Modelling the long-term behaviour of population processes

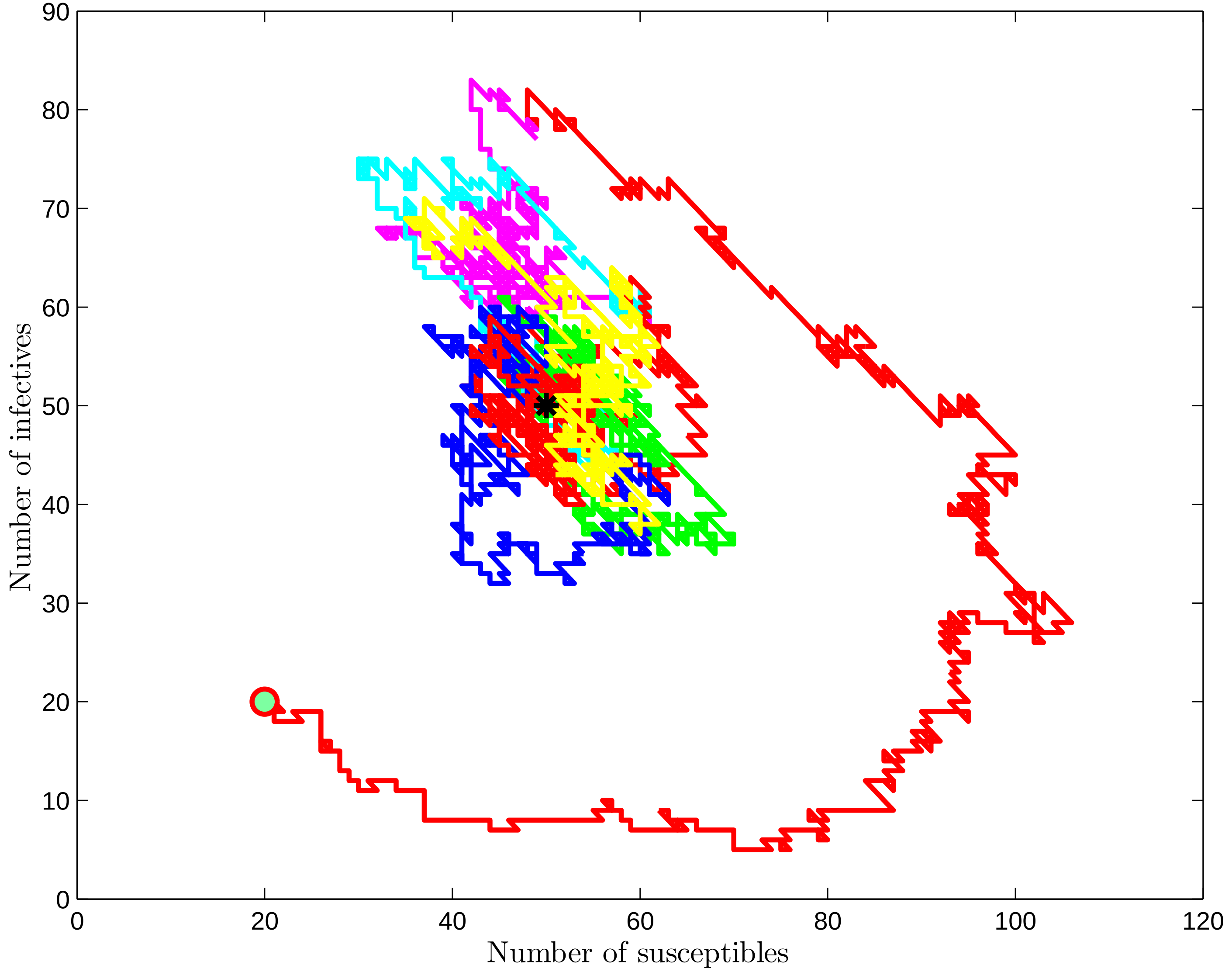
Phil Pollett

Department of Mathematics

University of Queensland

AUSTRALIAN RESEARCH COUNCIL
Centre of Excellence for Mathematics
and Statistics of Complex Systems

The progress of an epidemic



An autocatalytic reaction

Consider the reaction scheme $A \xrightarrow{X} B$, where X is a catalyst. Suppose that there are two stages, namely

$$A + X \xrightarrow{k_1} 2X \quad \text{and} \quad 2X \xrightarrow{k_2} B.$$

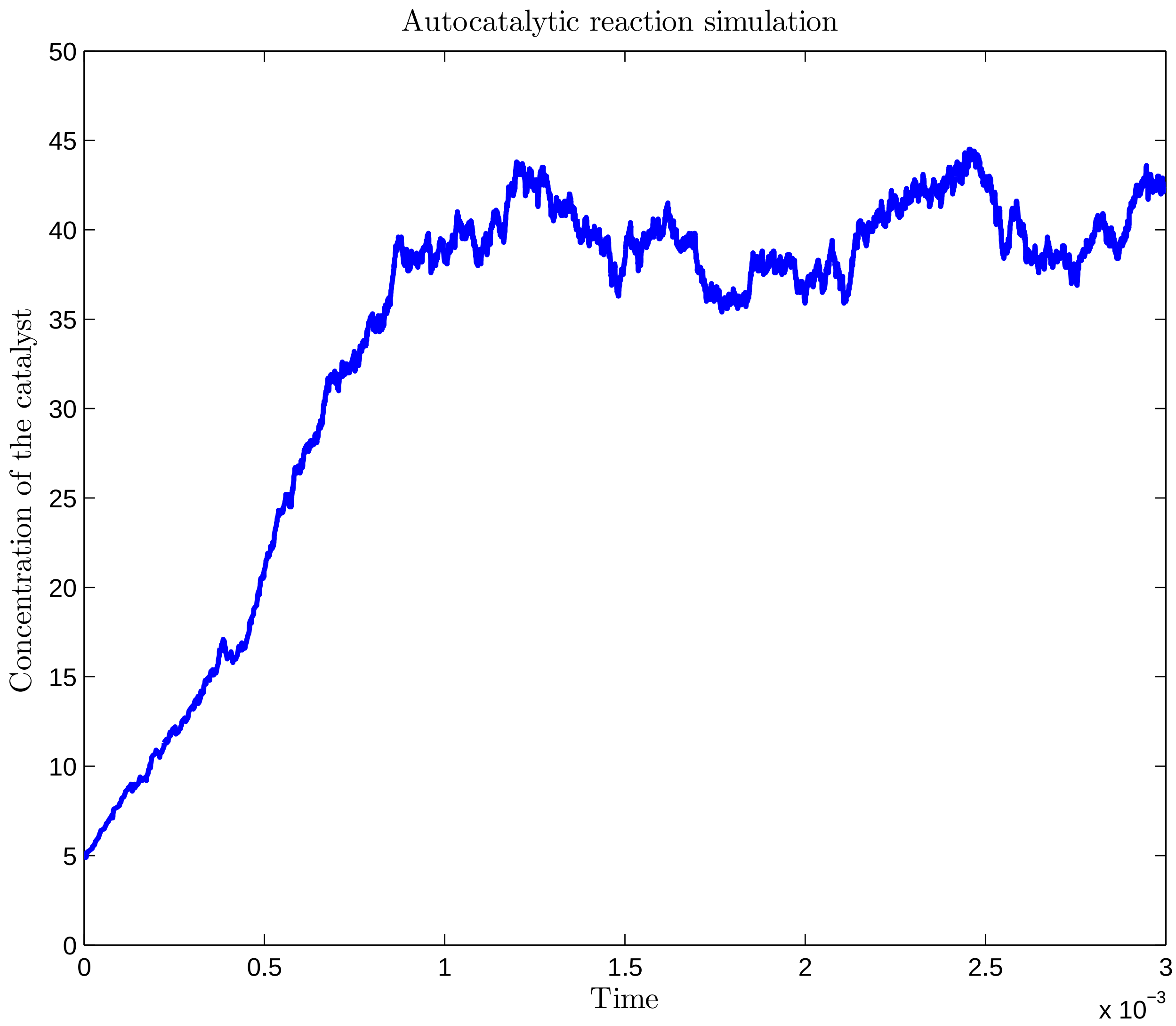
Let n_t be the number of X molecules at time t.

Let a be the number of A molecules. Suppose that the concentration of A is held constant.

The state space is $S = \{0, 1, 2, \dots\}$ and the transitions are:

$$\begin{array}{lll} n \to n+1 & \text{at rate} & \frac{k_1}{V} a n = k_1 [A] n \\ n \to n-2 & \text{at rate} & \frac{k_2}{V} \binom{n}{2} \quad (V \text{ is volume}) \end{array}$$

An autocatalytic reaction



A population network

There are N "patches" of habitat. Each occupied patch becomes empty at rate μ and colonization of empty patches by occupied patches occurs at rate λ/N for each suitable pair.

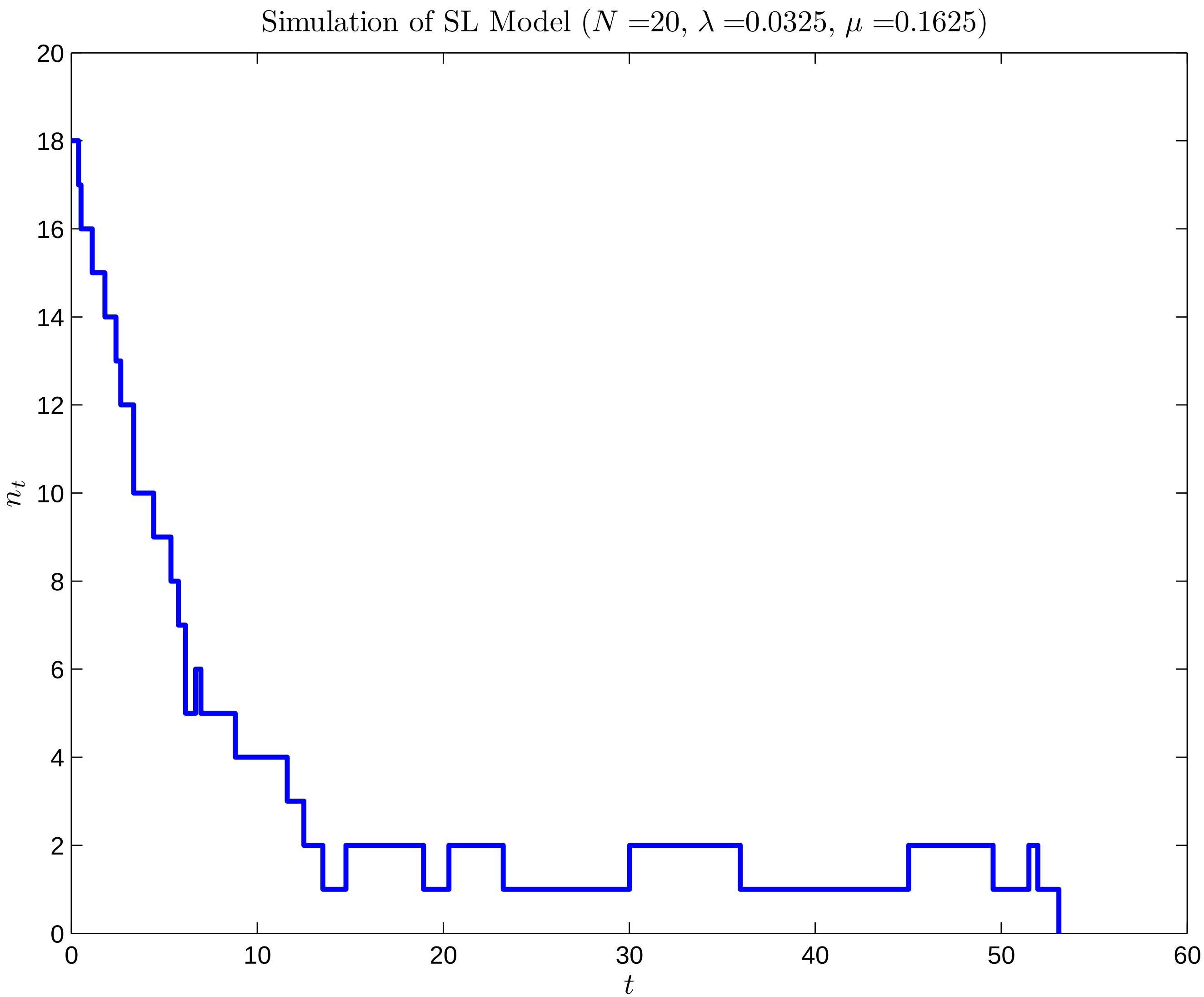
Let n_t be the number of occupied patches at time t. The state space is $S = \{0, 1, \ldots, N\}$ and the transitions are:

$n \to n+1$	at rate	$\frac{\lambda}{N} n\,(N-n)$
$n \to n-1$	at rate	μn

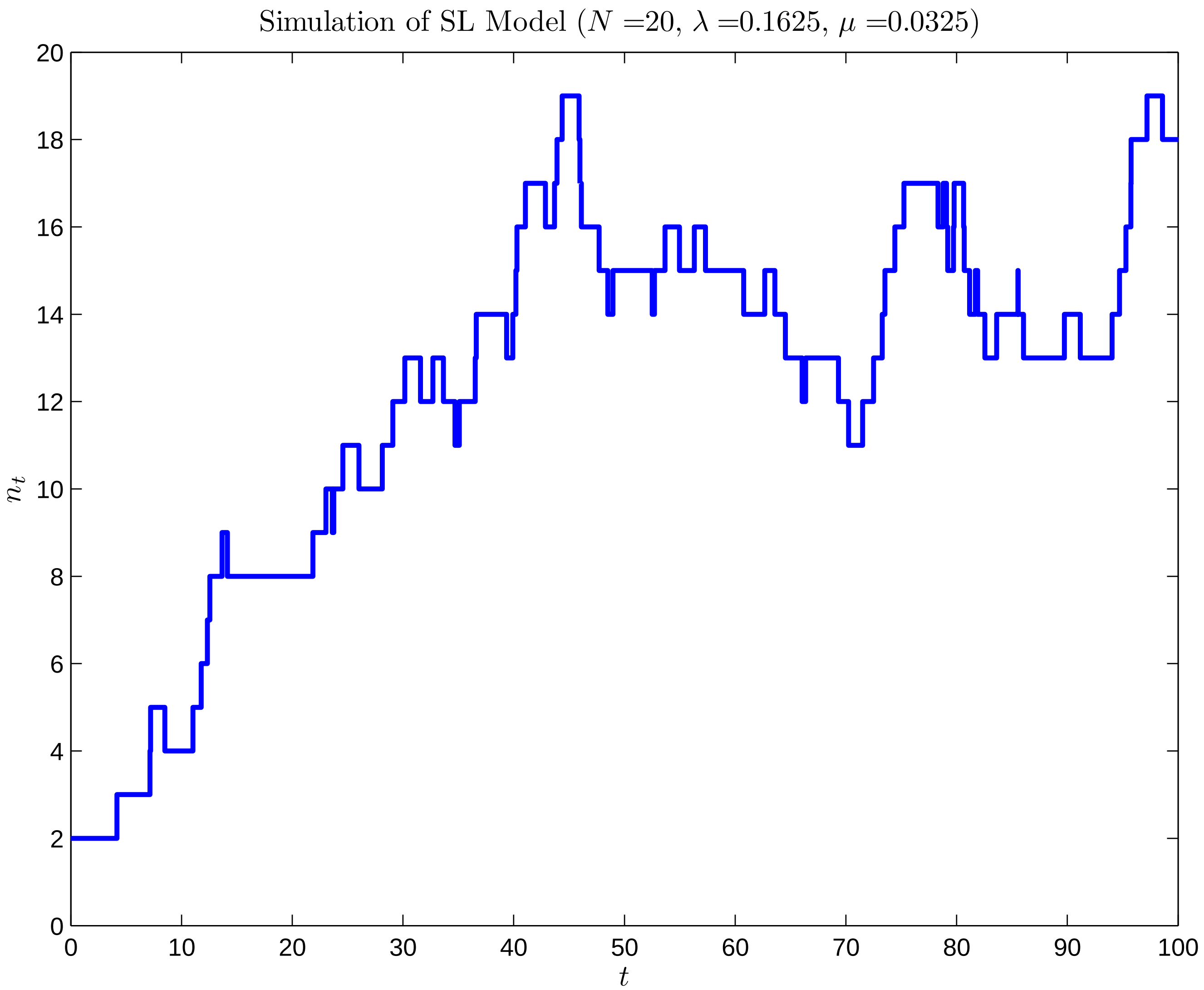
I will call this model the *stochastic logistic (SL) model*, though it has many names, having been rediscovered several times since Feller* proposed it.

*Feller, W. (1939) Die grundlagen der volterraschen theorie des kampfes ums dasein in wahrscheinlichkeitsteoretischer behandlung. *Acta Biotheoretica* 5, 11–40.

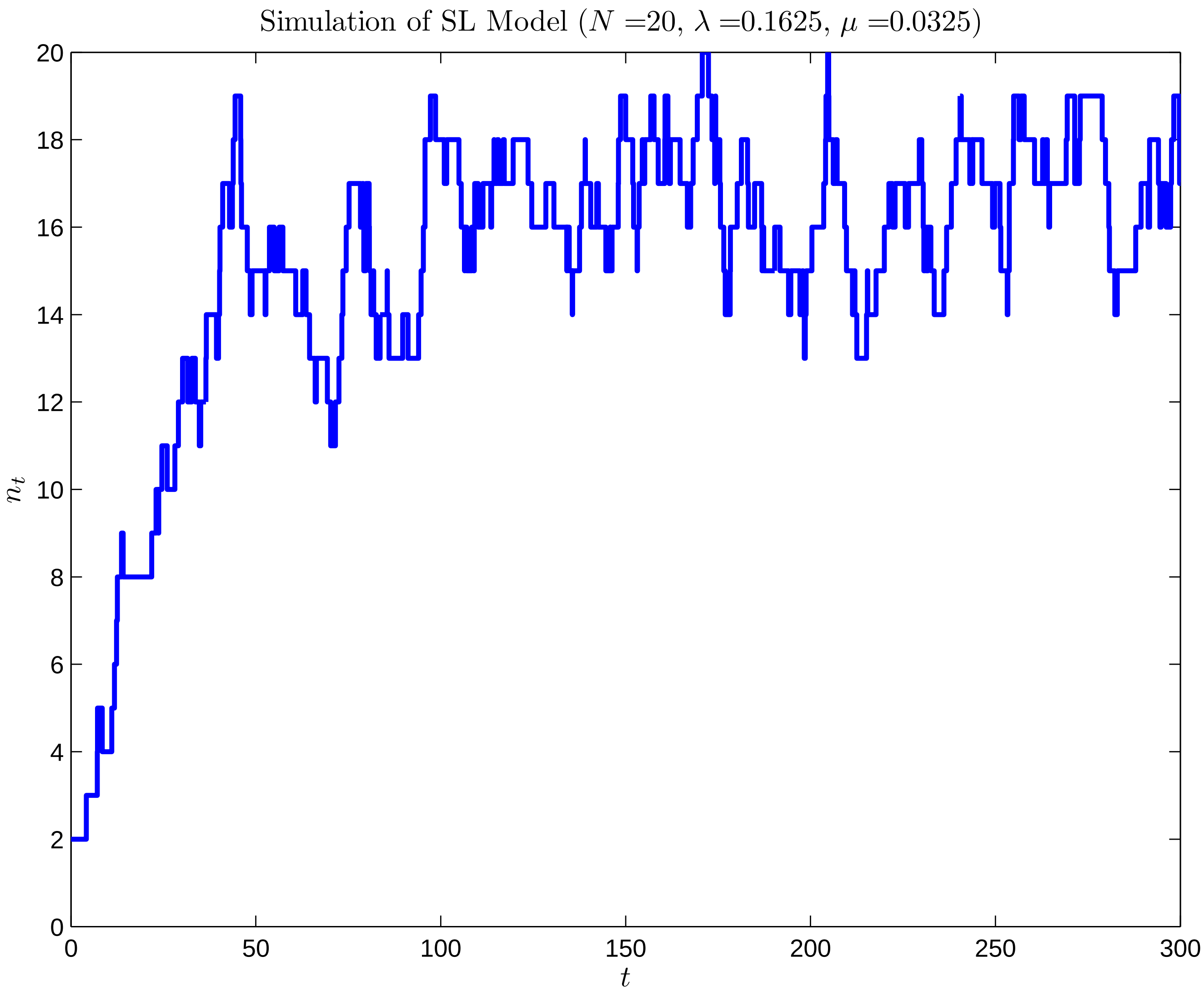
The SL model $(\lambda < \mu)$



The SL model $(\lambda > \mu)$



The SL model $(\lambda > \mu)$



Markov chains–ingredients

The *state* at time t : $n_t \in S$ (a countable set).

Transition rates $Q = (q_{nm},\, n, m \in S)$: $q_{nm}\,(\geq 0)$, for $m \neq n$, is the transition rate from state n to state m and $q_{nn} = -q_n$, where $q_n = \sum_{m \neq n} q_{nm}$, is the transition rate out of state n.

Markov chains–ingredients

The *state* at time t : $n_t \in S$ (a countable set).

Transition rates $Q = (q_{nm},\, n, m \in S)$: $q_{nm}\,(\geq 0)$, for $m \neq n$, is the transition rate from state n to state m and $q_{nn} = -q_n$, where $q_n = \sum_{m \neq n} q_{nm}$, is the transition rate out of state n.

Example. The autocatalytic reaction $A + X \xrightarrow{k_1} 2X$, $2X \xrightarrow{k_2} B$

$$Q = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & -\frac{k_1}{V}a & \frac{k_1}{V}a & 0 & 0 & \dots \\ \frac{k_2}{V} & 0 & -\frac{1}{V}\left(2k_1 a + k_2\right) & 2\frac{k_1}{V}a & 0 & \dots \\ 0 & 3\frac{k_2}{V} & 0 & -\frac{3}{V}\left(k_1 a + k_2\right) & 3\frac{k_1}{V}a & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

$\Big(n \to n+1$ at rate $\frac{k_1}{V}an$ and $n \to n-2$ at rate $\frac{k_2}{V}\binom{n}{2}\Big)$

Markov chains–ingredients

The *state* at time t : $n_t \in S$ (a countable set).

Transition rates $Q = (q_{nm},\, n, m \in S)$: $q_{nm}\,(\geq 0)$, for $m \neq n$, is the transition rate from state n to state m and $q_{nn} = -q_n$, where $q_n = \sum_{m \neq n} q_{nm}$, is the transition rate out of state n.

Example. The autocatalytic reaction $A + X \xrightarrow{k_1} 2X$, $2X \xrightarrow{k_2} B$

$$Q = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & \ldots \\ 0 & -\frac{k_1}{V}a & \frac{k_1}{V}a & 0 & 0 & \ldots \\ \frac{k_2}{V} & 0 & -\frac{1}{V}\left(2k_1 a + k_2\right) & 2\frac{k_1}{V}a & 0 & \ldots \\ 0 & 3\frac{k_2}{V} & 0 & -\frac{3}{V}\left(k_1 a + k_2\right) & 3\frac{k_1}{V}a & \ldots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

$\Big(n \to n+1$ at rate $\frac{k_1}{V}an$ and $n \to n-2$ at rate $\frac{k_2}{V}\binom{n}{2} \Big)$

More ingredients

Assumptions: take 0 to be the sole absorbing state (that is, $q_{0n} = 0$). For simplicity, suppose that $C = S - \{0\}$ is "irreducible" and that we reach 0 from C with probability 1.

State probabilities: $\mathbf{p}(t) = (p_n(t),\, n \in S)$, $p_n(t) = \Pr(n_t = n)$.

Initial distribution: $\mathbf{p}(0) = \mathbf{a} = (a_n,\, n \in S)$ $(a_0 = 0)$.

Forward equations (FEs): the state probabilities satisfy

$$\mathbf{p}'(t) = \mathbf{p}(t)Q, \qquad \mathbf{p}(0) = \mathbf{a}.$$

In particular, since $q_{0n} = 0$,

$$p_n'(t) = \sum_{m \in C} p_m(t) q_{mn} \quad (n \in S,\ t > 0).$$

Or, written as a *master equation*:

$$p_n'(t) = \sum_{m \in C} \{p_m(t) q_{mn} - p_n(t) q_{nm}\} \quad (n \in S,\ t > 0).$$

More ingredients

Assumptions: take 0 to be the sole absorbing state (that is, $q_{0n} = 0$). For simplicity, suppose that $C = S - \{0\}$ is "irreducible" and that we reach 0 from C with probability 1.

State probabilities: $\mathbf{p}(t) = (p_n(t),\, n \in S)$, $p_n(t) = \Pr(n_t = n)$.

Initial distribution: $\mathbf{p}(0) = \mathbf{a} = (a_n,\, n \in S)$ $(a_0 = 0)$.

Forward equations (FEs): the state probabilities satisfy

$$\mathbf{p}'(t) = \mathbf{p}(t)Q, \qquad \mathbf{p}(0) = \mathbf{a}.$$

In particular, since $q_{0n} = 0$,

$$p_n'(t) = \sum_{m \in C} p_m(t) q_{mn} \quad (n \in S,\ t > 0).$$

Or, written as a *master equation*:

$$p_n'(t) = \sum_{m \in C} \{p_m(t) q_{mn} - p_n(t) q_{nm}\} \quad (n \in S,\ t > 0).$$

Solution to FEs?

If S is a finite set (or, more generally, if $\sup_n q_n < \infty$), then the forward equations $\mathbf{p}'(t) = \mathbf{p}(t)Q$, with $\mathbf{p}(0) = \mathbf{a}$, have the unique solution $\mathbf{p}(t) = \mathbf{a}\,\mathbf{exp}(Qt)$, $t \geq 0$, where $\mathbf{exp}$ is the *matrix exponential*:

$$\mathbf{exp}(A) = I + A + \frac{A^2}{2!} + \cdots + \frac{A^n}{n!} + \ldots$$

Solution to FEs?

If S is a finite set (or, more generally, if $\sup_n q_n < \infty$), then the forward equations $\mathbf{p}'(t) = \mathbf{p}(t)Q$, with $\mathbf{p}(0) = \mathbf{a}$, have the unique solution $\mathbf{p}(t) = \mathbf{a}\,\mathbf{exp}(Qt)$, $t \geq 0$, where $\mathbf{exp}$ is the *matrix exponential*:

$$\mathbf{exp}(A) = I + A + \frac{A^2}{2!} + \cdots + \frac{A^n}{n!} + \ldots$$

So, if we have at our disposal good methods for evaluating $\mathbf{exp}$, we can in principle answer any question about our model.

Solution to FEs?

If S is a finite set (or, more generally, if $\sup_n q_n < \infty$), then the forward equations $\mathbf{p}'(t) = \mathbf{p}(t)Q$, with $\mathbf{p}(0) = \mathbf{a}$, have the unique solution $\mathbf{p}(t) = \mathbf{a}\,\mathbf{exp}(Qt)$, $t \geq 0$, where $\mathbf{exp}$ is the *matrix exponential*:

$$\mathbf{exp}(A) = I + A + \frac{A^2}{2!} + \cdots + \frac{A^n}{n!} + \ldots$$

So, if we have at our disposal good methods for evaluating $\mathbf{exp}$, we can in principle answer any question about our model.

Moler, C.B. and Van Loan, C.F. (1978) Nineteen dubious ways to compute the exponential of a matrix, *SIAM Rev.* 20, 801–836.

Solution to FEs?

If S is a finite set (or, more generally, if $\sup_n q_n < \infty$), then the forward equations $\mathbf{p}'(t) = \mathbf{p}(t)Q$, with $\mathbf{p}(0) = \mathbf{a}$, have the unique solution $\mathbf{p}(t) = \mathbf{a}\,\mathbf{exp}(Qt)$, $t \geq 0$, where $\mathbf{exp}$ is the *matrix exponential*:

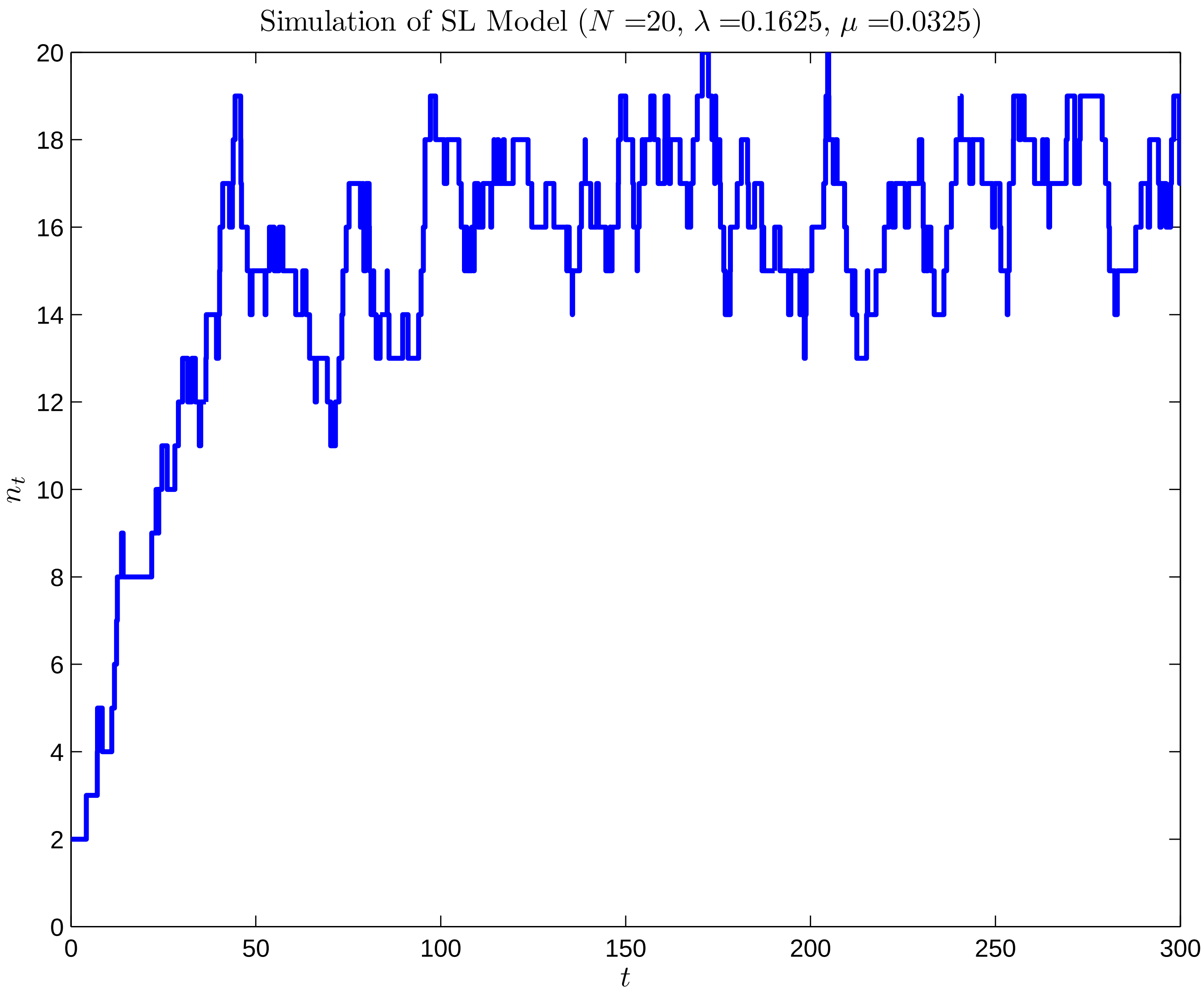
$$\mathbf{exp}(A) = I + A + \frac{A^2}{2!} + \cdots + \frac{A^n}{n!} + \ldots$$

So, if we have at our disposal good methods for evaluating $\mathbf{exp}$, we can in principle answer any question about our model.

Moler, C.B. and Van Loan, C.F. (1978) Nineteen dubious ways to compute the exponential of a matrix, *SIAM Rev.* 20, 801–836.

Use Matlab's `expm` or, better (especially if Q is sparse), Roger Sidje's *expokit*: `www.maths.uq.edu.au/expokit/`

The SL model $(\lambda > \mu)$



Exercise 1

Suppose that at any given time during your office hours there are n students waiting with probability $p_n := (1-p)p^n$ where say $p = 0.1$, so that, for example, the chance that there are no students waiting is $p_0 = 1 - p = 0.9$.

There is a knock at the door. What is the probability that there are n students waiting?

Exercise 1

Suppose that at any given time during your office hours there are n students waiting with probability $p_n := (1-p)p^n$ where say $p = 0.1$, so that, for example, the chance that there are no students waiting is $p_0 = 1 - p = 0.9$.

There is a knock at the door. What is the probability that there are n students waiting?

Answer : $p_n/(1-p_0) = (1-p)p^{n-1} = (0.9) \times (0.1)^{n-1}$ ($n \geq 1$).

Modelling quasi stationarity

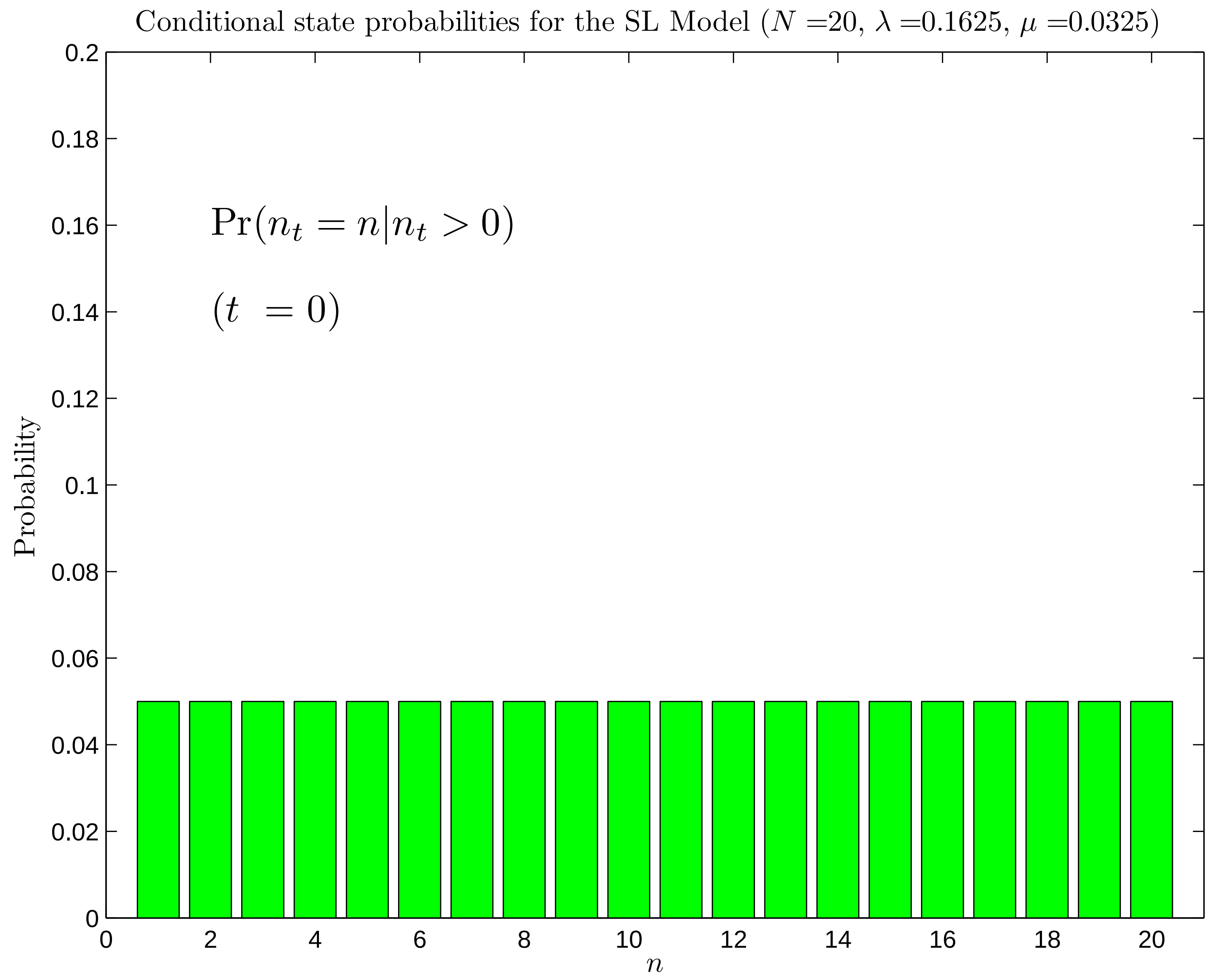
Recall that $S = \{0\} \cup C$, where 0 is an absorbing state and C is the set of transient states.

Define *conditional state probabilities* $\mathbf{r}(t) = (r_n(t),\, n \in C)$ by

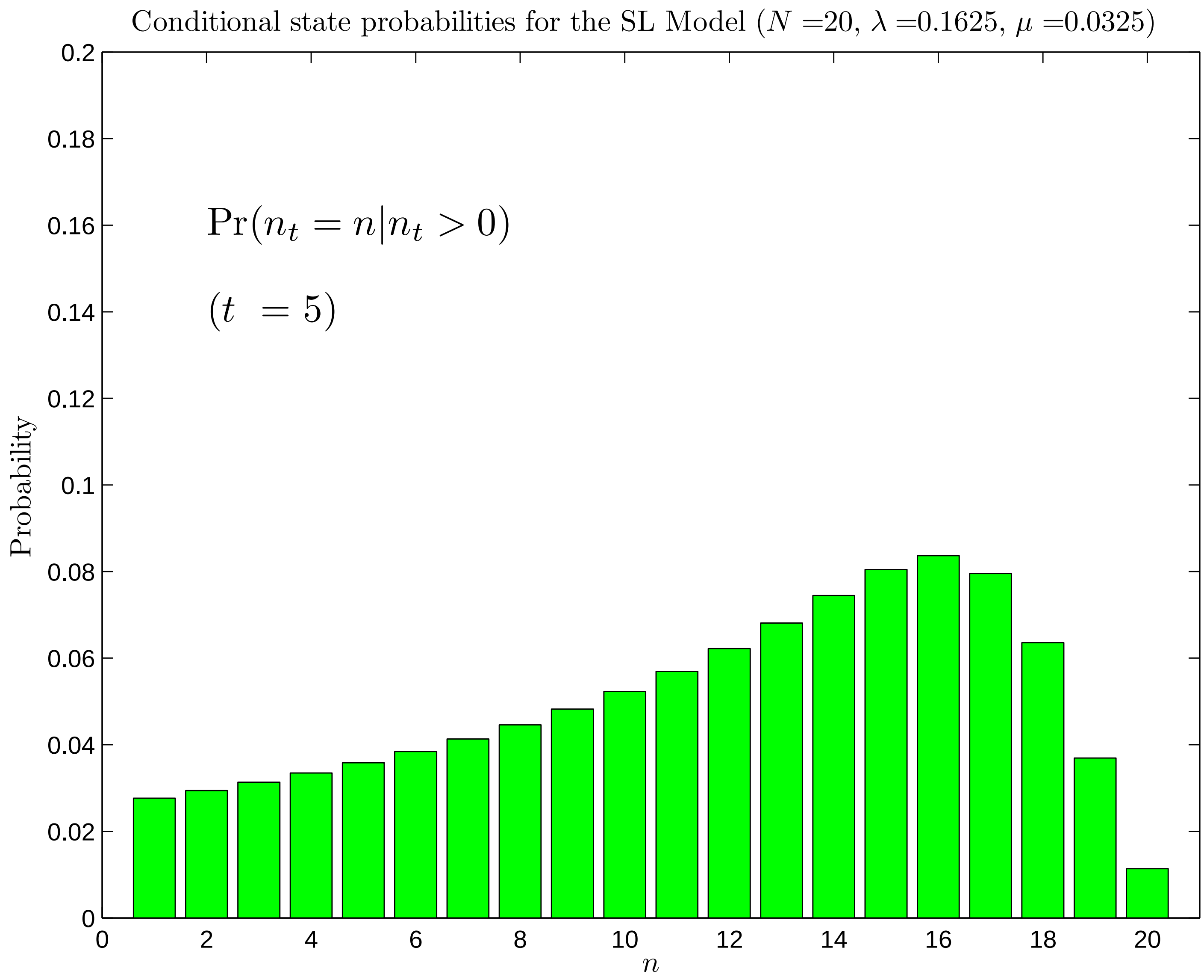
$$r_n(t) = \Pr(n_t = n \,|\, n_t \in C),$$

the chance of being in state n *given that the process has not reached* 0.

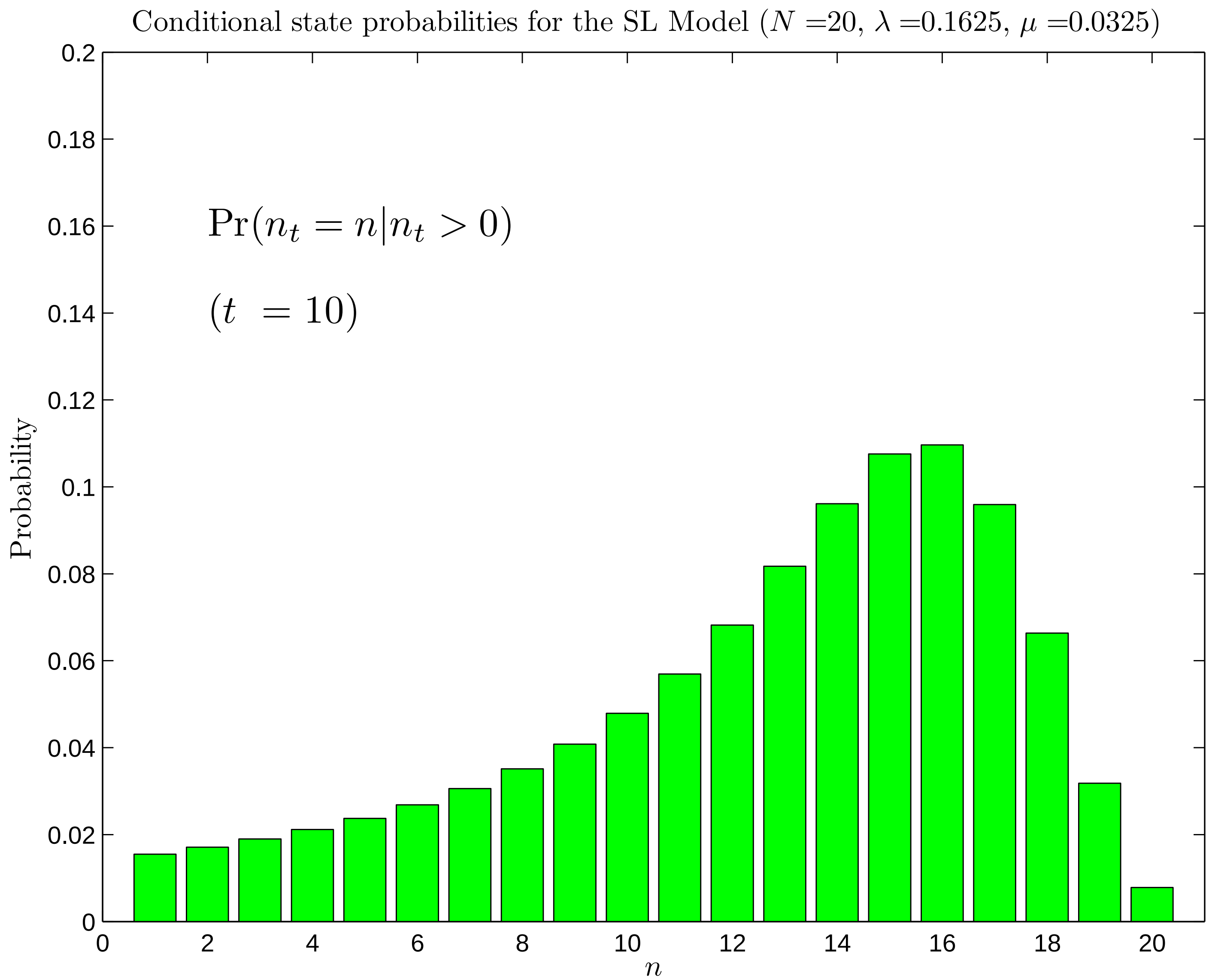
Conditional state probabilities



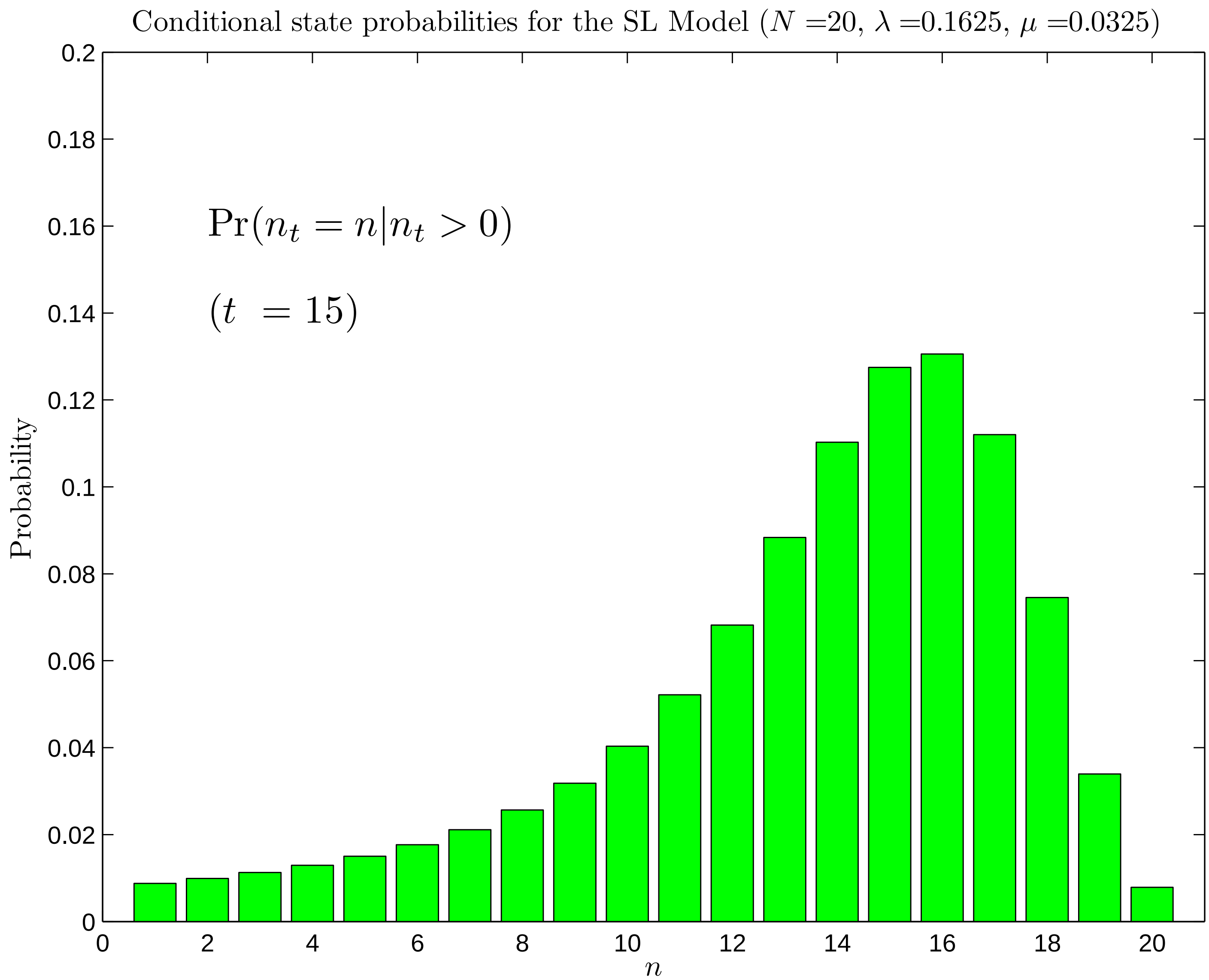
Conditional state probabilities



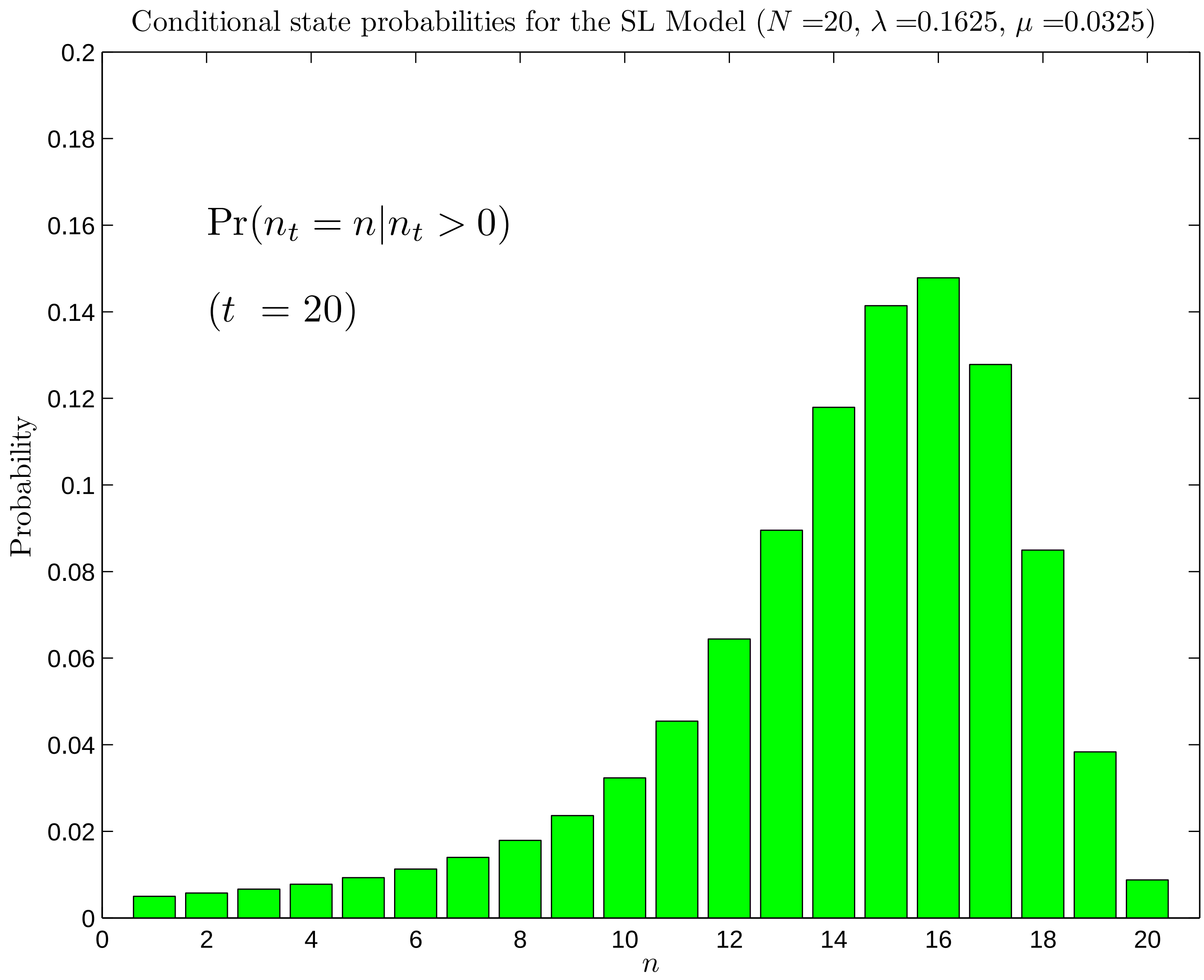
Conditional state probabilities



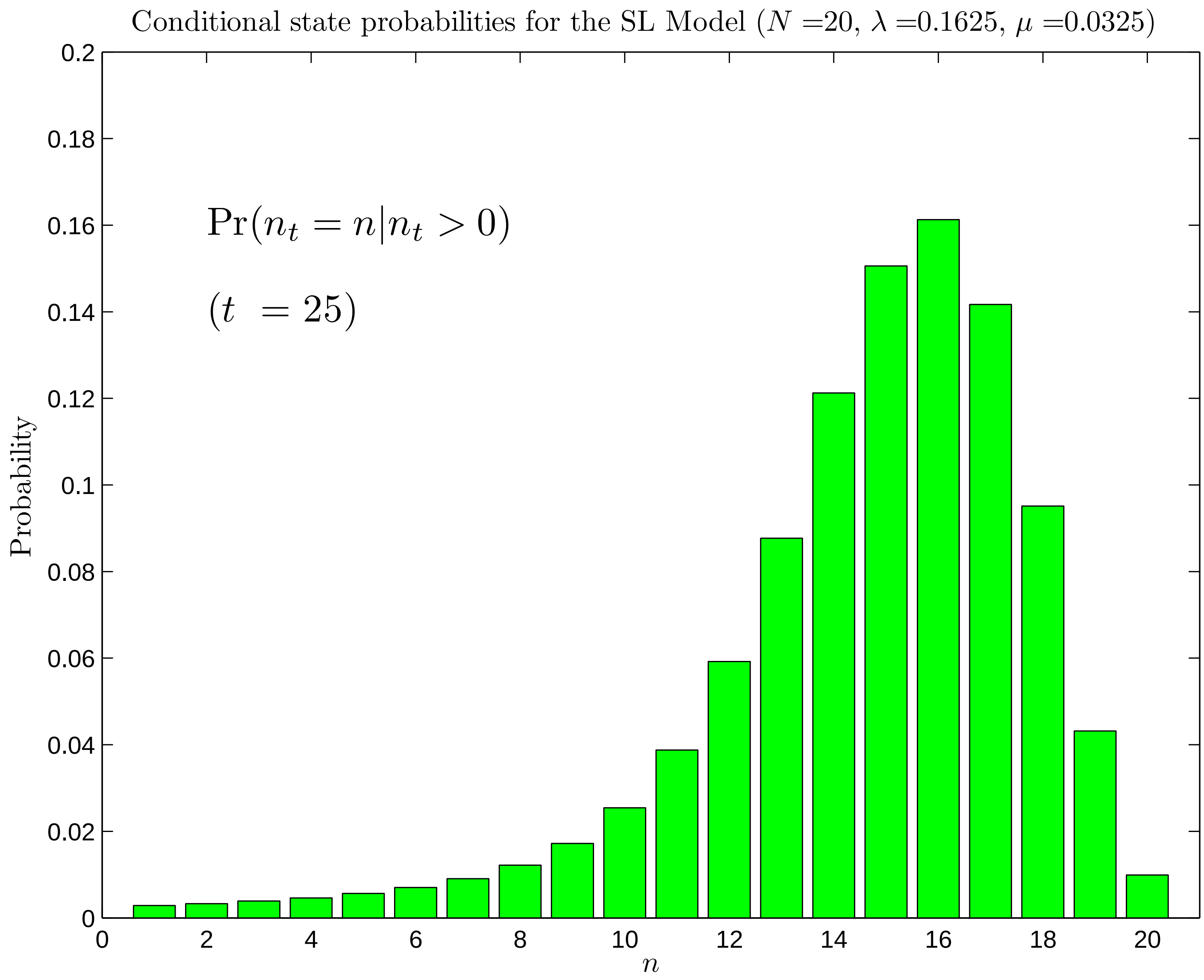
Conditional state probabilities



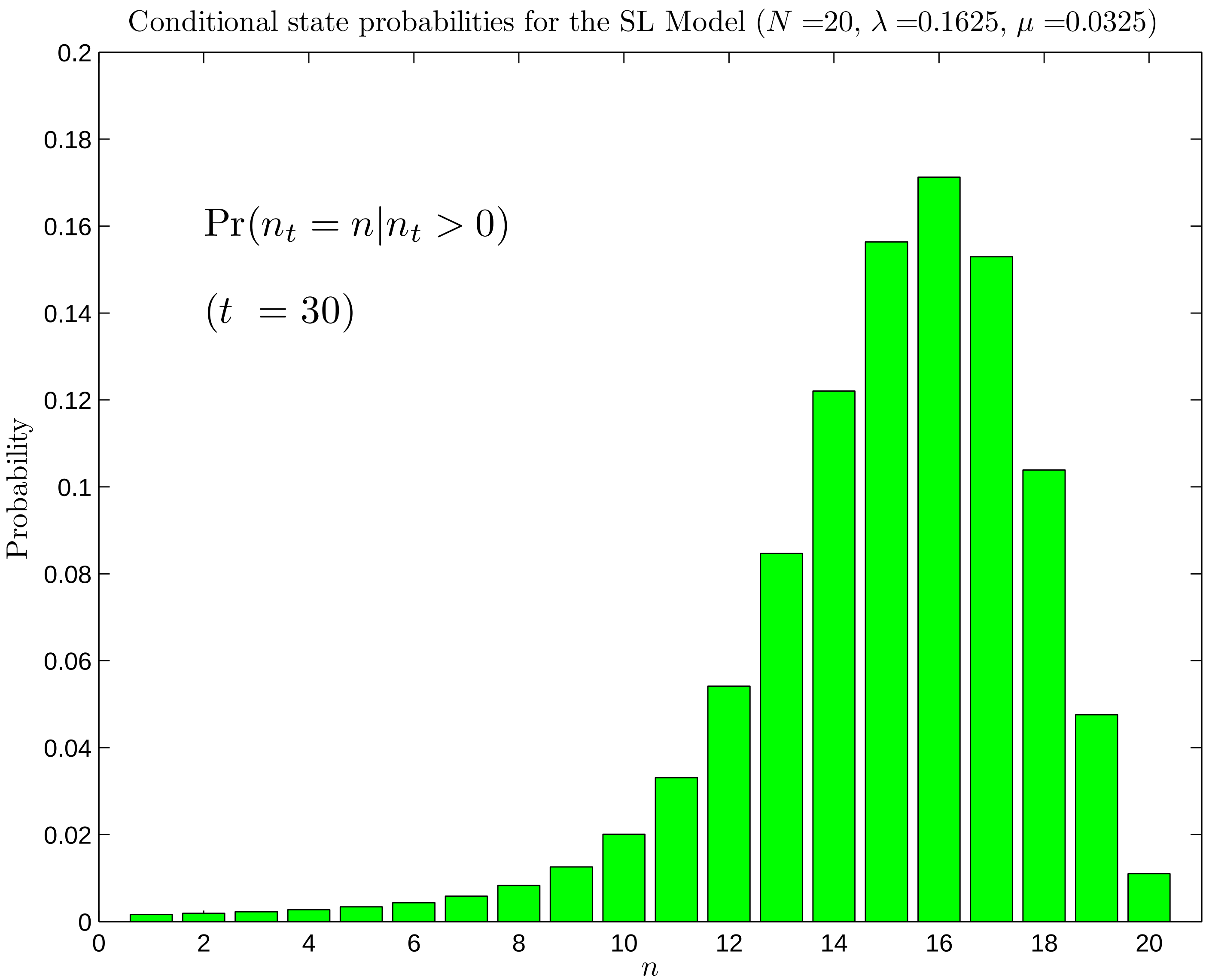
Conditional state probabilities



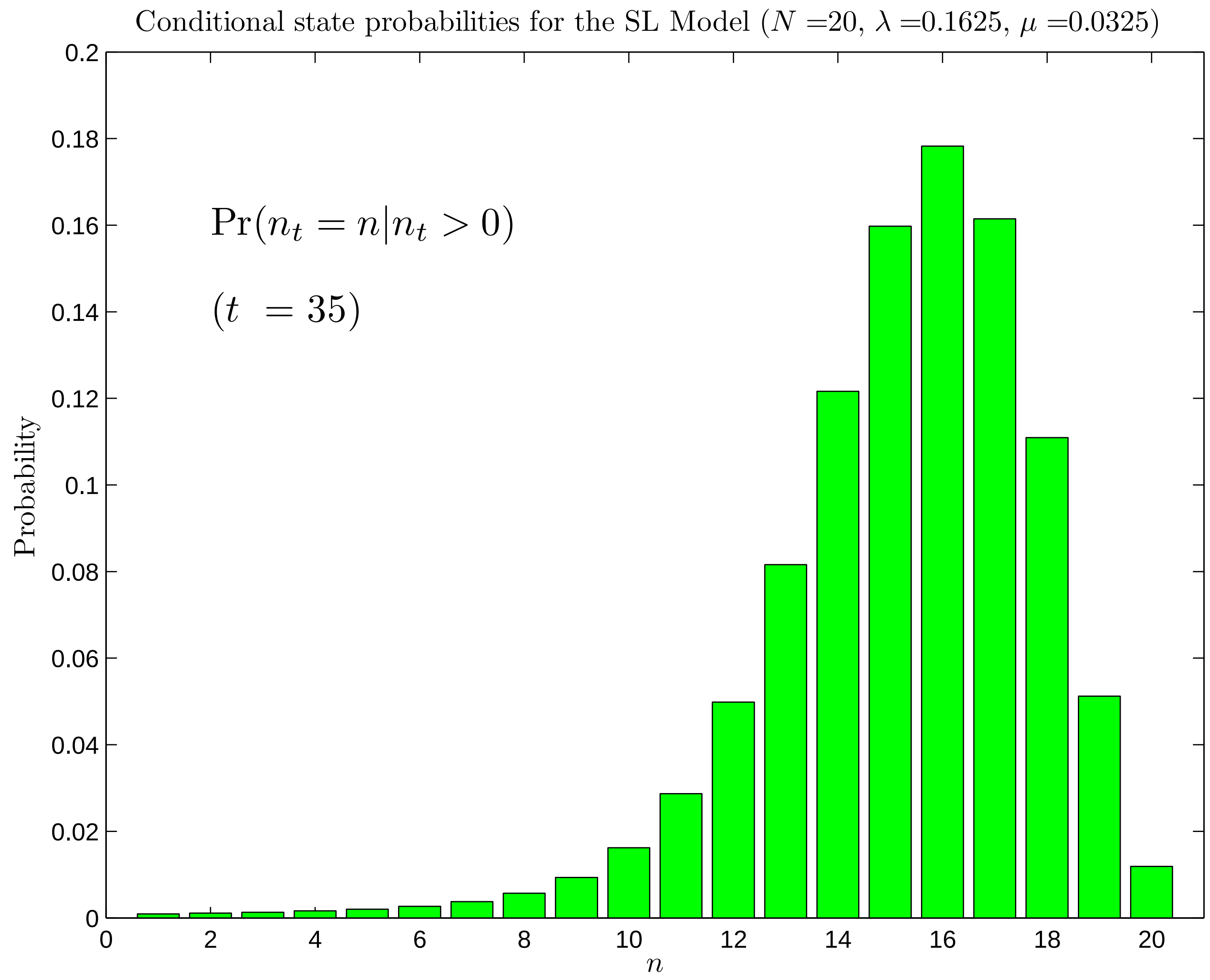
Conditional state probabilities



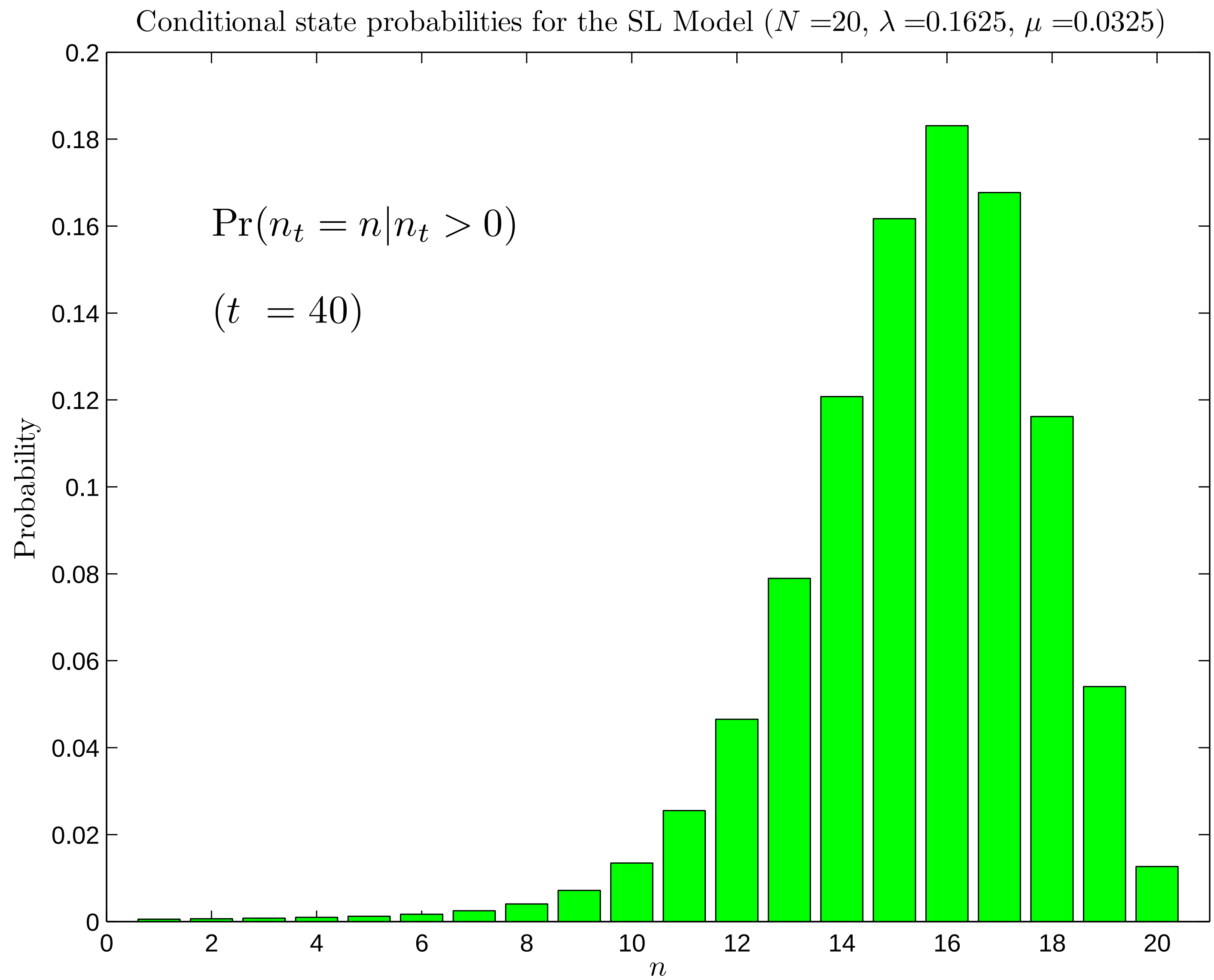
Conditional state probabilities



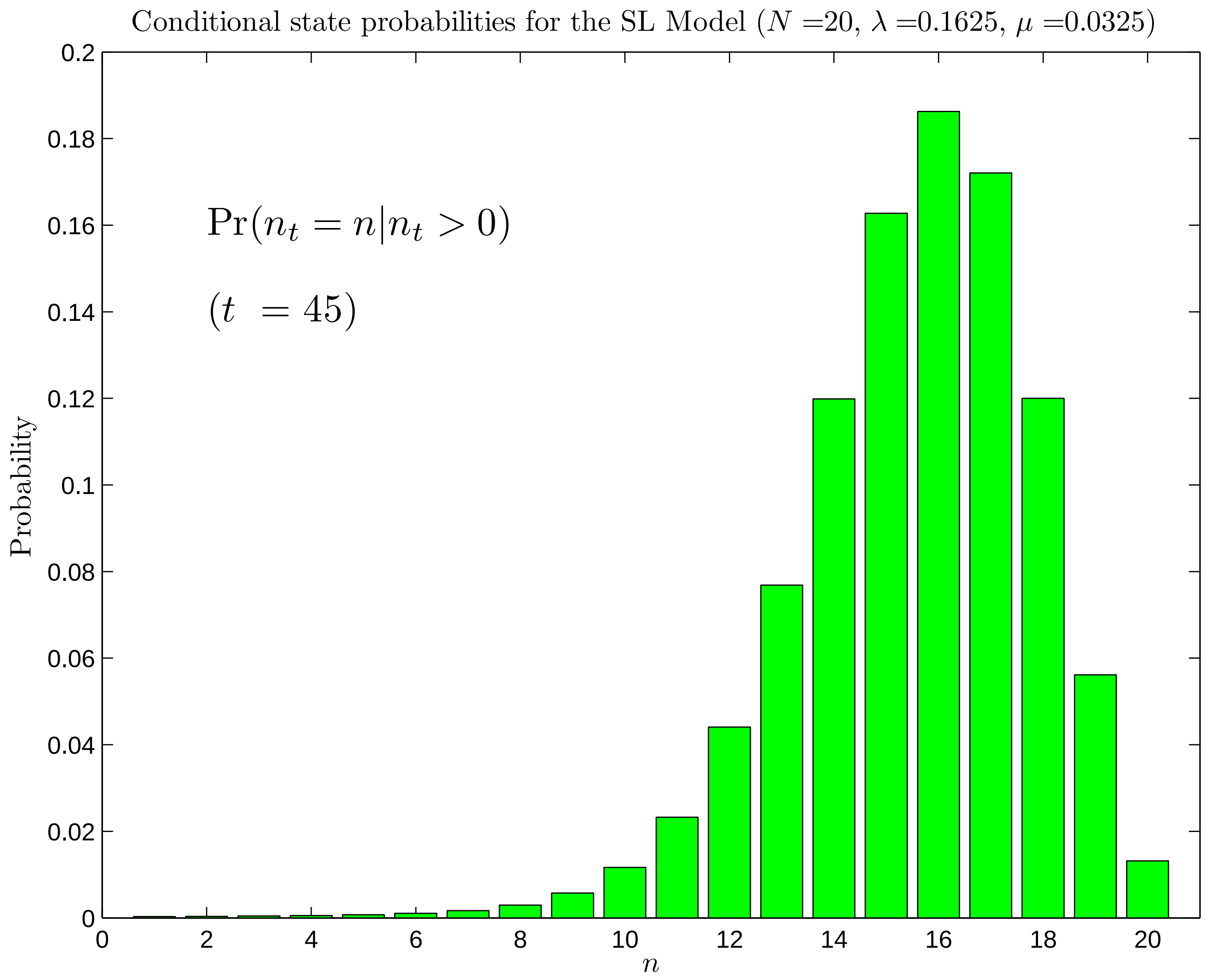
Conditional state probabilities



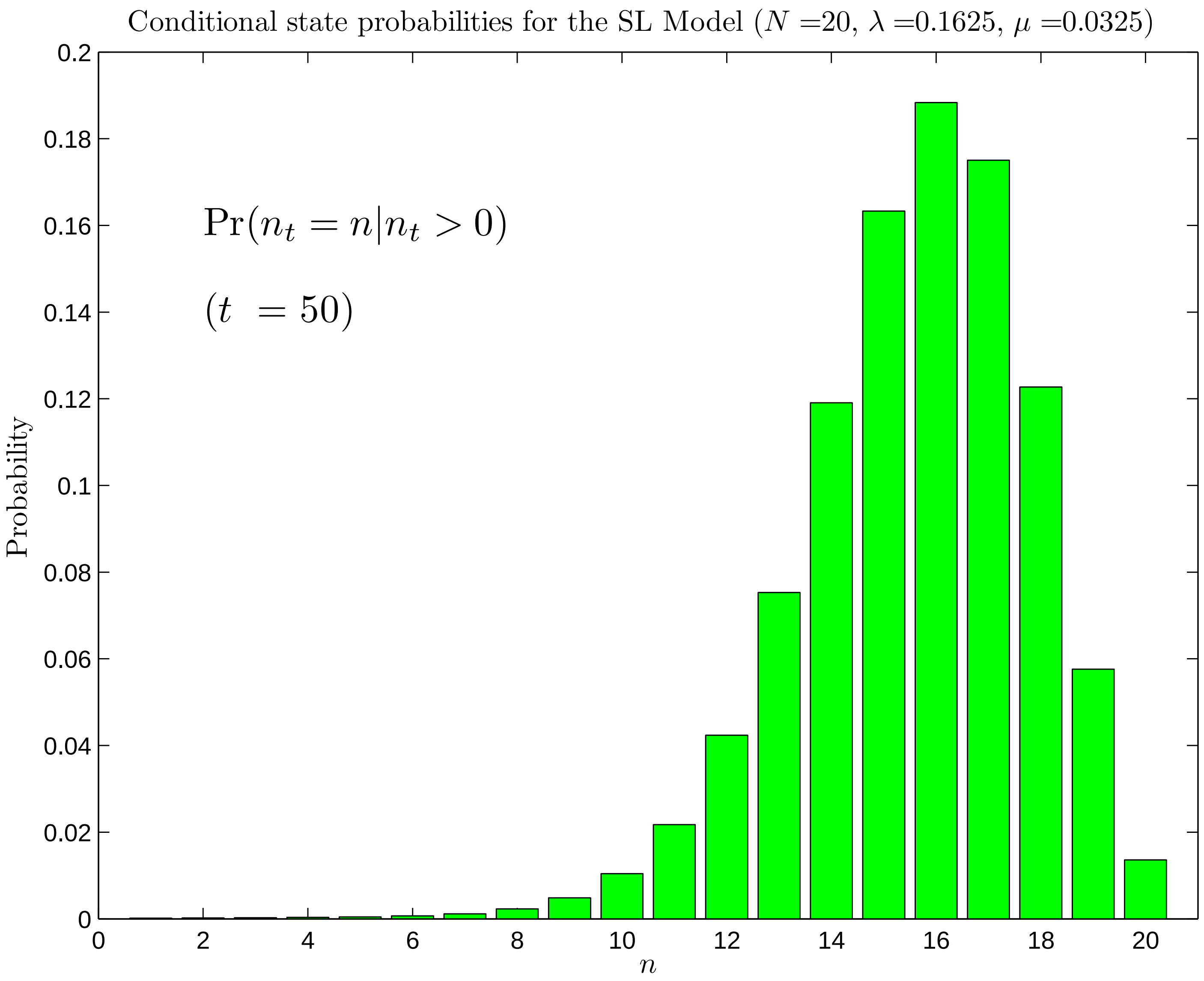
Conditional state probabilities



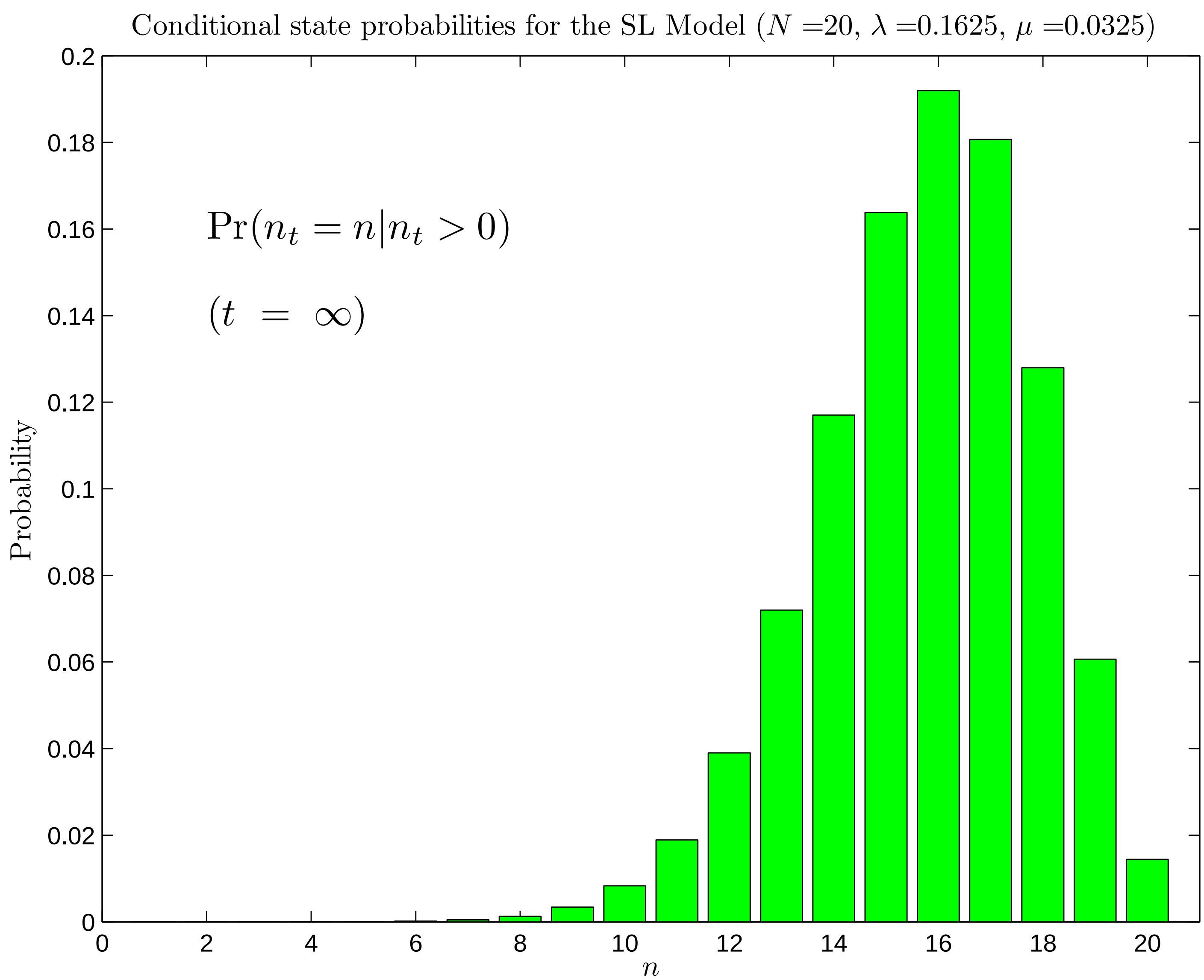
Conditional state probabilities



Conditional state probabilities



Conditional state probabilities



Modelling quasi stationarity

Recall that $S = \{0\} \cup C$, where 0 is an absorbing state and C is the set of transient states.

Define *conditional state probabilities* $\mathbf{r}(t) = (r_n(t),\, n \in C)$ by

$$r_n(t) = \Pr(n_t = n \,|\, n_t \in C),$$

the chance of being in state n *given that the process has not reached* 0.

Modelling quasi stationarity

Recall that $S = \{0\} \cup C$, where 0 is an absorbing state and C is the set of transient states.

Define *conditional state probabilities* $\mathbf{r}(t) = (r_n(t),\, n \in C)$ by

$$r_n(t) = \Pr(n_t = n \,|\, n_t \in C),$$

the chance of being in state n *given that the process has not reached* 0.

Does $\mathbf{r}(t) \to \mathbf{r}$ as $t \to \infty$?

Modelling quasi stationarity

Recall that $S = \{0\} \cup C$, where 0 is an absorbing state and C is the set of transient states.

Define *conditional state probabilities* $\mathbf{r}(t) = (r_n(t),\, n \in C)$ by

$$r_n(t) = \Pr(n_t = n \,|\, n_t \in C),$$

the chance of being in state n *given that the process has not reached* 0.

Does $\mathbf{r}(t) \to \mathbf{r}$ as $t \to \infty$?

Definition. A distribution $\mathbf{r} = (r_n,\, n \in C)$ satisfying $\mathbf{r}(t) = \mathbf{r}$ for all $t > 0$ is called a *quasi-stationary distribution* (QSD). If $\mathbf{r}(t) \to \mathbf{r}$ then $\mathbf{r}$ is a *limiting-conditional distribution* (LCD).

Modelling quasi stationarity

Definition. A distribution $\mathbf{r} = (r_n,\, n \in C)$ satisfying $\mathbf{r}(t) = \mathbf{r}$ for all $t > 0$ is called a *quasi-stationary distribution* (QSD). If $\mathbf{r}(t) \to \mathbf{r}$ then $\mathbf{r}$ is a *limiting-conditional distribution* (LCD).

So, we may think of a QSD as being an *equilibrium point* $\mathbf{r}$ of the master equation governing the evolution of the *conditional state probabilities* $\mathbf{r}(t) = (r_n(t),\, n \in C)$, where recall that

$$r_n(t) = \Pr(n_t = n \,|\, n_t \in C) \qquad (n \in C,\ t > 0).$$

And, if $\mathbf{r}$ is *asymptotically stable*, then $\mathbf{r}$ is an LCD.

Modelling quasi stationarity

Definition. A distribution $\mathbf{r} = (r_n,\, n \in C)$ satisfying $\mathbf{r}(t) = \mathbf{r}$ for all $t > 0$ is called a *quasi-stationary distribution* (QSD). If $\mathbf{r}(t) \to \mathbf{r}$ then $\mathbf{r}$ is a *limiting-conditional distribution* (LCD).

So, we may think of a QSD as being an *equilibrium point* $\mathbf{r}$ of the master equation governing the evolution of the *conditional state probabilities* $\mathbf{r}(t) = (r_n(t),\, n \in C)$, where recall that

$$r_n(t) = \Pr(n_t = n \,|\, n_t \in C) \qquad (n \in C,\; t > 0).$$

And, if $\mathbf{r}$ is *asymptotically stable*, then $\mathbf{r}$ is an LCD.

So, what is the master equation for $\mathbf{r}(t)$?

Some calculations

For $n \in C$,

$$r_n(t) = \Pr(n_t = n \,|\, n_t \in C)$$
$$= \frac{\Pr(n_t = n)}{\Pr(n_t \in C)} = \frac{p_n(t)}{\sum_{m \in C} p_m(t)} = \frac{p_n(t)}{1 - p_0(t)}.$$

Therefore,

$$r_n'(t) = \frac{p_n'(t)}{1 - p_0(t)} + p_n(t)\,\frac{p_0'(t)}{(1 - p_0(t))^2}$$
$$= \frac{p_n'(t)}{1 - p_0(t)} + r_n(t)\,\frac{p_0'(t)}{1 - p_0(t)} \quad \text{(now use FEs for } p_n(t)\text{)}$$
$$= \sum_{m \in C} r_m(t) q_{mn} + r_n(t) \sum_{m \in C} r_m(t) q_{m0}.$$

Modelling quasi stationarity

We arrive at

$$r_n'(t) = \sum_{m \in C} r_m(t) q_{mn} + r_n(t) \sum_{m \in C} r_m(t) q_{m0}.$$

Since $\sum_{n \in S} q_{mn} = 0$, this can be written

$$\mathbf{r}'(t) = \mathbf{r}(t) Q_C - \nu(t) \mathbf{r}(t),$$

where $\nu(t) = \mathbf{r}(t) Q_C \mathbf{1}$, and Q_C is the restriction of Q to C.

Modelling quasi stationarity

We arrive at

$$r'_n(t) = \sum_{m \in C} r_m(t) q_{mn} + r_n(t) \sum_{m \in C} r_m(t) q_{m0}.$$

Since $\sum_{n \in S} q_{mn} = 0$, this can be written

$$\mathbf{r}'(t) = \mathbf{r}(t) Q_C - \nu(t) \mathbf{r}(t),$$

where $\nu(t) = \mathbf{r}(t) Q_C \mathbf{1}$, and Q_C is the restriction of Q to C.

Formally we have $\mathbf{r}(t) \to \mathbf{r}$, where $\mathbf{r}$ satisfies

$$\mathbf{r} Q_C = \nu \mathbf{r},$$

so that $\mathbf{r} = (r_n,\, n \in C)$ is a left eigenvector of Q_C corresponding to a (strictly negative) real eigenvalue ν. Postmultiplying by $\mathbf{1}$ gives $\nu = \mathbf{r} Q_C \mathbf{1}$, or, written out, $\nu = -\sum_{n \in C} r_n q_{n0}$.

Modelling quasi stationarity

If the state space is finite, this can be justified using classical *Perron-Frobenius* theory.

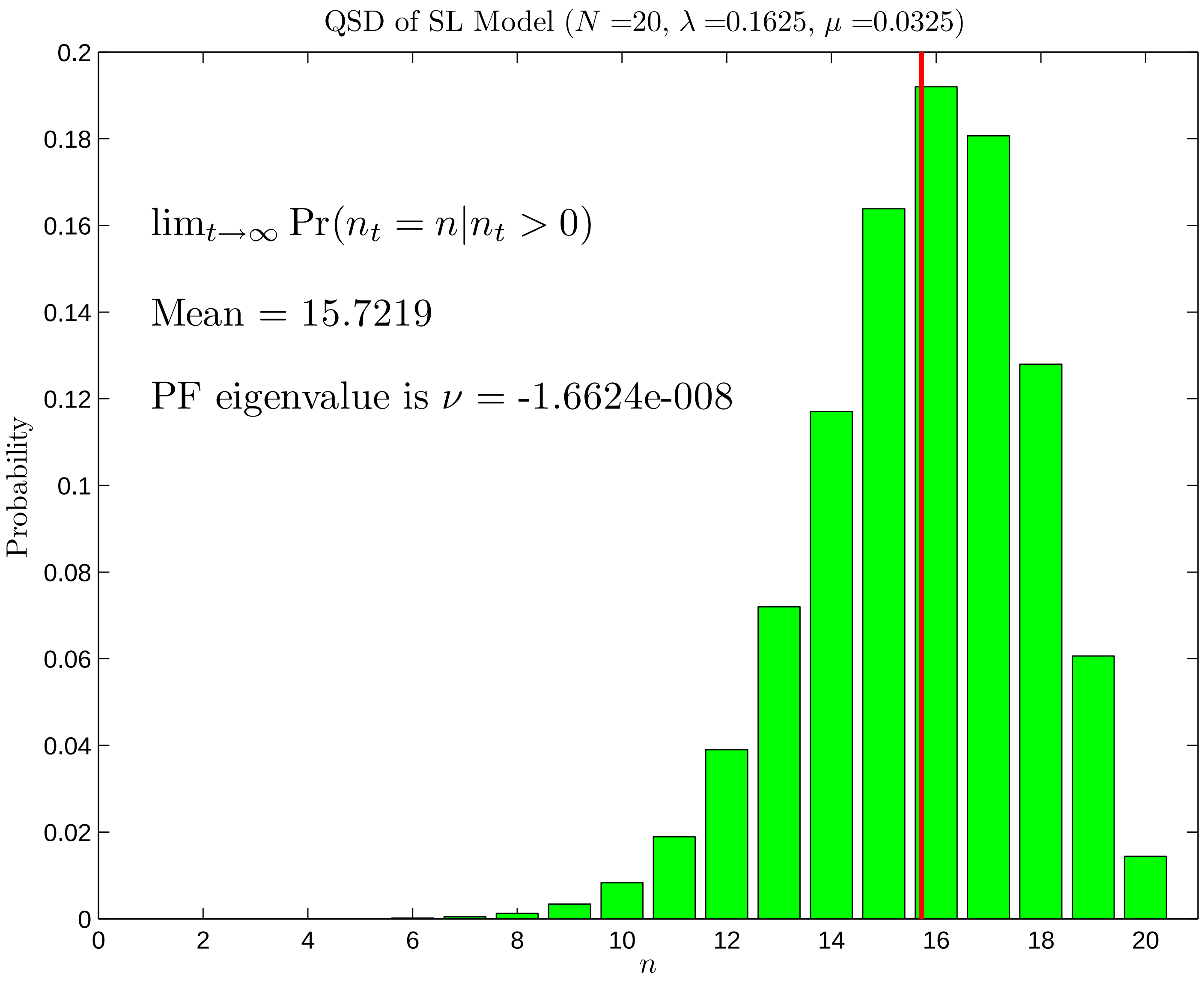
Theorem The restriction Q_C of Q to C has eigenvalues with strictly negative real parts and the one with maximal real part (called ν above) is real and has multiplicity 1, and, the corresponding left eigenvector $\mathbf{x} = (x_n,\, n \in C)$ has strictly positive entries.

The quasi-stationary distribution $\mathbf{r} = (r_n,\, n \in C)$ exists uniquely and is given by $r_n = x_n / \sum_{m \in C} x_m$. Moreover, $\mathbf{r}$ is the limiting-conditional distribution. In particular, if $\Pr(n_0 \in C) = 1$,

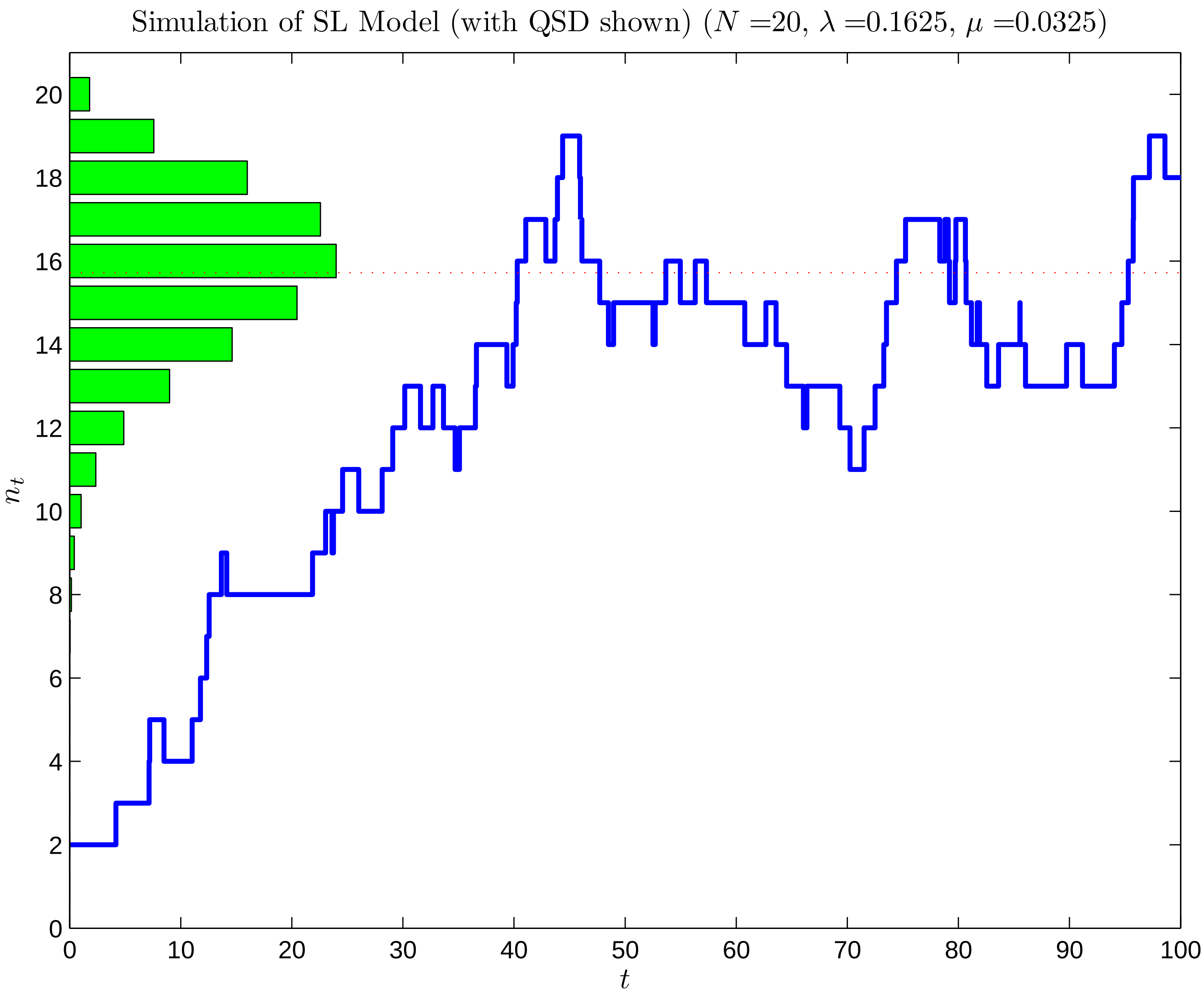
$$\Pr(n_t = n \,|\, n_t \in C) \to r_n \qquad \text{as} \quad t \to \infty,$$

the limit being the *same* for all initial distributions.

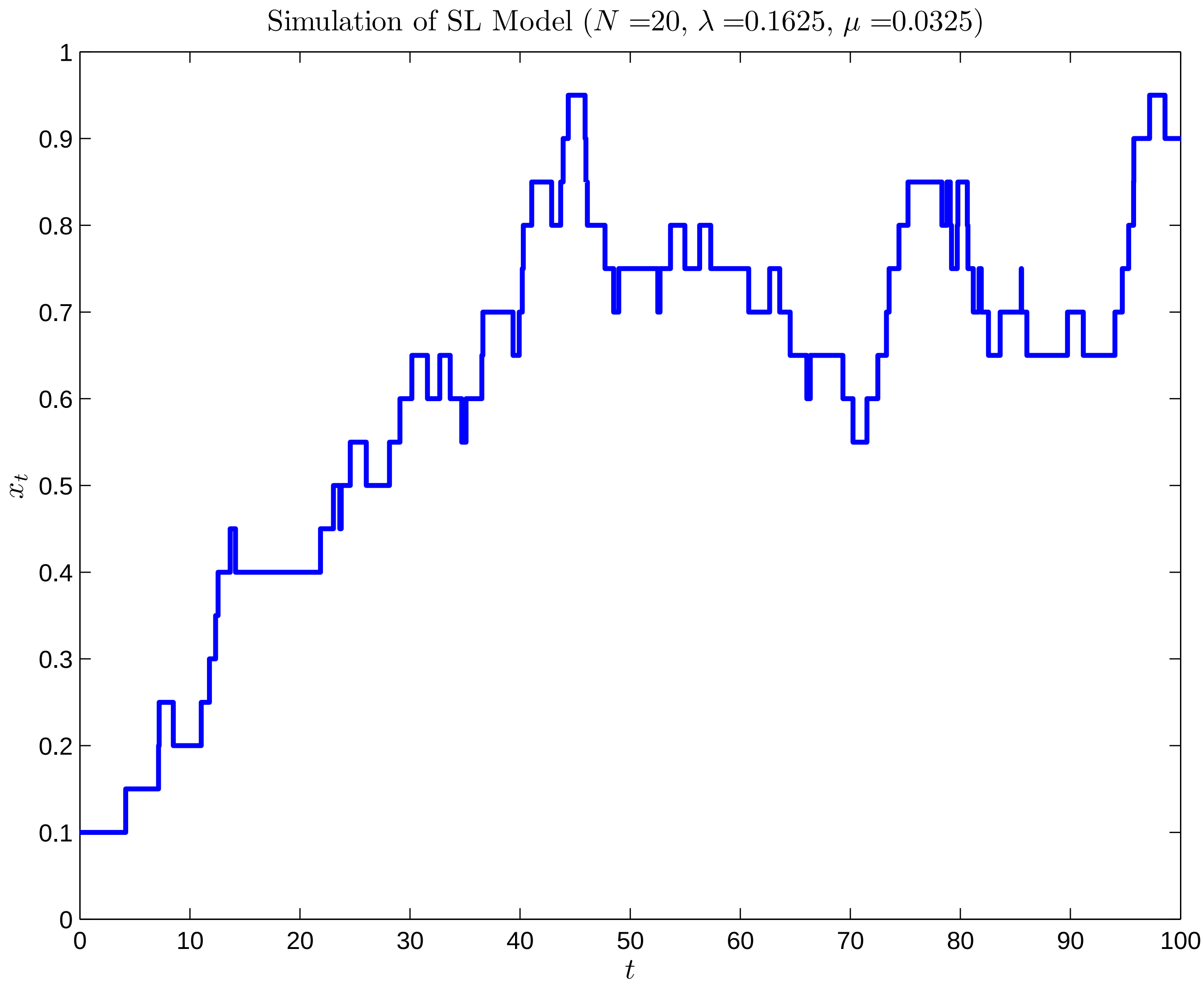
QSD of the SL model



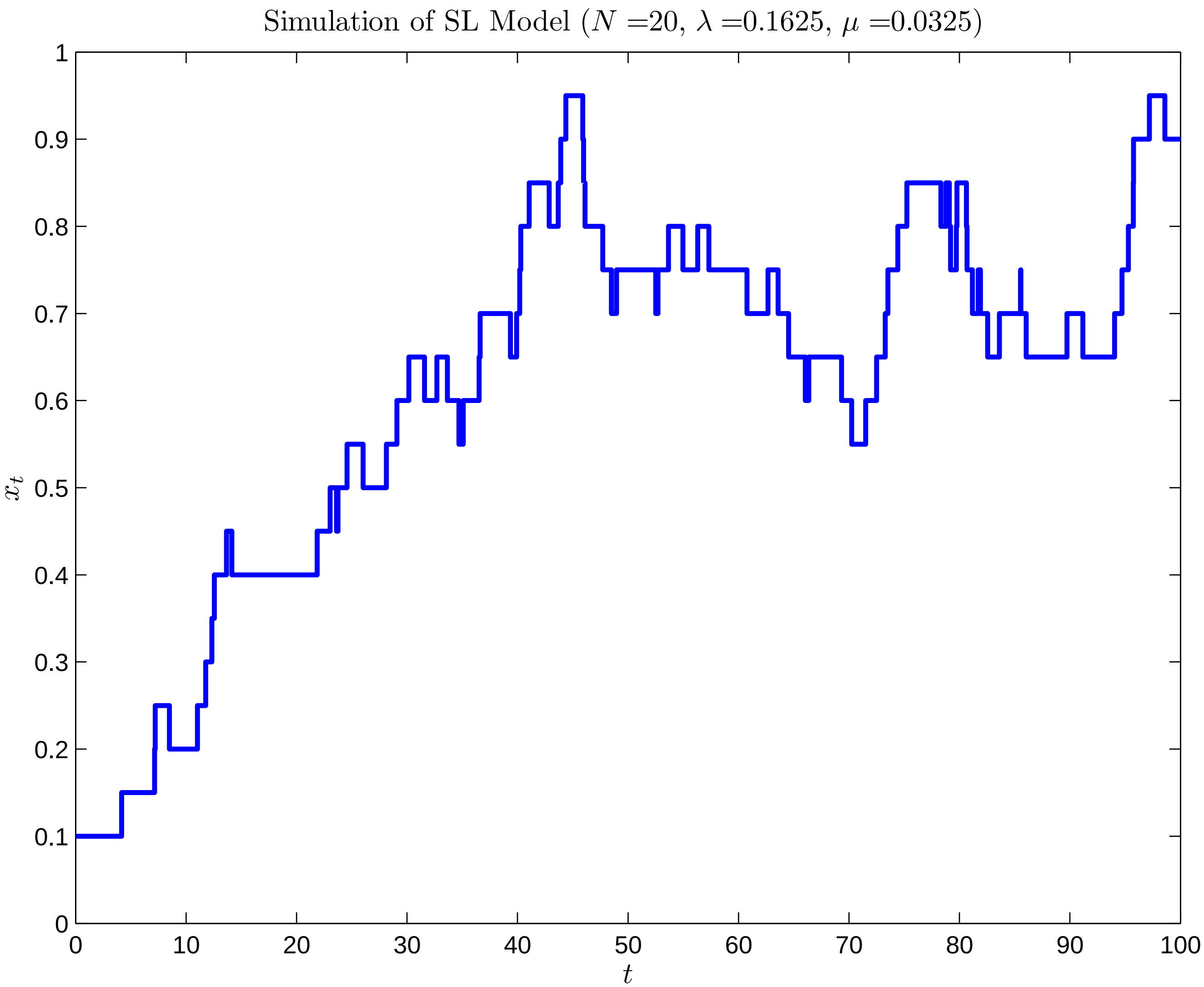
QSD of the SL model



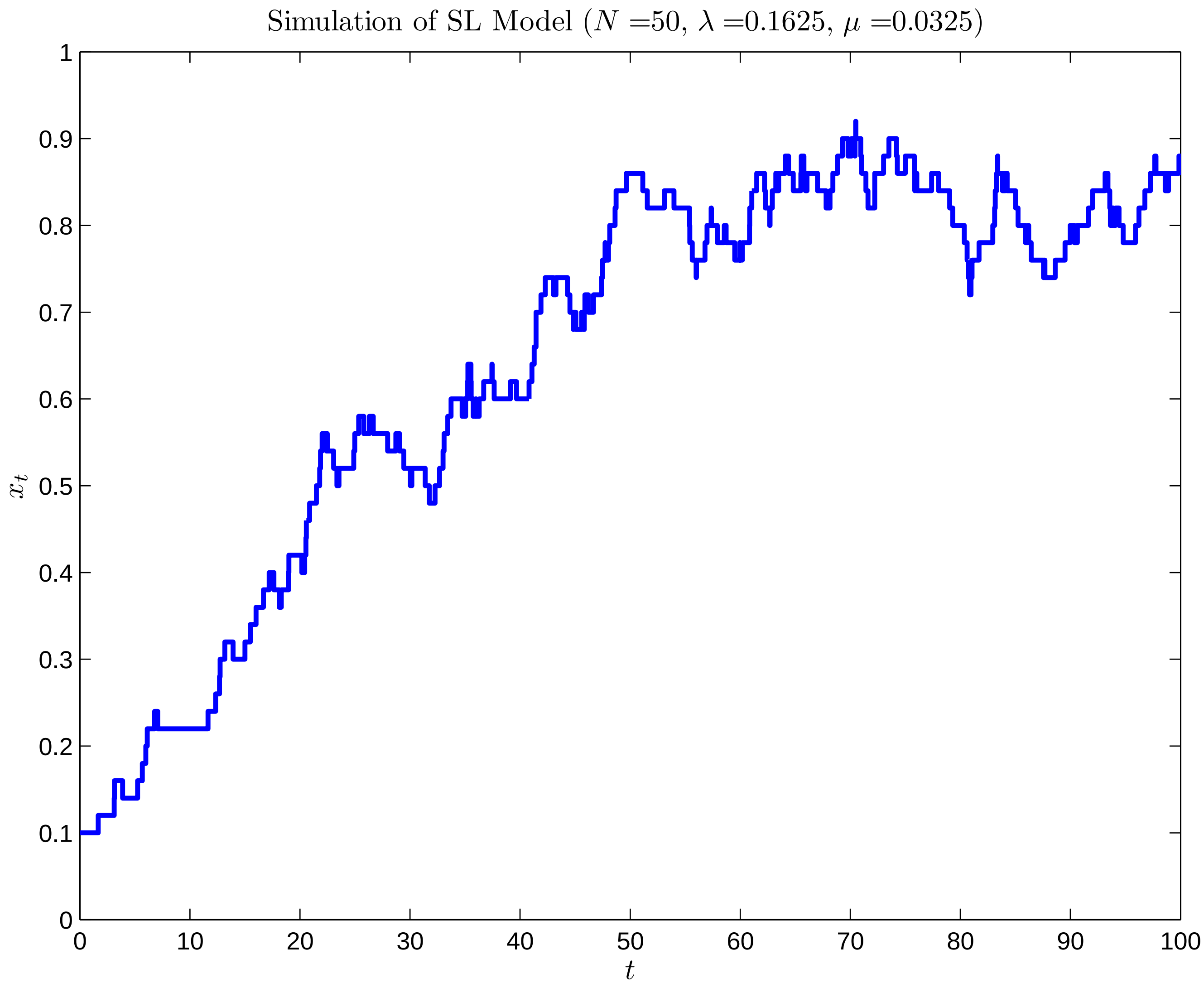
Proportion of patches occupied



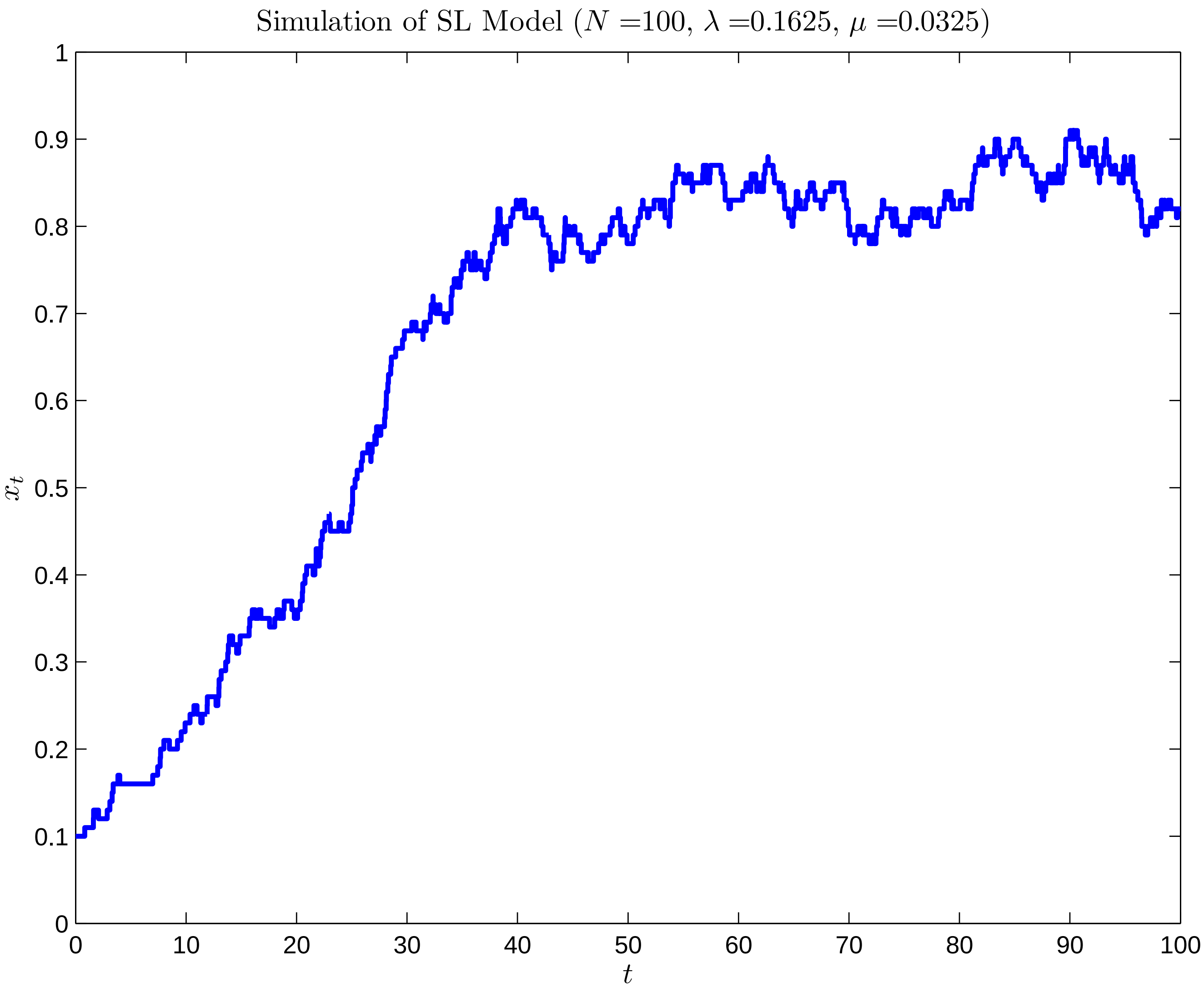
The SL model $(N = 20)$



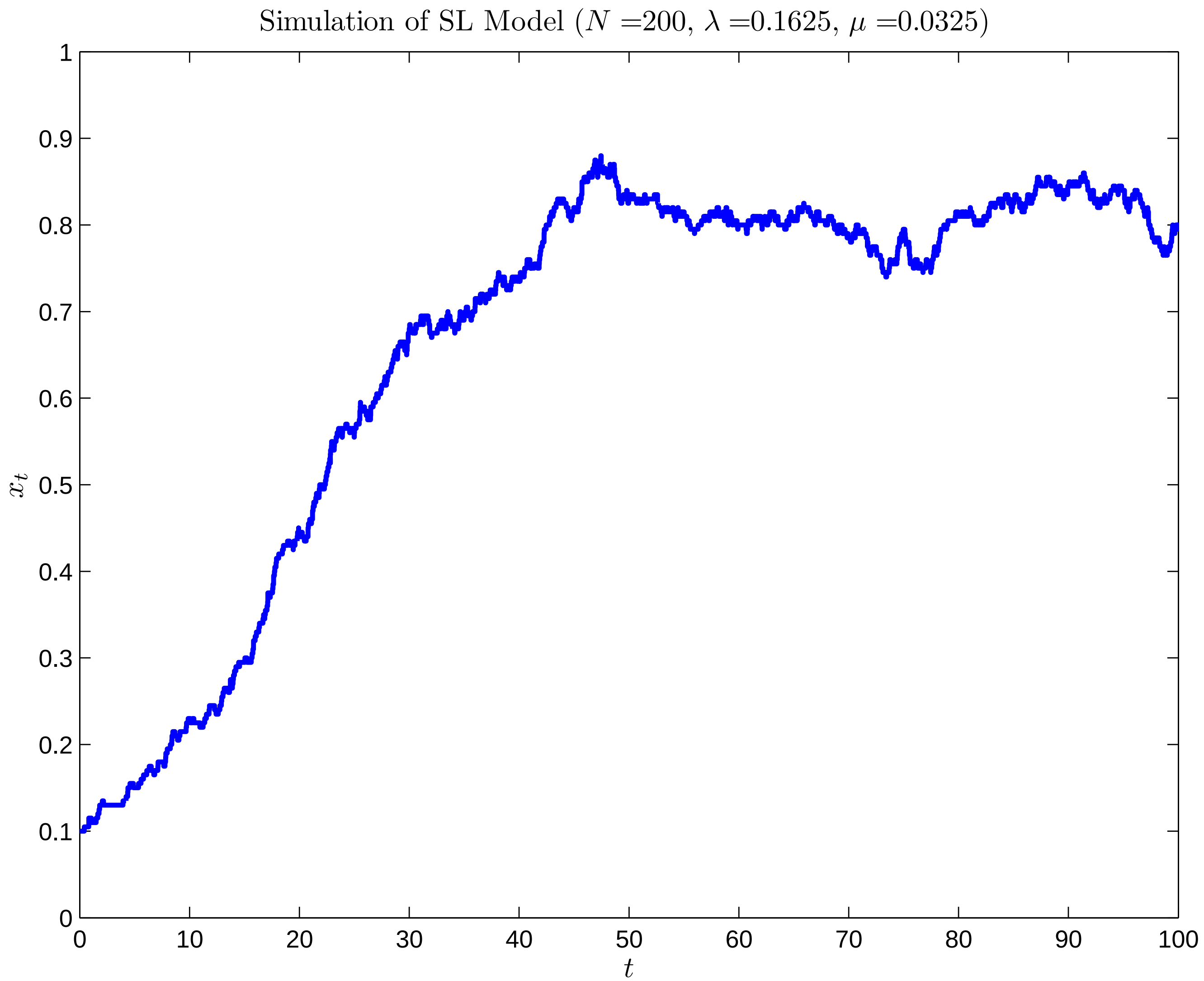
The SL model ($N = 50$)



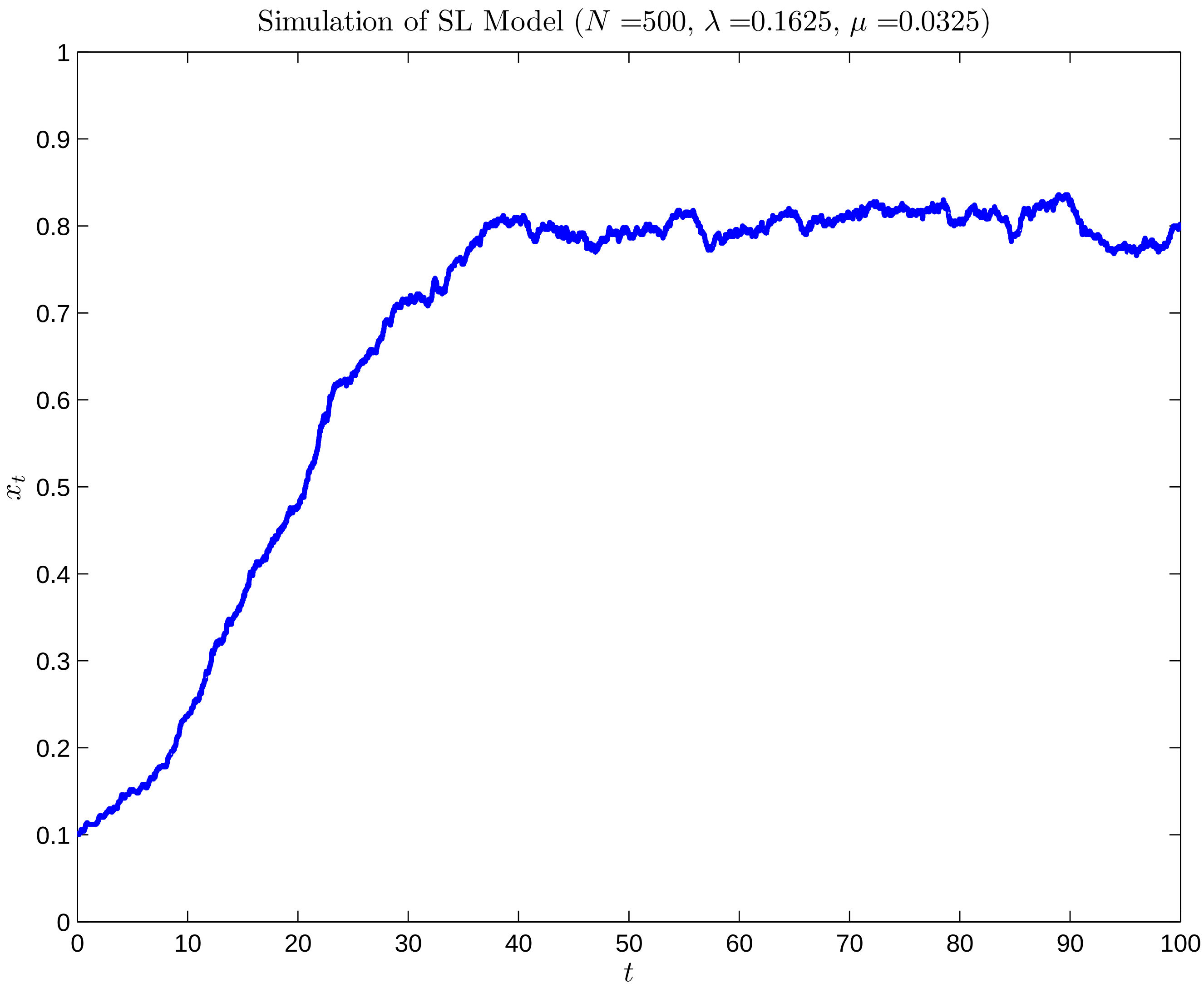
The SL model $(N = 100)$



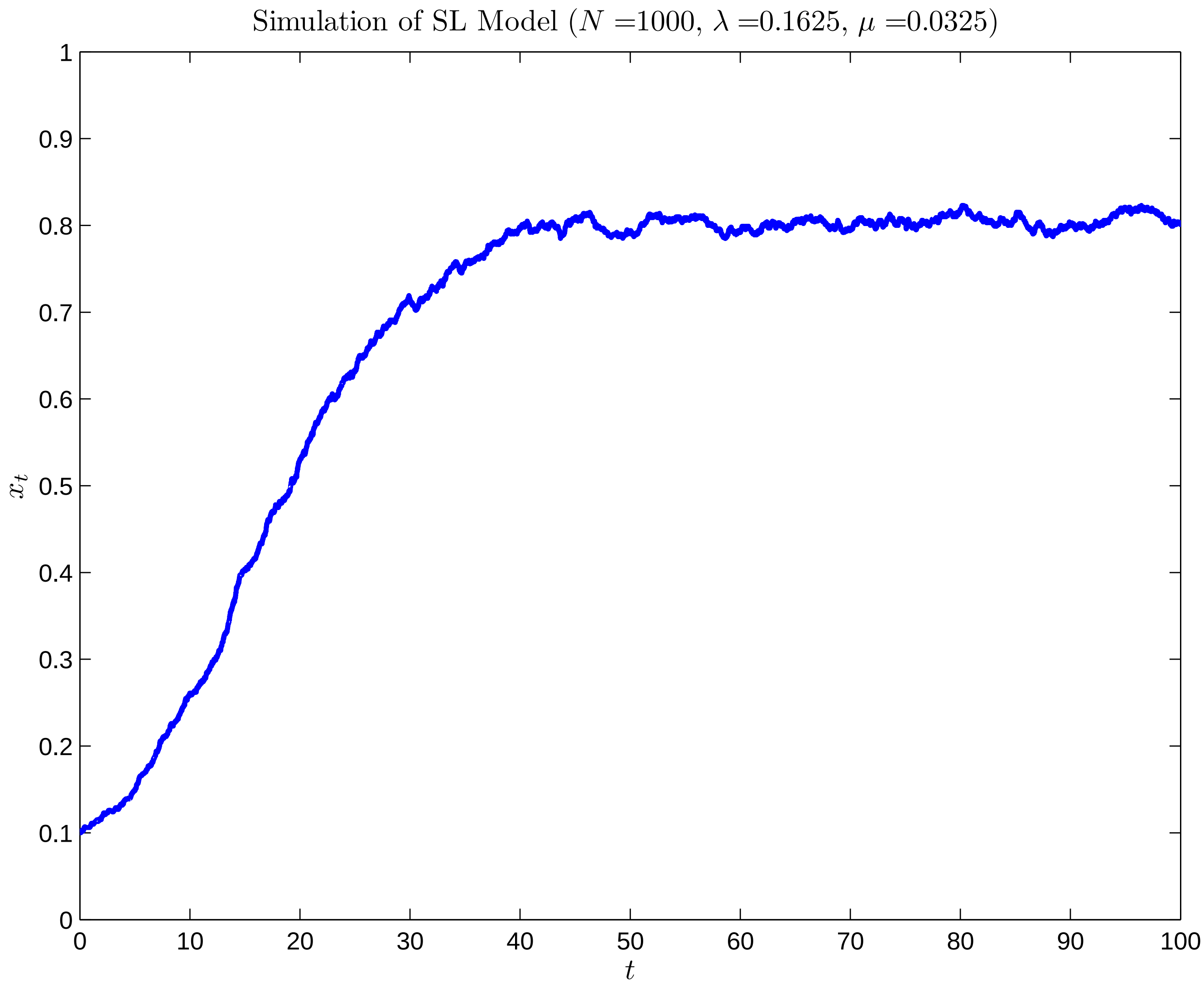
The SL model $(N = 200)$



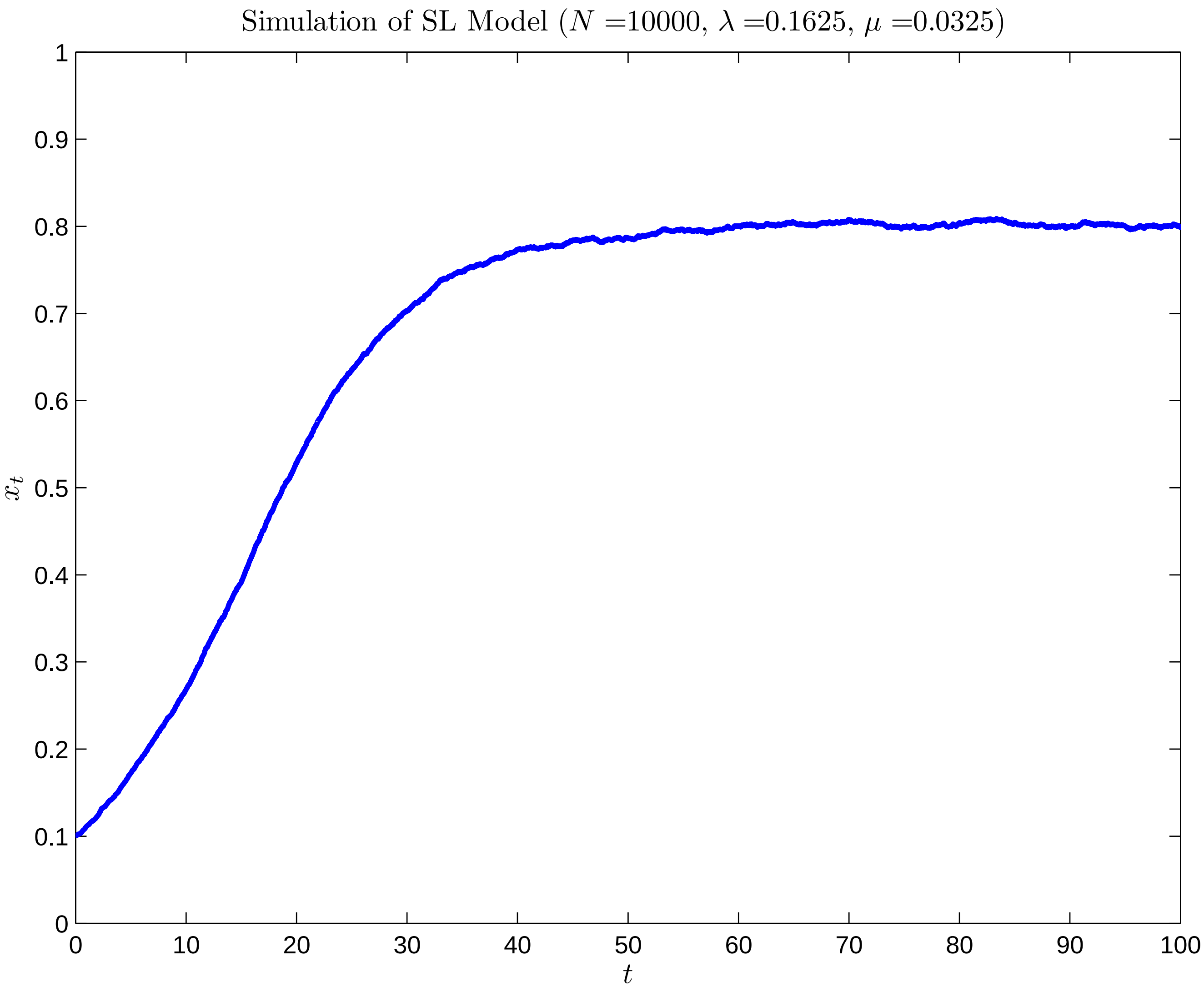
The SL model $(N = 500)$



The SL model $(N = 1\,000)$



The SL model $(N = 10\,000)$



Density dependence

The idea is the same as for deterministic models: the rate of change of n_t depends on n_t only through the “density” n_t/N:

$$n \to n + l \qquad \text{at rate} \qquad N f_l\left(\frac{n}{N}\right) \qquad (l \neq 0)$$

for suitable functions $f_l(x)$.

The analogous (approximating!) deterministic model for the “density” $x_t := n_t/N$ is

$$\frac{dx}{dt} = F(x) := \sum_{l \neq 0} l\, f_l(x).$$

The SL model

For the SL model we have $S = \{0, 1, \dots, N\}$ and transitions:

$$\begin{aligned} n \to n+1 \quad &\text{at rate} \quad \frac{\lambda}{N} n\,(N-n) &= N\lambda \frac{n}{N}\left(1 - \frac{n}{N}\right) \\ n \to n-1 \quad &\text{at rate} \quad \mu n &= N\mu \frac{n}{N} \end{aligned}$$

Therefore, $f_{+1}(x) = \lambda x\,(1-x)$ and $f_{-1}(x) = \mu x$, $x \in E := [0, 1]$, and so $F(x) = \lambda x\,(q - x)$, $x \in E$, where $q = 1 - \mu/\lambda$.

The SL model

For the SL model we have $S = \{0, 1, \ldots, N\}$ and transitions:

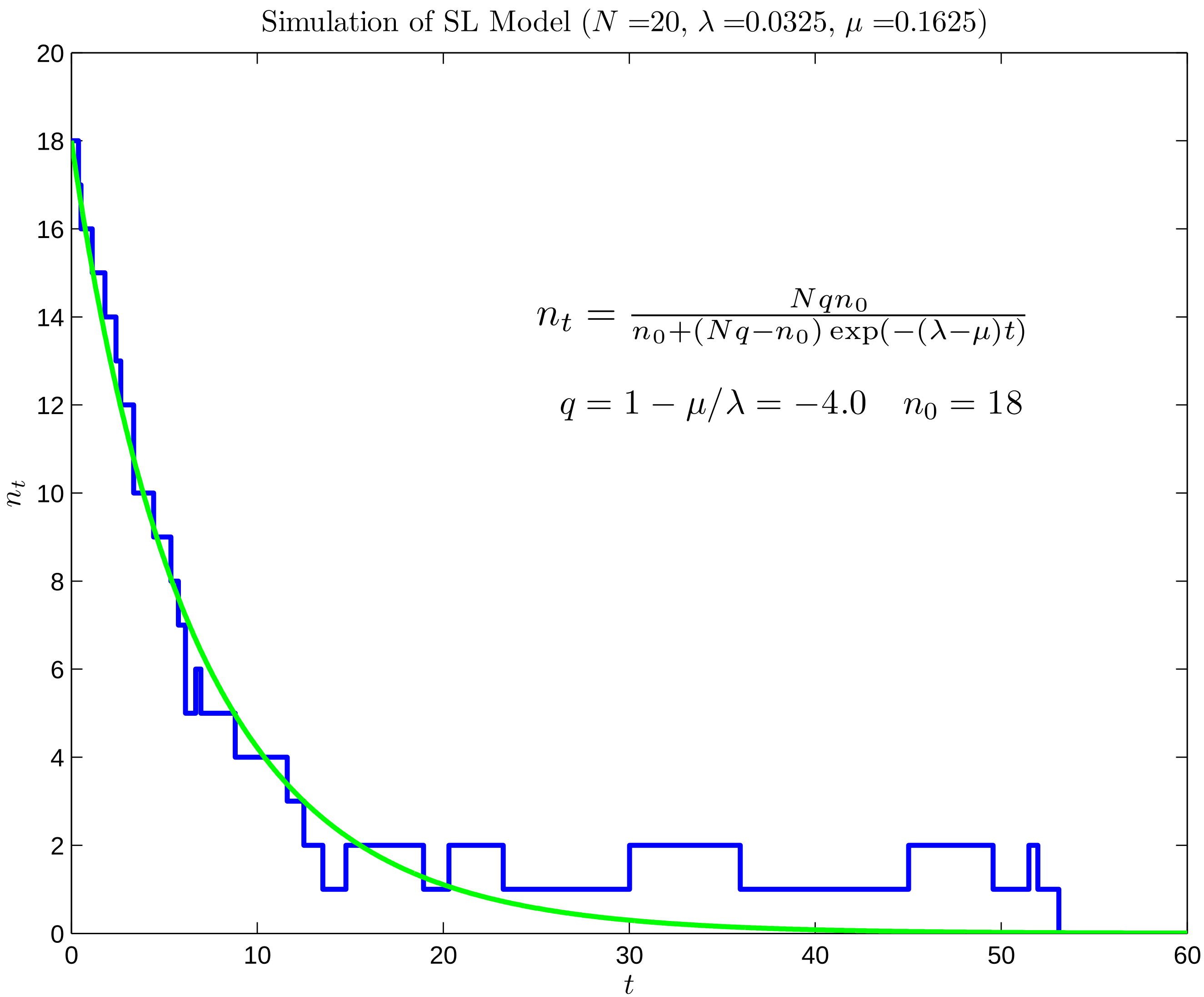
$$\begin{array}{llll} n \to n+1 & \text{at rate} & \frac{\lambda}{N} n\,(N-n) & = N\lambda \frac{n}{N}\left(1 - \frac{n}{N}\right) \\ n \to n-1 & \text{at rate} & \mu n & = N\mu \frac{n}{N} \end{array}$$

Therefore, $f_{+1}(x) = \lambda x\,(1-x)$ and $f_{-1}(x) = \mu x$, $x \in E := [0,1]$, and so $F(x) = \lambda x\,(q-x)$, $x \in E$, where $q = 1 - \mu/\lambda$.

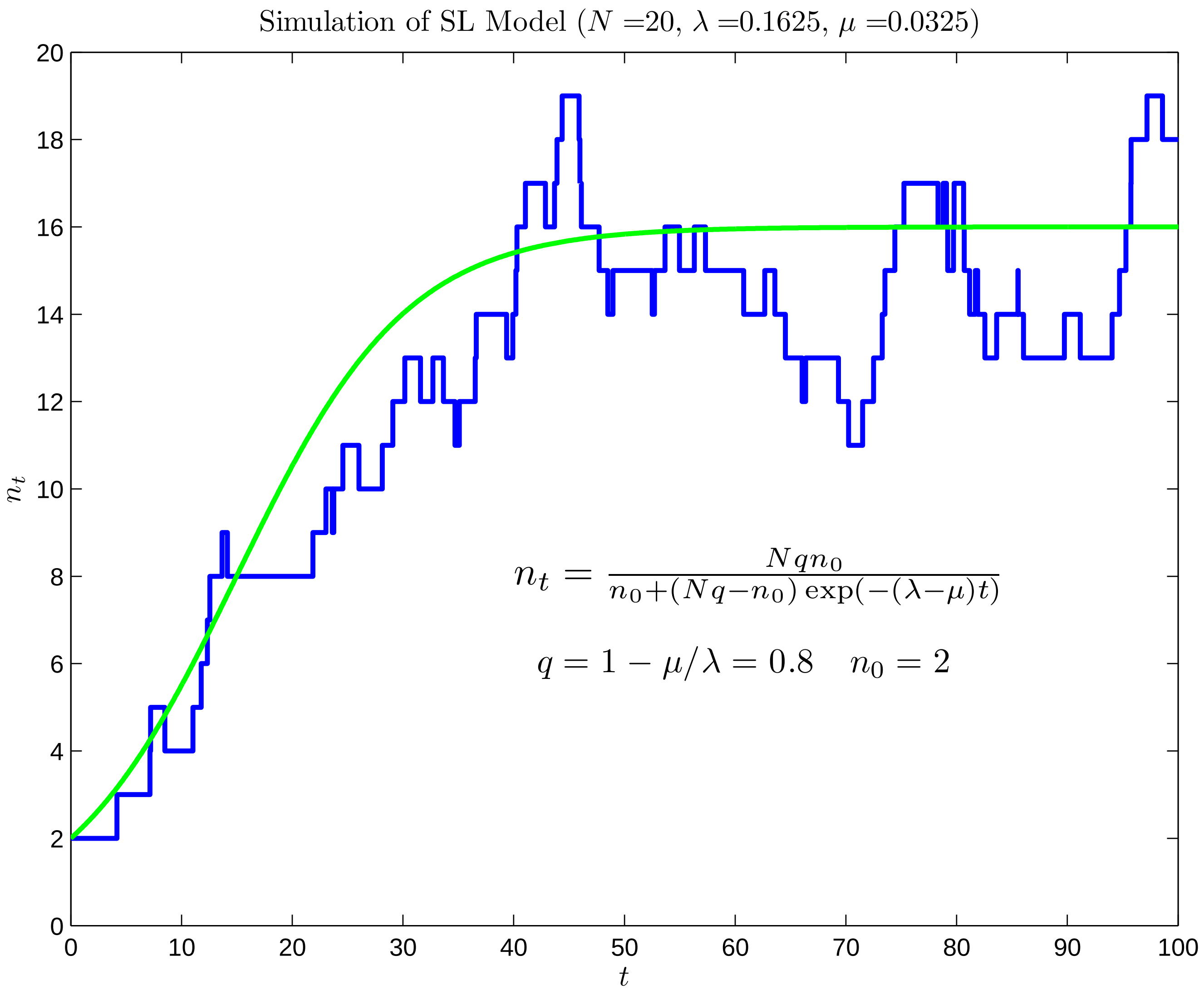
We arrive at the classical Verhulst (1838) model $x_t' = \lambda x_t\,(q - x_t)$, which for us describes the proportion of occupied patches. It has the unique solution

$$x_t = \frac{q\,x_0}{x_0 + (q - x_0)\,e^{-(\lambda-\mu)t}} \qquad (t \geq 0).$$

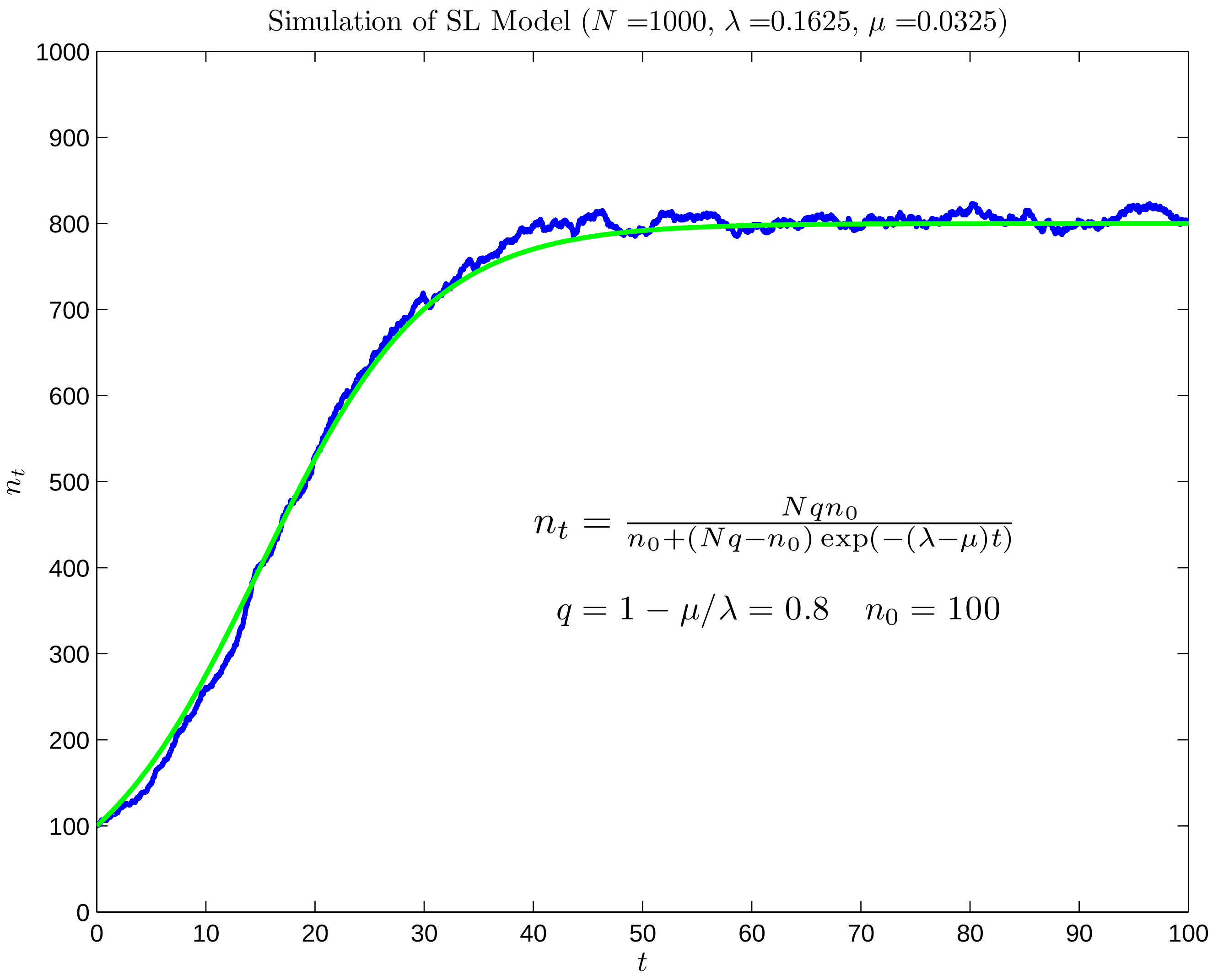
The SL model $(\lambda < \mu)$



The SL model $(\lambda > \mu)$



The SL model $(N = 1000)$



Density dependence of MCs

Let $(n_t,\, t \geq 0)$ be a continuous-time Markov chain taking values in $S \subseteq \mathrm{Z}^k$ with transition rates $Q = (q_{nm},\, n, m \in S)$.

We identify a quantity N, usually related to the size of the system being modelled (for example, volume, area, number of patches, population ceiling).

Definition (Kurtz*) The model is *density dependent* if there is a subset E of R^k and a continuous function $f : \mathrm{Z}^k \times E \to \mathrm{R}$, such that

$$q_{n,n+l} = N f_l\left(\tfrac{n}{N}\right), \quad l \neq 0 \quad (l \in \mathrm{Z}^k).$$

*Kurtz, T. (1970) Solutions of ordinary differential equations as limits of pure jump Markov processes, *J. of Appl. Probab.* 7, 49–58.

Density dependence of MCs

Let $(n_t,\, t \geq 0)$ be a continuous-time Markov chain taking values in $S \subseteq \mathrm{Z}^k$ with transition rates $Q = (q_{nm},\, n, m \in S)$.

We identify a quantity N, usually related to the size of the system being modelled (for example, volume, area, number of patches, population ceiling).

Definition (Kurtz*) The model is *density dependent* if there is a subset E of R^k and a continuous function $f : \mathrm{Z}^k \times E \to \mathrm{R}$, such that

$$q_{n,n+l} = N f_l\left(\frac{n}{N}\right), \quad l \neq 0 \quad (l \in \mathrm{Z}^k).$$

*Kurtz, T. (1970) Solutions of ordinary differential equations as limits of pure jump Markov processes, *J. of Appl. Probab.* 7, 49–58.

Density dependence of MCs

We now formally define the *density process* $(X_t^{(N)})$ by

$$X_t^{(N)} = n_t/N \qquad (t \geq 0).$$

This is a Markov chain that takes values in the lattice $S_N := S/N$ and has transition rates $q_{x,x+l/N}$, $x \in S_N$, $l \in \mathrm{Z}^k$.

We hope that $(X_t^{(N)})$ becomes more deterministic as N gets large. Moreover, we anticipate that the limiting deterministic trajectory satisfies $x_t' = F(x_t)$, where

$$F(x) = \sum_{l \neq 0} l f_l\,(x) \qquad (x \in E).$$

Density dependence of MCs

We now formally define the *density process* $(X_t^{(N)})$ by

$$X_t^{(N)} = n_t/N \qquad (t \geq 0).$$

This is a Markov chain that takes values in the lattice $S_N := S/N$ and has transition rates $q_{x,x+l/N}$, $x \in S_N$, $l \in \mathrm{Z}^k$.

We hope that $(X_t^{(N)})$ becomes more deterministic as N gets large. Moreover, we anticipate that the limiting deterministic trajectory satisfies $x_t' = F(x_t)$, where

$$F(x) = \textstyle\sum_{l \neq 0} l f_l\,(x) \qquad (x \in E).$$

To simplify the statement of results, I'm going to assume that the state space is finite.

A law of large numbers

The following *functional law of large numbers* establishes convergence of the family $(X_t^{(N)})$ to the unique trajectory of the appropriate approximating deterministic model.

Theorem (Kurtz*) Suppose F is Lipschitz on E (that is, $|F(x) - F(y)| < M_E|x - y|$). If $\lim_{N\to\infty} X_0^{(N)} = x_0$, then the density process $(X_t^{(N)})$ converges uniformly in probability on $[0, t]$ to (x_t), the unique (deterministic) trajectory satisfying

$$\frac{d}{ds}x_s = F(x_s) \quad (x_s \in E,\ s \in [0, t]).$$

*Kurtz, T. (1970) Solutions of ordinary differential equations as limits of pure jump Markov processes, *J. of Appl. Probab.* 7, 49–58.

A law of large numbers

Convergence *uniformly in probability* on $[0, t]$ means that for every $\epsilon > 0$,

$$\lim_{N\to\infty} \Pr\left(\sup_{s\leq t} \left|X_t^{(N)} - x_t\right| > \epsilon\right) = 0.$$

A law of large numbers

Convergence *uniformly in probability* on $[0, t]$ means that for every $\epsilon > 0$,

$$\lim_{N\to\infty} \Pr\left(\sup_{s\le t}\left|X_t^{(N)} - x_t\right| > \epsilon\right) = 0.$$

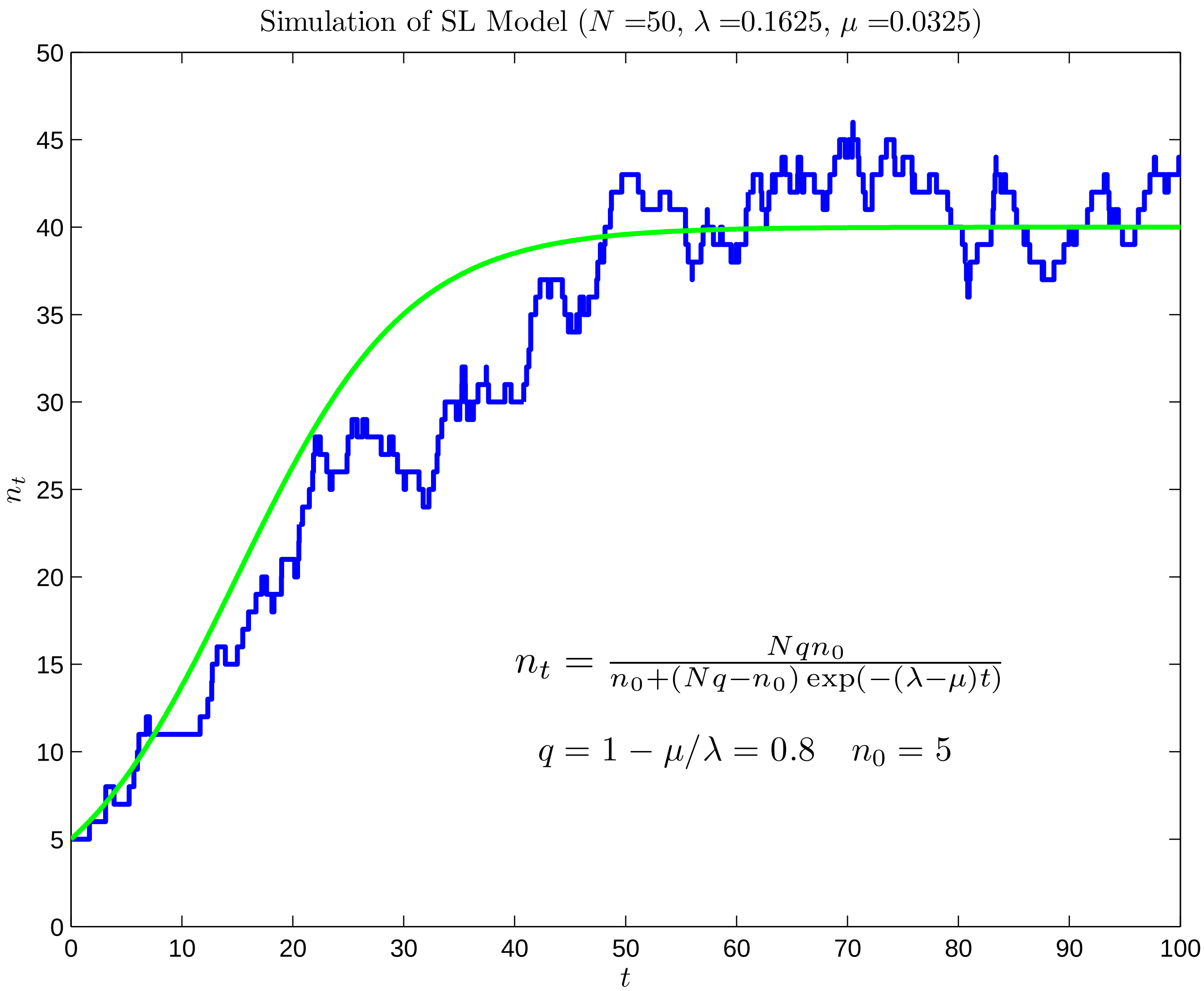
The conditions of the theorem hold for the SL model: since $F(x) = \lambda x(q - x)$, we have, for all $x, y \in E = [0, 1]$, that

$$|F(x) - F(y)| = \lambda|x - y||q - (x + y)| \le (2 - q)\lambda|x - y|.$$

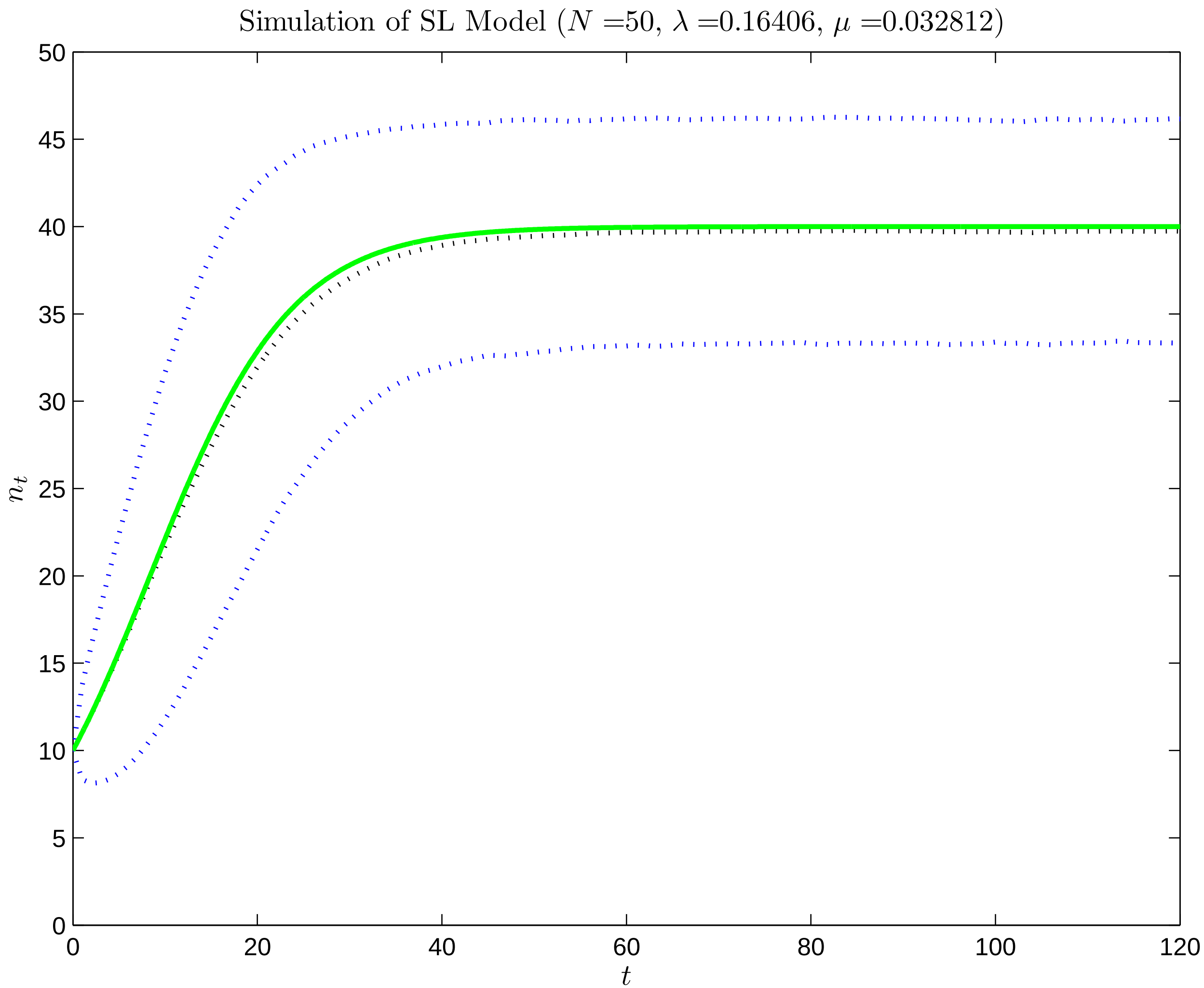
So, provided $X_0^{(N)} \to x_0$ as $N \to \infty$, the proportion $(X_t^{(N)})$ of occupied patches converges (uniformly in probability *on finite time intervals*) to deterministic trajectories in E:

$$x_t = \frac{q\,x_0}{x_0 + (q - x_0)\,e^{-(\lambda-\mu)t}} \qquad (x_0 \in E,\ t \ge 0).$$

The SL model $(N = 50)$

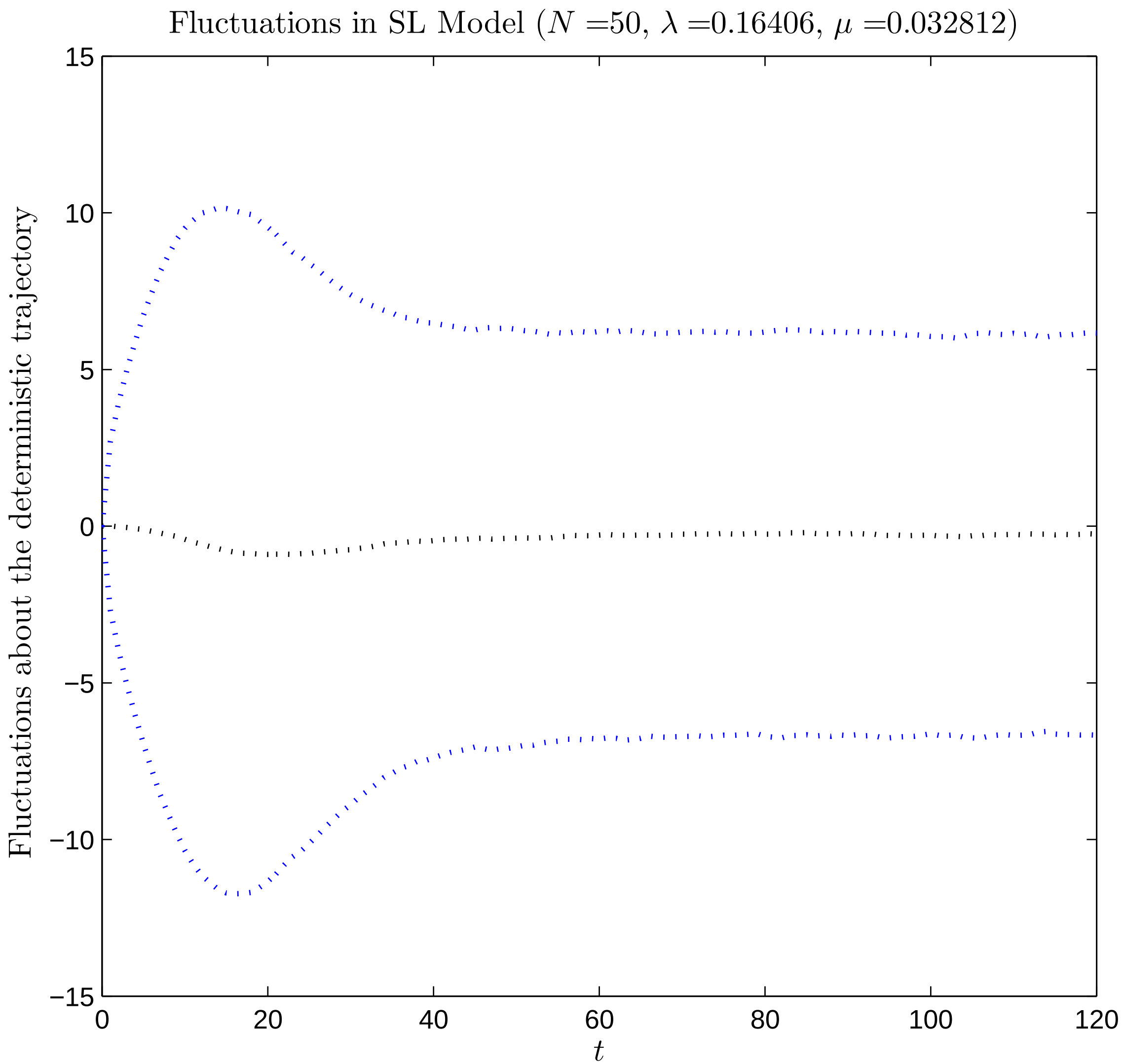


Variation in SL model



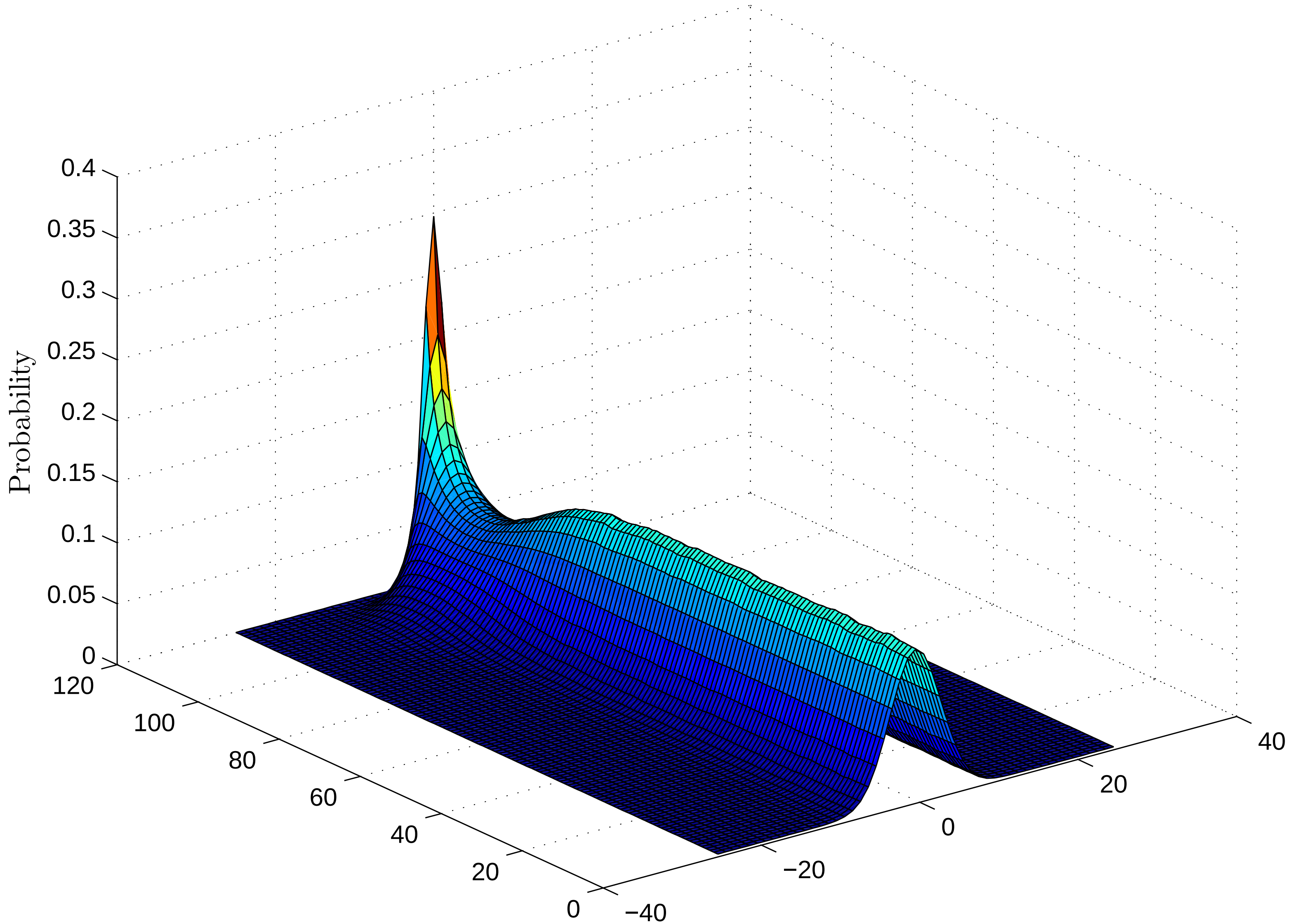
Mean path plus or minus two standard deviations

Variation in SL model

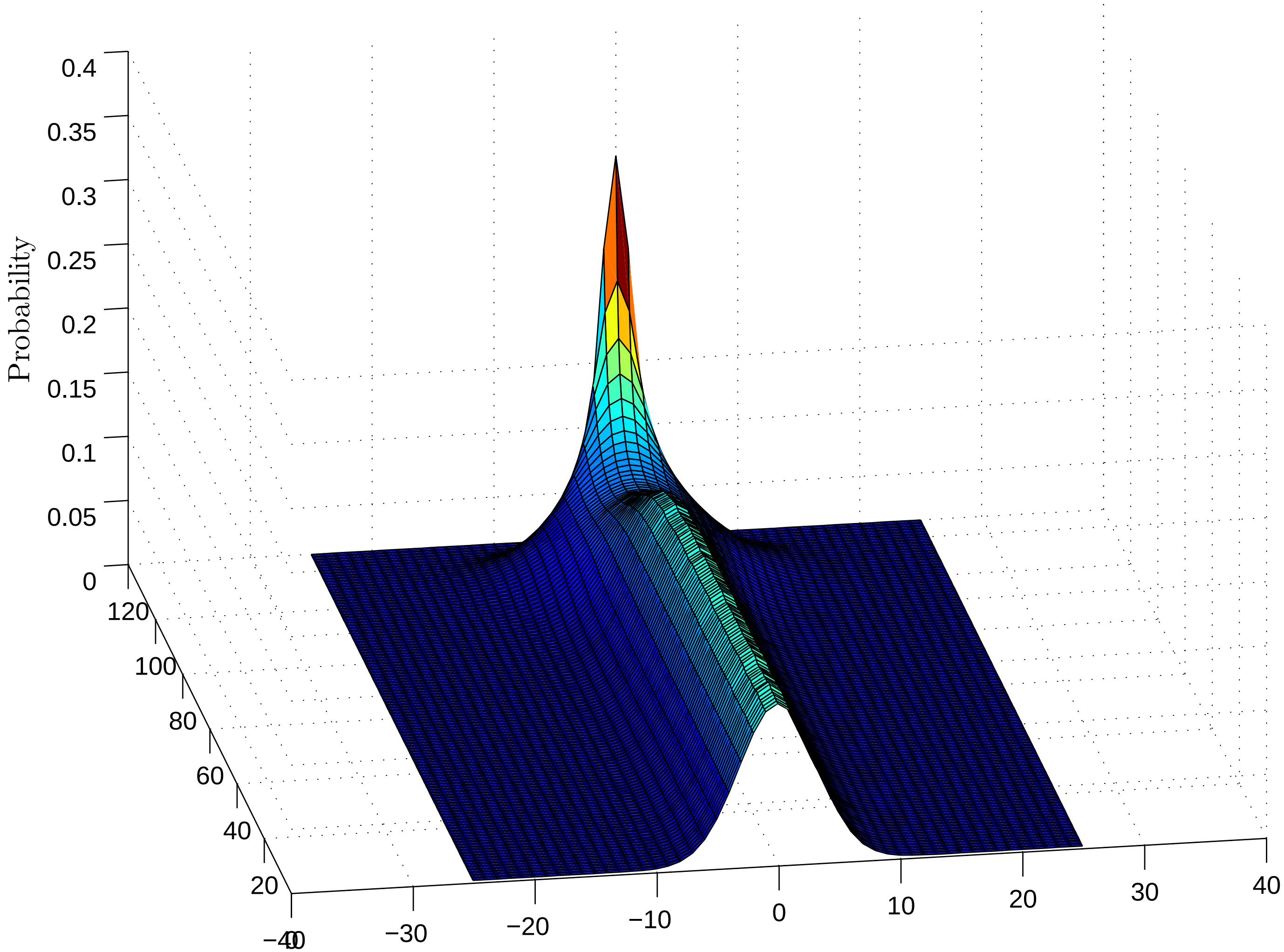


Deterministic trajectory subtracted

Variation in SL model



Variation in SL model



Modelling variation

We will consider the family of processes $\{(Z_t^{(N)})\}$, indexed by N, and defined by

$$Z_t^{(N)} = \sqrt{N}\left(X_t^{(N)} - x_t\right) \qquad (t \geq 0),$$

where recall that $(X_t^{(N)})$ is the *density process*, defined by $X_t^{(N)} = n_t/N$, and (x_t) is the limiting deterministic trajectory, which satisfies $x_t' = F(x_t)$, where

$$F(x) = \sum_{l \neq 0} l f_l(x) \qquad (x \in E).$$

I will call $\{(Z_t^{(N)})\}$ the *scaled density process*.

Modelling variation

We will consider the family of processes $\{(Z_t^{(N)})\}$, indexed by N, and defined by

$$Z_t^{(N)} = \sqrt{N}\left(X_t^{(N)} - x_t\right) \qquad (t \geq 0),$$

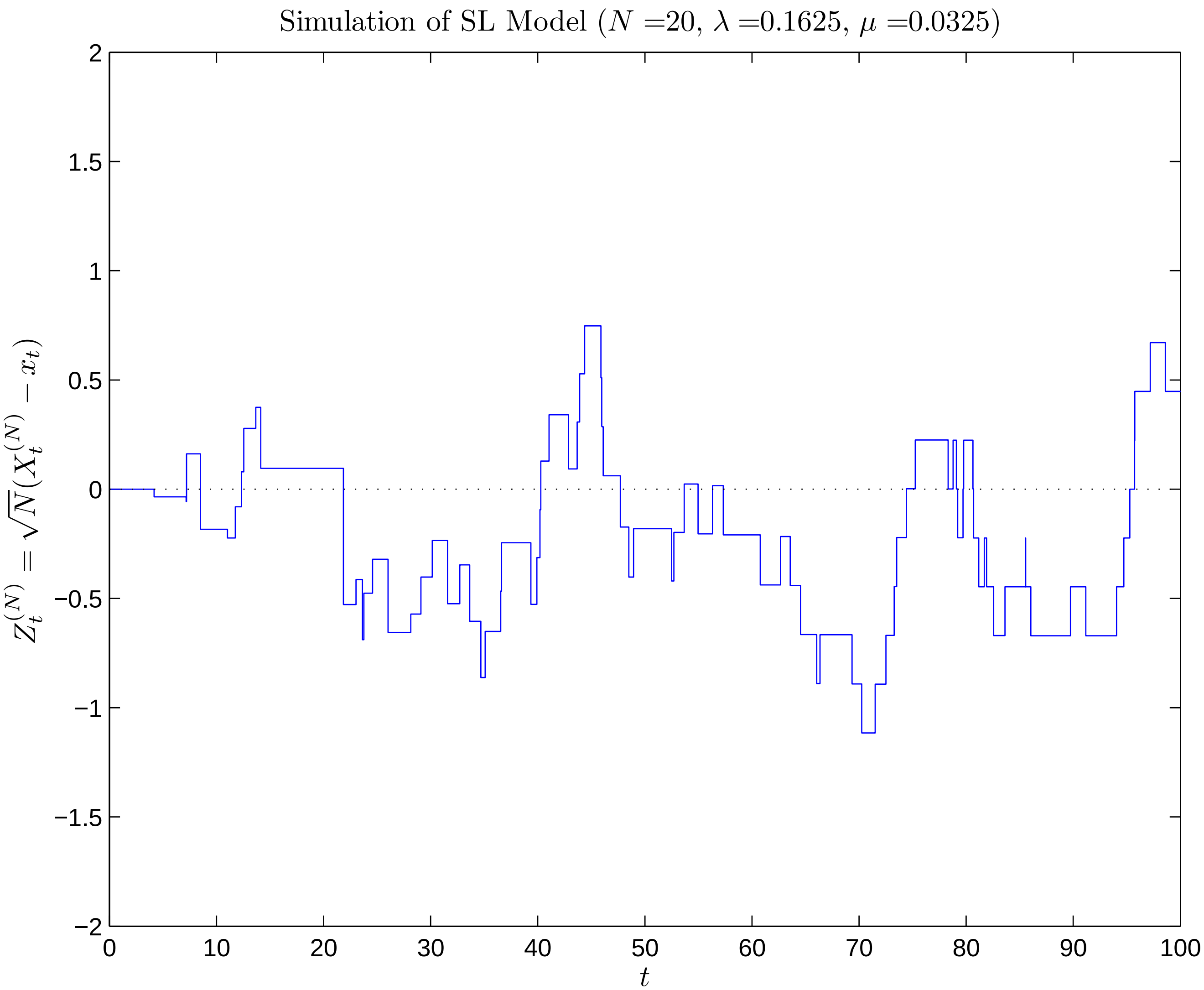
where recall that $(X_t^{(N)})$ is the *density process*, defined by $X_t^{(N)} = n_t/N$, and (x_t) is the limiting deterministic trajectory, which satisfies $x_t' = F(x_t)$, where

$$F(x) = \textstyle\sum_{l \neq 0} l f_l(x) \qquad (x \in E).$$

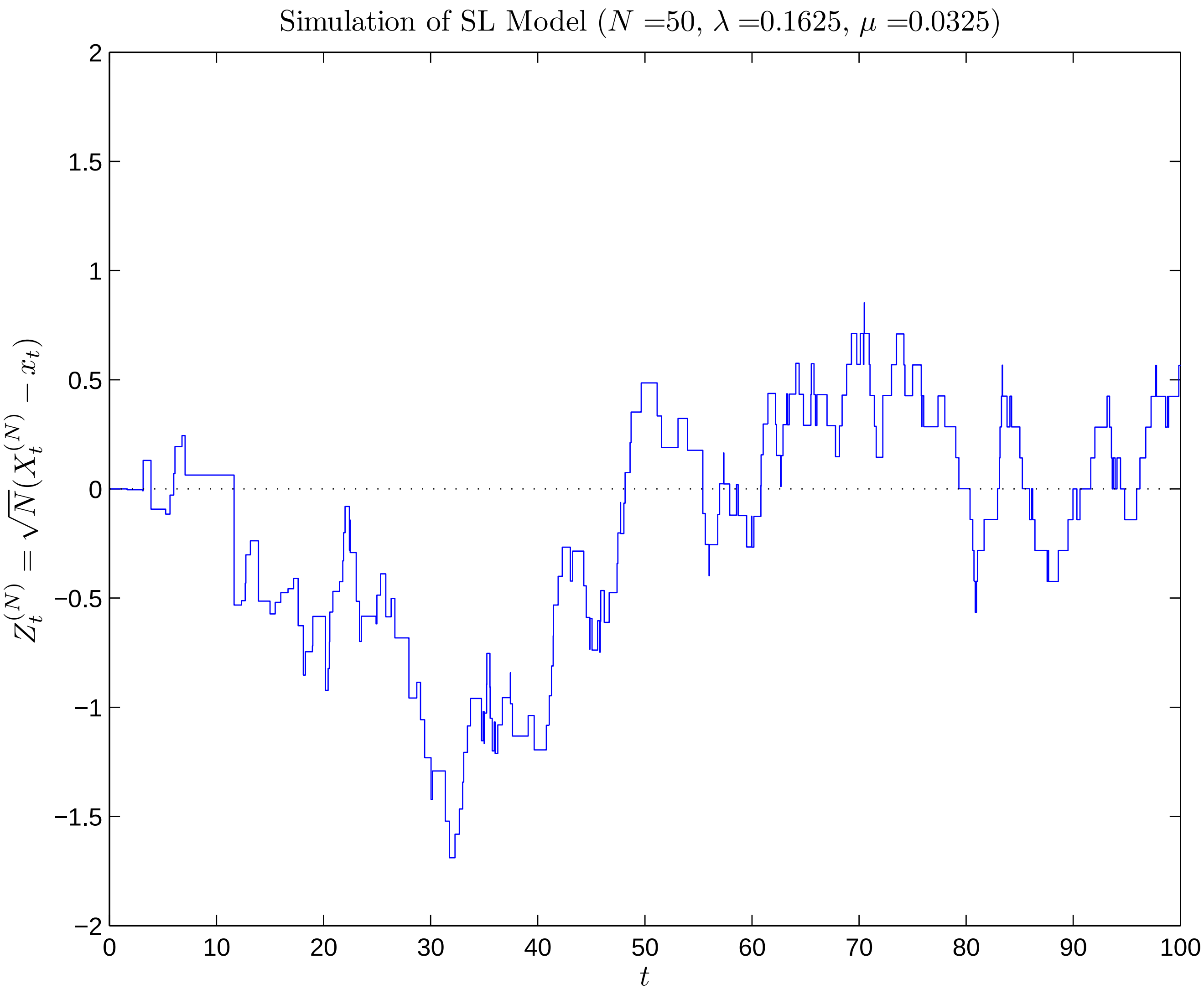
I will call $\{(Z_t^{(N)})\}$ the *scaled density process*.

In view of the *Central Limit Theorem* we might expect $\{(Z_t^{(N)})\}$ to become more "Gaussian" as N gets large; in particular, for each fixed t, $Z_t^{(N)} \xrightarrow{D} \mathrm{Normal}\left(\mu_t, V_t\right)$ as $N \to \infty$.

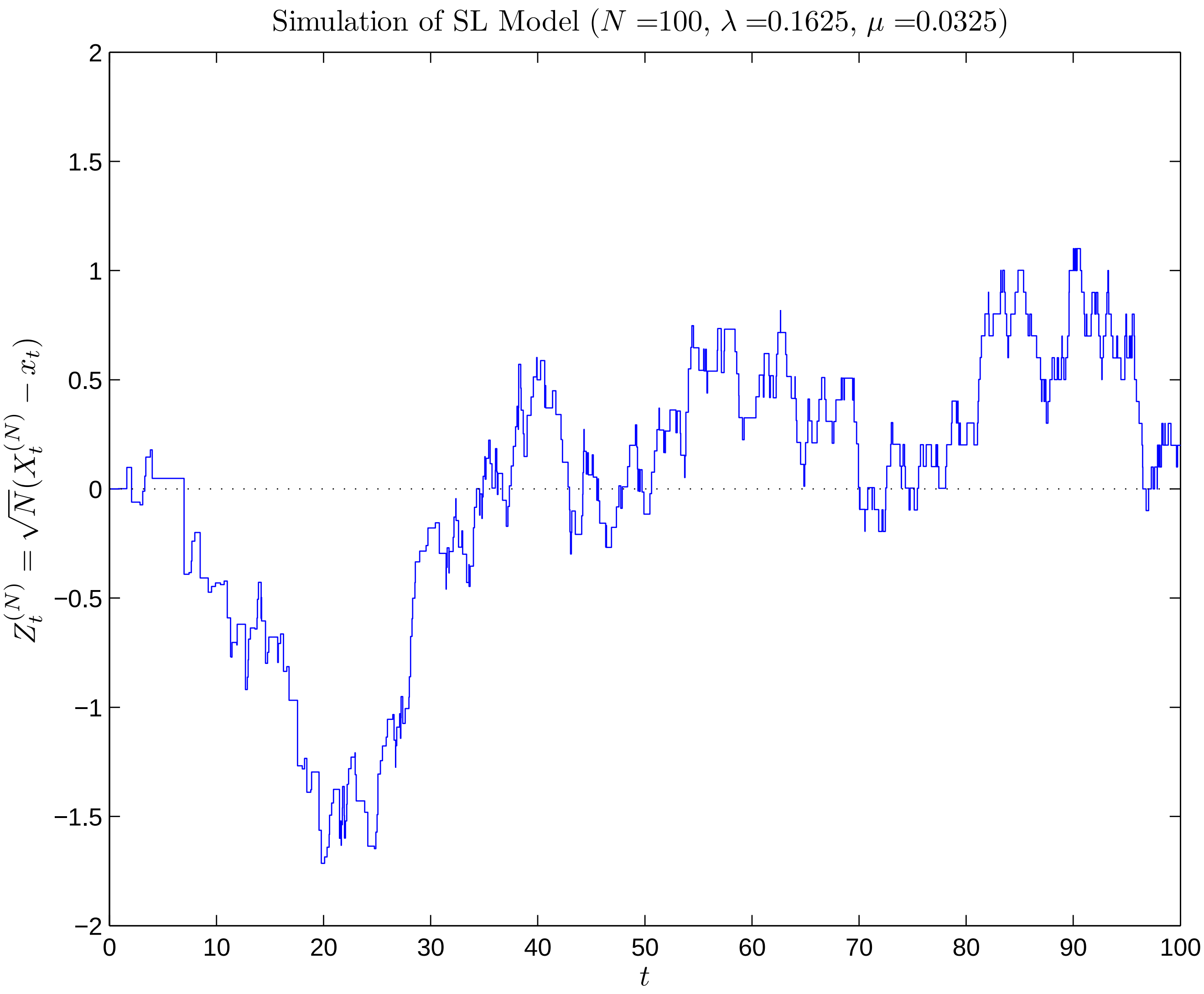
The SL model $(N = 20)$



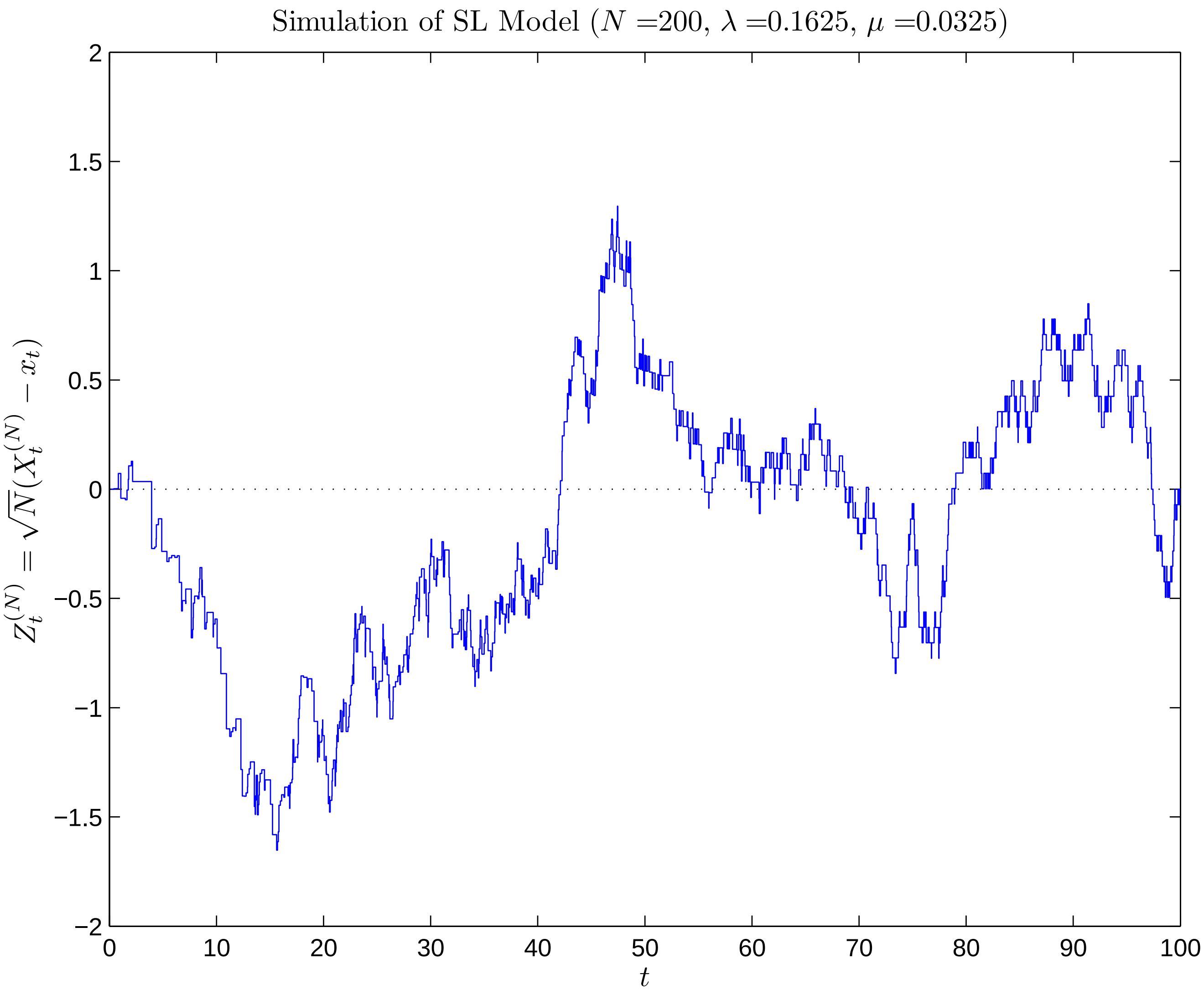
The SL model $(N = 50)$



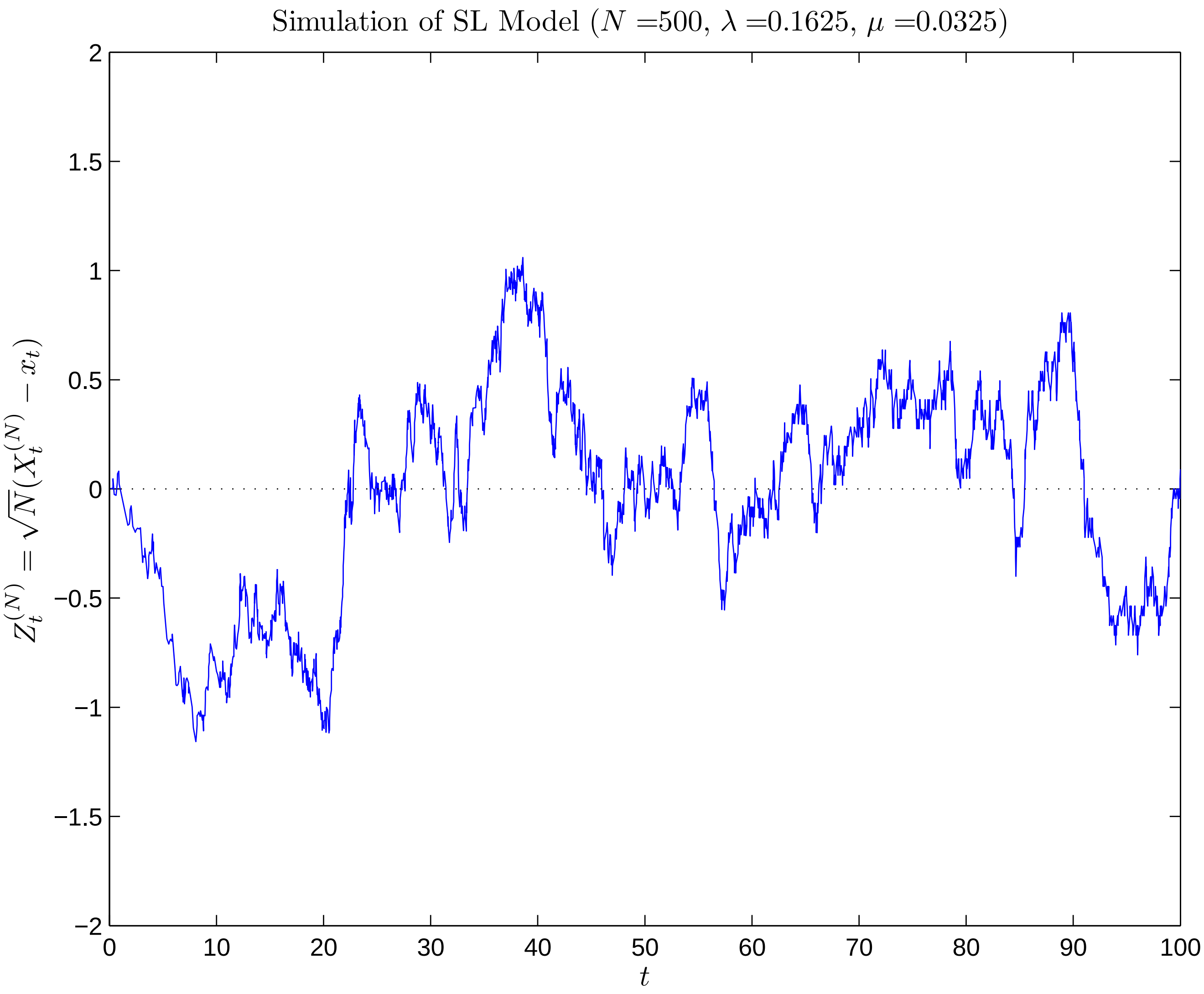
The SL model $(N = 100)$



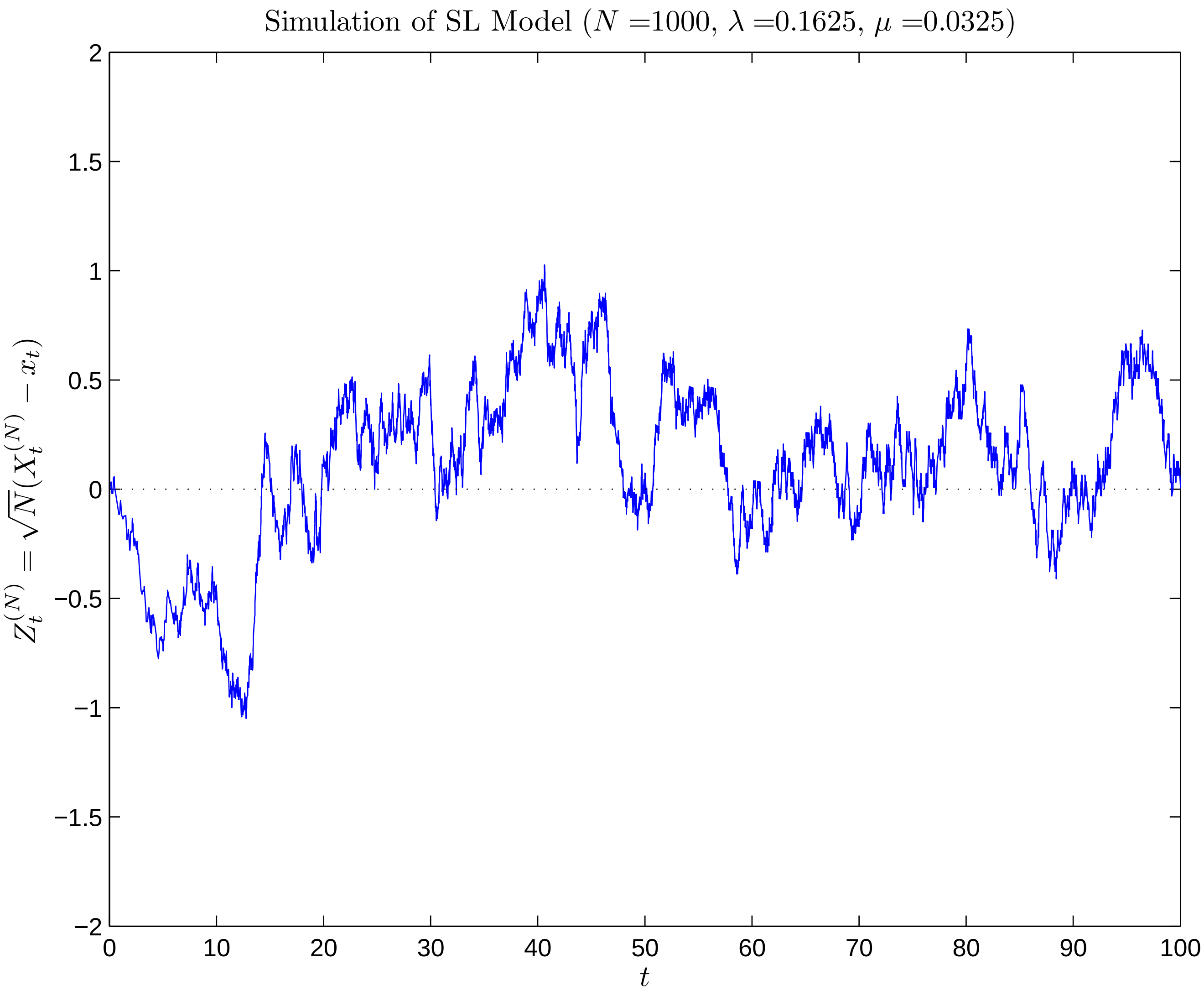
The SL model $(N = 200)$



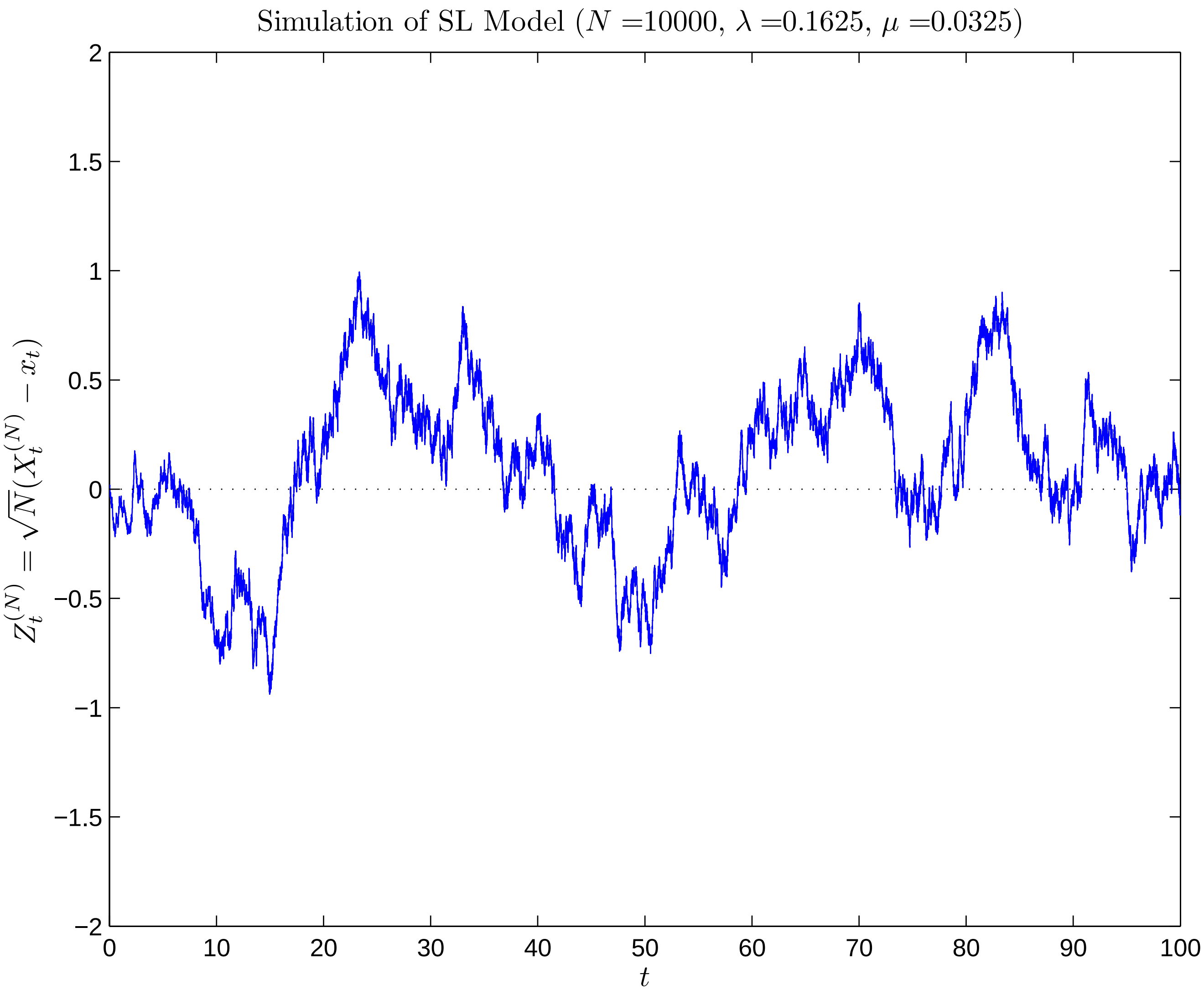
The SL model $(N = 500)$



The SL model $(N = 1\,000)$



The SL model $(N = 10\,000)$



Kurtz's theorem

In a later paper Kurtz* proved a *functional central limit law* which establishes that, for large N, the fluctuations about the deterministic trajectory do indeed follow a *Gaussian diffusion*, provided that some mild extra conditions are satisfied.

*Kurtz, T. (1971) Limit theorems for sequences of jump Markov processes approximating ordinary differential processes. *J. Appl. Probab.* 8, 344–356.

A central limit law

Theorem (Kurtz) Suppose that F is Lipschitz and has uniformly continuous first derivative on E, and that the $k \times k$ matrix $G(x)$ defined by $G_{ij}(x) = \sum_{l \neq 0} l_i l_j f_l(x)$, for each $x \in E$, is uniformly continuous on E.

Let (x_t) be the unique deterministic trajectory starting at x_0 and suppose that $\lim_{N\to\infty} \sqrt{N}\left(X_0^{(N)} - x_0\right) = z$.

Then, $\{(Z_t^{(N)})\}$ converges weakly in $D[0,t]$ (the space of right-continuous, left-hand limits functions on $[0,t]$) to a Gaussian diffusion (Z_t) with initial value $Z_0 = z$ and with mean and covariance given by $\mu_s := \mathbb{E}(Z_s) = M_s z$, where $M_s = \exp(\int_0^s B_u \, du)$ and $B_s = \nabla F(x_s)$, and

$$V_s := \mathrm{Cov}(Z_s) = M_s \left(\int_0^s M_u^{-1} G(x_u) (M_u^{-1})^T \, du \right) M_s^T \,.$$

A central limit law

Theorem (Kurtz) Suppose that F is Lipschitz and has uniformly continuous first derivative on E, and that the $k \times k$ matrix $G(x)$ defined by $G_{ij}(x) = \sum_{l \neq 0} l_i l_j f_l(x)$, for each $x \in E$, is uniformly continuous on E.

Let (x_t) be the unique deterministic trajectory starting at x_0 and suppose that $\lim_{N\to\infty} \sqrt{N}\left(X_0^{(N)} - x_0\right) = z$.

Then, $\{(Z_t^{(N)})\}$ converges weakly in $D[0,t]$ (the space of right-continuous, left-hand limits functions on $[0,t]$) to a Gaussian diffusion (Z_t) with initial value $Z_0 = z$ and with mean and covariance given by $\mu_s := \mathbb{E}(Z_s) = M_s z$, where $M_s = \exp(\int_0^s B_u\, du)$ and $B_s = \nabla F(x_s)$, and

$$V_s := \mathrm{Cov}(Z_s) = M_s \left(\int_0^s M_u^{-1} G(x_u) (M_u^{-1})^T\, du\right) M_s^T .$$

The SL model

For the SL model we have $F(x) = \lambda x(q - x)$, and the solution to $dx/dt = F(x)$ is

$$x(t) = \frac{qx_0}{x_0 + (q - x_0)e^{-(\lambda-\mu)t}}.$$

We also have $F'(x) = \lambda(q - 2x)$ and

$$G(x) = \sum_l l^2 f_l(x) = \lambda x(2 - q - x) = F(x) + 2\mu x,$$

giving

$$M_t = \exp\left(\int_0^t F'(x_s)\, ds\right) = \frac{q^2 e^{-(\lambda-\mu)t}}{(x_0 + (q - x_0)e^{-(\lambda-\mu)t})^2}.$$

We can evaluate

$$V_t := \mathrm{Var}(Z_t) = M_t^2 \left(\int_0^t G(x_s)/M_s^2\, ds\right)$$

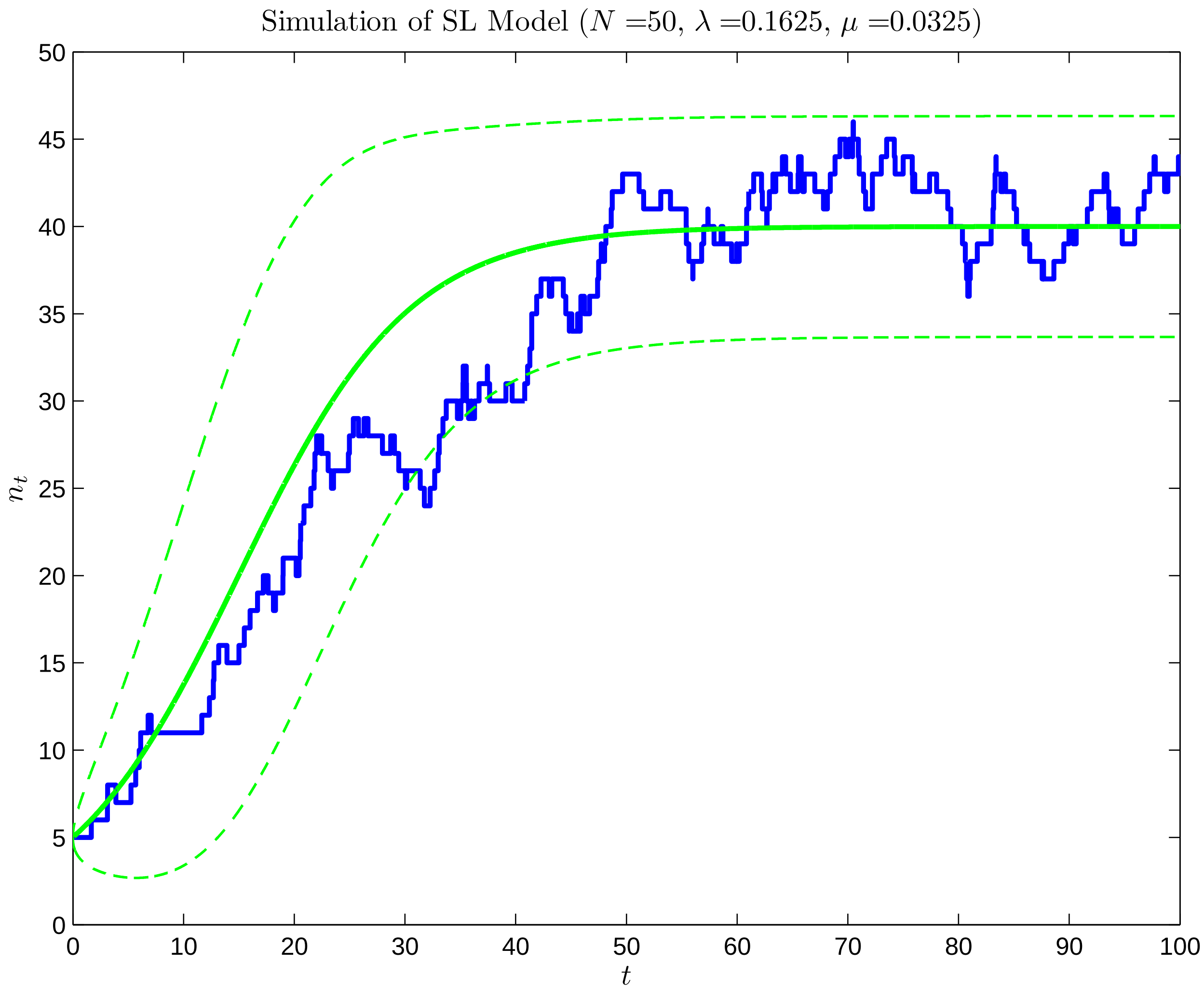
numerically, or ...

Or ....

$$
\begin{aligned}
V_t = x_0\Big((1+q)x_0^3 + x_0^2(6+5q)(q-x_0)e^{-\alpha t} \\
+ 2x_0(3+2q)(q-x_0)^2\alpha t e^{-2\alpha t} \\
- \big((q-x_0)[3(1+q)x_0^2 + (3+q)qx_0 - (3+2q)q^2] \\
+ (1+q)q^3\big)e^{-2\alpha t} \\
- (2+q)(q-x_0)^3 e^{-3\alpha t}\Big)\Big/\Big(x_0 + (q-x_0)e^{-\alpha t}\Big)^4,
\end{aligned}
$$

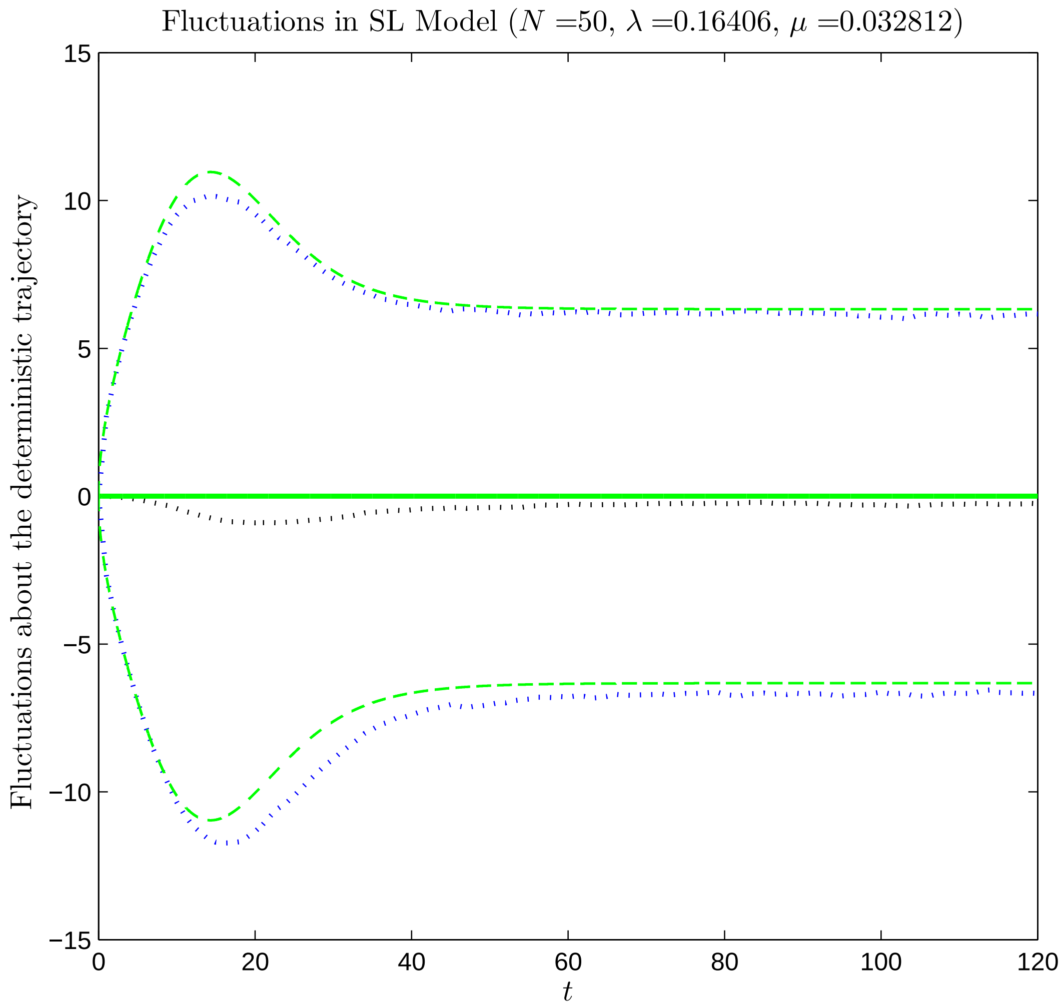
where $\alpha = \lambda - \mu$.

The SL model



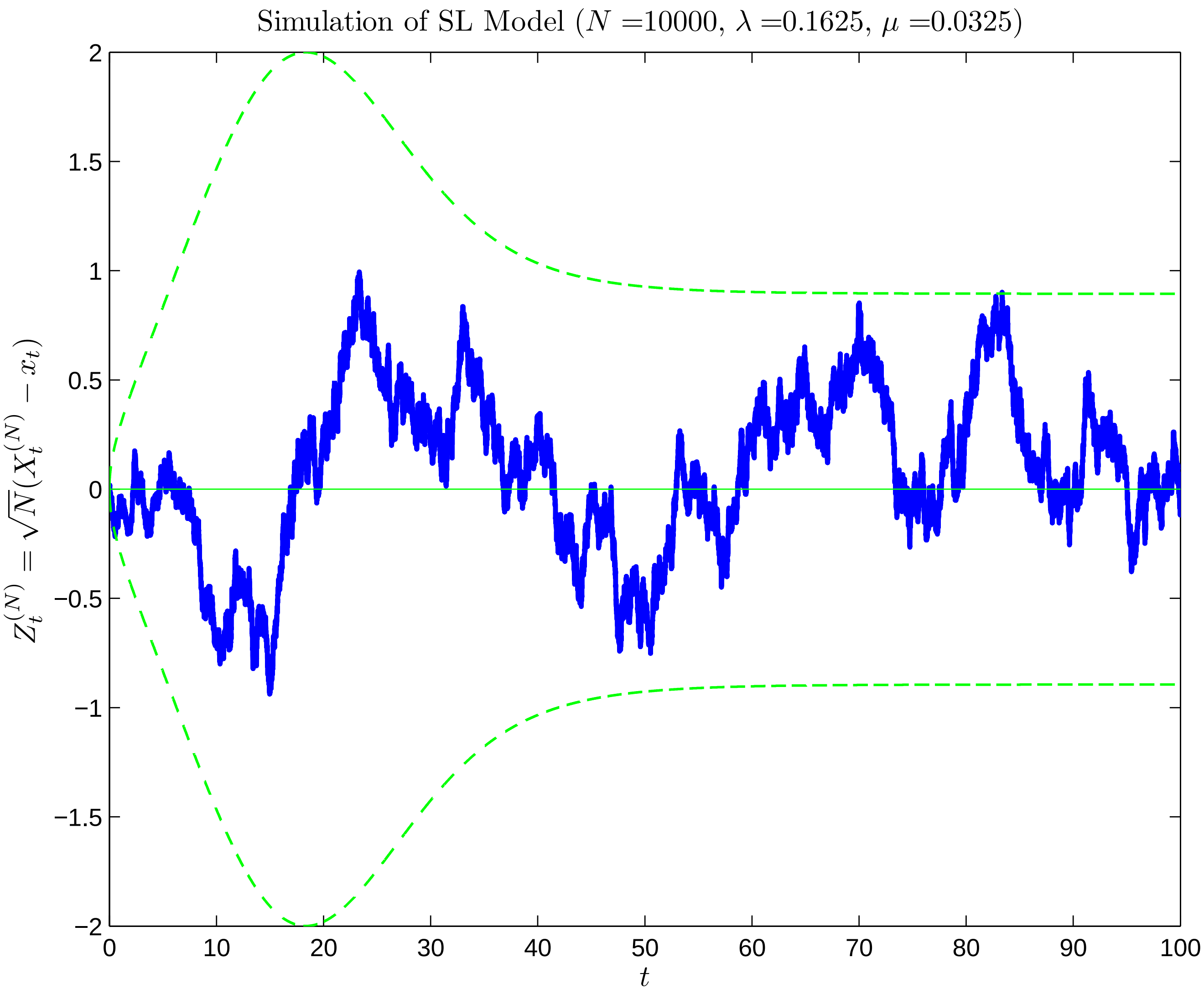
Deterministic trajectory plus or minus two standard deviations

The SL model

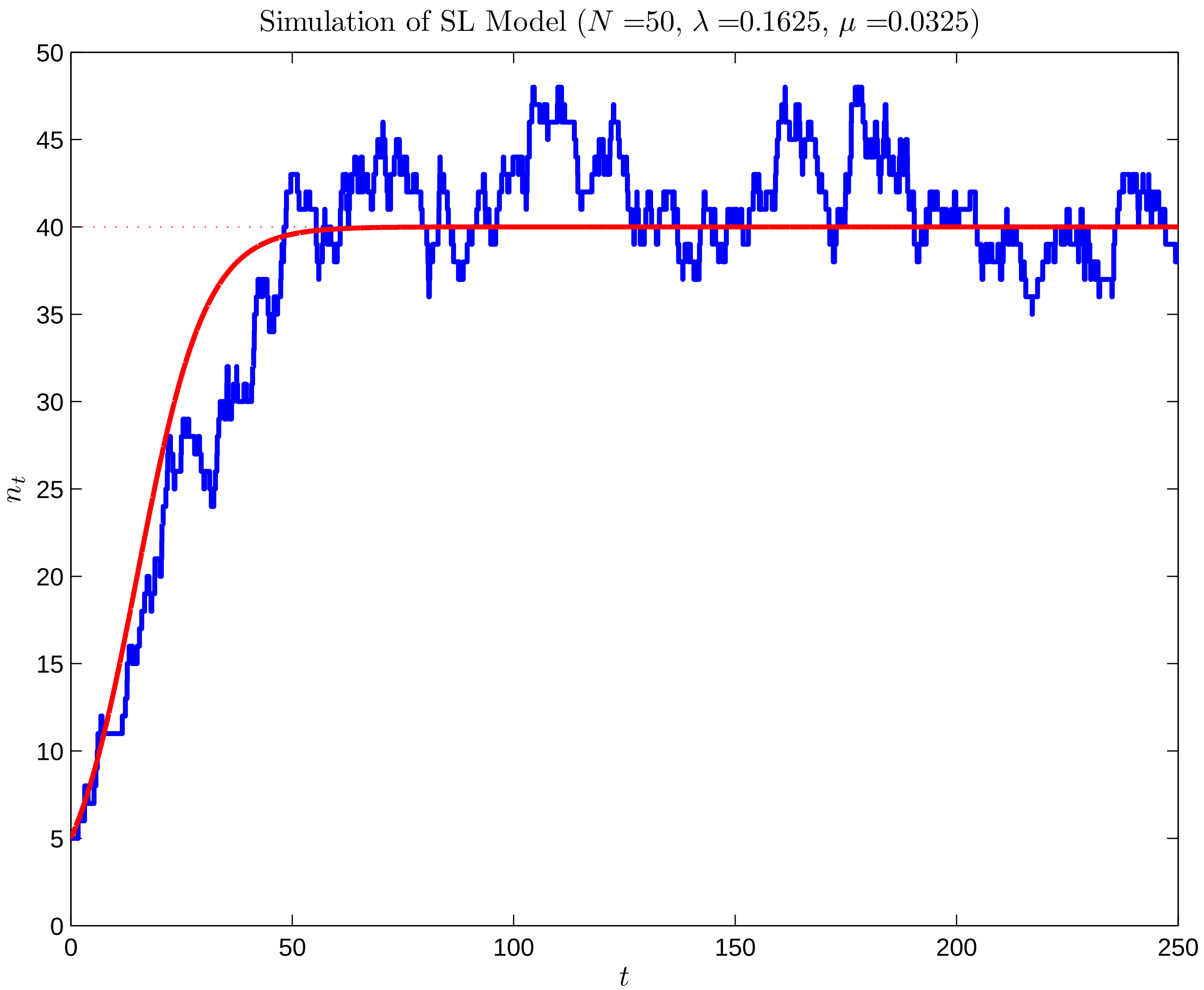


Deterministic trajectory plus or minus two standard deviations
(Empirical variance in blue and diffusion approximation in green)

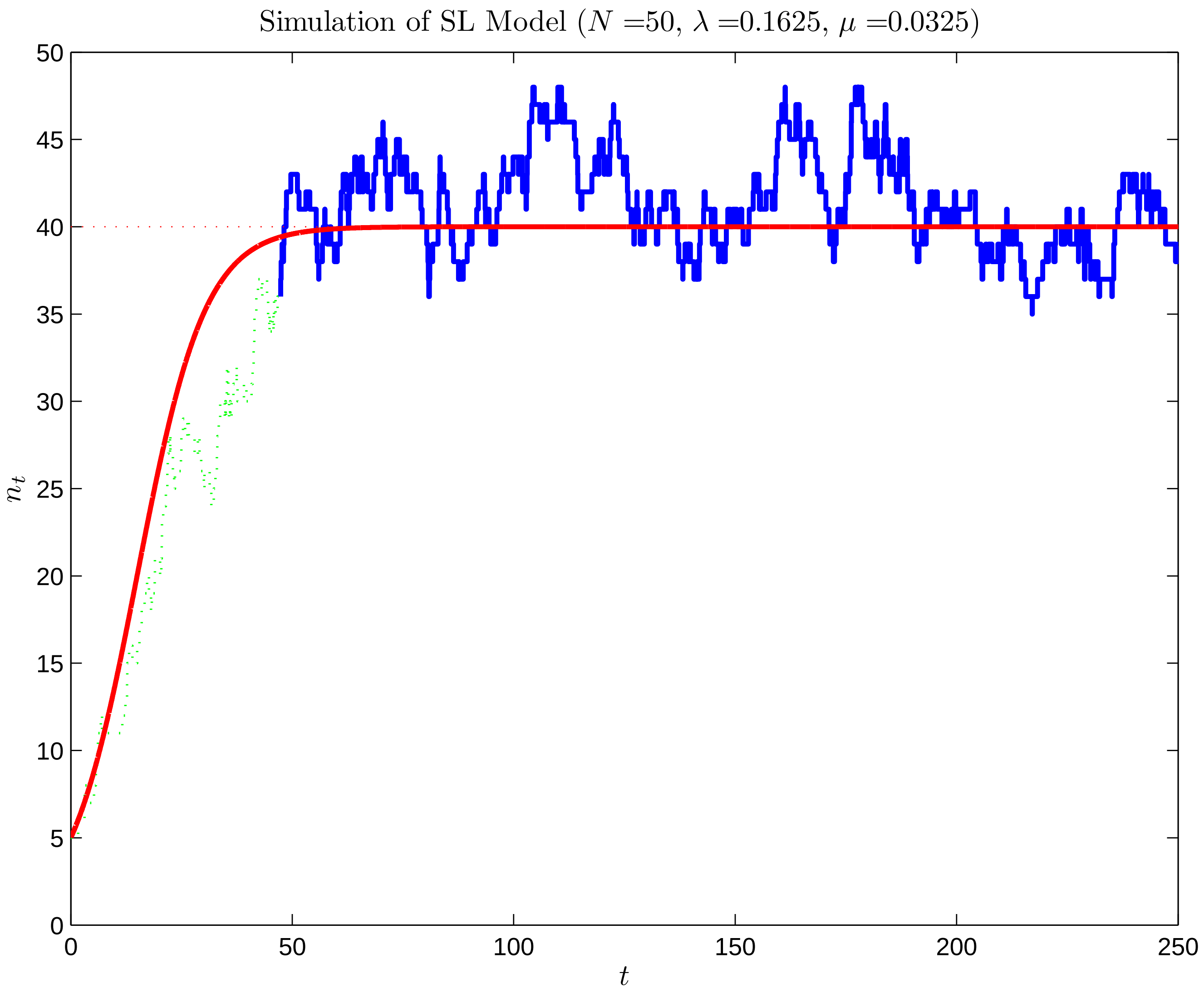
Scaled density process



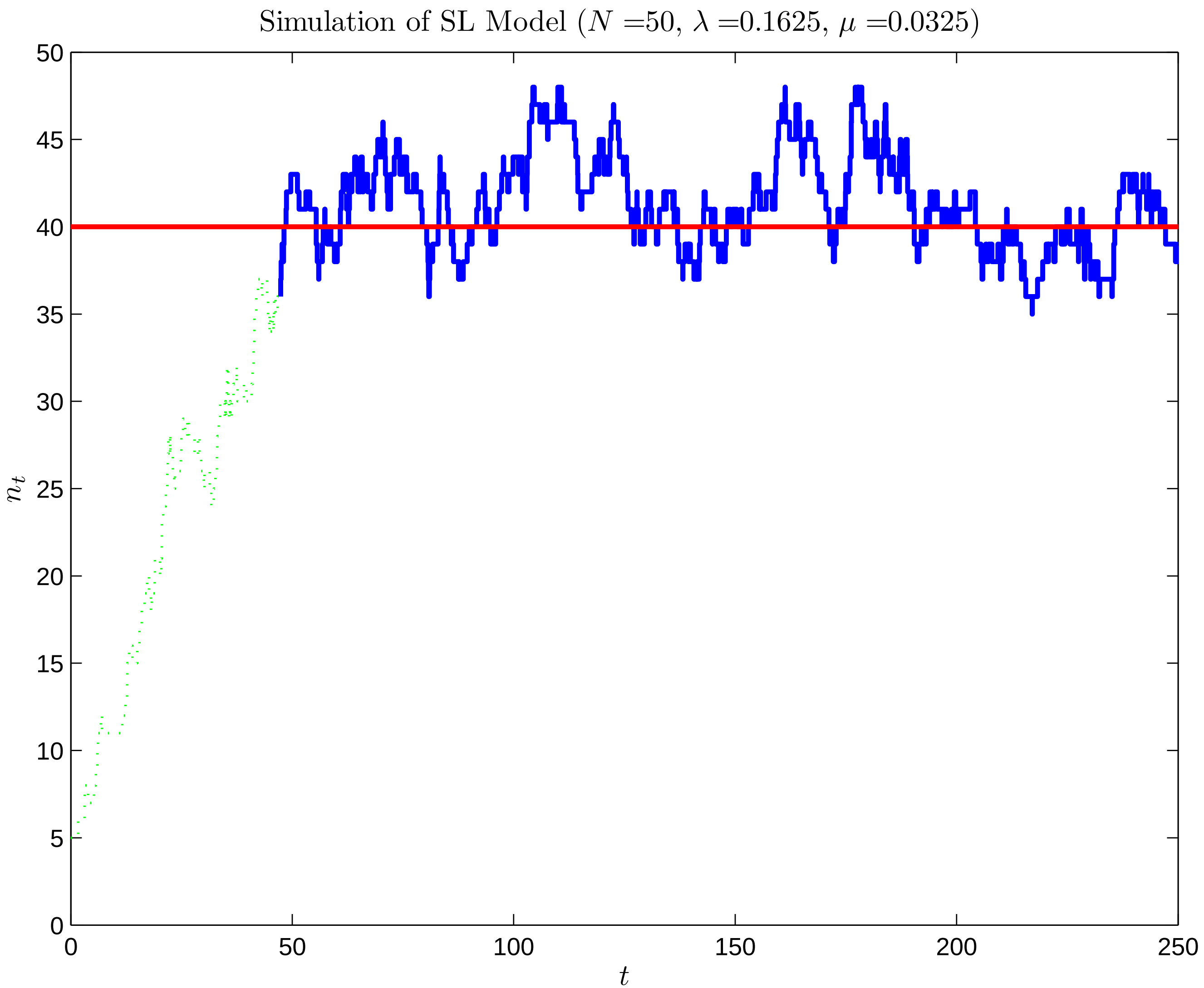
Equilibrium phase



Equilibrium phase



Equilibrium phase



Equilibrium

If we are only interested in the equilibrium phase of the process, then it is simpler to consider the family of processes $\{(Z_t^{(N)})\}$ defined by $Z_t^{(N)} = \sqrt{N}\left(X_t^{(N)} - x_{\mathrm{eq}}\right)$, where x_{eq} is an equilibrium point of the deterministic model. We can now be far more precise about the approximating diffusion.

Corollary If x_{eq} satisfies $F(x_{\mathrm{eq}}) = 0$, then, under the conditions of the theorem, $\{(Z_t^{(N)})\}$ converges weakly in $D[0,t]$ to an *Ornstein-Uhlenbeck (OU) process* (Z_t) with initial value $Z_0 = z$, local drift matrix $B := \nabla F(x_{\mathrm{eq}})$ and local covariance matrix $G(x_{\mathrm{eq}})$. In particular, Z_s is normally distributed with mean and covariance given by $\mu_s := \mathbb{E}(Z_s) = e^{Bs} z$ and

$$V_s := \mathrm{Cov}(Z_s) = \int_0^s e^{Bu} G(x_{\mathrm{eq}}) e^{B^T u}\, du\,.$$

Equilibrium

If we are only interested in the equilibrium phase of the process, then it is simpler to consider the family of processes $\{(Z_t^{(N)})\}$ defined by $Z_t^{(N)} = \sqrt{N}\left(X_t^{(N)} - x_{\mathrm{eq}}\right)$, where x_{eq} is an equilibrium point of the deterministic model. We can now be far more precise about the approximating diffusion.

Corollary If x_{eq} satisfies $F(x_{\mathrm{eq}}) = 0$, then, under the conditions of the theorem, $\{(Z_t^{(N)})\}$ converges weakly in $D[0,t]$ to an *Ornstein-Uhlenbeck (OU) process* (Z_t) with initial value $Z_0 = z$, local drift matrix $B := \nabla F(x_{\mathrm{eq}})$ and local covariance matrix $G(x_{\mathrm{eq}})$. In particular, Z_s is normally distributed with mean and covariance given by $\mu_s := \mathbb{E}(Z_s) = e^{Bs}z$ and

$$V_s := \mathrm{Cov}(Z_s) = \int_0^s e^{Bu} G(x_{\mathrm{eq}}) e^{B^T u}\, du\,.$$

The OU approximation

Note that

$$V_s = \int_0^s e^{Bu} G(x_{\mathrm{eq}}) e^{B^T u}\, du = V_{\mathsf{st}} - e^{Bs} V_{\mathsf{st}} e^{B^T s},$$

where V_{st}, the stationary covariance matrix, satisfies

$$B V_{\mathsf{st}} + V_{\mathsf{st}} B^T + G(x_{\mathrm{eq}}) = 0.$$

The OU approximation

Note that

$$V_s = \int_0^s e^{Bu} G(x_{\text{eq}}) e^{B^T u}\, du = V_{\text{st}} - e^{Bs} V_{\text{st}} e^{B^T s},$$

where V_{st}, the stationary covariance matrix, satisfies

$$BV_{\text{st}} + V_{\text{st}} B^T + G(x_{\text{eq}}) = 0.$$

We conclude that, for N large, $X_t^{(N)}$ has an approximate Gaussian distribution with $\text{Cov}(X_t^{(N)}) \simeq V_t/N$ (which for large t is approximately V_{st}/N).

The OU approximation

Note that

$$V_s = \int_0^s e^{Bu} G(x_{\mathrm{eq}}) e^{B^T u}\, du = V_{\mathsf{st}} - e^{Bs} V_{\mathsf{st}} e^{B^T s},$$

where V_{st}, the stationary covariance matrix, satisfies

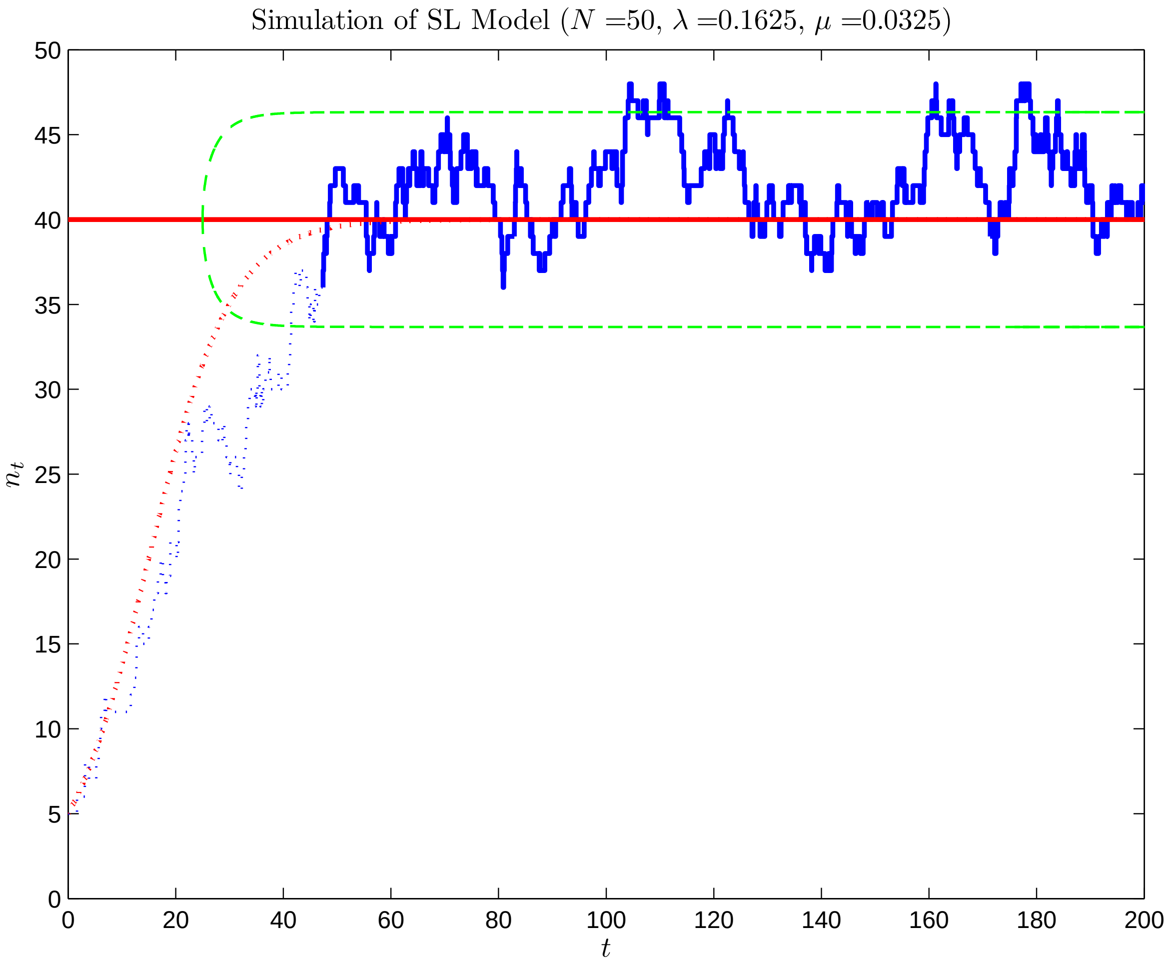
$$B V_{\mathsf{st}} + V_{\mathsf{st}} B^T + G(x_{\mathrm{eq}}) = 0.$$

We conclude that, for N large, $X_t^{(N)}$ has an approximate Gaussian distribution with $\mathrm{Cov}(X_t^{(N)}) \simeq V_t/N$ (which for large t is approximately V_{st}/N).

For the SL model,

$$\mathrm{Var}(X_t^{(N)}) \simeq \frac{1}{N}\left(\frac{\mu}{\lambda}\right)\left(1 - e^{-2(\lambda-\mu)t}\right) \quad \left(\simeq \frac{\mu}{N\lambda} \text{ for large } t\right).$$

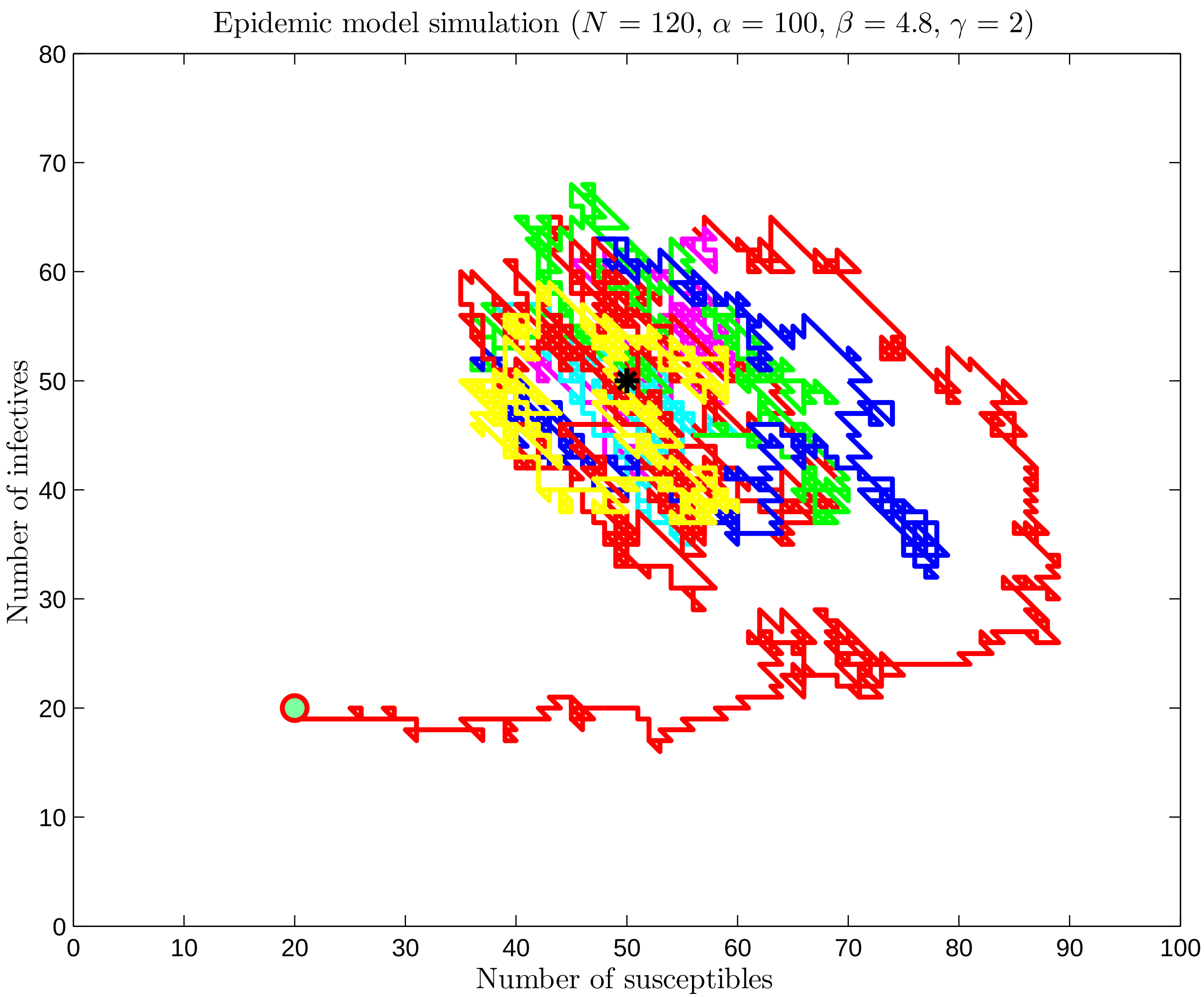
The SL model



Deterministic equilibrium plus or minus two standard deviations

(Deterministic trajectory in red and OU approximation in green)

An epidemic model



An epidemic model

The state at time t is (s_t, i_t), where s_t is the number of susceptibles and i_t is the number of infectives.

The state space is $S = \{(s,i) : s,i = 0,1,2,\ldots\}$.

The transitions are:

$(s,i) \to (s+1,i)$	at rate	α	($\rightarrow$ immigration)
$(s,i) \to (s,i-1)$	at rate	γi	($\downarrow$ death or removal)
$(s,i) \to (s-1,i+1)$	at rate	$\frac{\beta}{N} si$	($\nwarrow$ infection)
			(N is system size)

An epidemic model

The state at time t is (s_t, i_t), where s_t is the number of susceptibles and i_t is the number of infectives.

The state space is $S = \{(s,i) : s,i = 0,1,2,\dots\}$.

The transitions are:

$(s,i) \rightarrow (s+1,i)$	at rate	α	($\rightarrow$ immigration)
$(s,i) \rightarrow (s,i-1)$	at rate	γi	($\downarrow$ death or removal)
$(s,i) \rightarrow (s-1,i+1)$	at rate	$\frac{\beta}{N} si$	($\nwarrow$ infection)
			(N is system size)

Is the model density dependent?

An epidemic model

Is the Markov chain density dependent?

$$(s,i) \to (s+1,i) \quad \text{at rate} \quad N\left(\frac{\alpha}{N}\right)$$

$$(s,i) \to (s,i-1) \quad \text{at rate} \quad N\gamma\left(\frac{i}{N}\right)$$

$$(s,i) \to (s-1,i+1) \quad \text{at rate} \quad N\beta\left(\frac{s}{N}\right)\left(\frac{i}{N}\right)$$

An epidemic model

Is the Markov chain density dependent?

$(s,i) \longrightarrow (s+1,i)$	at rate	$N\left(\frac{\alpha}{N}\right)$
$(s,i) \longrightarrow (s,i-1)$	at rate	$N\gamma\left(\frac{i}{N}\right)$
$(s,i) \longrightarrow (s-1,i+1)$	at rate	$N\beta\left(\frac{s}{N}\right)\left(\frac{i}{N}\right)$

An epidemic model

Is the Markov chain density dependent?

$$(s,i) \to (s+1,i) \quad \text{at rate} \quad N\left(\frac{\alpha}{N}\right)$$

$$(s,i) \to (s,i-1) \quad \text{at rate} \quad N\gamma\left(\frac{i}{N}\right)$$

$$(s,i) \to (s-1,i+1) \quad \text{at rate} \quad N\beta\left(\frac{s}{N}\right)\left(\frac{i}{N}\right)$$

The α/N term is a *problem*. Since α is a constant, the immigration term will vanish when N becomes large.

An epidemic model

Is the Markov chain density dependent?

$$(s,i) \to (s+1,i) \qquad \text{at rate} \qquad N\left(\frac{\alpha}{N}\right)$$

$$(s,i) \to (s,i-1) \qquad \text{at rate} \qquad N\gamma\left(\frac{i}{N}\right)$$

$$(s,i) \to (s-1,i+1) \qquad \text{at rate} \qquad N\beta\left(\frac{s}{N}\right)\left(\frac{i}{N}\right)$$

The α/N term is a *problem*. Since α is a constant, the immigration term will vanish when N becomes large.

For density dependence we must have $\alpha = O(N)$ (say $\alpha \sim aN$). Is this reasonable?

An epidemic model

$$(s,i) \to (s,i) + (+1,0) \qquad \text{at rate} \qquad N\left(\frac{\alpha}{N}\right)$$

$$(s,i) \to (s,i) + (0,-1) \qquad \text{at rate} \qquad N\gamma\left(\frac{i}{N}\right)$$

$$(s,i) \to (s,i) + (-1,+1) \qquad \text{at rate} \qquad N\beta\left(\frac{s}{N}\right)\left(\frac{i}{N}\right)$$

An epidemic model

$$(s,i) \to (s,i) + (+1,0) \quad \text{at rate} \quad N\left(\frac{\alpha}{N}\right)$$

$$(s,i) \to (s,i) + (0,-1) \quad \text{at rate} \quad N\gamma\left(\frac{i}{N}\right)$$

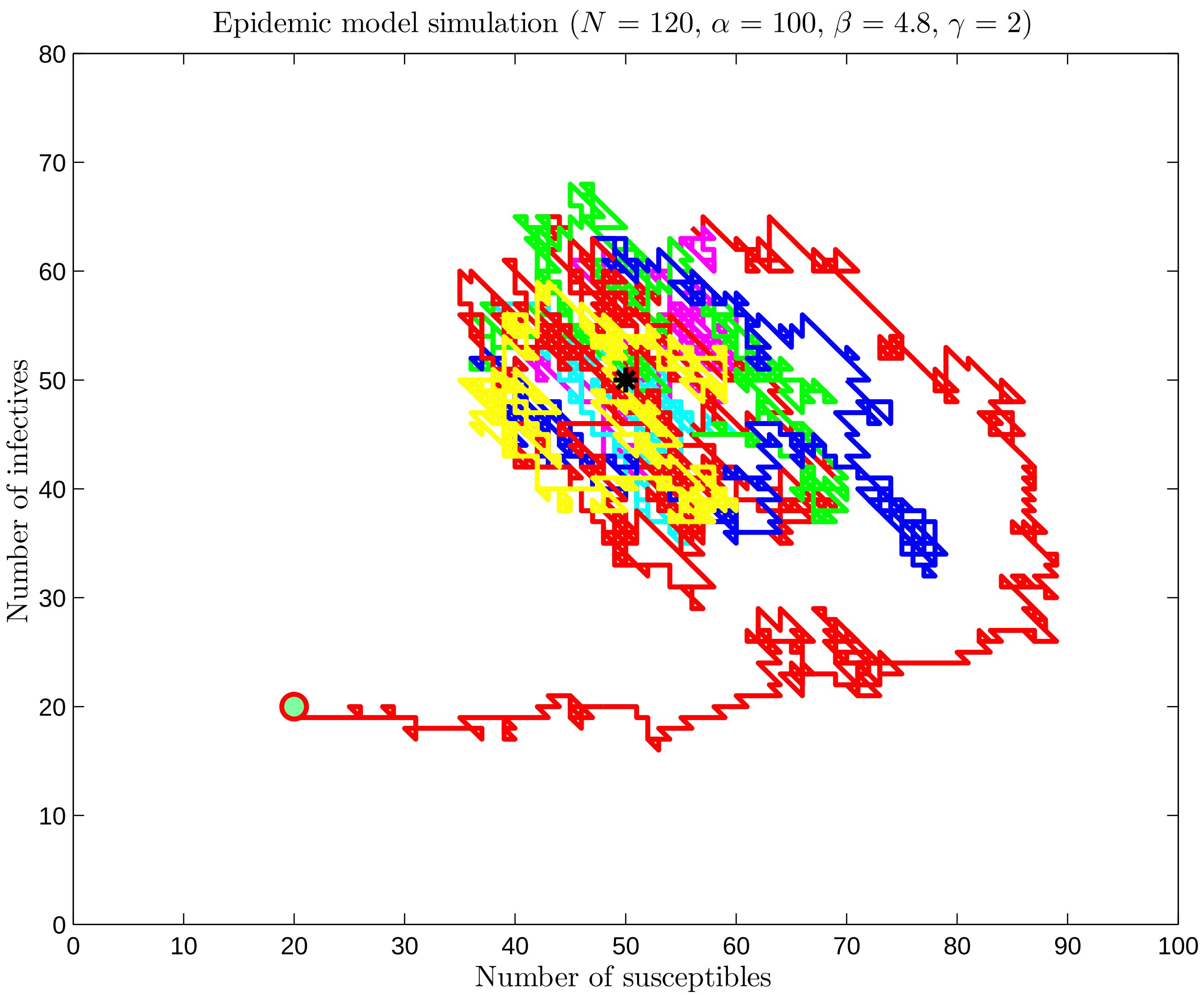
$$(s,i) \to (s,i) + (-1,+1) \quad \text{at rate} \quad N\beta\left(\frac{s}{N}\right)\left(\frac{i}{N}\right)$$

$$f_{(+1,0)}(\mathbf{x}) = a \quad f_{(0,-1)}(\mathbf{x}) = \gamma x_2 \quad f_{(-1,+1)}(\mathbf{x}) = \beta x_1 x_2$$

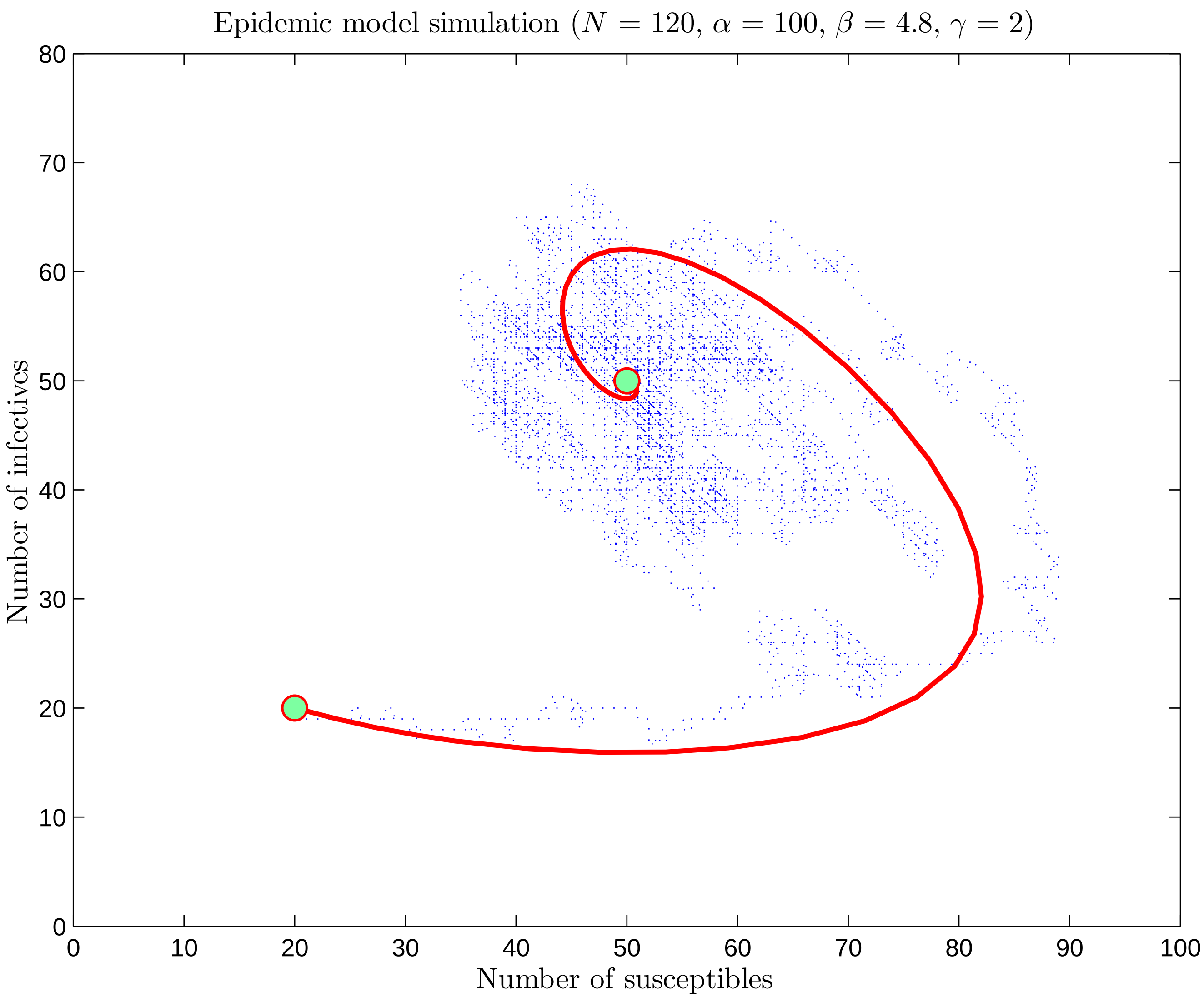
$$F(\mathbf{x}) = \sum_{l \neq 0} l f_l(\mathbf{x}) = \begin{pmatrix} a - \beta x_1 x_2 \\ -\gamma x_2 + \beta x_1 x_2 \end{pmatrix} \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

(The deterministic model is $\mathbf{x}_t' = F(\mathbf{x})$)

An epidemic model



An epidemic model



An epidemic model

$F(\mathbf{x}_{\rm eq}) = 0$ gives $\mathbf{x}_{\rm eq} = (\gamma/\beta, a/\gamma)$. Also,

$$\nabla F(\mathbf{x}) = \begin{pmatrix} -\beta x_2 & -\beta x_1 \\ \beta x_2 & \beta x_1 - \gamma \end{pmatrix} \; B := \nabla F(\mathbf{x}_{\rm eq}) = \begin{pmatrix} -a\beta/\gamma & -\gamma \\ a\beta/\gamma & 0 \end{pmatrix}.$$

The eigenvalues of B are both negative if $4\gamma^2 \leq a\beta$, and complex if $4\gamma^2 > a\beta$.

$$G_{ij}(\mathbf{x}) = \textstyle\sum_{l \neq 0} l_i l_j f_l(\mathbf{x}).$$

So,

$$G(\mathbf{x}) = \begin{pmatrix} a + \beta x_1 x_2 & -\beta x_1 x_2 \\ -\beta x_1 x_2 & \gamma x_2 + \beta x_1 x_2 \end{pmatrix}.$$

An epidemic model

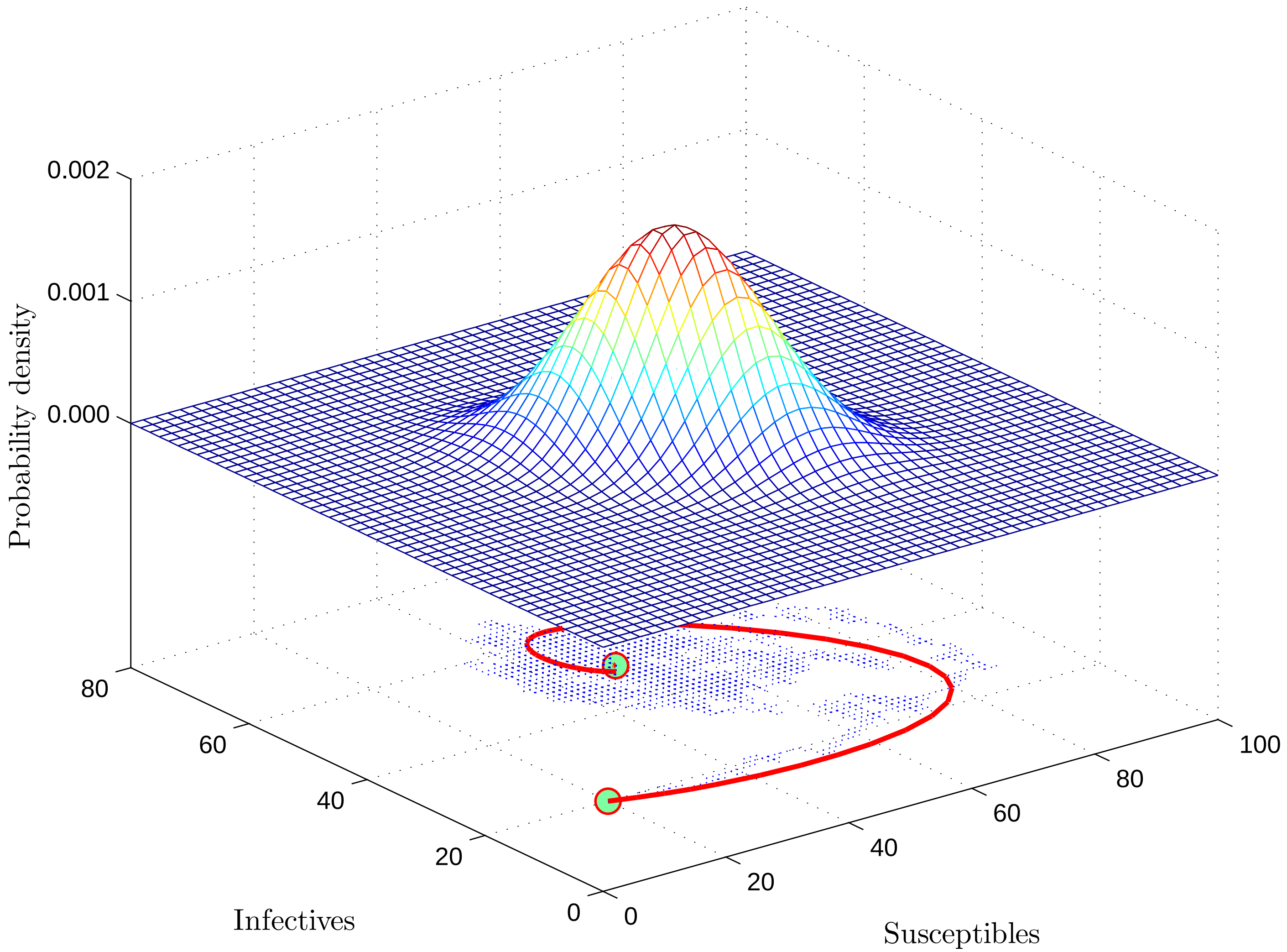
$$B = \begin{pmatrix} -a\beta/\gamma & -\gamma \\ a\beta/\gamma & 0 \end{pmatrix}$$

$$G(\mathbf{x}_{\mathrm{eq}}) = \begin{pmatrix} 2a & -a \\ -a & 2a \end{pmatrix}$$

$$V_t := \mathrm{Cov}(Z_t) = V_{\mathsf{st}} - e^{Bt} V_{\mathsf{st}} e^{B^T t}$$

$$V_{\mathsf{st}} = \begin{pmatrix} \frac{\gamma}{\beta}\left(1 + \frac{\gamma^2}{a\beta}\right) & -\frac{\gamma}{\beta} \\ -\frac{\gamma}{\beta} & \frac{\gamma}{\beta} + \frac{a}{\gamma} \end{pmatrix}$$

The OU approximation



Van Kampen's method

Van Kampen* considered the "Kramers-Moyal expansion" of the *master equation* (aka the forward equation) for the jump process (n_t). He transformed n_t by introducing a new variable Z_t so that $n_t = Nx_t + \sqrt{N}Z_t$.

He then derived the corresponding master equation for (Z_t), noting that if (x_t) obeys $x_t' = F(x_t)$, then terms of order $N^{1/2}$ cancel, and only a single term in the expansion survives in the limit as $N \to \infty$: arriving at the *Fokker-Planck* equation

$$\frac{\partial}{\partial t} P_z(t) = -\alpha(x_t) z \frac{\partial}{\partial z} P_z(t) + \frac{1}{2}\beta(x_t)\frac{\partial^2}{\partial^2 z} P_z(t),$$

where $\alpha(\cdot)$ and $\beta(\cdot)$ are determined for the particular model. So, the variable Z_t is indeed Gaussian.

*Van Kampen, N.G. (1961) A Power series expansion of the master equation. *Canadian J. Phys.* 39, 551–567.