Ancestral lineages in locally regulated populations

Matthias Birkner

Johannes-Gutenberg-Universität Mainz

based on joint work, in part in progress, with
Jiří Černý, Andrej Depperschmidt and Nina Gantert

EURANDOM, August 2014

General aim

Study/understand the space-time embedding of ancestral lineages in spatial models for populations with local density regulation (in particular, with non-constant local population sizes).

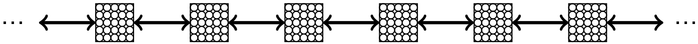
A step towards combining ecological and population genetics aspects in a stochastic spatial population model

Caveat: Most results so far are more of conceptual than practical interest.

Outline

Stepping stone model (in discrete time)

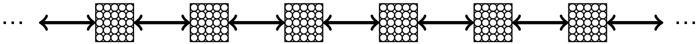
Colonies of *fixed* size N are arranged in a geographical space, say $\mathbb{Z}^d$



($d = 1$ in this picture)

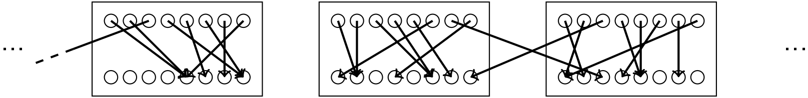
Stepping stone model (in discrete time)

Colonies of *fixed* size N are arranged in a geographical space, say $\mathbb{Z}^d$



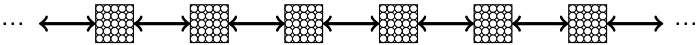
($d = 1$ in this picture)

For each child: Assign a random parent in same colony with probability $1 - \nu$, in a neighbouring colony with probability ν



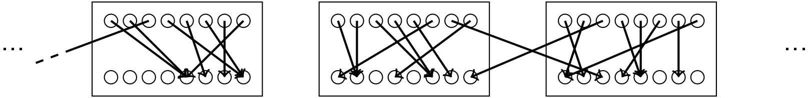
Stepping stone model (in discrete time)

Colonies of *fixed* size N are arranged in a geographical space, say $\mathbb{Z}^d$



($d = 1$ in this picture)

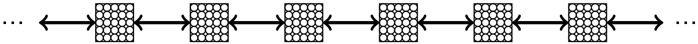
For each child: Assign a random parent in same colony with probability $1-\nu$, in a neighbouring colony with probability ν



More generally, at for each individual in colony x, with probability $p(x,y) = p(y-x)$ assign a random parent in previous generation from colony y

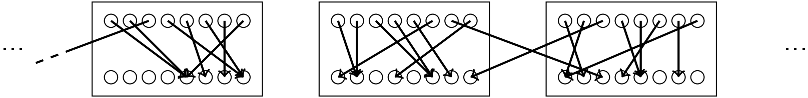
Stepping stone model (in discrete time)

Colonies of *fixed* size N are arranged in a geographical space, say $\mathbb{Z}^d$



($d = 1$ in this picture)

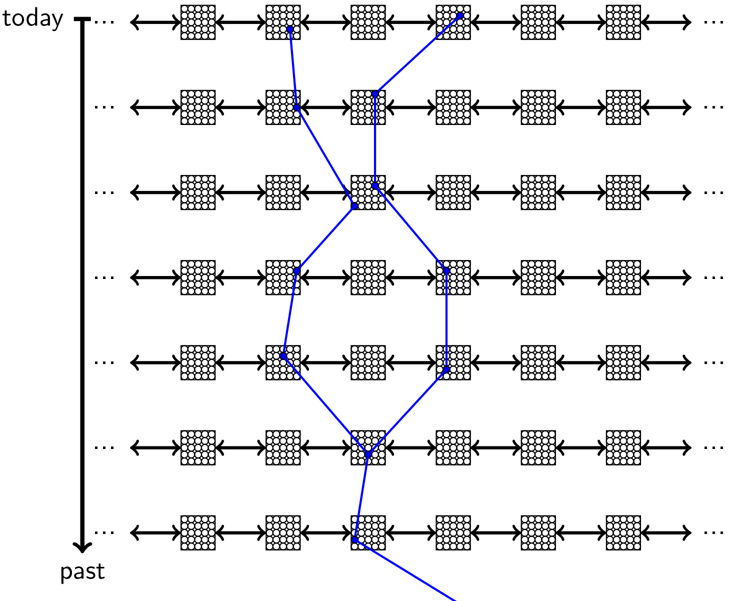
For each child: Assign a random parent in same colony with probability $1 - \nu$, in a neighbouring colony with probability ν



More generally, at for each individual in colony x, with probability $p(x, y) = p(y - x)$ assign a random parent in previous generation from colony y

"Trivial" demographic structure, but paradigm model for evolution of *type distribution* in space

Stepping stone model: Ancestral lines



Stepping stone model: Spatial embedding of ancestral lines

Sample one individual from colony x and one from colony y.

The spatial positions of the ancestral lines are random walks with (delayed) coalescence:

While not yet merged, each takes an independent step according to the random walk transition kernel p,

every time they are in the same colony, the two lines merge with probability $1/N$.

Stepping stone model: Spatial embedding of ancestral lines

Sample one individual from colony x and one from colony y.

The spatial positions of the ancestral lines are random walks with (delayed) coalescence:

While not yet merged, each takes an independent step according to the random walk transition kernel p,

every time they are in the same colony, the two lines merge with probability $1/N$.

Write T_{merge} for the number of steps until the two ancestral lines/walkers merge.

Ancestral lines are coalescing random walks

An Application:

Assume each child mutates with probability u (to a completely new type), let

$$\psi(x, y) := \text{probability in equilibrium that two individuals randomly drawn from colonies } x \text{ and } y \text{ have same type}$$

Ancestral lines are coalescing random walks

An Application:

Assume each child mutates with probability u (to a completely new type), let

$$\psi(x,y) := \text{probability in equilibrium that two individuals randomly drawn from colonies } x \text{ and } y \text{ have same type}$$

Then (assuming p is symmetric)

$$\psi(x,y) = \mathbb{E}_{x,y}\left[(1-u)^{2T_{\text{merge}}}\right] = \frac{G_u(x,y)}{N + G_u(0,0)}$$

(where p_k is the k-step transition kernel) with

$$G_u(x,y) = \sum_{k=1}^{\infty}(1-u)^{2k} p_{2k}(x,y).$$

The stepping stone model (Kimura, 1953)

A very popular model for spatial population genetics :

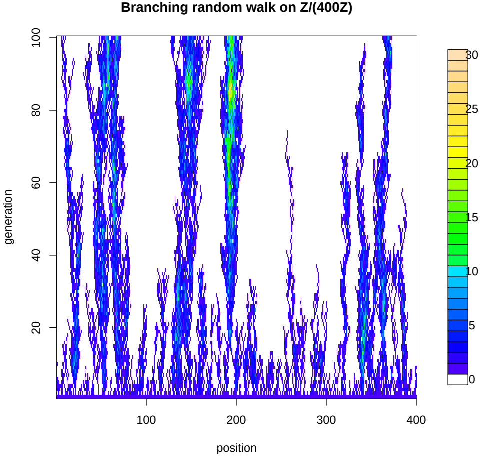
Fixed local population size N in each patch (arranged on $\mathbb{Z}^d$), patches connected by (random walk-type) migration

Pros:
- $+$ Stable population, no local extinction, nor unbounded growth
- $+$ Ancestral lineages are (delayed) coalescing random walks, this makes detailed analysis feasible, in particular via duality: long-time behaviour of (neutral) type distribution

Cons:
- $-$ An 'ad hoc' simplification, effects of local size fluctations no longer explicitly modelled
- $-$ N is an 'effective' parameter, relation to 'real' population dynamics is unclear

Remark: A problem with branching random walk

(Critical) branching random walks, where particles move and produce offspring independently, explicitly model fluctuations in local population size, but do not allow stable populations in $d \leq 2$:



(Felsenstein's "pain in the torus" 1975; Kallenberg 1977)

Branching random walk with local density-dependent feedback

- Possible and natural extension of the stepping stone model (and of branching random walks)
- Offspring distribution supercritical when there are few neighbours, subcritical when there are many neighbours

e.g. Bolker & Pacala (1997), Murrell & Law (2003), Etheridge (2004), Fournier & Méléard (2004), Hutzenthaler & Wakolbinger (2007) Blath, Etheridge & Meredith (2007), B. & Depperschmidt (2007), Pardoux & Wakolbinger (2011), Le, Pardoux & Wakolbinger (2013), ...

Branching random walk with local density-dependent feedback

- Possible and natural extension of the stepping stone model (and of branching random walks)
- Offspring distribution supercritical when there are few neighbours, subcritical when there are many neighbours

e.g. Bolker & Pacala (1997), Murrell & Law (2003), Etheridge (2004), Fournier & Méléard (2004), Hutzenthaler & Wakolbinger (2007) Blath, Etheridge & Meredith (2007), B. & Depperschmidt (2007), Pardoux & Wakolbinger (2011), Le, Pardoux & Wakolbinger (2013), ...

Challenges:

- Mathematical analysis harder (population sizes are now a space-time random field; feedback mechanism makes different families dependent)
- Dynamics of ancestral lineages?

A spatial logistic model

Particles 'live' in $\mathbb{Z}^d$ in discrete generations,
$\eta_n(x) = \#$ particles at $x \in \mathbb{Z}^d$ in generation n.

Given η_n,

each particle at x has Poisson$\big(\big(m - \sum_z \lambda_{z-x}\eta_n(z)\big)^+\big)$ offspring,
$m > 1$, $\lambda_z \geq 0$, $\lambda_0 > 0$, symmetric, finite range.

(Interpretation as local competition:
Ind. at z reduces average reproductive success of focal ind. at x by λ_{z-x})

A spatial logistic model

Particles 'live' in $\mathbb{Z}^d$ in discrete generations,
$\eta_n(x) = \#$ particles at $x \in \mathbb{Z}^d$ in generation n.

Given η_n,

each particle at x has Poisson$\big((m - \sum_z \lambda_{z-x}\eta_n(z))^+\big)$ offspring,
$m > 1$, $\lambda_z \geq 0$, $\lambda_0 > 0$, symmetric, finite range.

(Interpretation as local competition:
Ind. at z reduces average reproductive success of focal ind. at x by λ_{z-x})

Children take an independent random walk step to y with probability p_{y-x},
$p_{xy} = p_{y-x}$ symmetric, aperiodic finite range random walk kernel on $\mathbb{Z}^d$.

A spatial logistic model

Particles 'live' in $\mathbb{Z}^d$ in discrete generations,
$\eta_n(x) = \#$ particles at $x \in \mathbb{Z}^d$ in generation n.

Given η_n,
each particle at x has Poisson$\big(\big(m - \sum_z \lambda_{z-x}\eta_n(z)\big)^+\big)$ offspring,
$m > 1$, $\lambda_z \geq 0$, $\lambda_0 > 0$, symmetric, finite range.
(Interpretation as local competition:
Ind. at z reduces average reproductive success of focal ind. at x by λ_{z-x})
Children take an independent random walk step to y with probability p_{y-x},
$p_{xy} = p_{y-x}$ symmetric, aperiodic finite range random walk kernel on $\mathbb{Z}^d$.

Given η_n,

$$\eta_{n+1}(y) \sim \mathrm{Poi}\Big(\sum_x p_{y-x}\eta_n(x)\big(m - \sum_z \lambda_{z-x}\eta_n(z)\big)^+\Big), \quad \text{independent}$$

Remarks

$$\eta_{n+1}(y) \sim \mathrm{Poi}\Big(\sum_x p_{y-x}\eta_n(x)\big(m - \sum_z \lambda_{z-x}\eta_n(z)\big)^+ \Big), \quad \text{independent}$$

- For $\lambda \equiv 0$, (η_n) is a branching random walk.
- (η_n) is a spatial population model with local density-dependent feedback:
  Offspring distribution supercritical when there are few neighbours, subcritical when there are many neighbours
- System is in general *not* attractive.
- Conditioning[1] on $\eta_n(\cdot) \equiv N$ for some $N \in \mathbb{N}$ ("effective local population size") yields a discrete version of the stepping stone model

[1] and considering types and/or ancestral relationships

Remarks, 2

- Poisson offspring distribution is a somewhat artificial (though technically very convenient) choice, one could take any family $\nu(a) \in \mathcal{M}_1(\mathbb{Z}_+)$ parametrised by

$$a = \sum_k k\nu_k(a) \quad \text{satisfying} \quad \sum_k (k-a)^2 \nu_k(a) \le \text{Const.} \times a$$

- Logistic term $x(1-x)$ could be replaced by another suitable function $h(x)$, e.g. $h(x) = x\exp(a - bx)$.
- We have little "explicit" information on the system, e.g. no closed formulas for means, variances/covariances, etc.
- Related continuous-mass models (Etheridge 2004, Blath et al 2007) can be obtained as scaling limit

Remark: Rescaling

$(p_{y-x})_{x,y\in\mathbb{Z}^d}$ a translation invariant stochastic kernel on $\mathbb{Z}^d$, $a>0$, $M>0$, $(\lambda_{y-x})_{x,y\in\mathbb{Z}^d}$ a translation invariant positive kernel.

N-th system has parameters $m^{(N)}=1+M/N$, $p^{(N)}_{xy}=\frac{a}{N}p_{xy}+(1-a/N)\delta_{xy}$, $\lambda^{(N)}_{xy}=\lambda_{xy}/N^2$, starts from

$$\eta_0^{(N)}(x)=[N\mu(x)] \qquad (\mu \text{ is a finite measure}).$$

Then

$$X_t^{(N)}(x):=\frac{1}{N}\eta^{(N)}_{[Nt]}(x), \quad t\geq 0,\ x\in\mathbb{Z}^d.$$

converges in distribution on $D_{[0,\infty)}(\mathcal{M}_f(\mathbb{Z}^d))$ to X, solution of

$$\begin{aligned} dX_t(x) &= a\int_0^t \sum_y p_{x-y}\big(X_t(y)-X_t(x)\big)dt \\ &\quad + X_t(x)\big(M-\sum_z \lambda_{x-z}X_t(z)\big)dt+\sqrt{X_t(x)}dW_t(x) \end{aligned}$$

Survival and complete convergence

Theorem (B. & Depperschmidt, 2007)

Assume $m \in (1,3)$, $0 < \lambda_0 \ll 1$, $\lambda_z \ll \lambda_0$ for $z \neq 0$.

(η_n) survives for all time globally and locally with positive probability for any non-trivial initial condition η_0. Given survival, η_n converges in distribution to its unique non-trivial equilibrium.

Survival and complete convergence

Theorem (B. & Depperschmidt, 2007)

Assume $m \in (1, 3)$, $0 < \lambda_0 \ll 1$, $\lambda_z \ll \lambda_0$ for $z \neq 0$.

(η_n) survives for all time globally and locally with positive probability for any non-trivial initial condition η_0. Given survival, η_n converges in distribution to its unique non-trivial equilibrium.

Proof uses

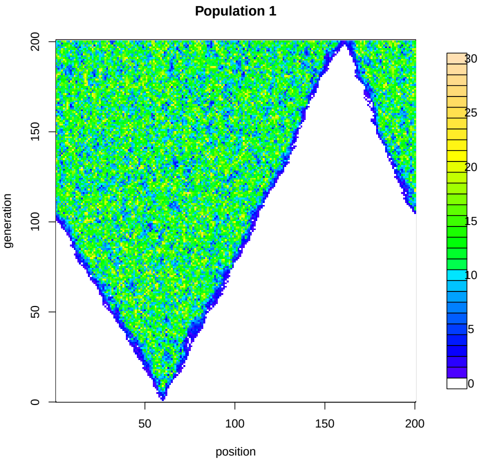
- corresponding deterministic system

$$\zeta_{n+1}(y) = \sum_x p_{y-x}\zeta_n(x)\Big(m - \sum_z \lambda_{z-x}\zeta_n(z)\Big)^+$$

  has unique (and globally attracting) non-triv. fixed point
- strong coupling properties of η
- coarse-graining and comparison with directed percolation

Restriction $m < 3$ is "inherited" from logistic iteration $w_{n+1} = mw_n(1 - w_n)$.

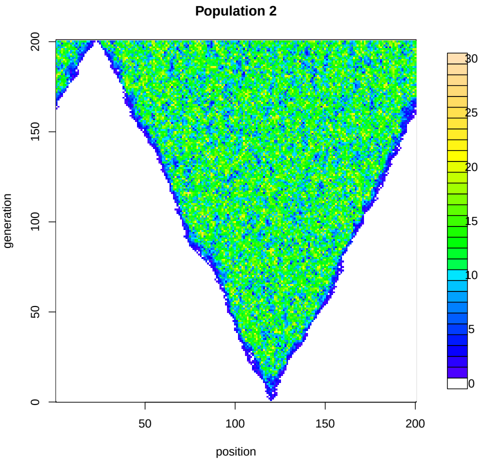
Coupling: An essential proof ingredient



$m = 1.5,\ p = (1/3, 1/3, 1/3),\ \lambda = (0.01, 0.02, 0.01)$

Starting from any two initial conditions η_0, η'_0, copies (η_n), (η'_n) can be coupled such that if both survive, $\eta_n(x) = \eta'_n(x)$ in a space-time cone.

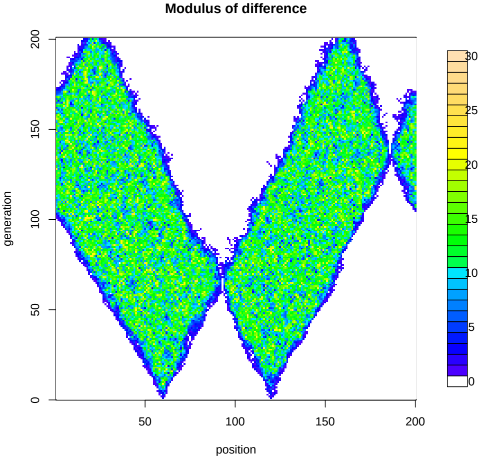
Coupling: An essential proof ingredient



$m = 1.5,\ p = (1/3, 1/3, 1/3),\ \lambda = (0.01, 0.02, 0.01)$

Starting from any two initial conditions η_0, η_0', copies (η_n), (η_n') can be coupled such that if both survive, $\eta_n(x) = \eta_n'(x)$ in a space-time cone.

Coupling: An essential proof ingredient



$m = 1.5,\ p = (1/3, 1/3, 1/3),\ \lambda = (0.01, 0.02, 0.01)$

Starting from any two initial conditions η_0, η'_0, copies (η_n), (η'_n) can be coupled such that if both survive, $\eta_n(x) = \eta'_n(x)$ in a space-time cone.

A stronger form of coupling: "Flow version"

Given $\eta_n = (\eta_n(x))_{x \in \mathbb{Z}^d} \in \mathbb{Z}_+^{\mathbb{Z}^d}$, obtain η_{n+1} (again with values in $\mathbb{Z}_+^{\mathbb{Z}^d}$) via

$$\eta_{n+1}(y) = N_n^{(y)}\Big(\sum_x p_{y-x} \eta_n(x) \big(m - \sum_z \lambda_{z-x} \eta_n(z)\big)^+ \Big) \quad y \in \mathbb{Z}^d$$

with $N_n^{(y)}$, $y \in \mathbb{Z}^d$, $n \in \mathbb{Z}_+$ independent standard Poisson processes

This defines η_n simultaneously for *all* initial conditions $\eta_0 \in \mathbb{Z}_+^{\mathbb{Z}^d}$, write $\Phi_n(\eta_0)$ for conf. at time n starting from η_0

A stronger form of coupling: "Flow version"

Given $\eta_n = (\eta_n(x))_{x \in \mathbb{Z}^d} \in \mathbb{Z}_+^{\mathbb{Z}^d}$, obtain η_{n+1} (again with values in $\mathbb{Z}_+^{\mathbb{Z}^d}$) via

$$\eta_{n+1}(y) = N_n^{(y)}\Big(\sum_x p_{y-x}\eta_n(x)\big(m - \textstyle\sum_z \lambda_{z-x}\eta_n(z)\big)^+ \Big) \quad y \in \mathbb{Z}^d$$

with $N_n^{(y)}$, $y \in \mathbb{Z}^d$, $n \in \mathbb{Z}_+$ independent standard Poisson processes

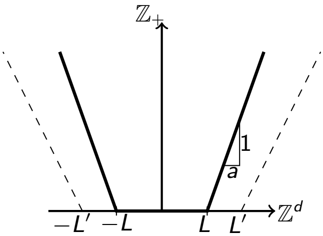
This defines η_n simultaneously for *all* initial conditions $\eta_0 \in \mathbb{Z}_+^{\mathbb{Z}^d}$, write $\Phi_n(\eta_0)$ for conf. at time n starting from η_0

There exist $L' > L$ ($\gg 1$), $a > 0$ and $\mathscr{G} \subset \mathbb{Z}_+^{B_{L'}(\mathbf{0})}$ "good local configurations" such that with very high probability

$\forall\, \eta_0, \eta_0'$:

$$\eta_0|_{B_{L'}(\mathbf{0})} = \eta_0'|_{B_{L'}(\mathbf{0})} \in \mathscr{G}$$

$$\Rightarrow \Phi_n(\eta_0)(x) = \Phi_n(\eta_0)(x') \;\forall (x, n) \in \mathrm{cone}(a, L)$$



and $\mathscr{G}$ has very high prob. under η^{stat}

Dynamics of an ancestral line

Given stationary $(\eta_n^{\mathrm{stat}}(x), n \in \mathbb{Z}, x \in \mathbb{Z}^d)$, cond. on $\eta_0^{\mathrm{stat}}(\mathbf{0}) > 0$ (and "enrich" suitably to allow bookkeeping of genealogical relationships), sample an individual from space-time origin $(\mathbf{0}, 0)$ (uniformly)

Let X_n = position of her ancestor n generations ago:

Given η^{stat} and $X_n = x$, $X_{n+1} = y$ w. prob.

$$\frac{p_{x-y}\eta_{-n-1}^{\mathrm{stat}}(y)\big(m - \sum_z \lambda_{z-y}\eta_{-n-1}^{\mathrm{stat}}(z)\big)^+}{\sum_{y'} p_{x-y'}\eta_{-n-1}^{\mathrm{stat}}(y')\big(m - \sum_z \lambda_{z-y'}\eta_{-n-1}^{\mathrm{stat}}(z)\big)^+}$$

Dynamics of an ancestral line

Given stationary $(\eta_n^{\mathrm{stat}}(x), n \in \mathbb{Z}, x \in \mathbb{Z}^d)$, cond. on $\eta_0^{\mathrm{stat}}(\mathbf{0}) > 0$ (and "enrich" suitably to allow bookkeeping of genealogical relationships), sample an individual from space-time origin $(\mathbf{0}, 0)$ (uniformly)

Let X_n = position of her ancestor n generations ago:

Given η^{stat} and $X_n = x$, $X_{n+1} = y$ w. prob.

$$\frac{p_{x-y}\eta_{-n-1}^{\mathrm{stat}}(y)\big(m - \sum_z \lambda_{z-y}\eta_{-n-1}^{\mathrm{stat}}(z)\big)^+}{\sum_{y'} p_{x-y'}\eta_{-n-1}^{\mathrm{stat}}(y')\big(m - \sum_z \lambda_{z-y'}\eta_{-n-1}^{\mathrm{stat}}(z)\big)^+}$$

Question:
(X_n) is a random walk in a – relatively complicated – random environment. Is it similar to an ordinary random walk when viewed over large enough space-time scales?

Dynamics of an ancestral line

$$\mathbb{P}\big(X_{n+1} = y \,\big|\, X_n = x, \eta^{\mathrm{stat}}\big) = \frac{p_{x-y}\eta^{\mathrm{stat}}_{-n-1}(y)\big(m - \sum_z \lambda_{z-y}\eta^{\mathrm{stat}}_{-n-1}(z)\big)^+}{\sum_{y'} p_{x-y'}\eta^{\mathrm{stat}}_{-n-1}(y')\big(m - \sum_z \lambda_{z-y'}\eta^{\mathrm{stat}}_{-n-1}(z)\big)^+}$$

Remarks

- Analysis of random walks in random environments (also in dynamic random environments) is today a major industry.
  Yet as far as we know, none of the general techniques developed so far in this context is applicable.

  In particular: The natural "forwards" time direction for the walk is "backwards" time for the environment.
- Observation: (X_n) is close to ordinary rw in regions where relative variation of $\eta_{-n-1}(x)$ is small.

Large scale dynamics of an ancestral line

X_n = position of ancestor n generations ago of an individual sampled today at origin in equilibrium

Theorem: LLN and (averaged) CLT

If $m \in (1,3)$, $0 < \lambda_0 \ll 1$, $\lambda_z \ll \lambda_0$ for $z \neq 0$,

$$\mathbb{P}\Big(\frac{1}{n}X_n \to 0 \,\Big|\, \eta_0(0) \neq 0\Big) = 1 \text{ and } \mathbb{E}\Big[f\big(\tfrac{1}{\sqrt{n}}X_n\big) \,\Big|\, \eta_0(0) \neq 0\Big] \underset{n\to\infty}{\longrightarrow} \mathbb{E}\big[f(Z)\big]$$

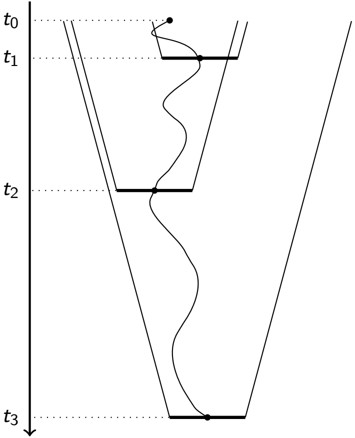
for $f \in C_b(\mathbb{R}^d)$, where Z is a (non-degenerate) d-dimensional normal rv.

The proof uses a *regeneration* construction
(and *coarse-graining* and *coupling*, in particular with directed percolation).

Idea for constructing regeneration times

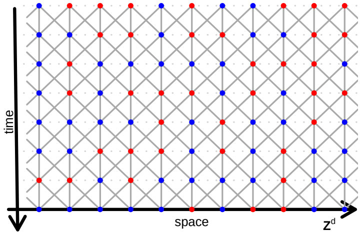
Find time points along the path such that:

- a cone (with fixed suitable base diameter and slope) centred at the current space-time position of the walk covers the path and everything it has explored so far (since the last regeneration)
- configuration η^{stat} at the base of the cone is "good"
- "strong" coupling for η^{stat} occurs inside the cone

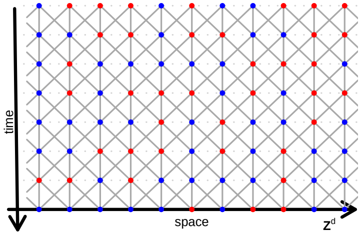


Then, the conditional law of future path increments is completely determined by the configuration η^{stat} at the base of the cone (= a finite window around the current position)

$\omega(x, n)$, $x \in \mathbb{Z}^d$, $n \in \mathbb{Z}$, i.i.d. Bernoulli(p)
Interpretation: $\omega(x, n) = 1$: (x, n) is open, otherwise closed



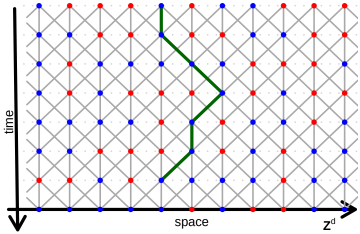
$\omega(x,n)$, $x \in \mathbb{Z}^d$, $n \in \mathbb{Z}$, i.i.d. Bernoulli(p)
Interpretation: $\omega(x,n) = 1$: (x,n) is open, otherwise closed



Open paths:

$n < m$, $x, y \in \mathbb{Z}^d$: $(x,n) \to_\omega (y,m)$ if there exist $x = x_0, x_1, \dots, x_{m-n} = y$ such that $\|x_i - x_{i-1}\| \le 1$ and $\omega(x_i, n+i) = 1$ for $i = 0, \dots, m-n$

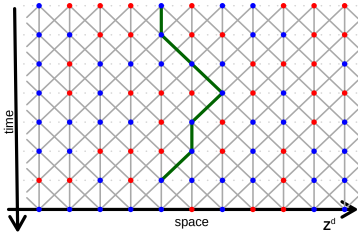
$\omega(x,n)$, $x \in \mathbb{Z}^d$, $n \in \mathbb{Z}$, i.i.d. Bernoulli(p)
Interpretation: $\omega(x,n) = 1$: (x,n) is open, otherwise closed



Open paths:

$n < m$, $x, y \in \mathbb{Z}^d$: $(x,n) \to_\omega (y,m)$ if there exist $x = x_0, x_1, \dots, x_{m-n} = y$ such that $\|x_i - x_{i-1}\| \le 1$ and $\omega(x_i, n+i) = 1$ for $i = 0, \dots, m-n$

$\omega(x,n)$, $x \in \mathbb{Z}^d$, $n \in \mathbb{Z}$, i.i.d. Bernoulli(p)
Interpretation: $\omega(x,n) = 1$: (x,n) is open, otherwise closed



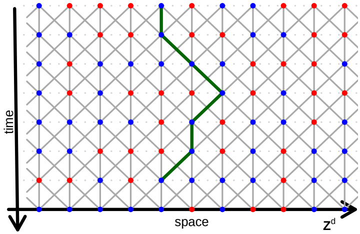
Open paths:

$n < m$, $x, y \in \mathbb{Z}^d$: $(x,n) \to_\omega (y,m)$ if there exist $x = x_0, x_1, \ldots, x_{m-n} = y$ such that $\|x_i - x_{i-1}\| \leq 1$ and $\omega(x_i, n+i) = 1$ for $i = 0, \ldots, m-n$,

$\mathcal{C}_{(x,n)} := \{(y,m) : y \in \mathbb{Z}^d, m \geq n, (x,n) \to_\omega (y,m)\}$ is the (directed) cluster of (x,n)

Write $\xi(x,n) := \mathbf{1}(\#\mathcal{C}_{(x,n)} = \infty)$, i.e. $\xi(x,n) = 1 \iff (x,n) \to_\omega \infty$

$\omega(x,n)$, $x \in \mathbb{Z}^d$, $n \in \mathbb{Z}$, i.i.d. Bernoulli(p)
Interpretation: $\omega(x,n) = 1$: (x,n) is open, otherwise closed



Open paths:

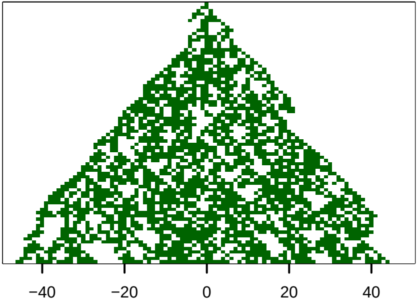
$n < m$, $x, y \in \mathbb{Z}^d$: $(x,n) \to_\omega (y,m)$ if there exist $x = x_0, x_1, \dots, x_{m-n} = y$ such that $\|x_i - x_{i-1}\| \leq 1$ and $\omega(x_i, n+i) = 1$ for $i = 0, \dots, m-n$,

$\mathcal{C}_{(x,n)} := \{(y,m) : y \in \mathbb{Z}^d, m \geq n, (x,n) \to_\omega (y,m)\}$ is the (directed) cluster of (x,n)

Write $\xi(x,n) := \mathbf{1}(\#\mathcal{C}_{(x,n)} = \infty)$, i.e. $\xi(x,n) = 1 \iff (x,n) \to_\omega \infty$

Remark $(\xi(x,-n))_{n \in \mathbb{Z}}$ is the discrete time contact process in its upper invariant measure

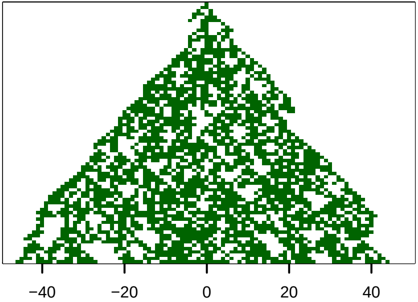
Critical value



There exists $p_c \in (0,1)$ such that

$$\mathbb{P}(|\mathcal{C}_{(\mathbf{0},0)}| = \infty) > 0 \quad \text{iff} \quad p > p_c.$$

Critical value



There exists $p_c \in (0,1)$ such that

$$\mathbb{P}(|\mathcal{C}_{(\mathbf{0},0)}| = \infty) > 0 \quad \text{iff} \quad p > p_c.$$

Theorem (Durrett 1984, "Folklore")
If $p > p_c$, $\mathbb{P}(\mathcal{C}_{(\mathbf{0},0)} \text{ reaches height } n \mid \#\mathcal{C}_{(\mathbf{0},0)} < \infty) \leq Ce^{-cn}$ for some $c, C \in (0,\infty)$.

Auxiliary model: Definitions

$K_n^\omega(x,y)$ probability kernels on $\mathbb{Z}^d$, finite range (say 1),

- $K_n^\omega(\cdot,\cdot)$ compatible with $\mathbb{Z}^d$-symmetries in distribution
- $K_n^\omega(x,\cdot)$ depends only on $\big(\omega(x+z,n), \xi(x+z,n) : \|z\| \le R\big)$ for some $R \in \mathbb{N}$
  (Recall that ξ's are a [very non-local] function of ω's.)
- $\{\xi(x,n)=1\} \subset \big\{\|K_n^\omega(x,\cdot) - K_{\mathrm{unif}}(x,\cdot)\|_{TV} < \varepsilon_K\big\}$ for some suitably small $\varepsilon_K > 0$, where $K_{\mathrm{unif}}(\cdot,\cdot)$ is the symmetric range 1 random walk kernel

(We will assume p sufficiently close to 1)

Consider (X_n) walk in random environment given by $K_n^\omega(\cdot,\cdot)$, i.e.

$$\mathbb{P}\big(X_{n+1} = y \,\big|\, X_n = x, \omega\big) = K_n^\omega(x,y)$$

Interpretation: $\{\xi(x,n)=1\} \mathrel{\widehat{=}}$ "η^{stat} has small fluctuations in a neighbourhood of (x,n)"

Regeneration times, LLN and CLT

Assume $\varepsilon_K \ll 1$, p sufficiently close to 1

There exist random times $0 = T_0 < T_1 < T_2 < \cdots$ such that with $\tau_n := T_n - T_{n-1}$, $Y_n := X_{T_n} - X_{T_{n-1}}$,

$$(Y_n, \tau_n)_{n \in \mathbb{N}} \quad \text{is an i.i.d. sequence,}$$

$\mathbb{E}\big[\tau_1^b\big] < \infty$, $\mathbb{E}\big[\|Y_1\|^b\big] < \infty$ for some $b > 2$, $\mathbb{E}\big[Y_1\big] = \mathbf{0}$, Y_1 is not concentrated on a subspace

Corollary

(X_n) satisfies the law of large numbers and a central limit theorem with non-trivial variance (when averaging over both ω and the walk).

Localising "negative" information

Let $\ell(x, n)$ = length (in steps) of longest directed open path starting in (x, n), with conventions $\ell(x, n) = -1$ if $\omega(x, n) = 0$ and $\ell(x, n) = \infty$ if $\xi(x, n) = 1$

Put $D_n = n + max\{\ell(y, n) + 2 : \|y - X_n\| \leq R, \ell(y, n) < \infty\}$

(Interpretation: At time D_n any "negative" future information deducible from $\xi(y, n) = 0$ that the path has explored at time n is decided and does not affect the future law any more.)

Localising “negative” information

Let $\ell(x,n)$ = length (in steps) of longest directed open path starting in (x,n), with conventions $\ell(x,n) = -1$ if $\omega(x,n) = 0$ and $\ell(x,n) = \infty$ if $\xi(x,n) = 1$

Put $D_n = n + max\{\ell(y,n) + 2 : \|y - X_n\| \leq R, \ell(y,n) < \infty\}$

(Interpretation: At time D_n any “negative” future information deducible from $\xi(y,n) = 0$ that the path has explored at time n is decided and does not affect the future law any more.)

Put $\sigma_0 := 0$, $\sigma_i := \min\Big\{m > \sigma_{i-1} : \max_{\sigma_{i-1} \leq n \leq m} D_n \leq m\Big\}$,

note: σ_i are stopping times w.r.t. $(\mathcal{F}_n)$ where
$\mathcal{F}_n = \sigma\big(\omega(\cdot,k), k \leq n,\ \xi$ in R-tube around X-path until step $n\big)$

Localising "negative" information

Let $\ell(x,n)$ = length (in steps) of longest directed open path starting in (x,n), with conventions $\ell(x,n) = -1$ if $\omega(x,n) = 0$ and $\ell(x,n) = \infty$ if $\xi(x,n) = 1$

Put $D_n = n + max\{\ell(y,n) + 2 : \|y - X_n\| \leq R, \ell(y,n) < \infty\}$
(Interpretation: At time D_n any "negative" future information deducible from $\xi(y,n) = 0$ that the path has explored at time n is decided and does not affect the future law any more.)

Put $\sigma_0 := 0$, $\sigma_i := \min\Big\{m > \sigma_{i-1} : \max\limits_{\sigma_{i-1} \leq n \leq m} D_n \leq m\Big\}$,
note: σ_i are stopping times w.r.t. $(\mathcal{F}_n)$ where
$\mathcal{F}_n = \sigma\big(\omega(\cdot,k), k \leq n,\ \xi \text{ in } R\text{-tube around } X\text{-path until step } n\big)$

Lemma $\mathbb{P}\big(\sigma_{i+1} - \sigma_i > k \,\big|\, \mathcal{F}_{\sigma_i}\big) \leq Ce^{-ck}$
Pf idea:
$\{\sigma_{i+1} - \sigma_i > k\}$ enforces existence of finite clusters of combined heights $\geq k$, "positive" information about $\xi = 1$ contained in $\mathcal{F}_{\sigma_i}$ is harmless by the FKG inequality for ξ

Dry clusters are small (when p suff. large)

Lemma

For any $V = \{(x_i, t_i) : 1 \le i \le k\} \subset \mathbb{Z}^d \times \mathbb{Z}$ with $t_1 < t_2 < \cdots < t_k$,

$$\mathbb{P}_p\left(\xi(x,t) = 0 \text{ for all } (x,t) \in V\right) \le \varepsilon(p)^k$$

with $\varepsilon(p) \to 0$ when $p \nearrow 1$.

Idea:
$\eta^{\mathrm{stat}} \equiv 0$ on V enforces existence of finite clusters of combined heights $\ge k$

Controlling "negative" information from "outside"

$\widetilde{U} \subset \mathbb{Z}^d$ finite, symmetric
$(\widetilde{\eta}_n)$ with values in $\{-1, 0, 1\}$, dynamics:

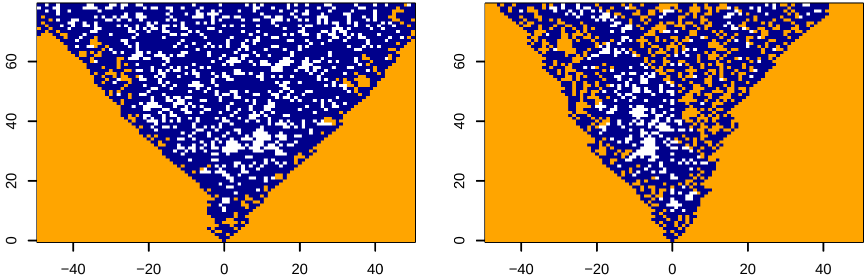
$$\widetilde{\eta}_{n+1}(x) = \omega(x, -n-1)\mathbf{1}_{\{\exists\, y \in U+x \,:\, \widetilde{\eta}_n(y)=1\}}\Big(1 + \mathbf{1}_{\{\exists\, y' \in \widetilde{U}+x \,:\, \widetilde{\eta}_n(y)=-1\}}\Big) - \mathbf{1}_{\{\exists\, y' \in \widetilde{U}+x \,:\, \widetilde{\eta}_n(y)=-1\}},$$

note:

- $\xi(x, -n) = \widetilde{\eta}_n(x) \vee 0$
- a site x that would become 0 in $(\xi(\cdot, -n))_{n\in\mathbb{Z}}$ becomes -1 in $\widetilde{\eta}$ if there was a -1 in the $\widetilde{U}$-neighbourhood
- can interpret this as a (rather particular) two-type contact process

“Everything is caught by the cluster started at $(\mathbf{0}, 0)$”

(no information comes from outside at late times when p suff. large)



$p = 0.77$ $\qquad$ $p = 0.68$

(Here, $U = \{-1, 0, 1\}$, $\widetilde{U} = \{-2, -1, 0, 1, 2\}$)

Lemma (Durrett 1992)

For p suff. close to 1 there exists $s = s(p)$ s.th. on $\{\eta^{\{\mathbf{0}\}}$ survives$\}$,

$$\widetilde{\eta}_n(x) = \eta_n^{\{\mathbf{0}\}}(x) \quad \text{for all} \;\; n \geq N_0, \; \|x\| \leq ns/2$$

(and N_0 has exponential tails).

An a priori bound on the speed

There is $\overline{s} > 0$ (which can be chosen small when $\varepsilon_K \ll 1$, $1 - p \ll 1$)

$$\text{such that} \quad \mathbb{P}\big(\|X_n - X_j\| > \overline{s}(n-j)\big) \le Ce^{-c(n-j)} \text{ for } 0 \le j \le n$$

(using the fact that the path cannot visit too many sites with $\xi = 0$ by finite cluster size bounds and standard large deviations on sum of increments from sites with $\xi = 1$)

An a priori bound on the speed and cone time points

There is $\overline{s} > 0$ (which can be chosen small when $\varepsilon_K \ll 1$, $1 - p \ll 1$)

$$\text{such that} \quad \mathbb{P}\big(\|X_n - X_j\| > \overline{s}(n-j)\big) \le Ce^{-c(n-j)} \text{ for } 0 \le j \le n$$

(using the fact that the path cannot visit too many sites with $\xi = 0$ by finite cluster size bounds and standard large deviations on sum of increments from sites with $\xi = 1$)

This yields for L sufficiently large and $\overline{s} < a$ that any given n is with high probability an (a, L)-cone time point for the (R-tube around the) path, i.e.

$$\|X_n - X_j\| \le L - R + a(n-j), \quad j = 0, 1, \dots, n$$

(The R-tube around the path up to time n is covered by a "cone" with base diameter L, slope a and base point (X_n, n).)

Combining: Constructing regeneration times

To construct T_1: $t_\ell \nearrow \infty$ a deterministic sequence $\subset \mathbb{N}$ with

$$\Theta^{(0,-t_\ell)}\big(\mathrm{cone}(a', L', t_\ell)\big) \subset \Theta^{(x,-t_{\ell+1})}\big(\mathrm{cone}(a, L, t_{\ell+1})\big)$$

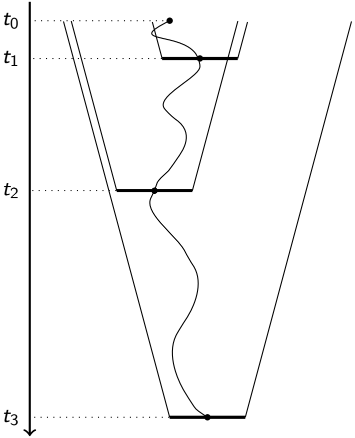
for $\ell \in \mathbb{N}$, $\|x\| \leq 2\bar{s}t_{\ell+1}$.

(This essentially enforces $t_\ell \approx c^\ell$ for a $c > 1$.)

If not previously successful, check at ℓ-th step if

- $\xi \equiv 1$ in L'-window around (X_{t_ℓ}, t_ℓ)
- ℓ-th cone covers previously considered cones and path,
  and successful coupling occurs inside

If yes, $T_1 = t_\ell$, otherwise check at $t_{\ell+1}$, etc.; no. of attempts bounded by geometric RV



Combining: Constructing regeneration times

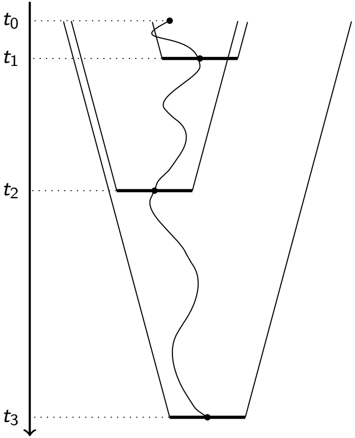
To construct T_1: $t_\ell \nearrow \infty$ a deterministic sequence $\subset \mathbb{N}$ with

$$\Theta^{(0,-t_\ell)}\big(\mathrm{cone}(a', L', t_\ell)\big) \subset \Theta^{(x,-t_{\ell+1})}\big(\mathrm{cone}(a, L, t_{\ell+1})\big)$$

for $\ell \in \mathbb{N}$, $\|x\| \le 2\bar{s}t_{\ell+1}$.
(This essentially enforces $t_\ell \approx c^\ell$ for a $c > 1$.)
If not previously successful, check at ℓ-th step if

- $\xi \equiv 1$ in L'-window around (X_{t_ℓ}, t_ℓ)
- ℓ-th cone covers previously considered cones and path,
  and successful coupling occurs inside

If yes, $T_1 = t_\ell$, otherwise check at $t_{\ell+1}$, etc.;
no. of attempts bounded by geometric RV



Translate to η^{stat} via percolation domination:
$(\xi(x,n) = 1 \leftrightarrow \eta^{\mathrm{stat}}_{-n}(x)$ "good")

Outlook

- Technique is robust (applies to many spatial population models in "high density" regime) but current result "conceptual" rather than practical
- We are hopeful that a "joint regeneration" construction can be implemented to analyse samples of size 2 (or even more) on large space-time scales.
- Meta-theorem: "Everything"[2] that is true for the neutral multi-type voter model is also true for the neutral multi-type spatial logistic model.
- Suitably controlled joint regeneration also allows to derive an a.s. version of the CLT, conditioned on a fixed realisation of η^{stat}.

[2] with a suitable interpretation of "everything".
Examples: Clustering of neutral types in $d = 1, 2$; multiype equilibria exist in $d \geq 3$, $\mathbb{P}\big(\text{two ind. sampled at distance } x \text{ have same type}\big) \sim C\, x^{2-d}$.

Outlook

- In fact, such a "joint regeneration" construction has been carried out for a simplified version of η^{stat}, the discrete time contact process. Then, (X_n) is a directed random walk on the "backbone" of an oriented percolation cluster.
- The diffusion rate $\sigma^2 = \sigma^2(p) = \mathbb{E}\big[Y_{1,1}^2\big]/\mathbb{E}[T_1] \in (0,\infty)$ is not very explicit (though in principle accessible by simulations),

  effective coalescence probability for two lineages still a "black box" (at least to me).

(Some) details can be found in

M. B., A. Depperschmidt, Ann. Appl. Probab. 17 (2007), 1777–1807

M. B., J. Černý, A. Depperschmidt, N. Gantert, Directed random walk on an oriented percolation cluster, Electron. J. Probab. 18 (2013), Article 80

(Some) details can be found in

M. B., A. Depperschmidt, Ann. Appl. Probab. 17 (2007), 1777–1807

M. B., J. Černý, A. Depperschmidt, N. Gantert, Directed random walk on an oriented percolation cluster, Electron. J. Probab. 18 (2013), Article 80

Thank you for your attention!