

Outline of the series

- Overview/Review "Equilibrium SM vs. Nonequilibrium SM"
- An Ising-like model in DDS "Shattered expectations"
- DDS in one-dimension "Bare bones NESM"
 - Interesting physics, despite just 1-D and "no interactions"
 - Potential applications - protein synthesis
 - Exact solutions and intractable extensions
- Systems with more than one driven species "American football, Barber poles, and Clouds"
- Summary and Outlook "Come and join in the fun!"



Driven Ising Lattice Gas

The surprises continue...

- $E = 0$ $J \neq 0$ $d = 1, 2$ (Lenz-Ising, Onsager, Lee-Yang, ...)
- $E > 0$ $J > 0$ $d = 2$ KLS
- $E > 0$ $J = 0$ $d = 1$ Asymmetric Simple Exclusion Process
- $E = \infty$ $J = 0$ $d = 1$ Totally ASEP (Spitzer 1970)
 - for PBC, P^* trivial, but dynamics non-trivial
 - for OBC, P^* non-trivial ("boundary induced" phases)
 - for OBC, dynamics even more exciting
 - extensions (still $E = \infty$ $J = 0$) for modeling protein synthesis



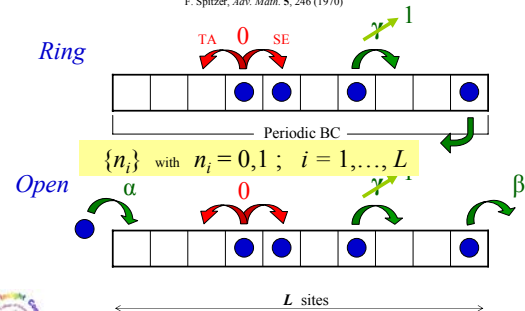
Outline of TASEP's

- Brief glimpse of "early" TASEP's
 - What's TASEP? - a "toy in pure mathematics" (1970)
 - A chemist's model for protein synthesis - the 1968&69 studies
- Physicists' TASEP
 - Phase transitions in 1-D, density profiles, anomalous diffusion...
 - Exact solutions and their limitations (for finding e.g., a simple quantity)
- Extensions motivated by protein synthesis
 - TASEP is too "simple" indeed
 - Effects of extended objects (particles covering more than one site)
 - Effects of inhomogeneities and their locations
 - Effects of competition with other TASEP's



Mathematician's TASEP

F. Spitzer, *Adv. Math.* 5, 246 (1970)



Questions of interest

- What is stationary distribution: $P^*(n_i)$?
- What can we say about the t dependence: $P(n_p, t)$?



Biochemist's TASEP

Working independently of, and most likely simultaneously as F. Spitzer,

J. H. Gibbs, a chemist + A. C. Pipkin, an applied mathematician studied a model for **protein synthesis**.

This study formed the PhD thesis of

C. T. MacDonald

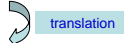
at Brown, and published as two papers:

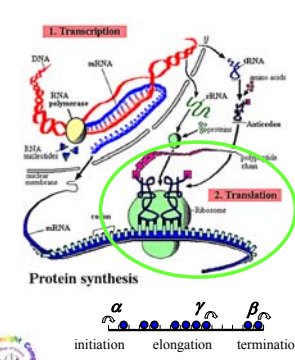
C.T. MacDonald, J.H. Gibbs, and A.C. Pipkin, *Biopolymers* (1968+69)

... brief detour into biology



Brief detour ... "Central Dogma"

- **DNA** ~1G base-pairs; 2% codes for ~25K genes
- **RNA** Many many types: mRNA, tRNA, rRNA, ...
- **mRNA** 10-3K codons, makes 
- **Proteins**
 - a string of amino-acids (an aa-chain), known as "polypeptide chain"
 - folds to have specific function
 - e.g., keratin (nail, hair), hemoglobin, receptor protein (nerve), ompA, ...

1. Transcription

2. Translation

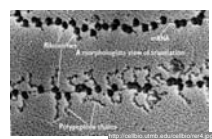
Protein synthesis

initiation elongation termination

A ribosome...

- starts at one end (**initiation**)
- goes to the other, "knitting" the aa-chain (**elongation**)
- releases aa-chain at the end and falls off mRNA (**termination**)

Before one falls off, another one starts!




(some) Questions of interest

- **What is production rate** in steady state?
 - related to protein levels assuming constant degradation rate
 - given by TASEP current J in steady state
- **What can we say about the overall occupation and density profile?**
 - of interest to competition ??
- Full dynamics is far, far down the present priority list!



Outline of TASEP's

- Brief glimpse of "early" TASEP's
 - What's TASEP? - a "toy in pure mathematics" (1970)
 - A chemist's model for protein synthesis - the 1968,69 studies
- **Physicists' TASEP**
 - Phase transitions in 1-D, density profiles, anomalous diffusion...
 - Exact solutions and their limitations (for finding e.g., a simple quantity)
- **Extensions** motivated by protein synthesis
 - TASEP is too "simple" indeed
 - Effects of extended objects (particles covering more than one site)
 - Effects of inhomogeneities and their locations
 - Effects of competition with other TASEPs



Questions for TASEP

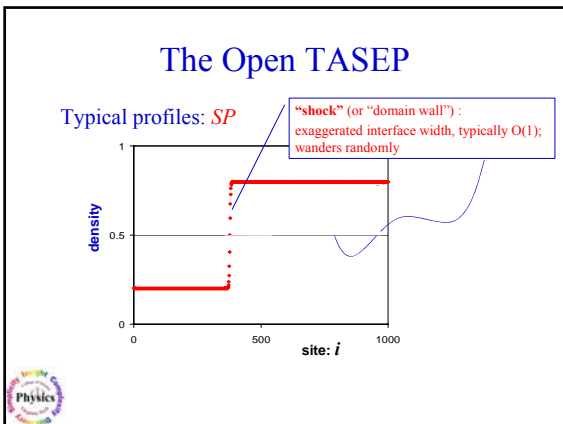
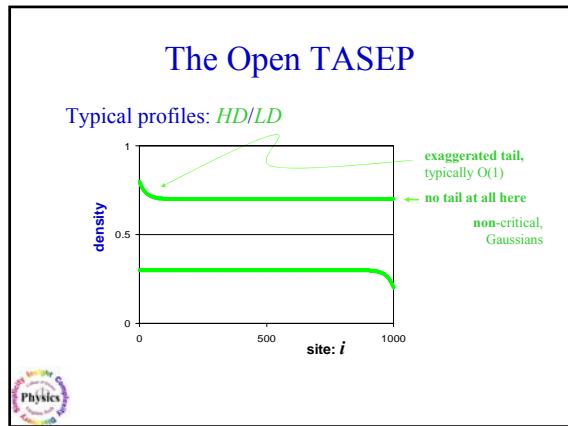
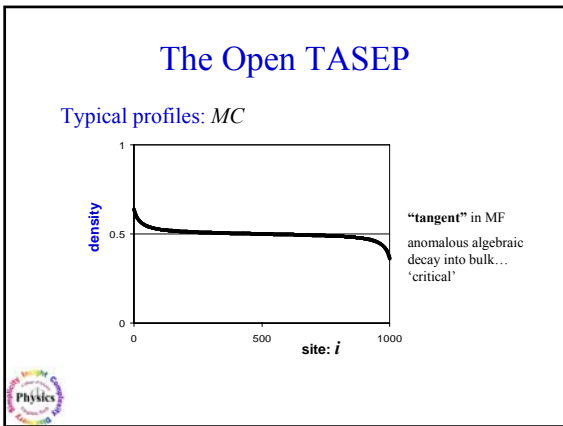
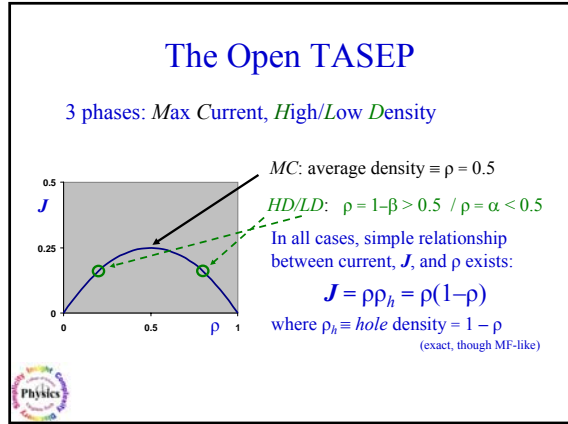
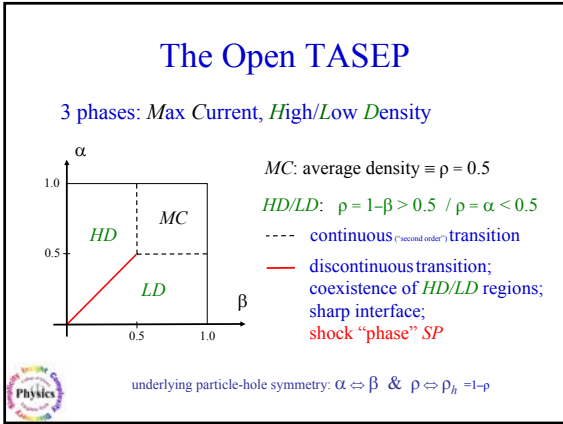
- What is stationary distribution: $P^*(n_i)$?
 - What is full dynamics $P(n_i, t)$?



(some) Answers for TASEP

- What is stationary distribution: $P^*(n_i)$?
 - What is full dynamics $P(n_i, t)$?
- For the *Ring*, $P^* \propto 1$ (Spitzer, 1970; pairwise DB)
 - but with non-trivial dynamics
- For the *Open* case, P^* non-trivial (1992)
 - with even more interesting dynamics





- ### The Open TASEP
- Much, much more is known *exactly*:
 - fluctuations and correlations, average J , anomalous dynamics, shock properties, etc.
 - using mapping to quantum chain, Bethe ansatz, and "matrix product" methods
(reviews by G.M. Schütz, Evans, Derrida, Mallick, ...)
 - lots of hard work, just to get (Bethe ansatz) *equations* for spectrum and eigenvectors!
 - Exact solutions have *limitations*:
 - very specialized (cannot be generalized, even for "small" changes)
 - ditto (even for "simple" quantities, such as ...)

Total Occupation in Open TASEP

- A simple, natural question:
- The stationary properties (e.g., average, fluctuations) of the total occupation, N , are known...
- What about its time dependence $N(t)$?
(i.e., correlations in time)
- ...especially for *finite* systems?
- Motivations come from biology.



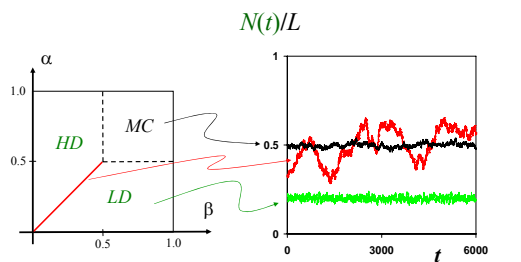
Motivations from biology

Many (copies of many different) mRNA's *compete* for one pool of ribosomes (and aa-tRNA's).

- Do some “win big” and others “lose out”?
- To quantify these, need $N_i(t)$ or their FT's.
- Baseline study: What's $N(t)$ like, for just *one* TASEP?
...especially for *finite* L !



Total Occupation in Open TASEP



$L = 1000; t_{\text{unit}} = 100\text{MCS}$



Power Spectra of Total Occupation

$$I(\omega) \equiv \langle |\tilde{N}(\omega)|^2 \rangle$$

To be precise:

Can be predicted, *in principle*, through the *exact solution*.

But, *in practice*, need

- the entire spectrum,
- special linear combinations of ... *all* the eigenvectors,
- sum over the above!

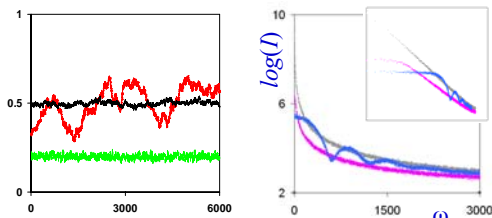
- take 10^8 MCS in steady state
- measure N every 100 MCS
- have 10^6 measurements
- cut into 100 samples of...
- 10^4 N 's and get \tilde{N}
- average over the 100 samples

D. Adams, R.K.P. Zia, and B. Schmittmann
PRL **99**, 020601 (2007)



Power Spectra of Total Occupation

$$I(\omega) \equiv \langle |\tilde{N}(\omega)|^2 \rangle$$



- Oscillations generic (for finite L):
 - Toy: single particle, constant v , random entrances
 - Effect of dispersion/diffusion: filling in minima
- Simple minded, linearized Langevin equation:

$$\partial_t \rho = D \partial^2 \rho - \partial [\rho(1-\rho)] + \partial \eta \quad (\eta(x,t)\eta(x',t') = \delta(x-x')\delta(t-t')$$

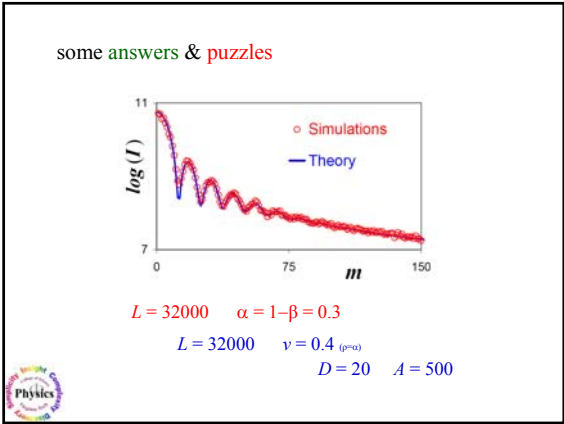
$$\rho = \bar{\rho} + \varphi$$

$$\partial_t \varphi = D \partial^2 \varphi - v \partial \varphi + \partial \tilde{\eta} \quad v = 1 - 2\bar{\rho}$$

linear diffusion conserved noise

$$N(t) = \int_0^L \rho(x,t) dx \Rightarrow \tilde{N}(\omega) = \int_k (e^{ikL} - 1) \frac{\tilde{\eta}(k, \omega)}{Dk^2 + ivk - i\omega}$$





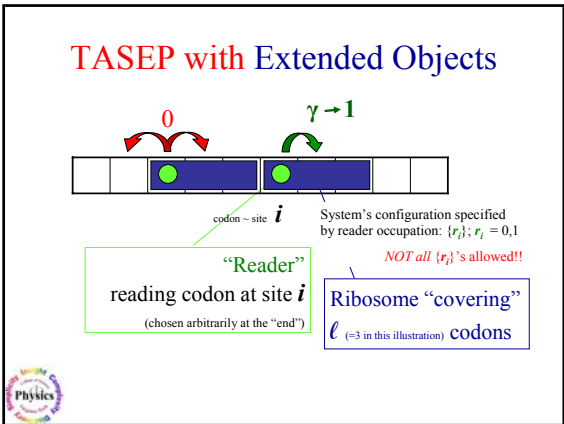
- ### Outline of TASEP bit
- Brief glimpse of “early” TASEP’s
 - What’s TASEP? - a “toy in pure mathematics” (1970)
 - A chemist’s model for protein synthesis - the 1968,69 studies
 - Physicists’ TASEP
 - Phase transitions in 1-D, density profiles, anomalous diffusion...
 - Exact solutions and their limitations (for finding e.g., a simple quantity)
 - Extensions motivated by protein synthesis
 - TASEP is too “simple” indeed
 - Effects of extended objects (particles covering more than one site)
 - Effects of inhomogeneities and their locations
 - Effects of competition with other TASEPs

Extensions (motivated by) protein synthesis

MacDonald, Gibbs, and Pipkin, *Biopolymers* (1968+69)

- To model translation in protein synthesis... where ribosomes bind to the mRNA, move along “reading” a string codons, and detach at the end,
- Gibbs introduced a 1-d lattice for the mRNA with
 - site \leftrightarrow codon particle \leftrightarrow ribosome

- ### Extensions (motivated by) protein synthesis
- But, many aspects of TASEP are too Simple!
 - Here, let’s focus on just *three* extensions:
 - A ribosome is a large molecule covering typically 10-12 codons, so our TASEP particles need to be *extended objects* covering ℓ (≈ 12 mostly) sites
 - Inhomogeneous jump rates
 - Competition for finite resources
- Gibbs & co. had these in '68+'69!



Open TASEP with extended objects first considered by

J. H. Gibbs

MacDonald’s PhD thesis 1968, 1969

good MFT for extended objects on open chain...
 ...solved numerically, mostly for $\ell = 2$

$J = \langle r_i(1-r_{i+1}) \rangle \rightarrow \langle r_i \rangle(1-r_{i+1})$ very poor, except for $\ell=1$. MGP more sophisticated

TASEP with Extended Objects

...on a homogeneous **Ring**: $\gamma_i = 1$

- $P^* \propto 1$ still holds...
- every configuration mapped to $\ell = 1$ case
...with a smaller ring: $L - (\ell - 1)M$
 $M = \text{the number of ribosomes}$



TASEP with Extended Objects

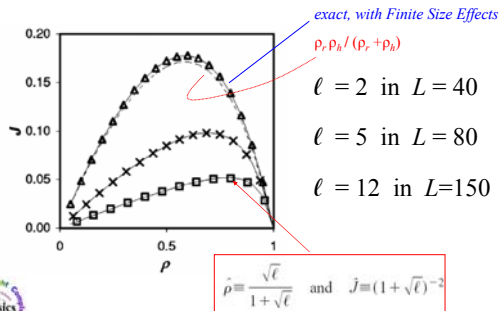
Homogeneous Ring: $\gamma_i = 1$

Yet, there are interesting aspects:

- two different overall densities:
 - ✓ "coverage": $\rho = 1 - \text{hole density} \equiv 1 - \rho_h \in [0, 1]$
 - ✓ "reader": $\rho_r = \text{particle density} = M/L = \rho/\ell \in [0, 1/\ell]$
- current-density: $J(\rho) = \rho_r \rho_h / (\rho_r + \rho_h)$
- non-zero, amusing r - r correlations!

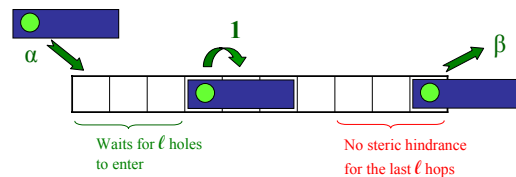


TASEP with Extended Objects



TASEP with Extended Objects

Homogeneous Open Lattice: $\gamma_i = 1; \alpha; \beta$



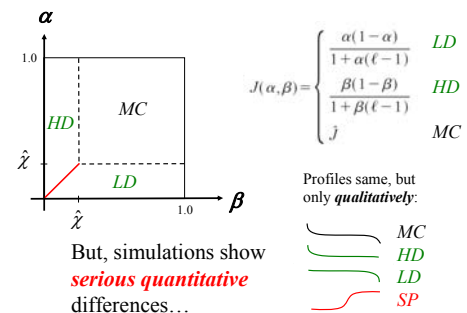
Lakatos and Chou, JPA 36, 2027 (2003): "complete entry" and "incremental exit"
Rules differ if "reader" placed elsewhere, but physics remains the same!



TASEP with Extended Objects

Homogeneous Open Lattice: $\gamma_i = 1; \alpha; \beta$

- $P^* = ???$
- mapping to $\ell=1$ case **not** simple
- many new interesting aspects:
 - o phase diagram and currents modified
 - o new features in density profiles
 - > expected: different ones for $(\rho_{\text{reader}})_i$ vs. ρ_i
 - > expected: period ℓ structures
 - > somewhat unexpected: **very long** tails (none for $\ell=1$)
 - > completely unexpected: yet to be discovered!




Gibbs & co. had these!

Recap of 1st Extension

Homogeneous TASEP's with *extended objects* are qualitatively similar to the $\ell=1$ case, but...

- **Ring:** Though $P^*=1$,
 - > there are non-trivial r - r correlations
- **Open:** $P^* = ???$
 - > phase diagram and $J(\rho)$ modified
 - > novel features in density profiles




Extensions (motivated by) protein synthesis

- But, many aspects of TASEP are too Simple!
- Here, let's focus on just *three* extensions:
 - Extended objects
 - 61 codons \leftrightarrow 61 aa-tRNA's, with widely varying concentrations. Ribosome waiting times may differ, so we should have *site-dependent jump rates*:

$$\{\gamma_i \neq 1\}$$
 - Competition for finite resources


Gibbs didn't have these.



TASEP with *extended objects* and *inhomogeneous rates*

TASEP \rightarrow T*AiSeP*


definitely **NOT** simple



T*AiSeP*

Inhomogeneous Open Lattice: $\alpha; \gamma_i; \beta$

- full problem of “quenched randomness”
 - Harris and Sinchombe ($\ell=1$), PRE 70, 016108 (2004)
 - ...every γ_i from the same distribution (e.g., Gaussian); our case even **more complex**
- Why study such a hard problem?
- Again, motivations from biology...
 - ...worse, *q.r.* averages may not be meaningful
- Indeed, there is a *new type* of problem: “Quenched distribution of distributions”




A little more detail...

... how does the ribosome move from codon to codon?

This leads us to consider $\{\gamma_i\}!!$

To advance, the ribosome must wait for the arrival of the appropriate aa-tRNA.


aa-tRNA concentrations [aa-tRNA] may vary greatly, so that the waiting times can differ seriously.



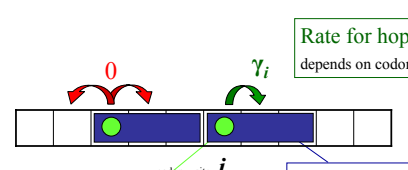
Our interests are many, e.g.,

Overall current, J , which controls protein levels and

Profiles, $(\rho_{reader})_i$, which controls which tRNA's are bound



TASEP




Rate for hopping depends on codon at site i

Ribosome “covering” ℓ (=3 in this illustration) codons

“Reader” reading codon at site i (chosen arbitrarily at the “end”)

System's configuration specified by $\{r_i\}$



TAiSeEP

Inhomogeneous Open Lattice: $\alpha; \gamma_i; \beta$

- full *quenched randomness* is tough!
- start simply...study one/two γ 's $\neq 1$
 - Fast ($\gamma > 1$) sites affect density profiles, but not J .
 - Slow ($\gamma < 1$) sites affect both.
 - Location of slow sites (e.g., clustering) is important.



Qualitative understanding from several mean field approaches:

Take two/three open TASEPs and match J across blockages
(details in DSZ, PRE-76, 2007+JID thesis)

Quantitative theory, good for all k 's, still lacking



Recap of 2nd Extension

- Slow “sites” reduce overall current (hardly surprising!)
- Two equally slow sites...
 - far apart affect current like just one blockage
 - close together reduces current much more
- Possible implications for translation and/or designer genes, exploiting *synonymous codons*
- Example with *dnaA* (from *E. coli*) shows promise.
- How to deal with problem of “Quenched distribution of distributions”?



Extensions (motivated by) protein synthesis

- But, many aspects of TASEP are too Simple!
- Here, let's focus on just *three* extensions:
 - Extended objects
 - Inhomogeneous jump rates
 - Typical cells have many copies of many types of mRNA's, all using ribosomes (and aa-tRNA's) from the same pool, so we should study the effects of a *finite reservoir of particles* on one or more TASEP's.



D. Adams, B. Schmittmann, and R.K.P. Zia JSTAT P06009 (2008)
L. J. Cook and R.K.P. Zia JSTAT P02012 (2009)
L. J. Cook, R. K. P. Zia, and B. Schmittmann arXiv: 0906.3730

David Adams Jon Cook

Effects of finite resources

The supply of ribosomes in a cell is finite!
All mRNA's compete for this resource.

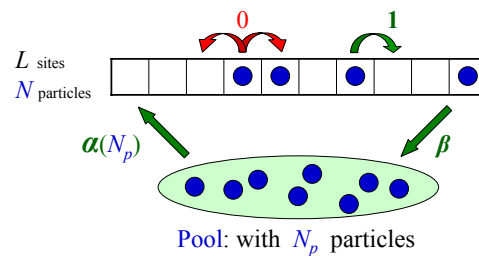
- Who “wins” and who “loses”?
- Should study many TASEP's competing for a pool of particles.

But, to begin with,

How does this affect a *single* TASEP?



Constrained TASEP



Constrained TASEP

$N_{tot} = N + N_p$
 is a **fixed** number of particles in the system
 (ribosomes in the cell)

We chose $\alpha(N_p) = \alpha \tanh(N_p/N_x)$
 ...numbers in TASEP feeds back to "on-rate", via $N_{tot} - N$

Some crossover scale, chosen as ρ^*L
 so that we can "see" an effect easily.

An intrinsic "on (initiation) rate",
 if the supply were unlimited.

Constrained TASEP

- Summary of parameters and **fixed choices**:
 - $L, \alpha, \beta, N_{tot}$
 - $N_x (= \rho^*L)$, *tanh* function
- Regions studied: **MC, HD, LD**
- Quantities of interest, as N_{tot} is varied:
 - steady state current: **J** (protein levels in a cell)
 - total occupancy: N_{or} or $\rho \equiv N/L$

No * here!!
constrained
TASEP

Constrained TASEP

Since the on-rate varies from 0 to α , the four possibilities, as N_{tot} increases from zero, are:

Constrained TASEP: **LD**

$\rho^* = 0.25$
 $J^* = \rho^*(1-\rho^*)$

◊ from theory!!

LD: $\alpha=1/4, \beta=3/4, L=1000, N_x/L=0.25$

Constrained TASEP: **LD**

- In ordinary *LD* TASEP, we have $\rho^* = \alpha; N_x = \alpha L$
- Set self-consistent condition:

$$\rho/\alpha = \tanh(\rho_{tot}/\alpha - \rho/\alpha)$$

$$\rho \rightarrow \begin{cases} \rho_{tot}/2 & \rho_{tot} \rightarrow 0 \\ \alpha & \rho_{tot} \rightarrow \infty \end{cases} \quad \rho_{tot} \equiv N_{tot}/L$$
- Current from the usual: $J = \rho(1-\rho)$

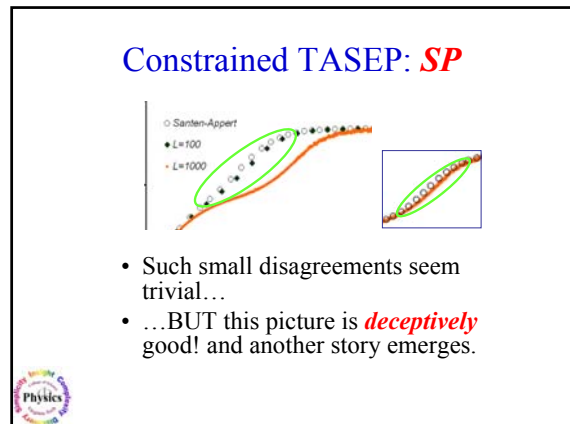
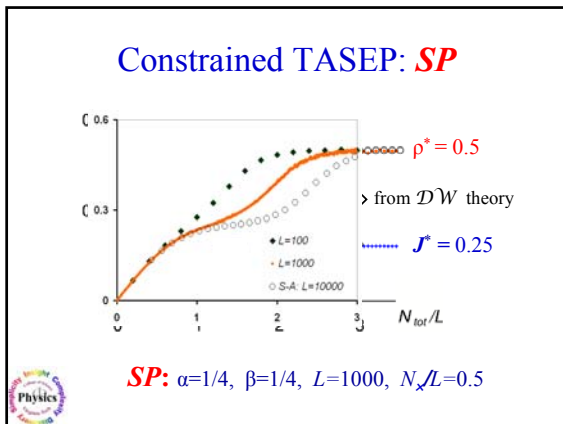
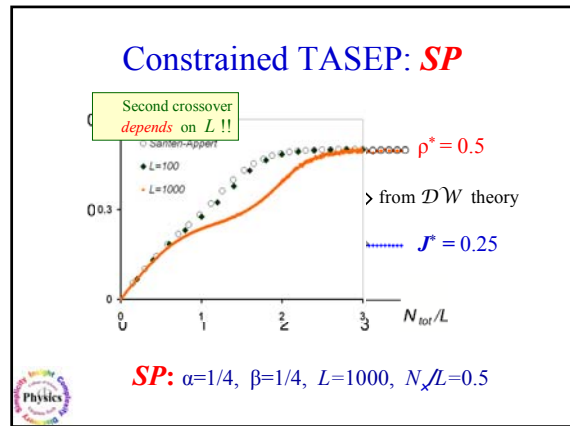
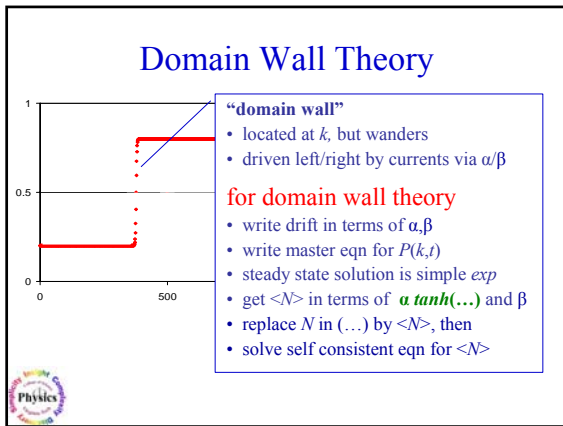
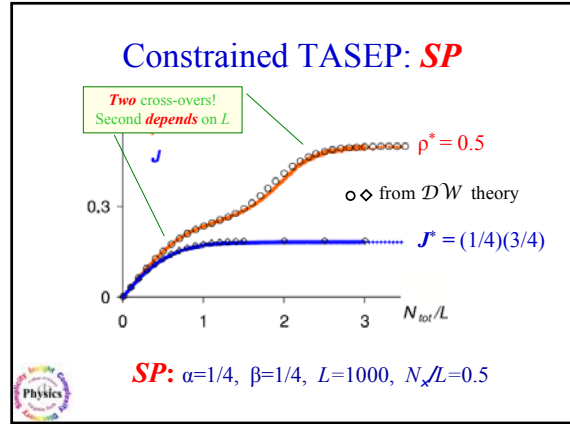
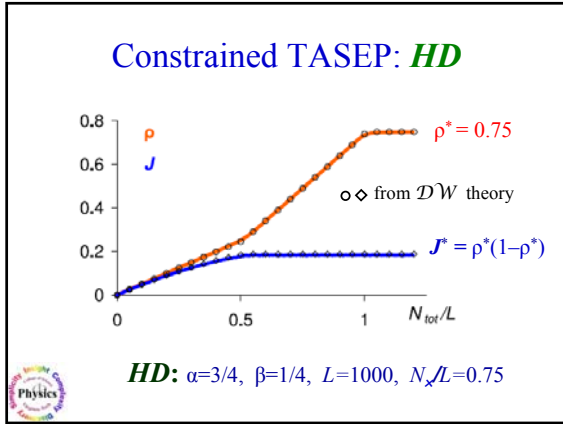
Constrained TASEP: **MC**

$\rho^* = 0.5$

on rate crosses phase boundary - into MC

◊ from theory

MC: $\alpha=3/4, \beta=3/4, L=1000, N_x/L=0.5$



Constrained TASEP: *SP*

- Fluctuations are clearly important for *SP*
- ...but they must be *suppressed* by the overall constraint.
- Suppression comes about through the feedback mechanism, i.e., $\alpha(N_{tot} - N)$
- Can see this even in the *LD* regime: Gaussians of $P(N)$ have narrower widths.
- Closer to *SP*, it leads to *shock localization*.



Constrained TASEP: *HD SP*

- Generalize DW to account for feedback...
- ...quick reminder of DW theory:

$$\partial_t P(k, t) = D_+ P(k-1, t) + D_- P(k+1, t) - (D_+ + D_-) P(k, t)$$

plus *reflecting* boundary conditions

position of DW: $\in [L, L]$

$$D_+ = \frac{j_+}{\rho_+ - \rho_-} = \frac{\beta(1-\beta)}{1-\beta-\alpha}$$

$$D_- = \frac{j_-}{\rho_+ - \rho_-} = \frac{\alpha(1-\alpha)}{1-\beta-\alpha}$$



Constrained TASEP: *HD SP*

- Generalize DW to account for feedback by a simple approximation:
- Promote $\alpha(N_{tot} - \langle N \rangle) \Rightarrow \alpha(N_{tot} - N)$
- Relate N to k , the DW position
- Have k dependent drift coefficients



Constrained TASEP: *HD SP*

$$\partial_t P(k) = D_{+,k-1} P(k-1) + D_{-,k+1} P(k+1) - (D_{+,k} + D_{-,k}) P(k)$$

$$P^*(k) = \left(\sum_{m=k_{\min}}^{k-1} \prod_{\ell=m+1}^k \frac{D_{-, \ell}}{D_{+, \ell-1}} + 1 + \sum_{m=k+1}^L \prod_{\ell=k+1}^m \frac{D_{+, \ell-1}}{D_{-, \ell}} \right)^{-1}$$

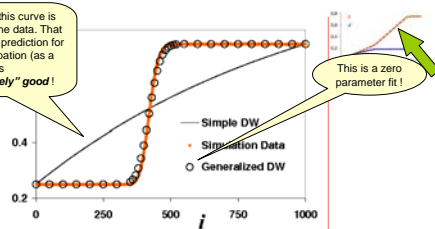
From here, we can find the density profile ρ_i ...

$$\rho_i = \sum_{k=k_{\min}}^i (1-\beta) P^*(k) + \sum_{k=i+1}^L \alpha_{\text{eff},k} P^*(k) \dots \text{and the result is}$$



Constrained TASEP: *HD*

The area under this curve is close to that of the data. That is why the SDW prediction for the overall occupation (as a function of N_{tot}) is "deceptively" good!



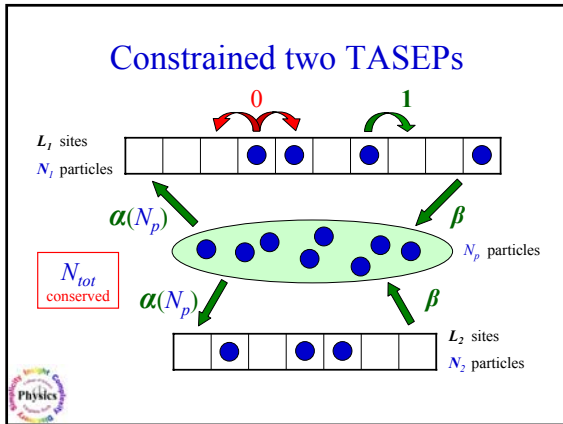
This GDW theory captures the essence of *shock localization*.



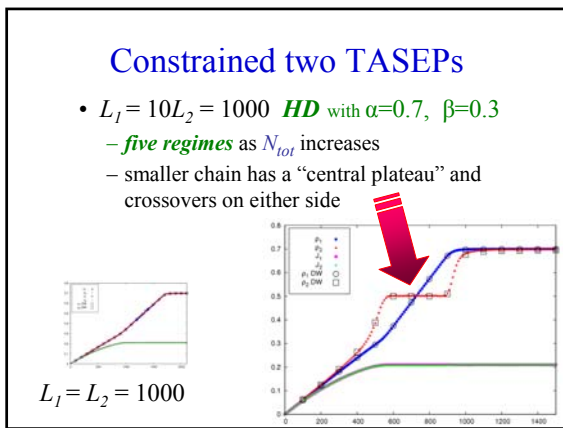
Constrained *single* TASEP

- Generalized DW theory is excellent at predicting properties of the stationary state.
- *Dynamics* remains more challenging, e.g., new *puzzles* associated with the *Power Spectrum* of $N(t)$.





- ### Constrained two TASEPs
- For $\alpha(N_p)$, keep *tanh* function with $N_x = \rho^*(L_1+L_2)/2$
 - $L_1 = L_2$ not much different from single case...
 - **except** for motion (anticorrelated) of DWs
 - de-localization; flat profiles recovered
 - “intrinsic” profile (from shifted averages) unchanged
 - $L_1 \neq L_2$ shows new feature...
 - for **HD**, there are **five regimes** as N_{tot} increases
 - smaller chain has a central plateau + 2 crossovers



- ### Constrained two TASEPs
- **Rough** intuitive picture for new feature
 - **Excellent** DW theory for two TASEPs
 - write master equation for $P(k_1, k_2; t)$
 - α_{eff} depends only on N_1+N_2 and so, $K \equiv k_1+k_2$
 - rates in ME satisfy detailed balance
 - stationary $P^*(k_1, k_2)$ “equilibrium-like” and so...
 - exactly solvable (with no fit parameters!)
 - **predicts** profiles, overall densities, etc.
 - generalized to **any number** of TASEPs

- ### Constrained many TASEPs
- DW theory **agrees well** with MC 3 TASEPs
 - No qualitatively new features ... at least so far!
 - Interesting analogy with canonical ensemble
 - constant $K (\equiv k_1+k_2+\dots+k_M)$ sheets have equal P^*
 - constant E sheets have equal $P^* : E \Leftrightarrow K ?$
 - T controls $E \Leftrightarrow N_{tot}$ controls $K ?$
 - If α of various TASEPs differ, then P^* is a genuinely **nonequilibrium** distribution.

- ### Recap of 3rd Extension
- “Finite resources” **provides many new interesting issues for TASEP**
 - **Some aspects understood; but puzzles remain**
 - **Many further extensions**, e.g., extended particles, non-uniform hopping rates, competition for other resources (e.g., aa-tRNA), ...
 - **Even wider afield**, e.g., multi-species, multi-lanes
 - **A long way to go, just to describe one cell**
 - ...let alone **many cells** that make up **one living being!**