SQUID: A Bayesian Approach for Physics-Informed Event Modeling

Sumantrak Mukherjee
DFKI*
Kaiserslautern, Germany
Sumantrak.Mukherjee@dfki.de

Sebastian Vollmer
DFKI*, RPTU
Kaiserslautern, Germany
Sebastian.Vollmer@dfki.de

Gerrit Großmann
DFKI*
Kaiserslautern, Germany
Gerrit.Grossmann@dfki.de

Abstract

We study the spatiotemporal evolution of event processes (e.g., gunshots, vehicle thefts, earthquakes). Classical approaches model the intensity of a spatio-temporal point process directly. Instead, we propose to infer a latent velocity field that transports the event intensity via a continuity equation as a more interpretable and mechanistic alternative description. The inference is performed in a Bayesian framework, using a Gaussian process to model the vector field. This provides calibrated uncertainty for both the vector field and the resulting forecasts. We evaluate the method on synthetic data to demonstrate efficient simulation and inference.

1 Introduction

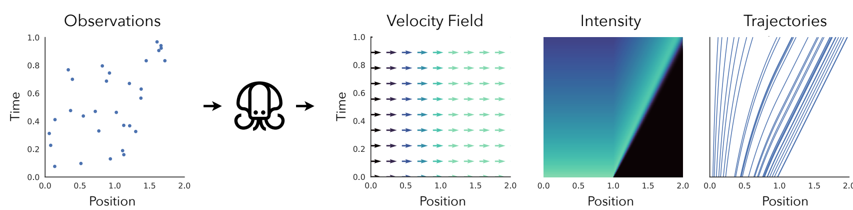

Figure 1: **Schematic.** SQUID takes spatio-temporal event data as input and regresses a velocity field (here: $v(x) = \min(x, 1)$, color indicates magnitude). This example is one-dimensional and the field does not change in time. Given an initial intensity, the continuity equation transports it over time, yielding a time-varying intensity. Estimating the field is equivalent to identifying each event with a particle trajectory. Note that SQUID computes a posterior distribution over vector fields.

Spatio-temporal *events* are ubiquitous across scientific domains. Applications span seismology, epidemiology (case onsets), ecology (species sightings), transportation (traffic collisions, taxi pickups), environmental hazards (lightning strikes, wildfire ignitions), urban analytics (911 calls, crime), and conflict studies (IED incidents, protests). Here, an event is any occurrence with a specific location and timestamp that can be treated as instantaneous.

Spatio-temporal point processes (STPPs) are the de facto standard for modeling event data [1]. Traditionally, statistical models have been used [2]; more recently, neural methods have become a

*DFKI: Deutsche Forschungszentrum für Künstliche Intelligenz GmbH (German Research Center for Artificial Intelligence).

1st Workshop on Differentiable Systems and Scientific Machine Learning @ EurIPS 2025.

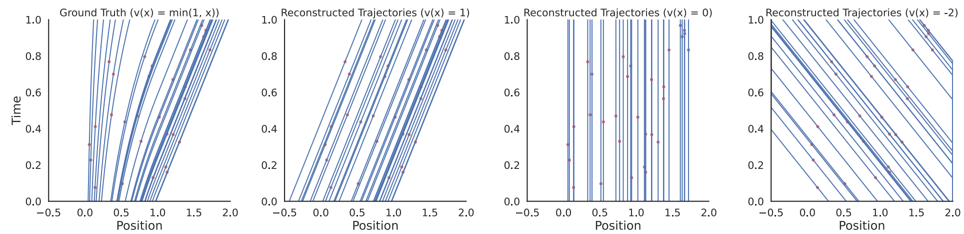

Figure 2: Consider again the event data from Figure 1. We reconstruct particle trajectories using different velocity fields: the ground-truth field (left) and three alternatives with increasing distance from the ground truth. The ground-truth field produces trajectories that locally concentrate more densely around observed events, yielding a higher (surrogate) data likelihood.

popular backbone for these models [3]. However, these approaches often struggle with interpretability and with providing reliable uncertainty estimates.

In this work, we take a slightly different view by focusing on *intensity drift*, a high-level shift in the spatial distribution of events over time (cf. Figure 1). Concretely, we aim to identify a velocity field that captures this drift while remaining consistent with the theory of spatio-temporal point processes which we estimate within a Bayesian framework.

Such drift appears across domains—ecology (seasonal migration corridors), epidemiology (advancing incidence fronts), urban analytics (shifting hotspots and displacement), crime (spatial diffusion driven by socioeconomic factors), and climate science (jet-stream shifts). Related work includes velocity estimation for spatio-temporal point patterns [4] and continuity-based transport or dynamic optimal transport [5, 6], as well as action matching [7] and normalizing flows [8].

Contribution. We introduce `SQUID` (Spatiotemporal QUantification of Intensity Drift), a Bayesian method to estimate a vector field.We employ a surrogate likelihood formulation for efficient inference and provide a fast simulation procedure to generate synthetic datasets for evaluation.

2 Background and Problem Setting

STPPs specify a probability law over event sequences, each sequence typically having the form $[(t_1, x_1, y_1), \ldots, (t_n, x_n, y_n)]$ with increasing times and $(x_i, y_i) \in \mathcal{S} \subset \mathbb{R}^2$ spatial coordinates within a domain $\mathcal{S}$ and for a given time horizon T. Inference seeks the data-generating mechanism, often from a single realization. The traditional parameterization uses an intensity that takes the history of past events as input and outputs an instantaneous rate for all locations for all future timepoints. For instance, given the event history $\mathcal{H}_t$ up to time t, the future rate at time $t' > t$ and location (x, y), denoted $\lambda_{t'}(x, y \mid \mathcal{H}_t) \geq 0$, is the instantaneous event rate. This method also yields an autoregressive generation mechanism. Free parameters are fitted by maximizing the standard log-likelihood $\sum_{i=1}^{n} \log \lambda_{t_i}(x_i, y_i \mid \mathcal{H}_{t_i}) - \int_0^T \int_{\mathcal{S}} \lambda_t(x, y \mid \mathcal{H}_t)\, d(x, y)\, dt$.

We likewise observe a single sequence but infer a latent velocity field $\mathbf{v}(x, y)$ interpreted as a transport map that advects the intensity via the mass-conserving continuity equation. Given an initial intensity (that does not depend on the event history) $\lambda_0 : \mathcal{S} \to \mathbb{R}_{\geq 0}$, we evolve λ forward in time (e.g., using semi-Lagrangian integration); thus a static vector field yields a time-varying intensity $\lambda_t : \mathcal{S} \to \mathbb{R}_{\geq 0}$.

Assumptions. We ignore event history and model only the evolution of the intensity over time (a Markov-type assumption). We omit event types/marks, though they could be added. We assume total intensity is conserved over time, a constraint that can be relaxed. We place a spatially smooth prior on $\mathbf{v}$ to avoid pathological, point-concentrated flows. The assumption of a time-invariant vector field is simplifying but can be extended to time-dependent fields. Jumping ahead, `SQUID` sidesteps specifying the initial intensity by learning the vector field nonparametrically from data.

3 Our Method: SQUID

3.1 Efficient Simulations

STPPs define probability distributions over event sequences from which we can sample given: an initial intensity field, a velocity field, a time horizon, and a spatial domain. Following the standard autoregressive interpretation of STPPs, one can proceed as follows:

1. Discretize the spatial and temporal domain into a $D \times D \times D$ grid.
2. Numerically solve the PDE describing the evolution of the intensity over time.
3. For each grid cell with value w, draw $u \sim \mathcal{U}(0, 1)$ and record an event if $u < w/V$, where V is the cell volume.

This procedure becomes exact as $D \to 0$, assuming smooth PDE behavior and appropriate boundary conditions. However, it is computationally inefficient: solving the PDE on fine grids is costly, and choosing D is nontrivial. Likewise, thinning-based methods do not require discretization but still require the PDE integration and yield potentially very high rejection rates.

We therefore propose a more efficient simulation scheme:

1. Sample the total number of events $n \sim \text{Poisson}\big(T \int_{\mathcal{S}} \lambda_0(x, y)\, dx\, dy\big)$.
2. For each event, sample surrogate spatial coordinates (x'_i, y'_i) independently from a distribution proportional to $\lambda_0(x, y)$ (e.g., via rejection sampling or a high-resolution 2D grid).
3. Sample event times independently $t_i \sim \mathcal{U}(0, T)$.
4. Compute the final spatial positions (x_i, y_i) by integrating (x'_i, y'_i) forward from 0 to t_i using the velocity field $\mathbf{v}$ with an ODE-based integrator.

We call this a particle-based interpretation of STPPs. However, it is important to notice that the particle trajectories (see also Figure 1 and 2) are a computational trick without physical meaning.

Correctness. We provide a short proof sketch for the equivalence of the two methods. First, note that all events are independent; thus sampling an event sequence is equivalent to sampling the total number of events and then sampling all individual events independently. Second, sampling the number of events is simple: by standard STPP theory, we know that it only depend on the overall integral of the intensity [9]. By exploiting that the PDE is mass-conserving, we can factor out T and easily sample the the total event count. Third, the occurrence time t_i is uniformly distributed on $[0, T]$, because the overall mass is preserved. Lastly, the location of an event i has density proportional to $\lambda_{t_i}(x, y)$; interpreting the intensity flow as a normalizing flow, sampling from the initial density and pushing the sample forward to t_i is equivalent to evolving the density to t_i and then sampling.

3.2 Vector Field Inference

To represent the vector field computationally, we discretize it into a tensor $T \in \mathbb{R}^{M \times M \times 2}$, where $T[x', y', :]$ stores the x and y component of the field at the discretized position (x', y'). For efficiency, we upscale the field to a higher resolution before solving the flow, for instance using bilinear interpolation. The prior distribution over the vector field follows a Gaussian process (conceptually a vector-valued GP), typically with an RBF kernel (cf. [10]). Hyperparameters, such as the kernel length scale, could also be treated as Bayesian variables.

A naive Bayesian inference approach would compute the log-likelihood of the data for a given vector field directly from the likelihood equation above. This, however, would require a Bayesian encoding of the initial intensity λ_0, followed by numerical PDE integration to obtain the likelihood of each event. While conceptually straightforward, this procedure is computationally demanding. First, estimating λ_0 dramatically increases the Bayesian search space. Second, numerically integrating the PDE is expensive, specifically if we wish to retain gradients for efficient samplers such as NUTS. A possible alternative would be to integrate all points backward in time and then estimate the initial intensity λ_0 directly via kernel density estimation (KDE). However, this approach remains computationally heavy, and the KDE introduces arbitrary design choices.

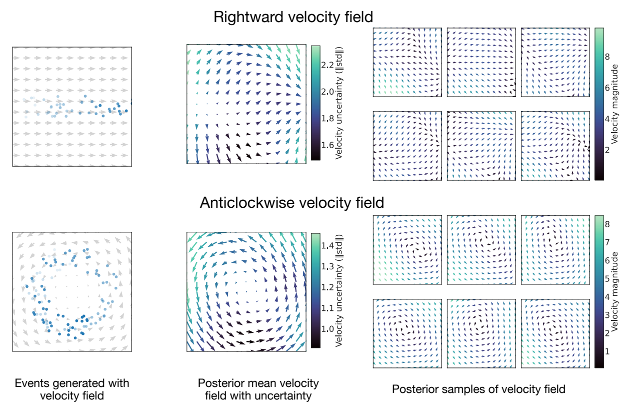

Figure 3: For two synthetic velocity fields — rightward horizontal (top) and anticlockwise rotating (bottom) — we show, from left to right, the ground truth events and velocity fields, the posterior mean with uncertainty and six posterior samples.

We propose a more practical solution based on the particle-based interpretation. Specifically, we first integrate all observed points backward to $t = 0$ and forward to $t = T$. For each event, we then estimate its likelihood based on the local density of nearby trajectories, for example using the inverse of the average distance to its ten nearest neighbors. This results in a surrogate likelihood that leverages the relationship between probability density (as expressed by the intensity) and the empirical density of trajectories (cf. Figure 2).

Overall, we can employ a wide range of Bayesian inference methods to sample from the posterior distribution of the vector field. `SQUID` further enables efficient mini-batch inference: one can subsample events or restrict training to short time intervals.

4 Experimental Results

Our experiments 3 show that the method is sensitive to hyperparameters, and although NUTS yields robust inference, it is slow. Hence, we employ our proposed inference approach, which efficiently retrieves velocity fields. These results serve as proof of concept for the proposed framework.

5 Related Work

Classical models for STPPs describe event intensities through autoregressive or self-exciting formulations, particularly in applications such as crime modeling [11]. Recent efforts have focused on estimating latent velocity fields underlying event propagation, for instance using log-Gaussian Cox processes to infer spatial spread velocities in epidemiological data [12]. Bayesian formulations of these models provide calibrated uncertainty and efficient inference [13, 14], while GP priors over vector fields enable nonparametric estimation of dynamical systems [15]. Complementary work leverages physics-informed regression constrained by partial differential equations [16] and integrates neural differential equations to recover underlying dynamics from sparse event observations [17]. From a broader transport perspective, approaches such as neural dynamic optimal transport and action matching formulate density evolution through the continuity equation [18, 7].

6 Conclusions and Future Work

We introduced `SQUID`, a Bayesian framework for inferring latent velocity fields that drive the evolution of event intensities through a velocity equation. The method provides an interpretable, uncertainty-aware view of spatio-temporal dynamics while remaining computationally efficient. Future work will focus on disentangling large-scale drift from local self-excitation effects. Separately, we will integrate domain-informed priors for specific applications and apply the approach to real-world datasets such as urban mobility or epidemiological event records.

7 Acknowledgment

This work was funded by the Bundesministerium für Bildung und Forschung - (BMFTR) under Grant 01IW23005. We thank the anonymous reviewers of the "1st Workshop on Differentiable Systems and Scientific Machine Learning @ EurIPS 2025" for their constructive feedback.

References

[1] Daryl J Daley and David Vere-Jones. *An introduction to the theory of point processes: volume I: elementary theory and methods*. Springer, 2003.

[2] Jonatan A González, Francisco J Rodríguez-Cortés, Ottmar Cronie, and Jorge Mateu. Spatio-temporal point process statistics: a review. *Spatial Statistics*, 18:505–544, 2016.

[3] Sumantrak Mukherjee, Mouad Elhamdi, George Mohler, David A Selby, Yao Xie, Sebastian Vollmer, and Gerrit Grossmann. Neural spatiotemporal point processes: Trends and challenges. *arXiv preprint arXiv:2502.09341*, 2025.

[4] Erin M Schliep and Alan E Gelfand. Velocities for spatio-temporal point patterns. *Spatial Statistics*, 29:204–225, 2019.

[5] Gabriel Peyré, Marco Cuturi, et al. Computational optimal transport: With applications to data science. *Foundations and Trends® in Machine Learning*, 11(5-6):355–607, 2019.

[6] Karthik Elamvazhuthi. Benamou-brenier formulation of optimal transport for nonlinear control systems on rd. *arXiv preprint arXiv:2407.16088*, 2024.

[7] Kirill Neklyudov, Rob Brekelmans, Daniel Severo, and Alireza Makhzani. Action matching: Learning stochastic dynamics from samples. In *International conference on machine learning*, pages 25858–25889. PMLR, 2023.

[8] Danilo Rezende and Shakir Mohamed. Variational inference with normalizing flows. In *International conference on machine learning*, pages 1530–1538. PMLR, 2015.

[9] Jakob Gulddahl Rasmussen. Lecture notes: Temporal point processes and the conditional intensity function. *arXiv preprint arXiv:1806.00221*, 2018.

[10] Mauricio A Alvarez, Lorenzo Rosasco, Neil D Lawrence, et al. Kernels for vector-valued functions: A review. *Foundations and Trends® in Machine Learning*, 4(3):195–266, 2012.

[11] George O Mohler, Martin B Short, P Jeffrey Brantingham, Frederic Paik Schoenberg, and George E Tita. Self-exciting point process modeling of crime. *Journal of the american statistical association*, 106(493):100–108, 2011.

[12] Fernando Rodriguez Avellaneda, Jorge Mateu, and Paula Moraga. Estimating velocities of infectious disease spread through spatio-temporal log-gaussian cox point processes. *arXiv preprint arXiv:2409.05036*, 2024.

[13] Peter J Diggle, Paula Moraga, Barry Rowlingson, and Benjamin M Taylor. Spatial and spatio-temporal log-gaussian cox processes: extending the geostatistical paradigm. 2013.

[14] Ming Teng, Farouk Nathoo, and Timothy D Johnson. Bayesian computation for log-gaussian cox processes: A comparative analysis of methods. *Journal of statistical computation and simulation*, 87(11):2227–2252, 2017.

[15] Markus Heinonen, Cagatay Yildiz, Henrik Mannerström, Jukka Intosalmi, and Harri Lähdesmäki. Learning unknown ode models with gaussian processes. In *International conference on machine learning*, pages 1959–1968. PMLR, 2018.

[16] Tongyu Li and Fang Yao. Physics-encoded spatio-temporal regression. *arXiv preprint arXiv:2410.04170*, 2024.

[17] Valerii Iakovlev and Harri Lähdesmäki. Learning spatiotemporal dynamical systems from point process observations. *arXiv preprint arXiv:2406.00368*, 2024.

[18] Alexander Tong, Jessie Huang, Guy Wolf, David Van Dijk, and Smita Krishnaswamy. Trajectorynet: A dynamic optimal transport network for modeling cellular dynamics. In *International conference on machine learning*, pages 9526–9536. PMLR, 2020.