Mean-field Methods for Large Stochastic Systems

with application to bike-sharing systems

Nicolas Gast (Inria, Grenoble)

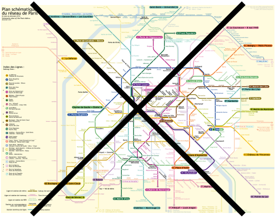
Inria Nancy seminar – May 7, 2015

We want to study large stochastic systems composed of many interacting objects

Example: (computer networks, biological models,...)...

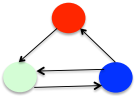
We want to study large stochastic systems composed of many interacting objects

Example: (computer networks, biological models,...)... bike-sharing systems

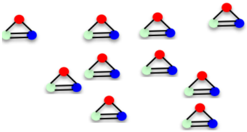


Markovian models suffers from the curse of dimensionality.

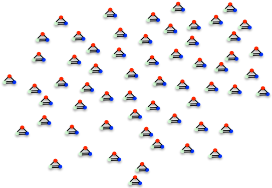
The state space grows exponentially with the number of objects.



1 object
3 states

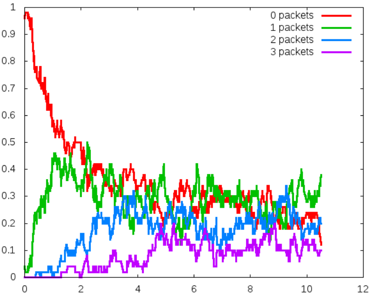


10 objects
$3^{10} \approx 10^5$ states

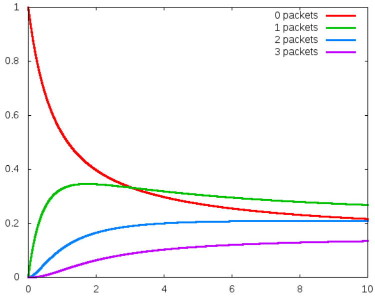


60 objects
$3^{60} \approx 10^{28}$ states

Dynamic systems can be modeled by using stochastic or deterministic models



CTMC



ODE

Outline

1. Mean-field interaction model
2. Finite time-horizon: convergence to ODE
3. Infinite time-horizon: steady-state and fixed-point method
4. Example: application to bike sharing systems
5. Conclusion

Outline

Mean-field interaction model

- Time is discrete
- N objects
- Object n as state $X_n(t)$
- $(X_1(t) \dots X_N(t))$ is Markov
- Objects are observable through their state only.

"Occupancy measure": $M^N(t)$ = distribution of object states at time t.
Theorem [G2012a]. $M^N(t)$ is Markov.

Example: Epidemics (SIR model)

Mobile nodes are:

- S: susceptible
- I: Infected
- R: Recovered

Occupancy measure is

$$M(t) : (S(t), I(t), R(t))$$

with $S(t) + I(t) + R(t) = 1$.

1. Direct Infection:

$$S \to I$$

2. Infection by others:

$$S + I \to I + I$$

3. Recovery:

$$I \to R$$

4.

$$R \to S$$

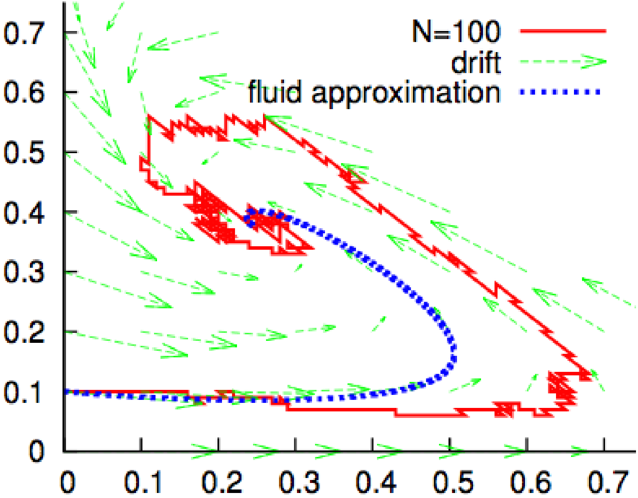
Example: Epidemics (SIR model)

Each time, a node is chosen:

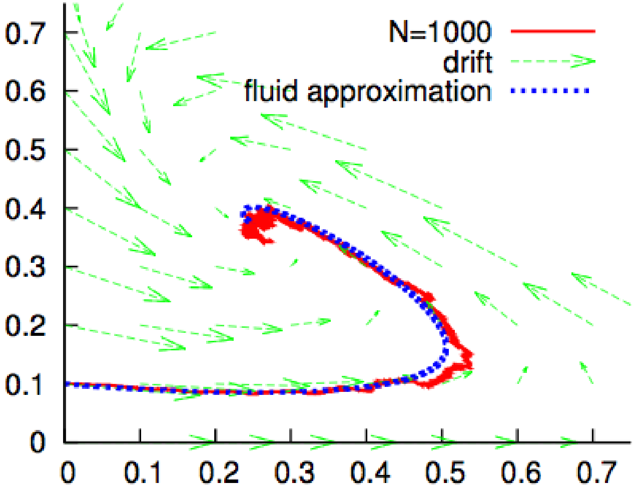
- If the node is in state 'S':
  1. He becomes I with probability α
  2. With probability $\beta N I(t)/(N-1)$ and becomes I.
- If the node is 'I':
  2. With probability $\beta N S(t)/(N-1)$, he meets an S and the S becomes I.
  3. With probability γ, he becomes R
- If the node is 'R':
  4. With probability δ, he becomes S.

1. Direct Infection:
$$S \to I$$
2. Infection by others:
$$S + I \to I + I$$
3. Recovery:
$$I \to R$$
4.
$$R \to S$$

Simulation with $N = 100$



Simulation with $N = 1000$



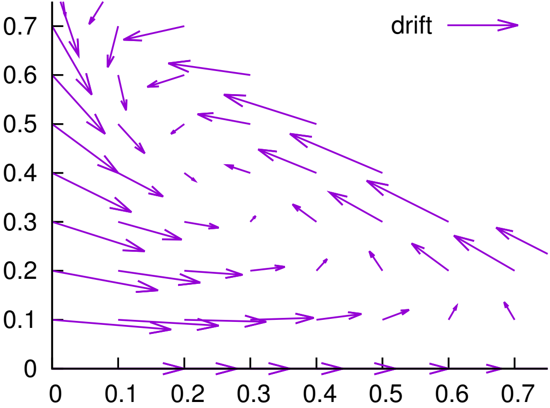
When are these approximation valid?

- Asymptotic for large N:
  - Fluid limit
  - Fast-simulation
- Asymptotic for large time-horizon:
  - Fixed-point method

Outline

We can construct the ODE by using the drift

The drift is $f^N(m) = \frac{1}{N}\mathbf{E}\left[M^N(t+1) - M^N(t) | M^N(t) = m\right]$.

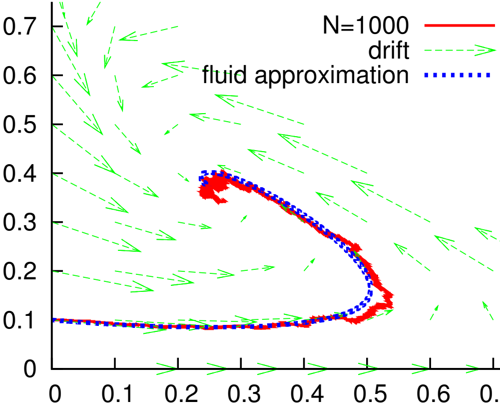


The mean-field limit

Under very general condition (given later), the occupancy measure $M^N(t)$ converges (in probability) to a deterministic process, $m(t)$, called the mean-field limit:

$$M^N(Nt) \to m(t),$$

Finite state space = ODE $dm/dt = f(m)$.



Sufficient convergence as verifiable by inspection

Theorem (Benaïm-Le Boudec 2008)

Assume that:

- *The number of object that change state has a bounded second moment.*
- *The drift converges uniformly to a Lipschitz function:* $f^N \to f$
- *The state space is finite.*

Then, uniformly for all t*:*

$$M^N(Nt) \to m(t),$$

in probability.

The proof is based on stochastic approximation $(x_{n+1} = x_n + \varepsilon\,(f(x_n) + u_{n+1}))$

$$M^N(t+1) = M^N(t) + \frac{1}{N}\left(f^N(M^N(t)) + \underbrace{N\left(M^N(t+1) - M^N(t)\right) - f^N(M^N(t))}_{\mathbf{E}[\cdot|\mathcal{F}_t]=0} \right)$$

The computation of the drift can be automated

$$Drift = \sum_{transitions} \text{Delta to } M^N(transition) \times Proba(transition)$$

Proba	Effect on $M^N = (S, I, R)$
αS	$\frac{1}{N}(-1, 1, 0)$
$\beta 2SI\frac{N}{N-1}$	$\frac{1}{N}(-1, 1, 0)$
γI	$\frac{1}{N}(0, -1, 1)$
δR	$\frac{1}{N}(1, 0, -1)$

$$\begin{aligned} \frac{dS}{dt} &= -\alpha S - 2\beta SI + \delta R \\ \frac{dI}{dt} &= \alpha S + 2\beta SI - \gamma I \\ \frac{dR}{dt} &= \gamma I - \delta R \end{aligned}$$

What is the relation between mean-field and the decoupling assumption?

- Decoupling = Objects are asymptotically independent.

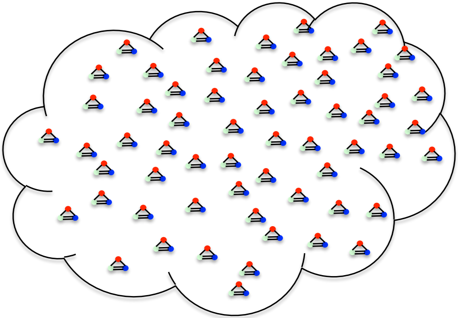
What is the relation between mean-field and the decoupling assumption?

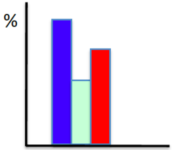
- Decoupling = Objects are asymptotically independent.

"Theorem" [Snitman 91]: For a mean-field interaction model, decoupling $\equiv M^N(t)$ converges to a deterministic limit.

The two sides of mean-field limit

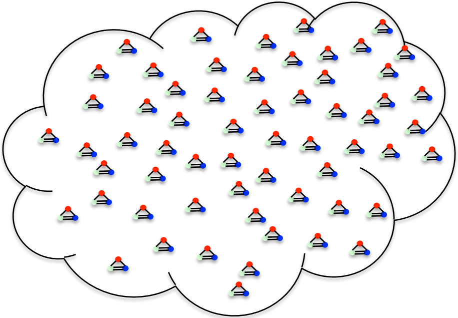
Side 1 : fluid limit

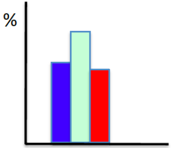




The two sides of mean-field limit

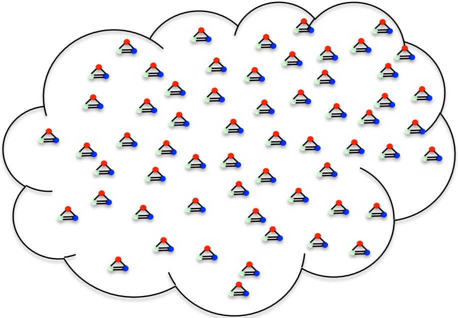
Side 1 : fluid limit

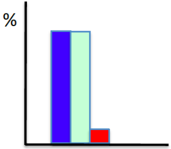




The two sides of mean-field limit

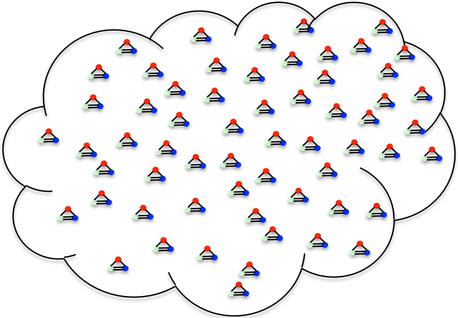
Side 1 : fluid limit

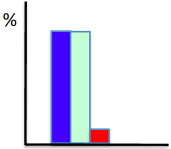




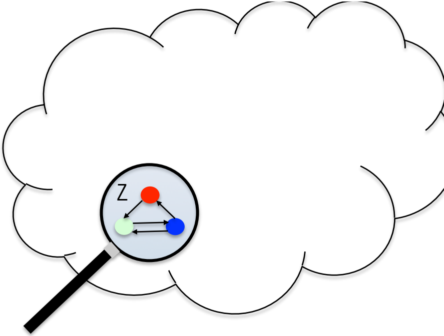
The two sides of mean-field limit

Side 1 : fluid limit





Side 2: decoupling



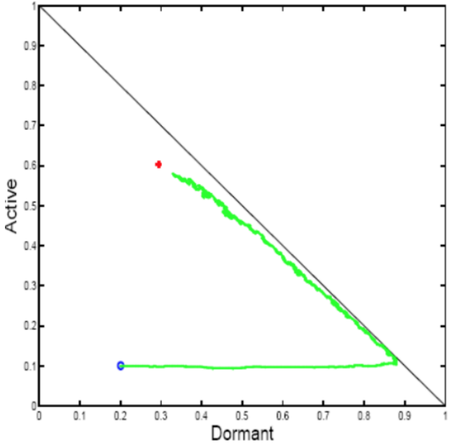
An object is considered in the mean field created by the rest (its dynamics is represented as a time-inhomogeneous CTMC.)

Outline

The fixed-point method

Method classically used:

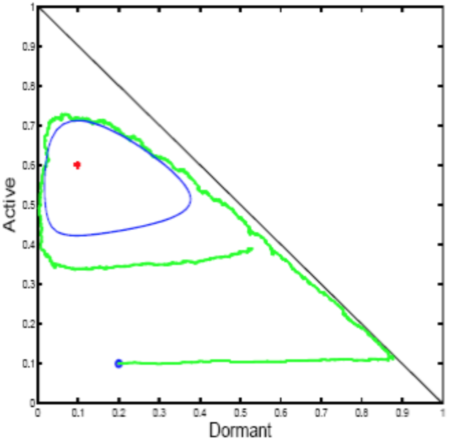
- We solve $f(m^*) = 0$.
- The steady-state is $\approx m^*$.



The fixed-point method does not work in general

Method classically used:

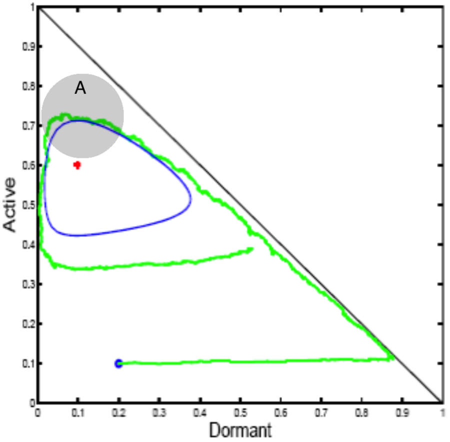
- We solve $f(m^*) = 0$.
- The steady-state is $\approx m^*$.



The fixed-point method does not work in general

Method classically used:

- We solve $f(m^*) = 0$.
- The steady-state is $\approx m^*$.



When the fixed point method fails, **the decoupling assumption does not hold**. : if we observe one node "active", then we are likely to be in region A. Another node is likely to be active.

A positive result

Theorem

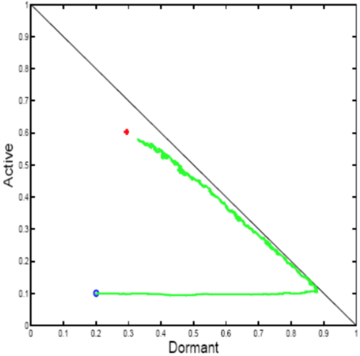
If the ODE has a unique fixed point m^* to which each trajectory converges, then the stationary measure concentrates on m^*

Problem: the asymptotic behavior of an ODE cannot be predicted from its structure.

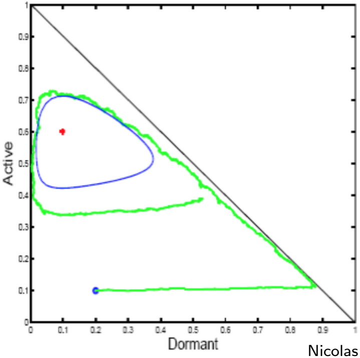
$$\begin{aligned}
\frac{\partial D}{\partial t} &= -\delta_D D - 2\lambda D^2 - \beta A \frac{D}{h+D} + (\alpha_0 + rD)S \\
\frac{\partial A}{\partial t} &= 2\lambda D^2 + \beta A \frac{D}{h+D} - \delta_A A + \alpha S \\
\frac{\partial S}{\partial t} &= \delta_D D + \delta_A A - (\alpha_0 + rD)S - \alpha S
\end{aligned}$$

$\beta = 0.01$,
$\delta_A = 0.005$,
$\delta_D = 0.00001$,
$\alpha = \alpha_0 = 0.00001$,
$r = 0.1$, $\lambda = 0.0001$

$h = 0.3$



$h = 0.1$



Mean-field approximation in short

Finite-horizon : In general: $M^N(t)$ converges to $dm/dt = f(m)$.

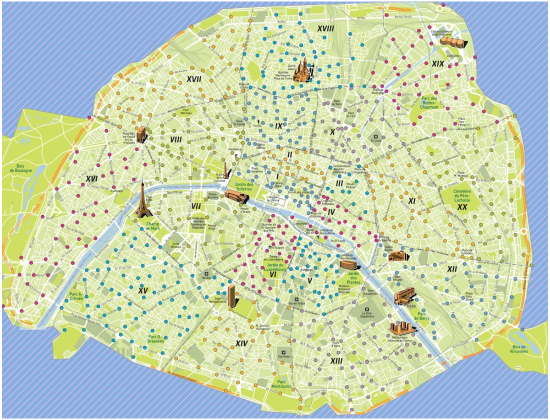
- Conditions can be verified by a direct inspection.
- Works for discontinuous f (differential inclusion $\dot{m} \in F(m)$ [Gast2012b])
- Works for controlled dynamics (HJB and Bellman equation [Gast2012a,Gast2014])
- Speed of convergence: $O(1/\sqrt{N})$

Infinite horizon : conditions are hard to verify.

- Fixed point works very well in practice but no guarantee.
- Fixed point always work when the process is reversible.
- No speed of convergence.

Outline

Bike-sharing are **large** stochastic systems



Map of Velib' stations in Paris (France).

Example of Velib':

- 20000 bikes
- 1200 stations.

The main problem is the lack of resource

(a) Empty station

(b) Full station

Problematic states

The system's operator want to anticipate and avoid those states.

State of the art

Visualization of existing systems

- Traces analysis, clustering (Borgnat et al. 10, Vogel et al. 11, Nair et al. 11, Côme et al. 13. . .)

Short-term / mid-term prediction of availability

- (Ji Won Yoon et al. 12, Guenther et al. 12)

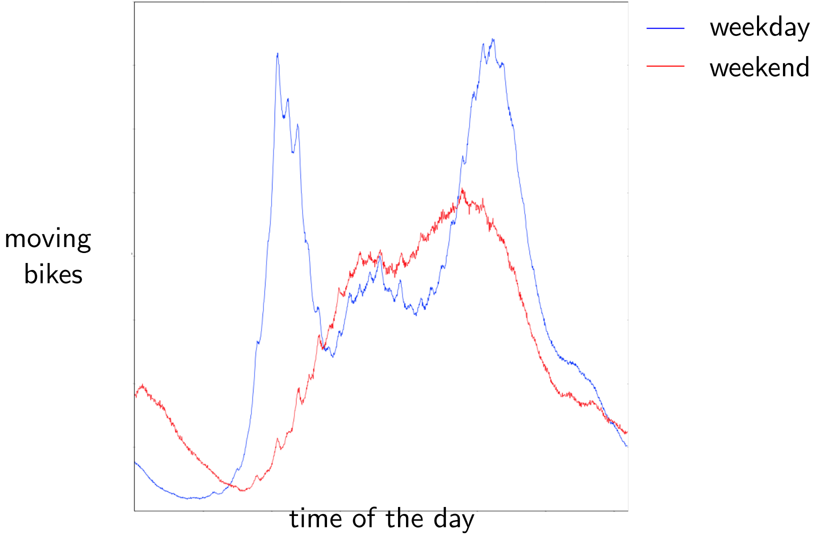
Bike re-positioning (classical RO problem)

- Redistribution based of forecast [Raviv et al. 11, Chemla et al. 13, Pfrommer 13,. . .]

Planing using macroscopic data

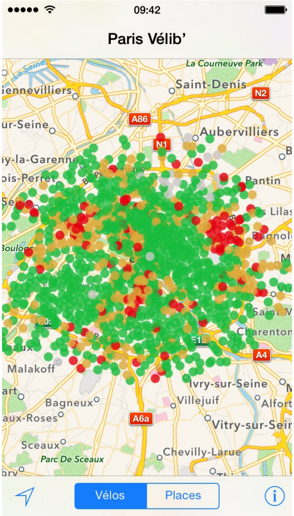
Visualizing the data: usage varies (data from paris, 2014)

Example : temporal variation



Visualizing the data: usage varies (data from paris, 2014)

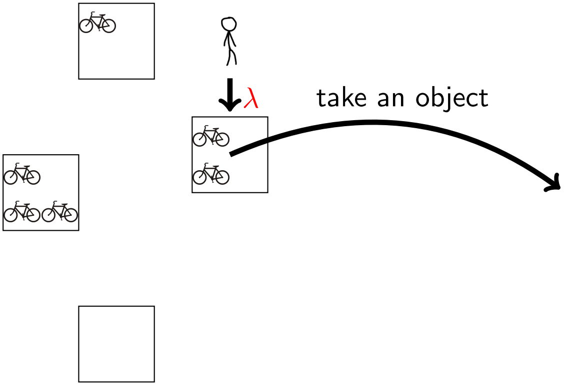
Example: spatial variation



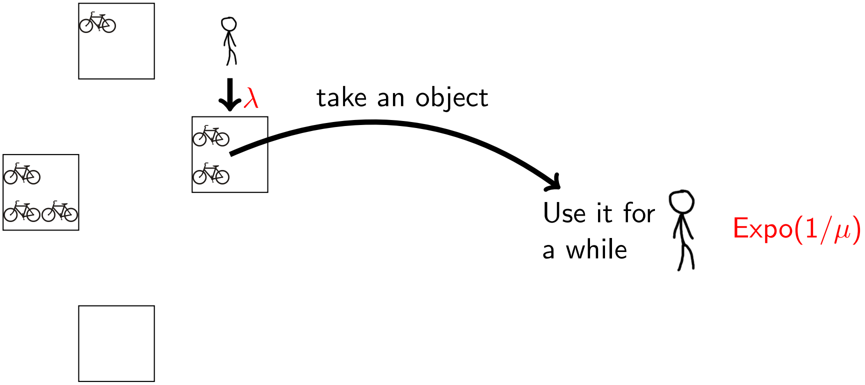
Solution: clustering?

Source: http://www.bicyclette-app.com/fr/

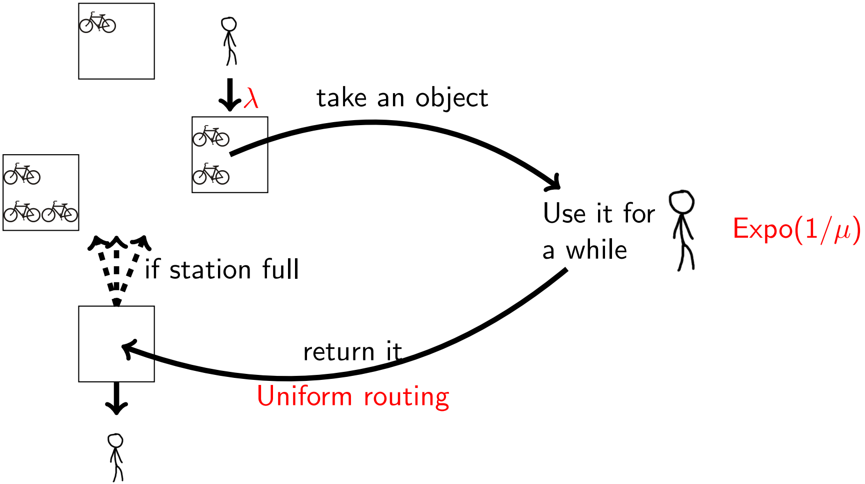
Uniform Bike-sharing systems as closed-queuing networks



Uniform Bike-sharing systems as closed-queuing networks

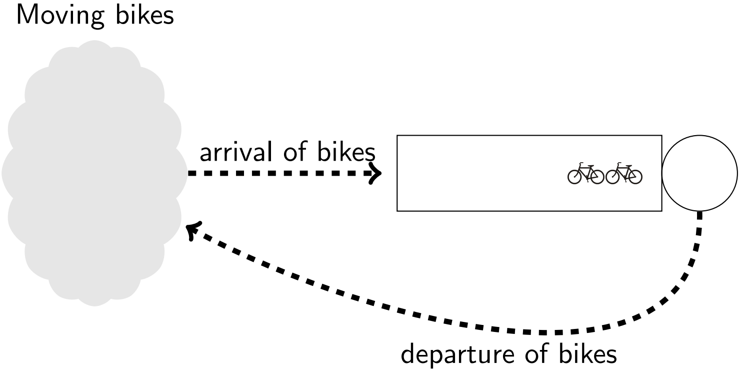


Uniform Bike-sharing systems as closed-queuing networks



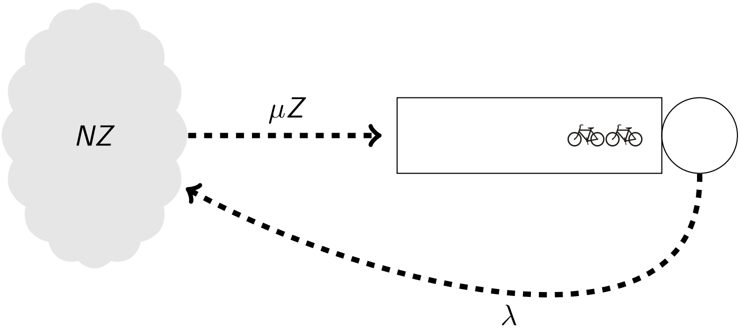
Scaling: $N \to \infty$ stations, s objects per station.

By using independence, the model boils down to the study of a single queue



By using independence, the model boils down to the study of a single queue

Moving bikes



$$i \mapsto i+1 \text{ at rate } \mu Z \qquad (i < K)$$
$$i \mapsto i-1 \text{ at rate } \lambda \qquad (i > 0)$$

Distribution of x_i, the fraction of station with i bikes

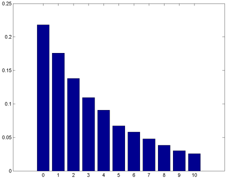
Theorem

There exists ρ, such that in steady state, as N goes to infinity:

$$x_i \propto \rho^i.$$

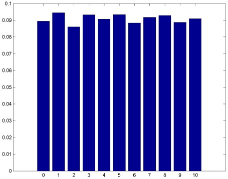
$\rho \leq 1$ *iff* $s \leq \dfrac{C}{2} + \dfrac{\lambda}{\mu}$ *where s be the average number of bikes per stations.*

$s < \frac{C}{2} + \frac{\lambda}{\mu}$



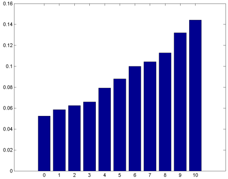
$\rho < 1$

$s = \frac{C}{2} + \frac{\lambda}{\mu}$



$\rho = 1$

$s > \frac{C}{2} + \frac{\lambda}{\mu}$



$\rho < 1$

Consequences: optimal performance for $s \approx C/2$

y-axis: Prop. of problematic stations. x-axis: number of bikes/station s.

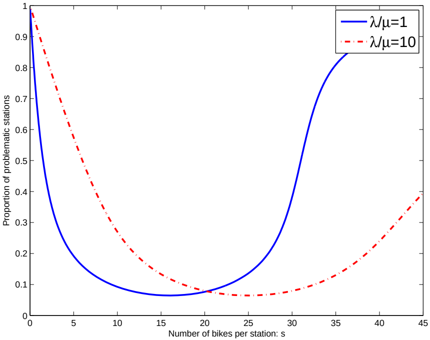
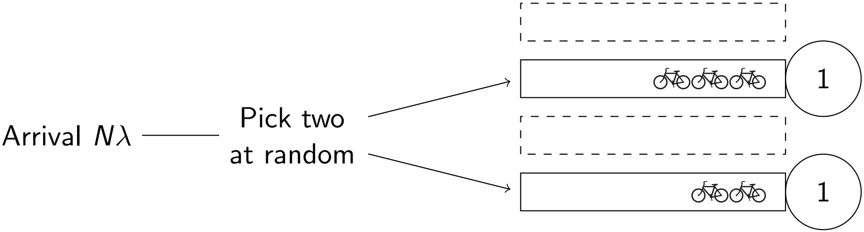


Figure : Capacity of 30 bikes

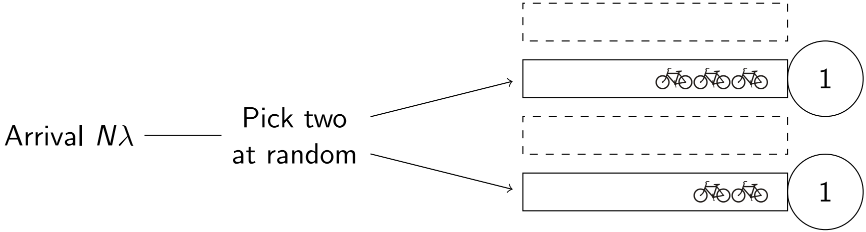
Fraction of problematic stations (=empty+full) minimal for $s = \lambda/\mu + C/2$

- Prop. of problematic stations is at least $2/(C+1)$ (6.5% for $C = 30$)

Two-choice improvement



Two-choice improvement



If x_j is the proportion of stations with j bikes.

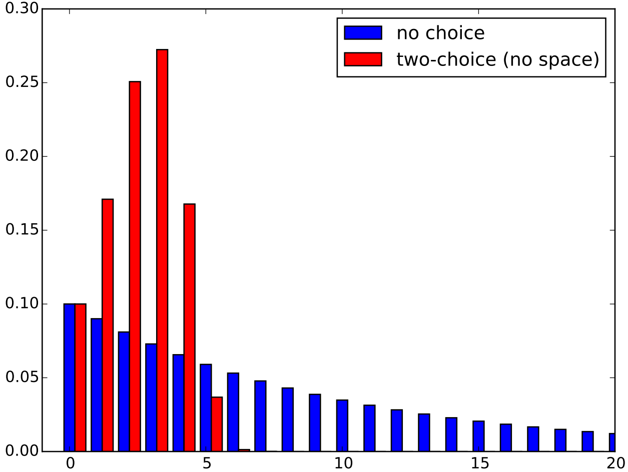
$$(i \mapsto i-1) \text{ at rate } 1$$

$$(i \mapsto i+1) \text{ at rate } \lambda(x_i + 2\sum_{j=i+1}^{\infty} x_j)$$

Note: the rate of change of x_i has to be multiplied by x_i.

With no geometry, we can solve the equation in close-form

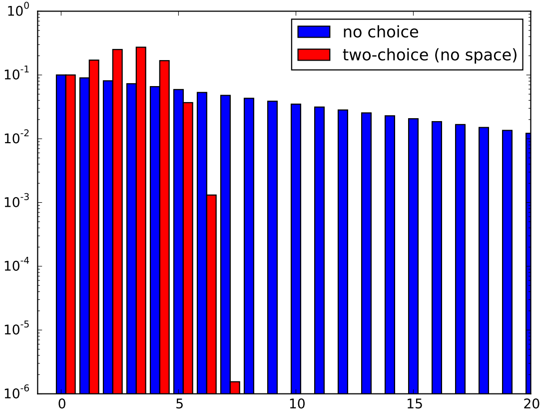
$$x_i = \lambda^{2^i} - \lambda^{2^{i+1}}$$



For velib, choosing two stations at random, improves perf. from $1/C$ to $\sqrt{C}2^{-C/2}$

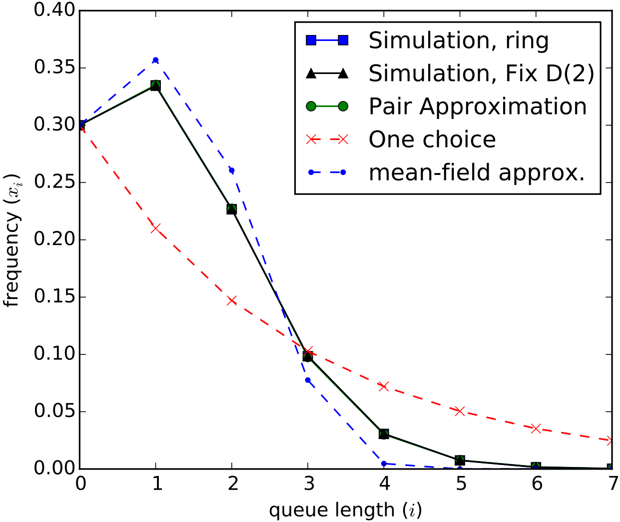
With no geometry, we can solve the equation in close-form

$$x_i = \lambda^{2^i} - \lambda^{2^{i+1}}$$

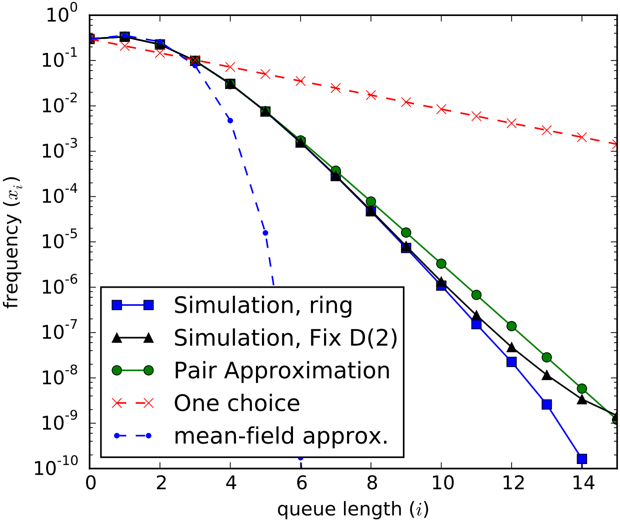


For velib, choosing two stations at random, improves perf. from $1/C$ to $\sqrt{C}2^{-C/2}$

To take geometry into account, we can use pair-approximation



To take geometry into account, we can use pair-approximation



Outline

Take-away message

Mean-field approximation makes possible the study of large systems. Beware of the decoupling assumption.

Take-away message

Mean-field approximation makes possible the study of large systems. Beware of the decoupling assumption.

Performance of bike-sharing is poor, even for homogeneous scenarios ($1/C$ of problematic stations). Incentives or frustration can help.

If an ideal symmetric system works poorly, do not expect perfect service in a real system ;)

To learn more: the slides are online

http://mescal.imag.fr/membres/nicolas.gast/

Mean-field models for performance evaluation

[Benaïm, Le Boudec, 2012] A class of mean field interaction models for computer and communication systems, Performance evaluation 2008

[G2012a] Mean Field for Markov Decision Processes: from Discrete to Continuous Optimization Nicolas Gast, Bruno Gaujal and Jean-Yves Le Boudec, Transaction on Automatic Control, 2012

[G2012b] Markov chains with discontinuous drifts have differential inclusion limits Nicolas Gast,Bruno Gaujal, Performance Evaluation, 2012.

[Sznitman 91]. Topic in Propagation of Chaos, 1991.

Bike-sharing systems

[Fricker Gast, 2014] – Incentives and redistribution in homogeneous bike-sharing systems with stations of finite capacity. EURO Journal on Transportation.

[Fricker, Gast, Mohamed (2012)]. *Mean field analysis for inhomogeneous bike sharing systems* DMTCS Proc.

[Waserhole, Jost (2012)] – Vehicle Sharing System Pricing Regulation : A Fluid Approximation

[Malyshev and Yakovlev]. *Condensation in large closed Jackson networks.* Ann. Appl. Proba. 1996.

[Côme et al (2013)] – Spatio-temporal analysis of Dynamic Origin-Destination data using Latent Dirichlet Allocation. Application to the Vélib' Bike Sharing System of Paris

[Ji Won Yoon et al. (2012)] Cityride: a predictive bike sharing journey advisor

Smart-grids

[G2013] Gast, Le Boudec, Proutière, Tomozei – Impact of Storage on the Efficiency and Prices in Real-Time Electricity Markets. ACM e-Energy '13,

[G2013b] Gast, Le Boudec, Tomozei – Impact of demand-response on the efficiency and prices in real-time electricity markets. ACM e-Energy '14,