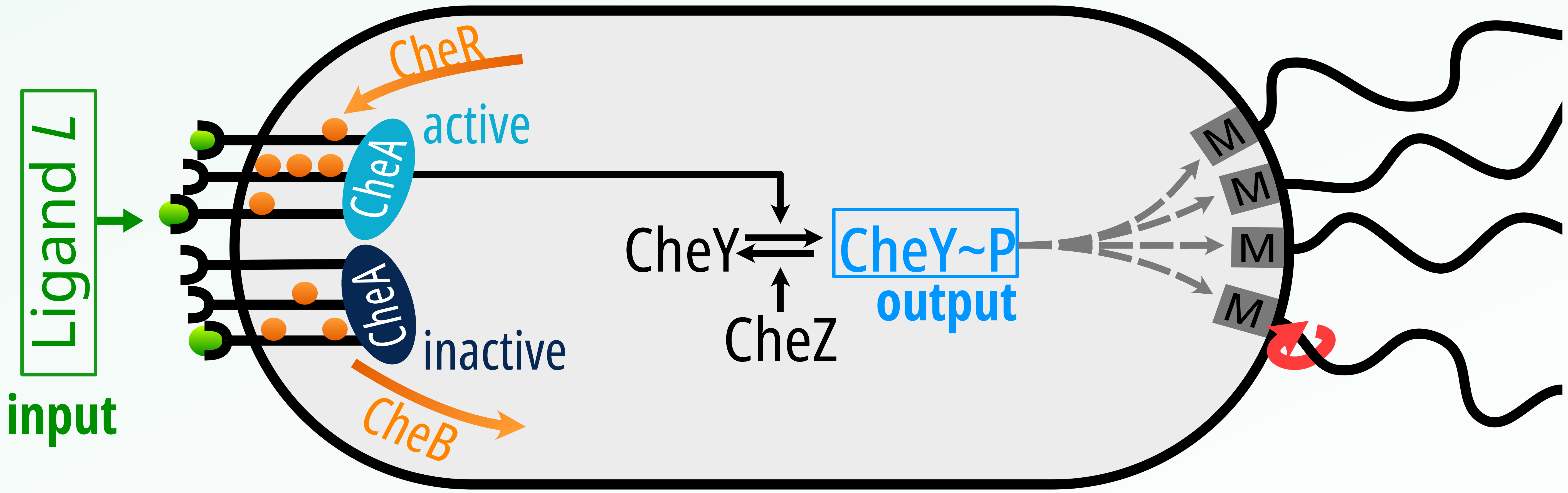
Computing the Information Flow Through Complex Systems

Manuel Reinhardt, Gašper Tkačik and Pieter Rein ten Wolde

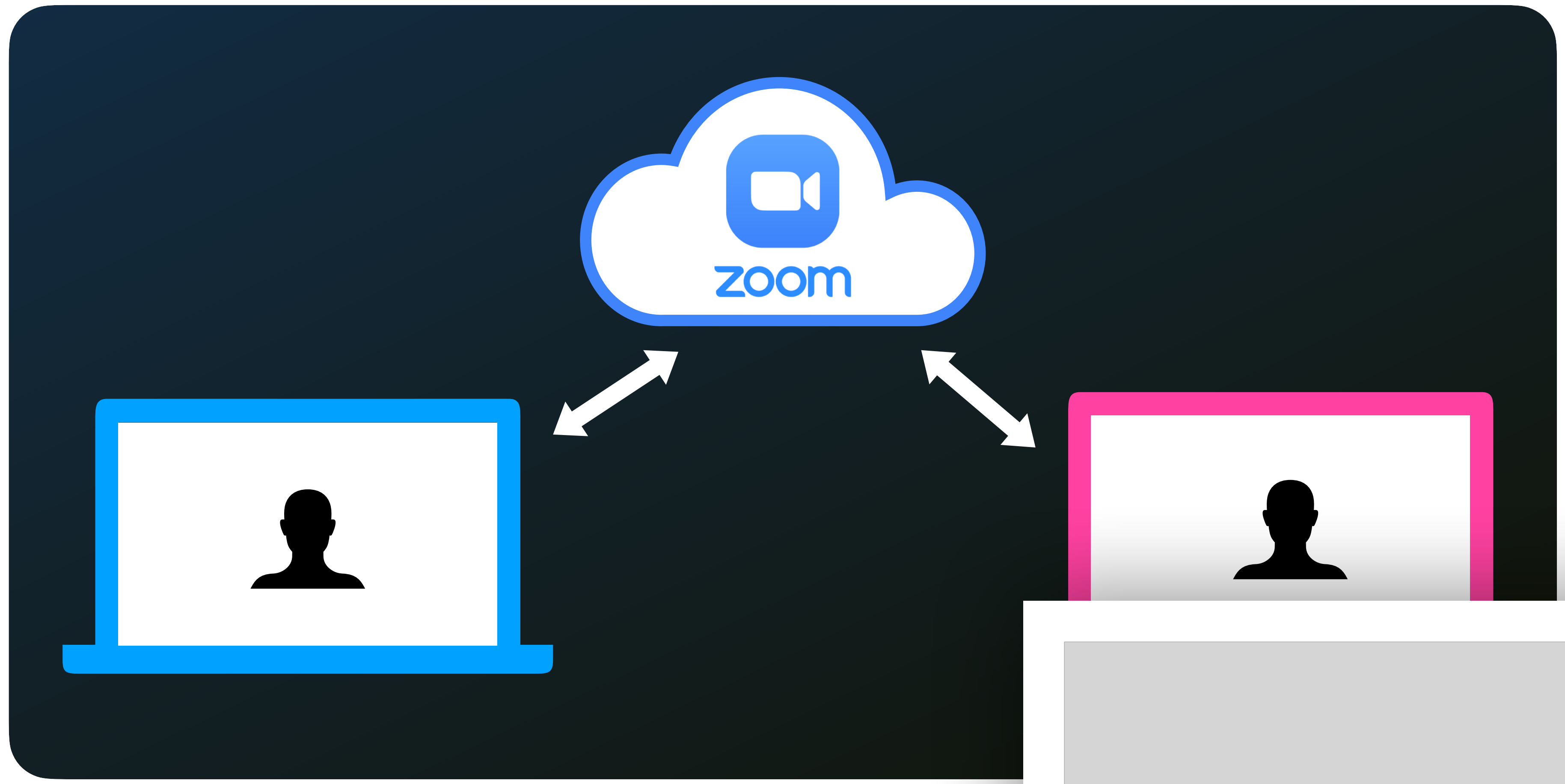


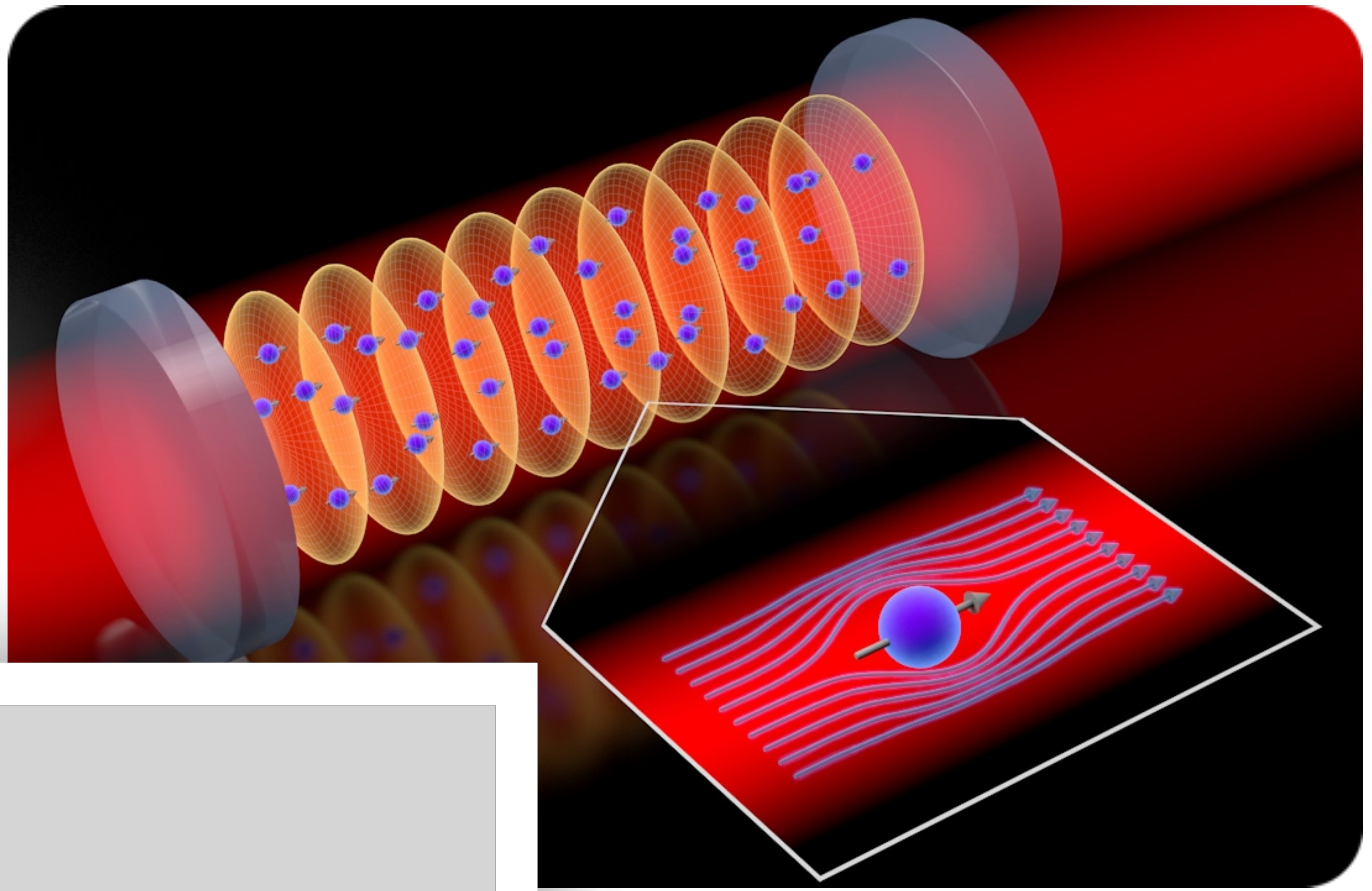
Information processing in bacterium *E. coli*

AMOLF physics of functional complex matter

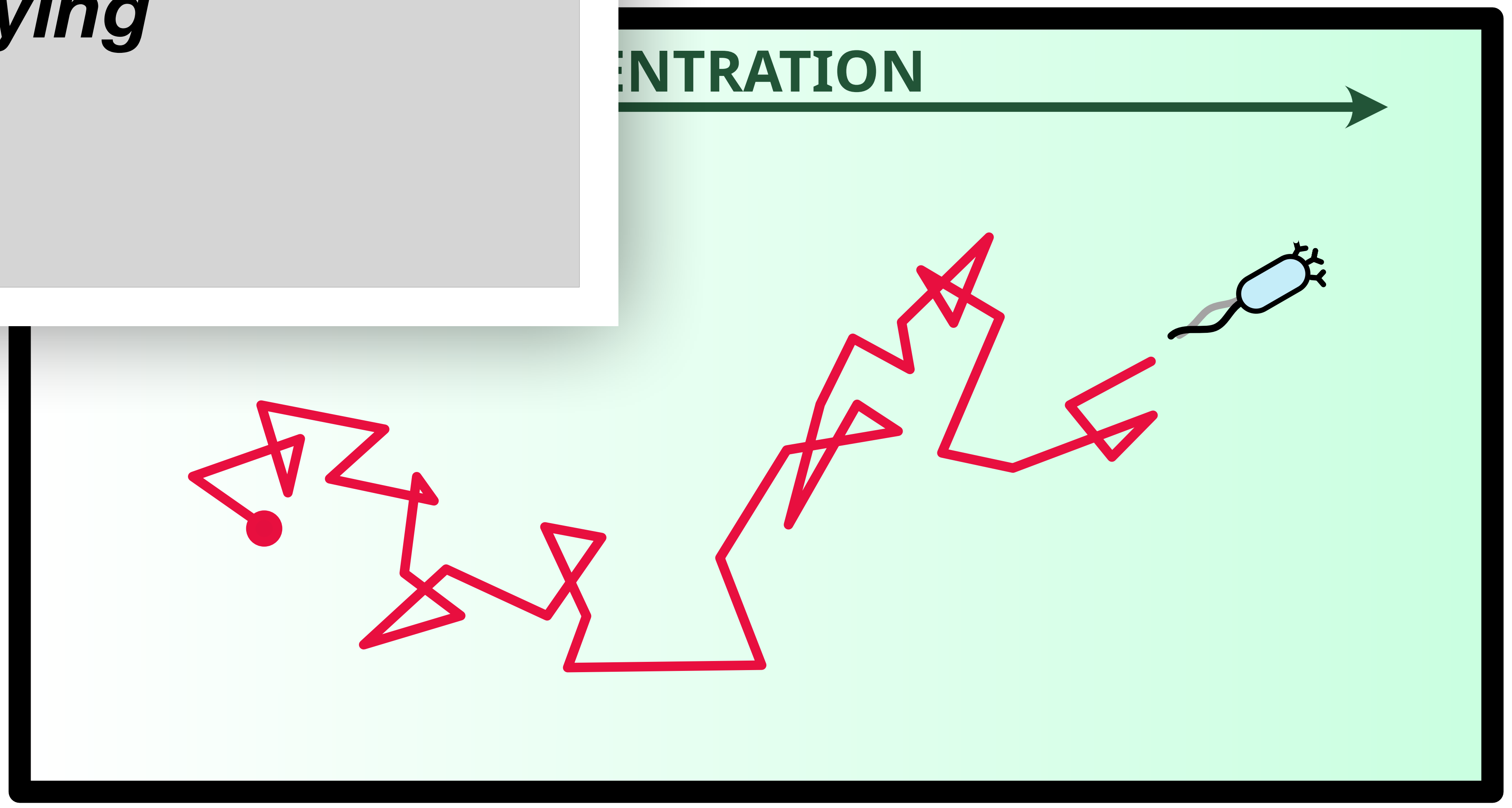
AMOLF Klein Colloquium, 07/02/2022

Information is all around us





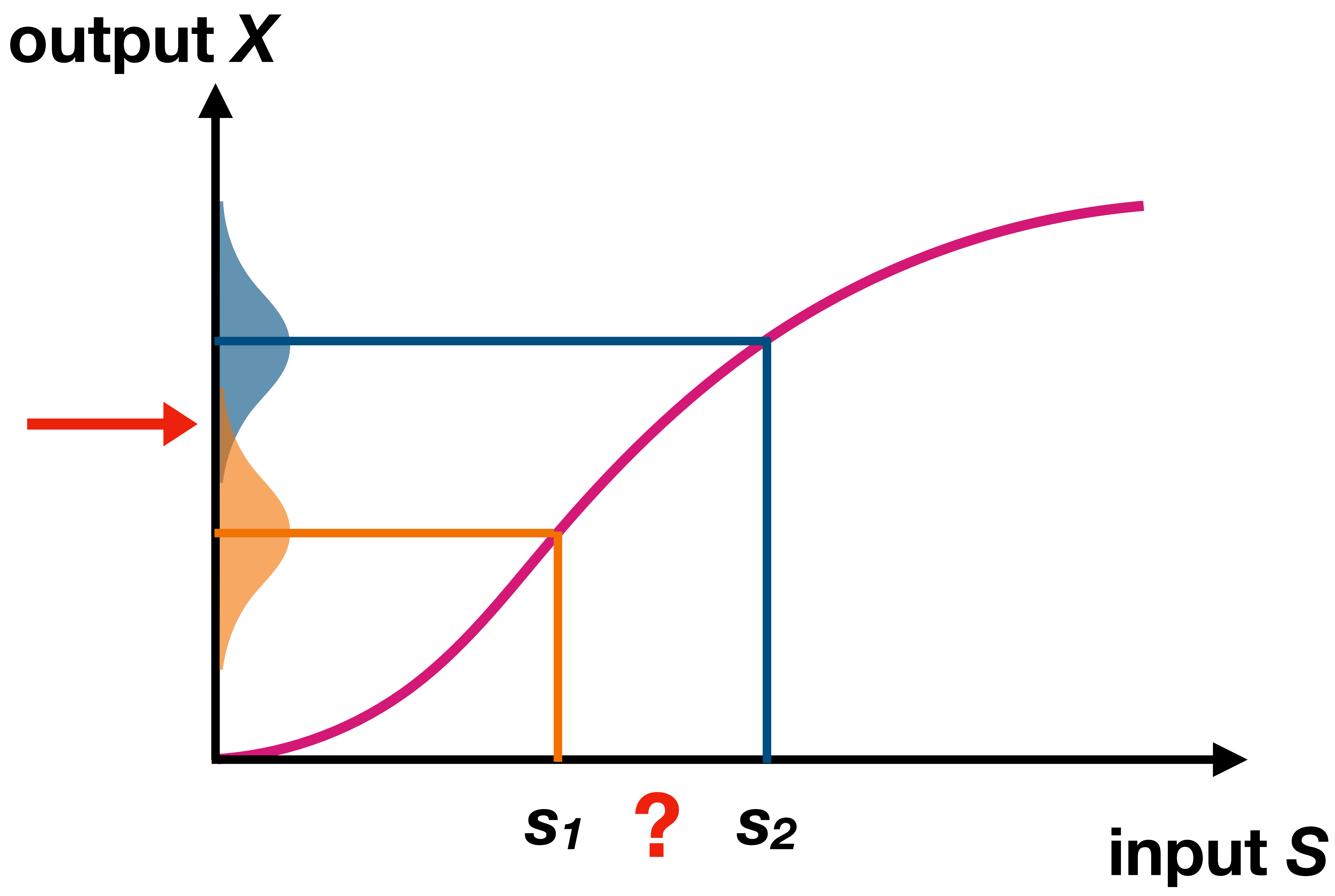
Understanding & Improving these Systems requires *Quantifying* Information



"The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point."

Claude E. Shannon (1948)

Input Signal *S* → Processing unit → Output Signal *X*



Mutual Information: I(*S*, *X*) = H(*S*) – H(*S*|*X*)

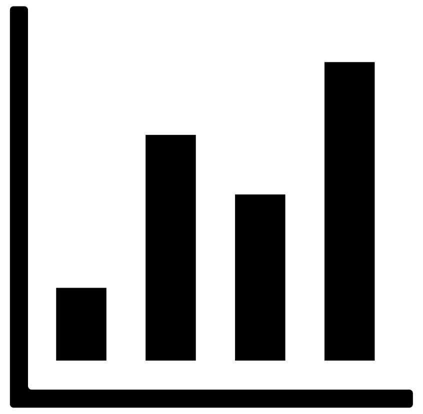
H(*S*) = $\langle -\log_2 P(s) \rangle$

→ Average input uncertainty *(Shannon entropy)*

H(*S*|*X*) = $\langle -\log_2 P(s\,|\,x) \rangle$

→ Average input uncertainty *after receiving output*

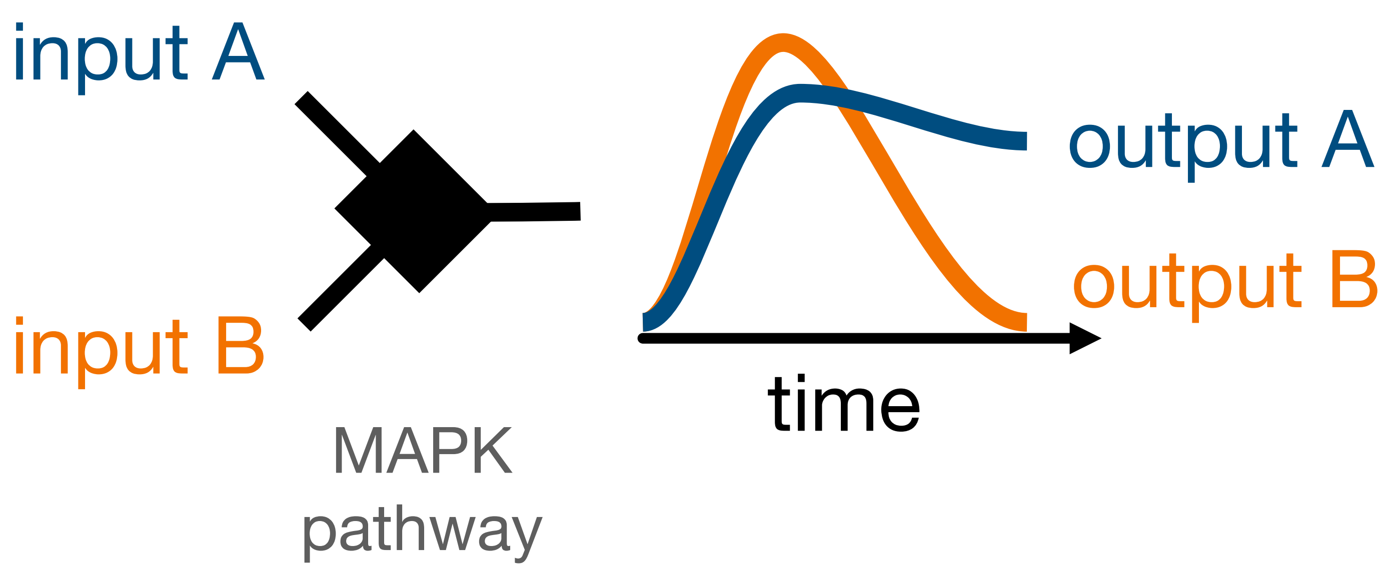
I(*S*, *X*) = I(*X*, *S*)



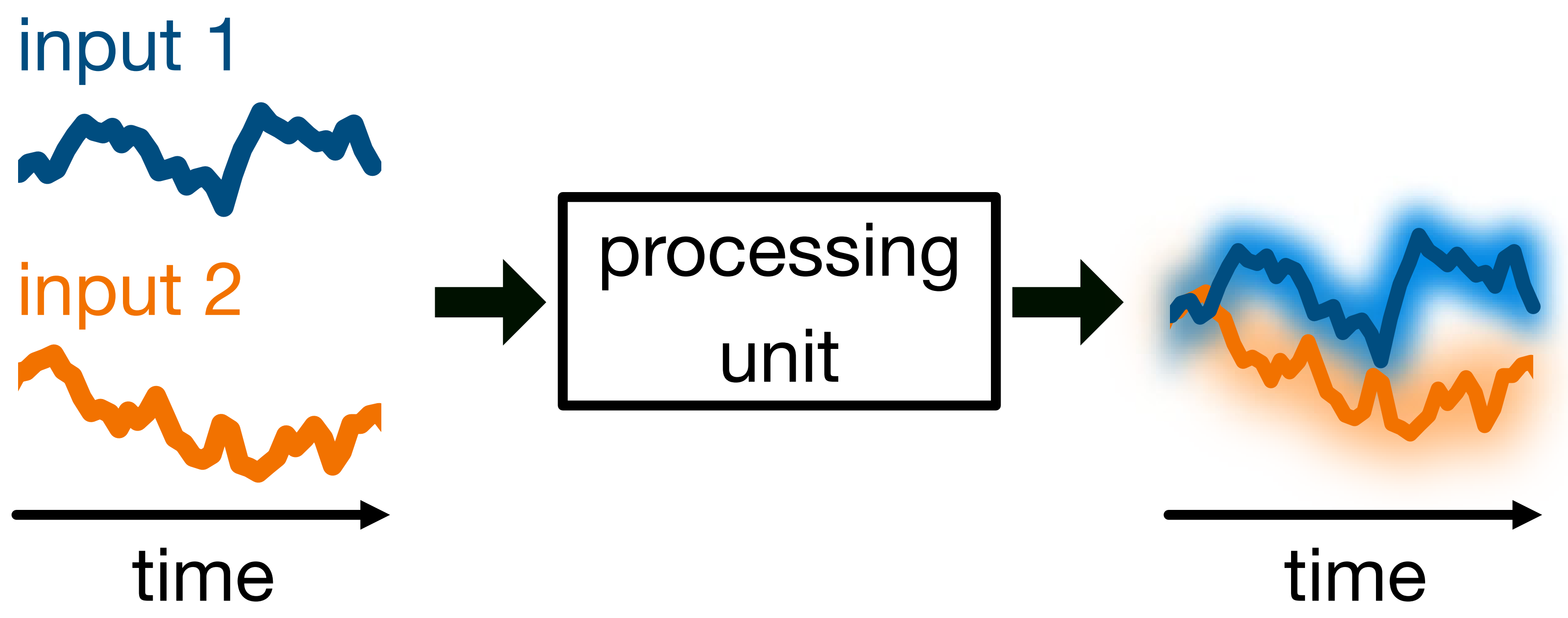
Conventionally, the mutual information can be computed using histograms.

Information is Encoded in *Trajectories*

Example from cellular signalling



General case

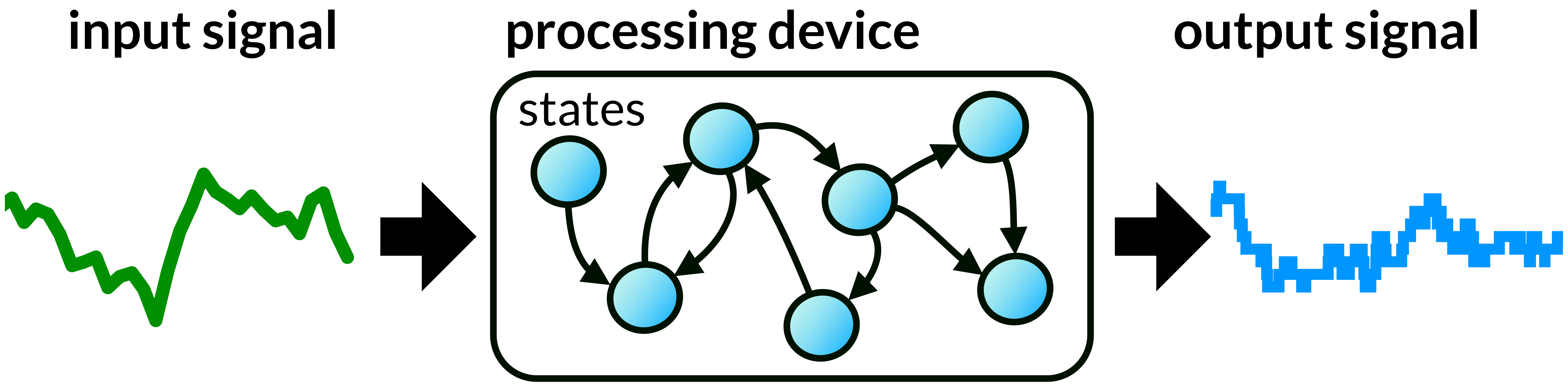


Information Rate $$R = \frac{1}{T}\,\mathrm{I}\left(\vec{S}_T, \vec{X}_T\right) \qquad [R] = \frac{\text{bits}}{s}$$

“Number of independent messages that can reliably be transmitted per unit time”

However, the conventional approach is not viable for trajectories!

Modeling a Complex System



Master Equation

$$\frac{\mathrm{d}\mathrm{P}_n(t)}{\mathrm{d}t} = \sum_k \mathrm{Q}_{nk}\mathrm{P}_k(t)$$

$\mathrm{P}_n(t)$: probability of being in state *n* at time *t*.

Solution

$$\mathrm{I}(\vec{S}, \vec{X}) = \mathrm{H}(\vec{X}) - \mathrm{H}(\vec{X} \mid \vec{S})$$

$$\mathrm{H}(\vec{X} \mid \vec{S}) = \left\langle -\log_2 \mathrm{P}\left(\vec{x} \mid \vec{s}\right) \right\rangle$$

Compute the average using **kinetic Monte Carlo algorithms**, based on the master equation.

$\mathrm{P}(\vec{x} \mid \vec{s})$ can be computed on-the-fly via **Master Equation**!

Solution, Part 2

$$\mathrm{I}(\vec{S}, \vec{X}) = \mathrm{H}(\vec{X}) - \mathrm{H}(\vec{X} \mid \vec{S})$$

$$\mathrm{H}(\vec{X}) = -\left\langle \log_2 \mathrm{P}\left(\vec{x}\right) \right\rangle$$

Requires the evaluation of a *marginalisation integral*

$$\mathrm{P}\left(\vec{x}\right) = \int_{\vec{s}} \mathrm{d}\vec{s}\ \mathrm{P}(\vec{s})\,\mathrm{P}(\vec{x} \mid \vec{s})$$

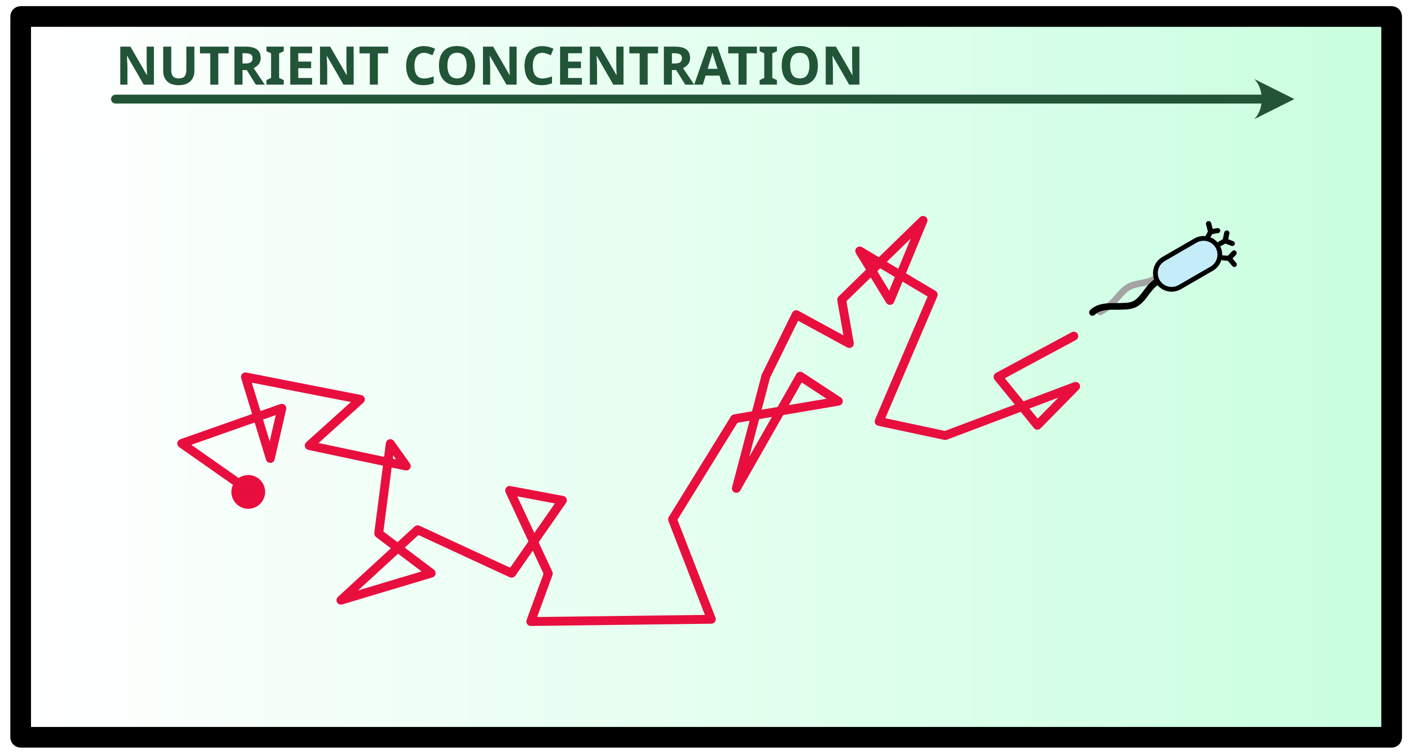
$$Z(\vec{x}) = \int_{\vec{s}} \mathrm{d}\vec{s}\ e^{-\beta U(\vec{s}, \vec{x})}$$

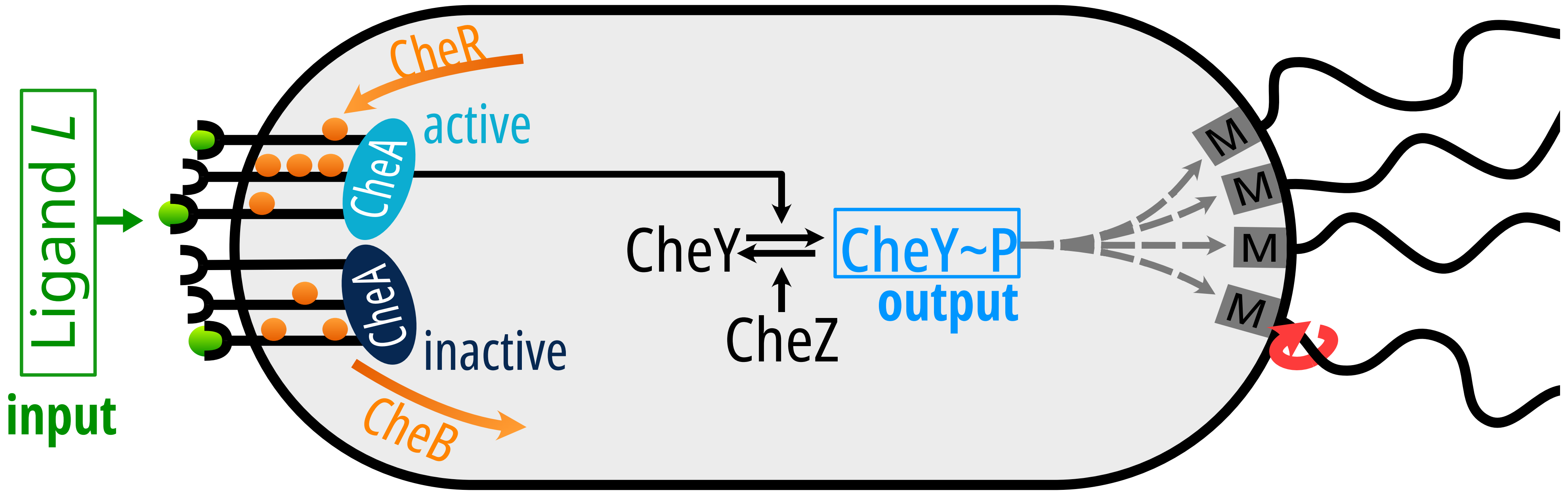
Inspired by classical techniques from Soft Condensed Matter Physics:

- Thermodynamic Integration
- Rosenbluth Sampling (PERM)
- Forward Flux Sampling (FFS)

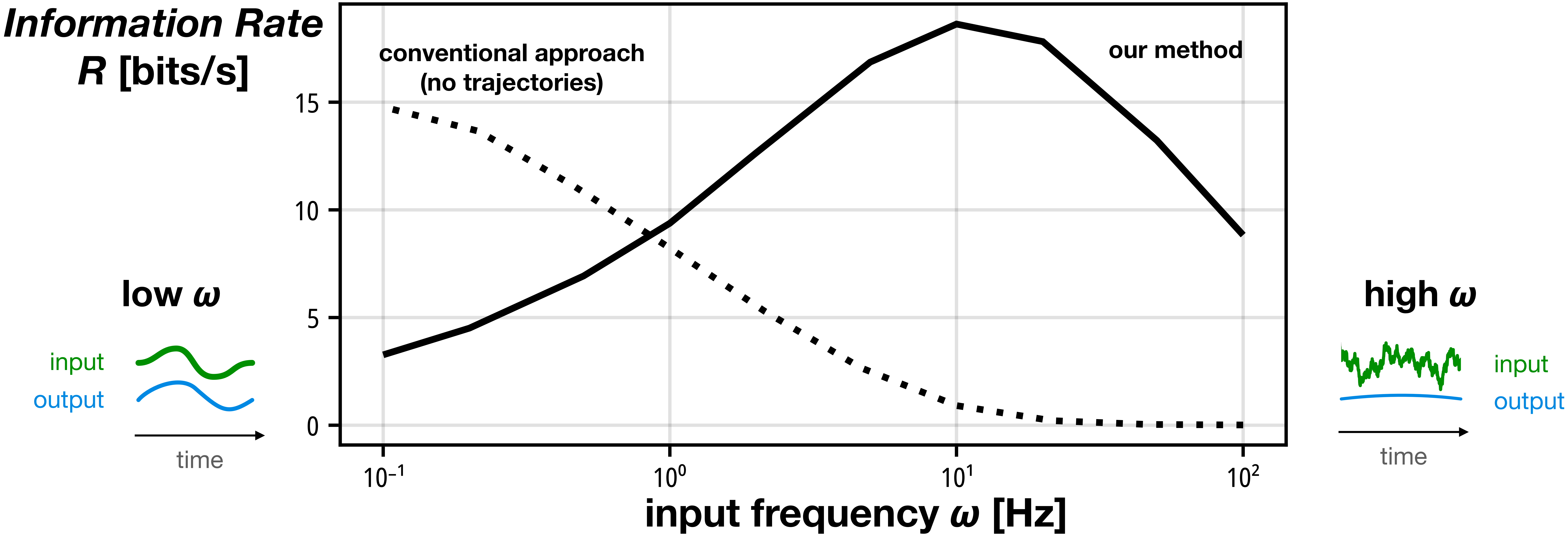
We can efficiently quantify Information using computational methods from polymer physics!

Example: Bacterium *E. Coli*

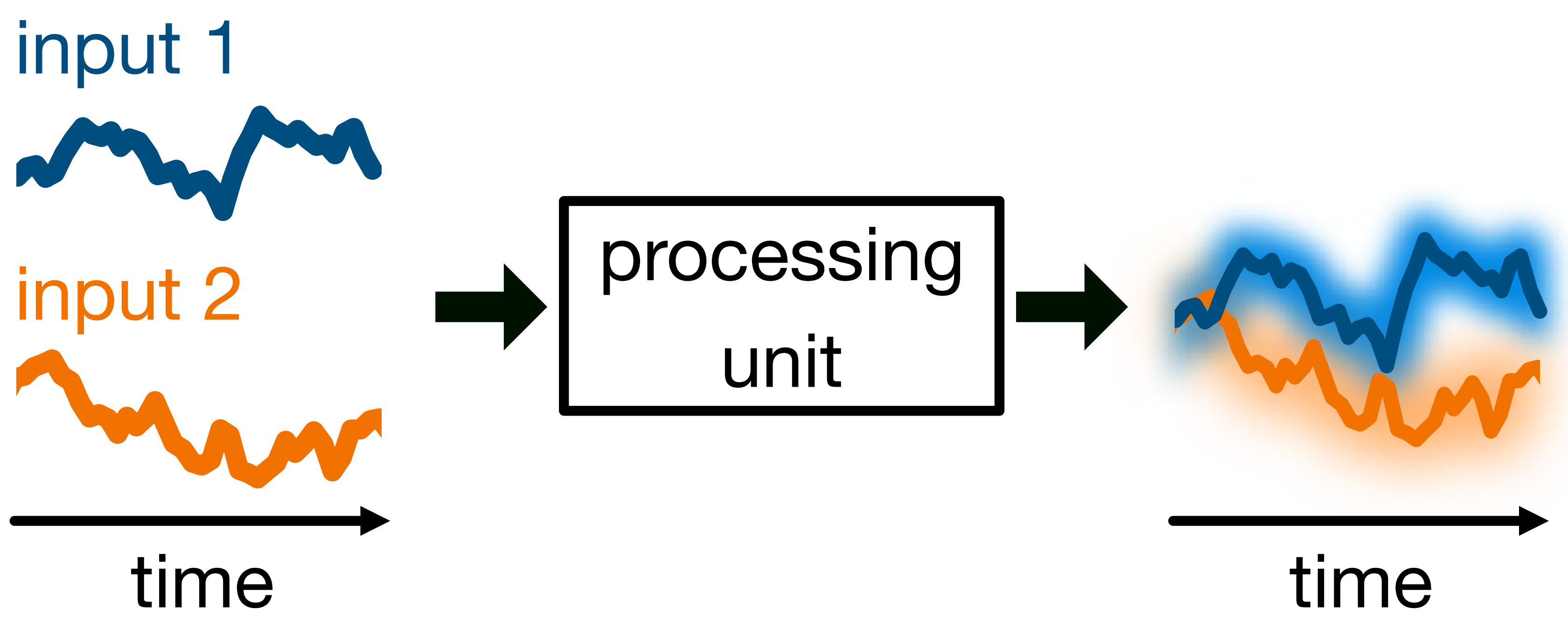




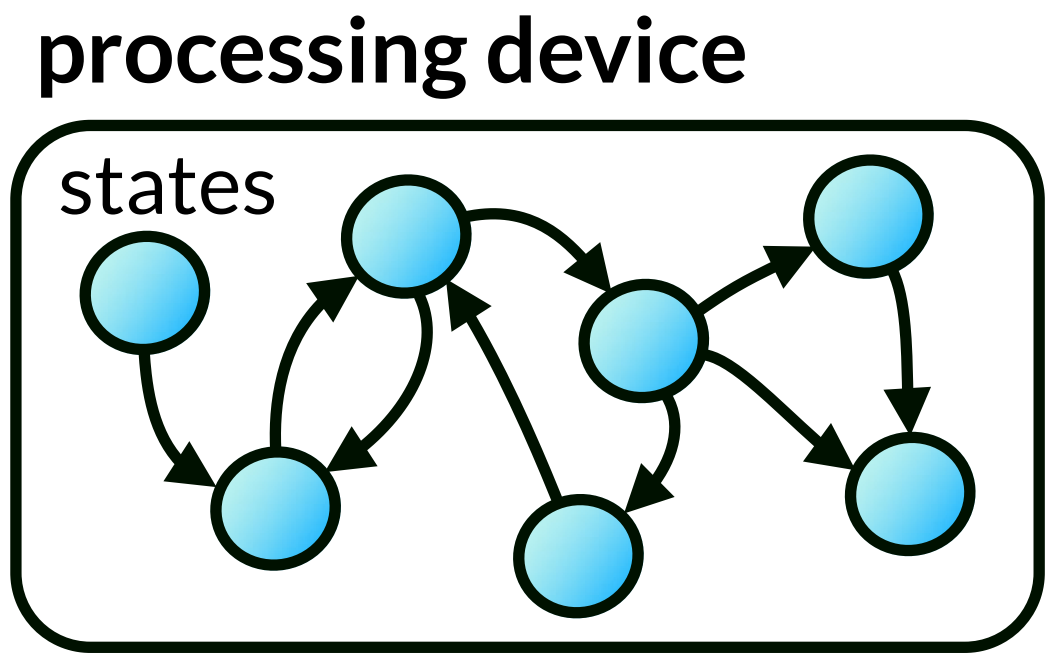
Complex biochemical network of 171 Reactions



Summary

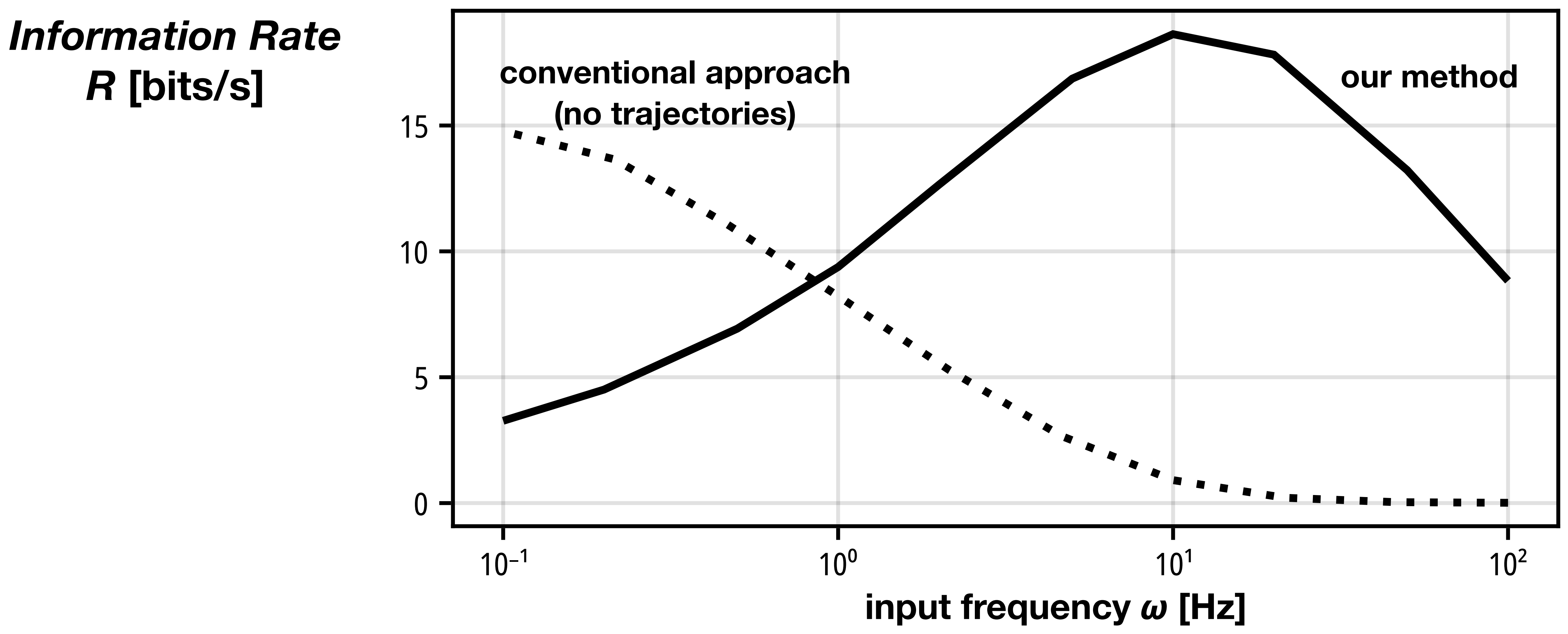


- The information rate is the essential quantity to describe information flow
- We have developed a *completely generic* method to compute it exactly *for the first time*



Next Step:

Going beyond the master equation



Thanks

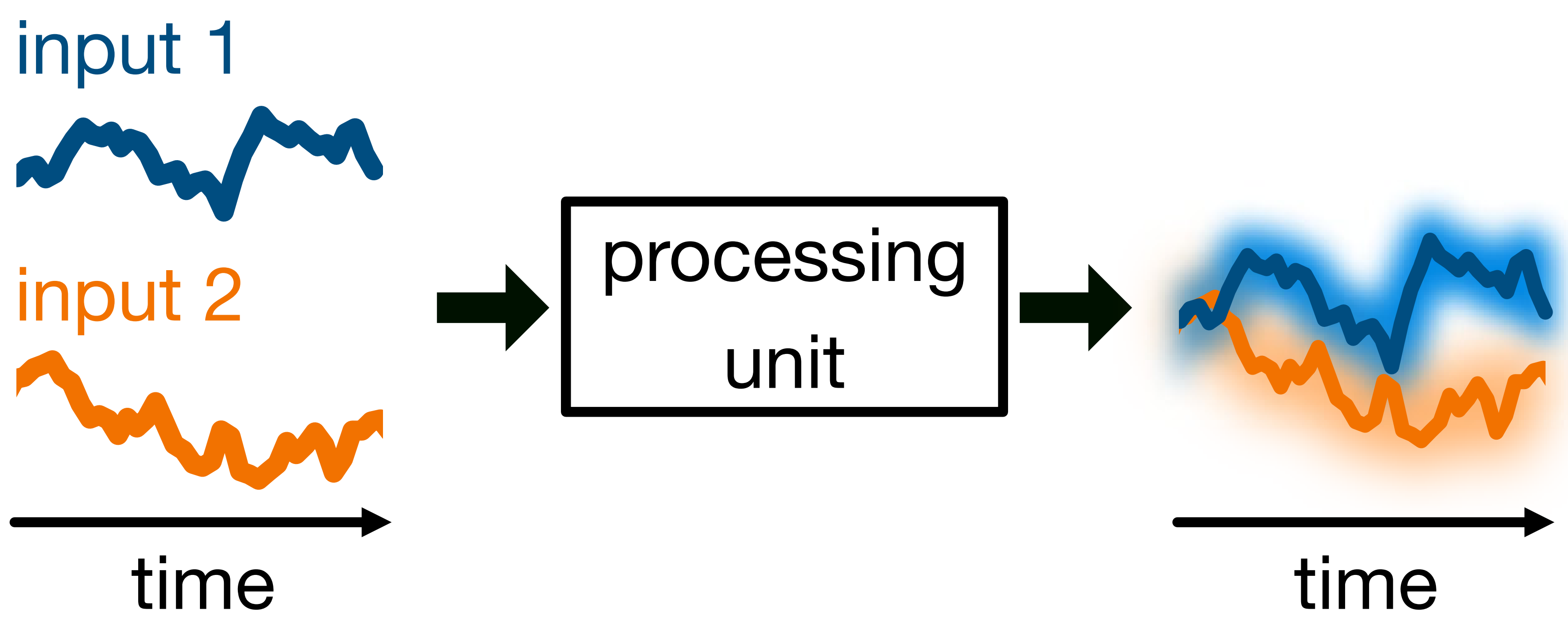
AMOLF physics of functional complex matter

BaSyC

IST AUSTRIA
Institute of Science and Technology

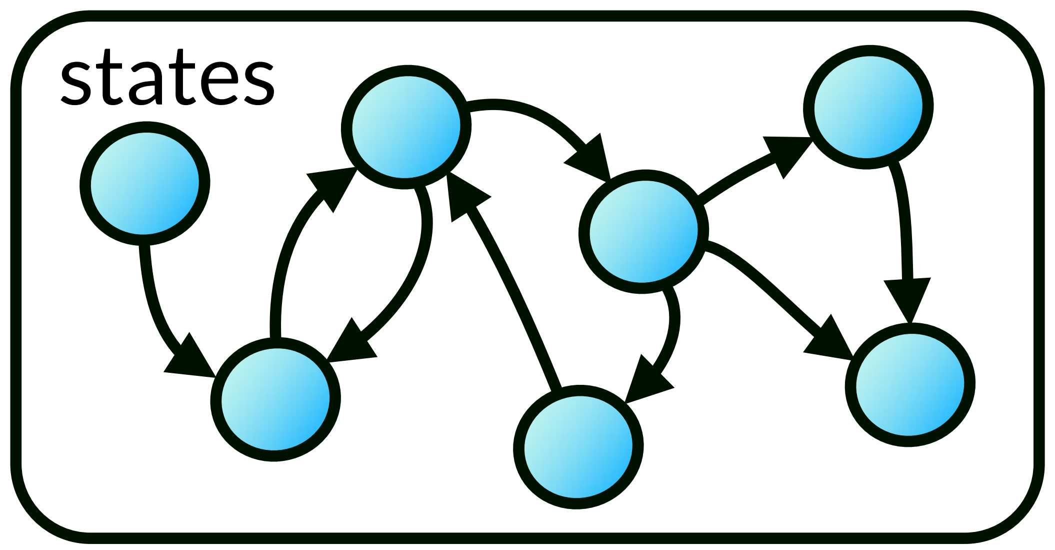
Gašper Tkačik

Summary



- The information rate is the essential quantity to describe information flow
- We have developed a *completely generic* method to compute it exactly *for the first time*

processing device



Next Step:

Going beyond the master equation