UC Irvine

UC Irvine Electronic Theses and Dissertations

Title

The Probability of Motion Sensor-Aware Digital Twins for In Silico Ethology

Permalink

https://escholarship.org/uc/item/9t10v9xv

ISBN

9798247966944

Author

Sayfoori, Reza

Publication Date

2026-06-04

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA,
IRVINE

The Probability of Motion

Sensor-Aware Digital Twins for In Silico Ethology

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Electrical Engineering and Computer Science

by

Reza Sayfoori

Dissertation Committee:
Associate Professor Hung Cao, Chair
Professor Chen Yu Sheu
Professor Nader Bagherzadeh

2026

Portions of Chapters 3 and 4 © 2025 IEEE.
Chapter 6 © 2026 Reza Sayfoori (MDPI).
All other materials © 2026 Reza Sayfoori.

DEDICATION

To Goli Vaisi,

with love and gratitude.

TABLE OF CONTENTS

Page

LIST OF FIGURES **vii**

LIST OF TABLES **viii**

ACKNOWLEDGMENTS **ix**

ABSTRACT OF THE DISSERTATION **x**

1 Introduction **1**
- 1.1 Overview and Motivation . . . 1
- 1.2 Challenges in High-Precision Animal Localization . . . 5
  - 1.2.1 Impact of Line-of-Sight and Non-Line-of-Sight Conditions . . . 6
  - 1.2.2 Multipath Propagation and Signal Distortion . . . 7
  - 1.2.3 Environmental Dynamics and Measurement Variability . . . 8
  - 1.2.4 Behavioral State Segmentation . . . 9
  - 1.2.5 Distortion of Behavioral Transition Structure . . . 10
- 1.3 Research Gap . . . 12
- 1.4 Contributions of This Dissertation . . . 14
- 1.5 Dissertation Organization . . . 16
- 1.6 Chapter Summary . . . 17

2 Background and Literature Review **18**
- 2.1 Introduction . . . 18
- 2.2 Conventional Animal Tracking Methods . . . 19
  - 2.2.1 Video-Based Tracking Systems . . . 19
  - 2.2.2 Radio-Frequency and Wireless Tracking Systems . . . 20
- 2.3 Ultra-Wideband Localization Fundamentals . . . 20
  - 2.3.1 Principles of UWB Ranging . . . 21
  - 2.3.2 Time-of-Flight and Two-Way Ranging . . . 22
  - 2.3.3 Trilateration and Position Estimation . . . 22
- 2.4 Challenges in Indoor UWB Localization . . . 23
  - 2.4.1 Line-of-Sight and Non-Line-of-Sight Conditions . . . 24
  - 2.4.2 Multipath Propagation and Signal Attenuation . . . 25
  - 2.4.3 Environmental and System-Level Variability . . . 26

2.5 Signal Processing for Localization Enhancement . . . 26
2.5.1 Filtering and Estimation Methods . . . 27
2.5.2 Error Mitigation Under NLoS Conditions . . . 28
2.6 Probabilistic Modeling of Behavioral Dynamics . . . 29
2.6.1 Behavioral State Segmentation . . . 29
2.6.2 Shannon Entropy and Uncertainty Quantification . . . 30
2.6.3 Bernoulli Modeling of Accuracy Thresholds . . . 31
2.6.4 Markov Chains for Behavioral Transitions . . . 31
2.7 Sensor-Aware Generative Modeling and Trajectory-Level Digital Twins . . . 32
2.8 Research Gap in Existing Literature . . . 33
2.9 Chapter Summary . . . 35

3 UWB-Based Real-Time Locating System 36
3.1 Introduction . . . 36
3.1.1 Neuroscientific Motivation and Experimental Need . . . 37
3.2 System Design and Experimental Framework . . . 37
3.2.1 System Architecture . . . 37
3.2.2 Hardware Platform . . . 38
3.2.3 Experimental Enclosure and Anchor Deployment . . . 39
3.2.4 Wearable Tag Design . . . 39
3.2.5 Localization Workflow . . . 40
3.2.6 Basic Ranging Principle . . . 40
3.2.7 Two-Way Ranging Formulation . . . 41
3.2.8 Time-of-Flight Representation . . . 42
3.2.9 Localization Uncertainty . . . 42
3.2.10 Path Loss and Signal Degradation . . . 43
3.2.11 Trilateration and Coordinate Estimation . . . 43
3.3 Localization Performance Under LoS and NLoS Conditions . . . 44
3.3.1 LoS Localization Accuracy . . . 46
3.3.2 NLoS Localization Degradation . . . 46
3.3.3 Implications for Behavioral Tracking . . . 47
3.4 Chapter Summary . . . 47

4 Signal Integrity and Localization Performance 49
4.1 Introduction . . . 49
4.2 Signal Propagation in Indoor UWB Environments . . . 50
4.2.1 Line-of-Sight and Non-Line-of-Sight Conditions . . . 50
4.2.2 Multipath Propagation and Attenuation . . . 51
4.2.3 Signal-to-Noise Ratio and Channel Quality . . . 52
4.3 Localization Performance Evaluation . . . 53
4.3.1 LoS Accuracy Assessment . . . 53
4.3.2 NLoS Error Characterization . . . 54
4.3.3 Comparative Error Analysis . . . 54
4.4 Signal Enhancement and Error Mitigation . . . 55
4.4.1 Filtering and Adaptive Estimation . . . 55

4.4.2 Phase-Based and Compensation Approaches . . . 56
4.4.3 Improvement Under Degraded Conditions . . . 57
4.5 Implications for Behavioral Tracking . . . 58
4.6 Chapter Summary . . . 59

5 Probabilistic Behavioral Modeling From UWB Trajectories 60
5.1 Introduction . . . 60
5.2 Behavioral State Segmentation . . . 61
5.3 Shannon Entropy and Uncertainty Quantification . . . 62
5.4 Bernoulli Modeling of Accuracy Thresholds . . . 63
5.5 Markov Chains for Behavioral Transitions . . . 64
5.6 Sensing Quality in Behavioral Modeling . . . 66
5.7 Chapter Summary . . . 67

6 Sensor-Aware Generative Modeling and Virtual Behavioral Simulation 69
6.1 Introduction . . . 69
6.2 Sensor-Aware Generative Framework . . . 70
6.3 Latent Behavioral Dynamics and State Modeling . . . 71
6.3.1 State Space Representation . . . 72
6.3.2 Run-Based Behavioral Dynamics . . . 73
6.3.3 Behavioral Occupancy and Stability . . . 74
6.4 State-Conditioned Kinematic Generation . . . 74
6.4.1 Reference Trajectory Construction . . . 75
6.4.2 Kinematic Feature Extraction . . . 76
6.4.3 State-Conditioned Motion Priors . . . 76
6.4.4 Trajectory Propagation . . . 77
6.4.5 Outlier Control and Long-Horizon Stability . . . 77
6.5 Sensor-Aware Observation Channel . . . 78
6.5.1 State- and Regime-Dependent Distortion . . . 78
6.5.2 Mixture Representation of Observation Error . . . 79
6.5.3 Residual-Based Parameter Estimation . . . 80
6.5.4 Sensor-Observed Trajectory Generation . . . 80
6.6 Evaluation of Synthetic Trajectory Realism . . . 81
6.6.1 Per-State Statistical Realism . . . 81
6.6.2 Occupancy-Weighted Realism and Residual Agreement . . . 82
6.6.3 Temporal Realism via Mean-Squared Displacement . . . 83
6.7 Robustness Under LoS/NLoS Domain Shift . . . 84
6.8 Qualitative Realism of Representative Trajectories . . . 86
6.9 Chapter Summary . . . 88

7 Conclusions, Contributions, and Scientific Significance 90
7.1 Introduction . . . 90
7.2 Central Thesis of the Dissertation . . . 91
7.3 Principal Contributions and Novelty . . . 91
7.4 Scientific and Methodological Significance . . . 93

7.5 Limitations and Scope 93
7.6 Future Directions 94
7.7 Final Synthesis 96

Bibliography **97**

LIST OF FIGURES

Page

1.1 Experimental UWB-based rodent tracking setup with anchor placement and tagged animal. Adapted from [1]. 3
1.2 Relationship between signal-to-noise ratio, theoretical channel capacity, and positioning error under LoS and NLoS conditions. Adapted from [2]. 4
1.3 Localization error distributions under LoS and NLoS. Adapted from [1]. . . . 6
1.4 Signal degradation and enhancement analysis under obstructed UWB conditions. Adapted from [1]. 8
1.5 Localization error entropy under LoS and NLoS. Adapted from [2]. 9
1.6 Heatmap of changes in behavioral transition probabilities relative to baseline under LoS and NLoS sensing conditions. Adapted from [2]. 11
1.7 Behavioral state dynamics and percentage of total time spent in each state. Adapted from [2]. 12

6.1 Graphical overview of the sensor-aware generative framework for in silico ethology. Real UWB sessions acquired under LoS and NLoS conditions are used to estimate latent behavioral dynamics and a proxy-referenced observation channel. The framework then generates observed-domain synthetic trajectories and evaluates them using occupancy, state-conditioned kinematics, mean-squared displacement, and downstream robustness under sensing-domain shift. Adapted from [3]. 71
6.2 Per-state statistical realism metrics for real and synthetic trajectories under LoS and NLoS conditions. Adapted from [3]. 81
6.3 Occupancy-weighted realism metrics and residual-domain agreement under LoS and NLoS conditions. Adapted from [3]. 83
6.4 Mean-squared displacement of real and synthetic trajectories under LoS and NLoS conditions. Adapted from [3]. 84
6.5 Representative real trajectories under (a) LoS and (b) NLoS sensing conditions, colored by behavioral state. Adapted from [3]. 87
6.6 Representative synthetic trajectory generated by the sensor-aware framework. Adapted from [3]. 88

LIST OF TABLES

Page

1.1 Operational definitions for behavioral states used in Markov modeling. States are defined using kinematic thresholds and spatial zones, enabling reproducible segmentation. 10

1.2 Transition probabilities between behavioral states. Exploring → Feeding and Feeding → Resting dominate cycles, while Resting → Burrowing is rare. . . 11

6.1 Classifier performance on the synthetic unperturbed-versus-perturbed benchmark under LoS/NLoS sensing-domain shift. The sensor-aware classifier includes sensing regime as an explicit covariate, whereas the condition-agnostic baselines omit this information and are evaluated under cross-domain transfer. AUC is computed from decision scores, while accuracy and balanced accuracy are computed using a fixed decision threshold. Adapted from [3]. 86

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my advisor and committee chair, Associate Professor Hung Cao, for his guidance, encouragement, and continued support throughout my doctoral studies at the University of California, Irvine. I am especially grateful for his kindness, patience, and generosity as a mentor. His confidence in me, thoughtful guidance, and willingness to support my ideas provided not only the academic foundation for this dissertation, but also an example of leadership that I deeply respect. His mentorship helped me grow as an independent researcher and shaped the way I approach scientific work.

I am also grateful to my dissertation committee members, Professor Chen Yu Sheu and Professor Nader Bagherzadeh, for their time, thoughtful feedback, and valuable suggestions. Their perspectives helped strengthen the technical depth and broader direction of this dissertation.

I would like to acknowledge the Department of Electrical Engineering and Computer Science at the University of California, Irvine, for providing the academic environment, resources, and support that made this work possible.

The text of Chapters 3 and 4 of this dissertation is an adaptation of the material as it appears in Reza Sayfoori, Mao-Hsiang Huang, Amir Naderi, Mehwish Bhatti, Ron D. Frostig, and Hung Cao, "High-Precision UWB Sensor-Based Real-Time Locating System for Rodent Behavioral Studies," *IEEE Sensors Journal*, vol. 25, no. 13, pp. 25551–25559, 2025, doi:10.1109/JSEN.2025.3572031, used with permission from IEEE. The coauthors listed in this publication are Mao-Hsiang Huang, Amir Naderi, Mehwish Bhatti, Ron D. Frostig, and Hung Cao. Hung Cao directed and supervised research which forms the basis for this dissertation.

The text of Chapter 6 of this dissertation is an adaptation of the material as it appears in Reza Sayfoori, Goli Vaisi, and Hung Cao, "Virtual Mice, Real Errors: A Sensor-Aware Generative Framework for In Silico Ethology," *Sensors*, vol. 26, no. 10, article 2977, 2026, doi:10.3390/s26102977, used under the Creative Commons Attribution License (CC BY 4.0). The coauthors listed in this publication are Goli Vaisi and Hung Cao. Hung Cao directed and supervised research which forms the basis for this dissertation.

I am thankful to my collaborators and colleagues who contributed to different stages of this research, including system development, experimentation, data analysis, and manuscript preparation. Their technical support, discussions, and encouragement were greatly appreciated.

I am deeply grateful to my family for their love, patience, and encouragement. I am especially thankful to my wife for her constant support, understanding, and belief in me throughout this journey.

ABSTRACT OF THE DISSERTATION

The Probability of Motion

Sensor-Aware Digital Twins for In Silico Ethology

By

Reza Sayfoori

Doctor of Philosophy in Electrical Engineering and Computer Science

University of California, Irvine, 2026

Associate Professor Hung Cao, Chair

Quantitative analysis of rodent behavior is essential for neuroscience, neuroengineering, and preclinical research, yet reliable long-duration tracking remains difficult in realistic laboratory environments where occlusion, non-line-of-sight propagation, multipath distortion, and sensor uncertainty can alter the measured trajectory. This dissertation develops a probabilistic view of animal motion in which observed trajectories are treated not as fixed representations of behavior, but as sensor-conditioned measurements shaped by uncertainty, signal integrity, and environmental distortion. Building on this perspective, the dissertation presents a sensor-aware framework for Ultra-Wideband (UWB)-based rodent behavioral inference that connects high-precision localization, probabilistic behavioral modeling, and trajectory-level digital twins within a unified measurement-to-inference pipeline.

The first contribution is the development and validation of a UWB-based real-time locating system for rodent behavioral studies. A compact wearable tag, an optimized anchor configuration, time-based ranging, phase-based localization, adaptive filtering, and Kalman-based correction are integrated to reconstruct rodent trajectories under both Line-of-Sight

(LoS) and Non-Line-of-Sight (NLoS) conditions. Experimental validation against synchronized video tracking demonstrates centimeter-level localization performance, with LoS tracking achieving a mean localization error of 2.28 centimeters and NLoS tracking maintaining bounded error under obstructed conditions. These results establish UWB sensing as a practical platform for continuous rodent tracking in conditions where vision-based systems may be limited by visibility, lighting, and occlusion.

The second contribution is a probabilistic framework that transforms UWB trajectories from position measurements into interpretable behavioral models. Rather than treating localization error only as an engineering metric, this dissertation shows that physical-layer signal integrity propagates into behavioral interpretation. Shannon entropy, Bernoulli accuracy modeling, and Markov transition analysis are used to quantify uncertainty, reliability, and sequential structure in rodent behavior. The analysis reveals a consistent LoS-to-NLoS sensing signature: localization-error entropy increases from 1.15 to 1.78 Bits, the probability of achieving sub-20-millimeter accuracy shifts from 63.2 percent to 27.7 percent, and average behavioral state persistence shifts from 68 percent to 52 percent. These shifts are not treated as isolated tracking failures; instead, they establish that sensing condition is a measurable component of behavioral inference, linking physical-layer signal quality to uncertainty, reliability, and sequential organization in rodent behavior.

The third contribution is a sensor-aware generative framework for in silico ethology. The trajectory-level digital twin separates latent behavioral dynamics from sensing-induced distortion by coupling a semi-Markov behavioral model with a regime-dependent UWB observation channel. This factorization enables generation of virtual rodent trajectories that are both behaviorally plausible and measurement-realistic under LoS and NLoS regimes. The generated trajectories reproduce key properties of real sessions, including behavioral occupancy, state-conditioned kinematics, residual-domain structure, and mean-squared displacement across time lags. In downstream classification, explicit sensor-aware conditioning

yields stable mixed-domain performance with an area under the curve of 0.995, whereas condition-agnostic baselines degrade under LoS/NLoS domain shift.

Taken together, this dissertation establishes a probability-of-motion framework for rodent behavioral inference: first by building and validating the UWB localization platform, then by linking signal integrity to probabilistic behavioral structure, and finally by extending the framework to trajectory-level digital twins for virtual cohort generation and robustness analysis. The work advances computational ethology by showing that behavioral inference is shaped not only by how animals move, but also by how sensing systems transform movement into data. By treating observed trajectories as sensor-conditioned probabilistic representations of behavior, this dissertation provides a foundation for scalable, reproducible, and measurement-aware behavioral studies in biomedical research.

Chapter 1

Introduction

1.1 Overview and Motivation

Quantitative analysis of animal behavior is central to neuroscience, biomedical engineering, pharmacology, and preclinical research. Rodent models are especially important because they provide experimentally controlled access to locomotor activity, exploratory behavior, disease progression, and responses to therapeutic intervention, including pharmacological studies in which behavioral readouts are used to evaluate neural and therapeutic effects [4, 5, 6]. In these studies, behavior is not merely a visual observation; it is a measurable output of underlying physiological, neurological, and environmental processes. Accurate behavioral inference therefore depends on the ability to reconstruct animal movement reliably over time and to interpret those trajectories within the limitations of the sensing system used to acquire them [4, 7].

A central challenge in rodent behavioral studies is that realistic laboratory environments are not ideal measurement environments. Animals may move rapidly, enter partially obstructed regions, interact with bedding or cage structures, and occupy spaces that are difficult to

image continuously. Vision-based tracking systems provide rich spatial and visual information, but their performance can be degraded by occlusion, illumination variability, camera placement, restricted fields of view, and extensive post-processing requirements. These limitations become especially important in long-duration experiments and in settings where animals move through concealed or cluttered environments [7, 8].

Wireless localization technologies provide an alternative route for continuous tracking. Bluetooth Low Energy (BLE), WiFi, Radio Frequency Identification (RFID), and other radio-frequency methods have been investigated for indoor positioning, but their spatial resolution is often insufficient for fine-grained rodent behavioral analysis. Ultra-Wideband (UWB) localization offers a stronger foundation because its large bandwidth and short pulse duration enable precise time-based ranging, improved resistance to multipath interference, and centimeter-level localization in controlled indoor environments [1].

The first part of this dissertation establishes this sensing foundation through the development and validation of a UWB-based real-time locating system for rodent behavioral studies. The system integrates a wearable tag, optimized anchor placement, time-based ranging, phase-based localization, adaptive filtering, and Kalman-based correction to support high-precision trajectory reconstruction under both Line-of-Sight (LoS) and Non-Line-of-Sight (NLoS) conditions. The fundamental ranging principle can be expressed as

$$d = c \cdot \Delta t \tag{1.1}$$

where c represents the speed of light and Δt denotes the measured signal propagation time. By combining multiple distance estimates from fixed anchors, the position of the mobile tag can be reconstructed through multilateration and filtering methods [1].

To establish both the physical and experimental foundations of the dissertation framework, the tracking environment and signal-quality model are introduced early in this chapter [1, 2].

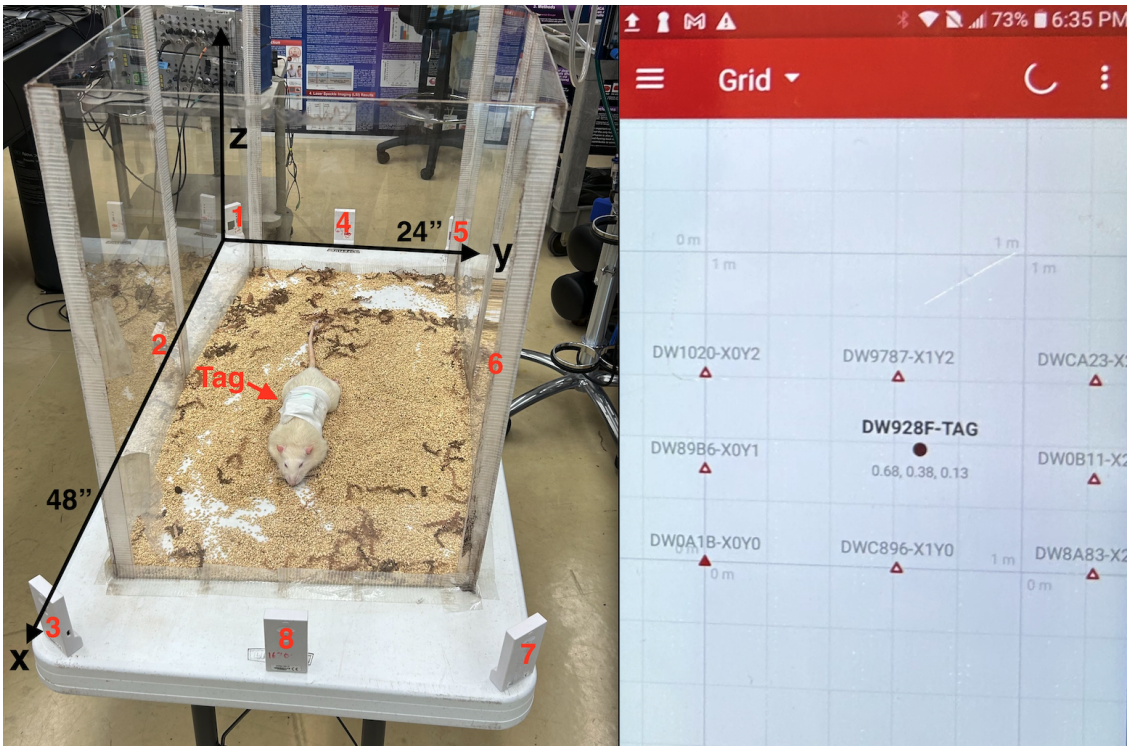


Figure 1.1: Experimental UWB-based rodent tracking setup with anchor placement and tagged animal. Adapted from [1].

Figure 1.1 illustrates the experimental UWB rodent tracking environment, including the enclosure, anchor placement, and wearable tag configuration.

However, localization accuracy alone is not sufficient for behavioral inference. In realistic deployments, UWB measurements are affected by propagation regime, signal-to-noise ratio, multipath reflections, attenuation, and NLoS distortion. These effects do not remain confined to engineering-level error metrics. They can alter reconstructed trajectories, reshape motion-derived features, and change the probabilistic structure inferred from behavioral state sequences. Therefore, this dissertation treats sensing condition not as an incidental nuisance variable, but as a structured component of the measurement-to-inference pipeline.

In this dissertation, the probability of motion refers to the probabilistic structure that emerges when animal movement is measured through an imperfect sensing system. The observed trajectory is therefore not treated as a purely deterministic record of behavior, but

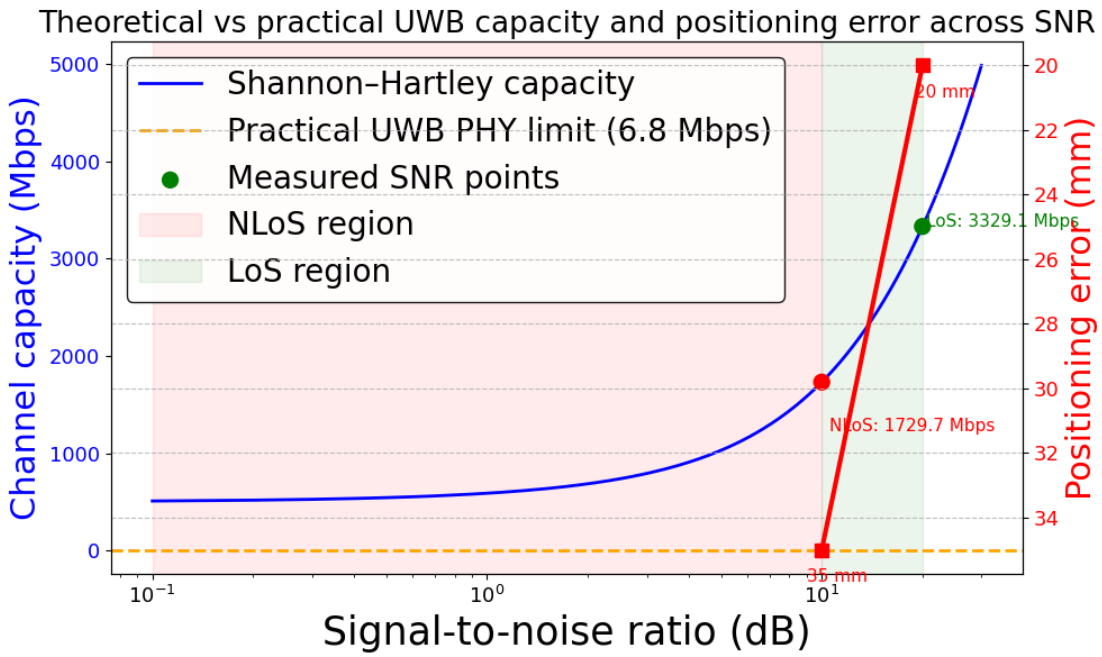


Figure 1.2: Relationship between signal-to-noise ratio, theoretical channel capacity, and positioning error under LoS and NLoS conditions. Adapted from [2].

as a sensor-conditioned representation shaped jointly by latent behavioral dynamics, localization uncertainty, and environmental distortion. This framing motivates the dissertation's progression from UWB-based trajectory acquisition to probabilistic behavioral inference and sensor-aware trajectory-level digital twins.

Figure 1.2 illustrates the relationship between signal-to-noise ratio, theoretical channel capacity, and positioning error under LoS and NLoS conditions.

The second part of this dissertation builds on the localization system by introducing a signal-aware probabilistic framework for behavioral interpretation. Instead of reducing UWB trajectories to deterministic metrics such as root mean square error or average positioning error, the framework uses Shannon entropy, Bernoulli reliability modeling, and Markov transition analysis to quantify uncertainty, threshold reliability, state persistence, and sequential organization in rodent behavior. This establishes a physics-to-behavior pathway in which signal integrity is directly linked to the reliability of behavioral inference [2].

The third part of this dissertation extends the framework from analysis to synthesis through a trajectory-level digital twin for in silico ethology. The sensor-aware generative model separates latent behavioral dynamics from sensing-induced distortion by coupling a semi-Markov behavioral process with a regime-dependent UWB observation channel. This enables the generation of virtual rodent trajectories that preserve ethological structure while reproducing LoS/NLoS-specific measurement distortion. In this dissertation, the term digital twin is used in this trajectory-level and sensor-aware sense: the model does not claim to reproduce the full biological organism, but instead represents the behavioral trajectory and measurement process needed for virtual cohort generation, robustness testing, and measurement-aware inference [3].

Together, these three components define the central motivation of this dissertation: to move from trajectory acquisition toward measurement-aware behavioral inference. The dissertation therefore develops a unified sensor-aware framework that connects UWB localization, probabilistic behavioral modeling, and trajectory-level digital twin simulation for reproducible rodent behavioral analysis.

1.2 Challenges in High-Precision Animal Localization

Achieving high-precision localization of small animals in indoor environments presents a unique set of technical challenges that extend beyond conventional positioning systems. While UWB technology provides a strong foundation for accurate ranging, its performance is highly sensitive to environmental conditions, signal propagation characteristics, and system-level constraints. In practical laboratory settings, these factors collectively introduce significant sources of error that must be carefully addressed to ensure reliable tracking [1].

1.2.1 Impact of Line-of-Sight and Non-Line-of-Sight Conditions

One of the primary challenges in UWB-based localization is the distinction between LoS and NLoS signal propagation. In LoS conditions, signals travel directly between the transmitter and receiver, enabling accurate distance estimation. However, in NLoS scenarios, signals are obstructed by physical objects such as cages, equipment, or structural elements, causing reflections, diffractions, and additional propagation delays [1]. This results in biased distance measurements, which may be expressed as

$$d_{\text{measured}} = d_{\text{true}} + \epsilon_{\text{NLoS}} \tag{1.2}$$

where ϵ_{NLoS} represents the bias introduced due to indirect signal paths.

Figure 1.3 compares localization error distributions under LoS and NLoS conditions.

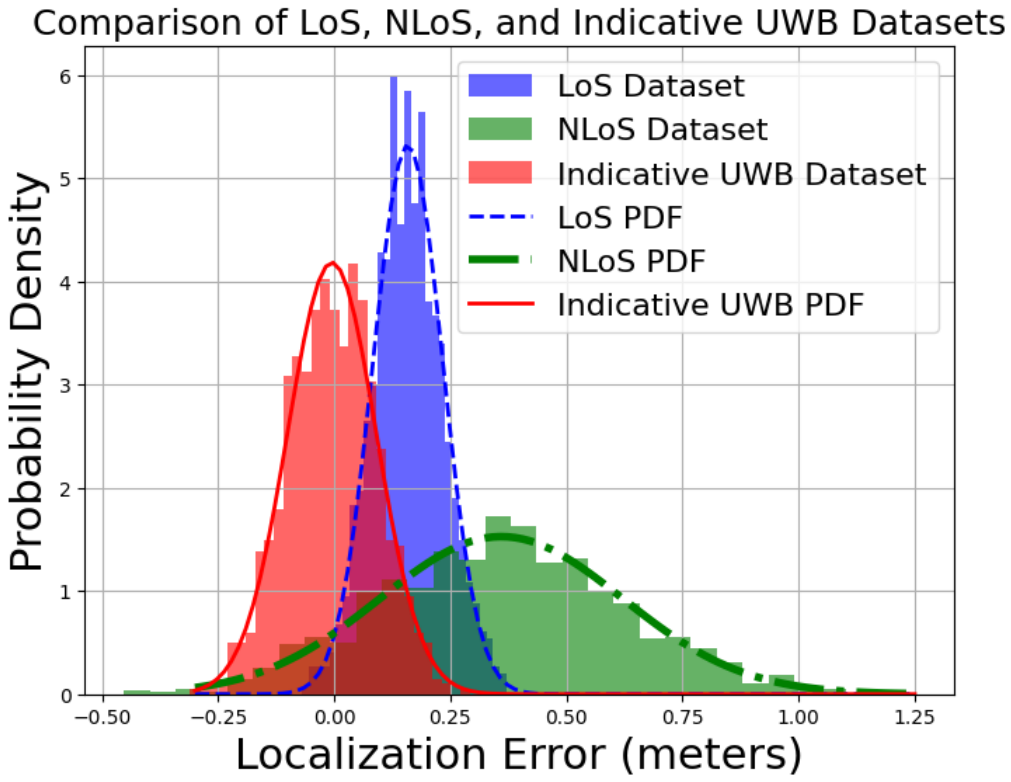


Figure 1.3: Localization error distributions under LoS and NLoS. Adapted from [1].

The difference between LoS and NLoS performance is not merely a signal-level issue; it directly influences the reliability of downstream behavioral interpretation. Increased positioning uncertainty under NLoS conditions can distort trajectories, alter inferred state transitions, and reduce confidence in quantitative behavioral metrics [2].

1.2.2 Multipath Propagation and Signal Distortion

Indoor environments are characterized by multipath propagation, where reflections from walls, floors, and surrounding objects create multiple signal paths that can interfere constructively or destructively, causing signal-strength fluctuations and inaccuracies in time-based measurements. The received signal can be modeled as

$$r(t) = \sum_{i=1}^{N} a_i s(t - \tau_i) \tag{1.3}$$

where a_i and τ_i represent the amplitude and delay of the ith path, respectively. The presence of multiple delayed components complicates the accurate estimation of the direct path, which is critical for precise ranging [1].

Figure 1.4 summarizes signal degradation and enhancement behavior under obstructed UWB conditions.

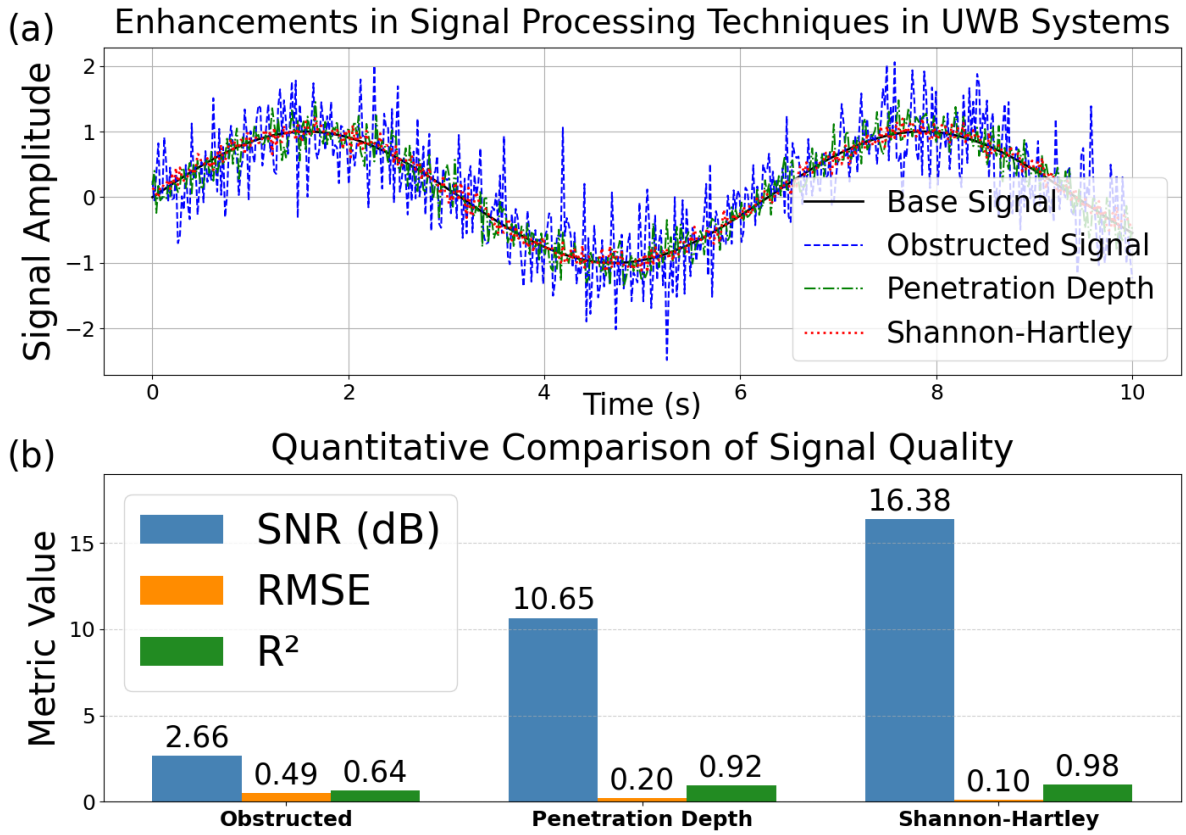


Figure 1.4: Signal degradation and enhancement analysis under obstructed UWB conditions. Adapted from [1].

1.2.3 Environmental Dynamics and Measurement Variability

Unlike static positioning systems, animal tracking involves dynamic and unpredictable motion within the environment. Rapid movement, changes in orientation, and interactions with surrounding structures introduce temporal variations in signal measurements. These dynamics result in non-stationary noise characteristics that cannot be effectively handled using simple averaging or static calibration techniques [1].

To address this issue, recursive filtering approaches are often adopted. A general state-space

model may be expressed as

$$x_k = Ax_{k-1} + w_k \tag{1.4}$$

$$z_k = Hx_k + v_k \tag{1.5}$$

where x_k represents the hidden state, z_k is the measurement, and w_k and v_k denote process and measurement noise, respectively. Such formulations support trajectory smoothing and improved robustness against noisy ranging measurements [1].

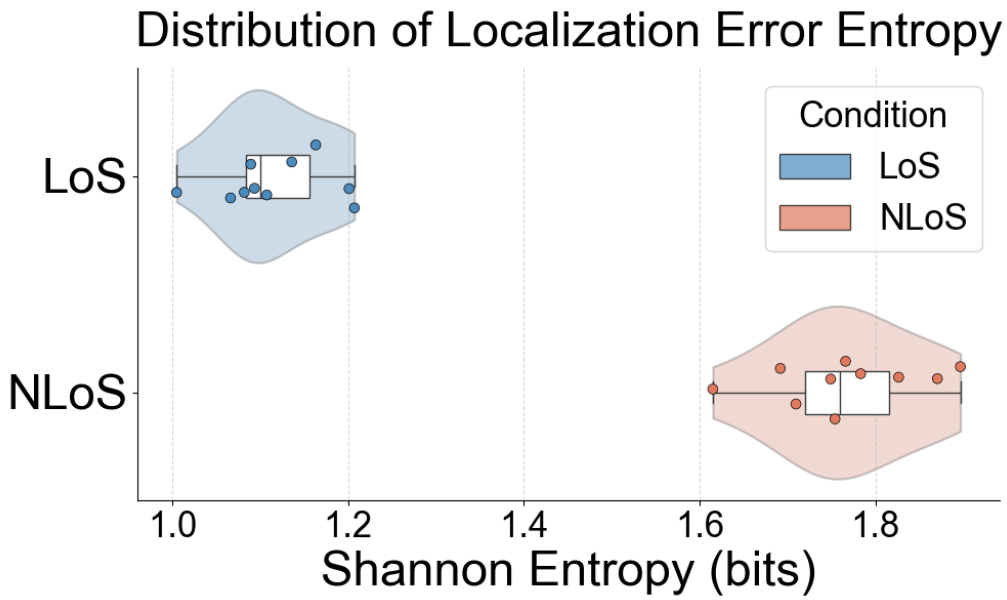


Figure 1.5: Localization error entropy under LoS and NLoS. Adapted from [2].

1.2.4 Behavioral State Segmentation

Figure 1.5 demonstrates that localization uncertainty increases under NLoS conditions. A major challenge in rodent tracking is that localization errors do not remain confined to engineering metrics such as RMSE; rather, they propagate into behavioral models. When signal integrity degrades, the uncertainty of inferred behavioral states increases, and the transition structure extracted from the trajectory may become more fragmented [2].

This effect can be quantified using Shannon entropy:

$$H = -\sum_{i=1}^{n} p_i \log_2 p_i \tag{1.6}$$

where p_i denotes the probability associated with the ith error bin or behavioral outcome. Higher entropy indicates greater uncertainty and lower predictability [2].

Table 1.1 summarizes the operational definitions used to segment continuous trajectories into discrete behavioral states.

Table 1.1: Operational definitions for behavioral states used in Markov modeling. States are defined using kinematic thresholds and spatial zones, enabling reproducible segmentation.

State	**Criteria**
Resting	Speed < 5 centimeters per Second maintained for ≥ 2 Seconds.
Exploring	Locomotion between arena zones without resource interaction.
Feeding	Occupancy in food-access zone for ≥ 1 Second.
Burrowing	Stationary near bedding zone with minimal displacement and visible substrate interaction.

1.2.5 Distortion of Behavioral Transition Structure

Behavioral studies often depend not only on position accuracy but also on the validity of inferred temporal structure. If localization quality deteriorates, the estimated sequence of behavioral states may no longer reflect the true organization of activity [2]. Figure 1.6 illustrates how transition probabilities change relative to baseline under LoS and NLoS conditions.

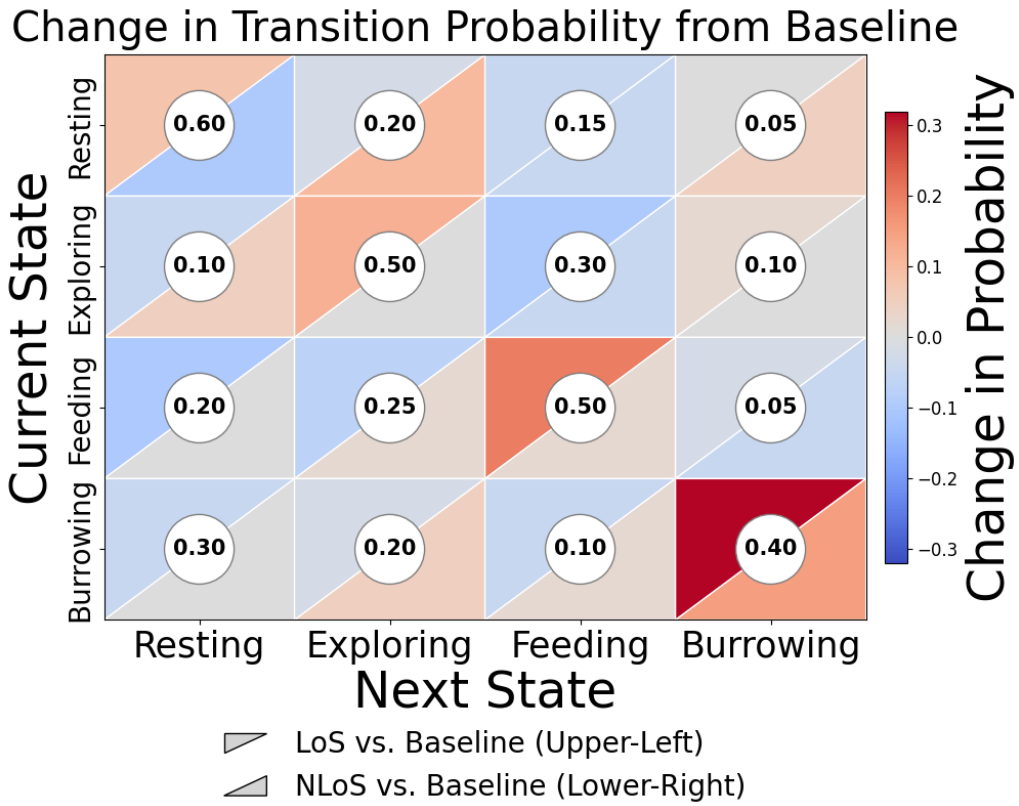


Figure 1.6: Heatmap of changes in behavioral transition probabilities relative to baseline under LoS and NLoS sensing conditions. Adapted from [2].

Table 1.2: Transition probabilities between behavioral states. Exploring → Feeding and Feeding → Resting dominate cycles, while Resting → Burrowing is rare.

Current State	**Resting**	**Exploring**	**Feeding**	**Burrowing**
Resting	0.60	0.20	0.15	0.05
Exploring	0.10	0.50	0.30	0.10
Feeding	0.20	0.25	0.50	0.05
Burrowing	0.30	0.20	0.10	0.40

Table 1.2 reports the transition probabilities between behavioral states used in the Markov representation.

Figure 1.7 presents both the temporal sequence of behavioral states and the proportion of

time allocated to each state.

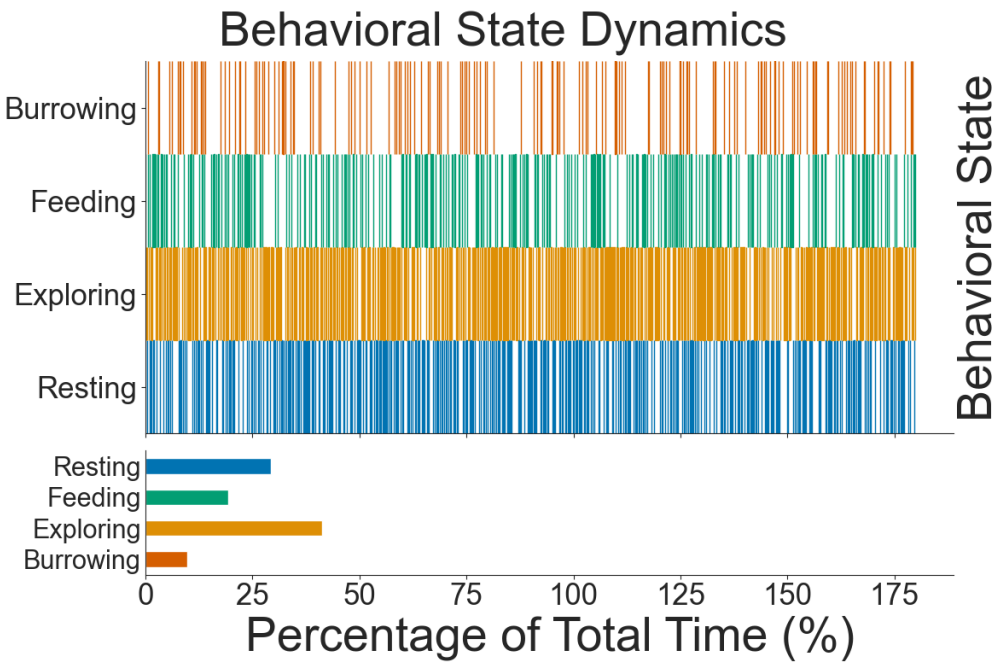


Figure 1.7: Behavioral state dynamics and percentage of total time spent in each state. Adapted from [2].

A first-order Markov model provides one way to represent these sequential relationships:

$$P_{ij} = \Pr(X_{t+1} = j \mid X_t = i) \tag{1.7}$$

where X_t denotes the behavioral state at time t, and P_{ij} is the transition probability from state i to state j [2].

1.3 Research Gap

Despite progress in animal tracking, indoor localization, and quantitative behavioral analysis, a methodological gap remains between sensing performance and behavioral interpretation. Vision-based systems provide visual information, but their reliability can be limited by occlusion, lighting variability, camera placement, bedding, enclosure geometry, and the

tendency of rodents to move through concealed or obstructed spaces [7, 8]. Radio-frequency approaches provide an attractive alternative, yet many conventional wireless systems lack the spatial resolution required for fine-grained rodent behavioral studies [9, 10].

UWB localization addresses these limitations by enabling centimeter-level tracking in indoor environments [11, 10, 1]. However, much of the existing UWB literature emphasizes deterministic localization metrics, such as root mean square error, average positioning error, and signal-to-noise ratio. These metrics are essential for validating the sensing platform, but they do not fully explain how sensing uncertainty propagates into trajectory-derived behavioral inference [1, 2]. In rodent ethology, this distinction is critical because localization error is not only an engineering artifact; it can alter inferred movement states, transition probabilities, uncertainty estimates, and downstream behavioral conclusions [2].

A second gap is that signal integrity and behavioral modeling are often treated as separate problems. In practice, they are tightly coupled. LoS and NLoS conditions do not merely change localization accuracy; they reshape the statistical properties of the observed trajectory. Multipath propagation, attenuation, and regime-dependent sensing distortion can induce domain shifts that affect probabilistic behavioral models, including entropy measures, Bernoulli reliability estimates, and Markov transition structures [2].

A third gap is the limited availability of sensor-aware generative frameworks for rodent behavioral analysis. Existing behavioral simulations often focus on plausible movement patterns without explicitly modeling the observation process that transforms latent motion into measured trajectories. Conversely, localization studies often model sensing error without extending the analysis toward virtual cohorts, robustness testing, or in silico ethology. As a result, there remains a need for a trajectory-level digital twin framework that represents latent behavioral dynamics and sensing-regime-dependent observation distortion [3].

This dissertation addresses these gaps through a probability-of-motion framework that treats

localization, signal integrity, behavioral modeling, and virtual trajectory generation as linked components of a measurement-to-inference pipeline. In this framework, animal motion is modeled as an ethological process and a sensor-conditioned probabilistic observation [1, 2, 3].

1.4 Contributions of This Dissertation

The dissertation builds on three connected publications by the author: the experimentally validated UWB real-time locating system for rodent tracking, the signal-integrity-to-behavior probabilistic modeling framework, and the sensor-aware generative framework for in silico ethology [1, 2, 3]. Together, these works establish a measurement-aware view of animal motion in which trajectories are not treated as deterministic records alone, but as probabilistic observations shaped by behavior, sensing uncertainty, and environmental distortion.

- **Measuring animal motion with high-precision UWB localization:** A UWB-based real-time locating system is developed and validated for continuous rodent tracking in a controlled laboratory environment. The system integrates a wearable UWB tag, optimized anchor placement, time-based ranging, phase-based localization, adaptive filtering, and Kalman-based correction to reconstruct centimeter-level trajectories under both LoS and NLoS sensing conditions.

- **Characterizing sensing uncertainty under LoS/NLoS regimes:** The dissertation quantifies how line-of-sight and non-line-of-sight propagation affect localization accuracy, signal integrity, and trajectory reliability. This characterization establishes sensing condition as a structured measurement variable that shapes the probability of the observed trajectory rather than as unmodeled background noise.

- **Quantifying the probabilistic structure of measured behavior:** A probabilistic analysis framework is introduced to transform reconstructed trajectories into interpretable behavioral models. Shannon entropy, Bernoulli accuracy modeling, and

Markov transition analysis are used to quantify uncertainty, threshold reliability, state persistence, and transition structure in sensor-derived rodent behavior.

- **Linking signal integrity to behavioral inference:** The dissertation establishes a physics-to-behavior bridge between UWB signal quality and quantitative ethology. Rather than treating UWB only as a position sensor, the work reframes UWB trajectories as probabilistic information sources from which behavioral markers can be extracted and interpreted under known sensing conditions.

- **Building trajectory-level digital twins for in silico ethology:** A sensor-aware generative framework is developed to separate latent behavioral dynamics from sensing-induced distortion. The trajectory-level digital twin combines semi-Markov behavioral modeling, occupancy calibration, state-conditioned kinematic priors, and a regime-dependent UWB observation channel to generate virtual trajectories that are both behaviorally plausible and measurement-realistic.

- **Generating virtual cohorts for robustness analysis:** The dissertation demonstrates that sensor-aware synthetic trajectories can support virtual cohort generation, statistical realism evaluation, and downstream robustness testing under LoS/NLoS domain shift. This provides a foundation for scalable in silico ethology, pilot experimentation, and future power-analysis workflows.

- **Establishing a unified probability-of-motion framework:** The dissertation integrates sensing hardware, localization accuracy, signal integrity, probabilistic behavioral modeling, and sensor-aware simulation into a single measurement-to-inference pipeline for modeling animal motion under uncertainty.

1.5 Dissertation Organization

The remainder of this dissertation is organized to reflect the progression from sensing hardware to measurement-aware behavioral inference and trajectory-level digital twin modeling.

Chapter 2 reviews literature relevant to indoor localization, rodent behavioral tracking, UWB sensing, LoS/NLoS propagation, probabilistic modeling, and computational ethology. This chapter establishes the technical and biological context for the dissertation and identifies the limitations of existing vision-based and radio-frequency-based tracking approaches.

Chapter 3 presents the UWB-based real-time locating system developed for rodent behavioral studies. It describes the hardware platform, wearable tag, anchor configuration, experimental enclosure, ranging workflow, and localization framework used to reconstruct rodent trajectories under controlled laboratory conditions. This chapter establishes the experimental sensing foundation of the dissertation.

Chapter 4 analyzes localization performance and signal integrity under LoS and NLoS conditions. It examines how multipath propagation, attenuation, signal-to-noise ratio, and environmental obstruction affect UWB positioning accuracy. This chapter also presents signal-processing and estimation methods used to improve robustness and demonstrates why sensing condition must be treated as a structured variable in behavioral studies.

Chapter 5 develops the probabilistic behavioral modeling framework derived from UWB trajectories. It introduces behavioral state segmentation and applies Shannon entropy, Bernoulli threshold modeling, and Markov transition analysis to quantify uncertainty, reliability, state persistence, and sequential organization. This chapter establishes the physics-to-behavior link by showing how signal integrity propagates into behavioral interpretation.

Chapter 6 extends the dissertation from analysis to synthesis through sensor-aware generative modeling and trajectory-level digital twins. It introduces a bounded digital-twin framework

in which latent behavioral dynamics are separated from sensing-induced distortion using a semi-Markov behavioral process and a regime-dependent UWB observation channel. This chapter demonstrates how measurement-realistic virtual trajectories can support in silico ethology, virtual cohort generation, and robustness testing under LoS/NLoS domain shift.

Chapter 7 concludes the dissertation by summarizing the major findings, contributions, and scientific significance of the work. It highlights the central conclusion that rodent behavioral inference is shaped not only by animal movement, but also by the sensing process through which movement is measured. The chapter also discusses limitations and future directions, including multimodal sensing, larger animal cohorts, and expanded trajectory-level digital twin models for biomedical behavioral research.

1.6 Chapter Summary

This chapter introduced the scientific motivation, technical context, and conceptual scope of the dissertation. It established the importance of high-precision rodent tracking for quantitative behavioral research and showed that conventional vision-based and narrowband wireless approaches remain limited under realistic laboratory conditions. It further introduced UWB sensing as a promising foundation for precise indoor localization and highlighted the major sources of error that arise under LoS and NLoS propagation, multipath distortion, and environmental variability.

The chapter also framed the central research problem of the dissertation: the need for a sensor-aware probability-of-motion framework that connects localization fidelity, sensing uncertainty, and behavioral interpretation. Finally, it outlined the organization of the remaining chapters. Together, these elements establish the intellectual foundation for the methodological developments that follow, where animal motion is progressively treated as measured trajectory, probabilistic behavioral structure, and trajectory-level digital twin.

Chapter 2

Background and Literature Review

2.1 Introduction

This chapter reviews the theoretical and methodological background needed to support the sensor-aware framework developed in this dissertation. The chapter first surveys conventional animal tracking approaches, including video-based and radio-frequency-based systems, and then introduces UWB technology as a high-resolution alternative for indoor localization. It next examines the principal challenges associated with LoS and NLoS propagation, multipath interference, attenuation, and environmental variability. The chapter then reviews signal-processing methods for localization enhancement and probabilistic tools for behavioral modeling, including entropy analysis, Bernoulli reliability modeling, and Markov transition analysis. Finally, it introduces sensor-aware generative modeling and trajectory-level digital twins as the conceptual bridge to the simulation framework developed later in the dissertation [3].

2.2 Conventional Animal Tracking Methods

2.2.1 Video-Based Tracking Systems

Video-based tracking systems have long been used in behavioral neuroscience because they provide direct visual access to animal movement, posture, and spatial occupancy, with early real-time and video path-analysis systems demonstrating the feasibility of automated animal-movement analysis [12, 13]. These systems typically combine cameras, image processing pipelines, and post hoc trajectory extraction to quantify locomotion, zone preference, and interaction patterns. Their widespread adoption has made them one of the most common tools for behavioral phenotyping and computational ethology [7, 8].

Despite their utility, video-based approaches exhibit several limitations that restrict their effectiveness in naturalistic and cluttered laboratory environments. Their performance is strongly influenced by illumination conditions, camera angle, spatial resolution, and the visibility of the subject. Occlusion is a particularly important limitation in rodent experiments, where animals may move under shelters, inside burrows, or near enclosure boundaries that partially obstruct the camera's field of view. In such cases, the resulting trajectories may become fragmented or require significant post-processing and manual correction [7, 8].

Another limitation is that video systems often demand stable acquisition conditions and substantial computational resources for frame-by-frame analysis. While recent computer vision and pose-estimation methods, including markerless deep-learning-based approaches, have greatly improved the richness of behavioral measurements and can achieve high accuracy in controlled settings, they may still depend on controlled imaging conditions and annotation-intensive workflows [14, 15]. These constraints motivate the exploration of complementary sensing modalities that can maintain reliable localization even when direct visual access is compromised [7, 8].

2.2.2 Radio-Frequency and Wireless Tracking Systems

To address the limitations of camera-based tracking, researchers have explored radio-frequency, inertial, and wireless sensing technologies for animal localization and movement monitoring [16]. These include Bluetooth Low Energy (BLE), WiFi, Radio Frequency Identification (RFID), and other wireless sensor network approaches. Such systems offer advantages in deployment flexibility and can operate without continuous visual contact, which is beneficial in enclosed or low-visibility environments [9, 10].

However, many conventional wireless localization systems provide only coarse positioning accuracy. BLE- and WiFi-based methods commonly rely on received signal strength measurements, which are highly sensitive to multipath effects, environmental interference, and signal attenuation. As a result, their positioning accuracy is often limited to the meter or submeter scale, which is insufficient for fine-grained rodent behavioral analysis [9]. RFID-based methods have also been explored for animal tracking, but their performance is constrained by limited range, sensitivity to tag-reader geometry, and reduced spatial precision in dynamic environments [17, 9, 10].

These limitations have motivated increasing interest in UWB technology, which offers substantially higher temporal resolution, wider bandwidth, and greater resistance to interference than many narrowband wireless alternatives. Compared with BLE, WiFi, and RFID, UWB is better suited for applications requiring centimeter-level localization and reliable operation in complex indoor environments [11, 10, 1].

2.3 Ultra-Wideband Localization Fundamentals

UWB technology has emerged as one of the most promising wireless solutions for high-precision indoor localization because of its wide bandwidth, fine temporal resolution, and robustness against multipath effects relative to many narrowband alternatives [11, 10, 18,

19]. Unlike conventional wireless systems that rely primarily on received signal strength for approximate positioning, UWB transmits very short pulses across a large frequency spectrum, enabling more accurate estimation of propagation time and, consequently, more precise ranging.

A major advantage of UWB is that its large bandwidth improves temporal discrimination between closely arriving signal components. This makes it possible to estimate direct-path arrival times with substantially higher resolution than is typically achievable with BLE, WiFi, or RFID-based systems. As a result, UWB has become attractive for applications requiring centimeter-level localization, including robotics, industrial tracking, healthcare monitoring, and animal behavior studies in controlled indoor environments [11, 1].

2.3.1 Principles of UWB Ranging

The basic principle of UWB ranging is to estimate the travel time of an electromagnetic signal between a transmitter and a receiver, and then convert that travel time into distance. In its simplest form, the relation between propagation time and distance is given by

$$d = c \cdot \Delta t \tag{2.1}$$

where d is the estimated distance, c is the speed of light, and Δt is the signal propagation time. Because electromagnetic waves travel extremely quickly, even nanosecond-scale timing errors can translate into significant localization error. Therefore, precise timestamping, synchronization, and accurate time-of-arrival estimation are central to UWB-based positioning in realistic environments [1, 11, 20].

In practice, UWB ranging is affected by both measurement noise and environmental distortion. Reflections from nearby surfaces, partial obstructions, and device-specific timing offsets may all perturb the measured propagation delay. Consequently, raw ranging values

are often refined using calibration, filtering, or statistical correction methods before being used for final position estimation [1].

2.3.2 Time-of-Flight and Two-Way Ranging

Among the most widely used ranging approaches in UWB systems are Time-of-Flight (ToF) and Two-Way Ranging (TWR). In ToF-based localization, the propagation time of a signal between two nodes is measured and used directly to infer distance. However, one-way ToF methods typically require very tight clock synchronization between devices, which can be difficult to maintain in practical systems [11].

To overcome this issue, many real-world UWB platforms use TWR, in which two devices exchange messages and estimate range from the round-trip propagation time. By measuring the total travel time over a two-way exchange, TWR reduces dependence on exact synchronization between independent clocks and is therefore well suited to embedded localization systems [1]. In experimental rodent-tracking applications, TWR has proven especially useful because it supports robust ranging with compact wearable tags and fixed anchors under realistic laboratory conditions.

The use of TWR, however, does not completely eliminate error sources. Residual clock drift, timestamp quantization, and environmental distortions may still degrade ranging accuracy, particularly under non-line-of-sight conditions. For this reason, TWR is often combined with signal processing strategies such as recursive filtering and error compensation to improve stability and precision [1].

2.3.3 Trilateration and Position Estimation

Once distances from a mobile tag to multiple anchors have been estimated, the spatial position of the tag can be recovered through trilateration or multilateration. In two-dimensional positioning, each anchor defines a circle centered at the anchor location with radius equal to

the estimated distance. The intersection of these circles provides the estimated tag position. In ideal conditions with exact ranging, the circles intersect at a single point. In real systems, however, measurement noise causes imperfect intersections, requiring numerical estimation methods to determine the most likely location [10, 11, 21].

In a simplified two-dimensional form, position estimation can be expressed geometrically using inter-anchor distances and tag-to-anchor ranges. More generally, practical localization systems solve an overdetermined estimation problem using three or more anchors and optimize the position estimate by minimizing residual ranging error. The resulting localization accuracy depends not only on ranging precision but also on anchor geometry, environmental conditions, and the mathematical stability of the estimation procedure [1].

These considerations are especially important in behavioral experiments involving rodents, where small-scale motion and rapid directional changes demand both high temporal fidelity and precise spatial reconstruction. Thus, the value of UWB localization lies not only in its ranging mechanism, but also in the full estimation pipeline that converts raw signal timing into robust trajectories suitable for quantitative behavioral analysis [1].

2.4 Challenges in Indoor UWB Localization

Although UWB technology offers substantial advantages for indoor positioning, its practical deployment remains constrained by several persistent challenges. In controlled laboratory environments, localization performance is influenced not only by the intrinsic ranging capability of the hardware, but also by signal propagation conditions, environmental clutter, anchor geometry, and the dynamic motion of the tracked subject. These factors can collectively degrade positioning accuracy and reduce the reliability of reconstructed trajectories [11, 10, 1].

In animal behavioral studies, these challenges are particularly significant because the re-

quired spatial precision is often much higher than that needed in conventional indoor navigation applications. Small localization errors may distort inferred locomotion paths, alter apparent transitions between behavioral zones, and reduce confidence in the interpretation of state-dependent behavioral metrics. Therefore, understanding the major sources of UWB localization error is essential for designing robust tracking systems [1, 2].

2.4.1 Line-of-Sight and Non-Line-of-Sight Conditions

One of the most important factors affecting UWB localization accuracy is whether the signal propagates under LoS or NLoS conditions. In LoS settings, the direct propagation path between the transmitter and receiver is unobstructed, which allows accurate estimation of signal travel time. Under NLoS conditions, however, physical barriers such as enclosure walls, equipment, shelters, or animal-induced obstructions may block the direct path and force the signal to propagate through reflections or diffractions [1, 22].

This introduces a positive ranging bias because the measured propagation time corresponds to an indirect path that is longer than the true geometric distance. A simplified expression for this effect is

$$d_{\text{measured}} = d_{\text{true}} + \epsilon_{\text{NLoS}} \tag{2.2}$$

where ϵ_{NLoS} denotes the error introduced by obstructed propagation. Unlike zero-mean random noise, this bias is systematic and can significantly affect downstream localization performance if not explicitly identified and mitigated [11, 23, 1].

The distinction between LoS and NLoS conditions is particularly important in behavioral experiments because rodents naturally occupy concealed areas, move near obstacles, and generate dynamic partial occlusions. As a result, the signal environment may change continuously during a single session, creating condition-dependent distortions in the observed

trajectory [2].

2.4.2 Multipath Propagation and Signal Attenuation

A second major challenge in indoor UWB localization is multipath propagation. In enclosed environments, transmitted signals may reflect off walls, floors, metallic structures, and experimental apparatus before reaching the receiver, a well-established characteristic of indoor radio propagation channels [24]. The received waveform can therefore contain multiple delayed versions of the transmitted signal, which complicates the identification of the direct path and makes UWB channel modeling important for realistic indoor localization analysis [11, 10, 25].

A general representation of the received signal is

$$r(t) = \sum_{i=1}^{N} a_i s(t - \tau_i) \tag{2.3}$$

where a_i and τ_i denote the amplitude and delay associated with the ith propagation path. When multiple reflected components arrive close together in time, their superposition may distort the observed waveform and reduce the precision of timing estimation.

In addition to multipath, attenuation also affects localization reliability. Signal strength decreases with distance and can further decline due to penetration through materials or partial obstruction by structures within the arena. Such attenuation reduces effective signal-to-noise ratio and may increase the variance of timing measurements, especially in cluttered environments [1]. Consequently, the combined effects of multipath and attenuation can substantially degrade positioning fidelity even when the system operates within its nominal communication range.

2.4.3 Environmental and System-Level Variability

Indoor localization performance is influenced by variability from the environment and the localization system. Environmental factors include arena geometry, anchor placement, enclosure material composition, and objects that may reflect or block the signal. System-level factors include timestamp resolution, synchronization error, hardware calibration, and the geometric arrangement of anchors relative to the tracked tag [1, 10].

Anchor geometry is especially important because poor spatial distribution of anchors can amplify ranging uncertainty during position estimation. Even with accurate pairwise distance estimates, unfavorable geometry may produce larger coordinate errors through geometric dilution of precision. For this reason, anchor placement must be carefully optimized to achieve stable localization across the full tracking region [1].

A further source of variability is the dynamic behavior of the animal itself. Rodents may move rapidly, change orientation suddenly, or interact with structures in ways that alter tag visibility and local propagation conditions. These effects create non-stationary measurement noise, meaning that the characteristics of the localization error may change over time rather than remain constant throughout an experiment [2].

Taken together, these challenges demonstrate that indoor UWB localization is not solely a matter of applying a high-resolution ranging technology. Rather, it requires an integrated understanding of propagation physics, experimental geometry, and signal-aware estimation in order to produce trajectories that are sufficiently accurate for behavioral interpretation.

2.5 Signal Processing for Localization Enhancement

Because raw UWB measurements are affected by timing noise, multipath propagation, attenuation, and non-stationary environmental effects, signal processing plays a central role in improving localization reliability. In practice, the quality of the final trajectory depends

not only on the underlying ranging mechanism, but also on the estimation methods applied after measurement acquisition. Filtering, calibration, and error-mitigation strategies are therefore essential for converting noisy distance observations into stable spatial tracks suitable for behavioral analysis [1, 11, 10].

In rodent-tracking applications, this requirement is particularly important because the tracked subject moves dynamically and often interacts with structures that alter propagation conditions over time. As a result, the signal-processing layer must be sufficiently robust to suppress short-term fluctuations while preserving the temporal fidelity of the reconstructed motion trajectory [1, 2].

2.5.1 Filtering and Estimation Methods

A common approach to improving localization stability is to formulate the tracking problem in state-space form and apply recursive estimation methods, an approach also widely used for modeling animal movement from noisy observations [26]. In this framework, the true position evolves over time according to a process model, while the observed measurements are treated as noisy projections of the hidden state. A general linear formulation may be written as

$$x_k = Ax_{k-1} + w_k \tag{2.4}$$

$$z_k = Hx_k + v_k \tag{2.5}$$

where x_k denotes the hidden system state at time step k, z_k is the measurement, A is the state-transition matrix, H is the observation matrix, and w_k and v_k represent process and measurement noise, respectively. Recursive filters based on this formulation can smooth trajectories, reduce sensitivity to outliers, and provide more stable coordinate estimates than raw multilateration alone [27, 1].

In practical UWB systems, Kalman-based estimation is especially useful because it combines sequential measurements in a statistically principled way while remaining computationally efficient. This makes it well suited for real-time localization applications in which the system must continuously update the position of a moving tag. In experimental rodent tracking, such filtering improves continuity of the estimated path and reduces the visual and quantitative effects of frame-to-frame ranging fluctuations [1].

2.5.2 Error Mitigation Under NLoS Conditions

Filtering alone does not fully resolve the challenges introduced by non-line-of-sight conditions, because NLoS errors are often biased rather than purely random. For this reason, localization systems frequently incorporate additional correction strategies that explicitly target NLoS-induced distortion. These strategies may include adaptive filtering, signal-strength-aware calibration, phase-based estimation, and model-based compensation methods that reduce the effect of indirect propagation paths [1].

In the context of the present dissertation, signal-aware mitigation is important because improved ranging accuracy directly enhances the quality of downstream behavioral analysis. If NLoS bias is left uncorrected, the resulting coordinate error may shift reconstructed trajectories, alter the inferred occupancy of behavioral zones, and distort the transition structure used in probabilistic modeling. Thus, error mitigation is not merely an engineering refinement; it is a necessary step in preserving biological interpretability [2].

More broadly, these considerations show that localization enhancement must be understood as part of a complete sensing-and-inference pipeline. Accurate behavioral tracking depends not only on the physical signal itself, but also on how that signal is filtered, corrected, and translated into trajectories that retain meaningful spatial and temporal structure.

2.6 Probabilistic Modeling of Behavioral Dynamics

Accurate localization alone is not sufficient for meaningful behavioral interpretation. Modern computational ethology emphasizes that behavior should be quantified across multiple spatial and temporal scales rather than reduced only to coarse summary measures, because systematic behavioral quantification provides a critical pathway for relating movement patterns to neural function and brain dynamics [28, 29, 30]. In rodent studies, the primary scientific objective is often to infer patterns of activity, state transitions, and behavioral variability from the recorded trajectories, including fine-scale behavioral structure that may occur over short temporal windows [31]. For this reason, trajectory analysis must extend beyond geometric reconstruction and incorporate probabilistic models that describe how animals occupy space, transition between behaviors, and express uncertainty in their movement patterns [2, 3].

Probabilistic modeling is especially valuable because real behavioral data are inherently variable and often influenced by both biological and sensing-related uncertainty. A deterministic interpretation of trajectories may obscure this variability, whereas probabilistic methods provide a more flexible and interpretable framework for quantifying uncertainty, persistence, and transition structure. In the present context, these approaches are particularly useful for linking localization quality to downstream behavioral interpretation [2].

2.6.1 Behavioral State Segmentation

A necessary first step in probabilistic behavioral analysis is the conversion of continuous trajectories into discrete behavioral states. This process, often referred to as behavioral state segmentation, enables the trajectory to be represented as a sequence of symbolic states that can be analyzed statistically. In rodent studies, such states may correspond to locomotor and context-dependent activities such as resting, exploring, feeding, and burrowing [2].

State segmentation typically relies on a combination of spatial zones, velocity thresholds, dwell time, and context-specific experimental definitions. Although this abstraction simplifies the raw trajectory, it provides a tractable representation for analyzing how behavior evolves over time. Related computational ethology approaches also use unsupervised and embedding-based discovery methods to identify spontaneous behavioral structure from movement data, further emphasizing the importance of reliable trajectory acquisition before downstream behavioral modeling [32, 33]. The reliability of the resulting state sequence depends on both the segmentation criteria and the positional accuracy of the localization system, since spatial or temporal distortions can lead to misclassification of behavioral events [2].

2.6.2 Shannon Entropy and Uncertainty Quantification

One of the most useful information-theoretic measures for behavioral analysis is Shannon entropy, which quantifies the uncertainty or dispersion of a probability distribution. In the context of localization and behavior, entropy can be used to evaluate the unpredictability of error distributions, the variability of state occupancy, or the uncertainty of state transitions. The entropy of a discrete distribution is given by

$$H = -\sum_{i=1}^{n} p_i \log_2 p_i \tag{2.6}$$

where p_i denotes the probability associated with the ith outcome. Higher entropy indicates greater uncertainty and a more dispersed distribution, whereas lower entropy reflects greater predictability [2].

In behavioral analysis, entropy is important because it provides a compact measure of disorder or fragmentation in the observed data. When signal quality deteriorates, the inferred behavioral sequence may become less stable, increasing the uncertainty associated with both localization errors and behavioral transitions. Thus, entropy provides a direct bridge between

sensing quality and behavioral interpretability [2].

2.6.3 Bernoulli Modeling of Accuracy Thresholds

Another useful probabilistic perspective is to treat localization success as a binary event relative to a chosen accuracy threshold. In this formulation, each localization estimate is classified as either successful or unsuccessful depending on whether its error falls within an acceptable tolerance. This can be modeled as a Bernoulli random variable:

$$X = \begin{cases} 1, & \text{if } |e| \leq \epsilon \\ 0, & \text{if } |e| > \epsilon \end{cases} \tag{2.7}$$

where e denotes localization error and ϵ is the application-specific tolerance threshold. The empirical success probability can then be estimated as

$$\hat{p} = \frac{1}{N} \sum_{k=1}^{N} X_k \tag{2.8}$$

where N is the total number of localization samples [2].

This perspective is valuable because it translates engineering performance into a directly interpretable statistical quantity. Rather than reporting only average error, Bernoulli modeling reveals the probability that the localization system satisfies the precision required for behavioral interpretation. In experiments where even small spatial deviations may alter behavioral labels, such threshold-based reliability analysis is especially informative [2].

2.6.4 Markov Chains for Behavioral Transitions

To model sequential dependencies in behavior, first-order Markov chains provide a useful and interpretable framework. In this representation, the probability of the next behavioral

state depends only on the current state, yielding the transition relation

$$P_{ij} = \Pr(X_{t+1} = j \mid X_t = i) \tag{2.9}$$

where X_t denotes the behavioral state at time t, and P_{ij} is the probability of transitioning from state i to state j [2].

Markov models are useful in behavioral studies because they summarize the structure of state transitions in a compact matrix form while remaining easy to interpret. They can reveal dominant behavioral cycles, state persistence, and differences in behavioral organization across experimental conditions. In the context of UWB-based tracking, Markov analysis is particularly relevant because degradation in localization quality may alter the inferred transition matrix, thereby changing the apparent structure of animal behavior [2, 3].

Together, behavioral state segmentation, entropy analysis, Bernoulli modeling, and Markov representations provide a probabilistic foundation for understanding how spatial trajectories encode behavioral structure. These tools are therefore central to the broader framework developed in this dissertation, which seeks to connect localization fidelity with robust behavioral interpretation.

2.7 Sensor-Aware Generative Modeling and Trajectory-Level Digital Twins

Beyond analyzing recorded trajectories, behavioral research increasingly requires methods for simulating realistic movement patterns, generating virtual cohorts, and evaluating the robustness of downstream inference under controlled conditions. Simulation-based robustness testing is especially useful when models must be evaluated under controlled variability or domain shifts between synthetic and real observations [34]. In this dissertation, this

need is addressed through sensor-aware generative modeling, where the goal is not only to produce plausible animal movement, but also to reproduce the measurement distortions introduced by the sensing system. A key distinction in this framework is the separation between latent behavioral dynamics and observed sensor measurements. The latent trajectory represents the underlying movement process, whereas the observed trajectory represents the same movement after it has been transformed by localization error, signal distortion, and sensing-regime-dependent uncertainty. This distinction is especially important for UWB-based behavioral studies because LoS and NLoS conditions can produce different error distributions and therefore different observed trajectory statistics. The trajectory-level digital twin concept used in this dissertation follows this sensor-aware interpretation while remaining aligned with the broader digital twin idea of representing a physical process through a virtual model that captures relevant system dynamics and measurement interactions [35]. It represents locomotor trajectories and the measurement process through which they are observed, rather than attempting to reproduce the complete biological organism. Sensor-aware generative modeling therefore provides a bridge between probabilistic behavioral analysis and in silico ethology by enabling virtual trajectories to be generated under controlled LoS and NLoS sensing regimes [3].

2.8 Research Gap in Existing Literature

Despite substantial progress in indoor localization and behavioral tracking, several important limitations remain unresolved in the existing literature. First, many conventional animal tracking systems, especially video-based approaches, remain sensitive to occlusion, illumination variability, camera geometry, and visual obstruction. These limitations restrict their reliability in naturalistic laboratory environments where rodents may move through concealed, cluttered, or partially obstructed regions [7, 8, 1].

Second, although UWB technology has demonstrated strong localization performance, many

prior studies emphasize deterministic engineering metrics such as mean error, root mean square error, and signal-to-noise ratio. These metrics are necessary for validating the sensing platform, but they do not fully explain how sensing degradation propagates into higher-level behavioral interpretation. As a result, the relationship between physical-layer signal integrity and probabilistic behavioral inference remains insufficiently addressed in much of the current literature [11, 10, 2].

Third, localization accuracy and behavioral modeling are often treated as separate problems. In practice, however, they are tightly coupled. Errors introduced by NLoS propagation, multipath interference, attenuation, and sensing-regime-dependent distortion may alter reconstructed trajectories, inferred state transitions, entropy estimates, threshold reliability, and behavioral persistence. Consequently, engineering-level localization errors may translate directly into behavioral misinterpretation if the sensing process is not explicitly modeled [2].

Fourth, relatively few studies connect robust UWB-based localization, signal-aware error analysis, probabilistic behavioral modeling, and sensor-aware virtual trajectory generation within a single framework. Existing work often focuses either on improving positioning accuracy or on analyzing behavior after trajectories have already been reconstructed. Fewer approaches treat the observation process itself as part of the behavioral inference problem. This gap becomes especially important for in silico ethology, where virtual trajectories should be not only behaviorally plausible, but also measurement-realistic under the sensing conditions in which real data are acquired [3].

Therefore, the central research gap addressed in this dissertation is the lack of an experimentally grounded probability-of-motion framework that treats sensing, localization, behavioral interpretation, and virtual trajectory generation as coupled components of a single measurement-to-inference pipeline. This gap motivates a sensor-aware view in which observed animal motion is modeled jointly through behavior, uncertainty, and measurement distortion.

2.9 Chapter Summary

This chapter reviewed the theoretical and methodological foundations relevant to high-precision indoor localization, rodent behavioral analysis, and sensor-aware virtual modeling. It surveyed conventional tracking methods, introduced the core principles of UWB ranging and position estimation, examined the main challenges associated with indoor signal propagation and localization error, and discussed the role of signal processing in improving tracking robustness. The chapter also reviewed probabilistic tools for interpreting behavioral trajectories, including state segmentation, entropy-based uncertainty analysis, Bernoulli success modeling, and Markovian state transitions.

In addition, this chapter introduced the motivation for sensor-aware generative modeling and trajectory-level digital twins. This framing is important because the dissertation does not treat behavioral trajectories as purely biological signals or localization errors as isolated engineering artifacts. Instead, it treats the measured trajectory as the product of both animal movement and the sensing process used to observe that movement.

Taken together, the reviewed literature motivates the probability-of-motion framework developed in the following chapters, where sensing quality, localization performance, behavioral interpretation, and virtual trajectory generation are treated as coupled components of a single measurement-to-inference pipeline. This framing supports the central dissertation view that observed animal motion is shaped jointly by latent behavior and the sensing process used to measure it.

Chapter 3

UWB-Based Real-Time Locating System

Portions of this chapter are adapted from the author's published article: Reza Sayfoori, Mao-Hsiang Huang, Amir Naderi, Mehwish Bhatti, Ron D. Frostig, and Hung Cao, "High-Precision UWB Sensor-Based Real-Time Locating System for Rodent Behavioral Studies," IEEE Sensors Journal*, vol. 25, no. 13, pp. 25551–25559, 2025.*

3.1 Introduction

This chapter presents the Ultra-Wideband (UWB)-based real-time locating system developed for rodent behavioral studies. The chapter describes the system architecture, hardware platform, experimental enclosure, anchor deployment strategy, wearable tag implementation, and localization workflow that together enable high-precision indoor tracking. Unlike the previous chapter, which focused on motivation and background, this chapter emphasizes the finalized design and implementation of the UWB tracking framework [1, 11, 10].

3.1.1 Neuroscientific Motivation and Experimental Need

The experimental motivation for this system was shaped by a neuroscience need to quantify rodent behavior in enriched and naturalistic environments. Prior work from the Frostig laboratory showed that naturalistic sensory and behavioral experience can reshape adult cortical representations, supporting the broader argument that laboratory rodent studies should preserve richer interactions between animals and their environment [36, 37]. More broadly, naturalistic behavior has become an important direction in neuroscience because it captures how animals move, explore, and interact under conditions that better reflect ethologically meaningful behavior [38].

The practical request for this tracking system emerged from collaboration with the Frostig laboratory, where naturalistic rodent behavior requires measurement in environments that include richer spatial structure, concealed regions, and behaviors such as exploration, tunneling, and burrowing. These behaviors are scientifically valuable because they better preserve the animal–environment interactions that motivate naturalistic neuroscience studies, but they also create major sensing challenges for camera-dependent tracking. In this context, the UWB system was developed as an engineering response to a neuroscience-driven measurement problem: how to obtain reliable, real-time rodent trajectories when direct visual access may be interrupted by enclosure geometry, bedding, tunnels, low-light conditions, or other occlusions [14, 15, 1].

3.2 System Design and Experimental Framework

3.2.1 System Architecture

The UWB tracking system was designed as an integrated sensing platform for real-time indoor localization of rodents in controlled laboratory environments. The system consists of five primary components: fixed UWB anchor nodes, a rodent-mounted mobile UWB tag, a

controlled experimental enclosure, a real-time data acquisition pipeline, and a localization engine. The fixed anchors provide spatial reference points for range estimation, while the mobile tag exchanges UWB signals with the anchor network during free movement. The 122 centimeters $\times$ 61 centimeters enclosure provides a controlled behavioral tracking environment, and the data acquisition pipeline synchronizes sequential ranging measurements for subsequent processing. Finally, the localization engine applies ranging, filtering, and estimation algorithms to reconstruct the spatial trajectory of the rodent [1].

In contrast to conventional tracking approaches that depend primarily on optical visibility, the developed framework was designed to maintain robust positioning under both line-of-sight (LoS) and non-line-of-sight (NLoS) conditions. This makes the system suitable for behavioral experiments in which rodents naturally occupy enclosed, partially obstructed, or visually challenging regions of the arena. The overall framework therefore combines hardware design, anchor geometry, ranging methodology, and signal-aware estimation into a unified real-time locating system [1].

3.2.2 Hardware Platform

The hardware platform was built around the DecaWave DWM1001C module, a compact UWB transceiver that integrates radio-frequency circuitry, embedded processing, and wireless communication capability into a single localization unit. This platform was selected because it supports indoor positioning with fine temporal resolution while maintaining low-power operation and a compact form factor, both of which are critical for rodent behavioral experiments [1, 39, 40].

Each DWM1001C unit can be configured either as a fixed anchor or as a mobile tag. In the implemented system, multiple anchors were installed around the boundary of the enclosure, while the tag was mounted on the animal using a lightweight wearable structure designed to minimize interference with natural movement. This hardware configuration enabled stable

communication between the tag and anchors while preserving the mobility and behavioral freedom of the animal during experiments [1].

3.2.3 Experimental Enclosure and Anchor Deployment

The enclosure measured 122 centimeters by 61 centimeters and provided a controlled environment for evaluating localization performance under behavioral conditions. The enclosure dimensions were selected to support rodent movement while remaining compatible with dense UWB anchor coverage and synchronized ground-truth comparison [1].

Anchor placement plays a role in localization accuracy because the spatial distribution of anchors affects geometric stability during position estimation. For this reason, anchors were deployed around the perimeter of the enclosure to improve coverage and reduce geometric dilution of precision across the tracking region. This arrangement supports reliable localization by limiting blind spots and improving the numerical stability of multilateration [1].

3.2.4 Wearable Tag Design

To support continuous rodent tracking, the mobile UWB unit was implemented as a lightweight wearable tag attached directly to the animal. The wearable design was developed to balance mechanical stability, electrical functionality, and minimal behavioral disruption. Because excessive payload or discomfort can alter locomotion and compromise experimental validity, the wearable unit was designed to remain compact and unobtrusive during movement [1].

The wearable tag enabled real-time communication with the anchor network and continuously provided ranging data for localization. This design was especially important for behavioral experiments involving free movement, where the localization system must track the subject without restricting motion or introducing substantial artificial interference [1].

3.2.5 Localization Workflow

The complete localization workflow begins with signal exchange between the mobile tag and the anchor nodes. These signal exchanges generate ranging estimates, which are then combined through position estimation algorithms to reconstruct the trajectory of the moving subject. Because raw distance measurements are affected by noise, propagation distortion, and environmental clutter, the localization workflow also incorporates estimation and correction procedures that improve stability and trajectory quality [1].

This system-level workflow provides the foundation for the mathematical ranging and localization models developed in the next section. In particular, the following section formalizes the ranging equations, distance estimation process, and coordinate reconstruction strategy used throughout the dissertation.

3.2.6 Basic Ranging Principle

The most fundamental relationship in UWB ranging is the conversion of signal propagation time into distance. For one-way time-of-flight ranging, if an electromagnetic signal travels for a propagation interval Δt_{ToF}, the corresponding distance is

$$d = c \cdot \Delta t_{\mathrm{ToF}} \tag{3.1}$$

where d is the estimated distance, c is the speed of light, and Δt_{ToF} is the one-way propagation time. For round-trip ranging, the measured round-trip interval must be divided by two before conversion to distance. Because the speed of light is extremely large, accurate ranging requires highly precise timing measurements. Even very small timestamp errors can lead to noticeable localization inaccuracies [11, 1].

A related expression that explains the high localization capability of UWB systems is the

spatial resolution, given by

$$\Delta x = \frac{c}{2B} \tag{3.2}$$

where Δx is the spatial resolution, c is the speed of light, and B is the effective bandwidth. This relation shows that larger bandwidth leads to finer spatial resolution, which is one of the reasons UWB is well suited for high-precision indoor localization [1, 11].

In practical systems, this relationship is embedded within ranging protocols that account for timing, signal exchange structure, reply intervals, and synchronization constraints.

3.2.7 Two-Way Ranging Formulation

To reduce sensitivity to clock mismatch between the mobile tag and the anchors, the implemented system employs a Two-Way Ranging (TWR) strategy. In TWR, the tag and anchor exchange signals in both directions, and the round-trip timing is used to estimate the effective propagation delay. This approach reduces dependence on perfectly synchronized clocks and is therefore well suited to embedded indoor localization systems [1].

A generalized TWR-based range estimate may be written as

$$R = \frac{c}{4}\left[(T_{rr} - T_{sp}) - T_{rsp} + (T_{rf} - T_{sr}) - T_{rsp}\right] \tag{3.3}$$

where R denotes the range, T_{rr} and T_{rf} denote reception-related timestamps, T_{sp} and T_{sr} denote transmission-related timestamps, and T_{rsp} represents the reply interval used in the ranging exchange. In practice, this formulation allows the system to infer range using reciprocal timing information while reducing the synchronization error between devices [1].

Although TWR improves robustness relative to one-way timing methods, it does not eliminate all sources of error. Residual clock drift, timestamp quantization, and environmental

distortions may still affect the measured range, especially under NLoS conditions.

3.2.8 Time-of-Flight Representation

The range estimation process can also be expressed through the more direct Time-of-Flight (ToF) representation

$$\text{ToF} = \frac{d}{c} \tag{3.4}$$

where ToF denotes the signal travel time for a propagation distance d. This formulation highlights the physical basis of UWB ranging and is useful for understanding how signal timing errors translate directly into spatial uncertainty [1].

To support stable localization, ToF values must be processed with correction and filtering strategies that reduce the influence of noise and propagation distortion.

3.2.9 Localization Uncertainty

The final localization performance depends not only on range estimation itself, but also on the accumulation of multiple uncertainty sources. A simplified expression for overall localization uncertainty is

$$\sigma_{\text{pos}} = \sqrt{\sigma_{\text{ToF}}^2 + \sigma_{\text{sync}}^2} \tag{3.5}$$

where σ_{pos} is the total position uncertainty, σ_{ToF} represents the uncertainty associated with propagation-time estimation, and σ_{sync} reflects synchronization-related timing error [1]. This expression emphasizes that localization performance is inherently coupled to both signal timing fidelity and system-level synchronization quality.

3.2.10 Path Loss and Signal Degradation

Because indoor environments introduce attenuation and distortion, received signal quality varies with distance and environmental structure. A common path-loss representation is

$$S = S_0 - 10n \log_{10}(d) \tag{3.6}$$

where S is the received signal strength, S_0 is the reference signal strength, n is the path-loss exponent, and d is the propagation distance [1]. This model provides a useful abstraction for describing the reduction in signal strength as the separation between transmitter and receiver increases, especially in cluttered indoor environments.

A related and equally important quantity is the signal-to-noise ratio (SNR), which characterizes signal quality relative to background noise:

$$\text{SNR} = 10 \log_{10}\left(\frac{P_{\text{signal}}}{P_{\text{noise}}}\right) \tag{3.7}$$

where P_{signal} denotes the received signal power and P_{noise} denotes the background noise power [1]. Higher SNR generally corresponds to more reliable ranging performance, whereas lower SNR increases timing uncertainty and measurement variance.

Such degradation is important because weakened or distorted signals reduce ranging reliability, increase measurement variance, and amplify localization error under both multipath and NLoS conditions.

3.2.11 Trilateration and Coordinate Estimation

Once distances from the mobile tag to multiple anchors have been estimated, the system reconstructs the tag position through trilateration. In two-dimensional localization, each measured distance defines a circle centered at a known anchor position. The intersection of

multiple such circles determines the estimated coordinates of the tag [11, 10].

In simplified form, the two-dimensional coordinates may be expressed as

$$x = \frac{d_1^2 - d_2^2 + d_{12}^2}{2d_{12}} \tag{3.8}$$

$$y = \frac{d_1^2 - d_3^2 + d_{13}^2 - x^2}{2d_{13}} \tag{3.9}$$

For three-dimensional positioning, the vertical coordinate may be written as

$$z = \sqrt{d_1^2 - x^2 - y^2} \tag{3.10}$$

where d_1, d_2, and d_3 are the measured distances from the tag to three anchors, and d_{12} and d_{13} are the known distances between anchor pairs [1]. In practice, real systems often use more than three anchors and estimate coordinates through an optimization procedure that minimizes the discrepancy among all available range constraints.

These equations demonstrate that localization is a geometric reconstruction problem built upon the quality of the underlying range measurements. Thus, ranging accuracy, anchor geometry, and signal-processing quality all jointly determine the fidelity of the final trajectory.

3.3 Localization Performance Under LoS and NLoS Conditions

The practical effectiveness of the UWB-based real-time locating system depends on its ability to maintain accurate positioning under both Line-of-Sight (LoS) and Non-Line-of-Sight (NLoS) conditions. In controlled laboratory settings, LoS measurements generally provide

higher ranging fidelity because the direct propagation path between the tag and anchor remains unobstructed. Under NLoS conditions, reflections, attenuation, and indirect propagation introduce additional timing error that degrades localization performance [1].

Experimental evaluation demonstrated systematic differences between these two sensing regimes. Under LoS conditions, the system achieved a mean localization error of 2.28 centimeters with a standard deviation of 0.76 centimeters across 267 samples. Under NLoS conditions, the mean localization error increased to 4.40 centimeters with a standard deviation of 1.03 centimeters across 220 samples, with a maximum observed deviation of 6.2 centimeters. These results confirm that UWB tracking remains effective under both regimes, while also showing that obstruction and multipath propagation produce measurable changes in trajectory reliability [1].

The phase-based normalized-distance estimation strategy further improved localization robustness relative to traditional ToF estimation. In LoS conditions, the phase-based method achieved an error of 0.5 ± 0.1 centimeters, compared with 1.0 ± 0.2 centimeters for the ToF method. In NLoS conditions, the phase-based method achieved an error of 2.0 ± 0.4 centimeters, compared with 3.5 ± 0.8 centimeters for the ToF method. These results indicate that phase-based estimation provides improved resistance to multipath and environmental distortion [1].

Signal-quality analysis also showed that LoS and NLoS conditions differ at the physical layer. LoS measurements maintained higher signal-to-noise ratio values, whereas NLoS conditions produced reduced signal quality due to obstruction, attenuation, and multipath effects. Adaptive pulse repetition frequency control and filtering improved localization performance under degraded conditions, with adaptive pulse repetition frequency optimization improving positioning accuracy relative to static configurations [1].

3.3.1 LoS Localization Accuracy

Under LoS conditions, the implemented UWB system achieved high localization fidelity due to unobstructed direct-path signal propagation. The observed error distribution remained relatively narrow, indicating stable ranging performance and consistent coordinate reconstruction across the tracking region. This level of precision is particularly important for rodent behavioral studies, where small spatial displacements may correspond to meaningful behavioral events such as zone entry, resource interaction, or directional change [1].

The strong LoS performance also validates the anchor deployment strategy and the suitability of the hardware platform for real-time behavioral tracking. Because the anchor geometry was selected to reduce geometric dilution of precision, the resulting coordinate estimates remained stable throughout the enclosure, supporting reliable trajectory reconstruction during free movement [1].

3.3.2 NLoS Localization Degradation

In contrast, NLoS conditions introduced measurable degradation in localization performance. Signal obstruction increased the effective propagation path, reduced signal quality, and amplified uncertainty in range estimation. As a consequence, reconstructed coordinates exhibited greater variance and larger deviations from the expected path [1].

This degradation is especially relevant in animal experiments because rodents naturally interact with the environment in ways that create transient obstruction and dynamic multipath effects. Thus, NLoS distortion should not be viewed as an exceptional edge case, but rather as a realistic component of practical indoor tracking. For a localization system to remain useful in such experiments, it must therefore be evaluated not only under ideal LoS conditions but also under more challenging NLoS scenarios that reflect actual behavioral settings [1, 2].

3.3.3 Implications for Behavioral Tracking

The difference between LoS and NLoS performance has direct consequences for behavioral phenotyping. When localization noise increases, the inferred spatial trajectory may become fragmented or biased, which in turn affects the extraction of state-dependent behavioral features. Errors in coordinate estimation may shift the apparent time spent in a zone, alter the inferred sequence of behavioral states, and increase the uncertainty of transition-based models [2].

For this reason, localization performance should be interpreted not only in terms of engineering accuracy but also in terms of its impact on behavioral validity. A tracking system used in neuroscience or biomedical research must provide sufficiently robust measurements to preserve the structure of the behavior being studied. This requirement motivates the signal-aware and probabilistic extensions developed in later chapters of this dissertation [2, 3].

3.4 Chapter Summary

This chapter presented the Ultra-Wideband (UWB)-based real-time locating system developed for rodent behavioral studies. It described the overall system design, including the hardware platform, experimental enclosure, anchor deployment strategy, wearable tag implementation, and localization workflow used to support real-time indoor tracking. Together, these elements established the practical sensing framework through which rodent trajectories were acquired in controlled laboratory environments.

The chapter also introduced the mathematical basis of the localization system, including the ranging principle, TWR formulation, ToF representation, uncertainty modeling, path-loss effects, signal-to-noise ratio, and trilateration-based coordinate estimation. These models demonstrated how raw UWB signal timing is converted into spatial position estimates and

clarified the major factors that influence localization fidelity.

Finally, the chapter examined localization performance under both LoS and NLoS conditions. The results demonstrate that the implemented system achieves centimeter-level tracking accuracy while also revealing that sensing regime strongly affects trajectory reliability. This chapter therefore establishes the validated sensing foundation for the probability-of-motion framework developed in the following chapters, where measured trajectories are analyzed through signal integrity, probabilistic behavioral modeling, and sensor-aware trajectory-level digital twins.

Chapter 4

Signal Integrity and Localization Performance

Portions of this chapter are adapted from the author's published article: Reza Sayfoori, Mao-Hsiang Huang, Amir Naderi, Mehwish Bhatti, Ron D. Frostig, and Hung Cao, "High-Precision UWB Sensor-Based Real-Time Locating System for Rodent Behavioral Studies," IEEE Sensors Journal*, vol. 25, no. 13, pp. 25551–25559, 2025.*

4.1 Introduction

This chapter examines the signal integrity and localization performance of the Ultra-Wideband (UWB)-based real-time locating system under realistic indoor experimental conditions. While Chapter 3 established the hardware architecture and localization workflow, the present chapter focuses on how signal propagation, environmental obstruction, and measurement uncertainty influence tracking accuracy in practice. Particular emphasis is placed on the distinction between Line-of-Sight (LoS) and Non-Line-of-Sight (NLoS) conditions because these sensing regimes directly affect ranging fidelity, coordinate reconstruction, and the quality of

the resulting behavioral trajectories [1, 2].

This chapter evaluates signal performance as a behavioral-inference factor by examining how degradation in UWB signal quality affects localization error and the reliability of reconstructed trajectories.

The chapter first reviews major signal-propagation phenomena that influence UWB performance in indoor environments, including LoS/NLoS propagation, multipath effects, attenuation, and signal-to-noise ratio (SNR) degradation. It then evaluates localization performance under both LoS and NLoS conditions, highlighting the practical consequences of environmental distortion on tracking accuracy. Finally, the chapter discusses signal-enhancement and error-mitigation strategies that improve system robustness and preserve the validity of downstream behavioral interpretation [1, 2].

4.2 Signal Propagation in Indoor UWB Environments

Indoor UWB localization performance is governed not only by the nominal bandwidth and timing resolution of the hardware, but also by the propagation environment through which the signal travels. In controlled laboratory settings, signal behavior is shaped by the enclosure geometry, surrounding materials, experimental structures, and the motion of the tracked animal itself. These factors alter the received waveform and directly affect the stability of time-based ranging measurements [11, 10, 1].

4.2.1 Line-of-Sight and Non-Line-of-Sight Conditions

Under line-of-sight (LoS) conditions, the direct path between the transmitter and receiver remains unobstructed, allowing the system to estimate propagation time with high precision. In such settings, the first-arriving path is generally the physically correct path, which improves the reliability of both range estimation and final coordinate reconstruction [1].

Under non-line-of-sight (NLoS) conditions, however, physical obstacles such as enclosure boundaries, shelters, equipment, or other structures may block the direct path. In these cases, the receiver detects signals that have propagated through longer reflected or diffracted routes, thereby introducing a positive ranging bias. This effect can be represented as

$$d_{\text{measured}} = d_{\text{true}} + \epsilon_{\text{NLoS}} \tag{4.1}$$

where d_{measured} is the measured range, d_{true} is the true geometric distance, and ϵ_{NLoS} is the propagation bias introduced by obstructed transmission [1, 11]. Unlike zero-mean noise, this error is systematic and can significantly distort the reconstructed trajectory if not mitigated.

4.2.2 Multipath Propagation and Attenuation

In indoor environments, the received UWB waveform often includes multiple delayed components caused by reflections from walls, floors, and nearby objects. This phenomenon, known as multipath propagation, complicates identification of the true direct path and can degrade ranging precision. A general representation of the received signal is

$$r(t) = \sum_{i=1}^{N} a_i s(t - \tau_i) \tag{4.2}$$

where a_i and τ_i denote the amplitude and delay of the ith propagation path [11, 10]. When reflected components arrive within a short interval, the waveform may become distorted in a manner that shifts the detected arrival time and increases measurement variance.

Attenuation further compounds this problem. As signal distance increases, and as the signal penetrates or reflects from materials in the environment, the received signal strength decreases. This degradation is commonly approximated by a path-loss model of the form

$$S = S_0 - 10n \log_{10}(d) \tag{4.3}$$

where S is the signal strength, S_0 is the reference strength, n is the path-loss exponent, and d is the propagation distance [1]. Reduced signal strength lowers measurement quality and increases sensitivity to timing error, especially under NLoS conditions.

4.2.3 Signal-to-Noise Ratio and Channel Quality

SNR is a central indicator of UWB channel quality because it directly affects the clarity of the received signal and the stability of timing estimation. In general form,

$$\mathrm{SNR}(t) = \frac{P_{\mathrm{signal}}(t)}{P_{\mathrm{noise}}(t)} \tag{4.4}$$

where $P_{\mathrm{signal}}(t)$ and $P_{\mathrm{noise}}(t)$ represent the signal and noise power at time t, respectively [1, 2]. Higher SNR corresponds to stronger signal clarity and typically leads to lower positioning error, whereas reduced SNR increases uncertainty in range measurements.

The relationship between channel quality and theoretical information capacity can be expressed using the Shannon–Hartley theorem,

$$C = B \log_2 \left(1 + \frac{S}{N}\right) \tag{4.5}$$

where C is the channel capacity, B is the effective bandwidth, and S/N is the SNR. This relationship provides a physical interpretation of why reduced SNR under NLoS conditions can degrade localization reliability: lower channel quality reduces the available information for stable timing and position estimation [2].

In the implemented rodent-tracking system, LoS environments generally preserve stronger signal quality, while NLoS environments reduce effective SNR because of attenuation, multipath propagation, and partial obstruction. In the experimental results, LoS measurements maintained SNR values in the approximate range of 20–25 dB, whereas NLoS measurements dropped to approximately 8–15 dB. This reduction in channel quality is important because

it not only degrades engineering performance metrics such as ranging accuracy, but also propagates into the quality of reconstructed behavioral trajectories [1, 2].

4.3 Localization Performance Evaluation

The localization performance of the UWB-based real-time locating system was evaluated under both LoS and NLoS conditions to determine how propagation regime affects tracking fidelity in practical laboratory environments [1, 2].

Experimental observations showed that localization accuracy was substantially higher under LoS conditions than under NLoS conditions. Under LoS conditions, the system achieved a mean localization error of 2.28 centimeters with a standard deviation of 0.76 centimeters across 267 samples. Under NLoS conditions, the mean localization error increased to 4.40 centimeters with a standard deviation of 1.03 centimeters across 220 samples, with a maximum observed deviation of 6.2 centimeters. These results confirm that UWB tracking remains effective under both regimes, while also demonstrating that obstruction and multipath propagation produce measurable changes in trajectory reliability [1].

4.3.1 LoS Accuracy Assessment

Under LoS conditions, the implemented system demonstrated stable and accurate localization performance. Because the direct path between the tag and anchors remained unobstructed, the ranging process was dominated by the physically correct arrival path, enabling more reliable coordinate estimation. This produced relatively low error and reduced variance across the tracking region, supporting the use of UWB localization for fine-grained rodent movement analysis [1].

The strong LoS performance also validates the anchor deployment strategy and system design introduced in Chapter 3. In particular, the perimeter anchor geometry and real-time

ranging workflow provided a stable basis for reconstructing trajectories within the experimental enclosure. Such performance is especially important in behavioral studies, where small spatial errors may influence inferred zone occupancy, movement smoothness, event timing, and state classification [1, 2].

4.3.2 NLoS Error Characterization

Under NLoS conditions, localization performance degraded because the direct propagation path was partially or fully obstructed. In these settings, the received signal was more strongly influenced by reflections, attenuation, and indirect paths, which introduced additional uncertainty into the ranging process. As a result, reconstructed position estimates exhibited greater deviation and variance than those observed in LoS conditions [1].

This degradation is particularly relevant in rodent experiments because animals naturally move near walls, enter enclosed regions, and interact with structures that create temporary obstruction. Therefore, NLoS effects should be regarded as an expected component of realistic behavioral tracking rather than as rare exceptions. A practical UWB-based tracking system must therefore remain robust under such conditions if it is to provide behaviorally meaningful measurements [1, 2].

4.3.3 Comparative Error Analysis

The contrast between LoS and NLoS performance demonstrates that localization error in indoor UWB systems is regime-dependent rather than uniform across the experiment. LoS measurements are characterized by lower error and tighter distributions, whereas NLoS measurements exhibit larger deviations and higher dispersion due to propagation distortion [1]. This difference is important because average error alone may obscure the structure of the sensing regime.

From a behavioral perspective, comparative error analysis is essential because different sens-

ing conditions may produce different interpretations of the same physical motion. Increased error variance may fragment trajectories, shift apparent positions relative to behavioral zones, and alter inferred transition patterns. Consequently, comparative performance analysis provides a critical bridge between engineering validation and the biological reliability of trajectory-based behavioral metrics [2].

Taken together, these results show that localization performance must be evaluated not only by aggregate engineering metrics but also by its effect on the interpretability of reconstructed behavioral trajectories. This motivates the signal-enhancement and error-mitigation methods discussed in the following section.

4.4 Signal Enhancement and Error Mitigation

Because indoor UWB localization is strongly influenced by propagation distortion, raw ranging measurements alone are often insufficient for reliable behavioral tracking. To preserve trajectory quality under realistic laboratory conditions, the implemented system incorporates signal-enhancement and error-mitigation strategies that reduce the influence of timing noise, multipath effects, and non-line-of-sight (NLoS) bias. These methods are essential not only for improving engineering accuracy, but also for maintaining the behavioral validity of the reconstructed trajectory [1, 2].

4.4.1 Filtering and Adaptive Estimation

A primary component of the mitigation framework is recursive filtering, which improves localization stability by combining sequential measurements over time. Rather than treating each range estimate independently, recursive filtering exploits temporal continuity in the trajectory and suppresses short-term fluctuations caused by measurement noise. In practical rodent tracking, this is particularly useful because the animal moves continuously and the underlying trajectory is expected to vary smoothly over short time intervals even when raw

measurements contain abrupt perturbations [1].

Kalman-based estimation is especially effective in this context because it provides a computationally efficient mechanism for real-time trajectory smoothing while preserving temporal responsiveness. By updating the state estimate at each time step using both the current measurement and the prior prediction, the filter reduces variance in the reconstructed path and improves the reliability of downstream spatial analysis [1].

4.4.2 Phase-Based and Compensation Approaches

In addition to filtering, the framework incorporates signal-aware compensation approaches that improve ranging robustness under degraded channel conditions. In particular, phase-based normalized-distance estimation was used to reduce the effects of propagation distortion, especially in NLoS environments where conventional time-of-flight methods are more vulnerable to indirect-path bias [1].

The phase-based approach improved localization performance relative to traditional time-of-flight estimation. Under LoS conditions, phase-based estimation achieved an error of 0.5 $\pm$ 0.1 centimeters, compared with 1.0 $\pm$ 0.2 centimeters for the time-of-flight method. Under NLoS conditions, phase-based estimation achieved an error of 2.0 $\pm$ 0.4 centimeters, compared with 3.5 $\pm$ 0.8 centimeters for the time-of-flight method. These improvements indicate that phase-based estimation provides greater robustness to multipath and environmental distortion [1].

These approaches are important because NLoS error is not purely random; it often contains systematic bias associated with obstruction and indirect propagation. As a result, simple averaging is not sufficient to restore accuracy. Compensation methods that account for signal characteristics, propagation behavior, or channel quality can therefore provide meaningful improvement beyond basic filtering alone. In the broader context of this dissertation, such

approaches support the transition from raw signal acquisition to behaviorally interpretable trajectory reconstruction [1, 2].

4.4.3 Improvement Under Degraded Conditions

The practical value of signal-enhancement methods becomes most apparent under degraded conditions, where obstruction, attenuation, and multipath effects would otherwise produce unstable localization output. Experimental observations showed that adaptive estimation and signal-aware correction improved tracking robustness by reducing localization error and preserving greater consistency across challenging measurement regimes [1].

Adaptive pulse repetition frequency control was used to maintain signal integrity under obstructed environments and improve positioning accuracy. Together with Kalman-based correction and phase-based estimation, this approach provided a layered mitigation strategy for preserving trajectory quality under indoor conditions [1].

This improvement is significant for behavioral applications because degraded sensing conditions can propagate directly into behavioral uncertainty. If localization enhancement is not applied, the resulting trajectory may become noisier, more fragmented, or systematically biased, which in turn affects zone occupancy, inferred state segmentation, and transition-based behavioral models. Thus, signal-enhancement methods should be interpreted not merely as engineering optimizations, but as necessary steps for preserving the scientific interpretability of trajectory-derived behavioral measures [2].

Overall, the results indicate that robust indoor rodent tracking requires a layered mitigation strategy in which hardware capability, signal-aware compensation, adaptive estimation, and recursive filtering work together to sustain localization quality. This perspective is especially important in the present dissertation because the downstream behavioral framework depends on trajectory fidelity as its primary data source.

4.5 Implications for Behavioral Tracking

The localization results presented in this chapter show that signal integrity is directly linked to the quality of trajectory-based behavioral analysis. In rodent experiments, the recorded position sequence is not merely an engineering output; it is the primary observational basis for identifying behavioral states, quantifying movement patterns, and inferring transitions between activities. Consequently, degradation in localization quality may propagate into the behavioral interpretation derived from those trajectories [2, 3].

Under high-quality LoS conditions, stable ranging and low localization error support smoother and more reliable trajectory reconstruction. This improves the accuracy of downstream measures such as zone occupancy, movement continuity, and event timing. Under NLoS conditions, however, increased error variance and systematic bias can fragment the reconstructed path, alter spatial relationships, and reduce confidence in inferred behavioral states [1, 2].

This distinction is especially important because many behavioral metrics depend on precise spatial and temporal information. Small shifts in reconstructed position may move an animal across a decision boundary between behavioral zones, while noisy coordinates may introduce spurious transitions or reduce apparent state persistence. Thus, localization error should not be viewed only as a technical limitation; it is also a source of behavioral uncertainty [2].

The key implication is that LoS and NLoS sensing regimes should be treated as structured measurement conditions rather than incidental noise. This interpretation motivates the probabilistic and generative analyses developed in Chapters 5 and 6, where sensing quality is modeled as part of behavioral inference rather than as a separate engineering artifact [2, 3].

4.6 Chapter Summary

This chapter examined how indoor signal-propagation conditions affect the localization performance of the UWB-based real-time locating system. The analysis focused on LoS/NLoS sensing regimes, multipath effects, attenuation, and signal-to-noise ratio degradation as key factors shaping trajectory reliability.

The results demonstrated that LoS tracking achieved lower error and greater stability, whereas NLoS tracking introduced larger deviations due to obstruction, attenuation, and multipath distortion. Phase-based estimation, adaptive control, and Kalman-based correction improved robustness under degraded conditions, supporting centimeter-level trajectory reconstruction for rodent behavioral studies [1].

Most importantly, this chapter established that localization quality directly affects the reliability of trajectory-derived behavioral measures. Within the probability-of-motion framework, signal integrity is therefore not treated as an external engineering factor, but as part of the measurement process that shapes observed animal motion. This perspective provides the foundation for Chapter 5, where probabilistic models quantify the connection between signal integrity and behavioral dynamics.

Chapter 5

Probabilistic Behavioral Modeling From UWB Trajectories

5.1 Introduction

This chapter develops the probabilistic behavioral modeling framework derived from reconstructed Ultra-Wideband (UWB) trajectories. While the previous chapters established the tracking system, localization workflow, and LoS/NLoS signal-performance characteristics, this chapter examines how measured trajectories can be transformed into interpretable behavioral models and how sensing quality affects that interpretation.

This chapter treats UWB trajectories as probabilistic information sources from which behavioral state structure, uncertainty, and transition dynamics can be quantified [2].

The key contribution of this chapter is to show that LoS/NLoS signal differences propagate into behavioral inference by altering uncertainty, threshold reliability, state persistence, and transition structure [2, 1].

The chapter introduces four components of this framework: behavioral state segmentation, Shannon entropy for uncertainty quantification, Bernoulli modeling for threshold-based localization reliability, and Markov chains for sequential behavioral dynamics. Together, these tools transform UWB trajectories into interpretable probabilistic models and provide the foundation for the sensor-aware trajectory-level digital twin developed in Chapter 6 [3].

5.2 Behavioral State Segmentation

An initial step in probabilistic behavioral analysis is the conversion of continuous trajectories into discrete behavioral states. This transformation allows the position sequence to be represented as a symbolic time series, making it possible to apply information-theoretic and stochastic models to behavior. In this framework, the behavioral trajectory is segmented into four ethologically meaningful states: resting, exploring, feeding, and burrowing [2].

The segmentation process relies on a combination of kinematic thresholds and spatial-zone definitions. Velocity, dwell time, and arena context are used together to distinguish sustained inactivity from active locomotion and to separate general exploration from resource-oriented or shelter-oriented activity. The operational definitions used to assign the four behavioral states were introduced earlier in Table 1.1 and are used consistently throughout this chapter.

This representation preserves the major structure of rodent behavior while reducing the trajectory to a form that is tractable for probabilistic analysis. The resulting state sequence enables estimation of occupancy, uncertainty, threshold reliability, transition probability, state persistence, and long-term behavioral organization.

Because state segmentation depends on reconstructed positions, its reliability is influenced by localization accuracy. Coordinate perturbations under degraded sensing conditions may shift zone occupancy, fragment trajectories, or introduce spurious state switches [2, 3].

This state representation provides the basis for the probabilistic analyses developed in the following sections, where uncertainty, threshold-based success, and transition dynamics are quantified directly from the segmented trajectory sequence.

5.3 Shannon Entropy and Uncertainty Quantification

A useful way to quantify uncertainty in trajectory-derived behavioral analysis is through Shannon entropy. In this framework, entropy measures the dispersion or unpredictability of a probability distribution and can characterize localization error variability as well as uncertainty in derived behavioral patterns. Because degraded sensing conditions broaden the error distribution and destabilize the reconstructed trajectory, entropy provides a natural way to connect signal integrity with behavioral reliability [2].

For a discrete probability distribution, Shannon entropy is defined as follows, following the standard information-theoretic formulation of uncertainty [41, 42].

$$H = -\sum_{i=1}^{n} p_i \log_2 p_i \tag{5.1}$$

where p_i denotes the probability associated with the ith outcome. Larger entropy values correspond to greater uncertainty and more dispersed distributions, whereas smaller values indicate greater predictability and concentration [41, 2].

In the present context, entropy is useful at two levels. First, it characterizes uncertainty in localization error distributions under different sensing conditions. When LoS propagation is preserved, the resulting error distribution tends to remain narrower and more predictable. Under NLoS conditions, distortion from obstruction, attenuation, and multipath propagation increases the variability of the localization error, which raises entropy [2, 1].
Second, entropy provides insight into behavioral variability itself. If the reconstructed trajectory becomes unstable or fragmented because of sensing degradation, then the inferred

behavioral sequence may also become less predictable. In this way, entropy serves as a bridge between physical-layer measurement quality and higher-level behavioral interpretation.
In the experimental analysis, NLoS conditions produced a clear increase in localization-error entropy relative to LoS conditions. Specifically, localization-error entropy increased from 1.15 Bits under LoS to 1.78 Bits under NLoS. Transition entropy also increased under NLoS conditions, rising to 1.28 ± 0.04 Bits compared with 1.10 ± 0.03 Bits under LoS. These results demonstrate that degraded signal conditions measurably increase uncertainty in both localization and behavioral-transition structure [2].

Thus, Shannon entropy provides an interpretable metric for evaluating how much uncertainty is introduced into the behavioral analysis pipeline by sensing imperfections. This makes it a foundational tool in the probabilistic modeling framework developed in this chapter.

5.4 Bernoulli Modeling of Accuracy Thresholds

While average localization error provides a useful summary of system performance, it does not directly indicate how often the tracking system meets the precision required for behavioral analysis. In many rodent experiments, the practical question is whether the localization estimate is sufficiently accurate to preserve valid interpretation of behavioral zones, movement continuity, and event timing. For this reason, localization success is modeled as a binary outcome relative to a predefined accuracy threshold [2].

In this formulation, each localization estimate is classified as either successful or unsuccessful depending on whether its error remains within an acceptable tolerance. This can be expressed as a Bernoulli random variable:

$$X = \begin{cases} 1, & \text{if } |e| \leq \epsilon \\ 0, & \text{if } |e| > \epsilon \end{cases} \tag{5.2}$$

where e denotes the localization error and ϵ is the application-specific tolerance threshold [2]. In this dissertation, a 20-millimeter threshold is used as a behaviorally meaningful tolerance for evaluating localization reliability.

The empirical success probability is estimated as

$$\hat{p} = \frac{1}{N} \sum_{k=1}^{N} X_k \tag{5.3}$$

where N is the number of localization samples and X_k is the Bernoulli outcome associated with the kth sample [43, 2]. This quantity provides a measure of how often the tracking system satisfies the spatial precision needed for meaningful behavioral interpretation.

The experimental results show that sensing regime strongly affects threshold reliability. Under LoS conditions, the probability of achieving sub-20-millimeter accuracy was 63.2 percent. Under NLoS conditions, this probability shifted to 27.7 percent. This result is important because it translates signal degradation into an interpretable reliability measure: NLoS does not merely increase average error, but reduces the probability that the system remains within a behaviorally useful accuracy range [2].

Accordingly, Bernoulli modeling complements entropy analysis by focusing not on uncertainty in the full distribution, but on the probability of satisfying a behaviorally meaningful accuracy requirement. Together, these two views provide a richer description of localization reliability than deterministic error metrics alone.

5.5 Markov Chains for Behavioral Transitions

To model sequential dependencies in rodent behavior, this dissertation adopts a first-order Markov chain representation. In this framework, the probability of the next behavioral state depends on the current state, making it possible to summarize the temporal structure of the segmented trajectory in a compact and interpretable form [2, 3].

For a behavioral sequence, the transition probability from state i to state j is

$$P_{ij} = \Pr(X_{t+1} = j \mid X_t = i) \tag{5.4}$$

where X_t denotes the behavioral state at time t, and P_{ij} is the probability of transitioning from state i to state j. Collectively, these probabilities form a row-stochastic transition matrix that characterizes the short-term structure of the behavioral process [44, 2].

An example transition-probability matrix based on these behavioral states was presented earlier in Table 1.2. If $N(s_i \to s_j)$ denotes the number of observed transitions from state s_i to state s_j, then the corresponding transition probability may be estimated as

$$P_{ij} = \frac{N(s_i \to s_j)}{\sum_{k=1}^{n} N(s_i \to s_k)} \tag{5.5}$$

where the denominator represents the total number of observed transitions originating from state s_i [44, 2]. This construction allows the dynamics of exploration, resting, feeding, and burrowing to be represented quantitatively in matrix form.

Markov modeling is useful because it preserves the temporal organization of behavior while remaining directly interpretable. The transition matrix can reveal dominant cycles, preferred behavioral sequences, and the relative persistence of each state. In the experimental results, the strongest baseline behavioral transitions included Exploring $\to$ Feeding and Feeding $\to$ Resting, indicating a recurrent exploratory–feeding–resting behavioral structure [2].

A related quantity is the stationary distribution, which represents the long-term fraction of time the system spends in each behavioral state. It is defined by

$$\pi P = \pi, \qquad \sum_{i=1}^{n} \pi_i = 1 \tag{5.6}$$

where π is the stationary vector and P is the transition matrix [44, 2]. In the framework, the stationary distribution provides a probabilistic summary of long-term behavioral occupancy

and supports comparison between modeled and empirically observed state prevalence.

The Markov results show that sensing quality affects sequential behavioral interpretation. Average behavioral state persistence decreased from 68 percent under LoS to 52 percent under NLoS, indicating more fragmented behavioral sequences when tracking fidelity declined. In addition, the model achieved reproducible performance across sessions, with prediction accuracy of 82.9 $\pm$ 1.0 percent and mean Kullback–Leibler divergence of 0.037 $\pm$ 0.003 between model-predicted and empirical stationary distributions [2].

These results demonstrate Markov analysis does more than summarize behavioral transitions. It provides a mechanism for showing how signal degradation propagates into the sequential organization of inferred behavior. Together, the transition matrix and stationary distribution form part of the probabilistic behavioral framework developed in this dissertation.

5.6 Sensing Quality in Behavioral Modeling

The probabilistic framework in this chapter shows that sensing quality affects the segmented behavioral sequence and probabilistic structure inferred from trajectories [2, 3].

The combined results from entropy, Bernoulli reliability modeling, and Markov transition analysis establish a consistent LoS-to-NLoS sensing signature. Under NLoS conditions, uncertainty increases, threshold reliability decreases, and behavioral state persistence is reduced. These shifts are not treated as isolated tracking failures; rather, they demonstrate that sensing condition is a measurable component of behavioral inference.

This effect is especially important in rodent behavioral studies because many interpretations depend on subtle changes in persistence, transition frequency, and spatial preference. If the sensing process is not explicitly accounted for, behavioral differences may be attributed to biology when they are in fact driven partly by measurement distortion. Therefore, sensing

quality should be treated as an integral part of the behavioral modeling pipeline rather than as an independent engineering concern [2].

From this perspective, entropy, Bernoulli success modeling, and Markov transition analysis are not merely descriptive tools; they are diagnostic tools for evaluating how robustly behavior can be inferred from the sensed trajectory. This observation motivates the next chapter, which extends the probability-of-motion framework toward sensor-aware generative modeling and trajectory-level digital twins, where synthetic trajectories are generated under explicit LoS/NLoS observation conditions.

5.7 Chapter Summary

This chapter developed the probabilistic behavioral modeling framework derived from reconstructed UWB trajectories. It introduced behavioral state segmentation as the basis for converting continuous trajectories into symbolic behavioral sequences and then presented Shannon entropy, Bernoulli success modeling, and first-order Markov chains as complementary tools for quantifying uncertainty, threshold-based reliability, and transition structure.

The chapter showed that these probabilistic representations are closely linked to sensing quality. Under NLoS conditions, localization-error entropy increased from 1.15 bits to 1.78 bits, the probability of achieving localization accuracy within 20 millimeters decreased from 63.2 percent under LoS conditions to 27.7 percent under NLoS conditions, and average behavioral state persistence decreased from 68 percent under LoS conditions to 52 percent under NLoS conditions. These results demonstrate that sensing degradation propagates into behavioral uncertainty, reliability, and sequential organization [2].

The main achievement of this chapter is the establishment of a physics-to-behavior modeling framework that treats UWB trajectories as probabilistic information sources for extracting interpretable behavioral markers. Within the probability-of-motion framework, this chapter

shows that measured animal motion contains both behavioral structure and sensing-induced uncertainty. This perspective provides the foundation for the sensor-aware trajectory-level digital twin framework developed in the next chapter.

Chapter 6

Sensor-Aware Generative Modeling and Virtual Behavioral Simulation

Portions of this chapter are adapted from the published article: Reza Sayfoori, Goli Vaisi, and Hung Cao, "Virtual Mice, Real Errors: A Sensor-Aware Generative Framework for In Silico Ethology," Sensors*, vol. 26, no. 10, article 2977, 2026.*

6.1 Introduction

This chapter extends the probability-of-motion framework beyond physical tracking and probabilistic interpretation toward sensor-aware generative modeling, virtual cohort generation, and trajectory-level digital twin modeling. Here, motion is treated as a generative probabilistic process observed through a sensing channel rather than as a fixed coordinate sequence alone. This direction is also consistent with the broader motivation of the 3Rs framework, particularly the use of computational and simulation-based methods to reduce experimental burden while preserving scientific interpretability [45]. The previous chapters established the UWB-based tracking system, characterized its LoS/NLoS signal-performance

behavior, and developed probabilistic behavioral models from reconstructed trajectories. Building on these foundations, this chapter examines how experimentally grounded observations can be transformed into synthetic yet behaviorally meaningful virtual trajectories that preserve both latent behavioral structure and sensing-regime-dependent measurement distortion [3].

The central idea is that realistic behavioral simulation should preserve both the latent structure of animal motion and the sensing-related distortions through which that motion is observed. A useful virtual behavioral model must therefore distinguish ethological dynamics from measurement effects introduced by Line-of-Sight (LoS) and Non-Line-of-Sight (NLoS) sensing conditions. This distinction forms the basis of the sensor-aware generative framework developed in this chapter [3, 2].

In this dissertation, the term trajectory-level digital twin is used in a bounded and measurement-aware sense. It refers to a probabilistic model of behavioral trajectories and the sensing process through which they are measured, rather than a complete biological representation of the animal. This definition keeps the digital-twin claim focused on motion, uncertainty, and sensor-conditioned observation.

6.2 Sensor-Aware Generative Framework

The trajectory-level digital twin developed in this chapter separates latent behavioral dynamics from sensing-induced distortion so that synthetic rodent trajectories remain both behaviorally plausible and measurement-realistic [3]. Figure 6.1 illustrates the overall framework and shows how measured trajectories are transformed into sensor-aware virtual behavioral data.

At a high level, the framework begins with UWB trajectories collected under LoS and NLoS conditions. These trajectories are decomposed into two components: a latent behavioral

branch, which models underlying ethological dynamics, and an observation branch, which models how sensing conditions distort latent motion in the measured domain. The latent branch captures behavioral state evolution and kinematic structure, whereas the observation branch captures sensing-regime-dependent bias, variability, and outlier behavior [3].

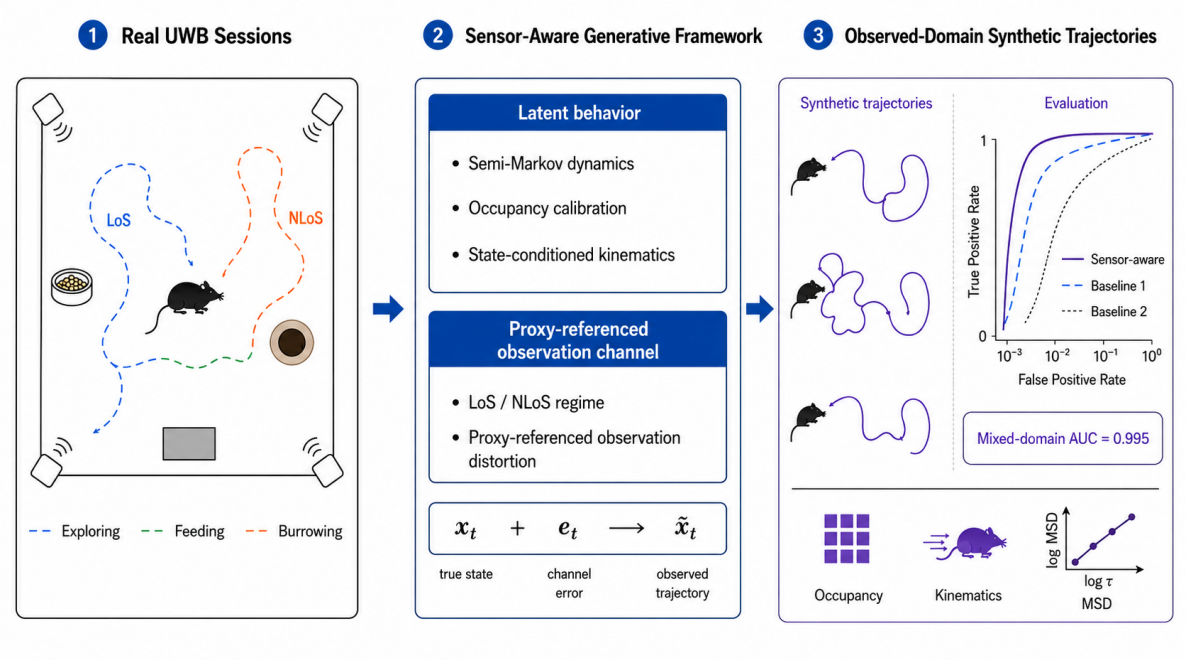


Figure 6.1: Graphical overview of the sensor-aware generative framework for in silico ethology. Real UWB sessions acquired under LoS and NLoS conditions are used to estimate latent behavioral dynamics and a proxy-referenced observation channel. The framework then generates observed-domain synthetic trajectories and evaluates them using occupancy, state-conditioned kinematics, mean-squared displacement, and downstream robustness under sensing-domain shift. Adapted from [3].

This separation allows synthetic trajectories to reflect both behavioral structure and sensing reality, providing the basis for the robustness analyses developed later in this chapter [3, 2].

6.3 Latent Behavioral Dynamics and State Modeling

The sensor-aware generative framework models latent behavioral motion separately from the sensing process through which motion is observed. This distinction allows the model to

preserve ethological structure while representing LoS/NLoS-dependent distortion [3].

To capture behavioral organization, the framework models the trajectory as a sequence of discrete behavioral states with state-dependent motion statistics. In the present study, the latent state space includes three principal behavioral modes: exploring, feeding, and burrowing [3]. These states are chosen because they capture the dominant trajectory-defining behavioral patterns in the available UWB dataset while providing a compact representation that remains stable under limited-data conditions.

At the latent level, the behavioral process is modeled as a stochastic state sequence that governs how the animal moves through time. Let z_t denote the latent behavioral state at time t, let $\mathbf{x}_t \in \mathbb{R}^2$ denote the latent planar position, and let $\tilde{\mathbf{x}}_t \in \mathbb{R}^2$ denote the observed UWB position, which is treated as a distorted version of the latent motion. This causal structure may be expressed as

$$z_t \rightarrow \mathbf{x}_t \rightarrow \tilde{\mathbf{x}}_t \tag{6.1}$$

where the latent state determines the behavioral dynamics, and the observation layer transforms those dynamics into the sensor-measured trajectory [3].

This factorization prevents sensing artifacts from being treated as intrinsic behavior and supports more principled simulation, interpretation, and robustness analysis [3, 2].

6.3.1 State Space Representation

The discrete latent state space is defined as

$$\mathcal{Z} = \{1, 2, 3\} \equiv \{\text{exploring}, \text{feeding}, \text{burrowing}\} \tag{6.2}$$

where each state corresponds to a behaviorally meaningful mode of motion [3]. This reduced representation differs from the earlier four-state descriptive framework because the present objective is generative robustness rather than direct one-to-one reproduction of the full behavioral taxonomy.

The reduced state space improves stability in transition estimation, occupancy calibration, and state-conditioned motion synthesis. In particular, it avoids sparsity in rarely represented or near-stationary modes that would otherwise degrade the reliability of the generative model under limited experimental data [3].

6.3.2 Run-Based Behavioral Dynamics

To model temporal structure more faithfully, the latent state sequence is represented at the level of runs rather than frame-by-frame labels. Consecutive identical state labels are grouped into a single run, so that the process is described by both the state identity and its duration. This enables the model to account explicitly for state persistence rather than relying only on repeated self-transitions, which is consistent with semi-Markov and run-duration-based representations of stochastic behavioral dynamics [46, 3].

If run boundaries are denoted by $\tau_1, \tau_2, \ldots, \tau_K$, the kth run is represented by

$$(s_k, d_k), \qquad s_k = z_{\tau_k}, \qquad d_k = \tau_{k+1} - \tau_k \tag{6.3}$$

where s_k is the state associated with the run and d_k is its duration measured in frames [3]. This representation supports semi-Markov modeling, where transition structure and dwell-time behavior are estimated separately.

6.3.3 Behavioral Occupancy and Stability

A useful summary of the latent behavioral process is the frame-level occupancy of each state. If the empirical occupancy of state i is denoted by π_i^{real}, then

$$\pi_i^{\text{real}} = \frac{1}{T_{\text{real}}} \sum_{t=1}^{T_{\text{real}}} \mathbf{1}\{z_t = i\} \tag{6.4}$$

where T_{real} is the total number of frames and $\mathbf{1}\{\cdot\}$ is the indicator function [3]. Matching this occupancy structure is important because it ensures that the synthetic trajectories preserve the broad time allocation of the real behavioral data.

Thus, latent behavioral modeling in this framework is not limited to generating plausible instantaneous movement. It is designed to preserve higher-level temporal structure, including state persistence, duration, and occupancy. These properties provide the foundation for the state-conditioned kinematic synthesis described in the next section.

6.4 State-Conditioned Kinematic Generation

Once the latent behavioral dynamics have been defined, the next step is to generate motion trajectories that are consistent with the kinematic characteristics of each behavioral state. In this framework, movement is not generated from a single global motion model. Instead, trajectories are synthesized using state-conditioned motion priors, allowing exploring, feeding, and burrowing to exhibit distinct spatial dynamics. This approach preserves behavior-specific movement structure while maintaining consistency with the higher-level state sequence described in the previous section [3].

The central idea is that each latent state is associated with its own distribution of local movement increments. These increments determine both the step length of the animal and the change in movement direction. By conditioning these quantities on the active behavioral state, the generator preserves interpretable differences between fast exploratory motion,

localized feeding activity, and spatially constrained burrowing behavior [3].

6.4.1 Reference Trajectory Construction

Because the observed UWB trajectory contains sensing distortion, the framework first constructs a smoothed reference trajectory that serves as a proxy for the latent behavioral motion. The transformation from the observed UWB position to the smoothed reference trajectory is expressed as

$$\tilde{\mathbf{x}}_t = \begin{bmatrix} \tilde{x}_t \\ \tilde{y}_t \end{bmatrix} \xrightarrow{\mathcal{R}_w} \mathbf{x}_t = \begin{bmatrix} x_t \\ y_t \end{bmatrix}, \tag{6.5}$$

where $\tilde{\mathbf{x}}_t$ denotes the observed UWB position at time t, $\mathbf{x}_t$ denotes the reference trajectory, and $\mathcal{R}_w(\cdot)$ represents a two-stage smoothing operator with window length w.
The smoothing operator is defined as

$$\mathcal{R}_w(s_t) = \mathrm{MA}_w(\mathrm{Med}_w(s_t)), \tag{6.6}$$

where $\mathrm{Med}_w(\cdot)$ denotes a centered rolling median and $\mathrm{MA}_w(\cdot)$ denotes a centered rolling mean [3]. Applied coordinate-wise, the reference trajectory is obtained as

$$x_t = \mathcal{R}_w(\tilde{x}_t), \qquad y_t = \mathcal{R}_w(\tilde{y}_t). \tag{6.7}$$

This smoothed trajectory is not interpreted as perfect ground truth. Rather, it serves as a least-distorted proxy for estimating behavioral motion priors and residual sensing effects.

6.4.2 Kinematic Feature Extraction

From the reference trajectory, the framework extracts local motion descriptors that characterize state-dependent movement. For $t \geq 2$, the displacement vector is defined as

$$\Delta \mathbf{x}_t = \mathbf{x}_t - \mathbf{x}_{t-1} = \begin{bmatrix} \Delta x_t \\ \Delta y_t \end{bmatrix} \tag{6.8}$$

and the corresponding step length is

$$\ell_t = \|\Delta \mathbf{x}_t\|_2 \tag{6.9}$$

The heading angle is given by

$$\phi_t = \text{atan2}(\Delta y_t, \Delta x_t) \tag{6.10}$$

and the turning angle is defined as

$$\theta_t = \text{wrap}(\phi_t - \phi_{t-1}) \tag{6.11}$$

where $\text{wrap}(\cdot)$ maps angular differences into the interval $(-\pi, \pi]$ [3]. Together, ℓ_t and θ_t summarize the local kinematic behavior of the animal and form the basis of the state-conditioned priors.

6.4.3 State-Conditioned Motion Priors

For each behavioral state $i \in \mathcal{Z}$, the generator constructs empirical priors over step length and turning angle. Let

$$L_i = \{\ell_t : z_t = i\}, \qquad T_i = \{\theta_t : z_t = i\} \tag{6.12}$$

denote the sets of step-length and turning-angle samples associated with state i [3]. During synthesis, the motion increments are sampled conditionally as

$$\ell_t \mid (z_t = i) \sim \hat{p}(L_i), \qquad \theta_t \mid (z_t = i) \sim \hat{p}(T_i), \tag{6.13}$$

where $\hat{p}(\cdot)$ denotes the empirical distribution induced by the corresponding sample set. This formulation allows each state to retain its characteristic movement statistics while remaining directly grounded in the real experimental data.

6.4.4 Trajectory Propagation

Given sampled motion increments, the latent trajectory is propagated recursively through the heading and position update equations. The heading evolves as

$$\phi_t \leftarrow \phi_{t-1} + \theta_t, \tag{6.14}$$

and the planar position is updated according to

$$\mathbf{x}_t \leftarrow \mathbf{x}_{t-1} + \ell_t \begin{bmatrix} \cos(\phi_t) \\ \sin(\phi_t) \end{bmatrix} \tag{6.15}$$

These equations generate the latent motion path corresponding to the behavioral state sequence. Because increments are conditioned on state identity, the resulting trajectory preserves interpretable differences in movement style across behavioral modes [3].

6.4.5 Outlier Control and Long-Horizon Stability

To improve stability during long-horizon synthesis, the framework limits extreme increments using percentile-based caps. This prevents rare but unusually large motion samples from destabilizing the trajectory. For each state i, high-percentile thresholds are defined for step

length and turning angle, and sampled increments are truncated when necessary [3].

This stabilization step is important because virtual trajectories support robustness studies and synthetic cohort generation over extended durations. Without control, a small number of extreme samples could dominate the simulation and reduce realism. By combining state-conditioned priors with percentile-based stabilization, the framework produces trajectories that are both behaviorally plausible and numerically stable.

Thus, state-conditioned kinematic generation provides the bridge between latent behavioral states and continuous spatial motion. The next section extends this formulation by modeling how the latent trajectory is transformed into a sensor-observed trajectory through a regime-dependent observation channel.

6.5 Sensor-Aware Observation Channel

After generating the latent trajectory, the model maps it into the observed Ultra-Wideband (UWB) measurement domain through a sensor-aware observation channel that represents LoS/NLoS-dependent distortion [3]. Using the trajectory notation introduced in the preceding section, the observation model is written as

$$\tilde{\mathbf{x}}_t = \mathbf{x}_t + \mathbf{e}_t, \tag{6.16}$$

where $\mathbf{x}_t$ denotes the latent planar trajectory, $\tilde{\mathbf{x}}_t$ denotes the observed UWB trajectory, and $\mathbf{e}_t$ is the sensing distortion term [3]. This distortion captures the discrepancy between the latent behavioral trajectory and the measured UWB trajectory.

6.5.1 State- and Regime-Dependent Distortion

The sensing distortion is modeled as a function of both behavioral state and sensing regime. This is important because different behaviors may occupy different spatial regions and may

therefore interact differently with the sensing environment. In addition, LoS and NLoS conditions produce distinct error characteristics, with NLoS observations generally exhibiting larger bias and heavier-tailed deviations [3, 2].

Conditioned on behavioral state $z_t = i$ and sensing regime $c \in \{\text{LoS}, \text{NLoS}\}$, the distortion is modeled as

$$\mathbf{e}_t \mid (z_t = i, c) = \boldsymbol{\mu}_{i,c} + \mathbf{u}_t \tag{6.17}$$

where $\boldsymbol{\mu}_{i,c}$ is a state- and regime-dependent bias term, and $\mathbf{u}_t$ is a zero-centered random component [3]. This formulation allows the model to represent both systematic distortion and random variability.

6.5.2 Mixture Representation of Observation Error

To capture both nominal measurement noise and occasional extreme deviations, the stochastic distortion component is modeled using an outlier-aware mixture distribution, consistent with robust statistical treatment of heavy-tailed deviations [47, 3].

$$\mathbf{u}_t \sim (1 - \rho_{i,c})\,\mathcal{N}(\mathbf{0}, \Sigma_{i,c}) + \rho_{i,c}\,\mathcal{T}_\nu(\mathbf{0}, \Lambda_{i,c}) \tag{6.18}$$

where $\rho_{i,c}$ is the outlier probability, $\Sigma_{i,c}$ is the covariance matrix of the Gaussian component, and $\mathcal{T}_\nu(\mathbf{0}, \Lambda_{i,c})$ denotes a multivariate Student-t distribution with degrees of freedom ν and scale matrix $\Lambda_{i,c}$ [3]. This structure models ordinary localization variability together with heavier-tailed outliers under degraded sensing conditions.

6.5.3 Residual-Based Parameter Estimation

Observation-channel parameters are estimated from residuals obtained by comparing the observed UWB trajectory with the smoothed reference trajectory. The residual process is

$$\mathbf{e}_t = \tilde{\mathbf{x}}_t - \mathbf{x}_t \tag{6.19}$$

where $\mathbf{x}_t$ is the reference latent proxy introduced earlier and $\tilde{\mathbf{x}}_t$ is the measured UWB position [3]. These residuals provide empirical information about sensing distortion under each behavioral state and sensing regime.

By estimating the bias, covariance, and outlier behavior from these residuals, the framework preserves the measurement characteristics of the real UWB system rather than imposing an abstract or idealized noise model.

6.5.4 Sensor-Observed Trajectory Generation

During simulation, the latent trajectory is generated from the behavioral state model and state-conditioned kinematic priors. At each time step, an observation distortion sample is drawn from the corresponding state- and regime-dependent distribution and added to the latent coordinate according to (6.16). In this way, the synthetic trajectory preserves the ethological structure of latent motion and the sensing realism of the observed domain [3].

This observation-layer formulation explains why the generated data are useful for robustness studies. It enables synthetic trajectories to reflect the sensing imperfections that affect real experimental measurements, thereby supporting virtual cohort generation, domain-shift analysis, and sensor-aware evaluation of downstream behavioral models.

6.6 Evaluation of Synthetic Trajectory Realism

Synthetic trajectories are compared against real UWB sessions using complementary realism metrics, including behavioral occupancy, local kinematic structure, residual-domain distortion, and multi-scale temporal dynamics [3].

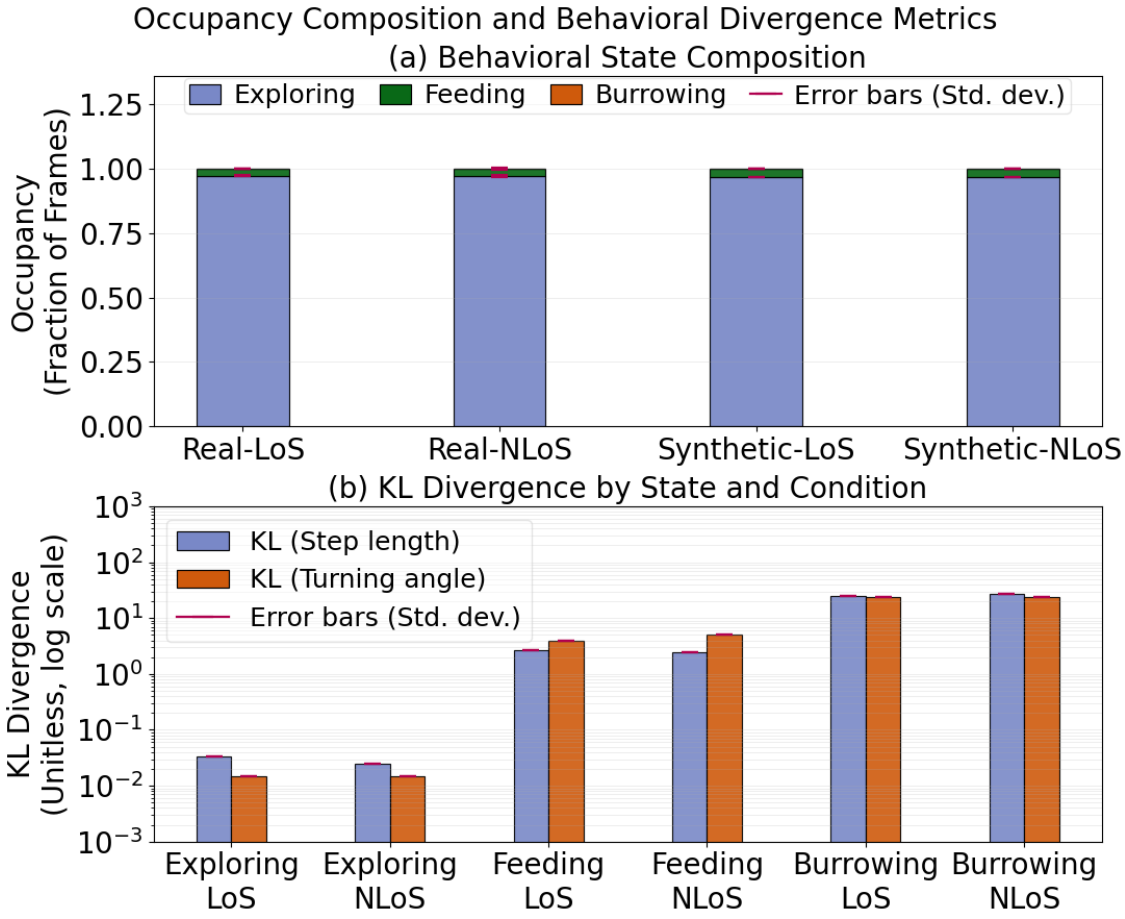


Figure 6.2: Per-state statistical realism metrics for real and synthetic trajectories under LoS and NLoS conditions. Adapted from [3].

6.6.1 Per-State Statistical Realism

A first level of evaluation examines per-state agreement between the real and synthetic trajectories. Two quantities are especially important: the frame-level occupancy of each behavioral state and the state-conditioned divergence of local motion statistics. Occupancy agreement tests whether the generator preserves the overall fraction of time spent in each

state, while state-conditioned divergence evaluates whether the local movement patterns within each state remain consistent with those observed in the real data [3].

Figure 6.2 presents the per-state statistical realism metrics under LoS and NLoS conditions. These results evaluate whether the framework preserves both the prevalence of each behavioral state and the state-dependent distributions of step length and turning angle. Close agreement indicates that the generator reproduces the main local motion characteristics of the observed data rather than merely generating visually plausible trajectories [3].

6.6.2 Occupancy-Weighted Realism and Residual Agreement

Because some states occur more frequently than others, per-state divergence may overemphasize discrepancies in rare states. To address this, the framework reports occupancy-weighted realism metrics, which aggregate divergence values by empirical state prevalence.
For a feature f, the weighted Kullback–Leibler divergence is defined using the standard information-theoretic divergence between probability distributions [48, 3].

$$D^{f}_{\mathrm{WKL}} = \sum_{i \in \mathcal{Z}} \pi_i^{\mathrm{real}} \, D_{\mathrm{KL}}\Big(p_i^{f} \,\|\, q_i^{f}\Big) \tag{6.20}$$

where π_i^{real} is the real-state occupancy and $D_{\mathrm{KL}}(\cdot\|\cdot)$ denotes the state-conditioned divergence between real and synthetic feature distributions [3]. This summary better reflects overall realism because it weights each behavioral state by its actual contribution to the trajectory. In addition to kinematic divergence, the framework evaluates residual-domain agreement by comparing the distribution of sensing distortion magnitudes between real and synthetic data. Figure 6.3 summarizes these occupancy-weighted and residual-domain realism metrics.

These metrics are important because the goal of the framework is not only to reproduce latent behavior, but also to preserve the sensing characteristics of the observed UWB domain. Agreement in residual magnitude therefore indicates that the synthetic trajectories remain

realistic from both a behavioral and measurement perspective [3].

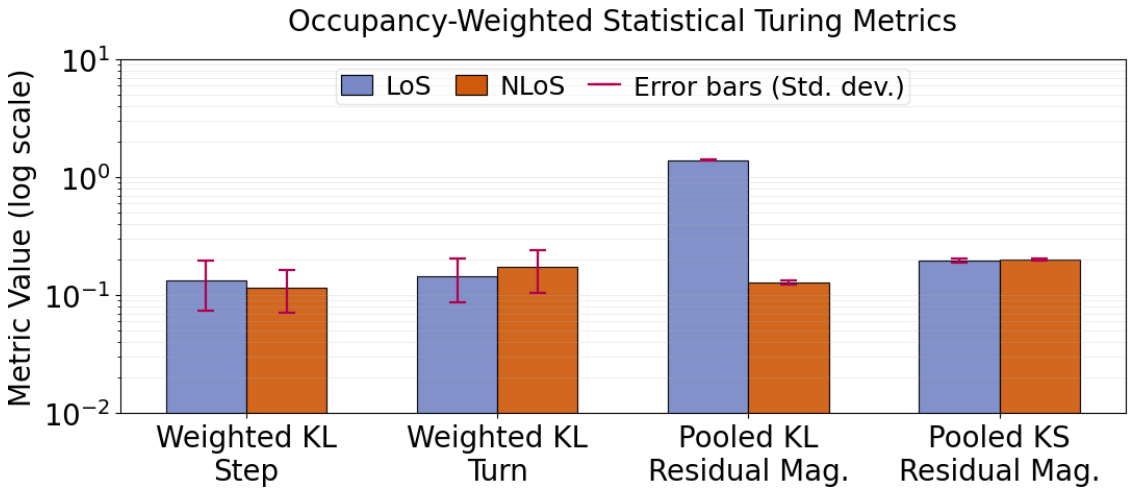


Figure 6.3: Occupancy-weighted realism metrics and residual-domain agreement under LoS and NLoS conditions. Adapted from [3].

6.6.3 Temporal Realism via Mean-Squared Displacement

Matching marginal distributions does not guarantee that a synthetic trajectory reproduces the temporal organization of real movement. To evaluate temporal realism, the framework computes the mean-squared displacement (MSD) across multiple time lags. For a discrete latent trajectory $\{\mathbf{x}_t\}_{t=1}^{T}$, the MSD at lag k is defined as

$$\mathrm{MSD}(k) = \frac{1}{T-k} \sum_{t=1}^{T-k} \|\mathbf{x}_{t+k} - \mathbf{x}_t\|_2^2 \tag{6.21}$$

where the result is expressed in square meters [3]. To summarize agreement between real and synthetic MSD curves, the root-mean-square error is computed as

$$\mathrm{RMSE}_{\mathrm{MSD}} = \sqrt{\frac{1}{K} \sum_{k=1}^{K} \left(\mathrm{MSD}_{\mathrm{syn}}(k) - \mathrm{MSD}_{\mathrm{real}}(k)\right)^2} \tag{6.22}$$

where smaller values indicate closer agreement in multi-scale displacement dynamics [3].

Figure 6.4 compares the mean MSD curves of real and synthetic trajectories under both

sensing regimes.

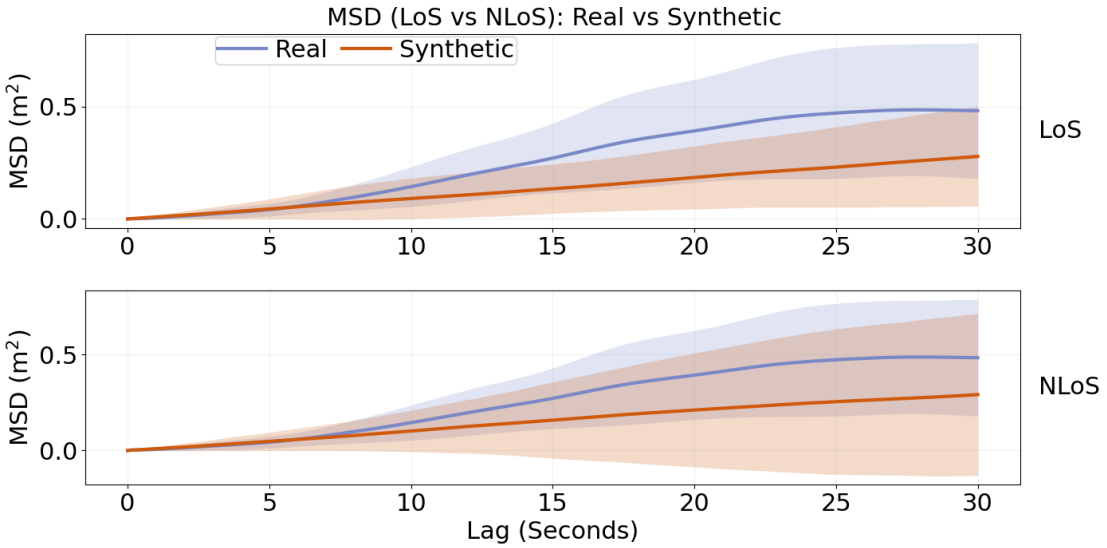


Figure 6.4: Mean-squared displacement of real and synthetic trajectories under LoS and NLoS conditions. Adapted from [3].

The close alignment of these curves indicates that the framework captures not only local motion statistics but also the broader temporal organization of the behavior. The root-mean-square error between real and synthetic MSD curves was 5.52×10^{-2} Square Meters under LoS and 6.06×10^{-2} Square Meters under NLoS, indicating useful agreement in multi-scale displacement dynamics. This is a particularly important result because it shows that the generator preserves realism across multiple time scales rather than matching only instantaneous kinematic features [3].

6.7 Robustness Under LoS/NLoS Domain Shift

Beyond trajectory realism, the framework is evaluated for robustness under LoS/NLoS sensing-domain shift. This analysis tests whether modeling the sensing regime improves inference when training and evaluation occur under different observation conditions [3].

To evaluate this effect, the framework considers a classification setting in which synthetic

trajectories are used to distinguish unperturbed sessions from mildly perturbed sessions, following the general machine-learning principle that held-out classification performance can be used to evaluate whether extracted trajectory features preserve discriminative structure [49]. The main comparison is between a condition-agnostic baseline, which ignores sensing regime, and a sensor-aware model, which incorporates the sensing condition as an explicit covariate. This design makes it possible to assess whether explicit awareness of line-of-sight (LoS) versus non-line-of-sight (NLoS) distortion improves cross-domain robustness [3].

When sensing condition is ignored, classifier performance degrades under cross-domain transfer. In other words, a model trained on trajectories generated under one sensing regime performs less reliably when evaluated on trajectories generated under the other regime. This behavior confirms that sensing distortion is not merely random measurement noise; rather, it induces a structured shift in the observed trajectory domain [3].

By contrast, when sensing regime is incorporated into the inference process, performance remains more stable across mixed-domain evaluation. This result indicates that sensor-aware conditioning mitigates degradation caused by LoS/NLoS mismatch and supports reliable downstream classification in the presence of observation-domain shift [3].

Table 6.1: Classifier performance on the synthetic unperturbed-versus-perturbed benchmark under LoS/NLoS sensing-domain shift. The sensor-aware classifier includes sensing regime as an explicit covariate, whereas the condition-agnostic baselines omit this information and are evaluated under cross-domain transfer. AUC is computed from decision scores, while accuracy and balanced accuracy are computed using a fixed decision threshold. Adapted from [3].

Evaluation setting	AUC	Accuracy	Balanced Accuracy
Sensor-aware: mixed-domain test (LoS+NLoS)	0.995	0.973	0.973
Sensor-aware: test on LoS only	0.998	0.973	0.973
Sensor-aware: test on NLoS only	0.993	0.973	0.973
Condition-agnostic: train LoS, test NLoS	0.974	0.594	0.594
Condition-agnostic: train NLoS, test LoS	0.901	0.500	0.500

Table 6.1 summarizes classifier performance under the different evaluation settings. These results indicate that sensing regime is a structured part of the data-generating process. The sensor-aware model achieved mixed-domain performance with an AUC of 0.995, whereas condition-agnostic baselines degraded under cross-regime transfer [3].

Thus, the framework provides a controlled basis for studying sensing-domain mismatch and evaluating mitigation strategies through explicit sensor-aware conditioning.

6.8 Qualitative Realism of Representative Trajectories

In addition to quantitative realism metrics, it is also important to evaluate whether the generated trajectories are visually consistent with the spatial structure of the real behavioral recordings. Qualitative realism is useful because it reveals whether the synthetic trajectories preserve interpretable arena coverage, state-dependent motion structure, and temporally coherent behavioral organization rather than merely matching summary statistics [3].

Figure 6.5 presents representative real trajectories under LoS and NLoS sensing conditions.

Under LoS conditions, the trajectory shows smoother spatial continuity and clearer state transitions. Under NLoS conditions, the trajectory exhibits greater irregularity and occasional abrupt deviations that are consistent with sensing-induced distortion, while still preserving the overall behavioral structure of the session [3].

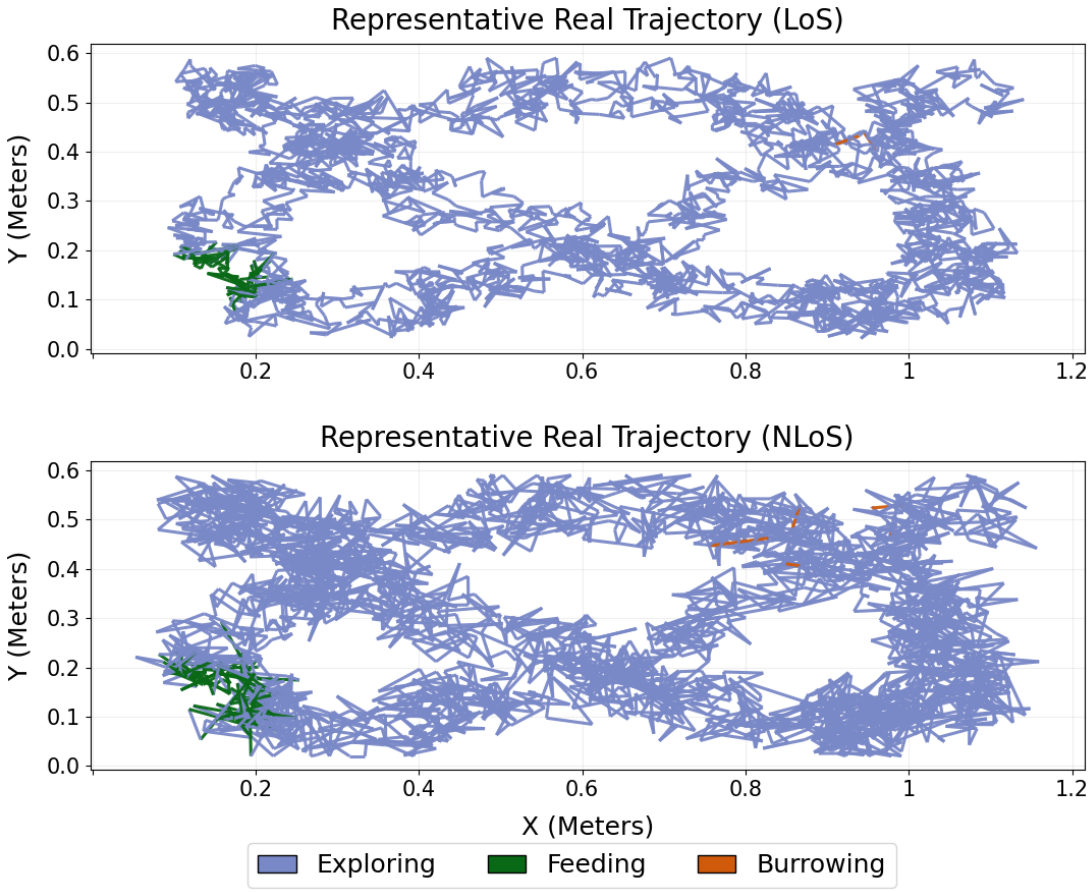


Figure 6.5: Representative real trajectories under (a) LoS and (b) NLoS sensing conditions, colored by behavioral state. Adapted from [3].

Figure 6.6 shows a representative synthetic trajectory generated by the sensor-aware framework. The trajectory exhibits realistic spatial coverage and state-dependent motion patterns, including exploratory motion with localized feeding and burrowing behavior. This visual consistency indicates that the generative model captures meaningful behavioral organization rather than arbitrary motion paths.

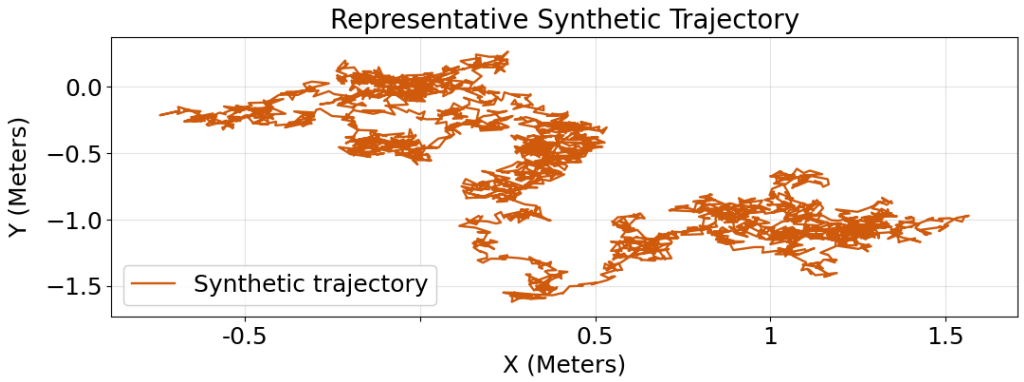


Figure 6.6: Representative synthetic trajectory generated by the sensor-aware framework. Adapted from [3].

Taken together, these visual comparisons complement the quantitative realism metrics presented earlier. They show that the framework reproduces not only distributional agreement, but also the broader spatial and temporal organization that makes the trajectories behaviorally interpretable. This strengthens the case for using the generated data in virtual cohort analysis, robustness testing, and sensor-aware behavioral simulation.

6.9 Chapter Summary

This chapter presented the sensor-aware generative framework and trajectory-level digital twin developed for in silico ethology. Within the probability-of-motion framework, the model treats animal movement as a latent probabilistic process and the measured trajectory as a sensor-conditioned observation shaped by LoS/NLoS-dependent distortion. The framework combines a latent state model, state-conditioned kinematic generation, and a LoS/NLoS-dependent observation channel to generate trajectories that are both behaviorally plausible and measurement-realistic [3].

The chapter also evaluated the realism and utility of the generated trajectories using statistical, temporal, qualitative, and downstream robustness analyses. The results showed that the framework preserves key properties of the real data, including behavioral occupancy, local

kinematic structure, residual-domain distortion, and multi-scale displacement dynamics. In downstream classification, explicit sensor-aware conditioning achieved stable mixed-domain performance with an AUC of 0.995, whereas condition-agnostic baselines degraded under LoS/NLoS domain shift [3].

The main achievement of this chapter is the development of a bounded and defensible trajectory-level digital twin for UWB-based rodent behavioral inference. The model does not claim to reproduce the complete biological animal. Instead, it represents the probabilistic relationship between latent behavioral trajectory generation and sensing-regime-dependent observation distortion. This distinction is essential because it makes the digital twin claim precise, experimentally grounded, and directly connected to the dissertation's probability-of-motion framework, where animal movement is modeled through both behavior and measurement uncertainty.

Taken together, these findings show that realistic behavioral simulation requires modeling not only how animals move, but also how sensing systems transform that motion into the observed data domain. Within the probability-of-motion framework, this chapter extends the dissertation from physical tracking and probabilistic interpretation toward trajectory-level digital twins, virtual cohorts, robustness testing, and sensor-aware computational ethology under measurement uncertainty.

Chapter 7

Conclusions, Contributions, and Scientific Significance

7.1 Introduction

This dissertation addressed a central challenge in quantitative animal behavior research: how to model animal motion when the measured trajectory is shaped by both behavior and the sensing system used to observe it. It developed a unified sensor-aware framework that connects Ultra-Wideband (UWB) localization, signal integrity, probabilistic behavioral modeling, and trajectory-level digital twin simulation for rodent studies [1, 2, 3].

The central argument of this dissertation is that rodent behavioral inference cannot be separated from the sensing conditions through which behavior is measured. LoS/NLoS propagation, multipath distortion, attenuation, and measurement uncertainty directly influence the reliability and interpretability of trajectory-derived behavioral models. In this sense, the probability of motion reflects both the stochastic organization of animal behavior and the uncertainty introduced by the observation process.

The dissertation follows a complete measurement-to-inference progression. First, it establishes a high-precision UWB real-time locating system for rodent behavioral tracking. Second, it shows that signal integrity propagates into probabilistic behavioral markers, including entropy, threshold reliability, and Markov state persistence. Third, it extends this framework into a trajectory-level digital twin that generates virtual rodent trajectories while preserving both ethological structure and sensing-regime-dependent measurement distortion.

7.2 Central Thesis of the Dissertation

The central thesis of this dissertation is that animal motion becomes scientifically meaningful when its probability structure is modeled together with the sensing uncertainty through which it is observed. Rather than treating a recorded trajectory as a fixed representation of behavior, this work treats the observed path as a sensor-conditioned outcome shaped by latent behavioral dynamics, localization performance, and environmental distortion.

This thesis is reflected in the dissertation's progression from UWB system validation, to signal-aware probabilistic behavioral modeling, to sensor-aware virtual trajectory generation. Across this progression, the dissertation establishes that the probability of motion is not only a property of animal behavior, but also a consequence of the measurement process that transforms movement into data.

7.3 Principal Contributions and Novelty

This dissertation makes four principal contributions.

First, it develops and validates a UWB-based real-time locating system for rodent behavioral studies. The system integrates compact wearable tagging, perimeter anchor deployment, time-based ranging, phase-based localization, adaptive filtering, and Kalman-based correction to support centimeter-level trajectory reconstruction in a controlled laboratory

environment. Experimental validation demonstrated a mean localization error of 2.28 centimeters under LoS conditions and 4.40 centimeters under NLoS conditions, confirming that UWB tracking can provide a practical sensing foundation for behavioral studies while also revealing structured regime-dependent error [1].

Second, the dissertation characterizes LoS and NLoS sensing regimes as structured measurement conditions that alter localization reliability and trajectory quality through obstruction, multipath propagation, attenuation, and reduced signal-to-noise ratio [1, 2].

Third, the dissertation develops a probabilistic behavioral modeling framework derived directly from reconstructed UWB trajectories. Through behavioral state segmentation, Shannon entropy, Bernoulli threshold modeling, and Markov chain analysis, the work demonstrates that sensing degradation propagates into behavioral uncertainty, reliability, and sequential organization. Under NLoS conditions, localization-error entropy increased from 1.15 bits to 1.78 bits, the probability of achieving localization accuracy within 20 millimeters decreased from 63.2 percent to 27.7 percent, and average behavioral state persistence decreased from 68 percent to 52 percent. These results establish a physics-to-behavior framework in which signal integrity is quantitatively linked to behavioral interpretation [2].

Fourth, the dissertation introduces a sensor-aware generative framework for in silico ethology and trajectory-level digital twin modeling. By separating latent behavioral dynamics from observation-domain distortion, the framework generates measurement-realistic virtual trajectories and supports downstream robustness testing under LoS/NLoS domain shift. Explicit sensor-aware conditioning achieved stable mixed-domain performance with an AUC of 0.995, whereas condition-agnostic baselines degraded under cross-regime transfer [3].

Taken together, these contributions define the core novelty of the dissertation: the establishment of a sensor-aware probability-of-motion framework for rodent behavioral inference. In this framework, sensing hardware, localization accuracy, signal integrity, probabilistic inter-

pretation, and trajectory-level digital twin simulation are treated as coupled components of a unified scientific architecture for modeling animal motion under uncertainty.

7.4 Scientific and Methodological Significance

The scientific significance of this dissertation lies in reframing rodent tracking as a probability-of-motion problem, where behavioral interpretation depends on both the latent structure of animal movement and the sensing process that preserves, distorts, or transforms reconstructed trajectories.

Methodologically, the dissertation contributes a multi-layer measurement-to-inference framework that connects validated UWB localization, LoS/NLoS signal-integrity analysis, probabilistic behavioral modeling, and sensor-aware trajectory-level digital twin generation.

A further significance of the work is that it bridges experimental rodent studies and computational modeling by grounding virtual trajectory generation in experimentally observed sensing behavior rather than idealized motion alone.

This dissertation also supports reproducibility and scalability by treating measured trajectories as products of both animal movement and sensing conditions, enabling more transparent comparison across sessions, sensing regimes, virtual cohorts, and future experimental configurations.

7.5 Limitations and Scope

Several limitations define the scope of the present work. First, the framework was developed and validated using a limited number of experimental sessions, which restricts the diversity of behavioral patterns, sensing conditions, and environmental configurations represented in the available dataset. Larger multi-animal datasets are needed to evaluate generalization

across animals, tasks, cage geometries, and behavioral contexts.

Second, the probabilistic and generative models were intentionally designed to remain interpretable and stable under limited-data conditions. This design choice supported transparent analysis of the relationship between signal integrity and behavioral inference, but it also limited the complexity of the modeled temporal structure relative to higher-capacity sequence-learning approaches [2, 3].

Third, the dissertation focused primarily on planar rodent localization and trajectory-derived behavioral inference in controlled environments. This scope is appropriate for the targeted experimental setting, but it does not exhaust the broader possibilities of multimodal, multi-animal, or three-dimensional behavioral analysis. Future extensions may incorporate additional sensing modalities, such as neural, physiological, or inertial measurements, to support richer brain–behavior and body–environment inference.

Finally, the generative framework should be interpreted as a trajectory-level digital twin focused on locomotor behavior and sensing-regime-dependent observation distortion, not as a complete biological digital twin of the animal [3].

These limitations define the scope of the dissertation: a methodological foundation for a sensor-aware probability-of-motion framework, rather than a complete solution to computational ethology or a complete biological digital twin of the animal.

7.6 Future Directions

Several future directions naturally follow from this dissertation. First, the UWB tracking system can be extended to larger and more complex environments, including enriched arenas, underground or partially occluded habitats, and three-dimensional tracking configurations. Vertical anchor arrangements, improved antenna placement, and sensor fusion with inertial

measurements may further improve localization fidelity in complex behavioral settings.

Second, the probabilistic behavioral framework can be expanded beyond first-order Markov models to include higher-order, semi-Markov, hidden Markov, or Bayesian hierarchical models. Such extensions would allow future versions of the framework to capture latent behavioral states, longer-range dependencies, animal-specific behavioral signatures, and uncertainty in rare transitions while preserving interpretability [50].

Third, future work can integrate UWB-based behavioral trajectories with neural and physiological measurements, including electroencephalography, electrophysiology, cardiovascular signals, or other wearable biosensors. This direction would allow the measurement-aware behavioral framework developed here to support joint brain–behavior modeling in freely moving animals.

Fourth, the trajectory-level digital twin framework can be expanded using larger multi-animal datasets and learned conditional sequence models while preserving the explicit sensor-aware observation layer. In future closed-loop or intervention-oriented extensions, such learned components may also benefit from formal verification and safety-certification tools for neural dynamical systems, including approaches based on neural control barrier functions and forward-invariant set extraction [51].

Together, these directions would extend the probability-of-motion framework from validated sensor-aware trajectory modeling toward a broader multimodal platform for quantitative, reproducible, and measurement-aware preclinical behavioral research across sensing modalities, animal cohorts, and experimental environments.

7.7 Final Synthesis

This dissertation has shown that animal motion should not be interpreted only as a sequence of measured coordinates, but as a probabilistic sensing-and-inference problem in which measurement conditions shape the trajectories from which behavior is understood. By connecting UWB localization, signal integrity analysis, probabilistic behavioral modeling, and sensor-aware trajectory-level digital twin simulation, this work provides a methodological foundation for quantitative rodent studies under realistic sensing conditions.

The lasting contribution of this dissertation is the establishment of a probability-of-motion framework in which sensing hardware, localization accuracy, behavioral uncertainty, and virtual trajectory simulation form a single sensor-aware scientific architecture. In this framework, sensing quality is not external to behavioral inference; it is one of the mechanisms through which movement becomes measurable, interpretable, and reproducible.

By making the measurement-to-inference relationship explicit, this dissertation contributes an interpretable and extensible foundation for future quantitative studies of animal behavior.

In summary, the dissertation establishes a sensor-aware pathway from physical localization to probabilistic behavioral interpretation and trajectory-level digital twin generation, providing a foundation for scalable, reproducible, and measurement-aware behavioral research in biomedical science.

Bibliography

[1] Reza Sayfoori, Mao-Hsiang Huang, Amir Naderi, Mehwish Bhatti, Ron D. Frostig, and Hung Cao. High-precision UWB sensor-based real-time locating system for rodent behavioral studies. *IEEE Sensors Journal*, 25(13):25551–25559, 2025.

[2] Reza Sayfoori and Hung Cao. Linking signal integrity to probabilistic models of behavioral dynamics. Under review at Behavior Research Methods, 2025.

[3] Reza Sayfoori, Goli Vaisi, and Hung Cao. Virtual mice, real errors: A sensor-aware generative framework for in silico ethology. *Sensors*, 26(10):2977, 2026.

[4] Colin Lever, Sarah Burton, and John O'Keefe. Rodent models of cognitive function and dysfunction. *Current Opinion in Neurobiology*, 20(6):712–718, 2010.

[5] Shitij Kapur and David Mamo. Half a century of antipsychotics and still a central role for dopamine d2 receptors. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 27(7):1081–1090, 2003.

[6] Mark A. Geyer and Bart Ellenbroek. Animal behavior models of the mechanisms underlying antipsychotic atypicality. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 27(7):1071–1079, 2003.

[7] David J. Anderson and Pietro Perona. Toward a science of computational ethology. *Neuron*, 84(1):18–31, 2014.

[8] Kristin Branson et al. High-throughput ethomics in large groups of Drosophila. *Nature Methods*, 6(6):451–457, 2009.

[9] Ramsey Faragher and Robert Harle. Analysis of the accuracy of Bluetooth Low Energy for indoor positioning applications. In *Proceedings of the International Technical Meeting of The Institute of Navigation*, 2015.

[10] Babak Alavi et al. Ultra-wideband (UWB) sensor networks for indoor localization. *IEEE Communications Surveys & Tutorials*, 18(4):3242–3265, 2016.

[11] Sinan Gezici et al. Localization via ultra-wideband radios: a look at positioning aspects. *IEEE Signal Processing Magazine*, 22(4):70–84, 2005.

[12] John B. Hoy, Philip G. Koehler, and Richard S. Patterson. A microcomputer-based system for real-time analysis of animal movement. *Journal of Neuroscience Methods*, 64(2):157–161, 1996.

[13] Wen-Harn T. Pan, Chih-Ren Lee, and Li-Hsien Lim. A new video path analyzer to monitor travel distance, rearing, and stereotypic movement of rats. *Journal of Neuroscience Methods*, 70(1):39–43, 1996.

[14] Alexander Mathis, Pranav Mamidanna, Kevin M. Cury, Taiga Abe, Venkatesh N. Murthy, Mackenzie W. Mathis, and Matthias Bethge. Deeplabcut: markerless pose estimation of user-defined body parts with deep learning. *Nature Neuroscience*, 21:1281–1289, 2018.

[15] Oliver Sturman, Lukas von Ziegler, Christian Schläppi, Furkan Akyol, Mattia Privitera, Damian Slominski, Jan Grimm, Loic Thieren, Valerio Zerbi, Benjamin F. Grewe, and Johannes Bohacek. Deep learning-based behavioral analysis reaches human accuracy and is capable of outperforming commercial solutions. *Neuropsychopharmacology*, 45:1942–1952, 2020.

[16] Danielle D. Brown, Roland Kays, Martin Wikelski, Rory Wilson, and A. Peter Klimley. Observing the unwatchable through acceleration logging of animal behavior. *Animal Biotelemetry*, 1(1):20, 2013.

[17] Eduardo Pereira et al. Rfid technology for animal tracking: A survey. *IEEE Journal of Radio Frequency Identification*, 7:609–620, 2023.

[18] Faheem Zafari, Athanasios Gkelias, and Kin K. Leung. A survey of indoor localization systems and technologies. *IEEE Communications Surveys & Tutorials*, 21(3):2568–2599, 2019.

[19] Abdulrahman Alarifi, AbdulMalik Al-Salman, Mansour Alsaleh, Ali Alnafessah, Suheer Al-Hadhrami, Mohammad A. Al-Ammar, and Hend S. Al-Khalifa. Ultra wideband indoor positioning technologies: Analysis and recent advances. *Sensors*, 16(5):707, 2016.

[20] Chiara Falsi, Davide Dardari, Lorenzo Mucchi, and Moe Z. Win. Time of arrival estimation for uwb localizers in realistic environments. *IEEE Transactions on Wireless Communications*, 5(3):520–529, 2006.

[21] Moe Z. Win, Andrea Conti, Santiago Mazuelas, Yuan Shen, Wesley M. Gifford, Davide Dardari, and Marco Chiani. Network localization and navigation via ultra-wideband signals: theory, algorithms, and experimentations. *IEEE Communications Magazine*, 48(4):56–64, 2010.

[22] Ismail Güvenç, Chia-Chin Chong, and Fumiyuki Watanabe. Nlos identification and mitigation for uwb localization systems. In *2007 IEEE Wireless Communications and Networking Conference*, pages 1571–1576. IEEE, 2007.

[23] Stefano Marano, Wesley M. Gifford, Henk Wymeersch, and Moe Z. Win. Nlos identification and mitigation for localization based on uwb experimental data. *IEEE Journal on Selected Areas in Communications*, 28(7):1026–1035, 2010.

[24] Homayoun Hashemi. The indoor radio propagation channel. *Proceedings of the IEEE*, 81(7):943–968, 1993.

[25] Andreas F. Molisch. *UWB Propagation Channels: Theory, Measurement, and Modeling*. Wiley-IEEE Press, Hoboken, NJ, USA, 2 edition, 2011.

[26] Toby A. Patterson, Len Thomas, Chris Wilcox, Otso Ovaskainen, and Jason Matthiopoulos. State-space models of animal movement. *Ecological Modelling*, 219(2):161–174, 2008.

[27] Rudolf E. Kalman. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1):35–45, 1960.

[28] Gordon J. Berman. Measuring behavior across scales. *BMC Biology*, 16:23, 2018.

[29] Greg J. Stephens, Leslie C. Osborne, and William Bialek. Searching for simplicity in the analysis of neurons and behavior. *Proceedings of the National Academy of Sciences of the United States of America*, 108(Supplement 3):15565–15571, 2011.

[30] Talmo D. Pereira, Joshua W. Shaevitz, and Mala Murthy. Quantifying behavior to understand the brain. *Nature Neuroscience*, 23(12):1537–1549, 2020.

[31] Alexander B. Wiltschko, Matthew J. Johnson, Giuliano Iurilli, Ralph E. Peterson, Jesse M. Katon, Stan L. Pashkovski, Victoria E. Abraira, Ryan P. Adams, and Sandeep Robert Datta. Mapping sub-second structure in mouse behavior. *Neuron*, 88(6):1121–1135, 2015.

[32] Austin I. Hsu and Eric A. Yttri. B-soid: An open-source unsupervised algorithm for discovery of spontaneous behaviors. *eLife*, 10:e63720, 2021.

[33] Kevin Luxem, Felix Fuhrmann, Tim Kürschner, and Stefan Remy. Identifying behavioral structure from deep variational embeddings of animal motion. *Nature Methods*, 19:1482–1492, 2022.

[34] Josh Tobin, Rachel Fong, Alex Ray, Jonas Schneider, Wojciech Zaremba, and Pieter Abbeel. Domain randomization for transferring deep neural networks from simulation to the real world. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 23–30. IEEE, 2017.

[35] Aidan Fuller, Zhong Fan, Charles Day, and Chris Barlow. Digital twin: Enabling technologies, challenges and open research. *IEEE Access*, 8:108952–108971, 2020.

[36] Daniel B. Polley, Elizabeth Kvasnak, and Ron D. Frostig. Naturalistic experience transforms sensory maps in the adult cortex of caged animals. *Nature*, 429:67–71, 2004.

[37] Ron D. Frostig. Functional organization and plasticity in the adult rat barrel cortex: moving out-of-the-box. *Current Opinion in Neurobiology*, 16(4):445–450, 2006.

[38] Ann Kennedy. The what, how, and why of naturalistic behavior. *Current Opinion in Neurobiology*, 74:102549, 2022.

[39] Qorvo. DWM1001C Ultra-Wideband Module. `https://www.qorvo.com/products/p/DWM1001C`, 2024. Accessed: May 17, 2026.

[40] Decawave Ltd. *DWM1001 Datasheet*. Decawave Ltd., 2020. DWM1001/DWM1001C Ultra-Wideband Module documentation.

[41] Claude E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27(3):379–423, 1948.

[42] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley-Interscience, Hoboken, NJ, USA, 2 edition, 2006.

[43] Sheldon Ross. *A First Course in Probability*. Pearson, 9 edition, 2014.

[44] J. R. Norris. *Markov Chains*. Cambridge University Press, 1997.

[45] William M. S. Russell and Rex L. Burch. *The Principles of Humane Experimental Technique*. Methuen, London, UK, 1959.

[46] Ronald A. Howard. *Dynamic Probabilistic Systems, Volume I: Markov Models*. Wiley, New York, NY, USA, 1971.

[47] Peter J. Huber and Elvezio M. Ronchetti. *Robust Statistics*. Wiley, Hoboken, NJ, USA, 2 edition, 2009.

[48] Solomon Kullback. *Information Theory and Statistics*. Wiley, New York, NY, USA, 1959.

[49] Christopher M. Bishop. *Pattern Recognition and Machine Learning*. Springer, New York, NY, USA, 2006.

[50] Lawrence R. Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257–286, 1989.

[51] Goli Vaisi, James Ferlez, and Yasser Shoukry. Extracting forward invariant sets from neural network-based control barrier functions. In *Proceedings of the 28th ACM International Conference on Hybrid Systems: Computation and Control*, HSCC '25, pages 20:1–20:11, New York, NY, USA, 2025. Association for Computing Machinery.