# Evolutionary features in a minimal physical system: diversity, selection, growth, inheritance, and adaptation

Guy Bunin[1] and Olivier Rivoire[2]

[1] *Department of Physics, Technion-Israel Institute of Technology, Haifa 32000, Israel*
[2] *Gulliver, CNRS, ESPCI, Université PSL, 75005 Paris, France*

# Supplementary Information

## Appendix A: Microscopic model

We derive the chemical reactions shown in Fig. 1 from physical principles, starting from a description of the geometry and interaction potentials of fundamental elements. This ensures full consistency not only with thermodynamics but also with the constraints of elementary catalysts. The interaction potentials are simplified to two parameters, a forward and a backward activation energy, which control the kinetic rates through the Arrhenius law. This approach follows previous work showing that it accounts for properties obtained from molecular dynamics simulations with spherical particles at sufficiently low temperatures [1]. For illustrative purposes, we draw the basic elements as squares or rectangles in two dimensions, with interaction patches at specific locations on their surface controlling the strength and specificity of their interaction with other elements.

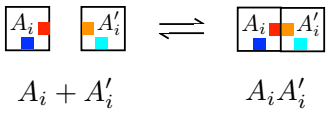
### 1. Elements and interactions

The physical model has 4 types of elements, $A_i$, $A'_i$, $B_i$ and $C_i$, each existing in $N$ different types $i$. Interactions occur only between elements of the same type $i$, except for a few binding interactions between some $A_i$ and $C_j$ with $j \neq i$. These interactions are reciprocal and are represented as $i \sim j$. An element $A_i$ has two interacting interfaces, one through which it interacts with $A'_i$, $B_i$ and $C_j$ with $j \sim i$, and one through which it interacts with $C_i$. Elements $A'_i$ and $B_i$ have a single interface each and interact only with $A_i$. Elements $C_i$ have three interfaces through which they interact with $A_i$, $A_j$ with $j \sim i$, and $A'_i$. For simplicity, we assume that all elements of the same type have equivalent physical properties.

We denote by $h^-$ the activation energy for the dissociation of $A_iA'_i$ into $A_i + A'_i$, and $h^+$ the activation energy for the association of $A_i + A'_i$ into $A_iA'_i$. These activation energies are represented as

$$A_i + A'_i \underset{h^-}{\overset{h^+}{\rightleftharpoons}} A_iA'_i, \tag{A1}$$

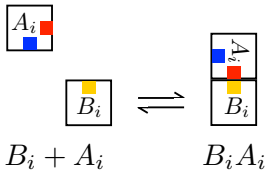
while the geometrical constraints are represented as




Similarly, the interaction of $A_i$ with $B_i$ is described as

$$B_i + A_i \underset{\phi_B^-}{\overset{\phi_B^+}{\rightleftharpoons}} B_iA_i \tag{A2}$$

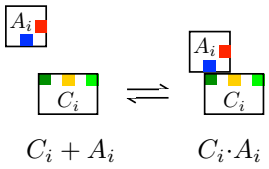
with an activation energy $\phi_B^+$ for association and an activation energy $\phi_B^-$ for dissociation, and represented as



The interaction of $A_i$ with $C_i$ is described as

$$C_i + A_i \underset{\epsilon}{\overset{0}{\rightleftharpoons}} C_i{\cdot}A_i \tag{A3}$$

with no activation energy for association and an activation energy $\epsilon$ for dissociation, and represented as

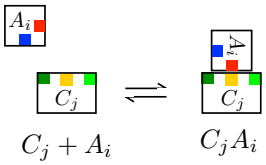



Note that $A_i$ interacts with $C_i$ through a different interface than with $A'_i$ or $B_i$. The difference is represented symbolically by denoting $C_i{\cdot}A_i$ with a dot.

The interaction of $A_i$ with $C_j$ with $j \sim i$ is described as

$$C_j + A_i \underset{\phi_C^-}{\overset{\phi_C^+}{\rightleftharpoons}} C_j A_i \tag{A4}$$

with an activation energy $\phi_C^+$ for association and an activation energy $\phi_C^-$ for dissociation, and represented as

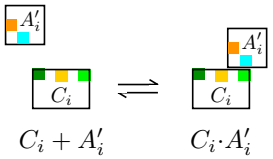



Note that $A_i$ interacts with $C_j$ through the same interface as with $A'_i$ and $B_i$. This ensures that $A_i$ can interact with $C_j$ but $A_iA'_i$ cannot.

Finally, the interaction of $A'_i$ with $C_i$ is described as

$$C_i + A'_i \underset{\epsilon}{\overset{0}{\rightleftharpoons}} C_i{\cdot}A'_i \tag{A5}$$

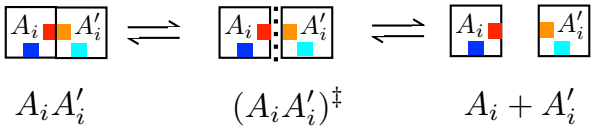
with the same activation energy $\epsilon$ as in Eq. (A3), and represented as




We ignore complexes $A_iB_iC_i$ (which could be excluded as a consequence of slightly modified geometries or by introducing repulsion between $B_i$ and $C_i$).

### 2. Detailed spontaneous reaction

To account for catalysis, a more detailed description of the reaction $A_iA'_i \underset{h^-}{\overset{h^+}{\rightleftharpoons}} A_i + A'_i$ is needed that includes the transition state $(A_iA'_i)^\ddagger$ when $A_i$ and $A'_i$ are at an intermediate distance of each other at which the potential energy is maximal,



This corresponds to the following sequence of events where the energy of each state is indicated below each symbol:

$$\underset{0}{A_i + A_i'} \rightleftharpoons \underset{h^+}{(A_iA_i')^\ddagger} \rightleftharpoons \underset{h^+-h-}{A_iA_i'} \tag{A6}$$

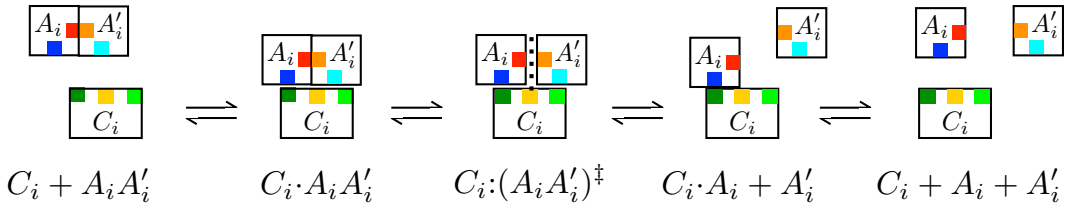
Taking differences of energies between successive states to define activation barriers, this is represented as

$$A_i + A_i' \underset{0}{\overset{h^+}{\rightleftharpoons}} (A_iA_i')^\ddagger \underset{h^-}{\overset{0}{\rightleftharpoons}} A_iA_i' \tag{A7}$$

$(A_iA_i')^\ddagger$ is an unstable intermediate and eliminating it leads to Eq. (A1).

### 3. Catalysis of bond cleavage

The geometry of $C_i$ is chosen to enable the catalysis of the spontaneous reaction $A_iA_i' \underset{h^+}{\overset{h^-}{\rightleftharpoons}} A_i + A_i'$. The mechanism is represented by




corresponding to a catalytic cycle described by

$$\underset{h^+-h^-}{C_i + A_iA_i'} \rightleftharpoons \underset{h^+-h^--\epsilon}{C_i{\cdot}A_iA_i'} \rightleftharpoons \underset{h^+-2\epsilon}{C_i{:}(A_iA_i')^\ddagger} \rightleftharpoons \underset{-\epsilon}{C_i{\cdot}A_i + A_i'} \rightleftharpoons \underset{0}{C_i + A_i + A_i} \tag{A8}$$

where an alternative path is for $A_i$ to be released before $A_i'$. This design is such that the two products $A_i$ and $A_i'$ cannot be fully attached to $C_i$, i.e., only in the transition state does $C_i$ has two interacting interfaces. This is achieved by proper positioning of the interacting interfaces [1].

Assuming $0 < \epsilon < h^\pm$, the cycle is equivalently described by activation energies as

$$C_i + A_iA_i' \underset{\epsilon}{\overset{0}{\rightleftharpoons}} C_i{\cdot}A_iA_i' \underset{0}{\overset{h^--\epsilon}{\rightleftharpoons}} C_i{:}(A_iA_i')^\ddagger \underset{h^+-\epsilon}{\overset{0}{\rightleftharpoons}} C_i{\cdot}A_i + A_i' \underset{0}{\overset{\epsilon}{\rightleftharpoons}} C_i + A_i + A_i' \tag{A9}$$

Eliminating the unstable intermediate, we obtain

$$C_i + A_iA_i' \underset{\epsilon}{\overset{0}{\rightleftharpoons}} C_i{\cdot}A_iA_i' \underset{h^+-\epsilon}{\overset{h^--\epsilon}{\rightleftharpoons}} C_i{\cdot}A_i + A_i' \underset{0}{\overset{\epsilon}{\rightleftharpoons}} C_i + A_i + A_i' \tag{A10}$$

In this mechanism, catalysis is achieved by replacing a single barrier $h^-$ by two smaller barriers $h^- - \epsilon$ and $\epsilon$. The catalytic rate is controlled by the largest of these two barriers, corresponding to an effective barrier $\max(h^- - \epsilon, \epsilon)$ which is minimal for $\epsilon = h^-/2$ (a more detailed analysis shows that this reasoning is valid only if $h^+ > h^-$, otherwise the effective barrier is larger because of frequent recrossing events) [2]. In what follows we take $h^+ = \infty$ and simply denote $h^-$ by $h$, so that the requirement for catalysis reads

$$0 < \epsilon < h \tag{A11}$$

### 4. All elementary reactions

We consider a chemostat where some species are injected and all species are diluted at a constant rate $\delta$. In the well-mixed case (where the diffusion constant $D$ is infinite), the model is described by the following reactions where, as above, we indicate above or below the arrows the reaction rate constants $k$ as $-(\ln k)/k_BT$ with $k_BT = 1$ to fix the unit of energy, where $T$ is the temperature and $k_B$ the Boltzmann constant.

Injection:

$$\emptyset \xrightarrow{-\ln(\delta[R])} A_i A_i' \tag{A12}$$

$$\emptyset \xrightarrow{-\ln(\delta[B]_0)} B_i \tag{A13}$$

$$\emptyset \xrightarrow{-\ln(\delta[C]_0)} C_i \tag{A14}$$

where $[R]$, $[B]_0$ and $[C]_0$ can be interpreted as the concentrations at which the molecules are injected with rate $\delta$.

Spontaneous reaction:

$$A_i A_i' \xrightarrow{h} A_i + A_i' \tag{A15}$$

Catalysis:

$$C_i + A_i A_i' \underset{\epsilon}{\overset{0}{\rightleftharpoons}} C_i{\cdot}A_i A_i' \tag{A16}$$

$$C_i{\cdot}A_i A_i' \xrightarrow{h-\epsilon} C_i{\cdot}A_i + A_i' \tag{A17}$$

$$C_i{\cdot}A_i A_i' \xrightarrow{h-\epsilon} C_i{\cdot}A_i' + A_i \tag{A18}$$

$$C_i{\cdot}A_i \underset{0}{\overset{\epsilon}{\rightleftharpoons}} C_i + A_i \tag{A19}$$

$$C_i{\cdot}A_i' \underset{0}{\overset{\epsilon}{\rightleftharpoons}} C_i + A_i' \tag{A20}$$

Inhibition:

$$B_i + A_i \underset{\phi_B^-}{\overset{\phi_B^+}{\rightleftharpoons}} B_i A_i \tag{A21}$$

$$C_j + A_i \underset{\phi_C^-}{\overset{\phi_C^+}{\rightleftharpoons}} C_j A_i \qquad (j \sim i) \tag{A22}$$

Dilution:

$$A_i \xrightarrow{-\ln\delta} \emptyset \tag{A23}$$

$$A_i' \xrightarrow{-\ln\delta} \emptyset \tag{A24}$$

$$A_i A_i' \xrightarrow{-\ln\delta} \emptyset \tag{A25}$$

$$B_i \xrightarrow{-\ln\delta} \emptyset \tag{A26}$$

$$C_i \xrightarrow{-\ln\delta} \emptyset \tag{A27}$$

$$B_i A_i \xrightarrow{-\ln\delta} \emptyset \tag{A28}$$

$$C_i{\cdot}A_i A_i' \xrightarrow{-\ln\delta} \emptyset \tag{A29}$$

$$C_i{\cdot}A_i \xrightarrow{-\ln\delta} \emptyset \tag{A30}$$

$$C_i{\cdot}A_i' \xrightarrow{-\ln\delta} \emptyset \tag{A31}$$

$$C_j A_i \xrightarrow{-\ln\delta} \emptyset \tag{A32}$$

Note that we do not distinguish between $C_i{\cdot}A_i A_i'$ and $C_i{\cdot}A_i' A_i$. These reactions are translated into ordinary differential equations by associating a kinetic rate with each activation barrier via the Arrhenius equation and ignoring differences in prefactors. For instance, the reaction $B_i + A_i \underset{\phi_B^-}{\overset{\phi_B^+}{\rightleftharpoons}} B_i A_i$ has a forward first-order kinetic rate $k_+ = \kappa_0 e^{-\phi_B^+}$ and backward second-order kinetic rate $k_- = \kappa_0 \Omega_0 e^{-\phi_B^-}$ where $\kappa_0^{-1}$ represents an unit of time and $\Omega_0$ a unit of volume (the unit of energy is $k_B T = 1$). We take $\kappa_0 = 1$ to fix the unit of time, while leaving $\Omega_0$ as a parameter. With these

conventions, the equations describing a large well-mixed system are

$$\frac{d[A_iA_i']}{dt} = \delta[R] - e^{-h}[A_iA_i'] + e^{-\epsilon}[C_i{\cdot}A_iA_i'] - \Omega_0[C_i][A_iA_i'] - \delta[A_iA_i'] \tag{A33}$$

$$\frac{d[A_i]}{dt} = e^{-h}[A_iA_i'] + e^{-\epsilon}[C_i{\cdot}A_i] - \Omega_0[C_i][A_i] + e^{-h+\epsilon}[C_i{\cdot}A_iA_i'] + e^{-\phi_B^-}[B_iA_i] - e^{-\phi_B^+}\Omega_0[B_i][A_i] + \sum_{j\sim i}(e^{-\phi_C^-}[C_jA_i] - e^{-\phi_C^+}\Omega_0[C_j][A_i]) - \delta[A_i] \tag{A34}$$

$$\frac{d[A_i']}{dt} = e^{-h}[A_iA_i'] + e^{-h+\epsilon}[C_i{\cdot}A_iA_i'] - \delta[A_i'] + e^{-\epsilon}[C_i{\cdot}A_i'] - \Omega_0[C_i][A_i'] \tag{A35}$$

$$\frac{d[B_i]}{dt} = \delta[B]_0 + e^{-\phi_B^-}[B_iA_i] - e^{-\phi_B^+}\Omega_0[B_i][A_i] - \delta[B_i] \tag{A36}$$

$$\frac{d[C_i]}{dt} = \delta[C]_0 - \Omega_0[C_i][A_iA_i'] + e^{-\epsilon}[C_i{\cdot}A_iA_i'] + e^{-\epsilon}[C_i{\cdot}A_i] - \Omega_0[C_i][A_i] + e^{-\epsilon}[C_i{\cdot}A_i'] - \Omega_0[C_i][A_i'] + \sum_{j}(e^{-\phi_C^-}[C_iA_j] - e^{-\phi_C^+}\Omega_0[C_i][A_j]) - \delta[C_i] \tag{A37}$$

$$\frac{d[B_iA_i]}{dt} = e^{-\phi_B^+}\Omega_0[B_i][A_i] - e^{-\phi_B^-}[B_iA_i] - \delta[B_iA_i] \tag{A38}$$

$$\frac{d[C_i{\cdot}A_iA_i']}{dt} = \Omega_0[C_i][A_iA_i'] - e^{-\epsilon}[C_iA_iA_i'] - 2e^{-h+\epsilon}[C_i{\cdot}A_iA_i'] - \delta[C_i{\cdot}A_iA_i'] \tag{A39}$$

$$\frac{d[C_i{\cdot}A_i]}{dt} = \Omega_0[C_i][A_i] - e^{-\epsilon}[C_i{\cdot}A_i] + e^{-h+\epsilon}[C_i{\cdot}A_iA_i'] - \delta[C_i{\cdot}A_i] \tag{A40}$$

$$\frac{d[C_i{\cdot}A_i']}{dt} = \Omega_0[C_i][A_i'] - e^{-\epsilon}[C_i{\cdot}A_i'] + e^{-h+\epsilon}[C_i{\cdot}A_iA_i'] - \delta[C_i{\cdot}A_i'] \tag{A41}$$

$$\frac{d[C_jA_i]}{dt} = e^{-\phi_C^+}\Omega_0[C_j][A_i] - e^{-\phi_C^-}[C_jA_i] - \delta[C_jA_i] \tag{A42}$$

## Appendix B: Coarse-graining

The set of reactions is described by a smaller number of effective reactions when the different reaction rates are separated in time.

### 1. Microscopic model

To perform the reduction, we rewrite all the reactions from slowest to fastest. We indicate above and below the arrows the reaction rate constants (rather than their activation barrier as above). The indices refer to the time scale of these rate constants.

$$A_iA_i' \xrightarrow{k_{-1}} A_i + A_i' \qquad (\tau \sim \varepsilon^{-1}) \tag{B1}$$

$$\emptyset \xrightarrow{\delta[R]_0} A_iA_i', \qquad \emptyset \xrightarrow{\delta[C]_0} C_i \qquad (\tau \sim 1) \tag{B2}$$

$$A_iA_i',\ C_i,\ A_i,\ A_i',\ C_i{\cdot}A_iA_i',\ C_i{\cdot}A_i,\ C_i{\cdot}A_i',\ C_jA_i \xrightarrow{\delta} \emptyset \qquad (\tau \sim 1) \tag{B3}$$

$$C_i{\cdot}A_iA_i' \xrightarrow{k_0} C_i{\cdot}A_i + A_i', \qquad C_i{\cdot}A_iA_i' \xrightarrow{k_0} C_i{\cdot}A_i' + A_i \qquad (\tau \sim 1) \tag{B4}$$

$$A_i + C_j \underset{k_1^-}{\overset{k_1^+}{\rightleftharpoons}} C_jA_i \qquad (\tau \sim \varepsilon) \tag{B5}$$

$$A_i + B_i \underset{k_2^-}{\overset{k_2^+}{\rightleftharpoons}} B_iA_i \qquad (\tau \sim \varepsilon^2) \tag{B6}$$

$$A_i + C_i \underset{k_3^-}{\overset{k_3^+}{\rightleftharpoons}} C_i{\cdot}A_i, \qquad A_i' + C_i \underset{k_3^-}{\overset{k_3^+}{\rightleftharpoons}} C_i{\cdot}A_i', \qquad A_iA_i' + C_i \underset{k_3^-}{\overset{k_3^+}{\rightleftharpoons}} C_i{\cdot}A_iA_i' \qquad (\tau \sim \varepsilon^3) \tag{B7}$$

The time scale of each of these reactions is given in parentheses on the right, with a parameter $\varepsilon$ that must be send to zero for the coarse-graining to be mathematically exact [3] (this adimensional $\varepsilon$ is distinct from the $\epsilon$ that denotes an activation energy). The correspondence to the previous notations is as follows

$$k_{-1} = e^{-h}, \quad k_0 = e^{-(h-\epsilon)}, \quad k_1^+ = \Omega_0 e^{-\phi_C^+}, \quad k_1^- = e^{-\phi_C^-}, \quad k_2^+ = \Omega_0 e^{-\phi_B^+}, \quad k_2^- = e^{-\phi_B^-}, \quad k_3^+ = \Omega_0, \quad k_3^- = e^{-\epsilon} \tag{B8}$$

### 2. Principle of the reduction

We consider the following hierarchy of timescales

$$k_{-1} \ll k_0, \delta \ll k_1^\pm \ll k_2^\pm \ll k_3^\pm \tag{B9}$$

Formally, this is controlled by the small quantity $\varepsilon$ by considering $k_n^\pm = \varepsilon^{-n} \bar{k}_n^\pm$ where $\bar{k}_n^\pm$ are of order 1.

In addition, to simplify the formula, we also need $k_n^+/k_n^-$ to be either large or small: an assumption of strong or weak binding that corresponds to a separation of timescale within the previous one. This is controlled by an other small quantity $\varepsilon'$ larger than $\varepsilon$: $\varepsilon \ll \varepsilon' \ll 1$. Formally, we take the limits $\varepsilon \to 0$ and $\varepsilon' \to 0$ with $\varepsilon/\varepsilon' \to 0$. This is achieved, for example, when $\varepsilon' = \varepsilon^{1/2}$ with $\varepsilon \to 0$.

Overall, we assume

$$k_{-1} \ll k_0, \delta \ll k_1^+ \ll k_1^- \ll k_2^- \ll k_2^+ \ll k_3^+ \ll k_3^- \tag{B10}$$

The constraints $k_{-1} \ll k_0 \ll k_3^- \ll k_3^+$ impose

$$0 < \epsilon < h/2 \tag{B11}$$

$$\epsilon < -\ln \Omega_0 \tag{B12}$$

The other kinetic rates each involve an independent parameter which ensures that Eq. (B10) can be satisfied.

### 3. Elimination of processes on the slowest timescale $\tau \sim \varepsilon^{-1}$

We ignore the spontaneous reaction which occurs on a timescale much slower than the other ones.

### 4. Elimination of processes on the fastest timescale $\tau \sim \varepsilon^3$

On the fastest timescale, we have a quasi-equilibrium of the reactions

$$A_i + C_i \underset{k_3^-}{\overset{k_3^+}{\rightleftharpoons}} C_i{\cdot}A_i, \qquad A_i' + C_i \underset{k_3^-}{\overset{k_3^+}{\rightleftharpoons}} C_i{\cdot}A_i', \qquad A_iA_i' + C_i \underset{k_3^-}{\overset{k_3^+}{\rightleftharpoons}} C_i{\cdot}A_iA_i' \tag{B13}$$

The slow variables are

$$[A_i]_3 = [A_i] + [C_i{\cdot}A_i] \tag{B14}$$

$$[A_i']_3 = [A_i'] + [C_i{\cdot}A_i'] \tag{B15}$$

$$[A_iA_i']_3 = [A_iA_i'] + [C_i{\cdot}A_iA_i'] \tag{B16}$$

$$[C_i]_3 = [C_i] + [C_i{\cdot}A_i] + [C_i{\cdot}A_i'] + [C_i{\cdot}A_iA_i'] \tag{B17}$$

They are defined to be unchanged by the fast processes. Assuming equilibrium on the fast timescale, we obtain $[C_i]$, $[A_i]$ and $[C_iA_iA_i']$ as function of $[A_i]_3$, $[C_i]_3$ etc. The slow variables are simply expressed in terms of the fast variables, for instance

$$[A_i]_3 = [A_i](1 + k_3^+/k_3^-[C_i]) \tag{B18}$$

$$[C_i]_3 = [C_i](1 + k_3^+/k_3^-([A_i] + [A_i'] + [A_iA_i'])) \tag{B19}$$

but inverting these relationships may not lead to simple expressions.

On the slower timescale, the effective dynamics is given by

$$\partial_t[C_i]_3 = \delta([C]_0 - [C_i]_3) + \sum_{j\sim i}(k_1^-[C_iA_j] - k_1^+[C_i][A_j]) \tag{B20}$$

$$\partial_t[A_i]_3 = k_0[C_i{\cdot}A_iA_i'] - \delta[A_i]_3 - \sum_{j\sim i}(k_1^-[C_jA_i] - k_1^+[C_j][A_i]) - (k_2^-[B_iA_i] - k_2^+[A_i][B_i]) \tag{B21}$$

$$\partial_t[A_iA_i']_3 = \delta([R]_0 - [A_iA_i']_3) - 2k_0[C_i{\cdot}A_iA_i'] \tag{B22}$$

where $[C_i]$, $[A_i]$ and $[C_i{\cdot}A_iA_i']$ are understood as functions of $[A_i]_3$, $[A_iA_i']_3$ and $[C_i]_3$ so that we have effectively eliminated them.

To simplify the formula, we take the limit $k_3^+/k_3^- \ll 1$ where $[A_i]_3 \simeq [A_i]$, $[A_iA_i']_3 \simeq [A_iA_i']$, $[C_i]_3 \simeq [C_i]$ and $[C_iA_iA_i'] \simeq k_3^+/k_3^-[C_i][A_iA_i']$, leading to

$$\partial_t[C_i] = \delta([C]_0 - [C_i]) + \sum_{j\sim i}(k_1^-[C_iA_j] - k_1^+[C_i][A_j]) \tag{B23}$$

$$\partial_t[A_i] = k_0k_3^+/k_3^-[C_i][A_iA_i'] - \delta[A_i] - \sum_{j\sim i}(k_1^-[C_jA_i] - k_1^+[C_j][A_i]) - (k_2^-[B_iA_i] - k_2^+[A_i][B_i]) \tag{B24}$$

$$\partial_t[A_iA_i'] = \delta([R]_0 - [A_iA_i']) - 2k_0k_3^+/k_3^-[C_i][A_iA_i'] \tag{B25}$$

More precisely, $k_3^+/k_3^- \sim \varepsilon' \gg \varepsilon$ so this limit preserves the (other) timescale separation.

### 5. Elimination of processes on the second fastest timescale $\tau \sim \varepsilon^2$

On the second fastest timescale, we have a quasi-equilibrium of the reaction

$$A_i + B_i \underset{k_2^-}{\overset{k_2^+}{\rightleftharpoons}} B_iA_i \tag{B26}$$

The slow variables are

$$[A_i]_2 = [A_i] + [B_iA_i] \tag{B27}$$

$$[B_i]_2 = [B_i] + [B_iA_i] \tag{B28}$$

in term of which we can express the others using the equilibrium constants. We have an effective dynamics given by

$$\partial_t[A_i]_2 = k_0k_3^+/k_3^-[C_i][A_iA_i'] - \delta[A_i]_2 - \sum_{j\sim i}(k_1^-[C_jA_i] - k_1^+[C_j][A_i]) \tag{B29}$$

$$\partial_t[B_i]_2 = \delta([B]_0 - [B_i]_2) \tag{B30}$$

where $[A_i]$ is understood as function of the slower variables $[A_i]_2$ and $[B_i]_2$ so that we effectively eliminated $[A_i]$, $[B_i]$ and $[B_iA_i]$. The equations for $[C_i]$ and $[A_iA_i']$ are unchanged.

A first simplification is to replace $[B_i]_2$ by $[B]_0$, which is justified since if $[B_i]_2 = [B]_0$ at any time, it remains constant.

A second simplification is to remark that the equation for $[A_i]$ as a function of $[A_i]_2$ and $[B_i]_2$ simplifies in the limit $k_2^-/k_2^+ \ll 1$ to become

$$[A_i] \simeq \max(0, [A_i]_2 - [B]_0) \tag{B31}$$

This equation has a simple interpretation: in the limit of strong affinity, $B_i$ is saturated by any available $A_i$. More precisely, we consider $k_2^-/k_2^+ = \varepsilon'$; note that it is a limit of strong binding, opposite to the limit of weak binding made with $k_3^+/k_3^- = \varepsilon'$.

### 6. Elimination of processes on the third fastest timescale $\tau \sim \varepsilon$

On the third fastest timescale, we have a quasi-equilibrium of the reaction

$$A_i + C_j \underset{k_1^-}{\overset{k_1^+}{\rightleftharpoons}} C_jA_i \qquad (j \sim i) \tag{B32}$$

The slow variables are

$$[A_i]_1 = [A_i] + \sum_{j\sim i}[C_j A_i] \tag{B33}$$

$$[C_i]_1 = [C_i] + \sum_{j\sim i}[C_i A_j] \tag{B34}$$

in term of which we can express the others. We have an effective dynamics given by

$$\partial_t [C_i]_1 = \delta([C]_0 - [C_i]_1) \tag{B35}$$

$$\partial_t [A_i]_1 = k_0 k_3^+/k_3^- [C_i][A_i A_i'] - \delta[A_i]_1 \tag{B36}$$

where $[C_i]$ is understood as a function of the slow variables $[A_i]_1$ and $[C_i]_1$.

We have

$$[C_i] = \frac{[C_i]_1}{1 + k_1^+/k_1^- \sum_{j\sim i}[A_j]} \simeq \frac{[C_i]_1}{1 + k_1^+/k_1^- \sum_{j\sim i} \max(0, [A_j]_2 - [B]_0)} \tag{B37}$$

A first simplification is to assume $[A_i]_1 \simeq [A_i]_2$ which is achieved provided $k_1^+/k_1^- = \Gamma\varepsilon'$ where $\Gamma$ is of order 1 (limit of weak binding).

A second simplification is to assume $[C_i]_1 \simeq [C]_0$ which is justified since $[C_i]_1 = [C]_0$ remains constant.

A third simplification is to remark that $k_0 k_3^+/k_3^- \ll \delta$ since we assumed $k_3^+/k_3^- = \varepsilon'$ so that

$$\partial_t [A_i A_i'] = \delta([R]_0 - [A_i A_i']) - 2k_0 k_3^+/k_3^- [C_i][A_i A_i'] \tag{B38}$$

can be approximated by

$$\partial_t [A_i A_i'] = \delta([R]_0 - [A_i A_i']) \tag{B39}$$

and we can consider $[A_i A_i'] \simeq [R]_0$.

### 7. Coarse-grained model

The final result is

$$\partial_t [A_i]_1 = \frac{k_3^+/k_3^- k_0 [C]_0 [R]_0}{1 + k_1^+/k_1^- \sum_{j\sim i} \max(0, [A_j]_1 - [B]_0)} - \delta[A_i]_1 \tag{B40}$$

Or in terms of order 1 quantities only, since $k_3^+/k_3^- = \varepsilon'$ and $k_1^+/k_1^- = \varepsilon'\Gamma$,

$$\partial_t [A_i]_1 = \frac{\varepsilon' k_0 [C]_0 [R]_0}{1 + \varepsilon'\Gamma \sum_{j\sim i} \max(0, [A_j]_1 - [B]_0)} - \delta[A_i]_1 \tag{B41}$$

We can eliminate $\varepsilon'$ from this formula by rescaling the volume so that $[A_i]_1' = \varepsilon'[A_i]_1$, $[B]_0' = \varepsilon'[B]_0$, $[C]_0' = \varepsilon'[C]$, $[R]_0' = \varepsilon'[R]_0$. This leads to

$$\partial_t [A_i]_1' = \frac{k_0 [C]_0' [R]_0'}{1 + \Gamma \sum_{j\sim i} \max(0, [A_j]_1' - [B]_0')} - \delta[A_i]_1' \tag{B42}$$

Denoting $\Lambda = k_0 [C]_0' [R]_0'$, we arrive at Eq. (1). The remaining parameters are free to take any value of order 1 relative to $\varepsilon'$ and $\varepsilon$.

The coarse-grained model corresponds to two effective reactions,

$$\emptyset \xrightarrow{k_i} A_i \tag{B43}$$

$$A_i \xrightarrow{\delta} \emptyset \tag{B44}$$

with an effective rate $k_i$ of the form

$$k_i = \frac{\Lambda}{1 + \Gamma \sum_{j\sim i} \max(0, [A_j] - B)} \tag{B45}$$

## Appendix C: Demographic noise

We add to the deterministic description fluctuations arising from the finite number of molecules, leading to Poissonian noise in the actual realized number of reactions of each type. We follow standard and general procedures [4] that can be applied to both the microscopic and coarse-grained models, with or without diffusion.

### 1. Reaction rates and the Master equation

Without loss of generality, any system of reactions involving $S$ species $i$ can be decomposed into a set of unidirectional reactions of the form

$$\sum_i \nu_{\ell i}^- X_i \xrightarrow{k_\ell(N)} \sum_i \nu_{\ell i}^+ X_i \tag{C1}$$

where $X_i$ is a symbol for species $i$, $\nu_{\ell i}^{\pm}$ are stoichiometric coefficients and $k_\ell(N)$ is a reaction rate constant. Here $N$ is an $S$-dimensional vector representing the molecular number of each species $i$ and $\ell$ indexes the reaction. This is associated with a master equation of the form

$$\partial_t P(N) = \sum_\ell [a_\ell(N - r_\ell)P(N - r_\ell) - a_\ell(N)P(N)] \tag{C2}$$

with, assuming mass action kinetics,

$$r_\ell = \sum_i (\nu_{\ell i}^+ - \nu_{\ell i}^-)e_i \tag{C3}$$

$$a_\ell(N) = k_\ell(N)\Omega \prod_i \frac{N_i!}{(N_i - \nu_{\ell i}^-)!\Omega^{\nu_{\ell i}^-}} \tag{C4}$$

where $e_i$ is an $S$-dimensional vector with components $(e_i)_j = \delta_{ij}$ and where $\Omega$ is the volume.

### 2. Reduction using timescale separation

In the deterministic coarse-graining procedure (Appendix B), we used timescale separation to reduce the equations to simpler ones. A key result that we now show is that the noise of the fast processes can be neglected, and that the noise of the remaining slowest processes corresponds to the Poisson noise on the fluxes of the coarse-grained rate equations.

To show this, we follow the framework of [5] and take a timescale $\delta t$ which is short compared to the rates, $k\delta t \ll 1$, so that the concentrations do not change much, but where the absolute number changes in this time are large, $k\Omega\delta t \gg 1$ where $\Omega$ represents the volume. The deterministic rates of the processes in Eqs. (A12-A32) are given by $\Omega k_\ell(N)$ for process $\ell$. The number of events $\lambda_\ell(t)$ for process $\ell$ between $t$ and $t + \delta t$ is Poisson-distributed with mean $\Omega k_\ell(N)\delta t$, so that, for large $\Omega$, the probability $P(\lambda_\ell(t))$ satisfies

$$\ln P(\lambda_\ell(t)) \simeq -\Omega k_\ell(N)\delta t \phi\left(\frac{\lambda_\ell(t)}{\Omega k_\ell(N)\delta t}\right) \tag{C5}$$

where $\phi(x) \equiv x \ln x - x + 1$ has a single minimum at $x = 1$. For $\delta t$ chosen as prescribed, the $\lambda_\ell(t)$ are independent [5], so the probability of all rates is simply

$$\ln P[\{\lambda_\ell\}_\ell] = \sum_\ell \ln P(\lambda_\ell(t)) = -\Omega\delta t \sum_l k_\ell(N)\phi\left(\frac{\lambda_\ell(t)}{\Omega k_\ell(N)\delta t}\right) \tag{C6}$$

We now consider a situation where there is timescale separation, so that the rates $k_{\ell'}$ of the fast processes $\ell' \in \mathcal{F}$ are larger by a factor of $1/\epsilon$ compared to others $\ell'' \in \mathcal{S}$ ($k_{\ell'} = \tilde{k}_{\ell'}/\epsilon$ with $\tilde{k}_{\ell'} \sim k_{\ell''}$),

$$\ln P[\{\lambda_\ell\}_\ell] = -\Omega\delta t \left[\frac{1}{\epsilon}\sum_{\ell' \in \mathcal{F}} \tilde{k}_{\ell'}(N)\phi\left(\frac{\lambda_{\ell'}(t)}{\Omega k_{\ell'}(N)\delta t}\right) + \sum_{\ell'' \in \mathcal{S}} k_{\ell''}(N)\phi\left(\frac{\lambda_{\ell''}(t)}{\Omega k_{\ell''}(N)\delta t}\right)\right] \tag{C7}$$

Then, deviations of $x = \lambda_{\ell'}(t)/(\Omega k_{\ell'}(N)\delta t)$ from $x = 1$, where $\phi(x) = 0$, are small, since they are highly unlikely, with contributions scaling with $\epsilon$ as, $P \sim e^{-\kappa\epsilon^{-1}}$. We can therefore assume that the fast processes happen at the deterministic rate given by $\lambda_{\ell'}(t) = \Omega k_{\ell'}(N)\delta t$. More precisely, this assumes that the fluctuations of the fast processes are at most of the order of the typical contribution of the slow terms: since $k_\ell = O(1/\epsilon)$, we allow $\phi\left(\frac{\lambda_{\ell'}(t)}{\Omega k_{\ell'}(N)\delta t}\right) = O(\epsilon)$, i.e., $\frac{\lambda_i(t)}{\Omega k_{\ell'}(N)\delta t} = 1 + O(\sqrt{\epsilon})$. From the deterministic derivation of Appendix B, this implies that the fast processes are kept at a quasi-equilibrium.

In summary, accounting for demographic noise in the coarse-grained model can be done at the level of the effective slow reactions given by Eqs. (B43)-(B44) without considering the noise associated with the faster reactions that they encompass.

### 3. Langevin equation for typical fluctuations

For the purpose of numerical simulations, it is useful to make a further approximation, by assuming that the fluctuations of $\lambda_\ell(t)$ deviates from its mean as the standard deviation of the distribution, $[\Omega k_\ell(N)\delta t]^{1/2}$. In this case, the Poisson distribution can be approximated by a Gaussian, giving the chemical Langevin equation [6] where

$$\frac{dN_i}{dt} = \sum_\ell \left[ r_{\ell i} a_\ell(N) + r_{\ell i}\sqrt{a_\ell(N)}\xi_\ell(t)\right] \tag{C8}$$

with $\langle \xi_\ell(t)\xi_{\ell'}(t)\rangle = \delta_{\ell\ell'}\delta(t-t')$, in the Itô convention, corresponding to an independent contribution to the noise from each chemical reaction.

We verify that this approximation is justified for the range of parameters that we consider. In particular, flip events occur when one active variable decreases to allow for another to increase, as argued in Appendix E. These events happen when an active variable $[A_1]$ spontaneously decreases to about $B + O(1/\Gamma)$ (and in fact even at higher $[A_1]$, since for finite $\Gamma$ the inactive variable also increases at the same time). In the relevant range, $[A_1] \gtrsim B + O(1/\Gamma)$ and the distribution of $N_1 = \Omega[A_1]$ shows almost no difference between the two noises, as shown in Fig. S9. Since sampling Poisson random variables is computationally much more expensive than sampling Gaussian ones, we use Gaussian noise in our simulations.

### 4. The chemical Langevin equation for the well-mixed coarse-grained model

At the coarse-grained level, the fluctuations of all constituents but $A_i$ can be neglected, as argued above. We have only two effective reactions

$$C_i \to C_i + A_i \tag{C9}$$
$$A_i \to \emptyset \tag{C10}$$

Let $N$ be the vector whose components $N_i$ are the number of $A_i$ and $C_i(N)$ be defined by

$$C_i(N) = \frac{\Lambda\Omega}{1 + \Gamma\sum_{j\sim i}\max(0, N_j/\Omega - [B_j])} \tag{C11}$$

where $\Omega$ is the volume. The master equation has the form

$$\partial_t P_t(N) = \sum_i C_i(N - e_i)P_t(N - e_i) + \delta(N + e_i)P_t(N + e_i) - (C_i(N) - \delta N)P_t(N) \tag{C12}$$

and the chemical Langevin equation is therefore

$$\partial_t N_i = C_i(N) - \delta N_i + \sqrt{C_i(N)}\ \xi_i^{(c)}(t) + \sqrt{\delta N_i}\ \xi_i^{(d)}(t) \tag{C13}$$

which simplifies to

$$\partial_t N_i = C_i(N) - \delta N_i + \sqrt{C_i(N) + \delta N_i}\ \xi_i(t) \tag{C14}$$

where $\xi_i^{(c)}(t)$, $\xi_i^{(d)}(t)$ and $\xi_i(t)$ are Gaussian white noises with unit variance.

With $n_i = N_i/\Omega$ and $c_i(n) = C_i(N = \Omega n)/\Omega$, i.e.,

$$c_i(n) = \frac{\Lambda}{1 + \Gamma \sum_{j\sim i} \max(0, n_j - [B_j])} \tag{C15}$$

we obtain

$$\partial_t n_i = c_i(n) - \delta n_i + \omega\sqrt{c_i(n) + \delta n_i}\ \xi_i(t) \tag{C16}$$

where

$$\omega = \frac{1}{\sqrt{\Omega}} \tag{C17}$$

By redefining time as $t' = \delta t$ and concentrations as $[A_i]' = \delta[A_i]/\Lambda$, we can consider $\Gamma' = \Lambda\Gamma/\delta$ and $\rho' = \rho\delta/\Lambda$ so as to obtain effectively $\Lambda = 1$ and $\delta = 1$. Only three parameters are left, $\Gamma$, $[B_j] = [B]_0$ and $\omega$.

## Appendix D: Model with diffusion

When space is discretized into cells of extension $\Delta x$, diffusion can be formally described by additional "reactions" describing the transfer of particles from one cell to the next. In the context of a one-dimensional system, the additional reactions for $A_i$ are

$$A_i(x) \xrightarrow{D/\Delta x^2} A_i(x + \Delta x) \tag{D1}$$

$$A_i(x) \xrightarrow{D/\Delta x^2} A_i(x - \Delta x) \tag{D2}$$

where $x \pm \Delta x$ is understood modulo the total space length when considering periodic boundary conditions.

### 1. Coarse-grained model with diffusion

Since $A_i$ is involved in 4 diffusion reactions, each with its own noise, the extension of Eq. (C16) to include one-dimensional diffusion takes the form

$$\begin{aligned}\partial_t n_i(x) = c_i(n(x)) - \delta n_i(x) + D/\Delta x^2(n_i(x + \Delta x) + n_i(x - \Delta x) - 2n_i(x)) + \omega\sqrt{c_i(n) + \delta n_i}\xi_i(x, t)\\ + \omega\sqrt{D}/\Delta x\left(\sqrt{n_i(x + \Delta x)}\xi_i^-(x + \Delta x) + \sqrt{n_i(x - \Delta x)}\xi_i^+(x - \Delta x) - \sqrt{n_i(x)}(\xi_i^-(x, t) + \xi_i^+(x, t))\right)\end{aligned} \tag{D3}$$

In the continuous limit $\Delta x \to 0$, it is more concisely written

$$\partial_t n_i(x) = c_i(n(x)) - \delta n_i(x) + D\partial_x^2 n_i(x) + \omega\sqrt{c_i(n) + \delta n_i}\xi_i(x, t) + \omega\sqrt{2D}\partial_x\left(\sqrt{n_i(x)}\zeta_i(x, t)\right) \tag{D4}$$

where $\xi_i(x, t)$ and $\zeta_i(x, t)$ are Gaussian white noises. By rescaling space, we can always assume $D = 1$.

## Appendix E: Steady-state distribution of well-mixed systems

The results of numerical simulations reported in Fig. 2C suggest that the productivity increases on average with time. Here we explain the mechanism behind this behavior for well-mixed systems. To do so, we calculate the transition probability between states in a limiting case that can be analyzed analytically and leads to a simple picture, the low-noise limit where $\omega = \Omega^{-1/2}$ is small. This limit ensures that the "active" and "inactive" variables are well defined and that transitions are rare. In this limit, we expect an Arrhenius-like behavior for the transition rates, scaling as $\exp\left[-\frac{1}{\omega^2}g(\Gamma, B, \delta, \Lambda)\right]$ (this is called "to exponential accuracy" below.) We are interested in the form of the function $g$ . In addition, we consider the limit where $\Gamma$ is large and $B$ is small while keeping $1/\Gamma \ll B \ll 1$. This is in line with the requirement for bistability of two variables, where $\Gamma$ needs to be large enough, and $1/\Gamma < B$. After this limit is taken, we obtain $g(\delta, \Lambda)$ independent of $\Gamma$ and $B$.

The derivation below reveals the basic mechanism for the growth in productivity: when $\Gamma$ is large, the inhibited variables remain at low values, since it requires very strong noise for a variable to rise while it is inhibited. Transitions

are therefore more likely to occur due to a spontaneous decline of the values of active variables below $B$, at which point they no longer inhibit others. This makes transitions in which less active variables are inactivated more likely than their opposite counterparts, and leads to a preference for states with more active variables. This mechanism extends to models with unequal values of $\Lambda_i$.

Here we first derive the transition probability for one variable, then for two variables, before generalizing to any number of variables. From the transition probabilities, we obtain the stationary probability of each state.

### 1. System with a single variable

For a single isolated variable, the slow effective reactions Eqs (B43)-(B44) are simply

$$\emptyset \xrightarrow{\Lambda} A_1 \tag{E1}$$

$$A_1 \xrightarrow{\delta} \emptyset \tag{E2}$$

Let $N$ be the number of $A_1$ particles. $N \geq 0$ satisfies a Master equation with rates $w_{N\to N+1} = \Lambda\Omega$ and $w_{N+1\to N} = \delta N$ whose stationary distribution is Poissonian,

$$P_{\text{alone}}(N) = \frac{(\Lambda\Omega/\delta)^N}{N!} e^{-\Lambda\Omega/\delta} \simeq \exp\left[-\frac{\Lambda\Omega}{\delta}\phi\left(\frac{\delta N}{\Lambda\Omega}\right)\right] \ . \tag{E3}$$

where the second equality is to exponential accuracy, with $\phi(x) \equiv x\ln x - x + 1$. Since $\phi(x)$ has a single minimum at $x = 1$, $P_{\text{alone}}(N)$ is maximal when $\delta N^*/(\Lambda\Omega) = 1$, i.e. when $[A]^* = \Lambda/\delta$ as expected from the deterministic equations.

### 2. System with two variables

Now consider a system with two variables that reciprocally inhibit each other. If $N_1$ fluctuates around the noiseless value $N_1^* = \Omega\Lambda/\delta$ ("active" state), then $N_2$ fluctuated around the deterministic value

$$N_2^* = \frac{\Omega\Lambda/\delta}{1 + \Gamma\max(0, N_1^*/\Omega - B)} = \frac{\Omega\Lambda/\delta}{1 + \Gamma(\Lambda/\delta - B)} \tag{E4}$$

where we assumed that $\Gamma > 1/B$ and $\Lambda/\delta > B$.

How does a transition occur? When $\Gamma$ is large, it is very unlikely that the noise on $N_2$ will push $N_2$ up against the strong deterministic downward force. Therefore, a transition event at low noise typically requires $N_1$ to go down first, until the downward force acting on $N_2$ no longer scales with $1/\Gamma$, namely when $\Gamma\max(0, N_1/\Omega - B) = \Gamma\,(N_1/\Omega - B) \lesssim 1$. If $N_1$ reaches $N_1 < B\Omega$, this is enough, since then it has no inhibitory effect anymore. All in all, $N_1$ must go down to

$$N_1^{(\text{flip})} = \Omega(B + O(1/\Gamma)) \ . \tag{E5}$$

What is the probability of $N_1$ going from $\Omega\Lambda/\delta$ to $N_1^{(\text{flip})}$? For small noise, $N_1$ spends most of the time fluctuating close to its stable fixed point $N_1^*$. The probability of an excursion from there at small noise, is to exponential accuracy, the stationary probability $N_1^{(\text{flip})}$

$$\ln P\left(N_1^{(\text{flip})}\right) = -\frac{\Lambda\Omega}{\delta}\phi\left(\frac{\delta B}{\Lambda} + O(1/\Gamma)\right) \ . \tag{E6}$$

This can also be derived from a first-passage time calculation [7].

After reaching $N_1^{(\text{flip})} \simeq \Omega B$, $N_1$ must stay at this value for long enough to allow for $N_2$ to grow above $\Omega B$, at which point it will inhibit $N_1$, allowing the transition to be completed. Since $N_2$ starts close to zero, and grows as $\partial_t N_2 = \Omega\Lambda - \delta N_2$, its evolution from this point on is $N_2(t) = \Omega\Lambda/\delta(1 - e^{-\delta t})$, and the time it takes to reach $\Omega B$ is $t = \delta^{-1}\ln\frac{1}{1-B\delta/\Lambda} \simeq B/\Lambda + O\left(B^2\right)$. The probability to remain below $N_1^{(\text{flip})}$ decays exponentially with time and so contributes a term of the form $e^{-\Omega\kappa t} = e^{-\Omega\kappa B + O(B^2)}$ where $\kappa$ is a constant. We now take $B$ to be small and this term is negligible. All together, taking $B \to 0$ (but keeping $1/\Gamma \ll B$).

$$\ln P_{\text{one transition}} \sim -\Omega\left[\frac{\Lambda}{\delta}\phi\left(\frac{\delta}{\Lambda}B\right) + O(B)\right] \to -\Omega\frac{\Lambda}{\delta} = -\frac{\Lambda}{\omega^2\delta} \tag{E7}$$

where we used $\phi(x \to 0) = 1$.

### 3. Generalization and effective detailed-balance

The above argument generalizes to any transition between states with

$$P_{\text{transition}} = P_{\text{one transition}}^{k} \tag{E8}$$

where $k$ is the number of variables that have to go from active to inactive. From this it follows that the jump process satisfies a detailed balance condition. Indeed, consider two states with $F_1$ and $F_2$ active variables, respectively, with $F_{\text{overlap}}$ active variables in common. Then the transition $1 \to 2$ is $P_{\text{transition}}(1 \to 2) = P_{\text{one transition}}^{F_1 - F_{\text{overlap}}}$ and similarly in the other direction. So

$$\frac{P_{\text{transition}}(1 \to 2)}{P_{\text{transition}}(2 \to 1)} = \frac{P_{\text{one transition}}^{-F_2}}{P_{\text{one transition}}^{-F_1}} \tag{E9}$$

This means that for the coarse-grained process of jumps between stable states satisfies detailed balance with an equilibrium distribution

$$P_{\text{eq}} = \left(P_{\text{one transition}}\right)^{-F} \tag{E10}$$

where here $F$ is the number of active variables in a state. This holds to the same level of approximation as above (in particular, $\ln P_{\text{eq}}$ to order $\omega^{-2}$), so that

$$P_{\text{eq}} \simeq \exp\left(\frac{\Lambda F}{\omega^2 \delta}\right) \tag{E11}$$

This equilibrium distribution corresponds to the frequency with which each stable state is sampled in the nonequilibrium steady state of our model.

### 4. Different values of $\Lambda_i$

When $\Lambda_i$ depends on $i$, the above arguments generalize to lead to

$$P_{\text{eq}} \simeq \exp\left[\frac{1}{\omega^2 \delta} \sum_{i \in \text{active}} \Lambda_i\right] \tag{E12}$$

where the productivity $\sum_{i \in \text{active}} \Lambda_i$ reduces to $\Lambda F$ as in Eq. (E11) if all the $\Lambda_i$ are equal.

## Appendix F: Numerical methods

### 1. Numerical implementation

We simulate the chemical Langevin equation using the Euler-Maruyama algorithm, to which we add a maximum function to enforce positivity. The generic Langevin equation Eq. (C8) is thus discretized as

$$x_i(t + \delta t) = \max\left(0, x_i(t) + \sum_\ell \left[r_{\ell i} a_\ell(x) \delta t + r_{\ell i} \sqrt{a_\ell(x)} w_\ell(t) (\delta t)^{1/2}\right]\right), \qquad w_\ell(t) \sim \mathcal{N}(0,1) \tag{F1}$$

We take $\delta t = 10^{-2}$ in a context where other parameters are of order 1.

We discretize space by considering a one-dimensional system with periodic conditions divided into $M$ cells with $D = 1$ and run simulations for a given total time $T$. We record $K < T/\Delta t$ time points distributed in $[0, T]$ on a logarithmic scale. The output of a simulation is therefore a $K \times M \times N$ array $X$ of concentrations $X_{kmn} \geq 0$.

### 2. Visual representation

To represent time evolution and spatial diffusion, we collapse $X_{kmn}$ into the $K \times M$ array $Y_{km}$ and represent the third dimension by a color code. Following approaches developed in analytical chemistry [8], we use a coloring scheme based on similarity between compositions. To do this, we first perform a low-dimensional projection of the trajectories by UMAP (Uniform Mapping Approximation and Projection) [9]. Specifically, we first transform the $K \times M \times N$ output into a $KM \times N$ matrix and then apply an implementation of UMAP in Julia, UMAP.jl [10], using 3 components, the metric `corr_dist`, and otherwise default parameters. The result is a $KM \times 3$ output that we map to the CIELAB color space, which is designed so that numerical differences between values correspond to the amount of change humans see between colors [11].

### 3. States and productivity

Stable states are defined as combinations of active and inactive $A_i$, represented by 1 and 0, respectively. A stable state formally corresponds to an independent set of the graph of inhibitory interactions, where $A_i$ is active if and only if all $A_j$ to which it is connected by inhibitory interactions are suppressed. All stable states can be enumerated using the Bron-Kerbosch algorithm [12, 13], which is used for panels C and D of Fig. 3.

There are two ways to assign a state to a given composition $(X_{km1}, \ldots, X_{kmN})$. One is to note that the distribution of concentrations $[A_i]$ is bimodal (Fig. S4A) and define all $X_{kmi}$ exceeding a threshold, e.g., $1/2$, as active. Another is to first run a deterministic simulation (with $\omega = 0$) and only after a given time ($T = 10^2$) apply the threshold to define the state. We verify that the two approaches lead to equivalent stable states in most cases (Fig. S4B) and we adopt the second approach. The similarity between states is defined as the fraction of active or inactive $A_i$ that they have in common. In Figs. 3-5, the state at the end of a trajectory is obtained by taking the median state over the $M$ cells.

The productivity of a state is simply the fraction of its active $A_i$.

To quantify the number of states reached at some time, we calculate the frequency (or prevalence) $f_k$ of each state $k$ and compute the Shannon diversity, defined as $\exp(-\sum_k f_k \ln f_k)$.

### 4. Figures

Unless otherwise stated, we fix the following parameters: $N = 50$, $\Lambda_i = 1$, $\Gamma = 10$, $[B]_{\text{tot}} = 0.25$, $\omega = 0.1$, $\delta = 1$. We discretize the space in $M = 10$ or 100 cells, keeping a constant diffusion $D = 1$. Since the dynamics are stochastic whenever $\omega > 0$, different sample trajectories are obtained starting from the same initial condition, which we take by default to be $[A_i](x, t = 0) = 0$ for all $i$ and $x$. A particular graph of inhibitory interactions was chosen, which is typical of the ensemble of regular graph of size $N = 50$ and connectivity $c = 3$. Results with different graphs from the same random ensemble are shown in Fig. S3.

In Fig. 2, $\mathcal{N} = 20$ sample trajectories are considered with $M = 100$ cells and a total time $T = 10^5$, sampled logarithmically in $K = 10^4$ time points. One of these sample trajectories is used as an illustration in all panels (see Fig. S2 for the other 19 sample trajectories). The average productivity and the average number of different states over the 20 samples are also shown in panels C and D.

In Fig. 3, $\mathcal{N} = 10^5$ sample trajectories with $M = 10$ and $T = 10^2$ are considered. In panel B, the distribution of similarity is computed using a subsample of $10^4$ sample trajectories. The prevalence of a state is the frequency with which it is obtained as a final result in the $\mathcal{N} = 10^5$ samples and is therefore lower bounded by $10^{-5}$ (dotted line in panel D).

In Fig. 4, $\mathcal{N} = 5.10^4$ competitions are performed, concatenating as input the final compositions at the $M = 5$ first positions of two of the $\mathcal{N} = 10^5$ sample trajectories of Fig. 3. The simulations are run for a time period of $T = 10^2$. For each competition, the state associated with the final evolution, denoted 1+2, is compared with the two states associated with the initial condition, denoted 1 and 2. In panel B, the similarity to the closest of these two states is shown in red. In panel C, pairs of initial states are ranked by the absolute value of their difference of productivity, and the fraction of times the final state is closer to the state with the highest productivity is reported.

In Fig. 5, $\mathcal{N} = 10^3$ sample trajectories are drawn, each with different values of $\Lambda_i$, see Eq. (1), taken uniformly at random in $[0, 1]$. Competitions are then run as in Fig. 4, using the environments in which one of the two initial states

was obtained. The distributions of similarities between the final state and the two initial states are shown, in red for the state evolved in the same environment and in green for the other.

---


[1] M. Muñoz-Basagoiti, O. Rivoire, and Z. Zeravcic, Computational design of a minimal catalyst using colloidal particles with programmable interactions, Soft Matter 19(21), 3933-3939. (2023).

[2] O. Rivoire, How Flexibility Can Enhance Catalysis, Physical Review Letters **131**, 088401 (2023).

[3] S. Bo and A. Celani, Multiple-scale stochastic processes: Decimation, averaging and beyond, Physics Reports **670**, 1 (2017), 1612.04999.

[4] C. W. Gardiner, Handbook of stochastic methods for physics, chemistry and the natural sciences, Springer series in synergetics (1985).

[5] A. Lazarescu, T. Cossetto, G. Falasco, and M. Esposito, Large deviations and dynamical phase transitions in stochastic chemical networks, The Journal of Chemical Physics **151** (2019).

[6] D. Schnoerr, G. Sanguinetti, and R. Grima, Approximation and inference methods for stochastic biochemical kinetics : a tutorial review, Journal of Physics A: Mathematical and Theoretical **50**, 093001 (2017), NoStop

[7] C. R. Doering, K. V. Sargsyan, and L. M. Sander, Extinction times for birth-death processes: Exact results, continuum asymptotics, and the failure of the fokker–planck approximation, Multiscale Modeling & Simulation **3**, 283 (2005).

[8] W. Gardner, R. Maliki, S. M. Cutts, B. W. Muir, D. Ballabio, D. A. Winkler, and P. J. Pigram, Self-organizing map and relational perspective mapping for the accurate visualization of high-dimensional hyperspectral data, Analytical chemistry **92**, 10450 (2020).

[9] L. McInnes, J. Healy, and J. Melville, Umap: Uniform manifold approximation and projection for dimension reduction, arXiv preprint arXiv:1802.03426 (2018).

[10] D. Daudert, Umap.jl, a pure Julia implementation of the Uniform Manifold Approximation and Projection dimension reduction algorithm.

[11] International Commission on Illumination, *Colorimetry – Part 4: CIE 1976 L*a*b* Colour Space*, Joint ISO/CIE Standard ISO 11664-4:2008(E)/CIE S 014-4/E:2007 (CIE Central Bureau, Vienna, Austria, 2007).

[12] C. Bron and J. Kerbosch, Algorithm 457: finding all cliques of an undirected graph, Communications of the ACM **16**, 575 (1973).

[13] D. Eppstein, M. Löffler, and D. Strash, Listing all maximal cliques in sparse graphs in near-optimal time, in *Algorithms and Computation: 21st International Symposium, ISAAC 2010, Jeju Island, Korea, December 15-17, 2010, Proceedings, Part I 21* (Springer, 2010) pp. 403–414.

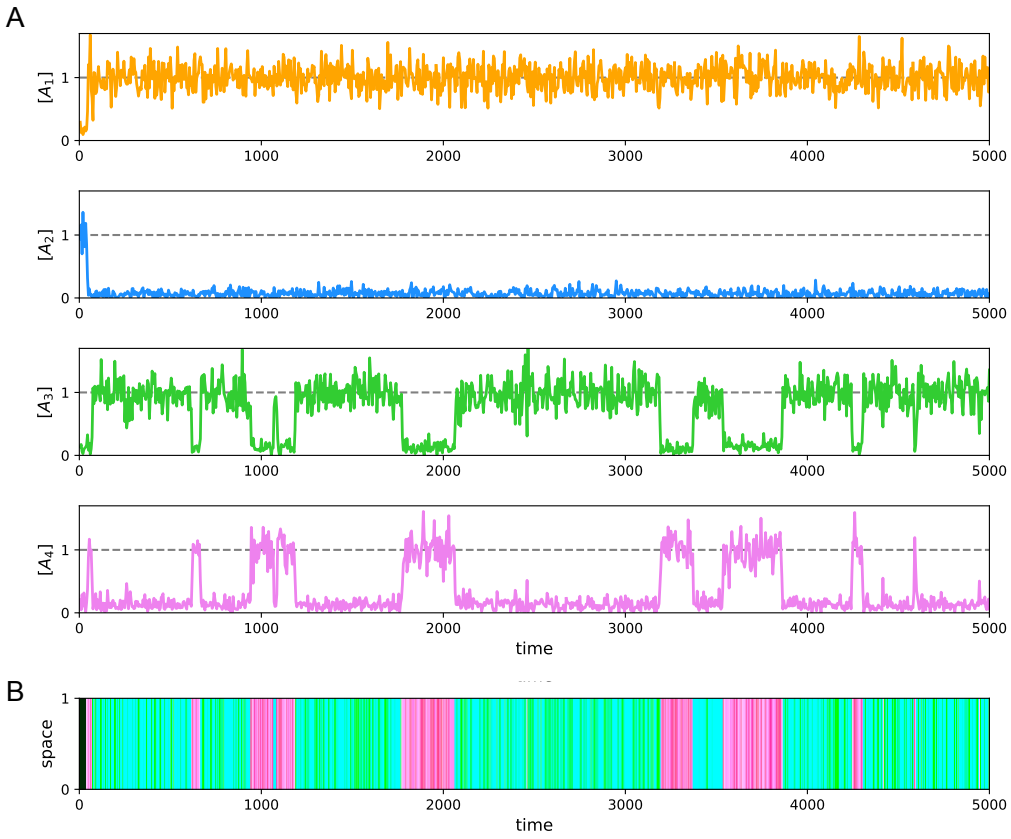



FIG. S1. **A.** Example trajectory for a well-mixed system with $N = 4$ nodes and the inhibition graph shown in Fig. 1C. The initial condition is the state with productivity $P = 1/4 = 0.25$, where the blue node is active and all others are inactive (see Fig. 1D), but the system quickly leaves this state to alternate between the two other states with productivity $P = 2/4 = 0.5$, where the concentrations $[A_i]$ associated with the green and pink nodes alternate between values close to 0 and 1. The parameters are $\Gamma = 10$, $[B_i]_{\text{tot}} = 0.25$ and $\omega = 0.2$. **B.** Representation of the same trajectory using UMAP, showing the alternation between two states except for the very beginning (the colors of the nodes for $A_3$ and $A_4$ in Fig. 1C are chosen to correspond to the colors of the states where they are at high concentration). The system being well-mixed, the composition is uniform across space.

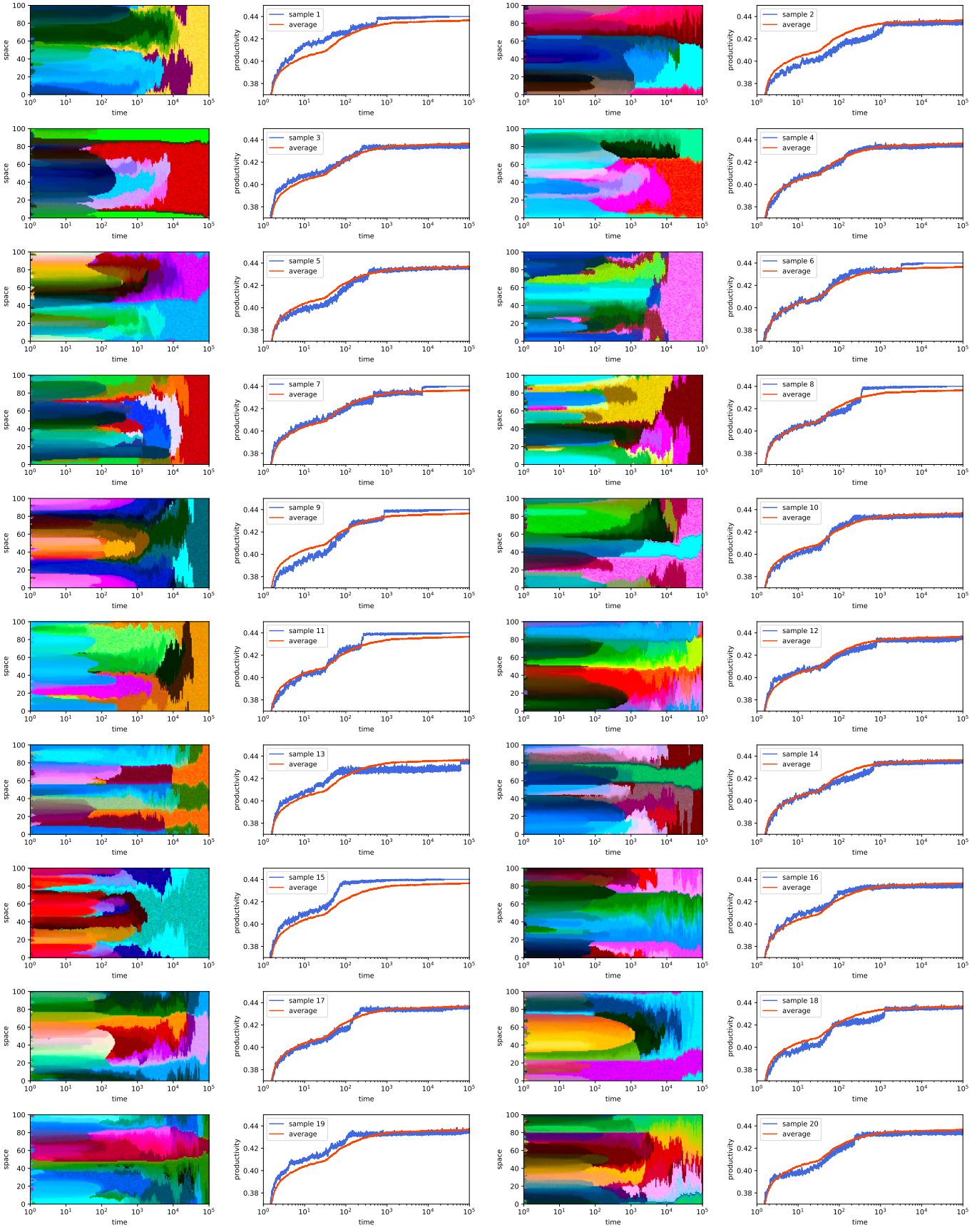



FIG. S2. 20 sample trajectories from the same inhibitory graph as used in Fig. 2 (which shows sample 7).

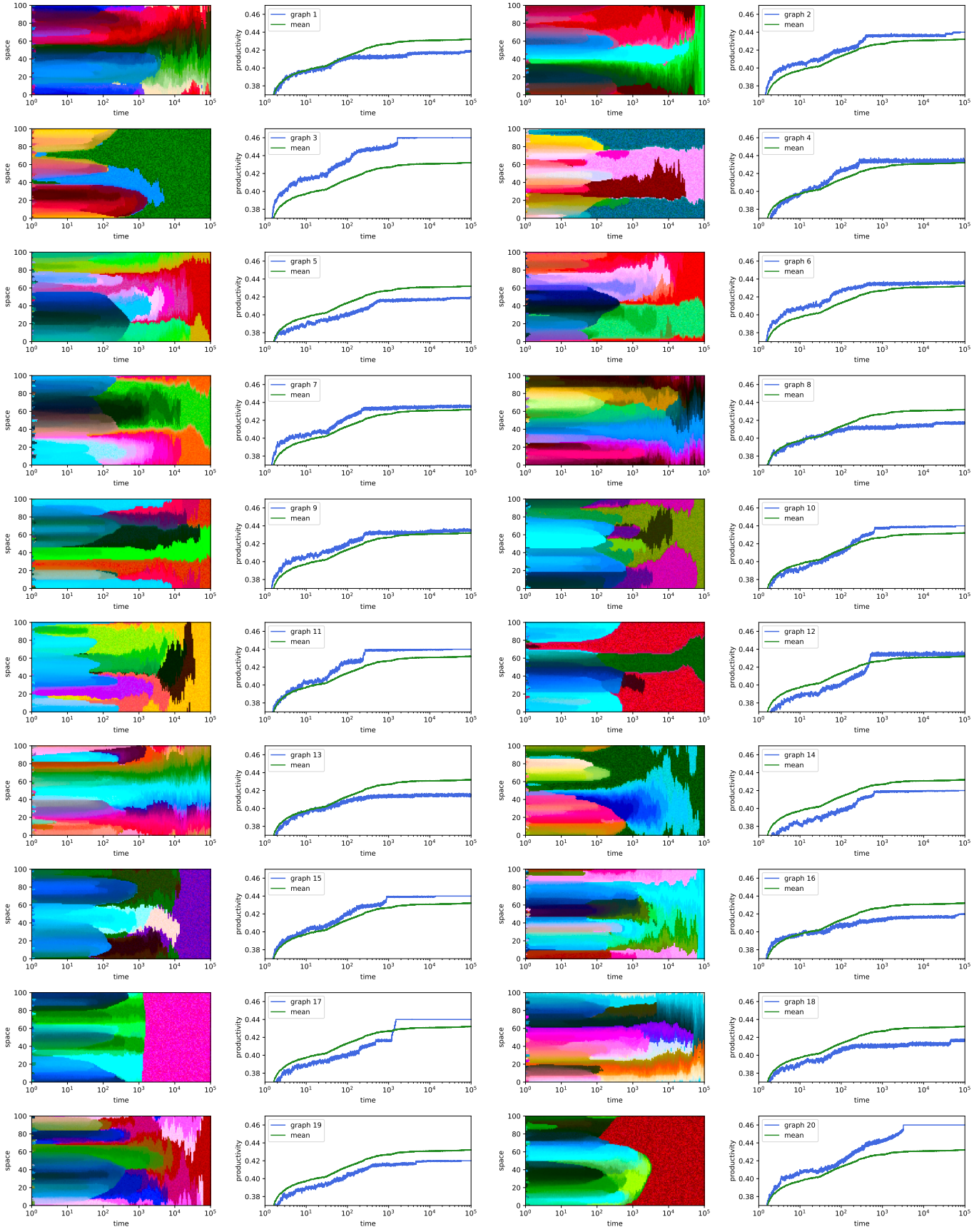



FIG. S3. 20 sample trajectories using different inhibitory graphs (Fig. 2 and Fig. S2 use graph 11). Note that the scale on the y-axis is different than in Fig. S2 to account for graphs that allow for higher levels of productivity, e.g. graphs 3 and 20.

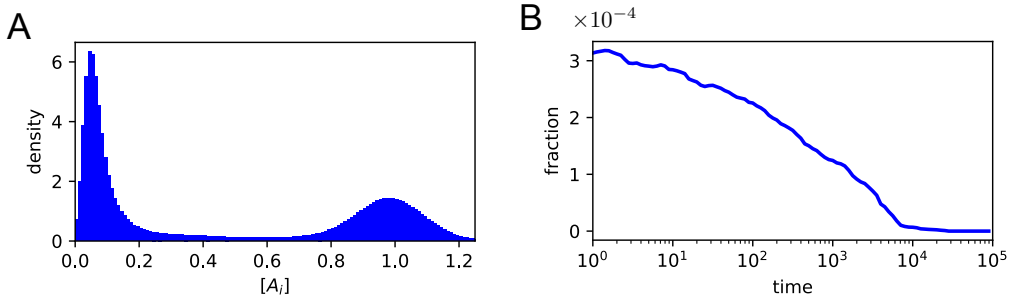



FIG. S4. **A.** Distribution of concentration $[A_i]$ over space and (log) time for the sample trajectory shown in Fig. 2. The distribution is bimodal, reflecting either active or suppressed $A_i$. **B.** An alternative to classifying as active an $A_i$ with concentration $[A_i] > 0.5$ is to use the concentrations of $[A_i]$ as the starting point of a deterministic simulation (without demographic noise) and to identify the stable state to which it necessarily converges. In most cases, the two procedures coincide. The graph shows the fraction of cases where there is a difference, which decreases with time and is in any case of the order of only 0.01%.

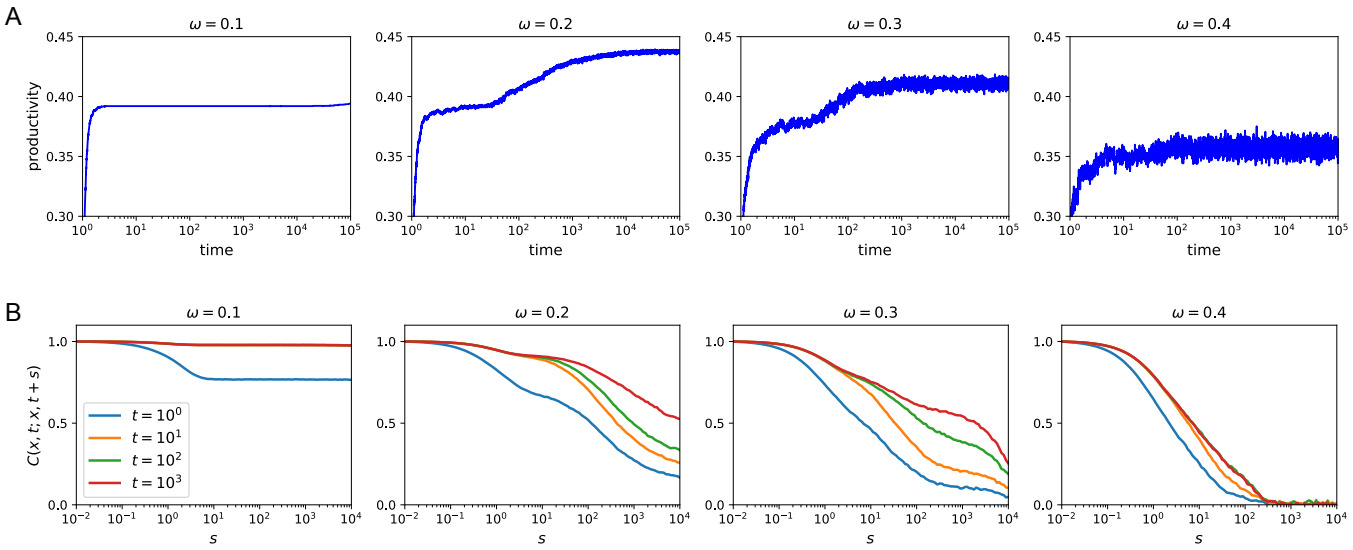



FIG. S5. **A.** Productivity in the well-mixed case for different values of $\omega$, averaged over 100 sample trajectories. **B.** Temporal correlation functions, showing "aging" for small enough values of $\omega$: the correlations grow with increasing values of $t$.

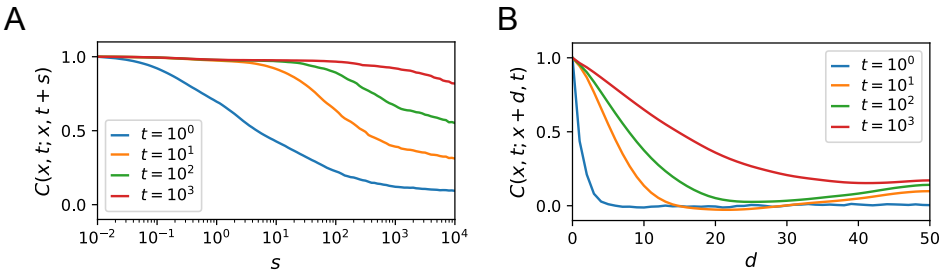



FIG. S6. Temporal and spatial correlation functions for the 20 trajectories shown in Fig. S2. **A.** Temporal correlation function $C(x, t; x, t+s)$ as a function of time interval $s$ for different times $t$, averaged over space $x$ and over the 20 samples. **B.** Spatial correlation function $C(x, t; x+d, t)$ as a function of the spatial interval $d$ for different times $t$, averaged over the 20 samples. The correlations are computed as Pearson correlations using the concentrations $[A_i]$. Both the temporal and spatial correlations increase with time $t$.

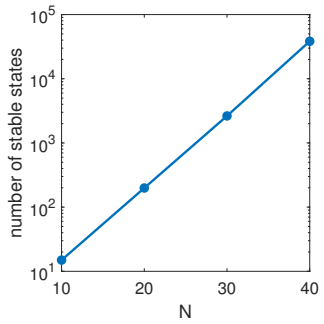



FIG. S7. The number of stable states grows exponentially with graph size $N$. Here shown for random regular graphs with connectivity $c = 3$. Results obtained by averaging over exhaustive counts, using the algorithm of [13]. Errorbars are smaller than marker size.

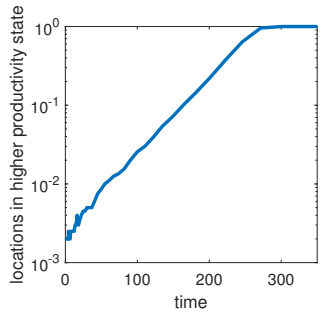



FIG. S8. Exponential growth in the number of spatial locations in a given state is observed when the connections between spatial locations are ordered in appropriate ways. For example, here migration between locations in space forms a random regular graph with degree 3. The number of variables is $N = 4$, and the interacting pairs of variables in a given location are $(1, 2), (1, 3), (1, 4)$. This gives two possible states: only variable 1 is "on" and the rest "off", or $2, 3, 4$ are "on". Four connected locations (a central location and its three neighbors) in the latter state, and the rest in the former. The number of locations is $M = 2000$, and $\omega = 0.14$, $B = 0.25$, $\Gamma = 10$, $D = 0.89$.

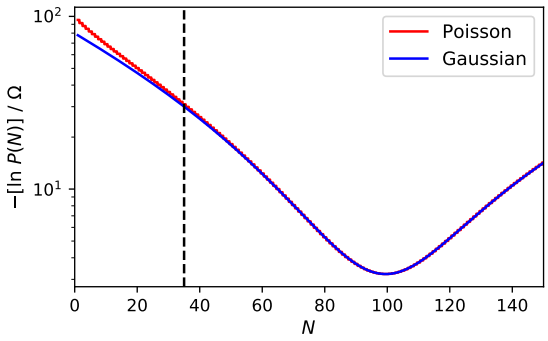



FIG. S9. Probability distribution of $N = \Omega[A_1]$ for a single isolated variable $A_1$ subject to Eqs. (B43)-(B44) with either Gaussian or Poisson noise. The vertical dashed line indicates the threshold $\Omega(B + 1/\Gamma)$ below which the variable would have no inhibitory effect. The parameters $B, \Gamma, \Lambda, \delta$ and $\Omega = 1/\omega^2$ are as in the main text. In the range of values of $N$ relevant to the simulations (roughly, to the right of the vertical dashed line), the two distributions coincide up to differences comparable to rounding of the continuous $N$ to an integer. Values where the two distributions differ more significantly (bottom left) are extremely rare ($P \lesssim 10^{-20}$).