# Quasi-stationary distributions: then and now

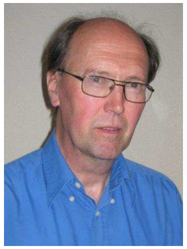
Phil Pollett

Department of Mathematics
The University of Queensland
http://www.maths.uq.edu.au/~pkp

ACEMS
AUSTRALIAN RESEARCH COUNCIL CENTRE OF EXCELLENCE FOR
MATHEMATICAL AND STATISTICAL FRONTIERS

# Erik!

Erik A. van Doorn
Applied Mathematics Department
University of Twente



*Van Doorn, E.A. and Pollett, P.K. (2013) Quasi-stationary distributions for discrete-state models (with web appendix). Invited paper. European J. Operat. Res. 230, 1–14.

*Van Doorn, E.A. and Pollett, P.K. (2009) Quasi-stationary distributions for reducible absorbing Markov chains in discrete time. Markov Process. Related Fields 15, 191–204.

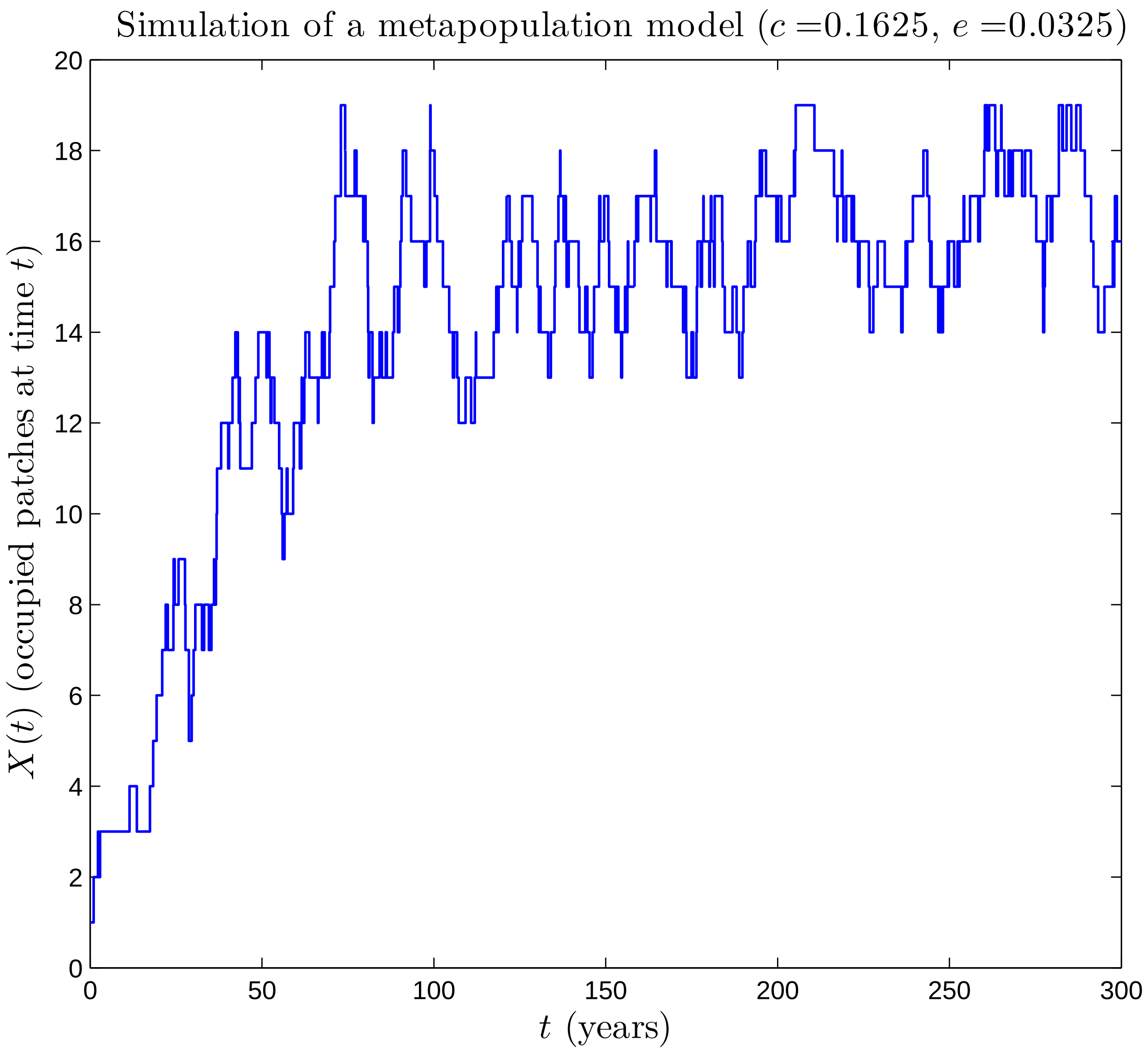
*Van Doorn, E.A. and Pollett, P.K. (2008) Survival in a quasi-death process. Linear Alg. Appl. 429, 776–791.

# Collaborators

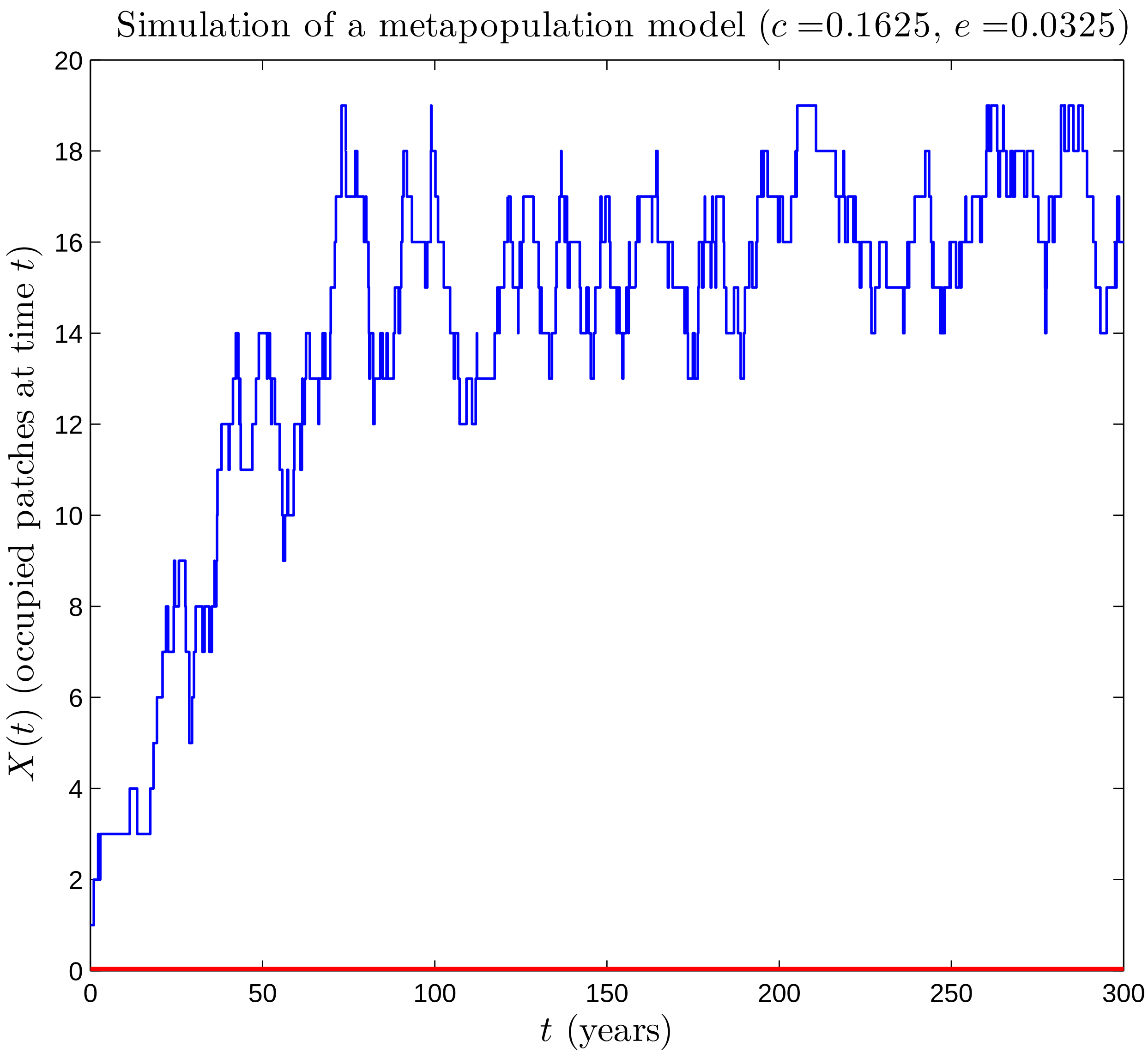
## Co-authors (by number of collaborations)

Adan, Ivo Jean-Baptiste François   Blanc, J. P. C.   Boucherie, Richard J.   Coolen-Schrijner, Pauline   de Wit, J. S. J.   Jagers, A. A.   Kijima, Masaaki   König, Dieter   Lenin, R. B.   Mehta, Madan Lal   Meijer, T. M. J.   Nair, M. Gopalan   Panfilova, Tatyana   Parthasarathy, Panamali Ramarao   Pollett, Philip K.   Regterschot, G. J. K.   Resing, Jacques   Scheinhardt, Werner R. W.   Schmidt, Volker[1]   van Arem, B.   van Foreest, Nicky D.   Zeifman, A. I.

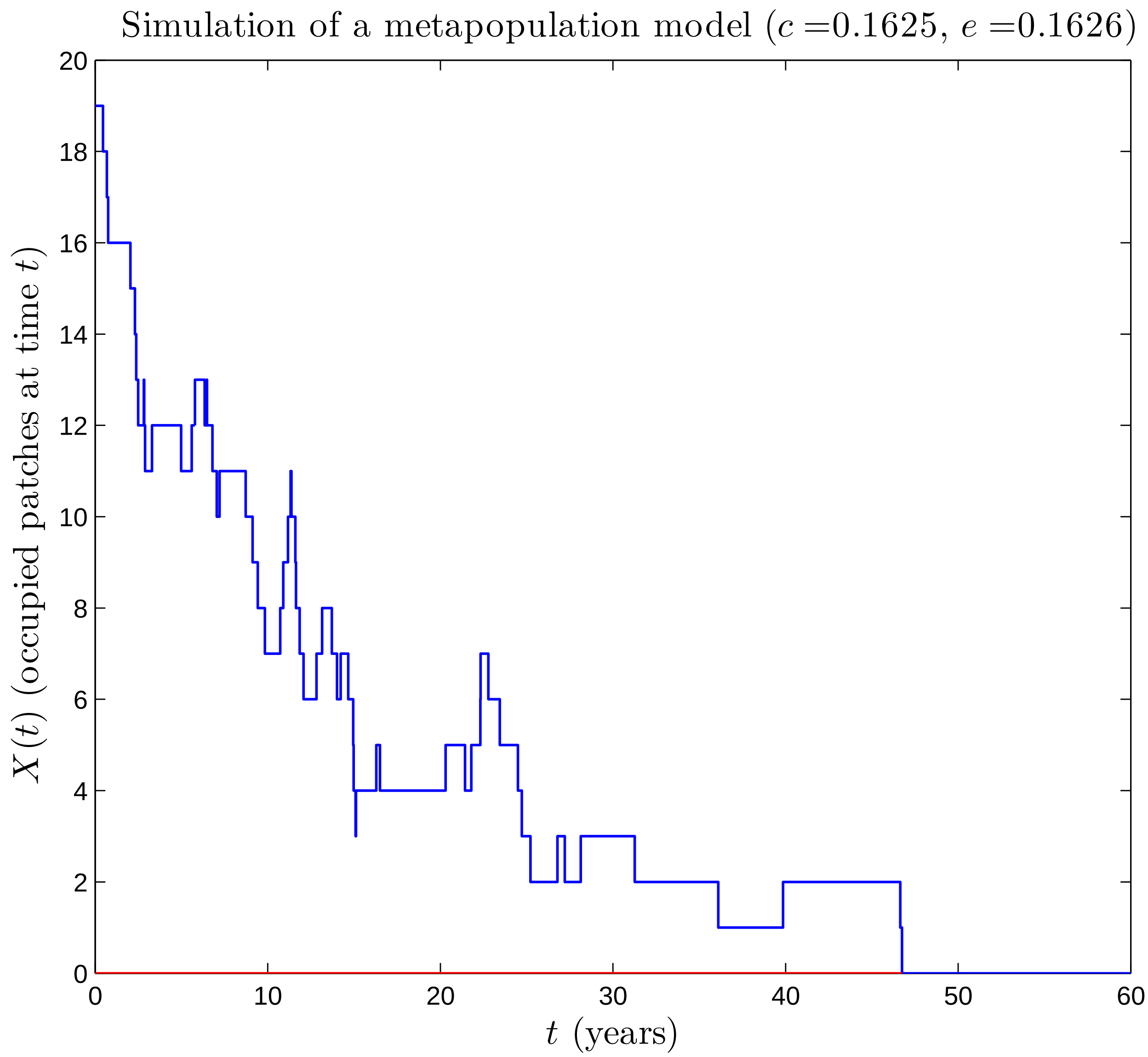
# Quasi stationarity



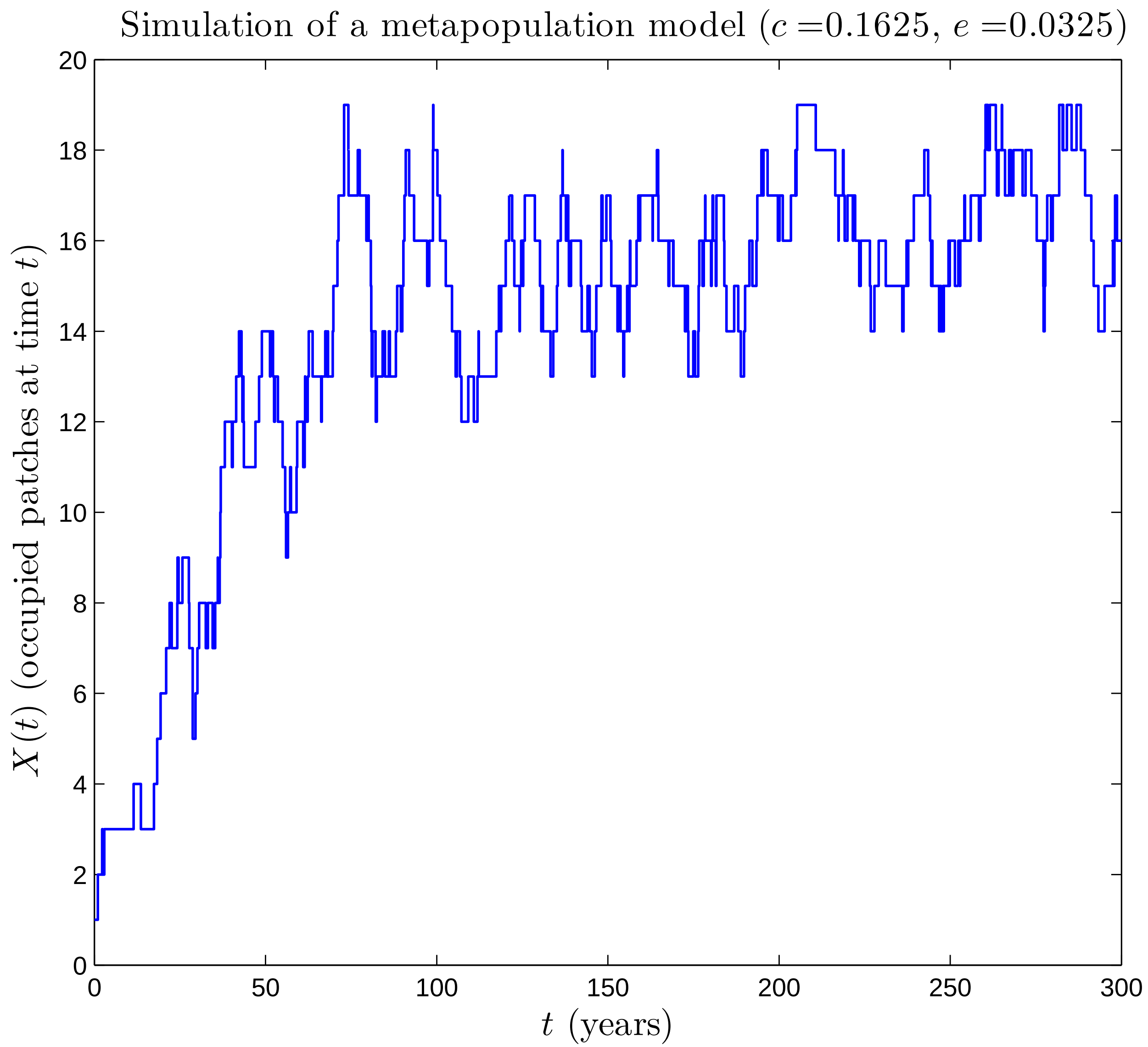
# Quasi stationarity



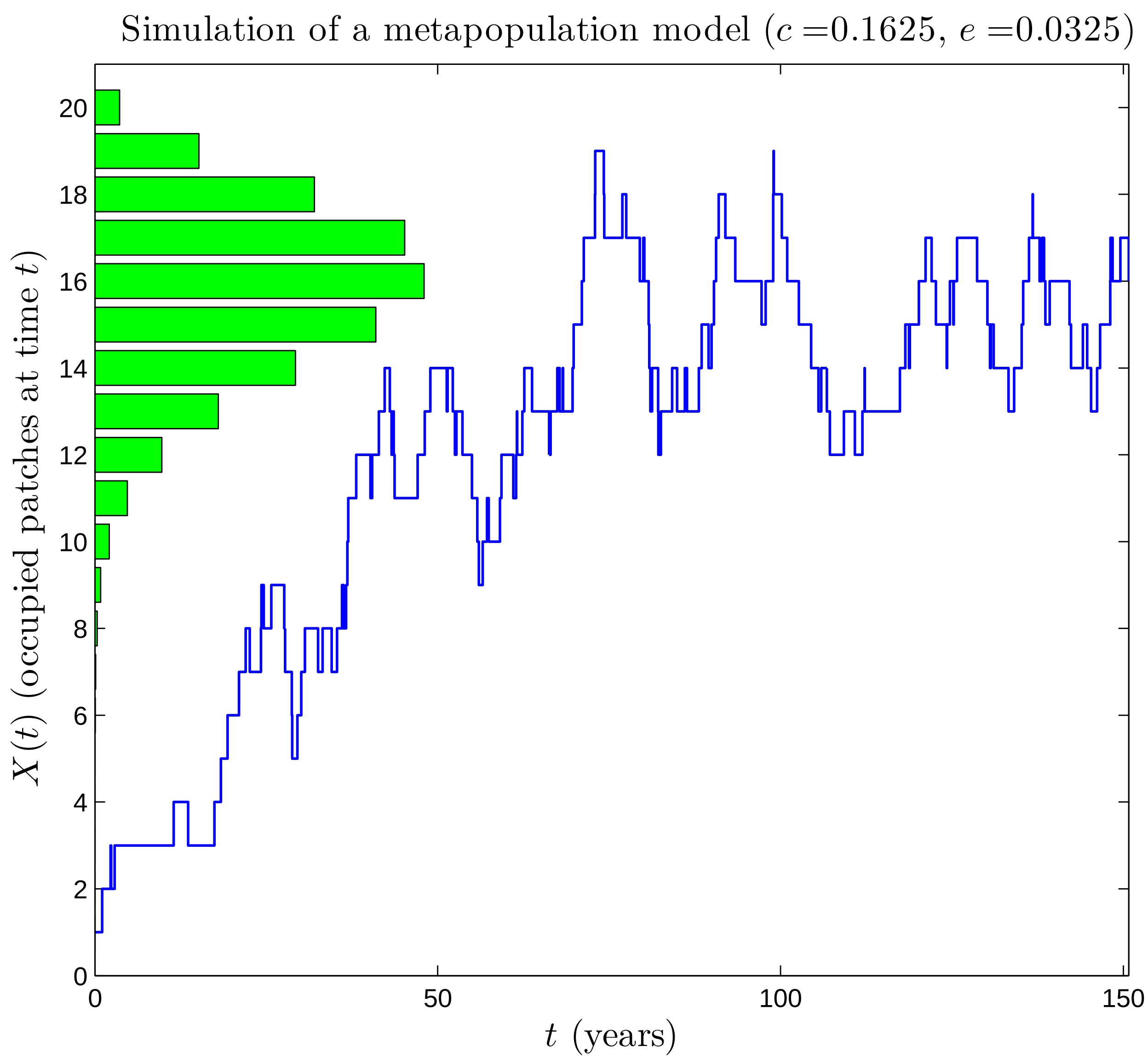
# Evanescence



# Quasi stationarity



# Quasi stationarity



# A “quasi-stationary” distribution

Think of an observer who at some time $t$ is *aware of the occupancy of some patches*, yet cannot tell exactly which of $n$ patches are occupied.

# A “quasi-stationary” distribution

Think of an observer who at some time $t$ is *aware of the occupancy of some patches*, yet cannot tell exactly which of $n$ patches are occupied.

What is the chance of there being precisely $i$ patches occupied?

# A “quasi-stationary” distribution

What is the chance of there being precisely $i$ patches occupied?

# A "quasi-stationary" distribution

What is the chance of there being precisely $i$ patches occupied?

If we were equipped with the full set of state probabilities

$$p_i(t) = \mathbb{P}(X(t) = i), \qquad i \in \{0, 1, \ldots, n\},$$

we would evaluate the *conditional probability*

$$u_i(t) = \mathbb{P}(X(t) = i | X(t) \neq 0) = \frac{p_i(t)}{1 - p_0(t)},$$

for $i$ in the set $S = \{1, \ldots, n\}$ of transient states.

# A “quasi-stationary” distribution

$$u_i(t) = \mathbb{P}(X(t) = i | X(t) \neq 0) = \frac{p_i(t)}{1 - p_0(t)}, \qquad i \in S.$$

Then, in view of the behaviour observed in our simulation, it would be natural for us to seek a distribution $\boldsymbol{u} = (u_i,\ i \in S)$ over $S$ such that if $u_i(t) = u_i$ for a particular $t > 0$, then $u_i(s) = u_i$ for all $s > t$.

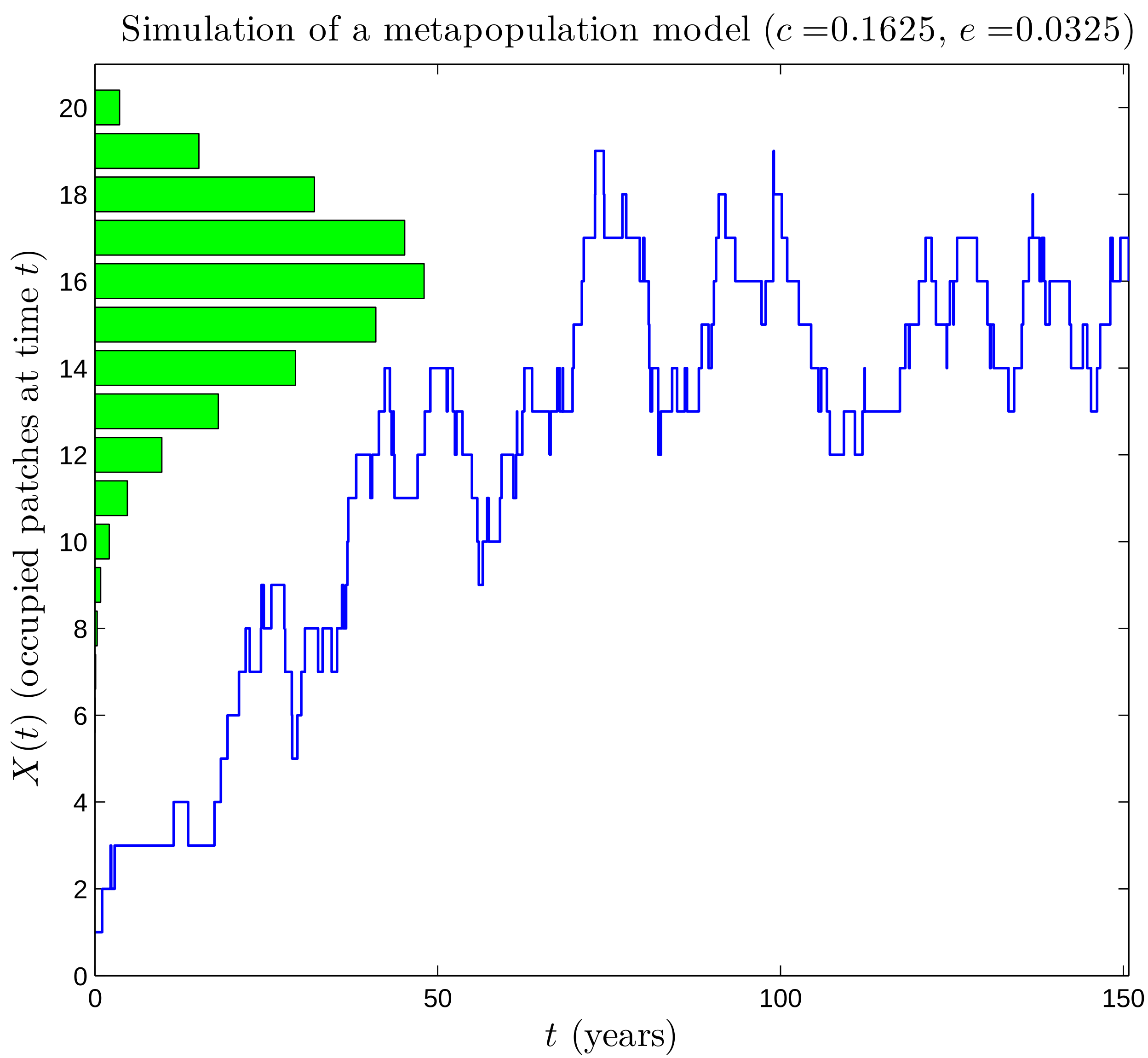
# A quasi-stationary distribution

$$u_i(t) = \mathbb{P}(X(t) = i | X(t) \neq 0) = \frac{p_i(t)}{1 - p_0(t)}, \qquad i \in S.$$

Then, in view of the behaviour observed in our simulation, it would be natural for us to seek a distribution $\boldsymbol{u} = (u_i,\ i \in S)$ over $S$ such that if $u_i(t) = u_i$ for a particular $t > 0$, then $u_i(s) = u_i$ for all $s > t$.

Such a distribution is called a *stationary conditional distribution* or *quasi-stationary distribution* (QSD).

# A quasi-stationary distribution



# Quasi-stationary distributions

We seek a distribution $\boldsymbol{u} = (u_i,\, i \in S)$ over $S$ such that if $u_i(t) = u_i$ for a particular $t > 0$, then $u_i(s) = u_i$ for all $s > t$.

Such a distribution $\boldsymbol{u}$ is called a *stationary conditional distribution* or *quasi-stationary distribution* (QSD).

# Quasi-stationary distributions

We seek a distribution $\boldsymbol{u} = (u_i,\, i \in S)$ over $S$ such that if $u_i(t) = u_i$ for a particular $t > 0$, then $u_i(s) = u_i$ for all $s > t$.

Such a distribution $\boldsymbol{u}$ is called a *stationary conditional distribution* or *quasi-stationary distribution* (QSD).

**Key message**: $\boldsymbol{u}$ can usually be determined from the transition rates of the process and $\boldsymbol{u}$ might then also be a *limiting conditional distribution* (LCD) in that $u_i(t) \to u_i$ as $t \to \infty$, and thus be of use in modelling the long-term behaviour of the process.

# Quasi-stationary distributions

Thanks to Erik, we have a complete picture for *birth-death processes* and *birth-death chains*. For example:

*Van Doorn, E.A. (1991) Quasi-stationary distributions and convergence to quasi-stationarity of birth-death processes. Adv. Appl. Probab. 23 683–700.

# Quasi-stationary distributions

Thanks to Erik, we have a complete picture for *birth-death processes* and *birth-death chains*. For example:

*Van Doorn, E.A. (1991) Quasi-stationary distributions and convergence to quasi-stationarity of birth-death processes. Adv. Appl. Probab. 23 683–700.

Conditions are given to delineate three possible cases:

(i) no QSD, and $u_i(t) \to 0$ (fixed initial state).

(ii) a unique QSD $\boldsymbol{u}$, and $u_i(t) \to u_i$ (fixed initial state).

(iii) a one-parameter family of QSDs, and we get convergence to the *extremal* QSD.

# Quasi-stationary distributions

Thanks to Erik, we have a complete picture for *birth-death processes* and *birth-death chains*. For example:

*Van Doorn, E.A. (1991) Quasi-stationary distributions and convergence to quasi-stationarity of birth-death processes. Adv. Appl. Probab. 23 683–700.

Conditions are given to delineate three possible cases:

(i) no QSD, and $u_i(t) \to 0$ (fixed initial state).

(ii) a unique QSD $\boldsymbol{u}$, and $u_i(t) \to u_i$ (fixed initial state).

(iii) a one-parameter family of QSDs, and we get convergence to the *extremal* QSD.

Furthermore, we are told *how fast* $u_i(t)$ converges to $u_i$.

# Erik in action

# Quasi-stationary distributions

Thanks to Erik, we have a complete picture for *birth-death processes* and *birth-death chains*. For example:

*Van Doorn, E.A. (1991) Quasi-stationary distributions and convergence to quasi-stationarity of birth-death processes. Adv. Appl. Probab. 23 683–700.

Conditions are given to delineate three possible cases:

(i) no QSD, and $u_i(t) \to 0$ (fixed initial state).

(ii) a unique QSD $\boldsymbol{u}$, and $u_i(t) \to u_i$ (fixed initial state).

(iii) a one-parameter family of QSDs, and we get convergence to the *extremal* QSD.

Furthermore, we are told *how fast* $u_i(t)$ converges to $u_i$.

# Domain of attraction problem

Let $T = \inf\{t \geq 0 : X(t) = 0\}$ be the *absorption time* (or *survival time*), and recall that a distribution $\boldsymbol{u}$ is a QSD if, for all $t$, $\mathbb{P}_{\boldsymbol{u}}(X(t) = j \,|\, T > t) = u_j$, $j \in S$.

# Domain of attraction problem

Let $T = \inf\{t \geq 0 : X(t) = 0\}$ be the *absorption time* (or *survival time*), and recall that a distribution $\boldsymbol{u}$ is a QSD if, for all $t$, $\mathbb{P}_{\boldsymbol{u}}(X(t) = j \,|\, T > t) = u_j$, $j \in S$.

Let $\boldsymbol{u} = (u_i,\ i \in S)$ be a given QSD. If $\boldsymbol{u}$ is a LCD for some initial distribution $\boldsymbol{w} = (w_i,\ i \in S)$, that is

$$\lim_{t\to\infty} \mathbb{P}_{\boldsymbol{w}}(X(t) = j \,|\, T > t) = u_j, \quad j \in S,$$

we say that $\boldsymbol{w}$ is *in the domain of attraction of* $\boldsymbol{u}$.

# Domain of attraction problem

Let $T = \inf\{t \geq 0 : X(t) = 0\}$ be the *absorption time* (or *survival time*), and recall that a distribution $\boldsymbol{u}$ is a QSD if, for all $t$, $\mathbb{P}_{\boldsymbol{u}}(X(t) = j \,|\, T > t) = u_j$, $j \in S$.

Let $\boldsymbol{u} = (u_i,\ i \in S)$ be a given QSD. If $\boldsymbol{u}$ is a LCD for some initial distribution $\boldsymbol{w} = (w_i,\ i \in S)$, that is

$$\lim_{t \to \infty} \mathbb{P}_{\boldsymbol{w}}(X(t) = j \,|\, T > t) = u_j, \quad j \in S,$$

we say that $\boldsymbol{w}$ is *in the domain of attraction of* $\boldsymbol{u}$.

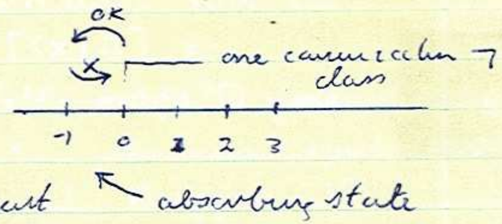
**Problem**: Identify the domains of attraction.

# QSDs - first contact - 21 July 1978

<u>Quasi stationary Distributions</u>

We consider a single transient class of a markov chain states 0, 1, 2, 3, .... Probability is ultimately absorbed (with prob 1) in a state −1.

If we shifted state labels up by 1. absorbing state 0
~~states~~ class $\{1, 2, 3, \dots\} = T$




Then this could for instance represent a population which ultimately becomes extinct.

We consider the chain as being described by $(Q_{ij} : i, j \geq 0)$

so $Q_{ij} \geq 0$ $\sum_j Q_{ij} \leq 1$

ie we consider <u>sub-stochastic matrices</u>

For simplicity we will assume that the states have period 1.

<u>Theorem D.V.J</u>

Suppose we have a substochastic irreducible chain $(Q_{ij})$. Then for all $i, j$, the series $\sum_{n=1}^{\infty} Q_{ij}^{(n)} z^n$ have a common radius of convergence. Suffices to prove that for any $i, j$ the series $\sum_{n=1}^{\infty} Q_{ij}^{(n)} z^n$, $\sum_{n=1}^{\infty} Q_{ii}^{(n)} z^n$, $\sum_{n=1}^{\infty} Q_{ji}^{(n)} z^n$, $\sum_{n=1}^{\infty} Q_{jj}^{(n)} z^n$ have same radius of convergence.

# The Yaglom limit

Yaglom* was the first to identify explicitly a LCD, establishing the existence of such for the subcritical Bienaymé-Galton-Watson branching process.

*Yaglom, A.M. (1947) Certain limit theorems of the theory of branching processes. Dokl. Acad. Nauk SSSR 56, 795–798 (in Russian).

# The Yaglom limit

Yaglom* was the first to identify explicitly a LCD, establishing the existence of such for the subcritical Bienaymé-Galton-Watson branching process.

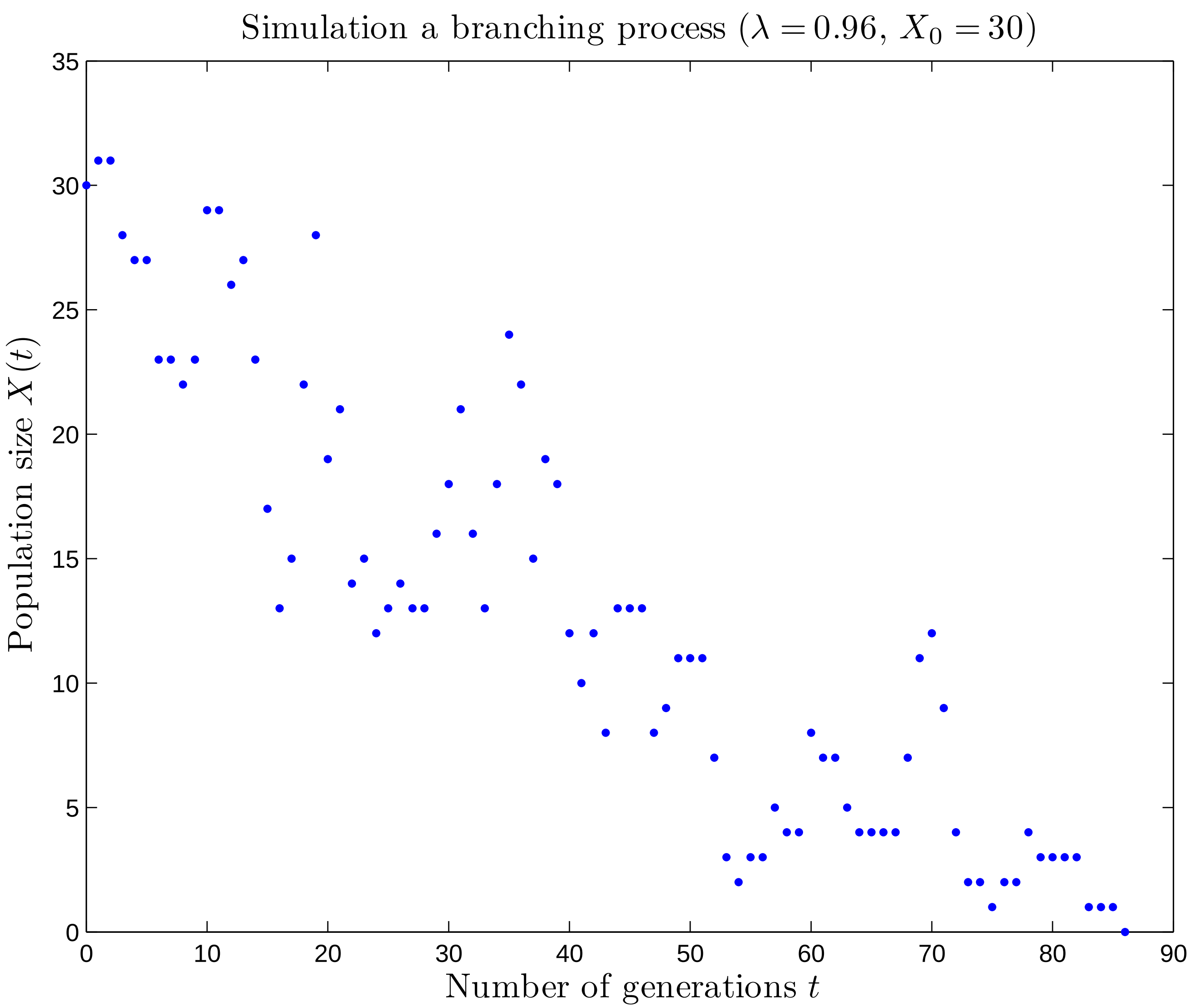
*Yaglom, A.M. (1947) Certain limit theorems of the theory of branching processes. Dokl. Acad. Nauk SSSR 56, 795–798 (in Russian).

If the expected number $\lambda$ of offspring is less than $1$, then

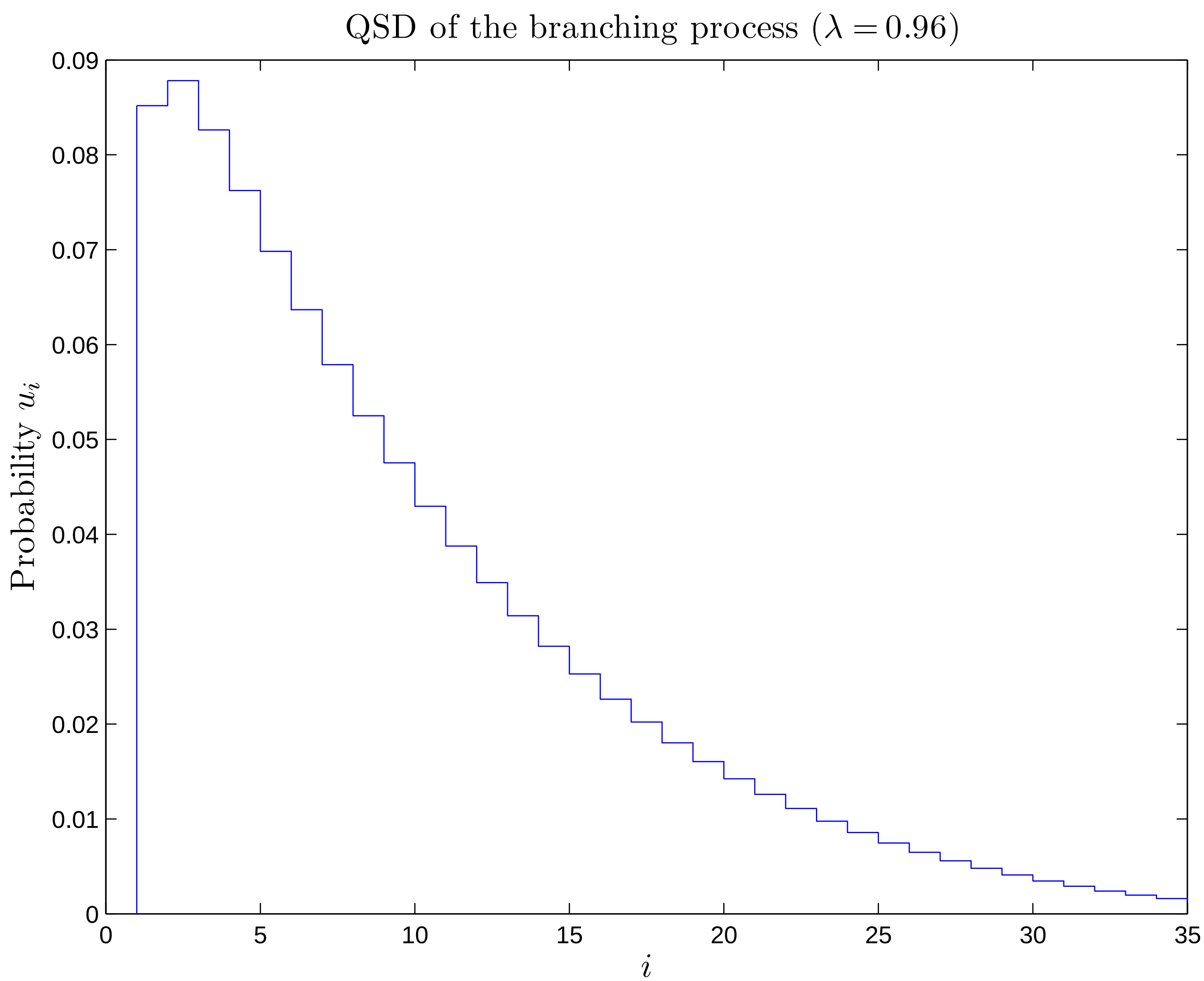
$$u_i = \lim_{n \to \infty} \mathbb{P}(X_n = i | X_n \neq 0, \, X_0 = 1), \qquad i \in S,$$

exists and defines a proper probability distribution $\boldsymbol{u} = (u_i, \, i \in S)$ over $S$.

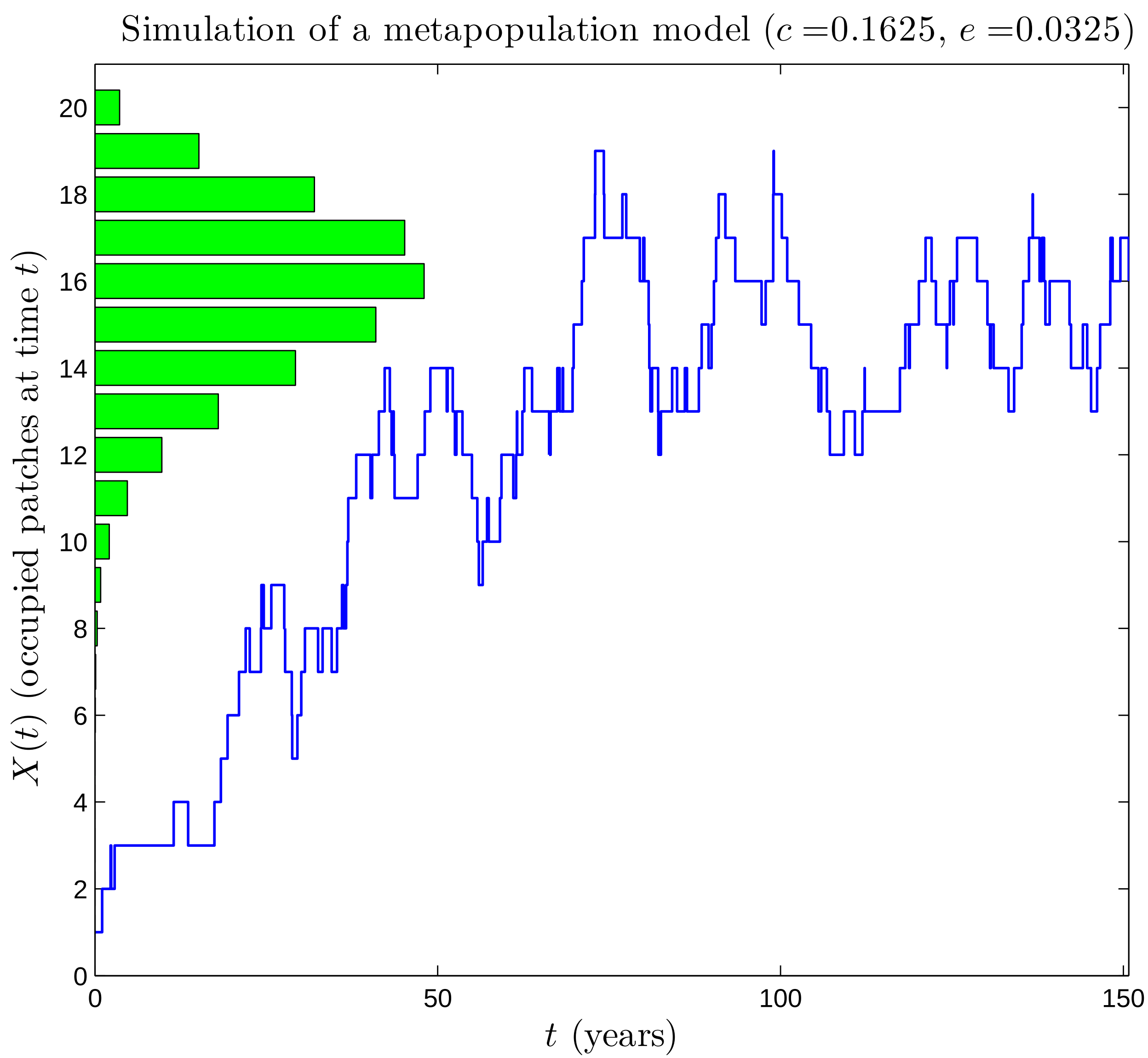
# Subcritical - quasi stationarity?



# The Yaglom limit



# The quasi-stationary distribution



# Origins of the idea

The idea of a limiting conditional distribution goes back much further than Yaglom, at least to Wright* in his discussion of gene frequencies in finite populations:

> "As time goes on, divergences in the frequencies of factors may be expected to increase more and more until at last some are either completely fixed or completely lost from the population. The distribution curve of gene frequencies should, however, approach a definite form if the genes which have been wholly fixed or lost are left out of consideration."

*Wright, S. (1931) Evolution in Mendelian populations. Genetics 16, 97–159.

# Origins of the idea

The idea of "quasi stationarity" was crystallized by Bartlett*:

"While presumably on the above model [for the interactions between active and passive forms of flour beetle] extinction of the population will occur after a long enough time, this may (for a deterministic 'ceiling' population not too small, but fluctuations relatively small) be so long delayed as to be negligible and an effective or quasi-stationarity be established."

*Bartlett, M.S. (1957) On theoretical models for competitive and predatory biological systems. Biometrika 44, 27–42.

# Origins of the idea

Bartlett* later coined the term “quasi-stationary distribution”:

> “It still may happen that the time to extinction is so long that it is still of more relevance to consider the effectively ultimate distribution (called a ‘quasi-stationary’ distribution) of [the process] $N$.”

*Bartlett, M.S. (1960) Stochastic Population Models in Ecology and Epidemiology. Methuen, London.

# The setting of our most recent work

We consider a time-homogeneous *finite-state* Markov process $(X(t),\, t \geq 0)$ taking values in $\{0\} \cup S$, where $0$, the sole absorbing state, is reached with probability $1$.

**Note**: $S$ is not necessarily irreducible.

*Van Doorn, E.A. and Pollett, P.K. (2009) Quasi-stationary distributions for reducible absorbing Markov chains in discrete time. Markov Process. Related Fields 15, 191–204.

*Van Doorn, E.A. and Pollett, P.K. (2008) Survival in a quasi-death process. Linear Alg. Appl. 429, 776–791.

# Structure - continuous time

Communicating classes: $S$ comprises $S_1, S_2, \ldots, S_L$.

# Structure - continuous time

Communicating classes: $S$ comprises $S_1, S_2, \ldots, S_L$.

Partial ordering: $S_i \prec S_j$ means $S_i$ is *accessible from* $S_j$.

# Structure - continuous time

**Communicating classes:** $S$ comprises $S_1, S_2, \ldots, S_L$.

**Partial ordering:** $S_i \prec S_j$ means $S_i$ is *accessible from* $S_j$.

**Assume:** $S_i \prec S_j \;\Rightarrow i \;\leq j$, so that

$$Q = \begin{pmatrix} \mathcal{Q}_1 & \mathsf{O} & \mathsf{O} & \cdots & \mathsf{O} \\ \mathcal{Q}_{21} & \mathcal{Q}_2 & \mathsf{O} & \cdots & \mathsf{O} \\ \mathcal{Q}_{31} & \mathcal{Q}_{32} & \mathcal{Q}_3 & \cdots & \mathsf{O} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathcal{Q}_{L1} & \mathcal{Q}_{L2} & \mathcal{Q}_{L3} & \cdots & \mathcal{Q}_L \end{pmatrix}$$

# Decay parameters

Eigenvalues. Clearly $\mathrm{Sp}(Q) = \cup_i \mathrm{Sp}(Q_i)$.

# Decay parameters

Eigenvalues. Clearly $\mathrm{Sp}(Q) = \cup_i \mathrm{Sp}(Q_i)$. Also, we know (Theorem 2.6 of Seneta's book*) that the eigenvalue $-\alpha_k$ of $Q_k$ with *maximal real part* is unique, simple (multiplicity $1$), and strictly negative.

*Seneta, E. (1981) Non-negative Matrices and Markov Chains. Revised Edition. Springer, New York.

# Decay parameters

Eigenvalues. Clearly Sp$(Q) = \cup_i$Sp$(Q_i)$. Also, we know (Theorem 2.6 of Seneta's book*) that the eigenvalue $-\alpha_k$ of $Q_k$ with *maximal real part* is unique, simple (multiplicity $1$), and strictly negative.

*Seneta, E. (1981) Non-negative Matrices and Markov Chains. Revised Edition. Springer, New York.

Hence, $-\alpha$, where $\alpha = \min_k \alpha_k > 0$, is the (possibly degenerate) eigenvalue of $Q$ with maximal real part.

# Decay parameters

Eigenvalues. Clearly $\mathrm{Sp}(Q) = \cup_i \mathrm{Sp}(Q_i)$. Also, we know (Theorem 2.6 of Seneta's book*) that the eigenvalue $-\alpha_k$ of $Q_k$ with *maximal real part* is unique, simple (multiplicity $1$), and strictly negative.

*Seneta, E. (1981) Non-negative Matrices and Markov Chains. Revised Edition. Springer, New York.

Hence, $-\alpha$, where $\alpha = \min_k \alpha_k > 0$, is the (possibly degenerate) eigenvalue of $Q$ with maximal real part.

Note that the $\alpha_k$ and $\alpha$ are *decay parameters*:

$$P_{ij}(t) \le C_{ij} e^{-\alpha_k t} \le C_{ij} e^{-\alpha t}, \qquad i, j \in S_k.$$

# Example: two competing species

Two species $A$ and $B$ affect one another's ability to survive on a habitat patch.

# Example: two competing species

Two species $A$ and $B$ affect one another's ability to survive on a habitat patch. *Which has survived given that the patch has been inhabited for a long time?*

# Example: two competing species

Two species $A$ and $B$ affect one another's ability to survive on a habitat patch. *Which has survived given that the patch has been inhabited for a long time?*

**State at time $t$**: $X(t) = (X_A(t), X_B(t))$, where $X_A(t)$ and $X_B(t)$ are the numbers of $A$ and $B$.

# Example: two competing species

Two species $A$ and $B$ affect one another's ability to survive on a habitat patch. *Which has survived given that the patch has been inhabited for a long time?*

State at time $t$: $X(t) = (X_A(t), X_B(t))$, where $X_A(t)$ and $X_B(t)$ are the numbers of $A$ and $B$.

Extinction state: $\mathbf{0} = (0, 0)$.

# Example: two competing species

Two species $A$ and $B$ affect one another's ability to survive on a habitat patch. *Which has survived given that the patch has been inhabited for a long time?*

State at time $\boldsymbol{t}$: $X(t) = (X_A(t), X_B(t))$, where $X_A(t)$ and $X_B(t)$ are the numbers of $A$ and $B$.

Extinction state: $\mathbf{0} = (0, 0)$.

$\boldsymbol{S}$ is not irreducible: Let $S_{AB}$, $S_A$ and $S_B$ be the communicating classes corresponding to the presence of both species, just $A$, and just $B$, respectively.

# Example: two competing species

Two species $A$ and $B$ affect one another's ability to survive on a habitat patch. *Which has survived given that the patch has been inhabited for a long time?*

**State at time $t$**: $X(t) = (X_A(t), X_B(t))$, where $X_A(t)$ and $X_B(t)$ are the numbers of $A$ and $B$.

**Extinction state**: $\mathbf{0} = (0, 0)$.

**$S$ is not irreducible**: Let $S_{AB}$, $S_A$ and $S_B$ be the communicating classes corresponding to the presence of both species, just $A$, and just $B$, respectively.

**Partial ordering**: $\{\mathbf{0}\} \prec S_A \prec S_{AB}$ and $\{\mathbf{0}\} \prec S_B \prec S_{AB}$.

# Example: two competing species

Transition rates:

$$\begin{pmatrix} 0 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \boldsymbol{a}_A^\top & Q_A & \mathsf{O} & \mathsf{O} \\ \boldsymbol{a}_B^\top & \mathsf{O} & Q_B & \mathsf{O} \\ \mathbf{0}^\top & Q_{AB\to A} & Q_{AB\to B} & Q_{AB} \end{pmatrix}$$

$$Q = \begin{pmatrix} Q_A & \mathsf{O} & \mathsf{O} \\ \mathsf{O} & Q_B & \mathsf{O} \\ Q_{AB\to A} & Q_{AB\to B} & Q_{AB} \end{pmatrix}$$