

Observability Thresholds for Damping and Stiffness Estimation in Stochastic Underdamped Oscillators

Denario

Anthropic, Gemini & OpenAI servers. Planet Earth.

Abstract

Accurately identifying physical parameters in underdamped systems from noisy position and velocity data, without direct acceleration measurements, poses a significant challenge. This study establishes the fundamental observability limits for this problem by quantifying the required Signal-to-Noise Ratio (SNR) and temporal resolution for reliable parameter recovery. Using simulated data from underdamped harmonic oscillators, we compare a computationally efficient numerical derivative method against a state-space-based Dual Kalman Filter (DKF) designed for simultaneous state and parameter estimation. Our findings demonstrate that the DKF is substantially more robust to noise, successfully estimating the spring constant (k) and damping coefficient (b) below a 5% error threshold at SNRs where the numerical derivative approach fails. Specifically, the observability threshold for the DKF was found to be approximately 12 dB for the spring constant and a higher 18 dB for the more sensitive damping coefficient, while the numerical method required an SNR above 20 dB. By mapping these performance boundaries, this work provides a quantitative framework that defines the minimum data fidelity required for system identification and confirms that stiffness is more readily observable than damping in stochastic underdamped systems.

1 Introduction

The accurate identification of physical parameters from observational data is a cornerstone of quantitative analysis across science and engineering. From characterizing the structural health of civil infrastructure to modeling intracellular transport, the ability to infer the intrinsic properties of a system is essential for prediction, control, and fundamental understanding. A canonical model for a wide range of oscillatory phenomena is the underdamped harmonic oscillator, whose dynamics are governed by its stiffness and damping coefficients. These parameters dictate the system's natural frequency and the rate at which energy dissipates, making their accurate estimation a critical task.

A significant practical challenge arises when these parameters must be inferred from time-series measurements of position and velocity that are inevitably corrupted by noise, particularly when direct acceleration data is unavailable. The governing second-order differential equation of motion, $m\ddot{x}(t) + b\dot{x}(t) + kx(t) = F(t)$, cannot be directly solved for the parameters without knowledge of the acceleration term, $\ddot{x}(t)$. Estimating this term by numerically differentiating noisy velocity data is an inherently unstable process that amplifies high-frequency measurement error. This sensitivity poses a fundamental question: at what point does noise overwhelm the underlying physical signal, rendering the estimated parameters unreliable? Without a quantitative answer, researchers risk placing false confidence in models derived from data of insufficient quality.

This study addresses this gap by establishing the quantitative observability thresholds for parameter estimation in stochastic underdamped systems. We move beyond the qualitative understanding that noise degrades accuracy to define the minimum Signal-to-Noise Ratio (SNR) and temporal resolution required to estimate the spring constant (k) and damping coefficient (b) within a predefined error tolerance. By mapping these performance boundaries, we provide a rigorous framework for assessing whether a given dataset possesses the fidelity needed to resolve its underlying physical characteristics. This work aims to equip researchers with a practical diagnostic tool to determine the feasibility of system identification before committing to complex modeling efforts.

To define these limits, we systematically compare two distinct estimation methodologies using simulated data. The first is a computationally efficient approach based on numerical differentiation to approximate acceleration, followed by linear least-squares regression to solve for the system parameters. The second is a state-space method employing a Dual Kalman Filter, which is explicitly designed to handle stochasticity by simultaneously estimating the system's state and its parameters. By evaluating these methods across a broad spectrum of noise levels and sampling rates, we chart the performance landscape for each. This investigation provides a quantitative guide to the minimum data requirements for system identification and empirically demonstrates that stiffness is a more readily observable parameter than damping in noisy dynamical systems.

2 Methods

2.1 Simulated Dataset and Experimental Design

The analysis was performed on a synthetic dataset generated from the numerical integration of the equation of motion for an underdamped harmonic oscillator:

$$m\ddot{x}(t) + b\dot{x}(t) + kx(t) = 0 \tag{1}$$

where m is the mass (assumed to be unity), b is the damping coefficient, and k is the spring constant. The dataset comprised 20 distinct oscillators, each with unique ground-truth values for b and k . For each oscillator, time-series data for position $x(t)$ and velocity $\dot{x}(t)$ were generated over a 20-second interval.

To investigate the limits of parameter observability, the pristine data were systematically degraded. We created an experimental grid by varying two key factors: data fidelity and temporal resolution. Zero-mean Gaussian white noise was injected into both the position and velocity time series to achieve six distinct Signal-to-Noise Ratio (SNR) levels, ranging from 5 dB to 40 dB. Concurrently, the temporal resolution was varied by downsampling the original trajectories to 500, 250, 100, 50, and 25 time steps, while preserving the total 20-second observation window. This process yielded a comprehensive set of test conditions for evaluating the robustness of the estimation algorithms.

2.2 Parameter Estimation Methodologies

We compared two distinct methodologies for estimating the parameters k and b from the noisy time-series data, without direct access to acceleration measurements.

2.2.1 Numerical Derivative with Least-Squares Regression

This approach approximates the acceleration term $\ddot{x}(t)$ to enable a direct algebraic solution for the parameters. The procedure involved three steps. First, to mitigate the noise amplification inherent in numerical differentiation, the noisy position $x(t)$ and velocity $\dot{x}(t)$ signals were smoothed using a Savitzky-Golay filter. The window size of the filter was dynamically adjusted to be proportional to the temporal resolution of the data. Second, the acceleration $\ddot{x}(t)$ was approximated by applying a central difference formula to the smoothed velocity data. Finally, with the state variables (x, \dot{x}, \ddot{x}) estimated, the parameters b and k were determined by solving the rearranged equation of motion using linear least-squares regression:

$$\ddot{x}(t) = - \left(\frac{b}{m} \right) \dot{x}(t) - \left(\frac{k}{m} \right) x(t) \quad (2)$$

2.2.2 Dual Kalman Filter

As a state-space alternative, we implemented a Dual Kalman Filter (DKF) for simultaneous state and parameter estimation. This method treats the unknown parameters as part of an augmented state vector. The system was modeled with the state vector $\mathbf{z} = [x, \dot{x}, k, b]^T$. The DKF architecture employs two interacting Kalman filters: a state filter estimates the kinematic state (x, \dot{x}) based on the current parameter estimates, while a parameter filter estimates the physical parameters (k, b) based on the output of the state filter. This iterative, online process allows the model to refine its estimates of both the system's state and its intrinsic parameters at each time step, making it inherently more robust to measurement noise than direct numerical differentiation. The filter was initialized with reasonable prior estimates for the parameters, rather than their ground-truth values, to simulate a realistic application scenario.

2.3 Performance Evaluation

To ensure that any observed estimation errors were attributable to noise and temporal sparsity rather than algorithmic flaws, a baseline validation was first conducted. Both methods were applied to the noise-free, high-resolution data, and both successfully recovered the ground-truth parameters with negligible error.

The primary metric for quantifying the accuracy of the parameter estimates was the Mean Absolute Percentage Error (MAPE), calculated for both k and b across all 20 oscillators for each experimental condition. The MAPE is defined as:

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N \left| \frac{\theta_i - \hat{\theta}_i}{\theta_i} \right| \times 100\% \quad (3)$$

where θ_i is the ground-truth value of a parameter for the i -th oscillator, $\hat{\theta}_i$ is its corresponding estimate, and N is the total number of oscillators.

Using this metric, we defined an “observability threshold” as the SNR level below which the MAPE for a given parameter exceeded 5%. This threshold serves as a quantitative boundary, delineating the minimum data fidelity required for reliable parameter recovery with each method.

3 Results

4 Conclusions