Online E-companion to:
"An Integrated Decomposition and Approximate Dynamic Programming Approach for On-demand Ride Pooling"

Xian Yu and Siqian Shen*

Abstract

Through smartphone apps, drivers and passengers can dynamically enter and leave ride-hailing platforms. As a result, ride-pooling is challenging due to complex system dynamics and different objectives of multiple stakeholders. In this paper, we study ride-pooling with no more than two passenger groups who can share rides in the same vehicle. We dynamically match available drivers to randomly arriving passengers and also decide pick-up and drop-off routes. The goal is to minimize a weighted sum of passengers' waiting time and trip delay time. A spatial-and-temporal decomposition heuristic is applied and each subproblem is solved using Approximate Dynamic Programming (ADP), for which we show properties of the approximated value function at each stage. Our model is benchmarked with the one that optimizes vehicle dispatch without ride-pooling and the one that matches current drivers and passengers without demand forecasting. Using test instances generated based on the New York City taxi data during one peak hour, we conduct computational studies and sensitivity analysis to show (i) empirical convergence of ADP, (ii) benefit of ride-pooling, and (iii) value of future supply-demand information.

This online companion provides proofs of some results omitted from the main body of the paper. We also report some additional computational results. The numbered references and citations correspond to those in the main paper, and all new expressions, results, figures and tables are numbered contiguously following those in the main paper.

Proof of Theorem 1

Theorem 1 *Let $\preceq$ be the generalized component-wise inequality over all dimensions of the state space. The optimal value function is monotone based on (19) and (20).*

Proof: Suppose that $R_t \leq R'_t,\ D_t = D'_t$. According to Proposition 3.1 in Jiang and Powell (2015), to show $V_t(S_t) \leq V_t(S'_t)$, we only need to verify

*Xian Yu and Siqian Shen are with Department of Industrial and Operations Engineering at the University of Michigan in Ann Arbor, USA. `yuxian@umich.edu, siqian@umich.edu`

(1) For every $S_t,\ S_t'$ with $S_t \preceq S_t',\ \mathbf{x}_t \in \mathcal{D}$ and W_{t+1}, the state transition function f (i.e., the composition of $S^{M,W}$ and $S^{M,x}$) satisfies

$$f(S_t, \mathbf{x}_t, W_{t+1}) \preceq f(S_t', \mathbf{x}_t, W_{t+1}).$$

(2) For every $t < T$, $S_t,\ S_t'$ with $S_t \preceq S_t',\ \mathbf{x}_t \in \mathcal{D}$,

$$C_t(S_t, \mathbf{x}_t) \leq C_t(S_t', \mathbf{x}_t),\ C_T(S_T) \leq C_T(S_T').$$

(3) For each $t < T$, R_t and W_{t+1} are independent.

First of all, (3) is true due to our assumption on W_{t+1}. Because $R_{ta} \leq R_{ta}'$, state S_t' has $R_{ta}' - R_{ta}$ more drivers than state S_t. To have $\mathbf{x}_t$ (defined for state S_t) also be feasible under state S_t' (i.e., it should satisfy the equality constraints (5)), we extend decision variable $\mathbf{x}_t$ to $\tilde{\mathbf{x}}_t = (\tilde{x}_{tad})_{a\in\mathcal{A}_t, d\in\mathcal{D}}$ for state S_t' as follows: for every $a \in \mathcal{A}_t$,

$$\tilde{x}_{tad} = \begin{cases} x_{tab}, & \text{if} \quad d = b \in \mathcal{B}_t \\ x_{tad} + R_{ta}' - R_{ta}, & \text{if} \quad d = \emptyset \\ 0, & \text{otherwise,} \end{cases}$$

which means that $\tilde{\mathbf{x}}_t$ is the same as $\mathbf{x}_t$ for the number of drivers within R_{ta} and the extended decision is to hold the extra $R_{ta}' - R_{ta}$ drivers (i.e., to assign no passenger to them).

We let

$$S_{t+1} = f(S_t, \mathbf{x}_t, W_{t+1}) = (R_{t+1}, D_{t+1})$$
$$S_{t+1}' = f(S_t', \tilde{\mathbf{x}}_t, W_{t+1}) = (R_{t+1}', D_{t+1}').$$

Since only the values of W_{t+1} determine D_{t+1}, we have $D_{t+1} = D_{t+1}'$.

According to (8),

$$R_{ta'}^x = \sum_{a\in\mathcal{A}_t} \sum_{d\in\mathcal{D}} \delta_{a'}(a, \mathbf{x}_t) x_{tad},\ \forall a' \in \mathcal{A}_{t,x}$$
$$R_{ta'}^x{}' = \sum_{a\in\mathcal{A}_t} \sum_{d\in\mathcal{D}} \delta_{a'}(a, \tilde{\mathbf{x}}_t) \tilde{x}_{tad},\ \forall a' \in \mathcal{A}_{t,\tilde{x}}.$$

Because $x_{tad} \leq \tilde{x}_{tad}$, we have $R_{ta'}^x \leq R_{ta'}^x{}'$, and furthermore $R_{t+1} \leq R_{t+1}'$ due to the monotonicity of transition function $a^{M,W}$.

Therefore, $S_{t+1} \preceq S_{t+1'}$, which shows the monotonicity of transition function f.

On the other hand, because the reward function $C_t(S_t, \mathbf{x}_t)$ is determined by $(x_{tab})_{a\in\mathcal{A}_t, b\in\mathcal{B}_t}$, while $(x_{ta\emptyset})_{a\in\mathcal{A}_t}$ does not contribute to $C_t(S_t, \mathbf{x}_t)$, we have $C_t(S_t, \mathbf{x}_t) = C_t(S_t', \tilde{\mathbf{x}}_t)$. This completes the proof.

Proof of Lemma 1

Lemma 1 *1. If $a \notin \mathcal{A}_t^0 \cup \mathcal{A}_t^1$ or $a \in \mathcal{A}_t^0 \cup \mathcal{A}_t^1,\ (a, b) \notin S,\ \forall b \in \mathcal{B}_t$, then $\nu_a = \gamma v_{a\emptyset}^n$.*

2. If $a \in \mathcal{A}_t^0$, then $\nu_a \geq \min_{b:(a,b)\in S} \{\lambda_1 w_{ab} - P + \gamma v_{ab}^n\}$;

3. *if* $a \in \mathcal{A}_t^1$*, then* $\nu_a \geq \min_{b:(a,b)\in S} \{\lambda_1 w_{ab} + \lambda_2 d_{ab} - P + \gamma v_{ab}^n\}$*.*

Proof:

1. If $a \notin \mathcal{A}_t^0 \cup \mathcal{A}_t^1$, then the constraints imposed on a are (25)–(27). Since M is a large positive number, constraint (25) is redundant. To maximize the objective, we must have $\nu_a = \gamma v_{a\emptyset}^n$. The case of $a \in \mathcal{A}_t^0 \cup \mathcal{A}_t^1,\ (a,b) \notin S,\ \forall b \in \mathcal{B}_t$ is similar and we omit the details here.
   Alternatively, we can show this fact from another perspective. If $a \notin \mathcal{A}_t^0 \cup \mathcal{A}_t^1$ or $a \in \mathcal{A}_t^0 \cup \mathcal{A}_t^1,\ (a,b) \notin S,\ \forall b \in \mathcal{B}_t$, then $x_{ta\emptyset} = 1$ because driver a is unavailable and we impose a large penalty on unavailable drivers. Then by complementary slackness, we have $\nu_a = \gamma v_{a\emptyset}^n$.

2. If $a \in \mathcal{A}_t^0$, then the constraints imposed on a are (22), (24), (26) and (27). Since N is a large positive number, constraint (24) is redundant, and we have

$$\nu_a \leq \min_{b:(a,b)\in S} \{\lambda_1 w_{ab} - P + \gamma v_{ab}^n - \mu_b\},\ \forall a \in \mathcal{A}_t^0.$$

   Because P is a large positive number, (22) implies (26). To maximize the objective, we have

$$\nu_a = \min_{b:(a,b)\in S} \{\lambda_1 w_{ab} - P + \gamma v_{ab}^n - \mu_b\},\ \forall a \in \mathcal{A}_t^0.$$

   As $\mu_b \leq 0$, we can further derive

$$\nu_a \geq \min_{b:(a,b)\in S} \{\lambda_1 w_{ab} - P + \gamma v_{ab}^n\}.$$

   Specifically, when there are unsatisfied demand, i.e., $\sum_{a\in\mathcal{A}_t} x_{tab} < D_{tb},\ \forall b \in \mathcal{B}_t$, by complementary slackness, $\mu_b = 0$. In this case,

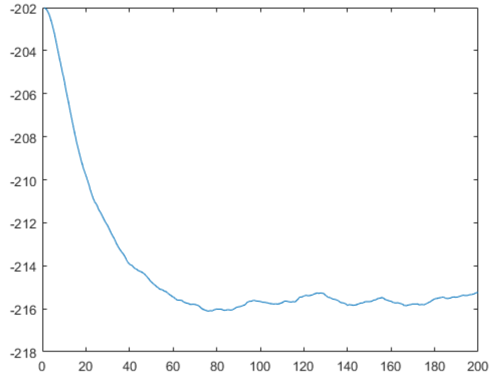
$$\nu_a = \min_{b:(a,b)\in S} \{\lambda_1 w_{ab} - P + \gamma v_{ab}^n\},\ \forall a \in \mathcal{A}_t^0.$$

3. If $a \in \mathcal{A}_t^1$, the constraints on a are (23), (24), (26) and (27). The rest of the analysis is similar to the one for $a \in \mathcal{A}_t^0$ and we omit the details. This completes the proof.

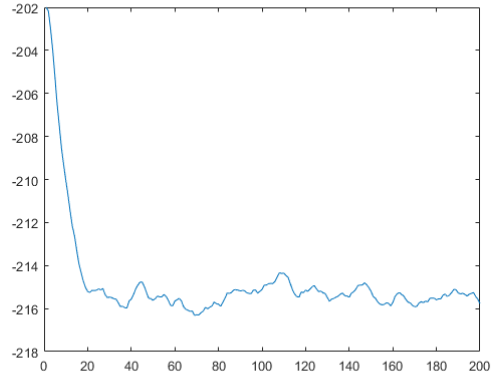
Parameter Configuration

We show the empirical performance of ADP on randomly generated instances with different parameter settings. We set the number of stages $T = 4$, and randomly distribute ten drivers on a 10×10 grid network following a uniform distribution. In every stage, we randomly generate five potential passenger origin-destination (O-D) pairs. We set penalty parameter $M = 1000,\ N = 1000,\ P = 500$ and weight $\lambda = (0.2, 0.8)$.

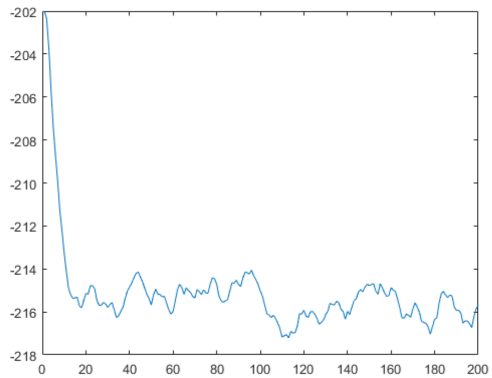
We first fix $\gamma = 0.9$ and vary $\alpha_n = 0.1, 0.2, 0.3, 1/n$ to depict the results in Figure 1, where n is the index of iteration. We observe that larger α-values lead to faster convergence but less stable objective values through iterations. When $\alpha_n = 1/n$, the algorithm converges fast and maintains stable performance.



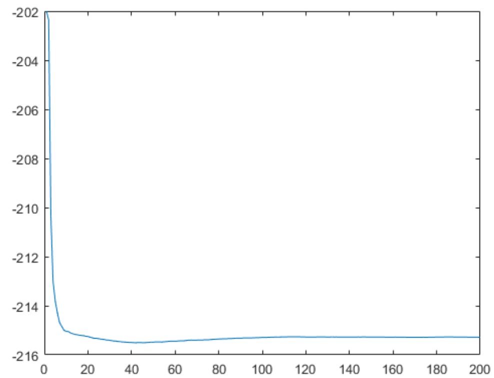
(a) $\alpha_n = 0.1$



(b) $\alpha_n = 0.2$



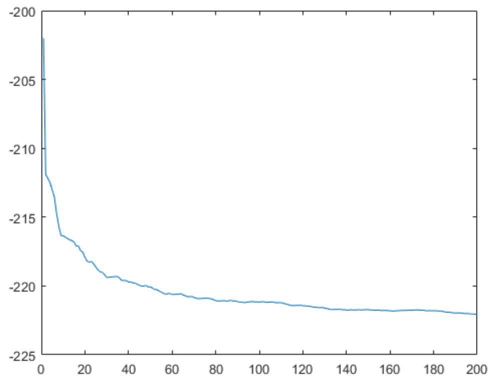
(c) $\alpha_n = 0.4$



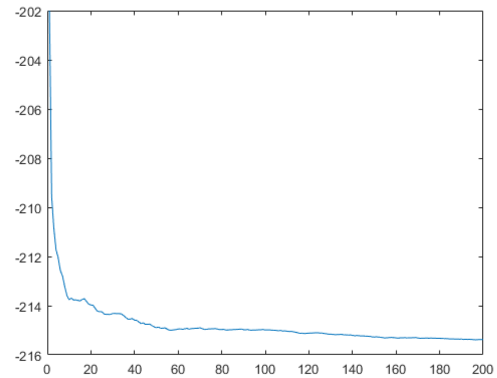
(d) $\alpha_n = 1/n$

Figure 1: Convergence performance of ADP under different α-values while x-axis represents the number of iterations and y-axis represents the objective value $F_0(S_0)$.

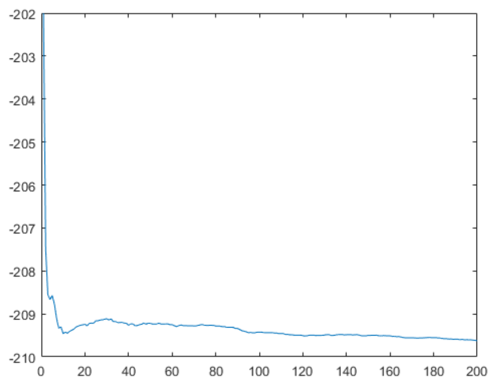
Next, we examine the effect of discount factor γ and fix $\alpha_n = 1/n$. We vary $\gamma = 0.3, 0.5, 0.7, 0.9$. From Figure 2, we observe that the value of discount factor γ could affect the optimal objective value. As γ represents the importance of future information when making decisions, we will fix $\gamma = 0.9$ in our later texts.



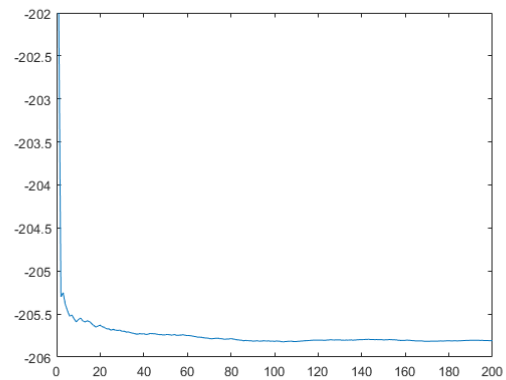
(a) $\gamma = 0.9$



(b) $\gamma = 0.7$



(c) $\gamma = 0.5$



(d) $\gamma = 0.3$

Figure 2: Convergence performance under different γ-values while x-axis represents the number of iterations and y-axis represents the objective value $F_0(S_0)$.

References

Jiang, D. R. and Powell, W. B. (2015). An approximate dynamic programming algorithm for monotone value functions. *Operations Research*, 63(6):1489–1511.