SIMULATION-FREE VARIATIONAL INFERENCE FOR LATENT CONTINUOUS-TIME MARKOV CHAINS

Anonymous authors
Paper under double-blind review

ABSTRACT

Latent continuous-time Markov chains (CTMCs) are a widely-used model class in systems biology, epidemiology, and spatial statistics, but learning them from noisy observations requires inference over unobserved continuous-time dynamics and can become prohibitively expensive in large state spaces. We introduce CTMC Matching, a simulation-free variational inference method for latent CTMCs that directly parameterizes the posterior's time marginals. Given a marginal curve, the KL-optimal path measure realizing those marginals has an explicit structure: its transition rates are an exponential tilt of the prior rates by a potential that solves a concave dual problem. We prove that this projection is exact: at the exact smoothing marginals it recovers the smoothing posterior. We develop efficient algorithms that couple this inner optimization with end-to-end learning of both the variational posterior, amortized with a neural network, and the model parameters. For structured CTMCs such as interacting particle systems and chemical reaction networks, we exploit model structure to obtain scalable implementations without enumerating the full joint state space. Across a range of systems, CTMC Matching achieves competitive inference and parameter-learning performance while avoiding latent-trajectory simulation during optimization.

1 INTRODUCTION

The evolution of many systems in science and engineering can be described as discrete events occurring at continuous, irregular times. Continuous-time Markov chains (CTMCs) are a natural model choice for such systems: discrete states encode possible configurations, and their evolution in continuous time is described by transition rates. This has made CTMCs a longstanding tool for statistical inference in stochastic chemical kinetics and systems biology (Gillespie, 1977; Boys et al., 2008; Wilkinson, 2018). This model class also includes interacting CTMCs, which model spatiotemporal propagation phenomena such as epidemics and wildfires (Lanchier, 2024). In many real-world systems, however, only limited noisy observations are available, and the underlying state must be inferred. For example, a patient's disease state may be observed only through irregular noisy measurements (Liu et al., 2015). Latent CTMCs, also known as continuous-time hidden Markov models (CT-HMMs), provide a natural model for this setting.

Inferring the latent states and the transition rates of a latent CTMC becomes harder for larger state spaces and long time horizons: exact forward–backward smoothing requires integrating a system of coupled ODEs whose dimension equals the size of the enumerated state space over the entire time horizon (Liu et al., 2015; Wildner & Koeppl, 2019; Eich et al., 2025). This can become extremely impractical for large datasets and long horizons, and downright impossible when the state space is so large that it cannot be enumerated. This is often the case in interacting CTMCs, where the state space grows exponentially with the number of interacting components Lanchier (2017; 2024). Sampling-based alternatives avoid explicit state enumeration, but must simulate latent event histories sequentially over time and build particle estimates of the posterior distribution over latent paths (Boys et al., 2008; Rao & Teh, 2013; Migliorini & Smyth, 2026). Variational methods replace posterior sampling with optimization, but existing approaches must still integrate the master equation, propagate moments, or sample paths at every gradient step (Opper & Sanguinetti, 2007; Cohn et al., 2010; Wildner & Koeppl, 2019; Seifner & Sánchez, 2023). Consequently, their optimization cost remains coupled to the time horizon and to the complexity of representing the latent process.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

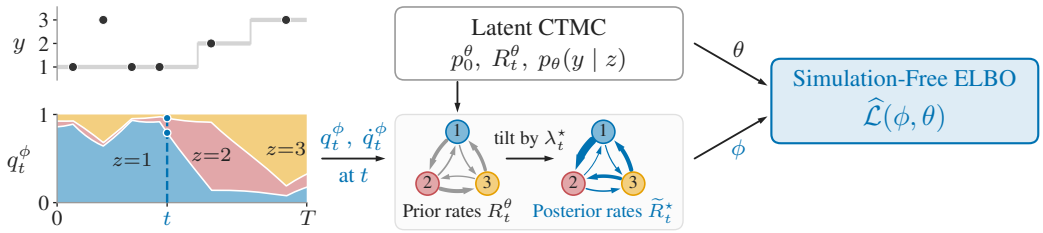


Figure 1: Overview of **CTMC Matching**. *Left:* a latent CTMC (gray) is observed through noisy observations y at irregular times. An amortized inference network represents the variational posterior marginals q_t^ϕ, and it can be queried at any time t without simulation. *Middle:* the prior rates R_t^θ of the latent CTMC are tilted by a potential $\lambda_t^\star$, obtained from a concave dual problem, into the posterior rates $\widetilde{R}_t^\star$ that satisfy the master equation for q_t^ϕ while minimizing the path-space KL divergence from the prior; edge widths denote the transition rates. *Right:* the latent CTMC's parameters (θ) and the posterior marginals and rates (ϕ) enter a simulation-free estimate of the ELBO.

Latent stochastic differential equations face an analogous problem in continuous state spaces, where conventional training similarly relies on numerical simulation and adjoint sensitivities (Li et al., 2020). Recent work has shown that variational inference for latent SDEs can be made *simulation-free* (Course & Nair, 2023a; Bartosh et al., 2025; Smith et al., 2026): rather than parameterizing a variational drift and simulating the resulting process, these methods directly parameterize its one-time marginal distributions. The training objective can then be estimated at randomly sampled times using states drawn directly from the corresponding marginals, removing sequential numerical simulation from the optimization loop. Marginals alone, however, do not determine a unique path law. In particular, Helmholtz-SDEs optimize over the variational posterior processes compatible with a prescribed marginal curve, selecting the one best suited to the variational objective (Smith et al., 2026).

In this work, we develop a discrete-state counterpart of this idea. We introduce **CTMC Matching**, a simulation-free variational inference method for latent CTMCs. We let an amortized inference network directly represent the posterior marginal curve $q_t^\phi(\cdot \mid \mathcal{Y})$. The network can be queried at any given time, and we then choose posterior transition rates that satisfy the master equation for q_t^ϕ while minimizing the instantaneous contribution to the path-space Kullback–Leibler divergence from the prior CTMC. The resulting rates are an exponential tilt of the prior generator by a state potential obtained from a concave dual problem, a structure known from the large-deviation theory of Markov chains (Mielke et al., 2014; Renger, 2018; Patterson & Renger, 2019). We prove that this projection is exact, in that it returns the smoothing posterior at the exact smoothing marginals, and we show that a stop-gradient at the inner solution removes the need to backpropagate through the inner solve. This enables a practical recipe for joint learning of the inference network and the latent CTMC, in a completely simulation-free manner.

Our main contributions are:

- We formulate the recovery of posterior CTMC dynamics from prescribed time marginals as a KL projection onto the generators compatible with the master equation, and we prove that the resulting family contains the smoothing posterior when the marginals are exact (Theorem 3.4).
- We derive a simulation-free stochastic variational objective whose gradients do not require differentiating through the inner optimization problem, enabling joint amortized inference and parameter learning.
- We develop structured implementations for large state spaces, including interacting particle systems and chemical reaction networks, and evaluate the resulting method across kinetic and spatiotemporal models.

The remainder of the paper is organized as follows. Section 2 introduces latent CTMCs, their variational objective, and the simulation-free perspective developed for latent SDEs. Section 3 derives CTMC Matching, its dual formulation and exactness, the simulation-free estimator, and extensions beyond explicit state enumeration. Section 4 evaluates posterior inference and parameter learning for a variety of systems: a discrete flashing ratchet, chemical reaction networks, and spatial wildfire dynamics. The appendix includes a primer on CTMCs, proofs, numerical details, and additional experimental details.

2 BACKGROUND

2.1 LATENT CONTINUOUS-TIME MARKOV CHAINS

Consider a sequence of noisy observations $\mathcal{Y} = \{(t_k, y_k)\}_{k=1}^K$, recorded at potentially irregular observation times. We model the trajectory as being generated from an unobserved process $Z_{[0,T]} = \{Z_t\}_{t\in[0,T]}$ on a discrete state space $\mathcal{Z}$. Specifically, we consider the latent CTMC model

$$Z_{[0,T]} \sim \mathrm{CTMC}(R_t^\theta, p_0^\theta), \qquad Y_{t_k} \mid Z_{t_k} = z \sim p_\theta(\cdot \mid z), \quad k = 1, \dots, K, \tag{1}$$

where θ collectively denotes the generative-model parameters. The latent path is initialized according to p_0^θ and evolves according to a time-dependent rate matrix R_t^θ. For $z' \neq z$, its off-diagonal entries are defined by

$$R_t^\theta(z, z') = \lim_{\Delta t \downarrow 0} \frac{\Pr_\theta(Z_{t+\Delta t} = z' \mid Z_t = z)}{\Delta t}. \tag{2}$$

By convention, $R_t^\theta(z, z) = -\sum_{z' \neq z} R_t^\theta(z, z')$. The initial distribution and rate matrix define a prior path law $\mathbb{P}_\theta$, while $p_\theta(y \mid z)$ denotes the emission distribution for the observations, that are conditionally independent given the latent path. For a primer on CTMCs see Appendix A.

Our goals are to infer the latent trajectory through the *smoothing distribution* $\mathbb{P}_\theta(Z_{[0,T]} \mid \mathcal{Y})$ and to estimate the parameters θ. Exact forward–backward smoothing is viable only for small state spaces because it requires solving the *master equation*, a system of coupled ordinary differential equations with dimension $|\mathcal{Z}|$. This limitation has motivated a variety of approximate-inference methods for latent CTMCs, including variational inference (Opper & Sanguinetti, 2007; Cohn et al., 2010; Zhang et al., 2017; Wildner & Koeppl, 2019; Seifner & Sánchez, 2023), Markov chain Monte Carlo (Boys et al., 2008; Hobolth & Stone, 2009; Rao & Teh, 2013), sequential Monte Carlo (Migliorini & Smyth, 2026), expectation propagation (Eich et al., 2025), and foundation models (Berghaus et al., 2024).

Variational inference. Let ϕ denote the parameters of an approximate posterior CTMC with path law $\mathbb{Q}_\phi$, initial distribution q_0^ϕ, time marginals q_t^ϕ, and transition rates $\widetilde{R}_t$. For notational simplicity, we suppress their dependence on $\mathcal{Y}$ when it is unambiguous. Variational inference fits $\mathbb{Q}_\phi$ by minimizing its reverse Kullback–Leibler divergence from the smoothing posterior. Under standard absolute-continuity and integrability conditions, this is equivalent to minimizing the negative evidence lower bound (nELBO)

$$\begin{aligned}\mathcal{L}(\phi, \theta) = &-\sum_{k=1}^{K} \mathbb{E}_{q_{t_k}^\phi}\left[\log p_\theta(y_k \mid Z_{t_k})\right] + D_{\mathrm{KL}}(q_0^\phi \,||\, p_0^\theta) \\ &+ \int_0^T \mathbb{E}_{q_t^\phi}\left[\sum_{z' \neq Z} \Phi\big(\widetilde{R}_t(Z, z'); R_t^\theta(Z, z')\big)\right] \mathrm{d}t,\end{aligned} \tag{3}$$

where $\Phi(f; a) = f \log(f/a) - f + a$. For a derivation see Appendix A.3; see also Opper & Sanguinetti (2007); Seifner & Sánchez (2023). We denote the final integral in equation 3, measuring the discrepancy between the posterior and prior transition rates, by $\mathcal{L}_{\mathrm{path}}(\phi, \theta)$. Existing variational methods obtain the quantities in equation 3 in different ways: mean-field methods propagate coupled forward–backward equations (Opper & Sanguinetti, 2007; Cohn et al., 2010), moment-based methods propagate selected expectations (Wildner & Koeppl, 2019), and NeuralMJP learns the posterior rates and integrates the master equation to obtain the marginals (Seifner & Sánchez, 2023).

2.2 SIMULATION-FREE VARIATIONAL INFERENCE

Latent SDEs (Archambeau et al., 2007; Li et al., 2020) are continuous-time models with latent paths represented by a process living in a continuous state space, and propagated forward by a deterministic drift component and a random Brownian motion one, modulated by a diffusion matrix (Oksendal, 2013). Inference and parameter-learning methods for latent SDEs faced many of the same computational bottlenecks as methods for latent CTMCs, until Course & Nair (2023b) spearheaded a novel type of simulation-free algorithm for latent SDEs: they proposed to parameterize directly a continuous-time sequence of variational time-marginal distributions $\{q_t^\phi\}_{t\in[0,T]}$, and only then build an SDE whose drift and diffusion are guaranteed to solve the Fokker–Planck equation with respect

to the parameterized marginals. The approach has been further extended to amortized families in Course & Nair (2023a), and generalized to non-Gaussian marginals in Bartosh et al. (2025).

A curve of marginals, however, does not uniquely determine a posterior path law: many drifts may satisfy the same Fokker–Planck equation. Existing simulation-free parameterizations implicitly select only one subset of these compatible dynamics, which can restrict the variational family. Smith et al. (2026) address this limitation with the Helmholtz-SDE, which optimizes the variational objective over path laws compatible with the prescribed marginals. This use of the freedom left unspecified by the marginals is the continuous-state analog of the construction we develop for CTMCs.

3 CTMC MATCHING

Below we introduce our method, which enables simulation-free learning of latent CTMCs. We focus on amortized inference: we aim at fitting an observation-conditioned *inference network* $\ell_\phi(t, \mathcal{Y})$, parameterizing time marginals q_t^ϕ which can be queried at any time $t \in [0, T]$. In the univariate case, we let $q_t^\phi = \mathrm{softmax}(\ell_\phi(t, \mathcal{Y}))$. However, as we have seen in equation 3, learning also requires a choice of posterior rates $\widetilde{R}_t$. Any rate matrix satisfying the *master equation* is a valid candidate:

$$\dot{q}_t^\phi(z) = \sum_{z' \neq z} \left[\widetilde{R}_t(z', z) q_t^\phi(z') - \widetilde{R}_t(z, z') q_t^\phi(z) \right], \qquad z \in \mathcal{Z}. \tag{4}$$

But this problem does not admit a unique solution: the master equation fixes only the net probability flux into each state, but it does not determine how that flux is distributed among individual transitions. Let E be the fixed set of off-diagonal transitions supported by the prior, and let $\mathcal{G}_E$ contain rate matrices with nonnegative off-diagonal entries and zeros outside E. Throughout this section, we work on a finite state space, take E to be strongly connected, and assume $R_t^\theta > 0$ on E and $q_t^\phi > 0$ on $\mathcal{Z}$. Appendix A.1 states the assumptions used throughout this work; proofs are in Appendix B.

Remark 3.1 (Freedom at fixed marginals)**.** *Let $R_t^0 \in \mathcal{G}_E$ satisfy equation 4. Another generator $\widetilde{R}_t \in \mathcal{G}_E$ satisfies the same equation if and only if $(\widetilde{R}_t - R_t^0)^\top q_t^\phi = 0$.*

For a proof, see Appendix B.1. We use this freedom to choose posterior transition rates that minimize the nELBO for fixed (ϕ, θ). This is the discrete-state analog of the freedom exploited by the Helmholtz-SDE framework (Smith et al., 2026).

At a time t and fixed marginals, the only term in the objective depending on the posterior rates is $\mathcal{L}_{\mathrm{path}}(\phi, \theta)$. Therefore, we choose the posterior rates solving the following optimization problem:

$$\mathcal{M}_\theta(q_t^\phi, \dot{q}_t^\phi) = \min_{\widetilde{R} \in \mathcal{G}_E :\, \widetilde{R}^\top q_t^\phi = \dot{q}_t^\phi} \sum_{(z,z') \in E} q_t^\phi(z) \Phi\big(\widetilde{R}(z, z'); R_t^\theta(z, z')\big). \tag{5}$$

The constraint is linear in the rates and, when q_t^ϕ is positive, the objective is strictly convex. Problems like equation 5 have been extensively studied in the literature on large deviations (Mielke et al., 2014; Renger, 2018; Patterson & Renger, 2019). In this work, we demonstrate that the same machinery can be applied to build a practical simulation-free VI method.

The following proposition shows that the optimal rates are obtained by exponentially tilting the prior rates using pairwise differences of a state potential, a form related to posterior-rate parameterizations used by Cohn et al. (2010) and Migliorini & Smyth (2026).

Proposition 3.2 (KL-optimal transition rates; Mielke et al., 2014; Renger, 2018; Patterson & Renger, 2019)**.** *Under the assumptions of Appendix A.1, equation 5 has a unique optimizer on its allowed off-diagonal entries:*

$$\widetilde{R}^\star(z, z') = R_t^\theta(z, z') \exp\big(\lambda_t^\star(z') - \lambda_t^\star(z)\big), \tag{6}$$

where $\lambda_t^\star$ maximizes the concave function

$$\mathcal{D}(\lambda; q_t^\phi, \dot{q}_t^\phi, R_t^\theta) = \sum_z \lambda(z) \dot{q}_t^\phi(z) - \sum_{(z,z') \in E} q_t^\phi(z) R_t^\theta(z, z') \left[e^{\lambda(z') - \lambda(z)} - 1 \right]. \tag{7}$$

The maximum equals $\mathcal{M}_\theta(q_t^\phi, \dot{q}_t^\phi)$, and the maximizing potential is unique up to a constant.

In practice, we remove this additive degree of freedom by choosing an arbitrary reference state z_{ref} and imposing $\lambda_t(z_{\mathrm{ref}}) = 0$ in the dual solve.

Computing the ELBO-optimal rates at each queried time requires solving the dual problem. Differentiating through solver iterations is costly, but the following proposition shows it is unnecessary:

Proposition 3.3. *Under the regularity conditions in Appendix A.1, the ϕ- and θ-gradients of $\mathcal{M}_\theta(q_t^\phi, \dot{q}_t^\phi)$ equal those of $\mathcal{D}(\texttt{stopgrad}(\lambda_t^\star); q_t^\phi, \dot{q}_t^\phi, R_t^\theta)$.*

In the following theorem, we show that CTMC Matching adds no approximation beyond the marginal family: exact marginals recover the full posterior path law and make the variational bound tight. Moreover, under exact marginals the dual potential is precisely the log backward message from forward–backward smoothing. Let $p_t^\star(z) = \mathbb{P}_\theta(Z_t = z \mid \mathcal{Y})$ be the exact smoothing marginals and $\beta_t(z) = p_\theta(\mathcal{Y}_{>t} \mid Z_t = z)$ the likelihood of future observations.

Theorem 3.4 (Exactness of CTMC Matching)**.** *Under the assumptions of Appendix A.1, for almost every t the optimizer of equation 5 at $(p_t^\star, \dot{p}_t^\star)$ is the generator of the exact smoothing posterior,*

$$\widetilde{R}_t^\star(z, z') = R_t^\theta(z, z') \frac{\beta_t(z')}{\beta_t(z)}, \qquad (z, z') \in E, \tag{8}$$

with dual potential $\lambda_t^\star = \log \beta_t$ up to a constant. Consequently, if $q_t^\phi = p_t^\star$ for all t, then $\mathbb{Q}_\phi = \mathbb{P}_\theta(\cdot \mid \mathcal{Y})$ and $\mathcal{L}(\phi, \theta) = -\log p_\theta(\mathcal{Y})$.

3.1 Estimating the Simulation-Free ELBO

Once compatible rates are available at queried times, the nELBO in equation 3 contains only expectations under known one-time marginals. Following simulation-free methods for latent SDEs (Course & Nair, 2023a; Bartosh et al., 2025; Smith et al., 2026), sampling an observation index $k \sim \mathcal{U}(\{1, \dots, K\})$ and a time $t \sim \mathcal{U}([0, T])$ gives the unbiased single-sequence estimator

$$\widehat{\mathcal{L}}(\phi, \theta) = D_{\mathrm{KL}}(q_0^\phi \| p_0^\theta) - K\, \mathbb{E}_{q_{t_k}^\phi} \log p_\theta(y_k \mid Z) + T\, \mathcal{D}(\texttt{stopgrad}(\lambda_t^\star); q_t^\phi, \dot{q}_t^\phi, R_t^\theta). \tag{9}$$

When the state space can be enumerated, the initial KL and reconstruction expectation are evaluated exactly. Otherwise, they may be estimated using discrete state samples, in which case their gradients with respect to ϕ generally require score-function estimators (Williams, 1992) or relaxed reparameterizations (Jang et al., 2017), unless additional structure makes the expectations analytically tractable. Multiple samples or numerical quadrature may also replace the single-sample estimates when lower variance is worth the additional computation.

To amortize inference over a dataset $\mathcal{D} = \{\mathcal{Y}^{(n)}\}_{n=1}^N$, we share the inference-network parameters ϕ across all observation sequences and condition the marginals on each sequence through $q_t^\phi(\cdot \mid \mathcal{Y}^{(n)})$. We optimize the empirical objective $\mathcal{J}(\phi, \theta) = \frac{1}{N} \sum_{n=1}^N \mathcal{L}(\phi, \theta; \mathcal{Y}^{(n)})$ using minibatches and average the corresponding per-sequence estimates from equation 9. Algorithm 1 summarizes the resulting **CTMC Matching** procedure for an enumerable state space. In this setting, the inference network outputs one logit per state, the marginals are obtained by a softmax, and all state-space expectations are evaluated exactly, so no discrete-gradient estimator is needed. The next section develops variants that bypass this enumeration requirement for specific structured state spaces.

3.2 Beyond State Enumeration

Interacting particle systems. We consider CTMCs where the state space has the form $\mathcal{Z} = \mathcal{V}^d$ for a *vocabulary* $\mathcal{V} = \{1, \dots, V\}$, and state transitions affect only a single coordinate at a time (i.e., interacting particle systems (IPSs), see Liggett (1985); Lanchier (2017; 2024); Migliorini & Smyth (2026)). Let $z = (z_1, \dots, z_d) \in \mathcal{V}^d$, and write $z^{i \to v'}$ for the configuration obtained by replacing z_i by v'. A transition at site i from $z_i = v$ to v' has prior local rate

$$r_{i,t}^\theta(v, v' \mid z_{N(i)}),$$

where $N(i) \subseteq \{1, \dots, d\} \setminus \{i\}$ is a spatial neighborhood of site i. Thus, for $z \neq z'$, the global rate matrix decomposes as

$$R_t^\theta(z, z') = \sum_{i=1}^d r_{i,t}^\theta(z_i, z_i' \mid z_{N(i)}) \prod_{j \neq i} \mathbf{1}\{z_j' = z_j\}, \tag{10}$$

with diagonal entries fixed by the row-sum convention.

Algorithm 1 CTMC Matching for enumerable state spaces

Require: Dataset $\mathcal{D} = \{\mathcal{Y}^{(n)}\}_{n=1}^{N}$; inference network ℓ_ϕ; generative model $(p_0^\theta, R_t^\theta, p_\theta)$
Ensure: Learned inference and generative-model parameters (ϕ, θ)

while not converged **do**
  Sample a minibatch $\mathcal{B} \subset \mathcal{D}$
  $\widehat{\mathcal{J}} \leftarrow 0$
  for each $\mathcal{Y}^{(n)} = \{(t_k^{(n)}, y_k^{(n)})\}_{k=1}^{K_n}$ in $\mathcal{B}$ **do**
    Sample $k_n \sim \mathcal{U}(\{1, \dots, K_n\})$ and $t_n \sim \mathcal{U}([0, T_n])$
    $q_s^{\phi,n} \leftarrow \mathrm{softmax}(\ell_\phi(s, \mathcal{Y}^{(n)}))$, for $s \in \{0, t_n, t_{k_n}^{(n)}\}$ ▷ *query marginals directly*
    $\dot{q}_{t_n}^{\phi,n} \leftarrow \mathrm{JVP}_t[q_t^{\phi,n}](t_n; 1)$ ▷ *differentiate in time*
    $\lambda_n^\star \leftarrow \underset{\lambda:\, \lambda(z_{\mathrm{ref}})=0}{\arg\max}\ \mathcal{D}(\lambda; q_{t_n}^{\phi,n}, \dot{q}_{t_n}^{\phi,n}, R_{t_n}^\theta)$ ▷ *concave inner solve*
    $\widehat{\mathcal{L}}_n \leftarrow D_{\mathrm{KL}}(q_0^{\phi,n} \| p_0^\theta) - K_n \mathbb{E}_{\substack{q^{\phi,n} \\ t_{k_n}^{(n)}}} \Big[\log p_\theta(y_{k_n}^{(n)} \mid Z)\Big] + T_n \mathcal{D}(\mathtt{stopgrad}(\lambda_n^\star); q_{t_n}^{\phi,n}, \dot{q}_{t_n}^{\phi,n}, R_{t_n}^\theta)$
    $\widehat{\mathcal{J}} \leftarrow \widehat{\mathcal{J}} + \widehat{\mathcal{L}}_n / |\mathcal{B}|$ ▷ *simulation-free nELBO*
  end for
  $(\phi, \theta) \leftarrow \mathrm{OptimizerStep}\Big((\phi, \theta), \nabla_{\phi,\theta} \widehat{\mathcal{J}}\Big)$
end while
return (ϕ, θ)

Local CTMC Matching. While CTMC Matching has overcome the need to simulate a posterior trajectory, the dual in equation 7 as stated still requires *enumerating* the entire state space. This can be intractable for large state spaces, e.g. IPSs. To overcome this limitation, we leverage posterior marginals with a product form, as in mean-field VI (Opper & Sanguinetti, 2007; Cohn et al., 2010): $q_t^\phi(z) = \prod_{i=1}^d \pi_{i,t}^\phi(z_i)$. The time derivative of $q_t^\phi(z)$ splits into one *contribution* per coordinate:

$$\dot{q}_t^\phi(z) = \sum_{i=1}^{d} q_t^{\phi,-i}(z^{-i})\, \dot{\pi}_{i,t}^\phi(z_i), \qquad q_t^{\phi,-i}(z^{-i}) = \prod_{j \neq i} \pi_{j,t}^\phi(z_j). \tag{11}$$

Here, $q_t^{\phi,-i}$ denotes the marginal over all coordinates except i. For each site i and neighborhood state $c \in \mathcal{V}^{N(i)}$, we solve equation 5 over the V local states:

$$\mathcal{M}_{i,\theta}(\pi_{i,t}, \dot{\pi}_{i,t}; c) = \min_{\substack{\widetilde{r}_{i,t}(c) \in \mathcal{G}_{i,c} \\ \widetilde{r}_{i,t}(c)^\top \pi_{i,t} = \dot{\pi}_{i,t}}} \sum_{v \neq v'} \pi_{i,t}(v) \Phi\big(\widetilde{r}_{i,t}(v, v' \mid c); r_{i,t}^\theta(v, v' \mid c)\big), \tag{12}$$

where $\mathcal{G}_{i,c}$ is the space of local rates preserving the local prior's allowed transitions.

Proposition 3.5. *If every local rate satisfies $\widetilde{r}_{i,t}(c)^\top \pi_{i,t}^\phi = \dot{\pi}_{i,t}^\phi$, the global rates satisfy $\widetilde{R}_t^\top q_t^\phi = \dot{q}_t^\phi$. If the local rates minimize equation 12,*

$$\mathcal{L}_{\mathrm{path}}(\phi, \theta) = \int_0^T \sum_i \mathbb{E}_{c \sim q_{N(i),t}^\phi} \big[\mathcal{M}_{i,\theta}(\pi_{i,t}^\phi, \dot{\pi}_{i,t}^\phi; c)\big] \mathrm{d}t, \quad \textit{where} \quad q_{N(i),t}^\phi(c) = \prod_{j \in N(i)} \pi_{j,t}^\phi(c_j).$$

Appendix B.5 proves the proposition. Appendix C gives the Monte Carlo gradient estimator and a training algorithm for the case where discrete neighborhood contexts require score-function gradients.

4 Experiments

We first validate CTMC Matching on the enumerable discrete flashing ratchet, showing competitive performance and agreement with exact smoothing. We then demonstrate its scalability on chemical reaction networks with large or unbounded state spaces and spatial wildfires with thousands of components, where enumeration is infeasible and simulation-based inference is costly. In high-dimensional settings, our method matches or exceeds the performance of the baselines.

4.1 Discrete Flashing Ratchet

We first consider the discrete flashing ratchet (DFR), a six-state CTMC used as a benchmark by Seifner & Sánchez (2023) and Berghaus et al. (2024). The model describes a particle hopping

Table 1: DFR parameter estimates (mean ± standard deviation over five seeds). ∘ denotes NeuralMJP values reported in Table 8 of Seifner & Sánchez (2023); † denotes published FIM estimates (Berghaus et al., 2024).

Method	V	r	b	RPE (%)
Truth	1.00	1.00	1.00	—
NeuralMJP$^{\circ}$	**0.98** $\pm$ 0.05	1.11 $\pm$ 0.06	1.13 $\pm$ 0.04	8.33
FIM†	1.11	**0.99**	**0.98**	**4.67**
Ours	0.95 $\pm$ 0.02	0.97 $\pm$ 0.02	0.93 $\pm$ 0.01	5.33

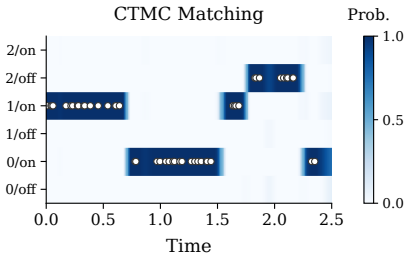


Figure 2: DFR posterior inferred by CTMC Matching for a held-out trajectory. White markers are irregular observations; blue is the inferred posterior probability of each state.

among three sites on a periodic lattice while an asymmetric potential switches on and off. Its joint state is $Z_t = (I_t, S_t) \in \{0, 1, 2\} \times \{\mathrm{ON}, \mathrm{OFF}\}$, where I_t is the particle position and S_t is the potential configuration. When the potential is active, the site energies are $U_i = Vi$, and jumps toward lower-energy sites are favored. When it is inactive, spatial jumps are symmetric with rate b. The potential switches in either direction at rate r without changing the particle position. At unit temperature, the nonzero off-diagonal rates are

$$\begin{aligned} R^{\theta}\big((i,\mathrm{ON}),(j,\mathrm{ON})\big) &= \exp\big[-\tfrac{V}{2}(j-i)\big], \qquad i \neq j, \\ R^{\theta}\big((i,\mathrm{OFF}),(j,\mathrm{OFF})\big) &= b, \qquad i \neq j, \\ R^{\theta}\big((i,s),(i,\bar{s})\big) &= r, \end{aligned} \tag{13}$$

where $\bar{s}$ denotes the opposite potential configuration and $\theta = (V, r, b)$ are the parameters to be learned. We use the irregular-grid setting of Seifner & Sánchez (2023), with both coordinates observed on a grid sampled independently for each trajectory. Although states at observation times are known, the intervening jump times and visited states are latent. We jointly learn the shared parameters θ and infer the posterior dynamics between observations. Table 1 shows competitive parameter recovery relative to the simulation-based NeuralMJP baseline (Seifner & Sánchez, 2023) and the FIM baseline (Berghaus et al., 2024). Figure 2 shows the inferred posterior on a held-out trajectory. The small state space permits exact forward–backward smoothing at fixed parameters, providing a reference for evaluating our approximate posterior. Appendix E.1 compares CTMC Matching with this smoother.

4.2 Chemical Reaction Networks

Stochastic chemical reaction networks (CRNs) are CTMCs whose states record the molecule counts of interacting species (Gillespie, 1977; Wilkinson, 2018). A reaction ρ changes the count vector $z \in \mathbb{N}_0^d$ by a stoichiometric increment $\nu_\rho \in \mathbb{Z}^d$ and occurs with propensity $a_\rho^\theta(z)$. The rates are

$$R^{\theta}(z, z') = \sum_{\rho:\, z+\nu_\rho = z'} a_\rho^{\theta}(z), \qquad z' \neq z. \tag{14}$$

Thus, a small collection of reactions specifies a process on a potentially large count space. We evaluate inference from noisy, partial observations and learning of the reaction-rate parameters. We use the

Table 2: Latent trajectory inference and parameter learning for the benchmark datasets from Eich et al. (2025). Posterior MSE is mean ± SE over the 100 benchmark trajectories, using our reproductions against an exact smoother. EP RPE uses the published estimates in Eich et al. (2025, Table 8).

	Gene		Enzyme		LV	
Method	MSE ↓	RPE (%) ↓	MSE ↓	RPE (%) ↓	MSE ↓	RPE (%) ↓
CTMC Matching	**0.105** $\pm$ 0.008	**0.029**	**0.001** $\pm$ 0.001	**0.330**	**0.618** $\pm$ 0.523	**3.645**
EP	0.184 $\pm$ 0.014	0.370	0.270 $\pm$ 0.032	4.000	0.679 $\pm$ 0.118	34.000
FFBS-ent	2.932 $\pm$ 0.286	—	1.326 $\pm$ 0.136	—	2.473 $\pm$ 0.208	—
G ADS	606.621 $\pm$ 92.174	—	14.802 $\pm$ 6.333	—	1.924 $\pm$ 0.178	—
MBVI	2.297 $\pm$ 0.138	—	0.434 $\pm$ 0.040	—	1.674 $\pm$ 0.094	—

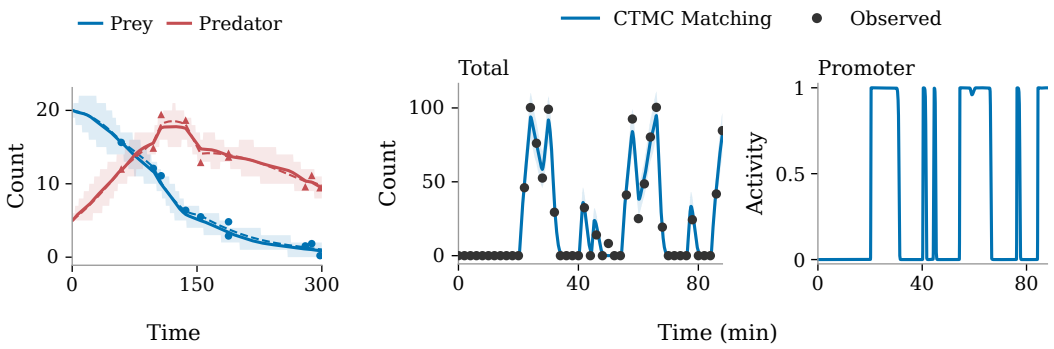


Figure 3: CTMC Matching posteriors for two CRNs. Markers denote noisy count observations, and shading denotes the central 90% credible intervals. *Left:* Lotka–Volterra; solid lines show means, dashed lines the exact truncated smoother. *Right:* reconstructed total transcript count and promoter-on probability for a held-out trajectory from Rullan et al. (2018).

gene-expression, enzyme-kinetics, and Lotka–Volterra benchmarks of Eich et al. (2025), generating 100 trajectories per model with their released code. Each trajectory has ten Gaussian-corrupted observations at irregular times. These benchmarks were used by Eich et al. (2025) to compare their expectation propagation (EP) method against a simplified version of their algorithm denoted FFBS-ent, a Gaussian assumed density smoother (ADS) (Cseke et al., 2016), and moment-based VI (Wildner & Koeppl, 2019). For details on each kinetic model, see Appendix E.2. Following the evaluation framework in Eich et al. (2025), we evaluate each method on two tasks: *latent trajectory inference*, where we fix the rates to their ground-truth values and compare the mean squared error (MSE) of posterior means against an exact smoother; and parameter learning, where we report the relative parameter error (RPE) metric over a selected number of parameters for each experiment. Our method outperforms all baselines on both metrics, see Table 2.

Real transcription data. We next consider the live-cell transcription measurements of Rullan et al. (2018) under low- and high-illumination conditions. We fit the promoter/stage CTMC of Mikelson & Khammash (2020), learning all four kinetic rates with a product of one Bernoulli promoter marginal and eight Poisson count marginals, without truncating the count space. Figure 3 shows representative held-out reconstructions for Lotka–Volterra and live-cell transcription. Appendix E.2.2 gives the model, protocol, and complete results.

4.3 Wildfires as Interacting Particle Systems

Wildfire spread is commonly modeled with stochastic cellular automata whose local transition probabilities depend on neighboring cells and spatial covariates (Clarke et al., 1994; Hargrove et al., 2000; Grieshop & Wikle, 2024). Latent interacting CTMCs provide a continuous-time alternative, and enable incorporating noisy satellite observations separated by several hours. Existing inference methods use sequential Monte Carlo, requiring multiple simulated trajectories during training (Migliorini & Smyth, 2026).

Following Migliorini & Smyth (2026), we use WildFire-SpreadTS (Gerard et al., 2023) and assign each pixel a latent state in $\{\mathsf{U}, \mathsf{A}, \mathsf{B}\}$, denoting unburned, actively burning, and burned out. Binary satellite detections are linked to these states through a Bernoulli emission. The 64×64 and 128×128 settings have 3^{4096} and $3^{16,384}$ global configurations, respectively, so we use local CTMC Matching with factorized variational marginals and neighbor-dependent ignition rates. Appendix E.3 specifies the local rates, support regularization, and evaluation protocols. We evaluate weekly smoothing and free-running forecasting on the same exper-

Table 3: One-day-ahead forecasting on all 3,337 test windows from 2021.

Metric	*Persistence*	UTAE	Ours
AP ↑	0.231	**0.596**	0.571
F1 ↑	0.445	0.502	**0.543**
Precision ↑	0.704	0.365	**0.708**
Recall ↑	0.326	**0.800**	0.440
IoU ↑	0.286	0.335	**0.373**
BCE ↓	0.042	0.015	**0.012**

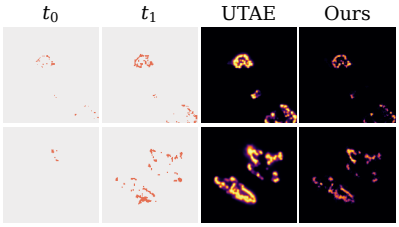


Figure 4: One-day-ahead forecasts on two held-out fires.

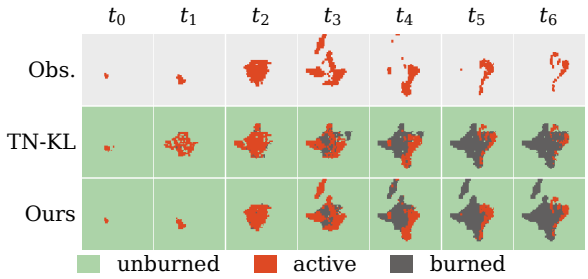


Figure 5: Weekly wildfire smoothing example. While TwistNet–KL misses some details, CTMC Matching perfectly reconstructs the active states over the entire week.

Table 4: Weekly wildfire results: mean per-pixel binary cross-entropy ± 2 SE over 26 held-out sequences. Baselines from Migliorini & Smyth (2026); italicized rows are observation-free predictors.

Method	Reconstruction ↓	Prediction ↓
Persistence	–	2.233 ± 0.608
All unburned	–	1.789 ± 0.526
TwistNet – KL	1.046 ± 0.389	1.424 ± 0.500
TwistNet – DRE	1.702 ± 0.532	1.649 ± 0.533
NAS-X	1.787 ± 0.531	1.576 ± 0.524
NASMC	3.350 ± 0.897	1.992 ± 0.563
Ours	$\mathbf{0.001} \pm 0.000$	$\mathbf{1.217} \pm 0.237$

imental setup as Migliorini & Smyth (2026), and perform an additional experiment where we leverage the scalability of our approach to train a model at larger scale, matching the setup in Gerard et al. (2023) and obtaining performance comparable to state-of-the-art approaches for one-day-ahead forecasting.

4.3.1 Daily forecasting

For the daily task, we train on fires from 2018–2020 and evaluate one-day-ahead predictions on 3,337 windows from 2021. Given five observed days, we roll the learned prior forward from the filtered state and compare against persistence and a UTAE model (Garnot & Landrieu, 2021), the strongest model benchmarked in Gerard et al. (2023), trained directly via binary cross-entropy minimization. Table 3 shows that CTMC Matching improves F1, precision, IoU, and binary cross-entropy, while UTAE attains higher average precision and recall. Figure 4 illustrates the qualitatively sharper, connected forecasts produced by the CTMC prior. For additional experimental results and details, see Appendix E.3.2.

4.3.2 Weekly forecasting

The weekly benchmark of Migliorini & Smyth (2026) contains seven daily 64×64 detection masks per sequence. Reconstruction conditions on the full week, whereas prediction rolls the learned prior forward for six days from the initial detections without further observations. Table 4 shows that CTMC Matching achieves the lowest binary cross-entropy in both settings, including near-zero reconstruction error against the observed masks; Figure 5 shows a representative reconstruction.

5 Conclusion

We have presented, to our knowledge, the first simulation-free variational inference method for latent CTMCs. CTMC Matching replaces sequential latent-trajectory simulation with a direct parameterization of the time marginals, constructs posterior rates from the exponential tilt of a concave dual problem, and is exact at the exact marginals. The resulting ELBO objective can be optimized end to end without simulating latent trajectories or differentiating through the inner solve. Across problems ranging from enumerable state spaces to chemical reaction networks and wildfire systems with thousands of interacting components, our method matches or surpasses existing baselines while eliminating latent-trajectory simulation from the optimization loop. We believe this makes latent CTMCs practical in regimes where existing inference methods are prohibitively costly or simply inapplicable.

Several limitations remain. Each objective evaluation requires an inner dual optimization, which may become costly when the dual is poorly conditioned; amortizing or simplifying this solve is an important direction for future work. Moreover, our treatment of non-enumerable state spaces uses product-form marginal distributions, which may fail to capture posterior dependencies induced by strong interactions. Extending local CTMC Matching to richer marginal families, such as autoregressive models, could address this limitation.

AI USE STATEMENT

In this work, we used generative AI tools for formulating mathematical claims, assisting in the writing of proofs design, and implementing methods. We have not used generative AI tools for helping develop theoretical models or conceptual frameworks, providing critical ingredients for proving mathematical claims, interpreting results, generating synthetic data sets, proposing or refining hypotheses, assisting with translation, cleaning and reformatting datasets, or supporting qualitative and thematic data analysis.

Additionally, we used generative AI tools for creating or modifying scientific figures or images, creating or editing software code, drafting parts of our research paper, summarizing or analyzing existing literature, discovering research topics or identify gaps, sourcing/searching for information, editing our research paper to improve readability, and identifying relevant literature. We have reviewed all AI-assisted work by verifying LLM-generated code, verifying information, and by checking for potential plagiarism for research ideas through manual literature survey. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

ETHICS STATEMENT

The authors do not feel that their paper submission raises any questions regarding the ICLR Code of Ethics beyond the general risks inherent to developing and distributing AI models.

REPRODUCIBILITY STATEMENT

We have included, to the best of our knowledge, all details necessary to replicate our experiments in Section E. We will publish all our code upon publication.

REFERENCES

Cedric Archambeau, Dan Cornford, Manfred Opper, and John Shawe-Taylor. Gaussian process approximations of stochastic differential equations. In *Gaussian Processes in Practice*, pp. 1–16. PMLR, 2007.

Larry Armijo. Minimization of functions having Lipschitz continuous first partial derivatives. *Pacific Journal of mathematics*, 16(1):1–3, 1966.

Grigory Bartosh, Dmitry Vetrov, and Christian A Naesseth. SDE matching: Scalable and simulation-free training of latent stochastic differential equations. In *International Conference on Machine Learning*, pp. 3054–3070. PMLR, 2025.

David Berghaus, Kostadin Cvejoski, Patrick Seifner, César Ojeda, and Ramses J Sanchez. Foundation inference models for Markov jump processes. *Advances in Neural Information Processing Systems*, 37:129407–129442, 2024.

Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

Richard J Boys, Darren J Wilkinson, and Thomas BL Kirkwood. Bayesian inference for a discretely observed stochastic kinetic model. *Statistics and Computing*, 18:125–135, 2008.

Andrew Campbell, Jason Yim, Regina Barzilay, Tom Rainforth, and Tommi Jaakkola. Generative flows on discrete state-spaces: Enabling multimodal flows with applications to protein co-design. In *Forty-first International Conference on Machine Learning*, pp. 5453–5512, 2024.

Raphaël Chetrite and Hugo Touchette. Nonequilibrium markov processes conditioned on large deviations. In *Annales Henri Poincaré*, volume 16, pp. 2005–2057. Springer, 2015.

Keith C Clarke, James A Brass, and Philip J Riggan. A cellular automaton model of wildfire propagation and extinction. *Photogrammetric Engineering and Remote Sensing. 60 (11): 1355-1367*, 60(11):1355–1367, 1994.

Ido Cohn, Tal El-Hay, Nir Friedman, and Raz Kupferman. Mean field variational approximation for continuous-time Bayesian networks. *The Journal of Machine Learning Research*, 11:2745–2783, 2010.

Marc Corstanje and Frank van der Meulen. Guided simulation of conditioned chemical reaction networks. *Statistical Inference for Stochastic Processes*, 28(2):8, 2025.

Kevin Course and Prasanth Nair. Amortized reparametrization: efficient and scalable variational inference for latent SDEs. *Advances in Neural Information Processing Systems*, 36:78296–78318, 2023a.

Kevin Course and Prasanth B Nair. State estimation of a physical system with unknown governing equations. *Nature*, 622(7982):261–267, 2023b.

Botond Cseke, David Schnoerr, Manfred Opper, and Guido Sanguinetti. Expectation propagation for continuous time stochastic processes. *Journal of Physics A: Mathematical and Theoretical*, 49(49): 494002, 2016.

Ron S Dembo, Stanley C Eisenstat, and Trond Steihaug. Inexact Newton methods. *SIAM Journal on Numerical analysis*, 19(2):400–408, 1982.

Yannick Eich, Bastian Alt, and Heinz Koeppl. Entropic matching for expectation propagation of Markov jump processes. In *The 28th International Conference on Artificial Intelligence and Statistics*, pp. 856–864. PMLR, 2025.

Vivien Sainte Fare Garnot and Loic Landrieu. Panoptic segmentation of satellite image time series with convolutional temporal attention networks. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 4872–4881, 2021.

Sebastian Gerard, Yu Zhao, and Josephine Sullivan. WildfirespreadTS: A dataset of multi-modal time series for wildfire spread prediction. *Advances in Neural Information Processing Systems*, 36: 74515–74529, 2023.

Daniel T Gillespie. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, 81(25):2340–2361, 1977.

Nicholas Grieshop and Christopher K Wikle. Data-driven modeling of wildfire spread with stochastic cellular automata and latent spatio-temporal dynamics. *Spatial Statistics*, 59:100794, 2024.

William W Hargrove, RH Gardner, MG Turner, WH Romme, and DG Despain. Simulating fire patterns in heterogeneous landscapes. *Ecological Modelling*, 135(2-3):243–263, 2000.

Asger Hobolth and Eric A Stone. Simulation from endpoint-conditioned, continuous-time Markov chains on a finite state space, with applications to molecular evolution. *The Annals of Applied Statistics*, 3(3):1204, 2009.

Lirong Huang, Loic Pauleve, Christoph Zechner, Michael Unger, Anders S Hansen, and Heinz Koeppl. Reconstructing dynamic molecular states from single-cell time series. *Journal of The Royal Society Interface*, 13(122):20160533, 2016.

Tobias Jahnke and Wilhelm Huisinga. Solving the chemical master equation for monomolecular reaction systems analytically. *Journal of mathematical biology*, 54(1):1–26, 2007.

Eric Jang, Shixiang Gu, and Ben Poole. Categorical reparameterization with Gumbel-softmax. In *International Conference on Learning Representations*, 2017.

John Frank Charles Kingman. *Poisson Processes*, volume 3. Clarendon Press, 1992.

Wouter Kool, Herke van Hoof, and Max Welling. Buy 4 REINFORCE samples, get a baseline for free! In *ICLR 2019 Workshop: Deep RL Meets Structured Prediction*, 2019. URL `https://openreview.net/forum?id=r1lgTGL5DE`.

Nicolas Lanchier. *Stochastic Modeling*. Springer, 2017.

Nicolas Lanchier. *Stochastic Interacting Systems in Life and Social Sciences*, volume 5. Walter de Gruyter GmbH & Co KG, 2024.

Xuechen Li, Ting-Kam Leonard Wong, Ricky TQ Chen, and David Duvenaud. Scalable gradients for stochastic differential equations. In *International Conference on Artificial Intelligence and Statistics*, pp. 3870–3882. PMLR, 2020.

Thomas Milton Liggett. *Interacting Particle Systems*, volume 2. Springer, 1985.

Lek-Heng Lim. Hodge laplacians on graphs. *Siam Review*, 62(3):685–715, 2020.

Yu-Ying Liu, Shuang Li, Fuxin Li, Le Song, and James M Rehg. Efficient learning of continuous-time hidden Markov models for disease progression. *Advances in Neural Information Processing Systems*, 28:3600–3608, 2015.

Alexander Mielke, Mark A Peletier, and DR Michiel Renger. On the relation between gradient flows and the large-deviation principle, with applications to markov chains and diffusion. *Potential Analysis*, 41(4):1293–1327, 2014.

Giosue Migliorini and Padhraic Smyth. Efficient inference for coupled hidden markov models in continuous time and discrete space. In *International Conference on Artificial Intelligence and Statistics*, Proceedings of Machine Learning Research, 2026.

Jan Mikelson and Mustafa Khammash. Likelihood-free nested sampling for parameter inference of biochemical reaction networks. *PLoS computational biology*, 16(10):e1008264, 2020.

Paul Milgrom and Ilya Segal. Envelope theorems for arbitrary choice sets. *Econometrica*, 70(2): 583–601, 2002.

J. Nocedal and S. Wright. *Numerical Optimization*. Springer Series in Operations Research and Financial Engineering. Springer New York, 2006. ISBN 9780387400655. URL `https://books.google.com/books?id=VbHYoSyelFcC`.

James R Norris. *Markov Chains*. Cambridge University Press, 1998.

Bernt Oksendal. *Stochastic Differential Equations: an Introduction with Applications*. Springer Science & Business Media, 2013.

Manfred Opper and Guido Sanguinetti. Variational inference for Markov jump processes. *Advances in Neural Information Processing Systems*, 20:1105–1112, 2007.

Robert IA Patterson and DR Michiel Renger. Large deviations of jump process fluxes. *Mathematical Physics, Analysis and Geometry*, 22(3):21, 2019.

Jason Rader, Terry Lyons, and Patrick Kidger. Optimistix: modular optimisation in jax and equinox. *arXiv:2402.09983*, 2024.

Vinayak Rao and Yee Whye Teh. Fast MCMC sampling for Markov jump processes and extensions. *Journal of Machine Learning Research*, 14:3295–3320, 2013.

D. R. M. Renger. Flux large deviations of independent and reacting particle systems, with implications for macroscopic fluctuation theory. *Journal of Statistical Physics*, 172(5), 09 2018. ISSN ISSN 0022-4715. doi: 10.1007/S10955-018-2083-0.

Marc Rullan, Dirk Benzinger, Gregor W Schmidt, Andreas Milias-Argeitis, and Mustafa Khammash. An optogenetic platform for real-time, single-cell interrogation of stochastic transcriptional regulation. *Molecular cell*, 70(4):745–756, 2018.

Patrick Seifner and Ramsés J Sánchez. Neural Markov jump processes. In *International Conference on Machine Learning*, pp. 30523–30552. PMLR, 2023.

Henry D Smith, Brian L Trippe, and Scott W Linderman. Closing the approximation gap in simulation-free latent SDEs. *arXiv preprint arXiv:2606.16138*, 2026.

Haoran Sun, Lijun Yu, Bo Dai, Dale Schuurmans, and Hanjun Dai. Score-based continuous-time discrete diffusion models. In *The Eleventh International Conference on Learning Representations*, 2023.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.

Christian Wildner and Heinz Koeppl. Moment-based variational inference for Markov jump processes. In *International Conference on Machine Learning*, pp. 6766–6775. PMLR, 2019.

Darren J Wilkinson. *Stochastic Modelling for Systems Biology*. Chapman and Hall/CRC, 2018.

Ronald J Williams. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine Learning*, 8(3):229–256, 1992.

Zeyu Xia and Sibo Cheng. Pytorchfire: A gpu-accelerated wildfire simulator with differentiable cellular automata. *Environmental Modelling & Software*, 188:106401, 2025.

Boqian Zhang, Jiangwei Pan, and Vinayak A Rao. Collapsed variational Bayes for Markov jump processes. *Advances in Neural Information Processing Systems*, 30:3752–3760, 2017.

Maksym Zhenirovskyy, Ion Matei, Rohit Vuppala, Takuya Kurihana, and Hon Yung Wong. Neural-parameterized cellular automata for wildfire spread. *Ecological Informatics*, pp. 103928, 2026.

A A PRIMER ON CONTINUOUS-TIME MARKOV CHAINS

A.1 ASSUMPTIONS

Let $Z_{[0,T]}$ be a CTMC on a finite state space $\mathcal{Z}$, with path law $\mathbb{P}$, marginal probabilities $p_t(z) = \mathbb{P}(Z_t = z)$, and time-dependent rate matrix R_t. Thus, $R_t(z, z') \geq 0$ for $z' \neq z$ and

$$R_t(z, z) = -\sum_{z' \neq z} R_t(z, z'). \tag{15}$$

Basic regularity. All rate functions are measurable and bounded on $[0, T]$, which ensures nonexplosion in the finite-state setting. When using the $O(\Delta t^2)$ short-time expansion below, we additionally assume that $t \mapsto R_t$ is locally Lipschitz. We write p_0 for the initial distribution.

Standing assumptions for CTMC Matching. The prior has a fixed off-diagonal support E: for every time and parameter value under consideration, $R_t^\theta(z, z') > 0$ on E and vanishes outside E. The directed graph $(\mathcal{Z}, E)$ is strongly connected. The prescribed marginal curve $t \mapsto q_t^\phi$ is positive and piecewise continuously differentiable; all statements involving $\dot{q}_t^\phi$ are understood at times where the derivative exists. These conditions guarantee that the master-equation constraint in equation 5 is feasible and that the KL-optimal rates and gauge-fixed dual potential are unique. In implementations, positivity and connectivity may be enforced with a small rate floor.

Path-space KL and exactness. Whenever a path-space KL divergence is used, we assume $q_0^\phi \ll p_0^\theta$ and restrict the variational rates to the prior support, so that $\mathbb{Q}_\phi \ll \mathbb{P}_\theta$; the displayed expectations are also assumed finite. The exactness result in Theorem 3.4 additionally assumes $p_\theta(y_k \mid z) > 0$ for every observation and state. Its conclusions hold almost everywhere in time, so derivatives need not exist at observation times. Proposition 3.3 further requires differentiable dependence of q_t^ϕ, $\dot{q}_t^\phi$, and R_t^θ on their parameters, with the support E held fixed.

A.2 MASTER EQUATION

In this section, we show that the master equation can be intuitively derived from the definition of CTMC given in Section 2.1, under the assumptions in Section A.1. For a more detailed treatment, see Norris (1998); Liggett (1985). As $\Delta t \downarrow 0$, the short-time transition kernel satisfies

$$\mathbb{P}(Z_{t+\Delta t} = z' \mid Z_t = z) = \delta_{z,z'} + \Delta t\, R_t(z, z') + O(\Delta t^2). \tag{16}$$

In particular, the probability of a jump from z to $z' \neq z$ is $\Delta t\, R_t(z, z') + O(\Delta t^2)$, whereas the probability of remaining in z is

$$1 + \Delta t\, R_t(z, z) + O(\Delta t^2) = 1 - \Delta t \sum_{y \neq z} R_t(z, y) + O(\Delta t^2). \tag{17}$$

Using the law of total probability and separating paths that remain in z from those that arrive from another state gives

$$p_{t+\Delta t}(z) = p_t(z) \left[1 - \Delta t \sum_{y \neq z} R_t(z, y)\right] + \Delta t \sum_{y \neq z} p_t(y) R_t(y, z) + O(\Delta t^2).$$

Subtracting $p_t(z)$, dividing by Δt, and taking $\Delta t \downarrow 0$ yields the master equation

$$\dot{p}_t(z) = \sum_{y \neq z} p_t(y) R_t(y, z) - p_t(z) \sum_{y \neq z} R_t(z, y) = \sum_y p_t(y) R_t(y, z). \tag{18}$$

Because every row of a rate matrix sums to zero,

$$\sum_z \dot{p}_t(z) = \sum_y p_t(y) \sum_z R_t(y, z) = 0, \tag{19}$$

so the master equation preserves total probability mass.

A.3 From Path Density to an ELBO

Next, we show that the path density of a CTMC, a crucial quantity for the nELBO objective in equation 3, can also be intuitively derived from first principles.

Survival probability. Conditionally on being in a state z at time s, let $H_z(s,t)$ denote the probability of remaining in z throughout $[s,t]$, i.e. the survival probability of the holding time in state z. If we want to expand the terminal time to $t+\Delta t$, by the Markov property and the probability of no state transitions being given by equation equation 17, we get

$$H_z(s,t+\Delta t) = H_z(s,t)[1 + R_t(z,z)\Delta t + o(\Delta t)].$$

Rearranging and taking a limit gives

$$\partial_t H_z(s,t) = R_t(z,z)H_z(s,t), \qquad H_z(s,s) = 1.$$

Solving this scalar equation gives

$$H_z(s,t) = \exp\left[\int_s^t R_u(z,z)\,\mathrm{d}u\right]. \tag{20}$$

Density of a jump trajectory. We can now use this survival probability to construct the density of a jump trajectory. Let a trajectory ω have ordered jump times

$$0 < s_1 < \cdots < s_N < T, \tag{21}$$

and corresponding states $(z_0,\ldots,z_N)$, where $z_{j-1} \neq z_j$. We set $s_0 = 0$ and $s_{N+1} = T$, so that the process is in state z_j on $[s_j, s_{j+1})$. Suppose the process is in state z_{j-1} at time s_{j-1}. The probability that it remains in z_{j-1} until time s_j and then jumps to z_j during an infinitesimal interval of length $\mathrm{d}s_j$ is

$$H_{z_{j-1}}(s_{j-1}, s_j) R_{s_j}(z_{j-1}, z_j)\,\mathrm{d}s_j. \tag{22}$$

Thus, by the Markov property, the probability associated with a trajectory having the specified state sequence and jump times in infinitesimal neighborhoods of $(s_1,\ldots,s_N)$ is

$$\mathbb{P}\left(Z_{[0,T]} \in \mathrm{d}\omega\right) = p_0(z_0) \prod_{j=1}^{N} \left[H_{z_{j-1}}(s_{j-1}, s_j) R_{s_j}(z_{j-1}, z_j)\,\mathrm{d}s_j\right] H_{z_N}(s_N, T). \tag{23}$$

To turn this expression into a path density, we introduce a reference measure μ as follows. For each N and each admissible state sequence $(z_0,\ldots,z_N)$, we use Lebesgue measure $\mathrm{d}s_1\cdots\mathrm{d}s_N$ on the space of N jump times $\{(s_1,\ldots,s_N) : 0 < s_1 < \cdots < s_N < T\}$, and use counting measure over N and over the discrete state sequences.

With respect to this reference measure, equation 23 gives the path density

$$\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mu}(\omega) = p_0(z_0) \prod_{j=0}^{N} H_{z_j}(s_j, s_{j+1}) \prod_{j=1}^{N} R_{s_j}(z_{j-1}, z_j). \tag{24}$$

Substituting the survival probability from equation 20, we obtain

$$\begin{aligned} \frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mu}(\omega) &= p_0(z_0) \prod_{j=0}^{N} \exp\left[\int_{s_j}^{s_{j+1}} R_t(z_j, z_j)\,\mathrm{d}t\right] \prod_{j=1}^{N} R_{s_j}(z_{j-1}, z_j) \\ &= p_0(z_0) \exp\left[\int_0^T R_t(Z_t, Z_t)\,\mathrm{d}t\right] \prod_{j=1}^{N} R_{s_j}(z_{j-1}, z_j). \end{aligned} \tag{25}$$

Kullback–Leibler divergence between CTMCs. Consider a second CTMC law $\mathbb{Q}$ with initial distribution q_0, marginals q_t, and rates $\widetilde{R}_t$. Using the same reference measure, its log-density ratio with respect to $\mathbb{P}$ is

$$\log \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}(\omega) = \log \frac{q_0(z_0)}{p_0(z_0)} + \sum_{j=1}^{N} \log \frac{\widetilde{R}_{s_j}(z_{j-1}, z_j)}{R_{s_j}(z_{j-1}, z_j)} + \int_0^T \left[\widetilde{R}_t(Z_t, Z_t) - R_t(Z_t, Z_t)\right] \mathrm{d}t. \tag{26}$$

If $q_0 \ll p_0$ and $\widetilde{R}_t(z, z') = 0$ whenever $R_t(z, z') = 0$, then $\mathbb{Q} \ll \mathbb{P}$, and the KL divergence is well defined:

$$D_{\mathrm{KL}}(\mathbb{Q} \,||\, \mathbb{P}) \coloneqq \mathbb{E}_{\mathbb{Q}} \left[\log \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}(\omega) \right] \tag{27}$$

We now turn each path-space expectation into an expectation with respect to the marginals of $\mathbb{Q}$. The initial-state contribution is

$$\mathbb{E}_{\mathbb{Q}} \left[\log \frac{q_0(Z_0)}{p_0(Z_0)} \right] = \sum_z q_0(z) \log \frac{q_0(z)}{p_0(z)} = D_{\mathrm{KL}}(q_0 \,||\, p_0). \tag{28}$$

For the staying contribution, Fubini's theorem and the diagonal convention give

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}} \left[\int_0^T \big(\widetilde{R}_t(Z_t, Z_t) - R_t(Z_t, Z_t) \big) \, \mathrm{d}t \right] &= \int_0^T \mathbb{E}_{q_t} \left[\widetilde{R}_t(Z, Z) - R_t(Z, Z) \right] \mathrm{d}t \\ &= \int_0^T \mathbb{E}_{q_t} \left[\sum_{z' \neq Z} \big(- \widetilde{R}_t(Z, z') + R_t(Z, z') \big) \right] \mathrm{d}t. \end{aligned} \tag{29}$$

Finally, Campbell's theorem (Kingman, 1992) converts the sum over random jump times into a time integral:

$$\mathbb{E}_{\mathbb{Q}} \left[\sum_{j=1}^N \log \frac{\widetilde{R}_{s_j}(z_{j-1}, z_j)}{R_{s_j}(z_{j-1}, z_j)} \right] = \int_0^T \mathbb{E}_{q_t} \left[\sum_{z' \neq Z} \widetilde{R}_t(Z, z') \log \frac{\widetilde{R}_t(Z, z')}{R_t(Z, z')} \right] \mathrm{d}t. \tag{30}$$

Combining these three contributions yields

$$D_{\mathrm{KL}}(\mathbb{Q} \,||\, \mathbb{P}) = D_{\mathrm{KL}}(q_0 \,||\, p_0) + \int_0^T \mathbb{E}_{q_t} \left[\sum_{z' \neq Z} \Phi\big(\widetilde{R}_t(Z, z'); R_t(Z, z') \big) \right] \mathrm{d}t, \tag{31}$$

where $\Phi(f; a) = f \log \frac{f}{a} - f + a$.

The ELBO objective. Finally, we consider the prior path law $\mathbb{P}_\theta$ and the variational law $\mathbb{Q}_\phi$. Conditional independence gives the path likelihood

$$p_\theta(\mathcal{Y} \mid \omega) = \prod_{k=1}^K p_\theta(y_k \mid Z_{t_k}), \tag{32}$$

and by Bayes' rule

$$\log \frac{\mathrm{d}\mathbb{Q}_\phi}{\mathrm{d}\mathbb{P}_\theta(\cdot \mid \mathcal{Y})}(\omega) = \log \frac{\mathrm{d}\mathbb{Q}_\phi}{\mathrm{d}\mathbb{P}_\theta}(\omega) - \sum_{k=1}^K \log p_\theta(y_k \mid Z_{t_k}) + \log p_\theta(\mathcal{Y}). \tag{33}$$

Taking expectations under $\mathbb{Q}_\phi$ gives

$$D_{\mathrm{KL}}(\mathbb{Q}_\phi \,||\, \mathbb{P}_\theta(\cdot \mid \mathcal{Y})) \coloneqq \mathbb{E}_{\mathbb{Q}_\phi} \left[\log \frac{\mathrm{d}\mathbb{Q}_\phi}{\mathrm{d}\mathbb{P}_\theta(\cdot \mid \mathcal{Y})}(\omega) \right] \tag{34}$$

$$= D_{\mathrm{KL}}(\mathbb{Q}_\phi \,||\, \mathbb{P}_\theta) - \sum_{k=1}^K \mathbb{E}_{q^\phi_{t_k}} \left[\log p_\theta(y_k \mid Z_{t_k}) \right] + \log p_\theta(\mathcal{Y}). \tag{35}$$

Substituting equation 31, the first two terms on the right-hand side are exactly $\mathcal{L}(\phi, \theta)$ in equation 3. For fixed θ, the log evidence does not depend on ϕ, so minimizing the nELBO is equivalent to minimizing the reverse KL divergence to the smoothing posterior.

B Proofs for CTMC Matching

This section collects the proofs of all formal results in Section 3. Appendix B.1 characterizes the freedom left by fixed marginals; Appendix B.2 establishes feasibility, primal and dual attainment, the exponential tilt, and uniqueness; Appendix B.3 justifies the stop-gradient estimator; Appendix B.4 proves exact posterior recovery; and Appendix B.5 lifts the construction to factorized interacting systems and decomposes the path cost.

B.1 Freedom at Fixed Marginals

Proof of Remark 3.1. If both rate matrices satisfy the master equation, subtraction gives

$$\widetilde{R}_t^\top q_t^\phi - (R_t^0)^\top q_t^\phi = \dot{q}_t^\phi - \dot{q}_t^\phi = 0. \tag{36}$$

Writing $S_t = \widetilde{R}_t - R_t^0$, this is $S_t^\top q_t^\phi = 0$.

Conversely, suppose $S_t^\top q_t^\phi = 0$. Then

$$\begin{aligned} \widetilde{R}_t^\top q_t^\phi &= (R_t^0 + S_t)^\top q_t^\phi \\ &= (R_t^0)^\top q_t^\phi + S_t^\top q_t^\phi = \dot{q}_t^\phi. \end{aligned} \tag{37}$$

The constructed process and the proposed curve therefore solve the same forward equation from the same initial distribution, and uniqueness implies that their marginals are equal. Notice that S_t itself need not be a generator: its entries may be negative, but the sum $R_t^0 + S_t$ must have nonnegative off-diagonal rates and respect the required transition support. □

B.2 KL-Optimal Transition Rates

Proof of Proposition 3.2. Fix a time t and abbreviate $q = q_t^\phi$, $\dot{q} = \dot{q}_t^\phi$, and $R = R_t^\theta$. The argument follows the convex-duality proofs of Mielke et al. (2014) and Patterson & Renger (2019).

Existence of a strictly positive feasible point. For an allowed transition $(z, z') \in E$, define its probability flux by

$$f(z, z') = q(z)\widetilde{R}(z, z').$$

The master equation can then be rewritten as a divergence of this flux on the transition graph (Lim, 2020):

$$\dot{q}(z) = \sum_{y:(y,z)\in E} f(y, z) - \sum_{y:(z,y)\in E} f(z, y). \tag{38}$$

We first construct a nonnegative flow satisfying equation 38. Fix a reference state z_0. For every $z \neq z_0$ with $\dot{q}(z) > 0$, choose a directed path $P_{z_0 \to z}$ and send $\dot{q}(z)$ units of flow from z_0 to z. For every $z \neq z_0$ with $\dot{q}(z) < 0$, choose a directed path $P_{z \to z_0}$ and send $-\dot{q}(z)$ units from z to z_0. Strong connectivity guarantees that all these paths exist. Explicitly, for an allowed edge $e \in E$, define

$$f^0(e) = \sum_{\substack{z \neq z_0 \\ \dot{q}(z) > 0}} \dot{q}(z)\, \mathbf{1}\{e \in P_{z_0 \to z}\} + \sum_{\substack{z \neq z_0 \\ \dot{q}(z) < 0}} [-\dot{q}(z)]\, \mathbf{1}\{e \in P_{z \to z_0}\}. \tag{39}$$

This flow is nonnegative and along each chosen path the incoming and outgoing contributions cancel at every intermediate state. At a non-reference state z, the path ending or beginning there therefore contributes exactly $\dot{q}(z)$ to its divergence. At the reference state, the divergence is

$$-\sum_{z \neq z_0} \dot{q}(z) = \dot{q}(z_0),$$

because $\sum_z \dot{q}(z) = 0$. Hence f^0 has divergence $\dot{q}$ at every state.

We next modify f^0 to make it positive without changing its divergence. For every allowed edge $e = (z, z')$, strong connectivity provides a directed path from z' back to z. Together with e, this path forms a directed cycle. Sending one unit of flow around a directed cycle has zero divergence, since each visited state receives and emits the same amount. Let c be the sum of one such unit cycle flow for every $e \in E$. Then $c(e) > 0$ on every allowed edge and c has zero divergence.

Consequently, for any $\alpha > 0$ the flow

$$f^+ = f^0 + \alpha c$$

is strictly positive on every allowed edge, while its divergence remains $\dot{q}$. Finally, since $q(z) > 0$, define

$$\widetilde{R}^+(z, z') = \frac{f^+(z, z')}{q(z)}, \qquad (z, z') \in E,$$

and set the diagonal entries by the row-sum convention. Then $\widetilde{R}^+$ is a strictly positive generator on the prescribed support and satisfies

$$(\widetilde{R}^+)^\top q = \dot{q}.$$

Why the minimum is attained and unique. For each allowed transition,

$$\frac{\partial^2}{\partial r^2}\left[q(z)\Phi\big(r; R(z,z')\big)\right] = \frac{q(z)}{r} > 0, \qquad r > 0. \tag{40}$$

With the convention $0 \log 0 = 0$, the cost extends continuously and strictly convexly to $r = 0$, and it tends to infinity when any rate tends to infinity. There are finitely many rates, so every bounded-cost set is bounded. The feasible set is closed and, by the preceding construction, nonempty. A minimum therefore exists and is unique.

Deriving the optimal rates. Attach a multiplier $\lambda(z)$ to the master equation at each state. The resulting Lagrangian is

$$\mathcal{A}(\widetilde{R}, \lambda) = \sum_{(z,z')\in E} q(z)\Phi\big(\widetilde{R}(z,z'); R(z,z')\big) - \lambda^\top\big(\widetilde{R}^\top q - \dot{q}\big). \tag{41}$$

To differentiate it, first expand the multiplier term. Using the diagonal convention,

$$\begin{aligned}\lambda^\top \widetilde{R}^\top q &= \sum_z q(z) \sum_{z'} \widetilde{R}(z,z')\lambda(z') \\ &= \sum_{(z,z')\in E} q(z)\widetilde{R}(z,z')\big[\lambda(z') - \lambda(z)\big].\end{aligned} \tag{42}$$

Substitution into equation 41 gives

$$\begin{aligned}\mathcal{A}(\widetilde{R}, \lambda) &= \lambda^\top \dot{q} \\ &\quad + \sum_{(z,z')\in E} q(z)\left[\Phi\big(\widetilde{R}(z,z'); R(z,z')\big) - \widetilde{R}(z,z')\big(\lambda(z') - \lambda(z)\big)\right].\end{aligned} \tag{43}$$

Using $\partial_r \Phi(r; s) = \log(r/s)$, for an allowed off-diagonal entry we obtain

$$\frac{\partial \mathcal{A}}{\partial \widetilde{R}(z,z')} = q(z)\left[\log\frac{\widetilde{R}(z,z')}{R(z,z')} - \lambda(z') + \lambda(z)\right]. \tag{44}$$

For fixed λ, the corresponding Lagrangian term is strictly convex and has a unique positive minimizer. Setting this derivative to zero gives

$$\widetilde{R}(z,z') = R(z,z')e^{\lambda(z')-\lambda(z)}. \tag{45}$$

At the Lagrange multiplier $\lambda^\star$ associated with $\widetilde{R}^\star$, this proves the claimed form of the optimal rates.

Obtaining the dual. For fixed λ, the Lagrangian separates over the allowed rates and we can hence consider each pair (z, z') separately. Denote by $\Delta = \lambda(z') - \lambda(z)$ the potential difference for that pair. Its minimizing rate is $r = R(z,z')e^{\Delta}$, and

$$\begin{aligned}\Phi\big(r; R(z,z')\big) - r\Delta &= r\log\frac{r}{R(z,z')} - r + R(z,z') - r\Delta \\ &= r\Delta - r + R(z,z') - r\Delta \\ &= R(z,z')\big(1 - e^{\Delta}\big).\end{aligned}$$

Hence minimizing equation 43 over all nonnegative allowed rates gives

$$\begin{aligned}\inf_{\widetilde{R}\in\mathcal{G}_E} \mathcal{A}(\widetilde{R}, \lambda) &= \lambda^\top \dot{q} - \sum_{(z,z')\in E} q(z)R(z,z')\left[e^{\lambda(z')-\lambda(z)} - 1\right] \\ &= \mathcal{D}(\lambda; q, \dot{q}, R).\end{aligned} \tag{46}$$

For any compatible $\widetilde{R}$, the multiplier term in equation 41 vanishes, and therefore every dual value is at most the minimum compatible cost. At $(\widetilde{R}^\star, \lambda^\star)$, stationarity and equation 45 show that $\widetilde{R}^\star$ minimizes the Lagrangian over the rates. Consequently,

$$\mathcal{D}(\lambda^\star; q, \dot{q}, R) = \mathcal{A}(\widetilde{R}^\star, \lambda^\star) = \mathcal{M}_\theta(q, \dot{q}). \tag{47}$$

Thus the dual maximum is attained and equals the primal minimum.

Why only potential differences are determined. Define

$$\widetilde{R}_\lambda(z, z') = R(z, z')e^{\lambda(z')-\lambda(z)} \tag{48}$$

on allowed transitions, with its diagonal defined by the row-sum convention. Direct differentiation gives

$$\begin{aligned}\frac{\partial \mathcal{D}}{\partial \lambda(z)} &= \dot{q}(z) - \sum_{y\neq z} q(y)\widetilde{R}_\lambda(y, z) + q(z)\sum_{y\neq z}\widetilde{R}_\lambda(z, y) \\ &= \dot{q}(z) - (\widetilde{R}_\lambda^\top q)(z).\end{aligned} \tag{49}$$

To study the curvature of the dual, consider an arbitrary direction a in potential space and the one-dimensional perturbation

$$\lambda_\varepsilon = \lambda + \varepsilon a, \qquad g(\varepsilon) = \mathcal{D}(\lambda_\varepsilon; q, \dot{q}, R). \tag{50}$$

The second directional derivative of the dual is the Hessian quadratic form $g''(0) = a^\top \nabla^2_\lambda \mathcal{D}\, a$. Direct differentiation gives

$$-a^\top \nabla^2_\lambda \mathcal{D}\, a = \sum_{(z,z')\in E} q(z)\widetilde{R}_\lambda(z, z')\big[a(z') - a(z)\big]^2 \geq 0. \tag{51}$$

Thus $g''(0) \leq 0$ in every direction, proving that the dual is concave. All weights $q(z)\widetilde{R}_\lambda(z, z')$ in equation 51 are positive, so equality requires $a(z') = a(z)$ along every allowed edge and hence, by strong connectivity, requires a to be constant throughout the graph.

This constant direction is an invariance, as adding a constant c to every potential changes neither the tilted rates, which depend only on $\lambda(z') - \lambda(z)$, nor the linear term in the dual objective, because

$$(\lambda + c\mathbf{1})^\top \dot{q} = \lambda^\top \dot{q} + c\sum_z \dot{q}(z) = \lambda^\top \dot{q}. \tag{52}$$

Accordingly, we may choose any reference state z_{ref} and solve the gauge-fixed problem

$$\max_{\lambda:\, \lambda(z_{\mathrm{ref}})=0} \mathcal{D}(\lambda; q, \dot{q}, R). \tag{53}$$

On this constrained space, the only constant perturbation is $a = 0$, so every nonzero feasible direction has strictly negative second derivative. The dual is therefore strictly concave after gauge fixing, and its maximizer is unique. □

B.3 Proof of Proposition 3.3

Proof. We follow a standard envelope-theorem argument (Milgrom & Segal, 2002): applying the chain rule to the optimal value gives

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\eta}\mathcal{M}_\theta(q_t^\phi, \dot{q}_t^\phi) &= \frac{\mathrm{d}}{\mathrm{d}\eta}\mathcal{D}(\lambda_t^\star(\eta); q_t^\phi, \dot{q}_t^\phi, R_t^\theta) \\ &= \sum_z \frac{\partial \mathcal{D}}{\partial \lambda(z)}\frac{\partial \lambda_t^\star(z)}{\partial \eta} + \left.\frac{\partial \mathcal{D}}{\partial \eta}\right|_{\lambda=\lambda_t^\star},\end{aligned} \tag{54}$$

where we let η denote any scalar component of ϕ or θ. At the exact maximizer, equation 49 gives

$$\frac{\partial \mathcal{D}}{\partial \lambda(z)} = \dot{q}_t^\phi(z) - [(\widetilde{R}_t^\star)^\top q_t^\phi](z) = 0.$$

Therefore, the term involving $\partial\lambda_t^\star/\partial\eta$ vanishes, leaving only the partial derivative of the dual with respect to η at the optimal $\lambda_t^\star$ – which is exactly the derivative obtained by applying `stopgrad` to $\lambda_t^\star$, proving the proposition. □

Note that, for a solver stopped at finite tolerance, the dual gradient in equation 49 is not exactly zero, so the detached-potential gradient approximates the exact gradient to the accuracy of the inner solve.

B.4 Proof of Theorem 3.4

Proof. Write $\mathcal{Y}_{>t} = \{(t_k, y_k) : t_k > t\}$ and $\mathcal{Y}_{\le t} = \mathcal{Y} \setminus \mathcal{Y}_{>t}$, so that $\beta_t(z) = p_\theta(\mathcal{Y}_{>t} \mid Z_t = z)$ and $\beta_t \equiv 1$ after the last observation. We start by stating known results on the backward likelihood (Huang et al., 2016; Eich et al., 2025) and the smoothing posterior, which is Doob's h-transform of the prior with the backward likelihood as space–time harmonic function (Chetrite & Touchette, 2015; Corstanje & van der Meulen, 2025), and then proceed to the part specific to CTMC Matching: feasibility and KL optimality, and how this results in a tight ELBO.

Backward likelihood. Fix t and $h > 0$ such that no observation lies in $(t, t+h]$. By the Markov property,

$$\beta_t(z) = \sum_{z'} \mathbb{P}_\theta(Z_{t+h} = z' \mid Z_t = z)\beta_{t+h}(z').$$

Using the short-time expansion equation 16 and letting $h \downarrow 0$ gives

$$\dot{\beta}_t(z) = -\sum_{z'} R_t^\theta(z, z')\beta_t(z'), \qquad t \notin \{t_1, \dots, t_K\}, \tag{55}$$

where the sum includes the diagonal entry. At an observation time,

$$\lim_{s \uparrow t_k} \beta_s(z) = p_\theta(y_k \mid z)\beta_{t_k}(z). \tag{56}$$

see Huang et al. (2016) for a complete proof. By the assumptions in A.1, $\beta_t(z)$ is positive, piecewise absolutely continuous, and has positive one-sided limits, hence the ratios in equation 8 are well defined and bounded.

The smoothing posterior is a tilted CTMC. Denote the right-hand side of equation 8 by $R_t^{\mathrm{post}}(z, z')$, with diagonal entries fixed by the row-sum convention, and let $\mathbb{P}^{\mathrm{post}}$ be the CTMC with these rates and initial distribution

$$p_0^\star(z) = \frac{p_0^\theta(z)\beta_0(z)}{p_\theta(\mathcal{Y})}. \tag{57}$$

Equation 55 and the diagonal convention give

$$R_t^{\mathrm{post}}(z, z) - R_t^\theta(z, z) = \frac{\mathrm{d}}{\mathrm{d}t} \log \beta_t(z) \tag{58}$$

between observation times. Consider a path ω with jump times $s_1 < \cdots < s_N$ and states $(z_0, \dots, z_N)$, as in Appendix A.3. Integrating equation 58 along the path and accounting for the state jumps and the discontinuities in equation 56 yields

$$\begin{aligned}\int_0^T \left[R_t^{\mathrm{post}}(Z_t, Z_t) - R_t^\theta(Z_t, Z_t)\right] \mathrm{d}t = & -\log \beta_0(z_0) - \sum_{j=1}^{N} \log \frac{\beta_{s_j}(z_j)}{\beta_{s_j}(z_{j-1})} \\ & + \sum_{k=1}^{K} \log p_\theta(y_k \mid Z_{t_k}),\end{aligned} \tag{59}$$

where $\beta_T \equiv 1$. Substituting equation 57 and equation 59 into the path-density ratio equation 26 cancels the backward-likelihood terms and gives

$$\frac{\mathrm{d}\mathbb{P}^{\mathrm{post}}}{\mathrm{d}\mathbb{P}_\theta}(\omega) = \frac{\prod_{k=1}^K p_\theta(y_k \mid Z_{t_k})}{p_\theta(\mathcal{Y})} = \frac{\mathrm{d}\mathbb{P}_\theta(\cdot \mid \mathcal{Y})}{\mathrm{d}\mathbb{P}_\theta}(\omega), \tag{60}$$

where the last equality follows from Bayes' rule and equation 32. Therefore $\mathbb{P}^{\mathrm{post}} = \mathbb{P}_\theta(\cdot \mid \mathcal{Y})$, and its marginals are $p_t^\star$.

Feasibility and KL optimality. For almost every $t > 0$, strong connectivity and positive emissions imply $p_t^\star(z) > 0$ for every state. Since $\mathbb{P}^{\mathrm{post}}$ has marginals $p_t^\star$, its master equation gives

$$\left(R_t^{\mathrm{post}}\right)^\top p_t^\star = \dot{p}_t^\star. \tag{61}$$

Thus R_t^{post} is feasible for equation 5. It is also the exponential tilt equation 6 with $\lambda_t = \log \beta_t$. Indeed, combining equation 49 and equation 61 gives

$$\nabla_\lambda \mathcal{D}\big(\log \beta_t; p_t^\star, \dot{p}_t^\star, R_t^\theta\big) = 0.$$

Concavity of the dual therefore makes $\log \beta_t$ a maximizer, and strict concavity after gauge fixing makes it unique up to a constant. Proposition 3.2 then gives $\widetilde{R}_t^\star = R_t^{\mathrm{post}}$ and

$$\mathcal{M}_\theta(p_t^\star, \dot{p}_t^\star) = \sum_{(z,z')\in E} p_t^\star(z)\Phi\big(R_t^{\mathrm{post}}(z,z'); R_t^\theta(z,z')\big). \tag{62}$$

Path-space KL and tightness. Applying equation 31 to the posterior CTMC and using equation 62 yields

$$D_{\mathrm{KL}}(\mathbb{P}_\theta(\cdot \mid \mathcal{Y}) \| \mathbb{P}_\theta) = D_{\mathrm{KL}}(p_0^\star \| p_0^\theta) + \int_0^T \mathcal{M}_\theta(p_t^\star, \dot{p}_t^\star)\, \mathrm{d}t. \tag{63}$$

Finally, if $q_t^\phi = p_t^\star$ for all t, CTMC Matching uses the same initial distribution and rates as the smoothing posterior. Hence $\mathbb{Q}_\phi = \mathbb{P}_\theta(\cdot \mid \mathcal{Y})$, and equation 35 gives $\mathcal{L}(\phi, \theta) = -\log p_\theta(\mathcal{Y})$. □

B.5 Proof of Proposition 3.5

Proof. Suppress time to simplify the notation, and assemble $\widetilde{R}$ from the posterior local rates $\widetilde{r}_i$ as in equation 10.

For a fixed joint configuration z, an incoming transition that changes site i must start at $z^{i\to v}$ for some $v \neq z_i$. The probability at that origin is then

$$q(z^{i\to v}) = \pi_i(v) \prod_{j\neq i} \pi_j(z_j) = \pi_i(v) q^{-i}(z^{-i}).$$

The outgoing probability has the same factorization, $q(z) = \pi_i(z_i) q^{-i}(z^{-i})$. The site-$i$ contribution to the global derivative is therefore

$$\begin{aligned}
&\sum_{v\neq z_i} q(z^{i\to v}) \widetilde{r}_i(v, z_i \mid z_{N(i)}) - q(z) \sum_{v'\neq z_i} \widetilde{r}_i(z_i, v' \mid z_{N(i)}) \\
&\quad = q^{-i}(z^{-i}) \left[\sum_{v\neq z_i} \pi_i(v) \widetilde{r}_i(v, z_i \mid z_{N(i)}) - \pi_i(z_i) \sum_{v'\neq z_i} \widetilde{r}_i(z_i, v' \mid z_{N(i)}) \right] \\
&\quad = q^{-i}(z^{-i}) [\widetilde{R}_i(z_{N(i)})^\top \pi_i](z_i) = q^{-i}(z^{-i}) \dot{\pi}_i(z_i).
\end{aligned} \tag{64}$$

Each local equation contributes exactly its term in the derivative of the product, and summing over all sites gives

$$(\widetilde{R}^\top q)(z) = \sum_i q^{-i}(z^{-i}) \dot{\pi}_i(z_i) = \frac{\mathrm{d}}{\mathrm{d}t} \prod_i \pi_i(z_i) = \dot{q}(z).$$

Proof of the cost decomposition. Every global transition changes exactly one site. The instantaneous path cost is thus

$$C_{\theta,\phi}(t) = \sum_i \sum_z q_t(z) \sum_{v'\neq z_i} \Phi\big(\widetilde{r}_{i,t}(z_i, v' \mid z_{N(i)}); r_{i,t}^\theta(z_i, v' \mid z_{N(i)})\big).$$

For a fixed site i, we use the fact that the rate term depends only on z_i and $z_{N(i)}$. Summing over all other coordinates removes their probabilities, since

$$\prod_{j\notin\{i\}\cup N(i)} \left[\sum_{z_j\in\mathcal{V}} \pi_{j,t}(z_j) \right] = 1.$$

The site contribution therefore reduces to

$$\sum_{c\in\mathcal{V}^{N(i)}} q_{N(i),t}^\phi(c) \sum_v \pi_{i,t}(v) \sum_{v'\neq v} \Phi\big(\widetilde{r}_{i,t}(v, v' \mid c); r_{i,t}^\theta(v, v' \mid c)\big).$$

For the optimal local rates, the inner two sums equal $\mathcal{M}_{i,\theta}(\pi_{i,t}, \dot{\pi}_{i,t}; c)$. Thus the site contribution is

$$\mathbb{E}_{c\sim q_{N(i),t}^\phi} \left[\mathcal{M}_{i,\theta}(\pi_{i,t}, \dot{\pi}_{i,t}; c) \right].$$

Summing over sites and integrating over time proves equation 3.5. □

C Local CTMC Matching: Details and Algorithm

Monte Carlo and envelope gradients. The expectation in equation 3.5 can be estimated by sampling a neighborhood context for each site, or by drawing one joint state $Z_t \sim q_t^\phi$ and setting $c_i = Z_{N(i),t}$ for every site simultaneously. The envelope argument from Proposition 3.3 applies independently to each local problem: conditional on a fixed context c,

$$\nabla_\eta \mathcal{M}_{i,\theta}(\pi_i, \dot{\pi}_i; c) = \nabla_\eta \mathcal{D}_{i,c}\big(\texttt{stopgrad}(\lambda^\star_{i,c}); \pi_i, \dot{\pi}_i, R_i^\theta(c)\big), \qquad \eta \in \{\phi, \theta\}. \tag{65}$$

Thus the local potential is detached, while gradients of the nELBO still pass through the explicit dependence of the dual on the site marginals, their time derivatives, and the prior rates.

Sampling the context introduces an additional dependence on ϕ through $q^\phi_{N(i),t}$. Define

$$\mathcal{D}^\star_{i,c}(t) \coloneqq \mathcal{D}_{i,c}\Big(\texttt{stopgrad}(\lambda^\star_{i,c}); \pi^\phi_{i,t}, \dot{\pi}^\phi_{i,t}, R_i^\theta(c)\Big). \tag{66}$$

The path-term gradient is then

$$\nabla_\phi \mathcal{L}_{\text{path}}(\phi, \theta) = \int_0^T \sum_i \mathbb{E}_{c \sim q^\phi_{N(i),t}} \Big[\nabla_\phi \mathcal{D}^\star_{i,c}(t) + \mathcal{D}^\star_{i,c}(t) \nabla_\phi \log q^\phi_{N(i),t}(c)\Big] \, \mathrm{d}t. \tag{67}$$

The first term differentiates the local dual at a fixed sampled context; the second is the score-function correction for sampling that context.

More generally, for $Z \sim q_\phi$ and a possibly parameter-dependent integrand $f_\phi(Z)$, we can define the score-gradient surrogate objective

$$\mathsf{SF}(f_\phi, q_\phi, Z) \coloneqq f_\phi(Z) + \texttt{stopgrad}\big(f_\phi(Z)\big) \log q_\phi(Z), \tag{68}$$

which satisfies

$$\mathbb{E}_{Z \sim q_\phi}\left[\nabla_\phi \mathsf{SF}(f_\phi, q_\phi, Z)\right] = \nabla_\phi \mathbb{E}_{Z \sim q_\phi}[f_\phi(Z)]. \tag{69}$$

Standard baselines may be subtracted inside the stopped score term to reduce variance (Williams, 1992), such as leave-one-out (Kool et al., 2019). Relaxed categorical reparameterizations provide an alternative (Jang et al., 2017), at the cost of biasing the gradient estimates.

In Algorithm 2, we present this construction when applied to all three nELBO contributions, for product-form marginals and one joint sample at the initial, observation, and path times. In practice, each term may be evaluated exactly whenever their structure permits.

D Solving the Inner Optimization Problem

Each evaluation of the simulation-free objective requires solving the concave dual in equation equation 7. For the remainder of this section, we fix a time $t \in [0, T]$ and abbreviate $q = q_t^\phi$, $\dot{q} = \dot{q}_t^\phi$, and $R = R_t^\theta$. We write

$$\widetilde{R}_\lambda(z, z') = R(z, z') \exp\big(\lambda(z') - \lambda(z)\big), \qquad g(\lambda) \coloneqq \nabla_\lambda \mathcal{D}(\lambda; q, \dot{q}, R) = \dot{q} - \widetilde{R}_\lambda^\top q. \tag{70}$$

Note that the dual gradient is exactly the master-equation residual, and vanishes when the tilted rates realize $(q, \dot{q})$. Since the rates are invariant to additive constants in λ, we fix one reference potential, $\lambda(z_{\text{ref}}) = 0$, throughout. Below, $n = |\mathcal{Z}|$ and $m = |E|$ denote the numbers of states and supported directed transitions in the matching problem. In the local construction of Section 3.2, n is the number of states of a single component rather than the size of the joint state space.

Note that, in the batched training strategies of Algorithms 1 and 2, the optimization problems generated by different trajectories, sampled times, sites, and sampled neighborhood configurations are independent. All of the solvers below therefore can be batched and parallelized at every gradient step.

Algorithm 2 Local CTMC Matching with score-function gradients

Require: Dataset $\mathcal{D}$; sitewise inference logits $\ell_{\phi,i}$; generative model $(p_0^\theta, r_{i,t}^\theta, p_\theta)$
Ensure: Learned inference and generative-model parameters (ϕ, θ)

while not converged **do**
  Sample a minibatch $\mathcal{B} \subset \mathcal{D}$
  $\widehat{\mathcal{J}}_{\mathrm{sur}} \leftarrow 0$
  for each $\mathcal{Y}^{(n)} = \{(t_k^{(n)}, y_k^{(n)})\}_{k=1}^{K_n}$ in $\mathcal{B}$ **do**
    Sample $k_n \sim \mathcal{U}(\{1, \dots, K_n\})$ and $t_n \sim \mathcal{U}([0, T_n])$
    $\pi_{i,s}^{\phi,n} \leftarrow \mathrm{softmax}(\ell_{\phi,i}(s, \mathcal{Y}^{(n)}))$ for all i and $s \in \{0, t_n, t_{k_n}^{(n)}\}$
    $q_s^{\phi,n}(z) \leftarrow \prod_i \pi_{i,s}^{\phi,n}(z_i)$ for $s \in \{0, t_n, t_{k_n}^{(n)}\}$
    $\dot{\pi}_{i,t_n}^{\phi,n} \leftarrow \mathrm{JVP}_t[\pi_{i,t}^{\phi,n}](t_n; 1)$ for all i
    Sample independently $z_0 \sim q_0^{\phi,n}$, $z_k \sim q_{t_{k_n}^{(n)}}^{\phi,n}$, and $z_t \sim q_{t_n}^{\phi,n}$
    $f_{0,n} \leftarrow \log q_0^{\phi,n}(z_0) - \log p_0^\theta(z_0)$
    $f_{y,n} \leftarrow -K_n \log p_\theta(y_{k_n}^{(n)} \mid z_k)$
    $\widehat{\mathcal{S}}_n \leftarrow \mathsf{SF}(f_{0,n}, q_0^{\phi,n}, z_0) + \mathsf{SF}(f_{y,n}, q_{t_{k_n}^{(n)}}^{\phi,n}, z_k)$
    for $i = 1, \dots, d$ **do**
      $c_{i,n} \leftarrow (z_t)_{N(i)}$
      $\lambda_{i,n}^\star \leftarrow \underset{\lambda:\, \lambda(v_{\mathrm{ref}})=0}{\arg\max}\ \mathcal{D}_{i,c_{i,n}}(\lambda; \pi_{i,t_n}^{\phi,n}, \dot{\pi}_{i,t_n}^{\phi,n}, R_i^\theta(c_{i,n}))$
      $d_{i,n} \leftarrow \mathcal{D}_{i,c_{i,n}}(\mathtt{stopgrad}(\lambda_{i,n}^\star); \pi_{i,t_n}^{\phi,n}, \dot{\pi}_{i,t_n}^{\phi,n}, R_i^\theta(c_{i,n}))$
      $\widehat{\mathcal{S}}_n \leftarrow \widehat{\mathcal{S}}_n + T_n \mathsf{SF}(d_{i,n}, q_{N(i),t_n}^{\phi,n}, c_{i,n})$
    end for
    $\widehat{\mathcal{J}}_{\mathrm{sur}} \leftarrow \widehat{\mathcal{J}}_{\mathrm{sur}} + \widehat{\mathcal{S}}_n / |\mathcal{B}|$
  end for
  $(\phi, \theta) \leftarrow \mathrm{OptimizerStep}((\phi, \theta), \nabla_{\phi,\theta} \widehat{\mathcal{J}}_{\mathrm{sur}})$
end while
return (ϕ, θ)

D.1 Binary Closed Form

For starters, we show a simple case where the optimum is available in closed form. For $q = (1-p, p)$, let $\alpha = R(0,1)$ and $\beta = R(1,0)$. Writing $\dot{p}$ for the derivative of the marginal of state 1, the matched rates satisfy

$$(1-p)\widetilde{\alpha} - p\widetilde{\beta} = \dot{p}, \qquad \widetilde{\alpha}\widetilde{\beta} = \alpha\beta,$$

where the second identity follows from $\widetilde{\alpha} = \alpha e^{\lambda(1)-\lambda(0)}$ and $\widetilde{\beta} = \beta e^{\lambda(0)-\lambda(1)}$. Solving the resulting quadratic gives

$$\widetilde{\alpha}^\star = \frac{d + \dot{p}}{2(1-p)}, \qquad \widetilde{\beta}^\star = \frac{d - \dot{p}}{2p}, \qquad d = \sqrt{\dot{p}^2 + 4p(1-p)\alpha\beta}. \tag{71}$$

The potential difference can then be recovered as $\lambda^\star(1) - \lambda^\star(0) = \log(\widetilde{\alpha}^\star/\alpha)$.

D.2 Damped Newton and Newton–CG

For larger state spaces, we resort to numerical methods for optimization. We use Newton's method (Nocedal & Wright, 2006; Boyd & Vandenberghe, 2004) to jointly update all free potentials. As shown in equation equation 51, the negative dual Hessian is the weighted graph Laplacian

$$H(\lambda) := -\nabla_\lambda^2 \mathcal{D}(\lambda) = \sum_{(z,z') \in E} q(z) \widetilde{R}_\lambda(z, z')(e_z - e_{z'})(e_z - e_{z'})^\top, \tag{72}$$

where e_z is the coordinate vector associated with state z. Let $U = \mathcal{Z} \setminus \{z_{\mathrm{ref}}\}$ index the free coordinates. Under the assumptions from A.1, H_{UU} is positive definite, and the Newton direction solves

$$H_{UU} s_U = g_U, \qquad s(z_{\mathrm{ref}}) = 0. \tag{73}$$

We can use backtracking (Armijo, 1966): starting from $a = 1$, a is reduced until

$$\mathcal{D}(\lambda + as) \geq \mathcal{D}(\lambda) + c_{\mathrm{ls}} a \, g_U^\top s_U, \qquad 0 < c_{\mathrm{ls}} < \tfrac{1}{2}, \tag{74}$$

followed by $\lambda \leftarrow \lambda + as$. For small state spaces, equation 73 can be solved by a dense Cholesky factorization, costing $O(n^3)$ time and $O(n^2)$ storage per nonlinear iteration.

For larger sparse graphs, Newton–CG (Dembo et al., 1982) avoids forming or factoring the Hessian. Its matrix–vector product is

$$Hv = \sum_{(z,z') \in E} q(z) \widetilde{R}_\lambda(z, z') \big(v(z) - v(z')\big)(e_z - e_{z'}), \tag{75}$$

which costs $O(m)$. Conjugate gradients is applied to the reduced system in equation 73, and the resulting direction uses the same line search. With J conjugate-gradient iterations, the direction costs $O(Jm)$ time and $O(n + m)$ storage.

D.3 Coordinate Ascent

An alternative is to optimize one potential exactly while holding the others fixed. For the current tilted rates, define the incoming and outgoing marginal fluxes at state z by

$$I_z = \sum_{y \neq z} q(y) \widetilde{R}_\lambda(y, z), \qquad O_z = q(z) \sum_{y \neq z} \widetilde{R}_\lambda(z, y). \tag{76}$$

Changing only a single state's $\lambda(z)$ by adding δ multiplies the incoming rates by e^δ and the outgoing rates by $e^{-\delta}$. Up to terms independent of δ, the dual becomes

$$\mathcal{D}(\lambda + \delta e_z) = \delta \dot{q}(z) - I_z e^\delta - O_z e^{-\delta} + \text{constant}.$$

This strictly concave scalar problem has the unique maximizer

$$\delta^\star = \operatorname{asinh}\left(\frac{\dot{q}(z)}{2\sqrt{I_z O_z}} \right) + \frac{1}{2} \log \frac{O_z}{I_z}. \tag{77}$$

Cycling through the non-reference states and applying equation 77 gives an exact coordinate-ascent solver. Each update increases the dual without a line search, although a later update at an adjacent state can disturb an equation balanced earlier in the sweep. With adjacency lists and cached rates, one sweep costs $O(m)$ time and $O(n + m)$ storage.

D.4 Solver Comparison and Acceptance Checks

Table 5 summarizes the cost of a single iteration for each method. However, note that the total cost also depends on the number of iterations required for convergence. A key choice in this algorithm is *when* to stop the inner solve. In our experiments, we use the following approach.

For each supported transition $(z, z') \in E$, let the prior and candidate posterior fluxes be

$$A(z, z') = q(z) R(z, z'), \qquad F_\lambda(z, z') = A(z, z') e^{\lambda(z') - \lambda(z)}. \tag{78}$$

The first check compares the primal rate cost with the dual value. Writing

$$\begin{aligned} C(F_\lambda) &= \sum_{(z,z') \in E} \Phi\big(F_\lambda(z, z'); A(z, z')\big), \\ D(\lambda) &= \lambda^\top \dot{q} - \sum_{(z,z') \in E} A(z, z') \left[e^{\lambda(z') - \lambda(z)} - 1 \right], \end{aligned} \tag{79}$$

Method	Unit of work	Arithmetic cost	Storage
Binary formula	Complete solve	$O(1)$	$O(1)$
Dense Newton	Newton iteration	$O(n^3 + (B+1)m)$	$O(n^2)$
Newton–CG	Newton iteration	$O((J+B+1)m)$	$O(n+m)$
Coordinate ascent	Coordinate sweep	$O(m)$	$O(n+m)$

Table 5: Costs for n states and m supported directed transitions. B is the number of line-search evaluations, and J the number of conjugate-gradient iterations.

where $\Phi(f;a) = f\log(f/a) - f + a$, we require

$$\Delta_C := |C(F_\lambda) - D(\lambda)| \le \varepsilon \max\{1, |D(\lambda)|\}\,. \tag{80}$$

The homogeneity of Φ gives $\Phi(F_\lambda(z,z'); A(z,z')) = q(z)\Phi(\widetilde{R}_\lambda(z,z'); R(z,z'))$, so $C(F_\lambda)$ is exactly the primal rate cost. Moreover, if the candidate flux satisfies the master equation, then $\lambda^\top \dot{q} = \sum_{(z,z')\in E} F_\lambda(z,z')[\lambda(z') - \lambda(z)]$, which implies $C(F_\lambda) = D(\lambda)$ at an exact solution.

The second check directly tests compatibility with the prescribed derivative. Let

$$F_{\mathrm{in}}(z) = \sum_{y\neq z} F_\lambda(y,z), \qquad F_{\mathrm{out}}(z) = \sum_{y\neq z} F_\lambda(z,y).$$

We accept the master-equation residual when

$$r_{\mathrm{ME}} := \sum_z q(z)\frac{|\dot{q}(z) - F_{\mathrm{in}}(z) + F_{\mathrm{out}}(z)|}{q(z) + F_{\mathrm{in}}(z) + F_{\mathrm{out}}(z)} \le \varepsilon. \tag{81}$$

We accept a solution only when both statistics are finite and both inequalities hold.

Both checks are necessary: at $\lambda = 0$, the candidate and prior fluxes coincide, so $C(F_\lambda) = D(\lambda) = 0$ even if the prior flux violates the requested master equation. Conversely, the second check verifies the constraint but not optimality.

E Experimental Details

E.1 Discrete Flashing Ratchet

Data and likelihood. We generate 5,000 trajectories using the irregular-grid protocol of Seifner & Sánchez (2023), with $(V, r, b) = (1, 1, 1)$ and horizon $T = 2.5$. Initial states are drawn from the stationary distribution of this generator. For each trajectory, 50 uniform observation times are sampled and we discard the final one, matching the data loader from Seifner & Sánchez (2023) [1] and leaving 49 observations. A fixed split assigns 4,480 trajectories to training and 520 to held-out evaluation.

Both state coordinates are observed without added noise. To keep the expected log likelihood finite for a nondegenerate categorical posterior, we use

$$p(y \mid z) = \begin{cases} 0.999, & y = z, \\ 0.001/5, & y \neq z. \end{cases}$$

The prior initial distribution is held fixed at the data-generating stationary law as (V, r, b) are learned.

Recognition model and training. A shared bidirectional transformer (Vaswani et al., 2017) maps the complete observation sequence to a categorical marginal at any query time. For an observation $y_k = (a_k, b_k)$, with position $a_k \in \{0, 1, 2\}$ and potential state $b_k \in \{0, 1\}$, we use sinusoidal embeddings to encode observations and normalized time t/T as follows:

$$e(y_k) = \left(\cos\frac{2\pi a_k}{3}, \sin\frac{2\pi a_k}{3}, E_{b_k}\right) \in \mathbb{R}^{18}.$$

$$\tau(t) = \left(\frac{t}{T}, \sin\frac{2\pi t}{T}, \cos\frac{2\pi t}{T}, \sin\frac{4\pi t}{T}, \cos\frac{4\pi t}{T}\right).$$

Here $E_{b_k} \in \mathbb{R}^{16}$ is learned. We prepend a learned boundary token at time zero, concatenate each token's state embedding, time features, and boundary flag, and project the resulting 24 features to width 256. Four pre-normalized self-attention blocks encode the 50 tokens. Each block has eight heads and a $256 \to 512 \to 256$ GELU feed-forward network. Attention uses rotary embeddings evaluated at the continuous observation times, a learned two-sided relative-time bias that distinguishes past and future observations, and one learned register key/value. Encoder dropout is 0.1 during training.

[1] https://github.com/pseifner/NeuralMJP

Solver	Trained checkpoint			Untrained checkpoint		
	Accepted	Steps (med./max.)	ms	Accepted	Steps (med./max.)	ms
Capped Newton (ours)	100%	7/12	2.1	100%	5/11	1.95
`optimistix.Newton`	97.3% (7 NaNs)	8/100	9.0	97.7% (6 NaNs)	6/72	6.4
`optimistix.LevenbergMarquardt`	99.6%†	63/100	13.3	100%	19/43	5.7
`optimistix.Dogleg`	100%	10/40	5.3	100%	7/23	3.2

Table 6: DFR inner-solver comparison on 256-cell batches. Runtime is in milliseconds per batch and excludes the outer training step. †Although accepted, the largest dual-value error was 3.4×10^{-2}.

At a query time t, a $5 \to 512 \to 256$ GELU MLP embeds $\tau(t)$. Two cross-attention blocks use this single query token to attend to the encoded observations. Concatenating the result with $\tau(t)$ gives a 261-dimensional vector, which a $261 \to 512 \to 6$ GELU decoder maps to direct state logits $h_\phi(t, \mathcal{Y})$.

We noticed that the logits had a hard time representing the "kinks" that the smoother naturally has at each observation time (Huang et al., 2016). Therefore, to represent the sharp posterior at each nearly exact observation, we add a likelihood readout "bias" to the logits

$$\rho_t(x) = \gamma \sum_{k=0}^{K} \bar{w}_k(t) r_k(x), \qquad \bar{w}_k(t) \propto e^{-\nu|t-t_k|},$$

where γ and ν are learned, $r_0 = 0$, and each r_k is the centered version of the normalized observation log-likelihood row, clipped below at -8. The final marginal is bounded with a $\tanh$ to ensure additional stability

$$q_t^\phi = \mathrm{softmax}\left[8 \tanh\left(\frac{h_\phi(t, \mathcal{Y}) + \rho_t}{8}\right)\right]. \tag{82}$$

The encoder is evaluated once per trajectory and is independent of the query time. We obtain the temporal score $s_t = \partial_t \log q_t^\phi$ by a forward-mode Jacobian–vector product through the query network, continuous-time attention, and likelihood readout, and set $\dot{q}_t^\phi = q_t^\phi \odot [s_t - \sum_x q_t^\phi(x) s_t(x)]$.

We optimize the stochastic nELBO using batches of 64 trajectories. Every observation term is evaluated at each update, while the path integral uses four stratified uniform times per trajectory, resampled at every update. The recognition network and log rates use separate Adam optimizers with a linear warm-up to learning rate 10^{-3}, followed by a constant phase and a cosine decay to 10^{-5} over the final 2,000 updates. The rates are initialized at $(0.4, 0.4, 0.4)$, frozen for the first 490 updates, and then learned jointly with the network. The parameter-recovery experiment uses 5,000 updates and five independent training seeds.

Inner optimization. Each trajectory–time pair is solved from $\lambda = 0$ by an adaptive capped-Newton method. At every iteration, we form the weighted graph Laplacian from the current tilted fluxes and solve the resulting dense 6×6 Newton system. We cap the Newton direction at $\|\delta\|_\infty \le 5$, and update $\lambda \leftarrow \lambda + \delta$. The cap prevents exponential overflow from a cold start while leaving the local quadratic convergence of Newton's method unchanged near the optimum.

The solver stops when both the cost discrepancy and mass-weighted master-equation residual of Appendix D.4 are at most 10^{-10}, or after 100 iterations. The 256 cells in a training batch—64 trajectories and four quadrature times—are solved together; cells that converge early are frozen while the remaining cells continue. The returned solutions are checked again at the training acceptance tolerance 10^{-3}, together with the zero-mass condition on the supplied derivative. Failed cells contribute a fixed penalty of 10^3 and no path gradient. For accepted cells, the potential is detached before evaluating the dual and its outer gradients.

We compared our custom solver with three Optimistix solvers[2] (Rader et al., 2024) on 256-cell batches from both trained and randomly initialized DFR checkpoints, using tolerance 10^{-10}. The step cap is the main practical difference from undamped Newton, which encounters nonfinite iterates at both checkpoints. `optimistix.Dogleg` is equally accurate. In our pure-JAX implementation,

[2] `https://docs.kidger.site/optimistix/`

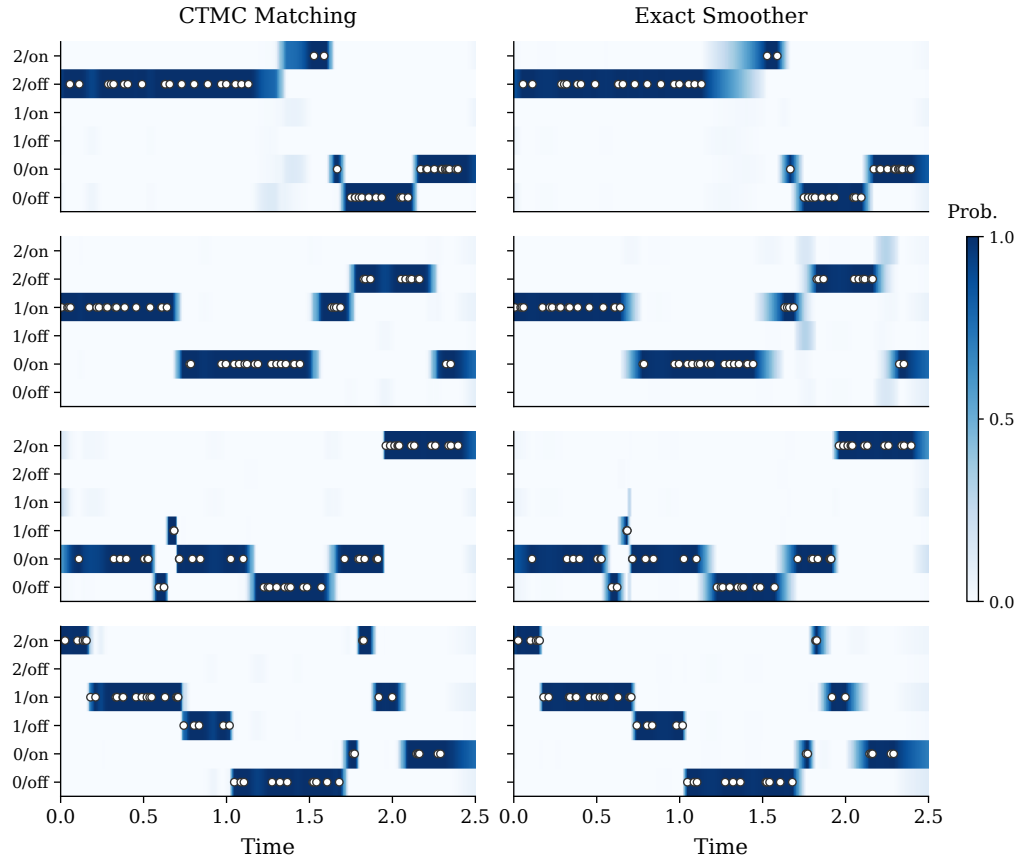


Figure 6: Comparison on held-out trajectories of the posterior marginals obtained via CTMC Matching with an exact smoother, for the discrete flashing ratchet experiment.

the 2.1 ms inner solve is negligible compared with a full training step, dominated by the transformer's forward and backward passes, which takes 845 ms on average.

E.2 Chemical Reaction Networks

A chemical reaction network (CRN) is a CTMC on a space of integer-valued molecule counts $z \in \mathcal{Z} \subseteq \mathbb{N}_0^d$ Wilkinson (2018). A reaction ρ changes the state by a stoichiometric increment ν_ρ and occurs with propensity $a_\rho^\theta(z)$, giving

$$R^\theta(z, z') = \sum_{\rho : z + \nu_\rho = z'} a_\rho^\theta(z), \qquad z' \neq z. \tag{83}$$

Although the number of reaction types is often small, the count state space can be very large or unbounded. To keep CTMC Matching tractable in this class of models, we use two approaches: when the count of each species can be bounded, so that the joint state space remains enumerable, we define a CTMC on the truncated state space. When this is not possible, one must find an ad hoc solution tailored to the problem at hand. In the remainder of this section, we present the first approach and its application to three of the synthetic benchmark systems considered in Eich et al. (2025), and a special case applicable to promoter cascade CRNs (Mikelson & Khammash, 2020) fit to experimental data from Rullan et al. (2018).

E.2.1 Synthetic benchmarks

We regenerate the gene-expression, enzyme-kinetics, and Lotka–Volterra datasets using the released scripts of Eich et al. (2025). Each dataset contains 100 trajectories from a common deterministic initial

state, observed at ten independently sampled and sorted uniform times with Gaussian measurement noise. The same saved observation times and values are used by CTMC Matching and all baselines, which we reproduce by re-running the public code from Eich et al. (2025)[3]. We fit two models independently for each system: one with the rates fixed at truth for posterior evaluation, and one with selected rates learned for parameter recovery.

The reaction schemes below use mass-action kinetics: c_ρ denotes a rate constant and z_A the molecule count of species A.

Gene expression. The species G, M, and P represent the gene, mRNA, and protein, respectively. Transcription and translation produce mRNA and protein, which also undergo degradation:

$$\mathsf{G} \xrightarrow{c_1} \mathsf{G} + \mathsf{M}, \qquad \mathsf{M} \xrightarrow{c_2} \mathsf{M} + \mathsf{P}, \qquad \mathsf{M} \xrightarrow{c_3} \emptyset, \qquad \mathsf{P} \xrightarrow{c_4} \emptyset. \tag{84}$$

The propensities, in reaction order, are $(c_1 z_G, c_2 z_M, c_3 z_M, c_4 z_P)$. We use $(c_1, c_2, c_3, c_4) = (200, 10, 25, 1)$, initial state $(z_G, z_M, z_P) = (1, 5, 20)$, and horizon $T = 8$. The gene count remains fixed at $z_G = 1$. Only protein is observed, through $Y \mid z \sim \mathcal{N}(z_P, 10)$, so the mRNA trajectory must be inferred indirectly.

Enzyme kinetics. A substrate S binds reversibly to a free enzyme E, forming a complex SE. The complex can dissociate or produce a product P while releasing the enzyme:

$$\mathsf{S} + \mathsf{E} \xrightarrow{c_1} \mathsf{SE}, \qquad \mathsf{SE} \xrightarrow{c_2} \mathsf{S} + \mathsf{E}, \qquad \mathsf{SE} \xrightarrow{c_3} \mathsf{P} + \mathsf{E}. \tag{85}$$

The propensities are $(c_1 z_S z_E, c_2 z_{\mathsf{SE}}, c_3 z_{\mathsf{SE}})$. We use $(c_1, c_2, c_3) = (0.05, 0.5, 0.5)$, initial state $(z_S, z_E, z_{\mathsf{SE}}, z_P) = (25, 5, 1, 1)$, and horizon $T = 20$. We observe substrate and product through $\mathcal{N}((z_S, z_P)^\top, 10 I_2)$; free enzyme and complex counts are latent. The reactions conserve $z_E + z_{\mathsf{SE}} = 6$ and $z_S + z_{\mathsf{SE}} + z_P = 27$.

Lotka–Volterra. The prey species X_1 reproduces, while predation converts one prey into an additional predator X_2, which can subsequently die:

$$\mathsf{X}_1 \xrightarrow{c_1} 2\mathsf{X}_1, \qquad \mathsf{X}_1 + \mathsf{X}_2 \xrightarrow{c_2} 2\mathsf{X}_2, \qquad \mathsf{X}_2 \xrightarrow{c_3} \emptyset. \tag{86}$$

The propensities are $(c_1 z_{X_1}, c_2 z_{X_1} z_{X_2}, c_3 z_{X_2})$. We use $(c_1, c_2, c_3) = (0.005, 0.001, 0.005)$, initial state $(z_{X_1}, z_{X_2}) = (20, 5)$, and horizon $T = 300$. Both populations are observed through $\mathcal{N}((z_{X_1}, z_{X_2})^\top, I_2)$. In particular, each predation event changes both populations simultaneously.

Finite inference graphs. We enable enumeration of the state space in each experiment by truncating the integer counts. Gene expression uses $0 \le z_M \le 30$ and $0 \le z_P \le 120$, giving 3,751 states. Enzyme kinetics retains the 174 states satisfying the two conservation laws within componentwise bounds $(26, 6, 6, 27)$. Lotka–Volterra uses $0 \le z_{X_1}, z_{X_2} \le 100$, giving 10,201 states. Boundary transitions follow the truncated smoother considered in Eich et al. (2025). To obtain strongly connected inference graphs, missing reverse transitions in the enzyme and Lotka–Volterra systems receive a fixed rate of 10^{-6}, and the exact smoother uses the same modified generator.

The deterministic initial state is slightly smoothed for CTMC Matching and the common reference:

$$p_0(z) = (1 - \varepsilon_0)\mathbf{1}\{z = z_0\} + \frac{\varepsilon_0}{|\mathcal{Z}|}, \qquad \varepsilon_0 = 10^{-12}. \tag{87}$$

This makes the initial KL finite while leaving the initial distribution effectively deterministic.

Posterior model and observation encoding. The recognition architecture follows the DFR model, with observation features adapted to continuous measurements. Each observed component is standardized using its training-set mean and standard deviation, and augmented with Fourier features:

$$\begin{aligned} z_{k,j} &= \frac{y_{k,j} - \mu_j}{\sigma_j + 10^{-9}}, \qquad e(z) = \big[z, \sin z, \cos z, \sin(2z), \cos(2z)\big], \\ e_y(y_k) &= \mathrm{concat}_{j=1}^{d_y}\, e(z_{k,j}). \end{aligned} \tag{88}$$

[3] https://github.com/yannickeich/ep4crns

Thus, the gene-expression observations have five features, while enzyme kinetics and Lotka–Volterra have ten. We concatenate these with the five time features used in the DFR experiment and a boundary indicator, then apply a learned projection to width 256. An additional learned boundary token is placed at $t = 0$. Four bidirectional transformer blocks encode the sequence, and a query at time t attends to it through two cross-attention blocks.

The output is a categorical distribution over the full finite state space, so dependencies between species are retained. Species-level means and intervals are obtained by marginalizing this joint distribution. We do not use the likelihood-readout bias of the DFR experiment, since the CRN smoothing distributions vary smoothly around the observation times.

Initial-distribution bias. To represent the rapid relaxation from the nearly deterministic initial law in equation equation 87, we add an exponentially decaying initial-distribution bias outside the bounded neural decoder:

$$b_0(t,z) = \gamma e^{-t/\tau}\,\mathrm{clip}(\log p_0(z), -C, C)\,, \qquad C = \max(8, 4\log|\mathcal{Z}|)\,,$$
$$\ell_\phi(t,z \mid \mathcal{Y}) = C\tanh\!\left(\frac{h_\phi(t,\mathcal{Y})_z}{C}\right) + b_0(t,z), \qquad q_t^\phi(z \mid \mathcal{Y}) = \mathrm{softmax}_z\,\ell_\phi(t,z \mid \mathcal{Y}). \tag{89}$$

The positive amplitude and decay timescale are learned. This soft inductive bias encourages agreement with the known initial law near $t = 0$ without enforcing $q_0^\phi = p_0$, and the initial KL term remains in the objective.

Training. All runs use 5,000 updates. Gene and enzyme use batches of 50 trajectories and eight stratified path times per trajectory; Lotka–Volterra uses batches of 25 and four path times. The Adam schedule warms up to 10^{-3} over 500 updates, remains constant for 2,500 updates, and decays to 10^{-5} over the final 2,000 updates. Inner problems are solved by matrix-free Newton–CG and must pass the acceptance checks of Appendix D.4. Note that, due to the large state space, a naive implementation of the inner solve takes much longer to converge: 38 ms for enzyme kinetics, 1,692 ms for gene expression, and 7,511 ms for LV. We leave an investigation of possible speedups in this setting, such as the use of mean-field variational families, to future work.

For parameter learning, rates are frozen for 500 updates and then optimized jointly with the recognition network. The parameter initializations follow the exact same setup as Eich et al. (2025): gene expression starts from $(200, 5, 25, 1)$ and learns only c_2; enzyme kinetics starts from $(0.05, 0.5, 0.1)$ and learns only c_3; Lotka–Volterra starts from $(0.002, 0.002, 0.002)$ and learns all three rates. All remaining rates stay fixed at their true values, and the learned rates are shared point estimates across the 100 trajectories.

Reference posterior and metrics. At the true rates, the reference posterior is computed by forward–backward smoothing on the finite graph using sparse matrix-exponential actions, using the public codebase of Eich et al. (2025). Posterior means are evaluated at 1,000 equally spaced times. For each trajectory i, we average squared errors between a method's posterior mean $\widehat{\mu}_{i,d}(t_j)$ and the reference mean $\mu^\star_{i,d}(t_j)$ over evaluation times and species:

$$e_i = \frac{1}{JD}\sum_{j=1}^{J}\sum_{d=1}^{D}\left[\widehat{\mu}_{i,d}(t_j) - \mu^\star_{i,d}(t_j)\right]^2, \qquad J = 1{,}000. \tag{90}$$

We report the mean and standard error of e_i across 100 trajectories. The relative parameter error is

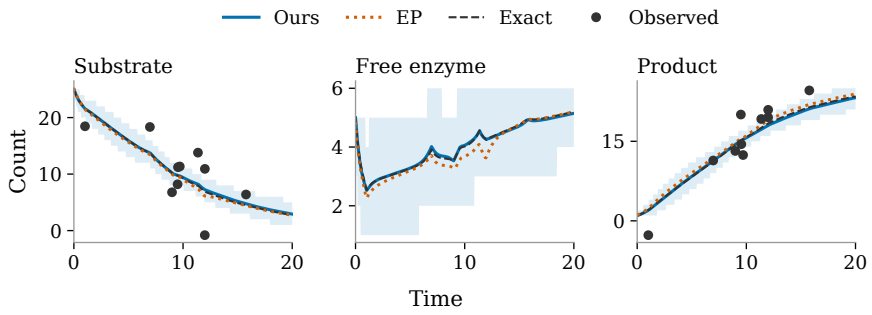
$$\mathrm{RPE} = \frac{100}{|\mathcal{I}|}\sum_{j\in\mathcal{I}}\left|\frac{\widehat{\theta}_j - \theta^\star_j}{\theta^\star_j}\right|,$$

where $\mathcal{I}$ contains the learned rates. Published EP rate estimates are used for its RPE, whereas the posterior baselines are rerun on the common observation arrays and reference grid.

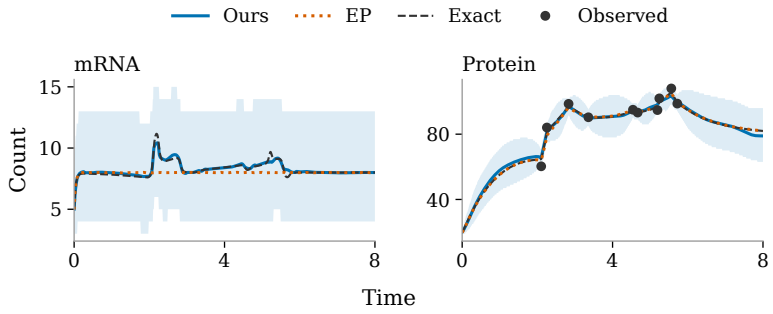
Figure 7 shows representative enzyme-kinetics and gene-expression reconstructions.

Model	Parameter	Truth	Initial	Ours	EP (Eich et al., 2025)	EP (reproduced)
Gene	c_2	10	5	9.9971	9.963	10.0003
Enzyme	c_3	0.5	0.1	0.501652	0.520	0.535551
LV	c_1	0.005	0.002	0.00463182	0.0077	0.00773192
	c_2	0.001	0.002	0.000969527	0.0012	0.00124123
	c_3	0.005	0.002	0.00497376	0.0064	0.00738512

Table 7: Synthetic CRN parameter estimates underlying the RPE comparison. Initial values and published EP estimates are transcribed from the original paper, Table 8. Only learned parameters are listed.



(a) Enzyme kinetics.



(b) Gene expression.

Figure 7: Representative posterior reconstructions for enzyme kinetics (top) and gene expression (bottom) on independently simulated trajectories. Lines show posterior means from CTMC Matching, EP, and the exact truncated smoother; shading denotes the central 90% posterior credible interval from CTMC Matching. In the enzyme system, only substrate and product are observed; in the gene-expression system, only protein is observed.

E.2.2 Live-cell transcription model

We use the reaction model of Mikelson & Khammash (2020). A single promoter switches between inactive and active states; active promoters initiate transcripts that pass through eight stages before leaving the measured nascent-RNA pool:

$$\begin{aligned} &\mathsf{G}_{\mathrm{off}} \xrightarrow{a} \mathsf{G}_{\mathrm{on}}, && \mathsf{G}_{\mathrm{on}} \xrightarrow{b} \mathsf{G}_{\mathrm{off}}, \\ &\mathsf{G}_{\mathrm{on}} \xrightarrow{c} \mathsf{G}_{\mathrm{on}} + \mathsf{R}_1, && \mathsf{R}_j \xrightarrow{k} \mathsf{R}_{j+1} \quad (j = 1, \dots, 7), \\ &\mathsf{R}_8 \xrightarrow{k} \emptyset. \end{aligned} \tag{91}$$

Writing $G_t \in \{0, 1\}$ for promoter activity and $N_{j,t}$ for the count of R_j, the propensities are $a(1-G_t)$, bG_t, cG_t, $kN_{j,t}$ for each stage transfer, and $kN_{8,t}$ for exit. The stages represent transcript residence time at the transcription site and are not separately observed. The initial state is $G_0 = 0$ and $N_{j,0} = 0$.

The Gaussian measurement model from Mikelson & Khammash (2020) is

$$Y_{t_\ell} \mid S_{t_\ell} \sim \mathcal{N}\big(S_{t_\ell},\, 16 + 0.3536^2 S_{t_\ell}\big)\,, \qquad S_t = \sum_{j=1}^{8} N_{j,t}. \tag{92}$$

Observations therefore constrain the total count but reveal neither individual stage counts nor promoter activity directly. We learn all four rates (a, b, c, k) separately for each illumination condition.

Variational family. We use a mean-field variational family

$$q_t^\phi(G, N) = \mathrm{Bern}(G; g_t) \prod_{j=1}^{8} \mathrm{Poi}(N_j; m_{j,t}), \tag{93}$$

where the nine scalars $(g_t, m_{1,t}, \ldots, m_{8,t})$ are the outputs of the inference network at $(t, \mathcal{Y})$. The recognition head enforces the known initial state, $g_0 = 0$ and $m_{j,0} = 0$, so that $q_0^\phi = p_0^\theta$ and the initial KL term vanishes. Equation 93 is an exponential family with sufficient statistic (G, N), so its time derivative is available by the same JVP as in Algorithm 1,

$$\frac{\dot q_t^\phi(G,N)}{q_t^\phi(G,N)} = \dot g_t\, \frac{G - g_t}{g_t(1-g_t)} + \sum_{j=1}^{8} \dot m_{j,t}\Big(\frac{N_j}{m_{j,t}} - 1\Big). \tag{94}$$

Under the family, $S_t = \sum_j N_{j,t}$ is Poisson with mean $\sum_j m_{j,t}$, so the reconstruction term in equation 92 is a one-dimensional Poisson expectation.

Given the promoter path, the stage counts form a first-order network: creation $\emptyset \to \mathsf{R}_1$ at rate cG_t, stage transfers $\mathsf{R}_j \to \mathsf{R}_{j+1}$ and exit $\mathsf{R}_8 \to \emptyset$ at per-capita rate k. We use a classical result from Jahnke & Huisinga (2007): first-order (monomolecular) reaction networks map product-Poisson laws to product-Poisson laws, with means that follow the reaction-rate equations, so the family contains the exact conditional prior of the counts given the promoter. The smoothing posterior given noisy count observations is not of product form, and equation 93 is a mean-field approximation of it. What the first-order structure buys for CTMC Matching is a compatible generator in closed form: the only reaction whose propensity couples G and N is creation, its effect on the balance of a state depends on N_1 alone, and the switching rates are free functions of the state, so a switching correction indexed by N_1 absorbs the coupling exactly. Fix t and write $D_{j,t} = \sum_{i \le j} \dot m_{i,t}$.

Lemma E.1 (Compatible completion for the transcription model)**.** *Suppose $g_t \in (0,1)$ and $m_{j,t} > 0$ for all j, and let $u_t \ge \max\{0, \max_j D_{j,t}\}$. Set*

$$\widetilde c_t = \frac{u_t}{g_t}, \qquad \widetilde k_{j,t} = \frac{u_t - D_{j,t}}{m_{j,t}}, \quad j = 1, \ldots, 8, \tag{95}$$

and let $\widetilde a_t(n_1), \widetilde b_t(n_1) \ge 0$ satisfy, for every $n_1 \in \mathbb{N}_0$,

$$(1-g_t)\,\widetilde a_t(n_1) - g_t\,\widetilde b_t(n_1) = h_t(n_1) := \dot g_t + (1-g_t)\,u_t\Big(1 - \frac{n_1}{m_{1,t}}\Big). \tag{96}$$

Then the generator with switching rates $\widetilde a_t(N_1)$ (off to on) and $\widetilde b_t(N_1)$ (on to off), creation rate $\widetilde c_t G$, and stage-transfer and exit rates $\widetilde k_{j,t} N_j$ satisfies $\widetilde R_t^\top q_t^\phi = \dot q_t^\phi$ on $\{0,1\} \times \mathbb{N}_0^8$, and its instantaneous path cost is

$$g_t\,\Phi(\widetilde c_t; c) + \sum_{j=1}^{8} m_{j,t}\,\Phi(\widetilde k_{j,t}; k) + \mathbb{E}_{n_1 \sim \mathrm{Poi}(m_{1,t})}\Big[(1-g_t)\,\Phi\big(\widetilde a_t(n_1); a\big) + g_t\,\Phi\big(\widetilde b_t(n_1); b\big)\Big]. \tag{97}$$

Among admissible switching pairs, the cost is minimized by solving, for each n_1, the two-state problem of Appendix D.1 with $\dot p$ replaced by $h_t(n_1)$. The remaining freedom is the scalar u_t.

Proof. Suppress time, and write $\Sigma(N) = \sum_j \dot m_j\,(N_j/m_j - 1)$, so that equation equation 94 becomes

$$\frac{\dot q(0,N)}{q(0,N)} = -\frac{\dot g}{1-g} + \Sigma(N), \qquad \frac{\dot q(1,N)}{q(1,N)} = \frac{\dot g}{g} + \Sigma(N).$$

Under a product-Poisson law, neighboring states have probability ratios

$$\frac{q(G,N-e_1)}{q(G,N)}=\frac{N_1}{m_1},\quad \frac{q(G,N+e_j-e_{j+1})}{q(G,N)}=\frac{m_j}{N_j+1}\,\frac{N_{j+1}}{m_{j+1}},\quad \frac{q(G,N+e_8)}{q(G,N)}=\frac{m_8}{N_8+1}.$$

Dividing the master equation at (G,N) by $q(G,N)$ and inserting these ratios, the count reactions contribute

$$C_G(N)=G\,\widetilde{c}\Big(\frac{N_1}{m_1}-1\Big)+\sum_{j=1}^{7}\widetilde{k}_j\Big(\frac{m_j}{m_{j+1}}N_{j+1}-N_j\Big)+\widetilde{k}_8\,(m_8-N_8),$$

which is affine in N, while the switching reactions contribute $\frac{g}{1-g}\widetilde{b}(N_1)-\widetilde{a}(N_1)$ at $G=0$ and $\frac{1-g}{g}\widetilde{a}(N_1)-\widetilde{b}(N_1)$ at $G=1$. The master equation therefore reads

$$\frac{g}{1-g}\widetilde{b}(N_1)-\widetilde{a}(N_1)+C_0(N)=-\frac{\dot g}{1-g}+\Sigma(N),\quad \frac{1-g}{g}\widetilde{a}(N_1)-\widetilde{b}(N_1)+C_1(N)=\frac{\dot g}{g}+\Sigma(N). \tag{98}$$

Count rates. Multiplying the first equation by $1-g$, the second by g, and adding, the switching terms cancel:

$$(1-g)\,C_0(N)+g\,C_1(N)=\Sigma(N). \tag{99}$$

Creation enters here only through the expected flux $g\widetilde{c}=u$. Both sides are affine in N, and comparing the coefficients of N_1 and of N_j for $j\ge 2$ gives

$$\widetilde{k}_1 m_1=u-\dot m_1,\qquad \widetilde{k}_j m_j=\widetilde{k}_{j-1}m_{j-1}-\dot m_j,$$

i.e. $\widetilde{k}_j m_j=u-D_j$, which is equation equation 95. The constant terms then agree as well, since $-u+\widetilde{k}_8 m_8=-D_8$.

Switching rates. With equation equation 99 in place, the second equation in equation 98, multiplied by g, becomes

$$\begin{aligned}(1-g)\,\widetilde{a}(N_1)-g\,\widetilde{b}(N_1)&=\dot g+g\,\Sigma(N)-g\,C_1(N)\\&=\dot g+g(1-g)\big(C_0(N)-C_1(N)\big)\\&=\dot g+(1-g)\,u\Big(1-\frac{N_1}{m_1}\Big),\end{aligned}$$

which is constraint equation 96, and the first equation in equation 98 is the difference of equation 99 and this one. Hence the rates of the lemma satisfy the master equation at every state.

Cost. Since $\Phi(\kappa n;kn)=n\,\Phi(\kappa;k)$, each transfer or exit reaction contributes $\sum_{G,N}q(G,N)\,N_j\,\Phi(\widetilde{k}_j;k)=m_j\,\Phi(\widetilde{k}_j;k)$, creation contributes $\sum_N q(1,N)\,\Phi(\widetilde{c};c)=g\,\Phi(\widetilde{c};c)$, and the switching reactions contribute $\sum_N q(0,N)\,\Phi(\widetilde{a}(N_1);a)+q(1,N)\,\Phi(\widetilde{b}(N_1);b)$, which depends on N only through N_1. This is equation 97. Both the constraint in equation 96 and the last term separate over the value of N_1, so the cost is minimized by taking each pair $(\widetilde{a}(n_1),\widetilde{b}(n_1))$ to be the two-state solution of Appendix D.1 with $\dot p=h(n_1)$, which exists because $ab>0$. □

Recognition network. We use the same observation features and continuous-time transformer architecture as in the synthetic experiments, with a smaller configuration. For each illumination condition, scalar observations are standardized using the training-set mean and standard deviation, $z_k=(y_k-\mu_{\text{train}})/\max(s_{\text{train}},1)$. The five observation features in equation 88 and the five time features are projected to width 64. Two bidirectional transformer encoder blocks process the observation sequence, and one cross-attention block computes the representation at each query time. All attention blocks use four heads, pre-layer normalization, and GELU feed-forward networks of hidden width 128. We retain the learned continuous-time rotary embeddings, relative-time attention biases, and register key/value of the shared architecture. No likelihood bias term in the logits is used. The 64-dimensional query representation is concatenated with the five query-time features and passed

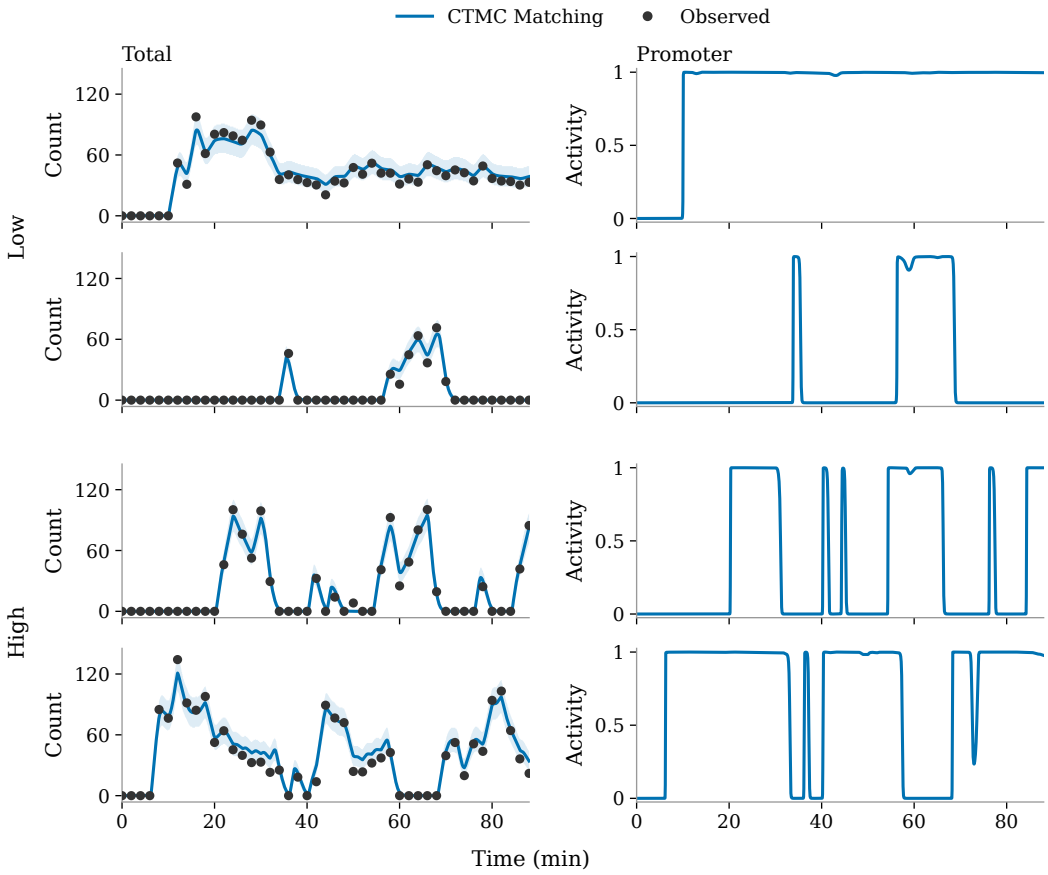


Figure 8: Posterior reconstructions for two held-out trajectories from each illumination condition: low (top two rows) and high (bottom two rows). The left column shows the inferred mean total nascent-RNA count and the observations, shading denotes pointwise 90% posterior credible intervals for the latent total count. The right column shows the inferred probability that the promoter is active. Each reconstruction conditions on the complete observation sequence.

	CRPS ↓	
Method	Low	High
CTMC Matching	14.606 ± 0.986	**19.249** ± 1.238
FFBS-ent	14.365 ± 1.389	24.830 ± 1.670
LFNS	**14.189** ± 1.697	—

Table 8: Transcription predictive CRPS (mean $\pm$ SE across 21 low- and 11 high-condition held-out trajectories). Predictions use the known initial state without conditioning on a trajectory prefix. The LFNS row uses the published posterior-mean parameters of Mikelson & Khammash (2020); all scores use the same normalized observation model.

through a GELU MLP of dimensions $69 \to 64 \to 9$, producing outputs $g_t, m_{1,t}, \ldots, m_{8,t}$ and their JVPs give $\dot{g}_t, \dot{m}_{1,t}, \ldots, \dot{m}_{8,t}$. We parameterize u_t as $\max\{0, \max_j D_{j,t}\}$ plus a positive slack, so that the constraints of Lemma E.1 hold by construction, and the slack is optimized numerically. The switching pairs are the closed-form two-state solutions, the count costs are the closed-form terms of equation 97, and the switching cost and the reconstruction term are finite Poisson sums. We train one network for each of the two illumination conditions.

Data and predictive evaluation. We use the processed measurements released by Mikelson & Khammash (2020),[4] treating the low- and high-illumination conditions as separate datasets. They contain 218 and 103 trajectories, respectively, each with 45 observations at two-minute intervals. We fit CTMC Matching and FFBS-ent Eich et al. (2025) using 195 low-condition and 92 high-condition training trajectories, whereas the published LFNS estimate from Mikelson & Khammash (2020) was fitted to only five low-condition trajectories. We assess the learned rates by predicting complete held-out trajectories from the known inactive-promoter, zero-count initial state. Table 8 reports the continuous ranked probability score of the noisy-observation predictive marginals, averaged over time and held-out trajectories. Note that, as this is real experimental data, latent trajectories and true rates are unavailable.

E.3 Wildfire Experiments

Latent dynamics and covariates. The WildFireSpreadTS observations are binary VIIRS fire-detection masks at a resolution of $375\mathrm{m} \times 375\mathrm{m}$ per pixel (Gerard et al., 2023). In both tasks, a task-specific UTAE encoder maps the accompanying covariate sequence X to a 16-dimensional feature field $H_t^\theta = (H_{t,i}^\theta)_i$ (the weekly and daily constructions are given below). Following Migliorini & Smyth (2026), we model the latent state of each pixel as a value in $\{\mathsf{U}, \mathsf{A}, \mathsf{B}\}$, denoting unburned, actively burning, and burned out, and link it to the detection mask through a Bernoulli emission. The allowed transitions are $\mathsf{U} \to \mathsf{A}$, $\mathsf{A} \to \mathsf{B}$, and $\mathsf{B} \to \mathsf{A}$.

We construct our architecture to mirror the dynamics model of a stochastic cellular automaton Clarke et al. (1994); Xia & Cheng (2025); Zhenirovskyy et al. (2026): a neural network shared across sites and times maps the neighboring latent states and local covariate features to three unconstrained rate scores:

$$\left(\eta_{i,t}^{\theta,\mathsf{UA}}, \eta_{i,t}^{\theta,\mathsf{AB}}, \eta_{i,t}^{\theta,\mathsf{BA}}\right) = f_\theta\left(z_{N(i)}, H_{t,N(i)\cup\{i\}}^\theta\right). \tag{100}$$

Writing

$$m_i(z) = \mathbf{1}\{\exists\, j \in \mathcal{N}(i) : z_j = \mathsf{A}\}$$

for the indicator that at least one of the eight Moore neighbors of pixel i is actively burning, the local transition rates are then

$$\begin{aligned} R_{i,t}^\theta(\mathsf{U},\mathsf{A}) &= \varepsilon_{\mathrm{ign}} + m_i(z)\,\mathrm{softplus}\left(\eta_{i,t}^{\theta,\mathsf{UA}}\right), \\ R_{i,t}^\theta(\mathsf{A},\mathsf{B}) &= \mathrm{softplus}\left(\eta_{i,t}^{\theta,\mathsf{AB}}\right), \\ R_{i,t}^\theta(\mathsf{B},\mathsf{A}) &= m_i(z)\,\mathrm{softplus}\left(\eta_{i,t}^{\theta,\mathsf{BA}}\right). \end{aligned} \tag{101}$$

[4] https://github.com/Mijan/LFNS_paper

We model f_θ as a small bidirectional transformer: the model processes the 3×3 Moore patch around each pixel as nine tokens, and the token for a neighboring site j is $[e(z_j), P(H^\theta_{t,j}), P(H^\theta_{t,j}) - P(H^\theta_{t,i})]$, combining a latent-state embedding with projections of the covariate features. The state embedding at the center site i is replaced by a learned mask token, and two pre-normalized transformer blocks of width 64 with four heads produce the scores in equation 100. To satisfy the connected support assumption in Appendix A.1, we additionally set $R^\phi_{i,t}(\mathsf{B}, \mathsf{U}) = \varepsilon_{\mathrm{ret}}$, with $0 < \varepsilon_{\mathrm{ign}}, \varepsilon_{\mathrm{ret}} \ll 1$. We evaluate the network only on the fire frontier (i.e., sites that are active or have an active neighbor) since the remaining rates are known directly from equation 101.

E.3.1 Weekly benchmark

Training and evaluation. The benchmark from Migliorini & Smyth (2026) provides 130 training and 26 test sequences of seven daily 64×64 detection masks, giving a six-day continuous-time horizon.[5] We hold out 26 of the 130 training trajectories for validation. We use $p(Y = 1 \mid Z = \mathsf{A}) = 0.99$, $p(Y = 1 \mid Z \neq \mathsf{A}) = 10^{-3}$, and $\varepsilon_{\mathrm{ign}} = \varepsilon_{\mathrm{ret}} = 10^{-3}$. For each trajectory and update, we draw one uniform time and two independent latent fields, using the other field's cost as a leave-one-out score-function baseline (Kool et al., 2019). The local matching problems use 40 coordinate-ascent sweeps and tolerance 10^{-8}, batched across sites and datapoints. We train with AdamW for 500 epochs using batch size 32, a peak learning rate of 3×10^{-4} after 500 warm-up updates, weight decay 10^{-6}, and gradient clipping at 1. The checkpoint is selected on the validation split. Reconstruction conditions on all seven observations, and due to our architecture can be evaluated in parallel. For prediction, we sample the initial latent field and run the prior forward for six days without further observations, following Migliorini & Smyth (2026). We use an Euler jump discretization with $\mathrm{d}t = 0.02$ days: transition probabilities are rate times $\mathrm{d}t$, with the diagonal assigned the residual probability, followed by one categorical draw per pixel (Campbell et al., 2024; Sun et al., 2023; Migliorini & Smyth, 2026). Active-state probabilities are estimated with 16 particles and averaged over three simulation seeds. We report mean per-pixel binary cross-entropy against the detections over all seven days, with intervals of ± 2 standard errors over the 26 test sequences.

Encoders and variational marginals. The generative encoder applies a UTAE (Garnot & Landrieu, 2021) to the three daily VIIRS reflectance channels. To obtain features for day d, inputs from later days are zeroed before encoding, so $H^\theta_{i,d} \in \mathbb{R}^{16}$ depends only on days $0{:}d$. At a continuous time t, the rate network uses the feature block for day $\lfloor t \rfloor$. The UTAE has channel widths $(32, 32, 64)$, model width 64, four attention heads, and group normalization.

Following Section 3.2, the variational posterior factorizes over sites into three-state categorical marginals. Its recognition UTAE has the same architecture as the generative encoder, and its five input channels are the three VIIRS channels, the fire detections, and an observation mask. For each pixel i and observation time τ_k, the UTAE produces a feature $h^\phi_{i,k} \in \mathbb{R}^{16}$. Before predicting the temporal coefficients, we concatenate

$$u^\phi_{i,k} = \left[h^\phi_{i,k};\, g_{i,k};\, o_{i,k};\, \tau_k/T\right] \in \mathbb{R}^{21}, \qquad g_{i,k}(v) = \log p(Y_{i,k} \mid Z_{i,k} = v), \tag{102}$$

where $g_{i,k} \in \mathbb{R}^3$ gives the emission log-likelihood under each latent state and $o_{i,k}$ indicates whether the observation is included in the conditioning set. These tokens determine time-independent coefficients for the posterior logits:

$$\ell^\phi_{i,v}(t \mid \mathcal{Y}) = \sum_{r=0}^{11} a^\phi_{i,r,v}(\mathcal{Y}) \cos\left(\frac{\pi r t}{T}\right) + \sum_k b^\phi_{i,k,v}(\mathcal{Y})\, \kappa_h(t - \tau_k), \tag{103}$$

where $\kappa_h(s) = \max(0, 1 - |s|)$. Thus all observation dependence is contained in $a^\phi(\mathcal{Y})$ and $b^\phi(\mathcal{Y})$, which are computed once per sequence, while query-time dependence enters only through the fixed basis. We set $q^\phi_{i,t}(v \mid \mathcal{Y}) = \mathrm{softmax}_v\, \ell^\phi_{i,v}(t \mid \mathcal{Y})$ and obtain $\dot{q}^\phi_{i,t}$ analytically from the basis derivatives and softmax Jacobian.

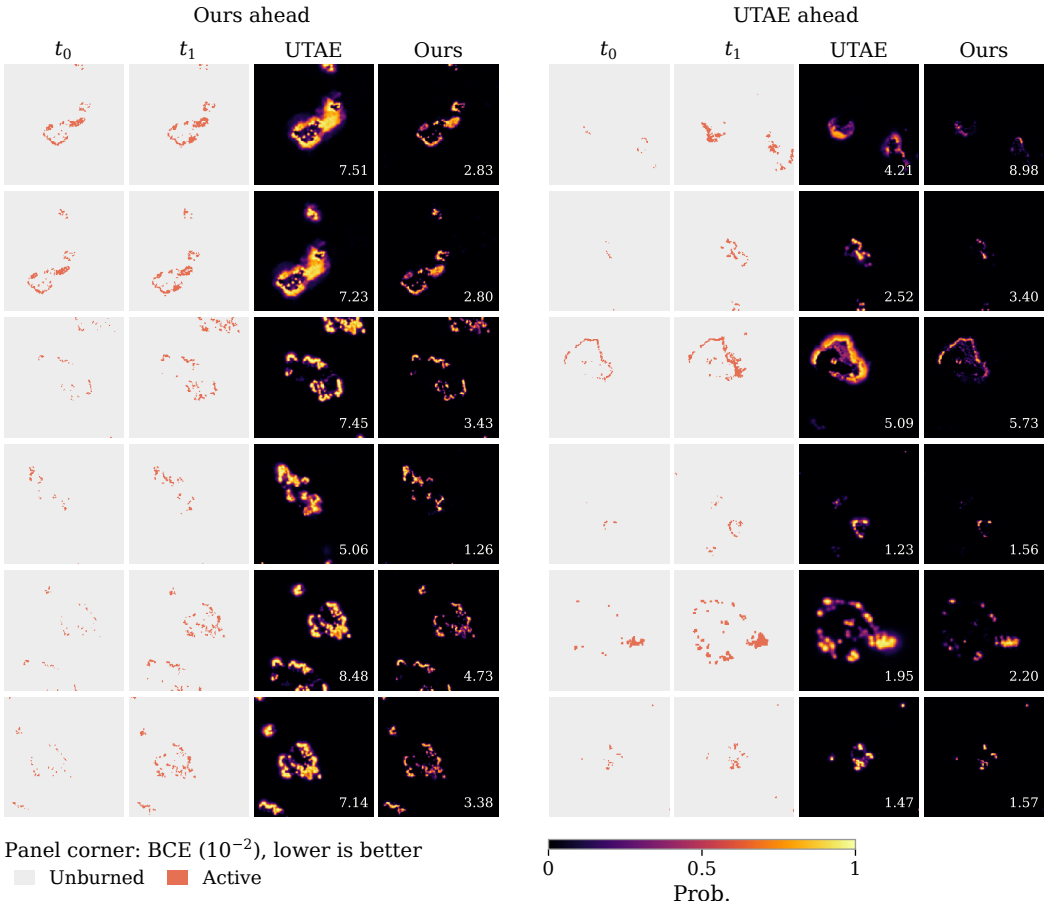


Figure 9: Selected held-out samples display UTAE's predicted probabilities against the predictions obtained from an ensemble of 32 rollouts from our latent IPS model. *Left:* examples where the latent IPS has lower BCE; *Right:* examples where the UTAE has lower BCE.

E.3.2 Daily WildFireSpreadTS forecasting

Training and evaluation. Each example contains five observed days and the next active-fire mask as its target.[6] We retain seven channels: VIIRS M11, I2, and I1, NDVI, EVI2, and the raw and binary active-fire channels. The 2018–2020 fires are split by fire identifier into 6,084 training and 1,151 validation windows, and the 3,337 windows from 2021 form the test set. Covariates are standardized using the dataset's fixed training statistics. We reuse the data-loading mechanism from Gerard et al. (2023): at training we sample fire-preferring random 128×128 crops, whereas evaluation uses each test window at its native spatial extent. The supervised UTAE baseline of Gerard et al. (2023) uses the same inputs, splits, preprocessing, and validation-AP checkpoint selection. We use the UTAE hyperparameters of Gerard et al. (2023) for the same task.

As in the weekly task, each update uses one uniform time, two independently sampled latent fields, and a leave-one-out baseline. The batched local solve uses 10 coordinate-ascent sweeps at tolerance 10^{-8}. The emission probabilities are initialized at $p_{\text{on}} = 0.99$ and $p_{\text{off}} = 10^{-4}$ and learned jointly with the dynamics. During training, we set $\varepsilon_{\text{ign}} = 10^{-3}$ and $\varepsilon_{\text{ret}} = 10^{-5}$. We use AdamW with constant learning rate 3×10^{-4}, batch size 48, weight decay 10^{-6}, gradient clipping at 1, and report the checkpoint after 50 epochs of training.

[5] Our weekly training implementation builds on code from the LatentIPS repository: `https://github.com/giosueio/LatentIPS`.

[6] Our daily training implementation builds on code from the WildfireSpreadTS repository: `https://github.com/SebastianGer/WildfireSpreadTS`.

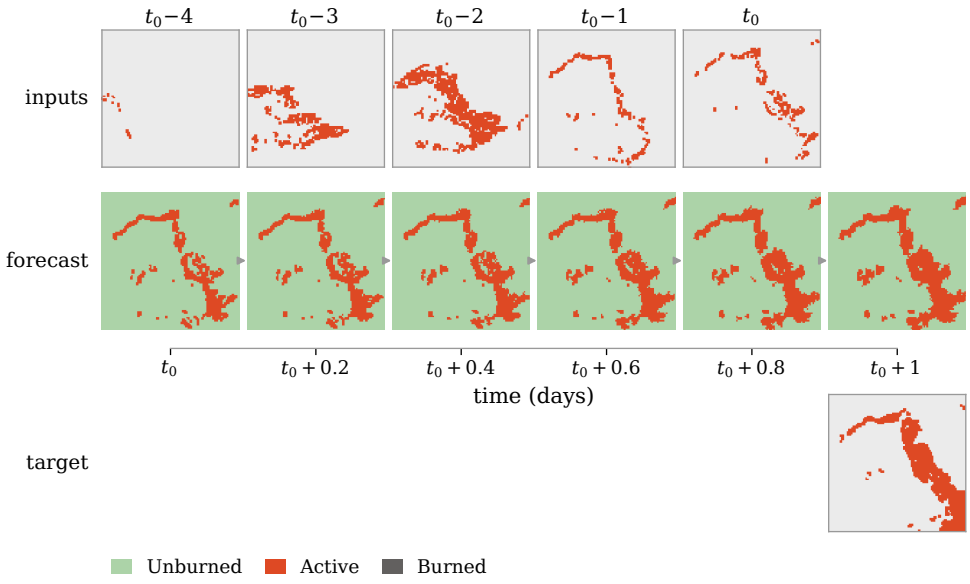


Figure 10: Continuous-time rollout from our latent IPS model for the daily forecasting task, on a held-out sample.

			Median		99th percentile		Maximum	
Grid	Model state	Pixels	Sweeps	ms (%)	Sweeps	ms (%)	Sweeps	ms (%)
64×64	Random init.	32,768	3	2.8 (0.91)	6	5.3 (1.74)	7	6.1 (1.97)
64×64	Converged	32,768	3	2.7 (0.89)	5	4.2 (1.34)	8	6.3 (2.03)
128×128	Random init.	131,072	4	4.6 (0.67)	6	6.8 (0.99)	6	6.8 (0.99)
128×128	Converged	131,072	2	2.5 (0.37)	3	3.5 (0.51)	6	6.2 (0.92)

Table 9: Counts of coordinate-ascent sweeps needed for wildfire training, with batch size 4. Sweep quantiles are aggregated over 100 microbatches. Runtime is the cost of the corresponding sweep budget in milliseconds, with its percentage relative to a full training step in parentheses.

At the forecast origin, the latent field is sampled from the learned filtering distribution conditioned on all five observed days, after which the learned prior is rolled forward. Forecasts use 32 ensemble members and the Euler jump scheme above with $\mathrm{d}t = 0.02$. The off-frontier ignition rate $\varepsilon_{\mathrm{ign}}$ is set to zero for daily evaluation.

Encoders and variational marginals. The generative UTAE encodes the five-day, seven-channel input window $(\mathcal{Y}_{t_0-5:t_0}, X_{t_0-5:t_0})$ into a single feature vector $H_i^\theta \in \mathbb{R}^{16}$ per pixel. Mirroring the weekly task, this feature is held fixed over the forecast interval $t \in [t_0, t_1]$. The recognition model receives the complete six-day, seven-channel window $(\mathcal{Y}_{t_0-5:t_1}, X_{t_0-5:t_1})$, i.e. it learns a one-day-ahead smoother. For simplicity, let $t_0 = 0$ and $t_1 = 1$. We found that a modification to the functional form of the output made convergence much faster: the model outputs endpoint logits $l_{0,i}^\phi(\mathcal{Y})$ and $l_{1,i}^\phi(\mathcal{Y})$ together with a 32-dimensional spatial feature $h_i^\phi(\mathcal{Y})$, and the logits interpolate through

$$\ell_i^\phi(t \mid \mathcal{Y}) = (1-t)l_{0,i}^\phi(\mathcal{Y}) + tl_{1,i}^\phi(\mathcal{Y}) + t(1-t)c_i^\phi\Big(t, h_i^\phi(\mathcal{Y})\Big), \tag{104}$$

where c_i^ϕ is a small local correction network conditioned on a Fourier embedding of t. We apply a state-wise softmax to obtain $q_{i,t}^\phi$ and compute its derivative by a forward-mode Jacobian–vector product. The filtering law $p_{t_0}^\theta$ used to initialize a forecast is obtained by applying the same trunk's terminal head to the five observed days, with the target withheld.

Coordinate-ascent sweeps. We additionally measured how many sweeps individual local matching problems require to pass the 10^{-8} acceptance tolerance. Table 9 aggregates 100 representative microbatches of four sequences, each using one sampled time and two latent-field samples. Across both resolutions (64×64, used in the weekly experiment, and 128×128, used in the daily experiment) and both initialization regimes (at initialization and converged), no observed solve requires more than eight sweeps. Under this adaptive regime, the wallclock time of a converged inner solve sits under 2% of the total for a full training step.

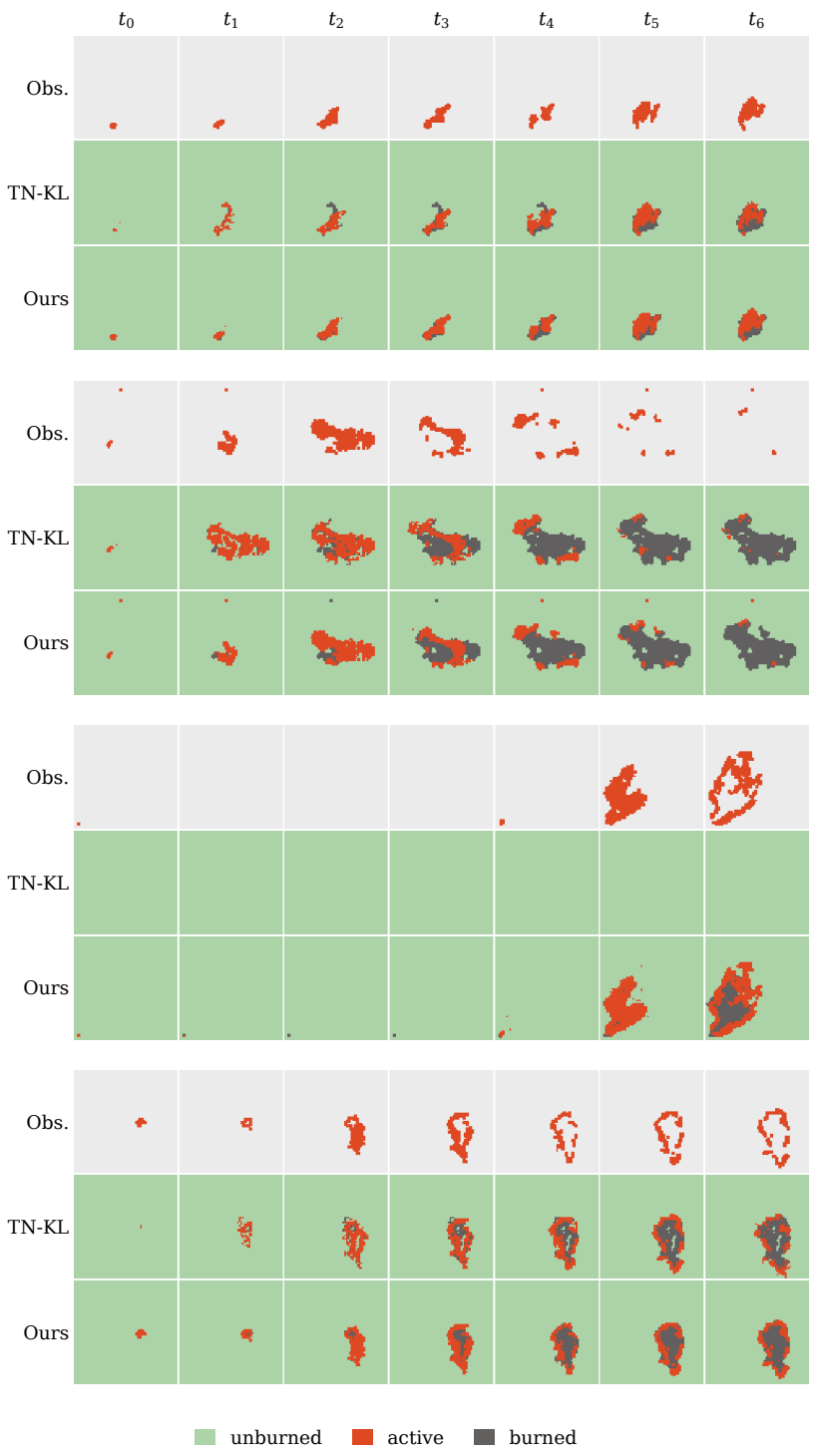


Figure 11: Weekly wildfire reconstructions. Each group shows the observations, TwistNet–KL reconstruction, and CTMC Matching reconstruction over seven days. Green, orange, and gray denote unburned, active, and burned states.

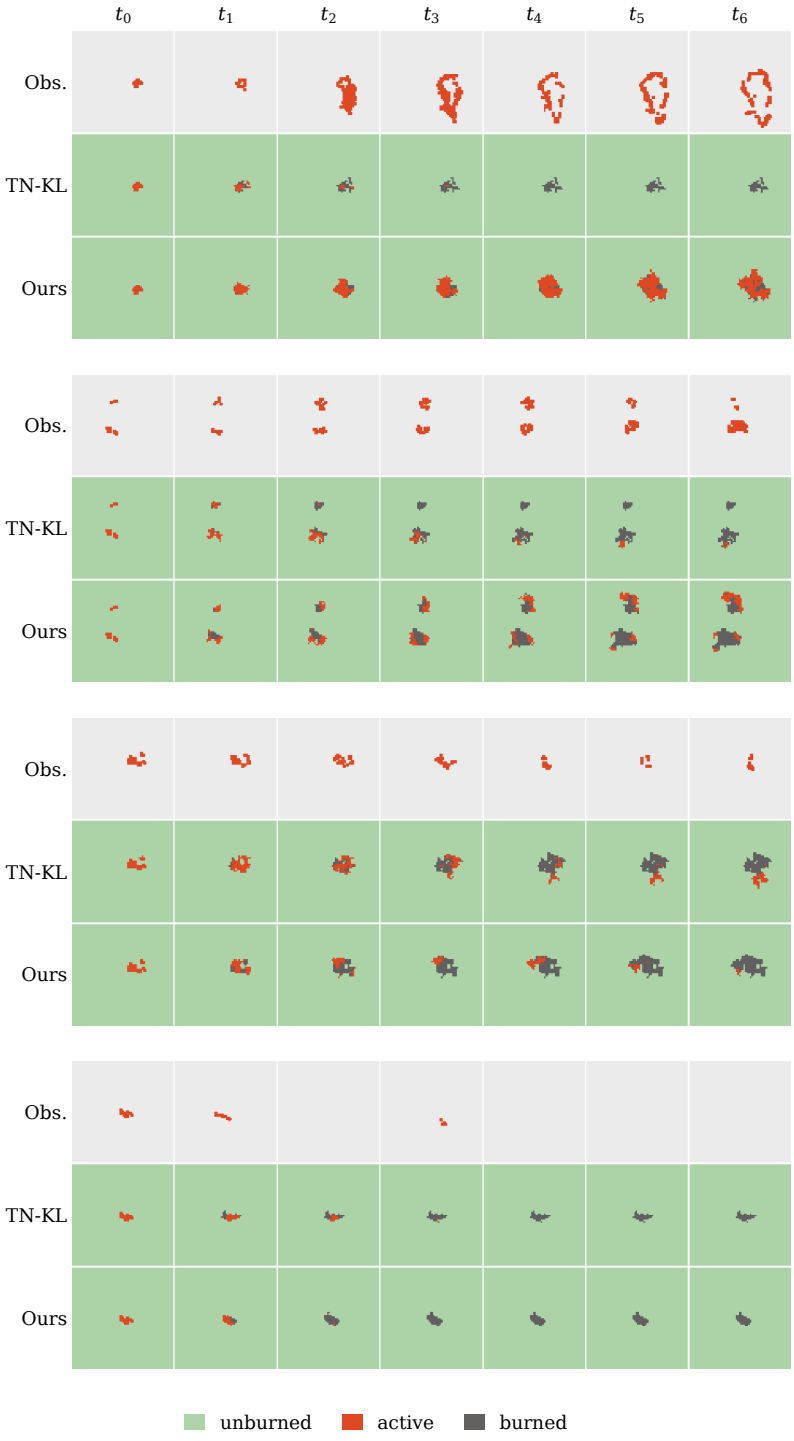


Figure 12: Weekly wildfire forecasts initialized at day 0. Rows compare the observations, TwistNet–KL, and CTMC Matching over the six-day forecast horizon.

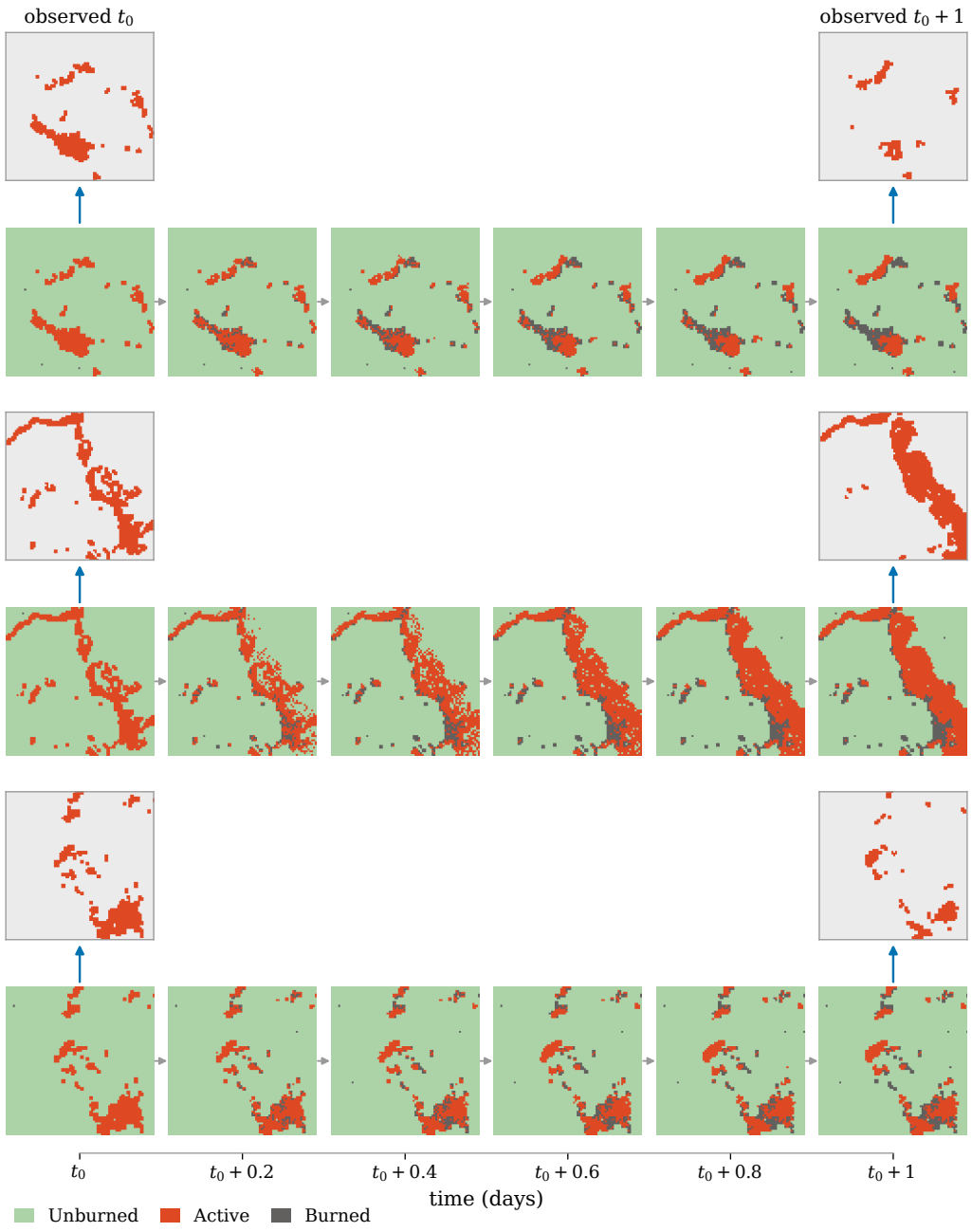


Figure 13: Sub-daily continuous-time trajectories simulated from the variational smoothing posterior CTMC, from the model trained for the weekly forecasting task.