Populations in Reality

Quasispecies truncated by integer particle numbers

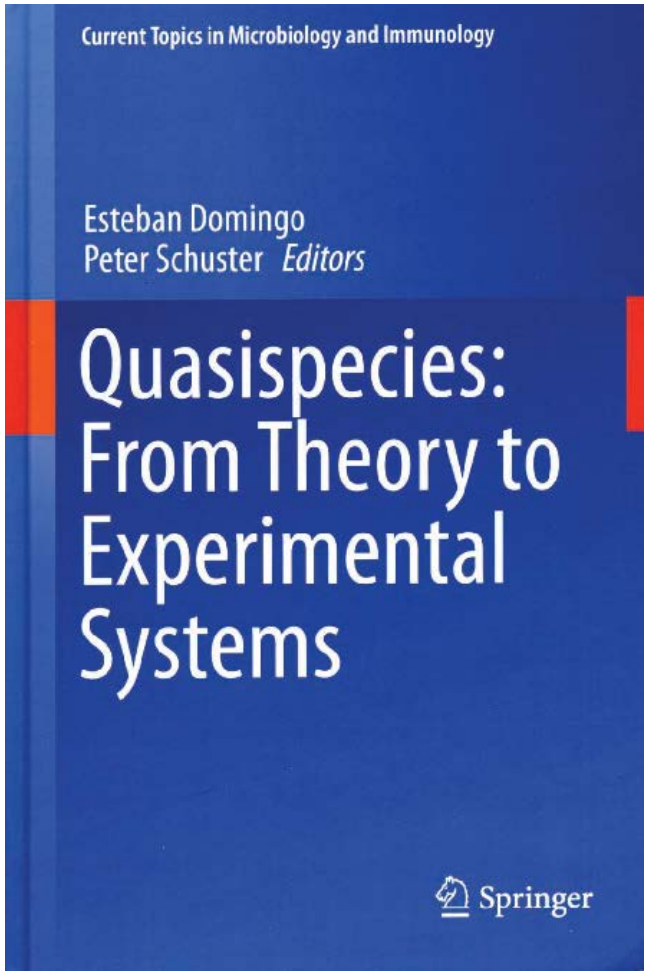
Peter Schuster

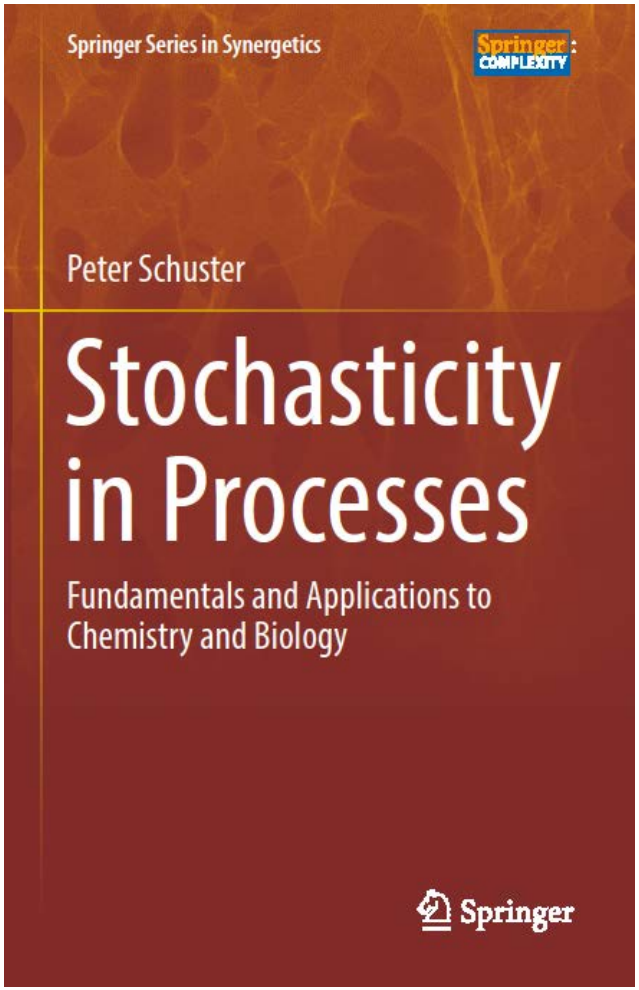
Institut für Theoretische Chemie, Universität Wien, Austria
and
The Santa Fe Institute, Santa Fe, New Mexico, USA

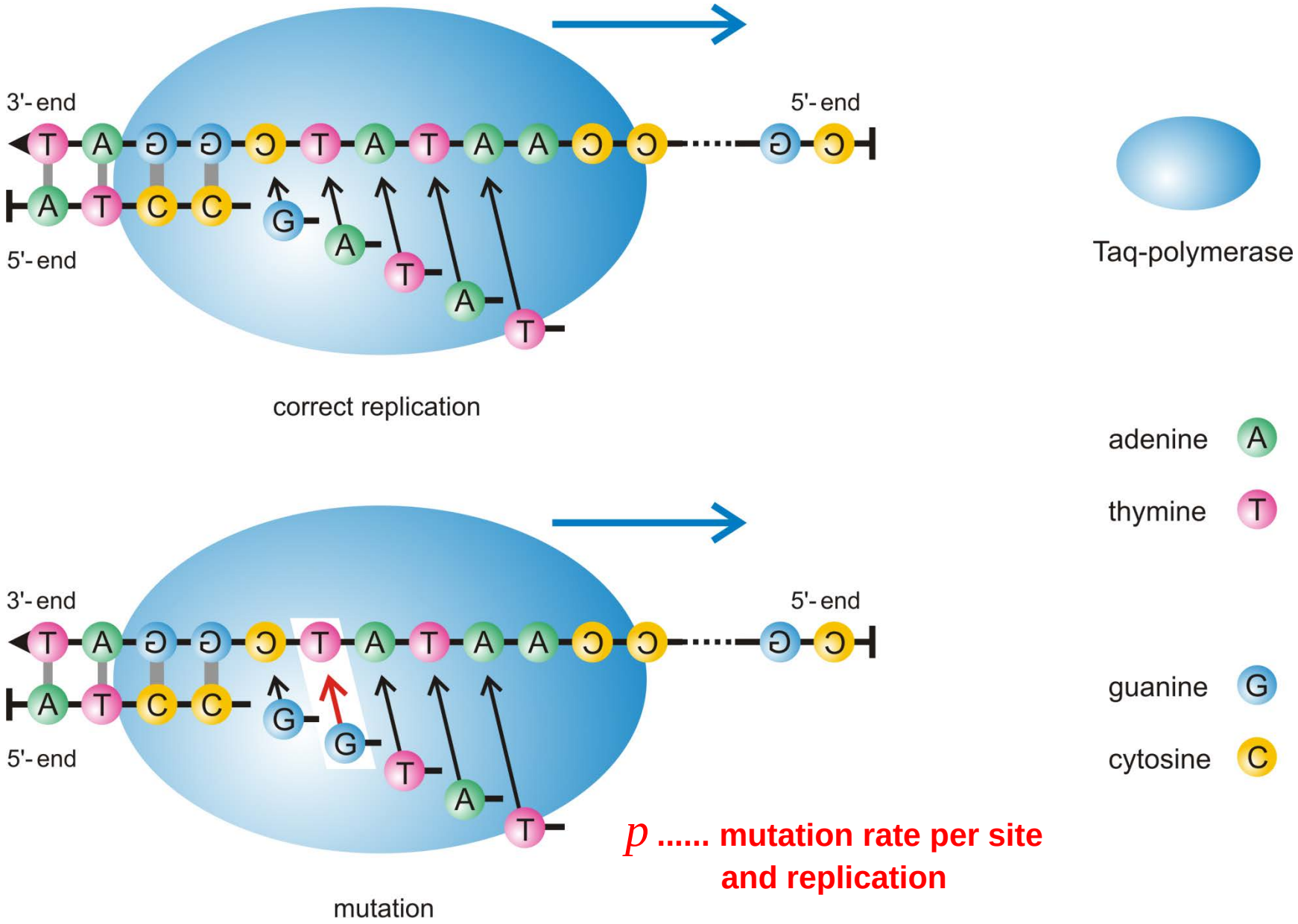
Joint Seminar TBI - KLI

Wien, 27.09.2017

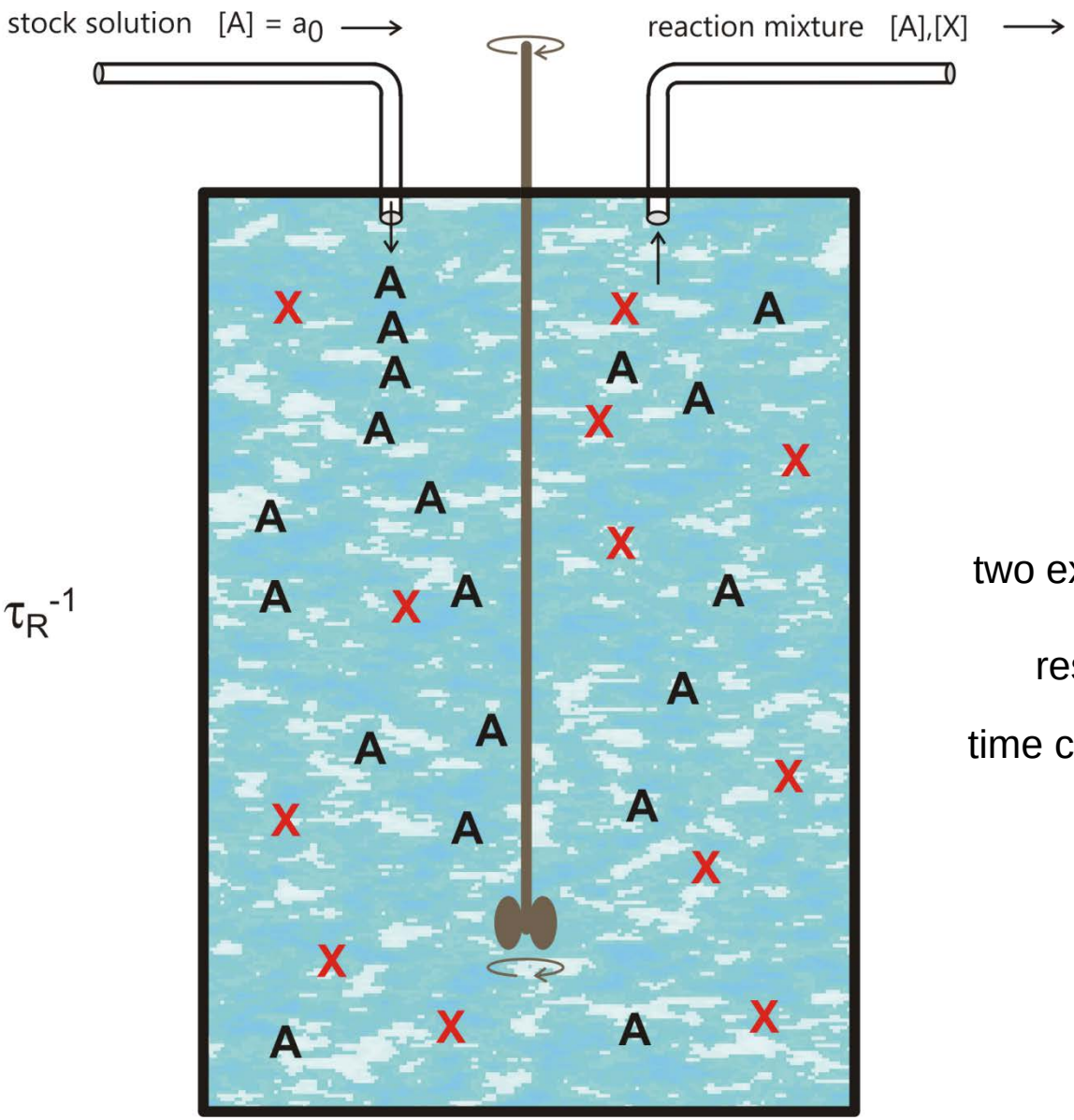
Web-Page for further information:
http://www.tbi.univie.ac.at/~pks







DNA replication and mutation



flow rate $r = \tau_R^{-1}$

two external parameters:

resources **A** … a_0

time constraint … $\tau_R = r^{-1}$

The continuously fed stirred tank reactor (CFSTR)

quasispecies in the flow reactor

$$
\begin{aligned}
* &\xrightarrow{a_0\, r} \mathbf{A}\ , \\
\mathbf{A} + \mathbf{X}_i &\xrightarrow{k_i\, Q_{ji}} \mathbf{X}_i + \mathbf{X}_j\ ;\quad i,j = 1,\ldots,n\ , \\
\mathbf{A} &\xrightarrow{r} \varnothing\ ,\quad \text{and} \\
\mathbf{X}_i &\xrightarrow{r} \varnothing\ ;\ i = 1,\ldots,n
\end{aligned}
$$

$$
[\mathbf{A}] \;=\; a \quad \text{and} \quad [\mathbf{X}_j] \;=\; x_j\ ;\quad j = 1,\ldots,n
$$

$$
\begin{aligned}
\frac{\mathrm{d}a}{\mathrm{d}t} &= -\,a\left(\sum_{j=1}^{n} k_j\, x_j \;+\; r\right) \;+\; a_0\, r \\
\frac{\mathrm{d}x_j}{\mathrm{d}t} &= \left(\sum_{i=1} Q_{ji} k_i\, x_i\right) a \;-\; x_j\, r\ ;\ j = 1,\ldots,n
\end{aligned}
$$

stationary solutions: (i) $\overline{a} = \overline{c}_0,\ \overline{x}_j = 0\,;\ c_0 = \sum_{j=1}^{n} x_j;\ j = 1,\ldots,n$

$$
\text{(ii)}\ \left(\sum_{i=1}^{n} Q_{ji} k_i\, \overline{x}_i\right) \overline{a} \;-\; \overline{x}_j\, r \;=\; 0\,;\ j = 1,\ldots,n
$$

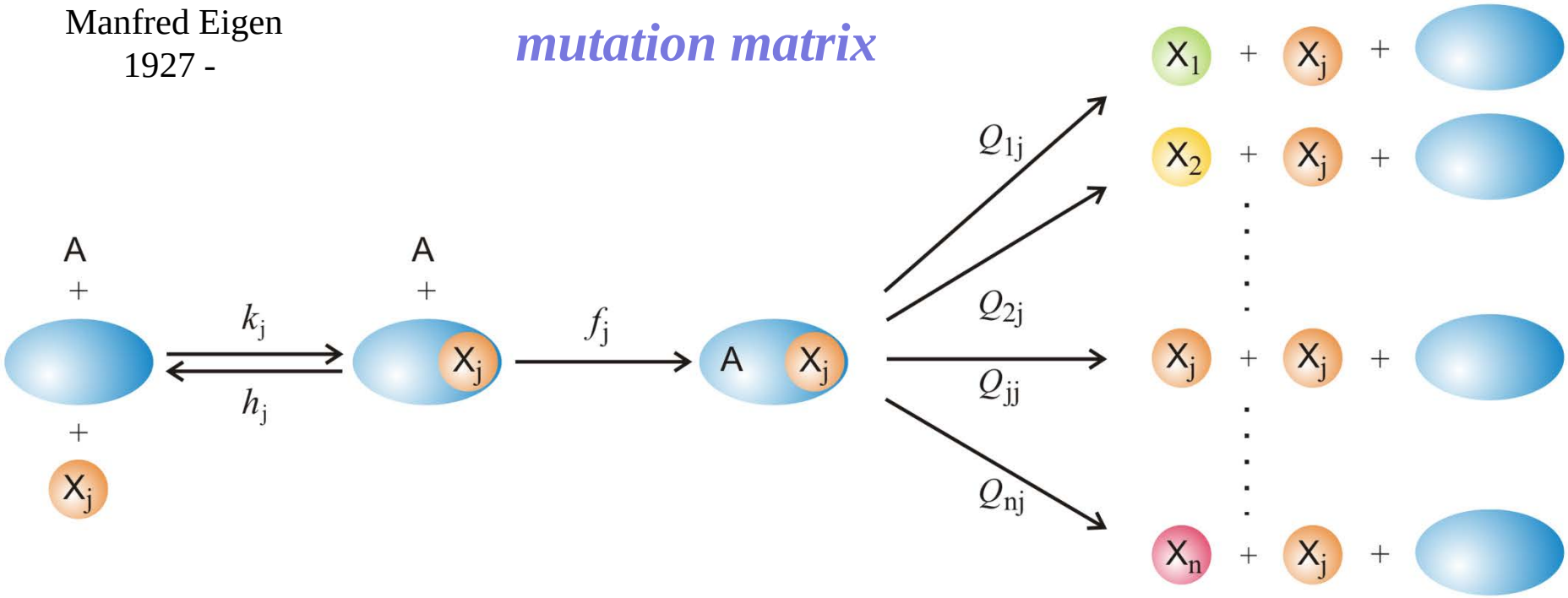
Manfred Eigen
1927 -

$$\frac{dx_j}{dt} = \sum_{i=1}^{n} W_{ji}\, x_i - x_j\, \Phi\,; \quad j = 1, 2, \ldots, n$$

$$W_{ji} = Q_{ji} \cdot f_i\,, \quad \sum_{i=1}^{n} x_i = 1, \quad \Phi = \sum_{i=1}^{n} f_i\, x_i$$

fitness landscape

mutation matrix



Mutation and (correct) replication as parallel chemical reactions

M. Eigen. 1971. *Naturwissenschaften* 58:465,
M. Eigen & P. Schuster.1977-78. *Naturwissenschaften* 64:541, 65:7 und 65:341

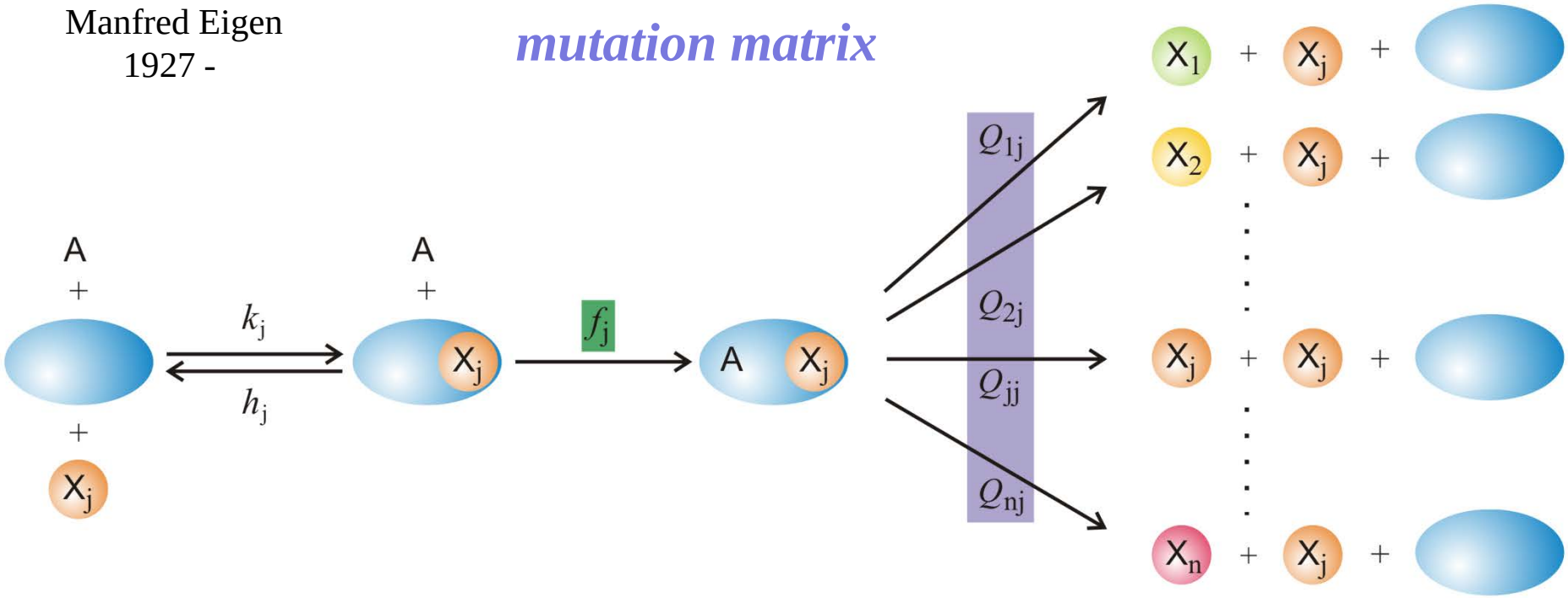
Manfred Eigen
1927 -

$$\frac{dx_j}{dt} = \sum_{i=1}^{n} W_{ji}\, x_i - x_j\, \Phi\,; \quad j = 1, 2, \ldots, n$$

$$W_{ji} = Q_{ji} \cdot f_i\,, \quad \sum_{i=1}^{n} x_i = 1, \quad \Phi = \sum_{i=1}^{n} f_i\, x_i$$

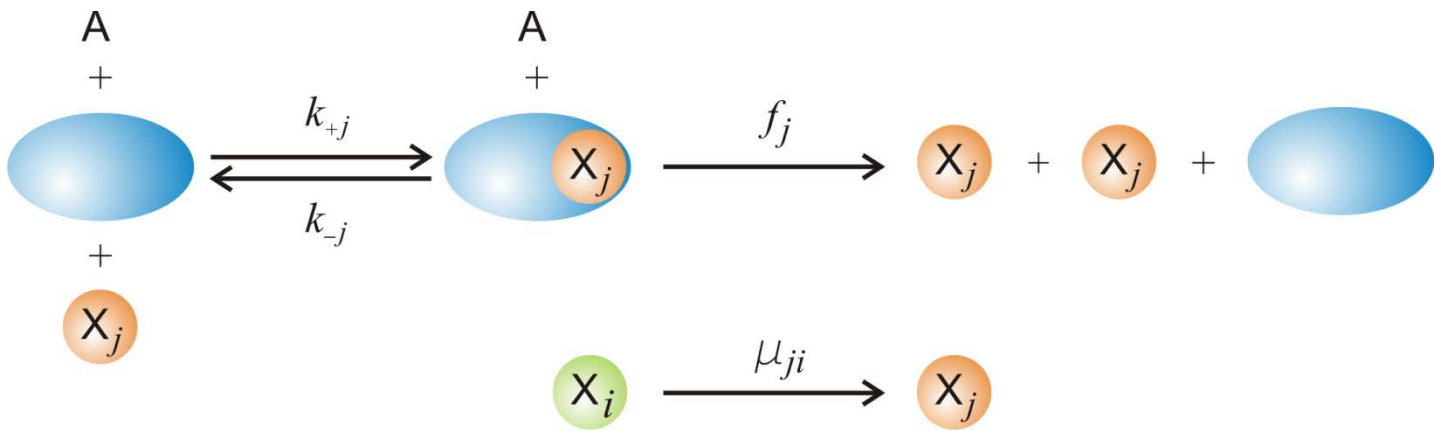
fitness landscape

mutation matrix



Mutation and (correct) replication as parallel chemical reactions

M. Eigen. 1971. *Naturwissenschaften* 58:465,
M. Eigen & P. Schuster.1977-78. *Naturwissenschaften* 64:541, 65:7 und 65:341



$$\frac{\mathrm{d}x_j}{\mathrm{dt}} = \left(f_j - \overline{f(t)}\right) x_j + \sum_{j=1}^{N} \mu_{ji}\, x_j \quad \text{or} \quad \frac{\mathrm{d}\mathbf{x}}{\mathrm{dt}} = \left(\mathrm{F} + \boldsymbol{\mu} - \overline{f(t)}\right) \mathbf{x}$$

paramuse – paralell mutation and selection model:

Ellen Baake, Michael Baake, Holger Wagner. 2001. *Ising quantum chain is equivalent to a model of biological evolution.* Phys.Rev.Letters 78:559-562.

James F. Crow and Motoo Kimura. 1970. *An introduction into population genetics theory*. Harper & Row, New York. Reprinted at the Blackburn Press, Cladwell, NJ, 2009, p.265.

The Crow-Kimura model of replication and mutation

$$\mathrm{Q} = \begin{pmatrix} q^3 & p & p & p & p^2 & p^2 & p^2 & p^3 \\ p & q^3 & p^2 & p^2 & p & p & p^3 & p^2 \\ p & p^2 & q^3 & p^2 & p & p^3 & p & p^2 \\ p & p^2 & p^2 & q^3 & p^3 & p & p & p^2 \\ p^2 & p & p & p^3 & q^3 & p^2 & p^2 & p \\ p^2 & p & p^3 & p & p^2 & q^3 & p^2 & p \\ p^2 & p^3 & p & p & p^2 & p^2 & q^3 & p \\ p^3 & p^2 & p^2 & p^2 & p & p & p & q^3 \end{pmatrix}$$

$$\mu = \begin{pmatrix} -3\gamma & \gamma & \gamma & \gamma & 0 & 0 & 0 & 0 \\ \gamma & -3\gamma & 0 & 0 & \gamma & \gamma & 0 & 0 \\ \gamma & 0 & -3\gamma & 0 & \gamma & 0 & \gamma & 0 \\ \gamma & 0 & 0 & -3\gamma & 0 & \gamma & \gamma & 0 \\ 0 & \gamma & \gamma & 0 & -3\gamma & 0 & 0 & \gamma \\ 0 & \gamma & 0 & \gamma & 0 & -3\gamma & 0 & \gamma \\ 0 & 0 & \gamma & \gamma & 0 & 0 & -3\gamma & \gamma \\ 0 & 0 & 0 & 0 & \gamma & \gamma & \gamma & -3\gamma \end{pmatrix}$$

$$q^3 = (1-p)^3$$

$$\mathrm{Q}: \sum_{i=1}^{N} Q_{ij} = 1;\ j = 1, \ldots, N$$

$$\mu: \sum_{i=1}^{N} \mu_{ij} = 0;\ j = 1, \ldots, N$$

$$\mathrm{W} = \mathrm{Q} \cdot \mathrm{F}$$

$$\mathrm{W} = \boldsymbol{\mu} + \mathrm{F}$$

The mutation matrix in the quasispecies and the Crow-Kimura model

Integrating factor transformation: $$z_j(t) \;=\; x_j(t)\,\cdot\,\exp\left(\int_0^t \overline{f}(\tau)\,\mathrm{d}\tau\right)\;;\; j=1,\ldots,N$$

Eigenvalue problem: $$\Lambda \;=\; \mathrm{H}\cdot\mathrm{W}\cdot\mathrm{H}^{-1} \;=\; \mathrm{H}\cdot\mathrm{W}\cdot\mathrm{B}$$

$$\Lambda = \{\Lambda_{ii}=\lambda_i \;\; i=1,\ldots,N\}\,,\;\; \mathrm{H}=\{h_{ij}\}\,,\;\; \mathrm{H}^{-1}=\mathrm{B}=\{b_{ij}\}$$

$$\mathrm{H}\cdot\mathrm{W} \;=\; \Lambda\cdot\mathrm{H}\,,\;\; \mathbf{h}_k=(h_{ki},\; i=1,\ldots,N) \text{ and}$$

$$\mathrm{W}\cdot\mathrm{B} \;=\; \mathrm{B}\cdot\Lambda\,,\;\; \mathbf{b}_j=(b_{ij},\; i=1,\ldots,N)^{\mathrm{t}}$$

Solution:

$$x_j(t) \;=\; \frac{\sum_{k=1}^N b_{jk}\,\sum_{l=1}^N h_{kl}\,x_l(0)\,\exp(\lambda_k\,t)}{\sum_{i=1}^N\sum_{k=1}^N b_{ik}\,\sum_{l=1}^N h_{kl}\,x_l(0)\,\exp(\lambda_k\,t)} \;=$$

$$=\; \frac{\sum_{k=1}^N b_{jk}\,\beta_k(0)\,\exp(\lambda_k\,t)}{\sum_{i=1}^N\sum_{k=1}^N b_{ik}\,\beta_k(0)\,\exp(\lambda_k\,t)} \quad \text{with} \quad \beta_k(0) \;=\; \sum_{l=1}^N h_{kl}\,x_l(0)$$

Solution of the quasispecies equation

Largest eigenvalue λ_1 and corresponding eigenvector $\mathbf{b}_1$:

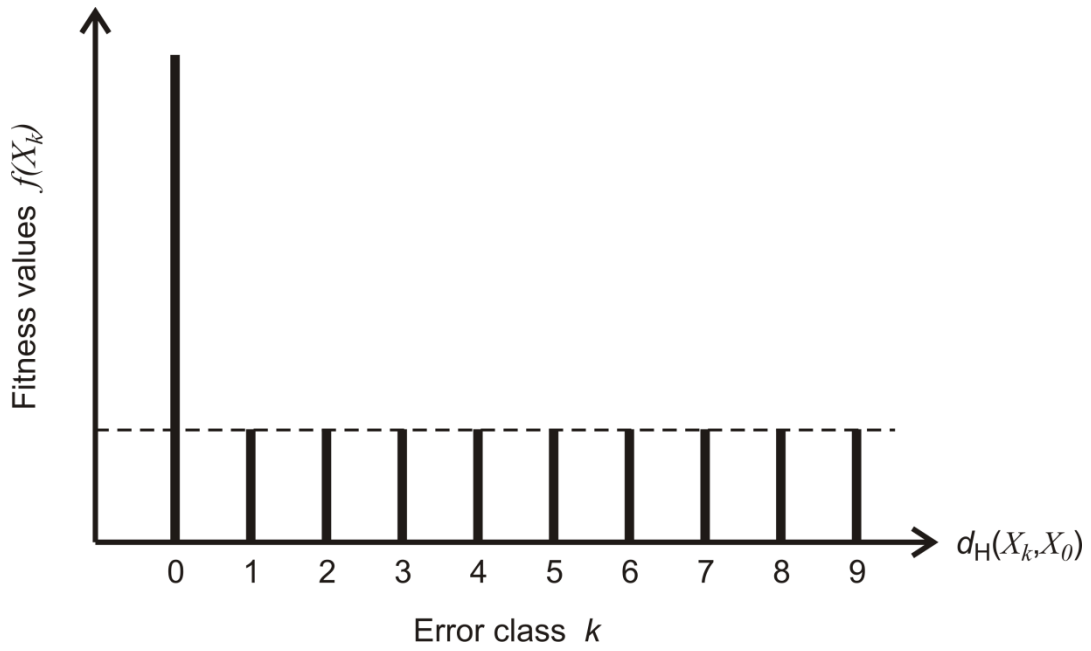
$$x_j(t) \approx \frac{b_{j1}\,\beta_1(0)\,\exp(\lambda_1\,t)}{\sum_{i=1}^{N} b_{i1}\,\beta_1(0)\,\exp(\lambda_1\,t)} = \frac{b_{j1}}{\sum_{i=1}^{N} b_{i1}} = \overline{x}_j \quad \text{for large } t$$

master sequence: X_m at concentration $\overline{x}_m$

mutant cloud: X_j at concentration $\overline{x}_j$; $j=1,\ldots,N$; $j\neq m$

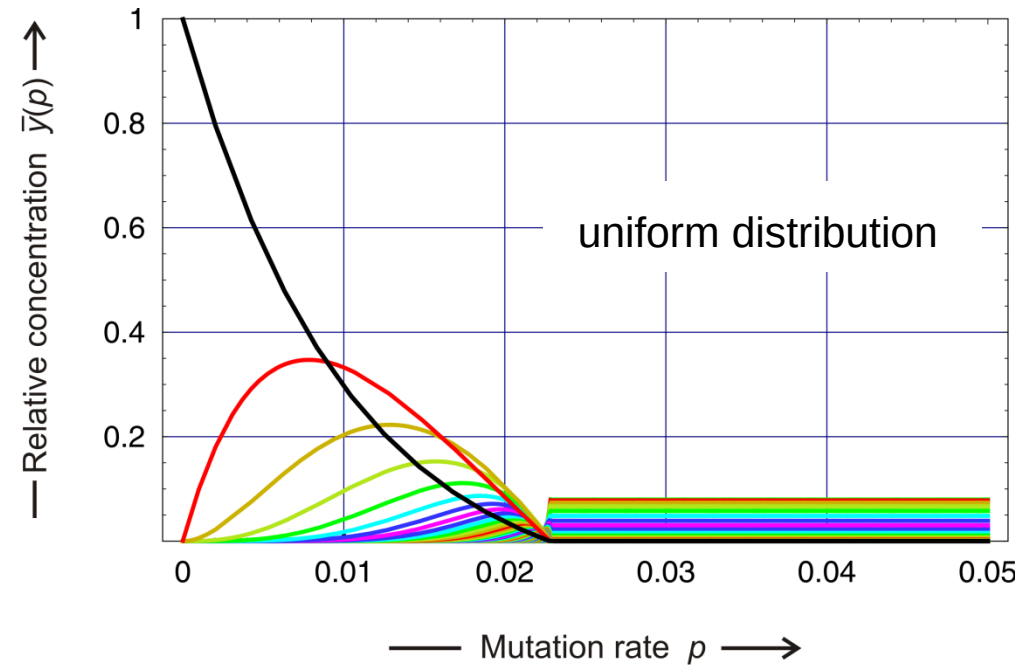
Stationary solution of the quasispecies equation

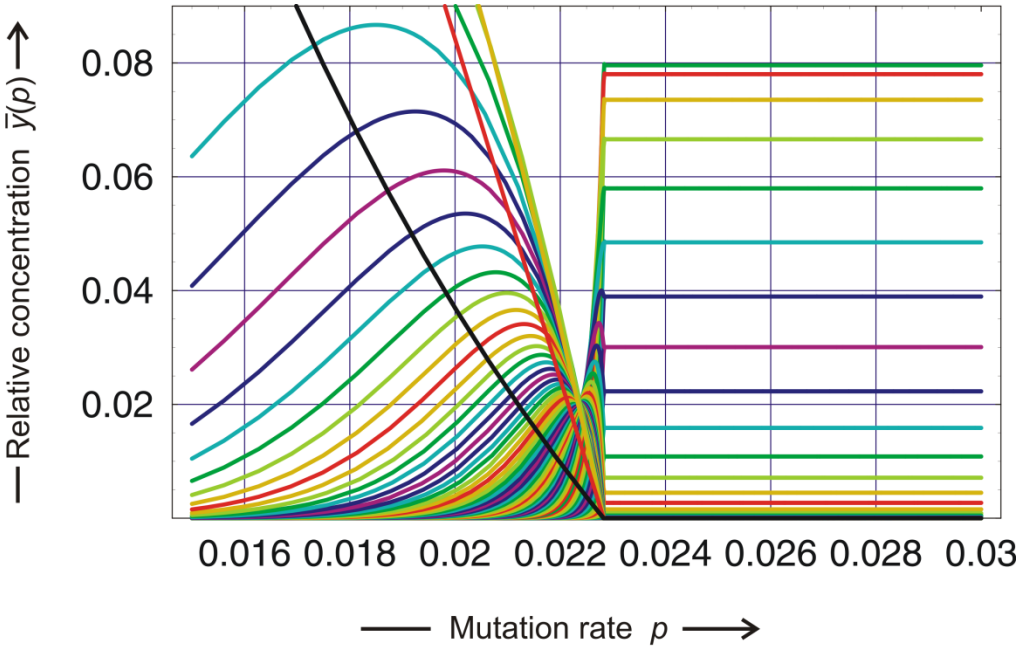
single peak landscape



A simple model fitness landscapes

$l = 100$, $f_m = 10$, $f = 1$,
$\sigma_m = 10$





error threshold on the
single peak landscape

phenomenological approach (Eigen, M., Naturwissenschaften 1971)

(i) zero mutational backflow (non consistently applied)

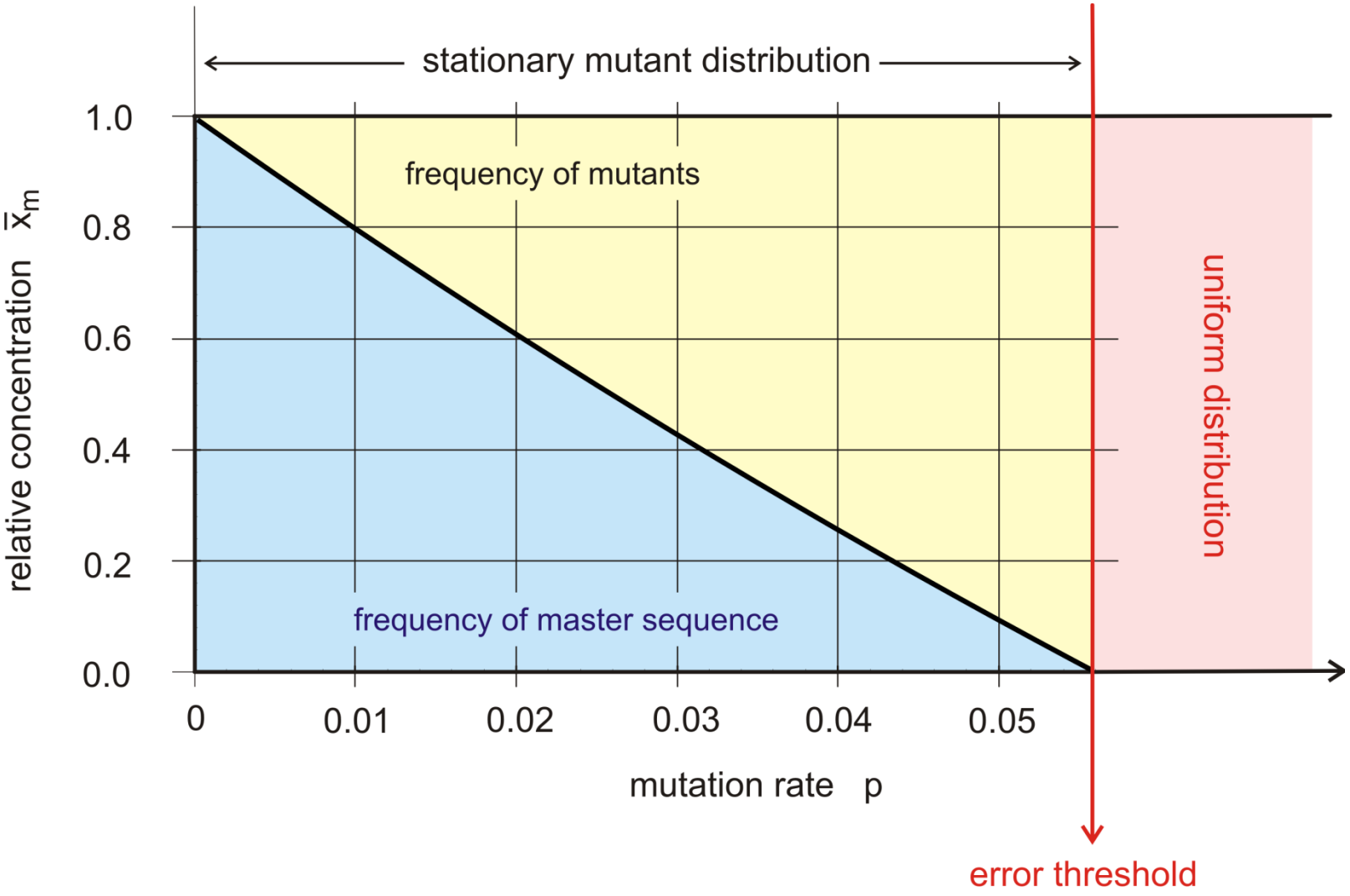
(ii) uniform error rate: p is independent of nature and position of nucleotide

$$Q_{ij} = \varepsilon^{d_{ij}^{(H)}} Q \text{ with } \varepsilon = \frac{p}{1-p} \text{ and } Q = (1-p)^l$$

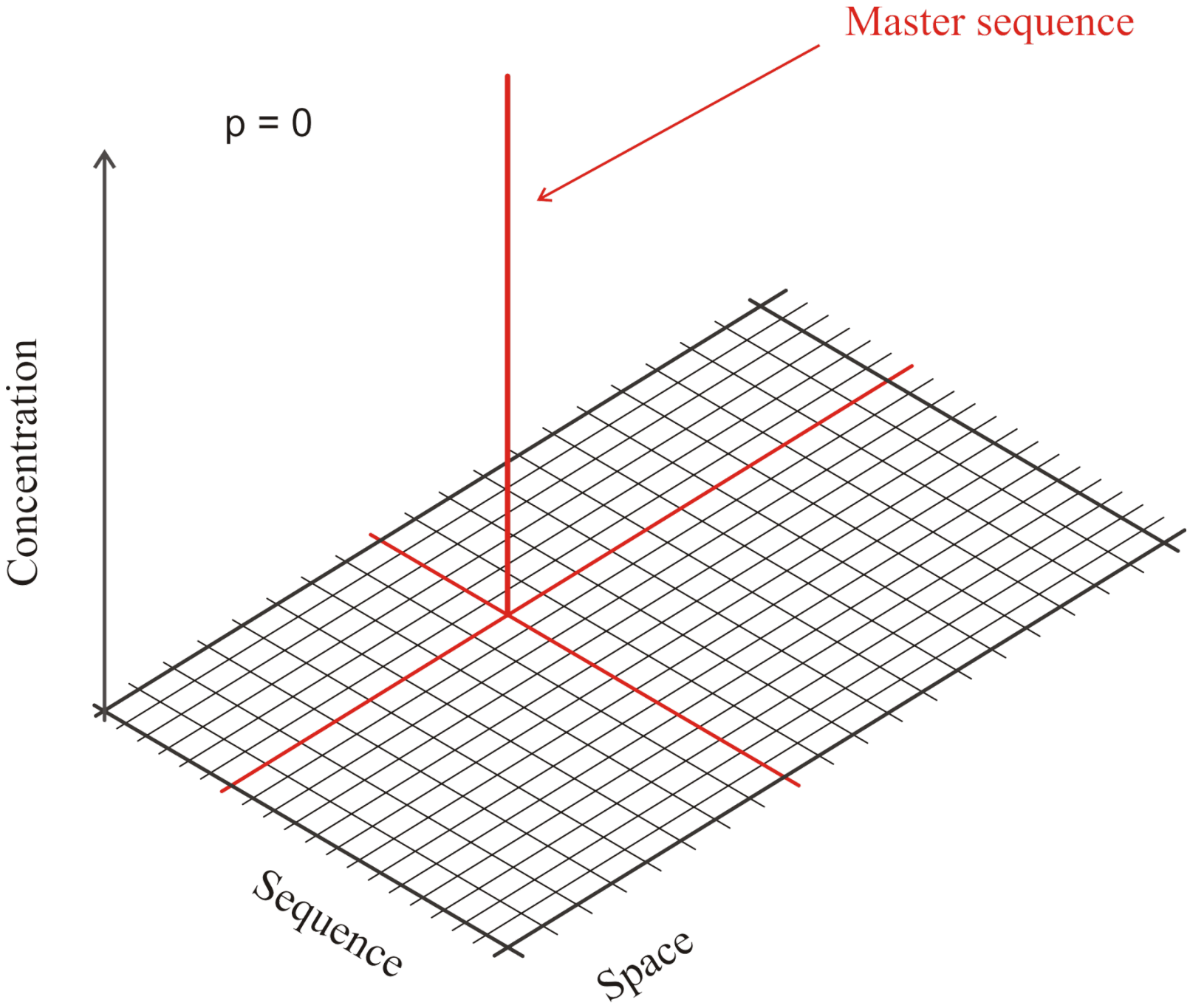
and $d_{ij}^{(H)}$ is the Hamming distance between $\mathbf{X}_i$ and $\mathbf{X}_j$.

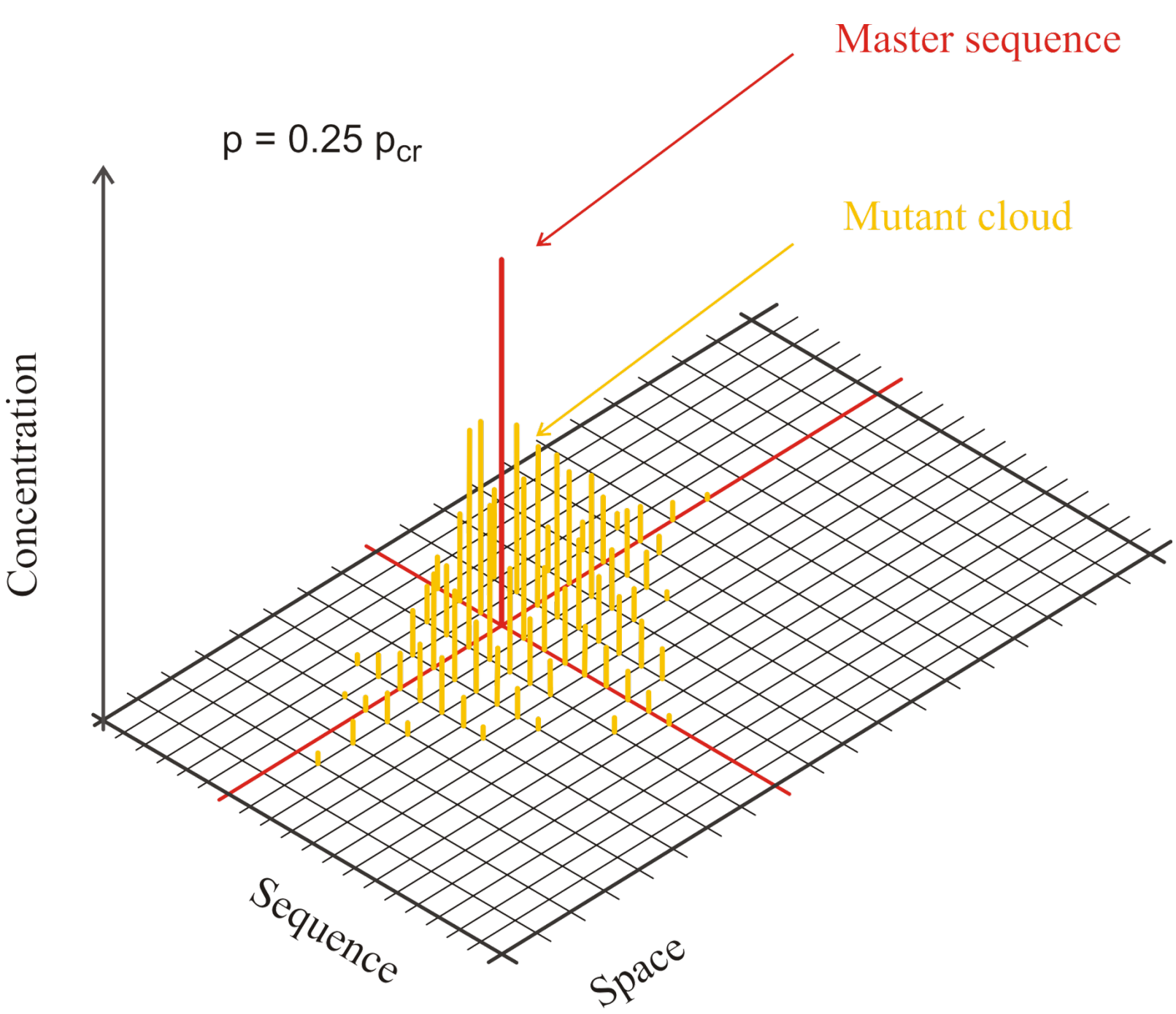
(iii) single peak landscape: $f_m = f_0$, $f_j = f \ \forall\ j \neq m$

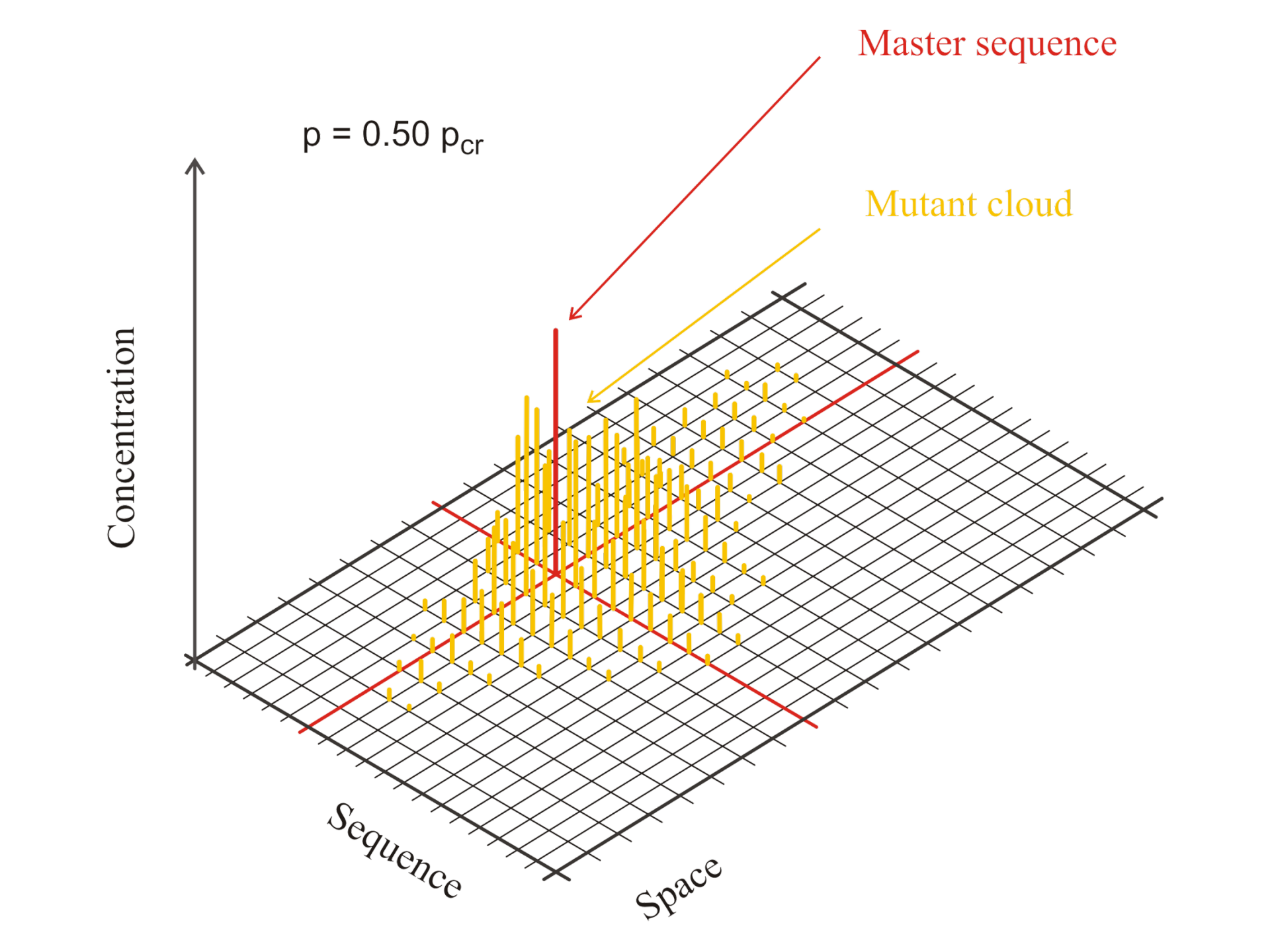
phenomenological approximation to the quasispecies equation

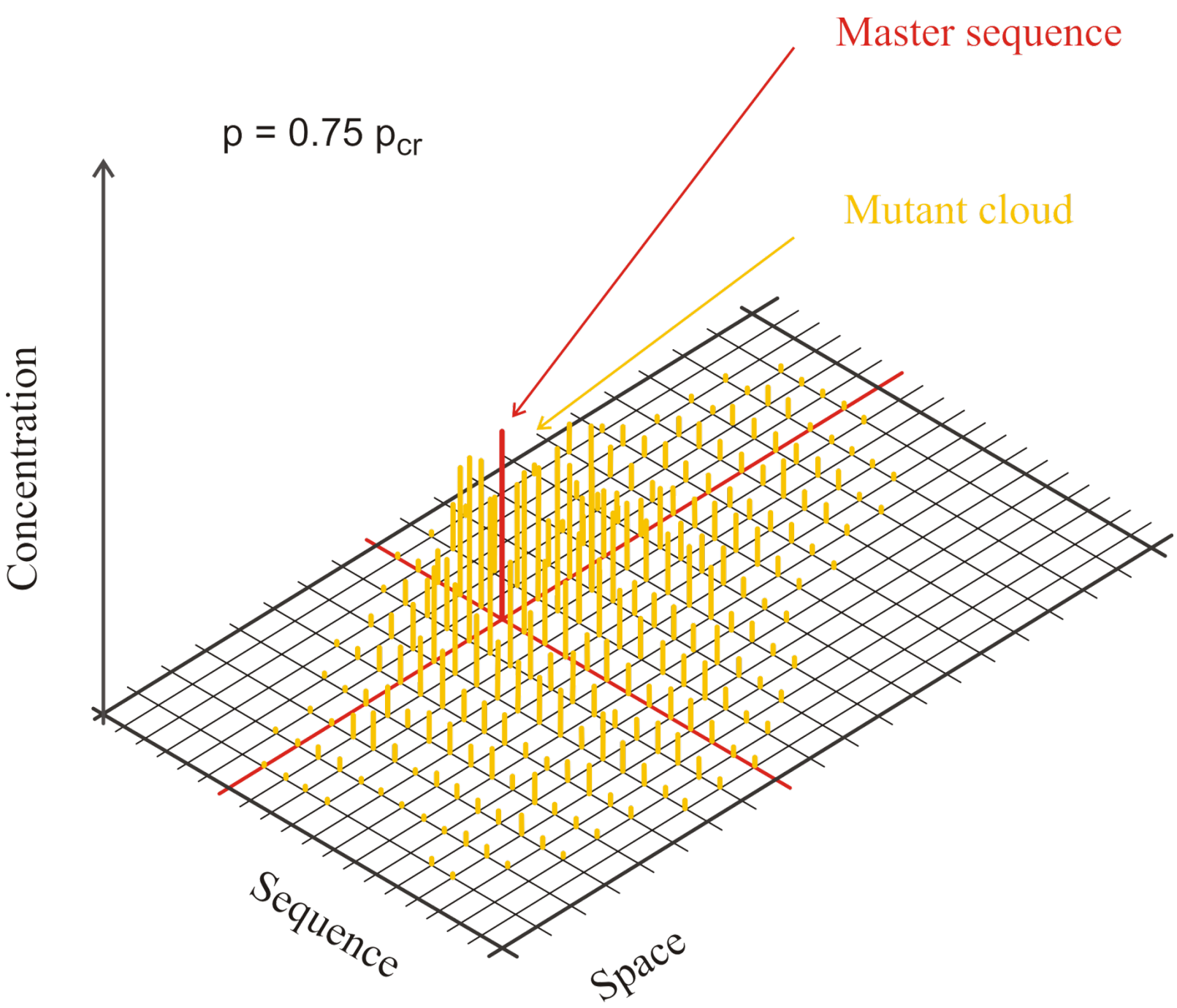


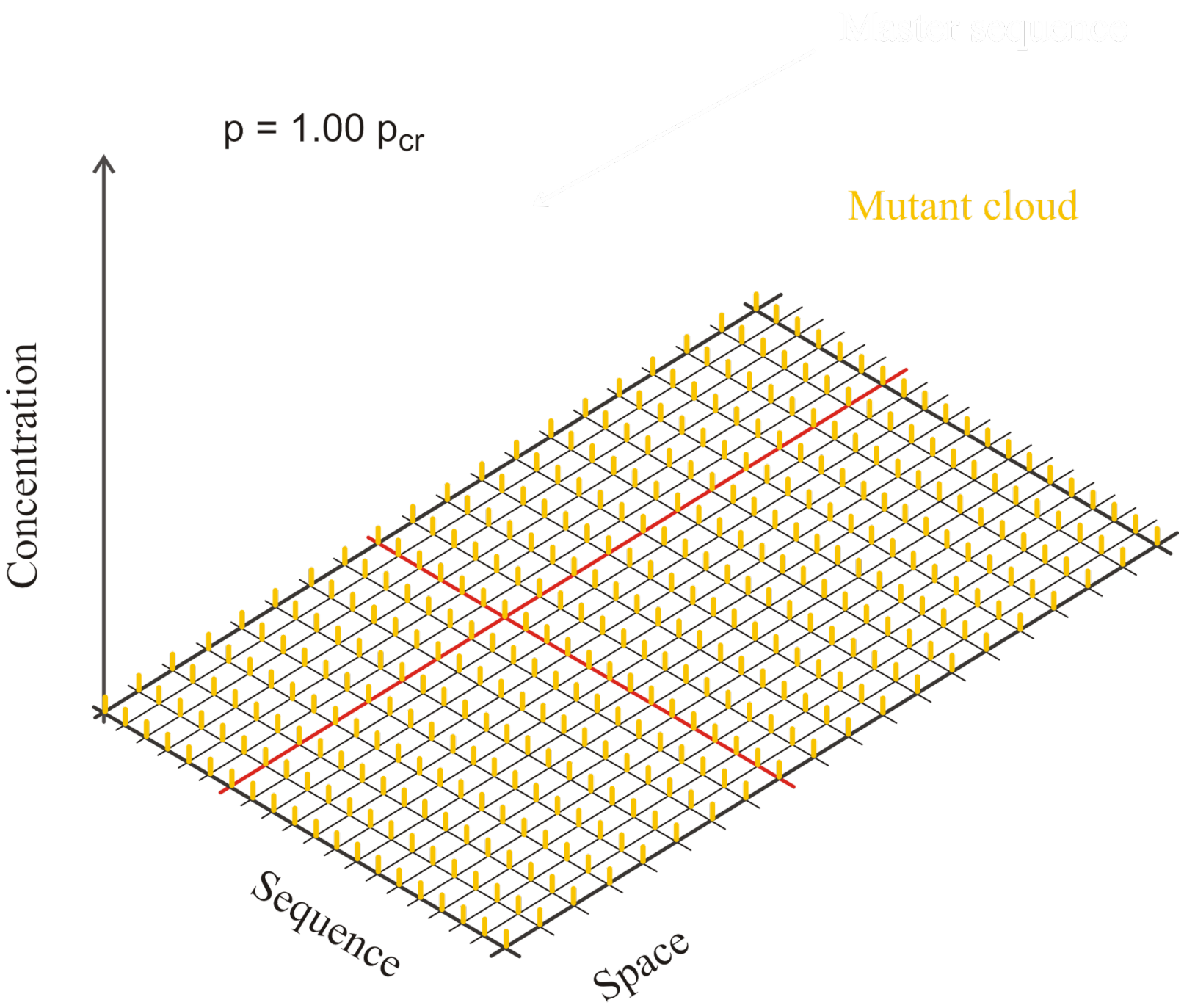
The error threshold in replication and mutation


Master sequence
p = 0
Concentration
Sequence
Space


Master sequence
p = 0.25 p_cr
Mutant cloud
Concentration
Sequence
Space


Master sequence
p = 0.50 p_cr
Mutant cloud
Concentration
Sequence
Space


Master sequence
p = 0.75 p_cr
Mutant cloud
Concentration
Sequence
Space


Master sequence
p = 1.00 p_cr
Mutant cloud
Concentration
Sequence
Space

$$\Pi(t) = x_1(t)\mathbf{X}_1 \oplus x_2(t)\mathbf{X}_2 \oplus \ldots \oplus x_n(t)\mathbf{X}_n$$

$$\lim_{t\to\infty} \Pi(t) = \overline{\Upsilon} = \overline{x}_1\mathbf{X}_1 \oplus \overline{x}_2\mathbf{X}_2 \oplus \ldots \oplus \overline{x}_n\mathbf{X}_n$$

continuous quasispecies

$$\widetilde{\Upsilon} = \overline{\eta}_1\mathbf{Y}_1 \oplus \overline{\eta}_2\mathbf{Y}_2 \oplus \ldots \oplus \overline{\eta}_n\mathbf{Y}_n \quad \text{with} \quad \overline{\eta}_i = \begin{cases} \lfloor \overline{x}_i \rfloor & \text{if } \overline{x}_i \geq 1 \\ 0 & \text{if } \overline{x}_i < 1 \end{cases}$$

discrete quasispecies

$$\widehat{x}_m = \frac{Q - \sigma_m^{-1}}{1 - \sigma_m^{-1}}\,\widehat{c}_0 \quad \text{and}$$

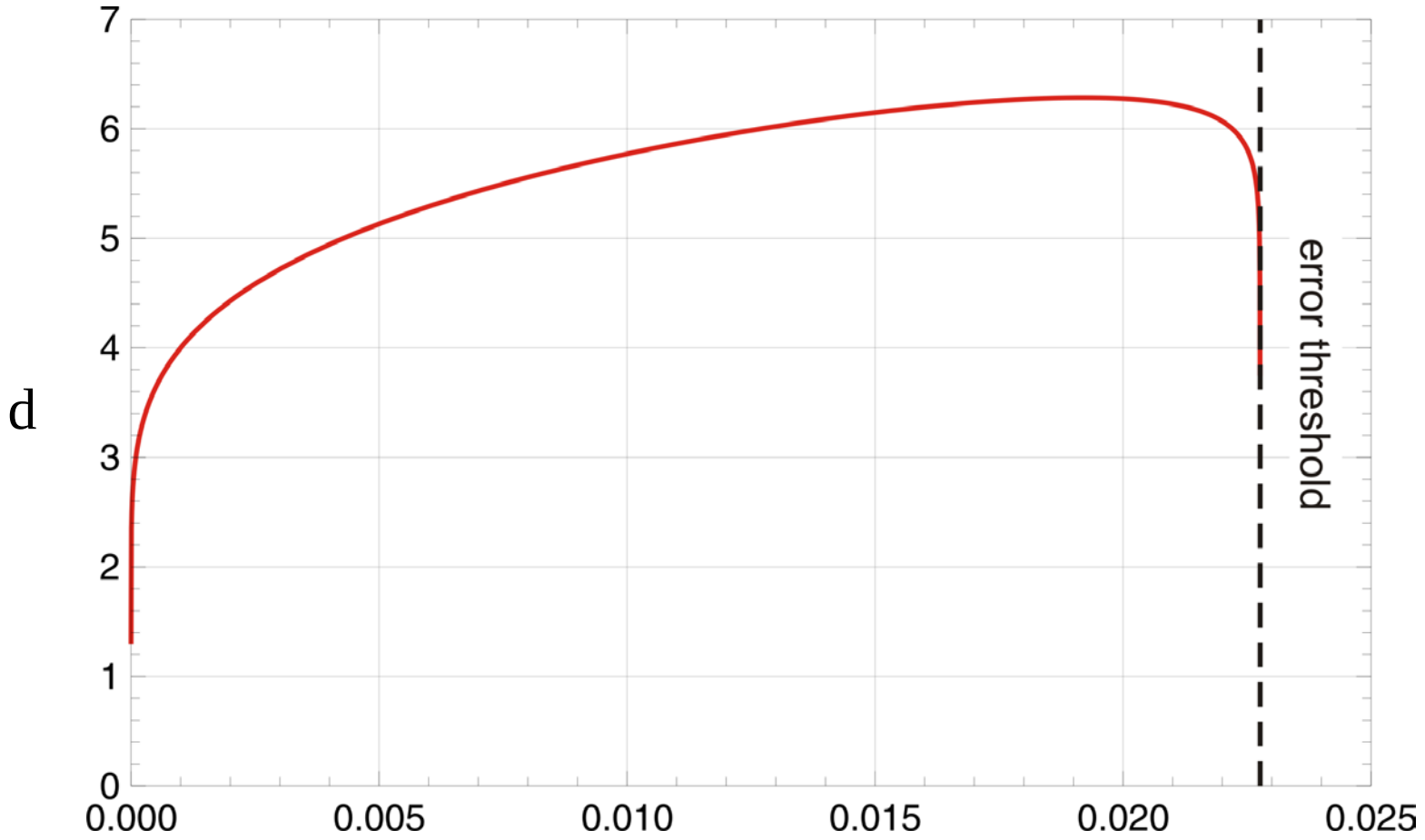
$$\widehat{x}_j = \varepsilon^{d_{mj}^{(H)}}\,\frac{Q - \sigma_m^{-1}}{(1 - \sigma_m^{-1})^2}\,\widehat{c}_0\,, \quad j \neq m$$

phenomenological approximation

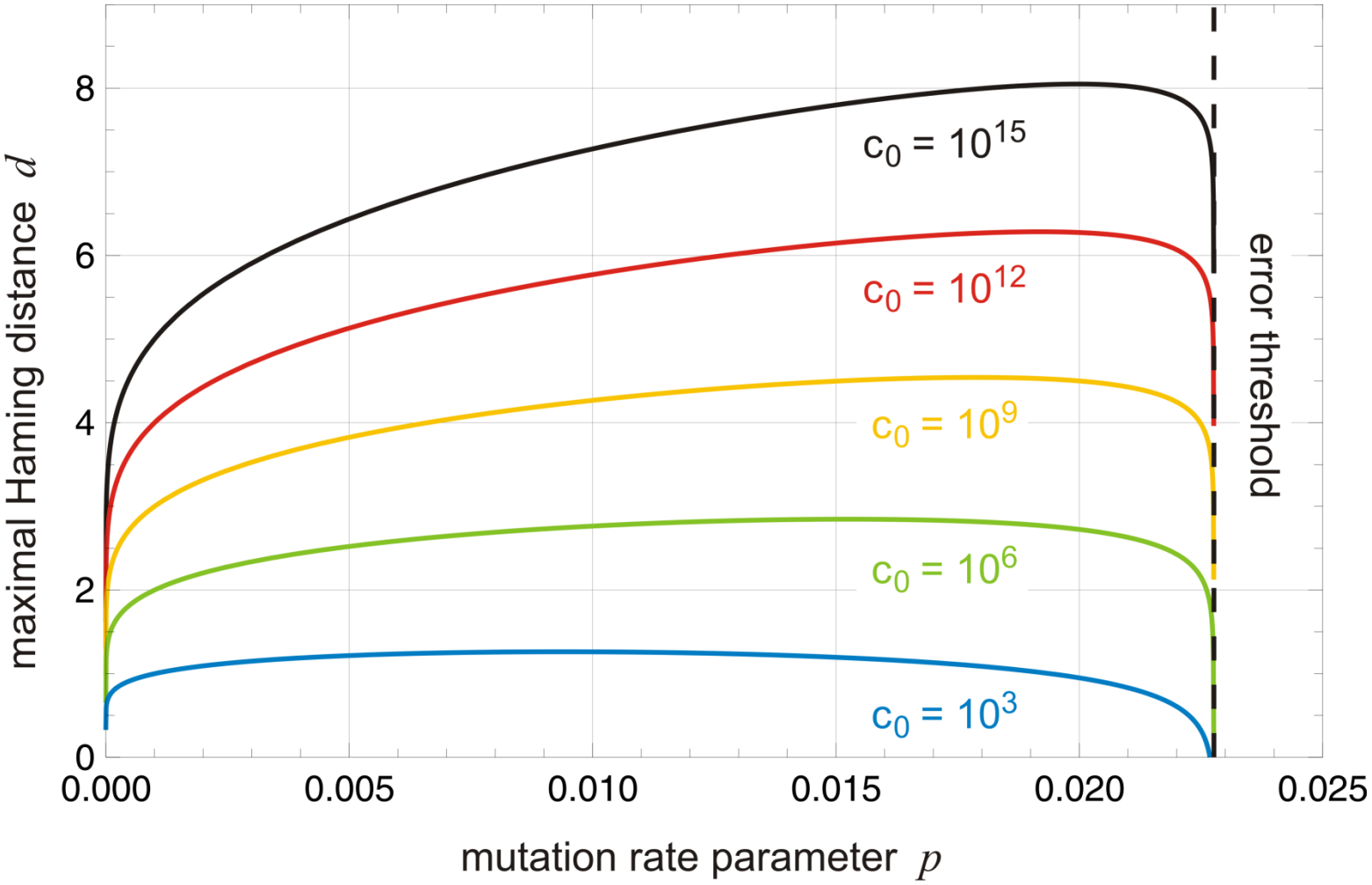
$$\hat{x}_j \geq k \implies d \leq \frac{1}{\log \varepsilon}\,\log \frac{k(1 - \sigma_m^{-1})^2}{(Q - \sigma_m^{-1})\widehat{c}_0}$$

discreteness condition

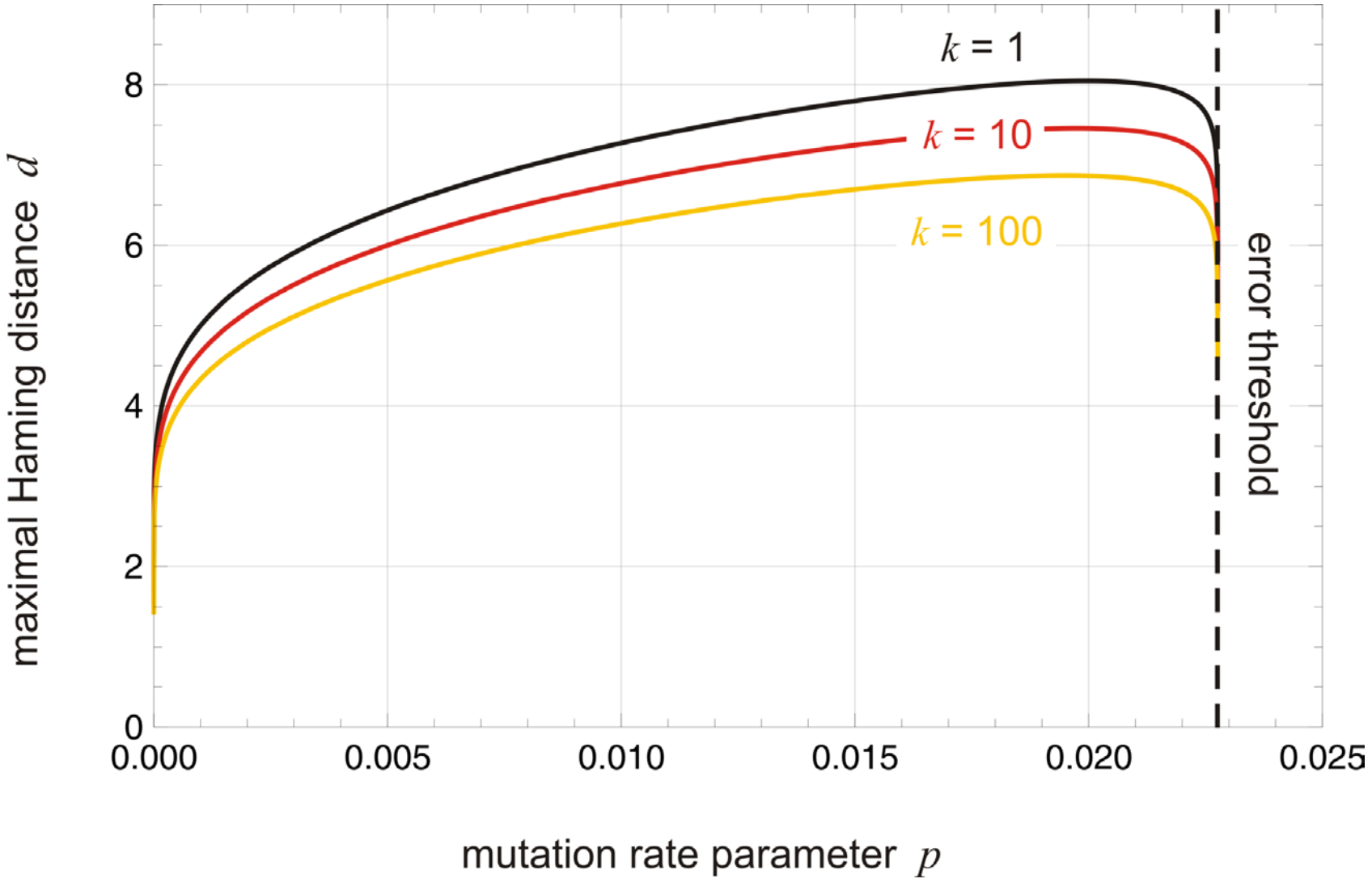
width of the discrete quasispecies: 2d



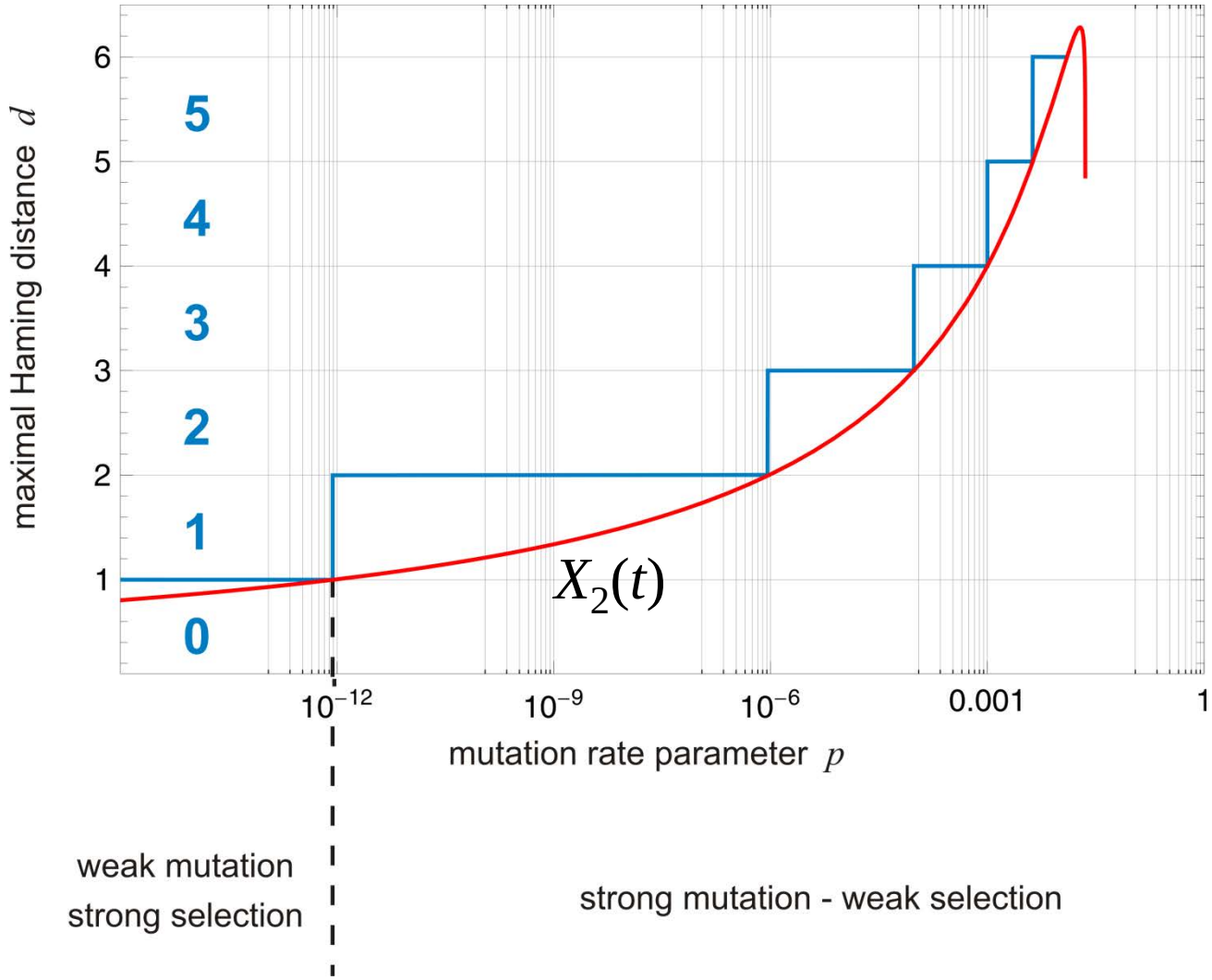
discrete quasispecies: $c_0 = 10^{12}$, $l = 100$, $k = 1$



discrete quasispecies: $l = 100$, $k = 1$



discrete quasispecies: $c_0 = 10^{12}$, $l = 100$



width of the discrete quasispecies

$$P_m(t) = \text{Prob}\Big(A(t) = m\Big) \quad \text{and} \quad P_{s_i}(t) = \text{Prob}\Big(X_i(t) = s_i\Big)$$

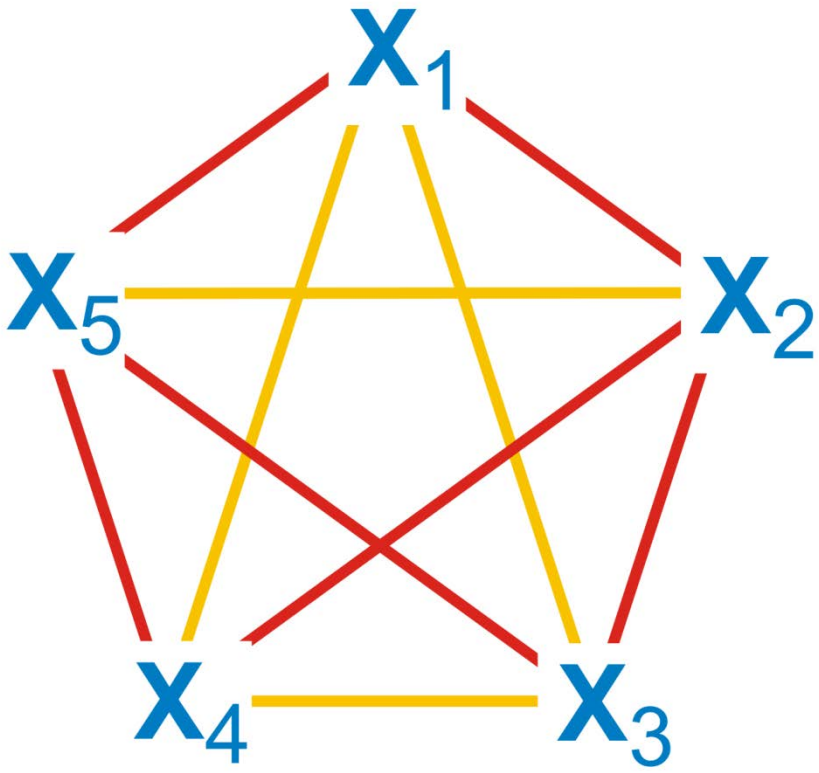
$$\mathbf{m} = (m, s_1, \ldots, s_n)$$

$$\begin{aligned}
\frac{\mathrm{d}P_{\mathbf{m}}}{\mathrm{dt}} &= a_0 r\Big(P_{(\mathbf{m};m-1)} - P_{\mathbf{m}}\Big) + r\Big((m+1)P_{(\mathbf{m};m+1)} - mP_{\mathbf{m}}\Big) + \\
&\quad + r\sum_{i=1}^{n}\Big((s_i+1)P_{(\mathbf{m};s_i+1)} - s_i P_{\mathbf{m}}\Big) + \\
&\quad + \sum_{i=1}^{n} Q_{ii}k_i\Big((m+1)(s_i-1)P_{(\mathbf{m};m+1,s_i-1)} - m\,s_i P_{\mathbf{m}}\Big) + \\
&\quad + \sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n} Q_{ij}k_j\,s_j\Big((m+1)P_{(\mathbf{m};m+1,s_i-1)} - mP_{\mathbf{m}}\Big)
\end{aligned}$$

$$\mathbf{m} = (m, s_1, \ldots, s_n)$$

$$\begin{aligned}
\mathbf{m}' &= (m \pm 1, s_1, \ldots, s_n) \equiv (\mathbf{m}; m \pm 1) \quad \text{or} \\
\mathbf{m}' &= (m, s_1, \ldots, s_k \pm 1, \ldots, s_n) \equiv (\mathbf{m}; s_k \pm 1)
\end{aligned}$$

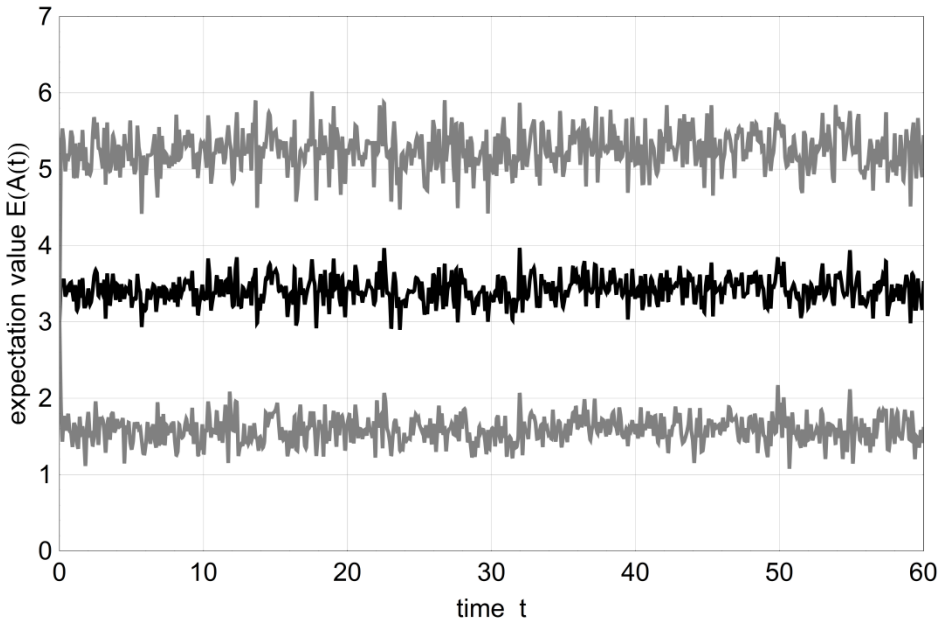
jumps: $S_{\mathbf{m}} \rightarrow S_{\mathbf{m}'}$ and $S_{\mathbf{m}'} \rightarrow S_{\mathbf{m}}$



mutation scheme *pentagram* for n = 5

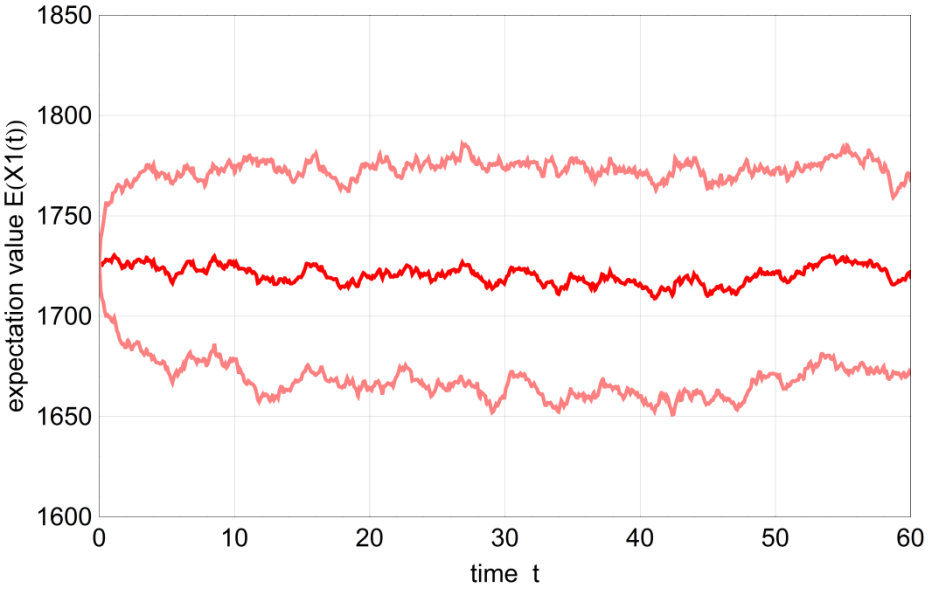
statistics of 100 trajectories:

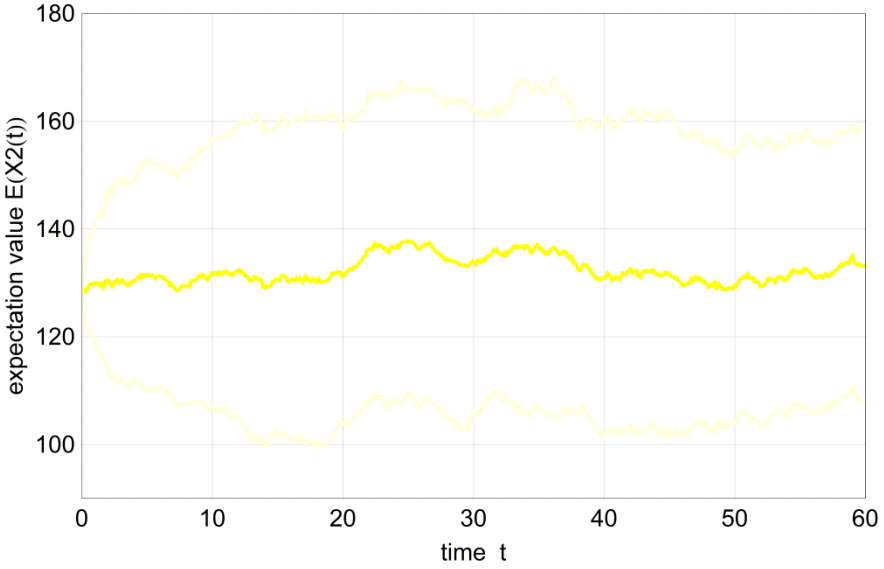
N = 2000, k = 5, p = 0.0124;
flow reactor: r = 0.5;
$k_1 = 0.150$, $k_2 = k_5 = 0.0125$,
$k_3 = k_4 = 0.100$

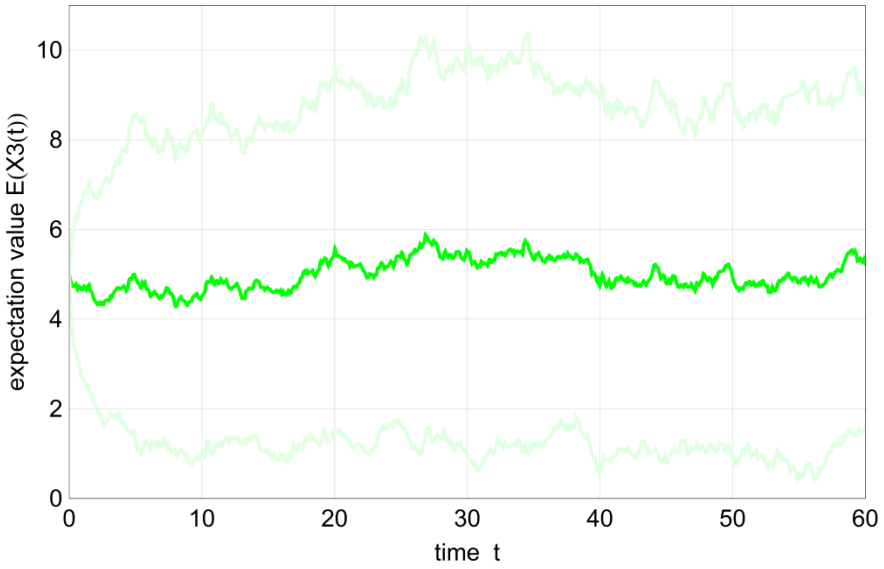


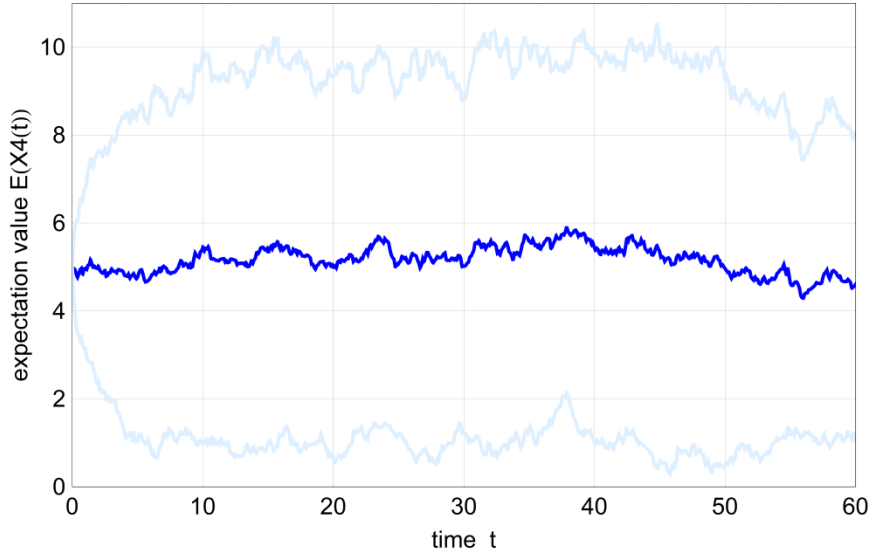
one standard deviation band:

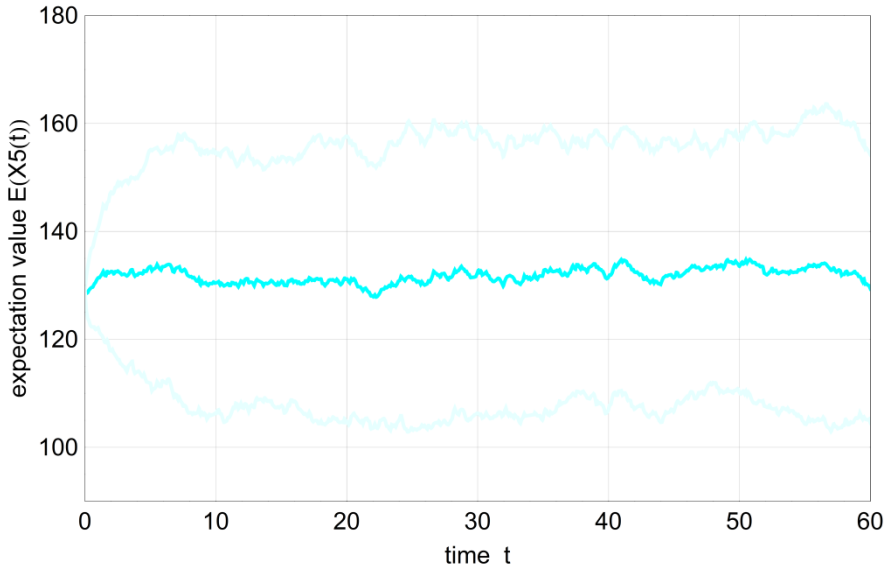
resource A, master sequence X_1







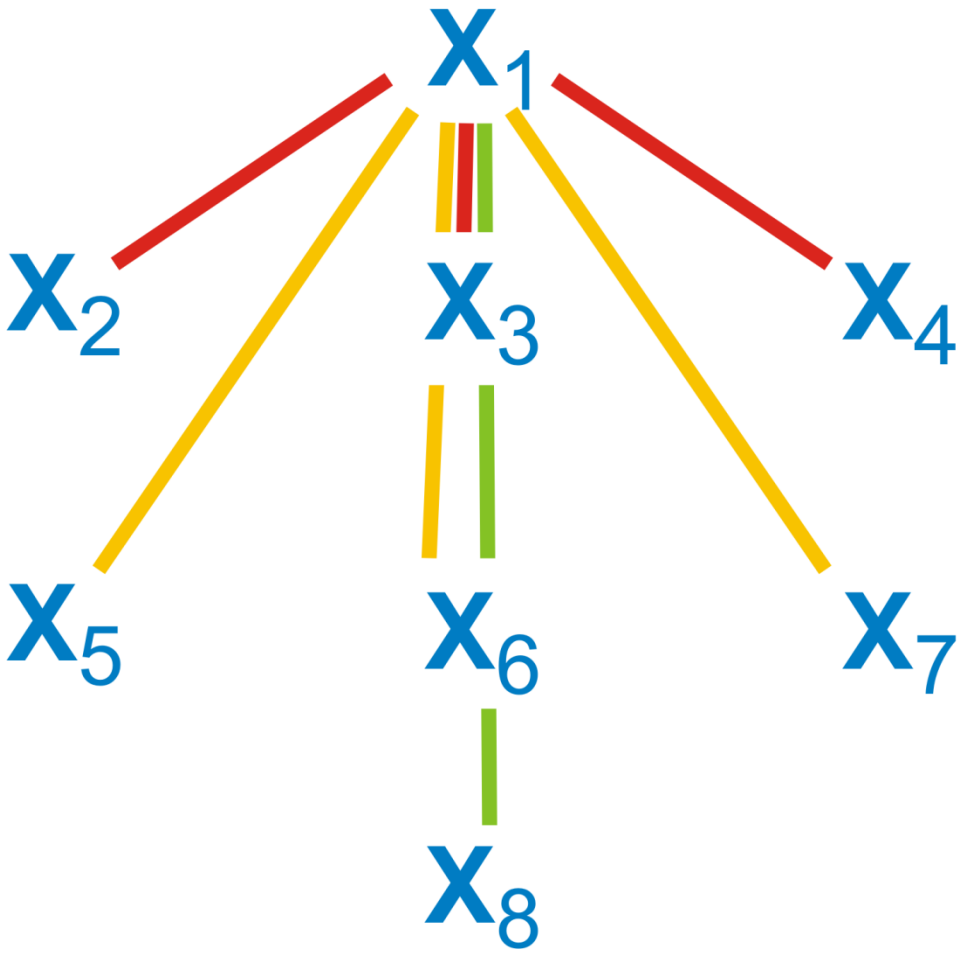




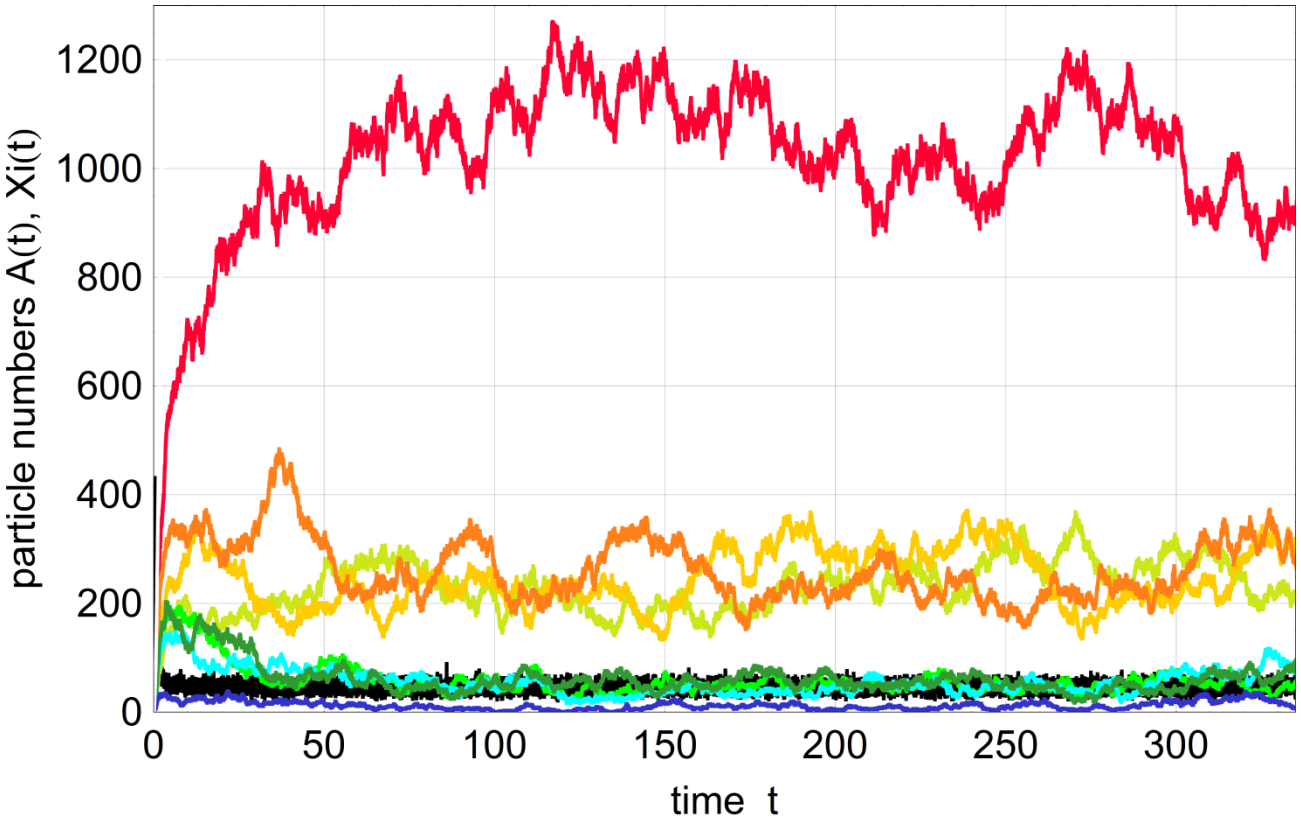
one standard deviation bands: mutants

	deterministic	expectation E	standard dev. σ	√E
$A(t)$	3.413	3.424	1.843	1.850
$X_1(t)$	1727.9	1720	54.68	41.47
$X_2(t)$	129.31	133.2	27.69	11.54
$X_3(t)$	5.0298	5.132	3.945	2.265
$X_4(t)$	5.0298	5.200	4.215	2.280
$X_5(t)$	129.31	132.1	25.59	11.49

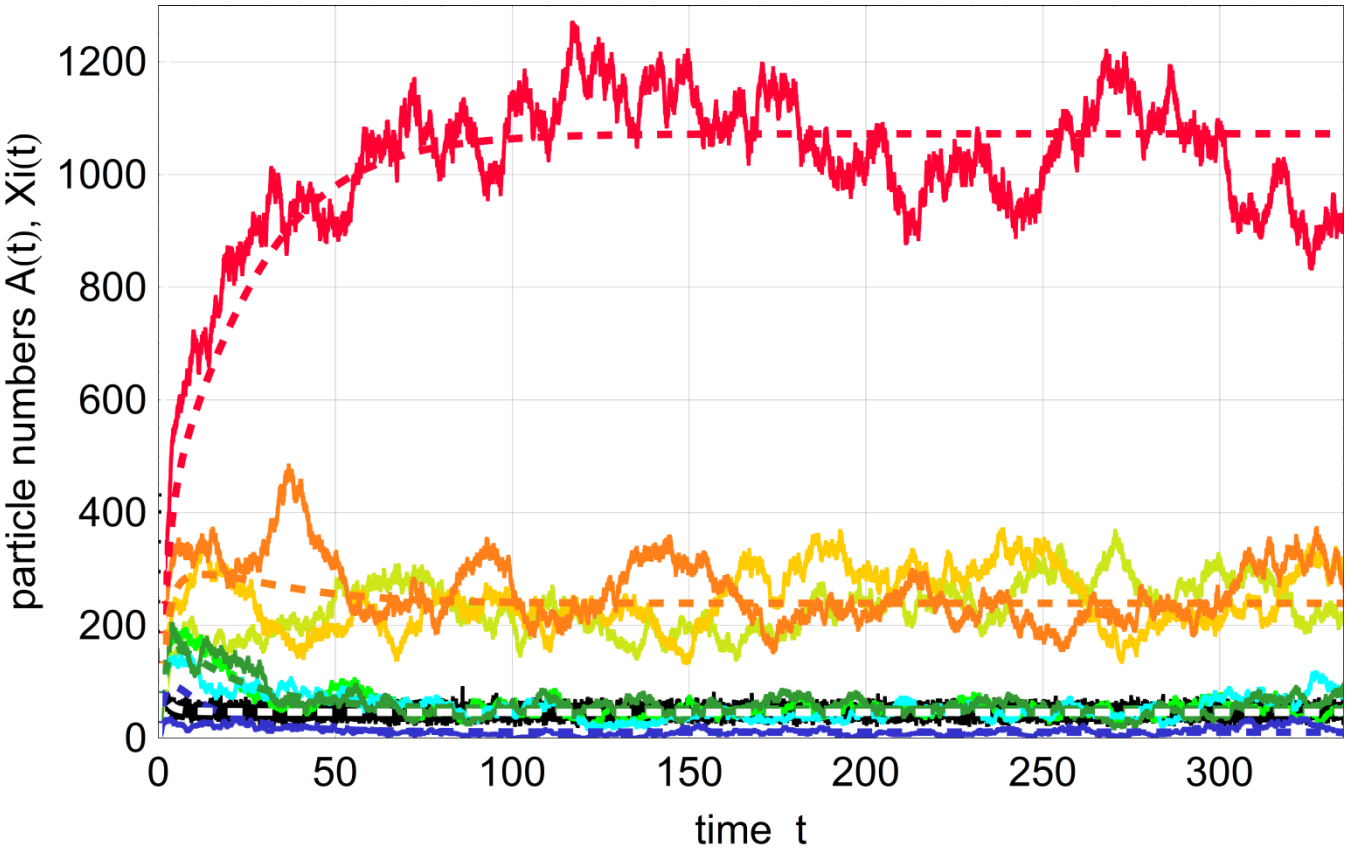
N = 2000, k = 5, p = 0.0124; flow reactor: r = 0.5; $k_1 = 0.150$, $k_2 = k_5 = 0.125$, $k_3 = k_4 = 0.100$



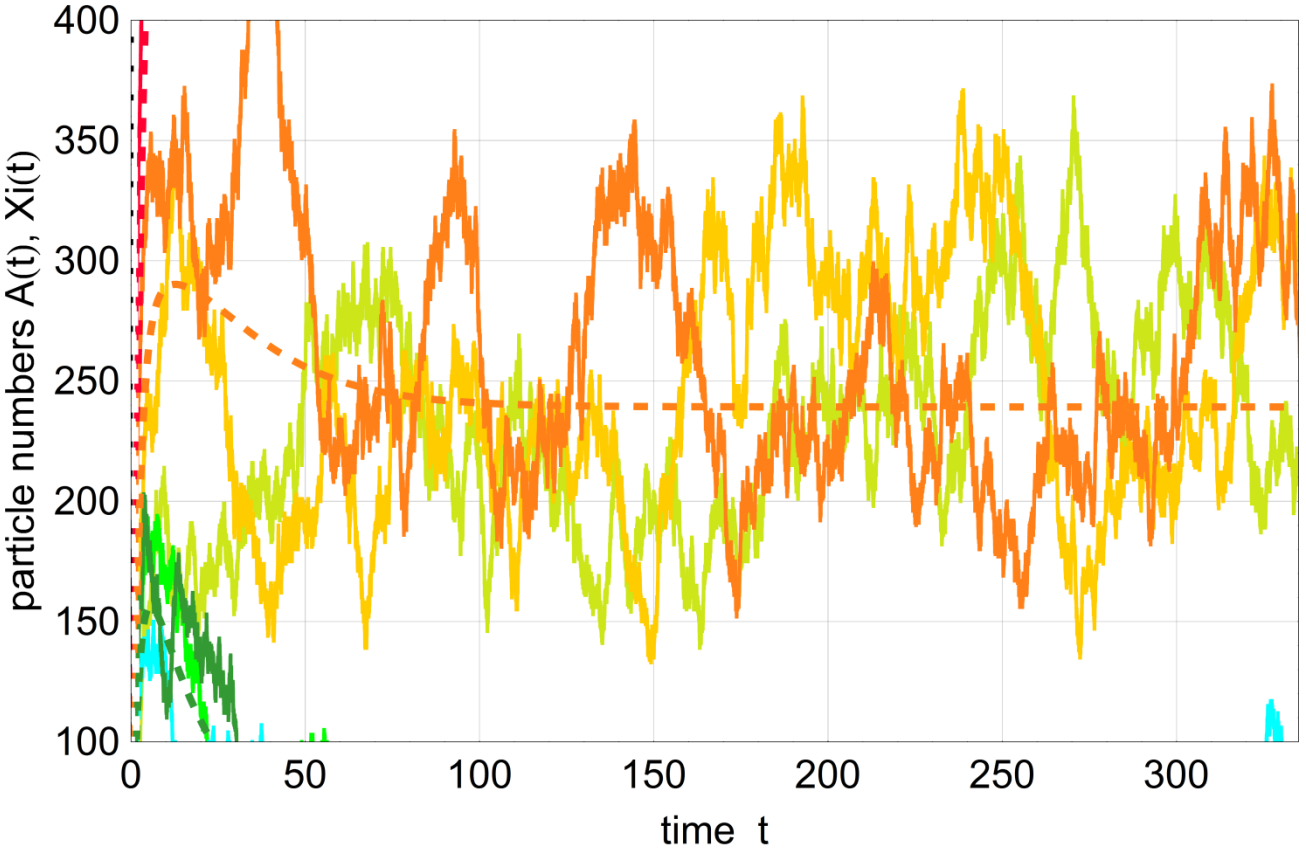
mutation scheme *sequence space* for $l = 3$ ($n = 8$)



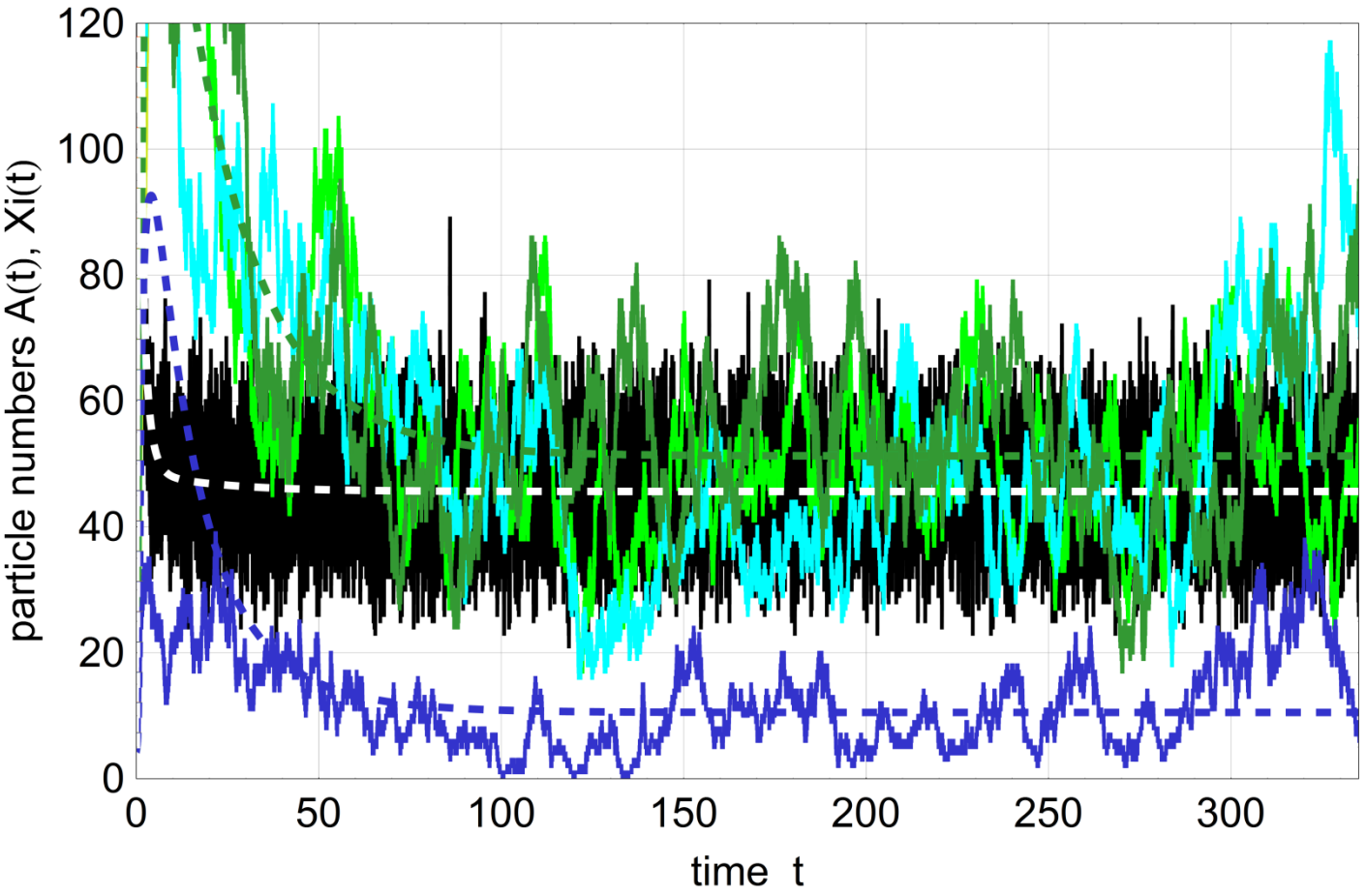
Gillespie simulation of stochastic quasispecies $l = 3$



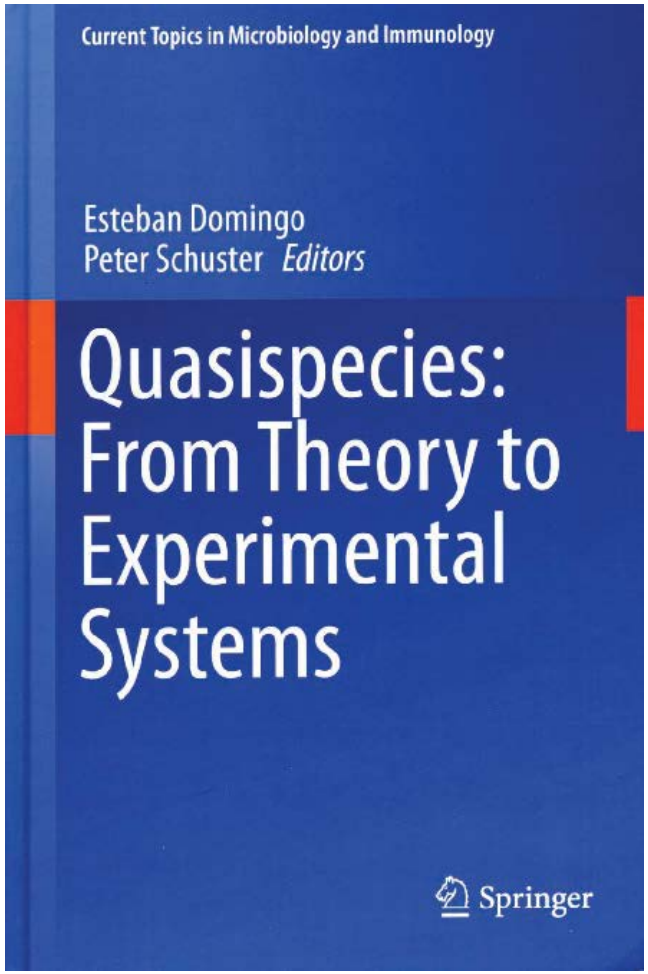
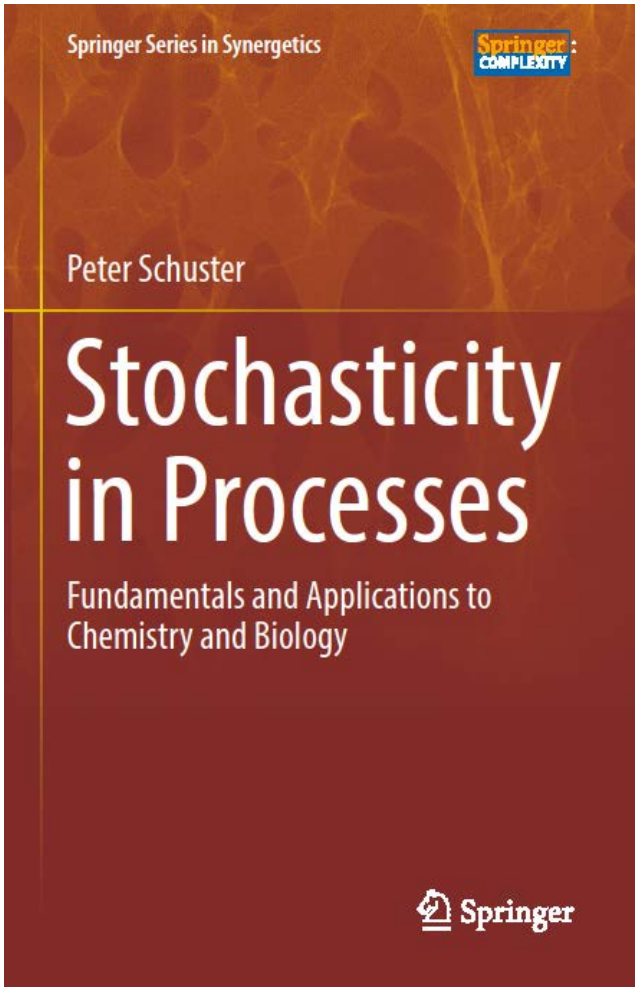
comparison with deterministic solution



comparison with deterministic solution



comparison with deterministic solution

Thank you for your attention!

Web-Page for further information:
http://www.tbi.univie.ac.at/~pks