Particle Filter Data Assimilation for Reconstructing Latent Hospital Care Pathways under Partial Observability

Laith Ahmad[0009-0001-3892-7747] and Sergey V. Kovalchuk[0000-0001-8828-4615]

ITMO University, Saint Petersburg, Russia
475790@niuitmo.ru, kovalchuk@itmo.ru

Abstract. Hospital information systems record large volumes of clinical events, yet these records provide only a partial and asynchronous view of patient movement through care processes. Logged timestamps typically reflect administrative completion rather than true inter-department transitions, making direct reconstruction of care pathways unreliable.

We model daily ambulatory cardiology care pathways as partially observed stochastic dynamical systems. Department-level dynamics are represented using data-driven first- and second-order Markov transition models learned from multi-year hospital data (2015–2019), comprising 5,308 visit-level episodes across 51 departments. Latent department trajectories are reconstructed from sparse and noisy observations using a particle filter data assimilation framework that integrates structural transition priors with partial evidence. The transition model incorporates empirical termination behavior derived from observed episode statistics.

Reconstruction performance is evaluated via temporal holdout validation, training on 2015–2018 data and testing on 2019 episodes. Under representative sparse-observation conditions (stride 4, correctness 0.8), sequential assimilation improves reconstruction accuracy from 0.274 to 0.551 in-sample and from 0.289 to 0.518 out-of-sample. Gains persist across alternative sparsity and noise regimes, with larger improvements observed under increased observational degradation.

Complementary Bluetooth Low Energy localization experiments conducted on the same hospital floor provide a physical-layer perspective on observation uncertainty. Filtering mitigates stochastic sensor noise but does not correct systematic calibration bias, mirroring the department-level setting. Together, the results provide empirical evidence that healthcare pathways can be effectively modeled as partially observed dynamical systems and demonstrate that probabilistic data assimilation offers a principled and computationally tractable framework for reconstructing latent institutional trajectories under uncertainty.

Keywords: Healthcare systems modeling, Data assimilation, Particle filtering.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29915-4_12

1 Introduction

Many real-world systems are only partially observable: their underlying state evolves according to structured dynamics, while available measurements provide incomplete, delayed, or noisy information. Reconstructing latent trajectories from sparse observations is therefore a central problem in computational science, with established applications in geophysics, engineering, and biological systems [1; 2; 3]. In such settings, probabilistic state-space modeling provides a principled framework for integrating structural dynamics with uncertain data.

Organizational systems exhibit similar characteristics. Their evolution is governed by institutional rules, operational constraints, and resource allocation mechanisms, yet available data typically consist of irregular event logs rather than direct state measurements. Inferring structured dynamics from such event-level observations requires modeling under partial observability rather than direct sequence analysis.

Hospital care pathways provide a concrete and operationally critical example. Patients move between departments according to clinical workflows and capacity constraints, but hospital information systems (MIS) record only discrete medical events, often timestamped at administrative completion rather than at physical transition. Process mining approaches reconstruct and analyze observed activity sequences directly from event logs [4; 5]. While effective for workflow discovery and deviation analysis, such methods generally treat recorded events as complete representations of the process and do not explicitly reconstruct latent department-level trajectories evolving under hidden dynamics.

In contrast, this work treats department sequences as latent stochastic processes and focuses on reconstruction under partial and noisy observation. The contribution lies in formulating and evaluating a state-space approach that integrates transition modeling with probabilistic inference for healthcare pathways.

Stochastic modeling offers a complementary perspective. First-order Markov chains are widely used in healthcare pathway and decision modeling due to their interpretability and tractability [6]. Higher-order extensions allow transition probabilities to depend on recent history, relaxing the memoryless assumption and increasing expressive capacity [10]. However, Markov modeling alone describes transition tendencies and does not address inference under incomplete and noisy observation.

The reconstruction of latent trajectories from partial data is closely related to state-space modeling and sequential Bayesian filtering. The Kalman filter provides optimal inference for linear Gaussian systems [1], while Sequential Monte Carlo (SMC) methods—commonly referred to as particle filters—extend inference to nonlinear and non-Gaussian settings [3; 7]. These methods approximate posterior distributions using weighted particle ensembles and are widely applied in geosciences and tracking problems [2]. Their systematic application to institutional process reconstruction in healthcare remains comparatively limited.

A related line of research arises in indoor localization. RSSI-based Bluetooth Low Energy (BLE) fingerprinting methods provide fine-grained spatial observations but are sensitive to environmental calibration drift and measurement noise [8; 9]. Particle filtering has been used to stabilize localization trajectories under stochastic fluctuations,

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29915-4_12

though filtering cannot compensate for systematic bias in the observation model. These findings illustrate a broader computational principle: structural dynamical constraints combined with probabilistic assimilation improve robustness under partial observability.

In this work, we integrate these strands by modeling daily ambulatory cardiology care pathways as partially observed stochastic dynamical systems within an agent-based representation. Each patient is represented as an agent whose latent state corresponds to the current department. State evolution follows a data-driven first- or second-order Markov transition model estimated from multi-year hospital data. Latent trajectories are reconstructed using a particle filter that assimilates sparse and noisy observations into the learned structural prior.

Although MIS event data and BLE localization data are analyzed separately, they represent complementary observational layers of the same institutional system. MIS records provide semantic but indirect information about departmental transitions, while BLE localization captures physical trajectories subject to sensor noise and calibration variability. By examining both layers, the study demonstrates that sequential probabilistic assimilation stabilizes inference under stochastic observation noise across semantic and physical representations of hospital processes.

The contributions of this paper are threefold:

1. A formulation of hospital care pathways as partially observed stochastic dynamical systems using data-driven first- and second-order Markov transition models within an agent-based representation.
2. A probabilistic data assimilation framework that integrates learned institutional transition dynamics with partial and noisy observational evidence to reconstruct latent department trajectories.
3. Quantitative in-sample and temporally held-out validation demonstrating substantial reconstruction gains over observation-only and model-only baselines, together with complementary physical-layer localization analysis illustrating the generality of the partial-observability principle.

The remainder of the paper presents the data sources and modeling framework, experimental results at both institutional and physical layers, and a discussion of implications for computational modeling of organizational systems.

2 Methods

We formulate hospital pathway reconstruction as a partially observed state-space model integrating structural transition dynamics with heterogeneous observational sources. Figure 1 provides an overview of the modeling architecture.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29915-4_12

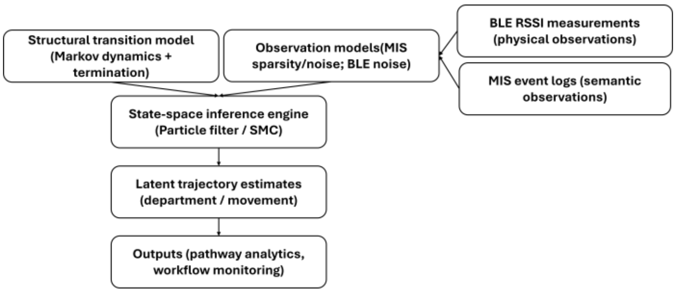

Fig. 1. State-Space Architecture for Hospital Pathway Reconstruction under Partial Observability.

The following subsections detail the data sources, transition modeling, and sequential inference framework.

2.1 Institutional Event Data

MIS Event Structure. The empirical data originate from the hospital Medical Information System (MIS) and consist of timestamped medical process events linked to internal patient identifiers. Each record contains the calendar date and time of event completion (CompletionDate and CompletionTime), the responsible Department, an ExecutionStatus flag (predominantly "completed"), and case-level boundaries (StartEpizode and EndEpizode).

Importantly, CompletionTime reflects administrative documentation rather than guaranteed physical transition time between departments. In routine clinical practice, events may be entered retrospectively or in batches, producing intra-day temporal compression and recording bias [11; 5]. For this reason, timestamps are treated as discrete observational markers rather than exact movement times.

The dataset spans 2015–2019 and includes ambulatory cardiology visits only.

Operational Episode Definition. An episode is defined as the ordered sequence of departments visited by a single patient during a single calendar day with cardiology-related activity.

Let i denote a visit identifier and d a calendar date. Let $E_{i,d}$ denote the set of MIS events associated with visit i on date d. Events are filtered to retain ambulatory admissions with completed status and valid department identifiers. The filtered events are sorted by CompletionTime. From the resulting ordered list, consecutive duplicate departments are removed to obtain a department trajectory

$$x_{i,d} = (x_1, x_2, \ldots, x_T),$$

where $x_t \in S$ denotes the department at position t, and T is the number of distinct department transitions observed during that calendar day.

Episodes are strictly day-bounded and independent of broader case-level boundaries. This operational definition aligns with outpatient workflow structure and avoids cross-day documentation artifacts [4].

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29915-4_12

After preprocessing, the dataset contains 5,308 episodes with a mean episode length of 3.66 departments over the period 2015–2019. Episodes of length one are retained for termination modeling but do not contribute to higher-order transition counts.

State Space Construction. Let S denote the department state space. Departments are ranked by frequency in the training period (2015–2018), and the 50 most frequent departments are retained explicitly. All remaining departments are aggregated into a residual state labeled OTHER, yielding $|S| = 51$.

This frequency-based truncation mitigates sparsity in higher-order transition estimation while preserving dominant cardiology care flows [10; 12].

Temporal Split. To evaluate temporal generalization, episodes are partitioned chronologically. Episodes from 2015–2018 (3,765 episodes) form the training set, and episodes from 2019 (1,543 episodes) form the test set. Transition probabilities and termination statistics are estimated exclusively from the training set, while the test set is reserved for out-of-sample reconstruction.

Chronological splitting avoids information leakage and reflects realistic deployment conditions in which historical institutional data are used to reconstruct future pathways [13].

Observational Characteristics. Two structural properties of the MIS data are central to modeling. First, completion-time bias may compress multiple events within short intervals, obscuring true temporal spacing [11]. Second, the MIS records medical actions rather than physical movement; department transitions are therefore inferred from successive event assignments rather than directly observed transfers [5].

These characteristics motivate modeling department sequences as latent stochastic trajectories observed through imperfect event-level proxies, rather than treating recorded event sequences as complete representations of patient movement.

2.2 Physical Localization Data (BLE)

To characterize the reliability of sensor-derived observations under realistic hospital conditions, Bluetooth Low Energy (BLE) indoor localization experiments were conducted on the second floor of the facility. The localization infrastructure consists of nine BLE access points (APs) and a predefined set of spatial key points (KPs) distributed across corridors and rooms.

Figure 2(a) illustrates the physical deployment of APs together with the dense grid of fingerprint KPs. The KPs serve as discrete reference locations for RSSI fingerprinting and define the measurement anchors for constructing the radio map. During a dedicated calibration session, RSSI measurements were collected at each KP to build a fingerprint-based radio map. This radio map specifies the observation model, providing the likelihood function that relates measured RSSI vectors to spatial hypotheses during inference.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29915-4_12

In addition to the dense fingerprint grid, a reduced set of navigational waypoints defines the topological motion structure of the floor. As shown in Figure 2(b), these waypoints are connected through a graph representation encoding feasible indoor movement paths. This graph is used to constrain particle propagation during localization, ensuring that motion respects architectural boundaries and navigable routes rather than unconstrained Euclidean transitions.

The deployment and calibration protocol follow prior experiments conducted on the same hospital floor [8; 9], but the present formulation explicitly separates the measurement layer (fingerprint KPs and RSSI likelihood) from the motion layer (waypoint graph and topology-aware constraints). The BLE dataset is analyzed independently from the MIS data and serves to quantify physical-layer observation uncertainty within the same institutional environment.

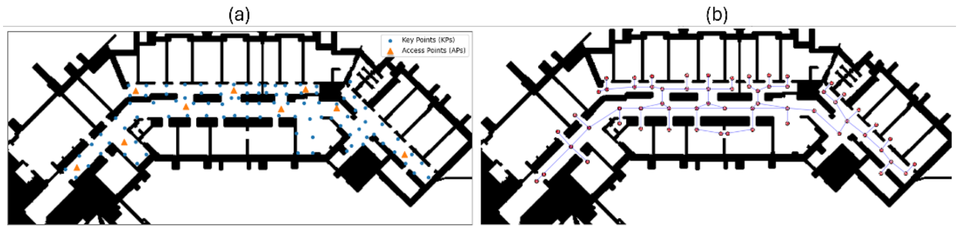

Fig. 2. BLE localization modeling components. (a) Access points (APs) and dense fingerprint key points (KPs) used for RSSI measurement modeling. (b) Reduced waypoint graph defining navigable indoor topology and constraining particle motion.

2.3 Markov Transition Modeling

Transition Structure. Let S denote the department state space with $|S| = 51$. Each episode is represented as a sequence

$$x_{1:T}, x_t \in S.$$

Department transitions are modeled in discrete time using a Markov framework. As a baseline, we consider a first-order process

$$P(x_{t+1} \mid x_t),$$

which assumes that the next department depends only on the current department. First-order Markov models are widely used in healthcare pathway and decision modeling due to their interpretability and tractability [6; 14].

Transition probabilities are estimated from the training set using empirical maximum-likelihood estimation. Let $N(b, c)$ denote the number of observed transitions from department b to department c, and let

$$N(b) = \sum_{c \in S} N(b, c).$$

The first-order transition probability is then estimated as

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29915-4_12

$$\hat{P}(\, c \mid b \,) = \frac{N(b,c)}{N(b)}, \text{ for } N(b) > 0.$$

To capture short-term temporal dependencies beyond the memoryless assumption, we additionally evaluate a second-order extension

$$P(\, x_{t+1} \mid x_t, x_{t-1} \,) \text{ [10].}$$

Let $N(a,b,c)$ denote the number of occurrences of the triplet $(x_{t-1} = a, x_t = b, x_{t+1} = c)$, and let

$$N(a,b) = \sum_{c \in S} N(a,b,c).$$

The second-order transition probability is estimated as

$$\hat{P}(\, c \mid a, b \,) = \frac{N(a,b,c)}{N(a,b)}, \text{ for } N(a,b) > 0.$$

When a second-order context (a,b) is unobserved, the model reverts to the first-order estimate $\hat{P}(c \mid b)$. All transition probabilities are computed via empirical maximum-likelihood estimation without additional smoothing or regularization, allowing the transition structure to reflect directly the observed institutional dynamics in the training data. Given the moderate state space (|S| = 51) and multi-year training corpus, sparsity effects were limited in practice.

The resulting transition model is time-homogeneous, i.e., transition probabilities are assumed to be stationary and do not vary over time within the modeled period. This assumption reflects aggregation of multi-year training data into a single empirical transition structure.

Empirical Termination Model. Episodes exhibit heterogeneous lengths. Rather than imposing a fixed stopping rule, termination probabilities are estimated empirically from training data.

For the first-order model, let $N_{\text{end}}(b)$ denote the number of times department $b$appears as the final state of an episode, and let $N_{\text{occ}}(b)$ denote the total number of occurrences of $b$in any position. The first-order termination probability is

$$\hat{p}_{\text{stop},1}(b) = \frac{N_{\text{end}}(b)}{N_{\text{occ}}(b)}.$$

For the second-order model, termination probabilities are defined analogously. Let $N_{\text{end}}(a,b)$denote the number of times the ordered pair (a,b) occurs as the final two states of an episode, and let $N_{\text{occ}}(a,b)$ denote the total number of occurrences of this pair. The second-order termination probability is

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29915-4_12

$$\hat{p}_{\text{stop},2}(a,b) = \frac{N_{\text{end}}(a,b)}{N_{\text{occ}}(a,b)}.$$

During simulation or filtering, termination is sampled according to the corresponding empirical probability.

2.4 Particle Filter Assimilation Framework

State-Space Formulation. The reconstruction of department trajectories is formulated as a discrete-time state-space model. The latent state at time t is denoted by $x_{\mathrm{t}} \in S$, where S is the department state space. State transitions follow either the first-order or second-order Markov dynamics previously introduced.

Observed medical events provide incomplete and potentially noisy information about the latent department sequence. Let y_{t} denote the observation at time t.

In the institutional setting, observations correspond to department-level proxies derived from recorded data sources. For MIS event data, each observation $y_{\mathrm{t}} \in S$ is obtained from the department attribute associated with a recorded medical event, representing an administrative indication of patient presence rather than a direct measurement of physical location. For BLE localization data, observations originate from RSSI-based position estimates, which are mapped to the nearest department or spatial zone, yielding a discrete department-level proxy consistent with the state space S.

In both cases, y_{t} should be interpreted as a noisy and incomplete observation arising from delayed or aggregated MIS recording and measurement uncertainty in BLE localization, motivating the use of a probabilistic observation model.

The resulting model corresponds to a partially observed Markov process (POMP), or equivalently, a hidden Markov-type system [15, 16]. The objective is to infer the posterior distribution of the latent trajectory given the observations.

Observation Model. To evaluate robustness under controlled partial observability, we adopt a stochastic observation mechanism.

Observations are available every s time steps (stride). When an observation is available, it equals the true latent state with probability p_{correct}. With probability $1 - p_{\text{correct}}$, the observation is replaced by a uniformly sampled alternative state from $S \setminus \{x_t\}$.

The observation likelihood is therefore defined as

$$P(y_t \mid x_t) = \begin{cases} p_{\text{correct}}, & \text{if } y_t = x_t, \\ \dfrac{1 - p_{\text{correct}}}{\mid S \mid - 1}, & \text{otherwise.} \end{cases}$$

This abstraction enables systematic evaluation of assimilation performance under varying observation sparsity and noise levels while keeping the structural transition model fixed.

Sequential Monte Carlo Approximation. We approximate the posterior distribution

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29915-4_12

$$P(x_{1:T} \mid y_{1:T})$$

using a particle filter with $N = 500$ particles.

At each time step, particles are first propagated according to the chosen transition model. For the first-order model,

$$x_{t+1}^{(i)} \sim P(x_{t+1} \mid x_t^{(i)}),$$

and for the second-order model,

$$x_{t+1}^{(i)} \sim P(x_{t+1} \mid x_t^{(i)}, x_{t-1}^{(i)}).$$

Importance weights are then updated according to the observation likelihood,

$$w_{t+1}^{(i)} \propto w_t^{(i)} \, P(y_{t+1} \mid x_{t+1}^{(i)}).$$

Weights are normalized at each step. Multinomial resampling is performed when the effective sample size falls below a predefined threshold.

Sequential Monte Carlo methods provide flexible inference for nonlinear and non-Gaussian state-space models [17; 18]. In the present setting, they enable integration of institutional transition regularities with sparse and noisy observations. The particle filtering approach follows a standard bootstrap formulation and is used here as a general inference mechanism within the proposed modeling framework.

Reconstruction Output. The reconstructed trajectory is obtained using the maximum a posteriori (MAP) state estimate at each time step. At each time step, the estimated state corresponds to the department with the highest total posterior weight across the particle ensemble, effectively selecting the most probable state given the observations and transition dynamics. Formally,

$$\hat{x}_t = \arg\max_{x \in S} \sum_{i=1}^{N} w_t^{(i)} \, \mathbf{1}\{x_t^{(i)} = x\},$$

where $\mathbf{1}\{\cdot\}$ denotes the indicator function.

This corresponds to selecting, at each time step, the most probable department under the particle-based posterior approximation. Reconstruction accuracy is evaluated as the proportion of time steps for which the reconstructed department exactly matches the ground-truth department.

2.5 Experimental Design

MIS part. Reconstruction performance is evaluated using per-time-step exact department match accuracy:

$$\text{Accuracy} = \frac{1}{\sum_i T_i} \sum_{i,t} \mathbf{1}\{\hat{x}_{i,t} = x_{i,t}\},$$

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29915-4_12

where $x_{i,t}$ denotes the true department at time $t$in episode i, $\hat{x}_{i,t}$ denotes the reconstructed estimate, and T_i is the episode length.

Transition probabilities are estimated from episodes in 2015–2018 (training set). Reconstruction is evaluated both in-sample (2015–2018) and out-of-sample (2019).

The primary experimental configuration uses stride $s = 4$, observation correctness $p_{correct} = 0.8$, and 500 particles. The parameter $p_{correct}$ represents the probability that an observed department label matches the true latent state and is used to control the level of observation noise.

In the absence of direct ground-truth estimates of observation reliability in MIS or BLE data, $p_{correct}$ was selected empirically to represent a moderate-noise regime. Values in the range 0.7–0.9 were explored during preliminary experiments and produced similar qualitative results, with 0.8 providing a representative balance between signal and noise. The reported performance trends were stable across this range, indicating that the conclusions are not sensitive to the exact choice of this parameter.

Sequential inference follows the particle filtering procedure described previously [16; 19]. Particle counts were selected to balance computational tractability and posterior stability; increasing particle number did not materially change accuracy estimates.

Reconstruction accuracy values are averaged over 500 randomly sampled evaluation episodes for each split. For particle filtering, results are further averaged over 10 independent runs with different random seeds. Reported means reflect aggregate performance across episodes; variability across episodes was comparable across models and did not affect relative performance ranking.

BLE Localization. Localization performance is evaluated on multiple controlled walking trajectories recorded on the same hospital floor. Ground truth is represented as a trajectory polyline aligned with the floor plan. Each evaluation trajectory was recorded using five independent devices, and reported errors correspond to averages across devices.

Two inference pipelines are compared. The baseline uses k-nearest neighbors (KNN) fingerprinting with $k = 6$, assigning positions based on similarity between observed RSSI vectors and the radio map. The sequential approach applies a particle filter with $N = 1000$ particles, sampling interval $\Delta t = 10$ seconds, and motion velocity parameter $v = 1.5\ m/s$. Particle propagation follows a bounded-motion model, and importance weights are computed from RSSI consistency with the fingerprint radio map.

To ensure physically valid trajectories, topology constraints are incorporated into the particle filter. Particles are restricted to accessible indoor areas, preventing wall crossings, and motion feasibility is evaluated using shortest-path distances on the floor graph computed via Dijkstra's algorithm. Consequently, spatial propagation respects navigable indoor routes rather than straight-line Euclidean distances.

Localization accuracy is measured as the mean Euclidean distance between estimated positions and the ground-truth trajectory polyline at each timestamp, consistent with the evaluation protocol described in prior work [8; 9]. The objective of this experimental design is not to introduce a new localization algorithm, but to quantify observation variability and to evaluate the effect of sequential filtering under realistic sensor noise conditions.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29915-4_12

3 Results

3.1 Department-Level Reconstruction

Baseline Comparison. Table 1 summarizes reconstruction accuracy under the primary configuration. The comparison isolates three effects: observation-only inference, structural modeling without assimilation, and sequential data assimilation. The largest improvement is obtained when structural transitions are combined with sequential particle filtering.

Table 1. Reconstruction accuracy (stride = 4, $p_{\text{correct}} = 0.8$)

Model	Train (≤2018)	Test (2019)
OBS (observation only)	0.274	0.289
ABM (first-order rollout)	0.402	0.408
PF (first-order)	0.551	0.503
PF (second-order)	0.547	0.518

Comparing OBS and ABM isolates the effect of structural transition modeling without assimilation. Model-only rollout improves reconstruction accuracy from 0.274 to 0.402 in the training period and from 0.289 to 0.408 in the test period.

Sequential assimilation yields the largest gains. In training, first-order particle filtering improves accuracy from 0.402 to 0.551. Out-of-sample, accuracy increases from 0.408 to 0.503 (first-order) and to 0.518 (second-order). Gains grow with larger stride values.

The first- and second-order formulations show comparable performance in training, with a modest out-of-sample advantage for the second-order model (0.518 vs 0.503).

Variability across evaluation episodes was comparable across models (standard deviation approximately 0.23–0.25) and did not alter the observed performance ordering.

3.2 Localization Accuracy

Table 2 reports mean localization error (m) for representative trajectories.

Table 2. Mean localization error (meters)

Session	Path	KNN (mean)	PF (mean)
Eval A	1	0.82	0.76
Eval A	2	0.52	0.79
Eval A	3	1.20	1.18
Eval B	1	0.82	0.39
Eval B	2	1.00	0.81
Eval B	3	8.06	8.15

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29915-4_12

Two distinct behaviors emerge. For several trajectories (e.g., Eval B, Path 1), particle filtering substantially improves localization accuracy, reducing mean error from 0.82 m to 0.39 m by mitigating stochastic RSSI fluctuations and producing smoother, topology-consistent trajectories.

In contrast, in Eval B, Path 3, both KNN and PF exhibit large errors (≈8 m), indicating systematic calibration mismatch rather than stochastic noise. In such cases, filtering does not yield improvement.

Figure 3 illustrates these behaviors: PF closely follows the ground truth in the improvement case, while both methods deviate substantially in the failure case, showing that filtering cannot compensate for structural bias in the observation model.

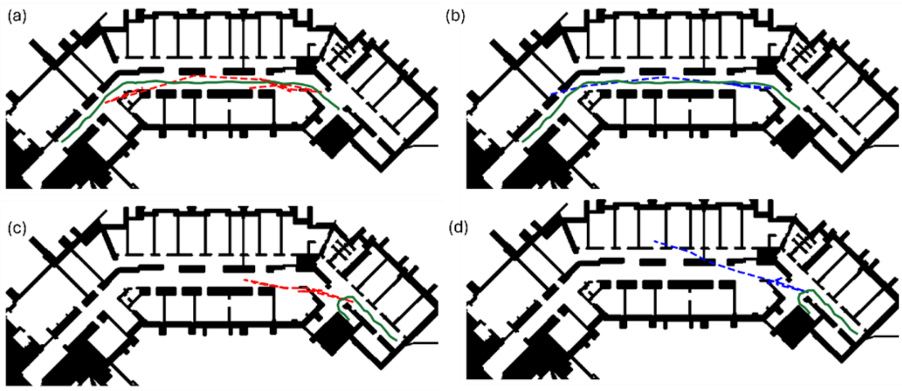

Fig. 3. Representative BLE localization trajectories. (a) KNN vs ground truth for the improvement case (Eval B, Path 1). (b) Particle filter vs ground truth for the same trajectory. (c) KNN vs ground truth for the failure case (Eval B, Path 3). (d) Particle filter vs ground truth for the same trajectory. Ground truth trajectories are shown in green; KNN estimates are shown as red dashed lines; particle filter estimates are shown in blue.

4 Discussion

This study models daily cardiology care pathways as partially observed dynamical systems. Instead of treating hospital event logs as complete representations of patient trajectories, department transitions are interpreted as latent states evolving under learned institutional transition regularities. The task therefore shifts from descriptive sequence analysis to probabilistic state inference under uncertainty.

The results indicate that structural transition modeling alone improves reconstruction relative to direct use of sparse and noisy observations, demonstrating that department flows exhibit statistically stable patterns rather than arbitrary sequencing. More importantly, sequential data assimilation yields the largest gains in both in-sample and temporally held-out evaluation. Together, these findings support modeling healthcare pathways as structured stochastic processes whose latent dynamics can be systematically exploited for probabilistic reconstruction. From a computational perspective, this illustrates that state-space modeling and sequential inference methods, traditionally

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29915-4_12

applied in physical systems, transfer effectively to organizational settings characterized by discrete and partially informative observations.

Across all evaluated configurations, the principal improvement arises from sequential integration of observations with the learned transition prior. The particle filter combines institutional structure with partial evidence to approximate posterior trajectory distributions that consistently outperform both observation-only inference and model-only rollout. The persistence of gains in the temporally held-out 2019 data indicates that learned transition tendencies retain predictive value beyond the training period, suggesting gradual institutional evolution rather than rapid structural drift.

The comparison between first- and second-order transition models further clarifies the source of improvement. While incorporating short-term memory yields only modest additional benefit, the dominant effect stems from probabilistic state estimation itself. Assimilation—rather than increased transition order—constitutes the primary methodological mechanism driving reconstruction performance.

The BLE localization experiments provide a complementary perspective on observation reliability within the same institutional environment. Even under topology-aware particle propagation and graph-based motion constraints, localization accuracy varies across sessions. Sequential filtering reduces stochastic RSSI fluctuations but fails under systematic calibration mismatch, illustrating a general property of Bayesian state estimation: filtering mitigates random noise but cannot correct structural bias or misspecified likelihood models. This behavior parallels the department-level reconstruction problem. MIS event records, like RSSI measurements, provide indirect and imperfect proxies of the underlying state, potentially affected by delayed entries, aggregation effects, or partial transition visibility. Structural transition models combined with sequential assimilation stabilize inference when uncertainty is predominantly stochastic, yet remain sensitive to systematic distortions in the observation process. The BLE analysis therefore reinforces the broader modeling principle of this work: healthcare pathways can be treated as partially observed dynamical systems, and probabilistic assimilation yields measurable benefits when observational noise is stochastic rather than structurally biased.

From an operational perspective, improved reconstruction of care pathways enables more reliable analysis of patient flow from incomplete event data. By recovering latent department transitions, the approach supports identification of bottlenecks and atypical routing patterns, informing workflow optimization and resource allocation. It may also support online monitoring of patient trajectories under uncertainty.

4.1 Limitations and Future Work

Several limitations warrant consideration. First, the transition model is frequency-based and assumes stationarity over the training period. Although temporal holdout results indicate stability, substantial workflow changes could reduce predictive performance. Adaptive or time-varying transition estimation may address this limitation.

Second, the department-level observation model uses a controlled stride-based corruption mechanism. Real hospital event noise may exhibit structured temporal dependencies or department-specific biases not captured in this abstraction.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29915-4_12

Third, BLE experiments were conducted on a single hospital floor and are intended to illustrate modeling principles rather than provide a comprehensive localization study.

Future work may explore adaptive transition modeling under evolving institutional dynamics, richer observation likelihoods derived directly from raw MIS event characteristics, and multi-scale assimilation frameworks that integrate physical localization and departmental state modeling within a unified inference architecture.

5 Conclusion

This study formulates daily ambulatory cardiology care pathways as partially observed stochastic processes and demonstrates that probabilistic data assimilation substantially improves reconstruction of department trajectories under sparse and noisy observations. Department dynamics are modeled using data-driven first- and second-order Markov transitions learned from multi-year hospital data, and latent trajectories are inferred through particle filtering within a discrete-time state-space framework.

Across both in-sample and temporally held-out evaluation, sequential assimilation consistently outperforms observation-only inference and model-only rollout. Higher-order transition memory yields only modest additional benefit; the primary improvement arises from integrating structural transition priors with partial observations via Sequential Monte Carlo inference. Temporal holdout results further indicate that learned transition dynamics retain predictive value beyond the training period, supporting the interpretation of healthcare pathways as structured stochastic processes rather than arbitrary event sequences.

Complementary BLE localization experiments provide a physical-layer validation of the same principle. Filtering mitigates stochastic observation noise but cannot compensate for systematic calibration bias, mirroring the department-level setting. Assimilation therefore enhances reconstruction when uncertainty is primarily stochastic, while remaining sensitive to model misspecification.

Overall, the results show that state-space modeling and particle-based data assimilation extend naturally to institutional and healthcare systems characterized by partial observability. When structured transition dynamics are present, probabilistic assimilation provides a principled and computationally tractable framework for reconstructing latent process trajectories.

Future work may investigate adaptive transition estimation under evolving workflows, richer observation models derived directly from MIS event characteristics, and multi-scale assimilation strategies integrating physical and departmental state representations.

Acknowledgments. The research was supported by The Russian Science Foundation, agreement №24-11-00272, https://rscf.ru/project/24-11-00272/.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29915-4_12

References

1. Kalman, R.E.: A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering* **82**(1), 35–45 (1960).
2. Evensen, G.: *Data Assimilation: The Ensemble Kalman Filter*. Springer (2009).
3. Doucet, A., de Freitas, N., Gordon, N. (eds.): *Sequential Monte Carlo Methods in Practice*. Springer (2001).
4. van der Aalst, W.M.P.: *Process Mining: Data Science in Action*, 2nd edn. Springer (2016).
5. Rojas, E., Muñoz-Gama, J., Sepúlveda, M., Capurro, D.: Process Mining in Healthcare: A Literature Review. *Journal of Biomedical Informatics* **61**, 224–236 (2016).
6. Sonnenberg, F.A., Beck, J.R.: Markov Models in Medical Decision Making: A Practical Guide. *Medical Decision Making* **13**(4), 322–338 (1993).
7. Arulampalam, M.S., Maskell, S., Gordon, N., Clapp, T.: A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking. *IEEE Transactions on Signal Processing* **50**(2), 174–188 (2002).
8. Prins, C.: *Advancing Healthcare Navigation: Indoor Localization in Hospitals Using Fingerprinting and Particle Filters*. Master's Thesis, University of Amsterdam (2024).
9. Vlasenko, V., Balakhontceva, M.: Implementation of Indoor Positioning Methods: Virtual Hospital Case. *Procedia Computer Science* **193**, 183–189 (2021).
10. Ching, W.-K., Ng, M.K.: *Markov Chains: Models, Algorithms and Applications*. Springer (2006).
11. Hripcsak, G., Albers, D.J.: Next-Generation Phenotyping of Electronic Health Records. *Journal of the American Medical Informatics Association* **20**(1), 117–121 (2013).
12. Bishop, C.M.: *Pattern Recognition and Machine Learning*. Springer (2006).
13. Bzdok, D., Meyer-Lindenberg, A.: Machine Learning for Precision Psychiatry: Opportunities and Challenges. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* **3**(3), 223–230 (2018).
14. Puterman, M.L.: *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley (1994).
15. Cappé, O., Moulines, E., Rydén, T.: *Inference in Hidden Markov Models*. Springer (2005).
16. Doucet, A., Johansen, A.M.: A Tutorial on Particle Filtering and Smoothing: Fifteen Years Later. *Journal of Statistical Software* **30**(12), 1–41 (2009).
17. Kantas, N., Doucet, A., Singh, S.S., Maciejowski, J., Chopin, N.: On Particle Methods for Parameter Estimation in State-Space Models. *Statistical Science* **30**(3), 328–351 (2015).
18. Naesseth, C.A., Lindsten, F., Schön, T.B.: Elements of Sequential Monte Carlo. *Foundations and Trends in Machine Learning* **12**(3), 307–392 (2019).
19. Chopin, N., Papaspiliopoulos, O.: *An Introduction to Sequential Monte Carlo*. Springer (2020).
20. Särkkä, S.: *Bayesian Filtering and Smoothing*. Cambridge University Press (2013).

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29915-4_12