Quantifying Rare Events in Stochastic Reaction-Diffusion Dynamics Using Tensor Networks

Schuyler B. Nicholson and Todd R. Gingrich

Department of Chemistry, Northwestern University, 2145 Sheridan Road, Evanston, Illinois 60208, USA

The interplay between stochastic chemical reactions and diffusion can generate rich spatiotemporal patterns. While the timescale for individual reaction or diffusion events may be very fast, the timescales for organization can be much longer. That separation of timescales makes it particularly challenging to anticipate how the rapid microscopic dynamics gives rise to macroscopic rates in the nonequilibrium dynamics of many reacting and diffusing chemical species. Within the regime of stochastic fluctuations, the standard approach is to employ Monte Carlo sampling to simulate realizations of random trajectories. Here, we present an alternative numerically tractable approach to extract macroscopic rates from the full ensemble evolution of many-body reaction-diffusion problems. The approach leverages the Doi-Peliti second-quantized representation of reaction-diffusion master equations along with compression and evolution algorithms from tensor networks. By focusing on a Schlögl model with one-dimensional diffusion between L otherwise well-mixed sites, we illustrate the potential of the tensor network approach to compute rates from many-body systems, here with approximately 3×10^{15} microstates. Specifically, we compute the rate for switching between metastable macrostates, with the expense for computing those rates growing subexponentially in L. Because we directly work with ensemble evolutions, we crucially bypass many of the difficulties encountered by rare event sampling techniques—detailed balance and reaction coordinates are not needed.

I. INTRODUCTION

Connecting microscopic kinetics to emergent rates has a long history for equilibrium and nonequilibrium processes. Complex systems often settle into metastable basins with rare transitions regulating the rates of switching between the emergent states. In equilibrium, transition state theory [1,2], variational transition state theory [3], Grote-Hynes theory [4,5], and transition path sampling (TPS) [6,7] have given strategies to estimate the transition rate. Away from equilibrium, theories built upon a free-energy landscape lose their applicability, yet trajectory sampling approaches can extract rates from microscopic dynamics, through a noise-guided TPS approach [8,9] or forward flux sampling (FFS) [10–12]. For all their merits, sampling approaches bring intrinsic challenges. How many samples are required to estimate the rate? Is a good approximate reaction coordinate needed, and if so, how is it found [13]? How does the expense grow with larger or more complicated microscopic systems? Motivated by those ubiquitous

concerns, we set out to consider a different approach— one that effectively samples all possible trajectories as an ensemble evolution instead of propagating individual realizations. Such a proposition may appear absurd given the exploding state space presented by the many-body problem, but we show that for the class of discrete state, continuous-time, reaction-diffusion master equations (RDMEs), we can evolve the full distribution of classical many-body systems in a numerically controllable manner by combining two well-developed methods, the Doi-Peliti (DP) formalism and tensor network (TN) algorithms.

While TPS and FFS have been developed for a variety of dynamical equations of motion (Hamiltonian, Langevin, Markov jump, etc.), we focus our attention solely on the chemical master equation (CME) dynamics, the setting where FFS was first developed [10]. More specifically, we study discrete reaction-diffusion dynamics of L sites or voxels as they are often called in the RDME literature [14,15]. The rationale for this focus is twofold. From a practical perspective, the RDME is important as a popular method to model chemical and biological processes [16–18]. From a theoretical perspective, the RDME is interesting because it provides a setting where one can analyze the influence of stochasticity on the emergence of patterns [19–21]. In many pattern-formation problems, it is possible to understand the underlying physics in terms of a

mean-field partial differential equation that tracks spatio-temporal evolution of deterministically evolving fields [22,23]. It is, however, well appreciated that stochastic effects can quantitatively and even qualitatively impact kinetics [24–29]. It is therefore important that the methodology we discuss captures more than just the typical dynamics. Accurately computing transition rates between metastable states requires quantitatively resolving even rare tails of the state space distribution since these rare fluctuations can be instrumental in triggering transitions [11,30].

Historically, similar schemes for the CME ensemble evolution have been hampered by the curse of dimensionality. As a result, exact solutions are rare [31] and numerical solutions often resort to sampling methods built upon the kinetic Monte Carlo stochastic simulation algorithm (SSA) [32,33], also known as the Gillespie algorithm. Another approach is to evolve a subset of the state space through the finite state projection (FSP) method [34]. FSP has previously been combined with TNs to approximately describe the CME [35,36], resembling our use of TNs to tame the many-body problem. Unlike those projective methods, however, the approach we describe uses the time-dependent variational principle (TDVP) to evolve the dynamics with probability conservation at all times, akin to prior work with diffusion and no reactions [37–39]. Since we estimate rare events by measuring the distribution's small probability fluxes, conservation of probability and high-fidelity dynamical evolution are materially important.

While the formalism we describe applies quite generally to different well-mixed and reaction-diffusion models, we center our paper around a single model problem: rate calculations for transitions between basins of the Schlögl model [40] in the bistable regime, both with and without diffusion [41]. In Sec. II, we first illustrate the rare event problem and extract rate constants from a propagator in the well-mixed regime. For more complicated systems (either systems with more chemical species or more voxels), the curse of dimensionality precludes the exact calculation of the propagator, but the remainder of the paper shows how to employ TNs to follow the same conceptual path. In Sec. III, we review the Doi-Peliti framework for writing reaction-diffusion dynamics in a second-quantized field theoretic form. Section IV then reviews the TN matrix product state (MPS) ansatz, a controllable approximation to the ensemble that can be evolved via TDVP. Those methods are applied to the many-body rate calculation in Sec. V. In Sec. VI, we show results for the RDME with diffusion between L well-mixed Schlögl voxels, demonstrating that the methodology can extend numerical transition rate calculations to regimes where the SSA becomes impractical. The approach effectively evolves the probability of occupying each microstate in an unfathomably large many-body state space, yet the calculations scale subexponentially with system size and do not assume detail-balanced equilibrium. Section VII discusses how the TDVP
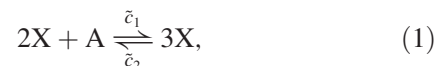
calculations yield not only a rate, but also a mechanism. Finally, in Sec. VIII, we discuss potential applications and challenges that should motivate future work.

II. ENSEMBLE RATE CALCULATIONS

Consider an archetypal two-state rate process between reactants and products R⇌P, initiated entirely in the reactant state. After some reaction timescale, the concentration of products grows before eventually leveling off at its steady-state value. The reaction rate can simply be defined and measured in terms of how quickly the product concentration rises. An alternative formulation is to study the fate of an individual reactant on its path to transitioning into a product. The TPS and FFS path sampling approaches extract the reaction rate by focusing on the statistical properties of these individual trajectories. Algorithms exist to propagate those single trajectories, algorithms that can be practically implemented even when the system grows large and complex. By contrast, propagating the entire ensemble of trajectories has typically been limited to simple low-dimensional problems. In this section, we illustrate an ensemble-evolution rate calculation for a simple case in order to develop notation and lay the groundwork for the many-body problem. The remainder of the paper will show that the methodology can be extended beyond such simple problems by employing TNs.

A. Well-mixed Schlögl model

Consider the well-mixed Schlögl model, a quintessential example of bistability in chemical reaction networks (CRNs) [42,43]. The model tends to be studied under chemostatted conditions that fix the concentration of two species A and B, while allowing the amount of a third species X to evolve according to the four reactions:

$$2X + A \underset{\tilde{c}_2}{\overset{\tilde{c}_1}{\rightleftharpoons}} 3X, \qquad (1)$$

$$B \underset{\tilde{c}_4}{\overset{\tilde{c}_3}{\rightleftharpoons}} X, \qquad (2)$$

where $\tilde{c}_1$, $\tilde{c}_2$, $\tilde{c}_3$, and $\tilde{c}_4$ are the stochastic rate coefficients. Our starting point is microscopic such that the rate $\tilde{c}_1$ should be interpreted as the reciprocal of the waiting time for a particular set of 2 X molecules to react and form 3 X molecules when mediated by a fixed number of A molecules. In the large-system, well-mixed limit of chemical kinetics, it is common to also define kinetic rate constants $k_r = \tilde{c}_r (N_A V)^m$ for each reaction r with m reactants. Here, Avogadro's number N_A serves to convert between microscopic measures (counting molecules) and macroscopic kinetic measures (concentrations in molarity) for a volume V. In this paper, we focus on the microscopic kinetic picture, viewing the $\tilde{c}_r$ as the fundamental parameters, but

these parameters can be converted into k_r for linear reactions or in the thermodynamic limit [44]. In that thermodynamic limit, the stochastic kinetics often tends to the solutions of the deterministic reaction-rate equations. For the Schlögl model, analysis of that deterministic limit reveals that the sign of the discriminant,

$$\Delta = 4k_1^3 k_3 A^3 B - k_1^2 k_4^2 A^2 + 4k_2 k_4^3 - 18k_1 k_2 k_3 k_4 AB + 27k_2^2 k_3^2 B^2, \tag{3}$$

determines the number of stable steady-state solutions [42], where the A and B denote the concentrations of species A and B. While holding the rate coefficients fixed, the sign of Δ changes as a function of A and B, allowing one to pass from $\Delta > 0$ with one steady-state solution, through $\Delta = 0$ with a diverging variance of the number of X molecules [45], and on to $\Delta < 0$ with one unstable and two stable solutions. In the $\Delta < 0$ regime, a finite (stochastic) system sees the regions around the two stable fixed points become metastable states. The number of X molecules stochastically switches between the two metastable values. A representative trajectory, shown in Fig. 1, oscillates between a state with $n \approx 5$ X molecules and another with $n \approx 50$. We first set out to compute the rate of bistable switching between those two metastable states in the well-mixed regime, making use of the fact that the dynamics is one dimensional—n executes a random walk between 0 and some large maximum number M, which we impose for the sake of calculation.

Rather than track $n(t)$ for a single trajectory as in Fig. 1, the ensemble approach tracks evolution of the distribution $p_t(n)$, measuring the probability that an individual trajectory would have n molecules at time t. Foreshadowing the analogies with quantum-mechanical methodologies, we choose to write that distribution as a ket: $|p_t\rangle = [p_t(0), p_t(1), \ldots, p_t(M)]^T$, where T indicates the transpose to give a column vector. The distribution evolves according to the master equation:

$$\frac{d|p_t\rangle}{dt} = H|p_t\rangle. \tag{4}$$

Here, H is a rate matrix whose off-diagonal elements $H_{n'n}$ are the rates of transitioning from n X molecules to n' and whose diagonal elements H_{nn} are minus the rate of escaping from the state with n X molecules. The rate matrix conserves probability, a fact compactly expressed as $\langle 1|H = 0\langle 1|$, with $\langle 1|$ denoting a vector of ones. The distribution $|p_t\rangle$ can always be formally evolved in terms of a propagator e^{Ht} as

$$|p_t\rangle = e^{Ht}|p_0\rangle. \tag{5}$$

Because of the analogy with quantum-mechanical time evolution, we equivalently call H an effective
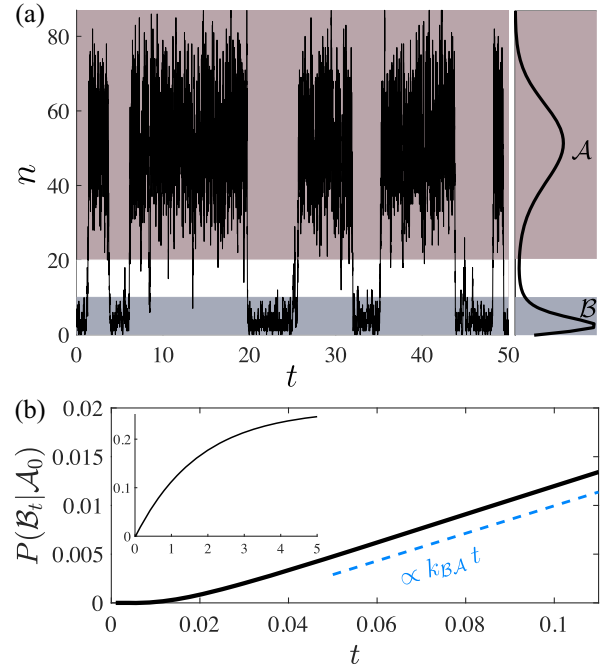

FIG. 1. (a) A typical trajectory in the bistable regime ($\tilde{c}_1 = 36.38$, $\tilde{c}_2 = 0.040$, $\tilde{c}_3 = 2200$, and $\tilde{c}_4 = 37.6$) of the well-mixed Schlögl model of Sec. II A. The system stochastically switches between the two metastable basins, where basin $\mathcal{A}$ is the top colored region and $\mathcal{B}$ is the bottom colored region. (b) Starting with a distribution localized in basin $\mathcal{A}$, the evolution has three timescales. Initially, no trajectories can reach $\mathcal{B}$. After the so-called molecular timescale τ_{mol}, the fastest trajectories begin to reach $\mathcal{B}$. Subsequently, probability leaches into region $\mathcal{B}$ at a constant rate with $P(\mathcal{B}_t|\mathcal{A}_0) \propto k_{\mathcal{B}\mathcal{A}}t$. The inset shows that growth on a much longer timescale, revealing the third scale when the system eventual approaches the steady state.

Hamiltonian. Assuming H is irreducible, the longtime limit for this distribution tends to the unique steady state $|\boldsymbol{\pi}\rangle = \lim_{t\to\infty} e^{Ht}|p_0\rangle$.

H can be decomposed into contributions from the two different reversible mechanisms of Eqs. (1) and (2), labeled H^A and H^B, respectively, based on whether the transitions are mediated by the chemostatted species A or B. Then, the Schlögl model at the level of the master equations corresponds to tridiagonal rate matrices with elements

$$H_{n'n}^A = \frac{\tilde{c}_1 n(n-1)}{2}\delta_{n'-1,n} + \frac{\tilde{c}_2 n(n-1)(n-2)}{6}\delta_{n'+1,n} - \frac{n(n-1)}{2}\left(\tilde{c}_1 + \frac{\tilde{c}_2(n-2)}{3}\right)\delta_{n',n} \tag{6}$$

and

$$H_{n',n}^B = \tilde{c}_3 \delta_{n'-1,n} + \tilde{c}_4 n\delta_{n'+1,n} - (\tilde{c}_3 + \tilde{c}_4 n)\delta_{n',n}, \tag{7}$$

where $H = H^A + H^B$ and $\delta_{i,j}$ is the Kronecker delta. Given the specific tridiagonal H and the restriction $0 \leq n \leq M$,

both H and the propagator e^{Ht} are $(M + 1) \times (M + 1)$ matrices, amenable to evolving the entire distribution, effectively averaging over all the realizations.

To extract a rate from the evolution of distribution, we must initialize the distribution in one of the two metastable states and measure the rate of relaxation. In analogy with free-energy basins, we will call the two metastable regions basins $\mathcal{A}$ and $\mathcal{B}$ (red and blue regions of Fig. 1, respectively) with $\mathcal{A} \cap \mathcal{B} = \emptyset$. We write $\hat{\mathcal{P}}_{\omega}$ for the non-negative idempotent $(\hat{\mathcal{P}}_{\omega}^2 = \hat{\mathcal{P}}_{\omega})$ operator that projects onto the density lying in a region ω, where $\omega \in \{\mathcal{A}, \mathcal{B}\}$. If $|p_t\rangle$ has support over a state space Ω, then the projection $\hat{\mathcal{P}}_{\omega}|p_t\rangle$ is an unnormalized distribution with support restricted to ω. The total probability in ω at time t can then be expressed as $\langle \mathbf{1}|\hat{\mathcal{P}}_{\omega}|p_t\rangle$.

B. Rate constants

In general, one can compute the probability to pass from any subset $\mathcal{A} \subset \Omega$ into an arbitrary nonoverlapping subset $\mathcal{B}$. For the current example, $\mathcal{A}$ and $\mathcal{B}$ are basins of attraction for the Schlögl model [see Fig. 1(a)], and we would like to measure the rate of going from $\mathcal{A} \to \mathcal{B}$. Rate constants can be defined either for the unidirectional flux from $\mathcal{A} \to \mathcal{B}$, the quantity we focus on in this work, or for the timescale of the relaxation to the steady state, which also involves the $\mathcal{B} \to \mathcal{A}$ transitions [46]. Additionally, the precise definition of the rate depends on whether one seeks trajectories that move from the boundary of one basin to another or from the interior of one basin to another. For instance, studying the time to transition from the boundary of $\mathcal{A}$ to the boundary of $\mathcal{B}$ yields the transition path theory rate associated with the committor function [47]. This unidirectional rate $k_{\mathcal{BA}}$ is related to the probability of having reached $\mathcal{B}$ from $\mathcal{A}$ in time t without returning to $\mathcal{A}$. If we focus on trajectories short enough that those reaching $\mathcal{B}$ will stay (no recrossings), then $k_{\mathcal{BA}}$ can be expressed in terms of the probability $P(\mathcal{A}_0)$ to start in $\mathcal{A}$ at time zero and the conditional probability $P(\mathcal{B}_t|\mathcal{A}_0)$ to subsequently be found in $\mathcal{B}$ at time t:

$$k_{\mathcal{BA}} \equiv \frac{d}{dt} P(\mathcal{B}_t|\mathcal{A}_0)P(\mathcal{A}_0)\bigg|_{t > \tau_{\mathrm{mol}}}. \tag{8}$$

The derivative is evaluated at some timescale exceeding τ_{mol}, a molecular timescale of the system that is long enough for the first trajectories to reach $\mathcal{B}$ from $\mathcal{A}$ but much shorter than the typical transit time $k_{\mathcal{BA}}^{-1}$ [see Fig. 1(b)].

We compute the rate of Eq. (8) by initializing in a distribution confined to $\mathcal{A}$ then propagating for time t and measuring the probability in $\mathcal{B}$. A suitable initial distribution is $\hat{\mathcal{P}}_{\mathcal{A}}|\boldsymbol{\pi}\rangle / \langle \mathbf{1}|\hat{\mathcal{P}}_{\mathcal{A}}|\boldsymbol{\pi}\rangle$, a normalized distribution constructed by projecting the steady-state distribution $|\boldsymbol{\pi}\rangle$ onto $\mathcal{A}$. With this initialization, $P(\mathcal{A}_0) = 1$, and Eq. (8) becomes

$$k_{\mathcal{BA}} = \frac{d}{dt} \frac{\langle \mathbf{1}|\hat{\mathcal{P}}_{\mathcal{B}}e^{Ht}\hat{\mathcal{P}}_{\mathcal{A}}|\boldsymbol{\pi}\rangle}{\langle \mathbf{1}|\hat{\mathcal{P}}_{\mathcal{A}}|\boldsymbol{\pi}\rangle}\bigg|_{t > \tau_{\mathrm{mol}}}. \tag{9}$$

To the extent that there is a separation of timescales between exploring a basin and transiting between basins, the transition event is rare and no transitions can occur faster than the molecular timescale τ_{mol}. Beyond that time, the number of transitions accumulates over time as $P(\mathcal{B}_t|\mathcal{A}_0) \propto k_{\mathcal{BA}}t$. At times long compared to $k_{\mathcal{BA}}^{-1}$, $P(\mathcal{B}_t|\mathcal{A}_0)$ must eventually stop growing linearly to level off at the stationary value $\langle \mathbf{1}|\hat{\mathcal{P}}_{\mathcal{B}}|\boldsymbol{\pi}\rangle$, but our interest is in extracting $k_{\mathcal{BA}}$ from the growth at much shorter times. A typical realization of the process will not require observation much beyond τ_{mol} units of time, to ascertain the rate, meaning our ensemble evolutions must be propagated for roughly τ_{mol}. Notice from the inset of Fig. 1(b) that at long times the recrossing events obscure the rate of growth, ultimately yielding the plateau. As discussed in Sec. V and Appendix D, for rare rates and long times, the rate calculations can be improved by removing these recrossing events via a modified dynamics in which region $\mathcal{B}$ is an absorbing state, forbidding dynamics out of that region.

Our ability to compute $k_{\mathcal{BA}}$ from the ensemble evolution traces back to the fact that it is practical to numerically calculate e^{Ht} for the well-mixed model with a single species and a modest M. While Eq. (9) generalizes naturally for harder problems (those with more than one species or with diffusion between multiple well-mixed voxels), the state space for $|p_t\rangle$ grows exponentially with additional stochastic degrees of freedom. We next show that the Doi-Peliti framework reexpresses the rate matrix in a second-quantized form that allows us to generalize beyond the case of a single well-mixed species.

III. DOI-PELITI

The Doi-Peliti formalism is a classical version of the second-quantization methods from quantum field theory [48,49]. Originally introduced by Doi [50] as a critique to the assumptions of the Smoluchowski equation, it was later developed by Peliti [51] and others [52] as a path integral formalism for solving chemical reaction networks [53–56]. The DP framework has also found success in other areas such as studying kinetically constrained lattice models in the context of glassy dynamics [57,58], cellular signaling [59], predator prey models [60], and active matter [61]. The DP representation builds operators which encode the dynamics onto the high-dimensional space that accompanies many-body problems. Crucially, writing the chemical reactions in an operator form guarantees it can be easily converted into a matrix product operator (MPO) [62,63], as we will do in Sec. IV.

To start, we go beyond Sec. II by considering well-mixed reactions with multiple dynamic species as well as heterogeneous systems where molecules can diffuse between

neighboring voxels. Let $X_1, X_2, ..., X_L$ represent each chemical species in each site. We could be talking about a single species X with different values in L different voxels, a single well-mixed voxel with L distinct dynamic species, or a mixture with multiple species and multiple voxels. The number of molecules in a given voxel can change due to a chemical reaction or a diffusion event, with both events being cast as

$$\sum_j \eta_j^r X_j \xrightarrow{\tilde{c}_r} \sum_j \mu_j^r X_j, \quad (10)$$
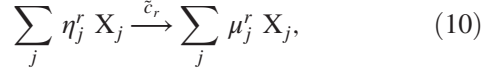
where $\tilde{c}_r$ are again the rate coefficients for reaction r. η_j^r and μ_j^r reflect the stochiometry of the reactions and give the stochiometric vector $\nu^r = \mu^r - \eta^r$ which evolves the system from state $\mathbf{n} \in \Omega$ to $\mathbf{n}' = \mathbf{n} + \nu^r$. In the previous section, these reactions induced changes between the $M + 1$ different microstates, specified by the scalar n. Now, the microstates $\mathbf{n}$ are specified by $\mathbf{n} = [n_1, n_2, ..., n_L]$, counting the occupation of each species within each voxel.

For a dilute well-mixed solution, where well mixed means molecules' locations within a voxel are unresolved, the change in probability of microstate $\mathbf{n}$ follows the gain-loss equation [14],

$$\frac{dp_t(\mathbf{n})}{dt} = \sum_r^N [\alpha_{\mathbf{n}-\nu^r}^r p_t(\mathbf{n} - \nu^r) - \alpha_{\mathbf{n}}^r p_t(\mathbf{n})], \quad (11)$$

where the sum is taken over N different reactions of the form Eq. (10), each with their own stochiometric vector ν^r. The transition rate between any two states is the sum of contributions from each reaction $H_{\mathbf{n}'\mathbf{n}} = \sum_r \alpha_{\mathbf{n}}^r$. Each propensity α^r is a product of a stochastic rate coefficient $\tilde{c}_r$ and the combinatorial number of ways to select either η_j or μ^r molecules from reaction r, depending on the direction of the reaction. For a "forward" reaction the propensity is $\alpha_{\mathbf{n}}^r = \tilde{c}_r \prod_r \binom{n_j}{\eta_j}$, where the product includes a term for each η_j participating in reaction r.

Encoding the reactions into a matrix form allows us to evolve the elements of the joint distribution $p_t(\mathbf{n})$ analogously to the previous section. While it is appealing to try to directly solve these more complicated dynamics as we did before, it is impractical. The curse of dimensionality demands a less onerous way to work with both H and $p_t(\mathbf{n})$. Since particles of the same species in the same voxel are indistinguishable, it is beneficial to work in an occupation number representation with

$$|p_t\rangle = \sum_{\mathbf{n}} p_t(\mathbf{n})|\mathbf{n}\rangle. \quad (12)$$

The basis vectors,

$$|\mathbf{n}\rangle = |n_1, n_2, ..., n_L\rangle$$
$$= |n_1\rangle \otimes |n_2\rangle \otimes \cdots \otimes |n_L\rangle, \quad (13)$$

correspond to a possible microstate, and the set of basis vectors forms a tensor product space called a Fock space $\mathcal{F}^L$ with $|\mathbf{n}\rangle \in \mathcal{F}^L$ and $\langle\mathbf{m}|\mathbf{n}\rangle = \delta_{\mathbf{mn}}$ [64]. Without loss of generality, we will discuss the case of a single chemical species and L physical voxels, in which case n_l has the interpretation of the number of molecules in voxel l. The second-quantization procedure rewrites the rate matrix as an operator $\hat{H}$ in terms of creation and annihilation operators, $x_l^\dagger$ and x_l, that respectively raise and lower the number of X molecules in voxel l. That is to say,

$$x_l^\dagger |n_l\rangle = |n_l + 1\rangle,$$
$$x_l |n_l\rangle = n_l |n_l - 1\rangle,$$
$$x_l |0\rangle = 0, \quad (14)$$

where $|0\rangle$ is the vacuum state, i.e., no molecules. By x_l or $x_l^\dagger$ we are implying that the operator only acts [on the lth voxel]. As shorthand one can use the same symbol x_l to act on a many-body state, in which case the symbol must be interpreted as:

$$x_l = \mathbb{I}_1 \otimes \mathbb{I}_2 \otimes \cdots \otimes \mathbb{I}_{l-1} \otimes x \otimes \mathbb{I}_{l+1} \otimes \cdots \otimes \mathbb{I}_L, \quad (15)$$

where $\mathbb{I}_1$ is the identity operator acting on voxel k. Because they act on the single-site space, the general one-site operators x and $x^\dagger$ can be written in a matrix form as

$$x^\dagger = \begin{bmatrix} 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \\ \vdots & \vdots & & \ddots \end{bmatrix} \quad \text{and} \quad x = \begin{bmatrix} 0 & 1 & 0 & \cdots \\ 0 & 0 & 2 & \cdots \\ 0 & 0 & 0 & \\ \vdots & \vdots & & \ddots \end{bmatrix}. \quad (16)$$

That these matrices yield Eq. (14) can be readily confirmed. More abstractly, the existence of such creation and annihilation operators in this classical setting originates from the classical commutation relation $[x_i, x_j^\dagger] = \delta_{ij}$ and the existence of a positive semidefinite number operator $\hat{n}_l = x_l^\dagger x_l$, so named because $\hat{n}_l |n_l\rangle = n_l |n_l\rangle$. Together, the set $\{\Omega, \mathcal{F}^L, x, x^\dagger\}$ are also sometimes called an interacting Fock space [48]. As in the construction of the number operator, we will be able to decompose contributions to $\hat{H}$ in terms of the $x^\dagger$ and x's.

In practice, the matrices of Eq. (16) will not be infinite in size but will rather be $(M + 1) \times (M + 1)$ matrices due to our enforcement of a maximum voxel occupancy of M. At the operator level, this maximum voxel occupancy supplements Eq. (14) with an additional requirement that

$$x_l^\dagger |M_l\rangle = 0, \quad (17)$$

a requirement that is met by the finite truncation of Eq. (16). Those truncated operators obey a commutation relation which is similar to that of their infinite counterpart, but the
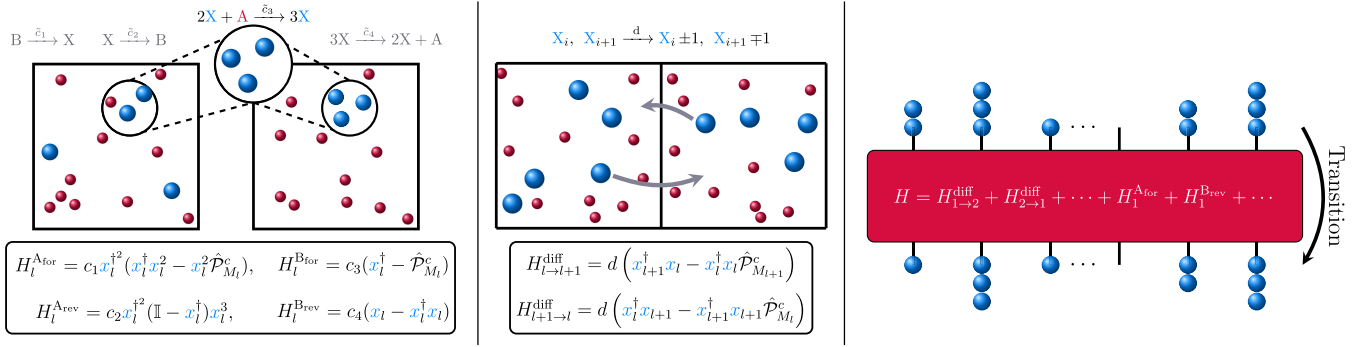
FIG. 2.   The reaction-diffusion master equation (RDME) consists of voxels in which well-mixed chemical reactions occur and between which species can diffuse. This work centers around a one-dimensional chain of L such voxels, each with a fluctuating number of molecules of species X in a Schlögl chemical reaction. The Doi-Peliti (DP) framework turns the reaction-diffusion evolution into a second-quantized effective Hamiltonian that governs the evolution of the distribution over classical reaction-diffusion microstates. That effective Hamiltonian is represented by a high-dimensional tensor, with each vertical leg taking one of $(M + 1)$ possible values corresponding to the number of molecules of species X which could occupy each site. The DP form of the effective Hamiltonian lends itself to a compact representation of the effective Hamiltonian tensor of rates for transitioning from one microstate to another.

usual commutation relation is broken at the maximal occupancy state. That is to say,

$$(x_l x_l^\dagger - x_l^\dagger x_l)|M_l\rangle = -M_l|M_l\rangle \neq |M_l\rangle. \qquad (18)$$

The commutation relation for the truncated operators is therefore

$$[x_i, x_j^\dagger] = \delta_{ij}(1 - (M + 1)|M_i\rangle\langle M_i|), \qquad (19)$$

an expression involving the projector operator onto site i's maximal occupancy state, $\hat{\mathcal{P}}_{M_i} = |M_i\rangle\langle M_i|$. As discussed in depth in Appendix A, this violation of the usual commutator relation must be accounted for when constructing a finite effective Hamiltonian that conserves probability.

One benefit of defining operators in terms of creation and annihilation operators is that expectation values for observables can be expressed easily in terms of inner products with operators. Consider, for example, the observable that counts all particles in all voxels, $\hat{n} = \sum_l \hat{n}_l$. The operator $\hat{n}$ acts on a many-body state $|\mathbf{n}\rangle$, yet each $\hat{n}_l$ is a single-body operator acting on voxel l and leaving all other voxels unchanged. It is therefore understood that when local operators such as $\hat{n}_l$ act on $|\mathbf{n}\rangle$, it is written as a shorthand for

$$\hat{n}_l = \mathbb{I}_1 \otimes \cdots \otimes x_l^\dagger x_l \otimes \cdots \otimes \mathbb{I}_L. \qquad (20)$$

In terms of that shorthand, $\hat{n}_l|\mathbf{n}\rangle = n_l|\mathbf{n}\rangle$ counts the number of particles in voxel l for the many-body state $|\mathbf{n}\rangle$. The time-dependent expected number of particles in the whole system is thus $\langle n(t)\rangle = \sum_l \langle \mathbf{1}|\hat{n}_l|p_t\rangle$, where the vector of ones $\langle \mathbf{1}|$ serves to sum over all microstates. The time dependence of a distribution can also be extracted from the action of a second-quantized $\hat{H}$ on $|p_t\rangle$, and this

operation can be practically computed when $\hat{H}$ is expressed in terms of the single-site operators.

Mirroring the discussion in Sec. II, the generator $\hat{H}_l$ for the Schlögl model reactions within the lth voxel is a sum of contributions from the elementary reactions $\hat{H}_l = \sum_r \hat{H}_l^r$. Figure 2(a) shows each reaction, expressed as a second-quantized contribution to the effective Hamiltonian, as derived in Appendix A. Probability conservation requires that this effective Hamiltonian include a term with a projector $\hat{\mathcal{P}}_{M_l}^c \equiv \mathbb{I} - \hat{\mathcal{P}}_{M_l}$. This projector is complementary to the one appearing in Eq. (19) in that it projects onto states which are *not* maximally occupied. Combining together the forward and reverse reactions mediated by species A, we show in Appendix A that the reversible reaction Eq. (1) contributes

$$\hat{H}_l^A = c_1[x_l^{\dagger 3} x_l^2 - x_l^{\dagger 2} x_l^2 \hat{\mathcal{P}}_{M_l}^c] + c_2[x_l^{\dagger 2} x_l^3 - x_l^{\dagger 3} x_l^3] \qquad (21)$$

to the effective Hamiltonian. For convenience we have absorbed the combinatorial terms into that rate coefficient, defining $c_r \equiv \tilde{c}_r/\gamma!$, where γ is either η^r or μ^r for forward and reversed reactions, respectively. Likewise, the contribution

$$\hat{H}_l^B = c_3(x_l^\dagger - \hat{\mathcal{P}}_{M_l}^c) + c_4(x_l - x_l^\dagger x_l) \qquad (22)$$

comes from Eq. (2). In the $M \to \infty$ limit, $\hat{\mathcal{P}}_{M_l}^c \to \mathbb{I}$ to give the more familiar expressions as in Table I.

The Hamiltonian operator for L independent, well-mixed Schlögl voxels is thus $\hat{H}^{\text{rxn}} = \sum_{l=1}^{L} \hat{H}_l^A + \hat{H}_l^B$. To add diffusion, consider the well-mixed voxels arranged in a 1D lattice as in Fig. 2(b). Each chemical species is able to move between neighboring voxels as regulated by a diffusion operator,
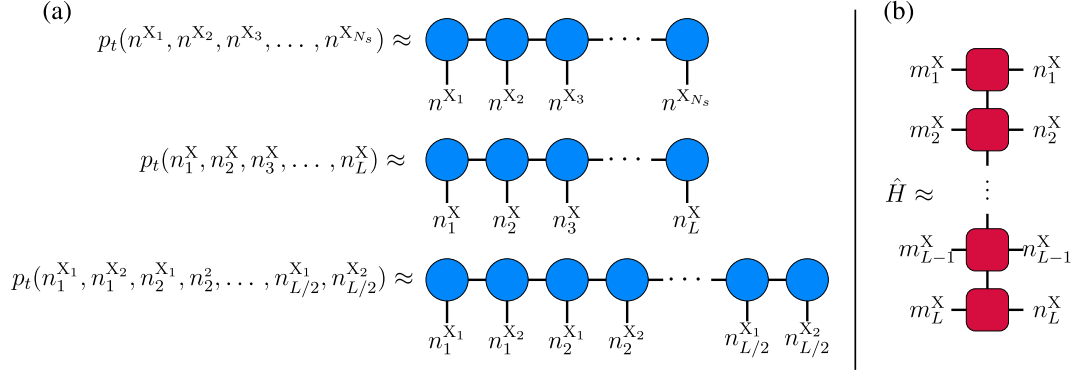
(a)

(b)

FIG. 3.   The ensemble evolution of the many-body state requires that one solve for the probability of every microstate at all times. That high-dimensional joint probability distribution can be well approximated by a matrix product state (MPS). An MPS approximation for the joint distribution can be advantageous whether there are N_s species in a well-mixed reaction [(a), top], one dynamical (unchemostatted) species with diffusion [(a), middle], or multiple species and diffusion [(a), bottom]. In all three cases, each circle represents a tensor, either rank 2 at the first and last site or rank 3 in the body. Horizontal lines represent the bond indices connecting tensors and vertical lines are physical indices. For an RDME, the physical indices represent the possible number of molecules at a given tensor. If two lines are connected, then the index is summed over. (b) An effective Hamiltonian H can be likewise decomposed into lower-rank tensors to form a matrix product operator (MPO). Now there are two physical indices per tensor, representing transitions from state n_l^X to m_l^X. Vertical connected lines are the bond indexes which are contracted over.

$$\hat{H}^{\text{diff}} = \sum_{l=1}^{L-1} d(x_{l+1}^\dagger x_l - x_l^\dagger x_l \hat{\mathcal{P}}_{M_{l+1}}^c)$$
$$+ d(x_l^\dagger x_{l+1} - x_{l+1}^\dagger x_{l+1} \hat{\mathcal{P}}_{M_l}^c), \qquad (23)$$

with hopping rate $d = D/h^2$, consistent with macroscopic diffusion constant D for voxel width h [18]. Combining reaction with diffusion gives

$$\hat{H} = \hat{H}^{\text{rxn}} + \hat{H}^{\text{diff}}, \qquad (24)$$

a straightforward linear superposition at the level of the effective Hamiltonian. We emphasize that the linear superposition of diffusion $\hat{H}^{\text{diff}}$ with reactions $\hat{H}_l^A$ and $\hat{H}_l^B$ does not imply a restriction to linear kinetics. The Schlögl model kinetics is nonlinear in that it involves elementary reactions which are not first order, and the diffusive coupling between neighbors can induce nontrivial interactions between voxels.

IV. TENSOR NETWORKS

We have now seen how to transform the matrix-form rate operator H of Sec. II into a Doi-Peliti form $\hat{H}$, expressed in terms of local creation and annihilation operators. That change of representation is not merely cosmetic. As the number of chemical species or number of lattice sites grows, the size of H grows exponentially. For N dynamical (unchemostatted) species and L lattice sites, the state space has $(M+1)^{NL}$ states. To put this scaling in perspective, the largest numerical result that follows uses $N = 1$, $M = 85$, and $L = 8$, meaning the system contains 2.99×10^{15} microstates. Calculations on a matrix H, even a very sparse matrix H, are untenable. By contrast, it is practical to construct a TN decomposition of the second-quantized

form of $\hat{H}$ in terms of an MPO, as shown in Fig. 3(b), and demonstrated explicitly for the Schlögl model in Appendix B. Given the second-quantized form, the MPO can be readily constructed in a practical form through singular value decomposition (SVD) decomposition using automated packages such as AutoMPO in ITensor [65].

Much like the operator $\hat{H}$ can be compressed as an MPO, the states that operators act upon can also be expressed more compactly in terms of TNs. Consider $p_t(\mathbf{n}) = p_t(n_1, n_2, \ldots, n_L)$, the probability of microstate $\mathbf{n}$ at time t. This $p_t(\mathbf{n})$ can be thought of as a rank L tensor. By specifying the occupancy of each of the L sites, $p_t(\mathbf{n})$ outputs a number—a probability—associated to that microstate. By leveraging a matrix product ansatz, $p_t(\mathbf{n})$ can be approximated as a partial contraction over L different low-rank tensors [see Fig. 3(a)]. The low-rank tensors of an MPS [66] are organized in a 1D chain of site-dependent tensors $Q_{s_1}^{n_1}, Q_{s_1 s_2}^{n_2}, \ldots, Q_{s_{L-2} s_{L-1}}^{n_{L-1}}$, and $Q_{s_{L-1}}^{n_L}$ whose upper index n_l specifies the occupancy of site l and whose lower indices are so-called bond indices, nonphysical indices which are contracted over to generate correlations between nearby sites. For compactness of notation, we suppress labels $Q^{(1)}$, $Q^{(2)}$, etc. that would distinguish the site-dependent tensors, implicitly carrying that site dependence through the labeling scheme for the physical indices. For suitably chosen Q tensors, we therefore approximate

$$|p_t\rangle = \sum_{n_1, n_2, \ldots, n_L} p_t(n_1, n_2, \ldots, n_L)|n_1, n_2, \ldots, n_L\rangle$$
$$\approx \sum_{\substack{n_1, n_2, \ldots, n_L \\ s_1, s_2, \ldots, s_{L-1}}} Q_{s_1}^{n_1} Q_{s_1 s_2}^{n_2} \cdots Q_{s_{L-1}}^{n_L}|n_1, n_2, \ldots, n_L\rangle$$
$$\equiv |q_t\rangle, \qquad (25)$$

where $|q_t\rangle$ is the approximate distribution over microstates generated from the MPS. Though not explicitly written, the Q tensors are time dependent.

There is a diagrammatic representation of the tensors, shown in Fig. 3, which aids in visualizing tensor operations. The tensors Q are given by circles, one per site, and vertical lines feeding into these circles are the physical indices that pick out a particular slice of the tensor based on the occupation number of that site. Circles also have horizontal lines emanating from them, representing the nonphysical bond indices. Connection of two circles by a line represents a shared index being contracted over. Figure 3 emphasizes that the TN decomposition is flexible enough to address varied CMEs. The MPS of Fig. 3(a), middle, has been the focus. Each site from 1 to L reflects the number of X molecules in that site, but Fig. 3(a) (top) and (bottom) show that the MPS decomposition can equally well be constructed when multiple species X_i, $i = 1, 2, ..., N_s$ can occupy a single site or when N_s different species diffuse between L/N_s sites to yield an MPS with L sites. What is practically important is that Q's be arranged such that the physical dynamics correlates the neighbors. For example, for the RDME Schlögl model, neighboring Q's correspond to neighboring voxels in the 1D Schlögl lattice.

An attractive feature of the MPS ansatz is that the size and accuracy of the MPS are controllable through the bond dimension of the auxiliary indices s. If the dimension of those indices is allowed to grow exponentially from s_1 through $s_{L/2}$, the ansatz in Eq. (25) is exact, but even when the bond dimension of the indices s are capped at a maximum dimension of κ, the ansatz can be a very good approximation. Crucially, capping that maximum bond dimension makes it practical to work with $|q_t\rangle$ because the rank-L tensor $|p_t\rangle$ has been replaced by the L different Q tensors, each with no more than $(M + 1)\kappa^2$ elements.

Having approximated both $|q_t\rangle$ and $\hat{H}$ by an MPS and MPO, respectively, we can revisit the otherwise intractable dynamics problem of Eq. (4), only now we wish to evolve $|q_t\rangle$ in lieu of $|p_t\rangle$. As $|q_t\rangle$ is an MPS state, we wish for the dynamics to be constrained for all times to the manifold $\mathcal{M}$ of possible MPS states. Merely replacing $|p_t\rangle$ by $|q_t\rangle$ in Eq. (4),

$$\frac{d|q_t\rangle}{dt} = \hat{H}|q_t\rangle, \tag{26}$$

would not impose such a constraint because $\hat{H}$ in general evolves $|q_t\rangle$ off $\mathcal{M}$. An appealing resolution is given by the time-dependent variational principle [67,68], which proposes to first project the right-hand side onto the tangent space of the MPS manifold:

$$\frac{d|q_t\rangle}{dt} = \hat{\mathcal{P}}_{T_\mathcal{M}} \hat{H}|q_t\rangle. \tag{27}$$

In this way, dynamics is restricted to the desired subspace of MPS states [69].

It turns out that the time-discretized numerical integration of this $|q_t\rangle$ is amenable to efficient TN operations. The technical details of the time-evolution algorithm have been reported elsewhere [68,70–72]. We give only a brief, high-level overview of that procedure (see also Appendix C for more details). The central step in deriving the TDVP time-evolution algorithm is to demonstrate that the projector $\hat{\mathcal{P}}_{T_\mathcal{M}}$ can be decomposed in terms of a sum over L terms (associated to each site of the MPS) and $L - 1$ terms (associated to each bond that connects MPS sites). In the limit of a small time step δt, the propagator $e^{\hat{\mathcal{P}}_{T_\mathcal{M}} \hat{H} \delta t}$ can thus be evaluated as a composition of time evolutions involving the $2L - 1$ individual terms. That composition of time evolutions is executed by a one-site TDVP algorithm that "sweeps" from lattice site 1 through L, advancing each term for the small time step.

The efficiency of the TDVP time-evolution algorithm relies on the flexibility to represent the same MPS state in redundant ways due to a gauge freedom. Consider, for example, neighboring tensors $Q_{s_{l-1}s_l}^{n_l}$ and $Q_{s_l s_{l+1}}^{n_{l+1}}$ in an MPS. For an invertible R, replacing those tensors by $Q^{n_l}R$ and $R^{-1}Q^{n_{l+1}}$, respectively, does not change the MPS state; upon contracting over the neighboring sites, the identity $\mathbb{I} = RR^{-1}$ would fall out. By leveraging the gauge freedoms, it is always possible to transform Eq. (25) into a so-called mixed canonical form [73]:

$$|q_t\rangle = \sum_{n_1, ..., n_L} A^{n_1} \cdots A^{n_{l-1}} Q^{n_l} B^{n_{l+1}} \cdots B^{n_L} |n_1, ..., n_L\rangle, \tag{28}$$

where each A is a left-orthogonal tensor satisfying $A^T A = \mathbb{I}$ and each B is a right-orthogonal tensor satisfying $BB^T = \mathbb{I}$. For notational compactness, we have suppressed implied summations over the bond indices between tensors in Eq. (28). We have retained the physical indices $n_1, ..., n_L$ and allow their presence to remind us that although tensors at each site are named merely A and B, they are actually distinct tensors $A^{(1)}, A^{(2)}, ..., B^{(1)}, B^{(2)}, \cdots$ that vary across sites. The mixed canonical form is said to be centered about the tensor Q^{n_l}. From that tensor's perspective, the chain of A's and chain of B's constitute the environment, and it is beneficial to contract over the links between A's and links between B's to obtain so-called environment tensors. These environment tensors inherit an orthogonality from the orthogonality of the A's and B's, significantly reducing the computational complexity to advance Q^{n_l} in time by δt provided the MPS was already centered about the Q^{n_l} we aim to evolve.

As made more explicit in Appendix C, the one-site TDVP algorithm initially centers the MPS around the first tensor, propagates that tensor forward in time by δt, changes the gauge by factorizing the tensor into a product of two tensors: QR, propagates the R term of that so-called QR
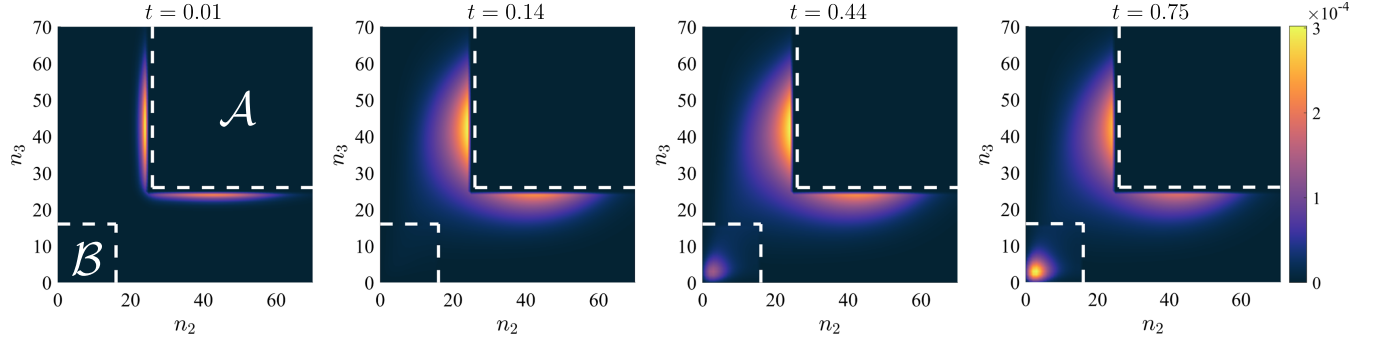
FIG. 4. Snapshots of the joint density in two neighboring voxels from an $L = 3$ site lattice of Schlögl voxels. Using $\tilde{c}_1 = 2.676$, $\tilde{c}_2 = 0.040$, $\tilde{c}_3 = 108.102$, $\tilde{c}_4 = 37.881$, and $d = 8.2207$, the steady-state distribution was computed with $\kappa = 30$ by evolving an initial uniform distribution according to $\hat{H}$ via TDVP. That steady-state distribution was projected into $\mathcal{A}$ at time 0, renormalized, then propagated according to the modified no-recrossing dynamics $\tilde{H}$ using TDVP. For clarity of visualization, the density within $\mathcal{A}$ is not plotted as it would overwhelm the small probability that leaks into $\mathcal{B}$. By $t = 0.01$ (left-hand image), probability has begun to leak from $\mathcal{A}$ but none has reached $\mathcal{B}$. Probability has just begun to reach $\mathcal{B}$ at $t = 0.14$, roughly the τ_{mol} timescale for this problem. By $t = 0.44$, probability is collecting in $\mathcal{B}$, and by $t = 0.75$, the flux passing from $\mathcal{A}$ to $\mathcal{B}$ has attained an approximately constant steady-state value corresponding to the linear growth phase where $P(\mathcal{B}_t|\mathcal{A}_0) \propto k_{\mathcal{B}\mathcal{A}}t$.

decomposition backward in time by δt, then passes that term to the neighboring site to recenter the MPS around the neighboring site. The neighboring site is then propagated forward in time as the algorithm iterates. After the complete sweep through the MPS, the probability distribution $|q_t\rangle$ has been propagated by a time step δt. Whether that propagation well approximates the evolution of the full distribution $|p_t\rangle$ depends on the time-step error but also on the severity of the MPS approximation, which can be systematically controlled through the maximum bond dimension κ.

Notably, the time-evolved MPS state $|q_t\rangle$ allows us to estimate the probability of *any* microstate of interest, even though it would be impossible to enumerate probabilities of all microstates. Alternatively, sets of microstates can be summed over to yield tractable marginal distributions. As an example, Fig. 4 shows the results of calculating $q_t(n_1, n_2, n_3)$ for a 1D chain of $L = 3$ Schlögl sites. Though there are three sites, site one is traced over to leave the two-dimensional joint probability of finding n_2 X molecules in site 2 and n_3 X molecules in site 3. At time zero, the distribution is initialized within region $\mathcal{A}$ according to $\hat{\mathcal{P}}_{\mathcal{A}}|\boldsymbol{\pi}\rangle / \langle \mathbf{1}|\hat{\mathcal{P}}_{\mathcal{A}}|\boldsymbol{\pi}\rangle$. As time progresses, a small flux of probability leaks from $\mathcal{A}$ to $\mathcal{B}$. Because the overwhelming preference is to stay in $\mathcal{A}$, Fig. 4 highlights the rare transitions by only plotting the joint distribution outside of $\mathcal{A}$. Notice that, consistent with a steady-state approximation, while the population in $\mathcal{B}$ gradually amasses, the distribution over microstates between $\mathcal{A}$ and $\mathcal{B}$ is nearly stationary. The slight asymmetry between sites is due to the fact that site 2 is in the bulk while site 3 lies on the boundary. The higher-dimensional joint $q_t(n_1, n_2, n_3)$ has the anticipated symmetry under interchange of n_1 and n_3. More generally, symmetry holds between interchanging sites l and $L - l + 1$, as reflected in Fig. 6. In the next section, we illustrate the fact that the MPS state's ability to

capture the rare time-dependent flux is key to calculating the rate of traversing between basins.

V. MANY-BODY RATE CALCULATION

Having established the TDVP dynamics of an MPS state, we can now repeat the strategy of Eq. (9) to compute a transition rate between states $\mathcal{A}$ and $\mathcal{B}$:

$$k_{\mathcal{B}\mathcal{A}} = \frac{d}{dt} \frac{\langle \mathbf{1}|\hat{\mathcal{P}}_{\mathcal{B}} e^{\hat{\mathcal{P}}_{T_{\mathcal{M}}} \hat{H} t} \hat{\mathcal{P}}_{\mathcal{A}}|\boldsymbol{\pi}\rangle}{\langle \mathbf{1}|\hat{\mathcal{P}}_{\mathcal{A}}|\boldsymbol{\pi}\rangle}\Bigg|_{t > \tau_{\text{mol}}}. \quad (29)$$

In doing so, there are a few important differences from Eq. (29). Most obviously, the TDVP dynamics includes the projection onto the MPS tangent space and $|\boldsymbol{\pi}\rangle$ must now be an MPS approximation to the steady-state distribution. Additionally, the meaning of $\hat{\mathcal{P}}_{\mathcal{A}}$ and $\hat{\mathcal{P}}_{\mathcal{B}}$ must be adapted to now define many-body analogs of the projectors onto the two metastable states. The single-voxel Schlögl model had an $\mathcal{A}$ state defined as all microstates with occupancy n above a threshold λ_A^*. Similarly, $\mathcal{B}$ included all microstates with occupancy less than some other threshold λ_B^*. The L-voxel Schlögl model generalizes to involve hypercubic metastable states with basins defined by $\mathcal{A} \equiv \{n_l \geq \lambda_A^* \ \forall \ l\}$ and $\mathcal{B} \equiv \{n_l \leq \lambda_B^* \ \forall \ l\}$.

With these generalizations, the transition rate of Eq. (29) would measure the flux crossing into $\mathcal{B}$ through its boundary, but if we are to compare with a transition path theory rate, we must exclude recrossing events. In other words, we want to include trajectories which exit $\mathcal{A}$ then some time later enter $\mathcal{B}$, but we want to exclude the flux that leaves $\mathcal{B}$ and reenters $\mathcal{B}$ some time later without having returned to $\mathcal{A}$. When $k_{\mathcal{B}\mathcal{A}}$ is sufficiently large, those recrossings can be ignored, but when $k_{\mathcal{B}\mathcal{A}}$ becomes small enough, those recrossings can become significant and must

be removed. We note that the TDVP dynamics can just as well be executed with a modified absorbing-boundary-condition dynamics with effective Hamiltonian

$$\tilde{H} = \tilde{H}^{\text{rxn}} + \tilde{H}^{\text{diff}}, \qquad (30)$$

where the reaction and diffusion Hamiltonians exclude events that exit $\mathcal{B}$. An explicit construction of that modified absorbing-boundary dynamics is provided in Appendix D, allowing us to directly remove the recrossing events.

VI. NUMERICAL EVALUATION OF THE *L*-VOXEL SCHLÖGL MODEL SWITCHING RATE

The formalism of the preceding sections has laid out a controllable approximation to extract rates between metastable states of RDMEs. Numerical experiments are necessary to evaluate whether that formalism can be practically useful. We set out to demonstrate that utility by computing the transition rates between high-occupancy and low-occupancy states of an L-voxel Schlögl model as a function of L. The model is parametrized by the four stochastic rate coefficients discussed in Sec. II. We sought rate coefficients that fall in a bistable regime, with sharp peaks of steady-state probability falling in high- and low-occupancy states. Since the dynamics would need to be truncated at some maximum occupancy M, we furthermore wanted to limit how many molecules would naturally accrue in the high-occupancy state. By scanning parameters with the $L = 1$ Schlögl model, we found parameters ($\tilde{c}_1 = 2.676$, $\tilde{c}_2 = 0.040$, $\tilde{c}_3 = 108.102$, and $\tilde{c}_4 = 37.881$) compatable with an $M = 85$ truncation and with a sharp bistability between $\mathcal{A}$ and $\mathcal{B}$ regions defined by the thresholds $\lambda_A^* = 15$ and $\lambda_B^* = 25$. By linking multiple voxels together in a 1D chain with a diffusive hopping rate d, it becomes rarer to switch between the two metastable states than it would be with a single isolated voxel. With multiple voxels, a switching event must flip the state of each voxel, and diffusion between neighboring voxels works to stabilize neighbors in identical high- or low-occupancy states. Because of that diffusive coupling between neighbors, the L-voxel Schlögl model's $\mathcal{A}$ to $\mathcal{B}$ transition is characterized by a correlated flip of all voxels, and the rarity of that correlated event depends on L, the number of correlated voxels. It is therefore straightforward for us to push the rate-calculation methodologies into challenging regimes; $k_{\mathcal{B}\mathcal{A}}$ can be decreased by growing L.

The most straightforward way to numerically detect the L dependence of the transition rate is via sampling of trajectories. Realizations of the continuous-time RDME, subject to the $M = 85$ maximum voxel occupancy at fixed L, were generated using the SSA. First, 5×10^5 independent SSA realizations were initialized from a uniform distribution and propagated in parallel up to $t = 5$ to approach the steady-state distribution. Sampled configurations that fell within $\mathcal{A}$ were then sampled as initial configurations for 400 independent trajectories which were evolved via the SSA until the trajectory first reached $\mathcal{B}$. From the 400 trajectories, the transition rate was computed from the mean of those waiting times to reach $\mathcal{B}$: $k_{\mathcal{B}\mathcal{A}} = \langle \tau_{\mathcal{B}\mathcal{A}} \rangle^{-1}$. Figure 5(a) (red line) shows that the rate exponentially decreases as L is increased, a result consistent with the exponential dependence of rate on volume in the well-mixed Schlögl model [74]. While the rate can be computed by trajectory sampling, the difficulty to do so (via brute-force sampling) grows in proportion to the typical $\tau_{\mathcal{B}\mathcal{A}}$, meaning the computational cost grows exponentially with L [see Fig. 5(c)].

The poor scaling of brute-force SSA sampled rate calculations has motivated so-called advanced sampling methods like FFS [10], which focus computational effort on the progress along a reaction coordinate. The idea is to introduce a reaction coordinate λ that resolves progress in transitioning from a value λ_A^* at the boundary of $\mathcal{A}$ to λ_B^* at
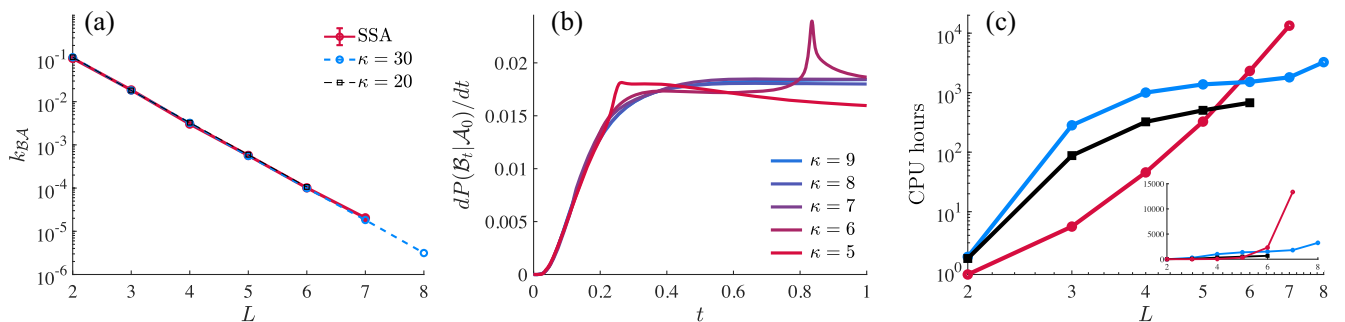

FIG. 5.   (a) The rate $k_{\mathcal{B}\mathcal{A}}$ for switching between metastable states of the L-voxel Schlögl model is calculated using the SSA (red line) and using TDVP (dashed lines) with varying bond dimensions κ using the same rate parameters as Fig. 4. For sufficiently large bond dimensions, TDVP agrees with an SSA rate estimate built using 400 transitions from $\mathcal{A}$ to $\mathcal{B}$ at each L. (b) As L grows a larger bond dimension is required to capture the dynamics. Signatures of numerical instability are apparent when κ is too small, shown here for $L = 3$, which stably converges when κ exceeds 6. For this system, $\kappa = 20$ is sufficient for numerical stability up to $L = 6$, while $\kappa = 30$ gives stable solutions up to $L = 8$. (c) The total number of CPU hours to estimate $k_{\mathcal{B}\mathcal{A}}$ grows exponentially with SSA but polynomially with TDVP.

the boundary of $\mathcal{B}$. Rather than seek long trajectories that snake from λ_A^* all the way to λ_B^*, it can be beneficial to break up the transition into smaller trajectories that make some progress in moving to states with values of λ closer to λ_B^* than they started. This FFS scheme is particularly clever in that it can be statistically unbiased, meaning any choice of reaction coordinate (any mapping taking a microstate to a value of λ) will yield the true rate in the infinite sampling limit. Despite that formal lack of bias, the computational cost of an FFS calculation depends strongly on the correlation between λ and the reaction's progress [11]. If one already has a good idea of how the reaction's rare transition events proceed, a smart λ can be constructed and the exponential scaling of the SSA can be avoided, but that identification of the reaction coordinate is seldom straightforward. For example, in the L-voxel Schlögl model, will $\mathcal{A}$ to $\mathcal{B}$ events typically be triggered by a high- to low-occupancy flip at a boundary voxel of the 1D chain or in the bulk? The boundary voxels have the benefit that they only have diffusion from one neighbor stabilizing the high-occupancy state, but the bulk voxels have the advantage of being more plentiful. In Appendix G we describe FFS calculations employing two reasonable choices of reaction coordinate. Those existing FFS methods can yield rates more cheaply than a brute-force SSA calculation, but the speedup depends strongly on the reaction coordinate, the choice of which is nontrivial even for this simple model.

The DP-TDVP approach we have described offers a radically different strategy for circumventing the SSA scaling problem. Because the reaction-coordinate selection problem is hard, we developed the DP-TDVP strategy to bypass it entirely, computing the rate and reaction mechanism without prior information about how the transition events move through the high-dimensional state space. Figure 5 shows that the reaction-coordinate-agnostic DP-TDVP approach indeed reproduces the SSA rates at lower computational expense, reducing the scaling from exponential to polynomial in L. The SSA rate calculations can be made more accurate by collecting more trajectories, and the DP-TDVP calculations likewise offer a systematic way to improve accuracy in exchange for a greater computational expense. The two most prominent ways to adjust that accuracy are through the time step δt for the TDVP evolution and through the maximum bond dimension κ for the MPS state. In this work, a time step of $\delta t = 10^{-4}$ was used throughout, and the trade-off between expense and accuracy was entirely tuned through the choice of κ. The single-site TDVP algorithm we used to propagate an MPS state does not alter the bond dimension of the MPS, so κ was entirely set by the bond dimension of the initial MPS. We therefore needed to approximate the steady state $|\boldsymbol{\pi}\rangle$ with different values of κ. Approximations were generated by starting with an initially uniform state using a bond dimension κ set to several different values between 5 and 30. Those uniform distributions were evolved

with single-site TDVP for 10 000 time steps allowing the natural dynamics of $\hat{H}$ to approach a stationary state with $\langle p_t|\dot{p}_t\rangle = \langle p_t|\hat{H}|p_t\rangle \approx 0$ by time $t = 1$. Irreducibility of $\hat{H}$ means the unique stationary state is the steady-state distribution $|\boldsymbol{\pi}\rangle$. While $\langle p_t|\hat{H}|p_t\rangle = 0$ would imply $|p_t\rangle = |\boldsymbol{\pi}\rangle$, $\langle p_t|\hat{H}|p_t\rangle \approx 0$ does not guarantee that $|p_t\rangle \approx |\boldsymbol{\pi}\rangle$, only that the distribution is evolving very slowly. When transitions between $\mathcal{A}$ and $\mathcal{B}$ are slow, one can imagine reaching a local steady state distribution within $\mathcal{A}$ and within $\mathcal{B}$ but having probability very slowly repartitioning between the two. For our purposes, what matters is not that $t = 1$ has exactly reached $|\boldsymbol{\pi}\rangle$ but rather that the $t = 1$ distribution confined to $\mathcal{A}$, upon renormalization, has approached $\hat{\mathcal{P}}_{\mathcal{A}}|\boldsymbol{\pi}\rangle/\langle\mathbf{1}|\hat{\mathcal{P}}_{\mathcal{A}}|\boldsymbol{\pi}\rangle$. Provided dynamics within $\mathcal{A}$ is comparatively fast, this local steady state within $\mathcal{A}$ is reached more rapidly than $|\boldsymbol{\pi}\rangle$ and is reasonably confirmed by checking that $\langle p_t|\hat{H}|p_t\rangle \approx 0$.

The renormalized $t = 1$ distribution was then propagated under single-site TDVP with the $\tilde{H}$ absorbing-boundary dynamics (see Appendix D) for another 10 000 time steps, and projected onto $\mathcal{B}$ to add up the total probability to have reached $\mathcal{B}$. Note that these 10 000 steps are not nearly enough to reach the steady state $|\boldsymbol{\pi}\rangle$ because the slow rate of transitioning from $\mathcal{A}$ to $\mathcal{B}$ is a bottleneck. As additional time steps of dynamics were taken, the growth of the transition flux from $\mathcal{A}$ to $\mathcal{B}$ was measured and plotted in Fig. 5(b). That figure shows that at very short times ($t < \tau_{\mathrm{mol}}$) the flux from $\mathcal{A}$ to $\mathcal{B}$ has not reached a steady state value because there has not been enough time for any transitions. After that initial time τ_{mol}, the plot plateaus at a value corresponding to the rate $k_{\mathcal{B}\mathcal{A}}$, a plateau that is reached orders of magnitude faster than the time for a typical SSA trajectory to make a transition. We find that if the bond dimension is insufficient [e.g., $\kappa = 5$ for $L = 3$ in Fig. 5(b)], then the dynamics becomes unstable, reflected in erratic estimates for the conditional probability $P(\mathcal{B}_t|\mathcal{A}_0)$. As the bond dimension is grown [e.g., $\kappa = 9$ for $L = 3$ in Fig. 5(b)], the estimates converge, allowing a stable rate to be extracted. That convergence as a function of κ is analyzed further in Appendix E. We find that setting the bond dimension to roughly 10 or 20 is often sufficient, but larger values of L can only converge with a larger bond dimension. For example, Fig. 5(a) shows that SSA rates agree with $\kappa = 20$ calculations up to $L = 6$, but $\kappa = 30$ was required to push up to $L = 8$. Provided that larger bond dimension is used, the DP-TDVP rates reproduce the SSA rates across 5 orders of magnitude, up to and beyond the point that the brute-force SSA rates are practical.

Though the required bond dimension grows with L, Fig. 5(c) shows that the expense to estimate the rate grows sub-exponentially. The blue and black dashed lines in that plot are the total number of CPU hours time to estimate the rate using TDVP with $\kappa = 30$ and $\kappa = 20$ respectively. Both SSA and TDVP were run using codes written in the

JULIA programming language [75] on equivalent CPUs with details of the runtime analysis described in Appendix F. While the required CPU hours, of course, depends on details of the hardware, parallelization, and the particular implementation, we illustrate the difference in scaling when comparing identical implementations run in identical ways across different system sizes. Figure 5(c) clearly demonstrates that difference in scaling between brute-force SSA and DP-TDVP. An increase in bond dimension increases the computation time, but it does not fundamentally alter the favorable scaling in L. Comparison of computational expense with FFS is complicated by the fact that one can optimize FFS in many different ways (how many interfaces, which reaction coordinate, etc.), but Appendix G shows that the DP-TDVP calculations remain favorable even compared to FFS with a reasonable guess of an order parameter.

VII. DEPLOYING THE TOOL

We have described a numerical tool to extract the rate of a rare switching event from a reaction-diffusion model, but the numerical value of that rate is often not the final goal. More than the value of the rate, it is desirable to know how the rate changes in response to controllable parameters. Figure 5(a) shows one such variation—the switching rate falls off exponentially with the number of voxels L. The essential physics of this exponential decay is consistent with a mechanism in which a boundary voxel of the 1D chain flips to create an interface between high- and low-occupation voxels. Subsequently, each voxel at the interface flips until the whole chain has flipped. Letting p_{end} denote the small rate of flipping a boundary voxel and p_{int} denote the small rate of flipping a voxel at the interface, it follows that at short times $P(\mathcal{B}_t|\mathcal{A}_0) \propto p_{\text{end}} p_{\text{int}}^{L-1}$. If $\langle \Delta t \rangle$ is the typical time to wait between interface flips, one therefore expects the exponential decay with $k_{\mathcal{B}\mathcal{A}}(L) \propto e^{-L \ln[p_{\text{int}} \langle \Delta t \rangle]}$.

The rate calculations as a function of L thus hint at a stepwise mechanism with a flipped voxel passing from one end of the chain to the other, but that mechanistic insight can also be more directly extracted from the DP-TDVP evolution. The time evolution of $|p_t\rangle$ contains information about the entire distribution that flows from $\mathcal{A}$ to $\mathcal{B}$, and we can track the average number of molecules in each lattice site as a function of time. For example, Fig. 6 shows $\langle n_l(t) \rangle = \langle \mathbf{1}|\hat{n}_l|p_t \rangle$, the average number of molecules in voxel l with l ranging from 1 through 4 for an $L = 8$ voxel lattice. At time $t = 0$, the plot shows the normalized steady state confined to $\mathcal{A}$, the high-occupation-number state. In the steady state, the six bulk voxels, 2 through 7, have essentially indistinguishable mean occupancies, while the boundary voxels, 1 and 8, have slightly depressed means reflecting that it is marginally more likely for those two voxels to fluctuate to low-occupancy states. As the time is allowed to evolve for a timescale $t < \tau_{\text{mol}}$, the mean
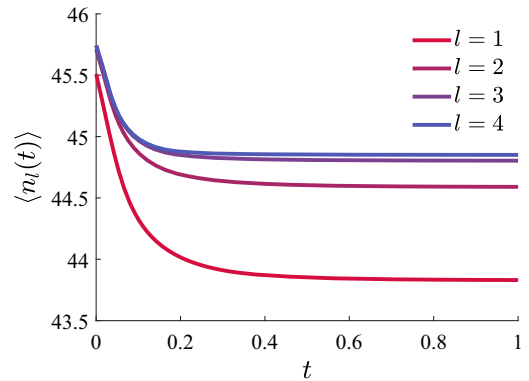

FIG. 6. The time evolution of the average number of molecules for different lattice sites when $L = 8$ and rate parameters are as in Fig. 4, computed with $\kappa = 30$. Different colors correspond to different voxels. The expected number of molecules $\langle n_l(t) \rangle = \langle \mathbf{1}|\hat{n}_l|p_t \rangle$ for voxels 1 to $\lfloor L/2 \rfloor$ (shown) are mirrored by voxels $\lfloor L/2 + 1 \rfloor$ to L. The initial nonlinear flow of probability out of $\mathcal{A}$ is shown by the steep drop in average number of molecules. The system passing τ_{mol} is shown by the transition to a linear change in particle number $\langle n_l(t) \rangle$.

occupation of the boundary voxels drops markedly due to the influence of rare trajectories in which one or both boundary voxels can flip. Subsequent to that drop in $\langle n_1(t) \rangle$, the mean occupancies of voxels 2 then 3 also drop, though less sharply, establishing the mechanistic sequence. After the molecular timescale τ_{mol}, the rate of decrease in $\langle n_l(t) \rangle$ becomes linear, reflecting the steady-state flux from $\mathcal{A}$ to $\mathcal{B}$.

In addition to varying the extent of the space, it is interesting to tune the balance between reaction and diffusion events through the diffusive hopping rate d. The sequential mechanism suggests that a limitation to the transition events emerges from p_{int}, the probability of flipping the state of an interior voxel at an interface between low- and high-occupation states. The transition events thus have low- and high-occupation blocks separated by an interface that diffuses until the low-occupation block covers the entire system. When molecules diffuse between interface voxels, the interface itself can diffuse, and the faster the interface sweeps through the system, the faster the transition events. One might therefore anticipate that the flipping rate grows with d. In fact, we find the exact opposite. Figure 7 shows the results from varying the diffusion over 20 evenly spaced values while holding the bond dimension fixed at $\kappa = 20$. The rate decreases for each value of L as d is increased due to diffusion acting as an effective stabilizer of high-occupation states.

At $d = 0$, each voxel is an independent well-mixed system. The switching rate depends on the probability that each voxel independently experiences a collapse of population, dislodging it from its metastable high-occupation state. As d grows, there is a diffusive mechanism to fight against that population collapse. When the population of a
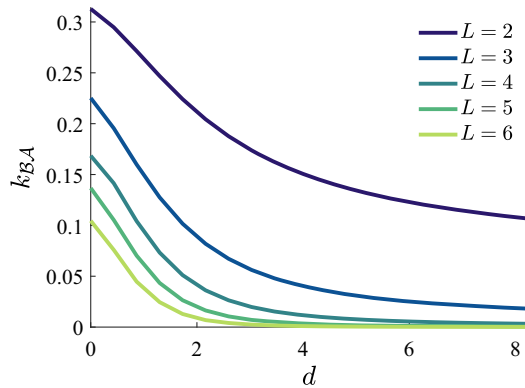
FIG. 7. The switching rate k_{BA} as a function of diffusion coefficient d for $2 \leq L \leq 6$, computed with $\kappa = 20$ for the rates from Fig. 4. As d is increased, the diffusion acts to inhibit the stochastic switching of each voxel. At large enough d the system acts similar to a large well-mixed system, resulting in a leveling off of the rate coefficients.

TABLE I. Examples of the mapping between a common chemical reaction and the Doi-Peliti form of the contribution to an effective Hamiltonian. The procedure to derive such correspondences is systematic [78] and typically discussed for a Fock space with occupation number ranging from 0 to infinity. To truncate to a finite Fock space, one must additionally add a projector to each effective Hamiltonian, as in this work, to ensure conservation of probability.

Reaction	Hamiltonian
$X \xrightarrow{\tilde{c}} Y$	$\tilde{c}[(y^\dagger - x^\dagger)x]$
$X + Y \xrightarrow{\tilde{c}} Z$	$\tilde{c}[(z^\dagger - x^\dagger y^\dagger)xy]$
$X + Y \xrightarrow{\tilde{c}} 2X$	$\tilde{c}[x^\dagger(x^\dagger - y^\dagger)xy]$

voxel wanders dangerously close to a tipping point, molecules could diffuse from a high-occupancy neighbor to restore the population. In the limit of very fast diffusion, the entire chain of voxels effectively behaves as a single large well-mixed system with a markedly lower switching rate because statistical fluctuations away from the metastable state get correspondingly smaller.

VIII. DISCUSSION

Using the 1D Schlögl model, we have demonstrated how TN calculations combined with the DP framework offer a new paradigm for numerical investigation of reaction-diffusion models, even when the physical bond dimension $(M + 1)$ is not small. That demonstration opens up a large class of RDME models which represent small well-mixed reactors, connected together by diffusive coupling. As illustrated in this work, these models can exhibit interesting physics when each reactor is small enough to exhibit significant statistical fluctuations and when the diffusion between reactors induces correlations between fluctuations in different reactors. In this work, our approach was to start with a well-studied kinetic toy model and to add spatial dynamics. Spatial extensions of other toy models like the Brusselator [76] or the Oregonator [77] would be amenable to similar numerical analysis because any reaction-diffusion model which can be written as a set of elementary reactions will have a DP representation. That representation can involve creation and annihilation operators for different particle types, not only for species X, but the mapping from reactions to effective Hamiltonians is systematic. Table I shows the effective Hamiltonians for several common (unchemostatted) reactions: unimolecular, bimolecular, and autocatalytic. Given the set of reactions, it is comparatively straightforward to derive the DP representation, and computational tools already exist to generate an

MPO [62] and propagate via TDVP from that DP form of the effective Hamiltonian.

As the TN techniques become increasingly well developed, one can imagine that the computational tools will offer generic "turn the crank" analysis of microscopic models of reaction-diffusion at the ensemble level. We anticipate these numerical approaches to be beneficial to both synthetic and biophysical reaction-diffusion systems. In the case of the former, the discretization of space into reactors can be explicitly realized, as in the microfluidic fabrication of a two-dimensional array of nanoliter-volume Belousov-Zhabotinsky (BZ) reactors [79]. In line with our 1D Schlögl model, the strength of coupling between those BZ nanoreactors is regulated by diffusion, and there is experimental interest in pattern formation as a function of that diffusive coupling. In the case of biophysics, sub-cellular dynamics is dominated by the generation and diffusion of proteins with small copy numbers. Much work has been devoted to understanding the logic of cellular decision making through stochastic CRN models, most famously gene regulatory models, which are most frequently studied in a well-mixed limit [80]. We see these TN tools as an intriguing pathway to numerically investigate the timescales and pathways for switching between metastable states in well-mixed CRNs like gene toggle switches [10] and in reaction-diffusion variants of those gene-regulatory models. Eventually, one can dream of bringing these methods to the broad array of systems biology applications where RDME models are widely simulated [81], including in large-scale GPU-accelerated simulations of systems with biologically relevant size [82]. At present those sorts of reaction-diffusion model can only be numerically accessed through sampling approaches, and we expect the ensemble perspective will some day complement those efforts.

To move from the toy models to the synthetic and biophysical systems, it will be important to expand the ability of TNs to approximate correlation between species and voxels that are not simply connected through a 1D

chain. Here, because our 1D chain of Schlögl reactors are connected only to their neighbors, it was natural to approximate the microscopic distribution with an MPS. That choice allowed us to capitalize on the efficient TDVP algorithm for MPS states, but a RDME with many dynamic species will not always be easily approximated by a 1D chain of tensors. If two chemical species are strongly correlated but far apart in the MPS structure, the MPS calculation will demand a very high bond dimension. It will be important for future work to dynamically adapt the bond dimension through subspace expansion [83] or by devising systematic ways to generate tractable TN architectures with connectivity mirroring the correlations between CRN species. Those technical advances may eventually lay the groundwork to move the methodology from this Schlögl study to the complex, biologically motivated models where SSA realizations presently reign supreme.

APPENDIX A: DOI-PELITI ENCODING

Assume a set of elementary reactions are given. These can be encoded into a Doi-Peliti representation which converts each elementary reaction into a term of an effective Hamiltonian. The operator representations can then be exactly written as an MPO, as shown below. Consider an L-site 1D lattice of well-mixed voxels. Inside each voxel the Schlögl model reactions can fire. A probability distribution will be written as a vector in the Fock space $\mathcal{F}^L$ spanned by orthonormal basis vectors $|\mathbf{n}\rangle = |n_1, n_2, \ldots, n_L\rangle$ which correspond to single microstates, with occupancy of each voxel given by $n_1, n_2, \ldots, n_L$. For any two microstates $|\mathbf{n}\rangle$ and $|\mathbf{m}\rangle$ we thus have $\langle \mathbf{m}|\mathbf{n}\rangle = \delta_{\mathbf{m},\mathbf{n}}$. Distributions over the microstates can be written as $|p_t\rangle = \sum_{\mathbf{n}} p_t(\mathbf{n})|\mathbf{n}\rangle$, where $p_t(\mathbf{n}) \equiv p(\mathbf{n}, t) \equiv p(n_1, n_2, \ldots, n_L, t)$ is the probability of microstate $\mathbf{n}$ at time t. When focusing on the lth voxel we will use the shorthand $p_t(\mathbf{n}) \equiv p_t(\mathbf{n}^c, n_l)$, where $\mathbf{n}^c = \{n_k, \ \forall \ k \neq l\}$

defines the state of the set of voxels that complement voxel l. Each reaction of the Schlögl model contributes to the set of differential equations which describe the time evolution of $p_t(\mathbf{n})$. We must convert the local change in probability of microstate $\mathbf{n}$ to the global action of the effective Hamiltonian on $|p_t\rangle$. The linear nature of the chemical master equation means we can consider one reaction at a time.

1. Reaction 1: $2X + A \xrightarrow{\tilde{c}_1} 3X$

We start with the reaction $2X + A \xrightarrow{\tilde{c}_1} 3X$. When this reaction fires inside site l, the change in probability is

$$\frac{dp_t(\mathbf{n})}{dt} = \frac{\tilde{c}_1}{2}[(n_l - 1)(n_l - 2)p_t(\mathbf{n}^c, n_l - 1) \\ - n_l(n_l - 1)p_t(\mathbf{n}^c, n_l)]. \tag{A1}$$

The first (positive) term captures the gain of probability into microstate $\mathbf{n}$ from microstate $(\mathbf{n}^c, n_l - 1)$. Two of the $n_l - 1$ possible X molecules must come together to make the third, and there are $(n_l - 1)(n_l - 2)/2$ total combinations. The second (negative) term of Eq. (A1) accounts for the probability of already being in state $\mathbf{n}$ at time t. Note that this reaction adds an X molecule, so if the system is already at the maximum cap of $n_l = M$ molecules, there can be no negative term at $M + 1$, to balance out conservation of probability. The loss term therefore requires extra care when there is an enforced cap since the loss term would come from a forbidden transition. Consequently, Eq. (A1) can only apply to microstates $\mathbf{n}$ whose lth voxel is not maximally occupied. For the general case, we could rewrite the evolution as

$$\frac{dp_t(\mathbf{n})}{dt} = \frac{\tilde{c}_1}{2}[(n_l - 1)(n_l - 2)p_t(\mathbf{n}^c, n_l - 1) \\ - (1 - \delta_{n_l,M})n_l(n_l - 1)p_t(\mathbf{n}^c, n_l)]. \tag{A2}$$

Summing each side of Eq. (A2) over the Fock states $|\mathbf{n}\rangle$ yields

$$\frac{d|p_t\rangle}{dt} = \frac{\tilde{c}_1}{2}\sum_{\mathbf{n}}[(n_l - 1)(n_l - 2)p_t(\mathbf{n}^c, n_l - 1)|\mathbf{n}\rangle \\ - (1 - \delta_{n_l,M})n_l(n_l - 1)p_t(\mathbf{n}^c, n_l)|\mathbf{n}\rangle]. \tag{A3}$$

It is useful to match the argument of p_t with the ket, which can be achieved by introducing the appropriate creation operator:

$$\frac{d|p_t\rangle}{dt} = \frac{\tilde{c}_1}{2}\sum_{\mathbf{n}}[(n_l - 1)(n_l - 2)p_t(\mathbf{n}^c, n_l - 1)x_l^\dagger|\mathbf{n}^c, n_l - 1\rangle \\ - (1 - \delta_{n_l,M})n_l(n_l - 1)p_t(\mathbf{n}^c, n_l)|\mathbf{n}\rangle]. \tag{A4}$$

The $x_l^\dagger$ term preceding the first ket is an operator, but all other terms preceding the kets are numbers that readily commute. We reorder those terms to get

$$\frac{d|p_t\rangle}{dt} = \frac{\tilde{c}_1}{2}\sum_{\mathbf{n}}[x_l^\dagger p_t(\mathbf{n}^c, n_l - 1)(n_l - 1)(n_l - 2)|\mathbf{n}^c, n_l - 1\rangle$$
$$- (1 - \delta_{n_l, M})p_t(\mathbf{n}^c, n_l)n_l(n_l - 1)|\mathbf{n}\rangle]. \tag{A5}$$

Equation (14) gives the identity $x_l^{\dagger 2}x_l^2|\mathbf{n}\rangle = n_l(n_l - 1)|\mathbf{n}\rangle$, allowing both the gain and loss terms to be reexpressed in terms of powers of the creation and annihilation operators:

$$\frac{d|p_t\rangle}{dt} = \frac{\tilde{c}_1}{2}\sum_{\mathbf{n}}[x_l^\dagger p_t(\mathbf{n}^c, n_l - 1)x_l^{\dagger 2}x_l^2|\mathbf{n}^c, n_l - 1\rangle$$
$$- (1 - \delta_{n_l, M})p_t(\mathbf{n}^c, n_l)x_l^{\dagger 2}x_l^2|\mathbf{n}\rangle]. \tag{A6}$$

Note that the creation and annihilation operators can now be factorized outside of the sum to yield two terms which look very similar to the Fock-space expansion $|p_t\rangle = \sum_{\mathbf{n}} p_t(\mathbf{n})|\mathbf{n}\rangle$:

$$\frac{d|p_t\rangle}{dt} = \frac{\tilde{c}_1}{2}\left[x_l^{\dagger 3}x_l^2\sum_{\mathbf{n}}p_t(\mathbf{n}^c, n_l - 1)|\mathbf{n}^c, n_l - 1\rangle \right.$$
$$\left. - x_l^{\dagger 2}x_l^2\sum_{\mathbf{n}}(1 - \delta_{n_l, M})p_t(\mathbf{n}^c, n_l)|\mathbf{n}\rangle\right]. \tag{A7}$$

The first term is like the expansion of $|p_t\rangle$ except that the bounds of summation are shifted, including kets with occupancy -1 through $M_l - 1$ rather than 0 through M_l. The nonphysical state with occupancy -1 has no probability, so it can be excluded from the sum. Furthermore, the occupancy M term can be added to the sum since $x_l^{\dagger 3}x_l^2|\mathbf{n}^c, M_l\rangle = 0$. Consequently, the first sum of Eq. (A7) is equivalent to $|p_t\rangle$. The second sum can also be expressed in terms of $|p_t\rangle$ if we recognize that the δ function constraint can be viewed as the eigenvalue of a projector operator:

$$\hat{\mathcal{P}}_{M_l}^c|\mathbf{n}\rangle = (\mathbb{I} - |\mathbf{n}^c, M_l\rangle\langle\mathbf{n}^c, M_l|)|\mathbf{n}\rangle = (1 - \delta_{n_l, M})|\mathbf{n}\rangle. \tag{A8}$$

Hence, both sums of Eq. (A7) simplify in terms of the same Fock state $|p_t\rangle$:

$$\frac{d|p_t\rangle}{dt} = \frac{\tilde{c}_1}{2}[x_l^{\dagger 3}x_l^2 - x_l^{\dagger 2}x_l^2\hat{\mathcal{P}}_{M_l}^c]|p_t\rangle. \tag{A9}$$

The factor of $1/2$ can be traced back to a combinatorial counting term for the unique number of ways to label the two X molecules that react. We absorb that counting term into the rate $c_1 = \tilde{c}_1/2$ and extract from Eq. (A9) the contribution to the effective Hamiltonian arising from the first reaction in site l:

$$\hat{H}_l^{A_{\text{for}}} = c_1[x_l^{\dagger 3}x_l^2 - x_l^{\dagger 2}x_l^2\hat{\mathcal{P}}_{M_l}^c]. \tag{A10}$$

We take that direction of the reaction to be "forward" and use the superscript A because this is the reaction mediated by species A. By applying Eq. (19), the projector term can be rewritten in terms of a commutator of x_l and $x_l^\dagger$ to give an alternative form of $\hat{H}_l^{A_{\text{for}}}$ expressed only in terms of the creation and annihilation operators, c_1, and the maximal voxel occupancy M:

$$\hat{H}_l^{A_{\text{for}}} = c_1\left[x_l^{\dagger 3}x_l^2 - \frac{x_l^{\dagger 2}x_l^3x_l^\dagger}{M + 1} + \frac{x_l^{\dagger 2}x_l^2x_l^\dagger x_l}{M + 1} - \frac{Mx_l^{\dagger 2}x_l^2}{M + 1}\right]. \tag{A11}$$

This form, shows that in the limit $M \to \infty$, Eq. (A11) recovers the more standard bosonic result, $\hat{H}_l^{A_{\text{for}}} = c_1[x_l^{\dagger 3}x_l^2 - x_l^{\dagger 2}x_l^2]$. For compactness, we typically leave our expressions for contributions to the effective Hamiltonians as expressions that explicitly involve the finite-voxel projector, as in Eq. (A10).

2. Reaction 2: $3X \xrightarrow{\tilde{c}_2} 2X + A$

The reverse reaction, which transitions from an occupancy of $n_l + 1$ to n_l, has $(n_l + 1)n_l(n_l - 1)/6$ different ways to combine three of the X molecules together. Consequently, the change in probability for the reverse reaction is

$$\frac{dp_t(\mathbf{n})}{dt} = \frac{\tilde{c}_2}{6}[(n_l - 1)n_l(n_l + 1)p_t(\mathbf{n}^c, n_l + 1)$$
$$- n_l(n_l - 1)(n_l - 2)p_t(\mathbf{n}^c, n_l)]. \tag{A12}$$

Unlike the forward reaction, the negative term does not include a factor of $\delta_{n_l, M}$ that treats the maximal occupancy state differently than the others. The distinction is that the loss terms from reaction 2 arise when probability is lost from occupancy n_l down to occupancy $n_l - 1$. Those loss terms are possible from $n_l = 3$ through M. If there are fewer than 3 molecules, the reaction 2 mechanism is not possible, but this restriction emerges naturally from the $n_l(n_l - 1)(n_l - 2)$ coefficients in Eq. (A12). Expanding in the Fock basis therefore gives

$$\frac{d|p_t\rangle}{dt} = \frac{\tilde{c}_2}{6}\sum_{\mathbf{n}}[(n_l - 1)n_l(n_l + 1)p_t(\mathbf{n}^c, n_l + 1)|\mathbf{n}\rangle$$
$$- n_l(n_l - 1)(n_l - 2)p_t(\mathbf{n})|\mathbf{n}\rangle]. \tag{A13}$$

Mirroring the treatment of the forward reaction, we rewrite the first $|\mathbf{n}\rangle$ using a lowering operator:

$$\frac{d|p_t\rangle}{dt} = \frac{\tilde{c}_2}{6}\sum_{\mathbf{n}}[(n_l - 1)n_l p_t(\mathbf{n}^c, n_l + 1)x_l|\mathbf{n}^c, n_l + 1\rangle$$
$$- n_l(n_l - 1)(n_l - 2)p_t(\mathbf{n})|\mathbf{n}\rangle]. \tag{A14}$$

Equation (14) implies $x_l^{\dagger 3}x_l^3|\mathbf{n}\rangle = n_l(n_l-1)(n_l-2)|\mathbf{n}\rangle$ and $x_l^{\dagger 2}x_l^3|\mathbf{n}^c,n_l+1\rangle = (n_l-1)n_l x_l|\mathbf{n}^c,n_l+1\rangle$, which allow us to simplify Eq. (A14) to

$$\frac{d|p_t\rangle}{dt} = \frac{\tilde{c}_2}{6}\left[x_l^{\dagger 2}x_l^3\sum_{\mathbf{n}}p_t(\mathbf{n}^c,n_l+1)|\mathbf{n}^c,n_l+1\rangle\right.$$
$$\left. - x_l^{\dagger 3}x_l^3\sum_{\mathbf{n}}p_t(\mathbf{n})|\mathbf{n}\rangle\right]. \qquad (A15)$$

We recognize the second sum over $\mathbf{n}$ as giving $|p_t\rangle$. The first sum ranges from $n_l=1$ through $M+1$, but the $M+1$ term vanishes due to the occupancy cap. Furthermore, an $n_l=0$ term can be added to the sum since that term also vanishes $(x_l^{\dagger 2}x_l^3|\mathbf{n}^c,0_l\rangle = 0)$. Thus the first sum can be reindexed to give $|p_t\rangle$. Performing both of those sums yields

$$\frac{d|p_t\rangle}{dt} = \frac{\tilde{c}_2}{6}[x_l^{\dagger 2}x_l^3 - x_l^{\dagger 3}x_l^3]|p_t\rangle. \qquad (A16)$$

The contribution to the effective Hamiltonian from the reverse A-mediated reaction in voxel l is thus

$$\hat{H}_l^{A_{rev}} = c_2[x_l^{\dagger 2}x_l^3 - x_l^{\dagger 3}x_l^3], \qquad (A17)$$

where we have again absorbed the combinatorial term into $c_2 = \tilde{c}_2/6$. By summing together the forward and reverse reactions, we arrive at the effective Hamiltonian associated with the reversible A-mediated reaction, quoted in Eq. (21):

$$\hat{H}_l^A = \hat{H}_l^{A_{for}} + \hat{H}_l^{A_{rev}}$$
$$= c_1[x_l^{\dagger 3}x_l^2 - x_l^{\dagger 2}x_l^2\hat{\mathcal{P}}_{M_l}^c] + c_2[x_l^{\dagger 2}x_l^3 - x_l^{\dagger 3}x_l^3]. \qquad (A18)$$

3. Reaction 3: $B \xrightarrow{\tilde{c}_3} X$

Deriving the effective Hamiltonian for the third reaction is even more straightforward than the previous two. As in the first reaction, the loss term is only present if the microstate is not maximally occupied, so the probability for microstate $\mathbf{n}$ evolves according to

$$\frac{dp_t(\mathbf{n})}{dt} = \tilde{c}_3 p_t(\mathbf{n}^c,n_l-1) - (1-\delta_{n_l,M})\tilde{c}_3 p_t(\mathbf{n}). \qquad (A19)$$

In the Fock basis, this evolution becomes

$$\frac{d|p_t\rangle}{dt} = \tilde{c}_3\sum_{\mathbf{n}}[p_t(\mathbf{n}^c,n_l-1)|\mathbf{n}\rangle - (1-\delta_{n_l,M})p_t(\mathbf{n})|\mathbf{n}\rangle]. \qquad (A20)$$

We again seek to make the first ket's microstate match the argument of p_t by introducing the creation operator:

$$\frac{d|p_t\rangle}{dt} = \tilde{c}_3\sum_{\mathbf{n}}[p_t(\mathbf{n}^c,n_l-1)x_l^{\dagger}|\mathbf{n}^c,n_l-1\rangle$$
$$- (1-\delta_{n_l,M})p_t(\mathbf{n})|\mathbf{n}\rangle]$$
$$= \tilde{c}_3\left[x_l^{\dagger}\sum_{\mathbf{n}}p_t(\mathbf{n}^c,n_l-1)|\mathbf{n}^c,n_l-1\rangle\right.$$
$$\left. - \sum_{\mathbf{n}}(1-\delta_{n_l,M})p_t(\mathbf{n})|\mathbf{n}\rangle\right]. \qquad (A21)$$

As in reaction 1, we reindex the first sum to get $|p_t\rangle$ and we introduce the projector operator into the second sum to get $\hat{\mathcal{P}}_{M_l}^c|p_t\rangle$, leaving

$$\frac{d|p_t\rangle}{dt} = \tilde{c}_3(x_l^{\dagger} - \hat{\mathcal{P}}_{M_l}^c)|p_t\rangle. \qquad (A22)$$

In this case, there is no combinatorial counting term, so $c_3 = \tilde{c}_3$ and

$$\hat{H}_l^{B_{for}} = c_3(x_l^{\dagger} - \hat{\mathcal{P}}_{M_l}^c). \qquad (A23)$$

4. Reaction 4: $X \xrightarrow{\tilde{c}_4} B$

The reverse reaction is like reaction 2 in that the reaction decreases X and therefore cannot cause an overflow beyond the maximal voxel occupancy. Consequently, the evolution of probability in microstate $\mathbf{n}$ is

$$\frac{dp_t(\mathbf{n})}{dt} = \tilde{c}_4(n_l+1)p_t(\mathbf{n}^c,n_l+1) - \tilde{c}_4 n_l p_t(\mathbf{n}^c,n_l). \qquad (A24)$$

Expanding in the Fock space gives

$$\frac{d|p_t\rangle}{dt} = \tilde{c}_4\sum_{\mathbf{n}}[(n_l+1)p_t(\mathbf{n}^c,n_l+1)|\mathbf{n}\rangle - np_t(\mathbf{n})|\mathbf{n}\rangle]$$
$$= \tilde{c}_4\sum_{\mathbf{n}}[p_t(\mathbf{n}^c,n_l+1)x_l|\mathbf{n}^c,n_l+1\rangle - np_t(\mathbf{n})|\mathbf{n}\rangle]. \qquad (A25)$$

As before, we expand into two sums. Upon checking the two boundary terms of the first sum, we reindex it and sum over microstates, recovering $|p_t\rangle$ from both sums to get

$$\frac{d|p_t\rangle}{dt} = \tilde{c}_4(x_l - x_l^{\dagger}x_l)|p_t\rangle. \qquad (A26)$$

The combinatorial term is again trivial, so $\tilde{c}_4 = c_4$, and

$$\hat{H}_l^{B_{rev}} = c_4(x_l - x_l^{\dagger}x_l). \qquad (A27)$$

Combining Eqs. (A23) and (A27) yields the effective Hamiltonian arising from the reversible B-mediated reaction, Eq. (22):

$$\hat{H}_l^{\mathrm{B}} = c_3(x_l^\dagger - \hat{\mathcal{P}}_{M_l}^c) + c_4(x_l - x_l^\dagger x_l). \qquad \text{(A28)}$$

The diffusion operator $\hat{H}^{\mathrm{diff}}$ of Eq. (23) is derived similarly by adding the maximal voxel occupancy constraint to the derivation of the bosonic diffusion operator [84].

APPENDIX B: MATRIX PRODUCT OPERATORS

Note that the terms contributing to the effective Hamiltonian [Eqs. (21)–(23)] each take a "product" form in which the creation and annihilation operators at a site multiply other creation and annihilation operators at that same site or at a nearest-neighbor site. For systems with local and nearest-neighbor interactions, there exists a relatively simple representation of the effective Hamiltonian. Let $\hat{H}_l = \hat{H}_l^{\mathrm{A}} + \hat{H}_l^{\mathrm{B}}$ be the local effective Hamiltonian of the reversible reactions occurring in voxel l. The total effective Hamiltonian $\hat{H}$ includes this local contribution summed over all voxels l plus the $\hat{H}^{\mathrm{diff}}$. Owing to the nearest-neighbor interactions, these terms can be elegantly summed via a product of operator-valued vectors and matrices. We define the operator-valued vectors associated with voxels 1 and L as

$$W^{[1]} = [\hat{H}_1, dx_1, d\hat{\mathcal{P}}_{M_1}^c, dx_1^\dagger, -x_1^\dagger x_1, \mathbb{I}],$$

$$W[L] = \begin{bmatrix} \mathbb{I} \\ x_L^\dagger \\ -x_L^\dagger x_L \\ x_L \\ d\hat{\mathcal{P}}_{M_L}^c \\ H_L \end{bmatrix}. \qquad \text{(B1)}$$

and the operator-valued matrix associated with voxels 2 through $L-1$ as

$$W^{[l]} = \begin{bmatrix} \mathbb{I} & & & & & \\ x_l^\dagger & & & & & \\ -x_l^\dagger x_l & & & \mathbf{0} & & \\ x_l & & & & & \\ d\hat{\mathcal{P}}_{M_l}^c & & & & & \\ \hat{H}_l, & dx_l, & d\hat{\mathcal{P}}_{M_l}^c, & dx_l^\dagger, & -x_l^\dagger x_l, & \mathbb{I} \end{bmatrix}. \qquad \text{(B2)}$$

With these definitions, one can show that $\hat{H} = W^{[1]} \cdot W^{[2]} \dots W^{[L-1]} \cdot W^{[L]}$, with the centered dot representing a contraction.

This decomposition of $\hat{H}$ offers a highly efficient way to numerically encode the effective Hamiltonian as an MPO. While we have shown an explicit MPO decomposition for the nearest-neighbor situation, in practice one uses powerful automated packages such as AutoMPO in ITensor [65] to generate compressed forms of the MPO from the second-quantized expressions like Eqs. (21)–(23).

APPENDIX C: TDVP ALGORITHM

As described in Sec. IV, we wish to approximate the evolution of the distribution

$$\frac{d|p_t\rangle}{dt} = \hat{H}|p_t\rangle, \qquad \text{(C1)}$$

by approximating $|p_t\rangle$ as

$$|p_t\rangle = \sum_{n_1, n_2, \dots, n_L} p_t(n_1, n_2, \dots, n_L)|n_1, n_2, \dots, n_L\rangle$$

$$\approx \sum_{\substack{n_1, n_2, \dots, n_L \\ s_1, s_2, \dots, s_{L-1}}} Q_{s_1}^{n_1} Q_{s_1 s_2}^{n_2} \cdots Q_{s_{L-1}}^{n_L} |n_1, n_2, \dots, n_L\rangle$$

$$\equiv |q_t\rangle. \qquad \text{(C2)}$$

Here $|q_t\rangle$ is a distribution over microstates that can be expressed as an MPS using the (time-dependent) set of tensors $\{Q^{n_1}, Q^{n_2}, \dots, Q^{n_L}\}$. As in the main text, the distinct Q tensors are implicitly distinguished based on the labeling scheme for the physical index. Only a subset of possible distributions $|p_t\rangle$ can be written in the form of $|q_t\rangle$, but as the bond dimension κ increases, that subset necessarily grows. We call these distributions which can be represented by MPS states the MPS manifold $\mathcal{M}$. Because $\hat{H}$ generally evolves $|q_t\rangle$ off $\mathcal{M}$, the time-dependent variational principle [67,68] seeks the evolution of $|q_t\rangle$ along the MPS manifold:

$$\frac{d|q_t\rangle}{dt} = \hat{\mathcal{P}}_{T_{\mathcal{M}}} \hat{H}|q_t\rangle, \qquad \text{(C3)}$$

where $\hat{\mathcal{P}}_{T_{\mathcal{M}}}$ is a projector mapping the infinitesimal change in probability $\hat{H}|q_t\rangle$ onto the MPS manifold's nearest (with respect to the l^2 norm) tangent vector. The projector in this equation of motion ensures that $|q_t\rangle$ stays on $\mathcal{M}$ for all times.

Remarkably, in the case of the MPS manifold, an explicit form for the projector can be written down as a sum of $2L-1$ terms, L of which have a correspondence to the L tensors of the MPS and $L-1$ of which have a correspondence to the bonds. To write down that decomposition, it is beneficial to first define the environment tensors which were obliquely referenced in

Sec. IV. Through gauge transformations, $|q_t\rangle$ is expressed as

$$|q_t\rangle = \sum_{\substack{n_1,n_2,\ldots,n_L \\ s_1,s_2,\ldots,s_{L-1}}} A_{s_1}^{n_1} A_{s_1 s_2}^{n_2} \cdots A_{s_{l-2} s_{l-1}}^{n_{l-1}} Q_{s_{l-1} s_l}^{n_l} B_{s_l s_{l+1}}^{n_{l+1}} \cdots$$
$$\cdot B_{s_{L-2} s_{L-1}}^{n_{L-1}} B_{s_{L-1}}^{n_L} |n_1, \ldots, n_L\rangle, \qquad \text{(C4)}$$

where the A tensors are left orthogonal and B tensors are right orthogonal, meaning

$$\sum_{n_i,s_{i-1}} A_{s_{i-1} s_i}^{n_i} A_{s_{i-1} s_i'}^{n_i} = \delta_{s_i s_i'} \qquad \text{(C5)}$$

and

$$\sum_{n_i,s_i} B_{s_{i-1} s_i}^{n_i} A_{s_{i-1}' s_i}^{n_i} = \delta_{s_{i-1} s_{i-1}'}. \qquad \text{(C6)}$$

Summing the A tensors found in Eq. (C4) up to $l-1$ gives the left environment tensor,

$$\left| \Phi_{L,s_{l-1}}^{[n_1 : n_{l-1}]} \right\rangle = \sum_{\substack{n_1,\ldots,n_{l-1} \\ s_1,\ldots,s_{l-2}}} A_{s_1}^{n_1} A_{s_1,s_2}^{n_2} \cdots A_{s_{l-2} s_{l-1}}^{n_{l-1}} |n_1, n_2, \ldots, n_{l-1}\rangle,$$
$$\text{(C7)}$$

with physical indices counting the occupancy of sites 1 through $l-1$ and a single uncontracted bond index s_{l-1}. A similar sum over the s indices of the B tensors gives the right environment tensor,

$$\left| \Phi_{R,s_l}^{[n_{l+1} : n_L]} \right\rangle = \sum_{\substack{n_{l+1},\ldots,n_L \\ s_{l+1},\ldots,s_{L-1}}} B_{s_l,s_{l+1}}^{n_{l+1}} \cdots$$
$$\times B_{s_{L-2} s_{L-1}}^{n_{L-1}} B_{s_{L-1}}^{n_L} |n_{l+1}, n_{l+2}, \ldots, n_L\rangle, \qquad \text{(C8)}$$

with physical indices counting the occupancy of sites $l+1$ through L and a single uncontracted bond index s_l.

In terms of the left and right environment tensor projectors,

$$P_L^{[1:l-1]} = \sum_{s_{l-1}} \left| \Phi_{L,s_{l-1}}^{[n_1 : n_{l-1}]} \right\rangle \left\langle \Phi_{L,s_{l-1}}^{[n_1:n_{l-1}]} \right| \qquad \text{(C9)}$$

and

$$P_R^{[l+1:L]} = \sum_{s_l} \left| \Phi_{R,s_l}^{[n_{l+1}:n_L]} \right\rangle \left\langle \Phi_{R,s_l}^{[n_{l+1}:n_L]} \right|, \qquad \text{(C10)}$$

we can now write out the $2L-1$ terms of the projector for the MPS manifold [72,85]:

$$\hat{\mathcal{P}}_{T_{\mathcal{M}}} = \sum_{l=1}^{L} P_L^{[1:l-1]} \otimes \mathbb{I}_l \otimes P_R^{[l+1:L]} - \sum_{l=1}^{L-1} P_L^{[1:l]} \otimes P_R^{[l+1:L]}$$
$$\equiv \sum_{l=1}^{L} P_l^+ + \sum_{l=1}^{L-1} P_l^-. \qquad \text{(C11)}$$

The positive terms P_l^+ are associated to the L tensors and the negative terms P_l^- are associated to the $L-1$ bonds.

Integrating Eq. (C3) for a time step δt gives

$$|q_{t+\delta t}\rangle = \exp\left(\delta t \left[\sum_{l=1}^{L} P_l^+ \hat{H} + \sum_{l=1}^{L-1} P_l^- \hat{H} \right] \right) |q_t\rangle. \qquad \text{(C12)}$$

In the limit of an infinitesimal time step, we can factorize that exponential to act with one l at a time. For example,

$$|q_{t+\delta t}\rangle = \exp\left(\delta t \left[\sum_{l=2}^{L} P_l^+ \hat{H} + \sum_{l=2}^{L-1} P_l^- \hat{H} \right] \right) e^{\delta t P_1^- \hat{H}} e^{\delta t P_1^+ \hat{H}} |q_t\rangle. \qquad \text{(C13)}$$

The operator $e^{\delta t P_1^+ \hat{H}}$ has the physical interpretation of mapping site 1 forward in time by a time step δt. To see this interpretation, we illustrate the action of $e^{\delta t P_2^+ \hat{H}}$ on $|q_t\rangle$, where we assume the MPS has already be written in a mixed canonical form centered on site 2. The operator $e^{\delta t P_2^+ \hat{H}}$ acts globally on the MPS state, but as long as $|q_t\rangle$ is expressed in the mixed canonical form, only the tensor at site 2 will be altered. In that way, the P_2^+ term is implemented as a *local* update to a single tensor, a local update with the structure of a time evolution with a local effective Hamiltonian. The reduction from global to local operations is seen if the global exponential operator is expanded in a Taylor series:

$$e^{\delta t P_2^+ \hat{H}} |q_t\rangle = \sum_{\substack{n_1,n_2,\ldots,n_L \\ s_1,s_2,\ldots,s_{L-1}}} \sum_{k=0}^{\infty} \frac{\delta t^k}{k!} (P_2^+ \hat{H})^k Q_{s_1}^{n_1} Q_{s_1 s_2}^{n_2} \cdots Q_{s_L}^{n_L} |\mathbf{n}\rangle.$$
$$\text{(C14)}$$

Each term of the sum over $\mathbf{n}$ can be expressed graphically, where the blue triangular MPS tensors indicate their left or right orthogonality by the direction they point toward the black diamond tensor at the center of the mixed canonical form:
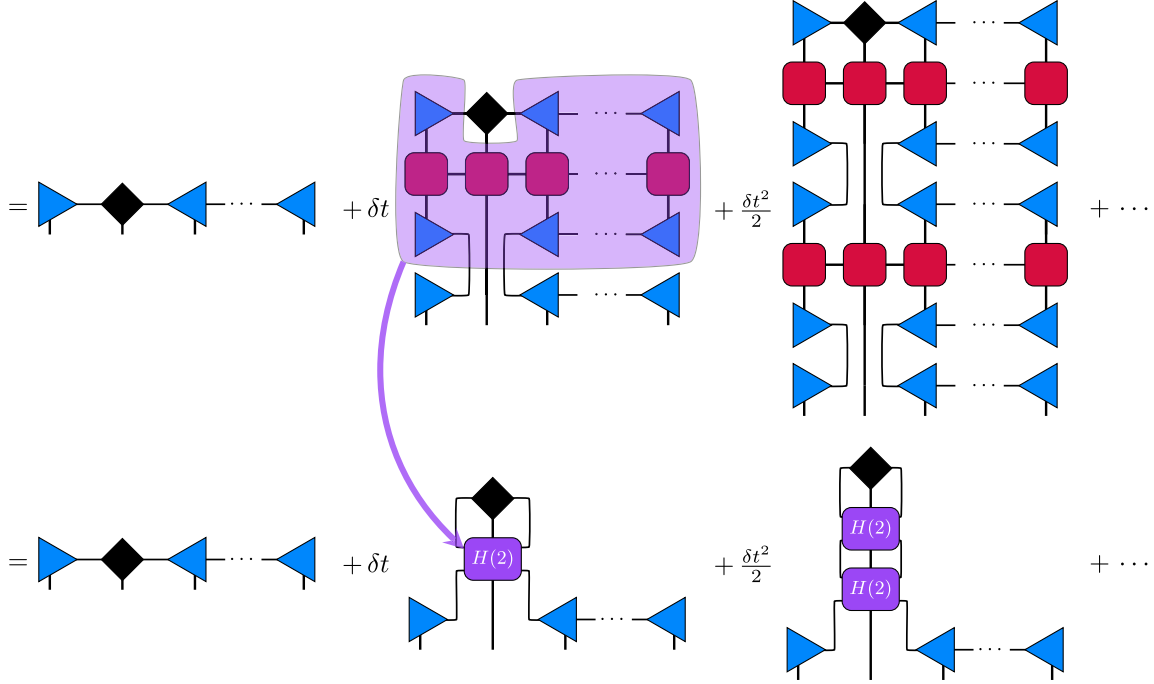
By contracting over all of the tensors in the shaded purple region, we obtain the local tensor we call $H(2)$. Note that the same tensor appears repeatedly in the higher order terms of the Taylor series. These terms can be resummed to give back a *local* matrix exponential acting on site 2. Hence, by contracting $Q^{n_2}(t)$ (the black diamond) with $e^{\delta t H(2)}$ we get $Q^{n_2}(t + \delta t)$.

Following the $P_2^+ \hat{H}$ propagation, the newly updated tensor $Q^{n_2}(t + \delta t)$ is then decomposed into two blocks $Q^{n_2}_{s_1 s_2} = A^{n_2}_{s_1 s_2'} R_{s_2' s_2}$ via QR or SVD decomposition. Like how $P_2^+ \hat{H}$ acted locally on the tensor at site 2, P_2^- will act locally to propagate $R_{s_2' s_2}$, but the interpretation is that the propagation is a δt time step backward in time, owing to the negative sign in Eq. (C11). $R_{s_2' s_2}$ is then contracted with the next tensor in the chain, $B^{n_3}_{s_2 s_3}$, which is also still at time t. In doing so, we shift gauge such that the new $t + \delta t$ tensor assigned to site 2 is $A^{n_2}_{s_1 s_2'}$ and the new center site tensor at site 3 becomes $Q^{n_3}_{s_2' s_3} = \sum_{s_2} R_{s_2' s_2} B^{n_3}_{s_2 s_3}$. Having shifted the center to site 3, one then iterates: acting with $P_3^+ \hat{H}$, performing another QR or SVD decomposition, acting with $P_3^- \hat{H}$, shifting the gauge, then continuing to sweep down the line. Sweeping over the entire MPS from n_1 to n_L evolves $|q_t\rangle$ by δt.

APPENDIX D: MODIFIED ABSORBING BOUNDARY CONDITION DYNAMICS

To measure a rate from $\mathcal{A}$ into $\mathcal{B}$ without multiple crossings of $\mathcal{B}$'s boundary, we define a modified dynamics

via an effective Hamiltonian that zeroes out the rates of any reaction or diffusion event that would depart the region $\mathcal{B}$. Deriving the forms of $\tilde{H}^{\text{rxn}}$ and $\tilde{H}^{\text{diff}}$ in Eq. (30) requires that we first identify the set of states and reactions that depart $\mathcal{B}$:

$$\mathcal{B}^{\text{out}} \equiv \{\mathbf{n}, r : \mathbf{n} + \nu^r \in \mathcal{B}^c | \mathbf{n} \in \mathcal{B}\}, \qquad \text{(D1)}$$

where $\mathcal{B}^c$ is the complement of $\mathcal{B}$. For the Schögl model, only diffusion and the forward reactions can map out of $\mathcal{B}$. The states in $\mathcal{B}^{\text{out}}$ are those where particles in the lth voxel are at the boundary to $\mathcal{B}$, $n_l = q_{\mathcal{B}}^*$ and $n_k \in \mathcal{B}$ $\forall$ $k \neq l$. The local boundary for $\mathcal{B}$ in one voxel can be projected onto using $\mathcal{Q} = |q_{\mathcal{B}}^*\rangle\langle q_{\mathcal{B}}^*|$. Combining $\mathcal{Q}$ with the projector $\mathcal{N}_l = \sum_{n_l \in \mathcal{B}} |n_l\rangle\langle n_l|$ lets us define the projection operator,

$$\hat{\mathcal{P}}_l^{\text{out}} = \mathcal{N}_1 \otimes \mathcal{N}_2 \otimes \cdots \mathcal{N}_{l-1} \otimes \mathcal{Q}_l \otimes \mathcal{N}_{l+1} \otimes \cdots \mathcal{N}_L, \qquad \text{(D2)}$$

that maps onto states which are at the boundary in voxel l but anywhere in $\mathcal{B}$ in the other voxels. To build a projector that isolates all boundary configurations, we require that at least one voxel is on the boundary. We can build such a projector from the powerset expansion,

$$\hat{\mathcal{P}}^{\text{out}} = \sum_{l=1}^{L} (-1)^{l+1} \sum_{\{s_l\}} \prod_{s_l} \hat{\mathcal{P}}_{s_l}^{\text{out}}, \qquad \text{(D3)}$$

where $\{s_l\}$ is the set of combinations from L elements taken l at a time.

For $L = 3$, Eq. (D3) becomes

$$\hat{\mathcal{P}}^{out} = [\hat{\mathcal{P}}_1^{out} + \hat{\mathcal{P}}_2^{out} + \hat{\mathcal{P}}_3^{out}]$$
$$- [\hat{\mathcal{P}}_1^{out}\hat{\mathcal{P}}_2^{out} + \hat{\mathcal{P}}_1^{out}\hat{\mathcal{P}}_3^{out} + \hat{\mathcal{P}}_2^{out}\hat{\mathcal{P}}_3^{out}]$$
$$+ \hat{\mathcal{P}}_1^{out}\hat{\mathcal{P}}_2^{out}\hat{\mathcal{P}}_3^{out}.$$

Each product over $\hat{\mathcal{P}}_{s_l}^{out}$ at multiple voxels can be simplified noting that $\mathcal{N}_l\mathcal{Q}_l = \mathcal{Q}_l$, which leads to

$$\hat{\mathcal{P}}^{out} = \sum_{i=1}^{L}(-1)^{l+1}\sum_{\{c_k\}}\prod_{c_k}\{\mathcal{N},\mathcal{Q}\}_{c_k}, \qquad (D4)$$

where the sum over c_k is now all combinations of $\mathcal{N}$ and $\mathcal{Q}$ where $\mathcal{Q}$ appears k times. The advantage of Eq. (D4) is that effective Hamiltonians can be projected onto only using local (single-voxel) projectors. In this work, $\mathcal{B}$ consists of all molecule numbers less than $q_{\mathcal{B}}^* + 1$. Then, projecting onto reactions that can map out of $\mathcal{B}$ the effective Hamiltonians for the Schlögl model become

$$\tilde{H}^{rxn} = \sum_{l=1}^{L}[\hat{H}_l^{A_{rev}} + \hat{H}_l^{B_{rev}} + (\hat{H}_l^{A_{for}} + \hat{H}_l^{B_{for}})(\mathbb{I} - \hat{\mathcal{P}}^{out})], \quad (D5)$$

$$\tilde{H}^{diff} = \hat{H}^{diff}(\mathbb{I} - \hat{\mathcal{P}}^{out}). \qquad (D6)$$

APPENDIX E: ERROR IN TDVP RATE CALCULATIONS

When calculating observables using the ensemble method, sampling error is avoided, but there are several different sources for numerical error. Each source of error is controllable in that a parameter can be increased or decreased to systematically improve accuracy at the expense of a greater numerical cost. That trade-off between accuracy and cost is not straightforward, particularly because it can switch depending on which source of error dominates. The factorization of the time propagation in Eq. (C13) relied upon a infinitesimal time step, but the computational expense grows essentially linearly with the number of time steps [86]. Consequently, one seeks a time step which is small enough but not too small. We highlight that the SSA has an effective time step that emerges naturally as the typical time between subsequent events, and that this time step shrinks as the system size grows. By contrast, the TDVP approach's time step is a parameter one can choose irrespective of the system size. In addition to the time-step error, the propagation of the local tensors discussed in Appendix C is implemented via a Krylov expansion. We found that 30 Krylov basis vectors were sufficient, but the model system here is not very sensitive to
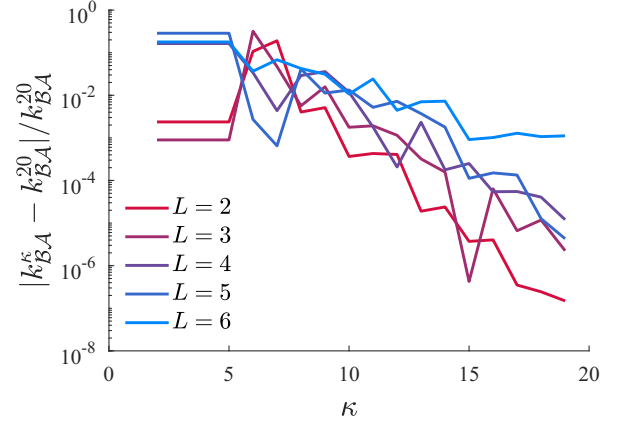

FIG. 8. The rate $k_{\mathcal{B}\mathcal{A}}$ is calculated at different values of L and different bond dimensions κ. The relative difference in rates is plotted against the "reference" rate calculated at $\kappa = 20$. For large enough bond dimension errors shrink exponentially.

the number used. We have not presented a thorough analysis of time-step or Krylov errors because it is clear that our error is dominated by the bond dimension κ, which can severely limit the variational search space of allowable MPS distributions.

In order to get an estimate of how the bond dimension impacts the numerical errors, we repeated the TDVP rate for varying κ and L. We would ideally compare each rate against an exact rate, but getting those exact rates is not numerically feasible for all system sizes considered. We have shown [see Fig. 5(c)] that for $L = 2$ through 6 the bond dimension $\kappa = 20$ was sufficient to reproduce the SSA rates to within the size of the plot markers. For the purposes of analyzing the bond-dimension convergence, we thus take that $\kappa = 20$ calculation as the "converged" reference value and measure the approach to that value as κ increases from 2 to 20. Figure 8 shows the absolute fractional error comparing the rate at a given bond dimension κ and the $\kappa = 20$ reference rate, all with the same time step and Krylov dimension. We see that errors can already be very small at modest bond dimension (where the smaller tensors make the calculations substantially faster), and the errors shrink exponentially with growing κ for all values of L.

APPENDIX F: RUN-TIME ANALYSIS

Figure 5(c) shows the total computational expense to estimate rates from the initial $\mathcal{A}$ ensemble. The dominant purpose of communicating the number of CPU hours is to illustrate the scaling of computational expense with the system size. That is to say, the brute-force SSA rates grow exponentially more costly as L increases while the TDVP cost grows subexponentially. It is enticing to compare the costs to each other. At a fixed L, was the SSA calculation or the TDVP calculation cheaper? Figure 5(c) appears to answer that question, showing that the two methods have a

crossover with TDVP becoming cheaper for sufficiently large L. We argue that the existence of the crossover is a generic consequence of the different scalings with L, but the exact location of the crossover depends on specific details of the implementations of both SSA and TDVP. Considering the difference between the $\kappa = 20$ (black) and $\kappa = 30$ (blue) lines in Fig. 5(c), one must be especially careful to compare one method versus the other. Each method has parameters that can be adjusted to improve accuracy, and to compare different CPU hour timings, one should have a sense that the two calculations are returning comparable accuracies. Even more importantly, one should recognize that although we have used the methods to compute the rate, the computational expense goes into two very different objects. In the case of SSA, that expense results in a collection of representative trajectories, while for TDVP it yields time-dependent joint distributions. If one were using the calculations to compute different quantities (say, waiting time distributions rather than just the rate), the same TDVP joint distributions that were generated to compute the rates could be used at essentially no additional computational expense. By contrast, generating a converged waiting time distribution from SSA samples would require many more than the 400 trajectories used to estimate the rate. In a similar vein, in Fig. 6 we show how the TDVP calculations give converged insight into the mechanism of the rare event, but the SSA approach would return only 400 representative events from which one would extrapolate.

Having laid out the caveats, we spell out explicitly how we computed the total computational expense for these two classes of rate calculations (further information about forward flux sampling calculations follows in the next appendix). The TN calculations are not trivial to parallelize since the TDVP algorithm proceeds from one tensor to the next in serial, with the output of one tensor operation feeding into the next one. We did not seek a sophisticated parallelization, but we did note some imperfect parallelization could be easily accessed from the way the ITensor software in the JULIA language handles the organization and contraction of tensors. Figure 9 shows that this imperfect parallelization gave a roughly twofold speedup if 8 processors were used instead of a single processor. The computational expense reported in Fig. 5(c) is the total number of CPU hours for those 8 processors to complete the single-site TDVP dynamics with time step $\delta t = 10^{-4}$ for 10000 time steps with 30 Krylov dimensions starting from an already computed distribution in $\mathcal{A}$. In a sense, this timing is artificially worse than it would need to be. If we were willing to run for twice the duration of real time, we could decrease the total number of CPU hours for TDVP rates in Fig. 5(c) by a factor of 4 with a serial implementation.

Whereas the TDVP timing would have room for improvement with more sophisticated parallelization schemes, the
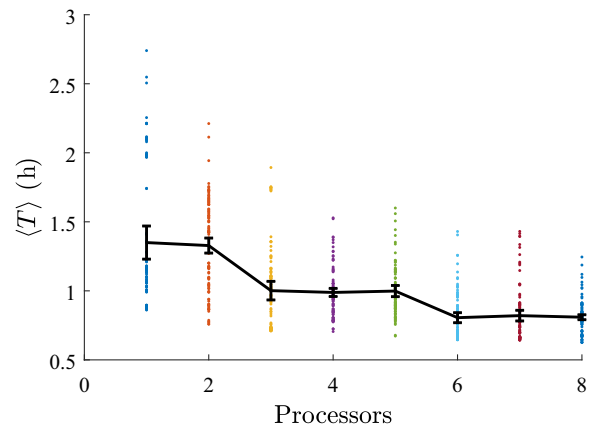

FIG. 9.   The wall time to complete 100 time steps of TDVP dynamics with $L = 3$ and $\kappa = 30$ as a function of the number of processors. Parallelization was not explicitly designed into the TDVP sweep, so the limited parallelism emerged only from the built-in capabilities of ITensor. If one could reach perfect parallelization, either through a clever parallelization scheme or by just running TDVP in serial on a single processor, the number of CPU hours needed to compute the rates in Fig. 5 would decrease by up to a factor of 4.

SSA calculations are already fully optimized in that they are perfectly parallelizable. Each stochastic trajectory can and should be run on a separate CPU. The CPU hours reported in Fig. 5(c) is thus the sum of the wall times for 400 independent realizations, initialized in $\mathcal{A}$ and run until $\mathcal{B}$ was first reached. The cost for SSA calculations would shift up and down the log scale if one decided to use more or less than the 400 independent trajectories. Our choice here was to seek a number of trajectories that could make the SSA error bars on the same order as the plot markers of Fig. 5(a).

APPENDIX G: COMPARISON WITH FORWARD FLUX SAMPLING

It is interesting to see how the computational expense of our rate calculation compares to advanced sampling rate calculations. Here we compare against the original FFS algorithm [10]. The key idea of FFS is that a (typically one-dimensional) progress coordinate is defined to decompose the single rare event from $\mathcal{A}$ to $\mathcal{B}$ down into less rare excursions that make partial progress along that order parameter. The progress coordinate is discretized to define a collection of N nonoverlapping interfaces $\lambda_1, \lambda_2, \ldots, \lambda_N$. The FFS procedure decomposes the transition rate as

$$k_{\mathcal{B}\mathcal{A}} = \Phi_{1,\mathcal{A}} \prod_{k=1}^{N-1} P(\lambda_{k+1}|\lambda_k), \tag{G1}$$

where $\Phi_{1,\mathcal{A}}$ is the flux to pass from $\mathcal{A}$ to the λ_1 interface and $P(\lambda_{k+1}|\lambda_k)$ is the probability of reaching λ_{k+1} before first returning to $\mathcal{A}$ conditioned upon initialization at the λ_k interface. To the extent that the chosen progress coordinate

acts as a good reaction coordinate, the method can radically improve the efficiency of a rate calculation compared to brute-force SSA, but if the order parameter is not "close enough" to the reaction coordinate, the conditional probability terms can become very expensive to compute [87]. It is tempting to think that $P(\lambda_{k+1}|\lambda_k)$ requires a trajectory starting at λ_k and run until it hits either the next or the previous interface, but the actual procedure requires the trajectory to hit either the next interface or $\mathcal{A}$. When the progress coordinate poorly approximates the reaction coordinate, a tremendous amount of time can be wasted simulating trajectories slowly wandering back to $\mathcal{A}$ from one of the final interfaces. The efficiency of FFS calculations additionally depends on the number of interfaces, their distance from each other, and the number of trajectories needed to resolve the conditional distributions. Given all the tunable knobs that go into a FFS rate calculation, it only makes sense to compare the numerical expense of TDVP rate calculations against a prudently designed FFS with a reasonable choice of progress coordinate.

Without *a priori* knowing the mechanism of the Schlögl reaction-diffusion switching events, we designed two different order parameters and performed calculations with both. For the first order parameter, we focus on the global number of X molecules, summed over all voxels. The interfaces for this order parameter measuring global progress are $\lambda^G \equiv \{\mathcal{A}, \lambda_1^G, \ldots, \lambda_8^G, \mathcal{B}\}$ with $\lambda_i^G = L\lambda_{\mathcal{A}}^* + \lfloor i\Delta \rfloor$, where $\Delta = (\lambda_{\mathcal{B}}^* - \lambda_{\mathcal{A}}^*)L/9$. Remember that $\lambda_{\mathcal{A}}^* = 15$ and $\lambda_{\mathcal{B}}^* = 25$, so this order parameter tracks the progress of the whole system in moving from $25L$ or more molecules to at least $15L$ molecules. The second order parameter, inspired by the demonstration in Fig. 6 that the voxels on the end decrease first, uses only the number of molecules in the first voxel as the sign of progress. The interfaces for this order parameter measuring voxel 1 are $\lambda^1 \equiv \{\mathcal{A}, \lambda_1^1, \ldots, \lambda_8^1, \mathcal{B}\}$ with each λ_i^1 interface recording the addition of another X molecule in voxel 1 ($\lambda_i^1 = 15 + i$), so that $\lambda_1^1 = 16$ and $\lambda_8^1 = 23$.

There are many ways to optimize both FFS calculations, for example, changing the number of interfaces, the number of trajectories to estimate $\Phi_{1,\mathcal{A}}$, and the number of trajectories to estimate the conditional probabilities. These changes impact the computational expense, and it is not generally knowable how to tune those parameters to most cheaply decrease the error bars of the rate calculation. We found that 4000 trajectories were sufficient to get decent estimates of the flux $\Phi_{1,\mathcal{A}}$ for both order parameters, though the expense of computing that flux was much greater for λ^G than for λ^1. While the flux was more costly to compute for λ^G, the conditional probabilities were easier. Each conditional probability could be well estimated with only 4000 additional trajectories per interface. Converging those conditional probability terms for λ^1 required 10^7 trajectories per interface, a cost that
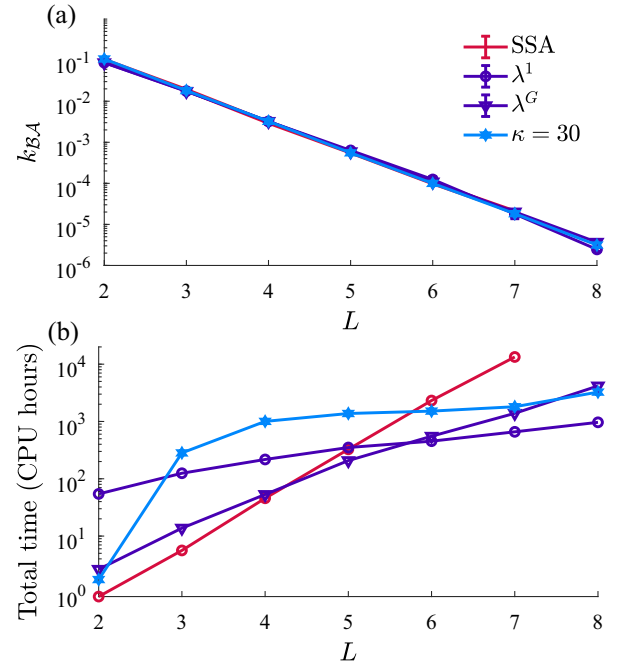

FIG. 10. (a) The rate $k_{\mathcal{BA}}$ from SSA and TDVP using $\kappa = 30$ bond dimension compared to FFS using different order parameters. Though FFS can correctly estimate the rate even when the order parameter is not exactly the reaction coordinate, the choice of order parameter significantly impacts the computational expense. It is notable that the TDVP rate calculation can bypass the introduction of an order parameter; in fact, by analyzing the TDVP-evolved joint distribution, one can reveal the reaction mechanism as an output of the calculation.

was at least partially offset by the cheaper $\Phi_{1,\mathcal{A}}$ calculation. For each order parameter we generated these samples to calculate ten independent estimates of the rate, allowing us to measure both the mean and standard error. From these standard errors, we determined that the FFS calculations were converged in line with the SSA errors in that plot.

A notable feature of FFS is that it will return the correct transition rates even when the order parameter is not a perfect reaction coordinate. Figure 10(a) reflects that in that both λ^G and λ^1 reproduce the rates given by the SSA and TDVP methods. However, the computational expense of the two are not identical. Figure 10(b) plots the CPU hours needed to generate the samples that yield the means and standard errors of Fig. 10(a). At small L, λ^1 is faster; at large L, λ^G is faster. The FFS calculations have many tunable parameters. They could likely be improved through further refinement of order parameters, changing the number of interfaces, fine-tuning the number of samples for each conditional probability, etc. Still, we find it notable that our reasonable but perhaps unoptimized FFS implementations perform with similar (and slightly greater) expense as the TDVP rate calculations. This observation

is particularly notable since the TDVP calculations do not need one to guess an order parameter and since they offer converged statistical information about the mechanism in addition to just giving the rate.

---

[1] H. Eyring, *The Activated Complex in Chemical Reactions*, J. Chem. Phys. **3**, 107 (1935).

[2] K. J. Laidler and M. C. King, *The Development of Transition-State Theory*, J. Phys. Chem. **87**, 2657 (1983).

[3] D. G. Truhlar and B. C. Garrett, *Variational Transition-State Theory*, Acc. Chem. Res. **13**, 440 (1980).

[4] R. F. Grote and J. T. Hynes, *The Stable States Picture of Chemical Reactions. II. Rate Constants for Condensed and Gas Phase Reaction Models*, J. Chem. Phys. **73**, 2715 (1980).

[5] J. T. Hynes, *Chemical Reaction Dynamics in Solution*, Annu. Rev. Phys. Chem. **36**, 573 (1985).

[6] C. Dellago, P. G. Bolhuis, F. S. Csajka, and D. Chandler, *Transition Path Sampling and the Calculation of Rate Constants*, J. Chem. Phys. **108**, 1964 (1998).

[7] P. G. Bolhuis, D. Chandler, C. Dellago, and P. L. Geissler, *Transition Path Sampling: Throwing Ropes over Rough Mountain Passes, in the Dark*, Annu. Rev. Phys. Chem. **53**, 291 (2002).

[8] G. E. Crooks and D. Chandler, *Efficient Transition Path Sampling for Nonequilibrium Stochastic Dynamics*, Phys. Rev. E **64**, 026109 (2001).

[9] T. R. Gingrich and P. L. Geissler, *Preserving Correlations between Trajectories for Efficient Path Sampling*, J. Chem. Phys. **142**, 234104 (2015).

[10] R. J. Allen, P. B. Warren, and P. R. ten Wolde, *Sampling Rare Switching Events in Biochemical Networks*, Phys. Rev. Lett. **94**, 018104 (2005).

[11] R. J. Allen, D. Frenkel, and P. R. ten Wolde, *Simulating Rare Events in Equilibrium or Nonequilibrium Stochastic Systems*, J. Chem. Phys. **124**, 024102 (2006).

[12] R. J. Allen, C. Valeriani, and P. R. ten Wolde, *Forward Flux Sampling for Rare Event Simulations*, J. Phys. Condens. Matter **21**, 463102 (2009).

[13] A. Ma and A. R. Dinner, *Automatic Method for Identifying Reaction Coordinates in Complex Systems*, J. Phys. Chem. B **109**, 6769 (2005).

[14] S. A. Isaacson and C. S. Peskin, *Incorporating Diffusion in Complex Geometries into Stochastic Chemical Kinetics Simulations*, SIAM J. Sci. Comput. **28**, 47 (2006).

[15] R. Erban, J. Chapman, and P. Maini, *A Practical Guide to Stochastic Simulations of Reaction-Diffusion Processes*, arXiv:0704.1908.

[16] D. Fange, O. G. Berg, P. Sjöberg, and J. Elf, *Stochastic Reaction-Diffusion Kinetics in the Microscopic Limit*, Proc. Natl. Acad. Sci. U.S.A. **107**, 19820 (2010).

[17] D. J. Wilkinson, *Stochastic Modelling for Systems Biology* (CRC Press, Boca Raton, FL, 2018).

[18] S. Smith and R. Grima, *Spatial Stochastic Intracellular Kinetics: A Review of Modelling Approaches*, Bull. Math. Biol. **81**, 2960 (2019).

[19] G. Craciun, Y. Tang, and M. Feinberg, *Understanding Bistability in Complex Enzyme-Driven Reaction Networks*, Proc. Natl. Acad. Sci. U.S.A. **103**, 8697 (2006).

[20] A. L. Krause, E. A. Gaffney, P. K. Maini, and V. Klika, *Modern Perspectives on Near-Equilibrium Analysis of Turing Systems*, Phil. Trans. R. Soc. A **379**, 20200268 (2021).

[21] S. Kondo and T. Miura, *Reaction-Diffusion Model as a Framework for Understanding Biological Pattern Formation*, Science **329**, 1616 (2010).

[22] A. Turing, *The Chemical Basis of Morphogenesis*, Phil. Trans. R. Soc. B **237**, 37 (1952).

[23] M. C. Cross and P. C. Hohenberg, *Pattern Formation Outside of Equilibrium*, Rev. Mod. Phys. **65**, 851 (1993).

[24] D. A. McQuarrie, *Stochastic Approach to Chemical Kinetics*, J. Appl. Probab. **4**, 413 (1967).

[25] R. E. Lee DeVille, C. B. Muratov, and E. Vanden-Eijnden, *Non-Meanfield Deterministic Limits in Chemical Reaction Kinetics*, J. Chem. Phys. **124**, 231102 (2006).

[26] H. H. McAdams and A. Arkin, *Stochastic Mechanisms in Gene Expression*, Proc. Natl. Acad. Sci. U.S.A. **94**, 814 (1997).

[27] M. Thattai and A. van Oudenaarden, *Intrinsic Noise in Gene Regulatory Networks*, Proc. Natl. Acad. Sci. U.S.A. **98**, 8614 (2001).

[28] J. M. G. Vilar, H. Y. Kueh, N. Barkai, and S. Leibler, *Mechanisms of Noise-Resistance in Genetic Oscillators*, Proc. Natl. Acad. Sci. U.S.A. **99**, 5988 (2002).

[29] J. Paulsson, *Summing Up the Noise in Gene Networks*, Nature (London) **427**, 415 (2004).

[30] T. Tian and K. Burrage, *Stochastic Models for Regulatory Networks of the Genetic Toggle Switch*, Proc. Natl. Acad. Sci. U.S.A. **103**, 8372 (2006).

[31] W. J. Heuett and H. Qian, *Grand Canonical Markov Model: A Stochastic Theory for Open Nonequilibrium Biochemical Networks*, J. Chem. Phys. **124**, 044110 (2006).

[32] D. T. Gillespie, A. Hellander, and L. R. Petzold, *Perspective: Stochastic Algorithms for Chemical Kinetics*, J. Chem. Phys. **138**, 170901 (2013).

[33] C. Yang, D. T. Gillespie, and L. R. Petzold, *Efficient Step Size Selection for the Tau-Leaping Simulation Method*, J. Chem. Phys. **124**, 044109 (2006).

[34] B. Munsky and M. Khammash, *The Finite State Projection Algorithm for the Solution of the Chemical Master Equation*, J. Chem. Phys. **124**, 044104 (2006).

[35] V. Kazeev, M. Khammash, M. Nip, and C. Schwab, *Direct Solution of the Chemical Master Equation Using Quantized Tensor Trains*, PLoS Comput. Biol. **10**, e1003359 (2014).

[36] H. D. Vo and R. B. Sidje, *An Adaptive Solution to the Chemical Master Equation Using Tensors*, J. Chem. Phys. **147**, 044102 (2017).

[37] P. Helms, U. Ray, and G. K.-L. Chan, *Dynamical Phase Behavior of the Single- and Multi-Lane Asymmetric Simple Exclusion Process via Matrix Product States*, Phys. Rev. E **100**, 022101 (2019).

[38] N. E. Strand, H. Vroylandt, and T. R. Gingrich, *Using Tensor Network States for Multi-Particle Brownian Ratchets*, J. Chem. Phys. **156**, 221103 (2022).

[39] N. E. Strand, H. Vroylandt, and T. R. Gingrich, *Computing Time-Periodic Steady-State Currents via the Time Evolution*

of Tensor Network States, J. Chem. Phys. **157,** 054104 (2022).

[40] F. Schlögl, *Chemical Reaction Models for Non-Equilibrium Phase Transitions*, Z. Phys. **253,** 147 (1972).

[41] C. Kim, A. Nonaka, J. B. Bell, A. L. Garcia, and A. Donev, *Stochastic Simulation of Reaction-Diffusion Systems: A Fluctuating-Hydrodynamics Approach*, J. Chem. Phys. **146,** 124110 (2017).

[42] M. Vellela and H. Qian, *Stochastic Dynamics and Non-Equilibrium Thermodynamics of a Bistable Chemical System: The Schlögl Model Revisited*, J. R. Soc. Interface **6,** 925 (2009).

[43] S. B. Nicholson, J. S. Greenberg, and J. R. Green, *Entrance and Escape Dynamics for the Typical Set*, Phys. Rev. E **97,** 012146 (2018).

[44] D. J. Higham, *Modeling and Simulating Chemical Reactions*, SIAM Rev. **50,** 347 (2008).

[45] G. Nicolis and J. W. Turner, *Stochastic Analysis of a Nonequilibrium Phase Transition: Some Exact Results*, Physica (Amsterdam) **89A,** 326 (1977).

[46] B. Roux, *Transition Rate Theory, Spectral Analysis, and Reactive Paths*, J. Chem. Phys. **156,** 134111 (2022).

[47] W. E and E. Vanden-Eijnden, *Transition-Path Theory and Path-Finding Algorithms for the Study of Rare Events.*, Annu. Rev. Phys. Chem. **61,** 391 (2010).

[48] J. Ohkubo, *Algebraic Probability, Classical Stochastic Processes, and Counting Statistics*, J. Phys. Soc. Jpn. **82,** 084001 (2013).

[49] U. C. Täuber, *Critical Dynamics: A Field Theory Approach to Equilibrium and Non-Equilibrium Scaling Behavior* (Cambridge University Press, Cambridge, England, 2014).

[50] M. Doi, *Stochastic Theory of Diffusion-Controlled Reaction*, J. Phys. A **9,** 1479 (1976).

[51] L. Peliti, *Path Integral Approach to Birth-Death Processes on a Lattice*, J. Phys. (Paris) **46,** 1469 (1985).

[52] P.-A. Rey and J. Cardy, *Asymptotic Form of the Approach to Equilibrium in Reversible Recombination Reactions*, J. Phys. A **32,** 1585 (1999).

[53] A. Reyes-Velázquez, A. Molgado, J. Berra-Montiel, and J. A. Martinez-Gonzalez, *A General Path-Integral Monte Carlo Method for Exact Simulations of Chemical Reaction Networks*, J. Phys. Chem. A **127,** 4363 (2023).

[54] J. J. Vastola, *Solving the Chemical Master Equation for Monomolecular Reaction Systems and Beyond: A Doi-Peliti Path Integral View*, J. Math. Biol. **83,** 48 (2021).

[55] X. Yang, P. Liu, and J. Zhang, *Analytical Time-Dependent Distributions for Two Common Signaling Systems*, Chin. J. Phys. **55,** 289 (2017).

[56] M. Harsh and P. Sollich, *Accurate Dynamics from Self-Consistent Memory in Stochastic Chemical Reactions with Small Copy Numbers*, arXiv:2303.00029.

[57] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, and F. van Wijland, *Dynamical First-Order Phase Transition in Kinetically Constrained Models of Glasses*, Phys. Rev. Lett. **98,** 195702 (2007).

[58] J. P. Garrahan, R. L. Jack, V. Lecomte, E. Pitard, K. van Duijvendijk, and F. van Wijland, *First-Order Dynamical Phase Transition in Models of Glasses: An Approach Based on Ensembles of Histories*, J. Phys. A **42,** 075007 (2009).

[59] Y. Lan, P. G. Wolynes, and G. A. Papoian, *A Variational Approach to the Stochastic Aspects of Cellular Signal Transduction*, J. Chem. Phys. **125,** 124106 (2006).

[60] E. DeGiuli and C. Scalliet, *Dynamical Mean-Field Theory: From Ecosystems to Reaction Networks*, J. Phys. A **55,** 474002 (2022).

[61] M. Bothe and G. Pruessner, *Doi-Peliti Field Theory of Free Active Ornstein-Uhlenbeck Particles*, Phys. Rev. E **103,** 062105 (2021).

[62] U. Schollwöck, *The Density-Matrix Renormalization Group in the Age of Matrix Product States*, Ann. Phys. (Amsterdam) **326,** 96 (2011).

[63] L. Causer, M. C. Bañuls, and J. P. Garrahan, *Finite Time Large Deviations via Matrix Product States*, Phys. Rev. Lett. **128,** 090605 (2022).

[64] M. J. del Razo, D. Frömberg, A. V. Straube, C. Schütte, F. Höfling, and S. Winkelmann, *A Probabilistic Framework for Particle-Based Reaction-Diffusion Dynamics Using Classical Fock Space Representations*, Lett. Math. Phys. **112,** 49 (2022).

[65] M. Fishman, S. R. White, and E. M. Stoudenmire, *The ITensor Software Library for Tensor Network Calculations*, SciPost Phys. Codebases 4 (2022).

[66] R. Orús, *A Practical Introduction to Tensor Networks: Matrix Product States and Projected Entangled Pair States*, Ann. Phys. (Amsterdam) **349,** 117 (2014).

[67] P. A. Dirac, *Note on Exchange Phenomena in the Thomas Atom*, in *Mathematical Proceedings of the Cambridge Philosophical Society*, Vol. 26 (Cambridge University Press, Cambridge, England, 1930), pp. 376–385.

[68] J. Haegeman, J. I. Cirac, T. J. Osborne, I. Pižorn, H. Verschelde, and F. Verstraete, *Time-Dependent Variational Principle for Quantum Lattices*, Phys. Rev. Lett. **107,** 070601 (2011).

[69] L. Vanderstraeten, J. Haegeman, and F. Verstraete, *Tangent-Space Methods for Uniform Matrix Product States*, SciPost Phys. Lect. Notes 7 (2019).

[70] S. Paeckel, T. Köhler, A. Swoboda, S. R. Manmana, U. Schollwöck, and C. Hubig, *Time-Evolution Methods for Matrix-Product States*, Ann. Phys. (Amsterdam) **411,** 167998 (2019).

[71] J. Haegeman, C. Lubich, I. Oseledets, B. Vandereycken, and F. Verstraete, *Unifying Time Evolution and Optimization with Matrix Product States*, Phys. Rev. B **94,** 165116 (2016).

[72] C. Lubich, I. V. Oseledets, and B. Vandereycken, *Time Integration of Tensor Trains*, SIAM J. Numer. Anal. **53,** 917 (2015).

[73] G. Vidal, *Efficient Classical Simulation of Slightly Entangled Quantum Computations*, Phys. Rev. Lett. **91,** 147902 (2003).

[74] M. Vlysidis and Y. N. Kaznessis, *On Differences between Deterministic and Stochastic Models of Chemical Reactions: Schlögl Solved with ZI-Closure*, Entropy **20,** 678 (2018).

[75] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah, *Julia: A Fresh Approach to Numerical Computing*, SIAM Rev. **59,** 65 (2017).

[76] J. J. Tyson, *Some Further Studies of Nonlinear Oscillations in Chemical Systems*, J. Chem. Phys. **58,** 3919 (1973).

[77] R. J. Field and R. M. Noyes, *Oscillations in Chemical Systems. IV. Limit Cycle Behavior in a Model of a Real Chemical Reaction*, J. Chem. Phys. **60**, 1877 (1974).

[78] M. Schulz and P. Reineker, *Exact Substitute Processes for Diffusion-Reaction Systems with Local Complete Exclusion Rules*, New J. Phys. **7**, 31 (2005).

[79] T. Litschel, M. M. Norton, V. Tserunyan, and S. Fraden, *Engineering Reaction-Diffusion Networks with Properties of Neural Tissue*, Lab Chip **18**, 714 (2018).

[80] P. B. Warren and P. R. ten Wolde, *Enhancement of the Stability of Genetic Switches by Overlapping Upstream Regulatory Domains*, Phys. Rev. Lett. **92**, 128101 (2004).

[81] D. J. Wilkinson, *Stochastic Modelling for Quantitative Description of Heterogeneous Biological Systems*, Nat. Rev. Genet. **10**, 122 (2009).

[82] M. J. Hallock, J. E. Stone, E. Roberts, C. Fry, and Z. Luthey-Schulten, *Simulation of Reaction Diffusion Processes over Biologically Relevant Size and Time Scales Using Multi-GPU Workstations*, Parallel Comput. **40**, 86 (2014).

[83] C. Hubig, I. P. McCulloch, U. Schollwöck, and F. A. Wolf, *Strictly Single-Site DMRG Algorithm with Subspace Expansion*, Phys. Rev. B **91**, 155115 (2015).

[84] J. Cardy, *Field Theory and Non-Equilibrium Statistical Mechanics, 2006*, University of Warwick, https://www-thphys.physics.ox.ac.uk/people/JohnCardy/warwick.pdf.

[85] J. Haegeman, T. J. Osborne, and F. Verstraete, *Post-Matrix Product State Methods: To Tangent Space and Beyond*, Phys. Rev. B **88**, 075133 (2013).

[86] We say essentially linearly rather than linearly because as time progresses the distribution can build richer correlations which can result in singular value decompositions that utilize the full bond dimension κ. Distributions with less rich correlations can be described by smaller tensors if the singular values are truncated at a threshold.

[87] S. Hussain and A. Haji-Akbari, *Studying Rare Events Using Forward-Flux Sampling: Recent Breakthroughs and Future Outlook*, J. Chem. Phys. **152**, 060901 (2020).