Hydrodynamics of the inhomogeneous ℓ-TASEP

Đan Daniel Erdmann-Pham, Khanh Dao Duc & Yun S. Song

University of California, Berkeley

June 25, 2019

Tel Aviv University

Modeling Translation Dynamics
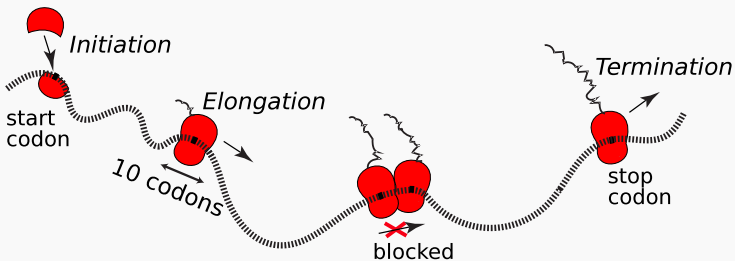
Which key parameters govern translation efficiency?

Figure: Ribosomes move along mRNA, translating it into a polypeptide

What determines translation efficiency?

▶ What is the local *density* of ribosomes?

▶ What controls the *production rate* of protein?

▶ How to *optimize* translation efficiency?

Complexity of translation

Elongation rates are locally regulated by a variety of factors:

- ▶ tRNA availability
- ▶ decoding and dissociation rates
- ▶ mRNA secondary structure
- ▶ co-translational folding
- ▶ signal recognition particle (SRP) binding
- ▶ interactions between polypeptide and ribosome exit tunnel

> - ▶ Do faster rates correspond to more efficient translation?
> - ▶ What are the implications of codon usage bias?
> - ▶ Does elongation or initiation limit production rates? What about termination rates?

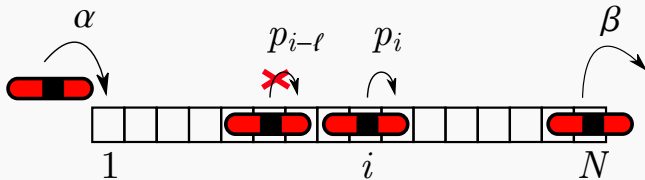# Totally Asymmetric Simple Exclusion Process

The model

Figure: Totally Asymmetric Simple Exclusion Process (TASEP) with inhomogeneous jump rates and particles of size $\ell = 3$

Inhomogeneous ℓ-TASEP

- continuous time Markov chain $\tau(\cdot) \in \{0, 1\}^N$
- particles enter at rate α if first ℓ sites are empty
- particles move from site i to $i + 1$ at rate p_i if site $i + \ell$ is empty
- particles exit through site N at rate β

The homogeneous 1-**TASEP**
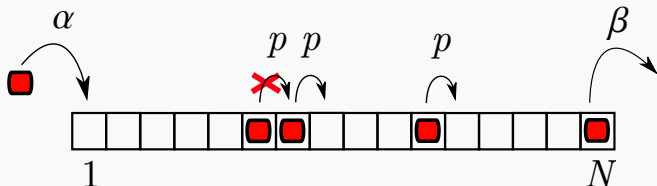
Modeling transport phenomena

▶ first introduced by MacDonald et al. (1968) to model translation

▶ solved by Derrida, Evans, Hakim, Pasquier (1993)

▶ since then used to model traffic flow, molecular transport, surface growth etc.

Hydrodynamics, KPZ and large deviations

▶ law of large numbers (Burger's Equation)

▶ fluctuations (mostly on $\mathbb{Z}$ and under specific initial data)
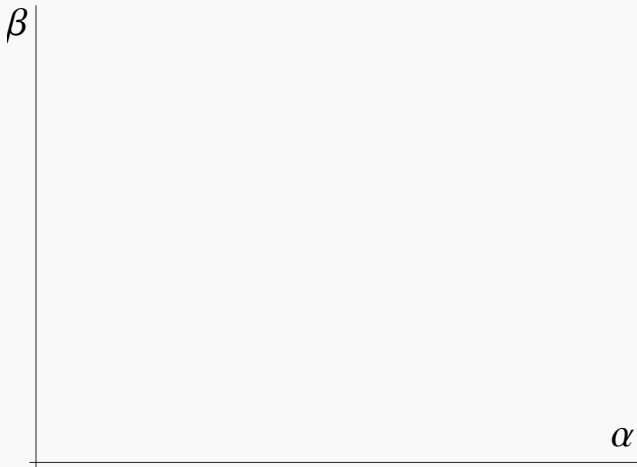
▶ rare events

Phase Diagram of the homogeneous 1-**TASEP**

Figure: We expect equilibrium density profiles to depend critically on α and β

Phase Diagram of the homogeneous 1-TASEP
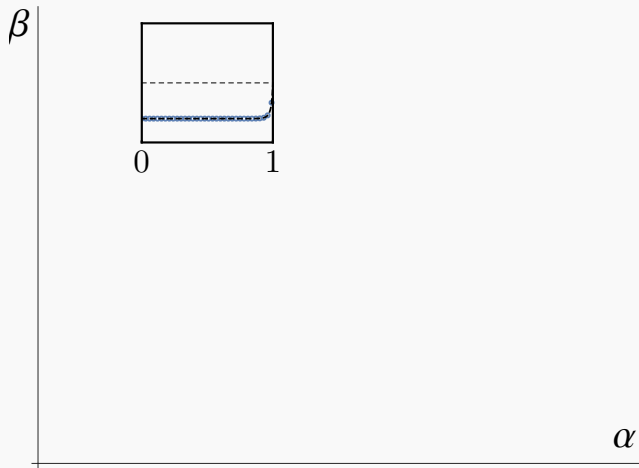

Figure: Low initiation rates and large termination rates leave the lattice sparsely populated
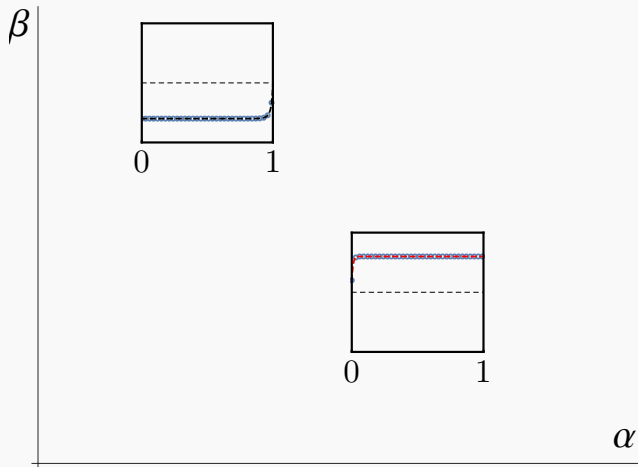
Phase Diagram of the homogeneous 1-TASEP

Figure: High initiation rate and low termination rates create traffic jams, resulting in a densely populated lattice

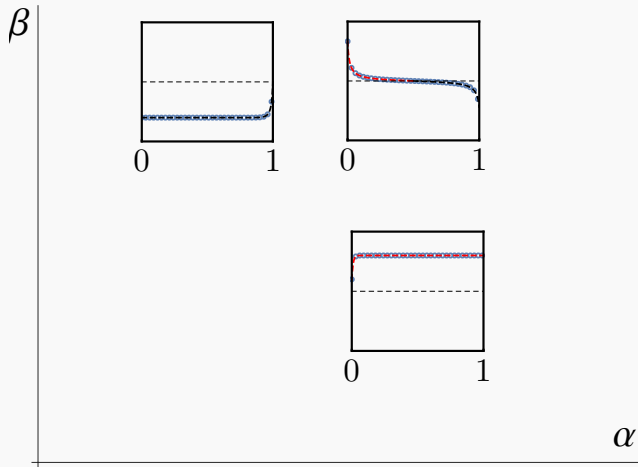## Phase Diagram of the homogeneous 1-TASEP

Figure: If particles both arrive and leave quickly, the results is a superposition of traffic jams and sparse regions
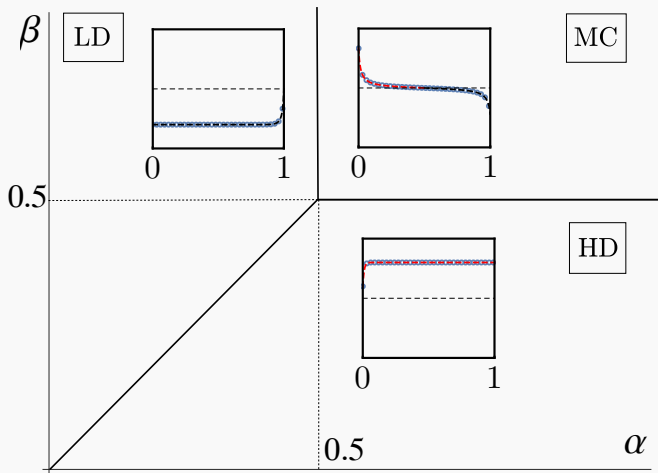
Phase Diagram of the homogeneous 1-TASEP

Figure: These three density profiles are representative of the low density, high density and maximum current regimes
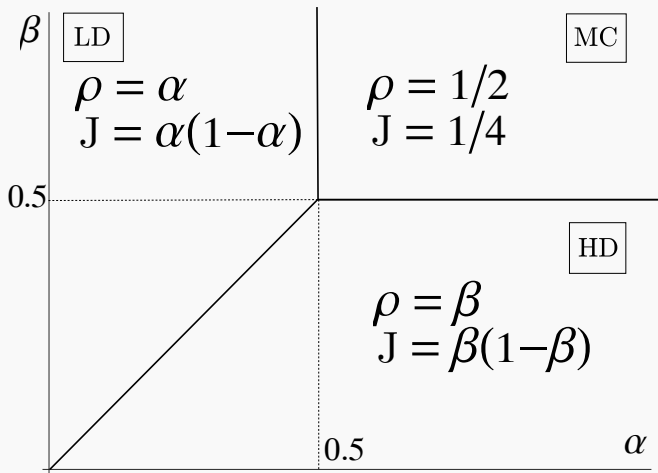
Phase Diagram of the homogeneous 1-TASEP

Figure: These three density profiles are representative of the low density, high density and maximum current regimes

Can we extend these results to the inhomogeneous ℓ-TASEP?

Mean-field approaches

- ▶ ignore higher order correlations, e.g. approximate

$$\partial_t \langle \tau_i \rangle = \langle \tau_{i-1} (1 - \tau_i) \rangle - \langle \tau_i (1 - \tau_{i+1}) \rangle$$
$$\approx \langle \tau_{i-1} \rangle (1 - \langle \tau_i \rangle) - \langle \tau_i \rangle (1 - \langle \tau_{i+1} \rangle)$$

- ▶ approximate bulk behavior by dynamics on a ring
- ▶ treat local defects as separating two decoupled copies of the homogeneous TASEP

Exact results through coupling

- ▶ global effects of local defects
- ▶ random elongation rates
- ▶ hydrodynamic limits
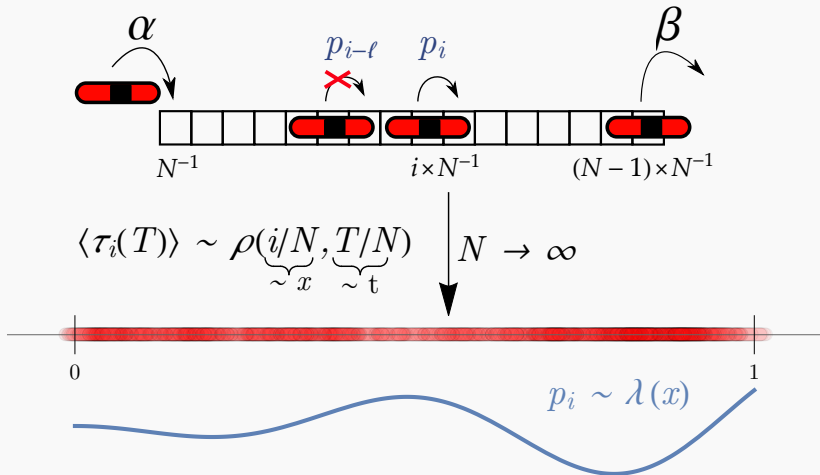- ▶ power series expansions in α

Hydrodynamic Limit

Hydrodynamic Limit

$$\langle \tau_i(T) \rangle \sim \rho(\underbrace{i/N}_{\sim\,x}, \underbrace{T/N}_{\sim\,\mathrm{t}})$$

$N \to \infty$

$p_i \sim \lambda(x)$

Figure: Continuum limit under Euler scaling

What dynamics does ρ follow?

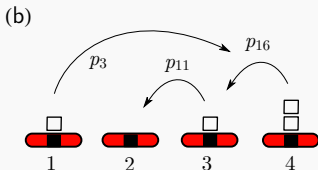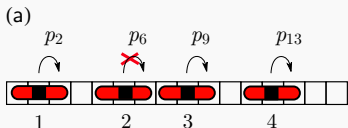1. Identify candidate PDE through Zero Range Process (ZRP):

(a)

(b)

Figure: Mapping from inhomogeneous ℓ-TASEP to ZRP

Result is a conservation law with current $J = \lambda \rho G(\rho)$:

$$\partial_t \rho = -\partial_x \left(\lambda \rho G(\rho) \right),$$

where

$$G(\rho) = \frac{1 - \ell \rho}{1 - (\ell - 1)\rho}.$$

2. Rigorous proof involves

▶ relative entropy techniques (*Rezakhanlou '97*)

▶ coupling with smoother system (*Bahadoran '12*)

▶ local stationarity

Solving the PDE

Boundary conditions & characteristic equations

α & β impose effective Dirichlet Boundary Conditions which propagate via characteristics

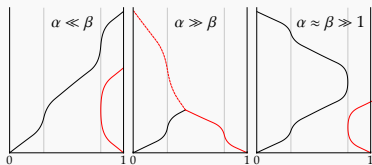Hydrostatics are determined by the characteristic curve that saturates the lattice
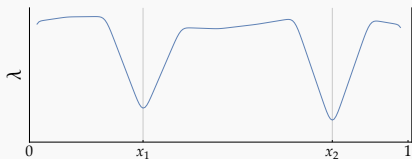

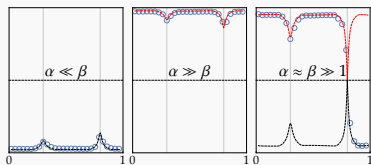**Figure:** Characteristic curves for different (α, β)

Figure: Simulated and theoretical static density profiles
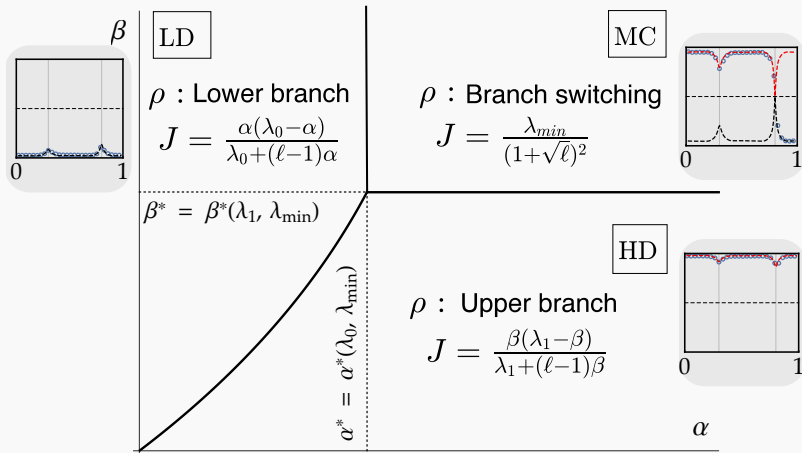
Phase diagram ●

Figure: Phase diagram depends only on $\lambda_0, \lambda_1, \lambda_{\min}$ and ℓ. Phase transitions are in ρ and J.

Application to Translation

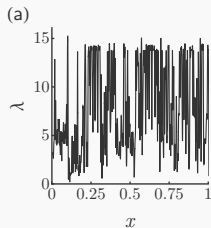# Caveat 1: Ribosome profiles are rarely smooth

(a)

Figure: Typical elongation profile of a yeast gene

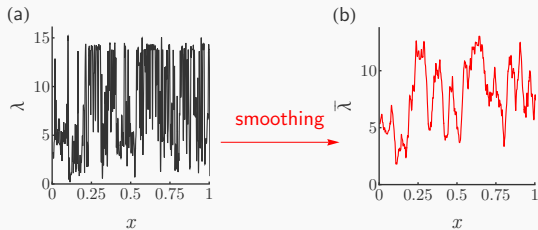# Caveat 1: Ribosome profiles are rarely smooth

Figure: Smoothing reveals clearer structure

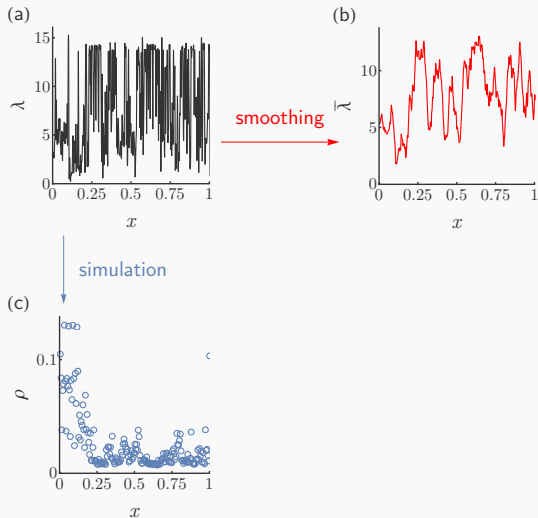# Caveat 1: Ribosome profiles are rarely smooth

Figure: Simulation under noisy elongation yields noisy densities

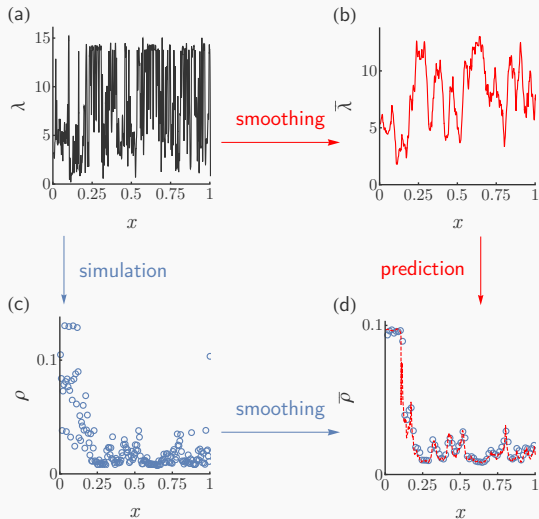# Caveat 1: Ribosome profiles are rarely smooth

Figure: Smoothed predictions match smoothed densities

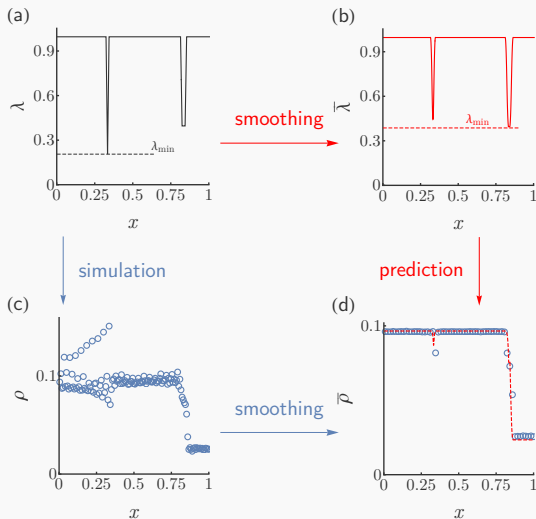# Caveat 1: Ribosome profiles are rarely smooth

Figure: Smoothed elongation rates predict translation dynamics

Caveat 2: Misdetection of stacked ribosomes

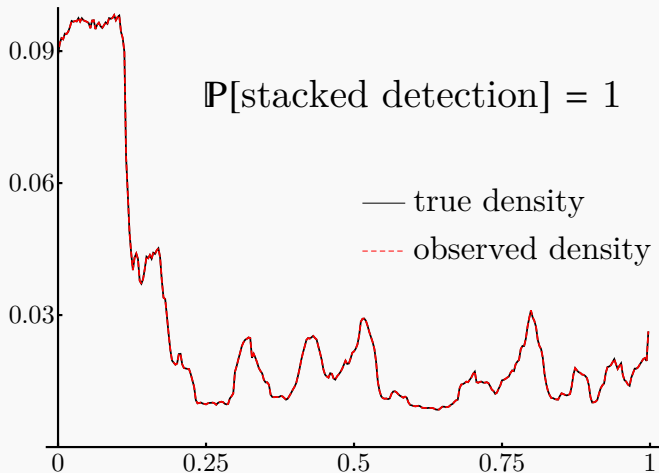**Figure:** Complete identifiability in case of full detection

Caveat 2: Misdetection of stacked ribosomes
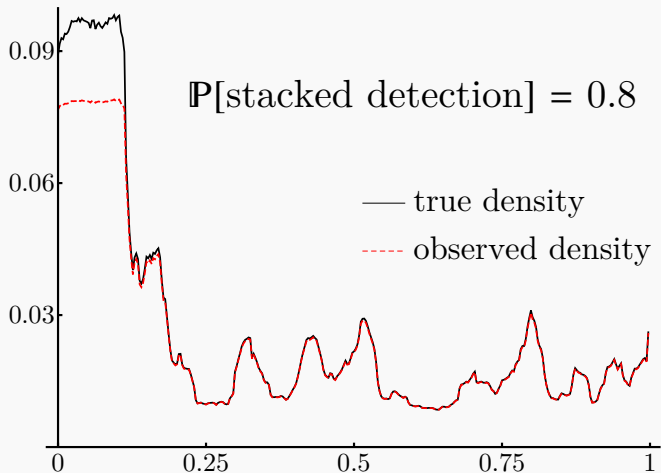
Figure: Mild misdetection affects HD, but preserves identifiability

Caveat 2: Misdetection of stacked ribosomes
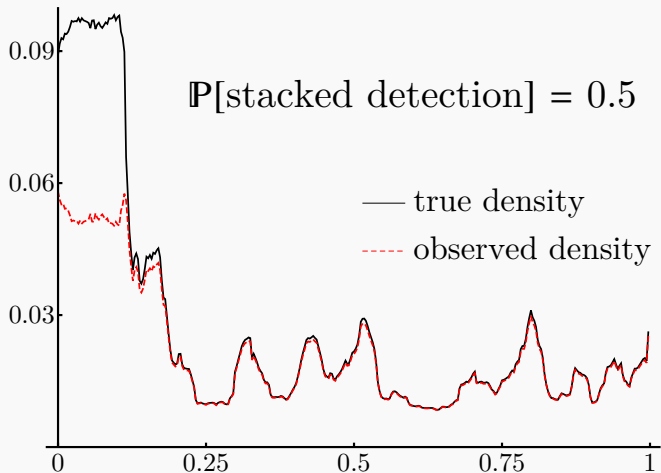
$$\mathbb{P}[\text{stacked detection}] = 0.5$$

—— true density

- - - - observed density

Figure: Identifiability is lost under moderate to severe misdetection

Caveat 2: Misdetection of stacked ribosomes

$\mathbb{P}[\text{stacked detection}] = 0.1$

—— true density
---- observed density

Figure: Identifiability is lost under moderate to severe misdetection

Caveat 3: Mixing time

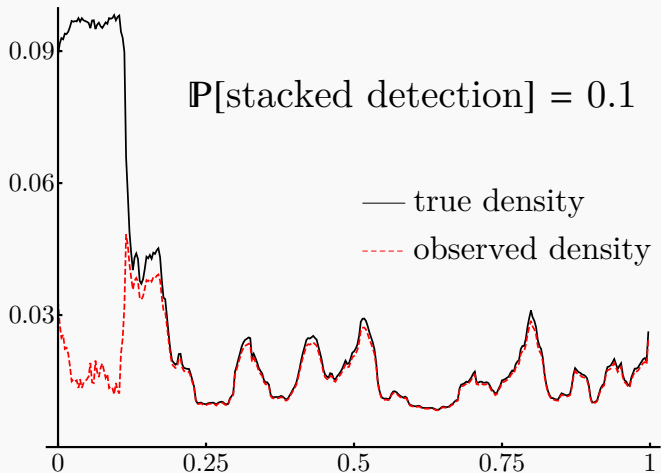- ▶ Analysis based on stationary distribution presupposes negligible transient behaviour
- ▶ How long does it take for the chain to mix?

Reassurance

- ▶ In LD, approximately

$$\int_0^1 \left\{ \lambda(x) H' \circ H_-^{-1} \left(\frac{J}{\lambda(x)} \right) \right\}^{-1} \, \mathrm{d}x$$

seconds

- ▶ In MC, roughly twice as long
- ▶ In yeast, the median mixing time is 68 seconds, compared to a ~ 30 minutes median half-life in a transcript

Utility of hydrodynamic theory in understanding translation

Benefits of closed-form formulas

1. Obviate the need for costly simulation studies

2. Identification of key parameters governing translation, namely $\alpha, \lambda_{\min}, \lambda_0$ and $x_{\min} = \arg\min_x \lambda(x)$

Design principles

1. Regulation: α should should act as main determinant of J

2. Production capacity: $J_{\max}$ fixed by $\lambda_{\min}$

3. Sensitivity: rate of change of J in α is modulated by λ_0

4. Ribosome cost: small $x_{\min}$ may save resources

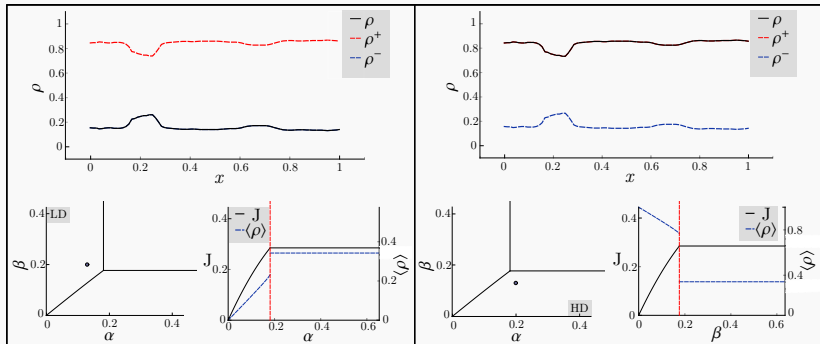# Regulation: Low Density (LD) vs. High Density (HD)

Figure: Any given level of production can be instantiated through either LD or HD. For the purposes of resource allocation, LD should be favored over HD, rendering α the major determinant of J.

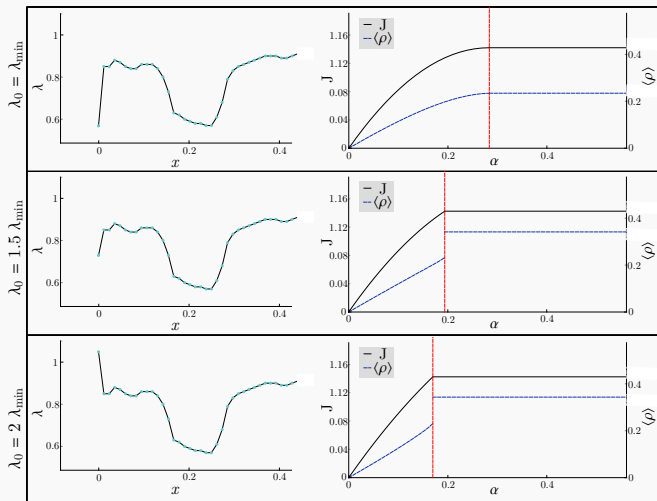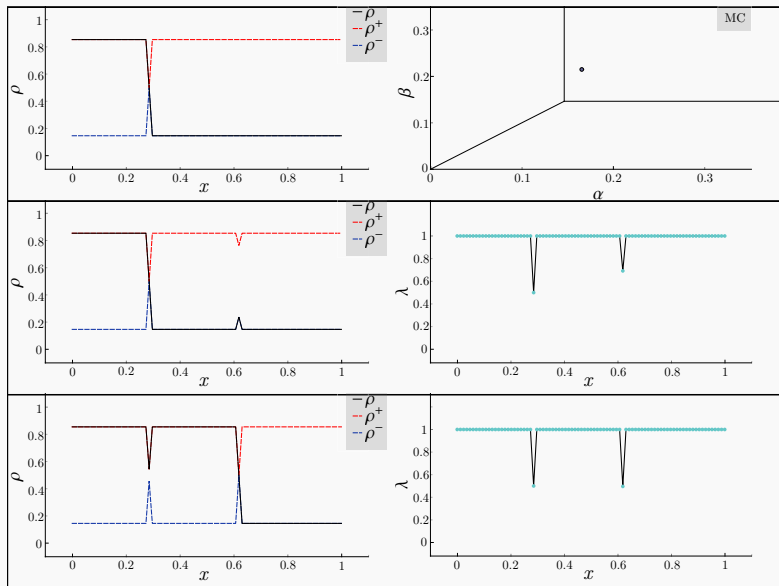# How and how quickly can production rates adjust?

Figure: J plateaus at $J_{\max} = \lambda_{\min}/(1 + \sqrt{\ell})^2$ with $\partial_\alpha J$ increasing in λ_0
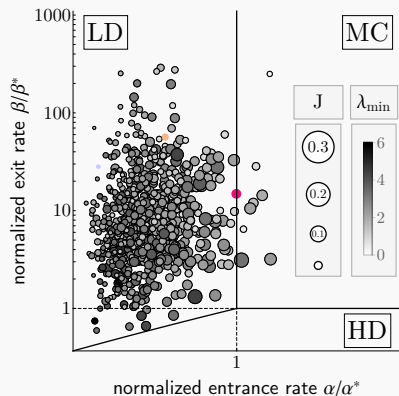
Optimizing ribosome cost at high production rates

Translational efficiency in yeast

What regimes do yeast genes operate in?

(a) Phase diagram

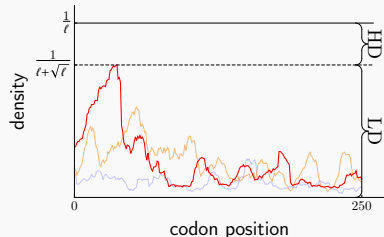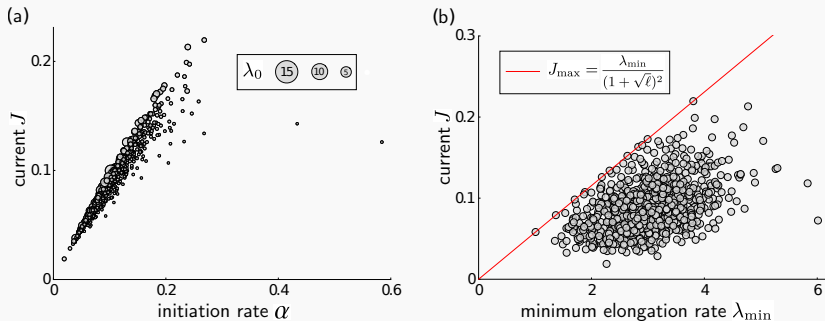(b) Sample densities

Figure: Analysis of 850 previously inferred elongation profiles: 841 genes are located in LD, and 9 genes have pushed into MC, rendering α, $\lambda_{\min}$ and λ_0 the main determinants of protein production.

Genes exploit full dynamic range of currents

(a)

current J vs initiation rate α

λ_0 (15) (10) (5)

(b)

current J vs minimum elongation rate $\lambda_{\min}$

$J_{\max} = \dfrac{\lambda_{\min}}{(1+\sqrt{\ell})^2}$

- ▶ Most genes are initiation rate limited
- ▶ Elongation limited profiles exist at high translation rates
- ▶ Constitutively highly expressed transcripts operate close to maximum capacity
- ▶ Genes with variable expression demands show large variation in J

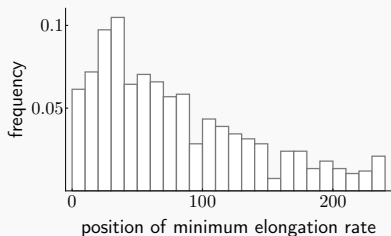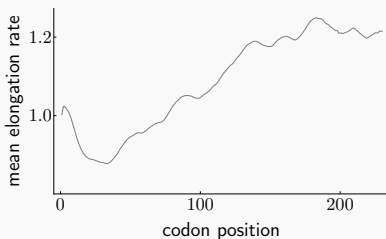# λ_0 facilitates realization and modulation of large currents
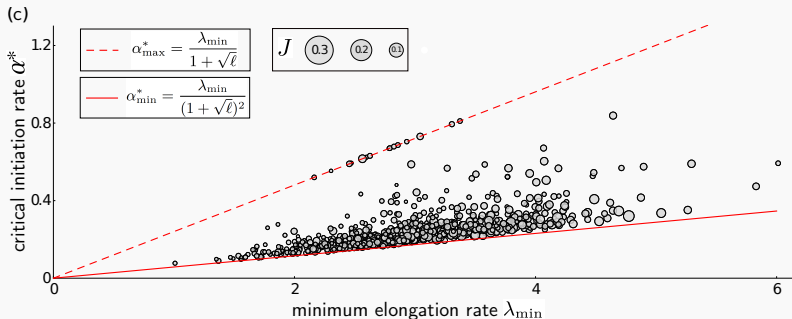
- ▶ Elevated initial rates close to maximum capacity
- ▶ Sensitivity ratio $\lambda_0/\lambda_{\min}$ increased for variable expression demands, and decreased for constitutive expression

Locations of minimum elongation rate are chosen judiciously

(b) concentration of $x_{\min}$

(b) 5' translational ramp

▶ Minimum elongation locations mirror 5′ translational ramp, balancing trade-off between resource and sensitivity considerations

▶ High sensitivity and proximity to MC lower $x_{\min}$ significantly

▶ Genes with large $x_{\min}$ are furthest from MC

Looking back and ahead

What happened?

1. Derived closed-form expressions for TASEP current and densities by passing to hydrodynamic limit
2. Identified $\alpha, \lambda_0, \lambda_{\min}$ and $x_{\min}$ as key parameters
3. Studied their consequences for codon usage bias
4. Developed four design principles governing translation
5. Provide evidence for implementation of principles in yeast

What might happen?

▶ Direct inference from ribosome profiling data
    ▶ incorporate stacked ribosomes
▶ Rate function regularity and smoothing
▶ Fluctuations, large deviations, mixing times, singular configurations
▶ Apply machinery to transcription

Acknowledgements

Khanh Dao Duc   Yun Song

Fraydoun Rezakhanlou

Son Won Jun

References

Bahadoran, C. (2012).
Hydrodynamics and hydrostatics for a class of asymmetric particle systems with open boundaries.
Communications in Mathematical Physics, 310(1):1–24.

Covert, P. and Rezakhanlou, F. (1997).
Hydrodynamic limit for particle systems with nonconstant speed parameter.
Journal of statistical physics, 88(1):383–426.

Erdmann-Pham, D. D., Duc, K. D., and Song, Y. S. (2018).
Hydrodynamic limit of the inhomogeneous ℓ-tasep with open boundaries: Derivation and solution.
arXiv preprint arXiv:1803.05609.

Son, W., Erdmann-Pham, D. D., Dao Duc, K., and Song, Y. S. (2018).
Current-inspired visualization and evaluation of the tasep (civet).
https://github.com/songlab-cal/civet.