

Does it pay to be consistent?

Peter Schuster

Institut für Theoretische Chemie, Universität Wien, Austria

and

The Santa Fe Institute, Santa Fe, New Mexico, USA



TBI Seminar

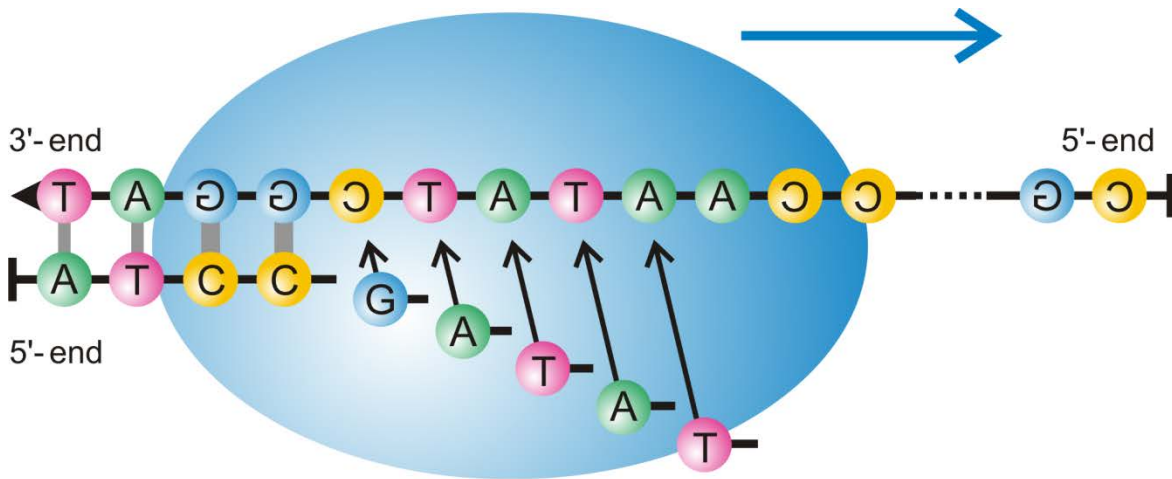
Wien, 22.04.2015

Web-Page for further information:

<http://www.tbi.univie.ac.at/~pks>

1. Quasispecies and Crow-Kimura model
2. Mutation flow analysis
3. Zero backflow and phenomenological approach
4. Error thresholds on realistic landscapes

1. Quasispecies and Crow-Kimura model
2. Mutation flow analysis
3. Zero backflow and phenomenological approach
4. Error thresholds on realistic landscapes

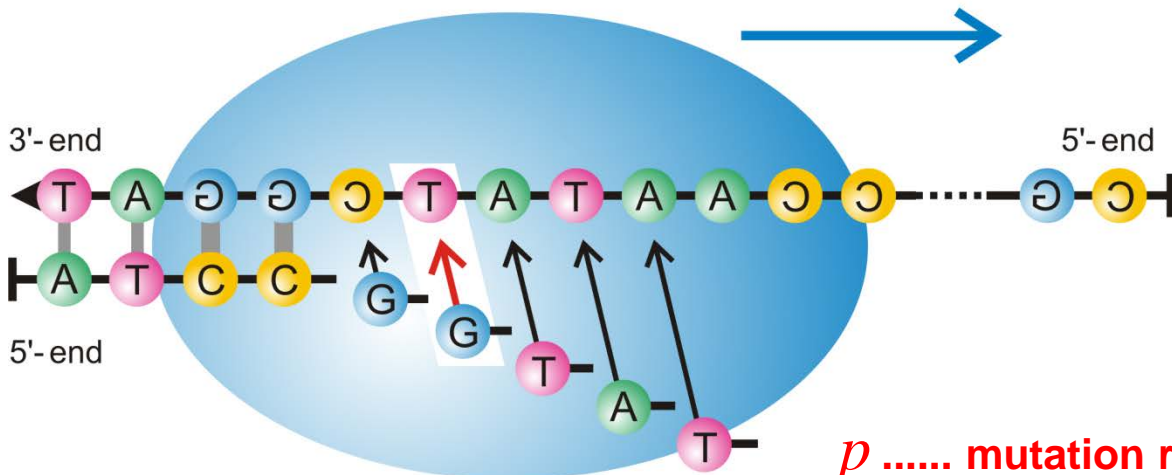


correct replication


Taq-polymerase

adenine 

thymine 



mutation

p mutation rate per site
and replication

guanine 

cytosine 

DNA replication and mutation



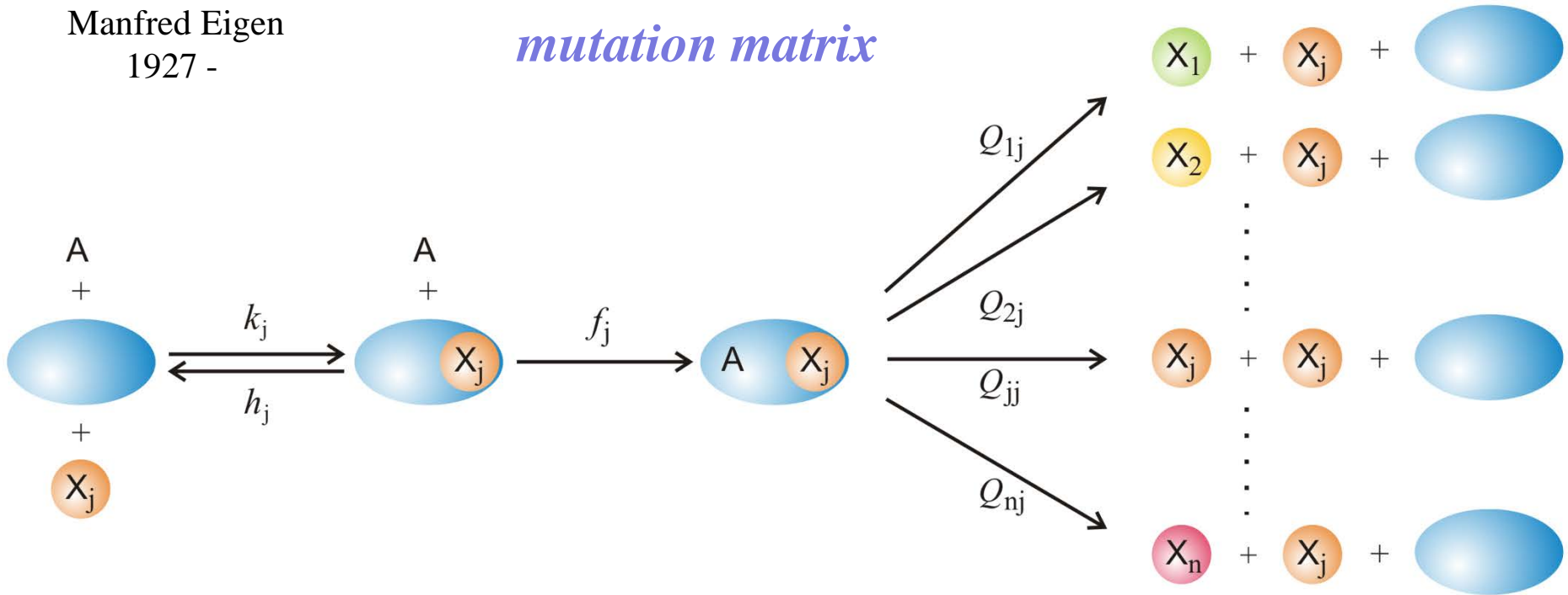
Manfred Eigen
1927 -

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

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

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



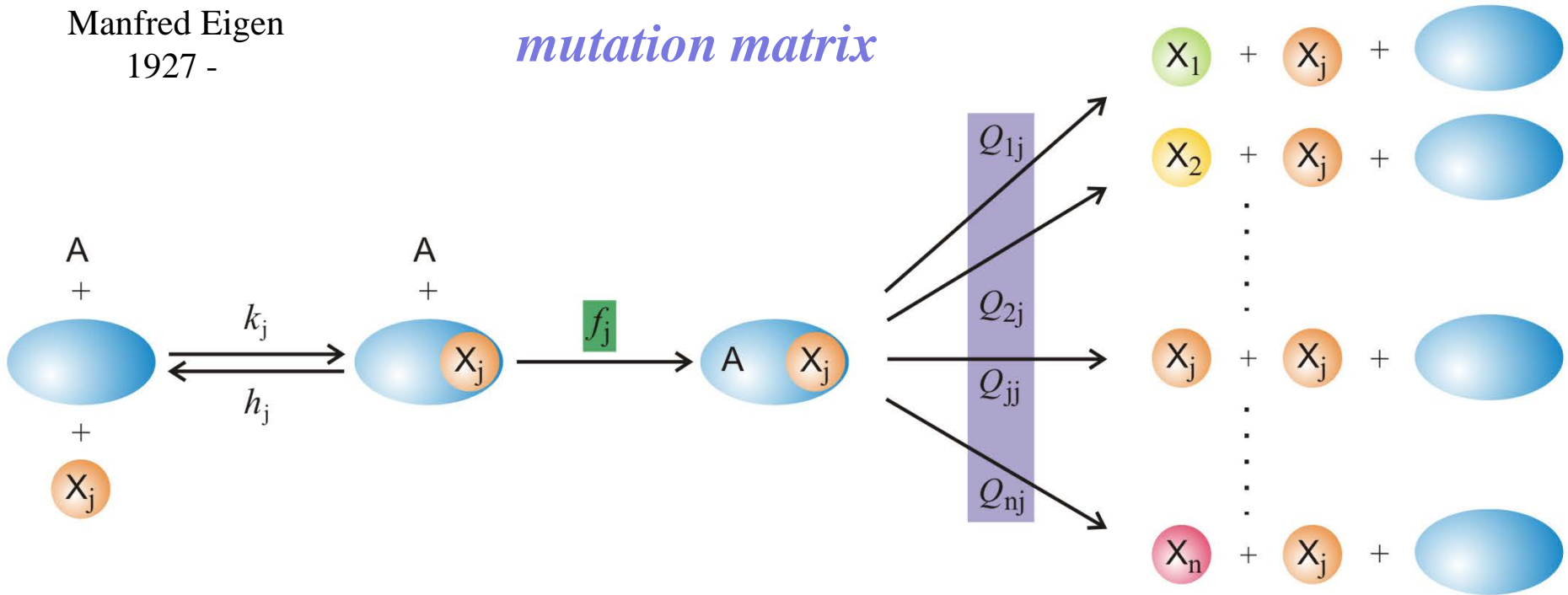
Manfred Eigen
1927 -

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

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

fitness landscape

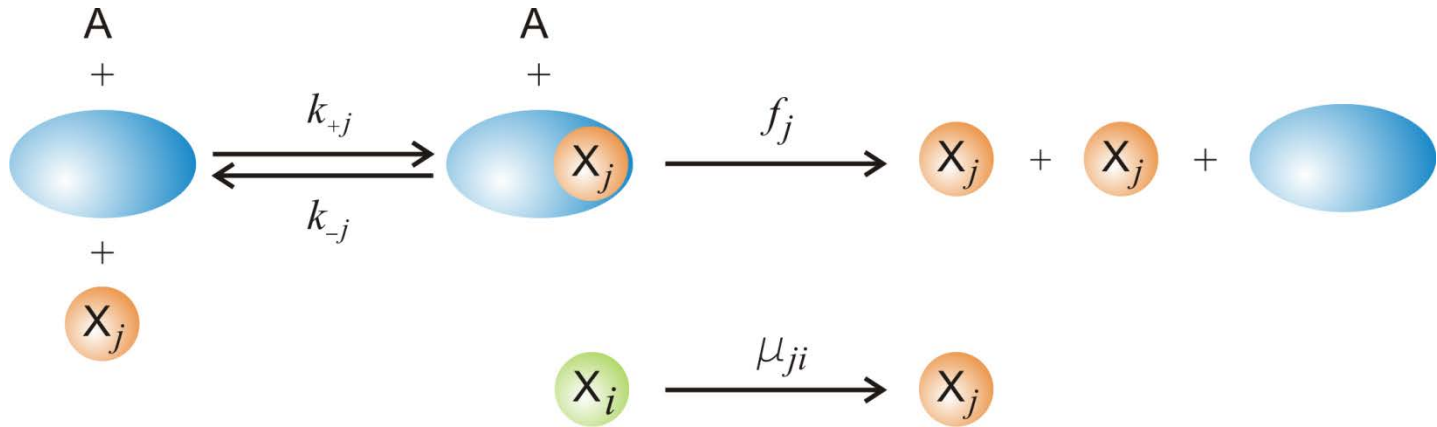
mutation matrix



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



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

paramuse – parallel 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

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

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

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

$$W = Q \cdot F$$

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

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

$$W = \mu + F$$

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

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

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

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

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

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

Solution:

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

Solution of the quasispecies equation

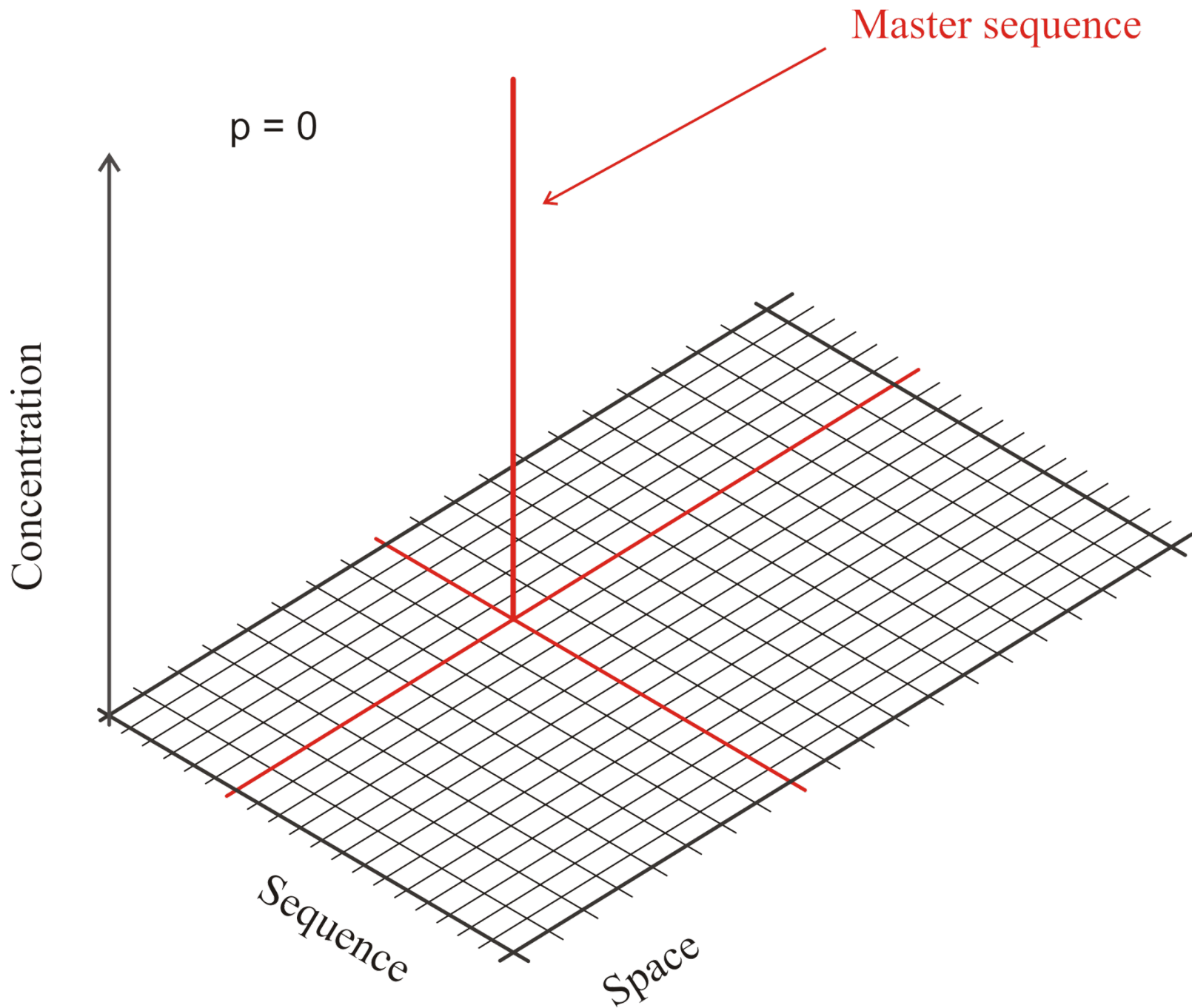
Largest eigenvalue λ_1 and corresponding eigenvector \mathbf{b}_1 :

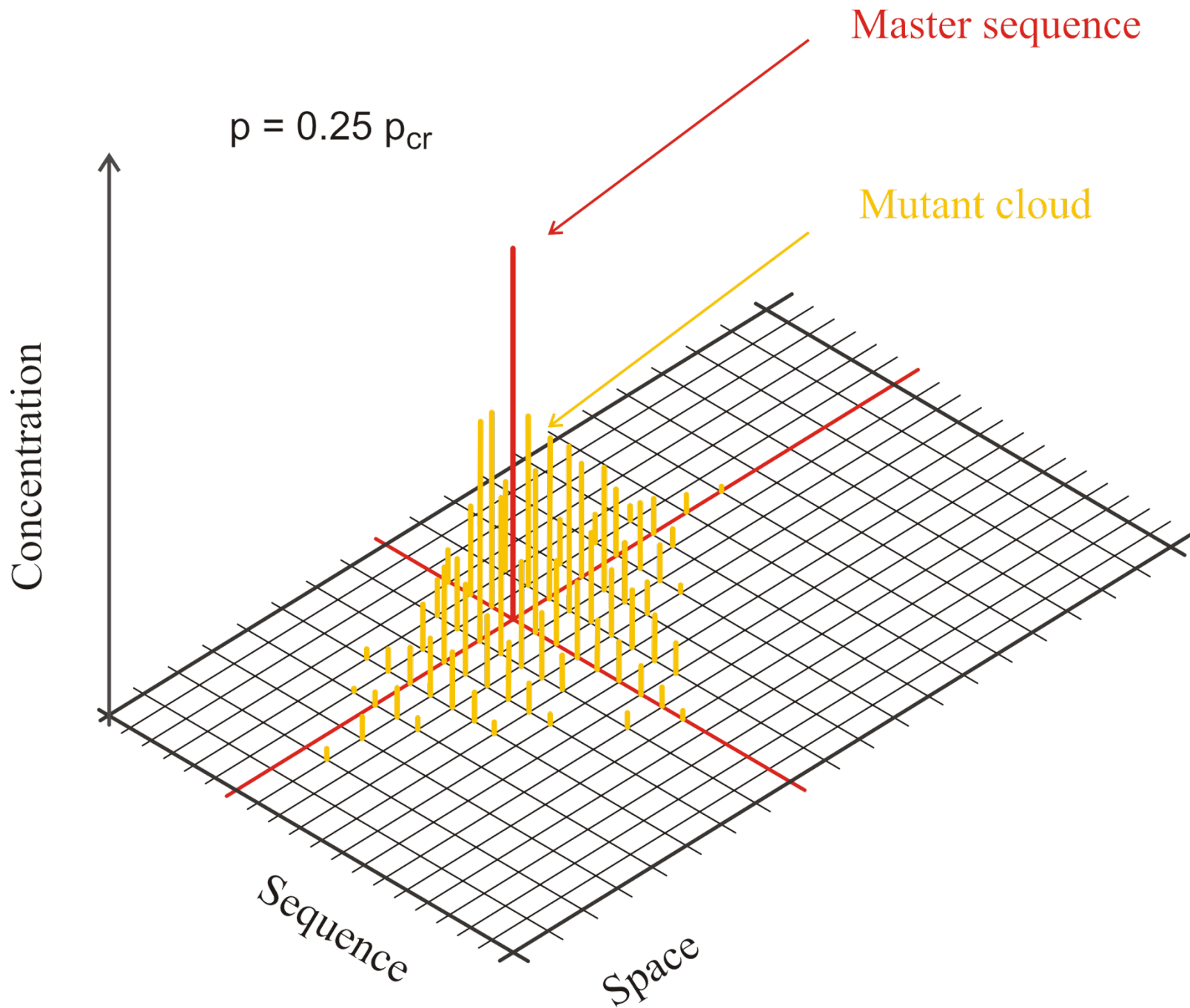
$$x_j(t) \approx \frac{b_{j1} \beta_1(0) \exp(\lambda_1 t)}{\sum_{i=1}^N b_{i1} \beta_1(0) \exp(\lambda_1 t)} = \frac{b_{j1}}{\sum_{i=1}^N b_{i1}} = \bar{x}_j \text{ for large } t$$

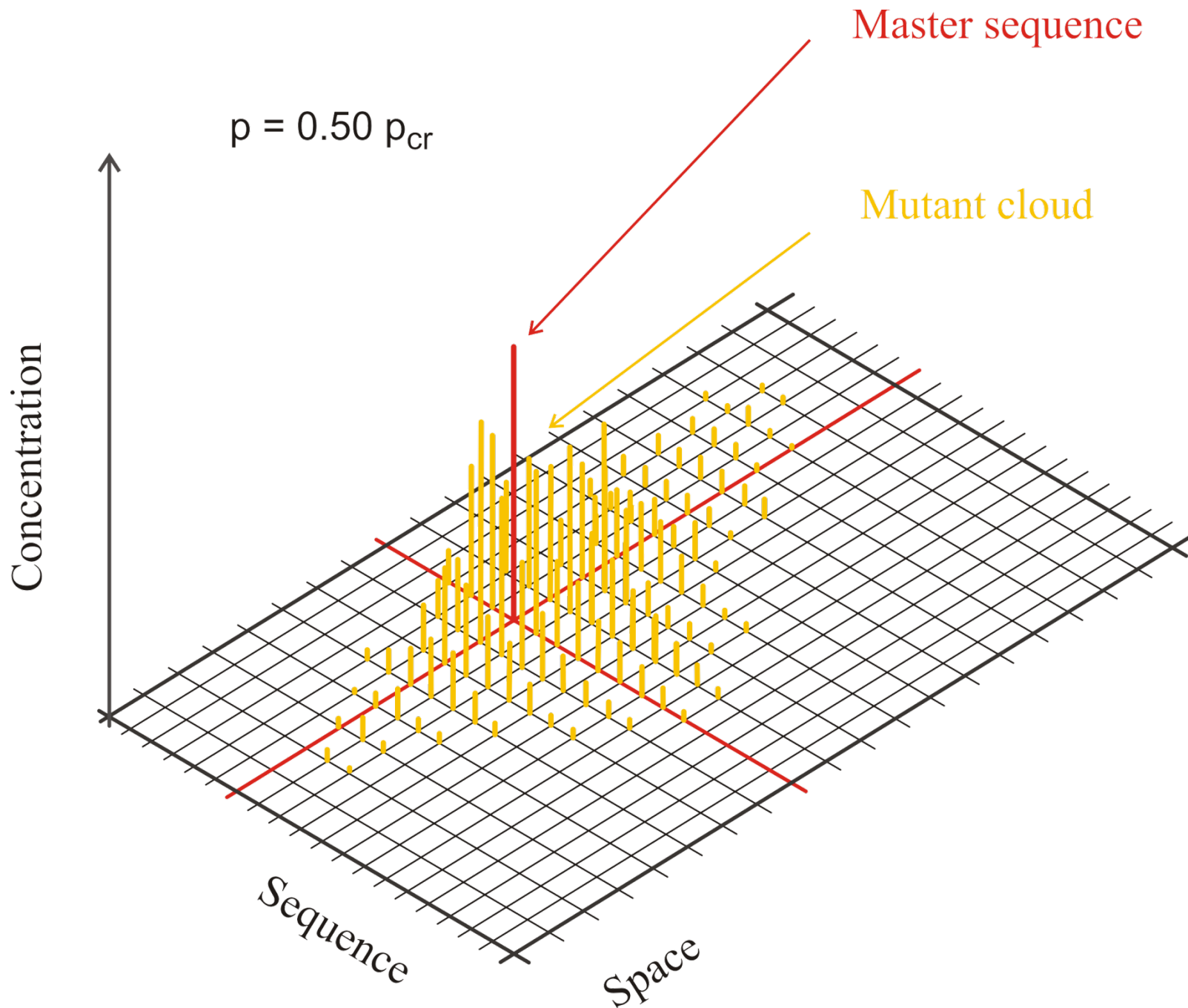
master sequence: \mathbf{X}_m at concentration \bar{x}_m

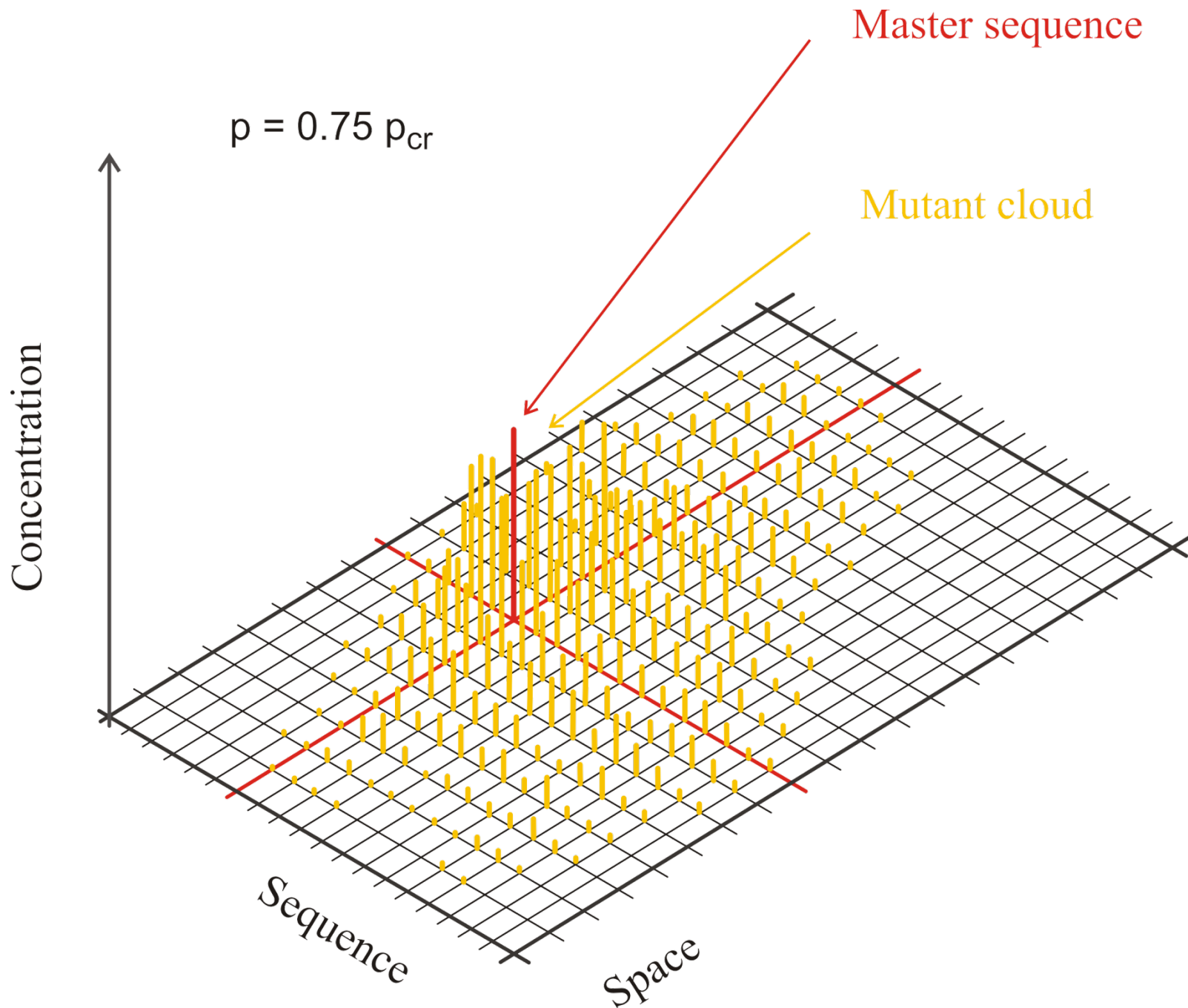
mutant cloud: \mathbf{X}_j at concentration \bar{x}_j ; $j=1, \dots, N$; $j \neq m$

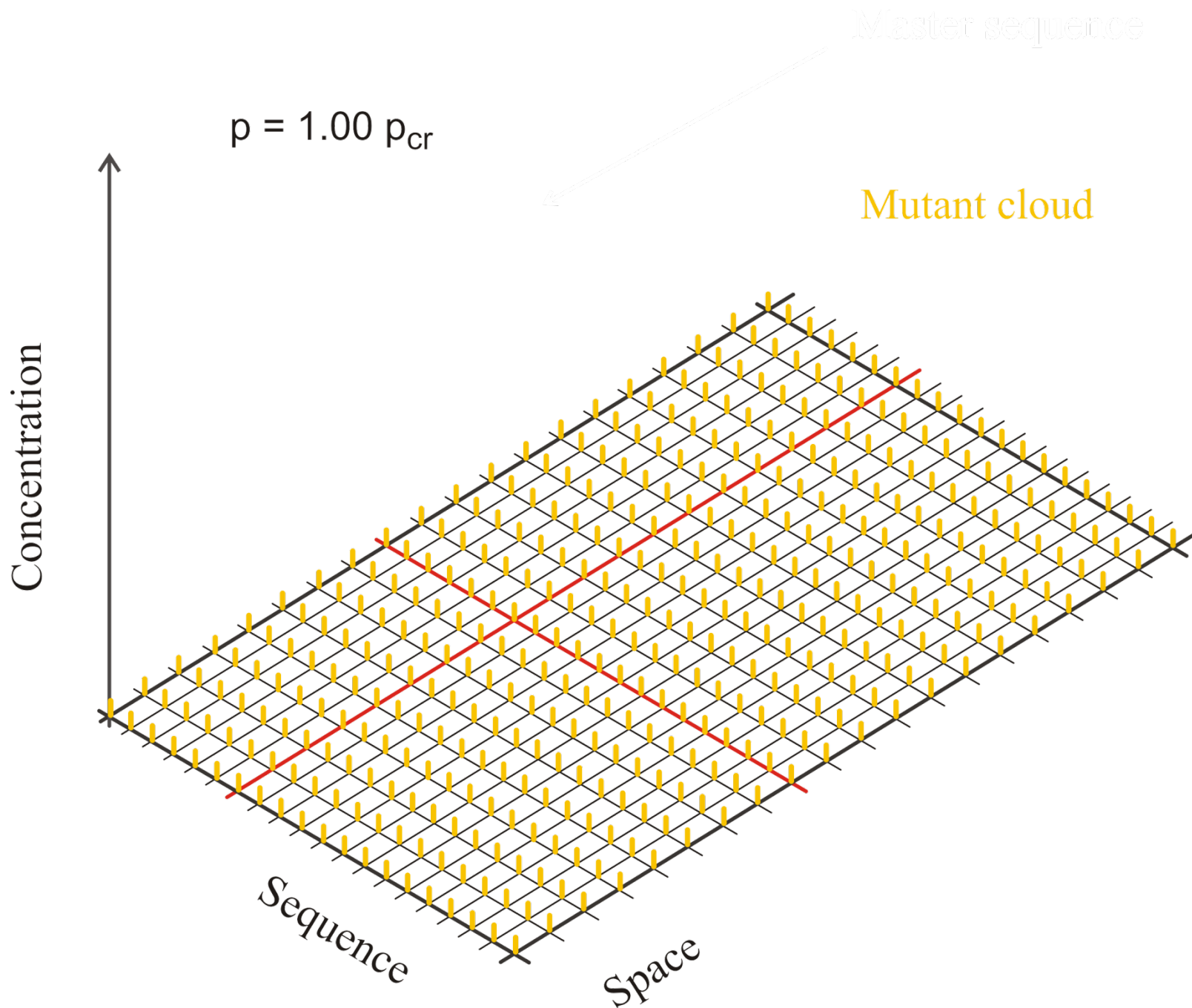
Stationary solution of the quasispecies equation

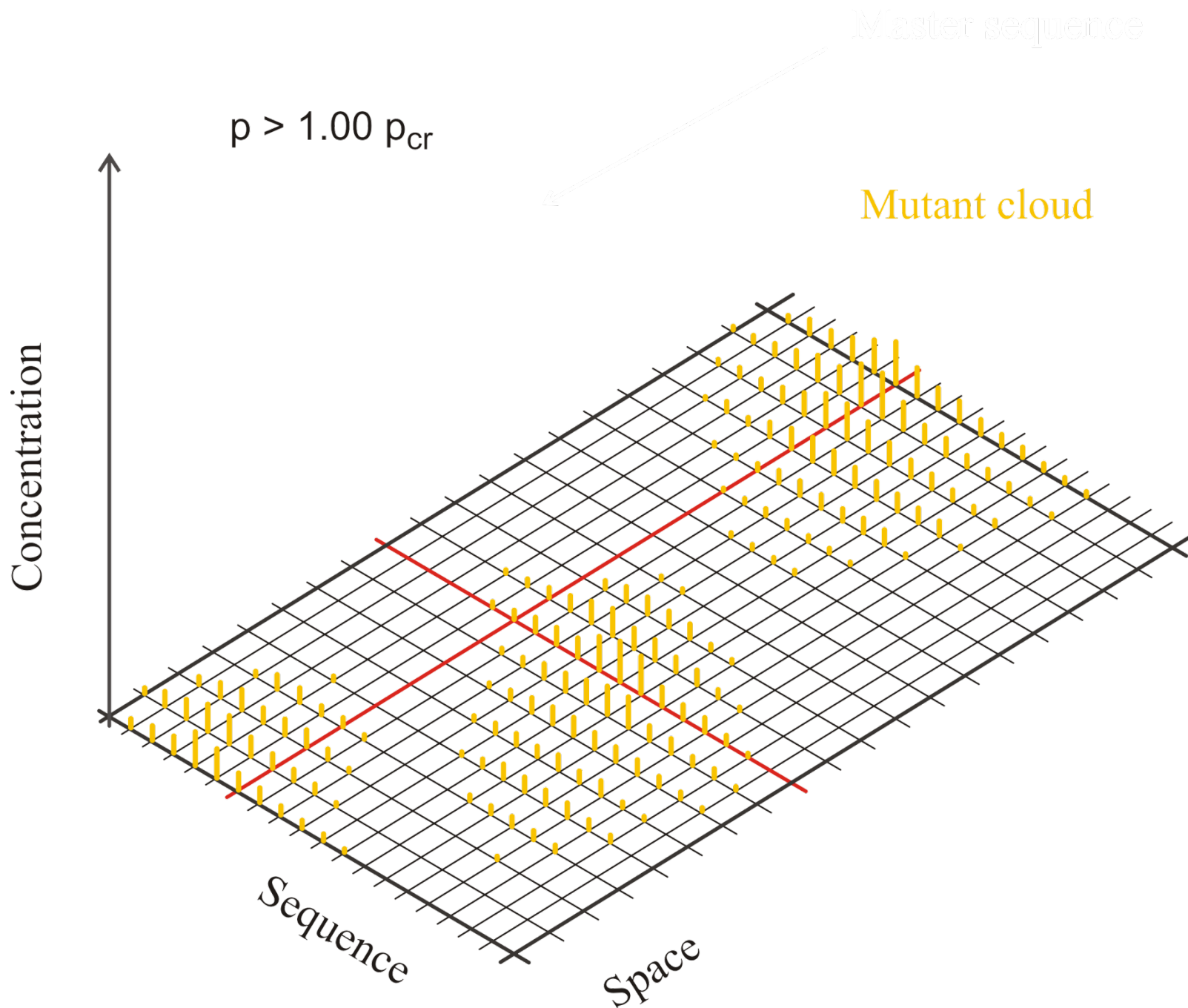


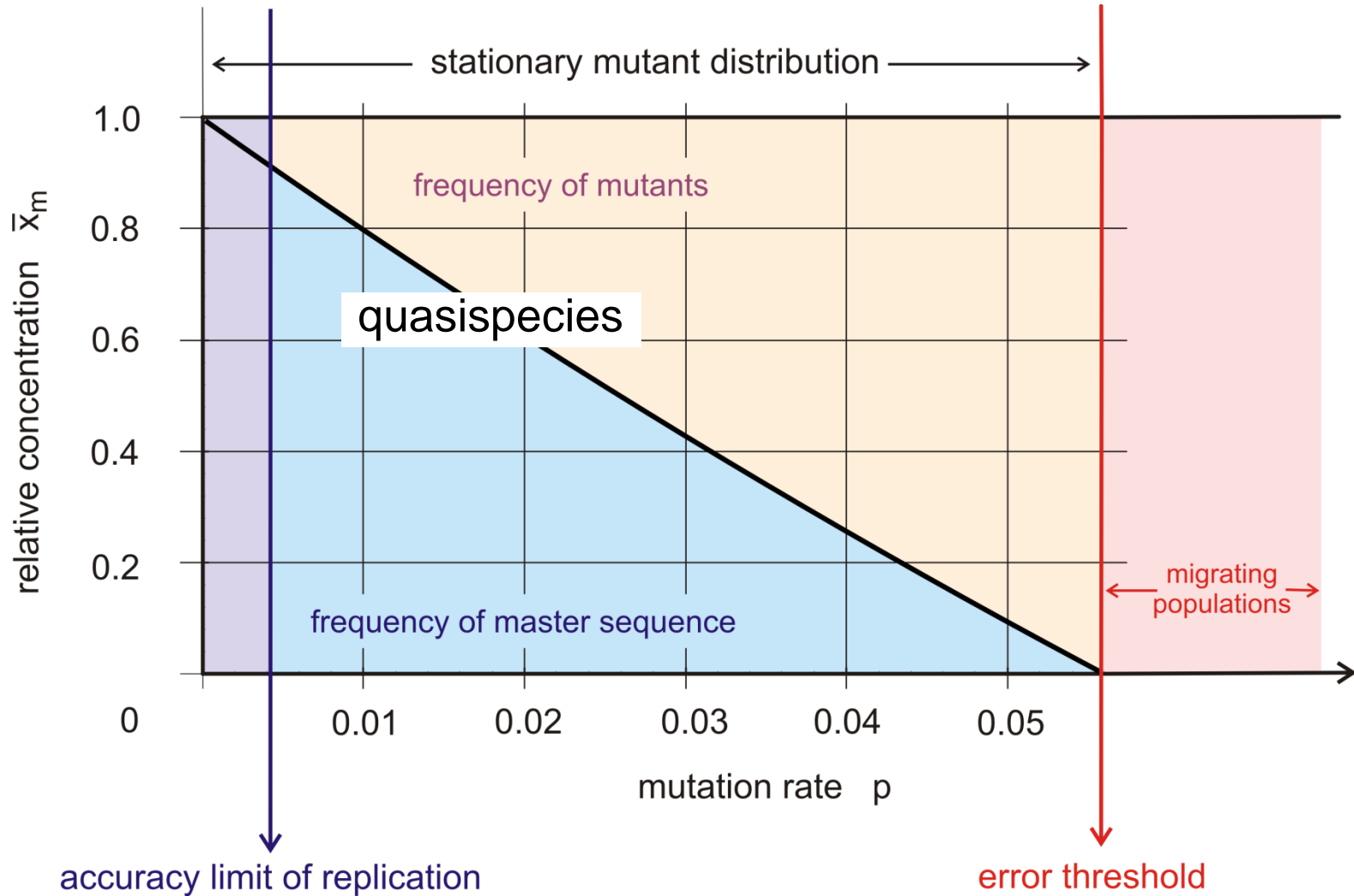




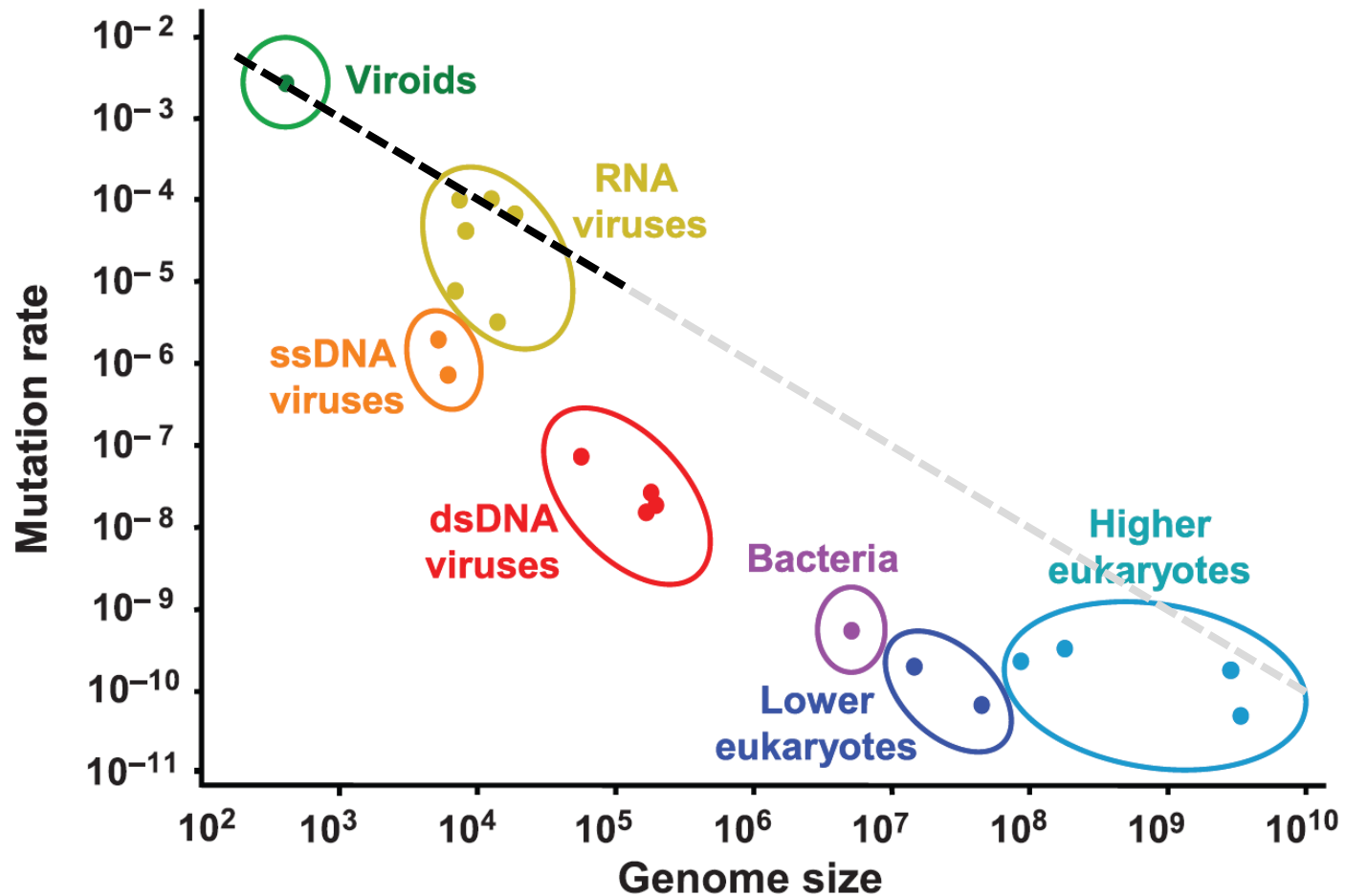








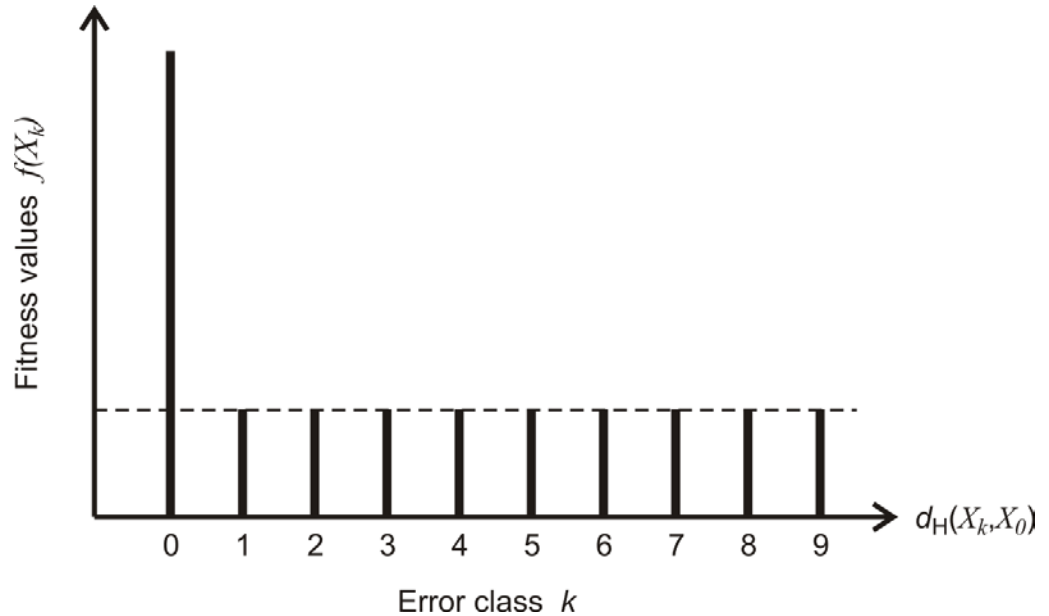
The error threshold in replication and mutation



Selma Gago, Santiago F. Elena, Ricardo Flores, Rafael Sanjuán. 2009, Extremely high mutation rate of a hammerhead viroid. *Science* 323:1308.

Mutation rate and genome size

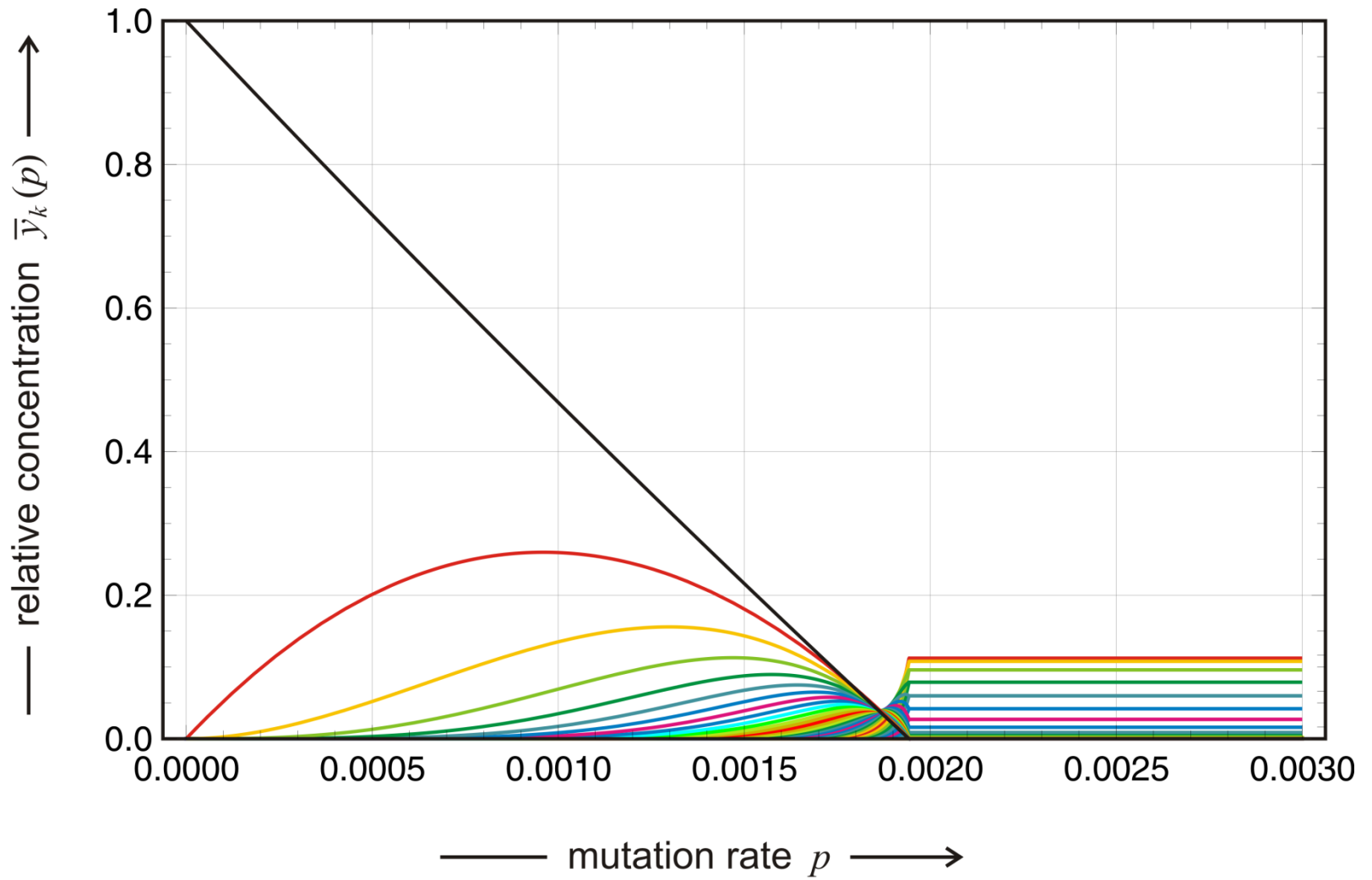
single peak fitness landscape



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

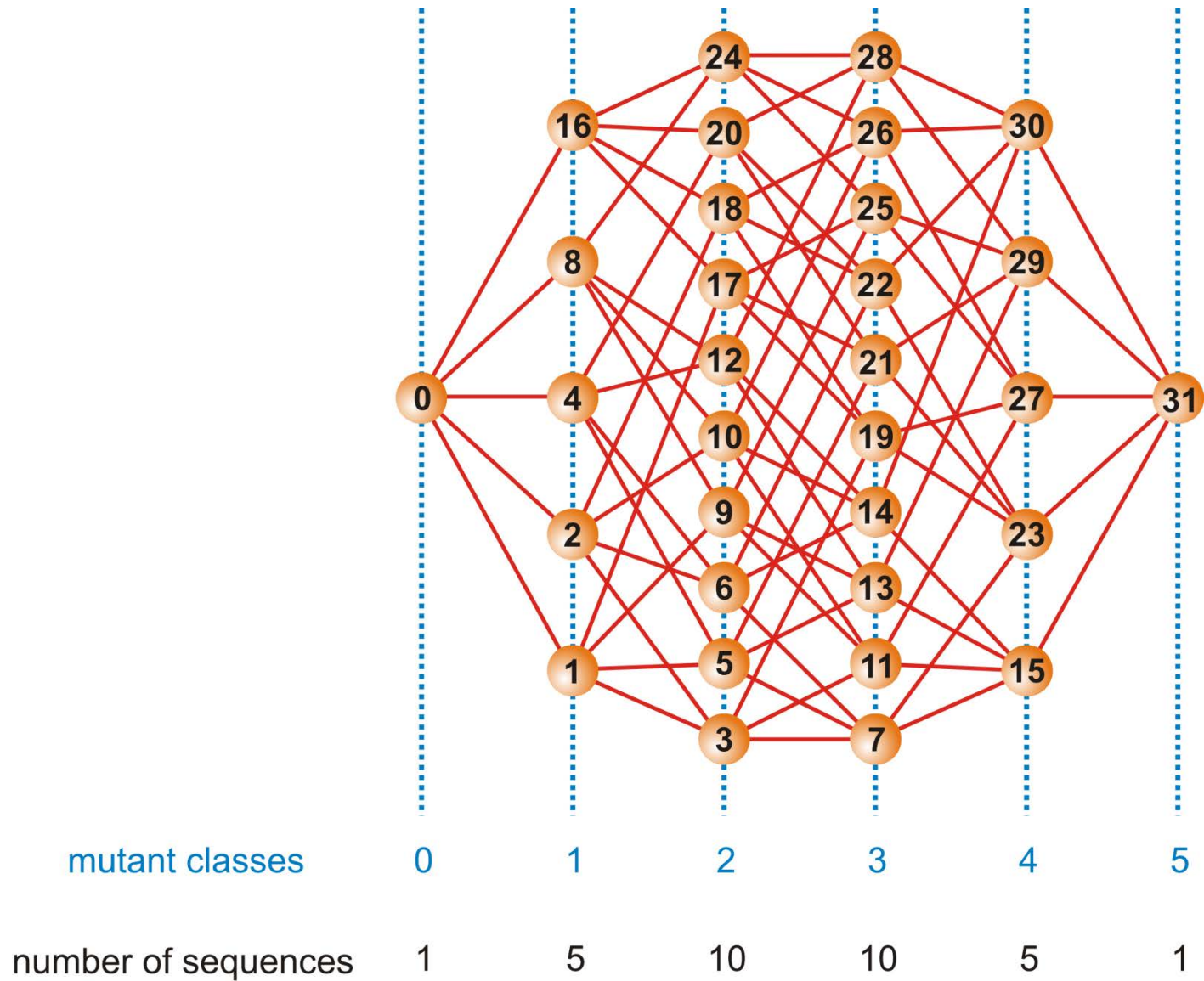
uniform error rate model

Approximations for handling realistic chain lengths

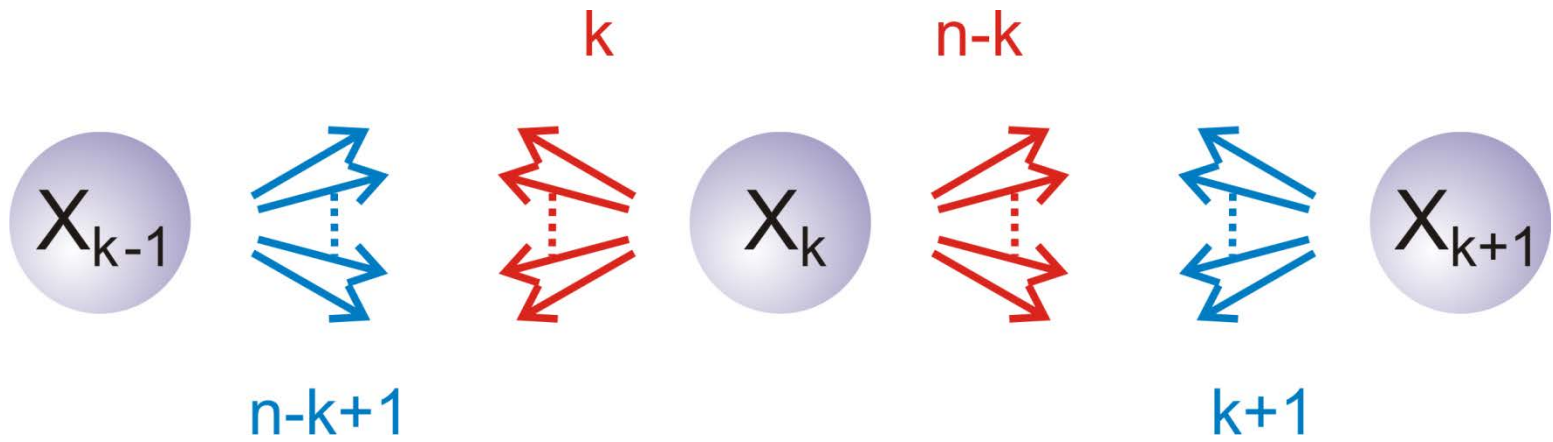


Quasispecies and error threshold

1. Quasispecies and Crow-Kimura model
2. **Mutation flow analysis**
3. Zero backflow and phenomenological approach
4. Error thresholds on realistic landscapes



The space of binary sequences



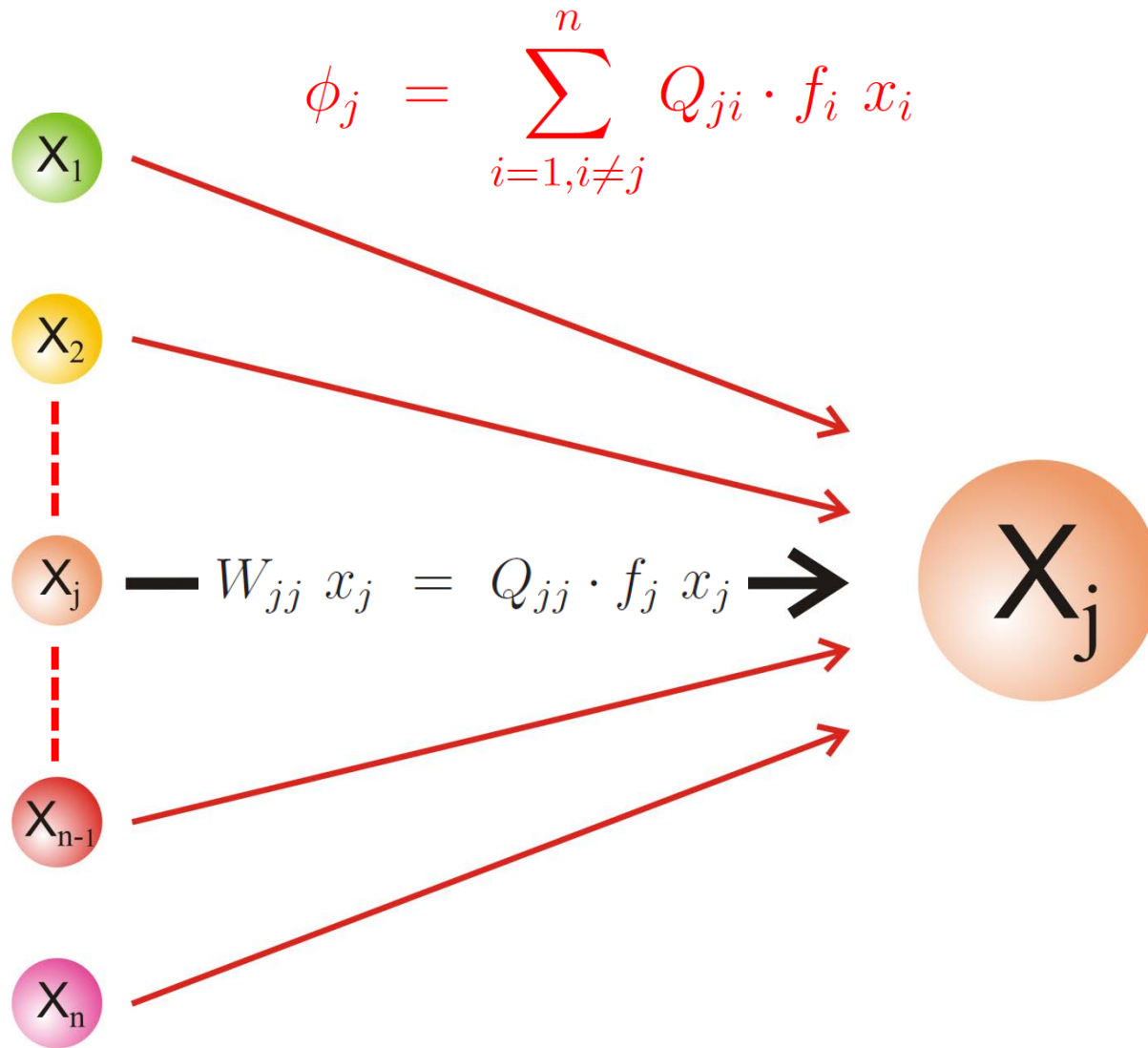
Neighbor distribution on binary sequence spaces

$$\varphi_{ji} = Q_{ji} \cdot f_i x_i; \quad Q_{ji} \approx (1 - p)^{l - d_{ji}^H} \cdot p^{d_{ji}^H}$$

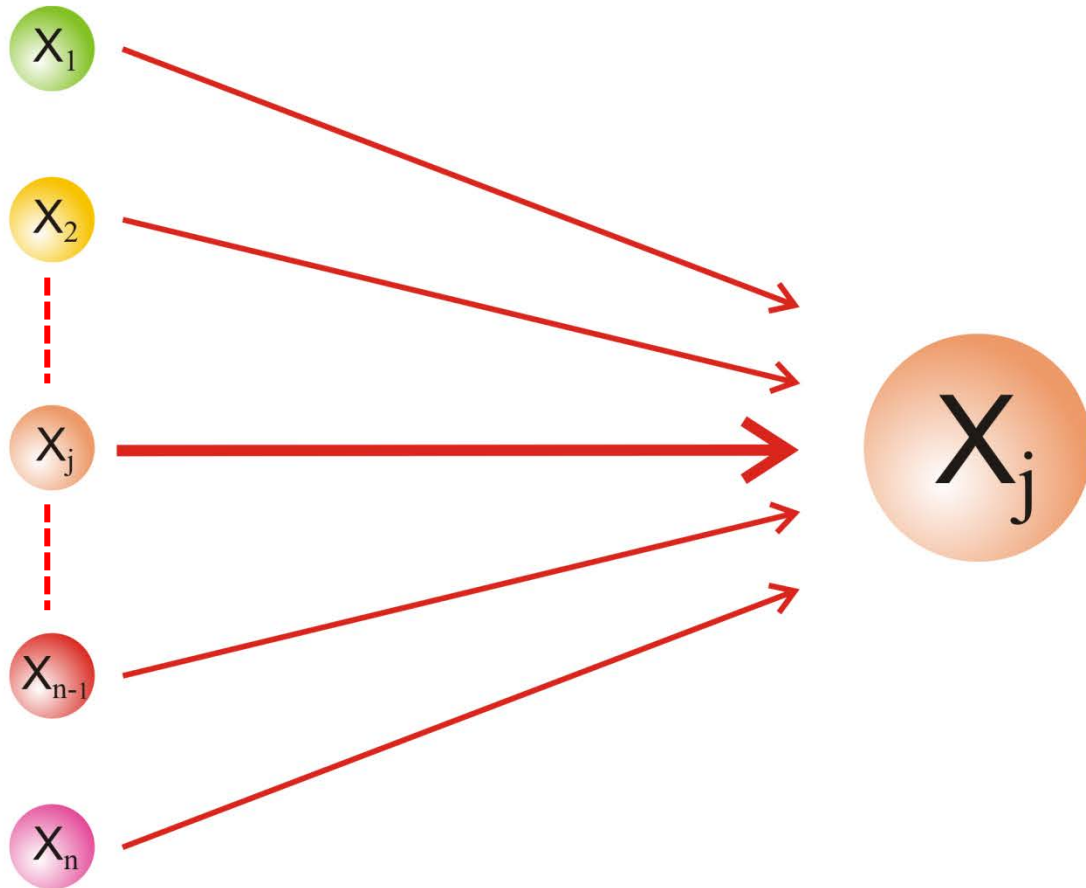


$$\phi_j = \sum_{i=1, i \neq j}^n \varphi_{ji} = \sum_{i=1, i \neq j}^n Q_{ji} \cdot f_i x_i$$

Mutation flow component and mutation flow



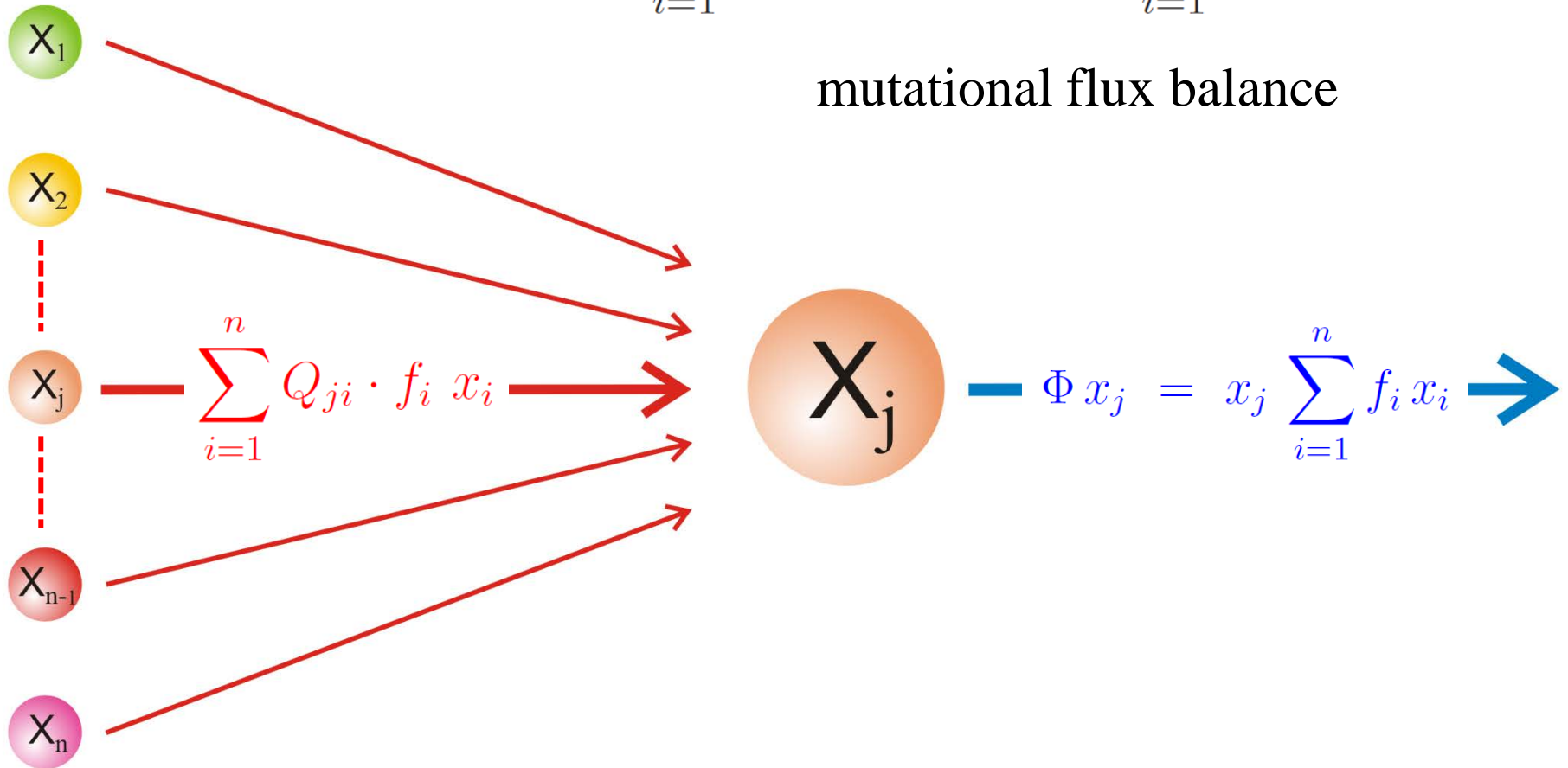
Definition of the mutation flow



Mutational flux balance and quasispecies

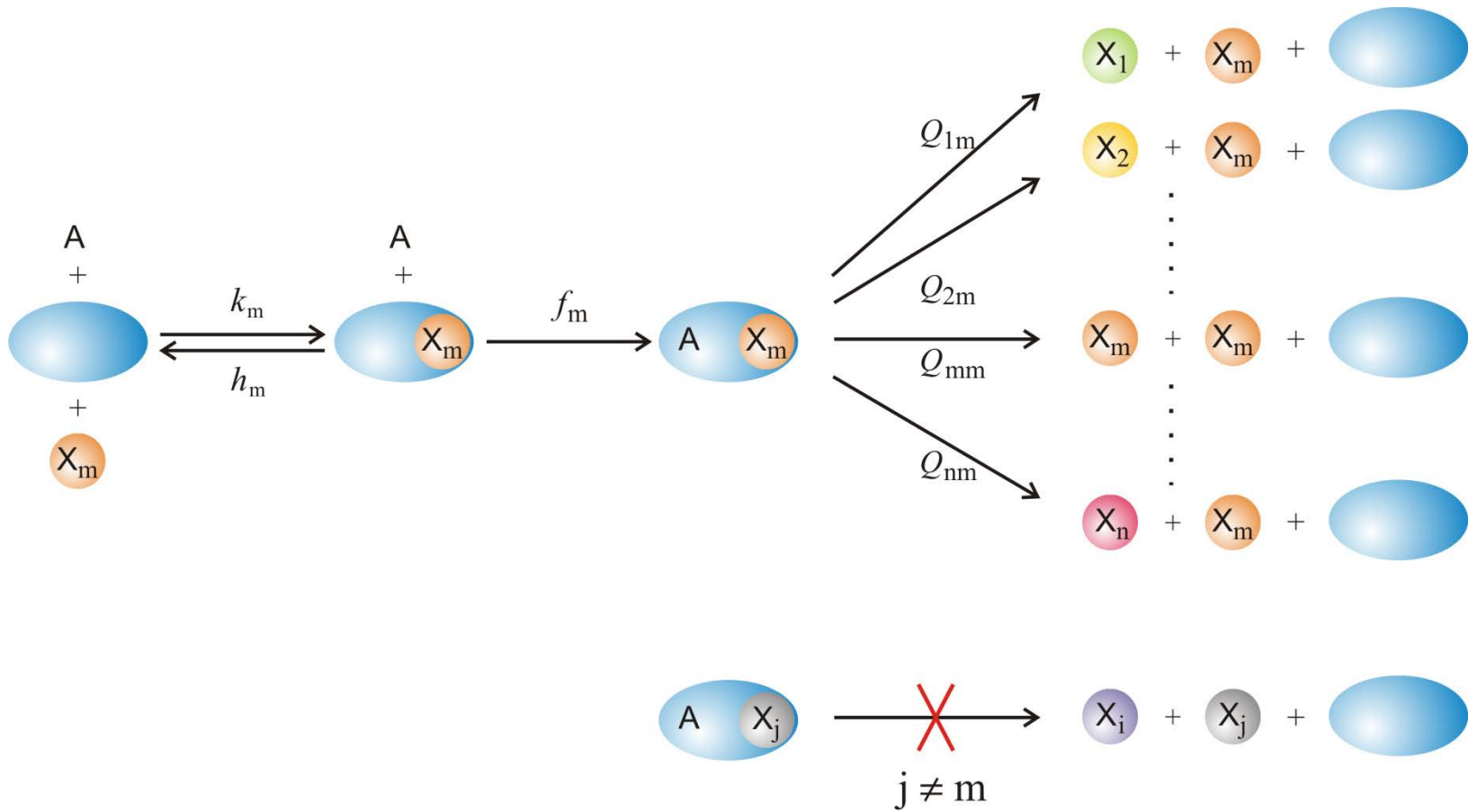
$$\sum_{i=1}^n Q_{ji} \cdot f_i x_i = x_j \sum_{i=1}^n f_i x_i = x_j \Phi$$

mutational flux balance



Mutational flux balance and quasispecies

1. Quasispecies and Crow-Kimura model
2. Mutation flow analysis
3. **Zero backflow and phenomenological approach**
4. Error thresholds on realistic landscapes



Zero mutation backflow

$$\frac{dx_m^{(0)}}{dt} = (W_{mm} - \phi) x_m^{(0)} = (Q_{mm} f_m - \phi) x_m^{(0)}$$

$$\phi = \sum_{i=1}^N Q_{ii} f_i x_i^{(0)} + f_m x_m^{(0)} \sum_{j=1, j \neq m}^n Q_{jm}$$

single peak, uniform error: $\phi = f_0 x_m^{(0)} + Q f (1 - x_m^{(0)})$

$$\begin{aligned} \frac{dx_j^{(0)}}{dt} &= (W_{jj} - \phi) x_j^{(0)} + W_{jm} x_m^{(0)} = \\ &= (Q_{jj} f_j - \phi) x_j^{(0)} + Q_{jm} f_m x_m^{(0)} \end{aligned}$$

Kinetic equations of the zero backflow approximation

$$\bar{x}_m^{(0)} = \frac{Q (1 - \sigma_m^{-1})}{1 - Q \sigma_m^{-1}}$$

$$Q = (1-p)^l \quad \text{and} \quad \sigma_m = f_m / \bar{f}_{-m} = f_m / f$$

$$\bar{x}_j^{(0)} = \frac{Q \varepsilon^{d_{jm}}}{1 - Q \sigma_m^{-1}}, \quad j = 1, \dots, n, \quad j \neq m$$

Solutions of the zero backflow approximation

$$Q_{mm} f_m \hat{c} - f_m \hat{x}_m^{(0)} - \bar{f}_{-m} (\hat{c} - \hat{x}_m^{(0)}) = 0$$

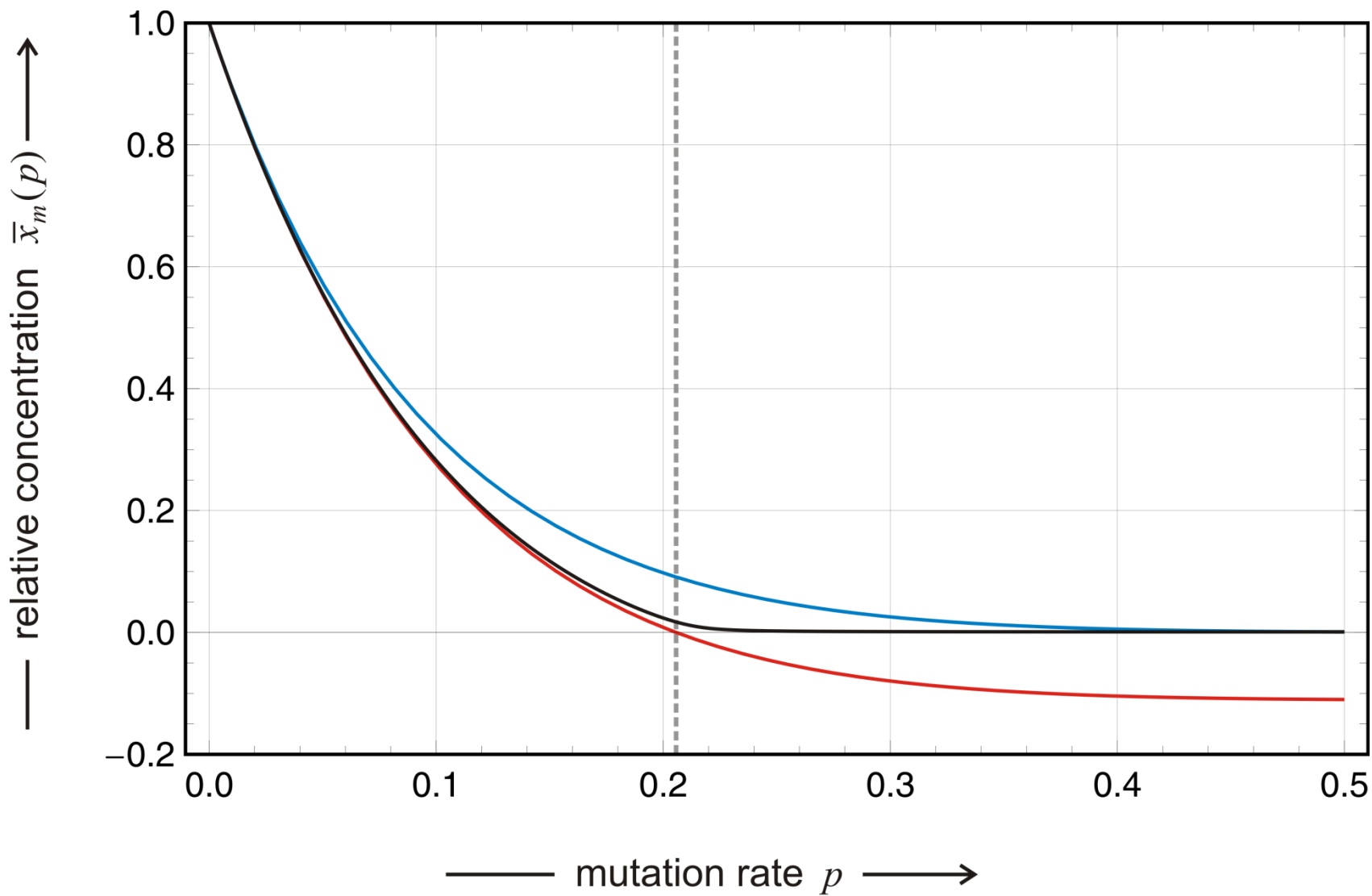
single peak, uniform error

$$\hat{x}_m^{(0)} = \frac{Q f_m - f}{f_m - f} = \frac{Q - \sigma_m^{-1}}{1 - \sigma_m^{-1}} \hat{c}^{(0)}$$

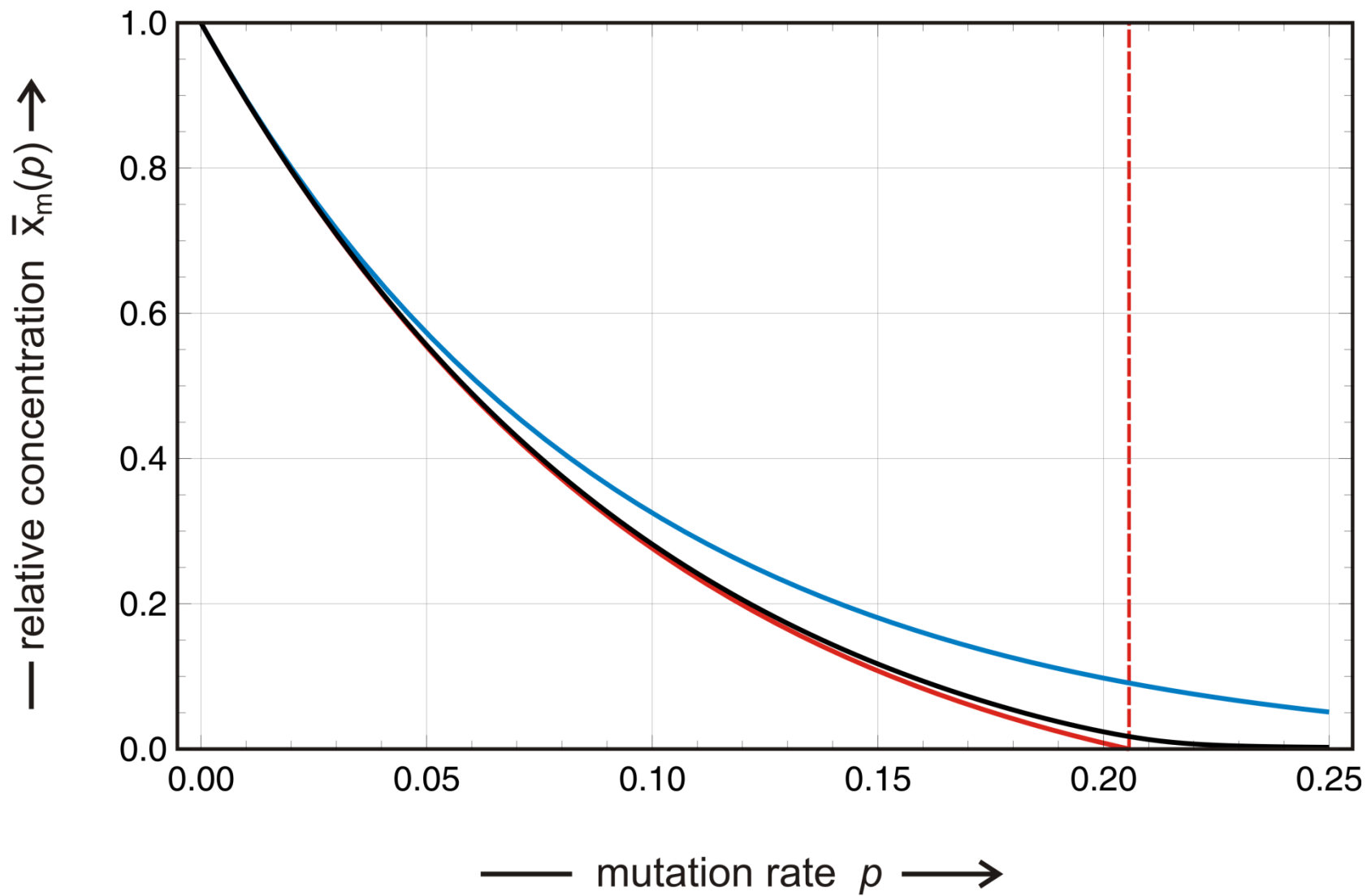
$$\hat{x}_j^{(0)} = \frac{\varepsilon^{d_{jm}^H} (Q - \sigma_m^{-1})}{(1 - \sigma_m^{-1})^2} ; j = 1, \dots, N$$

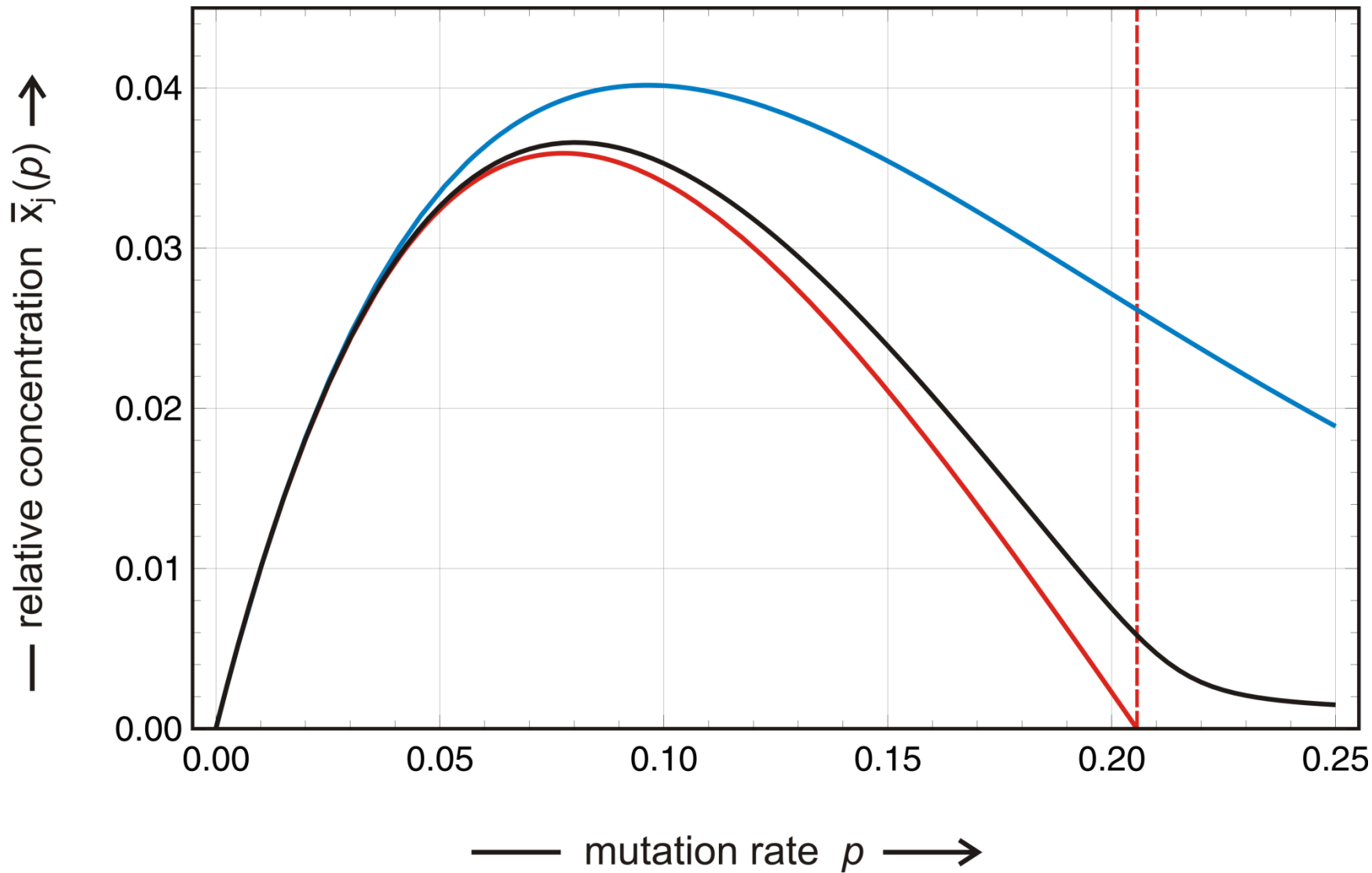
$$\hat{c}^{(0)} = \frac{(1 - Q\sigma_m^{-1})(Q - \sigma_m^{-1})}{Q(1 - \sigma_m^{-1})^2}$$

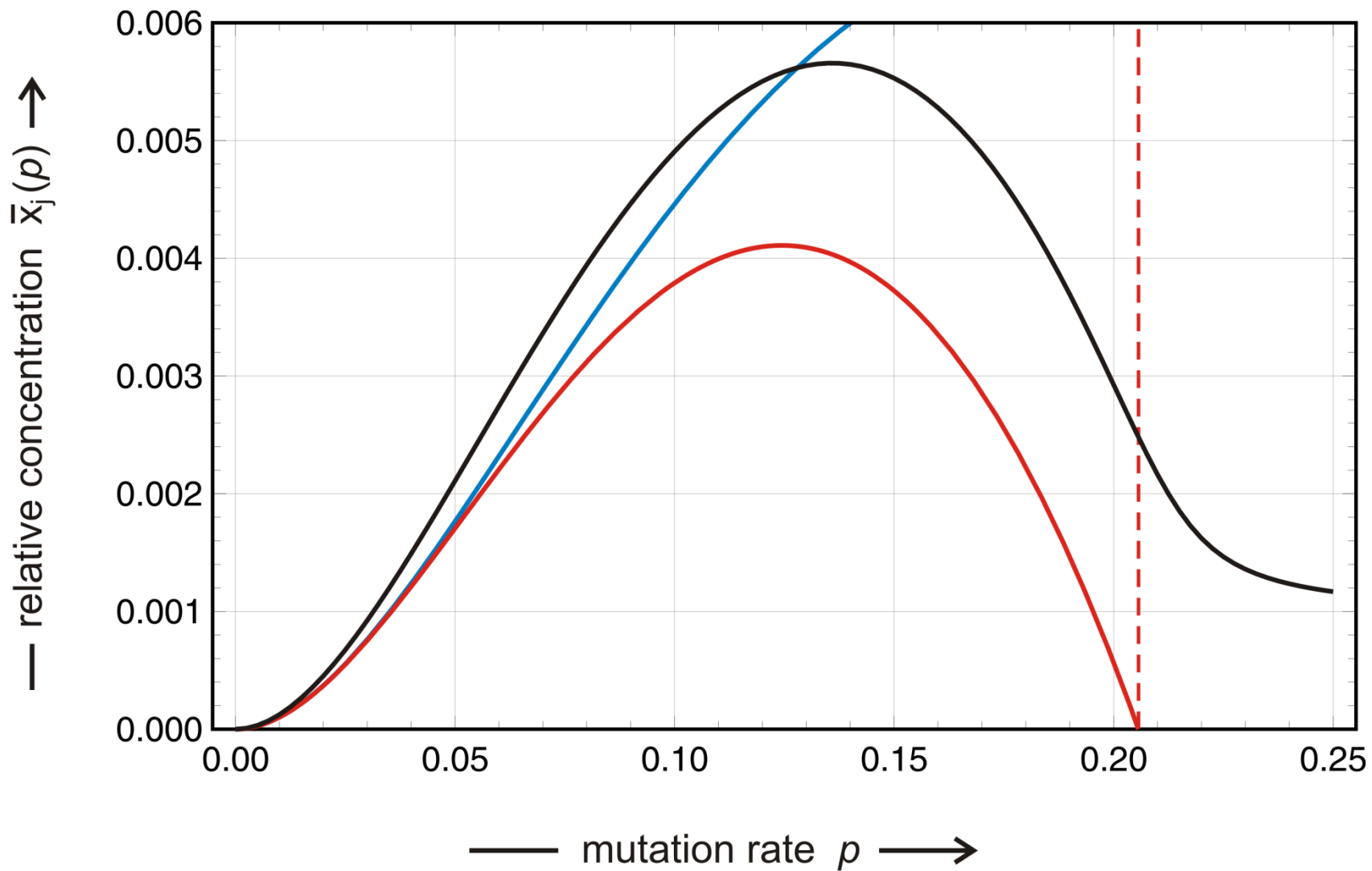
The phenomenological approach (Eigen, 1971)

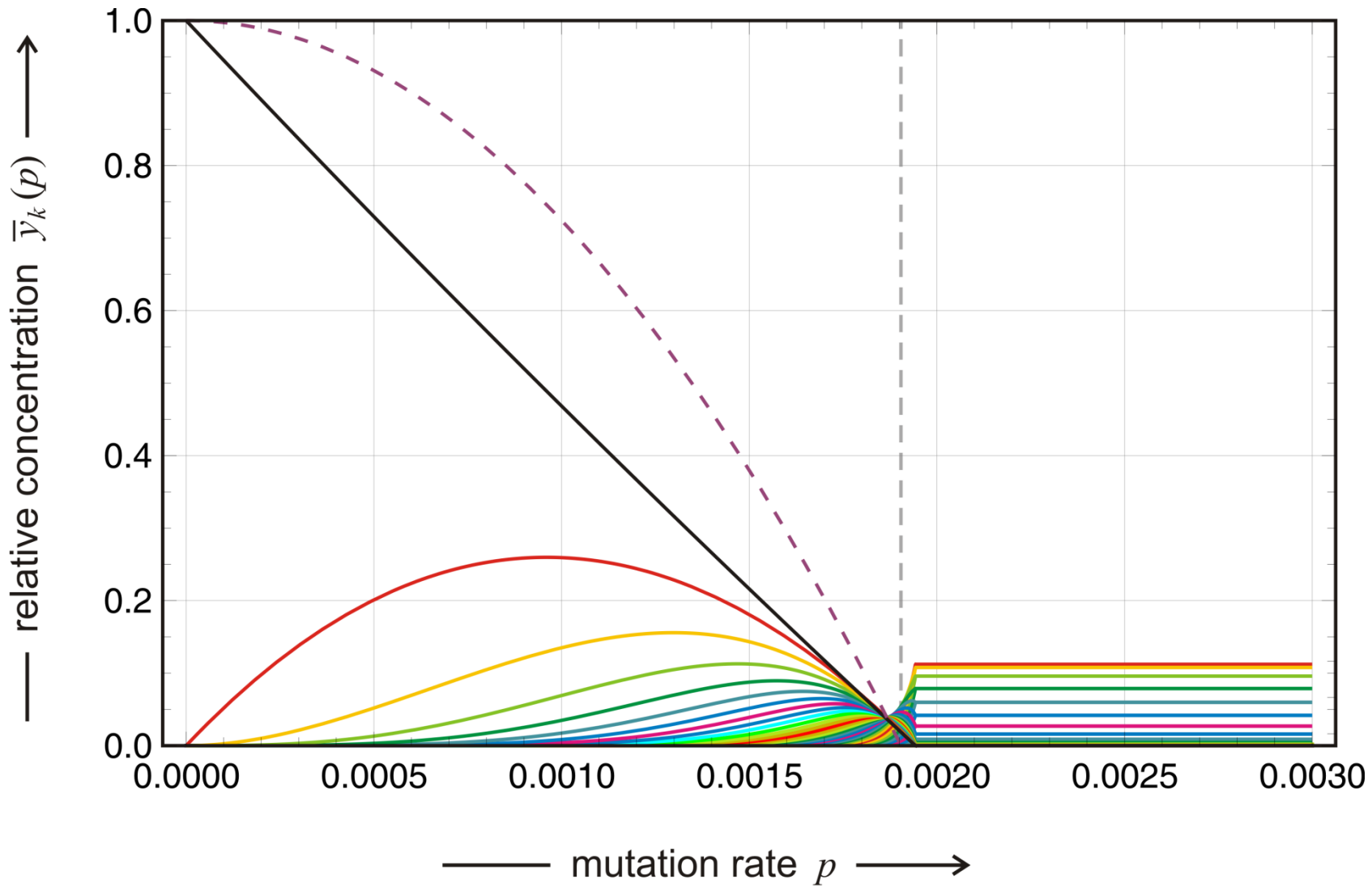


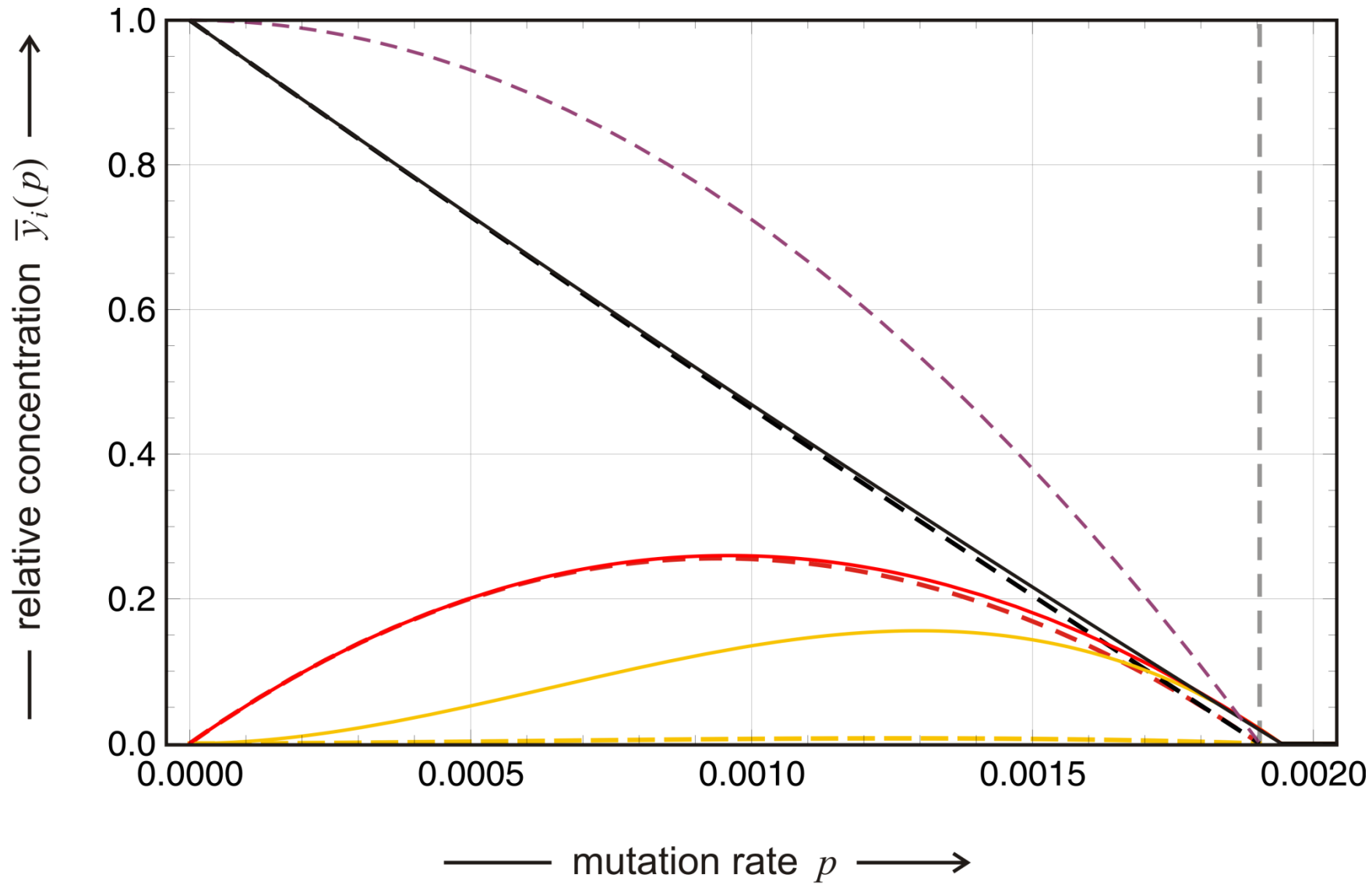
Comparison of exact, zero backflow and phenomenological solutions



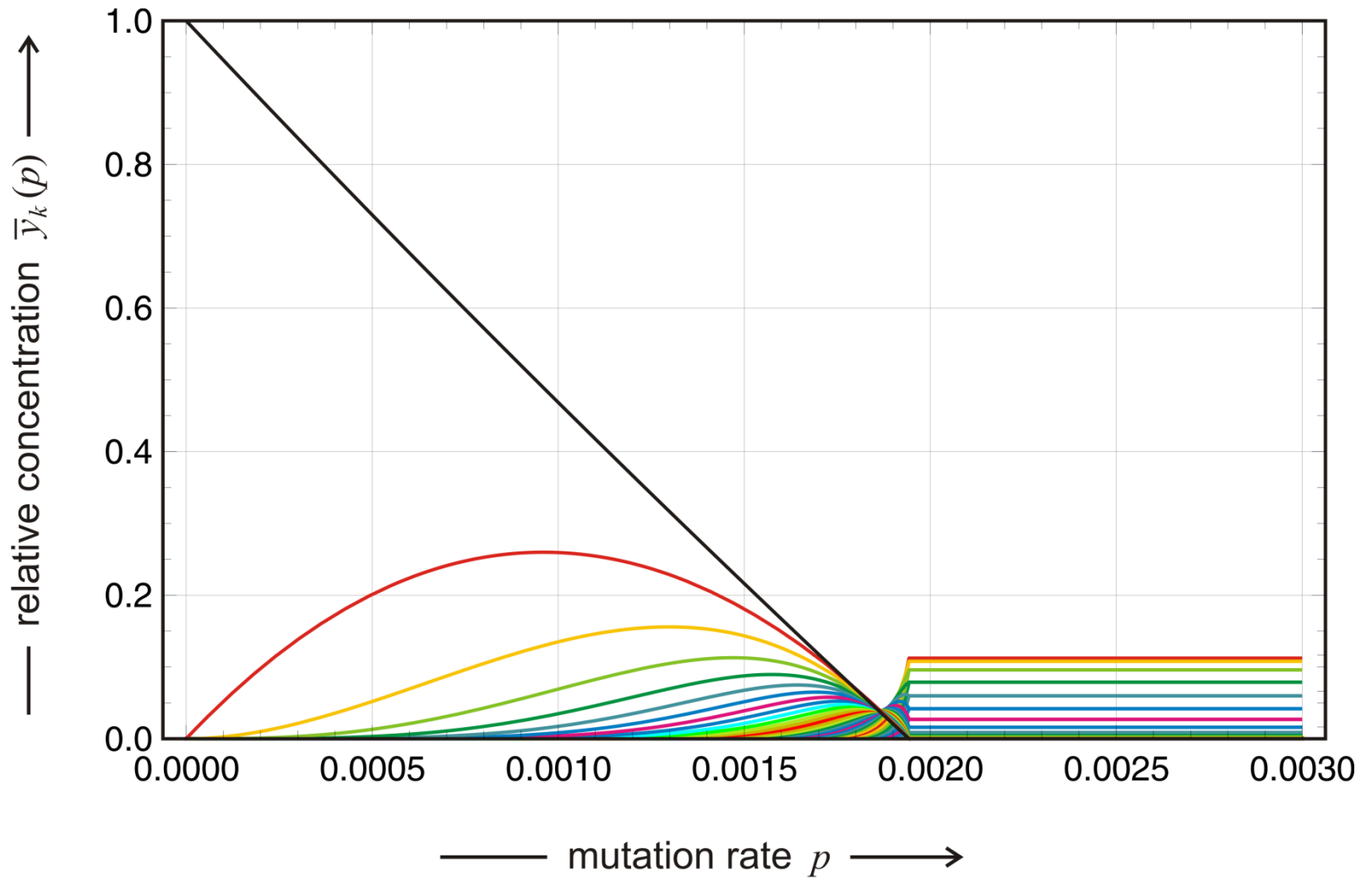




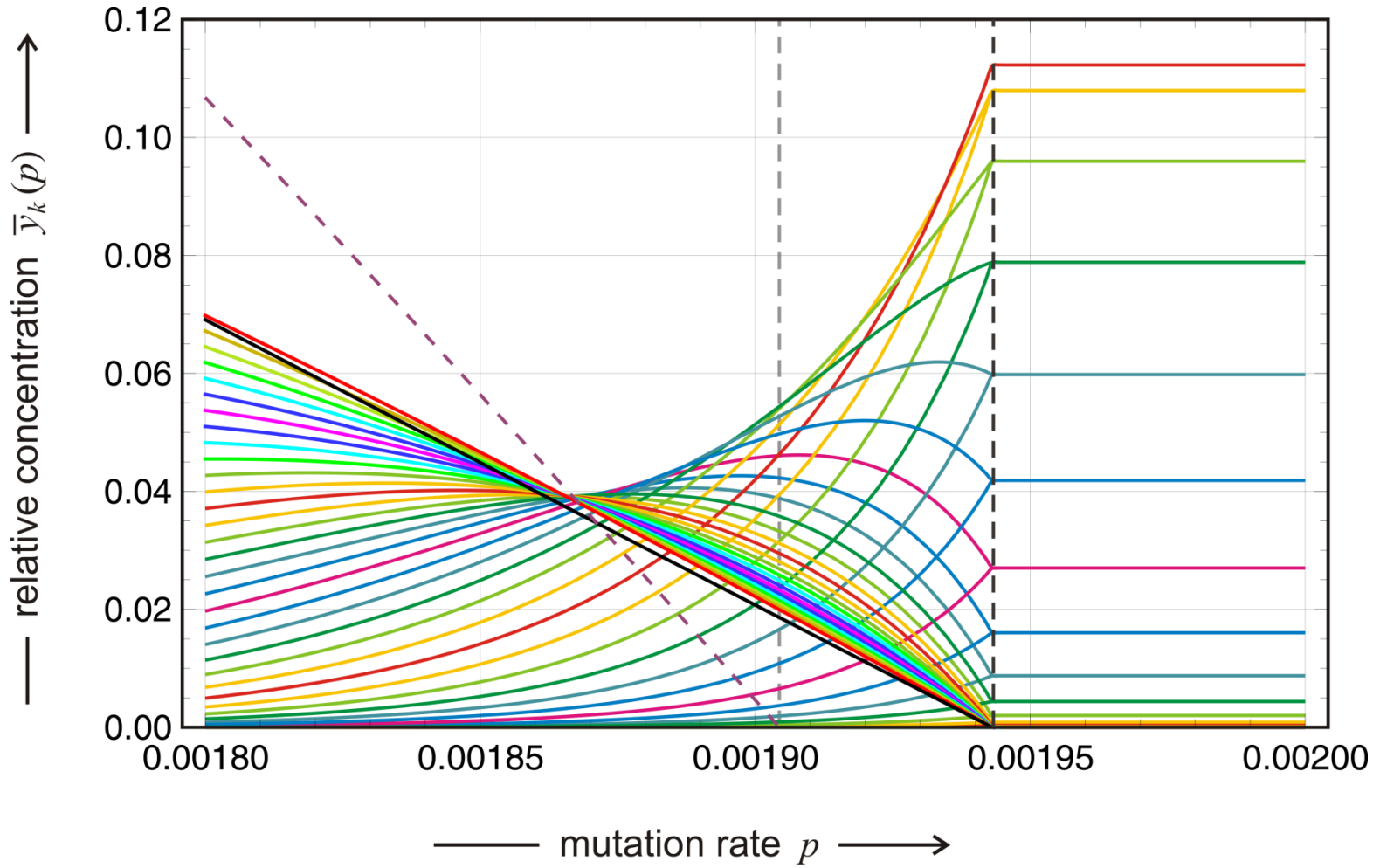




1. Quasispecies and Crow-Kimura model
2. Mutation flow analysis
3. Zero backflow and phenomenological approach
4. Error thresholds on realistic landscapes



Quasispecies and error threshold



Quasispecies and error threshold exact and in the phenotypic approach

level crossing of master sequence: $\bar{x}_m(p_{\text{tr}}^{(\mathcal{G})}) = \mathcal{G}$

complementary class merging: $\Delta_k = \left| \bar{y}_k - \bar{y}_{l-k} \right|$

$$(\Delta_k)_{\text{cr}} = \theta; \quad (p_{\text{mg}}^{(\theta)})_k; \quad k = 0, \dots, \left\lfloor \frac{l}{2} \right\rfloor$$

width of the transition $\Delta p_{\text{mg}}^{(\theta)} = \max(p_{\text{mg}}^{(\theta)})_k - \min(p_{\text{mg}}^{(\theta)})_k; \quad k = 0, \dots, \left\lfloor \frac{l}{2} \right\rfloor$

$$p_{\text{tr}}^{(\mathcal{G})} \approx (p_{\text{mg}}^{(\theta)})_0 \quad \text{for } \mathcal{G} = \theta$$

Quantitative analysis of error thresholds

Table 1: Concentration level crossing near the error threshold. The decline of the master class, $\bar{y}_0 = \bar{x}_0$, at p -values below the error threshold p_{cr} is illustrated by means of the points $p_{(1/M)}$ where $\bar{y}_0(p)$ crosses the level $1/M$ for the three fitness landscapes that sustain error thresholds. Parameters: $\nu = 100$, $f_0 = 10$, and $\bar{f}_{-0} = 1$.

Landscape	Level crossing			Error threshold p_{cr}
	$p_{(1/100)}$	$p_{(1/1000)}$	$p_{(1/10000)}$	
Single-peak	0.02198	0.02274	0.02282	0.02277
Hyperbolic	0.01450	0.01810	0.02036	0.02277
Step-linear	0.01067	0.01774	0.02330	0.02277

Level crossing on model landscapes

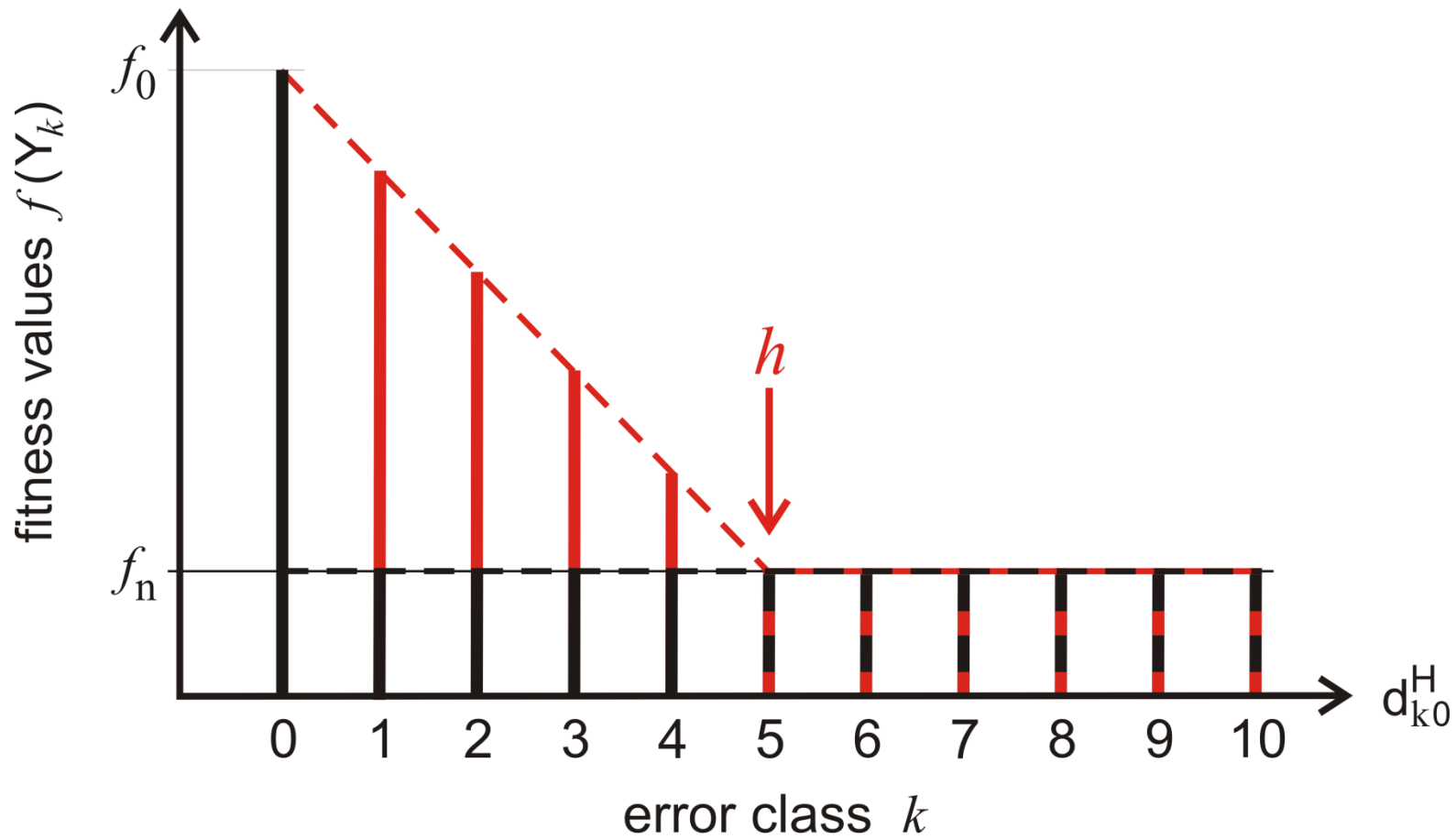
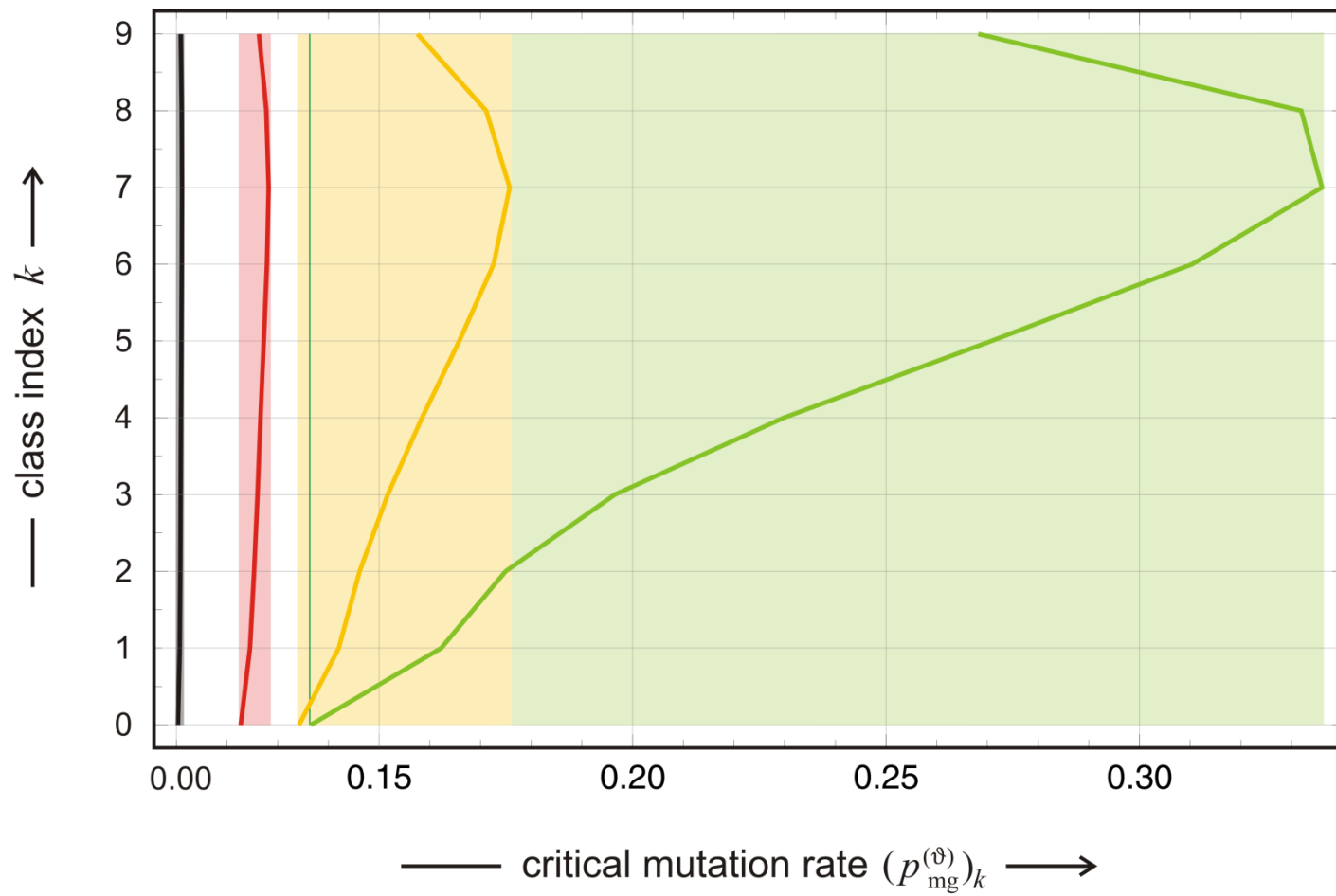
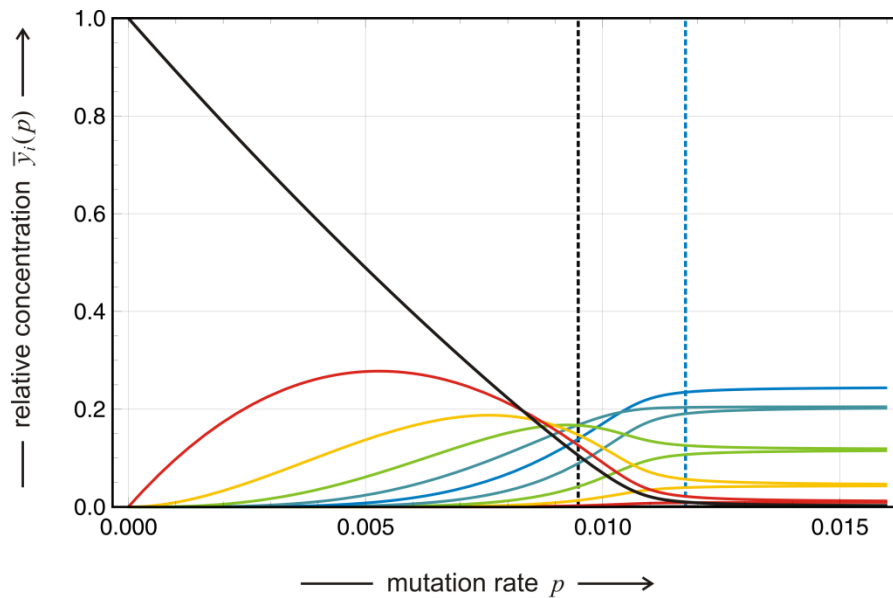


Table 1 Concentration level crossing and complementary class mergence near the error threshold. The decline of the master class, $\bar{y}_0 = \bar{x}_0$, at p -values near the error threshold p_{cr} is illustrated by means of the points $p_{\text{tr}}^{(\vartheta)}$ where the curves cross the level $\bar{x}_0(p) = \vartheta$. Complementary class mergence is characterized quantitatively by the band between the lowest and the highest $(p_{\text{tr}}^{(\vartheta)})_k$ -value. The lowest value is always observed with $k = 0$ (see Fig. 12). Parameters: $l = 20$, $f_0 = 10.0$, and $f_n = 1.0$ yielding an error threshold at $p_{\text{cr}} = 0.1088$.

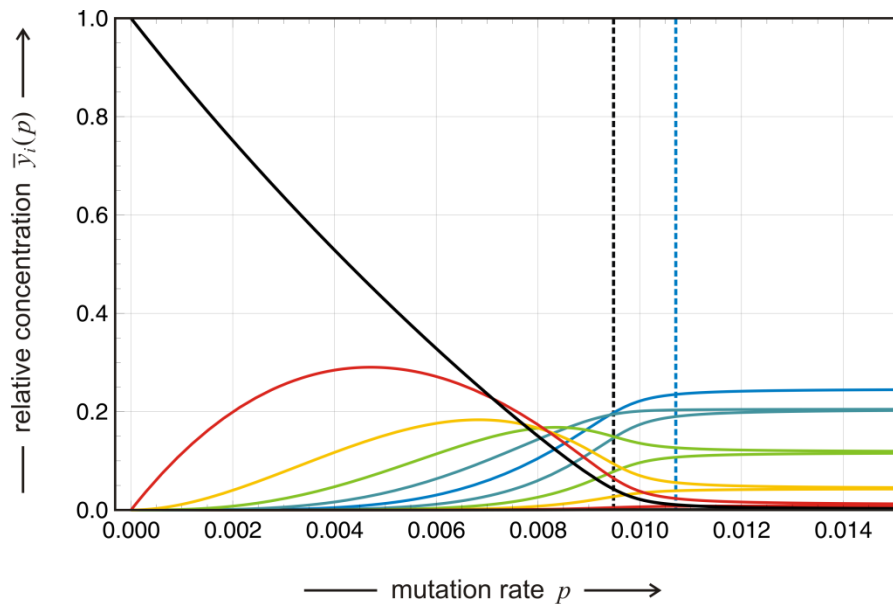
h	Level crossing $p_{\text{tr}}^{(\vartheta)}$			Class mergence $p_{\text{mg}}^{(\theta)}$
	$\vartheta = 1/100$	$\vartheta = 1/1000$	$\vartheta = 1/10000$	$\theta = 1/1000$
0	0.1067	0.1103	0.1110	0.1103 – 0.1111
2	0.1097	0.1227	0.1252	0.1227 – 0.1282
3	0.0999	0.1342	0.1428	0.1342 – 0.1758
4	0.0811	0.1365	0.1626	0.1365 – 0.3360
5	0.0638	0.1244	0.1777	0.1244 – 0.4453
6	0.0513	0.1053	0.1787	---
7	0.0426	0.0876	0.1650	---
8	0.0364	0.0737	0.1449	---



$d = 0.000$

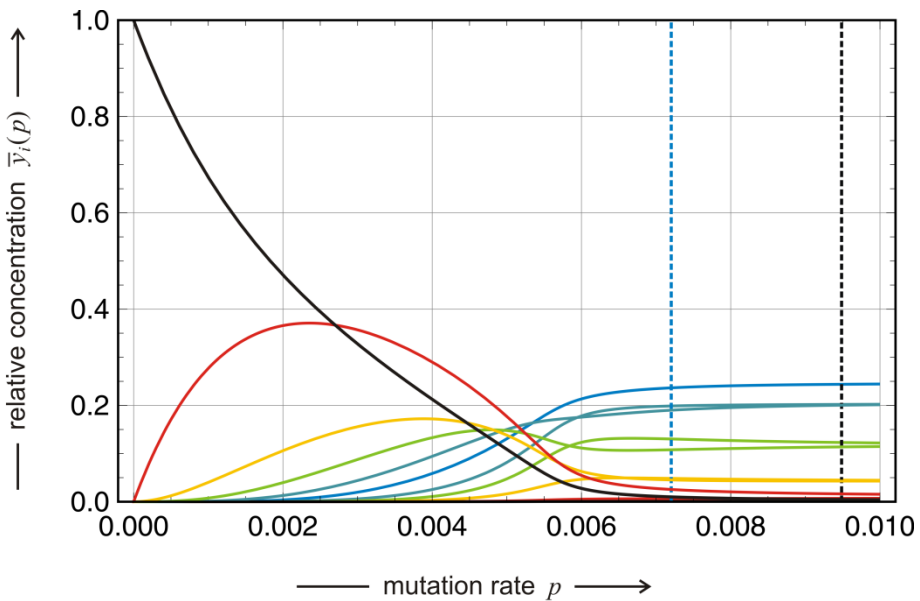


$d = 0.500$

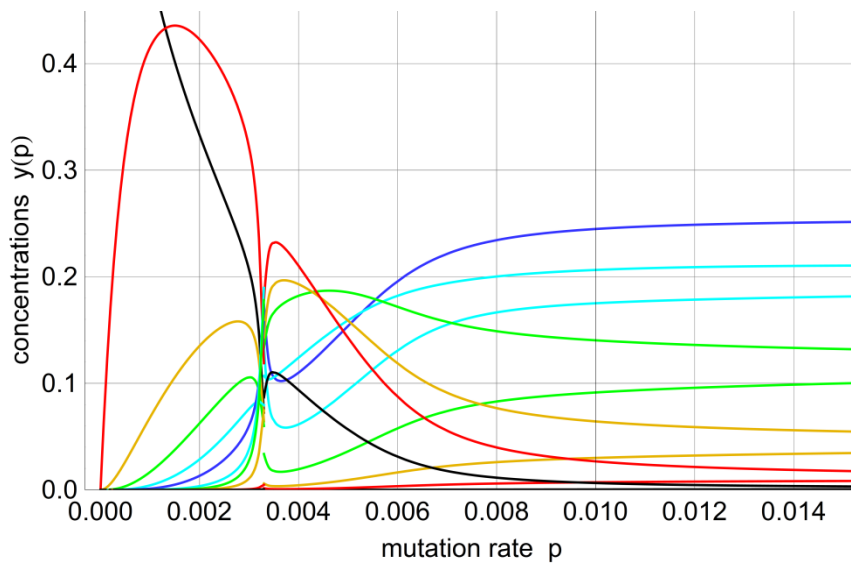


Quasispecies and error threshold on $\mathcal{L}(10,2,1.1,1.0;0.0,d,023)$

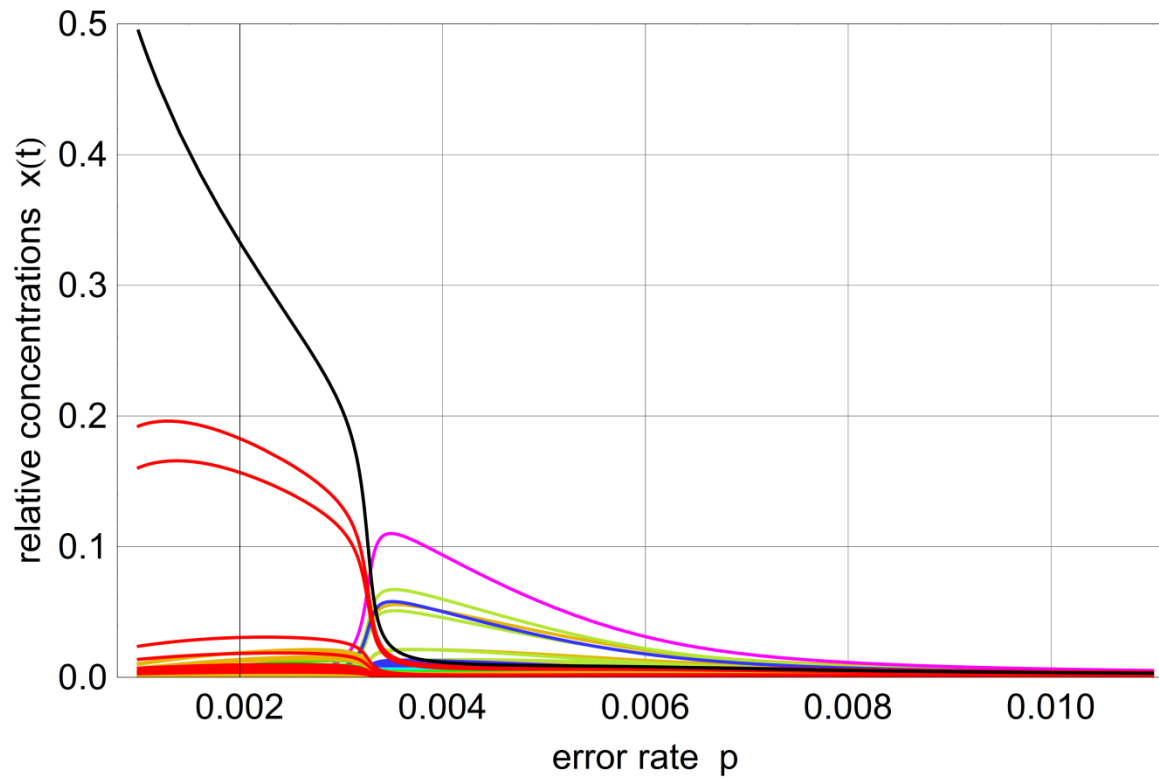
$d = 0.950$



$d = 1.000$

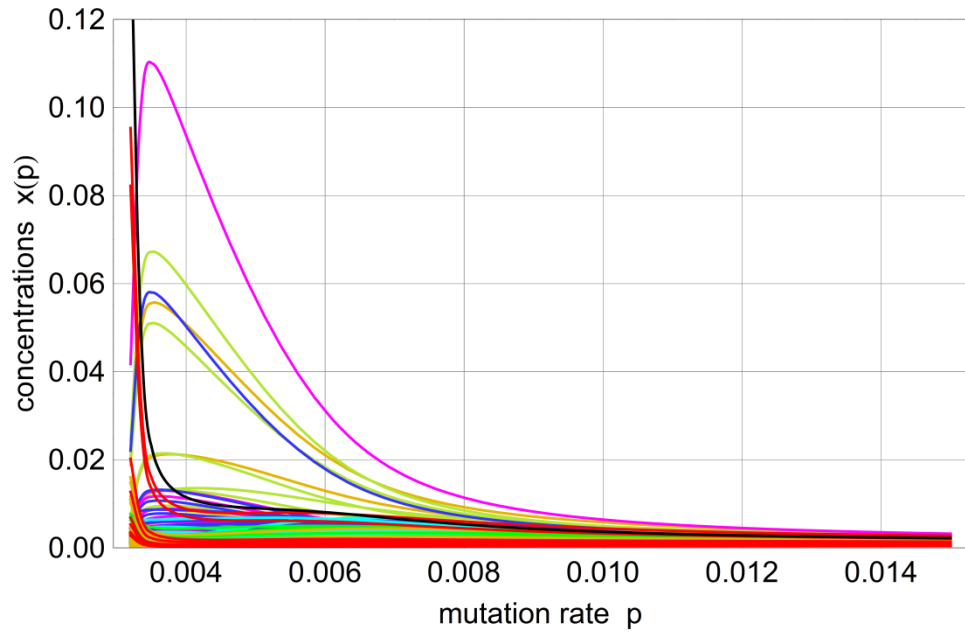


Quasispecies and error threshold on $\mathcal{L}(10,2,1.1,1.0;0.0,d,023)$

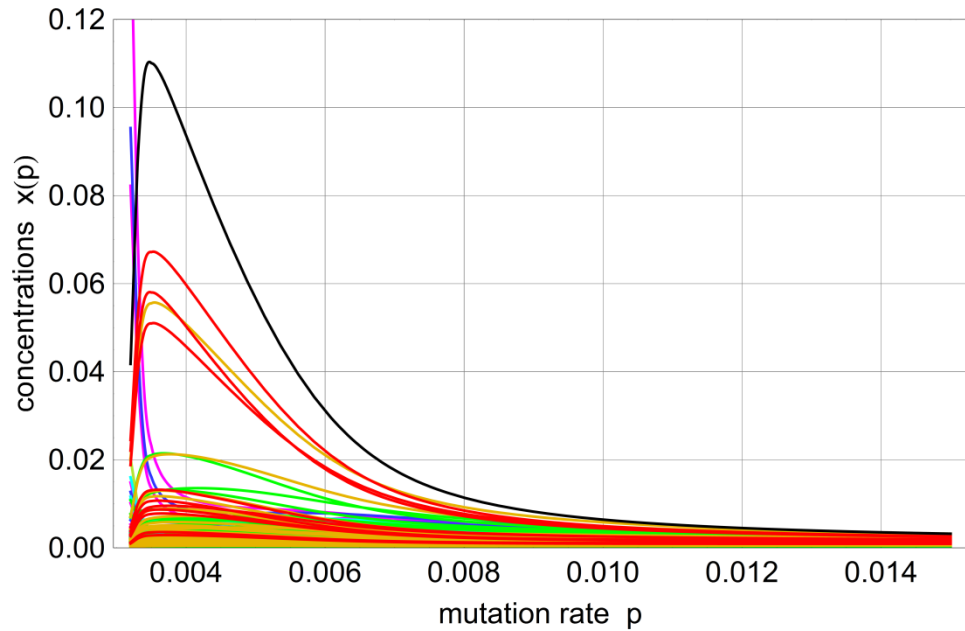


Quasispecies transition on $\mathcal{L}(10,2,1.1,1.0;0.0,1.000,023)$

centered around X_{000}



centered around X_{911}



Quasispecies transition on
 $\mathcal{L}(10,2,1.1,1.0;0.0,d,023)$

Thank you for your attention!

Web-Page for further information:

<http://www.tbi.univie.ac.at/~pks>

