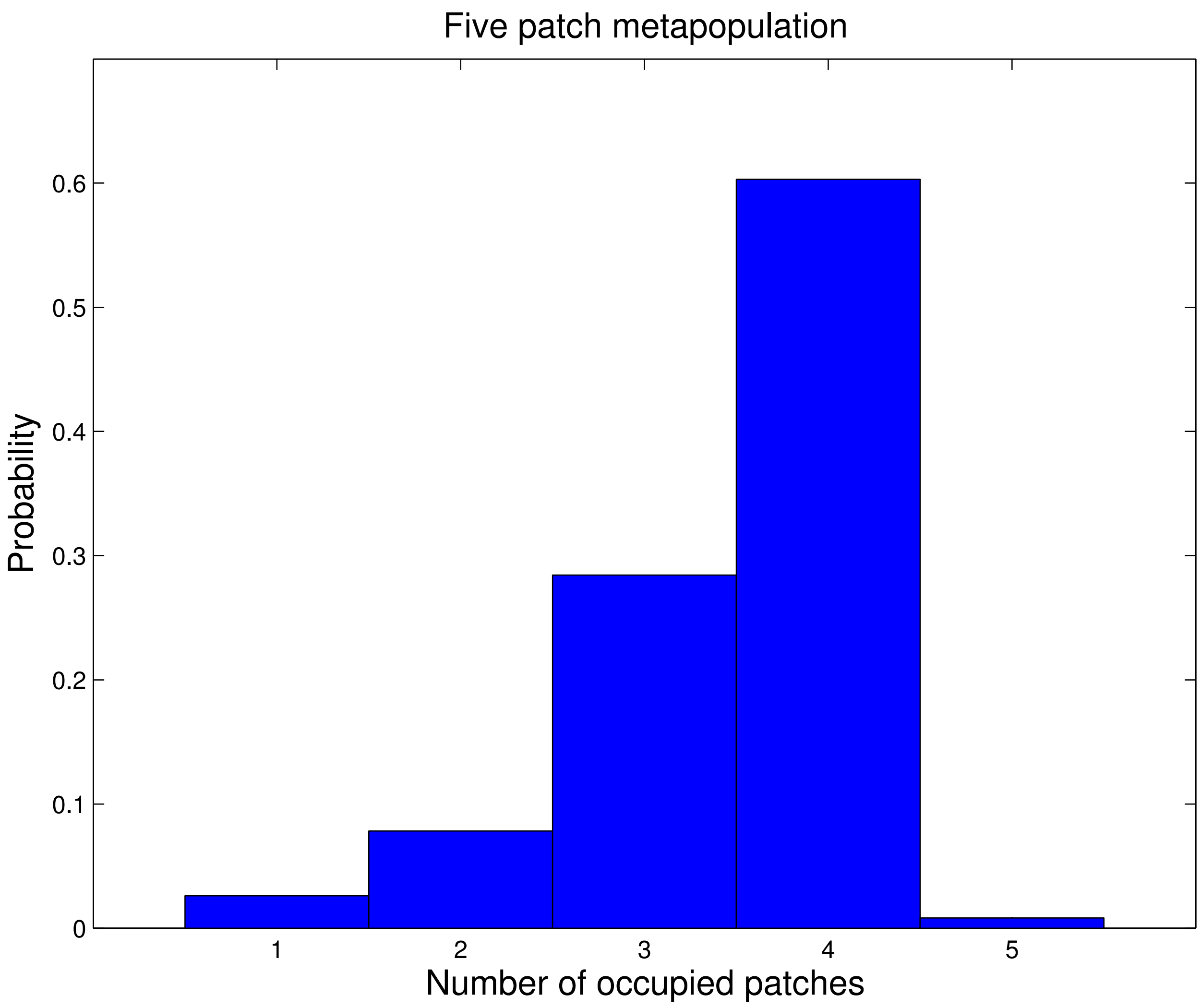
Limiting Conditional Distributions for Stochastic Metapopulation Models

Phil Pollett

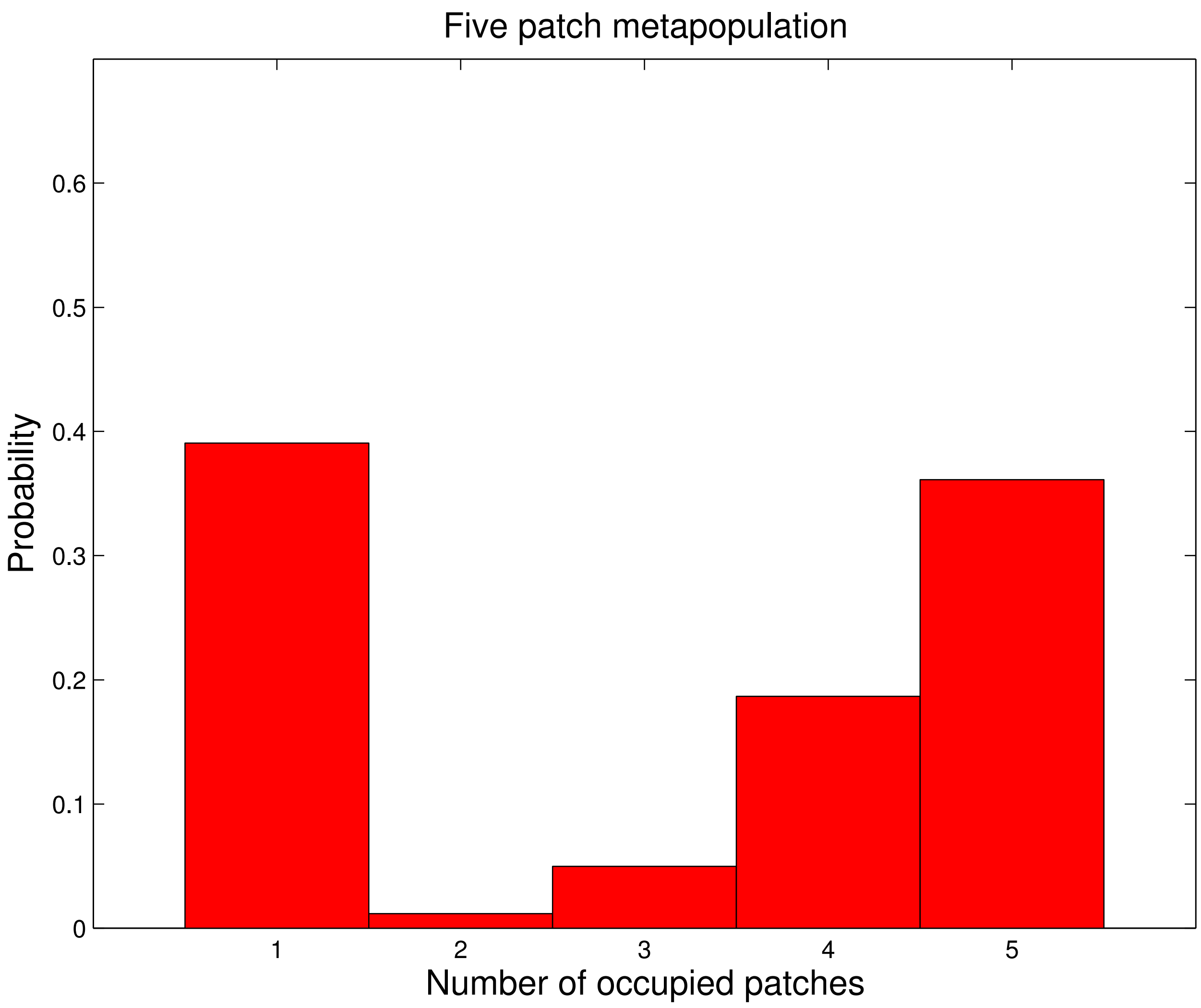
University of Queensland

ARC CENTRE OF EXCELLENCE FOR MATHEMATICS AND STATISTICS OF COMPLEX SYSTEMS

5-patch metapopulation



5-patch metapopulation



Conditional state distribution

$X(t)$ - state of the metapopulation at time t

We suppose that $(X(t),\, t \geq 0)$ is a discrete-time Markov chain with a discrete state space $\mathcal{S} = \{0\} \cup C$, where 0 is the state corresponding to extinction (of all patches) and C comprises the remaining states.

$p_x(t) = \Pr(X(t) = x)$ - state probabilities

Suppose these are given. We observe the population at an arbitrary time s and *extinction has not yet occurred*. How can we incorporate this information?

Conditional state distribution

We evaluate the state probabilities at time s conditioned on non-extinction:

$$m_x(s) = \Pr(X(s) = x | X(s) \neq 0)$$
$$= \frac{p_x(s)}{1 - p_0(s)}, \quad x \in C.$$

A metapopulation model

There are n separate geographical regions (patches): $\mathcal{N} = \{1, 2, \ldots, n\}$

Let $X = (X_1, X_2, \ldots, X_n)$, where $X_i(t)$ is 1 or 0 according as patch i is occupied or not at time t ($t = 0, 1, 2, \ldots$). Note that the state space is $\mathcal{S} = \{0, 1\}^n$.

$Q = (q_{ij},\ i, j \in \mathcal{N})$ - Interaction matrix:

> q_{ij}, for $j \neq i$, is the probability that patch j will *not* be colonized by migration from patch i, and q_{ii} is the probability that (in the absence of immigration) patch i will become extinct.

A metapopulation model

Assume (Gyllenberg and Silvestrov[a]) that

$$q_{ij} = \exp(-e^{-ad_{ij}} A_i), \quad i, j \in \mathcal{N},$$

where d_{ij} is the distance between patches i and j ($d_{ii} = 0$ and $d_{ij} = d_{ji}$), A_i is the area of patch i and $a (\geq 0)$ measures how badly individuals are at migrating.

[a] M. Gyllenberg and D.S. Silvestrov. Quasi-stationary distributions of a stochastic metapopulation model. *J. Math. Biol.*, 33:35–70, 1994.

Transition probabilities

Assume that the various colonization processes and local extinction processes are independent.

Define $q_i(x)$, where $x = (x_1, x_2, \ldots, x_n)$, by

$$q_j(x) = \prod_{i=1}^{n} q_{ij}^{x_i}, \quad j \in \mathcal{N},\ x \in \mathcal{S},$$

to be the probability that patch j will become extinct at the next time step given a present configuration x.

Transition probabilities

The transition matrix $P = (p(x,y),\ x,y \in \mathcal{S})$:

$$p(x,y) = \prod_{i=1}^{n} q_i(x)^{1-y_i}(1-q_i(x))^{y_i},\ x,y \in \mathcal{S}.$$

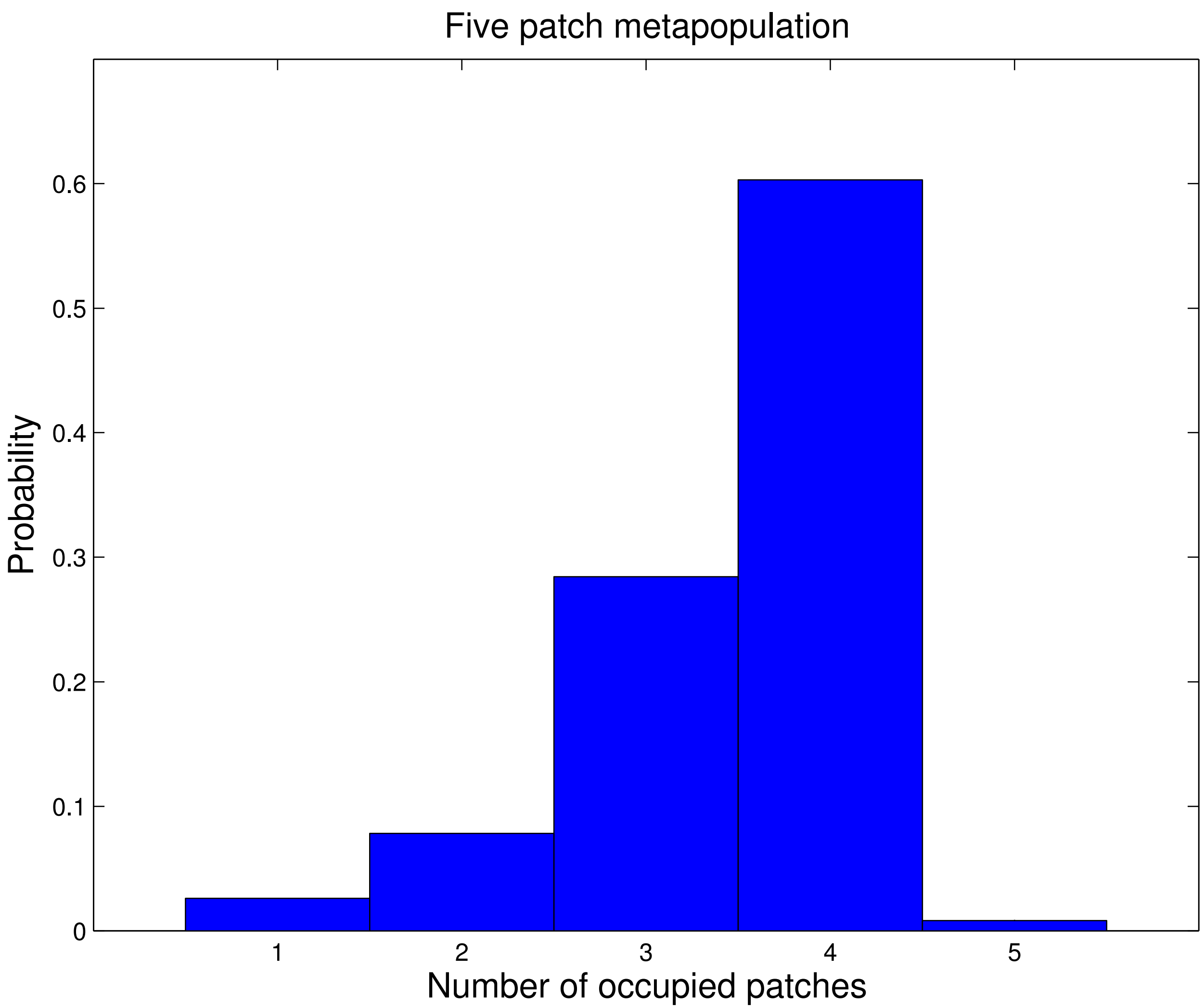
Note that, since $q_i(0) = 1$, $i \in \mathcal{N}$, state $0 = (0,0,\ldots,0)$ (corresponding to the extinction of all patches) is an absorbing state for the chain:

$$p(0,y) = \begin{cases} 1, & \text{if } y = 0, \\ 0, & \text{otherwise.} \end{cases}$$

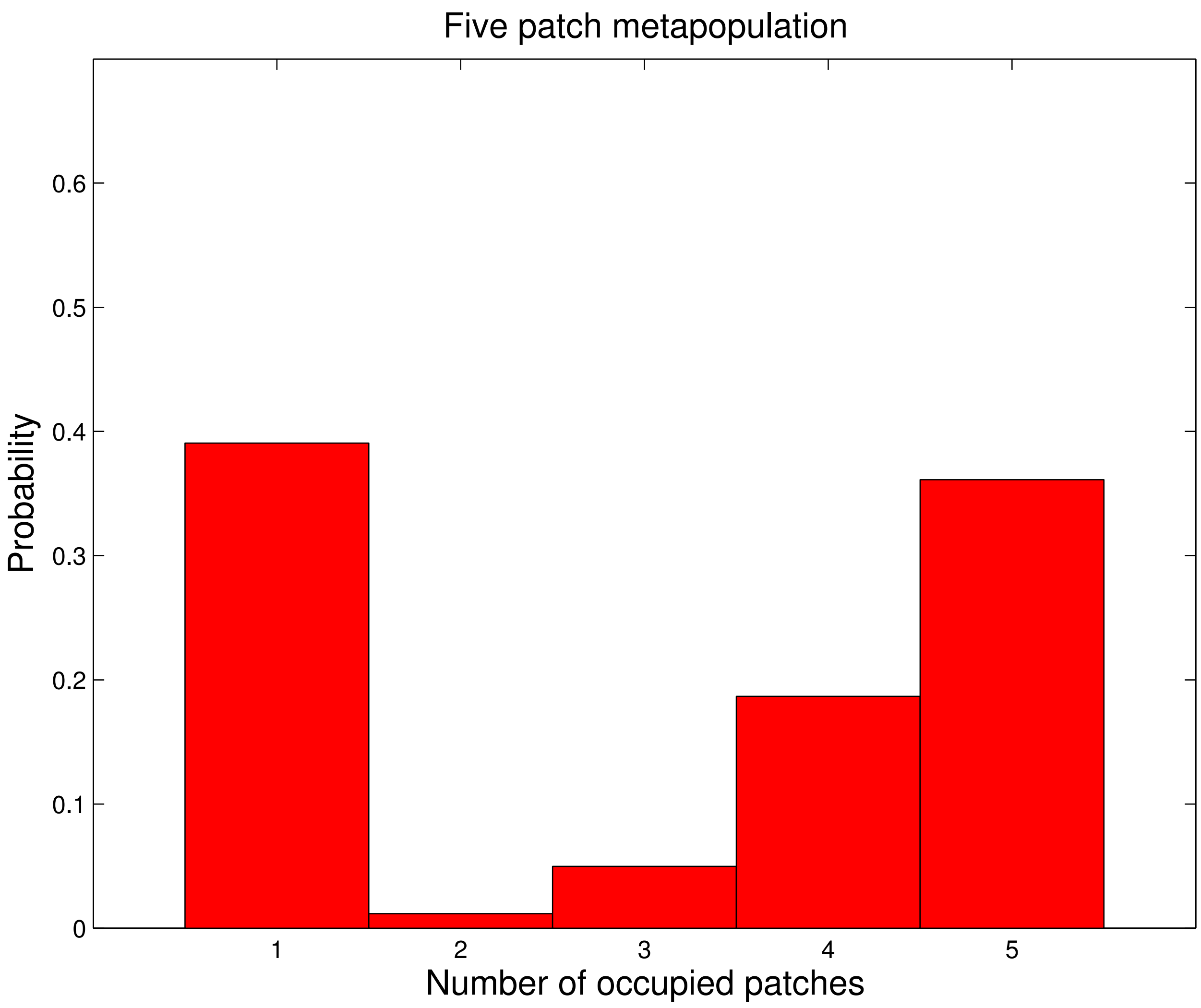
A 5-patch metapopulation

Patches 2, 3, 4 & 5 are equally spaced (a distance 0.1 apart). Patch 1 is 10 times that distance away from the others ($d_{1j} = 1$). All patches have the same area.

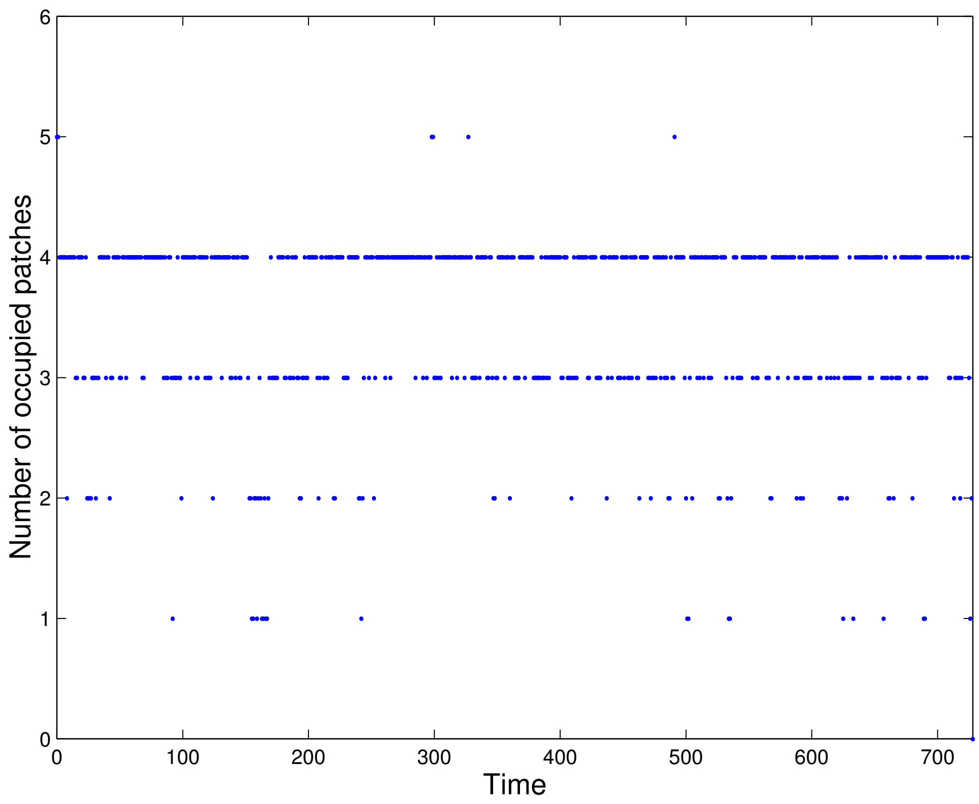
5-patch metapopulation



5-patch metapopulation

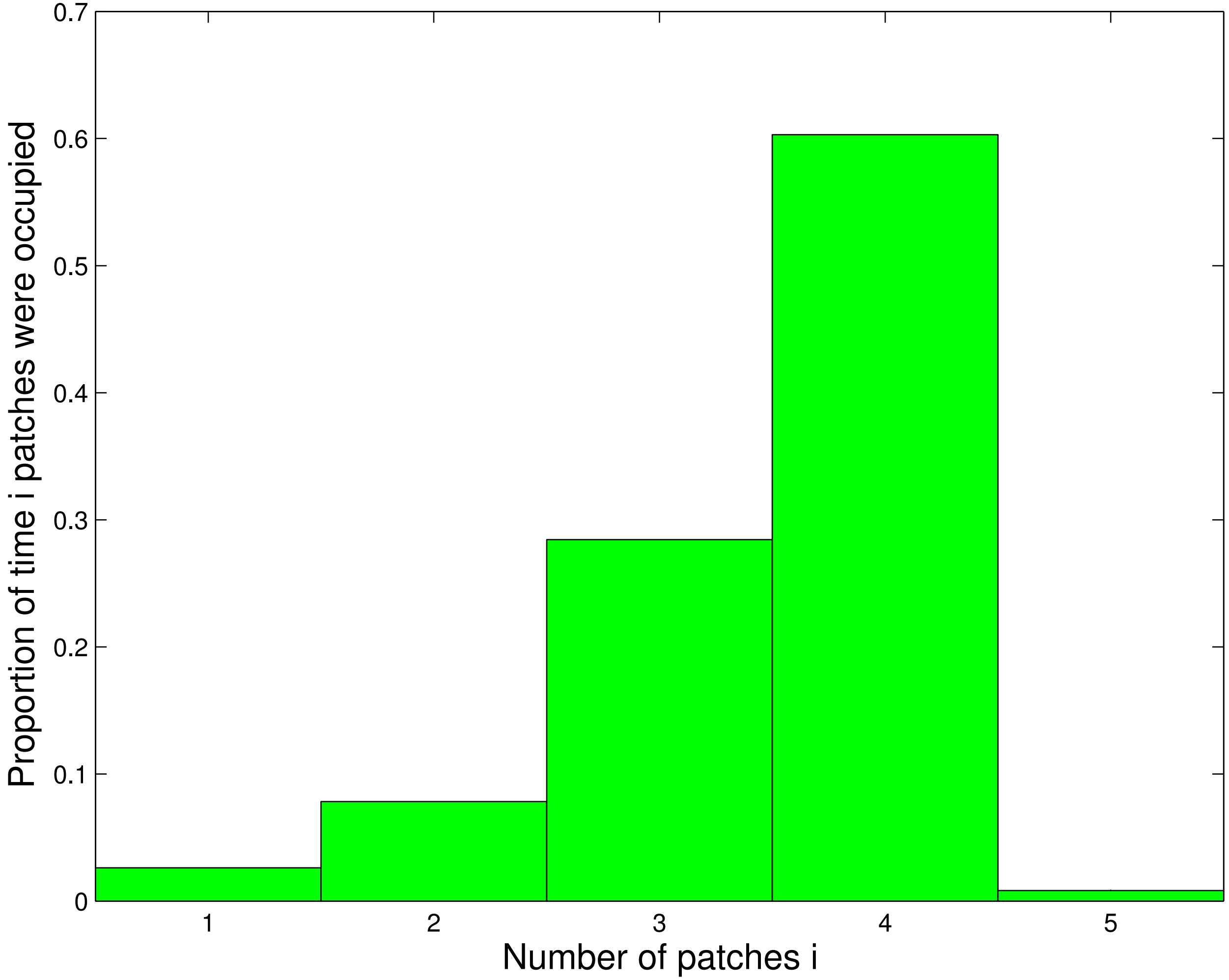


Persistence of metapopulations



Simulation of the 5-patch model with $a = 7$. The number of occupied patches is plotted against time (up to total extinction at $t = 728$).

Persistence of metapopulations



Persistence of metapopulations

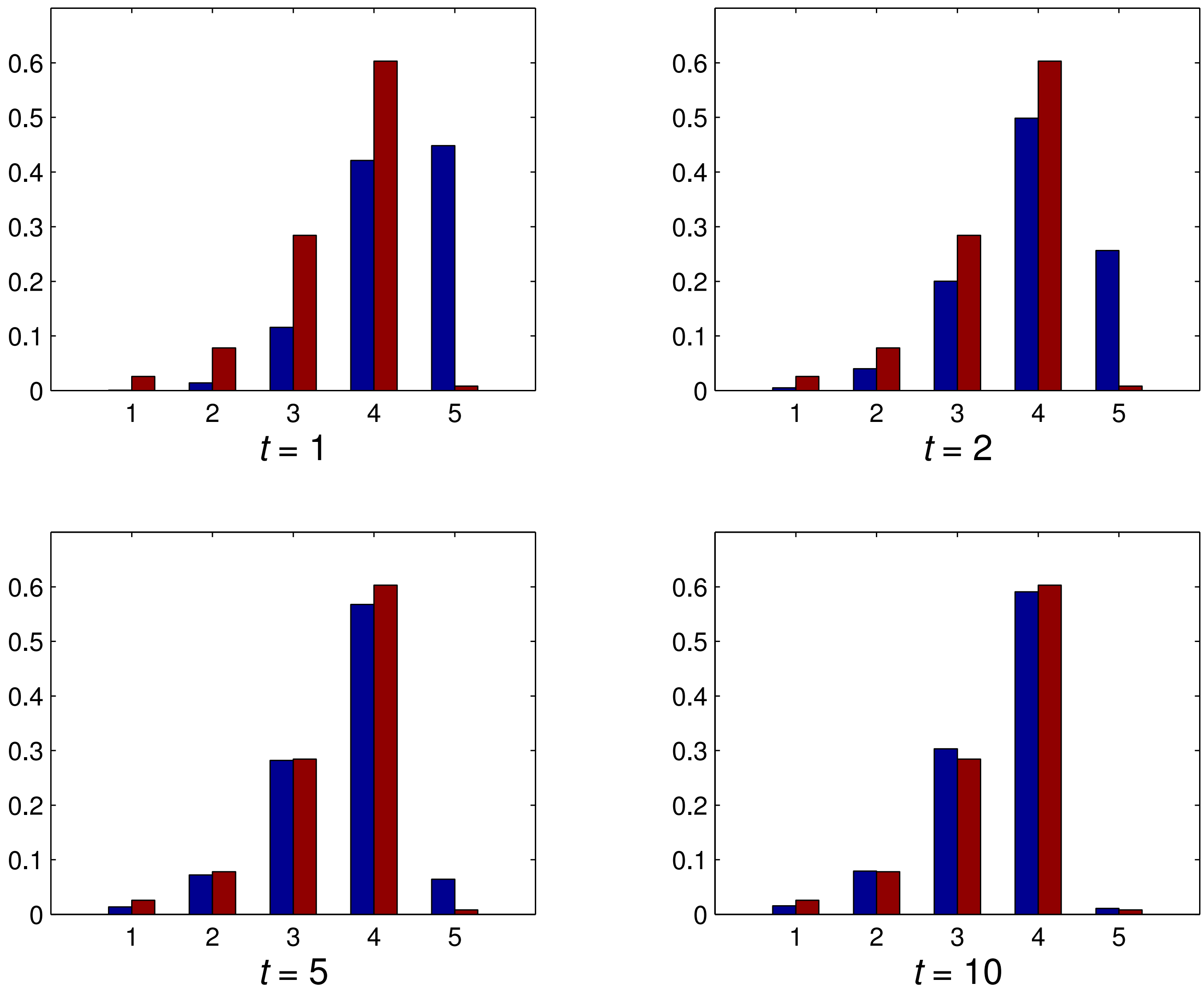
Recall that

$$m_x(s) = \Pr(X(s) = x | X(s) \neq 0) = \frac{p_x(s)}{1 - p_0(s)}, \quad x \in C.$$

Do these conditional state probabilities account for the observed behaviour?

We compare of the observed frequencies with the conditional state distribution $m_x(t)$ at $t = 1, 2, 5, 10$. The brown bar is the proportion of time for which i patches were occupied ($i = 1, 2, \ldots, 5$) during the period of the simulation. The blue bar is the distribution of the number of occupied patches evaluated using $m_x(t)$.

Persistence of metapopulations



Limiting conditional distributions

The observed trend has a simple theoretical explanation.

Since C is a finite set, the limit

$$\lim_{t \to \infty} m_x(t) = m_x$$

exists and defines a proper distribution $m = (m_x,\ x \in C)$, called a *limiting conditional distribution*, and m is the left eigenvector of P_C (P restricted to C) corresponding to the eigenvalue, ρ_1, with maximal modulus[a]. Note that the expected time till absorption, τ, is approximately $\rho_1/(1-\rho_1)$.

[a] J.N. Darroch and E. Seneta. On quasi-stationary distributions in absorbing discrete-time Markov chains. *J. Appl. Probab.*, 2:88–100, 1965.

Limiting conditional distributions

We can be precise about the *rate* of convergence by examining the eigenvalue, ρ_2, of P_C with *second-largest* modulus. It might not be real, and it has multiplicity $\kappa \geq 1$ (for simplicity, suppose $\kappa = 1$). It can be shown that

$$m_x(t) = m_x + O(\beta^t) \text{ as } t \to \infty,$$

where $\beta = |\rho_2|/\rho_1 (< 1)$.

Limiting conditional distributions

We can be precise about the *rate* of convergence by examining the eigenvalue, ρ_2, of P_C with *second-largest* modulus. It might not be real, and it has multiplicity $\kappa \geq 1$ (for simplicity, suppose $\kappa = 1$). It can be shown that

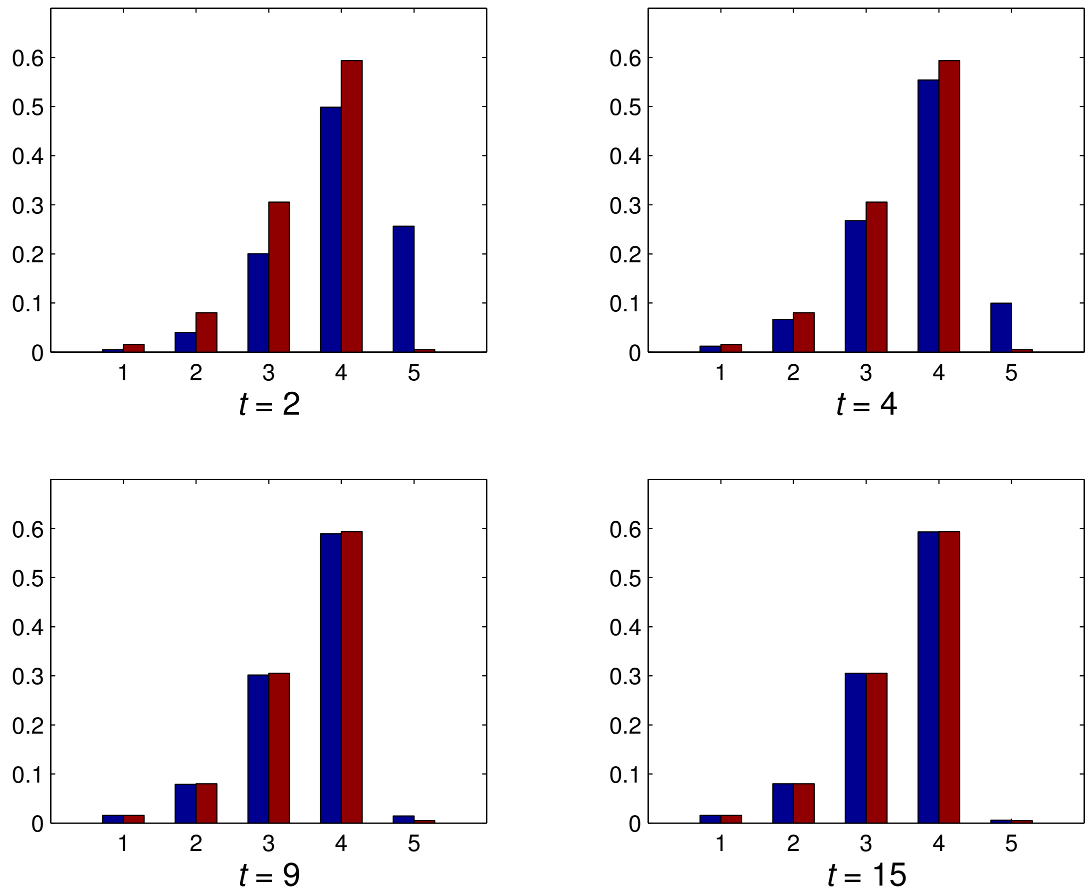
$$m_x(t) = m_x + O(\beta^t) \text{ as } t \to \infty,$$

where $\beta = |\rho_2|/\rho_1 (< 1)$.

Technical interlude

Conjecture. For a general absorbing Markov chain, $\beta < 1$ implies R-positive recurrence.

The convergence of $m_x(t)$ to m_x



Comparison between the conditional state distribution (blue) and the limiting conditional distribution (brown) of the number of occupied patches for the 5-patch model with $a = 7$.

The convergence of $m_x(t)$ to m_x

For the 5-patch metapopulation model with $a = 7$, we find that $\rho_1 \simeq 0.9979$, $\rho_2 \simeq 0.6312$ (real with multiplicity 1), $\beta(= \rho_2/\rho_1) \simeq 0.6325$ and $\tau \simeq 488$.

Pseudo-stationary distributions

(The method of Gyllenberg and Silvestrov)

Rationale: If we had assumed that Patch 1 (say) had a zero local extinction probability ($q_{11} = 0$), that patch would behave as a *mainland*.

A 5-patch metapopulation

Pseudo-stationary distributions

(The method of Gyllenberg and Silvestrov)

Rationale: If we had assumed that Patch 1 (say) had a zero local extinction probability ($q_{11} = 0$), that patch would behave as a *mainland*.

Pseudo-stationary distributions

(The method of Gyllenberg and Silvestrov)

Rationale: If we had assumed that Patch 1 (say) had a zero local extinction probability ($q_{11} = 0$), that patch would behave as a *mainland*. State 0 would no longer be accessible from all states: C would decompose into two classes, C_0 and C_1, consisting of those states in C which have $x_1 = 0$ and $x_1 = 1$ respectively; either the process would start in C_1 (mainland inhabited) and remain there, or, start in C_0 (mainland uninhabited) and eventually enter either C_1 or the absorbing state.

Pseudo-stationary distributions

(The method of Gyllenberg and Silvestrov)

Rationale: If we had assumed that Patch 1 (say) had a zero local extinction probability ($q_{11} = 0$), that patch would behave as a *mainland*. State 0 would no longer be accessible from all states: C would decompose into two classes, C_0 and C_1, consisting of those states in C which have $x_1 = 0$ and $x_1 = 1$ respectively; either the process would start in C_1 (mainland inhabited) and remain there, or, start in C_0 (mainland uninhabited) and eventually enter either C_1 or the absorbing state.

We identify a *"quasi-mainland"*, namely a single patch i with q_{ii} small (say Patch 1), and consider a sequence of processes indexed by $\epsilon = q_{ii}$, treating ϵ as a perturbation.

A 5-patch metapopulation

Perturbation theory

General idea:

$$\text{Model}(\epsilon) = \text{Model}(0) + \epsilon \times (\text{another bit}) \\ + \text{smaller order terms}$$

Perturbation theory

General idea:

$$\text{Model}(\epsilon) = \text{Model(0)} + \epsilon \times (\text{another bit})$$
$$+ \text{smaller order terms}$$

$$\text{Answer} = (\text{Answer when } \epsilon = 0) + \epsilon \times (\text{something})$$
$$+ \text{smaller order terms}$$

Perturbation theory

General idea:

$$\text{Model}(\epsilon) = \text{Model(0)} + \epsilon \times (\text{another bit}) \\ + \text{smaller order terms}$$

$$\text{Answer} = (\text{Answer when } \epsilon = 0) + \epsilon \times (\text{something}) \\ + \text{smaller order terms}$$

Indeed, we hope for

$$\text{Answer} = a_0 + a_1\epsilon + \sum_{n=2}^{\infty} a_n \epsilon^n.$$

Perturbation theory

Let $\epsilon \in (0,1]$ (now arbitrary) and suppose that our interaction matrix depends on ϵ in the following way:

$$q_{ij}^{(\epsilon)} = q_{ij} + \epsilon \hat{q}_{ij} + \circ(\epsilon), \qquad \text{as } \epsilon \to 0,$$

where

$$q_{ij} = \lim_{\epsilon \to 0} q_{ij}^{(\epsilon)} \quad \text{and} \quad \hat{q}_{ij} = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \left(q_{ij}^{(\epsilon)} - q_{ij} \right),$$

the latter assumed to be non-negative and finite, and, that $Q = (q_{ij},\ i,j \in \mathcal{N})$ satisfies $q_{11} = 0$.

Perturbation theory

Then, in an obvious notation,

$$p^{(\epsilon)}(x,y) = p(x,y) + \epsilon\hat{p}(x,y) + \circ(\epsilon),\ x,y \in \mathcal{S},$$

where $P^{(\epsilon)} = (p^{(\epsilon)}(x,y),\ x,y \in \mathcal{S})$ is the transition matrix corresponding to $Q^{(\epsilon)}$ and $P = (p(x,y),\ x,y \in \mathcal{S})$ is the transition matrix corresponding to Q.

The G&S limiting regime

Let $\epsilon \to 0$ and $t (= t_\epsilon) \to \infty$ in such a way that $\epsilon\, t_\epsilon \to s$, where $0 \le s \le \infty$.

Since the expected lifetime of the quasi-mainland is of order $1/\epsilon$, one is able to study the process on different time scales:

- $s = 0$ (smaller order)
- $s = \infty$ (larger order)
- $0 < s < \infty$ (same order)

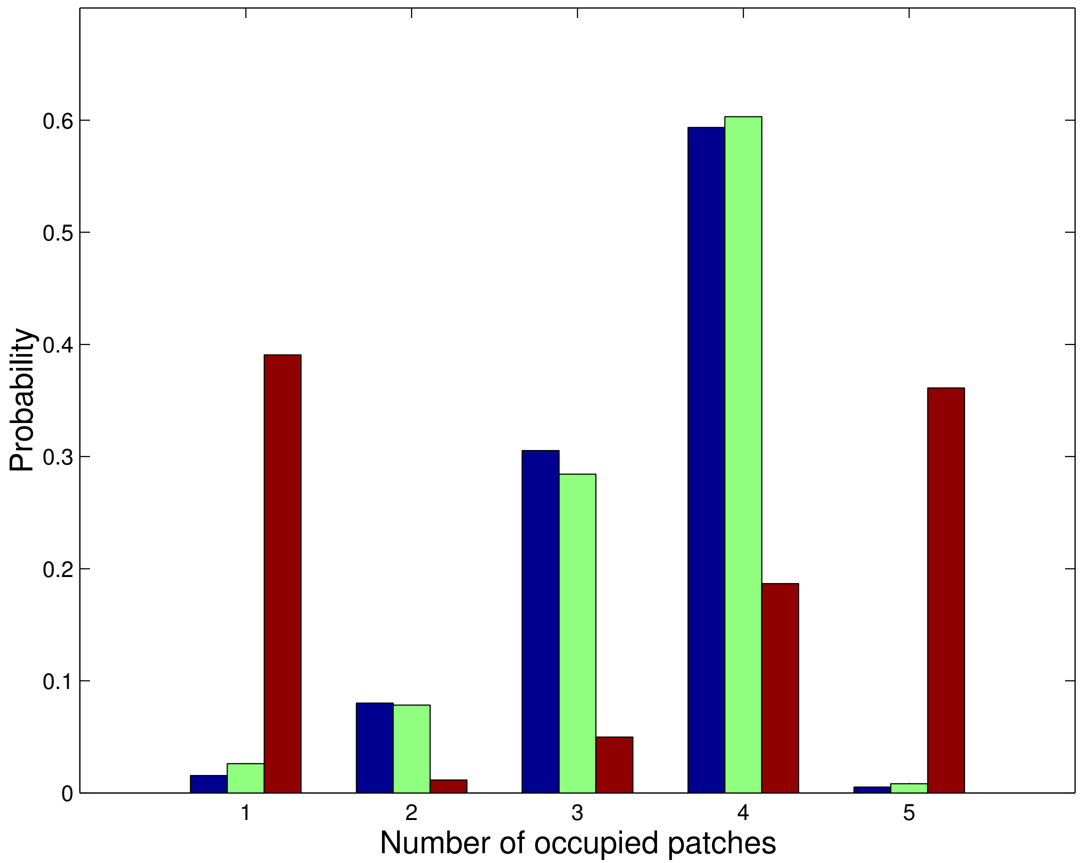
The G&S limiting regime

G&S showed that the limit

$$\lim_{\epsilon \longrightarrow 0} \Pr(X(t_\epsilon) = y | X(0) = x), \quad x, y \in C,$$

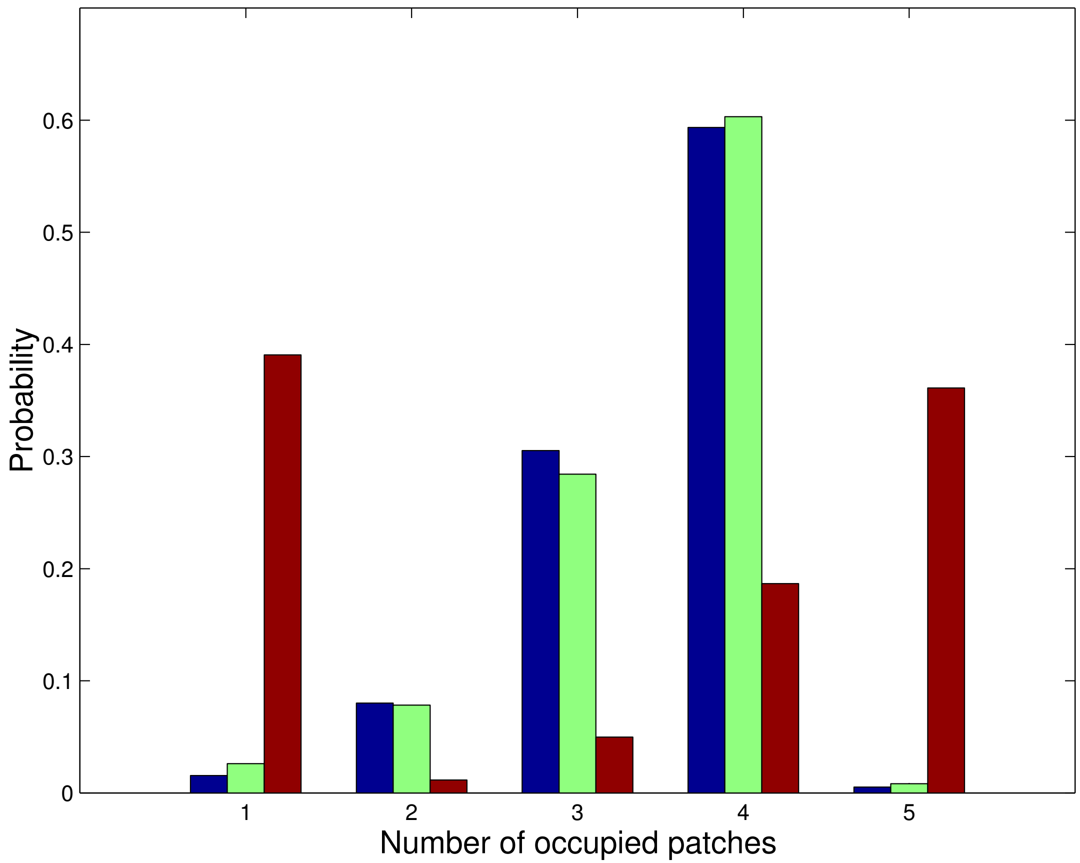
exists and is given by a mixture of the limiting probabilities $\pi(x, y)$ for the (ergodic) chain generated by Q and the degenerate distribution $\delta(y, 0)$ which assigns all its mass to state 0, the mixing probability being $e^{-\lambda s}$, where λ is a positive constant which is specified in terms of $\hat{p}(x, y)$.

Comparison using 5-patch model



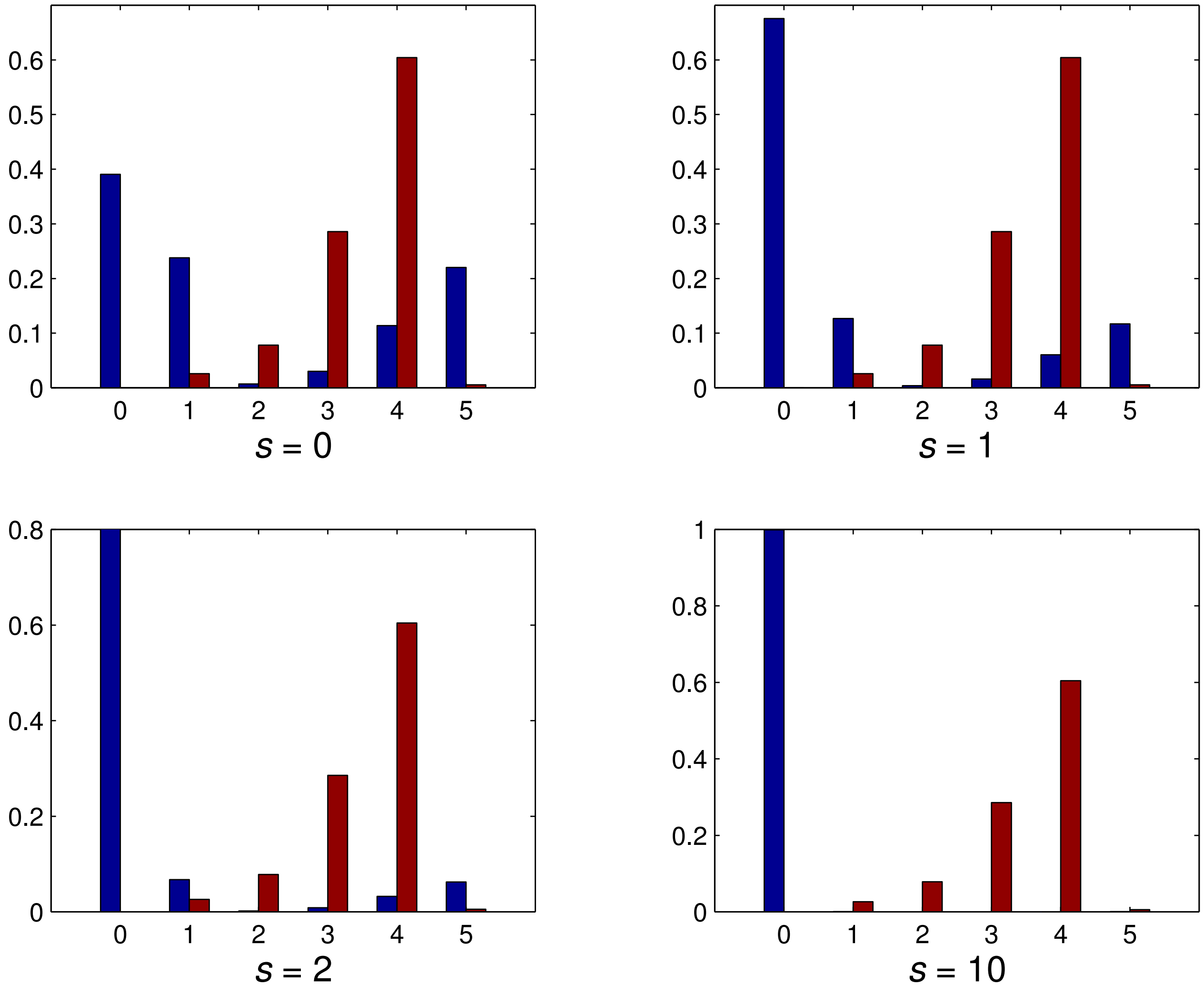
Comparison between the limiting conditional distribution (blue), the simulated proportions (green) and the pseudo-stationary distribution (brown).

Comparison using 5-patch model



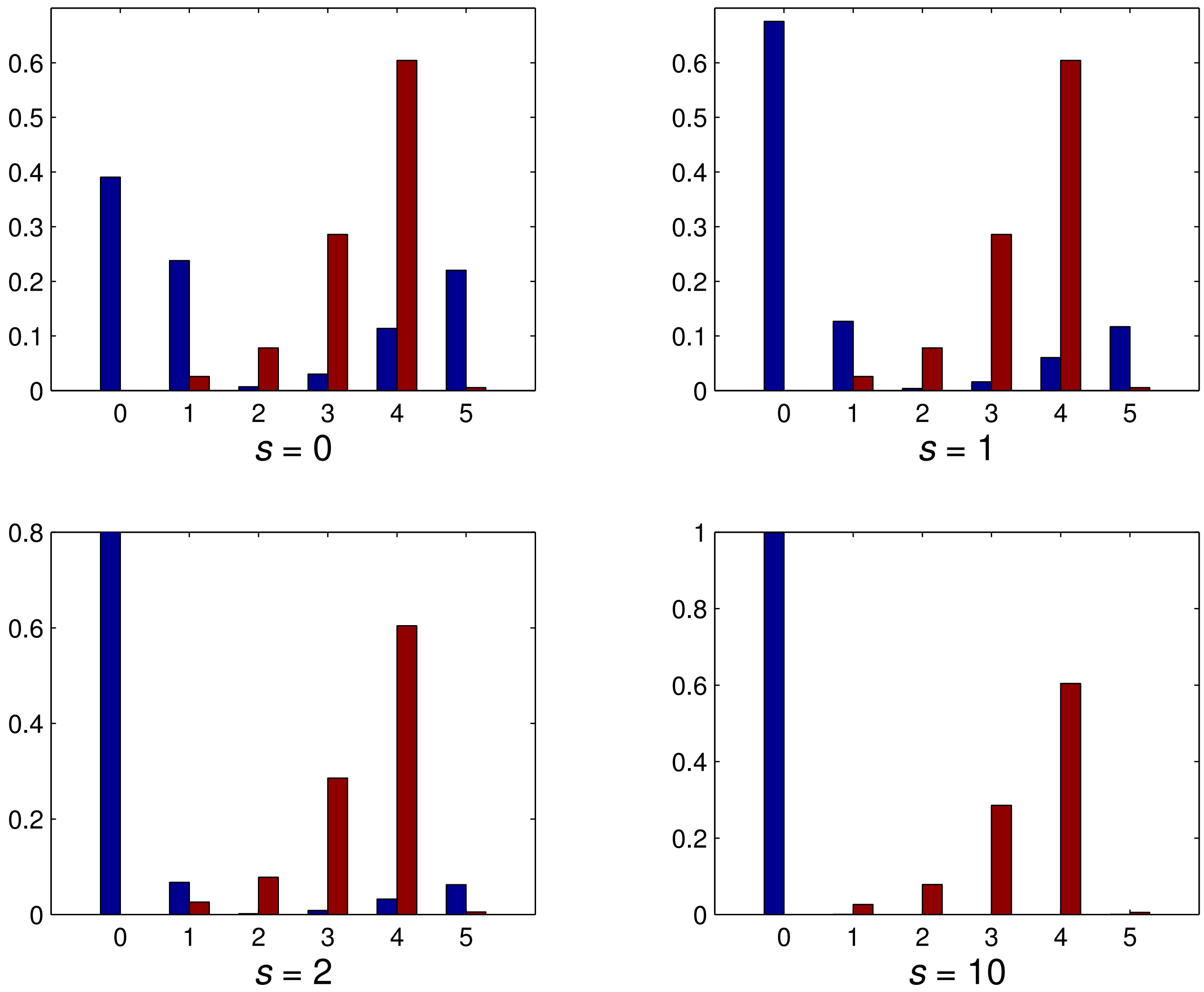
The disparity is marked: for this example, the two ways of analysing the model lead to quite different predictions.

Effect of varying s



Pseudo-stationary distribution (blue). Simulated proportions (brown).

Effect of varying s



The disparity becomes worse as the time-scale parameter s increases.

Reconciliation

Denote the state probabilities corresponding to $P^{(\epsilon)}$ by $p^{(\epsilon)}(t) = (p_x^{(\epsilon)}(t),\ x \in \mathcal{S})$, and denote the corresponding conditional probabilities by $m_x^{(\epsilon)}(t)$. Gosselin (1997) proved that

$$\lim_{\epsilon \to 0} \lim_{t \to \infty} m_x^{(\epsilon)}(t) = \begin{cases} \pi(x), & \text{if } x \in C_1, \\ 0, & \text{if } x \in C_0, \end{cases}$$

which he compared with Theorem 6.2 of G&S:

$$\lim_{\epsilon \to 0} m_x^{(\epsilon)}(t_\epsilon) = \begin{cases} \pi(x), & \text{if } x \in C_1, \\ 0, & \text{if } x \in C_0. \end{cases}$$

Reconciliation

Thus, in the important case $s = 0$ (where we are observing the process over a time scale of smaller order than the expected time to extinction of the quasi-mainland), the limiting conditional distribution and the pseudo-stationary agree when ϵ is small.

The problem with the 5-patch model is that $\epsilon (= q_{11}) \simeq 0.3679$ (not small enough).

Remarks

Quasi-stationarity is a *property of the model* and *not* the means of analysing it.

The 5-patch model exhibits quasi-stationarity, demonstrated emphatically using simulation, yet q_{11} is not small. The pseudo-stationary distribution does not capture this behaviour.

On the other hand, the conditional state distribution $m(t)$ does: after all, it is the *most information our model can provide* at any time t given that we know extinction has not occurred by time t. In cases when the convergence of $m(t)$ to the limiting conditional distribution m is rapid, *this* distribution can be used instead.