Minicourse: Multiscale behaviour in selection-mutation systems

D. Dawson, A. Greven

EINDHOVEN 2017

March 29-31,2017

Part 2: The multitype case

Goals

Our next task is now:

- extensions to multitype case, in particular say L low fitness types and K fit rare mutants
- extension to *infinite* geographic space.

Exhibit in that context the **wave of advance**.

The scenario of successive invasions.

The scenario of the transitions of populations concentrated on types $E_j, j \in \{0, 1, \cdots\}$ in infinite geographic space is occurring in steps labelled with j as follows. For each j there is a time scale such that

$$E_0 \to E_1 \to E_2 \cdots E_j \to E_{j+1} \to \cdots \tag{0.1}$$

due to rare mutations occurring somewhere within balls with center 0

$$B_1(0), B_2(0), \cdots \text{ of size } 1, 2, \cdots , \tag{0.2}$$

which can then expand via selection and migration in this respective ball to reach fixation.

- The first stage consists of one phase called phase 0 and involves in the $j-$th step two features.

This is the time scale in which selection and mutation approach a *quasi-equilibrium* on types in E_j and in which types from E_{j+1},i.e. of higher fitness, have not yet appeared with positive probability at a tagged site or in $B_j(0)$.

The second feature in phase zero is that at a sparse subset of the space rare mutation events do occur with positive probability in $B_{j+1}(0)$ in the $j-$time scale, the so called *droplet formation*.

- The second stage of transition consists of three phases.

- In *phase* 1 which occurs in a slightly larger time scale types in E_{j+1} arises through (rare) mutation at rare sites within a ball of size $j+1$ forming a *growing droplet* spreads throughout the population to reach finally a size of the droplet with positive spatial intensity.
- The next *phase 2* in the $j-$th step which occurs after we had an emergence the individuals of type $j+1$ take over all but negligible part of the population which we call *fixation* in the $j-$th step.
- Then in the $j-$th step continues in a subsequent *phase 3* where the population evolves with types in E_{j+1} for a very long time on types from E_{j+1} as in a *neutral* population dynamic.

- In the final *third stage* of the $j-$th step we see its *phase 4*, where the higher level types "equilibrate" in a much larger time scale into a mutation-selection *quasi-equilibrium* on the E_{j+1}-types.

Then the process starts over again to move in the $(j+1)$-step from level $j+1$ to level $j+2$, etc.

We describe first the mathematical model, the interacting system of Fleming-Viot diffusions with selection and mutation, which is the basis for our studies of the scenario of cascades of rare mutants emerging and then fixating. This scenario we obtain by specializing the model we describe next.

Model

A population consisting of multitype individuals divided into *colonies* (demes) that are located at sites labelled by a

$$\text{countable group } \Omega \tag{0.3}$$

(modeling geographic space)

and whose *types* (genotypes) belong to a set

$$\text{the type space} \quad \mathbb{K}. \tag{0.4}$$

The state space of a single component (describing frequencies of types) will be

$$\mathcal{P}(\mathbb{K}) = \text{set of probability measures on } \mathbb{K}. \tag{0.5}$$

The set of colonies (sites or components) will be indexed by a set Ω, which is countable.
The state space $\mathcal{X}$ of the system is therefore

$$\mathcal{X} = (\mathcal{P}(\mathbb{K}))^{\Omega}, \tag{0.6}$$

$$X = (x_\xi)_{\xi \in \Omega} \text{ with } x_\xi \in \mathcal{P}(\mathbb{K}). \tag{0.7}$$

We consider a *spatial Fleming-Viot model* describing a population of individuals distributed in

- geographic space (locations)
- carrying a type

which is described as a process

$$(X_t)_{t \geq 0},$$

with values in

$$((\mathcal{P}(\mathbb{K}))^{\Omega}.$$

Dynamics

The dynamic entails

- resampling at rate d
- mutation at rate m
- selection at rate s
- migration at rate c.

The process is defined by a *martingale problem*.

Parameters

$$ca(\cdot,\cdot) \quad , \quad c \in \mathbb{R}^+ \text{ and } a \text{ probability transition kernel on } \Omega \times \Omega. \tag{0.8}$$

Then we can say that c is the migration rate and $a(\xi, \xi')$ the probability that a jump from ξ to ξ' occurs.

In addition to describe *mutation* and *selection* we need two further objects. Let

$$M(\cdot,\cdot) \text{ be a probability transition kernel on } \mathbb{K} \times \mathbb{K}, \tag{0.9}$$

modeling mutation probabilities from one type to another.
Furthermore the *fitness* function on types

$$\chi(\cdot) \text{ a bounded function on } \mathbb{K}, 0 \leq \chi(\cdot) \leq 1, \quad 0 = \min \chi, 1 = \sup \chi. \tag{0.10}$$

Test functions

Polynomials of the form:

$$F(x) = \int f(u_1, \cdots, u_n) x(du_1) \cdots x(du_n) \quad , \quad x \in \mathcal{M}_{\text{fin}}(\mathbb{K}), \qquad (0.11)$$

where $\mathcal{M}_{\text{fin}}(\mathbb{K})$ denotes finite signed measures and

$$f \in C_b(\mathbb{K}^n, \mathbb{R}), \quad n \in \mathbb{N}. \qquad (0.12)$$

Evaluation of a function of *n sampled individuals*.

Operator

Define the linear operator G acting on functions $F \in \mathcal{A} \subseteq C_b^2(\mathcal{P}(\mathbb{K}), \mathbb{R})$, with values in $C_b(\mathcal{P}(\mathbb{K}), \mathbb{R})$, as follows :

$$\begin{aligned} (GF)(x) = \quad & m \int_{\mathbb{K}} \left(\int_{\mathbb{K}} \frac{\partial F(x)}{\partial x}[v](M(u,dv) - \delta_u(dv)) \right) x(du) \\ & + s \int_{\mathbb{K}} \frac{\partial F(x)}{\partial x}[v](\chi(v) - \int_{\mathbb{K}} \chi(u)x(du))x(dv) \\ & + d \int_{\mathbb{K}} \int_{\mathbb{K}} \frac{\partial^2 F(x)}{\partial x^2}[u,v] Q_x(du,dv), \end{aligned} \tag{0.13}$$

$$Q_x(du,dv) = x(du)\delta_u(dv) - x(du)x(dv). \tag{0.14}$$

$$
\begin{aligned}
(LF)(x) \quad = \sum\nolimits_{\xi\in\Omega_N} \Bigg[\quad & c \sum_{\xi'\in\Omega} a(\xi,\xi') \int_{\mathbb{K}} \frac{\partial F(x)}{\partial x_\xi}(u)(x_{\xi'} - x_\xi)(du) \\
& + s \int_{\mathbb{K}} \left\{ \frac{\partial F(x)}{\partial x_\xi}(u) \left(\chi(u) - \int_{\mathbb{K}} \chi(w) x_\xi(dw)\right) \right\} x_\xi(du) \\
& + m \int_{\mathbb{K}} \left\{ \int_{\mathbb{K}} \frac{\partial F(x)}{\partial x_\xi}(v) M(u,dv) - \frac{\partial F(x)}{\partial x_\xi}(u) \right\} x_\xi(du) \\
& + d \int_{\mathbb{K}} \int_{\mathbb{K}} \frac{\partial^2 F(x)}{\partial x_\xi \partial x_\xi}(u,v) Q_{x_\xi}(du,dv) \Bigg],
\end{aligned}
$$

$$
x \in (\mathcal{P}(\mathbb{K}))^{\Omega}.
$$

Definition (Martingale problem)

(a) The law P on a space of E-valued path for a Polish space E, either $D([0,\infty), E)$ or $C([0,\infty), E)$, is a solution to the martingale problem for (L, ν) w.r.t. $\mathcal{A}$ if and only if

$$\left(F(X(t)) - \int_0^t (LF)(X(s))ds \right)_{t \geq 0} \text{ is a martingale under } P \text{ for all } F \in \mathcal{A} \tag{0.15}$$

and

$$\mathcal{L}(X(0)) = \nu. \tag{0.16}$$

The martingale problem is called wellposed, if the finite dimensional distributions of P are uniquely determined by the property (0.15) and (0.16). □

Theorem (Existence and Uniqueness)

Let ν be a probability measure on $(\mathcal{P}(\mathbb{K}))^{\Omega}$ specifying the initial state which is independent of the evolution.
(a) Then the $(L;\nu)$-martingale problem w.r.t. $\mathcal{A}$, on the space $C([0,\infty)(\mathcal{P}(\mathbb{K}))^{\Omega})$, is well-posed. The resulting canonical stochastic process is denoted

$$(X_t)_{t\geq 0}. \tag{0.17}$$

(b) The solution defines a strong Markov process with the Feller property. □

Specification of a specific model

We first introduce a representation for the *hierarchically organized countable set of types*

$$\mathbb{I} = (\mathbb{N}_0 \times \{1, 2, \dots, M\}) \cup (0, 0) \cup (\infty, 1), \tag{0.18}$$

where $\mathbb{N}_0 = \{0\} \cup \mathbb{N}$ and M is an integer satisfying $1 < M < N$.

Then we can write:

$$\mathbb{I} = \bigcup_{j=0}^{\infty} E_j, \tag{0.19}$$

where

$$\begin{aligned} E_j &= \{(j,\ell); \ell = 1, \cdots, M\}, j \in \mathbb{N}, \quad E_0 = \{(0,\ell), \ell = 0, 1, \cdots, M\}, \\ E_\infty &= \{(\infty, 1)\}. \end{aligned} \tag{0.20}$$

Remark

In the discussion below we will fix M and let $N \to \infty$. Next we introduce the state space of the process and the basic parameters of the stochastic evolution.

(α) The state space of a single component (describing frequencies of types) will be

$$\mathcal{P}(\mathbb{I}) = \text{ set of probability measures on } \mathbb{I}. \tag{0.21}$$

The set of colonies (sites or components) will be indexed by a set Ω_N. The *state space* $\mathcal{X}$ of the system is therefore

$$\mathcal{X} = (\mathcal{P}(\mathbb{I}))^{\Omega_N}, \tag{0.22}$$

with the product topology of the weak topology of probability measures on the compact discrete set $\mathbb{I}$. A typical element is written

$$\mathcal{X} = (x_\xi)_{\xi \in \Omega_N} \text{ with } x_\xi \in \mathcal{P}(\mathbb{I}). \tag{0.23}$$

(β) The hierarchical group Ω_N indexing the *colonies* of the geographic space is defined by:

$$\Omega_N = \left\{ \xi = (\xi^i)_{i \in \mathbb{N}_0} \mid \xi^i \in \mathbb{Z}, 0 \leq \xi^i \leq N-1, \exists k_0 : \quad \xi^j = 0 \quad \forall j \geq k_0 \right\}, \tag{0.24}$$

with group operation defined as component-wise addition modulo N.
Note that:

$$\Omega_N = \bigoplus_{i=0}^{\infty} Z_N, \quad Z_N = \{0, \cdots, N-1\} \text{ with addition mod } (N). \tag{0.25}$$

Introduce a metric on Ω_N

$$d(\xi, \xi') = inf\{k \mid \xi^j = (\xi')^j \quad \forall j \geq k\}. \tag{0.26}$$

Below we relate Ω_N and $\mathbb{Z}^2$.

(γ) The transition kernel $a(\cdot,\cdot)$ on $\Omega_N \times \Omega_N$, modeling *migration rates* has some specific properties: $a = a_N$

$$a_N(\xi,\xi') = a_N(0,\xi'-\xi) \quad \forall \xi,\xi' \in \Omega_N, \tag{0.27}$$
$$a_N(0,\xi) = \sum_{k\geq j} \left(\frac{c_{k-1}}{N^{(k-1)}}\right)\frac{1}{N^k} \qquad \text{if } d(0,\xi) = j \geq 1,$$

where

$$c_k > 0 \quad \forall k \in \mathbb{N}, \quad \sum_{k=0}^{\infty} \frac{c_k}{N^k} < \infty \text{ for all } N \geq 2. \tag{0.28}$$

The kernel $a_N(\cdot,\cdot)$ should be thought of as follows.
With rate $c_{k-1}/N^{(k-1)}$ we choose a hierarchical distance k, and then each point within distance at most k is picked with equal probability as the new location.

For $N \to \infty, \Omega_N$ and above random walk is a good approximation of $\mathbb{Z}^2$ as is shown in Dawson, Gorostiza, Wakolbinger [DGW01].

Different choices for the c_k correspond, from the potential theoretic point of view, to a geography analogous to $2, 2-$, respectively $2+$ dimensions.

Example

$c_k = c^k, c > 1-$transient, $0 < c \leq 1-$recurrent

(δ) To describe *mutation* and *selection* we need

$$M(\cdot,\cdot) \text{ be a probability transition kernel on } \mathbb{I} \times \mathbb{I}, \tag{0.29}$$

modeling mutation probabilities from one type to another.

$$\chi(\cdot) \text{ be a bounded function on } \mathbb{I}, 0 \leq \chi(\cdot) \leq 1, \quad 0 = min\chi, 1 = sup\chi, \tag{0.30}$$

modeling relative fitness of the different types.

Restriction to the ball $\{\xi = (\ell, 0, 0, \ldots), \ell \in \{0, \ldots, N\}$.

$$X^N = (X^N_t)_{t\geq 0} = ((x^N_i(j,t), \quad i = 1, \cdots, 2M), \quad j = 1, \cdots, N)_{t\geq 0}. \tag{0.31}$$

We assume the following about the parameters of the dynamics

$$\begin{aligned}
&\text{selection rate is } s \\
&\text{types } M+1, \ldots, 2M \text{ have fitness: } 1 - \frac{e_i}{N}, 0 \leq e_i < 1, i = M+1, \cdots, 2M \\
&\text{types } i = 1, \ldots, M-1 \text{ have fitness: } 0 \leq e_i < 1 - \frac{1}{N} \text{ and} \\
&\text{type } M \text{ has fitness: } 1 - \frac{1}{N},
\end{aligned} \tag{0.32}$$

$$\text{the mutation rates, } m_{ij} \text{ between types } 1, \ldots, M, \text{ are strictly positive,} \tag{0.33}$$

$$\text{the mutation rate from lower type } i \text{ to higher type } j \text{ is } \frac{m_{ij}}{N}. \tag{0.34}$$

Remark

Fitness difference between $(k-1, M)$ and $(k, 1)$ is $\frac{O(1)}{N^k}$ fitness difference between $(k-1, i)$ and $(k-1, j)$ is $\frac{O(1)}{N^{k-1}}$. Deleterious mutation is $0(\frac{1}{N^k})$.

Single site:

$$G = G^{FV} + G^0 + G^{1,N} + G^{2,N} \tag{0.35}$$

Let $\mathbf{x} \in \Delta_{2M-1}, \quad \mathbf{x} = (x_i)_{i=1,\cdots,2M}$. Then we set

$$\begin{aligned} G^0 f(\mathbf{x}) &= \sum_{i=1}^{M} \left(\sum_{j=1}^{M} (m_{ji} x_j - m_{ij} x_i) \right) \frac{\partial f(\mathbf{x})}{\partial x_i} \\ &+ \sum_{i=1}^{M} s x_i \left(e_i - \sum_{k=1}^{M} e_k x_k - \sum_{k=M+1}^{2M} x_k \right) \frac{\partial f(\mathbf{x})}{\partial x_i}, \\ &+ \sum_{i=M+1}^{2M} s x_i \left(1 - \sum_{k=1}^{M} e_k x_k - \sum_{k=M+1}^{2M} x_k \right) \frac{\partial f(\mathbf{x})}{\partial x_i}, \end{aligned} \tag{0.36}$$

$$G^{FV} = \text{genetic drift}$$

Slow dynamics

$$G^{1,N}f(\mathbf{x}) \tag{0.37}$$
$$= \frac{1}{N}\sum_{i=M+1}^{2M}\left[\left(\sum_{j=M+1}^{2M}(m_{ji}x_j - m_{ij}x_i)\right) + sx_i\left(e_i - \sum_{k=M+1}^{2M}e_k x_k\right)\right]\frac{\partial f(\mathbf{x})}{\partial x_i}$$
$$+\frac{1}{N}\sum_{i=1}^{M}\left(\sum_{j=M+1}^{2M}m_{ji}x_j\right)\frac{\partial f(\mathbf{x})}{\partial x_i} - \frac{1}{N}\sum_{i=M+1}^{2M}\left(\sum_{j=1}^{M}m_{ij}x_i\right)\frac{\partial f(\mathbf{x})}{\partial x_i},$$

$$G^{2,N}f(\mathbf{x}) = \frac{1}{N}\sum_{i=M+1}^{2M}\left(\sum_{j=1}^{M}m_{ji}x_j\right)\frac{\partial f(\mathbf{x})}{\partial x_i} - \frac{1}{N}\sum_{j=M+1}^{2M}\left(\sum_{i=1}^{M}m_{ji}x_j\right)\frac{\partial f(\mathbf{x})}{\partial x_j}. \tag{0.38}$$

As a starting point we assume that at time $t = 0$ only the least fit type is present, that is,

$$x_1^N(j, 0) = 1, j = 1, \ldots, N. \tag{0.39}$$

- *Phase* 0 : E_0 *population*: For times $t_N \to \infty, t_N = o(\log N)$ as $N \to \infty$, k tagged sites approach a product equilibrium measure concentrated on $(E_0)^k$.
- *Phase* 0 : E_1 *population*: Microscopic emergence of rare mutants in the form of a mutant droplet formation in times of the same order as above: exponential growth in time and limiting Palm measure

- *Phase 1*: Macroscopic emergence of rare mutants. This involves in time scales $\beta^{-1}\log N$ the emergence of essentially disjoint *islands* consisting of sites containing a single higher level type whose distribution is given by the limiting Palm measure. This requires two steps:
  - McKeanVlasov limit in macroscopic time scale, i.e. at times $\beta^{-1}\log N + t, t \in \mathbb{R}$ the macroscopic time scale as $N \to \infty$ on Δ_{2M-1} and existence and uniqueness of solutions with entrance law at $t = -\infty$
  - Identification of the random entrance law
- *Linking Phase* 0 *and Phase* 1
- *Phase* 2: Fixation in times $\beta^{-1}\log N + t_N, t_n \to \infty$ with $t_N = o(N)$.
- *Phase* 3: Neutral evolution after fixation and equilibration in the E_1-equilibrium.

McKean-Vlasov non-linear limit dynamic

$$G + c\,[E[\mathbf{x}(t)] - \mathbf{x}(t)] \cdot \nabla. \tag{0.40}$$

$$\begin{aligned} Gf(\mathbf{x}) &= \frac{d}{2}\sum_{i,j}^{2M} x_i(\delta_{ij} - x_j)\frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j} + \sum_{i=1}^{M}\left(\sum_{j=1}^{M}(m_{ji}x_j - m_{ij}x_i)\right)\frac{\partial f(\mathbf{x})}{\partial x_i} \\ &+ s\sum_{i=1}^{M} x_i\left(e_i - \sum_{k=1}^{M} e_k x_k - \sum_{k=M+1}^{2M} x_k\right)\frac{\partial f(\mathbf{x})}{\partial x_i} \\ &+ s\sum_{i=M+1}^{2M} x_i\left(1 - \sum_{k=1}^{M} e_k x_k - \sum_{k=M+1}^{2M} x_k\right)\frac{\partial f(\mathbf{x})}{\partial x_i}. \end{aligned} \tag{0.41}$$

Theorem (Ergodic theorem for $(M, M)-$system)

Assume that the rare mutation is absent, i.e.

$$m_{j,j'} = 0 \text{ if } j = 1, \dots, M, j' = M + 1, \dots, 2M, \tag{0.42}$$

the initial configuration satisfies

$$\sum_{i \in E_0} x_i^N(j, 0) = 1, j = 1, \dots, N \tag{0.43}$$

and since this property is preserved under (0.42) *we can consider* $(X_t^N)_{t \geq 0}$ *to be a system restricted to the type space* $\{1, \dots, M\}$ *and we use this point of view below in (a)-(c).*

Theorem (Ergodic theorem for $(M, M)-$system)

(a) *Let $N < \infty$ and consider the exchangeable interacting system*

$$X_t^N = (\mathbf{x}^N(1,t), \dots, \mathbf{x}^N(N,t)) \in (\Delta_{M-1})^N, \quad \text{with } c > 0, d \geq 0 \text{ and } m_{i,j} > 0 \text{ for all } (i,j) \in \{1, \dots, M\}. \tag{0.44}$$

Consider the distribution of ℓ tagged sites $\mathcal{L}[(\mathbf{x}^N(1,t), \dots, \mathbf{x}^N(\ell,t))]$. Then for $\ell \leq N$

$$\mathcal{L}\left[(\mathbf{x}^N(1,t), \dots, \mathbf{x}^N(\ell,t))\right] \Rightarrow \mu_{\text{eq}}^{N,\ell} \in \mathcal{P}((\Delta_{M-1})^\ell) \text{ as } t \to \infty, \tag{0.45}$$

where $\mu_{\text{eq}}^{N,\ell}$ is the $\ell-$dimensional marginal of $P_{\text{eq}}^{0,N}$,

$$\text{the unique invariant measure of the process } X^N, \tag{0.46}$$

which is in particular independent of $X_0^N \in (\Delta_{M-1})^N$.

Theorem (Ergodic theorem for $(M, M)-$system)

(b) *Consider the McKean-Vlasov dynamic $(\mathcal{L}_t)_{t\geq 0}$ (the limit of $N \to \infty$ of the law of a component in (0.44)) corresponding to above set-up on $\mathbb{I} = \{1, \dots, M\}$. Then*

$$\mathcal{L}_t \Rightarrow \mathcal{L}_\infty = \mu^\infty_{\text{eq}}, \text{ as } t \to \infty, \tag{0.47}$$

where $\mu^\infty_{\text{eq}} \in \mathcal{P}(\Delta_{M-1})$ does not depend on $\mathcal{L}_0$ and is the marginal distribution of the unique equilibrium of the McKean Vlasov process given by the martingale problem for the operator G.

(c) *Now consider the sequence of the $\ell-$dimensional marginals of the equilibrium measure $P^{0,N}_{\text{eq}}$, namely $\{\mu^{N,\ell}_{\text{eq}}\}_{N\in\mathbb{N}}$ from Part (a). Then for every $\ell \in \mathbb{N}, (\ell \leq N)$, as $N \to \infty$,*

$$\mu^{N,\ell}_{\text{eq}}(dx_1, \dots, dx_\ell) \Rightarrow \prod_{i=1}^{\ell} (\mu^\infty_{\text{eq}}(dx_i)), \tag{0.48}$$

where μ^∞_{eq} is the stationary distribution given in (0.47).

Theorem (Ergodic theorem for $(M, M)-$system)

If (0.44) *is replaced by*

$$\sum_{i \in E_0} x_i^N(j, 0) = 1 - \varepsilon, j = 1, \ldots, N \tag{0.49}$$

with $0 < \varepsilon < 1$*, then in the limit* $N \to \infty$*, then* $t \to \infty$

$$\mu_t^{N,\ell}(dx_1, \ldots, dx_\ell) \Rightarrow \prod_{i=1}^{\ell} (\mu_{\mathrm{eq}}^{\infty}(dx_i)), \tag{0.50}$$

where $\mu_{\mathrm{eq}}^{\infty}$ *is the unique McKean-Vlasov equilibrium measure on* Δ_{2M-1} *with uniform mean measure on* E_1*.*

Note: Non-ergodicity in the neutral case without mutation.

Theorem (Ergodic theorem for $(M, M)-$system)

(d) *Consider the geographic space Ω_N and the system $X^N(t) := \{x_i^N(\xi, t)\}_{i \in \{1,\dots,M\}, \xi \in \Omega_N}$ with $d > 0$, under the above assumptions on the $\{m(i,j); i,j \in \{1, \cdots, 2M\}\}$ and with $c_j > 0$ for all j. Assume we have a spatially homogeneous and shift ergodic initial condition. In this case*

$$\mathcal{L}[X_N(t)] \Rightarrow \mathcal{L}_{\mathrm{eq}}^{\Omega_N}, \text{ as } t \to \infty, \tag{0.51}$$

where $\mathcal{L}_{\mathrm{eq}}^{\Omega_N} \in \mathcal{P}((\Delta_{M-1})^{\Omega_N})$ is the law of a spatially homogeneous shift-ergodic random field, which is an invariant measure of the evolution and which is unique under all invariant measures which are translation-invariant.

Proposition (The McKean-Vlasov random entrance law from $-\infty$.)

(a) *Given* $\mathbf{A} \in \widetilde{\Delta}^{+}_{M-1} \subseteq \mathcal{M}_f(E_1)$ *of the form*

$$\mathbf{A} = A_{M+1}\delta_{M+1} + \cdots + A_{2M}\delta_{2M}, \tag{0.52}$$

there exists a unique solution $\{u(t) : -\infty < t < \infty\}$ *of the McKean-Vlasov equation and a unique* $\beta > 0$ *such that*

$$\lim_{t\to-\infty} e^{-\beta t} m(u, t, \cdot \cap E_1) = \mathbf{A}. \tag{0.53}$$

(b) *Moreover every deterministic solution of the McKean-Vlasov equation satisfying*

$$\limsup_{t\to-\infty} e^{-\beta t} m(t, E_1) < \infty, \liminf_{t\to-\infty} e^{-\beta t} m(t, E_1) > 0, \tag{0.54}$$

that is in particular $e^{-\beta t}\int \mathbf{x}(E_1)u(t,\mathbf{x})dx$ *is tight, satisfies* (0.53) *for some* $\mathbf{A}$ *of the form given in* (0.52).

Proposition (The McKean-Vlasov random entrance law from $-\infty$.)

(c) *Given any solution u of the McKean-Vlasov equation satisfying the tightness condition* (0.54),

$$\lim_{t\to-\infty} e^{-\beta t} m(u,t,E_1) = \bar{A} \tag{0.55}$$

$$\lim_{t\to-\infty} e^{-\beta t} m(u,t,\{i\}) = A_i, \quad i = M+1,\cdots,2M, \tag{0.56}$$

$$\bar{A} = A_{M+1} + \cdots + A_{2M}, \tag{0.57}$$

$$\lim_{t\to-\infty} m(u,t,\{i\}) = q_i \quad , \quad i = 1,\cdots,M, \tag{0.58}$$

with

$$q_i = \int x_i \mu_{\mathrm{eq}}(d\mathbf{x}). \tag{0.59}$$

In addition with $\widehat{\pi}_A$ denoting the image measure under the projection π_A on the set A:

$$\widehat{\pi}_{E_0} \circ \mathcal{L}_i \underset{t\to-\infty}{\Longrightarrow} \mu_{\mathrm{eq}}. \tag{0.60}$$

Proposition (The McKean-Vlasov random entrance law from $-\infty$.)

(d) *If we have a solution u satisfying* (0.54) *and for which* (0.56) *holds, then the corresponding mean curve* $(m(t,u,\cdot))_{t\in\mathbb{R}}$ *exist and satisfy:*

$$\lim_{t\to\infty} m(u,t,E_1) = 1, \tag{0.61}$$

$$\lim_{t\to\infty} m(u,t,\{i\}) = \frac{A_i}{\sum_{i=M+1}^{2M} A_i}, \text{ for } i = M+1,\cdots,2M. \tag{0.62}$$

Proposition (Convergence and Emergence)

(a) *As* $N \to \infty$ *the* $\mathcal{P}(\Delta_{2M-1})-$*valued empirical processes*

$$\left((\Xi_N^{\log,\beta}(t))\right)_{t\in\mathbb{R}} \tag{0.63}$$

converge (in the sense of weak convergence of laws on path of measure-valued processes) to a random solution $(\mathcal{L}_t)_{t\in\mathbb{R}}$ *of the McKean-Vlasov equation, that is*

$$\mathcal{L}_t = u(t,\mathbf{x})d\mathbf{x}, \tag{0.64}$$

with $u(t,\cdot)$ *solving the McKean-Vlasov equation with corresponding (random) mean process* $m(u,t,\cdot)$.
The limit is random in the sense that the mean process satisfies:

$$\mathrm{Var}[m(u,t_0,E_1)] > 0, \quad \forall t_0 \in \mathbb{R}. \tag{0.65}$$

Proposition (Convergence and Emergence)

(b) *We have the following asymptotics for the mean curve of the limit of* (0.64)

$$\lim_{t \to -\infty} (e^{-\beta t} m(u,t,\cdot)) =^* \mathcal{W}_{M+1}\delta_{M+1} + \cdots +^* \mathcal{W}_{2M}\delta_{2M},$$

in probability (0.66)

with random variables ${}^*\vec{\mathcal{W}} = ({}^*\mathcal{W}_i, i = M+1, \cdots, 2M)$, *satisfying:*

$$\{{}^*\mathcal{W}_i, i = M+1, \cdots, 2M\} \text{ are independent.} \quad (0.67)$$

If we start with the distribution $(\mu_{\mathrm{eq}})^{\otimes N}$ *initially, then furthermore*

$$E[{}^*\mathcal{W}_i] = m_i = \sum_{i \in E_0} \mu_{\mathrm{eq}}^{\infty}(i), m_{i,j}. \quad (0.68)$$

Proposition (Convergence and Emergence)

(c) *Define the hitting time of the higher level types, (here $\bar{x}^N$ is the empirical mean measure):*

$$\bar{T}_\varepsilon^N = \inf_{t \geq 0} (\bar{x}^N(t)[\{M+1, \cdots, 2M\}] \geq \varepsilon). \tag{0.69}$$

We have

$$\mathcal{L}[\bar{T}_\varepsilon^N - (\frac{1}{\beta} \log N)] \underset{N \to \infty}{\Longrightarrow} \nu_\varepsilon, \quad \nu_\varepsilon \in \mathcal{P}((-\infty, \infty)), \tag{0.70}$$

where ν_ε is non-trivial.

Proposition (Fixation)

(a) *The system fixates asymptotically in the emergence frequencies:*

$$\lim_{t\to\infty}\lim_{N\to\infty}\left(\mathcal{L}\left[\int\sum_{i=1}^{M}x_i\Xi_N^{log,\beta}(t)(d\mathbf{x})\right]\right)=\delta_0, \tag{0.71}$$

$$\lim_{t\to\infty}\lim_{N\to\infty}\mathcal{L}\left[\left(\int x_i\Xi_N^{\log,\beta}(t)(d\mathbf{x})\right)_{i=M+1,\cdots,2M}\right]=\mathcal{L}\left[{}^*\vec{\mathcal{W}}/{}^*\mathcal{W}(E_1)\right]. \tag{0.72}$$

(b) *If we choose* $t_N\to\infty \quad t_N=o(\log N)$ *as* $N\to\infty$, *then the above holds in a joint (of time and* N *) limit:*

$$\lim_{N\to\infty}\mathcal{L}\left(\left[\int\sum_{i=1}^{M}x_i\Xi_N^{\log,\beta}(t_N)(d\mathbf{x})\right]\right)=\delta_0, \tag{0.73}$$

$$\lim_{N\to\infty}\mathcal{L}\left[\left(\int x_i\Xi_N^{\log,\beta}(t_N)(d\mathbf{x})\right)_{i=M+1,\cdots,2M}\right]=\mathcal{L}\left[{}^*\vec{\mathcal{W}}/{}^*\vec{\mathcal{W}}(E_1)\right]. \tag{0.74}$$

Step 1: Droplet description

$$\mathbf{x}^N(j,t) = ((x_i^N(j,t)))_{i=1,\cdots,2M}, \quad \mathbf{x}^N(j,t) \in \mathcal{P}(E_0 \cup E_1), \qquad (0.75)$$

denote the proportion of type i at site j at time t. Let now $\{a(j) \in [0,1], j \in \mathbb{N}\}$ be independently randomly chosen points in $[0,1]$ (which are also chosen independently of the further dynamic), that serve as a labels for site $j \in \{1,\cdots,N\}$.

$$\daleth_t^{N,M} = \sum_{i=M+1}^{2M} \sum_{j=1}^{N} x_i^N(j,t)(\delta_{a(j)} \otimes \delta_i), \qquad (0.76)$$

is an *atomic finite measure* on $[0,1] \times \{M+1,\cdots,2M\}$. To describe the droplet we need the abbreviations:

$$\Delta_{2M-1}^{i+} = \{\mathbf{x} \in \Delta_{2M-1} : x(\{i\}) > 0, \mathbf{x}(\{i'\}) = 0, i' \neq i, i', \in \{M+1,\ldots,2M\}\}, \qquad (0.77)$$

$$\Delta_{2M-1}^{E_1+} = \{\mathbf{x} \in \Delta_{2M-1} : \sum_{i=M+1}^{2M} \mathbf{x}(\{i\}) > 0\}.$$

We set:

$$\mathbb{Q}^i_{(x_1,\dots,x_M)} := \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}P_{(1-\varepsilon)(x_1,\dots,x_{M,0,\cdots,0})+\varepsilon\delta_i}, \quad i = M+1,\cdots,2M, \tag{0.78}$$

$$\{\widetilde{\mu}_s\}_{x\geq 0} \text{ be a measurable function from } [0,\infty) \text{ to } \mathcal{P}(E_0), \tag{0.79}$$

$$N^i(ds,da,du,dw) \text{ on } [0,\infty)\times[0,1]\times[0,\infty)\times W_0^{2M}, \tag{0.80}$$

$$ds \;\; da \;\; du \;\; \left(\int_{\Delta_{M-1}} (\mathbb{Q}^i_{x_1,\dots,x_M}(dw))\widetilde{\mu}_s(dx_1,\dots,dx_M)\right). \tag{0.81}$$

Set of possible excursions in $\Delta^{E_1+}_{2M-1}$ starting at time 0:

$$W^{2M}_0 = \{w \in C([0,\infty), \Delta^{+}_{2M-1})), w(0) \in \{\mathbf{x} : x_i = 0; i = M+1, \dots, 2M\},$$
$$\exists \zeta < \infty \text{ with } \sum_{i=M+1}^{2M} w(t,i) = 0 \;\forall\; t \geq \zeta, \quad w(t) \in \Delta^{E_1+}_{2M-1} \text{ for } t \in (0,\zeta)\} \tag{0.82}$$

The Palm entrance law is defined by

$$\widehat{\mathcal{L}}_t(dx_{M+1},\cdots,dx_{2M}) = \frac{x_{M+1}+\cdots+x_{2M}}{m(u,t,E_1)}\mathcal{L}_t(dx_{M+1},\cdots,dx_{2M}). \tag{0.83}$$

Main ingredient to calculate the exponential growth rate, namely $w(s, x_1(s), \cdots, x_{2M}(s))$ is given by

$$w(s, \Delta^{E_1+}_{2M-1}) = \sum_{i=M+1}^{2M} x_i(s) \tag{0.84}$$

and then set:

$$f(s) = \int_{\Delta_{M-1}} \int_{W_0} w(s, \Delta^{E_1+}_{2M-1})(\mathbb{Q}_{(p_1,\ldots,p_M)}(dw))\mu_{\mathrm{eq}}(dp_1, \ldots, dp_M). \tag{0.85}$$

Proposition (A continuous atomic-measure-valued Markov process: limit droplet dynamic)

(a) *Define $q^i(s,a)$ as the non-negative predictable function*

$$q^i(s,a) := \left(\widetilde{m}_i(s) + c\beth^M_{s-}([0,1]\times\{i\})\right), \quad i = M+1,\cdots,2M, \tag{0.86}$$

where

$$\widetilde{m}_j(s) := \sum_{i=1}^{M} \widetilde{\mu}_s(i) m_{ij}, \quad j\in\{M+1,\cdots,2M\}. \tag{0.87}$$

Then the stochastic integral equation

$$\beth^M_t(dv, d\mathbf{x}) = \beth^0_*(t) + \int_0^t \int_{[0,1]} \sum_{i=M+1}^{2M} \int_0^{q^i(s,a)} \int_{W_0} \delta_{w(t-s)}(d\mathbf{x})\delta_a(dv) N^i(ds, da, du, dw), \tag{0.88}$$

has a unique continuous solution, $(\beth^M_t)_{t\geq 0}$.

Proposition (A continuous atomic-measure-valued Markov process: limit droplet dynamic)

(b) $(\daleth_t^M)_{t\geq 0}$ *is a continuous* $\mathcal{M}_a([0,1]\times\{M+1,\cdots,2M\})$ *respectively* $\mathcal{M}_a([0,1]\times\Delta_{M-1}^{+})-$*valued strong Markov process.*

(c) *The process* $(\daleth_t^M)_{t\geq 0}$ *has the following properties:*

- *the mass of each atom follows an excursion from zero, which is obtained at the intensity given by the excursion law* $\mathbb{Q}_{x_1,\dots,x_M}$,
- *new* $\Delta_{M-1}^{i+}-$*valued excursions are produced at time t at rate*

$$\widetilde{m}_i(t)+c\daleth_t^M([0,1]\times\{i\}), \tag{0.89}$$

- *each new* Δ_{M-1}^{+}*-valued excursion for a type* $i\in\{M+1,\cdots,2M\}$ *produces an atom located at a point in* $[0,1]\times\{i\}$*, the first component chosen according to the uniform distribution on* $[0,1]$,
- *at each* t *and* $\delta>0$ *there are at most finitely many atoms* $\{a_i\}\times\{i\}$ *with mass* $\geq\delta$*, for some* $i\in M+1,\dots,2M$,
- $t\to\daleth_t([0,1]\times\{i\})$ *is a.s. continuous,* $i=M+1,\dots,2M$.

Proposition (Exponential growth of droplet total mass, limiting droplet frequencies)

Assume that we start in the lowest type only at every site.

(a) *There exists* $\beta^* > 0$, *the Malthusian parameter, such that*

$$e^{-\beta^* t} E[\gimel_t^M([0,1] \times \{M+1, \dots, 2M\})] \to A \in (0, \infty), \text{ as } t \to \infty. \tag{0.90}$$

The parameter β^* *is given by the unique positive solution* β^* *of the equation*

$$c \int_0^\infty e^{-\beta^* s} \quad f(s) ds = 1. \tag{0.91}$$

We have that:

$$\beta^* = \beta. \tag{0.92}$$

where β *is the macroscopic entrance law growth.*

Proposition

(b) *Furthermore there exist $A_i \in (0, \infty)$ such that as $t \to \infty$,*

$$\mathcal{L}\left[e^{-\beta t}(\daleth_t^M([0,1] \times \{i\}]_{i,M+1,\cdots,2M})\right] \underset{t\to\infty}{\Longrightarrow} \mathcal{L}[\vec{\mathcal{W}}^*]. \tag{0.93}$$

$$e^{-\beta t}\left(E\left[\daleth_t^M([0,1] \times \{i\})\right]_{i=M+1,\cdots,2M}\right) \to (A_i)_{i=M+1,\cdots,2M}. \tag{0.94}$$

(c) *There exist non-degenerate random variables*

$$\vec{\mathcal{W}}^* = (\mathcal{W}_i^*)_{i=M+1,\cdots,2M} \tag{0.95}$$

such that

Proposition (Droplet growth)

Assume that at time $t = 0 : \widehat{x}_1^N(i,0) = 1$ *for* $i = 1, \cdots, N$ *and that* $t_N \to \infty, t_N = o(\log N)$ *as* $N \to \infty$.
Then

$$\mathcal{L}\left[e^{-\beta t_N}\left(\widehat{x}_{M+1}^N(t_N), \ldots, \widehat{x}_{2M}^N(t_N)\right)\right] \tag{0.96}$$

$$= \mathcal{L}\left[e^{-\beta t_N}\left(\beth_{t_N}^{N,m}([0,1]\times\{M+1\}), \ldots, \beth_{t_N}^{N,m}([0,1]\times\{2M\})\right)\right]$$

$$\underset{N\to\infty}{\Longrightarrow} \mathcal{L}\left[\sum_{i=M+1}^{2M} \mathcal{W}_i^* \delta_i\right],$$

where $\mathcal{W}_i^*, \quad i = M+1, \ldots, 2M$ *are non-degenerate* $\mathbb{R}^+$*−valued independent random variables.*

Proposition (Fine structure of advantageous droplet)

(a) *The limiting empirical distribution of emerging types satisfies*

$$\lim_{t\to-\infty} e^{-\beta t} \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} \delta_{(x^N_{M+1}(\frac{\log N}{\beta}+t,i),\dots,x^N_{2M}(\frac{\log N}{\beta}+t,i))} \tag{0.97}$$

$$= \int_0^1 {}^*\mathcal{W}_{M+1} U_{M+1}(\infty, da_{M+1})(\delta_{a_{M+1}(1,0,\dots,0)}) + \cdots +$$

$${}^*\mathcal{W}_{2M} U_{2M}(\infty, da_{2M})(\delta_{a_{2M}(0,\dots,0,1)}).$$

Here $U_i(\infty, dx)$ *is the* stable size distribution *of type* i *excursions. This means that as* $t\to-\infty$, *the number of sites with mass of type* $i\in\{M+1, M+2,\cdots,2M\}$ *in* $(b,1]$, $b>0$, *grows like* $e^{\beta t}{}^*\mathcal{W}_i \cdot U_i(\infty,(b,1])$.

Proposition

(b) *The following limits exist:*

$$\lim_{t\to\infty}\lim_{N\to\infty}\widehat{\mu}_t^{N,i}(\cdot)=\widehat{\mu}_\infty^i(\cdot). \tag{0.98}$$

(c) *Moreover type i Palm measure*

$$\frac{xU_i(\infty,dx)}{\int_0^1 xU_i(\infty,dx)}=\widehat{\mu}_\infty^i(dx)\in\mathcal{P}([0,1]), \tag{0.99}$$

where

$$\widehat{\mu}_\infty^i(dx)=\frac{x\widetilde{\mathbb{Q}}_\infty^i(dx)}{\int x\widetilde{\mathbb{Q}}_\infty^i(dx)}, \tag{0.100}$$

ζ_b and ζ_d denoting birth and extinction time of an excursion:

$$\widetilde{\mathbb{Q}}_\infty^i(dy)=\int_{-\infty}^0\int_{\Delta_{M-1}} e^{\beta s}\widetilde{\mathbb{Q}}_{\vec{x}}(\zeta_b\in ds,\zeta_d>0,w(0)\in dy)$$

$$\mu_{\rm eq}(d\vec{x}). \tag{0.101}$$

Proposition (Exit = Entrance: equality of growth constants)

We have for the initial state starting all in type 1*:*

$$\mathcal{L}[{}^{*}\overrightarrow{\mathcal{W}}] = \mathcal{L}[\overrightarrow{\mathcal{W}}^{*}]. \tag{0.102}$$

The tool to prove the statements for multitype emergence/fixation is

Duality

D. Dawson , L. Gorostiza, A. Wakolbinger: Occupation time fluctuations in branching systems. J. Theoretical Probab. 14 (2001) 729-796.