Thermodynamic inference:

Concepts and applications

Udo Seifert

II. Institut für Theoretische Physik, Universität Stuttgart

review: U.S., Annu. Rev. Condens. Matter Phys. **10**, 171 (2019).

- Intro: Stochastic th'dynamics for systems with discrete meso-states

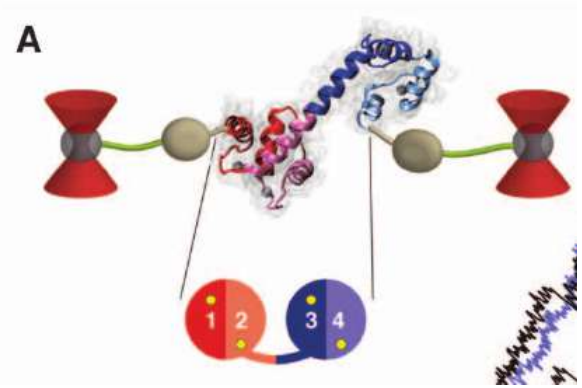
- Coarse-graining by observing

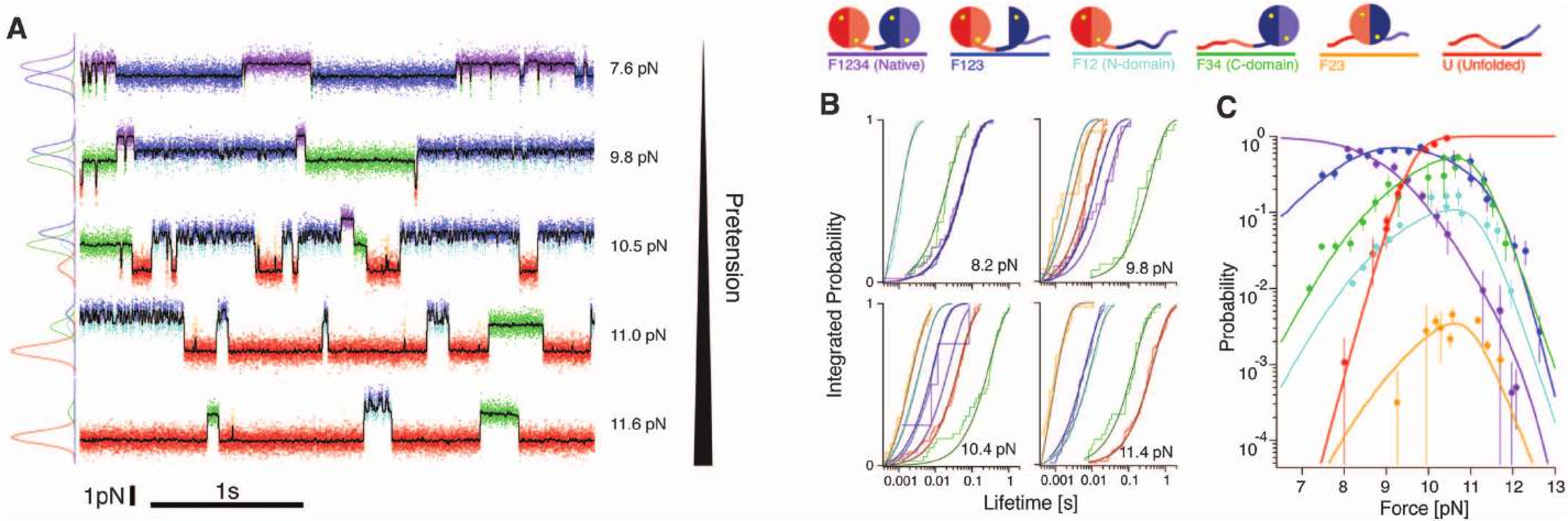
  - states
  - transitions
  - correlation functions

- Inference through

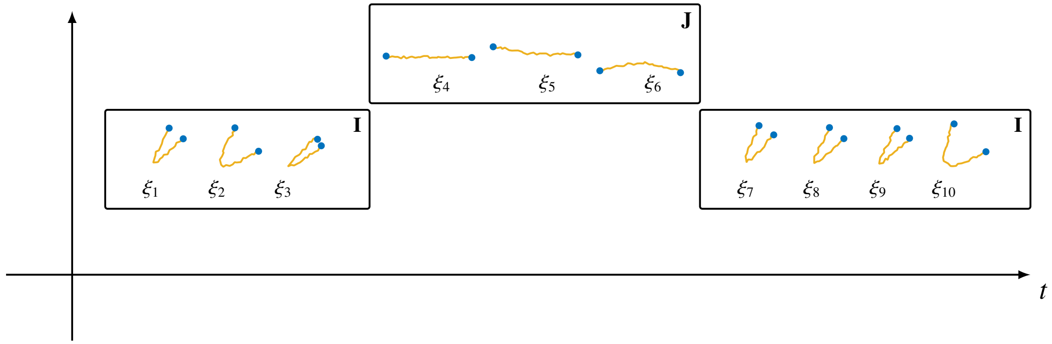
  - thermo'dyn unc' relation
  - waiting-time distributions
  - coherent oscillations

• Meso-states of a biomolecule: Calmodulin [J. Stigler et al, Science **334** 512 (2011)]





- Stochastic thermodynamics with discrete meso-states



  – crucial time-scale separation:
    * transitions between meso-states are slow
    * transitions between the micro-states belonging to one meso-state are fast

  – trajectory $i(t)$
  – master equation

$$\partial_t p_i(t) = \sum_J [p_j(t) k_{ji} - p_i(t) k_{ij}].$$

  – local detailed balance condition on the rates $\{k_{ij}\}$

$\Rightarrow$ $$\boxed{k_{ij}/k_{ji} = p_j^e/p_i^e = \tau_j^e/\tau_i^e = \exp(-\beta\Delta_{ij}F) = \exp(-\beta\Delta_{ij}E + \Delta_{ij}S)}$$

  – th'dyn potentials of meso-states operationally accessible from traj' data

- Driven systems
  - entropy production as quantitive measure of broken time reversal symmetry

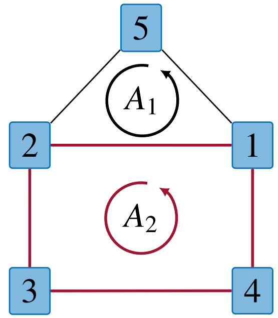
$$i(t) \to \tilde{i}(t) \equiv i(T-t) \qquad \boxed{\Delta s^{\text{tot}}[i(t)] \equiv \ln\{p[i(t)]/\tilde{p}[\tilde{i}(t)]\}}$$

  - in a NESS: three representations

$$\sigma = \langle \ln[p[I(t)]/p[\tilde{I}(t)]]\rangle/T = \sum_{ij} p_i^s k_{ij} \ln[p_i^s k_{ij}/p_j^s k_{ji}] \qquad \text{links}$$

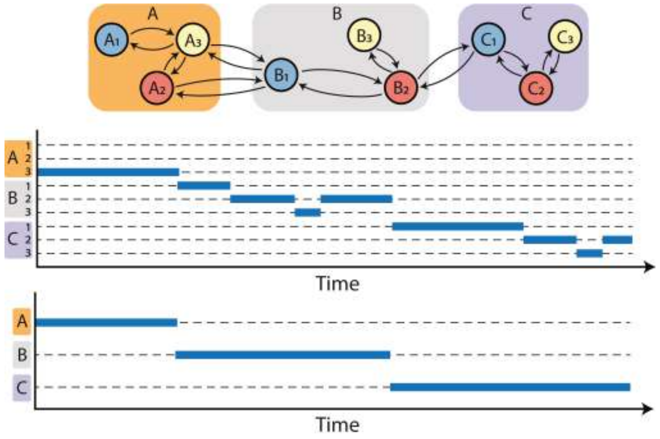
$$= \sum_c j_c A_c \qquad \text{cycles}$$

$$= \sum_K j_K A_K \qquad \text{physical driving "fields"}$$



  - cycle affinities $A_c \equiv \sum_{i<j\in c} \ln(k_{ij}/k_{ji})$
  - physical affinities $A_K = f, \Delta\mu, ...$ with $A_K = \sum_{i<j} d_{ij}^K \ln(k_{ij}/k_{ji})$

- Inference from coarse-graining states



Skinner and Dunkel, PNAS 2021

  - lower bound on entropy production

$$\sigma = (1/T) \ln[p(i(t))/\tilde{p}(\tilde{i}(t))]\rangle \geq \sigma_{\text{cgr}} \equiv (1/T)\langle \ln[p(I(t))/\tilde{p}(\tilde{I}(t))]\rangle$$

Kawai, Parrondo, van den Broeck PRL 2007, ...

  - from $p_I, \langle \nu_{IJ} \rangle \to K_{IJ} \equiv \langle \nu_{IJ} \rangle / p_I \to \sigma_1 \equiv \sum_{IJ} p_I K_{IJ} \ln(p_I K_{IJ} / p_j K_{JI}) \leq \sigma_{\text{cgr}}$ [Esposito, PRE 2012]

  - higher statistics $\langle \nu_{IJK} \rangle$, best Markov model leads to lower bound $\sigma_2 = ... \leq \sigma_{\text{cgr}}$
[Skinner and Dunkel, PNAS 2021]

  - discrete time-series with n-tupels $\{I(t_1), I(t_2), ...I(t_n)\}$ [Roldan and Parrondo, PRL 2010]

- Dual use of such inequalities

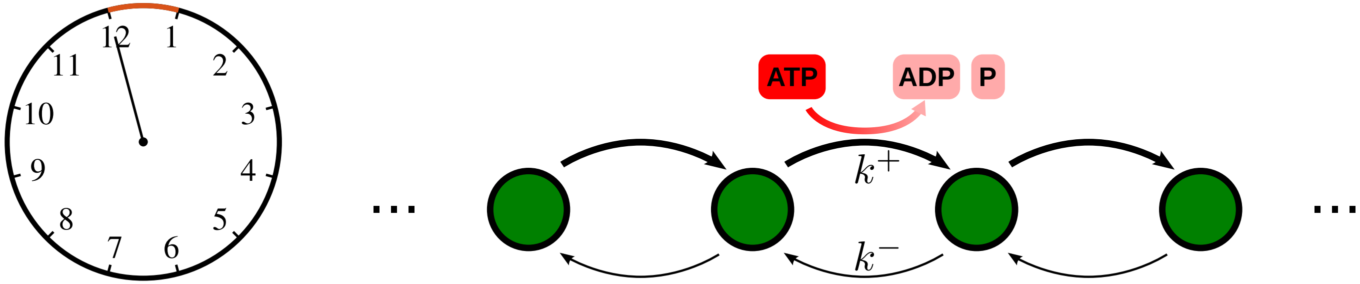
  $\sigma >$ ”something”

  – operational lower bound on ent’production

  – minimal thermodynamic cost of achieving ”something”

- Cost of running a simple clock: ARW

  [AC Barato and US, Phys. Rev. Lett. 114, 158101, 2015]



  – output $n(t)$ with $\langle n \rangle = Jt = (k^+ - k^-)t$

  – variance $\langle (n(t) - \langle n \rangle)^2 \rangle = 2Dt = (k^+ + k^-)t$

  – uncertainty $\epsilon^2 \equiv \text{var}/\text{output}^2 = 2D/J^2 t$

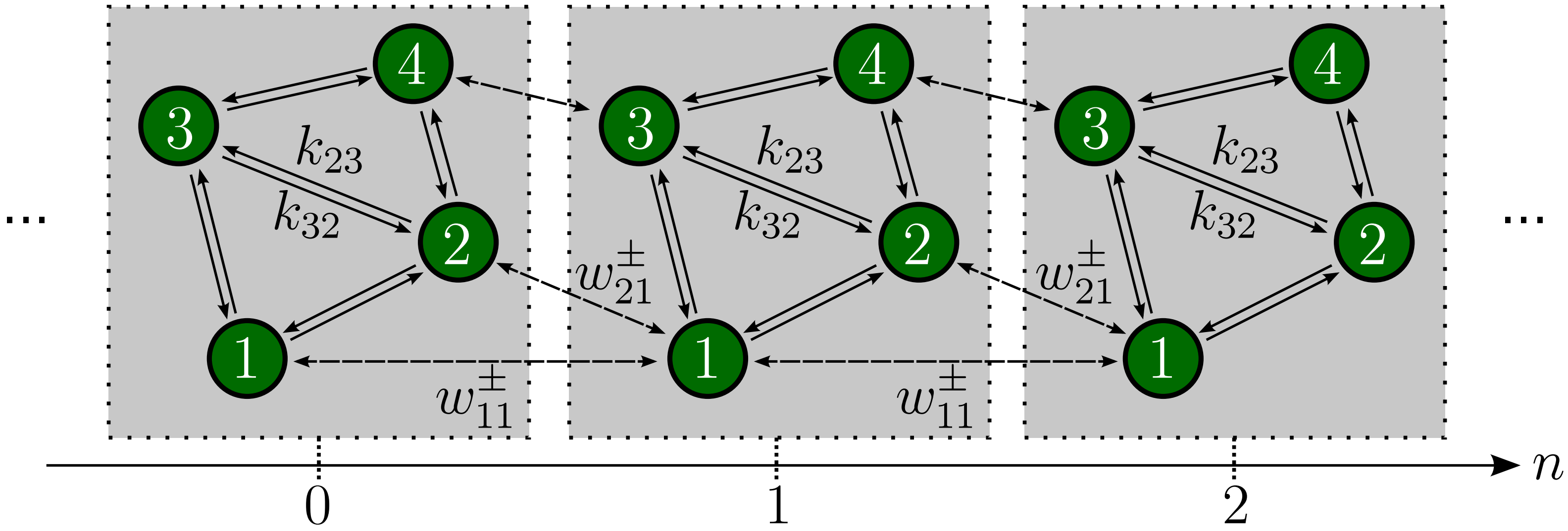
  – th'dyn cost $\mathcal{C} = \sigma t = (k^+ - k^-) \ln(k^+/k^-)t$ with $\sigma \equiv$ rate of entropy production

  – with affinity $\mathcal{A} = k_B T \ln(k^+/k^-) = \mu_{\text{ATP}} - \mu_{\text{ADP}} - \mu_{\text{P}}$

  – $\boxed{\mathcal{C}\epsilon^2 = 2\sigma D/J^2 = \mathcal{A} \coth[\mathcal{A}/2k_B T] \geq 2k_B T}$ independent of run time t

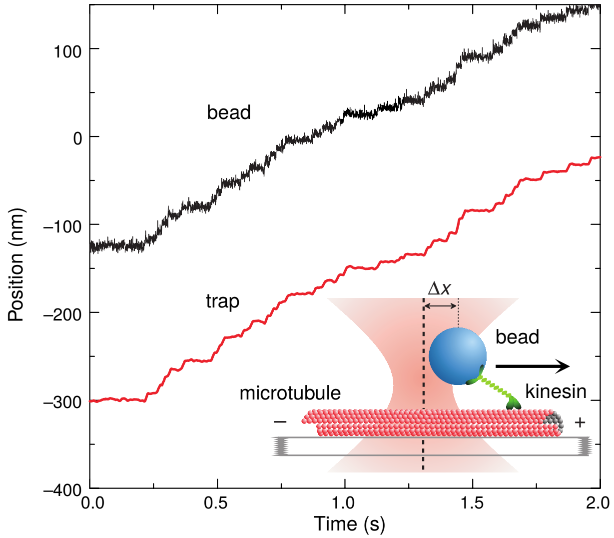
- Thermodynamic uncertainty relation holds for general multicyclic processes

  AC Barato and US, Phys. Rev. Lett. 114, 158101, 2015; proof by Gingrich et al, PRL 2016



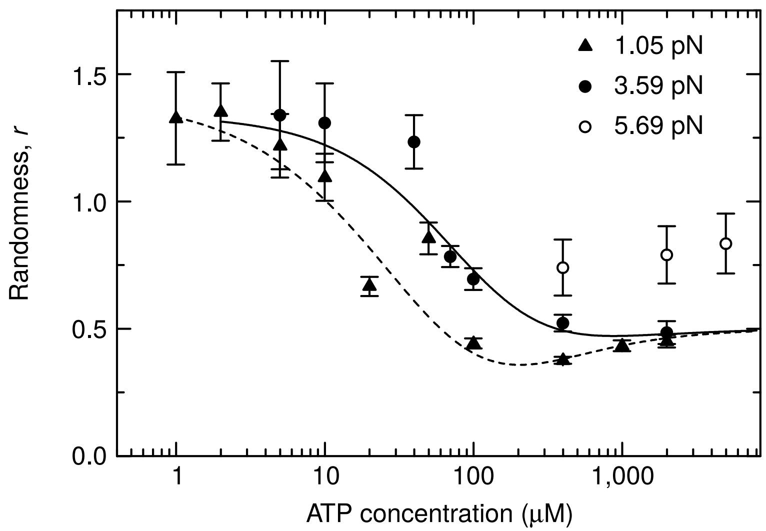
  - $\mathcal{C} \geq 2k_BT/\epsilon^2$ for any th'dyn consistent process at finite T
  - a precision of 1% costs at least 20.000 k_BT
  - inevitable, universal cost of temporal precision (within stationary Markov processes)
  - for any current $j = \sum_{ij} d_{ij} n_{ij}$ $\sigma \geq j^2/D_j$
  - violated for underdamped Langevin dynamics [P. Pietzonka, PRL 2022]

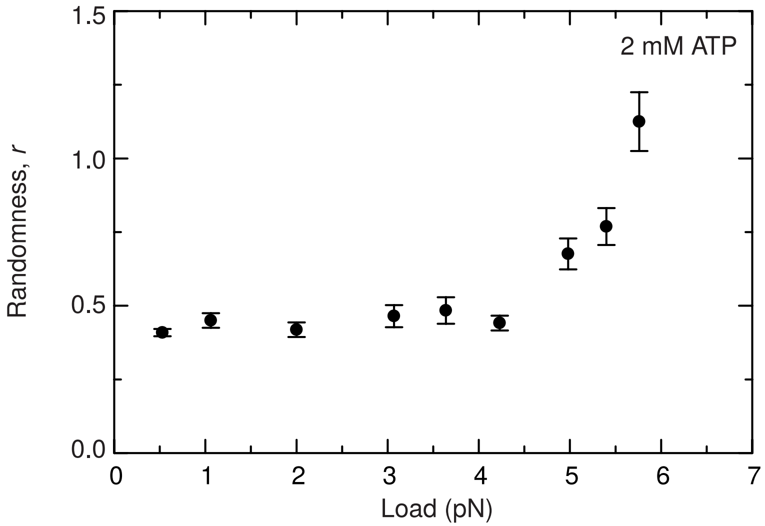
- Thermodynamic inference: Efficiency of a molecular motor



[Visscher et al, Nature, 1999]

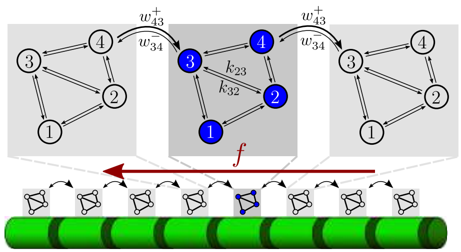
– experimental data on
  * velocity v
  * diffusion constant D
  * randomness parameter $r \equiv 2D/v\ell$





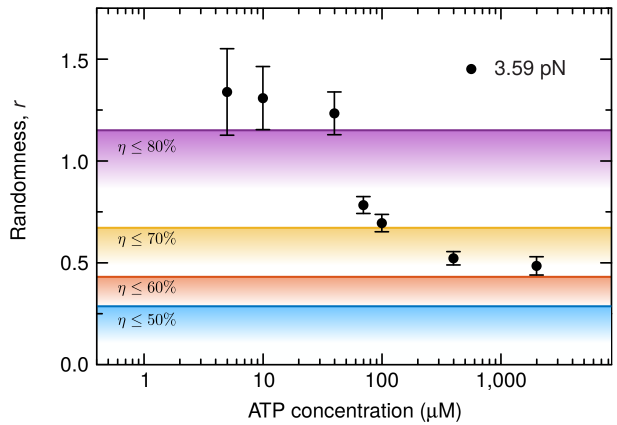
- Thermodynamic inference: Universal bound on the efficiency of molecular machines

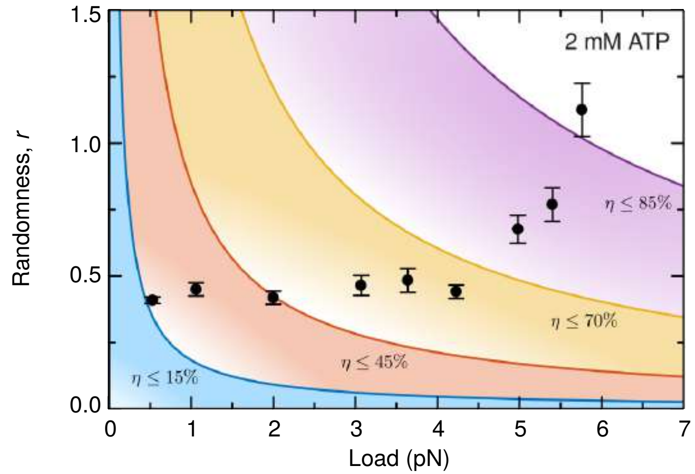
  [P. Pietzonka, AC Barato, U.S., J Stat Mech, 124004, 2016; U.S., Physica A 504, 176, 2018]



  - entropy production rate $\sigma = P^{\text{in}} - P^{\text{out}} =$ "chem energy" $- fv \geq v^2/D$
  - efficiency

$$\eta \equiv \frac{P^{\text{out}}}{P^{\text{in}}} = \frac{fv}{\text{unknown}} = \frac{fv}{fv+\sigma} \leq \frac{1}{1 + v k_B T/(Df)}$$



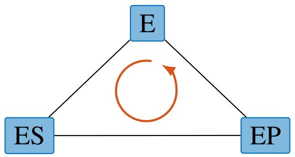


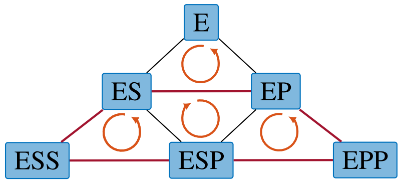
  - independent of the specific chemo-mechanical cycles and of $\Delta\mu$

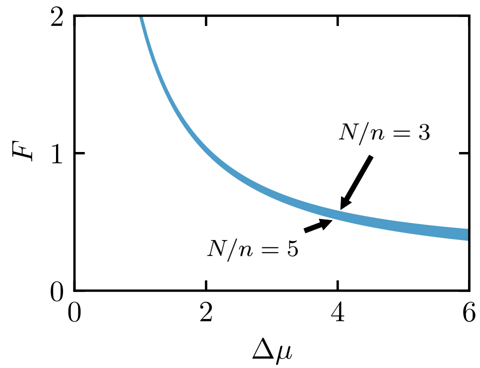
- Topology- and affinity-dependent bound in multicyclic networks

  [P Pietzonka, AC Barato and US, J. Phys. A 49 (2016) 34LT01, J. Phys. Chem. B, (2015) 119, 6555]

$$\boxed{2\sigma D/j^2 \geq (A/N)^* \coth[(A/N)^*/2]}$$ with $(A/N)^*$ min' of (A_c/N_c) over all cycles







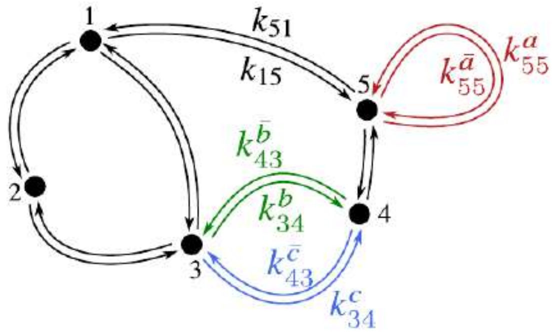
  - enzyme E transforms substrate S into product P using hydrolysis of one ATP with $\Delta\mu$

  - bound on Fano factor $F \equiv 2D/j^s \geq (n/N)^* \coth[(\Delta\mu/2)(n/N)^*]$

  - simple MM scheme (E$\rightarrow$ ES$\rightarrow$EP$\rightarrow$E, i.e., $N = 3, n = 1$) $\Rightarrow$ $F \geq [\coth(\Delta\mu/6)]/3$.

- Inference from observing states for time-dependent driving

  [T Koyuk and U.S., Phys. Rev. Lett. 125, 260604, 2020]



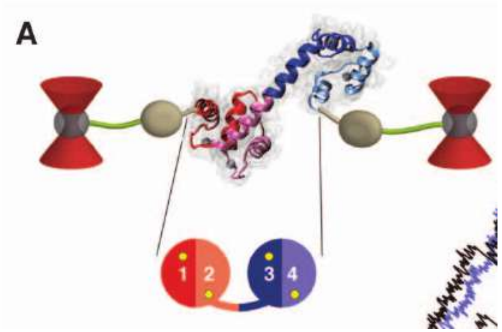
  – network with rates $k_{ij}(\lambda)$ that depend on a driving protocol $\lambda = \lambda(t)$
  – protocol $\lambda(t) = \lambda(vt)$ depends on an experimentally controlable speed parameter v
  – system is driven for a total (observation) time $t = \mathcal{T}$

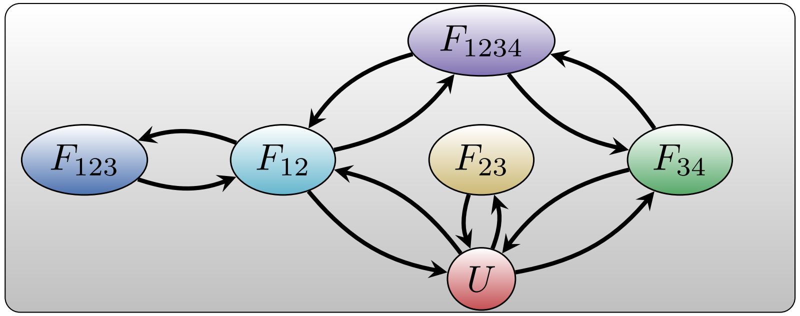
  – observable $A_i(\mathcal{T}, v) \equiv \tau_i/\mathcal{T}$ total time spent in state i

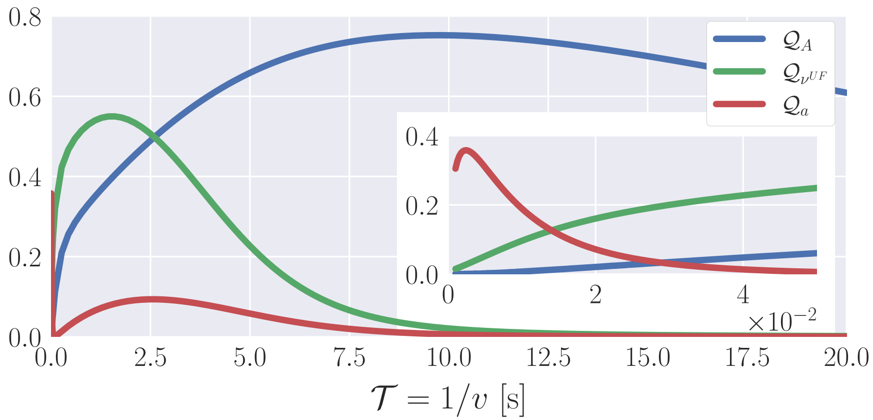
$$\{[\mathcal{T}\partial_{\mathcal{T}} - v\partial_v]\langle A_i(\mathcal{T}, v)\rangle\}^2/D_{A_i}(\mathcal{T}, v) \leq \sigma(\mathcal{T}, v)$$

  – same for observable $a_i(\mathcal{T}, v) \equiv \delta_{n(\mathcal{T})i}$ state at final time

- Real world example: Calmodulin unfolding [J. Stigler et al, Science **334** 512 (2011)]







– force ramp $f(t) = f_0 + vt$

– observable $a(\mathcal{T}, v) \equiv \delta_{n(\mathcal{T})12}$ at final time [red]

– observable $A(\mathcal{T}, v) \equiv \tau_U/\mathcal{T}$ total time spent in unfolded state [blue]

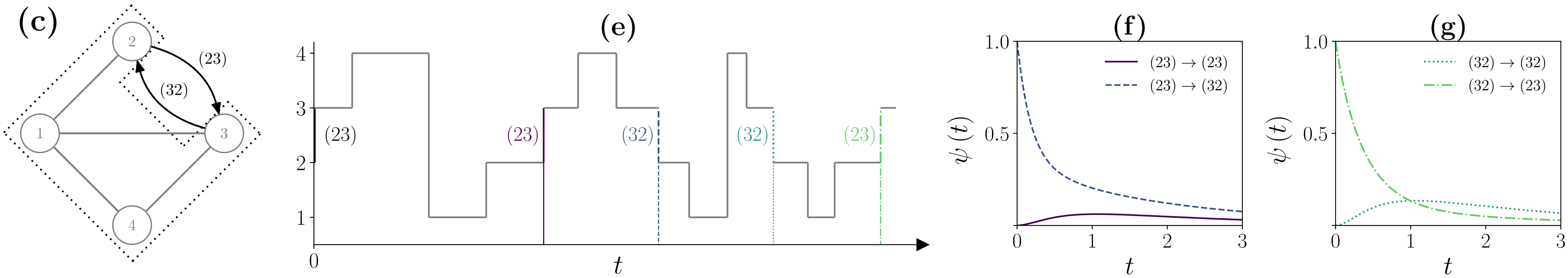
– current $\nu = [n_{UF}(\mathcal{T}, v) - n_{FU}(\mathcal{T}, v)]/\mathcal{T}$ [green]

– infer $40 \simeq 80\%$ of entropy production with no model input whatsoever

- NESS: Beyond the TUR (with the same data): Inference from waiting-time distributions

  Martinez et al, Nat Comm 2019; Skinner and Dunkel, PRL 2021; Ehrich, J Stat Mech 2021

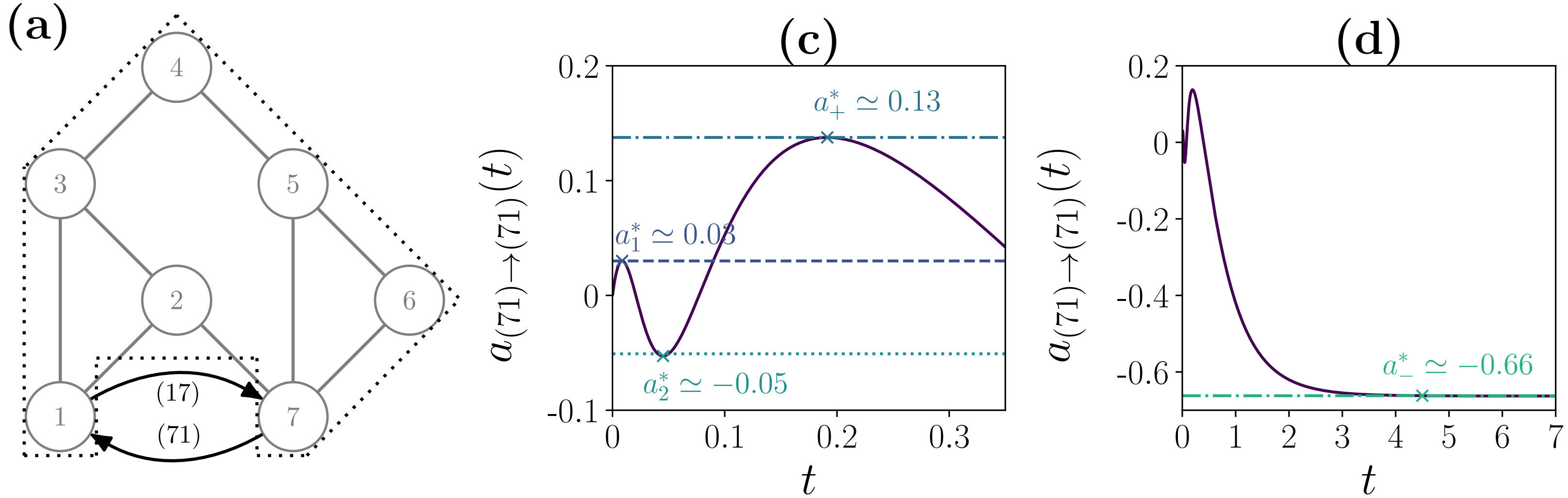
  Harunari, Dutta, Polettini, Roldan, arXiv:2203.07427; van der Meer, Ertel, U.S., arxiv:2203.12020



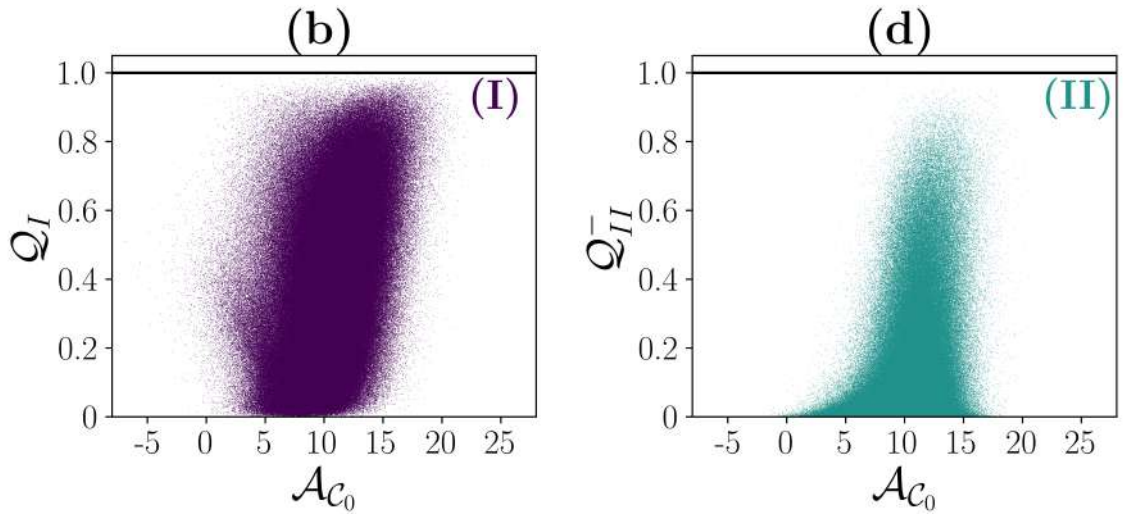
  - waiting time between two consecutive transition, e.g., $\psi_{23\to 23}(t) \equiv \psi_{++}(t)$
  - ratio of waiting times $a(t) \equiv \ln \psi_{++}(t)/\psi_{--}(t)$

- Unicyclic network
  - $a(t) = a = A_c$ $[= \ln p(+|+)/p(-|-)]$
  - one link is enough to get the affinity, better than TUR

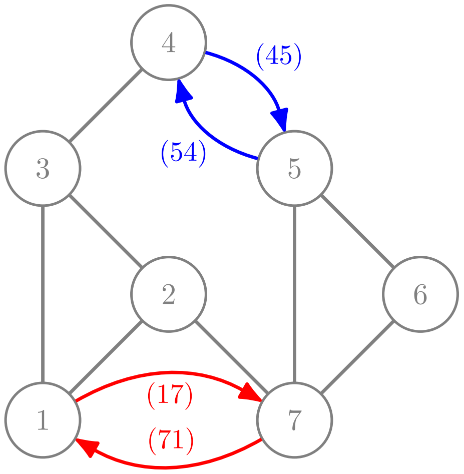
- One observed link in a multicycle system



- $a(t \to 0) = A_c$(shortest cycle)
- $\max a(t) \leq \max A_c$
- $\min a(t) \geq \min A_c$



- Several observed links in a multicyclic system



  - mean rate of an I-transition ν_I
  - wt-distributions for consecutive transitions J after I $\psi_{IJ}(t)$

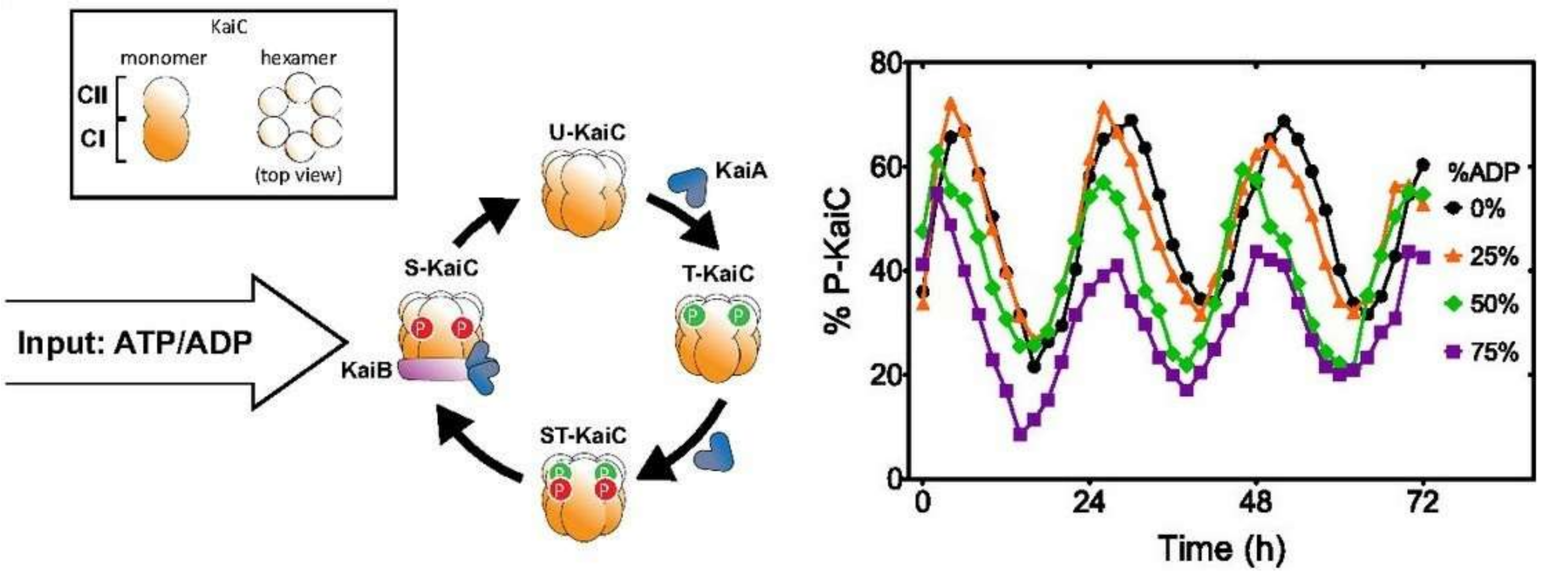
  - lower bound on ent'production

$$\boxed{\sigma \geq \sum_{IJ} \int_0^\infty dt \nu_I \psi_{IJ}(t) \ln[\psi_{IJ}(t)/\psi_{\tilde{J}\tilde{I}}(t)]}$$

  - equality if all $\psi_{IJ}(t)/\psi_{\tilde{J}\tilde{I}}(t)$ are time-independent

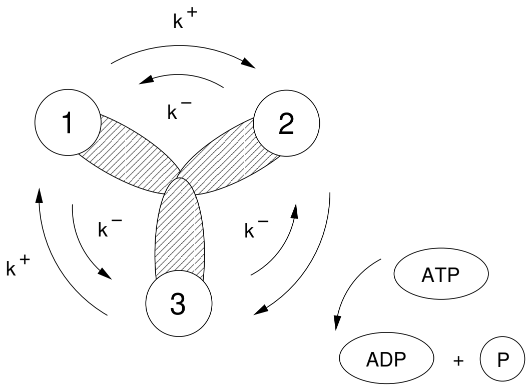
- Cost of/inference from coherent biochemical oscillators

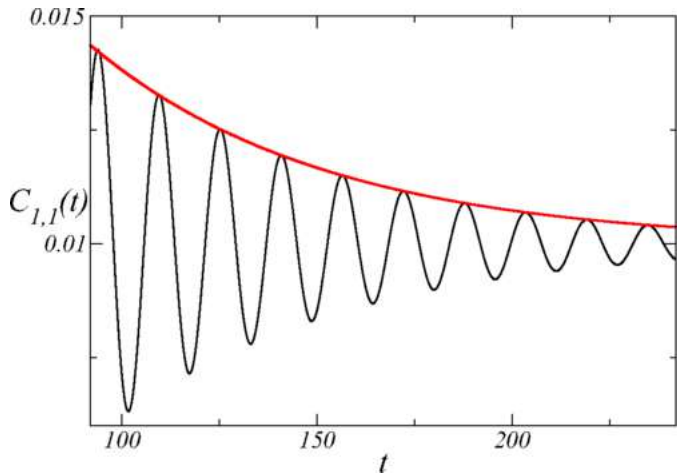
  - Kai-system circadian clock reconstructed from cyanobacterium



[C. Phong *et al.* PNAS **110**, 1124 (2013)]

- Universal minimal cost [AC Barato, L Oberreiter; Phys Rev E 95, 062409 (2017), arxiv 2121.01607]
  - unicycle with N equivalent states and driving affinity $\mathcal{A} = N \ln(k^+/k^-)$





$N = 100, \mathcal{A} = 200 \Rightarrow \mathcal{N} \simeq 4$

  - correlation function $C(1,t|1,0) = p_1^s + \sum_{j=2}^{N} c_j \exp[-\lambda_j t]$

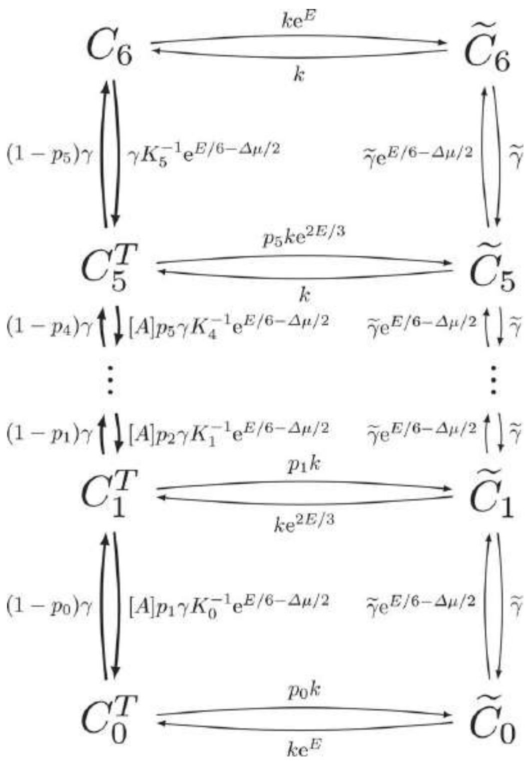
  - coherence lost after

$$\mathcal{N} = \frac{\Im\lambda_2}{2\pi\Re\lambda_2} = \frac{\tanh(\mathcal{A}/2N)}{2\pi\tan(\pi/N)}$$ oscillations

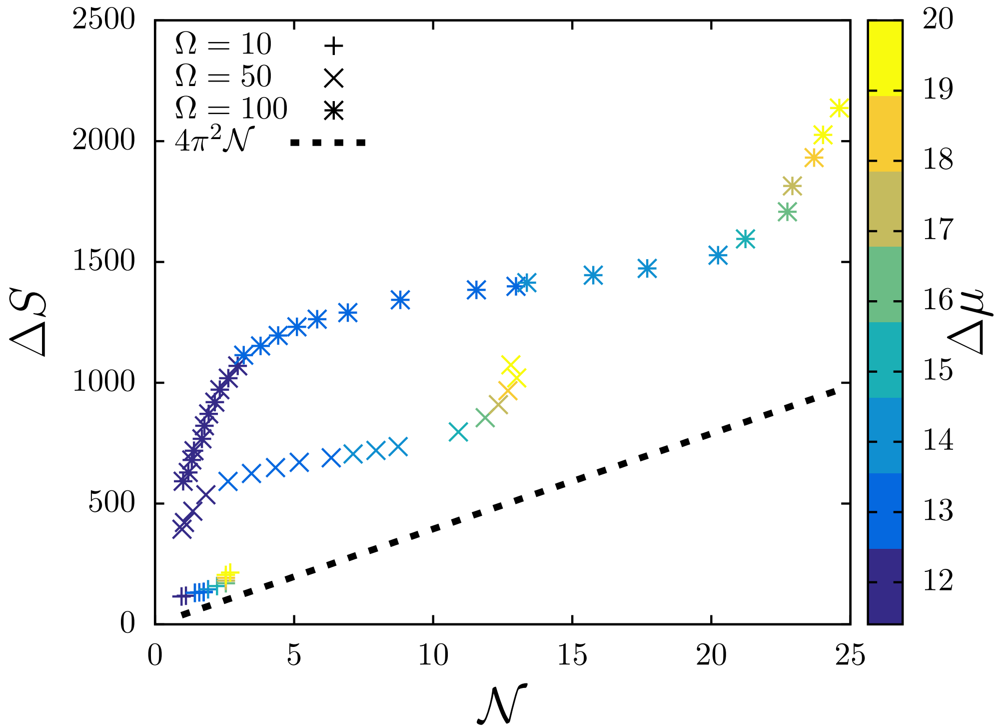
  - free energy cost/osci: $\Delta S = 2N\text{artanh}[2\pi\mathcal{N}\tan(\pi/N)] \geq 4\pi^2\mathcal{N}$
  - conjecture based on lots of numerics: $\boxed{\Delta S \geq 4\pi^2\mathcal{N}}$ in any (multicyclic) network

- Kai-C/Kai-A model

inspired by JS van Zon, DK Lubensky, PRH Altena and PR ten Wolde, PNAS 2007



active inactive



- Summary and acknowledgments
  - Inference through
    - thermodynamic uncertainty relation

      Andre Barato ($\rightarrow$ U Houston), Patrick Pietzonka ($\rightarrow$ MPI-PKS Dresden), Timur Koyuk

    - waiting-time distributions

      Benjamin Ertel, Jann van der Meer

    - coherent oscillations

      Andre Barato, Lukas Oberreiter