Designing Ergodic Markov Chains for Sampling Household Schedules

Masahiro Araki*[1]

[1]Researcher, Toyota Motor Corporation, Japan

Short summary

In activity-based models with household-level schedule choice, estimating discrete choice models requires realistic alternative schedules beyond those observed in finite samples. We propose a Metropolis–Hastings sampling algorithm for household schedules, grounded in Markov chain Monte Carlo theory. We define a finite schedule state space and a parsimonious set of elementary editing operations that jointly respect (i) spatiotemporal continuity, (ii) household-specific time-window constraints, (iii) asymmetry between licensed drivers and passengers in pick-up/drop-off behavior, and (iv) limits on the number of privately owned vehicles. Using tools from stochastic process theory, we prove that the Markov chain induced by these operations is irreducible. In our finite-state setting, together with the self-loop property of the Metropolis–Hastings kernel, this implies aperiodicity and positive recurrence, thereby establishing ergodicity. The framework provides a rigorous foundation for importance sampling of household schedules in OASIS-type models, enabling efficient estimation and robust policy analysis.
Keywords: Household schedule, Choice set generation, Markov chain Monte Carlo, Ergodicity proof

1 Introduction

Econometric activity-based models (ABMs) remain essential tools for transportation and urban policy analysis, as they offer interpretability and strong out-of-distribution (OOD) predictive performance (Hess & van Cranenburgh (2025)). A recent development in ABM frameworks is the OASIS approach (Pougala et al. (2023); Rezvany et al. (2023)), which provides an integrated discrete choice model representing entire schedules of individuals and households. When estimating the parameters of such discrete choice models, it is necessary to generate realistic alternative schedules that go beyond those actually observed in the data. In this study, we aim to provide theoretical guarantees for an algorithm that samples alternative household-level schedules. Figure 1 illustrates the problem addressed in this work.

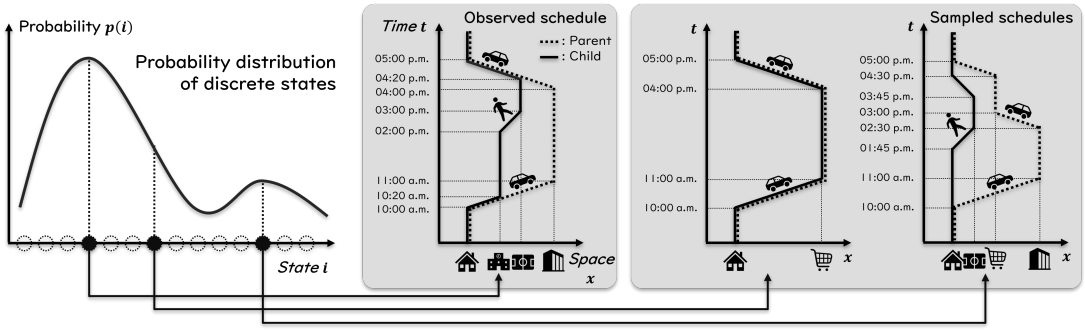


Figure 1: Conceptual diagram of the problem being addressed.

McFadden et al. (1977) showed that, even when sampled alternatives are used in the multinomial logit (MNL) model, the asymptotic properties of the estimated parameters—such as asymptotic unbiasedness and efficiency—are preserved if McFadden's correction factor is applied. More recently, Guevara & Ben-Akiva (2013a,b) demonstrated that similar asymptotic properties also hold for logit mixture models and multivariate extreme value (MEV) models. However, these results are

asymptotic. With finite datasets, empirical evidence indicates that importance sampling can substantially reduce estimator bias and variance compared to simple random sampling (Lemp & Kockelman (2012)). Because transportation and urban policy decisions rely on finite data to estimate model parameters and evaluate policies, sampling schemes must perform well in the finite-sample regime. In individual- or household-level schedule models, the number of alternative schedules grows combinatorially, making importance sampling particularly attractive. Unlike destination choice models, however, it is not straightforward to identify intuitive features for constructing importance weights. For this reason, several algorithms based on the Metropolis–Hastings (MH) method—a type of Markov chain Monte Carlo (MCMC) method—have recently been proposed (Pougala et al. (2021); Rezvany et al. (2024)).

Care is required when designing algorithms based on MCMC methods. By definition, general MCMC methods, including the MH method, must satisfy both the balance condition and the ergodic condition (Gelman et al., 2013, Chapter 11). Among these, satisfying the balance condition is relatively straightforward (Araki (2025)), whereas ensuring ergodicity is more challenging because schedules in physical space have specific structural properties. If the states and transitions for MCMC are not defined appropriately, the algorithm may fail to explore all feasible schedules. At the individual level, schedules must exhibit temporal and spatial continuity and remain within designated time windows. At the household level, additional constraints arise: licensed drivers may pick up and drop off unlicensed members, but not vice versa, and the household cannot use more vehicles than it owns.

Therefore, in this study, we return to the fundamentals of MCMC and build on Araki (2025), which established theoretical validity for individual-level schedule sampling. We propose an MH method for household-level schedules that comes with theoretical guarantees. Specifically, we define a finite schedule state space and a parsimonious set of operations that allow transitions only between feasible states, and we prove that the resulting Markov chain is ergodic in our finite-state setting. Ergodicity consists of irreducibility, aperiodicity, and positive recurrence; due to space limitations, this paper focuses on proving irreducibility, which is the most important component. Figure 2 summarizes the positioning of this study.

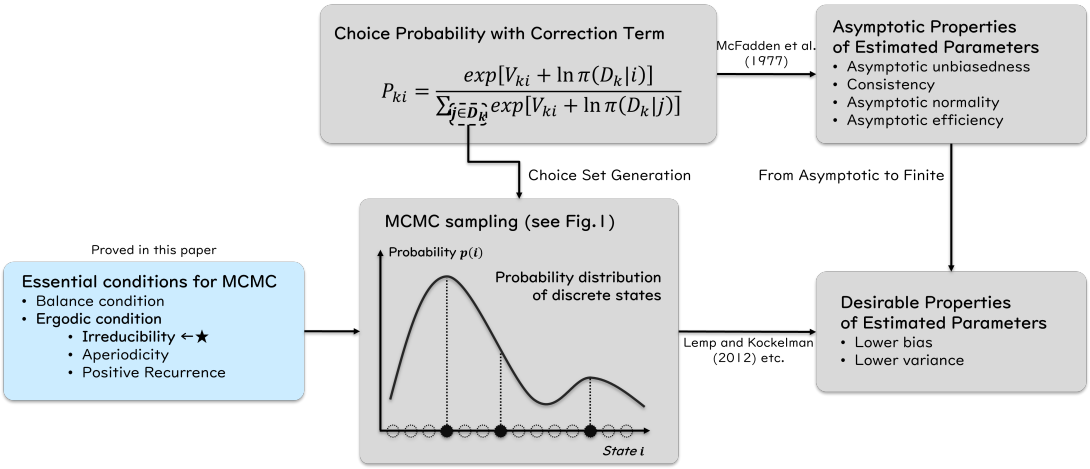


Figure 2: Diagram illustrating the positioning of this study

The contributions of this research are summarized as follows:

- We define an appropriate state space and a set of essential operations for household-level schedules that exhibit the following key features:
  - spatiotemporal continuity of each household member's schedule,
  - each schedule must fit within the designated time window,
  - asymmetry between licensed and unlicensed members in vehicle use and pick-up/drop-off behavior,
  - resource constraints due to the limited number of household vehicles.
- We prove that, using the proposed operations, the resulting Markov chain is ergodic in our finite-state setting.

2 Methodology

Definition of the State Space

The purpose of this study is to sample feasible household-level schedules. We treat the joint schedule of all household members as a discrete state and make it the target of sampling. A household consists of n_{pd} licensed members (potential drivers) and n_{pp} unlicensed members (potential passengers), for a total of $n(= n_{pd} + n_{pp})$ members, and is assumed to own m private vehicles. The state space is defined over the time, location (activity type), and mode of transport for each household member, as well as the time and location of each household vehicle.
The analysis period is specified by the analyst and divided into a finite number of discrete time units. Locations (activity types) and modes of transport are also assumed to be discrete and finite. Consequently, each individual has only a finite number of possible sequences of activities, locations, and modes of transport over the set of time intervals. Subject to the household-specific constraints described below, the household-level state space is given by the Cartesian product of these individual trajectories and the finite set of vehicle trajectories.
We impose the following household-specific constraints. First, if the number of licensed drivers operating household vehicles at a given time is n_d , then $n_d \leq n_{pd}$ and $n_d \leq m$ must hold. Second, whenever an unlicensed member is traveling by car, their spatiotemporal trajectory must fully overlap with that of a licensed driver operating a household vehicle; this corresponds to being picked up and dropped off by a household driver. Third, once a licensed driver leaves home operating a particular vehicle, they cannot drive a different vehicle until they return that vehicle home. By contrast, as passengers, members may ride in multiple household vehicles within a single home-based roundtrip. For example when a preschool child is driven from home to daycare by the father in the morning and back home by the mother later in the day.

Metropolis–Hastings Method

Following Pougala et al. (2021, 2023), we use a Metropolis–Hastings (MH) algorithm to sample household-level schedules (Figure 3). The target distribution over feasible household-level schedules X is taken to be proportional to the choice probability $P(X)(> 0)$ implied by the underlying discrete choice model. As illustrated in Figure 3, the proposal kernel $q(X, X')$ is constructed by selecting one of the nine elementary operations ω (introduced in the next section) with probability $P_\omega(> 0)$, and then choosing the internal parameters of that operation (e.g., activity, time interval, mode of transport) according to a prescribed rule. Since we only consider transitions between feasible states, $q(x, y)(> 0)$ holds whenever a transition from x to y is allowed.
MCMC methods generate a Markov chain whose stationary distribution equals a given probability distribution, thereby enabling sampling from that distribution. The MH algorithm is valid only when both the balance condition and the ergodic condition are satisfied (Ross (1995)). As shown in Araki (2025), it is straightforward to verify that the MH algorithm satisfies the balance condition, so in this paper we concentrate on establishing ergodicity. Issues related to the practical mixing behavior of the chain and the tuning of the proposal kernel are beyond the scope of this short paper and are left for future work.

Algorithm

$t \leftarrow 0$, initialise state with random schedule $X_t \leftarrow S_0$
Initialise utility function with random parameters $\widetilde{U_S}$
for $t = 1, 2, \ldots$ **do**
  Choose operation ω with probability P_ω
  $X^*, q(X_t, X^*) \leftarrow$ **ApplyChange**(ω, X_t)
    function **ApplyChange**$(\omega, state\ X)$
      return new state X', transition probability $q(X, X')$
  Compute target weight $p(X^*) = \widetilde{U_S}(X^*)$
  Compute acceptance probability $\alpha(X_t, X^*) = min\left(1, \frac{p(X^*)q(X_t|X^*)}{p(X_t)q(X^*|X_t)}\right)$
  With probability $\alpha(X_t, X^*)$, $X_{t+1} \leftarrow X^*$, else $X_{t+1} \leftarrow X_t$

Figure 3: The Metropolis–Hastings method for ABMs proposed by Pougala et al. (2021)

List of Essential Operations

To obtain an ergodic Markov chain $q(X, X')$, we propose nine elementary state-transition operations ω , depicted in Figure 4. Each operation maps feasible states to feasible states and can be applied only if all household-level constraints (spatiotemporal continuity, time windows, licensing, and vehicle availability) are satisfied. For collective operations, the term "spatiotemporal range" denotes the time intervals and locations in which a driver and accompanying passengers share the same origin, destination, and mode of transport, so that they travel together in the same vehicle.

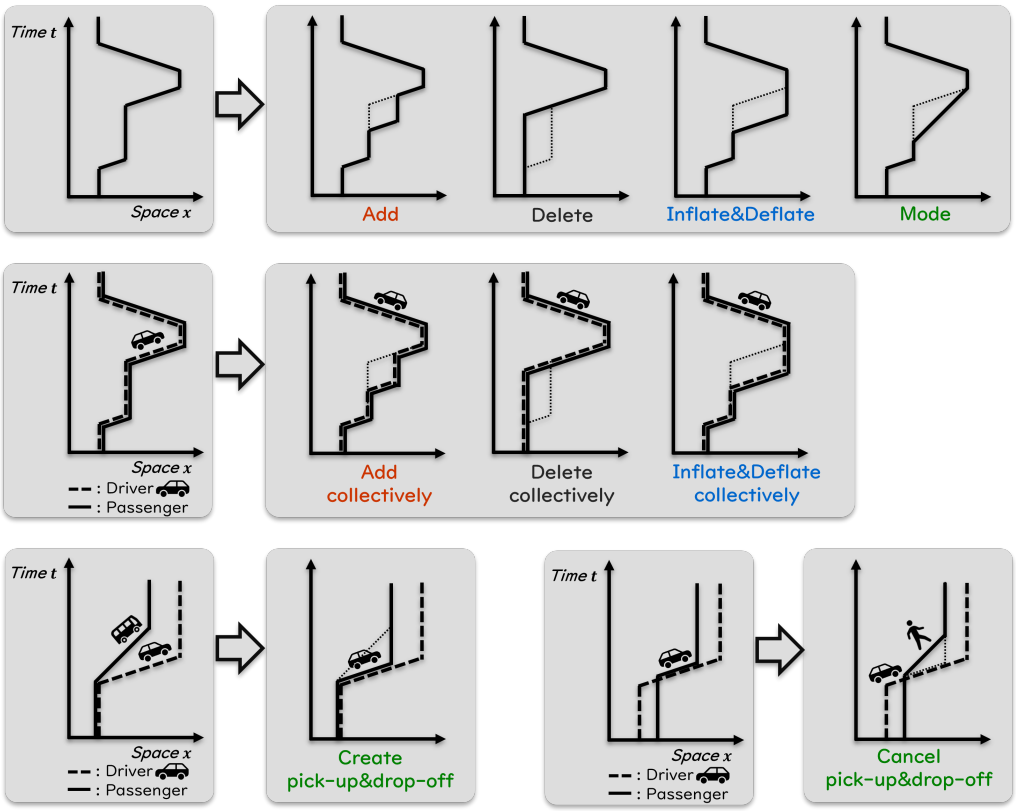


Figure 4: The nine essential state-transition operations proposed in this study, each defined to induce transitions only between feasible states.

A brief explanation of each operation is as follows:

- **Add**: Insert an activity for a household member at a given location.
- **Delete**: Remove an activity for a household member at a given location.
- **Inflate & Deflate**: Extend or shorten the duration of an activity for a household member by an integer multiple of the unit time.
- **Mode**: Change the mode of transport used for a specific trip by a household member.
- **Add collectively**: Insert an activity for a driver and accompanying members within a shared spatiotemporal range.
- **Delete collectively**: Remove an activity for a driver and accompanying members within a shared spatiotemporal range.
- **Inflate & Deflate collectively**: Extend or shorten the duration of an activity for a driver and accompanying members within a shared spatiotemporal range by an integer multiple of the unit time.
- **Create pick-up & drop-off**: When a member is traveling by a mode other than a household vehicle, arrange for a licensed driver to pick up and drop off that member.
- **Cancel pick-up & drop-off**: For a member currently being picked up and dropped off by a driver within a shared spatiotemporal range, cancel the pick-up and drop-off and switch to another mode of transport.

3 Proof

Verification of Ergodic Properties

Following Ross (1995) (Theorem 4.3.3), we call a Markov chain ergodic if it is irreducible and aperiodic and all states are positive recurrent; in this case, it has a unique stationary distribution.

Theorem 1 *THEOREM 4.3.3 in Ross (1995)*
An irreducible aperiodic Markov chain belongs to one of the following two classes
(i) Either the states are all transient or all null recurrent, in this case, $P_{ij}^n (= P(X_{n+m} = j \mid X_m = i)) \to 0$ *as* $n \to \infty$ *for all* i, j *and there exists no stationary distribution*
(ii) Or else, all states are positive recurrent, that is,

$$\pi_j = \lim_{n\to\infty} P_{ij}^n > 0$$

In this case, $\{\pi_j, j = 0, 1, 2, \ldots\}$ *is a stationary distribution and there exists no other stationary distribution*

We are interested in establishing ergodicity of the Markov chain generated by the operations defined in the previous section. Following Ross (1995) (Theorem 4.3.3), this amounts to showing that the chain is irreducible, aperiodic, and that all states are positive recurrent. Due to space constraints, we focus on proving irreducibility, which is the most important and nontrivial component. In our setting, the state space is finite and the Metropolis–Hastings kernel has a positive probability of remaining in the current state. Hence, once irreducibility is established, aperiodicity and positive recurrence follow as immediate corollaries (see Ross (1995), Chapter 4), and ergodicity is guaranteed.

Proof of Irreducibility

Due to space limitations, we provide a constructive proof sketch rather than a fully formal proof. The state space S is said to be irreducible if it consists of a single class; that is, for every pair of states $s_i, s_j \in S$, state s_i and state s_j are mutually reachable (Durrett, 2016, Chapter 1.3). Let $T_y = \min\{n \geq 1 : X_n = y\}$ be the first hitting time of y (i.e., being at y at time 0 does not count). We adopt the notion of communication between states as in Definition 1.1 of Durrett (2016):

Definition 1 *Definition 1.1 of Durrett (2016)*
We say that x *communicates with* y *and write* $x \to y$ *if there is a positive probability of reaching* y *starting from* x*, that is, the probability*

$$\rho_{xy} = P_x(T_y < \infty) > 0$$

As illustrated in Figure 5, any household-level schedule corresponding to a discrete state $s_i \in S$ can be represented as a finite combination of four patterns of home-based roundtrips. Formally, we define a home-based roundtrip for a household member as a maximal contiguous subsequence of their daily schedule that starts at the home location, ends at the home location, and does not contain any intermediate visit to home. A household-level schedule (i.e., a state $s_i \in S$) can then be decomposed into a finite collection of such home-based roundtrips, one or several per household member. We also define, for each member, a trivial home-based roundtrip z in which the member stays at home during the period of interest, and denote by $s_z \in S$ the special household-level state in which all members stay at home throughout the analysis period, each following this trivial roundtrip.

- **Pattern 1**: A home-based roundtrip in which the household member does not use a private vehicle as a mode of transport.
- **Pattern 2**: A home-based roundtrip in which a licensed driver uses a specific private vehicle as a mode of transport, but no pick-up & drop-off occurs.
- **Pattern 3**: A home-based roundtrip in which a licensed driver uses a specific private vehicle as the mode of transport, and other household members are picked up and dropped off by that vehicle during their respective home-based roundtrips.

- **Pattern 4**: A home-based roundtrip in which multiple licensed drivers each use their own private vehicle as their mode of transport, and other household members transfer between those vehicles during pick-up and drop-off as part of their respective home-based roundtrips.

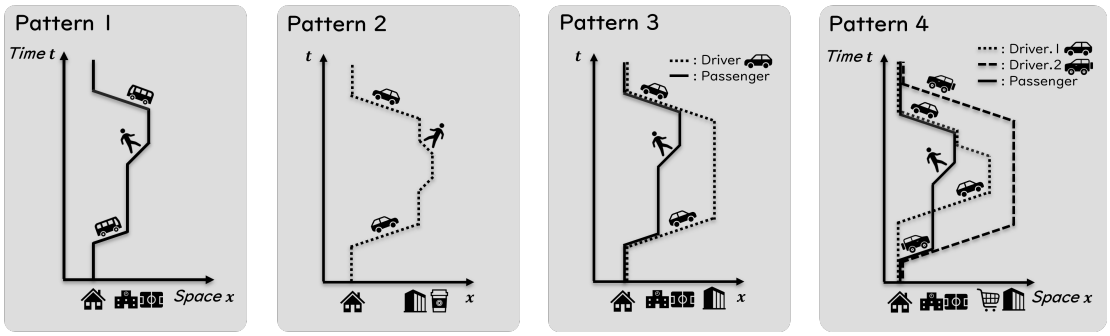


Figure 5: The four patterns of home-based roundtrips that make up household-level schedules $s_i (\in S)$

Here we rely on Lemma 1.4 of Durrett (2016):

Lemma 1 *Lemma 1.4. of Durrett (2016)*
If $x \to y$ and $y \to z$, then $x \to z$

With the above in mind, we focus on the trivial home-based roundtrip z and the special state s_z. If we can show that the four patterns of home-based roundtrips and the trivial roundtrip z are mutually reachable through a finite sequence of operations, then—since, as illustrated in Figure 5 and argued above, any household-level state $s_i, s_j \in S$ can be represented as a finite combination of these four patterns—any state s_i can reach any state s_j via the special state s_z . By Lemma 1, this implies that the state space S is irreducible.
First, consider Patterns 1 and 2. The main outline of the proof for these cases is illustrated in Figure 6. To begin with, any home-based roundtrip i consists of a finite number of activities. Therefore, by applying a finite number of Delete operations, it can reach home-based roundtrip z. Conversely, first, any home-based roundtrip j contains a finite number of activities, so by repeatedly applying the Add operation to home-based roundtrip z, we can obtain a state that contains all activities included in home-based roundtrip j. Second, by repeating the Mode operation a finite number of times, we can reach a state in which not only all activities but also all movements between them match those in home-based roundtrip j. Finally, by applying Inflate&Deflate operations a finite number of times, we can reach home-based roundtrip j.

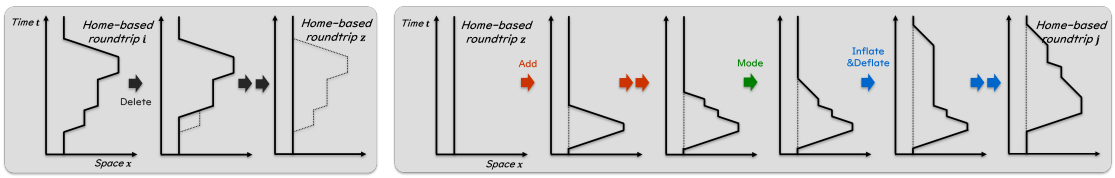


Figure 6: Overview of the approach for patterns 1 and 2 in Figure 5

Pattern 3 can be further divided into two cases. First, in Pattern 3-1 (Figure 7), consider situations that involve only short-distance travel, where the accompanying member does not need to be picked up and dropped off to complete the home-based roundtrip. In this case, any home-based roundtrip i contains a finite number of pick-up & drop-off events, and by applying Cancel pick-up & drop-off operations a finite number of times, it can reach a home-based roundtrip y with no pick-up & drop-off. This can be classified as Pattern 1 or 2 for each individual; by following the same steps used for Patterns 1 and 2, one can reach the home-based roundtrip z. The converse is also true.

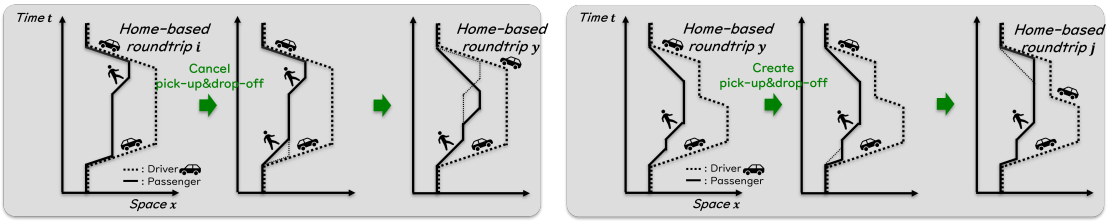


Figure 7: Overview of the approach for the cases within Pattern 3 of Figure 5 that consist solely of short-distance movements

Next, in Pattern 3-2 (Figure 8), consider cases where long-distance travel is involved and the accompanying member cannot complete the home-based roundtrip without being picked up and dropped off. A concrete example is a child taking a day trip by car with a parent to a remote mountain campsite; the child cannot complete the same home-based roundtrip alone. In this case, for any home-based roundtrip i, we modify the driver's schedule. Within a spatiotemporal range where the trajectories of the driver and a particular accompanying member do not overlap, the driver's schedule contains a finite number of activities and movements. By applying combinations of the proposed operations, we can make the spatiotemporal trajectories of the driver and that accompanying member completely overlap. Repeating this for all accompanying members, we reach a home-based roundtrip x in which the driver's trajectory completely overlaps with each accompanying member's trajectory. This home-based roundtrip x can be regarded as equivalent to Pattern 1 or 2 when the driver and all accompanying members are considered to be moving together. Then, by using the collective operations and following the same steps as in Pattern 1 or 2, we can reach the home-based roundtrip z. The converse also holds.

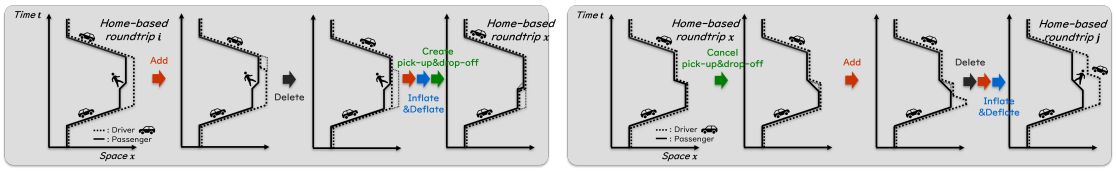


Figure 8: Overview of the approach for the cases within Pattern 3 of Figure 5 that include long-distance movements

Finally, in Pattern 4 (Figure 9), any home-based roundtrip i contains a finite number of pick-up & drop-off trajectories involving different vehicles. At this point, focus on a private vehicle from which all passengers have disembarked after completing transfers. By applying the proposed operations in an appropriate sequence to adjust that driver's trajectory to match the trajectory of the driver to whom they transferred, we can reach a home-based roundtrip w in which all pick-ups & drop-offs are handled by a single vehicle. This can be regarded as Pattern 3, and from there, following the same steps as in Pattern 3, the home-based roundtrip z can be reached. The converse also holds.

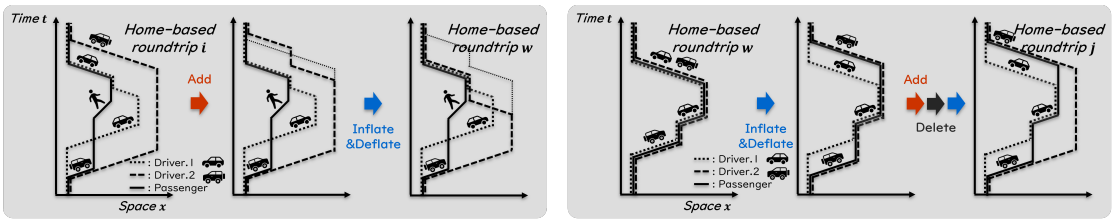


Figure 9: Overview of the approach for patterns 4 in Figure 5

In summary, since we have shown that the four patterns of home-based roundtrips and home-based roundtrip z are mutually reachable through a finite number of operations, any state s_i can reach any state s_j via the special state s_z. Therefore, according to the Lemma 1, the state space S is irreducible.

4 Conclusions

This study developed a Metropolis–Hastings sampling algorithm for the highly constrained state space of household-level schedules. We defined a compact state representation and a set of elementary operations that preserve spatiotemporal continuity of each household member, household-specific time-window constraints, asymmetries between licensed and unlicensed individuals in car-based travel and escorting, and constraints on the number of household vehicles. We showed that any feasible schedule can be reached by a finite sequence of these operations, implying irreducibility and, in our finite-state setting, ergodicity of the resulting Markov chain. This framework extends Araki (2025) from individual- to household-level schedules and provides a theoretical basis for generating alternative choice sets via importance sampling in OASIS-type activity-based models (Pougala et al. (2023); Rezvany et al. (2023)), thereby helping to reduce estimation bias and variance in finite samples.

Future work will address practical issues such as mixing and sampling efficiency, and implement the algorithm in large-scale household-level ABMs to evaluate its computational cost and the quality of generated schedules relative to existing sampling approaches. Another direction is to integrate the proposed sampler with household-level schedule models that combine discrete choice structures with deep learning components, aiming to retain interpretability while improving out-of-distribution predictive performance.

References

Araki, M. (2025). Metropolis-hastings sampling of individual schedules. In *Bridging transportation researchers (btr) conference.* Online. (Peer-reviewed international conference; no formal proceedings published)

Durrett, R. (2016). Essentials of stochastic processes. *Springer texts in statistics*.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis* (3rd ed.). Chapman and Hall/CRC.

Guevara, C. A., & Ben-Akiva, M. E. (2013a). Sampling of alternatives in logit mixture models. *Transportation Research Part B: Methodological*, *58*, 185–198.

Guevara, C. A., & Ben-Akiva, M. E. (2013b). Sampling of alternatives in multivariate extreme value (mev) models. *Transportation Research Part B: Methodological*, *48*, 31–52.

Hess, S., & van Cranenburgh, S. (2025). Combine and conquer: model averaging for out-of-distribution forecasting. *arXiv preprint arXiv:2506.03693*.

Lemp, J. D., & Kockelman, K. M. (2012). Strategic sampling for large choice sets in estimation and application. *Transportation Research Part A: Policy and Practice*, *46*(3), 602–613.

McFadden, D., Tye, W. B., & Train, K. (1977). *An application of diagnostic tests for the independence from irrelevant alternatives property of the multinomial logit model.* Institute of Transportation Studies, University of California Berkeley.

Pougala, J., Hillel, T., & Bierlaire, M. (2021). Choice set generation for activity-based models..

Pougala, J., Hillel, T., & Bierlaire, M. (2023). Oasis: Optimisation-based activity scheduling with integrated simultaneous choice dimensions. *Transportation Research Part C: Emerging Technologies*, *155*, 104291.

Rezvany, N., Bierlaire, M., & Hillel, T. (2023). Simulating intra-household interactions for in-and out-of-home activity scheduling. *Transportation Research Part C: Emerging Technologies*, *157*, 104362.

Rezvany, N., Hillel, T., & Bierlaire, M. (2024). Household-level choice-set generation and parameter estimation in activity-based models. In *24th swiss transport research conference (strc).*

Ross, S. M. (1995). *Stochastic processes.* John Wiley & Sons.