nature machine intelligence

Article

https://doi.org/10.1038/s42256-023-00632-6

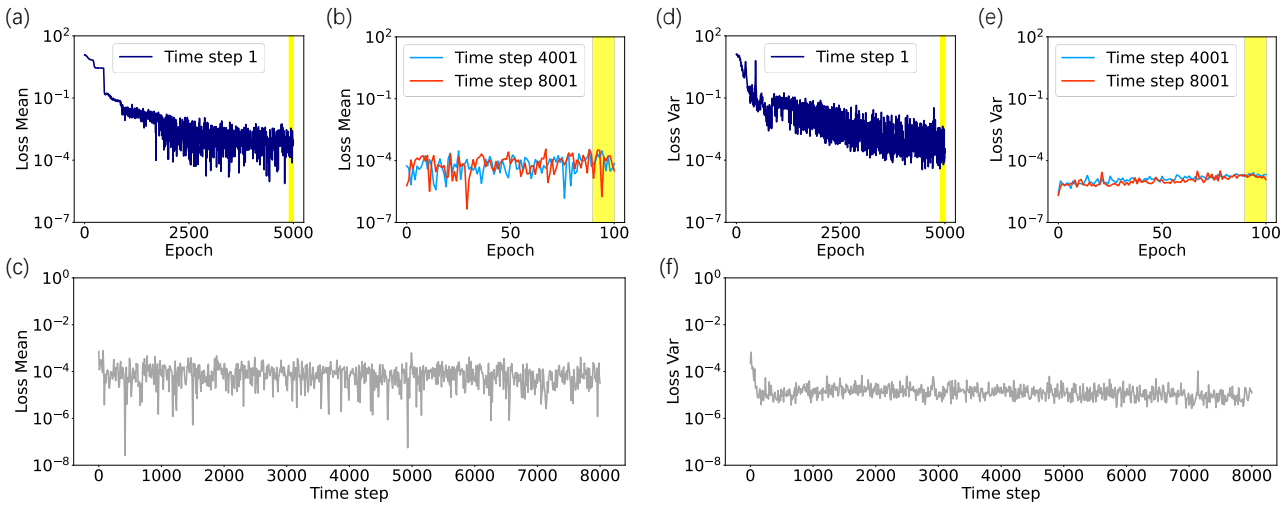
Neural-network solutions to stochastic reaction networks

In the format provided by the authors and unedited

CONTENTS

In the Supplementary Information, we first give the detail on evaluating the accuracy of training the VAN by the loss function. We also provide the application of the method to more examples in addition to those in the main text, including the birth-death process, gene expression without regulation, an autoregulatory feedback loop, the early life self-replicator, the epidemic model and more cases of the intracellular signaling cascade.

I. SUPPLEMENTARY NOTE 1: EVALUATING THE ACCURACY OF TRAINING



Supplementary Figure 1. The loss function during the training on the VAN for the genetic toggle switch. (a) The loss mean during the training at the first time step with 5000 epochs. The highlighted area denotes the converged regime with the last 10 training epochs, where we take the samples to evaluate statistics of the learnt distribution. (b) The loss mean during the training at 4001 and 8001 time steps with 100 epochs. The last 10 training epochs are also highlighted. (c) The converged value of the loss mean over time. (d-f) The loss variance under the same layout as (a-c).

In this section, we demonstrate a procedure to evaluate the accuracy of the VAN and the error accumulation of tracking the time-evolved distribution, by the loss function. This procedure is especially useful when no other methods of tracking the distribution is at hand for a comparison. We take the genetic toggle switch as an example for the demonstration.

We show the loss function during the training and the time evolution of the system. At the first time step, with the increase of training epoch, the average KL-divergence (Supplementary Fig. 1a) and the sample variance of the loss function (Supplementary Fig. 1d) decrease and gradually converge. For the later time of tracking the distribution such as at the time steps 4001 and 8001, the average KL-divergence (Supplementary Fig. 1b) and the sample variances (Supplementary Fig. 1e) also remains around $\mathcal{O}(10^{-4})$. In addition, we highlight the last 10 epochs for each chosen time step, and take the samples in this converged region to estimate the statistics.

We further provide the mean value of the converged loss over time points, which is typically around $\mathcal{O}(10^{-4})$ (Supplementary Fig. 1c). The converged value of the loss variance is around $\mathcal{O}(10^{-5})$ (Supplementary Fig. 1f). These results indicate that the variational training effectively learns the distribution over time. Therefore, we can examine the KL-divergence to quantify the convergence of the numerical calculation and the accuracy of the trained VAN. To further evaluate the accuracy of the VAN, we can compare the result with other methods when they are available.

To estimate statistics of the learnt distribution, we use samples over the last 10 epochs where the loss function already converges, as illustrated in Supplementary Fig. 1. Alternatively, it is common to use samples from the same probability distribution at the last training step. To compare the accuracy of these two methods, we also evaluate the statistics by drawing all the samples from the final distribution after the training. The results from the two ways of sampling are similar, such as for the genetic toggle switch in Fig. 2 and Supplementary Fig. 7, validating the accuracy of our sampling method. After the loss function converges, the learnt distributions over the last few epochs have similar statistics. Thus, using samples at these epochs not only increases the number of samples to evaluate the statistics, but also save the additional computational time of drawing a large batch of samples from the final distribution after the training.

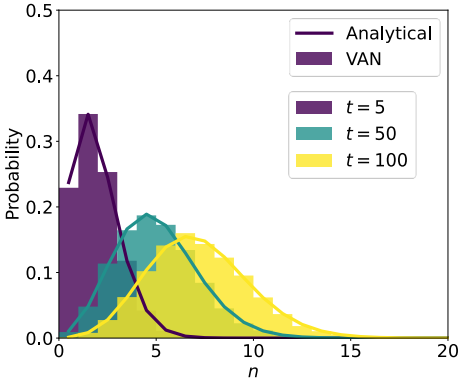
II. SUPPLEMENTARY NOTE 2: MORE EXAMPLES

A. Birth-death process

As the first example, we consider the time-homogeneous birth-death process [1] with the count of the species in the interval $X \in [0, N]$. The reactions are $\emptyset \xrightarrow{k_2} X$, and $X \xrightarrow{k_1} \emptyset$. Its dynamics are described by a time-dependent probability distribution $P_t(n)$ with the count n at time t. The chemical master equation is:

$$\partial_t P_t(n) = B(n+1)P_t(n+1) + F(n-1)P_t(n-1) - [B(n) + F(n)]P_t(n), \quad (0 \leq n \leq N), \tag{S2.1}$$

where $B(n)$ and $F(n)$ are backward (death) and forward (birth) propensities, respectively. We consider the case of propensities: $F(n) = k_2$, $B(n) = k_1 n$ with k_1, k_2 as rates, for $0 \leq n \leq N$ and zero otherwise, with the boundary conditions $F(N) = 0$, $B(0) = 0$.



Supplementary Figure 2. The result from the VAN for the birth-death process. The distributions of the species between the VAN and the analytical solution match over time. The bar is the result from the VAN, and the line is the analytical solution Eq. (S2.3). The color specifies the time points $t = 5, 50, 100$, and the transparency is used to visualize the overlap. Parameters: $k_1 = 0.01$, $k_2 = 0.1$. The hyperparameters of the VAN are in Table I, and the initial distribution is the Poisson distribution.

The birth-death process has an analytical solution. With the Poisson initial distribution:

$$P_0(n) = \frac{e^{-\alpha_0}\alpha_0^n}{n!}, \tag{S2.2}$$

where α_0 is a parameter, the time-dependent distribution is [1]:

$$P_t(n) = \frac{e^{-\alpha_t}\alpha_t^n}{n!}, \tag{S2.3}$$

$$\alpha_t = \alpha_0 e^{-k_1 t} + (k_2/k_1)(1 - e^{-k_1 t}). \tag{S2.4}$$

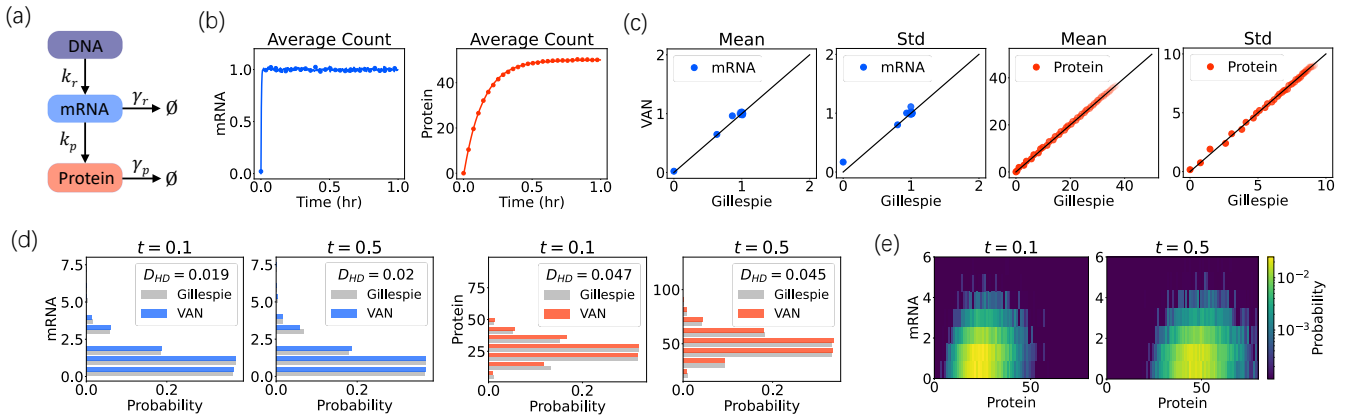
The analytical solution makes it convenient to test the present numerical approach. In Supplementary Fig. 2, the distributions from the analytical solution and the present approach match well.

B. Gene expression

We consider the example of gene expression without regulation, as one of the standard examples in the systems biology. The system involves two species, mRNA and protein. The four reactions include the transcription from DNA to mRNA, the transcription from mRNA to protein, and the decay of the mRNA and protein. The parameters are the transcription rate k_r, transcription rate k_p, and decay rates γ_r, γ_p. It has the chemical reaction:

$$DNA \xrightarrow{k_r} r, \quad r \xrightarrow{k_p} r + p, \quad r \xrightarrow{\gamma_r} \emptyset, \quad p \xrightarrow{\gamma_p} \emptyset. \tag{S2.5}$$

where r denotes the mRNA, p denotes the protein. The parameter values used in the simulation are listed in Supplementary Fig. 3.



Supplementary Figure 3. The result of gene expression without regulation. (a) A schematic of the chemical reaction. (b) The time series of the average counts for the mRNA and protein specified by color, from the VAN (dots) and the Gillespie simulation (lines). (c) Comparison of the means and standard deviations of the mRNA and protein between the VAN and the Gillespie simulation, at time points $t = 0, 0.004, 0.008, ..., 0.196, 0.2$ denoted by the transparency. (d) The marginal distributions of the mRNA and protein at time points $t = 0.1, 0.5$, from the Gillespie simulation (gray) and the VAN with the same color in (b). The inset shows the Hellinger distance between the two distributions. (e) The joint distribution of mRNAs and proteins at time points $t = 0.1, 0.5$ from the VAN, with the color as the probability values in the logarithmic scale. The parameters are $k_r = 0.1$, $k_p = 0.1$, $\gamma_r = 0.1$, and $\gamma_p = 0.002$ with the unit of time in seconds, and the Gillespie simulation has 10^4 trajectories. The hyperparameters of the VAN are in Table I. The initial distribution is the delta distribution with zero mRNA and protein.

In Supplementary Fig. 3, we show the mean, the marginal distribution and the joint distribution of the counts of the mRNA and protein over time, by using Gillespie simulation or the VAN method. The marginal statistics match well between the two methods. We quantify the similarity of the two distributions by the Hellinger distance, which is below 0.05. This validates the accuracy of the VAN in tracking the distribution.

C. Autoregulatory feedback loop

We consider an autoregulatory feedback loop [2] as illustrated in Supplementary Fig. 4a. The gene has two promoter states G_b, G_u switching to each other, with a binding rate σ_b and an unbinding rate σ_u. The two states of bond and unbound have different translation rates. The model has the chemical reactions:

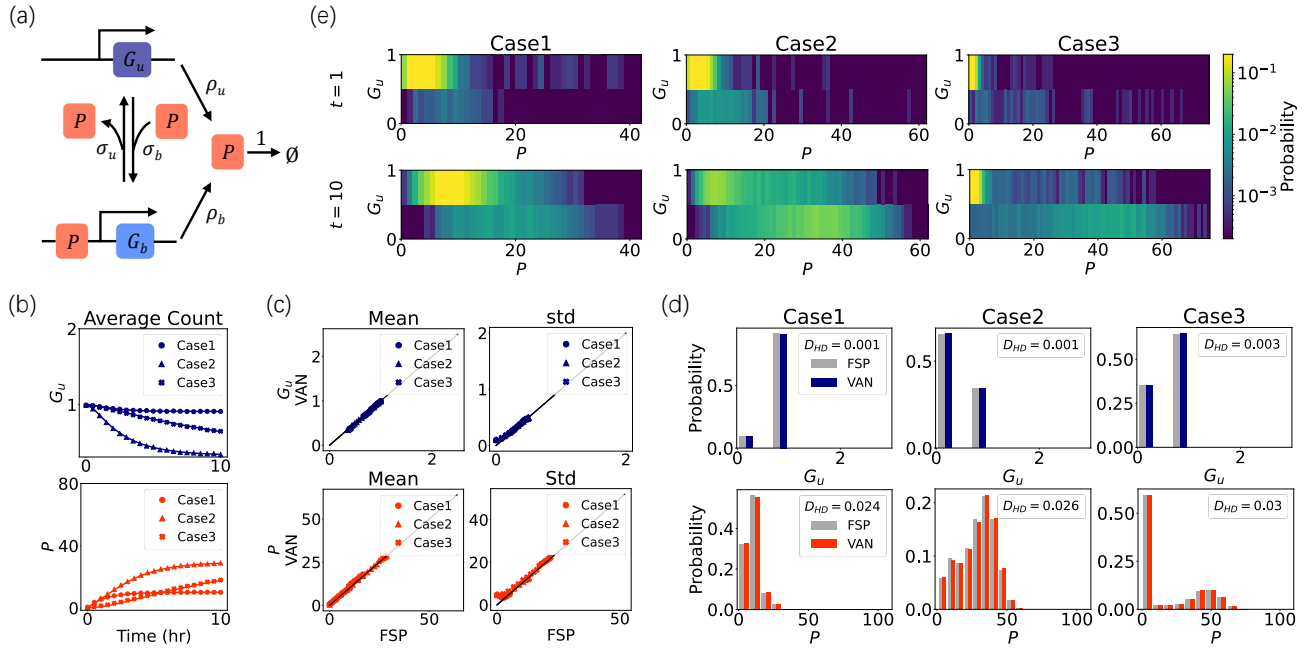
$$G_u + P \xrightarrow{\sigma_b} G_b, \quad G_b \xrightarrow{\sigma_u} G_u + P, \quad G_u \xrightarrow{\rho_u} P, \quad G_b \xrightarrow{\rho_b} P, \quad P \xrightarrow{1} \emptyset. \tag{S2.6}$$

where $\sigma_b, \sigma_u, \rho_b, \rho_u$ are rate constants with the unit of time in hours. The total count of the two promoter states is conserved: $G_b + G_u = 1$. This conservation effectively reduces a variable by having $G_u = 1 - G_b$, and sets a constraint on the counts of the two promoters: $G_u = 0, 1$, $G_b = 0, 1$. We have put this constraint on the neural network of the VAN. The three groups of parameters under the consideration [2] are listed in Supplementary Table I.

Parameter	σ_u	σ_b	ρ_u	ρ_b
Case 1	0.94	0.01	8.40	28.1
Case 2	0.69	0.07	7.20	40.6
Case 3	0.44	0.08	0.94	53.1

Supplementary Table I. The values of the parameters used in the three cases of the autoregulatory feedback loop.

By the present approach, the means (Supplementary Fig. 4b), the standard deviations (Supplementary Fig. 4c) and the marginal distributions (Supplementary Fig. 4d) of the counts of the protein and promoter match those from the finite state projection, for the three sets of parameters. As shown in the marginal distribution (Supplementary Fig. 4d) and the joint distribution (Supplementary Fig. 4e), the three cases include unimodal and bimodal distributions for the count of the protein. The results demonstrate that the VAN generates accurate marginal statistics and can effectively produce the bimodal probability distribution for the reaction network with the feedback loop.



Supplementary Figure 4. The result of the autoregulatory feedback loop. (a) A schematic of the reactions. (b) The time evolution of the average counts for the promoter (blue) and protein (red), from the VAN (symbols) and the finite state projection (FSP) (lines). Three cases with different parameters in Supplementary Table I are marked with different symbols (circle, triangle and cross). (c) Comparison on the means and standard deviations of the promoter and protein between the VAN and the FSP, at time points $t = 0, 0.5, ..., 9.5, 10$. (d) The marginal distributions of the promoter and protein at time points $t = 10$ from the FSP (gray) and the VAN with the same color in (b). The inset shows the Hellinger distance between the two distributions. (e) The joint distribution of the promoter and protein calculated from the VAN at time points $t = 1, 10$ in different cases, with the color bar for the probability values in the logarithmic scale. The hyperparameters of the VAN are in Table I, and the initial distribution is the delta distribution with one G_u and zero protein.

D. Early life self-replicator

We consider a chemical reaction modeling the biological homochirality of an early life self-replicator [3], which is an extension of the original model of spontaneous asymmetric synthesis [4]. One property of the model is that the species intrinsically has a total count conservation. To maintain count conservation, we use the implementation of the VAN in the Methods section, which effectively learns the distribution with conservation.

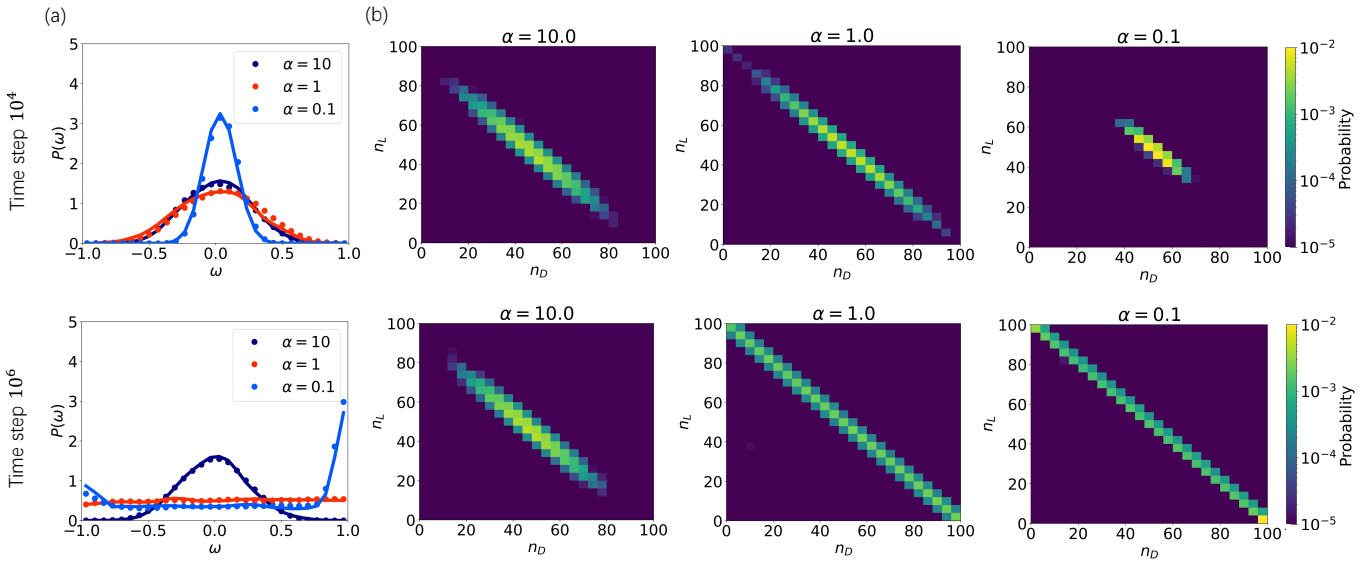
The system contains three species A, D, L, six reactions and three rate constants. The enantiomers D, L of a chiral molecule are produced autocatalytically from an achiral molecule A. It has the following chemical reactions.

$$A + D \xrightarrow{k_a} 2D, \quad A + L \xrightarrow{k_a} 2L, \quad A \xrightarrow{k_n} D, \quad D \xrightarrow{k_d} A, \quad A \xrightarrow{k_n} L, \quad L \xrightarrow{k_d} A. \tag{S2.7}$$

The rate constants k_a, k_n, k_d have the subscript identify the specific reaction. Another parameter $\alpha = Vk_n/k_a$ is introduced, corresponding to the volume V in which the reactions occur.

The state of this system is described by a chiral order parameter $\omega = (n_D - n_L)/(n_D + n_L)$ from the counts of D, L. The bifurcation on the stability at the steady state was found, which is controlled by the parameter α. The steady-state distribution is given analytically by $P_{ss}(\omega) = \mathcal{N}(1-\omega^2)^{\alpha-1}$, where $\mathcal{N} = \Gamma(\alpha + 1/2)/[\sqrt{\pi}\Gamma(\alpha)]$ and V denotes the volume [3]. At the critical value of the parameter α_c, the system changes between a unimodal system with the order parameter $\omega = 0$ in the racemic state, and a bimodal system with $\omega = \pm 1$ in the homochiral states, which was confirmed by the Gillespie simulation. The bimodal distribution was found to have two probability peaks concentrated on the two boundaries of the order parameter, $\omega = \pm 1$.

We consider three cases of the critical parameter: $\alpha_c = 10, 1, 0.1$ as above, equal, and below the critical value $\alpha_c \approx 1$. In the long-time regime, the probability distribution $P(\omega)$ by the VAN has unimodal and bimodal shapes for $\alpha_c = 10, 0.1$ separately (Supplementary Fig. 5a), which matches the Gillespie simulation. The bimodal distribution has a higher peak at the right side of the boundary, because the initial distribution has more $n_D = 50$ than $n_L = 47$, and will converge to the steady state with an equal height of two peaks after longer simulations. The two sharp peaks at the boundary $\omega = \pm 1$ are sensitive to the collected samples for performing the histogram, which affects



Supplementary Figure 5. The result of the early life self-replicator. (a) The order parameter $\omega = (n_D - n_L)/(n_D + n_L)$ from the VAN (dots) and the Gillespie simulation (line), at the time steps 10^4 and 10^6, where the latter reaches close to the steady state. The values of the critical parameter $\alpha_c = 10, 1, 0.1$ separately lead to unimodal, flat, and bimodal distributions in the long-time regime, as denoted by the color. (b) The joint distribution of D, L at time steps 10^4 and 10^6. The long-time joint distribution also shows unimodal, flat, and bimodal shapes (higher probabilities at the corner), where the color bar denotes the probability values in the logarithmic scale. The parameters are $k_a = k_n = k_d = 1$, and $\alpha = Vk_n/k_a$ with the volume V. The initial distribution is the delta distribution at $n_A = 3$, $n_D = 50$, $n_L = 47$ with the total count $N_0 = 100$. The Gillespie simulation has 10^3 trajectories. The hyperparameters of the VAN are in Table I, with the recorded time for the case of $\alpha_c = 1$. The time step length for the VAN is $\delta t = 4 * 10^{-3}, 4 * 10^{-4}, 4 * 10^{-5}$ for $\alpha = 10, 1, 0.1$, as the three cases have different time scales.

both the Gillespie simulation and the VAN approach. Increasing the batch samples in the VAN and the trajectories in the Gillespie simulation can lead to a more accurate estimation. Different from simulating trajectories by the Gillespie algorithm, the VAN approach generates the evolving joint probability distribution in the state space of D, L (Supplementary Fig. 5b). Starting from a unimodal distribution, the VAN approach reveals the emergence of unimodal and bimodal joint distributions over time.

E. Epidemic model

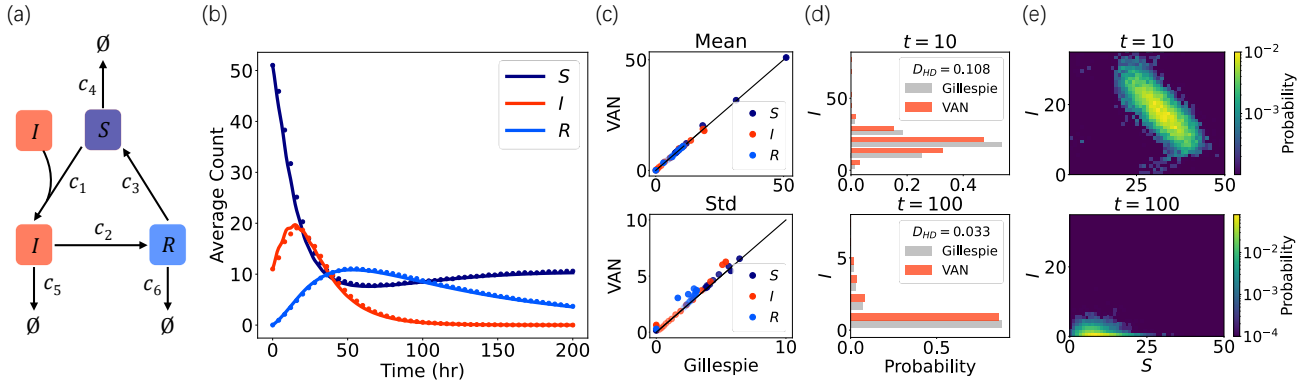
We consider the epidemic SIR model with a time-dependent contact rate [5], which demonstrates the effectiveness of our method in the system with time-dependent rates. The system models the transmission of an infectious disease, and contains three species: susceptible S, infected I and recovered R. A susceptible population S can be infected when it comes into contact with infected population I. The infected population I may become the recovered population R, which can become susceptible again. Each population dies at a certain rate.

A schematic of the chemical reaction is illustrated in Supplementary Fig. 6a. It has the following chemical reactions:

$$S + I \xrightarrow{c_1} 2I, \quad I \xrightarrow{c_2} R, \quad R \xrightarrow{c_3} S, \quad S \xrightarrow{c_4} \emptyset, \quad I \xrightarrow{c_5} \emptyset, \quad R \xrightarrow{c_6} \emptyset, \tag{S2.8}$$

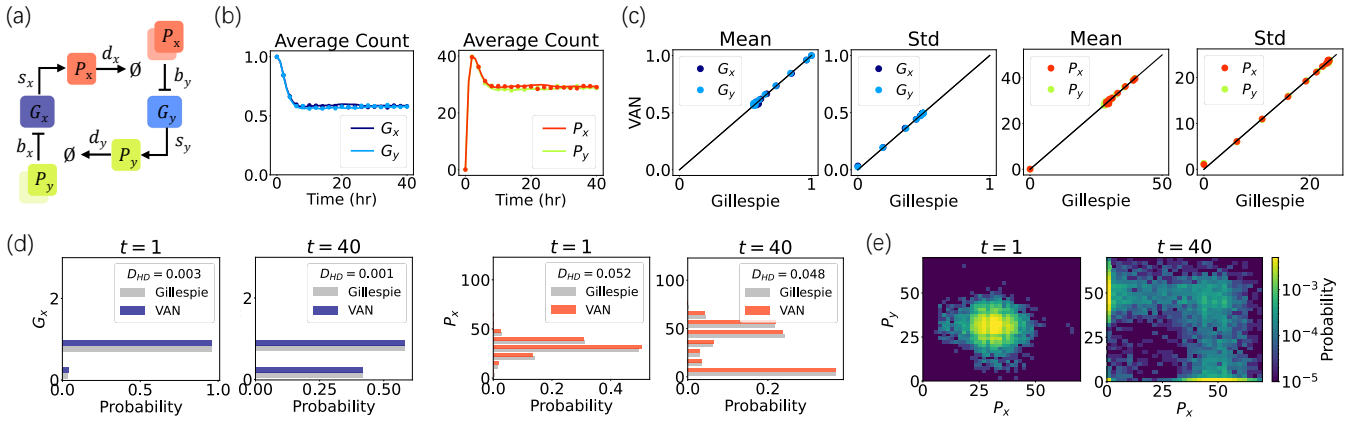
where c_1 is a time-dependent contact rate modeled as a periodic function as $c_1 = c_0(1 + \epsilon)\sin(\omega t)$ and the remaining parameters are constant. This time-dependent rate models the infection in a periodic environment.

We apply the present approach to estimate the evolution of joint probability distribution over time. As a comparison, the average count of the three species estimated from the VAN match with the Gillespie simulation (Supplementary Fig. 6b). The means, the standard deviations (Supplementary Fig. 6c) and marginal distributions (Supplementary Fig. 6d) of the two methods are consistent with each other. We further show the joint probability distribution of S and I and its evolution over time points (Supplementary Fig. 6e). The results demonstrate that the VAN method can accurately track the joint probability distribution for systems with time-dependent rates. Thus, it is applicable to more realistic situations where the parameters vary over time.



Supplementary Figure 6. The result of the epidemic model. (a) A schematic of the chemical reaction. (b) The time series of the average count of species from the VAN (dots) and the Gillespie simulation (lines). The color specifies the chemical species. (c) Comparison of the means and standard deviations of the chemical species between the VAN and the Gillespie simulation, at time $t = 0, 10, ..., 190, 200$ for chemical species specified by color. (d) The marginal count distributions of various chemical species are plotted horizontally at time $t = 10, 100$. The color in (d) specifies the result from the VAN for the species I, and gray denotes the Gillespie method. The inset contains the Hellinger distance between the two distributions. (e) The joint distribution of S and I at time $t = 10, 100$ from the VAN, with the color bar denoting the probability values in the logarithmic scale. The values of the parameters: $c_0 = 0.003$, $\omega = \pi/3$, $\epsilon = 0.2$, $c_2 = 0.02$, $c_3 = 0.007$, $c_4 = 0.002$, $c_5 = 0.05$ and $c_6 = 0.002$, with the unit of time in hours [5]. The Gillespie simulation has 10^4 trajectories. The initial distribution is the delta distribution with $S = 51$, $I = 11$ and $R = 0$.

F. Details of the genetic toggle switch



Supplementary Figure 7. The result of the genetic toggle switch by using a different method of drawing samples. The layout is the same as Fig. 2 in the main text. At every time step, 10^4 samples are drawn from the learnt distribution at the last epoch. The parameters of the system are $s_x = s_y = 50$, $d_x = d_y = 1$, $b_x = b_y = 10^{-4}$, $u_x = u_y = 0.1$ with the unit of time in hours, for both Fig. 2 and this figure.

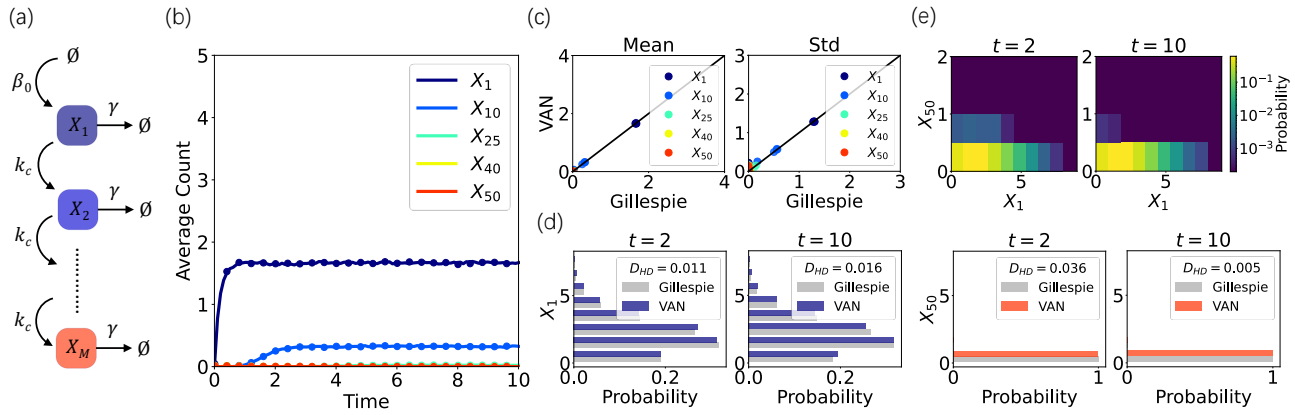
The genetic toggle switch [6, 7] has a multimodal distribution on the protein counts, depending on the genetic states of the two mutually inhibited genes. Thus, the system is suitable for testing the flexibility of the VAN to learn the multimodal distribution. The model has the chemical reactions:

$$\begin{aligned} &G_x \xrightarrow{s_x} G_x + P_x, \quad G_y \xrightarrow{s_y} G_y + P_y, \quad P_x \xrightarrow{d_x} \emptyset, \quad P_y \xrightarrow{d_y} \emptyset, \\ &2P_x + G_y \xrightarrow{b_y} \bar{G}_y, \quad 2P_y + G_x \xrightarrow{b_x} \bar{G}_x, \quad \bar{G}_y \xrightarrow{u_y} 2P_x + G_y, \quad \bar{G}_x \xrightarrow{u_x} 2P_y + G_x, \end{aligned} \tag{S2.9}$$

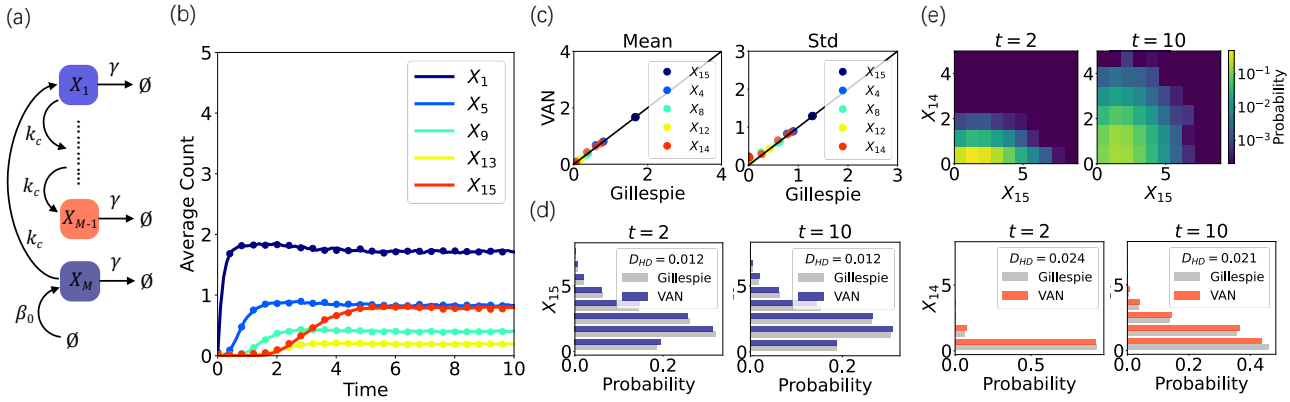
where s_x, s_y are the synthesis rates of the proteins, and p_x, p_y are the degradation rates of the proteins. The transition rate b_y (b_x) is the binding rate of two copies of protein P_X (P_Y) to the G_y (G_x), to form the complex $\bar{G}_y$ ($\bar{G}_x$). The unbinding of the complex $\bar{G}_y$ ($\bar{G}_x$) has the rate u_y (u_x).

The total count of the two forms for each gene is conserved: $G_x + \bar{G}_x = 1$. This conservation sets a constraint on the count of the two genes: $G_x = 0, 1$, $G_y = 0, 1$, and effectively reduces two variables by $\bar{G}_x = 1 - G_x$, $\bar{G}_y = 1 - G_y$. We have put this constraint on the count of the two genes in the VAN. We consider the parameter regime with weak promoter binding, because in this regime the joint distribution is multimodal with four probability peaks, posing a challenge to accurately track the joint distribution. The results are shown in Fig. 2 and Supplementary Fig. 7.

G. More cases of the intracellular signaling cascade



Supplementary Figure 8. The result of the linear signaling cascade (case 1) with 50 species. (a) A schematic of the chemical reaction. (b) The time series of the average count of species from the VAN (dots) and the Gillespie simulation (lines). The color specifies the chemical species. (c) Comparison of the means and standard deviations of the chemical species between the VAN and the Gillespie simulation, at time points $t = 1, 2, ..., 9, 10$ for chemical species denoted by color. (d) The marginal count distributions of various chemical species are plotted horizontally at time points $t = 2, 10$. The color in (d) specifies the result from the VAN for two species, and gray denotes the Gillespie method. The inset contains the Hellinger distance between the two distributions. (e) The joint distribution of the first and the last species at time points $t = 2, 10$ from the VAN, where the color bar denotes the probability values in the logarithmic scale. The Gillespie simulation has 10^4 trajectories.

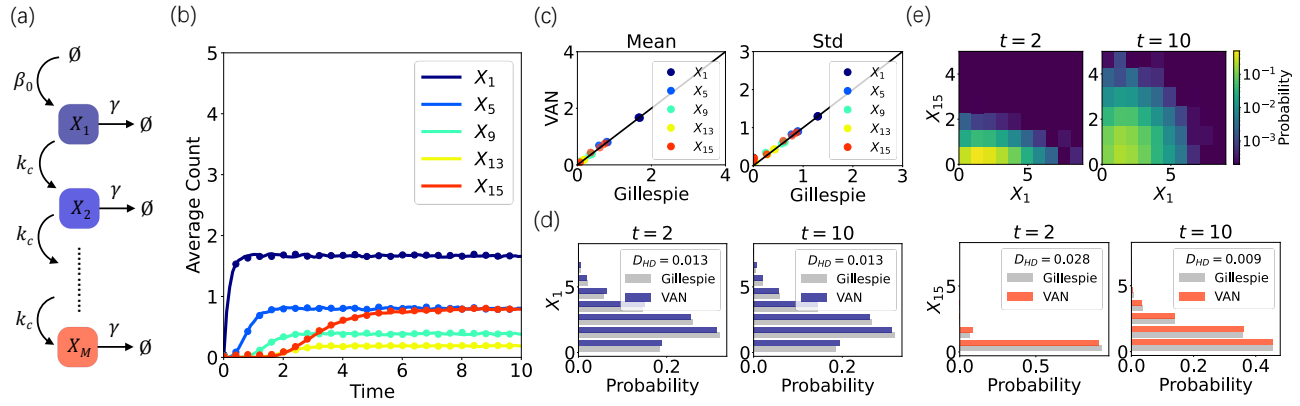


Supplementary Figure 9. The result of the linear signaling cascade (case 1) with the switched order of the 15 species. The signaling starts from the last species X_{15}, activating X_1 and so on. The result is as accurate as that without switching the order. The layout of the figure is the same as that of Supplementary Fig. 8.

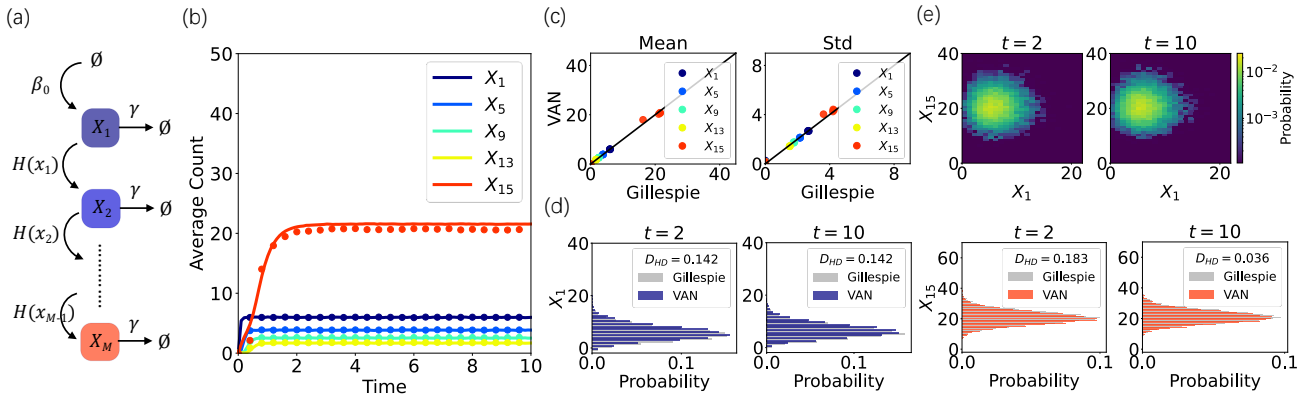
We consider three cases of the intracellular signaling cascade [8]: (1) a linear signaling cascade without feedback regulation; (2) a nonlinear signaling cascade without feedback regulation; and (3) a linear signaling cascade with feedback regulation.

The first case contains chemical reactions:

$$\emptyset \xrightarrow{\beta_0} X_1, \quad X_i \xrightarrow{k_c} X_{i+1} \quad (i = 1, \ldots, M-1), \quad X_i \xrightarrow{\gamma} \emptyset \quad (i = 1, \ldots, M). \tag{S2.10}$$



Supplementary Figure 10. The result of the linear signaling cascade (case 1) with 15 species by using the transformer, with hyperparameters of the VAN and the corresponding computational time in Supplementary Table II. The layout is the same as that of Supplementary Fig. 8.



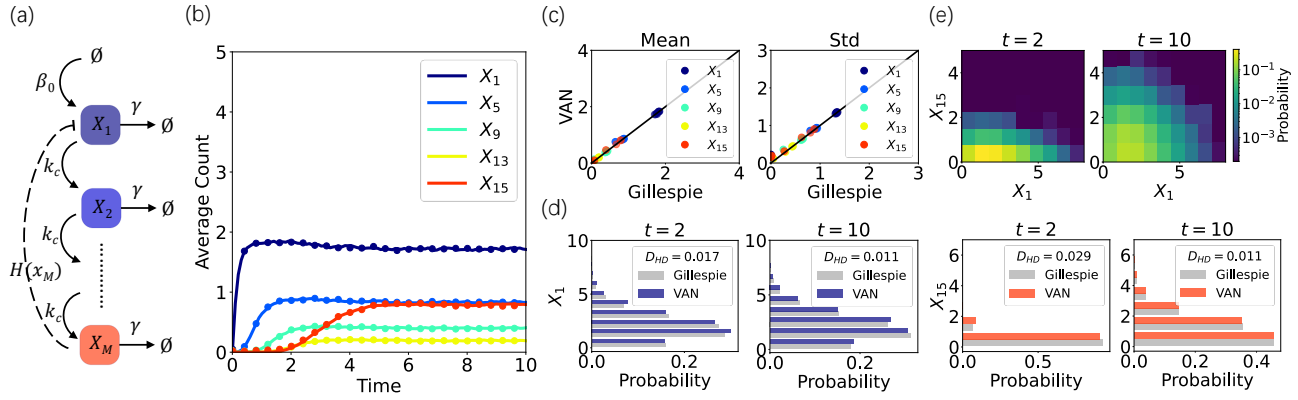
Supplementary Figure 11. The result of the nonlinear signaling cascade (case 2) for 15 species. The captions of each panel are the same as those of Supplementary Fig. 8. The parameters $\beta_0 = 10$, $\gamma = 0.1$, $b = 1$, $k_m = 100$, $k_0 = 10$, $h = 1$. The upper count limit is chosen as $N = 40$, and the result with $N = 100$ has a similar accuracy.

The parameter β_0 is the synthesis rate, k_c is the catalytic rate, and γ is the decay rate.

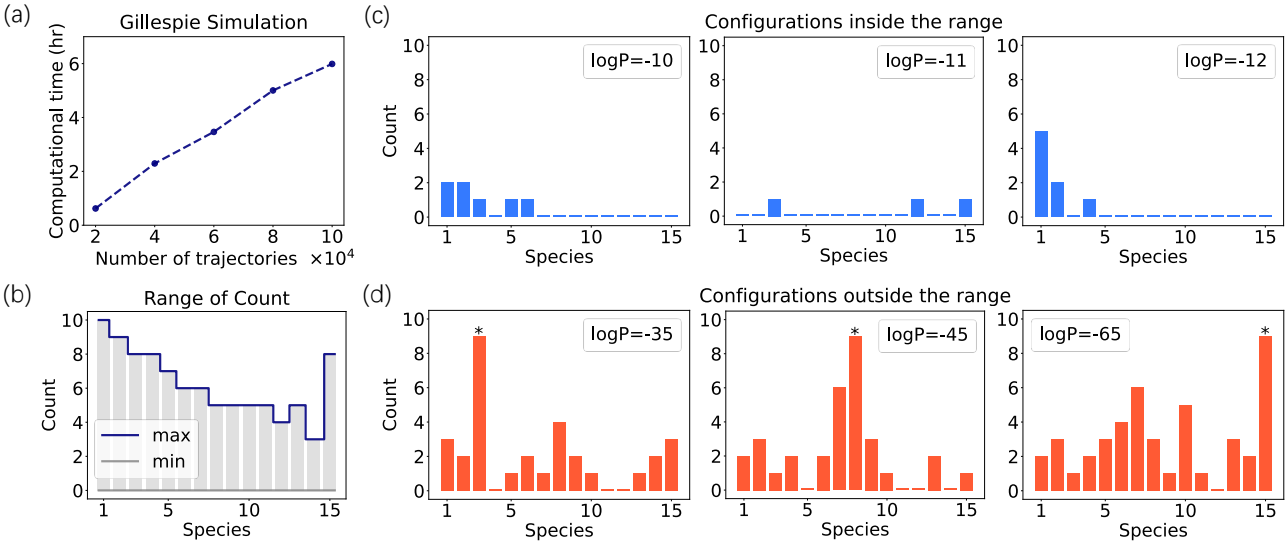
The result for the case 1 with 15 species is presented in the main text, and that with $M = 50$ species is in Supplementary Fig. 8. It shows a good match of the marginal statistics between the Gillespie simulation and the VAN method, demonstrating that the present approach is applicable to high-dimensional systems. Besides, we have switched the order of species, by starting the signaling from the last species X_M that activates X_1 and so on. The result (Supplementary Fig. 9) has the same high accuracy as that of the original order. Furthermore, we have considered a set of transformer hyperparameters (Supplementary Table II), where the hyperparameters include the number of features, the number of heads, the number of encoder and decoder layers, and the dimension of the feed-forward network. The results (Supplementary Fig. 10) show a similar accuracy in generating the distribution.

We next provide the details of cases 2, 3. In the case 2, the signaling cascade with nonlinear activation has the reaction scheme in Eq. (S2.10), and replaces the constant catalytic rate by a Hill function with a basal rate. For each reaction $X_i \longrightarrow X_{i+1}$, the catalytic production is: $H(x_i) = b + (k_m x_i^h)/(k_0 + x_i^h)$, where b is the basal rate, k_m controls the strength of the Hill activation, k_0 corresponds to the affinity for the substrate, and h is the Hill coefficient. For various M, the count of the last species x_M changes dramatically. For a better comparison, we have rescaled the parameters to be M-dependent, such that the count of the last species is kept similar for various M. We have also rescaled the time scale by 25-fold, such that the system reaches the steady state with a similar time as the case 1. In the case 3, the signaling cascade with negative feedback has the same reactions in Eq. (S2.10), and an additional negative feedback from the last species to the first one: $H(x_M) = b + k_m/(k_0 + x_M^h)$, where the parameters have the same meaning of the case 2.

The results of cases 2, 3 with $M = 15$ species are separately shown in Supplementary Figs. 11, 12. The accuracy of the VAN can be further improved by using more epochs, especially for the case 2 where the count of each species



Supplementary Figure 12. The result of the linear signaling cascade with feedback (case 3) for 15 species. The captions of each panel are the same as Supplementary Fig. 8. The parameters $\beta_0 = 1$, $k_c = 5$, $\gamma = 1$, $b = 1$, $k_m = 100$, $k_0 = 10$, $h = 1$ [8].



Supplementary Figure 13. The occupancy of the simulated trajectories in the state space from the Gillespie algorithm. The system is the first case of the signaling cascade with $M = 15$ species. (a) The computational time of simulating certain numbers of trajectories by using the Gillespie algorithm, based on Intel Xeon Gold 6240 CPU @2.6GHz. (b) The range of species counts at $t = 5$ from the simulated 10^5 trajectories by the Gillespie algorithm. The blue line is the maximum of counts and the gray line is the minimum. (c) The configurations with the species counts inside the range of the Gillespie simulation. The bars in each panel denote the count of each species, and the legend denotes the logarithmic joint probability of the configuration computed by the VAN. (d) The configurations outside the range of species counts, which are not captured by the Gillespie algorithm. The bar with an asterisk identifies the count that is out of the range. The probability of configuration can still be obtained by the VAN, as shown by the legend.

depends sensitively on the precedent species under the nonlinear activation. The results suggest that the present approach is applicable to a high-dimensional system with nonlinear interactions and feedback.

To demonstrate the advantage of the VAN on learning the joint distribution over the Gillespie algorithm, we provide a coarse estimation on the occupancy of the simulated trajectories in the state space, by considering the first case of the signaling cascade with $M = 15$ species. If we assume that there are only 10 possible counts for each species, $[0, 9]$, the size of the entire state space is 10^{15}. Then, 10^5 simulated trajectories only account for 10^{-10} of the state space. Even when simulating such a small proportion of trajectories, the computational time of the Gillespie algorithm (Supplementary Fig. 13a) has a same order as that of evolving the VAN (Table I). Moreover, many configurations are not captured by the 10^5 simulated trajectories, but the joint probability of any configuration in the state space can always be obtained by the VAN (Supplementary Fig. 13). It implies that the VAN provides a complete joint distribution of the high-dimensional space, whereas the Gillespie algorithm is practically unaffordable to cover the full state space.

	Species	Reactions	N	Time steps	δt	Epochs	n_{layers}	n_{heads}	d_{model}	d_{ff}	Comput. time (hr)
Signaling cascade 1	15	30	10	$1 * 10^3$	10^{-2}	100	2	2	16	32	3.96
Signaling cascade 1	15	30	10	$1 * 10^3$	10^{-2}	100	6	2	16	32	10.11
Signaling cascade 1	15	30	10	$1 * 10^3$	10^{-2}	100	2	4	16	32	4.50
Signaling cascade 1	15	30	10	$1 * 10^3$	10^{-2}	100	2	2	32	32	3.97
Signaling cascade 1	15	30	10	$1 * 10^3$	10^{-2}	100	2	2	16	128	4.31

Supplementary Table II. The computational time of the first case of the signaling cascade with $M = 15$ species. A set of hyperparameters of the transformer are considered, and the transformer typically needs longer computational time than the RNN. The other settings of hyperparameters are the same as shown in Table I.

[1] C. W. Gardiner, *Handbook of Stochastic Methods*, 3rd ed. (Springer-Verlag, Berlin, 2004).
[2] A. Sukys, K. Öcal, and R. Grima, Approximating solutions of the chemical master equation using neural networks, iScience **25**, 105010 (2022).
[3] F. Jafarpour, T. Biancalani, and N. Goldenfeld, Noise-induced mechanism for biological homochirality of early life self-replicators, Phys. Rev. Lett. **115**, 158101 (2015).
[4] F. C. Frank, On spontaneous asymmetric synthesis, Biochim. Biophys. Acta **11**, 459 (1953).
[5] V. H. Thanh and C. Priami, Simulation of biochemical reactions with time-dependent rates by the rejection-based algorithm, J. Chem. Phys. **143**, 054104 (2015).
[6] T. S. Gardner, C. R. Cantor, and J. J. Collins, Construction of a genetic toggle switch in escherichia coli, Nature **403**, 339 (2000).
[7] A. Terebus, C. Liu, and J. Liang, Discrete and continuous models of probability flux of switching dynamics: Uncovering stochastic oscillations in a toggle-switch system, J. Chem. Phys. **151**, 185104 (2019).
[8] A. Gupta, C. Schwab, and M. Khammash, Deepcme: A deep learning framework for computing solution statistics of the chemical master equation, PLoS Comput. Biol. **17**, e1009623 (2021).