Nonequilibrium Transitions in a Template Copying Ensemble

LPTMC seminar

Arthur Genthon
Max Planck Institute for the Physics of Complex Systems, Dresden

07/03/2025

My background

2015 - 2019 ENS Cachan, Physics

2017 - 2018 Master 2 Physics of Complex Systems (PCS), Paris

2018 - 2019 Master 2 Logique, Philosophie, Histoire et Sociologie des Sciences (LOPHISS), Paris

Statistical Physics

Statistical Field Theory

Stochastic processes

Information theory

Complex Networks

Non-linear dynamics

Inference, Machine Learning

My background

2015 - 2019 ENS Cachan, Physics

2017 - 2018 Master 2 Physics of Complex Systems (PCS), Paris

2018 - 2019 Master 2 Logique, Philosophie, Histoire et Sociologie des Sciences (LOPHISS), Paris

2019 - 2022 PhD in theoretical physics
with David Lacoste, Gulliver lab, ESPCI

Teaching @Université de Paris
(total 168,5 h ETD):

- Introductory computing, L1
TD/TP

- Lab work, L1
TP

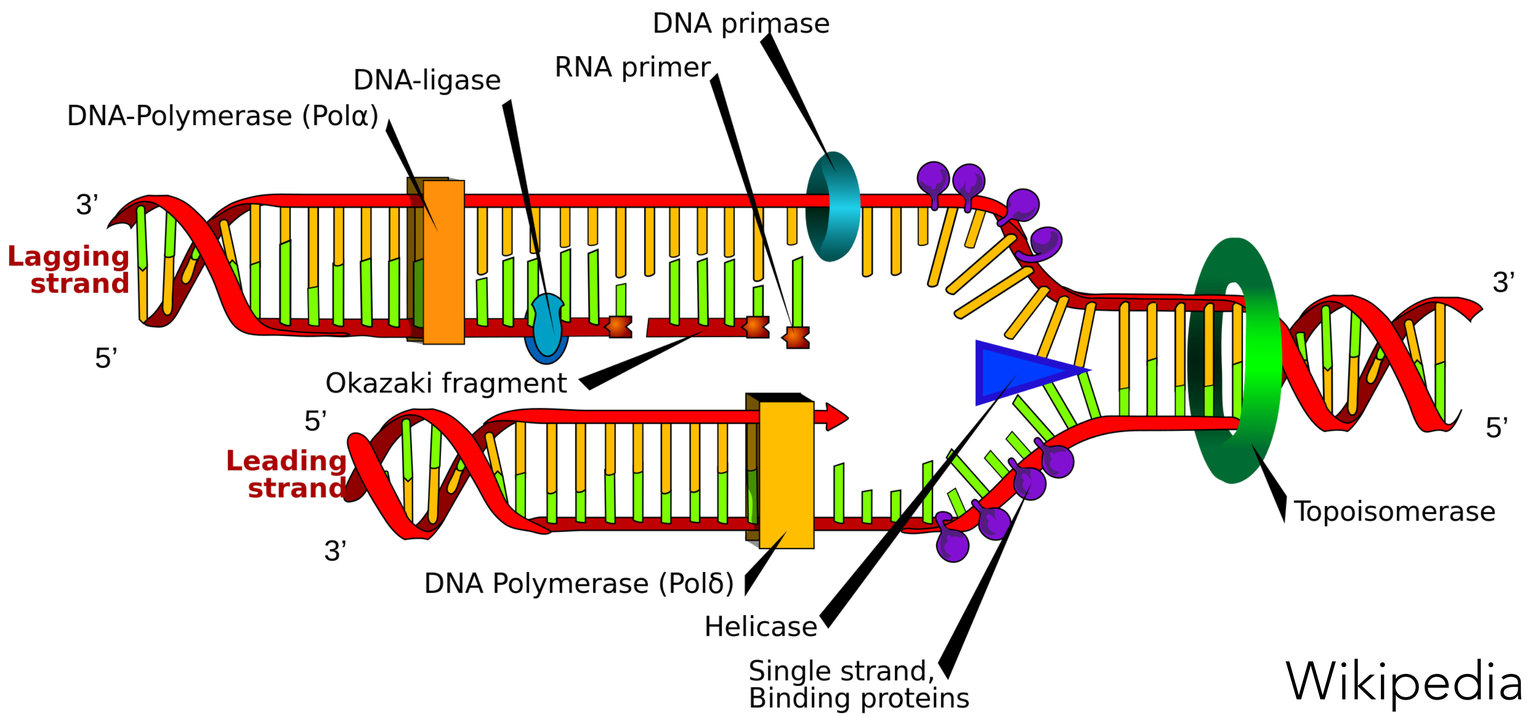
- Thermodynamics, L2
TD

- Mathematics, L3
Lectures + TD

My background

2015 - 2019	ENS Cachan, Physics
2017 - 2018	Master 2 Physics of Complex Systems (PCS), Paris
2018 - 2019	Master 2 Logique, Philosophie, Histoire et Sociologie des Sciences (LOPHISS), Paris
2019 - 2022	PhD in theoretical physics with David Lacoste, Gulliver lab, ESPCI
2023 - Now	Independent postdoc (Guest scientist status) Max Planck Institute for the Physics of Complex Systems, Dresden, Germany

Motivation: DNA replication (DNA → DNA)



Wikipedia

This molecular machinery consumes **energy** in the form of ATP molecules (monomer activation)

Impressively low **error rates**

$\eta \sim 10^{-9}$ DNA replication

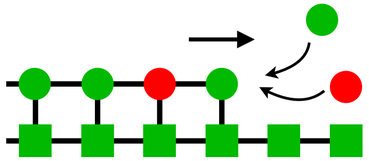
$\eta \sim 10^{-4}$ | RNA transcription (DNA → RNA)
RNA translation (RNA → proteins)

Which conditions for accuracy?
What is the cost of accuracy?

Previous works

Polymer elongation models:

[Andrieux, Gaspard, *PNAS* (2008)]
[Sartori, Pigolotti, *Phys. Rev. Lett.* (2013)]
[Poulton et al., *PNAS* (2019)]
…



Mean elongation velocity

$$\frac{d_{\mathrm{i}}S}{dt} = vA \geq 0,$$

Conditional entropy

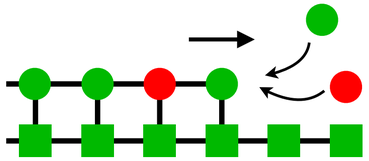
$$A = \varepsilon + D(\mathrm{polymer}|\mathrm{template})$$

[Andrieux, Gaspard, *PNAS* (2008)]

Previous works

Polymer elongation models:

[Andrieux, Gaspard, *PNAS* (2008)]
[Sartori, Pigolotti, *Phys. Rev. Lett.* (2013)]
[Poulton et al., *PNAS* (2019)]
…



Mean elongation velocity

$$\frac{d_{\mathrm{i}}S}{dt} = vA \geq 0,$$

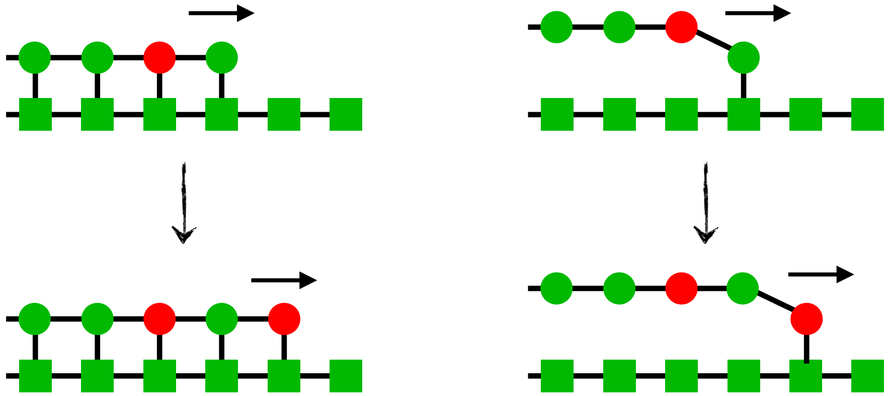
Conditional entropy

$$A = \varepsilon + D(\mathrm{polymer}|\mathrm{template})$$

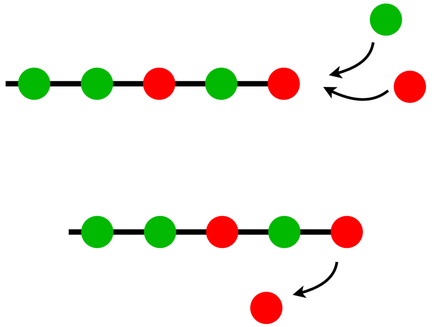
[Andrieux, Gaspard, *PNAS* (2008)]

Limitations:

- Dependent on molecular details



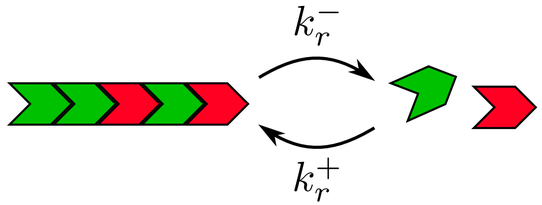
- Spontaneous fluctuations (passives, without template) ignored



⟶ More general formalism needed

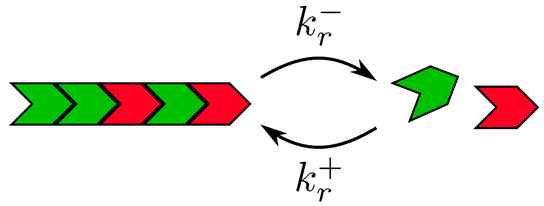
Our approach

- Coarse-graining over molecular details



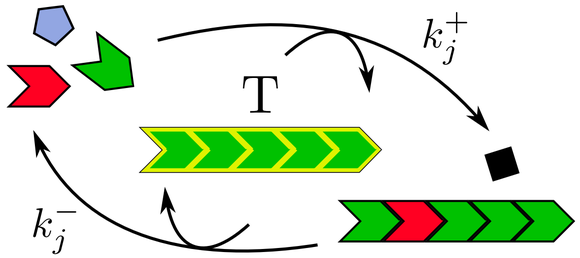
Our approach

- Coarse-graining over molecular details



- Two competing pathways: spontaneous and templated

VS

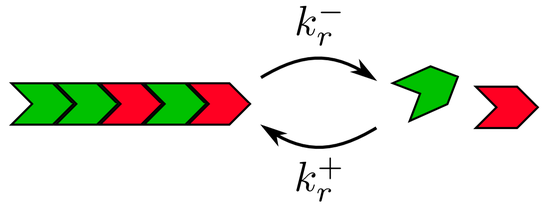


Our approach

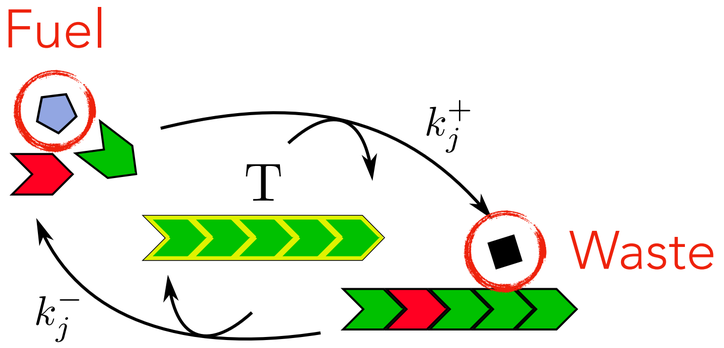
- Coarse-graining over molecular details

- Two competing pathways: spontaneous and templated

- Explicit description of energy consumption and dissipation



vs



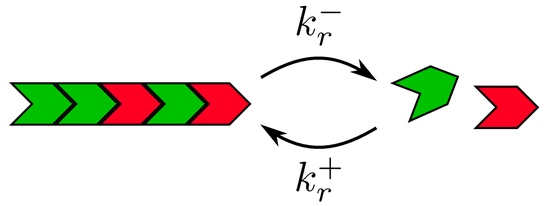
Our approach

- Coarse-graining over molecular details

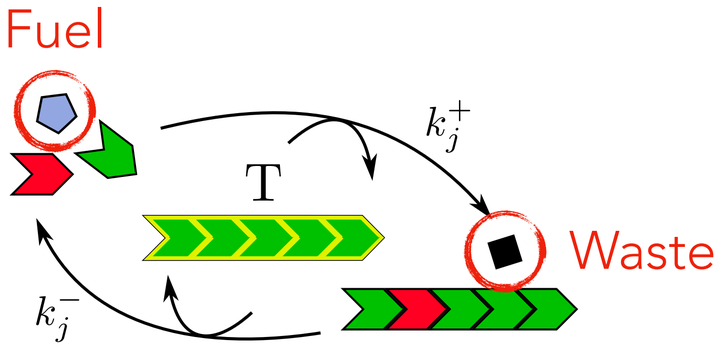
- Two competing pathways: spontaneous and templated

- Explicit description of energy consumption and dissipation

- Population level description

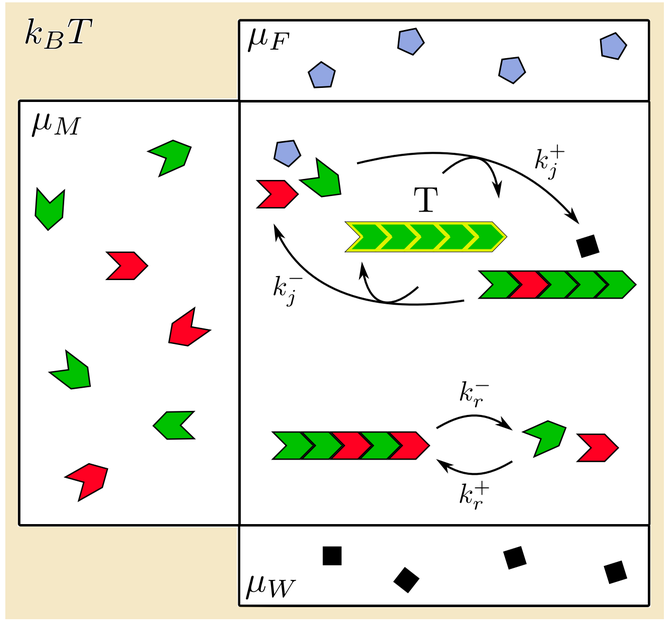


vs



Statistical mechanics ensemble

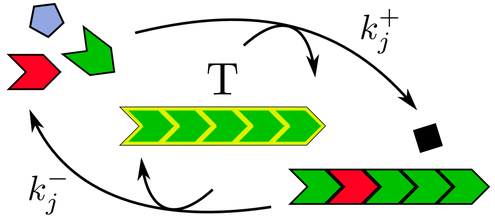
Schematic
$m = 2$



T	Template
L	Template length
M_i	Monomer of type i
m	**Number of monomer types**
n_{ij}	Stoichiometric coefficient
F	Fuel molecule
W	Waste molecule
S_j	Sequence $j = 1,\ldots,m^L$

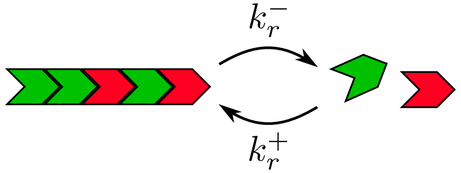
- Templated assembly / disassembly

$$\sum_{i=1}^{m} n_{ij}\mathrm{M}_i + \mathrm{T} + L\mathrm{F} \underset{k_j^-}{\overset{k_j^+}{\rightleftharpoons}} \mathrm{S}_j + \mathrm{T} + L\mathrm{W}$$

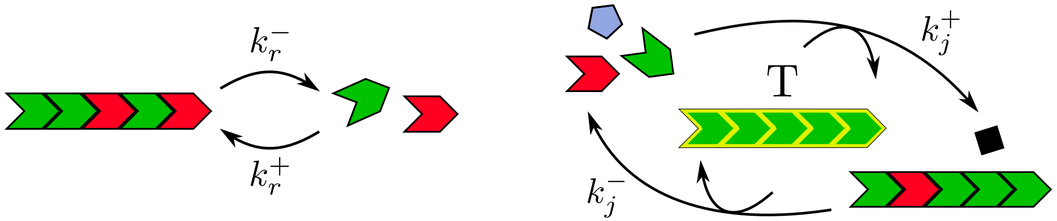


- Spontaneous disassembly / assembly

$$\mathrm{S}_j \underset{k_r^+}{\overset{k_r^-}{\rightleftharpoons}} \sum_{i=1}^{m} n_{ij}\mathrm{M}_i$$



Sequence selectivity



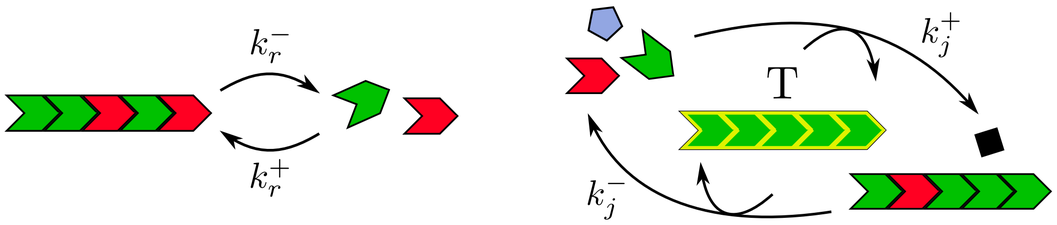
Micro-reversibility

$$\frac{k_r^+}{k_r^-} = e^{-\Delta\mu_r L} \qquad \frac{k_j^+}{k_j^-} = e^{-(\Delta\mu_r - \Delta\mu_F)L}$$

Energy changes per monomer

$$\Delta\mu_r = \epsilon_S/L - \mu_M \qquad \Delta\mu_F = \mu_F - \mu_W > 0$$

Sequence selectivity



Micro-reversibility

$$\frac{k_r^+}{k_r^-} = e^{-\Delta\mu_r L} \qquad \frac{k_j^+}{k_j^-} = e^{-(\Delta\mu_r - \Delta\mu_F)L}$$

Energy changes per monomer

$$\Delta\mu_r = \epsilon_S/L - \mu_M \qquad \Delta\mu_F = \mu_F - \mu_W > 0$$

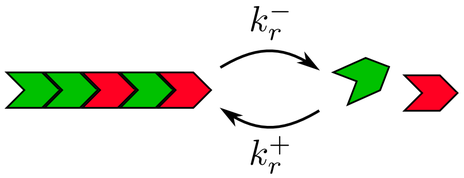
[Ouldridge and Rein ten Wolde, *PRL* (2017)]

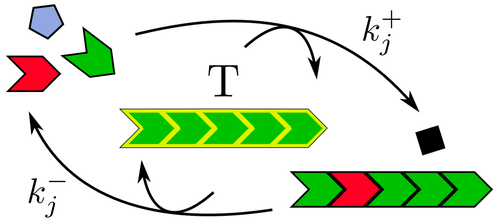
Copies must be **persistent**

→ No lasting copy/template interactions

→ For **autonomous** systems, selectivity must be kinetic

Sequence selectivity





Micro-reversibility

$$\frac{k_r^+}{k_r^-} = e^{-\Delta\mu_r L} \qquad \frac{k_j^+}{k_j^-} = e^{-(\Delta\mu_r - \Delta\mu_F)L}$$

Energy changes per monomer

$$\Delta\mu_r = \epsilon_S/L - \mu_M \qquad \Delta\mu_F = \mu_F - \mu_W > 0$$

[Ouldridge and Rein ten Wolde, *PRL* (2017)]

Copies must be **persistent**

→ No lasting copy/template interactions

→ For **autonomous** systems, selectivity must be kinetic

Kinetic sequence selectivity:

$$k_j^+, k_j^- \propto e^{-aq}$$

$0 \leq q \leq L$ Number of errors

$q = |\mathrm{T} - \mathrm{S}_j|$ Hamming distance

a Specificity

\+ Sequence-independent energetics:

$$\mu_M, \mu_F, \mu_W \perp\!\!\!\perp M$$
$$\epsilon_S \perp\!\!\!\perp S$$

Statistics of errors

$$\partial_t p(N_S,t) = k_a p(N_S-1,t) - (k_a + N k_d) p(N_S,t) + (N_S+1) k_d p(N_S+1,t)$$

$$\partial_t p(0,t) = -k_a p(0,t) + k_d p(1,t)$$

$k_a = k_j^+ + k_r^+$ Total assembly rate

$k_d = k_j^- + k_r^-$ Total disassembly rate

Solution

$$p(N_S,t) = \frac{\lambda_q^{N_S}}{N_S!} e^{-\lambda_q}$$

$$\lambda_q = \frac{k_a}{k_d}\left(1 - e^{-k_d t}\right)$$

Statistics of errors

$$\partial_t p(N_S,t) = k_a p(N_S-1,t) - (k_a + N k_d) p(N_S,t) + (N_S+1) k_d p(N_S+1,t)$$

$$\partial_t p(0,t) = -k_a p(0,t) + k_d p(1,t)$$

$k_a = k_j^+ + k_r^+$ Total assembly rate

$k_d = k_j^- + k_r^-$ Total disassembly rate

Solution

$$p(N_S,t) = \frac{\lambda_q^{N_S}}{N_S!} e^{-\lambda_q}$$

$$\lambda_q = \frac{k_a}{k_d}\left(1 - e^{-k_d t}\right)$$

$$\langle N_S \rangle \underset{t\to\infty}{\to} \frac{k_r^+ + k_j^+}{k_r^- + k_j^-}$$

Rate parametrisation matters:

$$k_r^+ = k_r e^{-\Delta\mu_r L} \qquad k_j^+ = k_0 e^{-aq} e^{-(\Delta\mu_r - \Delta\mu_F)L}$$

$$k_r^- = k_r \qquad k_j^- = k_0 e^{-aq}$$

Choice: Polymers disassemble only spontaneously for large L

Statistics of errors

$$\partial_t p(N_S,t) = k_a p(N_S-1,t) - (k_a + N k_d) p(N_S,t) + (N_S+1) k_d p(N_S+1,t)$$

$$\partial_t p(0,t) = -k_a p(0,t) + k_d p(1,t)$$

$k_a = k_j^+ + k_r^+$ Total assembly rate

$k_d = k_j^- + k_r^-$ Total disassembly rate

Solution

$$p(N_S,t) = \frac{\lambda_q^{N_S}}{N_S!} e^{-\lambda_q}$$

$$\lambda_q = \frac{k_a}{k_d}\left(1 - e^{-k_d t}\right)$$

$$\langle N_S \rangle \underset{t\to\infty}{\to} \frac{k_r^+ + k_j^+}{k_r^- + k_j^-}$$

Rate parametrisation matters:

$$k_r^+ = k_r e^{-\Delta\mu_r L} \qquad k_j^+ = k_0 e^{-aq} e^{-(\Delta\mu_r - \Delta\mu_F)L}$$

$$k_r^- = k_r \qquad k_j^- = k_0 e^{-aq}$$

Choice: Polymers disassemble only spontaneously for large L

Number of copies with an error fraction $x = \frac{q}{L}$

$$\langle N_x \rangle = \lambda_{xL} \Omega_{xL}$$

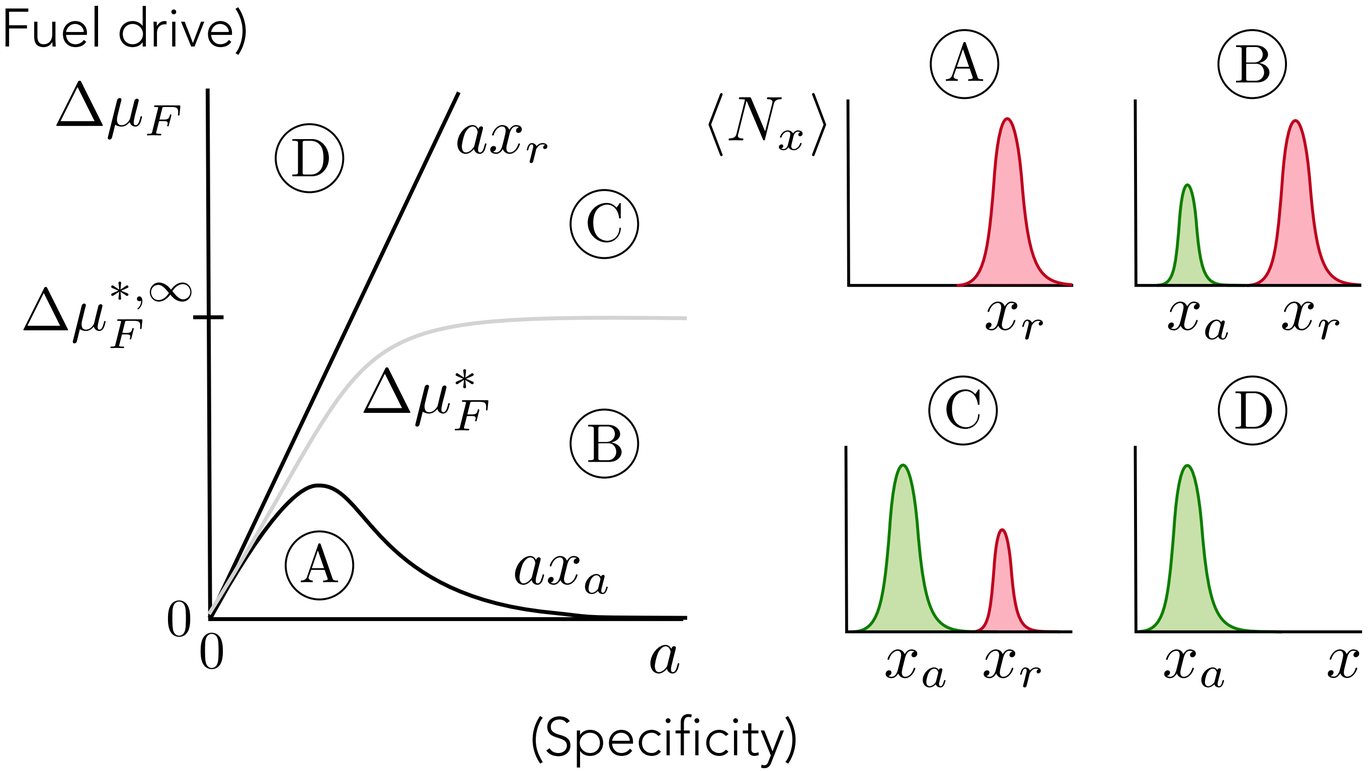
Number of sequences with q errors

$$\Omega_q = \binom{L}{q}(m-1)^q$$

Phase diagram

Finite L

$x = q/L$: monomeric error fraction



Random copies:

$$x_r = \frac{m-1}{m} + O(L^{-1})$$

Accurate copies:

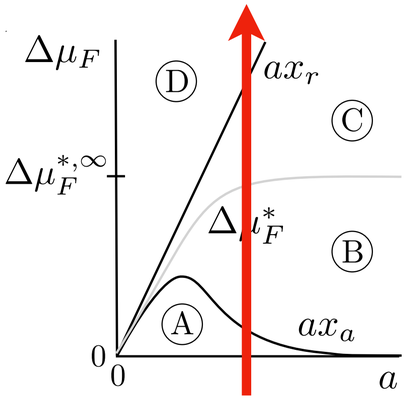
$$x_a = \frac{1}{1 + e^a/(m-1)} + O(L^{-1})$$

$$0 \le x_a \le x_r$$

$$\Delta\mu_F^{*,\infty} = \ln m + O(L^{-1})$$

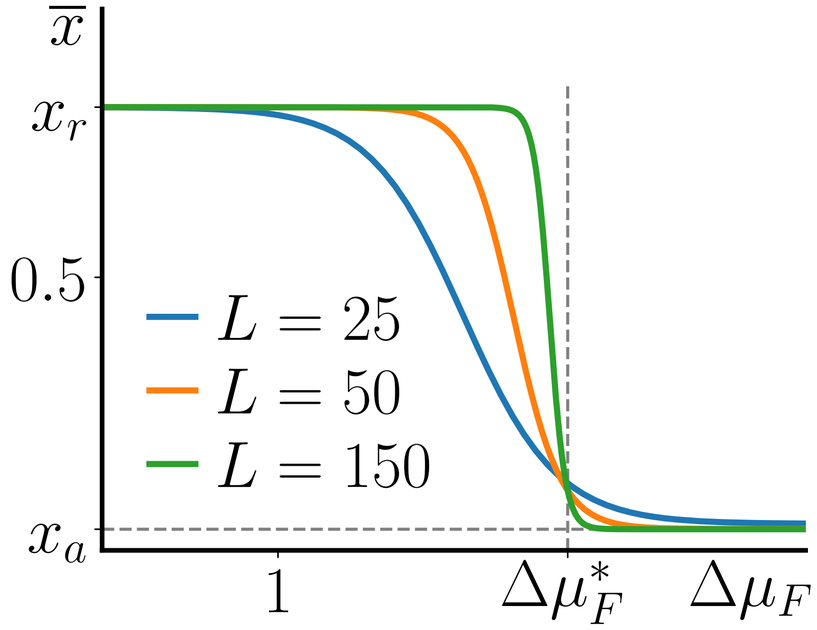
Phase transition from randomness to accuracy

When varying energy drive



Mean error fraction

$$\bar{x} = \frac{\sum_x x\langle N_x\rangle}{\sum_x \langle N_x\rangle}$$



Accurate copies dominate in the large L limit for all values of specificity if:

$$\Delta\mu_F \geq \max_a \Delta\mu_F^* = \ln m$$

Akin to Landauer's principle

$$\Delta\mu_F^* = \ln\left(\frac{m}{1+(m-1)e^{-a}}\right) + O\left(\frac{1}{L}\right)$$

Phase diagram

Can the population participating in this phase transition vanish?

Large L

Energetic contribution

Entropic contribution

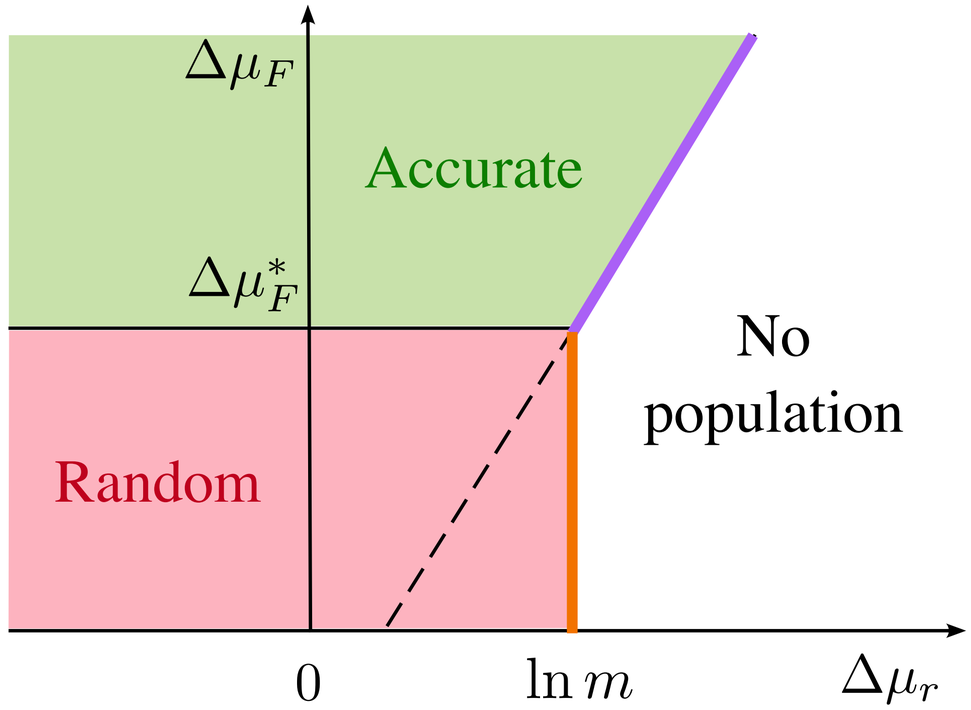
$$\ln\langle N_x\rangle = \ln\lambda_{xL} + \ln\Omega_{xL}$$

Random copies

$$\ln\langle N_{x_r}\rangle = (-\Delta\mu_r + \ln m)L$$

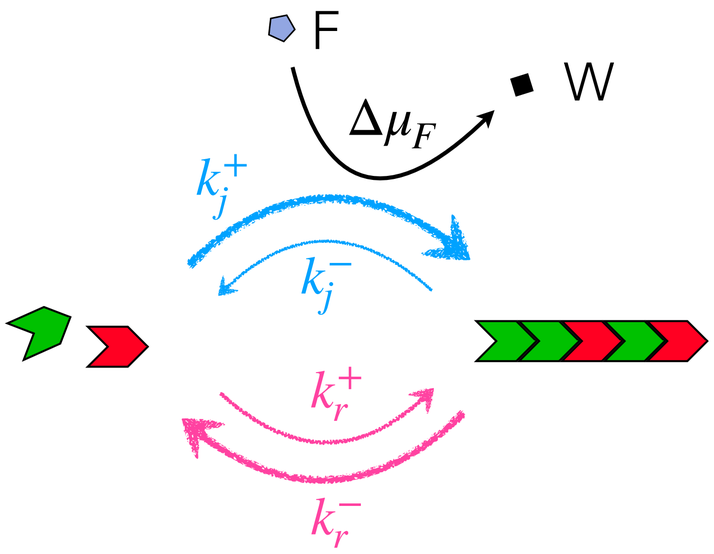
Accurate copies

$$\ln\langle N_{x_a}\rangle = [\Delta\mu_F - \Delta\mu_r + \ln(1 + (m-1)e^{-a})]L$$



For fixed specificity a

Non-equilibrium steady-state current

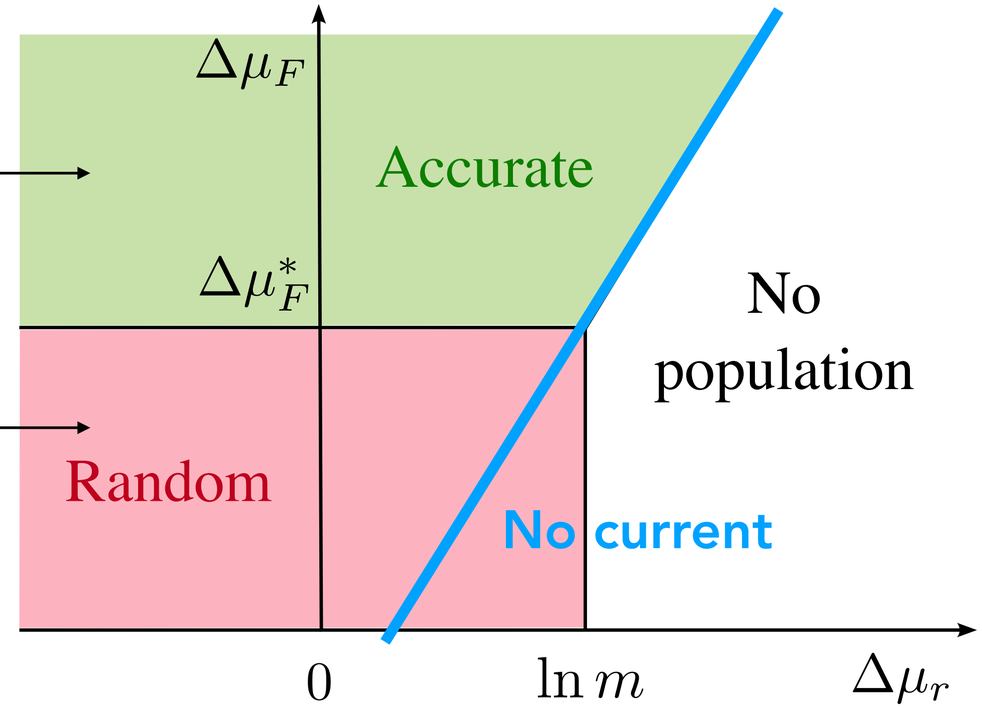


$$\langle J\rangle = L\sum_{j=1}^{m^L}(k_j^+ - \langle N_{S_j}\rangle k_j^-)$$

$$\sim Lk_0\exp\{[\Delta\mu_F - \Delta\mu_r + \ln(1+(m-1)e^{-a})]L\}$$

$\langle J\rangle \sim k_r\langle N_{x_a}\rangle$

Dissipation but no useful information transmitted from template to copy



Freedom in the parametrisation of kinetic rates

Kinetic barriers matter even in steady-state!

$$\langle N_S \rangle \underset{t\to\infty}{\to} \frac{k_r^+ + k_j^+}{k_r^- + k_j^-}$$

	Templated	Spontaneous
Constraints on ratios:	$\frac{k_j^+}{k_j^-} = e^{-(\Delta\mu_r - \Delta\mu_F)L}$	$\frac{k_r^+}{k_r^-} = e^{-\Delta\mu_r L}$

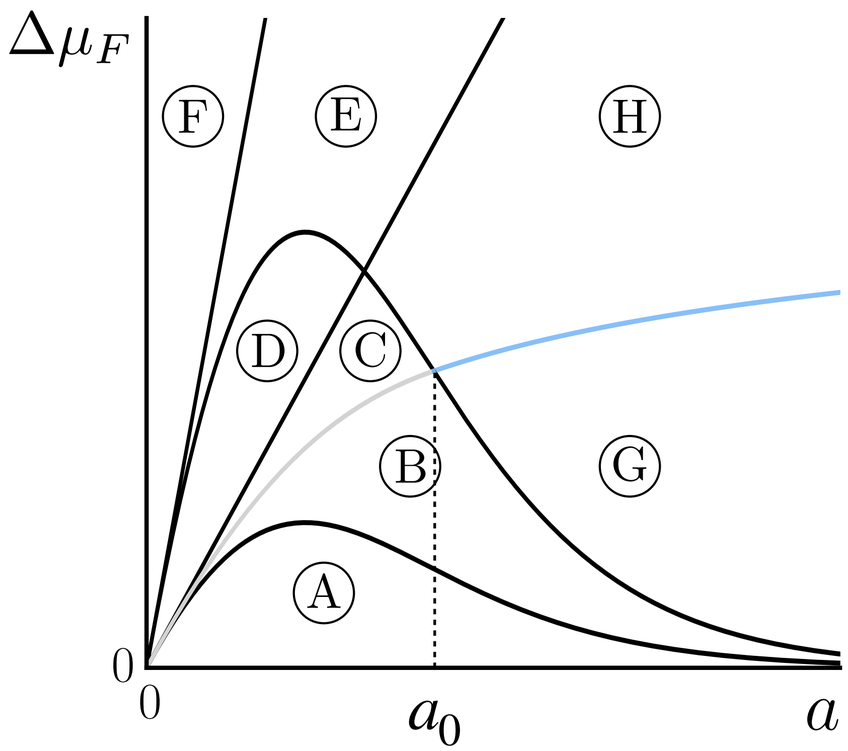
$\gamma \in \mathbb{R}$ Changes the relative timescale of the two processes

	Templated	Spontaneous
Assembly	$k_j^+ = k_0 e^{-axL} e^{-\Delta\mu_r L + (1+\gamma)\Delta\mu_F L}$	$k_r^+ = k_r e^{-\Delta\mu_r L}$
Disassembly	$k_j^- = k_0 e^{-axL} e^{\gamma\Delta\mu_F L}$	$k_r^- = k_r$

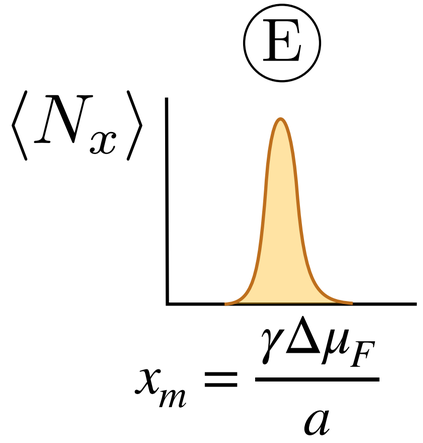
Templated disassembly dominates spontaneous disassembly for $ax < \gamma\Delta\mu_F$

Speeding up templated disassembly: finite L

$\gamma > 0$



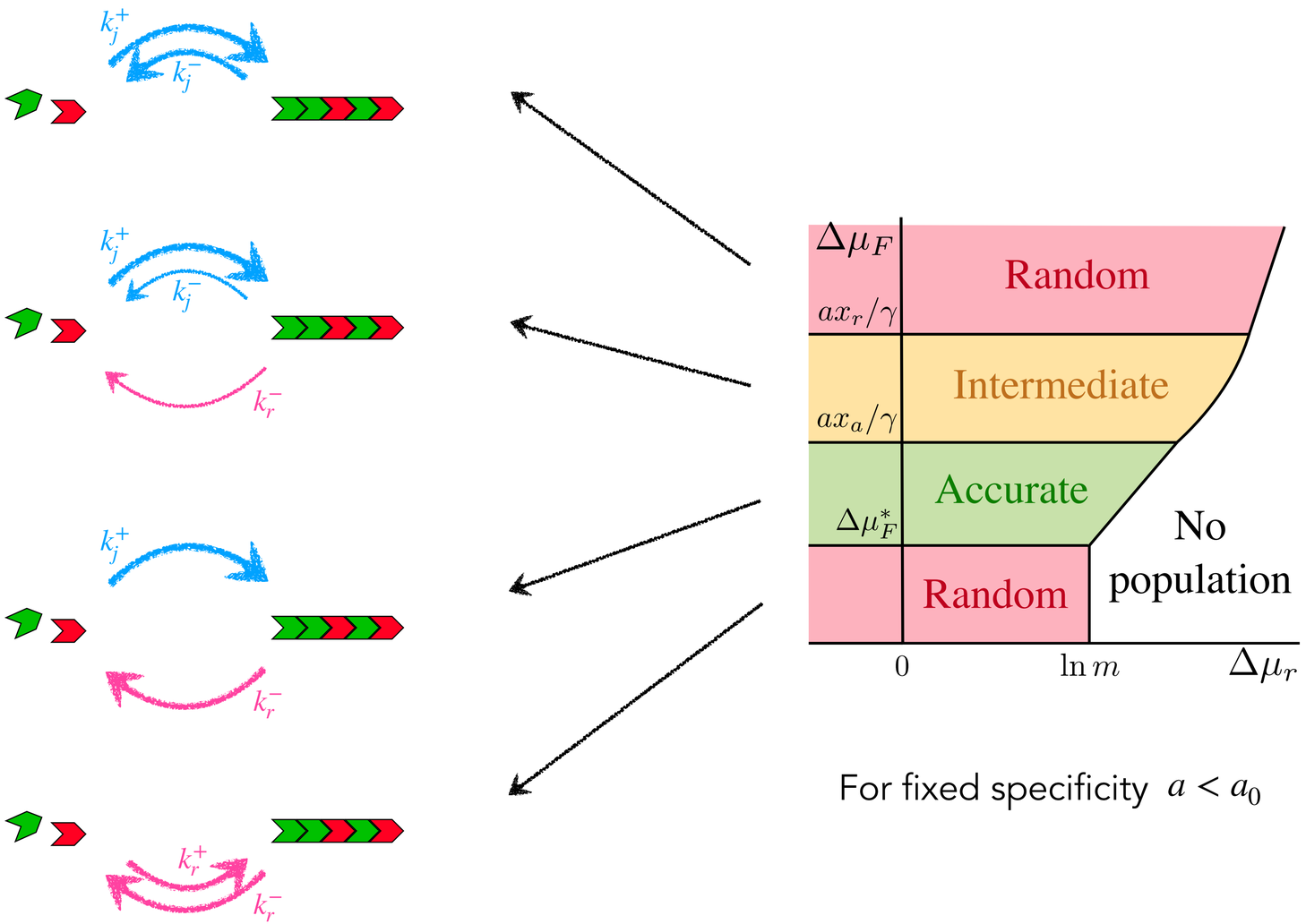
4 new phases: E-H



New error fraction

Speeding up templated disassembly: large L

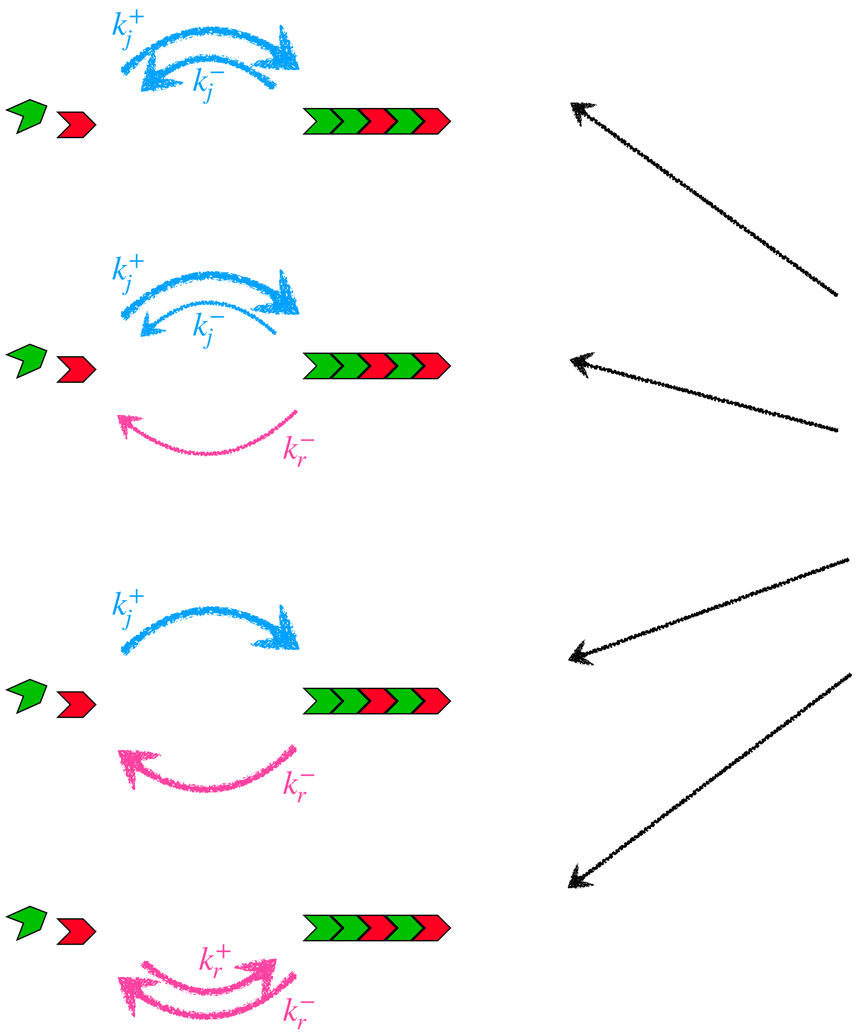
$\gamma > 0$

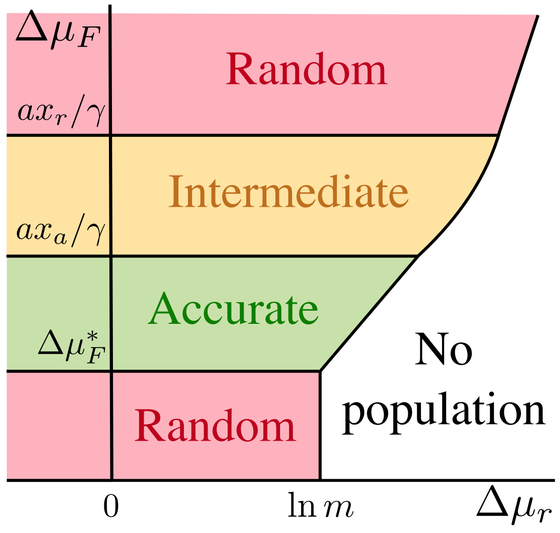


For fixed specificity $a < a_0$

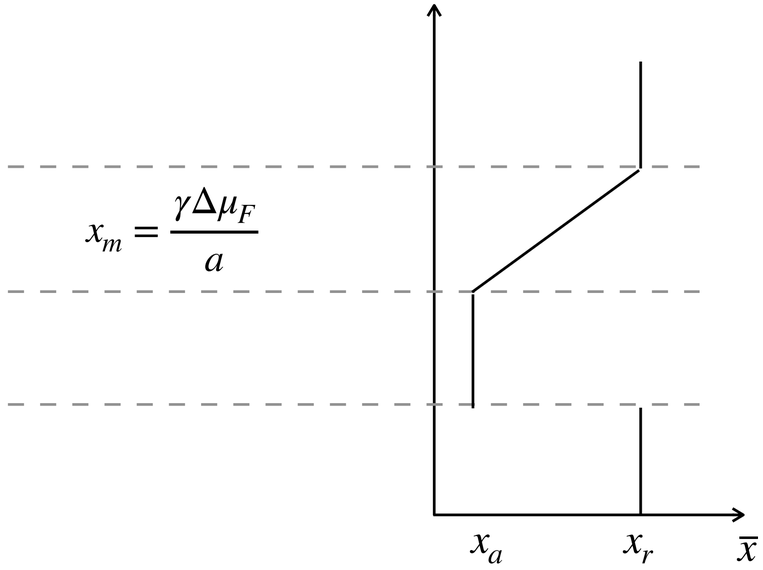
Speeding up templated disassembly: large L

$\gamma > 0$





For fixed specificity $a < a_0$



Take-home messages

[**Genthon**, Modes, Jülicher, Grill, *Phys. Rev. Lett.* (2025)]

- General framework to investigate energy-consuming information copying
- Allows for a discussion of **cost-accuracy trade-offs** $L\Delta\mu_F^*$ vs x_a
  (in particular with number m of building blocks)
- Discussion of processes that are not tightly-coupled is possible (e.g. **kinetic proofreading**)

Stephan Grill

Carl Modes

Frank Jülicher

MAX PLANCK INSTITUTE
OF MOLECULAR CELL BIOLOGY
AND GENETICS

mpipks

Backup slides

Cost-accuracy trade-off in number of monomer types m (Large L)

Accuracy

$$x_a = \frac{1}{1 + e^a/(m-1)}$$ Error fraction of accurate copies

Cost

Energy cost per monomer

$$\Delta\mu_F^* = \ln\left(\frac{m}{1 + (m-1)e^{-a}}\right)$$

Total energy cost

$$E_{\text{tot}}^* = L(m)\Delta\mu_F^* = \left(1 - \frac{\ln\left(1 + (m-1)e^{-a}\right)}{\ln m}\right)\ln\Omega$$

$\Omega = m^L$ Number of "messages" of length L in base m

Cost-accuracy trade-off in number of monomer types m (Large L)

Accuracy

$$x_a = \frac{1}{1 + e^a/(m-1)}$$ Error fraction of accurate copies

Cost

Energy cost per monomer

$$\Delta\mu_F^* = \ln\left(\frac{m}{1 + (m-1)e^{-a}}\right)$$

Total energy cost

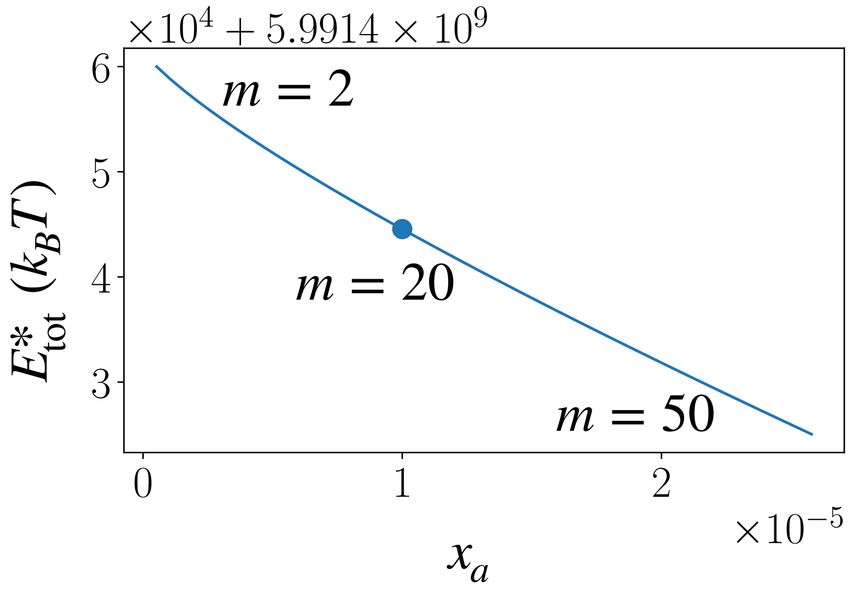
$$E_{\text{tot}}^* = L(m)\Delta\mu_F^* = \left(1 - \frac{\ln\left(1 + (m-1)e^{-a}\right)}{\ln m}\right)\ln\Omega$$

$\Omega = m^L$ Number of "messages" of length L in base m

Trade-off:

$$\frac{\partial E_{\text{tot}}^*}{\partial m} < 0 \qquad \frac{\partial x_a}{\partial m} > 0$$

For translation:



General phase diagram

Templated:

$$k_j^- = k_0 e^{-axL} e^{-((1+\alpha)\Delta\mu_r - (1+\gamma)\Delta\mu_F)L}$$

$$k_j^- = k_0 e^{-axL} e^{-(\alpha\Delta\mu_r - \gamma\Delta\mu_F)L}$$

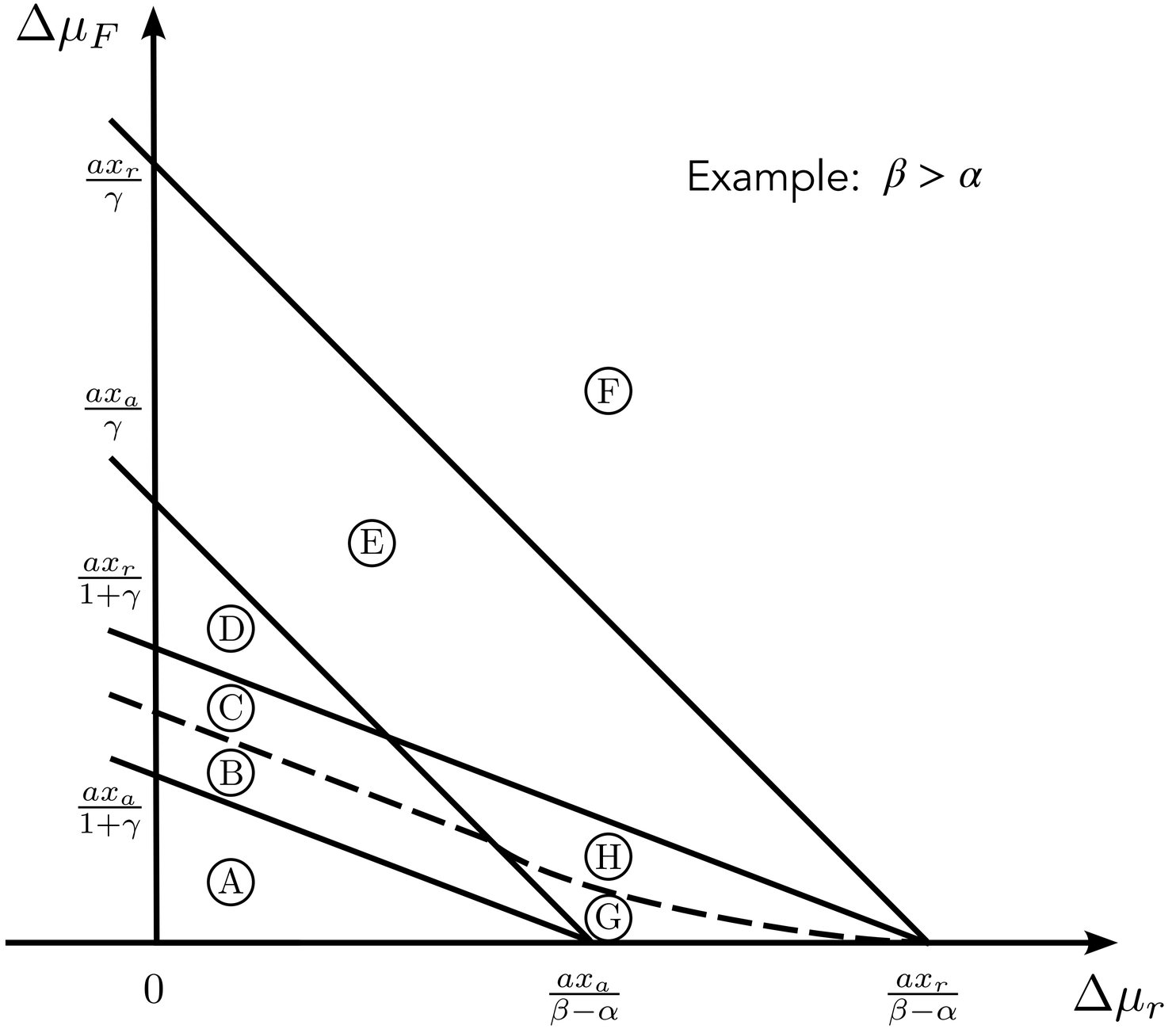
Spontaneous:

$$k_r^- = k_r e^{-\beta\Delta\mu_r L}$$

$$k_r^- = k_r e^{-(1+\beta)\Delta\mu_r L}$$

Arbitrary splittings:

$$\alpha, \beta, \gamma \in \mathbb{R}$$

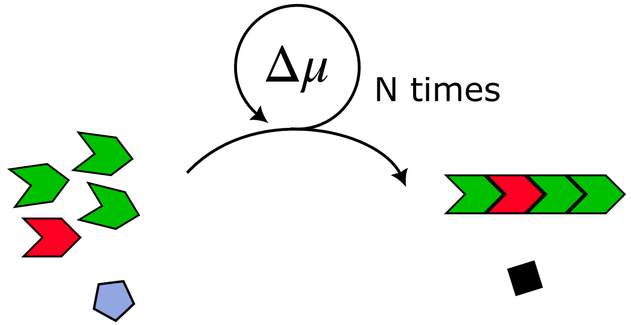


Example: $\beta > \alpha$

Kinetic proofreading

- Random number N of fuel-consuming proofreading cycles

  → distribution of values for $\Delta\mu_F$



$$\Delta\mu_F = \Delta\mu_0 + N\Delta\mu$$

$$p(N) \rightarrow p(\Delta\mu_F)$$

- Explicit coarse-graining over proofreading pathways

  → effective specificity $a(x, L)$
  → effective per-monomer energy $\Delta\mu_F(L)$

- Weak proofreading limit (single-state polymerase backtracking)

  → increase of the specificity $a_{KP} > a$
  → decrease of the error fraction $(x_a)_{KP} < x_a$