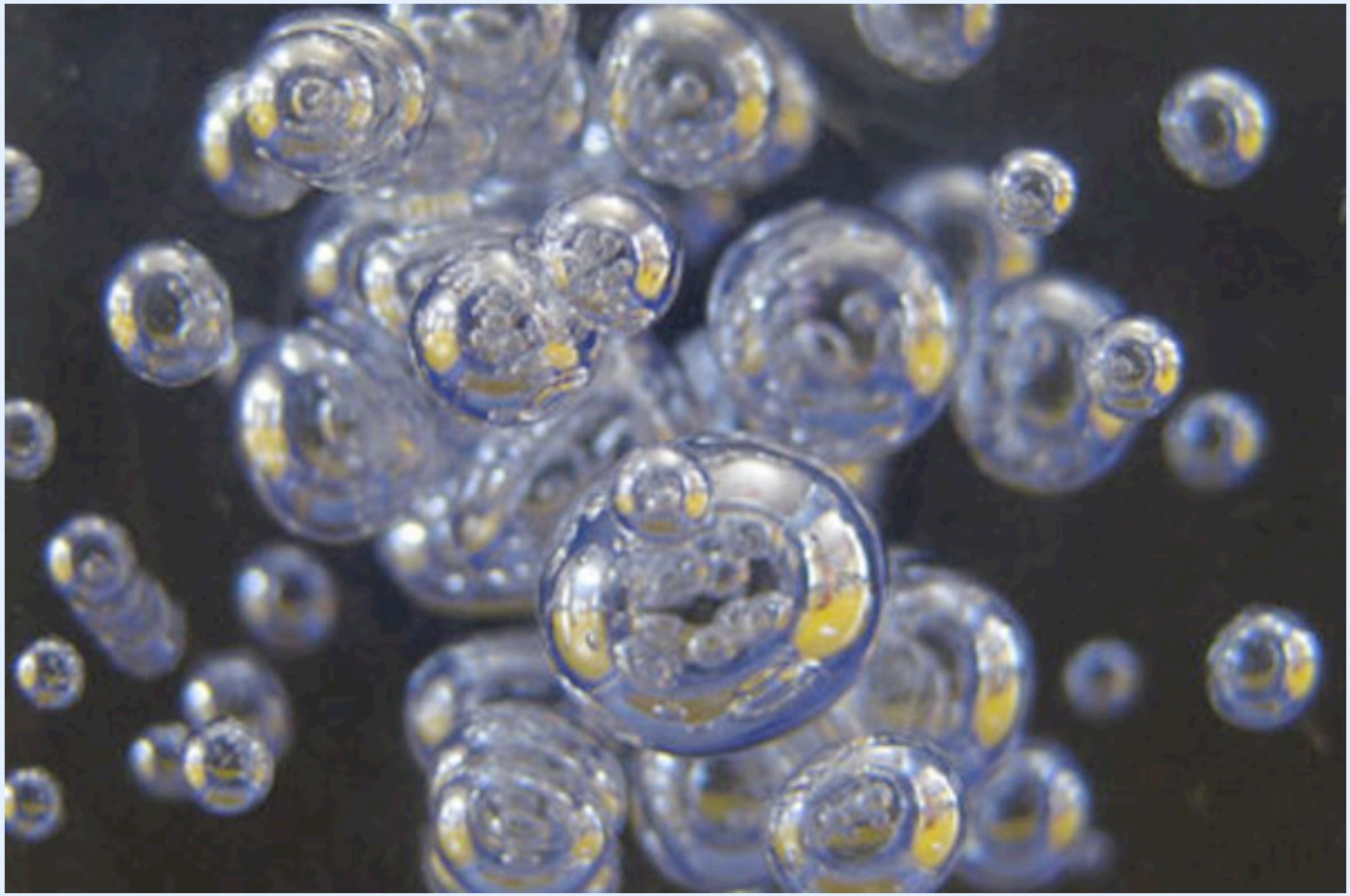
Stochastic Nucleation and Growth



Maria R. D'Orsogna, CSUN

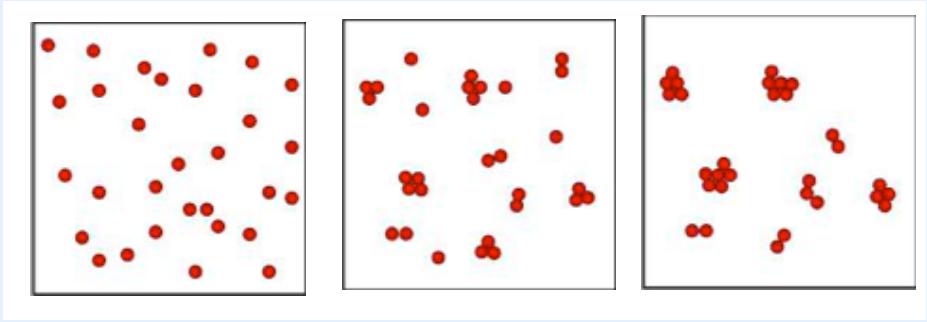
Tom Chou, UCLA - Greg Lakatos UBC

Nucleation and self assembly:

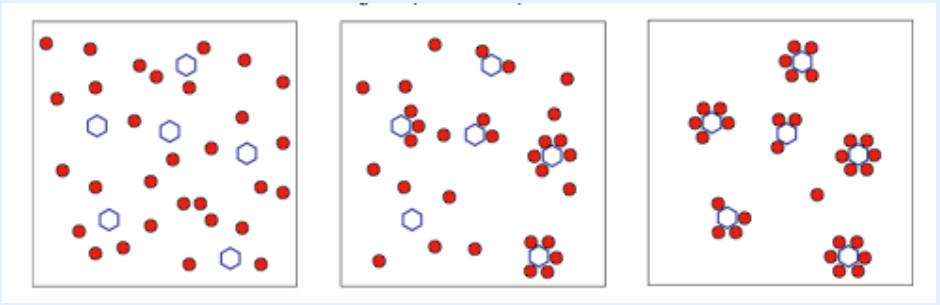
Bubbles, Crystals, Droplets, Glasses

Cloud condensation, Polymers, Semiconductors, Epitaxial Growth, Nanotechnology

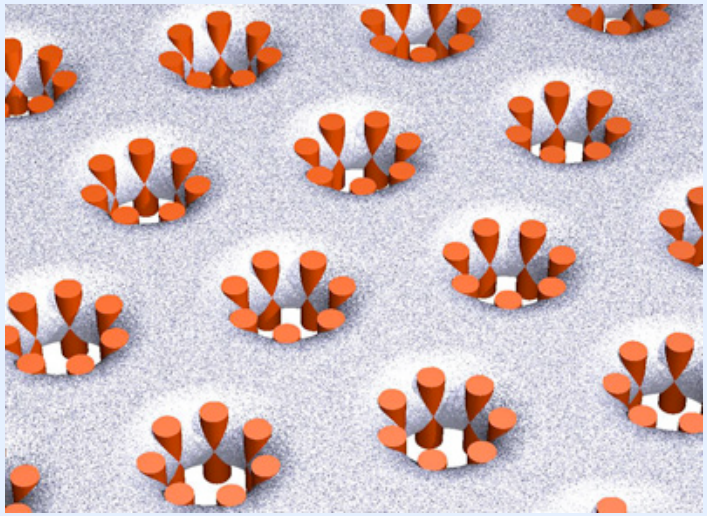
Homogeneous nucleation: spontaneous aggregation of particles



Heterogeneous nucleation: aggregation of particles triggered by impurities, boundaries, or special nucleation sites

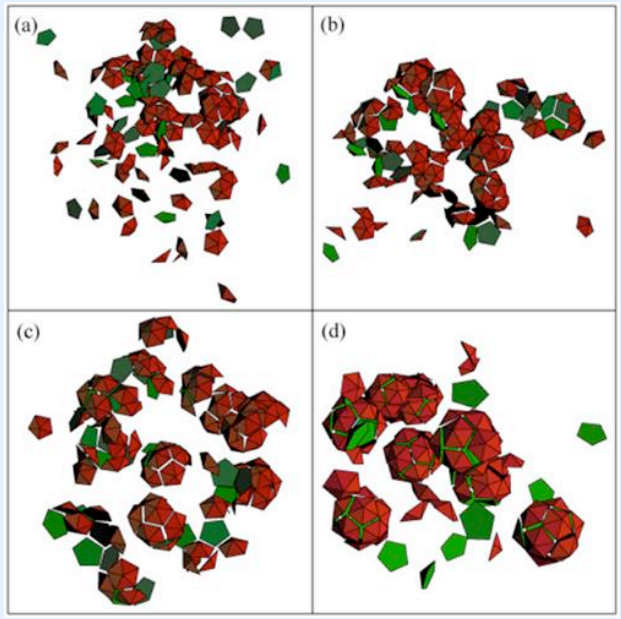


In Biology



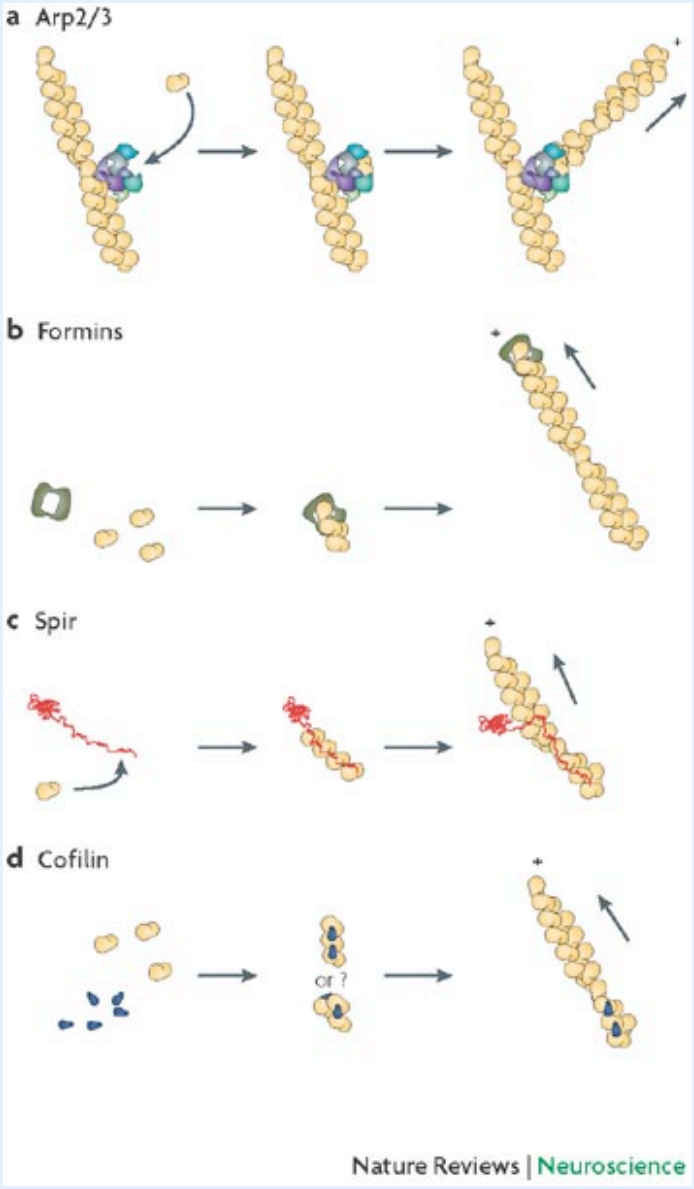
Membrane associated peptides
Self-assemble into pores
Transport

L.Yang et al.
Biophys J 2000



Viral capsid assembly
HIV, influenza

I.G. Johnston et al.
J Phys Cond Matt 2010



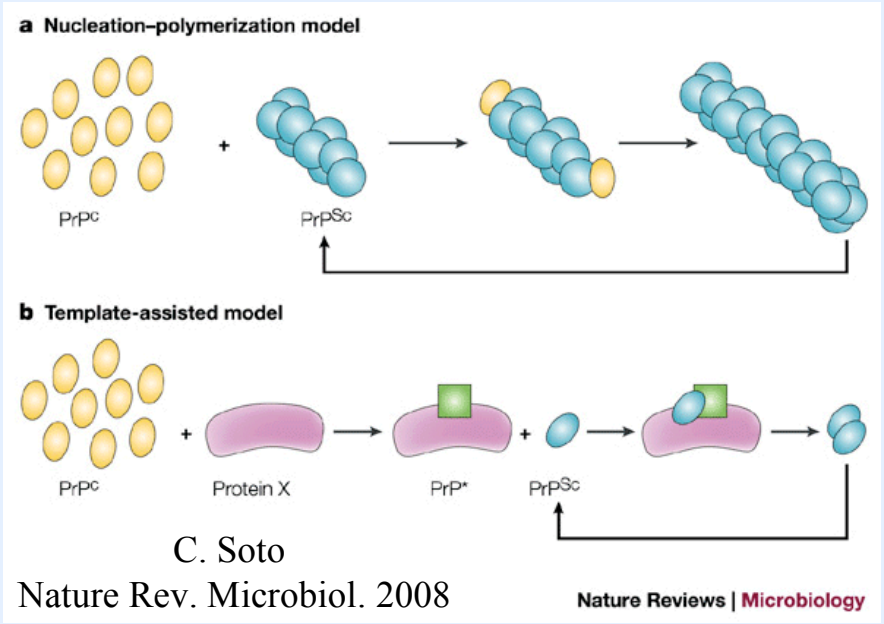
W. P. Chi et al.
Nature Rev. Neurosci. 2008

Actin nucleators to form branches, filaments

Cell motility, plaque formation

Misfolded proteins PrP^{Sc}
nucleate and grow into amyloids,
possibly via chaperones

Mad cow, Creutzfeldt-Jacob diseases



C. Soto
Nature Rev. Microbiol. 2008

Homogeneous nucleation:

Becker-Doering Mass Equations

$c_1, c_2, c_3 \dots c_k \dots c_N$	concentration of clusters with k particles
p, q	Monomer attachment, detachment rates
M	number of available monomers
N	maximum cluster size

$$0 \leq k \leq N$$

Within biology:

Aggregates have a maximum size and do not grow indefinetely

attachment and detachment are usually faster than production or degradation

Homogeneous nucleation:

Becker-Doering Mass Equations

$$\frac{dc_1}{dt} = -pc_1^2 - pc_1\sum_{j=2}^{N-1} c_j + 2qc_2 + q\sum_{j=3}^{N} c_j$$

$$\frac{dc_2}{dt} = -pc_1c_2 + \frac{1}{2}pc_1^2 - qc_2 + qc_3$$

$$\frac{dc_k}{dt} = -pc_1c_k + pc_1c_{k-1} - qc_k + qc_{k+1}$$

$$\frac{dc_N}{dt} = pc_1c_{N-1} - qc_N$$

We will rescale time (divide all by p)
and use $q/p = \varepsilon$

Homogeneous nucleation:

Becker-Doering Mass Equations

$$\frac{dc_1}{dt} = -c_1^2 - c_1 \sum_{j=2}^{N-1} c_j + 2\varepsilon c_2 + \varepsilon \sum_{j=3}^{N} c_j$$

$$\frac{dc_2}{dt} = -c_1 c_2 + \frac{1}{2} c_1^2 - \varepsilon c_2 + \varepsilon c_3$$

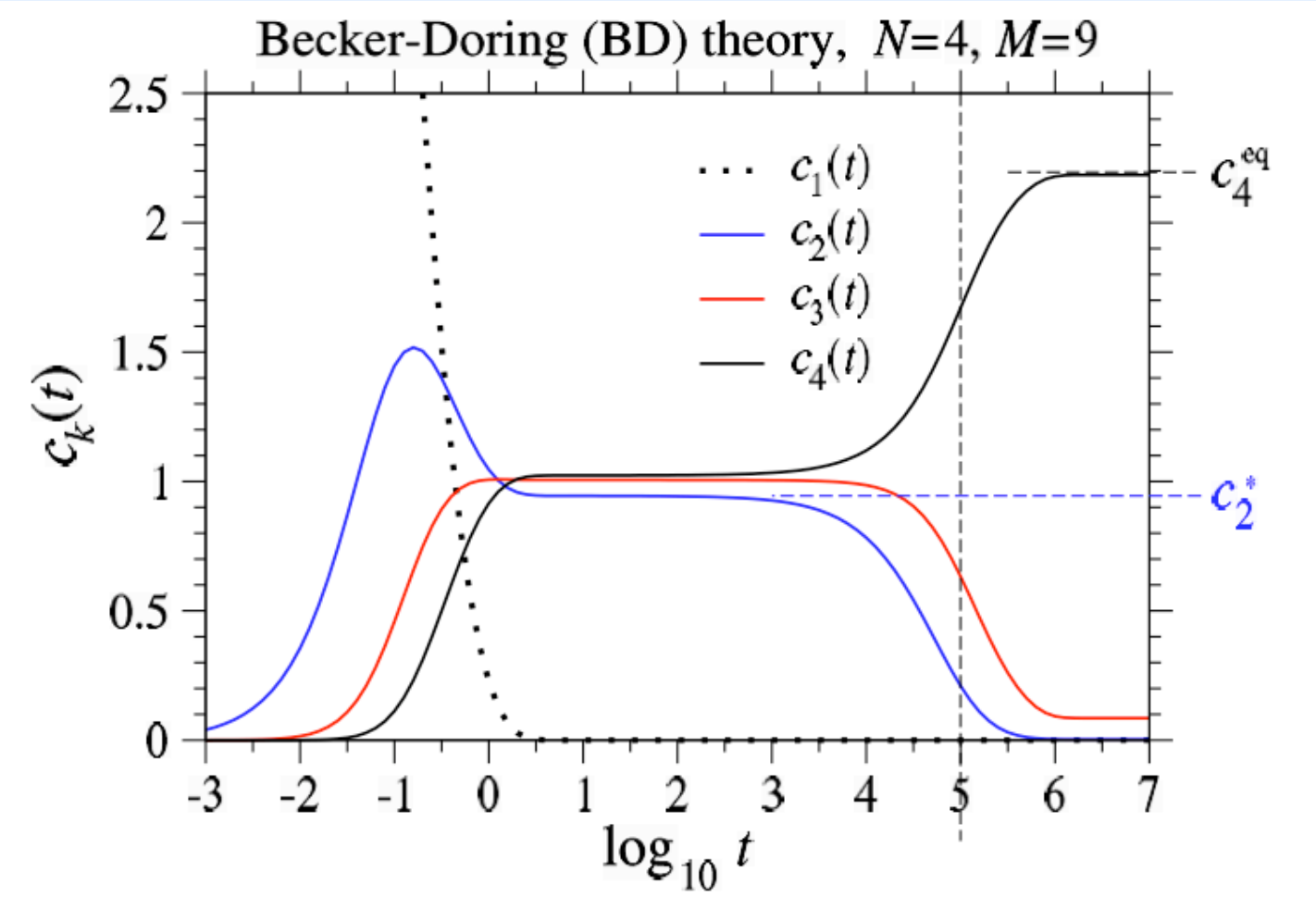
$$\frac{dc_k}{dt} = -c_1 c_k + c_1 c_{k-1} - \varepsilon c_k + \varepsilon c_{k+1}$$

$$\frac{dc_N}{dt} = c_1 c_{N-1} - \varepsilon c_N$$

We will rescale time (divide all by p) and use $q/p = \varepsilon$

Biologically relevant regime: $\varepsilon << 1$

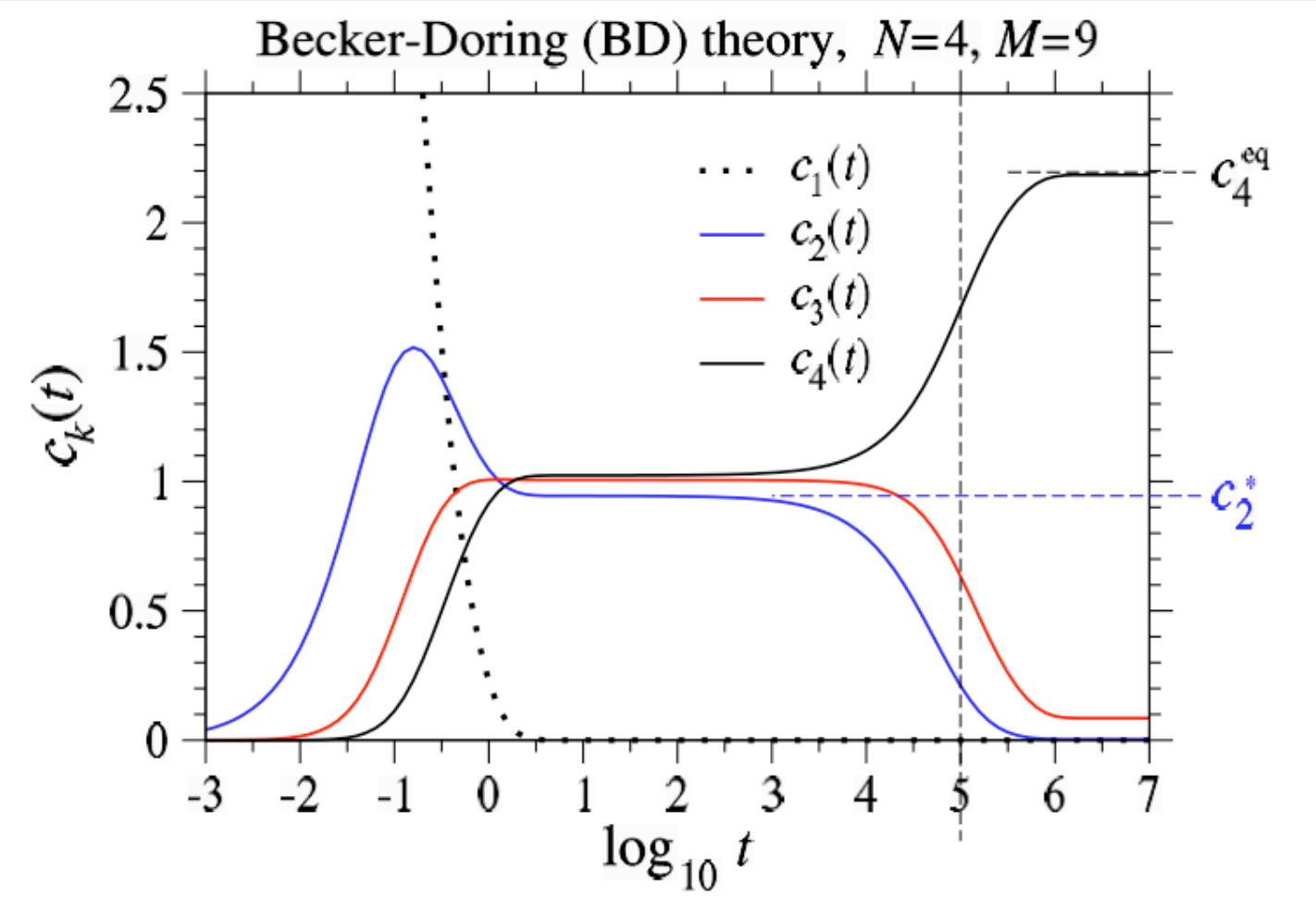
A Simple case:



M=9 monomers binding,
N=4 maximum cluster size

$\varepsilon = 10^{-5}$

A Simple case:



$\varepsilon = 10^{-5}$

Initially, attachment dominates, and we settle into a metastable regime for $t_c \sim 1/\varepsilon$

$$0 \rightarrow c_k^* \rightarrow c_k^{eq}$$

Estimates:

$$c_k^{\mathrm{eq}} \approx \frac{\varepsilon}{2}\left(\frac{2M}{\varepsilon N}\right)^{k/N} + O(\varepsilon^{1-k/N-1/N})$$

The equilibration values are small and scale as

$$c_k \sim \varepsilon^{1-k/N}$$

except

for the largest cluster c_N where

$$c_N \sim \frac{M}{N} + O(\varepsilon^{-1/N})$$

Is this all there is to it?

- Mass-action equations for mean size distribution used in many contexts:

  Virus capsid assembly: Morozov, Bruinsma, Rudnick, *J. Chem. Phys.*, **131**, 155101, 2009

  P. L. Krapivsky, E. Ben-Naim, and S. Redner, *Statistical Physics of Irreversible Processes*, CUP, 2010

- Extensive work on asymptotic analysis of mass-action, Becker-Döring eqs:

  P.-E. Jabin and B. Niethammer, *J. Differential Equations*, **191**, 518-543, 2003

  J. A. D. Wattis and J. R. King, *J. Phys. A: Math. Gen.*, **31**, 7169-7189, 1998

Stochastic treatment?

Very little on full stochastic analysis

J. S. Bhatt and I. J. Ford, *J. Chem. Phys.*, **118**, 3166-3176, 2003

F. Schweitzer *et al.*, *Physica A*, **150**, 261-279, 1988

Let's try to consider stochastic effects, discreteness, finite size by using a discrete Master Equation

Define $P(n_1, n_2, n_3, \ldots n_k, \ldots n_N, t)$

the probability of finding $n_1, n_2, n_3, \ldots n_k, \ldots n_N$, clusters of size 1, 2, …k, N concurrently and at time t.

n_1 = monomers

n_2 = dimers

…

Time evolution, leave state

$$\dot{P}(\{n\};t) = -\Lambda(\{n\})P(\{n\};t)$$

Leave state $P(\{n\},t) = P(n_1, n_2, n_3, \ldots n_k, \ldots n_N, t)$ in three ways:

1. Create a dimer by attachment of two monomers: $\binom{n_1}{2} = \frac{n_1(n_1-1)}{2}$ ways to do it

2. Create an n_{i+1}-mer by attachment of one monomer and one n_i-mer: $n_1 n_i$ ways to do it

3. Destroy an n_i-mer by detachment of a monomer: n_i ways to do it

$$\Lambda(\{n\}) = \frac{1}{2}n_1(n_1-1) + \sum_{i=2}^{N-1} n_1 n_i + \varepsilon \sum_{i=2}^{N} n_i,$$

Raising and lowering operators

W_i^+ particle attachment at cluster i

W_i^- particle detachment at cluster i

$$W_1^+ W_i^+ W_{i+1}^- P(\{n\},t)) = P(n_1+1, \ldots, n_i+1, n_{i+1}-1, \ldots n_{N,}t)$$

Use these operators to write down entry into state P({n},t)

Time evolution, populate state

$$\dot{P}(\{n\};t) = -\Lambda(\{n\})P(\{n\};t) + \tfrac{1}{2}(n_1+2)(n_1+1)W_1^+W_1^+W_2^-P(\{n\};t)$$
$$+\varepsilon(n_2+1)W_2^+W_1^-W_1^-P(\{n\};t)$$
$$+\sum_{i=2}^{N-1}(n_1+1)(n_i+1)W_1^+W_i^+W_{i+1}^-P(\{n\};t)$$
$$+\varepsilon\sum_{i=3}^{N}(n_i+1)W_1^-W_{i-1}^-W_i^+P(\{n\};t),$$

Time evolution, populate state

$$\dot{P}(\{n\};t) = -\Lambda(\{n\})P(\{n\};t) + \tfrac{1}{2}(n_1+2)(n_1+1)W_1^+W_1^+W_2^-P(\{n\};t)$$
$$+\varepsilon(n_2+1)W_2^+W_1^-W_1^-P(\{n\};t)$$
$$+\sum_{i=2}^{N-1}(n_1+1)(n_i+1)W_1^+W_i^+W_{i+1}^-P(\{n\};t)$$
$$+\varepsilon\sum_{i=3}^{N}(n_i+1)W_1^-W_{i-1}^-W_i^+P(\{n\};t),$$

Start from state $W_2^+ W_1^- W_1^- P(\{n\},t))$:
with one extra dimer n_2+1,
two less monomers n_1-2

any dimer can detach: form a state with n_2 dimers, n_1 monomers

(n_2+1) ways to pick a dimer to split, detachment rate ε

Full Stochastic equation:

$$\dot{P}(\{n\};t) = -\Lambda(\{n\})P(\{n\};t) + \tfrac{1}{2}(n_1+2)(n_1+1)W_1^+W_1^+W_2^-P(\{n\};t)$$
$$+\varepsilon(n_2+1)W_2^+W_1^-W_1^-P(\{n\};t)$$
$$+\sum_{i=2}^{N-1}(n_1+1)(n_i+1)W_1^+W_i^+W_{i+1}^-P(\{n\};t)$$
$$+\varepsilon\sum_{i=3}^{N}(n_i+1)W_1^-W_{i-1}^-W_i^+P(\{n\};t),$$

Initial condition $P(\{n\},t=0) = \delta_{n1,M}\,\delta_{n2,0}\ldots\delta_{nk,0}\,\delta_{nN,0}$

Mass Conservation

$$M = \sum_{j=1}^{N} jn_j$$

Connection to Becker Doering:

Define mean particle size

$$\langle n_k(t)\rangle \equiv \sum'_{\{n_j\}} n_k P(\{n\};t)$$

$$\begin{aligned}
\langle \dot{n}_1(t)\rangle &= -2\left\langle \frac{n_1(n_1-1)}{2}\right\rangle - \sum_{j=2}^{N-1}\langle n_1 n_j\rangle + 2\varepsilon\langle n_2(t)\rangle + \varepsilon\sum_{j=3}^{N}\langle n_j\rangle \\
\langle \dot{n}_2(t)\rangle &= -\langle n_1 n_2\rangle + \left\langle \frac{n_1(n_1-1)}{2}\right\rangle + \varepsilon\langle n_3\rangle - \varepsilon\langle n_2\rangle \\
\langle \dot{n}_k(t)\rangle &= -\langle n_1 n_k\rangle + \langle n_1 n_{k-1}\rangle - \varepsilon\langle n_k\rangle + \varepsilon\langle n_{k+1}\rangle \\
\langle \dot{n}_N(t)\rangle &= \langle n_1 n_{N-1}\rangle - \varepsilon\langle n_N\rangle.
\end{aligned}$$

Connection to Becker Doering:

Define mean particle size

$$\langle n_k(t)\rangle \equiv \sum'_{\{n_j\}} n_k P(\{n\};t)$$

$$\begin{aligned}
\langle \dot{n}_1(t)\rangle &= -2\left\langle \frac{n_1(n_1-1)}{2}\right\rangle - \sum_{j=2}^{N-1}\langle n_1 n_j\rangle + 2\varepsilon\langle n_2(t)\rangle + \varepsilon\sum_{j=3}^{N}\langle n_j\rangle \\
\langle \dot{n}_2(t)\rangle &= -\langle n_1 n_2\rangle + \left\langle \frac{n_1(n_1-1)}{2}\right\rangle + \varepsilon\langle n_3\rangle - \varepsilon\langle n_2\rangle \\
\langle \dot{n}_k(t)\rangle &= -\langle n_1 n_k\rangle + \langle n_1 n_{k-1}\rangle - \varepsilon\langle n_k\rangle + \varepsilon\langle n_{k+1}\rangle \\
\langle \dot{n}_N(t)\rangle &= \langle n_1 n_{N-1}\rangle - \varepsilon\langle n_N\rangle.
\end{aligned}$$

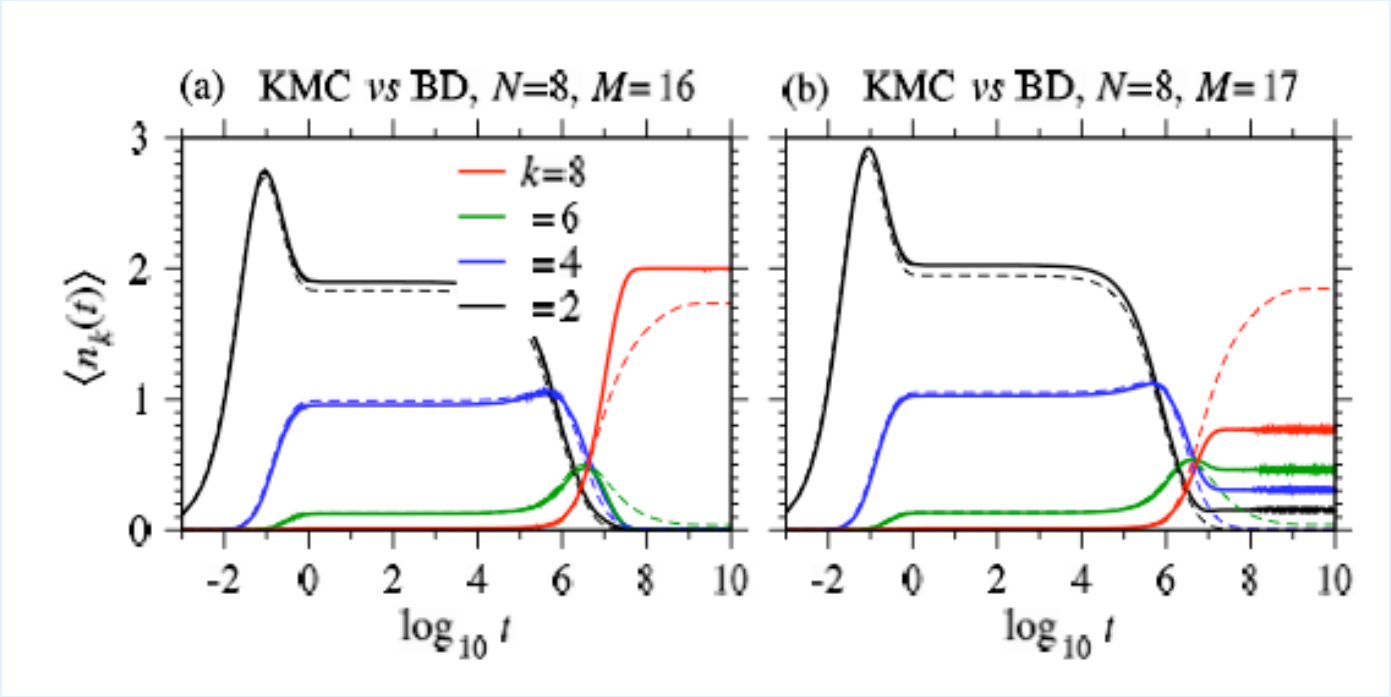
Reduces to BD for large n_1, mean field

$$\langle n_i n_j\rangle = \langle n_i\rangle\langle n_j\rangle \Rightarrow c_i c_j,$$
$$\langle n_1 - 1\rangle \approx \langle n_1\rangle \Rightarrow c_1.$$

Simulate and compare

Kinetic Monte Carlo /Master Equation vs. Becker Doering

$\varepsilon = 10^{-5}$



c_k (BD, dash) does not compare well with $\langle n_k \rangle$ (KMC, solid) at long times

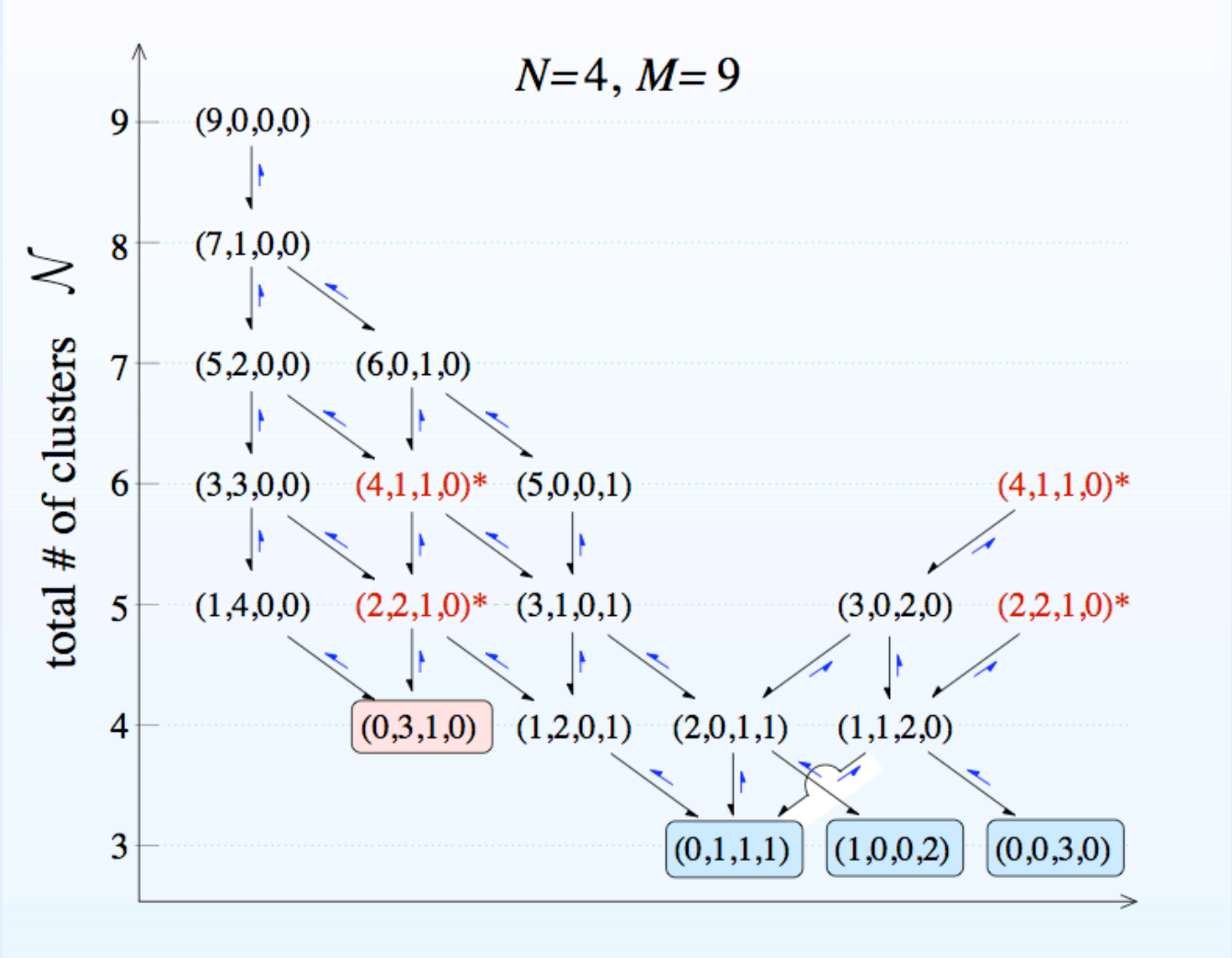
$$c_k^{eq} \sim \varepsilon^{1-k/N}$$

$$c_N^{eq} \sim \frac{M}{N} + O(\varepsilon^{-1/N})$$

$$c_k^{eq} \quad vs. \quad \left\langle n_k^{eq} \right\rangle$$

?

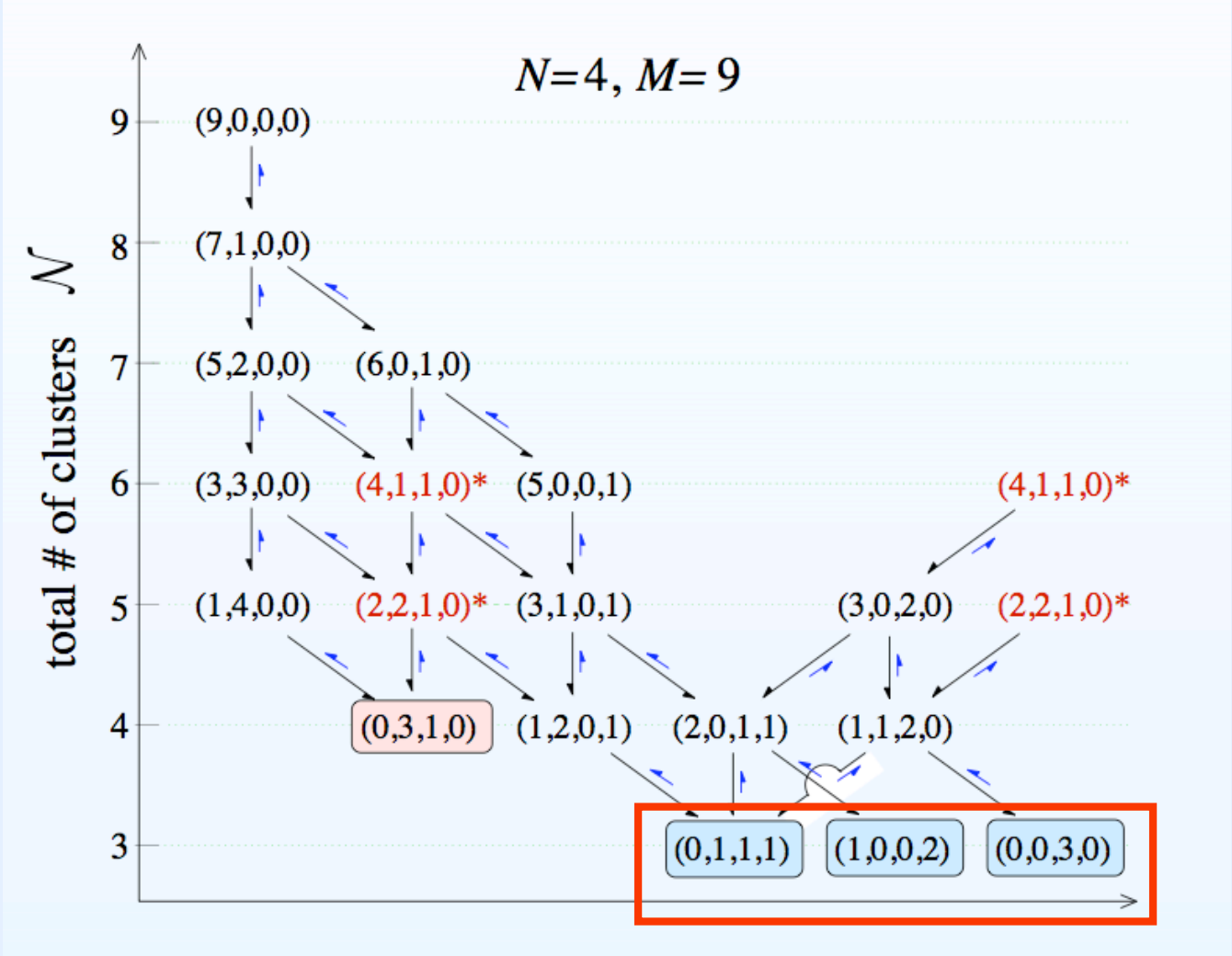
Let's calculate $\langle n_k^{eq} \rangle$ for small ε?



Detachment slow: will have few clusters at equilibrium

Find how many states with fewest number of clusters, $\mathscr{N} = 3$

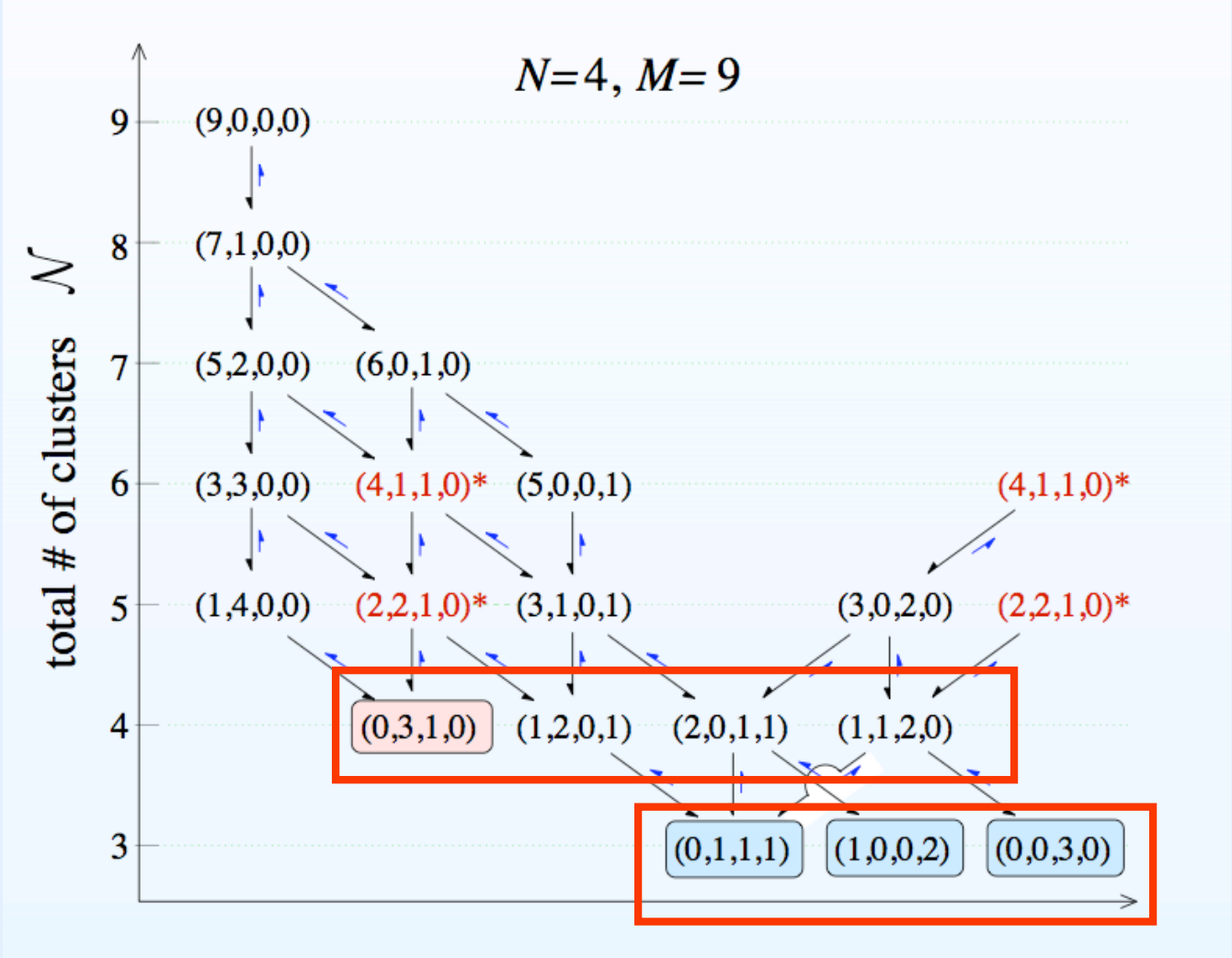
Can we calculate $\langle n_k^{eq} \rangle$ for small ε?



Detachment slow: will have few clusters at equilibrium

Find how many states with fewest number of clusters, $\mathscr{N} = 3$

Can we calculate $\langle n_k^{eq} \rangle$ for small ε?



Apply detailed balance with the $\mathcal{N}$+1 states, here $\mathcal{N}$ = 4

Can calculate highest order values of $P(n_1,n_2,n_3,n_4, t=\infty)$

Can do for general M,N

Exact $\langle n_k^{eq} \rangle$ for small ε!

$$M = \sigma N - j \;\; (0 \leq j \leq N-1)$$

Introduce σ

σ – 1 = largest integer divisor
j = reminder, incommensurability

$$0 \leq j < N-1$$

$$\langle n_N^{\rm eq} \rangle = \frac{\sigma(\sigma-1)}{(\sigma+j-1)} + O(\varepsilon)$$

$$\langle n_{N-k}^{\rm eq} \rangle = \frac{\sigma(\sigma-1)j(j-1)\ldots(j-k+1)}{(\sigma+j-1)(\sigma+j-2)\ldots(\sigma+j-k-1)} + O(\varepsilon)$$

Special case of one extra monomer

$$j = N - 1$$

$$M = \sigma N - N + 1 = (\sigma - 1)N + 1$$

$$\langle n_1^{\mathrm{eq}} \rangle = \frac{2(N-1)!}{D(\sigma, N-1)} + O(\varepsilon)$$

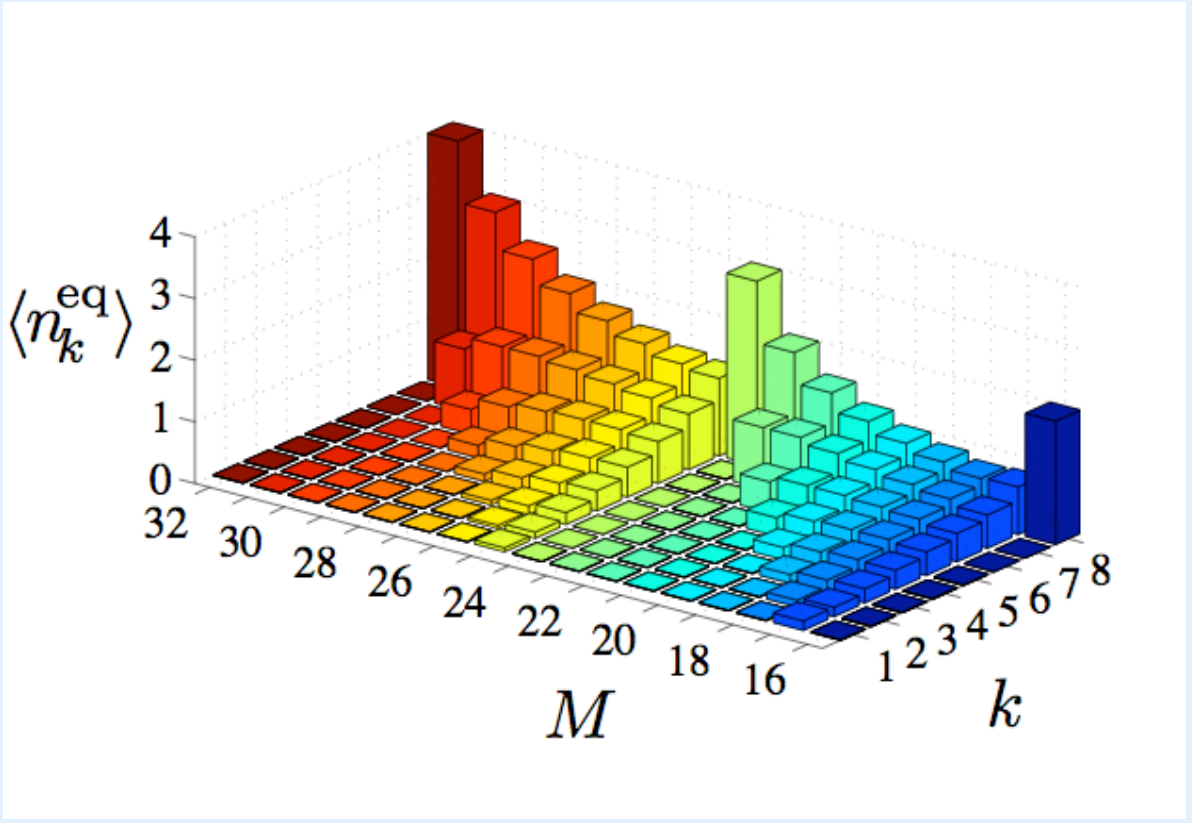
$$\langle n_{N-k}^{\mathrm{eq}} \rangle = \frac{\prod_{\ell=1}^{k}(N-\ell)\prod_{i=1}^{N-k-1}(\sigma-2+i)}{D(\sigma, N-1)} + O(\varepsilon)$$

$$\langle n_N^{\mathrm{eq}} \rangle = (\sigma-1)\frac{D(\sigma-1, N-1)}{D(\sigma, N-1)} + O(\varepsilon),$$

$$D(\sigma, j) = j! + \prod_{\ell=1}^{j-1}(\sigma + \ell)$$

KMC sims agree with analytics

cluster dispersal



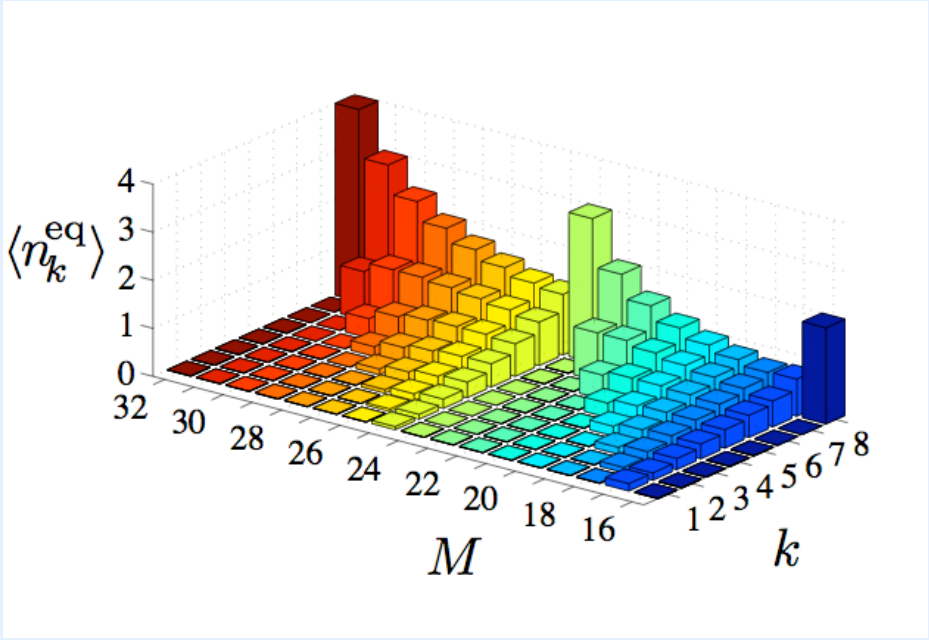
Mass in largest cluster only when M divisible by N

Why does this happen?

N=8

M=16 vs. M=17

small ε: few clusters



M=16 is exactly divisible by N=8,

fewest clusters: 2, no remainders

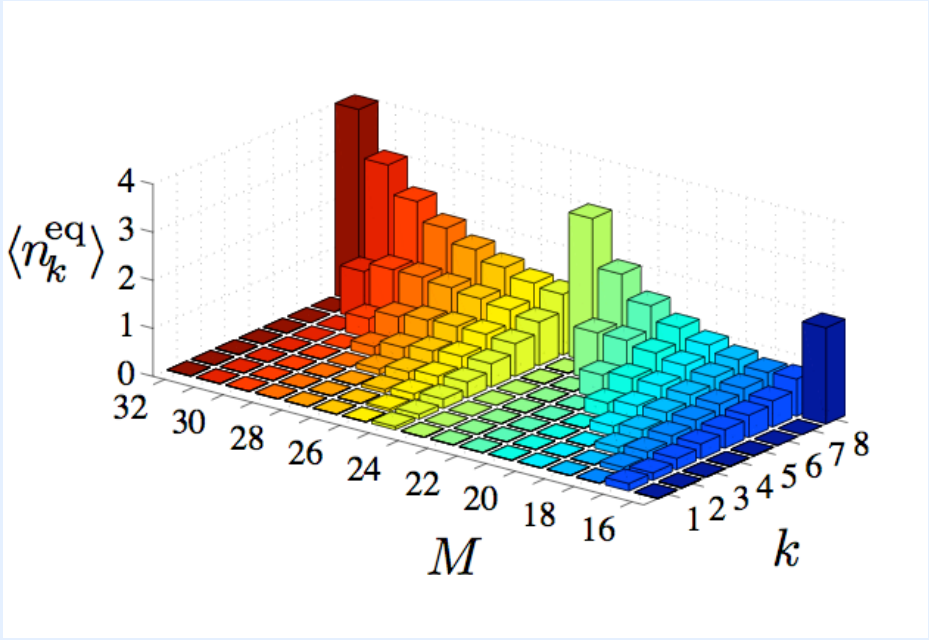
$(0,0,0,0,0,0,0,2)$ is the only relevant state, for small ε

Why does this happen?

N=8

M=16 vs. M=17

small ε: few clusters



M=17 is not divisible by 8, fewest clusters: 3

$$(1, 0, 0, 0, 0, 0, 0, 2)$$

But also:

Add one particle: emulsification

$(0,1,0,0,0,0,1,1)$ $(0,0,0,1,0,1,1,0)$ $(0,0,1,0,0,0,2,0)$ $(0,0,0,0,1,2,0,0)$

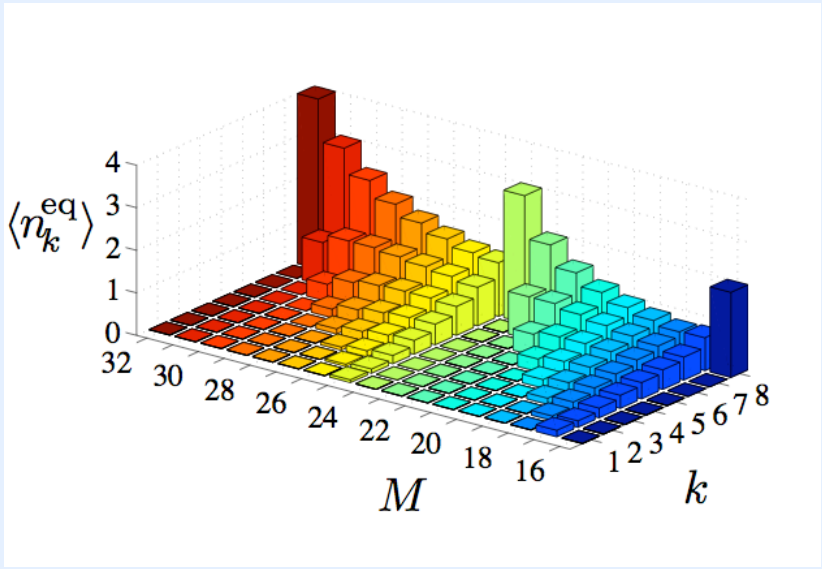
$(0,0,1,0,0,1,0,1)$ $(0,0,0,0,2,0,1,0)$

$(0,0,0,1,1,0,0,1)$

$+ \quad (1,0,0,0,0,0,0,2)$

8 STATES WITH 3 CLUSTERS

SIMILAR WEIGHTS ⇒ FLAT DISTRIBUTION!



Regimes of validity of Becker Doering?

equilibrium cluster numbers ($\varepsilon \ll 1$)	$\frac{M}{N} \to 0$	$\frac{M}{N}$ finite	$\frac{M}{N} \gg N$
BD ($N = \infty$)	MFT*	×	×
BD (finite N)	MFT*	MFT	MFT†
discrete model	exact*	exact	exact†

match

match

DO not match

Becker Doering not valid for M~N,
Even if both M,N are large

Summary and Applications:

- Full stochastic, discrete model for homogeneous nucleation and growth

- Derived exact equilibrium solution for average population sizes

- Mean cluster size distribution broadens if size N and mass M are incommensurate even if M,N are large, as long as M/N is finite

-Variance of cluster sizes

-Non ergodic case ($\varepsilon = 0$)

-Mean first passage times and first passage probability distributions to clusters of size N, relevant for plaque formation, viral dynamics

-Heterogeneous nucleation