Supplementary Material

The MaxCal approach to dynamical trajectories

We define $p_\Gamma(t)$, or simply p_Γ as the probability of observing trajectory Γ within some small time interval $[t, t+\delta t]$ [1]. In our toggle switch example, typical trajectories within interval δt involve protein number staying fixed or removal as well as addition of proteins of either type.

We assume that all possible events (birth, death, etc...) are equally likely for each particle within interval δt until a measurement says otherwise. In other words, we say our trajectory ensemble is maximally uncertain until data is used to constrain this ensemble thereby reducing the ensemble uncertainty. The uncertainty measure of the set of trajectories $\{p_\Gamma\}$ within interval δt is $-\Sigma_\Gamma p_\Gamma \ln p_\Gamma$. This uncertainty measure satisfies the following properties: it is a continuous function of the $\{p_\Gamma\}$, it satisfies a notion of additivity(actually the "composition property") and it is a monotonically increasing function of Γ(the total number of allowable trajectories) when all trajectories are equiprobable [2].

In order to compute the trajectory ensemble, we maximize the caliber, $\mathcal{C}(p_\Gamma)$, which is the uncertainty measure, $-\Sigma_\Gamma p_\Gamma \ln p_\Gamma$, subject to constraints on average observables, $\langle A_n \rangle$,

$$\mathcal{C} = -\Sigma_\Gamma p_\Gamma \ln p_\Gamma + \alpha \sum_\Gamma p_\Gamma + \sum_{n,\Gamma} \lambda_n A_{n,\Gamma} p_\Gamma \tag{1}$$

where λ_n and α are Lagrange multipliers. Average observables, to be discussed in more detail shortly, include birth and death rates for either protein type within interval δt as well as correlations between these quantities.

Maximizing $\mathcal{C}$ with respect to $\{p_\Gamma\}$, yields the following trajectory ensemble

$$p_\Gamma = Q^{-1} e^{\sum_n \lambda_n A_{n,\Gamma}} \text{ with } Q \equiv \sum_\Gamma e^{\sum_n \lambda_n A_{n,\Gamma}} \tag{2}$$

where Q is the dynamical analog of the equilibrium partition function. The dynamical partition function can be used to infer cumulants, subscripted c, of dynamical observables as follows

$$\langle A_n^m \rangle_c = \partial^m \log Q / \partial \lambda_n^m \tag{3}$$

where m is the cumulant order.

Now we discuss what quantities we can constrain in MaxCal. In our toggle switch example, we considered both one and two body terms in our dynamical partition function. The two body terms are the correlations between our lattice gas(or spin) variables. Our full dynamical partition function(including correlations), Eqns.(6) and (9) of the main body, is like an infinite range Ising model with 4 different types of variables. For simplicity, consider a generic infinite range Ising model with 2 types of variables with the following Hamiltonian $\beta\mathcal{H} = h\sum_i^N s_i + K\sum_{ij}^N s_i s_j$, where h and K are the magnetic field and spin coupling constant, respectively, and the sums over each spin variable index runs from 1 to N. The spin variables in the above Hamiltonian are Ising variables which take on values of ± 1.

In general, we want our Hamiltonian and, by extension, our partition function, to be valid for any N. In MaxCal we need our dynamical partition function to hold for all N as, when we simulate a particular MaxCal time trace using our dynamical partition function, the values of N_A and N_B change from one time interval to the next. This is equivalent to saying we want our Hamiltonian to be scalable with system size, or simply for $\beta\mathcal{H}$ to be extensive with particle number(recall that $\sum_{ij}^N s_i s_j$ is not extensive if the interaction is infinite range). As a consequence, K, the coupling constant, must scale as $1/N$ if the Hamiltonian is to be extensive. In other words $NK \equiv K_0$ where K_0 is a

"bare" coupling constant (independent of N). As a result, our generic infinite range Hamiltonian with lattice-gas variables, $\ell_i = (s_i + 1)/2$, takes the same form as our Hamiltonian expressed using Ising variables. Thus, our choice of whether to use lattice-gas or Ising variables to describe birth and death events for our toggle switch is arbitrary and should, in principle, not change the physical content of our trajectory ensemble [1].

In our toggle switch, the $K_{A\beta}$ and $K_{B\alpha}$ we measure approximately scale as $1/N_A$ and $1/N_B$ over the vast majority of the time trace. We verified this in our Gillespie time trace for $N_A = \{15, 20, 25\}$ and when a single birth event of B occurred within a particular small time interval. We found $K_{A\beta} = \{-0.506, -0.387, -0.291\}$ meaning that $K_0 = \{-7.59, -7.74, -7.27\}$ shows no particular upward or downward trend. As we move even further away from this range of N_A's, we find that K_0 does begin to drift with particle number. A more exact MaxCal model would thus require we find a coupling constant $K_{A\beta}$ for each possible N_A and a coefficient of $K_{B\alpha}$ for each N_B. In practice, a single value of the coupling constant is sufficient to achieve the semi-quantitative agreement we showed in our Figs.(5)-(7). In producing our MaxCal time traces, we did not allow the coupling constant to change with particle number. The time traces were all produced using a single value of $K_{A\beta}$ and $K_{B\alpha}$ evaluated when the high state has particle number k/d and the low state has particle number 0, i.e. the state encountered most often. Of course, the more details we put into the MaxCal model the better we expect the agreement to be with the simulated data. In addition, in experiments, while single birth-death events of a species in a low state might be easy to detect(the difference between no B or 1 B, say), fluctuations in the population of the species in the high state(23 versus 22 species A, say) might

[1] In lattice-gas variables, the Hamiltonian in the main body becomes $\beta\mathcal{H} = C + \tilde{h}\sum_i^N \ell_i + (\tilde{K_0}/N)\sum_{ij}^N \ell_i \ell_j$ where $\tilde{h} = 2h - 4K_0$, $\tilde{K_0} = 4K_0$ and the constant, $C = -hN + K_0 N$. The constant can be dropped from the Hamiltonian as it only redefines the zero of energy and does not change any of the expectation values for ℓ_i. In this way, both $\tilde{h}$ and $\tilde{K_0}$ are independent of N. As a sidenote, if N is zero then there is nothing to couple to anyways, so we can take the coupling constant as 0.

be more difficult to pick up. It is therefore conceivable that in experiments, only a single coupling constant between high and low states could be measured.

Deriving the master equations in the absence of correlation

Here, we simply confirm that the MaxCal partition-function-based approach, Eq.(3) of the main body, gives the expected master equation, Eq.(4) of the main body, when there are no correlations, $\mathcal{Z} \equiv 1$ in Eq.(3). The probability of having N_A particles of type A at time $t+\delta t$, $P(N_A, t+\delta t)$, is

$$P(N_A, t+\delta t) = (no\ event) \times P(N_A, t) + (production\ of\ A) \times P(N_A - 1, t)$$
$$+ (degradation\ of\ A) \times P(N_A + 1, t) \tag{4}$$

where the bracketed terms are statistical weights(or probabilities within δt) for the specified trajectories. Expanding the left hand side of the above, we find

$$\delta t \dot{P}(N_A, t) = ((no\ event) - 1) P(N_A, t) + (production\ of\ A) \times P(N_A - 1, t)$$
$$+ (degradation\ of\ A) \times P(N_A + 1, t). \tag{5}$$

When $\mathcal{Z} \equiv 1$ in Eq.(3) of the main body, the dynamical partition function for the production and degradation event for A is

$$Q = (1 + e^{h_\alpha})(1 + e^{h_A})^{N_A}. \tag{6}$$

The statistical weight for $(no\ event)$ is $e^{N_A h_A}/Q$. The statistical weight for $(production\ of\ A)$ - when the corresponding indicator variable for A production is 1- is $e^{h_\alpha} e^{N_A h_A}/Q$ while that for $(degradation\ of\ A)$- when the corresponding indicator variable for A degradation is 0 - is $e^{h_A (N_A - 1)}/Q$. To leading order, $Q \sim e^{h_A N_A} + e^{h_\alpha} e^{h_A N_A} + N_A e^{h_A (N_A - 1)}$. It then follows that $((no\ event) - 1) = (e^{N_A h_A} - Q)/Q \sim$

$-e^{h_\alpha} - N_A e^{-h_A}$, $(production\ of\ A) \sim e^{h_\alpha}$ and finally $(degradation\ of\ A) \sim N_A e^{-h_A}$. We can now re-write Eq.(5) as follows

$$\delta t \dot{P}(N_A,t) = -(e^{h_\alpha} + N_A e^{-h_A})P(N_A,t) + e^{h_\alpha} P(N_A - 1,t)$$

$$+ N_A e^{-h_A} P(N_A + 1,t). \tag{7}$$

We define statistical weights $\sigma_\alpha \equiv e^{h_\alpha}$ and $\sigma_A \equiv e^{-h_A}$(note the sign difference). By identifying $\sigma_A = \delta t d_A$ and $\sigma_\alpha = \delta t k_\alpha$, as they appear in the main body, we recover the master equation for the uncoupled ($\mathcal{Z} \equiv 1$) process.

References

[1] Stock, G., Ghosh, K., and K.A. Dill. 2008. Maximum Caliber: A variational approach applied to two-state dynamics. *J. Chem. Phys.*, 128:194102.

[2] Bretthorst, G.L., and E.T. Jaynes. 2003. Probability Theory: The Logic of Science. Cambridge University Press.

[3] Lipshtat, A., Loinger, A., Balaban, N.Q., and O. Biham. 2006. Genetic Toggle Switch without Cooperative Binding. *Phys. Rev. Lett.*, 96:188101.