Density dependent Markov chains and their approximations

András Horváth
Università di Torino Italy

Outline

1. Markov chains
2. Density dependent Markov chains (DDMC)
3. Approximations of DDMCs
   - 3.1 Deterministic approximation with ODEs
   - 3.2 Stochastic approximation with SDEs
4. Conclusions

1. Continuous time Markov chains

- discrete state space
- continuous time process, $Z(t)$
- enjoys the Markov property

$$Pr(Z(s+t) = j | Z(s) = i, \{Z(u) : 0 \leq u < s\}) =$$
$$Pr(Z(s+t) = j | Z(s) = i)$$

- this memoryless property implies that holding times are exponential

1. Continuous time Markov chains

- a CTMC is usually specified through its infinitesimal generator

$$Q = \begin{pmatrix} -q_1 & q_{1,2} & q_{1,3} & & q_{1,n-1} & q_{1,n} \\ q_{2,1} & -q_2 & q_{2,3} & \cdots & q_{2,n-1} & q_{2,n} \\ q_{3,1} & q_{3,2} & -q_3 & & q_{3,n-1} & q_{3,n} \\ & & & \ddots & & \\ q_{n-1,1} & q_{n-1,2} & q_{n-1,3} & \cdots & -q_{n-1} & q_{n-1,n} \\ q_{n,1} & q_{n,2} & q_{n,3} & \cdots & q_{n-1,n} & -q_n \end{pmatrix}$$

with

$$q_i = \sum_j q_{i,j}$$

1. M/M/1 queue

- queue fed by Poisson process with exponential server

$$Q = \begin{pmatrix} -\lambda & \lambda & 0 & \dots & & \\ \mu & -\lambda-\mu & \lambda & 0 & \dots & \\ 0 & \mu & -\lambda-\mu & \lambda & 0 & \dots \\ & & & \ddots & & \\ & \dots & 0 & \mu & -\lambda-\mu & \lambda \\ & & \dots & 0 & \mu & -\mu \end{pmatrix}$$

1. Analysis of CTMCs

- two ways of thinking what happens in a CTMC:
  - first choose sojourn time according to q_i and then the next state according to $q_{i,j}/q_i$
  - generate exponential random variables according to $q_{i,j}$ and then select the smallest of them to specify the next state
- transient probabilities calculated through matrix exponential

$$P(t) = [Pr(X(t) = j | X(0) = i)] = e^{tQ} = \sum_{n=0}^{\infty} \frac{Q^n t^n}{n!}$$

- steady state by linear system

$$\pi Q = 0, \quad \sum_i \pi_i = 1$$

1. Randomization

- several ways of calculating matrix exponential: Moler, C. and C. Van Loan. 2003. Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later. SIAM Review 45, 3V49.
- randomization is best suited to CTMCs

$$P(t) = \sum_{n=0}^{\infty} (I + Q/q)^n \frac{e^{-qt}(qt)^n}{n!}$$

with $q > max_i q_i$

2. Density dependent Markov chains

- we consider the class of **density dependent** Markov chains
- describe the **interaction of groups of identical objects**
- informally: the intensities of the interactions can be **expressed as a function of the density** of the objects present in the considered area or volume
- (instead of expressed as a function of the number of objects itself)

2. Density dependent Markov chains

- formally a sequence of **density dependent** Markov chains is:
  - indexed by a parameter, denoted by N (area or volume or total number of objects)
  - has state space $\mathcal{S}^{[N]} \subseteq \mathbb{Z}^k$ (k groups of identical objects)
  - the transition intensities are in the form:

$$q^{[N]}_{r,r+m} = N\, f\left(\frac{r}{N}, m\right)$$

- by relaxing the above form we obtain the class of **nearly density dependent** Markov chains with transition intensities in the form

$$q^{[N]}_{r,r+m} = N\, f\left(\frac{r}{N}, m\right) + N\, g\left(r/N, m, N\right)$$

  with $g(r/N, m, N) \in O\left(1/N\right)$

2. Example

- epidemic model with susceptible (S) and infected (I) individuals distributed over an area split into N equally sized cells
- a state is a pair (i, j)
- three kinds of transitions:
  - 1. susceptible individuals grows:

$$\emptyset \to S$$

with intensity

$$q^{[N]}_{(i,j),(i+1,j)} = N\lambda_1$$

because the larger the area the higher the intensity

2. Example

- three kinds of transitions:
  - 2. one susceptible individual becomes infected:

$$S + 2I \to 3I$$

with intensity

$$q^{[N]}_{(i,j),(i-1,j+1)} = \frac{ij(j-1)}{2}\frac{1}{N^3}N\lambda_2 = N\left(\frac{\lambda_2}{2}\frac{i}{N}\left(\frac{j}{N}\right)^2\right) - N\left(\frac{1}{N}\frac{\lambda_2}{2}\frac{i}{N}\frac{j}{N}\right)$$

because

$$\frac{ij(j-1)}{2}\frac{1}{N^3}$$

is the probability that one S and 2I meet in a given cell

2. Example

- three kinds of transitions:
  - 3. infected individuals can become immune:

$$I \to \emptyset$$

with intensity

$$q^{[N]}_{(i,j),(i,j-1)} = j\lambda_3 = q^{[N]}_{(i,j),(i,j-1)} = N\lambda_3 \frac{j}{N}$$

because every I individually gets immune with intensity λ_3

3. Fluid approximation

- the considered approximations are ***fluid***
- in order to compare models with different values of N we work with the **density process**:

$$Z^{[N]}(t) = X^{[N]}(t)/N$$

3.1 Deterministic approximation

- if the initial state that tends to z_0 as N tends to infinity:

$$\lim_{N \to \infty} Z^{[N]}(0) = z_0$$

- then the density process tends to the solution of

$$dz(t) = \sum_{l \in C} l\, f\,(z(t), l)\, dt, \quad z(0) = z_0$$

3.1 Deterministic approximation

- difference between the deterministic approximation and the original stochastic behavior is characterized by

$$\sup_{t \leq T} \left| Z^{[N]}(t) - z(t) \right| = O\left(1/\sqrt{N}\right) \quad a.s.$$

  i.e., the error of the deterministic approximation decreases as $1/\sqrt{N}$
- for any ϵ there exists M_ϵ such that

$$P\left(\frac{\sup_{t \leq T} \left| Z^{[N]}(t) - z(t) \right|}{1/\sqrt{N}} > M_\epsilon\right) < \epsilon$$

3.1 Deterministic approximation

- the deterministic approximation provides a **single trajectory**
- usually considered as the approximate mean
- important characteristics, like significant variance or bimodality or non-deterministic cycle times, can be lost
- these can be present even with very large values of N

3.1 Lotka-Volterra predator-prey reactions

- models predator pray interactions:

$$X \to 2X, X + Y \to 2Y, Y \to \emptyset$$

  with initial state

$$(N, N)$$

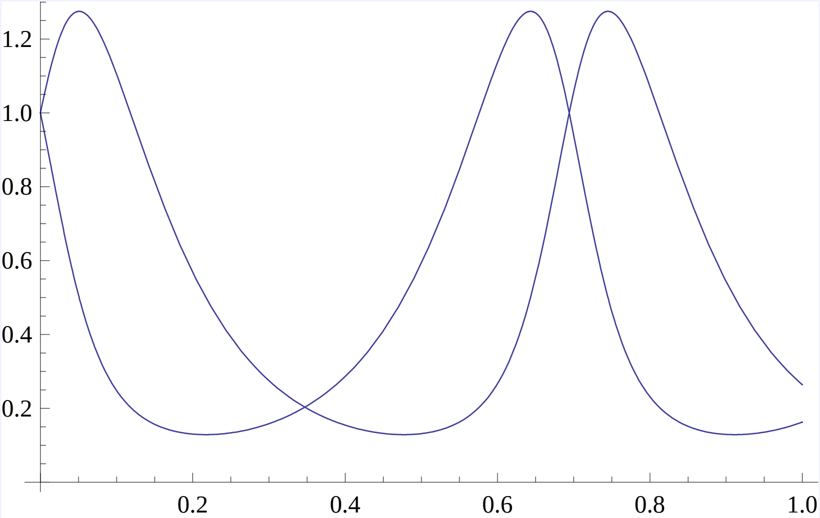
  and intensities

$$(10, 20/N, 10)$$

- accordingly the density process starts from $(1, 1)$

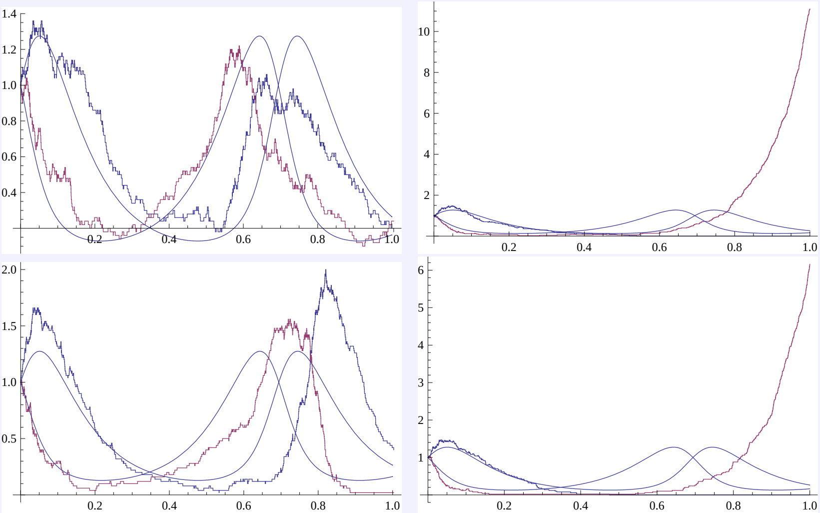
3.1 Lotka-Volterra predator-prey reactions

With the previous parameters the deterministic approximation oscillates regularly forever:



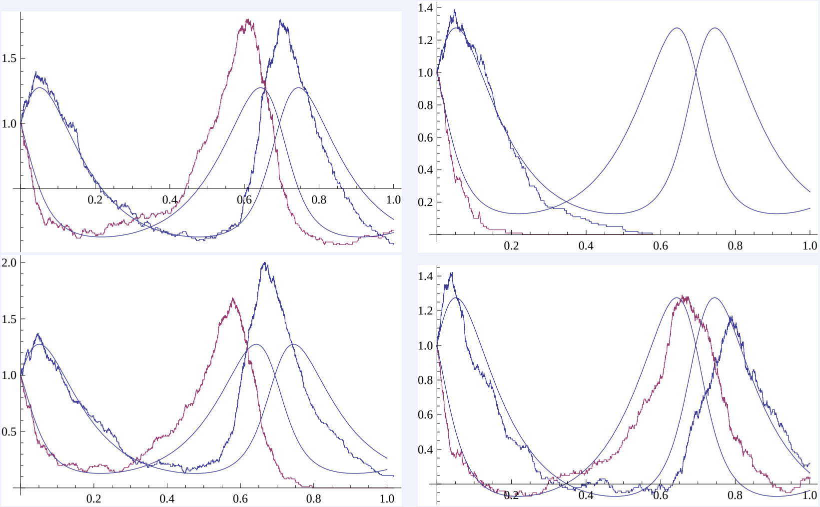
3.1 Lotka-Volterra predator-prey reactions

Compared to $N = 50$:



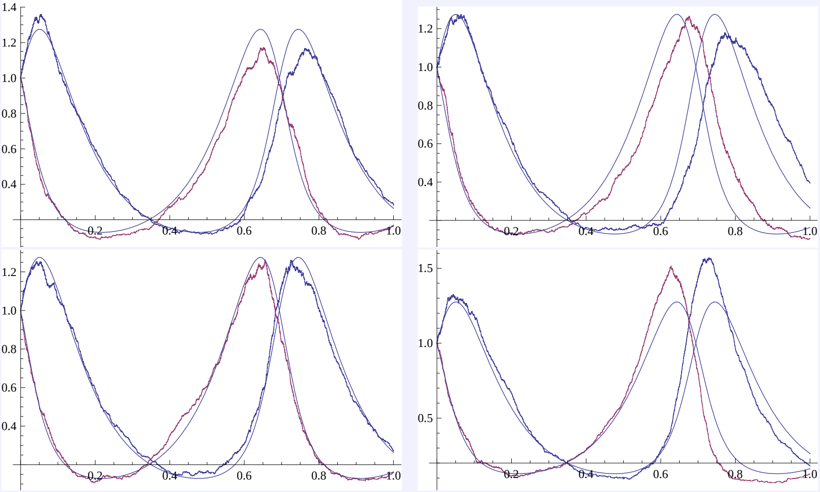
3.1 Lotka-Volterra predator-prey reactions

Compared to $N = 100$:



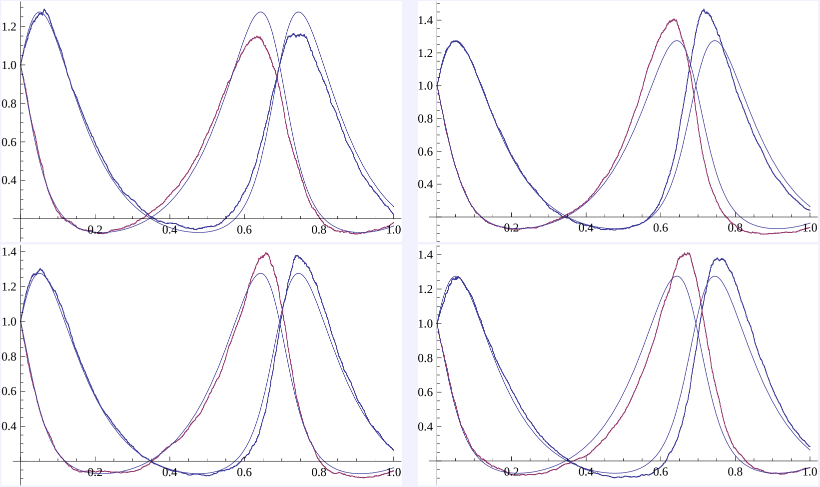
3.1 Lotka-Volterra predator-prey reactions

Compared to $N = 500$:



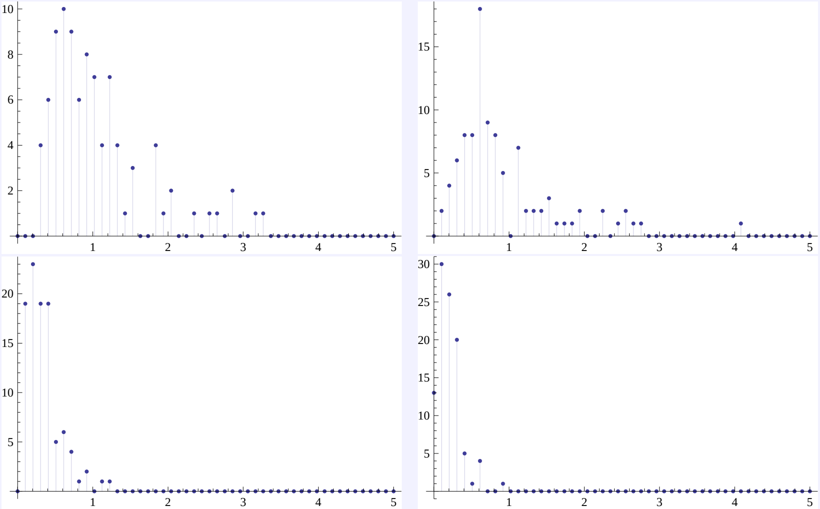
3.1 Lotka-Volterra predator-prey reactions

Compared to $N = 2000$:



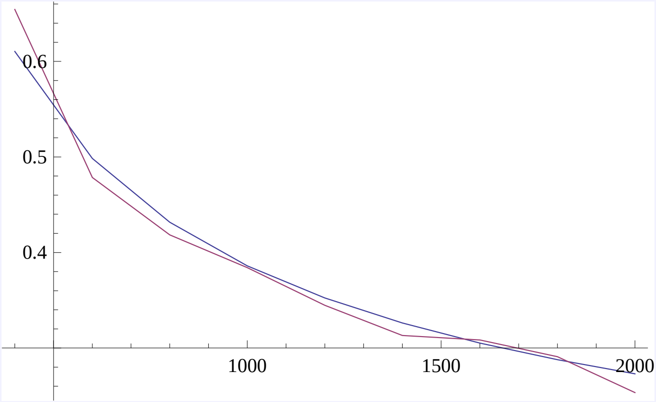
3.1 Lotka-Volterra predator-prey reactions

Distribution of the largest difference between the ODE and the original behaviour with $N = 100, 200, 1000, 2000$:



3.1 Lotka-Volterra predator-prey reactions

Mean error as function of N and its best least square $1/\sqrt{N}$ fit:

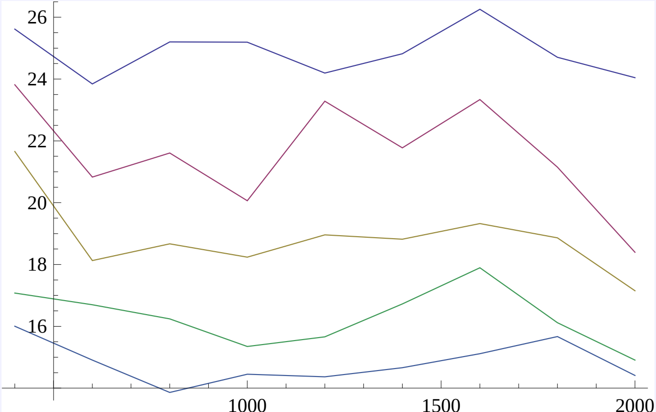


3.1 Lotka-Volterra predator-prey reactions

M_ϵ in function of N for which

$$P\left(\frac{\sup_{t\leq T}\left|Z^{[N]}(t)-z(t)\right|}{1/\sqrt{N}} > M_\epsilon\right) = \epsilon$$

for $\epsilon = 0.05, 0.1, 0.15, 0.2, 0.25$:



3.2 Stochastic approximation with SDEs

- approximation with stochastic differential equations:

$$dY^{[N]}(t) = \sum_{l \in C} l\, f\left(Y^{[N]}(t), l\right) dt + \sum_{l \in C} \frac{l}{\sqrt{N}} \sqrt{f\left(Y^{[N]}(t), l\right)}\, dW_l(t)$$

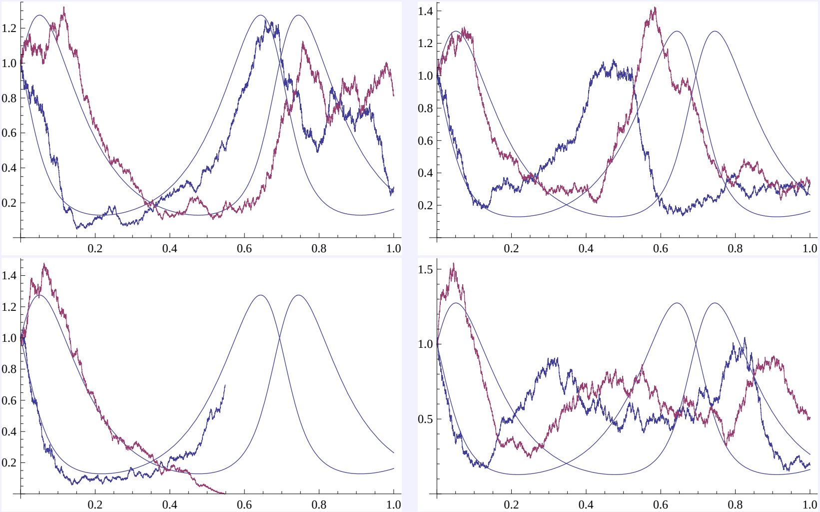
  where $W_l(t)$ with $l \in C$ are independent standard one-dimensional Brownian motions
- **maintains stochasticity**: provides distributions
- **explicitly uses N** (in case of the deterministic approximation $N = \infty$)
- has better convergence:

$$\sup_{t \leq T} \left| Z^{[N]}(t) - Y^{[N]}(t) \right| = O\left(\log N / N\right) \quad a.s.$$

  for *corresponding pairs* of trajectories

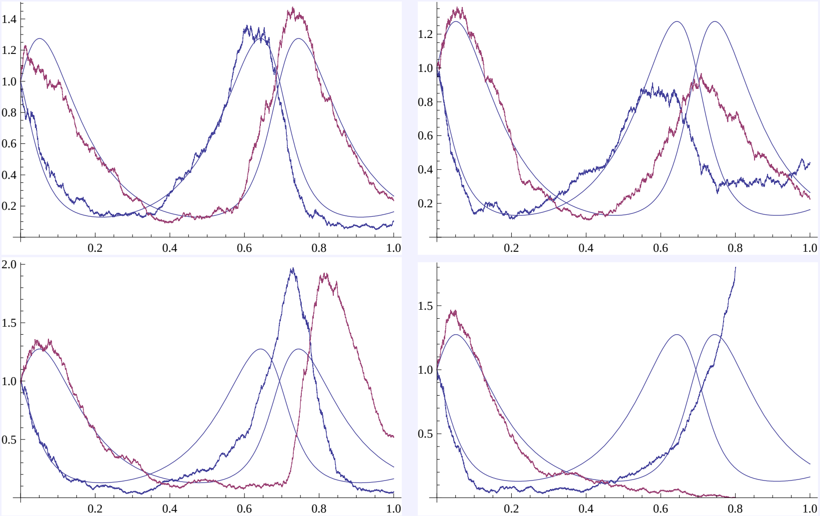
3.2 Lotka-Volterra predator-prey reactions

Trajectories with $N = 50$:



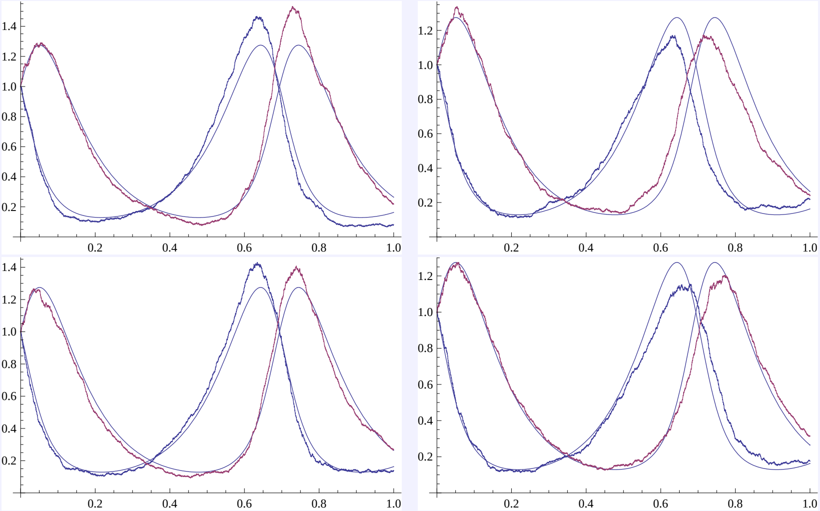
3.2 Lotka-Volterra predator-prey reactions

Trajectories with $N = 200$:



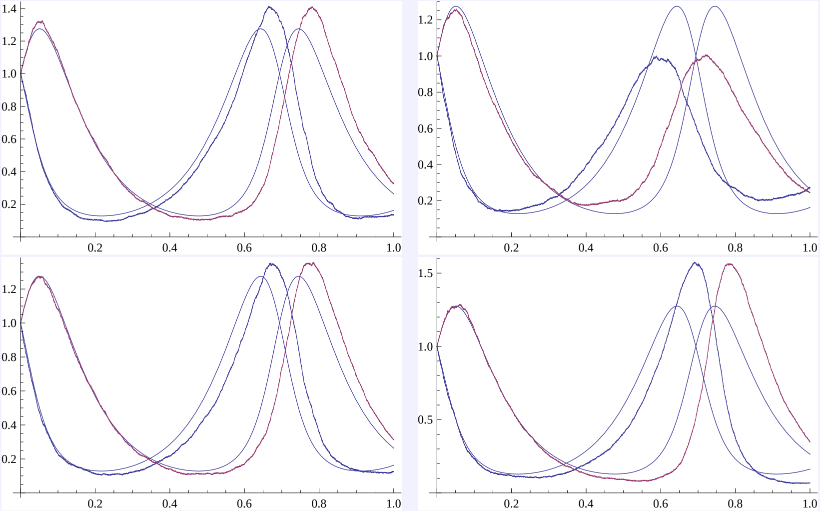
3.2 Lotka-Volterra predator-prey reactions

Trajectories with $N = 500$:



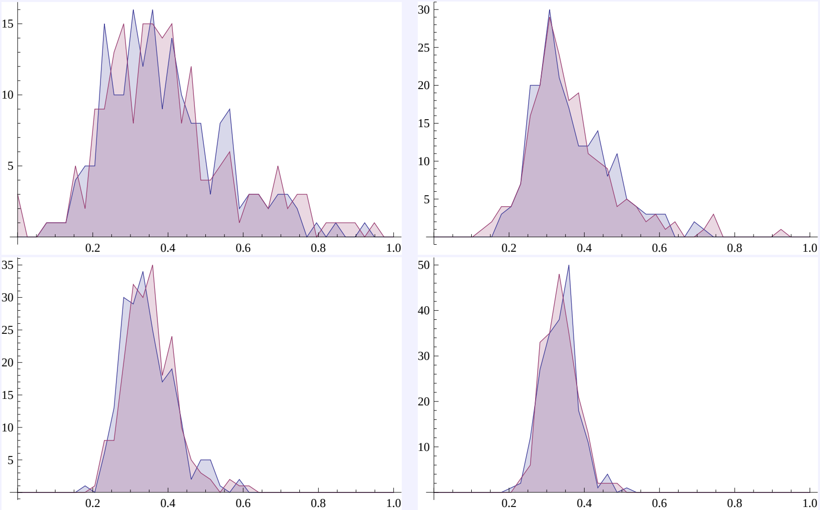
3.2 Lotka-Volterra predator-prey reactions

Trajectories with $N = 2000$:



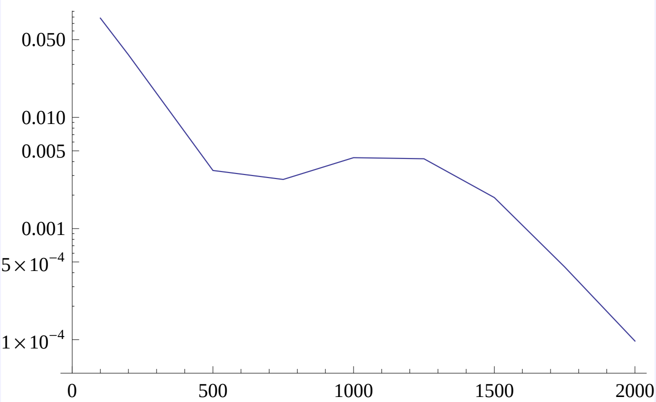
3.2 Lotka-Volterra predator-prey reactions

Pmf with $N = 200, 500, 1000, 2000$:



3.2 Lotka-Volterra predator-prey reactions

Difference in the mean as function of N:



4. Conclusions

- exact simulation of CTMC becomes slower with increasing N
- for fixed step size, simulation of SDE becomes more accurate as N increases
- ranges of N:
  - small N: use an analytical approach (randomization)
  - larger N: simulate the Markov chain
  - **even larger N but still important stochastic behavior: use diffusion approximation**
  - huge N, no stochasticity: use deterministic approximation