# Inference and Sampling of Point Processes from Diffusion Excursions

**Ali Hasan**[1,2] **Yu Chen**[2] **Yuting Ng**[1] **Mohamed Abdelghani**[3] **Anderson Schneider**[2] **Vahid Tarokh**[1]

[1]Department of Electrical and Computer Engineering, Duke University, Durham, North Carolina, USA
[2]Machine Learning Research, Morgan Stanley
[3]Department of Mathematics, University of Alberta, Edmonton, Alberta, Canada

## Abstract

Point processes often have a natural interpretation with respect to a continuous process. We propose a point process construction that describes arrival time observations in terms of the state of a latent diffusion process. In this framework, we relate the return times of a diffusion in a continuous path space to new arrivals of the point process. This leads to a continuous sample path that is used to describe the underlying mechanism generating the arrival distribution. These models arise in many disciplines, such as financial settings where actions in a market are determined by a hidden continuous price or in neuroscience where a latent stimulus generates spike trains. Based on the developments in Itô's excursion theory, we propose methods for inferring and sampling from the point process derived from the latent diffusion process. We illustrate the approach with numerical examples using both simulated and real data. The proposed methods and framework provide a basis for interpreting point processes through the lens of diffusions.

## 1 INTRODUCTION

Point processes are a powerful modeling tool for describing patterns of arrivals, with applications ranging from environmental and biological sciences to financial markets and social behavior [Björk et al., 1997, Rizoiu et al., 2017, Subramanian et al., 2020, Stoyan and Penttinen, 2000]. Often, a point process is represented through an *intensity function*, which is a function that describes the expected number of arrivals. This function is the primary mechanism for interpreting the properties of the process, with standard models such as the Poisson process and the Hawkes process as primary examples. However, considering only the intensity function may not provide a complete understanding of the

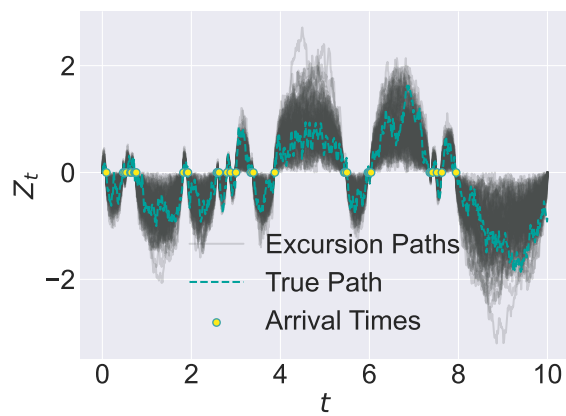



Figure 1: Decomposition of a sample path into distributions of Brownian excursions. Dashed line represents the true signal and solid lines represent possible excursions between observations of arrival times indicated by circle markers.

underlying cause of arrivals of points

In many cases, a point process may be related to a *continuous process*. The choice of the continuous process is also motivated by applications. To list a few examples, neuron spike trains may be related to the first passage time of an underlying chemical concentration surpassing a threshold [Sacerdote and Zucca, 2003]. Similarly, intracellular events are considered to be a function of a protein concentration exceeding a threshold and bursty transcription relates to the continuous movement of underlying molecules [Ghusinga et al., 2017, Lammers et al., 2020]. In an economic setting, one can think of information flow in a market, and the ensuing point process generated by orders of agents on an exchange as a function of the information flow [Babus and Kondor, 2018]. All of these models consider a multi-scale approach such that the point process is generated as a function of the unobservable continuous process. Developing inference methods for recovering a possible continuous process could provide additional insights into the point process

*Accepted for the 39th Conference on Uncertainty in Artificial Intelligence* (UAI 2023).

being studied. This leads us to the motivation of this work, where we focus on continuous stochastic processes defined by Itô diffusions and relate these paths to the arrival times of the point process. Specifically, we consider a decomposition into paths known as *excursions* – paths that begin and end at a reference state and are constrained to stay above or below the reference state for their entirety. The length of an excursion correspond to an interarrival time of the point process. This decomposition is illustrated in Figure 1 where the sample path is decomposed into excursions from the reference state of 0, marked by arrivals of the point process. The original idea for constructing such a process was introduced by Itô [1972], where the decomposition of sample paths into excursions was used as an alternative tool to stochastic calculus for studying diffusions. Itô additionally raised the question of what continuous process could represent observations of arrival times, which is what we contribute towards in this work [Watanabe, 2010].

**Motivating Example** Consider events in a market where individuals are buying and selling a group of assets by placing bids and asks. We can model this as a marked point process where the mark is given by the type of action (bids or asks) and the price at which the asset is requested. Then, we can assume that the different arrivals are associated with an unobserved, fair price that governs the asset — bids are placed when the asset is below its fair price and asks are placed when the asset is above the fair price. These *excursions* above and below the fair price give rise to the point process structure, and that is what we aim to model through this framework. The multi-dimensional case then corresponds to the point process of multiple, correlated assets in a market. The parameters recovered then correspond to the diffusion that models the unobserved fair price.

### 1.1 RELATED WORK

A number of research directions are related to understanding point processes through continuous processes, and, in the particular case of Brownian motion, the most relevant is the study of first hitting times (FHT).The FHT problem was originally posed by Albert Shiryaev regarding whether there exists a boundary such that the stopping time of a Brownian motion at such a boundary is distributed according to an exponential distribution [Potiron, 2021]. Anulova [1981] answered the question affirmatively for a series of barriers but did not make any consideration of the regularity of the barriers. Theoretical investigation of this problem has led to numerous computational approaches for computing the FHT density for arbitrary boundaries (e.g. [Jaimungal et al., 2014, Zucca and Sacerdote, 2009]). A feasible computational solution for estimating the drift of a process with a desired FHT density was provided by Ichiba and Kardaras [2011] who described a representation for the distribution in terms of an expectation but did not consider the problem of estimation. While these works consider methods of finding the appropriate boundary, they deviate from our goal of recovering a *continuous* latent process due to the fact that the state of the process must be reset after the arrival of each point. Since excursions begin and end at the same location, studying the excursion distribution allows one to reconstruct the full continuous path. We discuss the similarities between the proposed method and the FHT problem in greater detail in Section 5.3.

The second relevant line of work is based on Itô's description of Brownian paths through excursions [Itô, 2020]. Using this framework, Itô described point processes over the space of excursions as a technique for analyzing properties of diffusions. Watanabe [1987] considered a mathematical construction of semimartingales through their excursions. Recent approaches considered further analyses and applications in finance [Ananova et al., 2020]. Closely related is the Azéma martingale which provides an estimate of the value of a Brownian motion when only observing the sign of its excursion [Çetin, 2012]. This has found applications in pricing Parisian options and estimating firm default risk [Çetin, 2012]. See Watanabe [2010] for a more comprehensive history on excursions and the development of Itô's excursion theory. However, the related question regarding estimating a diffusion from its excursion lengths has not yet been answered through a computationally feasible framework, which is the main purpose of this work. A closely related line of work concerns filtering problems where the observation is a point process with an intensity function given by a diffusion. A Cox process can be understood as a Poisson process with a stochastic intensity function. Jaiswal et al. [2020] described a method for computing the posterior distribution of a Cox process with intensity given by diffusion through solving a stochastic PDE. These methods can also be seen as filtering problems where the observation model is a point process with a latent continuous process. Unfortunately, the approach generally requires solving computationally intractable equations for computing the posterior. To mitigate this issue Lloyd et al. [2015] describes modulating the intensity function with a Gaussian process and describes a variational inference approach for optimizing the parameters.

Other applications of these ideas have also been considered, particularly in the case of the FHT. In survival modeling, the first hitting time of a diffusion at some region of the domain is used to determine the end of the life of a particular process. Roberts and Sangalli [2010] proposes a method for recovering diffusions based on survival data with the assumption of underlying diffusion. The approach is based on a Markov-Chain Monte Carlo method for estimating the posterior density given the survival times and the hazard function is estimated with respect to the diffusion. The continuous process can represent the state of, for example, an engine throughout its lifetime, and can be useful for gaining

an interpretation of the stresses faced by the engine. Maystre and Russo [2022] considered modeling the survival distribution in terms of the first hitting time of a discrete time Markov process and related this procedure to a temporal difference learning problem in reinforcement learning. We review additional avenues of related work in the Appendix.

### 1.2 CONTRIBUTIONS

We propose a modeling framework based on Itô's excursion theory that represents a point process over the line as a decomposition of a diffusion in terms of excursions where the excursion length corresponds to the interarrival time. Our contributions are then the following:

1. We extend the point process modeling framework based on the diffusion process, where the time stamps are determined by excursions;
2. We provide an inference algorithm for the model;
3. We demonstrate the versatility of the framework by presenting applications to many classes of distributions.
4. We illustrate the framework's utility and interpretability on a variety of synthetic and real data experiments.

## 2 BACKGROUND

To provide the initial exposition, we will assume our class of continuous processes are solutions of the one-dimensional stochastic differential equations (SDEs) driven by Wiener processes. We suppose that the latent process $Z_t$ is the solution to the SDE given by

$$\mathrm{d}Z_t = \mu(Z_t, t)\mathrm{d}t + \sigma(t)\mathrm{d}W_t \tag{1}$$

where $W_t$ is a standard Brownian motion. The object of interest in this work is to model the drift function $\mu$.

### 2.1 DEFINITIONS

Here we provide the definitions that link the point process to the diffusion in (1). Our overall goal in estimation is to find a $\mu$ such that the corresponding excursion length distribution is the same as the interarrival time distribution of the point process. Therefore, we will study the properties of excursions to describe the method. We follow the terminology of Pitman and Yor [2007] to introduce the definitions and defer to that manuscript for a more comprehensive study on the implications of Brownian excursions.

Consider a sample path $Z_t$ satisfying (1). An excursion set can be thought of as the subsets of $Z_t$ that exceed a particular function $f(t)$. The length of the excursion is then related to the times that $Z_t$ first hits and surpasses $f(t)$ and the time that $Z_t$ returns to $f(t)$. Define the set of hitting times by

$$\mathbb{H}_t := \left\{ \sup_{r\in[0,s]} \{r \mid Z_r = f(r)\} \;\middle|\; s < t \right\} \tag{2}$$

and then consider the *local time* at $f(t)$ as

$$L_t = \lim_{\epsilon\to 0} \int_0^t \frac{1}{2\epsilon} \mathbb{1}_{|Z_s - f(s)|<\epsilon} ds.$$

The local time is an increasing function that, heuristically, describes the amount of time the process $Z_s, s < t$ has spent at $f(s), s < t$ up to time $t$. We next define the *inverse local time* as $\tau_\ell = \inf\{t > 0 : L_t > \ell\}, \ell \geq 0$ which describes the time at which $Z_t$ has spent $\ell$ time at $f(t)$. An excursion straddling $(\tau_{\ell^-}, \tau_\ell)$ is then defined as

$$e^\ell := \{Z_s : s \in (\tau_{\ell^-}, \tau_\ell)\}. \tag{3}$$

where $\tau_{\ell^-}$ is the left-sided limit of the inverse local time. We note that the space of all excursions is not relevant for our purposes of modeling due to the topological properties of (2). To give an example, taking $Z_t$ as standard Brownian motion starting at zero and $f(t) = 0$ results in $\mathbb{H}_1$ being a perfect set with properties that are not practical for a modeling task. From an applied perspective, very small excursions would not be observed due to limitations on the resolution of measuring devices used to collect data. Instead, one usually considers a subset of excursion paths that have some relevance, such as excursions of minimum length or minimum height[1]. In Ananova et al. [2020], excursions reaching a minimum height of $\delta$, described as $\delta$-excursions, were considered. Under this construction, $e^\ell$ allows us to decompose continuous sample paths given by $Z_t$ into different excursions with excursion lengths indexed by $\ell$. This generates a Poisson process where excursion lengths define the interarrival times of the point process.

To illustrate these concepts, we again refer to Figure 1 where the original sample path representing $Z_t$ (blue) is decomposed into excursions above and below the line $f(t) = 0$. The arrival times (yellow circles) describe the end of an excursion, the last point in $\mathbb{H}_t$. Finally, multiple samples of excursions (grey) with length $\tau_\ell - \tau_{\ell^-}$ are illustrated to describe the relationship to the true excursion of $Z_t$. For the remainder of the text, we will suppose that $f(t) = 0$ for all $t$ and consider the set of times $\mathbb{H}_t$ when $Z_t$ returns to 0.

### 2.2 ASSUMPTIONS

We state a few more properties to ensure we can compute valid excursion densities. These are conditions on the drift $\mu$ so that the interarrival time of the excursion is finite, which in turn guarantees that the measure is a valid density.

[1] This is the interpretation of the excursion measure given by D. Williams, see Yen and Yor [2013, Chapter 6] for a detailed description.

1. The diffusion must be recurrent; i.e. $\mathbb{P}(\tau = \infty) = 0$. This is guaranteed if $\lim_{a\to\infty} S(a,t) = \infty$ and $\lim_{a\to-\infty} S(a,t) = -\infty$ where

$$S(a,t) = \int_0^a \exp\left(\int_0^b \frac{-2\mu(x,t)}{\sigma(x,t)}\mathrm{d}x\right)\mathrm{d}b$$

   for all $t$,

2. The measure counting the number of excursions must be finite; that is, we do not consider excursions that have negligible length.

3. Novikov's condition $\mathbb{E}[e^{\frac{1}{2}\int_0^\tau |Z_t|^2\mathrm{d}t}] < \infty$ for Girsanov's theorem to hold.

4. Existence of a $t$-continuous strong solution to (1), that is $\mu, \sigma$ are Lipshitz [Øksendal, 2003, Theorem 5.2.1].

A final assumption that we consider is that $\sigma(t) = 1$. This is not necessary, but as noted later in the text, by introducing an additional parameter $\delta$ regarding the minimum height of the excursion, there exists an estimation ambiguity between the $\sigma$ and $\delta$ parameters. To circumvent the estimation ambiguity, we consider recovering the transformed process given by the Lamperti transform which results in the diffusion with unit volatility. Additional details are presented in the Appendix. We also note that the drift $\mu$ can depend on history or on an additional process, but we leave the drift in its standard form for ease of exposition.

# 3 METHOD

We now describe the inference method for finding $\mu$ given a set of interarrival times. We define the interarrival times as the set $\mathbb{T}_t = \{\tau_1, \ldots, \tau_N\}$ and relate them to the set $\mathbb{H}$ by $\tau_i = \mathbb{H}_t^{(i+1)} - \mathbb{H}_t^{(i)}$, $\mathbb{H}_t^{(i)}$ being the $i^{\text{th}}$ arrival time in ascending order up to time $t$. For example, $\mathbb{T}_t$ would contain elements that are exponentially distributed in the case of a Poisson process. As mentioned in the previous section, we remove small excursions by only considering excursions with a minimum height by redefining $\mathbb{H}_{t,\delta} = \{\tau_i \in \mathbb{H}_t \mid \max_{s\in(\tau_{i-1},\tau_i)} Z_s - f(s) \geq \delta\}$ in (2) for some $\delta > 0$. We consider a minimum height so that the density remains absolutely continuous with respect to the Lebesgue measure on the positive real line. To outline the method, we first state the excursion length distribution of standard Brownian motion with minimum height $\delta$. We then perform a change of measure to find the excursion length distribution of a diffusion with drift given by $\mu$. We represent the drift $\mu$ by a neural network and optimize for its parameters via stochastic gradient descent and maximum likelihood estimation on observations of excursion lengths given by interarrival times. In the remaining of the text, we will denote a general excursion as $e$ and the excursion at time $t$ as $e_t$.

## 3.1 EXCURSION LENGTH DENSITY OF BROWNIAN MOTION

Excursion times from 0 to $\delta$ and back to 0 have the distribution given by the inverse Laplace transform of $\mathbb{E}[e^{-\lambda\tau}] = e^{-2\sqrt{2\lambda}\delta}$. Taking the inverse Laplace transform, we obtain a zero shifted Lévy distribution with scale parameter as $4\delta^2$ and PDF of

$$p_e(\tau;\delta) = \delta\sqrt{\frac{2}{\pi\tau^3}}\exp\left(-\frac{2\delta^2}{\tau}\right). \tag{4}$$

Additional details regarding this derivation are provided in the Appendix. Note that if $\sigma$ is included in the density of (4) then $\delta$ and $\sigma$ are unidentifiable, motivating the previously stated assumption on $\sigma = 1$. In some cases, it may be easier to optimize one than the other, e.g. when simulating excursions with variance $\sigma$ is easier than excursions of a minimum height $\delta$, but we will focus on a minimum height $\delta$. With the Brownian excursion length density in mind, we now consider a change of measure for the drifted case.

## 3.2 CHANGE OF MEASURE FOR EXCURSION LENGTH DENSITY OF DIFFUSION

We consider an approach inspired by Ichiba and Kardaras [2011, Section A.1] where the authors use a change of measure technique to compute the density of the FHT of a diffusion. Let $e_t$ be the value of a Brownian excursion of length $\tau$ at time $t$. The excursion length density of a diffusion follows an expectation of a Radon-Nikodym derivative between the base measure on the space of $\delta$-excursions $\mathbb{Q}^\delta_\updownarrow$ and the diffusion measure $\mathbb{P}^\mu$.

**Proposition 1** (Diffusion Excursion Density)**.** *Let $Z_t$ satisfy an SDE with drift $\mu$ such that $Z_t$ is recurrent at zero. Then the density of the excursion lengths $\tau$ of $Z_t$ is given by:*

$$p_Z(\tau) = p_e(\tau;\delta)\times \mathbb{E}_{\mathbb{Q}^\delta_\updownarrow}\left[\exp\left(\int_0^\tau \mu(e_t,t;\theta)\mathrm{d}e_t - \frac{1}{2}\int_0^\tau \mu^2(e_t,t;\theta)\mathrm{d}t\right)\right]. \tag{5}$$

*Intuition of Proof.* The proof follows a change of measure argument. The full proof is in the Appendix. □

During optimization, $p_e$ does not need to be computed since it is a constant with respect to the input data $Z_t$. However, if we consider $\delta$ as a parameter for the optimization, we may include it in the computation of the likelihood. As noted, we may also consider non-unit $\sigma$, but this comes at the expense of identifiability of $\delta$. This results in a modification of the expectation over sample paths which should have the corresponding $\sigma$. Finally, we require that the drift must be recurrent for the density to integrate to 1. Numerically, we

enforce this condition by adding a regularizer that constrains the density to approximately integrate to 1.

**Evidence Lower Bound** Following Jensen's inequality, for maximizing (5) for given data, we can also optimize

$$\log p_z(\tau;\delta) \geq \log p_e(\tau;\delta) + \mathbb{E}_{\mathbb{Q}_\ddagger^\delta}\left[\int_0^\tau \mu(e_t,t;\theta)\mathrm{d}e_t - \frac{1}{2}\int_0^\tau \mu^2(e_t,t;\theta)\mathrm{d}t\right] \tag{6}$$

and therefore we need only to maximize (6) rather than (5). This may be beneficial in scenarios where numerical errors may make the calculation of the exponential unstable. We detail an algorithm for estimating the drift $\mu$ from data in Algorithm 1.

**Algorithm 1** Inference for latent diffusion from arrival times

1: **Input:** Sequences of interarrival times: $\mathbb{T} = \left\{\tau_1^{(j)}, \tau_2^{(j)}, \ldots, \tau_{n_j}^{(j)}\right\}_{j=1}^N$
2: **Initialize Parameters:** Parameters of drift $\mu(x,t)$, step size $\Delta t$, total time $T$, initial state $X_0$, minimum height $\delta$ and variance $\sigma$, number of points in expectation $K$.
3: Sample $K$ Brownian excursions using Vervaat transform [Vervaat, 1979] of a Brownian bridge and the Euler-Maruyama method $\mathbb{E} = \{e_i\}_{i=1}^K$.
4: Filter $\mathbb{E}$ by discarding $e_i$ where $\max e_i < \delta$.
5: Numerically compute (6) for the data $\mathbb{T}$ with excursions $\mathbb{E}$ computed above.
6: Maximize (6) using gradient decent with respect to the parameters of $\mu$ and $\delta$.
7: Repeat for $N$ iterations

### 3.3 MULTIDIMENSIONAL PROCESSES

The approach extends to multidimensional diffusions that have interacting components, where we consider excursions away from each axis. The idea is illustrated in Figure 2, where a diffusion in the 2-D plane generates excursions from either axis, each axis corresponding to a different mark. This leads to a dependence between the two classes through the drift function. Our approach for calculating the corresponding $\mu$ from the data is analagous to the 1-dimensional case where we compute expectations over Brownian excursions with the same lengths as the interarrival time observations and repeat over each dimension. In this case, $Z_t$ is a $d$-dimensional diffusion. From there, the likelihood of the data is maximized using a single multi-dimensional drift function that governs the relationships between the different marked processes. To estimate the drift from observations of the interarrival times of different coordinates, we compute the

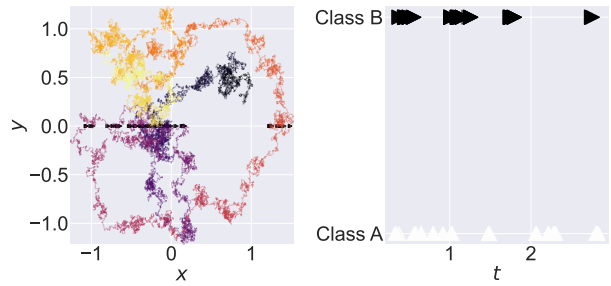



Figure 2: Schematic of the multidimensional framework. Left: the 2-dimensional latent diffusion crosses the axes producing points in the point process. Right: arrival times of the point process generated by the diffusion process on the left. Up triangles represent excursions from the $y$-axis and left triangles represent excursions from the $x$-axis.

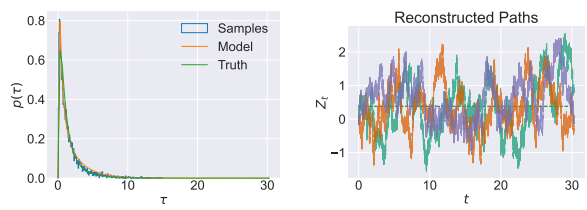



Figure 3: Example of estimated log-normal renewal process with samples generated from learned diffusion. Left: histogram of samples compared with the true density and estimated density. Right: learned sample paths with excursion lengths corresponding to the histogram. The dashed line corresponds to $\delta$.

change of measure as in (5) jointly over all components:

$$p_Z(\tau^{(1)},\ldots,\tau^{(d)}) = \prod_{k=1}^d p_e(\tau^{(k)};\delta)$$
$$\mathbb{E}_{\mathbb{Q}_\ddagger}\left[\exp\left(\int_0^{(\bigvee_{k=1}^d \tau^{(k)})\wedge T} \mu(\mathbf{e}_t,t;\theta)\mathrm{d}\mathbf{e}_t - \frac{1}{2}\int_0^{(\bigvee_{k=1}^d \tau^{(k)})\wedge T} \mu^\dagger\mu(\mathbf{e}_t,t;\theta)\mathrm{d}t\right)\right] \tag{7}$$

where $\mathbf{e}_t$ is a multidimensional excursion process with the excursion length of the $i^{\text{th}}$ component being $\tau^{(i)}$. Specifically, $\mathbf{e}_t$ has zeros only at the time points where the corresponding component has a realization.

## 4 SIMULATING POINT PROCESSES

Simulating realizations of the proposed point process follows from existing simulation techniques for diffusions stopped at a boundary. For example, when solving certain

linear elliptic PDEs, the solution is based on computing the first hitting time of diffusion on the boundary of the domain [Gobet and Menozzi, 2010]. We propose a method based on the Euler-Maruyama method where excursions are simulated by computing full sample paths and finding the times when an excursion occurs. Importantly, this allows for both sampling the full sample path based on the fitted drift and obtaining samples of interarrival times. We summarize the heuristic for the sampling procedure in Algorithm 2. Figure 3 shows an example of estimating a log-normal distribution. The figure on the left shows the histogram of samples in blue, the model probabilities computed using (5) in orange, and the true density for a log-normal distribution in green. On the right, samples of different trajectories whose excursions to the blue dotted line result in interarrival times distributed according to the left.

**Algorithm 2** Sampling arrival times

1: **Input:** Parameters of drift $\mu(x,t)$, step size $\Delta t$, total time $T$, initial state $X_0$, minimum height $\delta$ and variance $\sigma$
2: Sample using Euler-Maruyama a sample path from $X_0$ to $X_t$ using step size $\Delta t$ and variance $\sigma\Delta t$:

$$X_{t+1} \sim \mathcal{N}(X_t + \mu(X_t,t)\Delta t, \sigma\Delta t)$$

3: Compute the set $\tau_0 = \{X_s = 0 : s \in [0,T]\}$
4: Filter $\tau_0 \to \tau_\delta$ where $\tau_\delta = \tau_0 \setminus \{t_i \in \tau_0 : \max_{t_i<s<t_{i+1}} X_s < \delta\}$.
5: **Return.** Set of arrival times $\tau_\delta$.

The simulation algorithm for a more complicated process follows a similar procedure, for example, the drift term $\mu(x,t)$ can be replaced by the history-dependent function $\mu(t,\mathcal{H}_t)$, which we discuss further in the Appendix.

# 5 PRACTICAL CONSIDERATIONS

Here we discuss some practical considerations regarding the model. We first describe how partitioning the space of excursions can lead to a marked point process without resorting to a $d$-dimensional latent process. We then describe a result regarding the family of interarrival distributions the method can represent, We finally describe the relationship between the FHT problem and the excursions approach.

## 5.1 MULTI-DIMENSIONAL POINT PROCESS FROM A SINGLE ONE-DIMENSIONAL DIFFUSION

A unique property of the proposed method is the ability to represent a multi-dimensional point process with a single one-dimensional latent diffusion process. The main idea comes from partitioning the measure on the space of excursions to correspond to different classes. In the simplest case, the arrival can come from an excursion above or below the reference level. The structure that should be maintained is a natural ordering between the marks of the point process for discrete marks. For example, in the case of a two-dimensional process, one set of marks should always be greater than the other. This can then correspond to the running maximum and running minimum times. Note that this differs from the $d$-dimensional process that was described in Section 3.3 which assumes a $d$-dimensional noise source. We describe potential applications in a financial setting in the Appendix where bids and asks in an opaque market are generated by the running maximum or running minimum process.

## 5.2 EXPRESSIVENESS

A relevant question asks how expressive the class of interarrival times generated by excursions is. We characterize this in the following remark:

**Remark 1.** *Consider an excursion length distribution for a fixed $\delta$ as $p_{\updownarrow,\delta}$ with support on $\mathbb{R}_+$ and a distribution that we wish to approximate using the excursion length of an Ito diffusion as $p_\star$. Define the function space such that the excursion distribution is a density as a subset of Lipschitz functions $\mathrm{Exc} \subset \mathrm{Lip}(\mathbb{R}_+, \mathcal{D})$. For a fixed integration time $T$, the excursions of the diffusion with drift $\mu$ can represent $p_\star$ if the 1-Wasserstein distance between the two is less than*

$$\sup_{\mu \in \mathrm{Exc}(\mathbb{R}_+,\mathcal{D})} \sqrt{\frac{T^2}{2}\mathbb{E}_{X_t\sim\mathbb{Q}}\left[\int_0^T \mu dX_t - \frac{1}{2}\int_0^T \mu^2 dt\right]}.$$

*Intuition of Proof.* This follows bounding the Wasserstein distance using Pinsker's inequality. The full proof is in the Appendix. □

The remark allows for a simple condition on whether a distribution can be approximated by the proposed method for fixed integration time and excursion height. The 1-Wasserstein distance is easy to calculate since it is the difference between the CDFs, making the remark useful in practice since one can certify whether a distribution can be represented using the change of measure. More generally, Pitman and Yor [2007, Section 6] discusses the applicability of functions of Brownian excursions representing the full class of stable Lévy processes at the cost of relaxing many of the assumptions on the drift of the process.

## 5.3 COMPARING EXCURSIONS AND FIRST HITTING TIMES

We now describe in greater detail the similarities and differences between studying excursions and FHTs related

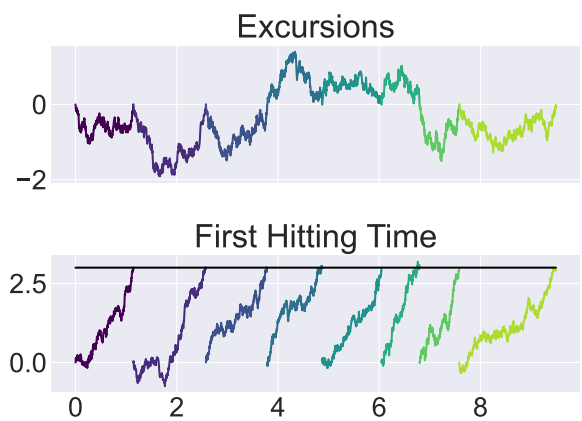



Figure 4: Schematic comparing excursions (top) with first hitting times (bottom). The excursions result in a continuous path whereas the first hitting time approach results in discontinuous paths that need to be reset after each arrival.

to the motivation of this work. Figure 4 provides a qualitative description of the difference between the excursion representation we consider here and the first hitting time approach.

**Interpretations of Excursions and Hitting Times** Excursions and hitting times share many similarities, since an excursion may be seen as the FHT to $\delta$ and back to $0$ again. This relationship is specified in detail in the proposition found in the Appendix where the Laplace transform of the first hitting time distribution is given. In that sense, the primary reason to consider one representation versus the other is the interpretation of the underlying phenomena being observed. The first hitting time density requires the assumption that the particle is returned to its original state at $t = 0$ for every subsequent arrival. This makes the full sample path discontinuous since the particle must hit a level $\delta \neq X_0$ and then instantly return to $X_0$. On the other hand, considering an excursion from a level yields a continuous sample path for a full sequence of observations.

**Densities Described by Laplace Transforms** Both have been historically studied through their Laplace transforms. For a univariate autonomous SDE, there exists a Sturm-Liouville problem associated with the Laplace transform of the excursion length distribution [Yen and Yor, 2013]. The FHT density has also been studied through the same mathematical formulation. However, working with the Laplace transform representation is difficult as it is necessary to invert the Laplace transform to obtain the density. Inverting the Laplace transform is numerically unstable and also prone to numerical errors. We present a detailed description of the approach in the Appendix.

**Connection to the Running-Maximum and the Drawdown Processes** Finally, one important property of excursions is the relationship between the running maximum of a Brownian motion and the zeros of a reflected Brownian motion. Specifically, the identity

$$\sup_{s<t} W_s - W_t \stackrel{d}{=} |W_t|$$

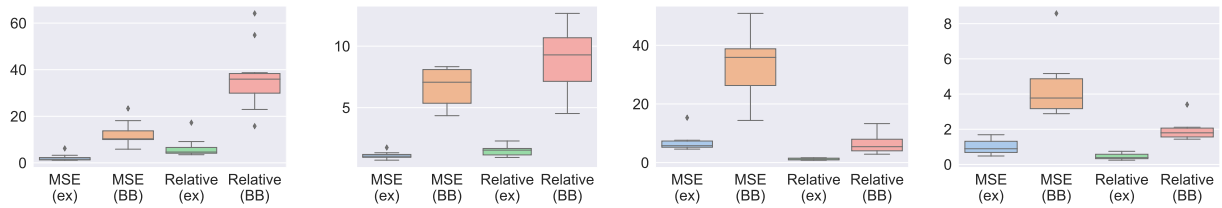
where $W_t$ is Brownian motion. This allows the interpretation where excursions are related to times the process reaches its running maximum. This interpretation is not possible when considering only FHT where the diffusion must be reset at each arrival.

# 6 EXPERIMENTS

We now consider the modeling framework in a number of synthetic and real data experiments. The first set of experiments is based on observing the zero times of different diffusions. These experiments examine how well the true drift can be recovered using the proposed estimator. The baseline for these experiments is a standard SDE regression algorithm that does not consider Brownian excursions but instead considers Brownian bridges. The second set of experiments analyzes the proposed estimator in representing the interarrival distributions of canonical renewal processes. Finally, we provide a real example regarding a physical process where the underlying behavior is posited to be related to a continuous process. In this case, we consider how well the learned diffusion agrees with the latent factor that's known to cause the point process.

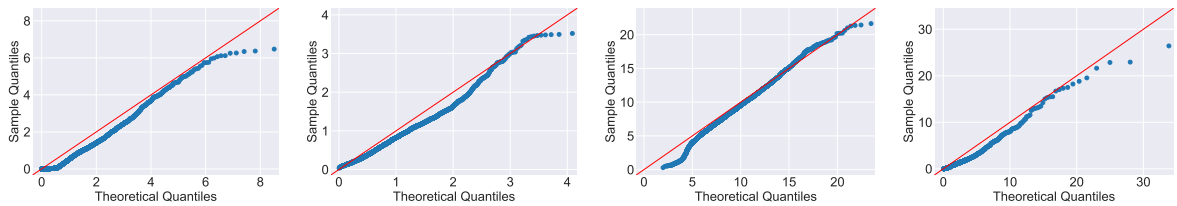
## 6.1 RECOVERY OF DRIFT FROM EXCURSION LENGTHS

In order to validate the method in the context of a latent diffusion, we consider a series of experiments on how well the method can recover the drift of the latent diffusion. This would correspond to a real scenario where the data are generated according to excursions of a diffusion. We observe $\{(\tau_i, m_i\}_{i=0}^N$ where $m \in \mathbb{N}$ is the mark corresponding to the dimension where the excursion occurs. We compute this for different choices of $\mu$ and compute the mean squared error (MSE) between the estimated $\hat{\mu}$ and the true $\mu$. The observation are generated by first simulating a diffusion and finding the zero times of the sample paths. We provide full details on the different models in the Appendix. The results are illustrated in Figure 5 where we compare the estimation based on maximizing the likelihood of diffusion with unknown drift based on a path integral estimator. The difference between the estimators is that the proposed one uses the expectation over excursions whereas the other considers Brownian bridges. The Brownian bridges do not use the complete information of the problem, and therefore result in




(a) 10-$d$ $\mu = -\tanh(x)$; (b) 10-$d$ $\mu = -x^3$; (c) 2-$d$ $\mu = circle$; (d) 5-$d$ $\mu = -x$;

Figure 5: Examples of point processes with interarrival distributions given by different $\mu$. Interarrival times are given as a crossing of 0. Results from 10 runs. Relative refers to the MSE divided by the norm of the solution. BB refers to the Brownian bridge estimator and Ex refers to the proposed excursion estimator.




(a) Exponential (b) Weibull (c) Gamma (d) Log-Normal

Figure 6: Comparison of QQ plots for Poisson process, Weibull, Gamma, and log-normal renewal processes. All are fit using MLE with 200 samples from the specified renewal process.

higher errors as well as higher variances than the Brownian excursions.

**History and Exogenous Signal Dependent Processes** We consider an extension of the previous experiment in the 1-dimensional case where we recover the coefficient of the drift when it is dependent on either the history of hitting times or on an observed exogenous signal. Specifically, we define a drift of the following form

$$\mu(X_t, \mathcal{S}_t) = -X_t + w\varphi(t - \mathcal{S}_t)$$

where $w$ is the coefficient of interest and $\varphi(\cdot)$ is a known kernel that influences the drift based on the observed history or exogenous process. We choose $\varphi$ to be the exponential kernel defined as $\varphi = \exp(-(t - \mathcal{S}_t)/\eta)$ with $\eta$ a fixed parameter. In the case of history dependence, $\mathcal{S}_t$ is given as $\mathbb{H}_t$ while the exogenous process is given by a randomly generated signal generated using uniform increments, full details are given in the Appendix. We compare the squared error of the estimated value of $\hat{w}$ to the true value of $w$ as recovered by the same Brownian bridge estimator and by the proposed excursion estimator in Table 1. The results again suggest that the proposed estimator achieves better performance than the relevant baseline of regressing a SDE to the data.

| | | $w = 0.5$ | $w = 1$ | $w = 2$ |
|---|---|---|---|---|
| History | BB | 0.247(0.003) | 0.999(0.012) | 3.996(0.021) |
| | Ex | 0.099(0.091) | 0.067(0.068) | 0.903(0.376) |
| Input | BB | 0.248(0.003) | 0.997(0.018) | 3.786(0.580) |
| | Ex | 0.148(0.090) | 0.069(0.113) | 1.301(2.012) |

Table 1: Squared error $(\hat{w} - w)^2$ of the history coefficient for the Brownian bridge estimator (BB) versus the Brownian excursion estimator (Ex).

## 6.2 ESTIMATING AND SAMPLING POINT PROCESSES

Next, we are interested in determining how well the proposed method can represent some canonical point processes. In this case, we consider the homogeneous Poisson process, a Gamma renewal process, a Weibull renewal process, and a log-normal renewal process with 40 samples with 5 realizations in each sample for a total of 200 points. Full parameters of the distributions are given in the Appendix. This experiment tests both inference (Algorithm 1) and the sampling (Algorithm 2). We plot a QQ plot of the samples generated by the excursion lengths versus the theoretical quantiles in Figure 6. The figures suggest that the estimation and sampling methods are able to capture the distributions of the point process.

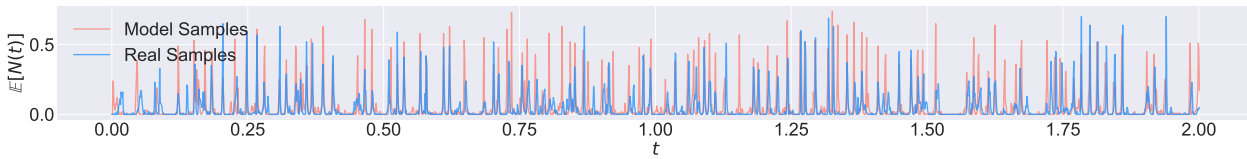



Figure 7: Comparison of the empirical histogram of arrival times versus the sampled histogram of arrival times. Sampled spikes generally align with true spikes.

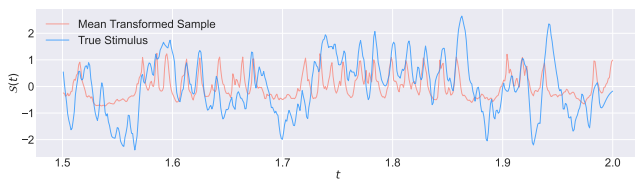



Figure 8: Average of learned sample on $t \in [1.5, 2]$ paths compared with the true stimulus. Spikes of the transformed learned stimulus generally align with the true stimulus.

### 6.3 REAL DATA

Finally, we consider a neuroscience dataset where the firing of mouse neurons is recorded as a function of an external stimulus as described in Tripathy et al. [2013]. Full details of the dataset are in the Appendix. Our main goal for this experiment is to determine how well the model can fit this data and more importantly determine whether we can use the estimated latent path $Z_t$ as a signal that relates the original stimulus to the observed neuron spike times. We illustrate these findings in Figures 7 and 8 where we compare the histograms of the averaged point processes generated by the true data samples and the estimated excursion process (Figure 7) and compare the learned stimulus to the true stimulus (Figure 8). The learned stimulus was obtained by transforming the sampled path $Z_t$ to $\tilde{Z}_t = a \log(\mathbb{E}[Z_t]) + b$ where $a, b$ are computed according to least squares with respect to the true stimulus. We use the $\log$ transformation since the peaks of the stimulus correspond to new arrivals whereas the zeros of the learned process correspond to new arrivals – applying the $\log$ then transforms the zeros to peaks. In both cases, the alignment of the spikes between the learned and the true signal is well maintained.

## 7 DISCUSSION

We proposed a framework that allows for interpreting the arrivals of a point process in terms of a latent diffusion. We described extensions to cases where the point process is multi-dimensional and depends on history or an exogenous signal. The numerical results suggest that the estimator and the framework is useful for modeling a variety of point processes and outperforms standard SDE regression techniques. Additionally, the results on neural data demonstrate the applicability of the proposed framework in scenarios where recovery of the unobserved continuous latent process is beneficial for analyzing a particular temporal point process.

**Limitations** The framework has a number of limitations as well. While we empirically validated the ability to recover the correct drift, we have no identifiability proof that guarantees the true drift will be recovered. In cases where the $\mu$, $\delta$, and $\sigma$ are all parameters, for example, identifiability does not hold, and restricting to a smaller class of parameters is necessary. Proving consistency of the estimator could be considered for follow-up work. On a practical front, there are many situations where having an interpretation in terms of a continuous process is not appropriate. In these cases, more traditional point process models, such as those that rely on intensities, should be considered for the modeling task.

## ACKNOWLEDGEMENT

We appreciate the great support from professor Nathan Urban by sharing the valuable neuroscience dataset. We thank Zaeem Burq and Kashif Rasul for invaluable discussions regarding the theory of excursions. We additionally thank Jessica Loo, Wei Deng, Volodymyr Volchenko, and Yuriy Nevmyvaka for helpful feedback on the manuscript. AH and VT were supported in part by the Air Force Office of Scientific Research under award number FA9550-20-1-0397. AH was also partially supported by NSF GRFP.

### References

Anna Ananova, Rama Cont, and Renyuan Xu. Excursion risk. *arXiv preprint arXiv:2011.02870*, 2020.

SV Anulova. On markov stopping times with a given distribution for a wiener process. *Theory of Probability & Its Applications*, 25(2):362–366, 1981.

Ana Babus and Péter Kondor. Trading and information diffusion in over-the-counter markets. *Econometrica*, 86 (5):1727–1769, 2018.

Tomas Björk, Yuri Kabanov, and Wolfgang Runggaldier. Bond market structure in the presence of marked point processes. *Mathematical Finance*, 7(2):211–239, 1997.

Umut Çetin. Filtered azéma martingales. *Electronic Communications in Probability*, 17:1–13, 2012.

Khem Raj Ghusinga, John J Dennehy, and Abhyudai Singh. First-passage time approach to controlling noise in the timing of intracellular events. *Proceedings of the National Academy of Sciences*, 114(4):693–698, 2017.

Emmanuel Gobet and Stéphane Menozzi. Stopped diffusion processes: boundary corrections and overshoot. *Stochastic Processes and Their Applications*, 120(2):130–162, 2010.

Tomoyuki Ichiba and Constantinos Kardaras. Efficient estimation of one-dimensional diffusion first passage time densities via monte carlo simulation. *Journal of applied probability*, 48(3):699–712, 2011.

Kiyosi Itô. Poisson point processes attached to markov processes. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971)*, volume 3, pages 225–239, 1972.

Kiyosi Itô. Poisson point processes attached to markov processes. In *Contributions to Probability Theory*, pages 225–240. University of California Press, 2020.

Sebastian Jaimungal, Alexander Kreinin, and Angel Valov. The generalized shiryaev problem and skorokhod embedding. *Theory of Probability & Its Applications*, 58(3): 493–502, 2014.

Prateek Jaiswal, Harsha Honnappa, and Vinayak Rao. Variational inference for diffusion modulated cox processes. 2020.

Nicholas C Lammers, Yang Joon Kim, Jiaxi Zhao, and Hernan G Garcia. A matter of time: Using dynamics and theory to uncover mechanisms of transcriptional bursting. *Current opinion in cell biology*, 67:147–157, 2020.

Chris Lloyd, Tom Gunter, Michael Osborne, and Stephen Roberts. Variational inference for gaussian process modulated poisson processes. In *International Conference on Machine Learning*, pages 1814–1822. PMLR, 2015.

Lucas Maystre and Daniel Russo. Temporally-consistent survival analysis. In *Advances in Neural Information Processing Systems*, 2022.

Bernt Øksendal. Stochastic differential equations. In *Stochastic differential equations*, pages 65–84. Springer, 2003.

J Pitman and M Yor. Itô's excursion theory and its applications. *Japanese Journal of mathematics*, 2(1):83, 2007.

Yoann Potiron. Existence in the inverse shiryaev problem. *arXiv preprint arXiv:2106.11573*, 2021.

Marian-Andrei Rizoiu, Young Lee, Swapnil Mishra, and Lexing Xie. Hawkes processes for events in social media. In *Frontiers of multimedia research*, pages 191–218. 2017.

Gareth O Roberts and Laura M Sangalli. Latent diffusion models for survival analysis. *Bernoulli*, 16(2):435–458, 2010.

Laura Sacerdote and Cristina Zucca. Threshold shape corresponding to a gamma firing distribution in an ornstein-uhlenbeck neuronal model. *Scientiae Mathematicae Japonicae*, 58(2):295–306, 2003.

Dietrich Stoyan and Antti Penttinen. Recent applications of point process methods in forestry statistics. *Statistical science*, pages 61–78, 2000.

Sandya Subramanian, Riccardo Barbieri, and Emery N Brown. Point process temporal structure characterizes electrodermal activity. *Proceedings of the National Academy of Sciences*, 117(42):26422–26428, 2020.

Shreejoy J Tripathy, Krishnan Padmanabhan, Richard C Gerkin, and Nathaniel N Urban. Intermediate intrinsic diversity enhances neural population coding. *Proceedings of the National Academy of Sciences*, 110(20):8248–8253, 2013.

Wim Vervaat. A relation between brownian bridge and brownian excursion. *The Annals of Probability*, pages 143–149, 1979.

Shinzo Watanabe. Construction of semimartingales from pieces by the method of excursion point processes. In *Annales de l'IHP Probabilités et statistiques*, volume 23, pages 297–320, 1987.

Shinzo Watanabe. Itô's theory of excursion point processes and its developments. *Stochastic processes and their applications*, 120:653–677, 2010.

Ju-Yi Yen and Marc Yor. Local times and excursion theory for brownian motion. *Lecture Notes in Mathematics*, 2088:x+–135, 2013.

Cristina Zucca and Laura Sacerdote. On the inverse first-passage-time problem for a wiener process. *The Annals of Applied Probability*, pages 1319–1346, 2009.