum_{t=1}^{k-1} \pi_{\omega(t)}^T (I - P) \left( \alpha_{0,\omega(t)} + \sum_{d=1}^{D} \text{diag}(\alpha_{d,\omega(t)}) \text{diag}(x(t))^{d-1} x(t) \right)
\]

\[
= \frac{m(1)}{k} + \frac{1}{k} \sum_{t=1}^{k-1} \pi_{\omega(t)}^T (I - P) \ell_{\omega(t)}(t) \left( \alpha_{0,\omega(t)} + \sum_{d=1}^{D} \text{diag}(\alpha_{d,\omega(t)}) \left( \text{diag}(x(t))^{d-1} x(t) - \text{diag}(\pi_{\omega(t)})^{d-1} \pi_{\omega(t)} \right) \right)
\]

\[
= \frac{m(1)}{k} + \frac{1}{k} \sum_{t=1}^{k-1} \pi_{\omega(t)}^T (I - P) \ell_{\omega(t)}(\pi_{\omega(t)})
\]

\[
= \frac{m(1)}{k} + \frac{1}{k} \sum_{t=1}^{k-1} \pi_{\omega(t)}^T (I - P) \ell_{\omega(t)}(\pi_{\omega(t)}) - \frac{1}{k} \sum_{t=1}^{k-1} \theta(t) \pi_{\omega(t)}^T (I - P) \sum_{d=1}^{D} \text{diag}(\alpha_{d,\omega(t)}) \sum_{i=0}^{d-1} \text{diag}(\pi_{\omega(t)})^{i} (I - P^T) \pi_{\omega(t)}
\]

\[
\leq \frac{m(1)}{k} + \frac{1}{k} \sum_{t=1}^{k-1} \pi_{\omega(t)}^T (I - P) \ell_{\omega(t)}(\pi_{\omega(t)})
\]

\[
= \frac{m(1)}{k} + \sum_{i,j \in [n]} P_{ij} \frac{1}{k} \sum_{t=1}^{k-1} \pi_{\omega(t),i} \left( \ell_{\omega(t),i}(\pi_{\omega(t)}) - \ell_{\omega(t),j}(\pi_{\omega(t)}) \right)
\]

(14)

\[11\] and the strong law of large numbers implies that, almost surely, lim$_{k \to \infty} \frac{1}{k} \sum_{t=1}^{k-1} \pi_{\omega(t),i} \left( \ell_{\omega(t),i}(\pi_{\omega(t)}) - \ell_{\omega(t),j}(\pi_{\omega(t)}) \right) \leq 0$ for all $i, j \in [n]$, and therefore, almost surely, lim sup$_{k \to \infty} m(k) \leq 0$, and hence lim$_{k \to \infty} \theta(k) = 0$ from \[4\]. The proposition then follows from \[5\].

D. Proof of Proposition \[2\]

\[7\] implies $e_\theta(k + 1) = (1 - \beta(k + 1)) e_\theta(k)$, and hence $e_\theta(k + 1) = \prod_{i=2}^{k+1} (1 - \beta(t)) e_\theta(1)$. The proposition then follows from the fact that $(1 - \beta_{\text{max}})^k e_\theta(1) \leq e_\theta(k + 1) \leq (1 - \beta_{\text{min}})^k e_\theta(1)$.

E. Proof of Proposition \[3\]

\[7\] implies

\[
e_\theta(k + 1) = (1 - \beta(k + 1)) e_\theta(k) + \delta(k)
\]

where $\delta(k) = \theta(k + 1) - \theta(k)$. Using the expression of $\theta$ from \[4\] and noting that $|a| - |b| \leq |a - b|$ for all $a, b \in \mathbb{R}$, we get

\[
\frac{-|u(k) - m(k)|^+}{(k + 1)m_{\text{max}}} \leq \delta(k) \leq \frac{|u(k) - m(k)|^+}{(k + 1)m_{\text{max}}}
\]
Since \([a - b]^+ \leq |a| + |b|\) for all \(a, b \in \mathbb{R}\), and both \(m(k)\) and \(u(k)\) are absolutely upper bounded by \(m_{\text{max}}\),

\[
\frac{-2}{k + 1} \leq \delta(k) \leq \frac{2}{k + 1} \tag{16}
\]

Since \(\delta(k)\) can only be bounded by a harmonic sequence whose partial sum sequence diverges, it is not easy to prove that \(\lim_{k \to \infty} e_\theta(k) = 0\) directly using (15). Instead, we lower and upper bound \(e(k)\) by \(\varepsilon_\theta(k)\) and \(\tau_\theta(k)\), respectively, and show that these converge to zero. (15) can be rewritten as

\[
e_\theta(k) = \left( \prod_{t=2}^{k} (1 - \beta(t)) \right) e_\theta(1) + \sum_{t=1}^{k-2} \left( \prod_{\tau=t+2}^{k} (1 - \beta(\tau)) \right) \delta(t) + \delta(k - 1) \tag{17}
\]

Substituting (16) into (17),

\[
\varepsilon_\theta(k) := \left( \prod_{t=2}^{k} (1 - \beta(t)) \right) e_\theta(1) - 2\tilde{\delta}(k) \leq e_\theta(k) \leq \left( \prod_{t=2}^{k} (1 - \beta(t)) \right) e_\theta(1) + 2\tilde{\delta}(k) =: \tau_\theta(k) \tag{18}
\]

where \(\tilde{\delta}(k) = \sum_{t=2}^{k} (1 - \beta_{\min}) \frac{k+1}{t} = \gamma^k \sum_{t=2}^{k} \frac{\gamma^{-t}}{t} \), with \(\gamma := 1 - \beta_{\min} \in (0, 1)\), \(\frac{\gamma^{-t}}{t}\) is decreasing over \((0, -\frac{1}{\ln \gamma})\), and increasing over \([-\frac{1}{\ln \gamma}, +\infty)\). Therefore, with \(k^* := \left[-\frac{1}{\ln \gamma}\right]\),

\[
\tilde{\delta}(k) = \gamma^k \left( \sum_{t=2}^{k^*} \frac{\gamma^{-t}}{t} \right) + \sum_{t=k^*+1}^{k} \frac{\gamma^{-t}}{t} \leq \gamma^k \left( \int_1^{k^*} \frac{\gamma^{-t}}{t} \, dt + \int_{k^*+1}^{k+1} \frac{\gamma^{-t}}{t} \, dt \right) \leq \gamma^k \int_{1}^{k+1} \frac{\gamma^{-t}}{t} \, dt
\]

Change of variable \(z = \gamma^{-t}\) gives \(I_{k+1} = \int_{\gamma^{-1}}^{\gamma^{-(k+1)}} \frac{dz}{\ln z} = \text{Li}(\gamma^{-(k+1)}) - \text{Li}(\gamma^{-1})\), where \(\text{Li}(\cdot)\) is the logarithmic integral function [8] whose asymptotic behavior is \(\text{Li}(r) = O\left(\frac{1}{\ln r}\right)\). Since \(\gamma \in (0, 1)\), \(\gamma^{-(k+1)} \to \infty\) as \(k \to \infty\). Therefore, as \(k \to \infty\),

\[
0 < \tilde{\delta}(k) \leq \gamma^k \left( \text{Li}(\gamma^{-(k+1)}) - \text{Li}(\gamma^{-1}) \right) = \gamma^k \left( O\left( \frac{\gamma^{-1}}{-(k+1) \ln \gamma} \right) - \text{Li}(\gamma^{-1}) \right) = O\left( \frac{\gamma^{-1}}{-(k+1) \ln \gamma} \right) - \gamma^k \text{Li}(\gamma^{-1}) \to 0
\]

which leads to \(\lim_{k \to \infty} \tilde{\delta}(k) = 0\). Using it in (18) therefore gives \(\lim_{k \to \infty} \varepsilon_\theta(k) = 0\) and \(\lim_{k \to \infty} \tau_\theta(k) = 0\), and hence \(\lim_{k \to \infty} e_\theta(k) = 0\).
F. Proof of Lemma 1

Plugging (4) into (3) and substituting (5) gives the following:

\[
m(k) = \frac{m(1)}{k} + \frac{1}{k} \sum_{t=1}^{k-1} \pi(t) T (I - P) \left( \alpha_{0,\omega(t)} + \sum_{d=1}^{D} \text{diag}(\alpha_{d,\omega(t)}) \text{diag}(x(t) + y(\hat{\theta}(t)))^{d-1} \left( x(t) + y(\hat{\theta}(t)) \right) \right)
\]

\[
= \frac{m(1)}{k} + \frac{1}{k} \sum_{t=1}^{k-1} \pi(t) T (I - P) \left( \alpha_{0,\omega(t)} + \sum_{d=1}^{D} \text{diag}(\alpha_{d,\omega(t)}) \text{diag}(\nu \pi_{\omega(t)} + y(0))^{d-1} (\nu \pi_{\omega(t)} + y(0)) \right)
\]

\[
+ \frac{1}{k} \sum_{t=1}^{k-1} \pi(t) T (I - P) \sum_{d=1}^{D} \text{diag}(\alpha_{d,\omega(t)}) \left( \text{diag}(x(t) + y(\hat{\theta}(t)))^{d-1} \left( x(t) + y(\hat{\theta}(t)) \right) \right)
\]

\[
= \frac{m(1)}{k} + \frac{1}{k} \sum_{t=1}^{k-1} \pi(t) T (I - P) \left( \alpha_{0,\omega(t)} + \sum_{d=1}^{D} \text{diag}(\alpha_{d,\omega(t)}) \text{diag}(\nu \pi_{\omega(t)} + y(0))^{d-1} (\nu \pi_{\omega(t)} + y(0)) \right)
\]

\[
+ \frac{1}{k} \sum_{t=1}^{k-1} \pi(t) T (I - P) \sum_{d=1}^{D} \text{diag}(\alpha_{d,\omega(t)}) \sum_{i=0}^{d-1} \text{diag}(x(t) + y(\hat{\theta}(t)))^{d-1-i} \left( \nu \pi_{\omega(t)} + y(0) \right)^i.
\]

(19)

Since \( m(k) \) is a bounded sequence, it contains a convergent subsequence \( \{m(k_s)\}_s \) with limit, say, \( m \). Considering (19) for this subsequence, using Proposition 3 we have, almost surely,

\[
m = \sum_{\omega \in \Omega} \mu_0(\omega) \pi_\omega T (I - P) \left( \alpha_{0,\omega} + \sum_{d=1}^{D} \text{diag}(\alpha_{d,\omega}) \text{diag}(\nu \pi_{\omega} + y(0))^{d-1} (\nu \pi_{\omega} + y(0)) \right)
\]

\[
+ \sum_{\omega \in \Omega} \mu_0(\omega) \pi_\omega T (I - P) \sum_{d=1}^{D} \text{diag}(\alpha_{d,\omega}) \sum_{i=0}^{d-1} \text{diag}(x(m, \pi_{\omega}) + y(\theta(m)))^{d-1-i} \left( \nu \pi_{\omega} + y(0) \right)^i.
\]

(20)

\[
= m_1 + m_2
\]

\[
= m_1 + m_2
\]

i.e., the limit \( m \) of every convergent subsequence of \( \{m(k)\}_k \) has to satisfy (20). We now show that (20) admits a unique solution in \( m \), and that this solution is non-positive. Toward that purpose, the following properties of \( m_1 \) and \( m_2 \) defined in (20) will be useful:

(a) \( m_1 \leq 0 \) for all \( m \). This is due to the obedience condition;

(b) \( m_2 = 0 \) for all \( m \leq 0 \). This follows from the definition of \( m_2 \), where \( \theta(m) = 0 \) for \( m \leq 0 \) from (4);

(c) \( m_2 \leq 0 \) for all \( m > 0 \), as follows, (10) for \( y(\theta(m)) \) and \( y(0) \), respectively, give:

\[
\sum_{\omega \in \Omega} \mu_0(\omega) (y(0) - y(\theta(m))) T \left( \alpha_{0,\omega} + \sum_{d=1}^{D} \text{diag}(\alpha_{d,\omega}) \text{diag}(x(m, \pi_{\omega}) + y(\theta(m)))^{d-1} (x(m, \pi_{\omega}) + y(\theta(m))) \right) \geq 0
\]

\[
\sum_{\omega \in \Omega} \mu_0(\omega) (y(\theta(m)) - y(0)) T \left( \alpha_{0,\omega} + \sum_{d=1}^{D} \text{diag}(\alpha_{d,\omega}) \text{diag}(\nu \pi_{\omega} + y(0))^{d-1} (\nu \pi_{\omega} + y(0)) \right) \geq 0
\]
Adding the two expressions gives
\[
\sum_{\omega \in \Omega} \mu_0(\omega) (y(0) - y(\theta(m)))^T \sum_{d=1}^{D} \text{diag}(\alpha_{d,\omega}) \left( \text{diag}(x(m, \pi_\omega) + y(\theta(m)))^{d-1} (x(m, \pi_\omega) + y(\theta(m))) \\
- \text{diag}(\nu \pi_\omega + y(0))^{d-1} (\nu \pi_\omega + y(0)) \right)
\]
\[
= \sum_{\omega \in \Omega} \mu_0(\omega) (y(0) - y(\theta(m)))^T \sum_{d=1}^{D} \text{diag}(\alpha_{d,\omega}) \sum_{i=0}^{d-1} \text{diag}(x(m, \pi_\omega) + y(\theta(m)))^{d-1-i} \text{diag}(\nu \pi_\omega + y(0))^i.
\]

Finally,
\[
m_2 = -\frac{1}{\nu \theta(m)} \sum_{\omega \in \Omega} \mu_0(\omega) (\nu \theta(m) \pi_\omega)^T (P - I) + y^T(\theta(m)) - y^T(0)) \sum_{d=1}^{D} \text{diag}(\alpha_{d,\omega}) \cdot \\
\sum_{i=0}^{d-1} \text{diag}(x(m, \pi_\omega) + y(\theta(m)))^{d-1-i} \text{diag}(\nu \pi_\omega + y(0))^i (\nu \theta(m)(P^T - I) \pi_\omega + y(\theta(m)) - y(0)) \geq 0 \quad (21)
\]

where we note that \(\theta(m) > 0\) for \(m > 0\).

(a), (b) and (c) imply that the right hand side of (20) is always non-positive. (a) and (b) also imply that the only non-positive solution to (20) is \(m = m_1\).

G. Proof of Lemma 2

Along the lines of Theorem 1 in [9], we just need to verify the following three conditions: Mangasarian-Fromowitz constraint qualification (MF), strong sufficient optimality condition of second order (SOC), and constant rank constraint qualification (CR). Plugging in (1) and rewriting the latency function in (8) gives:
\[
\ell_{\omega,i}(\hat{x}_i(\theta, \omega) + y_i) = \sum_{d=0}^{D} \alpha_{d,\omega,i}(\nu \pi_{\omega,i} + \nu \theta P^T \pi_\omega - \nu \theta \pi_{\omega,i} + y_i) = \alpha_{D,\omega,i} y_i^D + \sum_{d=0}^{D-1} \beta_{d,\omega,i} y_i^d \quad (22)
\]
where \(P_i \in \mathbb{R}^n\) is the \(i\)-th column of the row-stochastic matrix \(P\), and \(\beta_{d,\omega,i} = (\{\alpha_{k,\omega,i}\}_{k \in \{0,1,\ldots,D\}}, \pi_\omega, \theta, P_i)\) is the coefficient of monomial of degree \(d\). Taking expectation of (22) gives:
\[
E_{\omega \sim \mu_0}[\ell_{\omega,i}(\hat{x}_i(\theta, \omega) + y_i)] = E_{\omega \sim \mu_0}[\alpha_{D,\omega,i} y_i^D + \sum_{d=0}^{D-1} \beta_{d,\omega,i} y_i^d] = \bar{\alpha}_{D,\omega,i} y_i^D + \sum_{d=0}^{D-1} \bar{\beta}_{d,\omega,i} y_i^d
\]
where \( \tilde{\alpha}_{D,i} := E_{\omega \sim \mu_0}[\alpha_{D,i}], \tilde{\beta}_{d,i} := E_{\omega \sim \mu_0}[\beta_{d,i}] \), \( y(\theta) \) satisfies (8) if and only if it solves the following convex problem:

\[
\min_{y \in \mathbb{P}_n(1-\nu)} f(y, \theta) = \sum_{i \in [n]} \int_0^{y_i} E_{\omega \sim \mu_0}[\ell_{\omega,i}(\hat{x}_i(\theta, \omega) + s)] ds = \sum_{i \in [n]} \left( \frac{\tilde{\alpha}_{D,i}}{D} y_i^{D+1} + \sum_{d=0}^{D-1} \frac{\tilde{\beta}_{d,i}}{d+1} y_i^{d+1} \right)
\]

s.t. \( g_i(y, \theta) = -y_i \leq 0, \ i \in [n] \)

\[
h(y, \theta) = \sum_{i=1}^n y_i - 1 + \nu
\]

Moreover, such \( y \) is unique if \( \{\ell_{\omega,i}\}_{\omega,i} \) are strictly increasing over \( [0, 1] \). Let the Lagrange function be \( L(y, \theta, \lambda, \mu) = f(y, \theta) - \sum_{i \in [n]} \lambda_i y_i + \mu(\sum_{i=1}^n y_i - 1 + \nu) \). For a given \( \theta^0 \), \( y = y^0 \) is a locally optimal solution for (23) at \( \theta = \theta^0 \).

Now, we want to show that \( y(\theta) \) is Lipschitz continuous by verifying aforementioned three conditions.

**MF** From (23) we have:

\[
\nabla_y g_i(y^0, \theta^0) = \begin{bmatrix} 0 \\ \vdots \\ -1 \\ \vdots \\ 0 \end{bmatrix}_{n \times 1} \quad \text{← i-th entry} \quad \nabla_y h(y^0, \theta^0) = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}_{n \times 1}
\]

Obviously, \( \nabla_y h(y^0, \theta^0) \) is linearly independent. Let \( v \in \mathbb{R}^n \) be such that:

\[
v_i = \begin{cases} 1 & \text{if } g_i(y^0, \theta^0) = 0 \\ \frac{\sum_{i \in [n]} 1(g_i(y^0, \theta^0) = 0)}{\sum_{i \in [n]} 1(g_i(y^0, \theta^0) = 0) - n} & \text{if } g_i(y^0, \theta^0) \neq 0 \end{cases}
\]

\( v \) satisfies that:

\[
\nabla_y g_i(y^0, \theta^0)^T v < 0, \quad \text{for } g_i(y^0, \theta^0) = 0 \\
\nabla_y h(y^0, \theta^0)^T v = 0
\]

**SOC** Let \( \{H_{ij}\}_{i,j \in [n]} := \nabla_{yy}^2 L(y^0, \theta^0, \lambda, \mu) \), from (23) we have:

\[
H_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ D\tilde{\alpha}_{D,i} y_i^{D-1} + \sum_{d=1}^{D-1} d\tilde{\beta}_{d,i} y_i^{d-1} & \text{if } i = j \end{cases}
\]

which implies that for \( \forall v \in \mathbb{R}^n, \ v \neq 0, \ \text{we have } v^T \nabla_{yy}^2 L(y^0, \theta^0, \lambda, \mu) v > 0. \)

**CR** Let \( I_0 := \{i | g_i(y^0, \theta^0) = 0\} \), \( I \) is arbitrary but fixed subset \( I \subseteq I_0 \). Note that, \( \sum_{i \in [n]} 1(g_i(y^0, \theta^0) = 0) \leq n - 1 \), therefore the set of gradients \( \{\nabla_y g_i(y^0, \theta^0), \ i \in I\} \cup \{\nabla_y h(y^0, \theta^0)\} \) has constant rank \( |I| + 1 \), for all \( (y, \theta) \) in open neighborhood of \( (y^0, \theta^0) \).

**H. Proof of Theorem 7**

Following (5), it suffices to show that \( \lim_{k \to \infty} \theta(m(k)) = 0, \) almost surely.
Lemma 1 shows that there exists a convergent subsequence \{m(k_s)\}_s with limit \( m \leq 0 \), almost surely. That is, there exists \( \hat{s}_1 > 0 \) such that
\[
|m(k_s) - m| \leq -m/3 \quad \forall s > \hat{s}_1 \quad \text{a.s.} \tag{24}
\]

We now show that all the terms of the parent sequence \( m(k) \) between two consecutive subsequence terms \( m(k_s) \) and \( m(k_{s+1}) \), for \( s > \hat{s}_1 \), are negative. Proposition 3 implies, for all \( s > \hat{s}_1 \),
\[
m(k_s + 1) = \frac{k_s}{k_s + 1} m(k_s) + \frac{1}{k_s + 1} u_0(k_s) + \frac{1}{k_s + 1} \triangle u(k_s)
\]
i.e.,
\[
m(k_s + 1) - m = \frac{k_s}{k_s + 1} (m(k_s) - m) + \frac{1}{k_s + 1} (u_0(k_s) - m) + \frac{1}{k_s + 1} \triangle u(k_s)
\]
where
\[
u_0(k_s) := \pi_{\omega(k_s)}^T (I - P) \left( \alpha_{\omega(k_s)} + \sum_{d=1}^D \text{diag}(\alpha_{d,\omega(k_s)}) \text{diag}(\nu \pi_{\omega(k_s)} + y(0))^{d-1} (\nu \pi_{\omega(k_s)} + y(0)) \right)
\]
be the instantaneous mistrust assuming B-agents estimate that P-agents are obedient, i.e., assuming \( \hat{\theta}(k_s) = 0 \), and
\[
\triangle u(k_s) := u(k_s) - u_0(k_s)
\]
\[
= \pi_{\omega(k_s)}^T (I - P) \sum_{d=1}^D \text{diag}(\alpha_{d,\omega(k_s)}) \left( \text{diag}(x(k_s) + y(\hat{\theta}(k_s)))^{d-1} (x(k_s) + y(\hat{\theta}(k_s))) - \text{diag}(\nu \pi_{\omega(k_s)} + y(0))^{d-1} (\nu \pi_{\omega(k_s)} + y(0)) \right)
\]
\[
= \pi_{\omega(k_s)}^T (I - P) \sum_{d=1}^D \text{diag}(\alpha_{d,\omega(k_s)}) \sum_{t=0}^{d-1} \text{diag}(x(k_s) + y(\hat{\theta}(k_s)))^{d-1-t} \text{diag}(\nu \pi_{\omega(k_s)} + y(0))^{t} (y(\hat{\theta}(k_s)) - y(0))
\]
\[
\tag{26}
\]
In the above equations, we utilize (24) to get \( m(k_s) \leq 0 \), and hence \( \theta(m(k_s)) = 0 \), for \( s > \hat{s}_1 \). Using the Lipschitz property of \( y(\cdot) \) (cf. Remark 4) with (26) gives
\[
|\triangle u(k_s)| \leq L|\hat{\theta}(k_s)|, \quad s > \hat{s}_1
\]
for some \( L > 0 \). Proposition 3 implies that there exists \( \hat{k}_1 > 0 \) such that \( |\theta(m(k)) - \hat{\theta}(k)| \leq -m/3L \) for all \( k > \hat{k}_1 \).

Let \( \hat{s}_2 \) be the smallest integer greater than \( \hat{s}_1 \) such that \( k_{\hat{s}_2} > \hat{k}_1 \). Therefore, \( |\theta(m(k_s)) - \hat{\theta}(k_s)| = |\hat{\theta}(k_s)| \leq -m/3L \) for all \( s > \hat{s}_2 \). Combining with (27) gives
\[
|\triangle u(k_s)| \leq -\frac{m}{3}, \quad s > \hat{s}_2
\]
for some \( L > 0 \). Strong law of large numbers gives \( \lim_{\tau \to \infty} \frac{\sum_{t=1}^\tau u_0(t)}{\tau} = m \) almost surely, i.e., there exists \( \hat{k}_2 \) such that \( \left| \frac{\sum_{t=1}^\tau u_0(t)}{\tau} - m \right| \leq -m/6 \) for all \( \tau > \hat{k}_2 \). Therefore, for all \( \tau > \hat{k}_2 \) and a non-negative integer \( r \),
\[
\tau \frac{7m}{6} \leq \sum_{t=1}^{\tau} u_0(t) \leq (\tau + 1) \frac{5m}{6} \quad (\tau + 1 + r) \frac{7m}{6} \leq \sum_{t=1}^{\tau+1+r} u_0(t) \leq (\tau + 1 + r) \frac{5m}{6}
\]
Subtraction gives

\[ | \sum_{t=\tau+1}^{\tau+1+r} (u_0(t) - m) | \leq -(2\tau + r + 1) \frac{m}{6}, \quad \tau > \tilde{k}_2, r \geq 0 \] (29)

Let \( \tilde{s} \) be the smallest integer greater than \( \tilde{s}_2 \) such that \( k_2 > \tilde{k}_2 \). Using (24), (28) and (29) \((\tau = k_s - 1, r = 0)\) in (25) gives, for all \( s > \tilde{s} \),

\[ |m(k_s + 1) - m| \leq - \frac{k_s m}{k_s + 1} - \frac{(2k_s - 1) m}{k_s + 1} - \frac{1}{k_s + 1} \frac{m}{3} < - \frac{2m}{3} \]

which implies that \( m(k_s + 1) < \frac{m}{3} \leq 0 \).

Hereafter, we show that \( m(k_s + 2), m(k_s + 3), \ldots m(k_{s+1} - 1) \) are all negative, for \( s > \tilde{s} \), using induction. Let \( m(t) < 0 \) for all \( t \in \{k_s + 1, \ldots, k - 1\} \), for some \( k \in \{k_s + 2, \ldots, k_{s+1} - 1\} \). Repeated application of (25) gives

\[ m(k) - m = \frac{k_s}{k} (m(k_s) - m) + \frac{1}{k} \sum_{t=k_s}^{k-1} (u_0(t) - m) + \frac{1}{k} \sum_{t=k_s}^{k-1} \triangle u(t) \] (30)

Along similar lines as (27), one can show that \( |\triangle u(t)| \leq - \frac{m}{3} \) for all \( t \in \{k_s, \ldots, k - 1\} \). Using this, along with (24) and (29) \((\tau = k_s - 1, r = k - k_s - 1)\), in (30) gives

\[ |m(k) - m| \leq - \frac{k_s m}{k} - \frac{(k + k_s - 2) m}{k} - \frac{(k - k_s) m}{k} < - \frac{2m}{3} \]

which implies that \( m(k) < \frac{m}{3} \leq 0 \).