The Annals of Probability
2023, Vol. 51, No. 6, 2345–2386
https://doi.org/10.1214/23-AOP1654
This research was funded, in whole or in part, by [EPSEC/UKRI, EP/W026899]. A CC BY 4.0 license is applied to this article arising from this submission, in accordance with the grant's open access conditions.

SCALING LIMIT OF THE FLEMING–VIOT MULTICOLOR PROCESS

BY OLIVER TOUGH[a]

Department of Mathematics, University of Bath, [a]*okt24@bath.ac.uk*

We consider the N-particle Fleming–Viot process associated to a normally reflected diffusion with soft catalyst killing. The Fleming–Viot multicolor process is obtained by attaching genetic information to the particles in the Fleming–Viot process. We establish that, after rescaling time by $t \mapsto Nt$, this genetic information converges to the (very different) Fleming–Viot process from population genetics, as $N \to \infty$. An extension is provided to dynamics given by Brownian motion with hard catalyst killing at the boundary of its domain.

1. Introduction and main result. In this paper we study the behaviour of a system of interacting diffusion processes, known as a Fleming–Viot particle system, first introduced by Burzdy, Holyst and March in [16]. We will establish that if one attaches genetic information to the Fleming–Viot particle system and rescales time by $t \mapsto Nt$, this genetic information evolves for large N like the (very different) Fleming–Viot process from population genetics, which we refer to in this article as a *Wright–Fisher process* for the avoidance of confusion. This is our main theorem, Theorem 1.4. We emphasise that, despite sharing the same name, no link had previously been established between the Fleming–Viot particle system (or any similar particle system) and the Wright–Fisher process.

Throughout this paper $(X_t)_{0\leq t<\tau_\partial}$ will be defined to be a diffusion process evolving in the closure $\bar{D}$ of an open, connected, bounded domain $D \subseteq \mathbb{R}^d$, normally reflected at the C^∞ boundary ∂D, and killed at position dependent rate $\kappa(X_t)$ (*soft killing*). That is, prior to the killing time τ_∂, X_t evolves according to the SDE

$$\begin{aligned} &dX_t = b(X_t)\,dt + \sigma(X_s)\,dW_s + \vec{n}(X_t)\,d\xi_t \in \bar{D}, \quad 0 \leq t < \tau_\partial, \\ &\text{with } \mathbb{1}(\tau_\partial > t) + \int_0^t \kappa(X_s)\mathbb{1}(\tau_\partial > s)\,ds \quad \text{being a martingale,} \end{aligned} \tag{1.1}$$

whereby ξ_t is the boundary local time of X_t at ∂D and $\hat{n}(x)$ is the unit interior normal at $x \in \partial D$. A precise definition of such processes is given in Appendix A. We assume throughout that $\kappa \in C^\infty(\mathbb{R}^d; \mathbb{R}_{\geq 0})$ and is strictly positive somewhere on $\bar{D}$. We also assume that $b \in C^\infty(\mathbb{R}^d; \mathbb{R}^d)$ and $\sigma \in C^\infty(\mathbb{R}^d; \mathbb{R}^{d\times m})$ with $\sigma\sigma^T$ uniformly positive definite.

The Fleming–Viot particle system is defined as follows.

DEFINITION 1.1 (Fleming–Viot particle system). The Fleming–Viot particle system $(\vec{X}^N_t)_{t\geq 0}$ consists of $N \geq 2$ particles

$$\vec{X}^N_t = (X^{N,1}_t, \dots, X^{N,N}_t), \quad t \geq 0,$$

evolving independently in the domain $\bar{D}$ according to (1.1). When a particle is killed, we relocate it to the position of a different particle chosen independently and uniformly at random.

Received January 2022; revised July 2023.
MSC2020 subject classifications. Primary 60K35, 60J80; secondary 92D15.
Key words and phrases. Fleming–Viot process (particle system), Fleming–Viot process (population genetics).

In general, it is not clear that the Fleming–Viot particle system is well-posed due to the possibility of infinitely-many jumps in finite time. In the present setting, however, this is not an issue as the killing rate is bounded.

The Fleming–Viot particle system was introduced by Burdzy, Holyst and March [16] in the case of Brownian dynamics with instantaneous killing at the boundary (*hard killing*), where it was shown to provide an approximation method both for the heat equation with Dirichlet boundary conditions and the principal eigenfunction of the Dirichlet Laplacian. The Fleming–Viot particle system with soft killing was considered by Grigorescu in [28]. The Fleming–Viot particle system has been shown to provide a general approximation method for absorbed strong Markov processes by Villemonais [47] and has been shown to provide an approximation method for quasi-stationary distributions (QSDs) in a variety of settings [2, 3, 16, 45]. When a killed Markov process is Feller, quasi-stationary distributions correspond to left eigenmeasures of its infinitesimal generator [38], Proposition 4.

1.1. *The Fleming–Viot multicolor process.* We attach genetic information ("colors") to the Fleming–Viot particle system, resulting in the Fleming–Viot multicolor process, which was introduced by Grigorescu and Kang in [29], Section 5.1. Whereas the colours in the construction of [29], Section 5.1, are assumed to belong to a finite space, the present article develops this by instead assuming the colors belong to a complete, separable metric space. This space is referred to as the "color space" and is denoted by $\mathbb{K}$. The color $\eta_t^i \in \mathbb{K}$ gives the genetic information of the particle X_t^i, for $i = 1, \dots, N$. A precise definition of the Fleming–Viot multicolor process is given by the following.

DEFINITION 1.2 (Fleming–Viot multicolor process). We take $(\mathbb{K}, d)$ to be an arbitrary complete separable metric space, which we call the color space. We define $(\vec{X}_t^N, \vec{\eta}_t^N)_{0\le t<\infty} = \{(X_t^{N,i}, \eta_t^{N,i})_{0\le t<\infty} : i = 1, \dots, N\}$ as follows:

$$
(1.2)\quad \begin{cases}
(i) & \text{Initial condition: } \big((X_0^{N,1}, \eta_0^{N,1}), \dots, (X_0^{N,N}, \eta_0^{N,N})\big) \sim \upsilon^N \in \mathcal{P}\big((\bar{D} \times \mathbb{K})^N\big). \\
(ii) & \text{For } t \in [0,\infty) \text{ and between killing times, the particles } \big(X_t^{N,i}, \eta_t^{N,i}\big) \\
& \text{evolve and are killed independently, according to (1.1) in the first variable,} \\
& \text{and are constant in the second variable.} \\
(iii) & \text{We write } \tau_k^i \text{ for the death times of particle } \big(X^{N,i}, \eta^{N,i}\big)\big(\text{with } \tau_0^i := 0\big). \\
& \text{When particle } \big(X^{N,i}, \eta^{N,i}\big) \text{ is killed at time } \tau_k^i, \text{ it jumps to the location} \\
& \text{of particle } \big(X^{N,j}, \eta^{N,j}\big), \text{ with } j = U_k^i \in \{1, \dots, N\} \setminus \{i\} \text{ chosen} \\
& \text{independently and uniformly at random, at which time we set} \\
& \big(X_{\tau_k^i}^{N,i}, \eta_{\tau_k^i}^{N,i}\big) = \big(X_{\tau_k^i-}^{N,j}, \eta_{\tau_k^i-}^{N,j}\big). \text{ Moreover, we write } \tau_n \text{ for the } n\text{th time} \\
& \text{at which any particle is killed (with } \tau_0 := 0).
\end{cases}
$$

We then define

$$
(1.3)\qquad J_t^N := \frac{1}{N} \sup\{n > 0 : \tau_n \le t\}
$$

to be the number of deaths, up to time t normalised by $\frac{1}{N}$, and define the empirical measures

$$
(1.4)\qquad m_t^N := \frac{1}{N} \sum_{i=1}^{N} \delta_{X_t^{N,i}} \quad \text{and} \quad \chi_t^N := \frac{1}{N} \sum_{i=1}^{N} \delta_{\eta_t^{N,i}}.
$$

We will obtain a scaling limit for the colors as $N \to \infty$, and time is rescaled according to $t \mapsto Nt$. We now describe the scaling limit we will obtain.

1.2. *The Wright–Fisher process.* Given a gene with two neutral alleles, a and A, the SDE

$$dp_t = \sqrt{p_t(1-p_t)}\,dW_t$$

models the evolution of the proportion $p_t \in [0,1]$ of the population carrying the a-allele in a large population. This is the classical Wright–Fisher diffusion. Generalising this to n alleles, the driftless n-Type Wright–Fisher diffusion process of rate $\theta > 0$ takes values in the simplex $\Delta_n := \{p = (p_1, \dots, p_n) \in \mathbb{R}^n_{\geq 0} : \sum_j p_j = 1\}$ and is characterised by the generator

$$L_{\mathrm{WF}} = \frac{1}{2}\theta \sum_{i,j=1}^{n} p_i(\delta_{ij} - p_j)\frac{\partial^2}{\partial p_i \partial p_j}, \qquad \mathcal{D}(L) = C^2(\mathbb{R}^n). \tag{1.5}$$

This was generalised by Fleming and Viot [25] to a probability measure-valued process, which allows for the set of alleles to be infinite. This measure-valued process is typically called a Fleming–Viot process but is referred to as the Wright–Fisher process in the present article to avoid confusion. In particular, letting $\mathbb{K}$ be the (complete, separable) color space, we will consider the Wright–Fisher process on $\mathcal{P}(\mathbb{K})$. This corresponds to the set of possible alleles being $\mathbb{K}$ and will be our scaling limit.

The reader is directed toward [23] for a survey of the Wright–Fisher process due to Ethier and Kurtz. The Wright–Fisher process is defined as a solution of a martingale problem. This typically features additional terms representing mutation, selection and recombination, but we will not need this generality here. There are various possible formulations of this martingale problem, which can be found in [23], Section 3. The formulation we shall employ is given by [23], (3.20) and (3.21). This definition of the Fleming–Viot process as well as its well-posedness (which comes from [23], Theorem 7.1) are given in Appendix D.

As in (1.5), we parametrise the Wright–Fisher process on $\mathcal{P}(\mathbb{K})$ with a rate $\theta > 0$. The following proposition provides intuition for how one may think of the Wright–Fisher process and its relationship to the n-type Wright–Fisher diffusion. This proposition shall be used in the proof of our main theorem and is proven in Appendix D.

PROPOSITION 1.3. *We let $(\nu_t)_{t\geq 0}$ be a Wright–Fisher process on $\mathcal{P}(\mathbb{K})$ of rate $\theta > 0$, as defined in Appendix D. Then for all finite disjoint unions of measurable subsets, $\dot\bigcup_{j=1}^n \mathcal{A}_j = \mathbb{K}$, we have that*

$$\big(\nu_t(\mathcal{A}_1), \dots, \nu_t(\mathcal{A}_n)\big), \quad t \geq 0 \tag{1.6}$$

is an n-type Wright–Fisher diffusion of rate θ.

1.3. *Main result.* We will establish in Appendix A.2 the following. The absorbed process $(X_t)_{0\leq t<\tau_\partial}$ is Feller (there is no distinction between C_0-Feller and C_b-Feller, as $\bar{D}$ is compact). We write L for its infinitesimal generator. Then $(X_t)_{0\leq t<\tau_\partial}$ has a unique QSD, denoted by π, which is a left eigenmeasure of L. We denote the corresponding eigenvalue as $-\lambda < 0$. Furthermore, there exists a positive right eigenfunction $\phi \in \mathcal{D}(L) \cap C^2(\bar{D}; \mathbb{R}_{>0})$, which is both the unique nonnegative right eigenfunction and the unique right eigenfunction of eigenvalue $-\lambda$, up to rescaling. Throughout, we normalise ϕ so that $\langle \pi, \phi\rangle = 1$.

We may, therefore, define the constant

$$\Theta := \frac{2\lambda \|\phi\|^2_{L^2(\pi)}}{\|\phi\|^2_{L^1(\pi)}}. \tag{1.7}$$

We define the tilted empirical measure of the colors, denoted as $(\mathcal{Y}^N_t)_{0\leq t<\infty}$, by

$$\mathcal{Y}^N_t := \frac{\frac{1}{N}\sum_{i=1}^N \phi(X^i_t)\delta_{\eta^i_t}}{\frac{1}{N}\sum_{i=1}^N \phi(X^i_t)} \in \mathcal{P}(\mathbb{K}). \tag{1.8}$$

Whereas consideration of this quantity shall play a crucial role in our proof, for the purposes of our theorem statement its role is to provide the initial condition of our scaling limit. To the authors' knowledge, this process is original. The proof of Theorem 1.4 shall be outlined in Section 1.7, at which point we shall explain the role of $\mathcal{Y}_t^N$ in the proof.

Convergence will be stated in terms of the weak atomic metric on $\mathbb{K}$, denoted as W_a. The space of probability measures on $\mathbb{K}$, equipped with the weak atomic metric, is denoted by $\mathcal{P}_{\mathrm{W}_a}(\mathbb{K})$. This metric was introduced by Ethier and Kurtz [24] in the context of population genetics. Convergence in the weak atomic metric is equivalent to having both weak convergence of measures and convergence of the sizes and locations of the atoms. We provide a definition of the weak atomic metric in Appendix C.2.

Our main theorem is then the following.

THEOREM 1.4. *We take some deterministic initial profile $\nu^0 \in \mathcal{P}(\mathbb{K})$ and fix a Wright–Fisher process on $\mathcal{P}(\mathbb{K})$ of rate Θ and initial condition $\nu_0 = \nu^0$, which we denote as $(\nu_t)_{0\le t<\infty}$. We consider a sequence of Fleming–Viot multicolor processes, denoted by $(((\vec{X}_t^N, \vec{\eta}_t^N))_{0\le t<\infty} : 2 \le N < \infty)$, such that*

$$\mathcal{P}(\mathbb{K}) \ni \mathcal{Y}_0^N \to \nu^0 \in \mathcal{P}(\mathbb{K}) \quad \text{in } \mathrm{W}_a \text{ in probability as } N \to \infty. \tag{1.9}$$

We now rescale time according to $t \mapsto Nt$. Then $(\chi_{Nt}^N)_{t>0}$ converges to $(\nu_t)_{t>0}$ in finite-dimensional distributions in the following sense. We fix arbitrary $n < \infty$ and $\vec{t} = (t^1, \dots, t^n) \in [0,\infty)^n$ such that $t^1 \le \cdots \le t^n$. We consider arbitrary sequences $(\vec{t}^N)_{2\le N<\infty} := ((t_1^N, \dots, t_n^N))_{2\le N\le\infty}$ such that:

1. $t_1^N \le \cdots \le t_n^N$ *for all* $2 \le N < \infty$;
2. $t_i^N \to t_i$ *as* $N \to \infty$ *for all* $1 \le i \le n$;
3. $Nt_n^N \ge \cdots \ge Nt_1^N \to \infty$ *as* $N \to \infty$.

We then have that

$$\big(\chi_{Nt_1^N}^N, \dots, \chi_{Nt_n^N}^N\big) \to (\nu_{t_1}, \dots, \nu_{t_n}) \quad \text{in } \big(\mathcal{P}_{\mathrm{W}_a}(\mathbb{K})\big)^n \text{ in distribution as } N \to \infty. \tag{1.10}$$

REMARK 1.5. If we take constant killing rate $\kappa \equiv 1$ and consider the corresponding Fleming–Viot multicolor process, we recover the classical Moran model. This is well known to converge to the Wright–Fisher process of rate 2 [23], (4.12). On the other hand, we can check that $\Theta = 2$ when $\kappa \equiv 1$.

REMARK 1.6. We observe that, unless ϕ is constant (which only happens if κ is constant on $\bar{D}$), the empirical measures χ_0^N will, in general, not converge to the same limit as the tilted empirical measures $\mathcal{Y}_0^N$. We therefore no longer have (1.10) if we drop the requirement that $Nt_1^N \to \infty$ as $N \to \infty$. This represents the following separation of timescales phenomenon.

We will establish in the proof of Theorem 1.4 that the tilted empirical measure $\mathcal{Y}_t^N$ evolves slowly over an $\mathcal{O}(N)$ timescale, with $(\mathcal{Y}_{Nt}^N)_{t\ge0}$ converging to the Wright–Fisher process. We further establish that the empirical measure χ_t^N converges on a shorter $\mathcal{O}(1)$ timescale to the tilted empirical measure $\mathcal{Y}_t^N$. Theorem 1.4 then follows by combining these two facts.

Therefore, for large N the empirical measure χ_t^N quickly approaches ν^0 over an $\mathcal{O}(1)$ timescale before evolving like the Wright–Fisher process over the longer $\mathcal{O}(N)$ timescale.

1.4. *Background and related results.* A similar separation of timescales has been obtained by Méléard and Tran in [37]. They considered the evolution of traits in a population of individuals, where the individuals give birth (passing on their trait), die in an age-dependent manner and interact with each other through the effect of the common empirical measure of their traits upon their death rates (representing competition for resources). There the age component plays a similar role to spatial position in the present article. They found that the age component converges to a deterministic equilibria (which is dependent upon the traits) on a fast timescale, whilst the trait distribution evolves on a slow timescale, converging to a certain superprocess over the slow timescale as the population converges to infinity.

Aside from obtaining a different limiting process, they also employ a different proof strategy. In their setup individuals give birth and are killed at rates which ensure that the slow component does not have large drift terms on the fast timescale, whereas it does in the present setup. This necessitates the different proof strategy. In Section 1.7 we shall outline the proof strategy of Theorem 1.4, at which point we shall elaborate on the difference between this proof and the proof in [37].

The ancestral paths of both the Fleming–Viot particle system and similar particle systems have been considered by a number of authors, for instance by Méléard and Tran [37], Grigorescu and Kang [29] and Burdzy et al. [7, 15, 17, 18]. None of these make a link with the Wright–Fisher process. In a sequel to the present paper, we shall use Theorem 1.4 to link the ancestral paths of the Fleming–Viot particle system with a Wright–Fisher process on $\mathcal{P}(C([0,T];\bar{D}))$. This link was included in the original preprint version of this paper [42] and earlier in the author's Ph.D. thesis [43], Chapter 4.

In [29] Grigorescu and Kang constructed the immortal particle, also known as the spine, of the Fleming–Viot particle system—the unique ancestral path from time 0 to time ∞. They introduced the Fleming–Viot multicolor process, with the colors belonging to a finite set in order to construct this process. The construction of the spine of the Fleming–Viot particle system was later extended to a very general setting by Bieniek and Burdzy [7], Theorem 3.1. Bieniek and Burdzy [7], Section 5, established that, when the state space is finite, the distribution of the spine of the Fleming–Viot particle system converges as $N \to \infty$ to that of the driving Markov process $(X_t)_{0\leq t<\tau_\partial}$ conditioned never to be killed—referred to in the literature as the *Q-process* [19], Section 3. They conjectured that this is also true for general state spaces [7], page 3752. Since then, Burdzy, Kołodziejek and Tadić in [17, 18] have established a law of the iterated logarithm [18], Theorem 7.1, which, as they explain, hints that the conjecture of Bieniek and Burdzy should hold in the setting they consider. None of these articles draw a link with the Wright–Fisher process.

In a sequel to the present article, we shall prove Bieniek and Burdzy's conjecture, [7], page 3752, in the setting of the present paper. This proof was included in the original preprint version of this paper [42], and earlier in the author's Ph.D. thesis [43], Chapter 4. This was the first proof of the conjecture outside of the finite state space setting. Subsequent to [43], Chapter 4, and [42], Burdzy and Engländer have established this conjecture in [15], when the driving process is Brownian motion killed at the boundary of its bounded domain. We emphasise that the proof strategy due to Burdzy et al. in [7, 15] is completely different to the proof due to the present author in [43], Chapter 4, and [42], with no connection being made between the Fleming–Viot particle system and the Wright–Fisher process in [7, 15]. Bieniek and Burdzy's proof, when the state space is finite ([7], Section 5), used the finiteness of the state space in a seemingly essential way; they used the fact that if two particles are at the same location they must have the same probability of being the spine, and moreover, the particles can only be at a finite number of possible locations. Burdzy and Engländer were able to use the same argument in [15] when the driving process is Brownian motion killed at the boundary of its domain by dividing the domain up into cubes and using the form of the multidimensional Gaussian distribution to argue that any two particles in the same cube must have almost the same probability of being the spine. On the other hand, the proof appearing in [43], Chapter 4, and [42], which will appear in a sequel to the present article, instead leverages the connection between the Fleming–Viot particle system and the Wright–Fisher process established in Theorem 1.4.

The N-branching Brownian motion (N-BBM) consists of N particles evolving in between killing times as independent Brownian motions. At rate N, one kills the particle minimising or maximising a given fixed function. At the same time, as with the Fleming–Viot particle system, another particle chosen uniformly at random branches so that the number of particles remains fixed. Clearly, this particle system is similar to the Fleming–Viot particle system. Particle systems of this form were first introduced by Brunet and Derrida in [10]. Such particle systems have been studied, for instance, by Brunet and Derrida [11], Durrett and Remenik [22], Maillard [35], and Berestycki, Brunet, Nolen and Penington [6]. The genealogy of these particles systems has received particular attention; see also the work of Brunet, Derrida, Mueller and Munier [12, 13], Mallein [36] and Penington, Roberts and Talyigás [39].

For the N-BBM studied in [35], the particles are in *one* dimension with the leftmost particle being killed at each killing time. It is a hard open problem to show that the genealogy of this particle system is given by a Bolthausen–Sznitman coalescent ([35], page 1066) so we should not expect a Wright–Fisher process scaling limit. This conjecture has been proven for the similar near-critical branching Brownian motion by Berestycki, Berestycki and Schweinsberg in [5]. On the other hand, in the “Brownian bees” particle system considered in [6], it is the particle furthest away from 0 which is killed. In contrast to the N-BBM, we should expect the this particle system to have a Wright–Fisher process limit after rescaling time by $t \mapsto Nt$, as in Theorem 1.4, in the opinion of the present author. The key distinction between these two Brunet–Derrida-type particle systems is that the killing mechanism in the latter has the effect of constraining the mass of particles. However, the genealogy of the Brownian bees particle system has not yet been addressed nor has a Wright–Fisher process limit previously been established for any variant of this particle system.

A scaling limit for the geneaology of a sequential Markov chain Monte Carlo algorithm was established by Brown, Jenkins, Johansen and Koskela in [9], Theorem 3.2. This captures the phenomenon of ancestral degeneracy, which has a substantial impact on the performance of the algorithm. They established that the geneaology of an n-particle sample converges to Kingman’s n-coalescent, as the number of particles goes to infinity and time is suitably rescaled. This is suggestive of a Wright–Fisher process, since Kingman’s coalescent is dual to the Wright–Fisher process (see [32], Appendix A), but no such connection is made.

In the engineering literature, Mulatier, Dumonteil, Rosso and Zoia [21] considered a particle system whereby N Brownian particles branch at a rate λ, at which point another particle chosen uniformly at random is removed, conserving the number of particles. Clearly, this is very similar to the Fleming–Viot particle system, with the difference being that here particle births trigger another particle chosen uniformly at random to be killed rather than vice-versa. This is used as a toy model for neutrons in a nuclear reactor and their Monte Carlo simulation. They investigated the phenomenon of “clustering” in which particles cluster together in Monte-Carlo simulations of nuclear reactors, which has a substantial impact on the accuracy of these simulations. They explained this phenomenon as occurring when particle ancestries coalesce more quickly than particles are able to explore the space. They argued that this should occur on a timescale of $\frac{N}{\lambda}$. However, it is unknown how quickly ancestries coalesce for such systems (when the branching rate is nonconstant), even at the level of a conjecture. It should be straightforward to replicate the proof in the present paper for these systems, thereby quantifying how quickly ancestries coalesce via an analogue of Theorem 1.4. This would indicate how large N should be to avoid clustering. We will see in the following subsection that ancestral coalescence occurs more quickly when ϕ is nonconstant (but N and λ are the same) so that a larger N would be needed to avoid clustering.

1.5. *Effective population size.* In population genetics variance effective population size refers to the population of an idealised, spatially unstructured population with the same genetic drift per generation. For a variety of reasons, this effective population size is generally observed to be considerably less than the census population size [26].

We recall that $(\pi, -\lambda, \phi)$ is the principal eigentriple of the infinitesimal generator L. We obtained in Theorem 1.4 that, after rescaling time by $t \mapsto Nt$, the Fleming–Viot multicolor process converges to a Wright–Fisher process of rate $\Theta := \frac{2\lambda\|\phi\|^2_{L^2(\pi)}}{\|\phi\|^2_{L^1(\pi)}}$. It is straightforward to combine Theorem 1.8 with Theorem A.1 to establish that individuals in the Fleming–Viot multicolor process die, on average, λ times per unit time. If we remove space and instead assume that each individual is killed at fixed Poisson rate $\kappa \equiv \lambda$, we obtain the classical static Moran model. We, therefore, define the variance effective population here to be the size of an equivalent static Moran model.

The Wright–Fisher process is well known to arise as the limit of suitably rescaled Moran models [23], (4.12). If we let $(\vec{\eta}_t^{\mathrm{Moran},N})_{0\le t<\infty}$ be the N-individual static Moran model (where each individual dies at Poisson rate λ) and define the constant $c = \frac{\Theta}{2\lambda} = (\frac{\|\phi\|_{L^1(\pi)}}{\|\phi\|_{L^2(\pi)}})^2$, we have that $\vec{\eta}_{Nt}^{\mathrm{Moran},\lfloor cN\rfloor}$ converges to a Wright–Fisher process of rate Θ. It follows that

$$N_{\mathrm{eff}} \sim \left(\frac{\|\phi\|_{L^1(\pi)}}{\|\phi\|_{L^2(\pi)}}\right)^2 N. \tag{1.11}$$

We observe that $N_{\mathrm{eff}} \le N$, with equality if and only if ϕ is constant on $\bar{D}$, which is equivalent to κ being constant on $\bar{D}$.

We offer the following heuristic interpretation of (1.11). We have from Theorem A.1 that

$$\mathbb{P}_x(\tau_\partial > t) \sim \phi(x)e^{-\lambda t}.$$

On the other hand, the profile of the particles in the Fleming–Viot particle system will settle upon a close approximation of π. Therefore, if $\|\phi\|_{L^2(\pi)}$ is much larger than $\|\phi\|_{L^1(\pi)}$, then a small subset of individuals at any given time should be expected to subsequently survive for much longer than the average. These individuals will, therefore, have far more children than the average, having the effect of speeding up the coalescence time, hence reducing the effective population size.

1.6. *A hydrodynamic limit theorem for the Fleming–Viot multicolor process.* Both the proof of Theorem 1.4 and our heuristic explanation of it will make use of the following hydrodynamic limit theorem for the Fleming–Viot multicolor process. The hydrodynamic limit we obtain is given by the laws of the following killed Markov process.

DEFINITION 1.7. We define a $\bar{D} \times \mathbb{K}$-valued killed strong Markov process, denoted by $((X_t, \eta_t))_{0\le t<\tau_\partial}$, as follows. The process evolves in the first variable, like the killed normally-reflected diffusion $(X_t)_{0\le t<\tau_\partial}$ defined in (1.1), with the killing time of $((X_t, \eta_t))_{0\le t<\tau_\partial}$ being the same as the killing time of $(X_t)_{0\le t<\tau_\partial}$. In the second variable, η_t is a constant element of $\mathbb{K}$ up to the killing time τ_∂ so that $\eta_t = \eta_0$ for all $0 \le t < \tau_\partial$. After the killing time the process is sent to a fixed cemetery state.

THEOREM 1.8. *We consider the Fleming–Viot multicolor process* $((\vec{X}_t^N, \vec{\eta}_t^N))_{t\ge 0}$ *for* $N \ge 2$. *Then there exists constants* $C_{T,N}$ *for* $0 \le T < \infty$ *and* $N \ge 2$ *such that* $C_{T,N} \to 0$, *as* $N \to \infty$, *and such that for any initial condition* $(\vec{X}_0^N, \vec{\eta}_0^N)$ *and any* $f \in \mathcal{B}_b(\bar{D} \times \mathbb{K}; \mathbb{R})$, *we* *have that*

$$\mathbb{E}_{(\vec{X}_0^N,\vec{\eta}_0^N)}\Big[\sup_{t\le T}\Big|\Big(\frac{1}{N}\sum_{i=1}^N \delta_{(X_t^{N,i},\eta_t^{N,i})}-\mathcal{L}_{\frac{1}{N}\sum_{i=1}^N\delta_{(X_0^{N,i},\eta_0^{N,i})}}\big((X_t,\eta_t)\big)\Big)(f)\Big|\Big] \le C_{T,N}\|f\|_\infty, \tag{1.12}$$

$$\mathbb{E}_{(\vec{X}_0^N,\vec{\eta}_0^N)}\Big[\sup_{t\le T}\big|J_t^N-\ln\mathbb{P}_{\frac{1}{N}\sum_{i=1}^N\delta_{(X_0^{N,i},\eta_0^{N,i})}}(\tau_\partial>t)\big|\wedge 1\Big]\le C_{T,N}. \tag{1.13}$$

PROOF OF THEOREM 1.8. We take the Fleming–Viot particle system associated to the killed strong Markov process $((X_t,\eta_t))_{0\le t<\tau_\partial}$ defined in Definition 1.7 (which is well defined since the killing rate is bounded). We observe that its dynamics are identical to that of the Fleming–Viot multicolor process $(\vec{X}_t^N,\vec{\eta}_t^N)_{t\ge 0}$ associated to $(X_t)_{0\le t<\tau_\partial}$. We are, therefore, able to apply [47], Theorem 2.2, to the Fleming–Viot multicolor process.

The statement of [47], Theorem 2.2, only gives an estimate of the particle system at fixed times. However, its proof relied on a martingale decomposition, [47], Theorem 2.2, with L^2 controls obtained on the two martingales [47], (2.8) and (2.9). By applying Doob's L^2-martingale inequality, these controls become uniform over the time horizon $[0,T]$. We thereby make [47], Theorem 2.2, uniform over the time horizon $[0,T]$ at the cost of the estimate in [47], Theorem 2.2, being multiplied by 4. Applying this uniform estimate to the Fleming–Viot multicolor process, we obtain (1.12). We similarly obtain (1.13) from [47], $first$ equation on page 450. □

We prove in the Appendix that $((X_t,\eta_t))_{0\le t<\tau_\partial}$ has the following large-time limit.

PROPOSITION 1.9. *For arbitrary sequences $(x^i,\eta^i)_{1\le i\le n}$ in $\bar{D}\times\mathbb{K}$, we consider the process $(X_t,\eta_t)_{0\le t<\tau_\partial}$ with initial distribution given by the empirical measure $\frac{1}{n}\sum_{i=1}^n\delta_{(x^i,\eta^i)}$. Then there exists $c_t\to 0$ as $t\to\infty$ such that, for all sequences $(x^i,\eta^i)_{1\le i\le n}$ in $\bar{D}\times\mathbb{K}$ and all $n\in\mathbb{N}$, we have*

$$\Big\|\mathcal{L}_{\frac{1}{n}\sum_{i=1}^n\delta_{(x^i,\eta^i)}}\big((X_t,\eta_t)|\tau_\partial>t\big)-\frac{\sum_{i=1}^n\phi(x^i)\pi\otimes\delta_{\eta^i}}{\sum_{i=1}^n\phi(x^i)}\Big\|_{\mathrm{TV}}\le c_t,\quad 0\le t<\infty. \tag{1.14}$$

1.7. *Heuristics for the proof of Theorem 1.4.*

The principal difficulty to be addressed. Méléard and Tran considered in [37] the ancestries of a similar particle system in [37]. There the individuals in the population have a trait and an age, with the individuals giving birth (passing on their trait) and dying in an age-dependent manner. The age component plays a similar role to spatial position in the present article. However, aside from obtaining a different scaling limit, they also employed a different proof strategy.

The proof of Méléard and Tran in [37] extended to the particle system setting the strategy of Kurtz [31] and Ball, Kurtz, Popovic and Rempala [4], which concerned diffusions. In contrast, the proof in the present article extends to the particle system setting techniques of Katzenberger [30] (the author is not aware of this technique previously having been extended to the particle system setting), which also concerned diffusion processes. This is necessitated by the following qualitative difference between the two particle systems.

In [37] individuals have a trait x (the slow variable) and an age a (the fast variable). The speed-up of the timescale is given by the parameter n. On the fast timescale, they give birth at rate $nr(x,a)+b(x,a)$, whilst dying at rate $nr(x,a)+d(x,a)$. We observe that the fast term, $nr(x,a)$, is the same in both the former and the latter. Consequentially, when they formulate the corresponding martingale problem, the slow variable does not have a large drift term on the fast timescale. The terms $b(x,a)$ and $d(x,a)$ may change quickly due to the fast evolution of the age term a—this is dealt with via averaging—but they remain $\mathcal{O}(1)$ on the fast timescale.

We may contrast this with the Fleming–Viot multicolor process. We recall that the time change is $t \mapsto Nt$. We consider a test function $f \in C_b(\mathbb{K})$ and observe that, on the fast timescale, the empirical measure of the colors evaluated against f, $\chi^N_{Nt}(f)$, satisfies

$$d\chi^N_{Nt}(f) = \sum_{i=1}^{N} \kappa\big(X^i_{Nt}\big)\bigg[\frac{1}{N-1}\sum_{j\neq i}\big[f\big(\eta^j_{Nt}\big) - f\big(\eta^i_{Nt}\big)\big]\bigg]\,dt + d(\text{Martingale})_t. \tag{1.15}$$

We see that the drift term is of $\mathcal{O}(N)$ on the fast timescale. In particular, the change in position of an individual particle has an $\mathcal{O}(1)$ effect on the drift. A large deviations principle for the Fleming–Viot multicolor process would provide controls on the drift valid over a sufficiently large timescale (a LDP for the Fleming–Viot particle system driven by Brownian motion with soft killing was established by Grigorescu in [28]) but would only control the drift on a fast timescale to $\mathcal{O}(N)$. Since microscopic fluctuations in the position of individual particles have an $\mathcal{O}(1)$ effect on the drift, there would not seem to be any hope of obtaining adequate controls on the drift term in order to apply a compactness-uniqueness argument (in which one characterises the martingale problem solved by subsequential limits).

The key idea, allowing us to deal with these large drift terms, will be to consider the tilted empirical measure $\mathcal{Y}^N_t$, which we recall was given in (1.8) as

$$\mathcal{Y}^N_t := \frac{\frac{1}{N}\sum_{i=1}^N \phi(X^i_t)\delta_{\eta^i_t}}{\frac{1}{N}\sum_{i=1}^N \phi(X^i_t)} \in \mathcal{P}(\mathbb{K}).$$

Motivation for choosing $\mathcal{Y}^N_t$. We take inspiration from Katzeberger's approach in [30]. Consider a dynamical system in Euclidean space, $\dot{x}_t = b(x_t)$, with an attractive manifold of equilibrium $\mathcal{M}$ and flow map $\varphi(x,s)$. Katzenberger [30] established (under reasonable conditions) that the long-term dynamics of the randomly perturbed dynamical system,

$$dx^\epsilon_t = b\big(x^\epsilon_t\big)\,dt + \epsilon\,dW_t, \tag{1.16}$$

can be obtained by considering the following nonlinear projection onto the manifold of equilibria:

$$\varpi(x) := \lim_{s\to\infty} \varphi(x,s) \in \mathcal{M}. \tag{1.17}$$

We summarise Katzenberger's idea as follows. Since $\nabla\varpi \cdot b \equiv 0$, the Stratanovich chain rule implies that

$$d\varpi\big(x^\epsilon_t\big) = \epsilon\nabla\varpi\big(x^\epsilon_t\big) \circ dW_t.$$

In particular, the large drift term has been eliminated from the above expression. We may rescale time to see that $\varpi(x^\epsilon_{\frac{t}{\epsilon^2}})$ satisfies

$$d\varpi\big(x^\epsilon_{\frac{t}{\epsilon^2}}\big) = \nabla\varpi\big(x^\epsilon_{\frac{t}{\epsilon^2}}\big) \circ d\tilde{W}_t,$$

whereby $\tilde{W}_t$ is the Brownian motion $\tilde{W}_t := \epsilon W_{\frac{t}{\epsilon^2}}$. Since the dynamical system will be pushed toward the attractive manifold of equilibrium on a fast timescale, one can then argue that

$$x^\epsilon_{\frac{t}{\epsilon^2}} \approx \varpi\big(x^\epsilon_{\frac{t}{\epsilon^2}}\big). \tag{1.18}$$

We can, therefore, obtain a scaling limit for $x^\epsilon_{\frac{t}{\epsilon^2}}$. This scaling limit is a diffusion on $\mathcal{M}$.

Whilst Katzenberger's results in [30] were restricted to finite-dimensions, we may ask what the analogue of $\varpi(x_t^\epsilon)$ is in the present setting? We will see that $\mathcal{Y}_t^N$ can be thought of as being analogous to the quantity $\varpi(x_t^\epsilon)$ considered by Katzenberger.

We denote by $((X_t,\eta_t))_{0\le t<\tau_\partial}$ the killed Markov process defined in Definition 1.7. It follows from Theorem 1.8 that we can think of the Fleming–Viot multicolor process as a random perturbation of the dynamical system with flow map

$$\mathcal{P}(\bar{D}\times\mathbb{K})\times[0,\infty)\ni(\upsilon,s)\mapsto\mathcal{L}_\upsilon\big((X_s,\eta_s)|\tau_\partial>s\big)\in\mathcal{P}(\bar{D}\times\mathbb{K}). \tag{1.19}$$

Proposition 1.9 provides for the large-time limits of this flow. We therefore see from Proposition 1.9 that the analogue of $\varpi(x_t^\epsilon)$ is given by

$$\frac{\frac{1}{N}\sum_{i=1}^N\phi(X_t^{N,i})\pi\otimes\delta_{\eta_t^{N,i}}}{\frac{1}{N}\sum_{i=1}^N\phi(X_t^{N,i})}=\pi\otimes\mathcal{Y}_t^N.$$

We discard π, since it is constant, leaving only $\mathcal{Y}_t^N$.

There is a second heuristic reason for examining $\mathcal{Y}_t^N$. If $x(t)$ and $y(t)$ both satisfy the ODEs $\dot{x}=c(t)x$ and $\dot{y}=c(t)y$ for the same $c(t)$, then $\frac{y(t)}{x(t)}$ is constant. If we now instead consider the SDEs $dX_t=c_tX_t\,dt+\epsilon\,dW_t$ and $dY_t=c_tY_t\,dt+\epsilon\,dW_t$, $\frac{Y_t}{X_t}$ will satisfy an SDE with only $\mathcal{O}(\epsilon^2)$ drift terms, since the $\mathcal{O}(1)$ terms will cancel out as in the deterministic case (one can check this using Itô's lemma).

We now define for $\mathcal{E}\in\mathcal{B}(\mathbb{K})$ the following, which shall be used throughout the proof of Theorem 1.4:

$$\begin{aligned}&P_t^{N,\mathcal{E}}:=\frac{1}{N}\sum_{i=1}^N\mathbb{1}_{\eta_t^i\in\mathcal{E}}\phi(X_t^i),\qquad Q_t^N:=P^{N,\mathbb{K}}=\frac{1}{N}\sum_{i=1}^N\phi(X_t^i)\quad\text{and}\\&Y_t^{N,\mathcal{E}}:=\mathcal{Y}_t^N(\mathcal{E})=\frac{P_t^{N,\mathcal{E}}}{Q_t^N}.\end{aligned} \tag{1.20}$$

The important point is that, to leading order, both $P^{N,\mathcal{E}}$ and Q^N evolve with drift terms proportional to themselves with the same constant of proportionality. Indeed, on the slow timescale the killed process X_t satisfies

$$d\phi(X_t)=L\phi(X_t)+\text{martingale terms}=-\lambda\phi(X_t)+\text{martingale terms}.$$

Therefore, between jumps, and including the process of killing the particles, the quantities $P_t^{N,\mathcal{E}}$ and Q_t^N evolve with drift terms $-\lambda P_t^{N,\mathcal{E}}\,dt$ and $-\lambda Q_t^N\,dt$, respectively. Furthermore, if particle $(X^{N,i},\eta^{N,i})$ dies at time t, $\frac{1}{N}\phi(X_t^{N,i})\mathbb{1}(\eta_t^{N,i}\in\mathcal{E})$ (resp., $\frac{1}{N}\phi(X_t^{N,i})$) is added to the value of $P_t^{N,\mathcal{E}}$ (resp., $Q_t^{N,\mathcal{E}}$), the expected value of which is $P_{t-}^{N,\mathcal{E}}+\mathcal{O}(\frac{1}{N})$ (resp., $Q_{t-}^{N,\mathcal{E}}+\mathcal{O}(\frac{1}{N})$). This occurs at Poisson rate $\kappa(X_t^i)$. Thus, after the time-change $t\mapsto Nt$, we can write

$$\begin{aligned}dP_{Nt}^{N,\mathcal{E}}&=\Big[-\lambda N+\sum_{i=1}^N\kappa(X_t^i)\Big]P_{Nt}^{N,\mathcal{E}}\,dt+\mathcal{O}(1)\,dt+d(\text{martingale})_t,\\dQ_{Nt}^N&=\Big[-\lambda N+\sum_{i=1}^N\kappa(X_t^i)\Big]Q_{Nt}^N\,dt+\mathcal{O}(1)\,dt+d(\text{martingale})_t.\end{aligned} \tag{1.21}$$

In particular, on the fast timescale, given by $t\mapsto Nt$, both $P_{Nt}^{N,\mathcal{E}}$ and Q_{Nt}^N both evolve with drift proportional to themselves with the same constant of proportionality given by

$$-\lambda N+\sum_{i=1}^N\kappa(X_t^i). \tag{1.22}$$

We observe that the change in position of an individual particle has an $\mathcal{O}(1)$ effect on the constant of proportionality (1.22). However, these large effects cancel out by placing the normalisation at microscopic scale in the denominator, as the constant of proportionality in both the numerator and denominator must be the same.

From these considerations we see that, having rescaled time by $t \mapsto Nt$, $\mathcal{Y}^N_{Nt}$ should satisfy an SDE with $\mathcal{O}(1)$ drift terms. It is straightforward to see that the martingale terms will have $\mathcal{O}(1)$ quadratic variation on this timescale. It follows that $\mathcal{Y}^N_{Nt}$ should be susceptible to a compactness-uniqueness argument in which we establish tightness before uniquely characterising subsequential limits by characterising their drift and quadratic variation. We shall thereby obtain a scaling limit for $\mathcal{Y}^N_{Nt}$. We note that, since the leading order terms in (1.21) cancel out, we shall need to calculate the "$\mathcal{O}(1)\,dt$" higher order terms, which is responsible for much of the computational complexity in the proof of Theorem 1.4.

The relationship between χ^N_{Nt} and $\mathcal{Y}^N_{Nt}$. The above will allow us to characterise the limit in distribution of $(\mathcal{Y}^N_{Nt})_{t\geq 0}$. Our goal, however, is to characterise the limit in distribution of $(\chi^N_{Nt})_{t\geq 0}$. We would, therefore, like to relate $\mathcal{Y}^N_{Nt}$ with χ^N_{Nt}.

The key observation here is that, on the original slow timescale, the color of a particle and its spatial position become "independent" after an $\mathcal{O}(1)$ time. To be more precise, for any given $A \subseteq \mathbb{K}$, the spatial profile of particles whose colors belong to A,

$$\frac{\sum_{i=1}^N \mathbb{1}(\eta^i_t \in A)\delta_{X^i_t}}{|\{i : \eta^i_t \in A\}|},$$

converges over an $\mathcal{O}(1)$ timescale to the quasi-stationary distribution π, a deterministic profile. Thus, for different subsets $A, B \subseteq \mathbb{K}$, the number of particles with colors belonging to A and B may well be different, but the spatial profiles of the two sets of particles will be the same for large N. Since the particles corresponding to different colors have the same spatial profile, weighting the empirical measure of the colors according to the right eigenfunction evaluated at the corresponding spatial positions will have no effect. It follows that χ^N_t and $\mathcal{Y}^N_t$ will be close after an $\mathcal{O}(1)$ time. On the fast timescale, χ^N_{Nt} will, therefore, be close to $\mathcal{Y}^N_{Nt}$. This is analogous to the second step in Katzenberger's approach in [30], described above in (1.18).

The proof of Theorem 1.4 will, therefore, follow by establishing that $(\mathcal{Y}^N_{Nt})_{t\geq 0}$ converges to the Wright–Fisher process,] and showing that χ^N_{Nt} is close to $\mathcal{Y}^N_{Nt}$.

1.8. *Why is the limit a Wright–Fisher process?.* It follows from the above heuristic that χ^N_t should evolve over an $\mathcal{O}(N)$ timescale and that χ^N_{Nt} should converge to some $\mathcal{P}(\mathbb{K})$-valued process (at least on subsequences). In the proof of Theorem 1.4, we will calculate that the limit is a Wright–Fisher process. However, it is not readily apparent from this why the limit should necessarily be a Wright–Fisher process. We offer here a heuristic argument for why we should expect the limit to be a Wright–Fisher process.

We let $(\hat{\nu}_t)_{t\geq 0}$ be the limit to be determined of $(\chi^N_{Nt})_{t\geq 0}$ (perhaps along a subsequence). By the aforedescribed separation of timescales phenomenon, this will be a $\mathcal{P}(\mathbb{K})$-valued process, with the dependence on the spatial component "averaged out." We consider an arbitrary measurable map $\iota : \mathbb{K} \to \mathbb{K}$. We can think of ι as relabelling the colors. The key observation is that $\{(X^{N,i}_t, \iota(\eta^{N,i}_t)) : 1 \leq i \leq N\}$ is itself a Fleming–Viot multicolor process—the Fleming–Viot multicolor process remains one after relabelling the colors. It follows that whatever dynamics $(\hat{\nu}_t)_{t\geq 0}$ has, $(\iota_{\#}\hat{\nu}_t)_{t\geq 0}$ must have the same dynamics. This allows us both to exchange colors and to relabel different colors as the same color.

It follows that there should exist continuous functions

$$b, \sigma_{11} : [0,1] \to \mathbb{R} \quad \text{and} \quad \sigma_{12} : \{(p,q) \in [0,1]^2 : p + q \leq 1\} \to \mathbb{R}$$

such that the following are continuous martingales for all disjoint $A_1, A_2 \in \mathcal{B}(\mathbb{K})$:

$$\hat{\nu}_t(A_1) - \int_0^t b(\hat{\nu}_s(A_1))\,ds, \qquad (\hat{\nu}_t(A_1))^2 - \int_0^t \sigma_{11}(\hat{\nu}_s(A))\,ds, \quad \text{and}$$

$$(\hat{\nu}_t(A_1))(\hat{\nu}_t(A_2)) - \int_0^t \sigma_{12}(\hat{\nu}_s(A_1), \hat{\nu}_s(A_2))\,ds.$$

Moreover, since a color of mass p and a color of mass q can be relabelled to be a single color of mass $p+q$, it is clear that

$$b(p+q) = b(p) + b(q),$$

$$\sigma_{11}(p+q) = \sigma_{11}(p) + \sigma_{11}(q) + 2\sigma_{12}(p,q), \quad 0 \le p, q \le p+q \le 1$$

$$\sigma_{12}(p_1+p_2, q_1+q_2) = \sum_{1 \le i,j \le 2} \sigma_{12}(p_i, q_j), \quad 0 \le p_1, p_2, q_1, q_2 \le p_1+p_2+q_1+q_2 \le 1.$$

Furthermore, the whole color space $\mathbb{K}$ must have total mass 1, so $\hat{\nu}_t(\mathbb{K}) \equiv 1$. From these considerations we see that the only possibility is that, for some constant θ,

$$b \equiv 0, \qquad \sigma_{11}(p) = \frac{\theta}{2}(p - p^2) \quad \text{and} \quad \sigma_{12}(p,q) = -\frac{\theta}{2}pq.$$

We recognise the Wright–Fisher diffusion described in Section 1.2. In light of Proposition 1.3, it is, therefore, natural that our unknown limit $(\hat{\nu}_t)_{t \ge 0}$ should be a Wright–Fisher process.

1.9. *Hard catalyst killing.* The setting of the present paper—in which the Fleming–Viot particle system is driven by diffusions with soft killing—has been chosen to establish the connection between the Fleming–Viot process and the Wright–Fisher process with a minimum of technical difficulties. Nevertheless, in Section 5 we will extend this connection to the original setting, considered by Burdzy, Holyst and March [16], in which the Fleming–Viot particle system is driven by Brownian motion with instantaneous killing at the boundary (*hard killing*). To avoid switching back and forth between Fleming–Viot particle systems with different dynamics, we will only consider the case of hard killing in Section 5, the final section prior to the Appendix and in Appendix E. Our results in the case of hard killing are, therefore, stated and proved in Section 5.

We emphasise that the proof strategy employed in the present paper may be applied to the Fleming–Viot particle system driven by more general killed Markov processes. The principal requirements to apply this proof strategy are that:

1. The driving killed Markov process $(X_t)_{0 \le t < \tau_\partial}$ is Feller.
2. Its infinitesimal generator has a positive, continuous and bounded principal right eigenfunction ϕ.
3. $\mathcal{L}_\mu(X_t | \tau_\partial > t)$ converges to a unique quasi-stationary distribution for any initial condition μ.
4. We can constrain the empirical measure of the spatial positions of the particles m_t^N to a tight set of measures over any $\mathcal{O}(N)$ timescale, precluding in particular the mass from accumulating at the boundary.

In the case of hard killing at the boundary in a bounded domain, the main additional difficulty is to establish Requirement 4. We will obtain such controls for the Fleming–Viot particle system driven by Brownian motion with hard killing in Section 5. With these controls in hand, the extension of our results to this setting proceeds by essentially the same proof.

When the domain is unbounded, the situation is much more delicate. For diffusions on the positive real line $\mathbb{R}_{>0}$ with hard killing at 0, one could probably establish similar results for Ornstein–Uhlenbeck dynamics, using the strong negative drift to control the particles far away from 0 over an $\mathcal{O}(N)$ time scale. For this process the principal right eigenfunction is unbounded (it's given by $\phi(x) = x$); instead, strong controls on the mass of particles far away from 0 (where ϕ is large) over an $\mathcal{O}(N)$ timescale would be required to replace the boundedness of ϕ. To be more precise, we would need to show, for any $T < \infty$, that $\sup_{0\le t\le NT} \frac{1}{N}\sum_{1\le i\le N}(\phi(X_t^{N,i}))^2$ is bounded by some uniform constant with probability arbitrarily close to 1, uniformly in N. On the other hand, we should not expect the Fleming–Viot particle system driven by Brownian motion with drift -1 to have a Wright–Fisher process scaling limit, this drift being too weak to adequately control the particles. Indeed, it is a hard open problem to show that the genealogy of the very similar N-BBM is given by a Bolthausen–Sznitman coalescent [35], page 1066.

1.10. *Structure of the paper.* A summary of the notation, which we shall need for our proof, is given in Section 2. The proof of Theorem 1.4 shall rely on a number of calculations of the quantity $\mathcal{Y}_t^N$, defined in (1.8). To avoid obscuring our proof with calculations, we will carry out these calculations in Section 3. We shall then prove Theorem 1.4 in Section 4. We will extend our results to the Fleming–Viot multicolor process driven by Brownian motion with instantaneous killing at the boundary in Section 5. We conclude with the Appendix.

2. Notation for the proof of Theorem 1.4.

We recall from (1.20) that we define, for $\mathcal{E} \in \mathcal{B}(\mathbb{K})$,

$$P_t^{N,\mathcal{E}} := \frac{1}{N}\sum_{i=1}^N \mathbb{1}_{\eta_t^i\in\mathcal{E}}\phi(X_t^i), \qquad Q_t^N := P^{N,\mathbb{K}} = \frac{1}{N}\sum_{i=1}^N \phi(X_t^i) \quad \text{and}$$

$$Y_t^{N,\mathcal{E}} := \mathcal{Y}_t^N(\mathcal{E}) = \frac{P_t^{N,\mathcal{E}}}{Q_t^N}.$$

We recall the definition of m_t^N and χ_t^N from (1.4) and further define $m_t^{N,\mathcal{E}}$ for $\mathcal{E} \in \mathcal{B}(\mathbb{K})$,

$$(2.1)\qquad \begin{aligned} m_t^N &:= \frac{1}{N}\sum_{i=1}^N \delta_{X_t^{N,i}}, \qquad \chi_t^N := \frac{1}{N}\sum_{i=1}^N \delta_{\eta_t^i} \quad \text{and} \\ m_t^{N,\mathcal{E}} &:= \sum_{i=1}^N \mathbb{1}(\eta_t^i \in \mathcal{E})\delta_{X_t^i} = m_t^N(\mathcal{E}). \end{aligned}$$

We recall from Appendix A that the infinitesimal generator of the reflected diffusion with (resp., without) soft killing is denoted by L (resp., L_0). We further recall that the Carre du champs operator of the latter is denoted as Γ_0 and is defined on the algebra $\mathcal{A}$. This algebra contains the principal right eigenfunction ϕ of L, by Theorem A.1. We further define

$$(2.2)\qquad \begin{aligned} &\Lambda_t^{N,\mathcal{E}} := \langle m_t^{N,\mathcal{E}}, \Gamma_0(\phi) + \kappa\phi^2\rangle + \langle m_t^{N,\mathcal{E}}, \phi^2\rangle\langle m_t^N, \kappa\rangle \\ &\text{for } \mathcal{E} \in \mathcal{B}(\mathbb{K}), \text{ and } \Lambda_t^N := \Lambda_t^{N,\mathbb{K}}. \end{aligned}$$

2.1. *$\mathcal{O}$ notation.* The following notation shall significantly simplify our calculations.

For any finite variation process $(X_t)_{0\le t<\infty}$, we write $V_t(X)$ for the total variation

$$(2.3)\qquad V_t(X) = \sup_{0=t_0<t_1<\cdots<t_n=t} \sum_{i=0}^{n-1} |X_{t_{i+1}} - X_{t_i}|.$$

Moreover, for all càdlàg processes $(X_t)_{0\le t<\infty}$, we write

$$(2.4)\qquad \Delta X_t = X_t - X_{t-}.$$

Given some family of random variables $\{X^N : N \in \mathbb{N}\}$ and nonnegative random variables $\{Y^N : N \in \mathbb{N}\}$, we say that $X^N = \mathcal{O}(Y^N)$ if there exists a uniform constant $C < \infty$ such that $|X^N| \le CY^N$. Note that we shall abuse notation by using an equals sign rather than an inclusion sign.

We now define the notion of process sequence class. Given sequences of processes $\{(X_t^N)_{t\ge 0} : N \in \mathbb{N}\}$ and $\{(Y_t^N)_{t\ge 0} : N \in \mathbb{N}\}$, we say that:

1. $X_t^N = \mathcal{O}_t^{\mathrm{MG}}(Y^N)$ if for all $N \ge N_0$ (for some $N_0 < \infty$) and for some $C < \infty$, X_t^N is a martingale whose quadratic variation is such that

$$[X^N]_t - \int_0^t CY_s^N\,ds \quad \text{is a supermartingale.} \tag{2.5}$$

2. $X_t^N = \mathcal{O}_t^{\mathrm{FV}}(Y^N)$ if for all $N \ge N_0$ (for some $N_0 < \infty$) and for some $C < \infty$, X_t^N is a finite variation process whose total variation is such that

$$V_t(X^N) - \int_0^t CY_s^N\,ds \quad \text{is a supermartingale.} \tag{2.6}$$

3. $X_t^N = \mathcal{O}_t^{\Delta}(Y^N)$ if for all $N \ge N_0$ (for some $N_0 < \infty$) and for some $C < \infty$, X_t^N is such that

$$|\Delta X_t^N| \le CY_{t-}^N \quad \text{for all } 0 \le t < \infty, \text{ almost surely.} \tag{2.7}$$

4. $X_t^N = \mathcal{O}_t^{\mathrm{Cts}}$ or $X_t^N = \mathcal{O}_t^{\mathrm{Lip}}$ if for all $N \ge N_0$ (for some $N_0 < \infty$), X_t^N has continuous (resp., Lipschitz) sample paths, almost surely.

We refer to $\mathcal{O}_t^{\mathrm{MG}}(Y^N)$, $\mathcal{O}_t^{\mathrm{FV}}(Y^N)$ and $\mathcal{O}_t^{\mathrm{FV}}(Y^N)$ for $((Y_t^N)_{0\le t<\infty} : N \in \mathbb{N})$ a given sequence of processes, and $\mathcal{O}_t^{\mathrm{Cts}}$ and $\mathcal{O}_t^{\mathrm{Lip}}$ as process sequence classes. Note that as with sequences of random variables, we abuse notation by using an equals sign rather than an inclusion sign.

Suppose that we have constants $r_N > 0$ $(N \in \mathbb{N})$. For a given sequence of processes Y^N, write $Z_s^N := Y_{r_N s}^N$. The statements $X_t^N = \mathcal{O}_t^{\mathrm{MG}}(Y_{r_N \cdot}^N)$, $X_t^N = \mathcal{O}_t^{\mathrm{FV}}(Y_{r_N \cdot}^N)$ and $X_t^N = \mathcal{O}_t^{\Delta}(Y_{r_N \cdot}^N)$ should be interpreted as the statements $X_t^N = \mathcal{O}_t^{\mathrm{MG}}(Z_\cdot^N)$, $X_t^N = \mathcal{O}_t^{\mathrm{FV}}(Z_\cdot^N)$ and $X_t^N = \mathcal{O}_t^{\Delta}(Z_\cdot^N)$, respectively.

Given an index set $\mathbb{A}$, a family of sequences of processes $\{((X_t^{N,\alpha})_{0\le t<\infty})_{N=1}^\infty : \alpha \in \mathbb{A}\}$ and a family of process sequence classes $\{\mathcal{A}_t^{N,\alpha} : \alpha \in \mathbb{A}\}$, we say that $X_t^{N,\alpha} = \mathcal{A}_t^{N,\alpha}$ uniformly if the constants C^α and N_0^α used to define $X_t^{N,\alpha} = \mathcal{A}_t^{N,\alpha}$, as in 1–3, can be chosen uniformly in $\alpha \in \mathbb{A}$.

It will be useful to take the sum and intersection of process sequence classes and specific sequences of processes. To be more precise, for any process sequence classes $\mathcal{A}_t^N$ and $\mathcal{B}_t^N$ and the sequence of processes F_t^N, we say that:

1. $X_t^N = \mathcal{A}_t^N \cap \mathcal{B}_t^N$ if $X_t^N = \mathcal{A}_t^N$ and $X_t^N = \mathcal{B}_t^N$;
2. $X_t^N = F_t^N + \mathcal{A}_t^N$ if there exists a sequence of processes G_t^N such that $G_t^N = \mathcal{A}_t^N$ and $X_t^N = F_t^N + G_t^N$;
3. $X_t^N = \mathcal{A}_t^N + \mathcal{B}_t^N$ if there exists sequences of processes G_t^N and H_t^N such that $G_t^N = \mathcal{A}_t^N$, $H_t^N = \mathcal{B}_t^N$ and $X_t^N = G_t^N + H_t^N$;
4. $dX_t = dF_t^N + d\mathcal{A}_t^N + d\mathcal{B}_t^N$ if there exists sequences of processes G_t^N and H_t^N such that $G_t^N = \mathcal{A}_t^N$, $H_t^N = \mathcal{B}_t^N$ and $dX_t^N = dF_t^N + dG_t^N + dH_t^N$.

For example, if $X_t^N = \mathcal{O}_t^{\mathrm{MG}}(1) + \mathcal{O}_t^{\mathrm{FV}}(\frac{1}{N}) \cap \mathcal{O}_t^{\Delta}(\frac{1}{N^2})$, then there exists G_t^N and H_t^N such that $X_t^N = G_t^N + H_t^N$ whereby $G_t^N = \mathcal{O}_t^{\mathrm{MG}}(1)$ and $H_t^N = \mathcal{O}_t^{\mathrm{FV}}(\frac{1}{N}) \cap \mathcal{O}_t^{\Delta}(\frac{1}{N^2})$. Thus, $X_t^N =$

$\mathcal{O}_t^{\mathrm{MG}}(1)+\mathcal{O}_t^{\mathrm{FV}}(\frac{1}{N})\cap\mathcal{O}_t^{\Delta}(\frac{1}{N^2})$ means that, for some $0<C<\infty$, there exists for all N large enough martingales G_t^N and finite-variation processes H_t^N such that

$$\begin{aligned}
&X_t^N = G_t^N + H_t^N,\\
&\bigl[G^N\bigr]_t - Ct \quad \text{is a supermartingale since } G_t^N=\mathcal{O}_t^{\mathrm{MG}}(1),\\
&V_t\bigl(H^N\bigr)-\frac{t}{N} \quad \text{is a supermartingale since } H^N=\mathcal{O}_t^{\mathrm{FV}}\Bigl(\frac{1}{N}\Bigr)\\
&\quad\text{and}\quad \bigl|\Delta Z_t^N\bigr|\le\frac{C}{N^2} \quad \text{for all } 0\le t<\infty, \text{ almost surely, since } H_t^N=\mathcal{O}_t^{\Delta}\Bigl(\frac{1}{N^2}\Bigr).
\end{aligned}$$

3. Characterisation of $\mathcal{Y}_t^N$. In the proof of Theorem 1.4, we will obtain a scaling limit for the tilted empirical measure of the colors on a fast timescale, $(\mathcal{Y}_{Nt}^N)_{t\ge 0}$. This will rely on various calculations characterising its drift and quadratic variation. To avoid obscuring the proof of Theorem 1.4 with calculations, we perform these calculations here.

In this section we write $(\Omega,\mathcal{G},(\mathcal{G}_t)_{t\ge 0},\mathbb{P})$ for the underlying filtered probability space.

REMARK 3.1. In the present section, all statements as to processes belonging to various process sequence classes should be interpreted as being uniform over all choices $\mathcal{E},\mathcal{F}\in\mathcal{B}(\mathbb{K})$ (or over all sequences of $\mathcal{G}_0$-measurable random $\mathcal{E}^N,\mathcal{F}^N\in\mathcal{B}(\mathbb{K})$ in the case of Part 4 of Theorem 3.2).

We recall that

$$\mathcal{Y}_t^N := \frac{\frac{1}{N}\sum_{i=1}^N \phi(X_t^i)\delta_{\eta_t^i}}{\frac{1}{N}\sum_{i=1}^N \phi(X_t^i)}.$$

In this section we prove the following theorem.

THEOREM 3.2. *We have the following, uniformly over all choices of $\mathcal{E},\mathcal{F}\in\mathcal{B}(\mathbb{K})$:*

1. *The covariation $[Y^{N,\mathcal{E}},Y^{N,\mathcal{F}}]_t$ is such that $[Y^{N,\mathcal{E}},Y^{N,\mathcal{F}}]_t=\mathcal{O}_t^{\mathrm{FV}}(\frac{Y^{N,\mathcal{E}}Y^{N,\mathcal{F}}}{N})$ for disjoint $\mathcal{E},\mathcal{F}\in\mathcal{B}(\mathbb{K})$.*

2. *There exists martingales $\mathcal{K}_t^{N,\mathcal{E}}$ for $\mathcal{E}\in\mathcal{B}(\mathbb{K})$ such that $Y_t^{N,\mathcal{E}}$ satisfies*

$$\begin{aligned}
Y_t^{N,\mathcal{E}} = Y_0^{N,\mathcal{E}} + \int_0^t \Bigl[&-\frac{1}{(N-1)Q_s^N}\langle m_s^{N,\mathcal{E}} - Y_s^{N,\mathcal{E}} m_s^N,\kappa\phi\rangle\\
&-\frac{1}{NQ_s^N}\langle Y_s^{N,\mathcal{E}} m_s^N - m_s^{N,\mathcal{E}},\kappa\phi\rangle\\
&+\frac{1}{N(Q_s^N)^2}\bigl(Y_s^{N,\mathcal{E}}\Lambda_s^N-\Lambda_s^{N,\mathcal{E}}\bigr)\Bigr]ds + \mathcal{K}_t^{N,\mathcal{E}} + \mathcal{O}_t^{\mathrm{MG}}\Bigl(\frac{Y^{N,\mathcal{E}}}{N^3}\Bigr)\\
&+\mathcal{O}_t^{\mathrm{FV}}\Bigl(\frac{Y^{N,\mathcal{E}}}{N^2}\Bigr)\cap\mathcal{O}_t^{\Delta}\Bigl(\frac{1}{N^3}\Bigr)
\end{aligned}\tag{3.1}$$

for $\mathcal{E}\in\mathcal{B}(\mathbb{K})$ and such that

$$\begin{aligned}
[\mathcal{K}^{N,\mathcal{E}},\mathcal{K}^{N,\mathcal{F}}]_t = \int_0^t \frac{1}{N(Q_s^N)^2}\bigl[&\Lambda_s^{N,\mathcal{E}\cap\mathcal{F}} - Y_s^{N,\mathcal{E}}\Lambda_s^{N,\mathcal{F}} - Y_s^{N,\mathcal{F}}\Lambda_s^{N,\mathcal{E}}\\
&+Y_s^{N,\mathcal{E}}Y_s^{N,\mathcal{F}}\Lambda_s^N\bigr]ds + \mathcal{O}_t^{\mathrm{MG}}\Bigl(\frac{Y^{\mathcal{E}}Y^{\mathcal{F}}+Y^{\mathcal{E}\cap\mathcal{F}}}{N^3}\Bigr)\\
&+\mathcal{O}_t^{\mathrm{FV}}\Bigl(\frac{Y^{\mathcal{E}}Y^{\mathcal{F}}+Y^{\mathcal{E}\cap\mathcal{F}}}{N^2}\Bigr)\cap\mathcal{O}_t^{\mathrm{Cts}} \quad \textit{for all } \mathcal{E},\mathcal{F}\in\mathcal{B}(\mathbb{K}).
\end{aligned}\tag{3.2}$$

3. *Furthermore*, $Y_t^{N,\mathcal{E}}$ *satisfies*

$$Y_t^{N,\mathcal{E}} = \left[\mathcal{O}_t^{\mathrm{FV}}\left(\frac{Y^{N,\mathcal{E}}}{N}\right) + \mathcal{O}_t^{\mathrm{MG}}\left(\frac{Y^{N,\mathcal{E}}}{N}\right)\right] \cap \mathcal{O}_t^{\Delta}\left(\frac{1}{N}\right). \tag{3.3}$$

4. *Parts* 1–3 *remain true if* $\mathcal{E}$ *and* $\mathcal{F}$ *are replaced with a sequence of* σ_0*-measurable random sets* $\mathcal{E}^N$ *and* $\mathcal{F}^N$.

3.1. *Proof of Theorem* 3.2. We first introduce some definitions. We define

$$F(\vec{r}) = \frac{p}{q} \quad \text{for } \vec{r} = \begin{pmatrix} p \\ q \end{pmatrix} \in \mathbb{R}^2_{>0}. \tag{3.4}$$

We write $H = H(\vec{r})$ for the Hessian and calculate

$$\nabla F(\vec{r}) = \begin{pmatrix} \frac{1}{q} \\ -\frac{p}{q^2} \end{pmatrix} \quad \text{and} \quad H(\vec{r}) = \begin{pmatrix} 0 & -\frac{1}{q^2} \\ -\frac{1}{q^2} & 2\frac{p}{q^3} \end{pmatrix} \quad \text{for } \vec{r} = \begin{pmatrix} p \\ q \end{pmatrix} \in \mathbb{R}^2_{>0}. \tag{3.5}$$

We have the key property

$$\nabla F \cdot \vec{r} = 0 \quad \text{and} \quad \vec{r} \cdot H(F)\vec{r} = 0 \quad \text{for } \vec{r} = \begin{pmatrix} p \\ q \end{pmatrix} \in \mathbb{R}^2_{>0}. \tag{3.6}$$

We further define

$$\vec{R}_t^{N,\mathcal{E}} := \begin{pmatrix} P_t^{N,\mathcal{E}} \\ Q_t^N \end{pmatrix} \quad \text{so that } Y_t^{N,\mathcal{E}} = F(\vec{R}_t^{N,\mathcal{E}}). \tag{3.7}$$

We shall first establish the following proposition, which characterises $P^{N,E}$.

PROPOSITION 3.3. *We have for all* $\mathcal{E} \in \mathcal{B}(\mathbb{K})$ *that*

$$dP_t^{N,\mathcal{E}} = P_t^{N,\mathcal{E}}\left(-\lambda + \frac{N}{N-1}\langle m_t^N, \kappa\rangle\right)dt - \frac{1}{N-1}\langle m_t^{N,\mathcal{E}}, \kappa\phi\rangle\,dt + dM_t^{N,\mathcal{E}}, \tag{3.8}$$

whereby $M^{N,\mathcal{E}}$ *are martingales which satisfy, for all* $\mathcal{E}, \mathcal{F} \in \mathcal{B}(\mathbb{K})$,

$$\begin{aligned} [M^{N,\mathcal{E}}, M^{N,\mathcal{F}}]_t = &-\frac{1}{N}\int_0^t P_s^{N,\mathcal{E}}\langle m_s^{N,\mathcal{F}}, \kappa\phi\rangle + P_s^{N,\mathcal{F}}\langle m_s^{N,\mathcal{E}}, \kappa\phi\rangle\,ds \\ &+ \frac{1}{N}\int_0^t \Lambda_s^{N,\mathcal{E}\cap\mathcal{F}}\,ds + \mathcal{O}_t^{\mathrm{MG}}\left(\frac{P^{\mathcal{E}}P^{\mathcal{F}} + P^{\mathcal{E}\cap\mathcal{F}}}{N^3}\right) \\ &+ \mathcal{O}_t^{\mathrm{FV}}\left(\frac{P^{\mathcal{E}\cap\mathcal{F}}}{N^2}\right) \cap \mathcal{O}_t^{\mathrm{Lip}}. \end{aligned} \tag{3.9}$$

We write M_t^N *for* $M_t^{N,\mathbb{K}}$.

We will then establish Part 1 of Theorem 3.2, followed by the following version of Itô's lemma.

LEMMA 3.4 (Itô's lemma). *We have*

$$\begin{aligned} Y_t^{N,\mathcal{E}} = Y_0^{N,\mathcal{E}} &+ \int_0^t \nabla F(R_{s-}^{N,\mathcal{E}}) \cdot dR_s^{N,\mathcal{E}} + \frac{1}{2}dR_s^{N,\mathcal{E}} \cdot H(R_{s-}^{N,\mathcal{E}})\,dR_s^{N,\mathcal{E}} \\ &+ \mathcal{O}_t^{\mathrm{FV}}\left(\frac{Y^{N,\mathcal{E}}}{N^2}\right) \cap \mathcal{O}_t^{\Delta}\left(\frac{1}{N^3}\right). \end{aligned} \tag{3.10}$$

Combining (3.5) with Proposition 3.3 and Lemma 3.4, we obtain (3.1) by calculation, whereby

$$\mathcal{K}_t^{N,\mathcal{E}} := \int_0^t \frac{1}{Q_{s-}^N}\big(dM_s^{N,\mathcal{E}} - Y_{s-}^{N,\mathcal{E}}\, dM_s^N\big). \tag{3.11}$$

We then obtain (3.2) from (3.9) and (3.11).

Using the boundedness of ϕ and the fact that there are no simultaneous killing events, we obtain (3.3) from (3.1).

Since in parts 1–3 of Theorem 3.2 the statements of processes belonging to various process sequence classes are uniform over all choices $\mathcal{E}, \mathcal{F} \in \mathcal{B}(\mathbb{K})$, Part 4 is immediate.

It remains to prove Proposition 3.3, Part 1 of Theorem 3.2 and Lemma 3.4.

3.1.1. *Proof of Proposition* 3.3. Since N is fixed throughout this proof, we neglect the N superscript for the sake of notation, where it would not create confusion. We recall that τ_n^i represents the nth killing time of particle (X^i, η^i) ($\tau_0^i := 0$), τ_n is the nth killing time of any particle ($\tau_0 := 0$), and $J_t^N := \frac{1}{N} \sup\{n : \tau_n \le t\}$ is the number of killing times up to time t, renormalised by N.

We denote

$$\phi^{\mathcal{E}}(x, \eta) := \phi(x)\mathbb{1}(\eta \in \mathcal{E}), \quad \mathcal{E} \in \mathcal{B}(\mathbb{K}).$$

We define for $\mathcal{E} \in \mathcal{B}(\mathbb{K})$ the processes

$$\begin{aligned}
A_t^{\mathcal{E}} &= \frac{1}{N}\sum_{i=1}^N \sum_{\tau_n^i \le t} \phi^{\mathcal{E}}\big(X_{\tau_n^i}^i, \eta_{\tau_n^i}^i\big) - \frac{N}{N-1}\int_0^t P_s^{\mathcal{E}}\langle m_s^N, \kappa\rangle\, ds \\
&\quad + \frac{1}{N(N-1)}\sum_{i=1}^N \sum_{\tau_n^i \le t} \phi^{\mathcal{E}}\big(X_{\tau_n^i -}^i, \eta_{\tau_n^i -}^i\big), \\
B_t^{\mathcal{E}} &= \langle m_t^{N,\mathcal{E}}, \phi\rangle - \langle m_0^{N,\mathcal{E}}, \phi\rangle - \frac{1}{N}\sum_{i=1}^N \sum_{\tau_n^i \le t} \phi^{\mathcal{E}}\big(X_{\tau_n^i}^i, \eta_{\tau_n^i}^i\big) + \lambda \int_0^t \langle m_s^{N,\mathcal{E}}, \phi\rangle\, ds, \quad \text{and} \\
C_t^{\mathcal{E}} &= \frac{1}{N}\sum_{i=1}^N \sum_{\tau_n^i \le t} \phi^{\mathcal{E}}\big(X_{\tau_n^i -}^i, \eta_{\tau_n^i -}^i\big) - \int_0^t \langle m_s^{N,\mathcal{E}}, \kappa\phi\rangle\, ds.
\end{aligned} \tag{3.12}$$

We will first establish that $A_t^{\mathcal{E}}$, $B_t^{\mathcal{E}}$ and $C_t^{\mathcal{E}}$ are martingales so that

$$\begin{aligned}
M_t^{\mathcal{E}} = A_t^{\mathcal{E}} + B_t^{\mathcal{E}} - \frac{1}{N-1}C_t^{\mathcal{E}} &= -\frac{N}{N-1}\int_0^t P_s^{N,\mathcal{E}}\langle m_s^N, \kappa\rangle\, ds \\
&\quad + \langle m_t^{N,\mathcal{E}}, \phi\rangle - \langle m_0^{N,\mathcal{E}}, \phi\rangle + \lambda\int_0^t \langle m_s^{N,\mathcal{E}}, \phi\rangle\, ds + \frac{1}{N-1}\int_0^t \langle m_s^{N,\mathcal{E}}, \kappa\phi\rangle\, ds
\end{aligned} \tag{3.13}$$

is a martingale. We therefore have (3.8). We will then establish (3.9) by establishing it for $\mathcal{E} = \mathcal{F}$ and for $\mathcal{E}, \mathcal{F}$ disjoint:

$A_t^{\mathcal{E}}$ *is a martingale*. We have that if particle X^i dies at time t, then each $j \neq i$ is selected with probability $\frac{1}{N-1}$ so that the expected value of $\phi^{\mathcal{E}}(X_t^i, \eta_t^i)$ is given by

$$\frac{1}{N-1}\sum_{j \neq i} \phi^{\mathcal{E}}\big(X_{t-}^j, \eta_{t-}^j\big) = \frac{N}{N-1} \times \Big[P_{t-}^{\mathcal{E}} - \frac{1}{N}\phi^{\mathcal{E}}\big(X_{t-}^i, \eta_{t-}^i\big)\Big].$$

Therefore, summing over $\tau_n^i \le t$, we see that

$$\frac{1}{N}\sum_{i=1}^N\sum_{\tau_n^i\le t}\phi^{\mathcal{E}}(X^i_{\tau_n^i},\eta^i_{\tau_n^i})-\frac{N}{N-1}\int_0^t P^{\mathcal{E}}_{s-}\,dJ^N_s+\frac{1}{N(N-1)}\sum_{i=1}^N\sum_{\tau_n^i\le t}\phi^{\mathcal{E}}(X^i_{\tau_n^i-},\eta^i_{\tau_n^i-})$$

is a martingale. We finally note that $J^N_t-\int_0^t\langle m^N_s,\kappa\rangle\,ds$ is a martingale so that $A^{\mathcal{E}}_t$ is a martingale.

$B^{\mathcal{E}}_t$ *is a martingale*. Since $L\phi=-\lambda\phi$, we see that the following is a martingale:

$$B^{\mathcal{E},i,n}_t:=\mathbb{1}(\tau_n^i\le t<\tau_{n+1}^i)\phi^{\mathcal{E}}(X^i_t,\eta^i_t)-\phi^{\mathcal{E}}(X^i_{\tau_n^i},\eta^i_{\tau_n^i})\mathbb{1}(t\ge\tau_n^i)$$
$$+\lambda\int_0^t\mathbb{1}(\tau_n^i\le s<\tau_{n+1}^i)\phi^{\mathcal{E}}(X^i_s,\eta^i_s)\,ds.$$

Therefore, $\sum_{n<n_0}B^{\mathcal{E},i,n}_t$ is a martingale for all $n_0<\infty$. Since $\sum_{n<n_0}|B^{\mathcal{E},i,n}_t|\le C(1+t+NJ^N_t)$ for all $n_0<\infty$, for some $C<\infty$, $(\sum_{n<\infty}B^{\mathcal{E},i,n}_t)_{0\le t<\infty}$ is a martingale. Therefore,

$$B^{\mathcal{E}}_t=\frac{1}{N}\sum_{i=1}^N\sum_{n<\infty}B^{\mathcal{E},i,n}_t\quad\text{is a martingale.}$$

$C^{\mathcal{E}}_t$ *is a martingale*. We have that

$$C^{\mathcal{E},i,n}_t:=\mathbb{1}(t\ge\tau^i_{n+1})\phi^{\mathcal{E}}(X^i_{\tau^i_{n+1}-},\eta^i_{\tau^i_{n+1}-})-\int_0^t\mathbb{1}(\tau_n^i\le s<\tau_{n+1}^i)\kappa(X^i_s)\phi^{\mathcal{E}}(X^i_s,\eta^i_s)\,ds$$

is a martingale. Since for some $C<\infty$, $\sum_{n<n_0}|C^{\mathcal{E},i,n}_t|\le C(1+t+NJ^N_t)$ for all $n_0<\infty$, $\sum_{n<\infty}C^{\mathcal{E},i,n}$ is a martingale. Therefore,

$$C^{\mathcal{E}}_t=\frac{1}{N}\sum_{i=1}^N\sum_{n<\infty}C^{\mathcal{E},i,n}_t\quad\text{is a martingale.}$$

The Quadratic Variation of $M^{\mathcal{E}}$. We observe from (3.13) that $M^{N,\mathcal{E}}_t-\langle m^{N,\mathcal{E}}_t,\phi\rangle$ is a Lipschitz process. By considering seperately the quadartic variation of the continuous motion between jumps and at the jumps, it follows that

$$[M^{N,\mathcal{E}},M^{N,\mathcal{F}}]_t=\frac{1}{N}\int_0^t\langle\Gamma_0(\phi),m^{N,\mathcal{E}\cap\mathcal{F}}_s\rangle\,ds+H^{N,\mathcal{E},\mathcal{F}}_t,$$

whereby we define

$$H^{N,\mathcal{E},\mathcal{F}}_t:=\frac{1}{N^2}\sum_{i=1}^N\sum_{\tau_n^i\le t}[\phi^{\mathcal{E}}(X^i_{\tau_n^i},\eta^i_{\tau_n^i})-\phi^{\mathcal{E}}(X^i_{\tau_n^i-},\eta^i_{\tau_n^i-})][\phi^{\mathcal{F}}(X^i_{\tau_n^i},\eta^i_{\tau_n^i})-\phi^{\mathcal{F}}(X^i_{\tau_n^i-},\eta^i_{\tau_n^i-})].$$

To characterise $H^{N,\mathcal{E},\mathcal{F}}_t$, we split into the cases that $\mathcal{E}$ and $\mathcal{F}$ are disjoint and that $\mathcal{E}=\mathcal{F}$.

$\mathcal{E}=\mathcal{F}$. We write $H^{N,\mathcal{E}}_t$ for $H^{N,\mathcal{E},\mathcal{E}}_t$. At time τ_n^i-, the expected values of $\phi^{\mathcal{E}}(X^i_{\tau_n^i},\eta^i_{\tau_n^i})$ and $\phi^{\mathcal{E}}(X^i_{\tau_n^i},\eta^i_{\tau_n^i})^2$ are

$$P^{\mathcal{E}}_{\tau_n^i-}+\mathcal{O}\Big(\frac{P^{\mathcal{E}}_{\tau_n^i-}+\phi^{\mathcal{E}}(X^i_{\tau_n^i-},\eta^i_{\tau_n^i-})}{N}\Big)\quad\text{and}\quad\langle m^{N,\mathcal{E}}_{\tau_n^i-},\phi^2\rangle+\mathcal{O}\Big(\frac{P^{\mathcal{E}}_{\tau_n^i-}+\phi^{\mathcal{E}}(X^i_{\tau_n^i-},\eta^i_{\tau_n^i-})}{N}\Big),$$

respectively. Therefore, the expected value of $[\phi^{\mathcal{E}}(X^i_{\tau_n^i},\eta^i_{\tau_n^i})-\phi^{\mathcal{E}}(X^i_{\tau_n^i-},\eta^i_{\tau_n^i-})]^2$ at time τ_n^i- is

$$\langle m^{N,\mathcal{E}}_{\tau_n^i-},\phi^2\rangle-2P^{\mathcal{E}}_{\tau_n^i-}\phi^{\mathcal{E}}(X^i_{\tau_n^i-},\eta^i_{\tau_n^i-})+(\phi^{\mathcal{E}}(X^i_{\tau_n^i-},\eta^i_{\tau_n^i-}))^2+\mathcal{O}\Big(\frac{P^{\mathcal{E}}_{\tau_n^i-}+\phi^{\mathcal{E}}(X^i_{\tau_n^i-},\eta^i_{\tau_n^i-})}{N}\Big).$$

Then using the killing rate to characterise the rate at which killing events happen, we see that

$$H_t^{N,\mathcal{E}} - \frac{1}{N}\int_0^t \langle m_s^N,\kappa\rangle\langle m_s^{N,\mathcal{E}},\phi^2\rangle + \langle m_s^{N,\mathcal{E}},\kappa\phi^2\rangle - 2P_s^{N,\mathcal{E}}\langle m_s^{N,\mathcal{E}},\kappa\phi\rangle\,ds$$
$$= \mathcal{O}_t^{\mathrm{FV}}\left(\frac{P^{N,\mathcal{E}}}{N^2}\right)\cap\mathcal{O}_t^{\mathrm{Lip}} + \tilde{M}_t^{N,\mathcal{E}}$$

for some martingale $\tilde{M}_t^{N,\mathcal{E}}$. It is straightforward to then see that, for all N sufficiently large (which does not depend upon $\mathcal{E}$), $[\tilde{M}^{N,\mathcal{E}}]_t = [H^{N,\mathcal{E}}]_t = \mathcal{O}_t^{\mathrm{FV}}(\frac{P^{N,\mathcal{E}}}{N^3})$. We, therefore, obtain (3.9) in the case that $\mathcal{E}=\mathcal{F}$.

$\mathcal{E}\cap\mathcal{F}=\varnothing$. Since $\phi^{\mathcal{E}}(x,\eta)\phi^{\mathcal{F}}(x,\eta)=0$ for all $(x,\eta)\in\bar{D}\times\mathbb{K}$, we have that

$$[\phi^{\mathcal{E}}(X^i_{\tau^i_n},\eta^i_{\tau^i_n}) - \phi^{\mathcal{E}}(X^i_{\tau^i_n-},\eta^i_{\tau^i_n-})][\phi^{\mathcal{F}}(X^i_{\tau^i_n},\eta^i_{\tau^i_n}) - \phi^{\mathcal{F}}(X^i_{\tau^i_n-},\eta^i_{\tau^i_n-})]$$
$$= -\phi^{\mathcal{E}}(X^i_{\tau^i_n-},\eta^i_{\tau^i_n-})\phi^{\mathcal{F}}(X^i_{\tau^i_n},\eta^i_{\tau^i_n}) - \phi^{\mathcal{E}}(X^i_{\tau^i_n},\eta^i_{\tau^i_n})\phi^{\mathcal{F}}(X^i_{\tau^i_n-},\eta^i_{\tau^i_n-}).$$

The expected value of this at time τ^i_n- is then

$$-\frac{N}{N-1}[\phi^{\mathcal{E}}(X^i_{\tau^i_n-},\eta^i_{\tau^i_n-})P^{N,\mathcal{F}}_{\tau^i_n-} + \phi^{\mathcal{F}}(X^i_{\tau^i_n-},\eta^i_{\tau^i_n-})P^{N,\mathcal{E}}_{\tau^i_n-}].$$

It follows that

$$H_t^{N,\mathcal{E},\mathcal{F}} + \frac{1}{N-1}\int_0^t P_s^{N,\mathcal{E}}\langle m_s^{N,\mathcal{F}},\kappa\phi\rangle + P_s^{N,\mathcal{F}}\langle m_s^{N,\mathcal{E}},\kappa\phi\rangle\,ds \quad \text{is a martingale,}$$

which we denote as $\tilde{M}_t^{N,\mathcal{E},\mathcal{F}}$. It is then straightforward to see that $[\tilde{M}^{N,\mathcal{E},\mathcal{F}}]_t = [H^{N,\mathcal{E},\mathcal{F}}]_t = \mathcal{O}_t^{\mathrm{FV}}(P^{N,\mathcal{E}}P^{N,\mathcal{F}})$. We have, therefore, obtained (3.9) with $\mathcal{E}\cap\mathcal{F}=\varnothing$.

Having established (3.9) both in the case that $\mathcal{E}=\mathcal{F}$ and the case that $\mathcal{E}\cap\mathcal{F}=\varnothing$, the case of arbitrary $\mathcal{E},\mathcal{F}$ follows by linearity.

3.1.2. *Proof of part* 1 *of Theorem* 3.2. We recall that F, H and $\vec{R}$ were defined in (3.4), (3.5) and (3.7) as

$$F(\vec{r}) = \frac{p}{q},\qquad H(\vec{r}) = \begin{pmatrix} 0 & -\frac{1}{q^2} \\ -\frac{1}{q^2} & 2\frac{p}{q^3}\end{pmatrix}\quad \text{for } \vec{r} = \begin{pmatrix} p \\ q\end{pmatrix}\in\mathbb{R}^2_{>0},$$

$$\vec{R}_t^{N,\mathcal{E}} := \begin{pmatrix} P_t^{N,\mathcal{E}} \\ Q_t^N\end{pmatrix}\quad \text{so that } Y_t^{N,\mathcal{E}} = F(\vec{R}_t^{N,\mathcal{E}}).$$

We decompose

$$\vec{R}_t^{N,\mathcal{E}} = \vec{R}_t^{N,\mathcal{E},C} + \vec{R}_t^{N,\mathcal{E},J}\quad \text{and}\quad Y_t^{N,\mathcal{E}} = F(\vec{R}_t^{N,\mathcal{E}}) = Y_t^{N,\mathcal{E},C} + Y_t^{N,\mathcal{E},J}$$

for $\mathcal{E}\in\mathcal{B}(\mathbb{K})$, whereby

$$Y_t^{N,\mathcal{E},J} = \sum_{s\le t}\Delta Y_s^{N,\mathcal{E},J}\quad \text{and}\quad \vec{R}_t^{N,\mathcal{E},J} := \sum_{s\le t}\Delta\vec{R}_s^{N,\mathcal{E}} = \begin{pmatrix} P_t^{\mathcal{E},J} \\ Q_t^J\end{pmatrix}.$$

Then by Itô's lemma, we have

$$dY_t^{N,\mathcal{E},C} = \nabla F(\vec{R}_t^{N,\mathcal{E}})\cdot d\vec{R}_t^{N,\mathcal{E},C} + \frac{1}{2}d\vec{R}_t^{N,\mathcal{E},C}\cdot H(F)(\vec{R}_t^{N,\mathcal{E}})\,d\vec{R}_t^{N,\mathcal{E},C}. \tag{3.14}$$

We can, therefore, calculate

$$
\begin{aligned}
&d\big[Y^{N,\mathcal{E},C}, Y^{N,\mathcal{F},C}\big]_t \\
&\quad = \frac{1}{(Q_t^N)^2}\big(dP_t^{N,\mathcal{E},C} - Y_t^{N,\mathcal{E}}\,dQ_t^{N,C}\big)\cdot\big(dP_t^{N,\mathcal{F},C} - Y_t^{N,\mathcal{F}}\,dQ_t^{N,C}\big).
\end{aligned}
\tag{3.15}
$$

Proposition 3.3 implies that

$$
d\big[P^{N,\mathcal{E},C}, P^{N,\mathcal{F},C}\big]_t = \mathcal{O}_t^{\mathrm{FV}}\bigg(\frac{Y^{N,\mathcal{E}}Y^{N,\mathcal{F}}}{N} + \frac{Y^{N,\mathcal{E}\cap\mathcal{F}}}{N^2}\bigg)
\tag{3.16}
$$

for all $\mathcal{E}, \mathcal{F} \in \mathcal{B}(\mathbb{K})$. Combining (3.15) with (3.16), we have

$$
d\big[Y^{N,\mathcal{E},C}, Y^{N,\mathcal{F},C}\big]_t = \mathcal{O}_t^{\mathrm{FV}}\bigg(\frac{Y^{N,\mathcal{E}}Y^{N,\mathcal{F}}}{N}\bigg)
\tag{3.17}
$$

for all $\mathcal{E}, \mathcal{F} \in \mathcal{B}(\mathbb{K})$ disjoint. We also have that

$$
\big[Y^{N,\mathcal{E},J}, Y^{N,\mathcal{F},J}\big]_t = \sum_{\tau_n^i \le t} \Delta Y_{\tau_n^i}^{N,\mathcal{E}} \Delta Y_{\tau_n^i}^{N,\mathcal{F}}.
$$

Since Q_t^N is bounded below away from 0, by bounding the partial derivatives of F we can calculate for all $\mathcal{E}, \mathcal{F} \in \mathcal{B}(\mathbb{K})$ disjoint that

$$
\begin{aligned}
&\big|\Delta Y_{\tau_n^i}^{N,\mathcal{E}} \Delta Y_{\tau_n^i}^{N,\mathcal{F}}\big| \\
&\quad = \mathcal{O}\bigg(\big|\Delta P_{\tau_n^i}^{N,\mathcal{E}} \Delta P_{\tau_n^i}^{N,\mathcal{F}}\big| + \frac{P_{\tau_n^i-}^{N,\mathcal{E}}|\Delta P_{\tau_n^i}^{N,\mathcal{F}}| + P_{\tau_n^i-}^{N,\mathcal{F}}|\Delta P_{\tau_n^i}^{N,\mathcal{E}}|}{N} + \frac{P_{\tau_n^i-}^{N,\mathcal{E}} P_{\tau_n^i-}^{N,\mathcal{F}}}{N^2}\bigg) \\
&\quad = \mathcal{O}\bigg(\frac{|\phi^{\mathcal{E}}(X_{\tau_n^i}^i, \eta_{\tau_n^i}^i)\phi^{\mathcal{F}}(X_{\tau_n^i-}^i, \eta_{\tau_n^i-}^i)|}{N^2}\bigg) + \mathcal{O}\bigg(\frac{|\phi^{\mathcal{F}}(X_{\tau_n^i}^i, \eta_{\tau_n^i}^i)\phi^{\mathcal{E}}(X_{\tau_n^i-}^i, \eta_{\tau_n^i-})|}{N^2}\bigg) \\
&\qquad + \mathcal{O}\bigg(\frac{P_{\tau_n^i-}^{N,\mathcal{E}} P_{\tau_n^i-}^{N,\mathcal{F}}}{N^2}\bigg) + \mathcal{O}\bigg(\frac{|\phi^{\mathcal{E}}(X_{\tau_n^i}^i, \eta_{\tau_n^i}^i) - \phi^{\mathcal{E}}(X_{\tau_n^i-}^i, \eta_{\tau_n^i-}^i)|}{N^2} P_{\tau_n^i-}^{N,\mathcal{F}}\bigg) \\
&\qquad + \mathcal{O}\bigg(\frac{|\phi^{\mathcal{F}}(X_{\tau_n^i}^i, \eta_{\tau_n^i}^i) - \phi^{\mathcal{F}}(X_{\tau_n^i-}^i, \eta_{\tau_n^i-}^i)|}{N^2} P_{\tau_n^i-}^{N,\mathcal{E}}\bigg).
\end{aligned}
$$

Since κ is bounded, it is straightforward to then see that

$$
\sum_{\tau_n^i \le t} \big|\Delta Y_{\tau_n^i}^{N,\mathcal{E}} \Delta Y_{\tau_n^i}^{N,\mathcal{F}}\big| = \mathcal{O}_t^{\mathrm{FV}}\bigg(\frac{P^{N,\mathcal{E}}P^{N,\mathcal{F}}}{N}\bigg) = \mathcal{O}_t^{\mathrm{FV}}\bigg(\frac{Y^{N,\mathcal{E}}Y^{N,\mathcal{F}}}{N}\bigg)
$$

so that

$$
\begin{aligned}
\big[Y^{N,\mathcal{E},J}, Y^{N,\mathcal{F},J}\big]_t &= \sum_{\tau_n^i \le t} \big|\Delta Y_{\tau_n^i}^{N,\mathcal{E}} \Delta Y_{\tau_n^i}^{N,\mathcal{F}}\big| \\
&= \mathcal{O}_t^{\mathrm{FV}}\bigg(\frac{P^{N,\mathcal{E}}P^{N,\mathcal{F}}}{N}\bigg) = \mathcal{O}_t^{\mathrm{FV}}\bigg(\frac{Y^{N,\mathcal{E}}Y^{N,\mathcal{F}}}{N}\bigg)
\end{aligned}
\tag{3.18}
$$

for all $\mathcal{E}, \mathcal{F} \in \mathcal{B}(\mathbb{K})$ disjoint. Combining (3.17) with (3.18), we have Part 1 of Theorem 3.2.

3.1.3. *Proof of Lemma 3.4.* We take $0 \le t_0 \le t_1 \le t$ and write

$$Y^{N,\mathcal{E},J}_{t_1} - Y^{N,\mathcal{E},J}_{t_0} = \sum_{t_0 < s \le t_1} (Y^{N,\mathcal{E}}_s - Y^{N,\mathcal{E}}_{s-}).$$

We may calculate

$$\frac{\partial^3 F}{\partial p^3} = \frac{\partial^3 F}{\partial^2 p \partial q} = 0, \qquad \frac{\partial^3 F}{\partial p \partial^2 q} = \frac{2}{q^3}, \qquad \frac{\partial^3 F}{\partial^3 q} = \frac{-6p}{q^4}. \tag{3.19}$$

Thus, by Taylor's theorem, (3.19), the fact that almost surely there are no simultaneous killing events and the fact that Q^N_t is bounded above and below away from 0, we have

$$\begin{aligned}&\Big| Y^{N,\mathcal{E},J}_s - Y^{N,\mathcal{E},J}_{s-} - \nabla F(\vec{R}^N_{s-}) \cdot (\vec{R}^{N,J}_s - \vec{R}^{N,J}_{s-}) \\ &\qquad - \frac{1}{2}(\vec{R}^{N,J}_s - \vec{R}^{N,J}_{s-}) \cdot H(F)(\vec{R}^N_{s-})(\vec{R}^{N,J}_s - \vec{R}^{N,J}_{s-})\Big| \\ &\quad = \mathcal{O}(P^{N,\mathcal{E}}_{s-}|\Delta Q^N_{s-}|^3 + |\Delta P^{N,\mathcal{E}}_{s-}||\Delta Q^N_{s-}|^2).\end{aligned}$$

Since κ and ϕ are bounded, it is straightforward to then see that

$$\begin{aligned}&Y^{N,\mathcal{E},J}_t - Y^{N,\mathcal{E},J}_0 - \int_0^t \nabla F(\vec{R}^N_{s-}) \cdot d\vec{R}^{N,J}_s - \frac{1}{2}\int_0^t d\vec{R}^{N,J}_s \cdot H(F)(\vec{R}^N_{s-})\, d\vec{R}^{N,J}_s \\ &\quad = \mathcal{O}^{\mathrm{FV}}_t\Big(\frac{Y^{N,\mathcal{E}}}{N^3}\Big) \cap \mathcal{O}^{\Delta}\Big(\frac{1}{N^3}\Big).\end{aligned} \tag{3.20}$$

Combining this with (3.14), we have Lemma 3.4.

This completes the proof of Theorem 3.2.

4. Proof of Theorem 1.4. With the calculations of Section 3 in hand, we now prove Theorem 1.4. We shall make use of the Wasserstein distance W and the weak atomic metric W_a, which are defined in Appendix C.

We shall firstly prove the following proposition.

PROPOSITION 4.1. *For all $\mathcal{E} \in \mathcal{B}(\mathbb{K})$ and $f \in C_b(\bar{D})$, we have that*

$$(m^{N,\mathcal{E}}_t - Y^{N,\mathcal{E}}_t \pi)(f) \to 0 \quad \text{in probability as } t \wedge N \to \infty. \tag{4.1}$$

In particular, taking $f = 1$, we have for any $\mathcal{E} \in \mathcal{B}(\mathbb{K})$ that

$$\chi^N_t(\mathcal{E}) - \mathcal{Y}^N_t(\mathcal{E}) \to 0 \quad \text{in probability as } t \wedge N \to \infty. \tag{4.2}$$

Heuristically, this says that over an $\mathcal{O}(1)$ timescale, the number of particles whose color belongs to $\mathcal{E}$ is given by $Y^{N,\mathcal{E}}_t$, and the spatial distribution of these particles is given by π, for any $\mathcal{E} \in \mathcal{B}(\mathbb{K})$.

Using Proposition 4.1 and the calculations of Section 3, we will then establish that $(\mathcal{Y}^N_{Nt})_{0 \le t \le T}$ converges in distribution to the Wright–Fisher process of rate Θ.

PROPOSITION 4.2. *We take some deterministic initial profile $\nu^0 \in \mathcal{P}(\mathbb{K})$ and define $(\nu_t)_{0 \le t < \infty}$ to be a Wright–Fisher process of rate Θ and initial condition $\nu_0 := \nu^0$. We then consider a sequence of Fleming–Viot multicolor processes $(\vec{X}^N_t, \vec{\eta}^N_t)_{0 \le t < \infty}$. We assume that $\mathcal{Y}^N_0 \to \nu^0$ in W_a in probability.*

We fix $T < \infty$ and rescale time by $t \mapsto Nt$. We then have the convergence

$$(\mathcal{Y}^N_{Nt})_{0 \le t \le T} \to (\nu_t)_{0 \le t \le T} \quad \text{in } D([0,T]; \mathcal{P}_{\mathrm{W}}(\mathbb{K})) \text{ in distribution as } N \to \infty. \tag{4.3}$$

We recall, in particular, that $(\nu_t)_{0\leq t\leq T}\in C([0,T];\mathcal{P}_{\mathrm{W}}(\mathbb{K}))$ almost surely by Theorem D.2. We now take a sequence $(\vec{t}^N)_{2\leq N<\infty}=((t_1^N,\dots,t_n^N))_{2\leq t\leq N}$ converging to $\vec{t}=(t^1,\dots,t^n)$, as in the statement of Theorem 1.4. It follows that

$$(\mathcal{Y}^N_{Nt_1^N},\dots,\mathcal{Y}^N_{Nt_n^N})\to(\nu_{t_1},\dots,\nu_{t_n})\quad\text{in }\big(\mathcal{P}_{\mathrm{W}}(\mathbb{K})\big)^n\text{ in distribution as }N\to\infty.$$

Recalling the positivity and boundedness of ϕ from Theorem A.1, we observe that

$$\chi_t^N\leq C\mathcal{Y}_t^N\quad\text{for all }t\geq 0,N\in\mathbb{N},\text{ for some fixed uniform constant }C<\infty. \tag{4.4}$$

We now fix $1\leq k\leq n$. Since $(\mathcal{Y}^N_{Nt_k^N})_{N\geq 1}$ is a tight sequence of random measures, it follows from (4.4) that $(\chi^N_{Nt_k^N})_{N\geq 1}$ must also be a tight sequence of random measures. It, therefore, follows from (4.2) and Lemma C.2 that $\mathrm{W}(\mathcal{Y}^N_{Nt_k^N},\chi^N_{Nt_k^N})\to 0$ in probability as $N\to\infty$. We have, therefore, established that

$$(\chi^N_{Nt_1^N},\dots,\chi^N_{Nt_n^N})\to(\nu_{t_1},\dots,\nu_{t_n})\quad\text{in }\big(\mathcal{P}_{\mathrm{W}}(\mathbb{K})\big)^n\text{ in distribution as }N\to\infty.$$

We have left only to strengthen the notion of convergence to convergence in the weak atomic metric. After proving propositions 4.1 and 4.2, we shall establish the following proposition.

PROPOSITION 4.3. *We recall that* $\Psi(u):=(1-u)\vee 0$ *is the function used to define the* W_a *metric in Appendix* C.2. *For all* $\delta>0$, *there exists* $\epsilon>0$ *such that*

$$\inf_N\mathbb{P}\bigg(\sup_{0\leq t\leq T}\sum_{\substack{k,\ell\in\mathbb{K}\\ k\neq\ell}}\chi^N_{Nt}(\{k\})\chi^N_{Nt}(\{\ell\})\Psi\Big(\frac{d(k,\ell)}{\epsilon}\Big)\leq\delta\bigg)\geq 1-\delta.$$

Note that the above sum is well defined, as the terms are nonzero only for $k,\ell\in\operatorname{supp}(\chi_0^N)$.

We may, therefore, apply the compact containment condition, Lemma C.5, to conclude that $\{\mathcal{L}(\chi^N_{Nt_k^N})\}$ is tight in $\mathcal{P}(\mathcal{P}_{\mathrm{W}_a}(\mathbb{K}))$ for all $1\leq k\leq n$ so that we have Theorem 1.4.

We have left to prove propositions 4.1, 4.2 and 4.3

4.1. *Proof of Proposition* 4.1. We fix $\mathcal{E}\in\mathcal{B}(\mathbb{K})$ and $f\in C_b(\bar{D})$. We write

$$\psi_t^N:=\frac{1}{N}\sum_{i=1}^N\delta_{(X_t^i,\eta_t^i)}\in\mathcal{P}(\bar{D}\times\mathbb{K}),\quad 0\leq t<\infty,N\in\mathbb{N}\quad\text{and}\quad f^{\mathcal{E}}(x,\eta):=f(x)\mathbb{1}(\eta\in\mathcal{E}).$$

We take the $\bar{D}\times\mathbb{K}$-valued killed strong Markov process $((X_t,\eta_t))_{0\leq t<\tau_\partial}$ defined in Definition 1.7. It follows from Theorem 1.8 that there exists $c_t\to 0$ as $t\to\infty$ such that, for all $N<\infty$ and initial conditions $(\vec{X}_0^N,\vec{\eta}_0^N)$, we have that

$$\big\|\mathcal{L}_{\psi_0^N}\big((X_t,\eta_t)|\tau_\partial>t\big)-\pi\otimes\mathcal{Y}_0^N\big\|_{\mathrm{TV}}\leq c_t,\quad 0\leq t<\infty. \tag{4.5}$$

It follows from Theorem 1.8 that, for all $t<\infty$ and $N\geq 2$, there exists $C_{t,N}<\infty$ such that

$$\mathbb{E}_{(\vec{X}_0^N,\vec{\eta}_0^N)}\big[\big|\big(\psi_t^N-\mathcal{L}_{\psi_0^N}\big((X_t,\eta_t)\big|\tau_\partial>t\big)\big)(f^{\mathcal{E}})\big|\big]\leq C_{t,N}\|f\|_\infty, \tag{4.6}$$

for any initial condition $(\vec{X}_0^N,\vec{\eta}_0^N)$, with $C_{t,N}\to 0$ as $N\to\infty$ for fixed $t<\infty$. On the other hand, we observe that

$$\psi_t^N(f^{\mathcal{E}})=m_t^{N,\mathcal{E}}(f),\qquad(\pi\otimes\mathcal{Y}_t^N)(f^{\mathcal{E}})=\frac{\sum_{i=1}^N\phi(X_0^{N,i})\pi\otimes\delta_{\eta_0^{N,i}}}{\sum_{i=1}^N\phi(X_0^{N,i})}(f^{\mathcal{E}})=Y_0^{N,\mathcal{E}}\pi(f).$$

Therefore, combining (4.5) with (4.6), we obtain that

$$\mathbb{E}\big[\big|(m_t^{N,\mathcal{E}}-Y_0^{N,\mathcal{E}}\pi)(f)\big|\big]\leq(C_{t,N}+c_t)\|f\|_\infty.$$

Proposition 4.1 then follows by applying (3.3).

4.2. *Proof of Proposition* 4.2. Our proof proceeds in the following *two* steps:

1. We fix $\epsilon > 0$ and take $\{k_1, k_2, \dots\}$ to be a dense subset of $\mathbb{K}$. Then for all i we can find $\frac{\epsilon}{2} < r_i < \epsilon$ such that $\nu^0(\partial B(k_i, r_i)) = 0$. We set $A_i = B(k_i, r_i) \setminus (\bigcup_{j=1}^{i-1} A_j)$. Since the disjoint union of A_i is $\mathbb{K}$, we can find $n < \infty$ such that $\nu^0((\bigcup_{i=1}^{n} A_i)^c) < \epsilon$. We set $A_0 := (\bigcup_{i=1}^{n} A_i)^c$ and pick arbitrary $k_0 \in \mathbb{K}$.

We shall prove that $(Y_{Nt}^{N,A_0}, \dots, Y_{Nt}^{N,A_n})_{0 \le t \le T}$ converges in $D([0,T]; \mathbb{R}^{n+1})$ in distribution to a Wright–Fisher diffusion of rate Θ and initial condition $(\nu^0(A_0), \dots, \nu^0(A_n))$.

2. We then use this to prove that

$$(\mathcal{Y}_{Nt}^N)_{0 \le t \le T} \to (\nu_t)_{0 \le t \le T} \quad \text{in } D\big([0,T]; \mathcal{P}_{\mathrm{W}}(\mathbb{K})\big) \text{ in distribution.} \tag{4.7}$$

Step 1. We recall that the martingale $\mathcal{K}_t^{N,\mathcal{E}}$ was defined in Theorem 3.2, whilst $\Lambda_t^{N,\mathcal{E}}$ and Λ_t^N were defined in (2.2) to be given by

$$\Lambda_t^{N,\mathcal{E}} := \langle m_t^{N,\mathcal{E}}, \Gamma_0(\phi) + \kappa\phi^2\rangle + \langle m_t^{N,\mathcal{E}}, \phi^2\rangle\langle m_t^N, \kappa\rangle \quad \text{for } \mathcal{E} \in \mathcal{B}(\mathbb{K}), \text{ and } \Lambda_t^N := \Lambda_t^{N,\mathbb{K}}.$$

We further define

$$(\vec{Y}_{Nt}^N)_{0 \le t \le T} := \big((Y_{Nt}^{N,A_0}, \dots, Y_{Nt}^{N,A_n})\big)_{0 \le t \le T}.$$

We will now verify that $\{\mathcal{L}((\vec{Y}_{Nt}^N)_{0 \le t \le T})\}$ is tight in $\mathcal{P}(D([0,T]; \mathbb{R}^{n+1}))$ by using Aldous' criterion [1], Theorem 1. Since $0 \le Y_{Nt}^{N,A_i} \le 1$, [1], Condition (3), is satisfied.

We now take a sequence $(\tau_N, \delta_N)_{N=1}^{\infty}$ of stopping times τ_N and constants $\delta_n > 0$, satisfying [1], Condition (1), for the purpose of checking [1], Condition (A). In particular, we have by (3.3) that, for some $F^N = \mathcal{O}^{\mathrm{FV}}(1)$ and $M^N = \mathcal{O}^{\mathrm{MG}}(1)$, we have

$$Y_{N(\tau_N+\delta_N)}^{N,A_i} - Y_{N\tau_N}^{N,A_i} = F_{\tau_N+\delta_N}^N - F_{\tau_N}^N + Z_{\tau_N+\delta_N}^N - Z_N \to 0 \quad \text{in probability.}$$

Thus, $\{(\vec{Y}_{Nt}^N)_{0 \le t \le T}\}$ satisfies [1], Condition (A),

$$\vec{Y}_{N(\tau_N+\delta_N)}^N - \vec{Y}_{N\tau_N}^N \to 0 \quad \text{in probability,}$$

and hence, $\{\mathcal{L}((\vec{Y}_{Nt}^N)_{0 \le t \le T})\}$ is tight in $\mathcal{P}(D([0,T]; \mathbb{R}^{n+1}))$ by [1], Theorem 1. For all $\mathcal{E} \in \mathcal{B}(\mathbb{K})$, (4.1) implies that

$$\begin{aligned} &\Lambda_t^{N,\mathcal{E}} - Y_t^{N,\mathcal{E}}\big[\langle \pi, \Gamma_0(\phi) + \kappa\phi^2\rangle + \langle \pi, \phi^2\rangle\langle \pi, \kappa\rangle\big], \qquad Q_t^N - \langle \pi, \phi\rangle \xrightarrow{p} 0 \\ &\text{as } t \wedge N \to \infty. \end{aligned} \tag{4.8}$$

Then applying (4.1) and Fubini's theorem to (3.1), we obtain

$$\sup_{0 \le t \le T} \big|(Y_{Nt}^{N,\mathcal{E}} - Y_0^{N,\mathcal{E}}) - (\mathcal{K}_{Nt}^{N,\mathcal{E}} - \mathcal{K}_0^{N,\mathcal{E}})\big| \to 0 \quad \text{in probability as } N \to \infty. \tag{4.9}$$

We consider a subsequential limit in distribution of $\{(\vec{Y}_{Nt}^N)_{0 \le t \le T}\}$,

$$(\vec{Y}_t)_{0 \le t \le T} = \big((Y_t^{A_0}, \dots, Y_t^{A_n})\big)_{0 \le t \le T},$$

which by Part 3 of Theorem 3.2 must have continuous paths. Using (4.9), we conclude that $(\mathcal{K}_{Nt}^{N,A_0}, \dots, \mathcal{K}_{Nt}^{N,A_n})_{0 \le t \le T}$ converges in $D([0,T]; \mathbb{R}^{n+1})$ in distribution along this subsequence to

$$(\vec{Y}_t - \vec{Y}_0)_{0 \le t \le T}.$$

Since $(\mathcal{K}_{Nt}^{N,A_0},\dots,\mathcal{K}_{Nt}^{N,A_n})_{0\le t\le T}$ is a martingale for each N, $(Y_t^{A_0},\dots,Y_t^{A_n})_{0\le t\le T}$ is a martingale with respect to its natural filtration σ_t. We then obtain from (3.2) that, for all $0\le i,j\le n$,

$$\mathcal{K}_{Nt}^{N,A_i}\mathcal{K}_{Nt}^{N,A_j}-\int_0^t\frac{1}{(Q_s^N)^2}\big[\mathbb{1}(i=j)\Lambda_{Ns}^{N,A_i}-Y_{Ns}^{N,A_i}\Lambda_{Ns}^{N,A_j}$$
$$-Y_{Ns}^{N,A_j}\Lambda_{Ns}^{N,A_i}+Y_{Ns}^{N,A_i}Y_{Ns}^{N,A_j}\Lambda_{Ns}^{N}\big]\,ds-\mathcal{O}_t^{\mathrm{MG}}\bigg(\frac{Y_{N\cdot}^{N,A_i}Y_{N\cdot}^{N,A_j}+\mathbb{1}(i=j)Y_{N\cdot}^{N,A_i}}{N^2}\bigg)$$
$$-\mathcal{O}_t^{\mathrm{FV}}\bigg(\frac{Y_{N\cdot}^{N,A_i}Y_{N\cdot}^{N,A_j}+\mathbb{1}(i=j)Y_{N\cdot}^{N,A_i\cap A_j}}{N}\bigg)\cap\mathcal{O}_t^{\mathrm{Cts}}$$

is a martingale for all N so that, by (4.8) and (4.9),

$$Y_t^{A_i}Y_t^{A_j}-\int_0^t\frac{\langle\pi,\Gamma_0(\phi)+\kappa\phi^2\rangle+\langle\pi,\phi^2\rangle\langle\pi,\kappa\rangle}{\langle\pi,\phi\rangle^2}\big(\mathbb{1}(i=j)Y_s^{A_i}-Y_s^{A_i}Y_s^{A_j}\big)\,ds$$

is a $(\sigma_t)_{t\ge 0}$-martingale. Thus,

$$[Y^{A_i},Y^{A_j}]_t=\int_0^t\frac{\langle\pi,\Gamma_0(\phi)+\kappa\phi^2\rangle+\langle\pi,\phi^2\rangle\langle\pi,\kappa\rangle}{\langle\pi,\phi\rangle^2}\big(\mathbb{1}(i=j)Y_s^{A_i}-Y_s^{A_i}Y_s^{A_j}\big)\,ds.$$

We have that

$$\begin{aligned}\langle\pi,\Gamma_0(\phi)+\kappa\phi^2\rangle=\langle\pi,L(\phi^2)-2\phi L(\phi)\rangle=\lambda\langle\pi,\phi^2\rangle\quad\text{and}\\ \langle\pi,\kappa\rangle=\langle\pi,-L(1)\rangle=\lambda.\end{aligned}\tag{4.10}$$

Since $\nu^0(\partial A_i)=0$ for all $0\le i\le n$,

$$\vec{Y}_0^N\to\big(\nu^0(A_0),\dots,\nu^0(A_n)\big)\quad\text{in probability.}$$

Thus, each subsequential limit $(\vec{Y}_t)_{0\le t\le T}$ must be a solution of the $n+1$-type Wright–Fisher diffusion of rate Θ with initial condition $(\nu^0(A_0),\dots,\nu^0(A_n))$, which is unique in law. Therefore, we have convergence of the whole sequence in $D([0,T];\mathbb{R}^{n+1})$ in distribution to this Wright–Fisher diffusion.

Step 2. Whereas we use W to denote the Wasserstein metric on $\mathcal{P}(\mathbb{K})$ generated by $d\wedge 1$, we metrise $\mathcal{P}(D([0,T];\mathcal{P}_{\mathrm{W}}(\mathbb{K}))))$ using the Wasserstein-1 metric generated by the metric $d_{D([0,T];\mathcal{P}_{\mathrm{W}}(\mathbb{K}))}\wedge 1$, which we denote as $\overline{\mathrm{W}}$. We take $\epsilon_\ell\to 0$, giving $k_0^\ell,k_1^\ell,\dots,k_{n_\ell}^\ell\in\mathbb{K}$ for each $\ell\in\mathbb{N}$, as provided for in Step 1. We define, for each $\ell\in\mathbb{N}$, the projection

$$\mathbf{P}^\ell:\mathcal{P}(\bar{D})\ni\mu\mapsto\sum_{j=0}^{n_\ell}\mu(A_{k_j^\ell})\delta_{k_j^\ell}\in\mathcal{P}(\bar{D}).$$

We write $\nu_t^\ell:=\mathbf{P}^\ell(\nu_t)$ for all $0\le t<\infty$. It is immediate that

$$\mathrm{W}(\mathbf{P}^\ell(\mu),\mu)\le\epsilon+\mu(A_0^\ell)\quad\text{for all }\mu\in\mathcal{P}(\bar{D}).$$

Proposition 1.3 implies that we can write

$$\nu_t^\ell=\sum_{j=0}^{n_\ell}p_t^{\ell,j}\delta_{k_{n_\ell}^\ell},\quad 0\le t<\infty,$$

whereby $(p_t^{\ell,0},\dots,p_t^{\ell,n_\ell})_{0\le t<\infty}$ is an n_ℓ-type Wright–Fisher diffusion of rate Θ and initial condition $(\nu^0(A_{k_0^\ell}),\dots,\nu^0(A_{k_{n_\ell}^\ell}))$. Step 1, therefore, implies that

$$\bar{\mathrm{W}}\big(\mathcal{L}((\mathbf{P}^\ell(\mathcal{Y}_{Nt}^N))_{0\le t\le T}),\mathcal{L}((\nu_t^\ell)_{0\le t\le T})\big)\to 0\quad\text{as }N\to\infty.$$

Therefore, by the triangle inequality we have

$$
\begin{aligned}
&\limsup_{N\to\infty} \bar{\mathrm{W}}\big(\mathcal{L}((\mathcal{Y}^N_{Nt})_{0\le t\le T}), \mathcal{L}((\nu_t)_{0\le t\le T})\big) \\
&\quad\le \limsup_{N\to\infty} \bar{\mathrm{W}}\big(\mathcal{L}((\mathcal{Y}^N_{Nt})_{0\le t\le T}), \mathcal{L}((\mathbf{P}^\ell(\mathcal{Y}^N_{Nt}))_{0\le t\le T})\big) \\
&\qquad + \bar{\mathrm{W}}\big(\mathcal{L}((\nu^\ell_t)_{0\le t\le T}), \mathcal{L}((\nu_t)_{0\le t\le T})\big) \\
&\quad\le 2\epsilon_\ell + \limsup_{N\to\infty} \mathbb{E}\Big[\sup_{0\le t\le T} Y^{N,A^\ell_0}_{Nt}\Big] + \mathbb{E}\Big[\sup_{0\le t\le T} \nu_t(A^\ell_0)\Big].
\end{aligned}
\tag{4.11}
$$

Using again Step 1, we have that

$$
\limsup_{N\to\infty} \mathbb{E}\Big[\sup_{0\le t\le T} Y^{N,A^\ell_0}_{Nt}\Big] = \mathbb{E}\Big[\sup_{0\le t\le T} \nu_t(A^\ell_0)\Big].
$$

Since $(\nu_t(A_0))_{0\le t\le T}$ is a Wright–Fisher diffusion of rate Θ and initial condition $\nu^0(A^\ell_0) < \epsilon_l$,

$$
\mathbb{E}\Big[\sup_{0\le t\le T} \nu_t(A^\ell_0)\Big] \to 0 \quad \text{as } \ell\to\infty.
$$

Therefore, taking $\limsup_{\ell\to\infty}$ of both sides of (4.11), we obtain (4.7).

4.3. *Proof of Proposition* 4.3. It follows from (4.4) that it suffices to verify the following condition.

CONDITION 4.4. For every $\delta > 0$, there exists $\epsilon > 0$ such that

$$
\limsup_{N\to\infty} \mathbb{P}\bigg(\sup_{0\le t\le T} \sum_{\substack{k,\ell\in\mathbb{K}\\ k\ne\ell}} Y^{N,\{k\}}_{Nt} Y^{N,\{\ell\}}_{Nt} \Psi\Big(\frac{d(k,\ell)}{\epsilon}\Big) \le \delta\bigg) \ge 1-\delta.
\tag{4.12}
$$

We calculate using parts 1, 3 and 4 of Theorem 3.2 that

$$
\begin{aligned}
d\big(Y^{N,\{k\}}_t Y^{N,\{\ell\}}_t\big) &= Y^{N,\{k\}}_{t-}\, dY^{N,\{\ell\}}_t + Y^{N,\{\ell\}}_{t-}\, dY^{N,\{k\}}_t + d\big[Y^{N,\{k\}}, Y^{N,\{\ell\}}\big]_t \\
&= Y^{N,\{k\}}_{t-}\Big[d\mathcal{O}^{\mathrm{FV}}_t\Big(\frac{Y^{N,\{\ell\}}}{N}\Big) + d\mathcal{O}^{\mathrm{MG}}_t\Big(\frac{Y^{N,\{\ell\}}}{N}\Big)\Big] \\
&\quad + Y^{N,\{\ell\}}_{t-}\Big[d\mathcal{O}^{\mathrm{FV}}_t\Big(\frac{Y^{N,\{k\}}}{N}\Big) + d\mathcal{O}^{\mathrm{MG}}_t\Big(\frac{Y^{N,\{k\}}}{N}\Big)\Big] + d\mathcal{O}^{\mathrm{FV}}_t\Big(\frac{Y^{N,\{k\}}Y^{N,\{\ell\}}}{N}\Big) \\
&= d\mathcal{O}^{\mathrm{FV}}_t\Big(\frac{Y^{N,\{k\}}Y^{N,\{\ell\}}}{N}\Big) + d\mathcal{O}^{\mathrm{MG}}_t\Big(\frac{Y^{N,\{k\}}Y^{N,\{\ell\}}}{N}\Big),
\end{aligned}
$$

uniformly over all random $k, \ell \in \mathrm{supp}(\mathcal{Y}^N_0)$. Thus,

$$
\sum_{k,\ell\in\mathbb{K}} Y^{N,\{k\}}_{Nt} Y^{N,\{\ell\}}_{Nt} = \mathcal{O}^{\mathrm{FV}}_t\Big(\sum_{k,\ell\in\mathbb{K}} Y^{N,\{k\}}_{N\cdot} Y^{N,\{\ell\}}_{N\cdot}\Big) + (\text{martingale})_t.
$$

Therefore, using Gronwall's inequality, there exists uniform $C < \infty$ such that

$$
e^{-Ct} \sum_{\substack{k,\ell\in\mathbb{K}\\ k\ne\ell}} Y^{N,\{k\}}_{Nt} Y^{N,\{\ell\}}_{Nt} \Psi\Big(\frac{d(k,\ell)}{\epsilon}\Big)
\tag{4.13}
$$

is a supermartingale for all N large enough. Therefore, we have for all N large enough that

$$\mathbb{P}\Big(\sup_{0\le t\le T}\sum_{\substack{k,\ell\in\mathbb{K}\\ k\neq\ell}} Y_{Nt}^{N,\{k\}}Y_{Nt}^{N,\{\ell\}}\Psi\Big(\frac{d(k,\ell)}{\epsilon}\Big)\le\delta\Big)$$

$$\ge\mathbb{P}\Big(\sup_{0\le t\le T} e^{-Ct}\sum_{\substack{k,\ell\in\mathbb{K}\\ k\neq\ell}} Y_{Nt}^{N,\{k\}}Y_{Nt}^{N,\{\ell\}}\Psi\Big(\frac{d(k,\ell)}{\epsilon}\Big)\le e^{-CT}\delta\Big)$$

$$\ge 1-\frac{1}{e^{-CT}\delta}\mathbb{E}\Big[\sum_{\substack{k,\ell\in\mathbb{K}\\ k\neq\ell}} Y_{0}^{N,\{k\}}Y_{0}^{N,\{\ell\}}\Psi\Big(\frac{d(k,\ell)}{\epsilon}\Big)\Big].$$

We have assumed that the initial conditions $\mathcal{Y}_0^N$ converge in the weak atomic metric, so Lemma C.5 implies that

$$\sup_N\mathbb{E}\Big[\sum_{\substack{k,\ell\in\mathbb{K}\\ k\neq\ell}} Y_{0}^{N,\{k\}}Y_{0}^{N,\{\ell\}}\Psi\Big(\frac{d(k,\ell)}{\epsilon}\Big)\Big]\to 0\quad\text{as }\epsilon\to 0.$$

We have, therefore, verified Condition 4.4 and hence established Proposition 4.3.
This concludes the proof of Theorem 1.4.

5. Extension to the hard killing case. The Fleming–Viot particle system was first introduced by Burdzy, Holyst and March [16] in the case of Brownian dynamics with instantaneous killing at the boundary (*hard killing*). In this section we extend Theorem 1.4 to this setting.

We assume throughout this section that D is a bounded, connected, nonempty, open subset of $\mathbb{R}^d$ with C^∞ boundary. The process $(B_t)_{0\le t<\tau}$ evolves as a Brownian motion in D, killed at the time $\tau_\partial:=\inf\{t>0:B_{t-}\in\partial D\}$. The color space $\mathbb{K}$ remains an arbitrary complete, seperable metric space.

The Fleming–Viot particle system and Fleming–Viot multicolor process are then defined as before, except that the particles evolve as independent Brownian motions between killing times and are killed instantaneously upon contact with the boundary ∂D (i.e., when $B_{t-}^{N,i}\in\partial D$) rather than according to a Poisson clock. In particular, we define the following.

DEFINITION 5.1 (Fleming–Viot multicolor process with hard killing). The Fleming–Viot multicolor process, $((\vec{B}_t^N,\vec{\eta}_t^N))_{0\le t<\infty}=\{(B_t^{N,i},\eta_t^{N,i})_{0\le t<\infty}:i=1,\dots,N\}$, is a $(D\times\mathbb{K})^N$-valued process defined as follows:

(5.1)
- (*i*) Initial condition: $((B_0^{N,1},\eta_0^{N,1}),\dots,(B_0^{N,N},\eta_0^{N,N}))\sim\upsilon^N\in\mathcal{P}((D\times\mathbb{K})^N)$.
- (*ii*) For $t\in[0,\infty)$ and between killing times, the particles $(B_t^{N,i},\eta_t^{N,i})$ evolve as Brownian motions in the first variable, and are constant in the second variable.
- (*iii*) The particle $(B_t^{N,i},\eta_t^{N,i})$ is killed instantaneously whenever the first variable makes contact with the boundary, that is, when $B_t^{N,i}\in\partial D$. We write τ_k^i for the death times of particle$(B^{N,i},\eta^{N,i})$(with $\tau_0^i:=0$). When particle $(B^{N,i},\eta^{N,i})$ is killed at time τ_k^i, it jumps to the location of particle $(B^{N,j},\eta^{N,j})$, with $j=U_k^i\in\{1,\dots,N\}\setminus\{i\}$ chosen independently and uniformly at random, at which time we set $(B_{\tau_k^i}^{N,i},\eta_{\tau_k^i}^{N,i}):=(B_{\tau_k^i-}^{N,j},\eta_{\tau_k^i-}^{N,j})$. Moreover, we write τ_n for the nth time at which any particle is killed(with $\tau_0:=0$).

We further define as before

$$(5.2)\quad J_t^N := \frac{1}{N}\sup\{n > 0 : \tau_n \le t\}, \qquad m_t^N := \frac{1}{N}\sum_{i=1}^N \delta_{B_t^{N,i}} \quad \text{and} \quad \chi_t^N := \frac{1}{N}\sum_{i=1}^N \delta_{\eta_t^{N,i}}.$$

It is an open problem to establish the well-posedness of this Fleming–Viot particle system without imposing constraints upon the boundary regularity of ∂D [14]. The issue is the possibility of there being infinitely many jumps in finite time. Implicit in the proof of [16], Theorem 1.4, is a proof of the well-posedness of the particle system when the domain satisfies an interior ball condition. Another proof under this condition is due to Löbus [34]. A proof when the domain is Lipschitz with Lipschitz constant less than a given value (dependent upon the dimension and number of particles) is given in [8]. In particular, the assumption that D is bounded and the boundary ∂D is C^∞ certainly suffices to ensure the Fleming–Viot particle system (and, therefore, also the Fleming–Viot multicolor process) is well-posed.

Brownian motion with hard killing, $(B_t)_{0\le t<\tau_\partial}$, defines the C_0-Feller semigroup

$$P_t : C_0(D) \ni f \mapsto \big(x \mapsto P_t f(x) := \mathbb{E}_x\big[f(X_t)\mathbb{1}(\tau_\partial > t)\big]\big) \in C_0(D).$$

We write L for its infinitesimal generator, which is just the half Dirichlet Laplacian. We write $\phi \in C_0(D;\mathbb{R}_{>0}) \cap C^\infty(D)$ for the unique principal right eigenfunction of L of eigenvalue $-\lambda < 0$. In general, quasi-stationary distributions correspond to left eigenmeasures of the infinitesimal generator [38], Proposition 4, which in this case corresponds to the normalised right eigenfunction ϕ. Therefore, the unique QSD of $(X_t)_{0\le t<\tau_\partial}$, denoted as π, is given by

$$\pi(dx) = \frac{\phi(x)\,dx}{\int_D \phi(x')\,dx'}.$$

As in the soft killing case, the rate of the limiting Wright–Fisher process is given by

$$(5.3)\qquad \Theta := \frac{2\lambda\|\phi\|^2_{L^2(\pi)}}{\|\phi\|^2_{L^1(\pi)}}.$$

We again define the tilted empirical measure of the colors by

$$(5.4)\qquad \mathcal{Y}_t^N := \frac{\frac{1}{N}\sum_{i=1}^N \phi(B_t^i)\delta_{\eta_t^i}}{\frac{1}{N}\sum_{i=1}^N \phi(B_t^i)} \in \mathcal{P}(\mathbb{K}).$$

We prove the following analogue of Theorem 1.4.

THEOREM 5.2. *We take some deterministic initial profile $\nu^0 \in \mathcal{P}(\mathbb{K})$ and fix a Wright–Fisher process on $\mathcal{P}(\mathbb{K})$ of rate Θ and initial condition $\nu_0 = \nu^0$, which we denote as $(\nu_t)_{0\le t<\infty}$. We consider a sequence of Fleming–Viot multicolor processes, denoted by $(((\vec{B}_t^N,\vec{\eta}_t^N))_{0\le t<\infty} : 2 \le N < \infty)$, such that*

$$(5.5)\qquad \mathcal{P}(\mathbb{K}) \ni \mathcal{Y}_0^N \to \nu^0 \in \mathcal{P}(\mathbb{K}) \quad \text{in } \mathrm{W}_a \text{ in probability as } N \to \infty.$$

We further require the following condition:

$$(5.6)\qquad \limsup_{N\to\infty} \mathbb{E}\bigg[\frac{1}{N}\#\big\{i \in \{1,\dots,N\} : d\big(B_0^{N,i},\partial D\big) < \delta\big\}\bigg] \to 0 \quad \text{as } \delta \to 0.$$

We now rescale time according to $t \mapsto Nt$. Then $(\chi_{Nt}^N)_{t>0}$ converges to $(\nu_t)_{t>0}$ in finite-dimensional distributions, in the following sense. We fix arbitrary $n < \infty$ and $\vec{t} = (t^1,\dots,t^n) \in [0,\infty)^n$ such that $t^1 \le \dots \le t^n$. We consider arbitrary sequences $(\vec{t}^N)_{2\le N<\infty} := ((t_1^N,\dots,t_n^N))_{2\le N\le\infty}$ such that:

1. $t_1^N \leq \cdots \leq t_n^N$ *for all* $2 \leq N < \infty$;
2. $t_i^N \to t_i$ *as* $N \to \infty$ *for all* $1 \leq i \leq n$;
3. $Nt_n^N \geq \cdots \geq Nt_1^N \to \infty$ *as* $N \to \infty$.

We then have that

$$(\chi^N_{Nt_1^N}, \dots, \chi^N_{Nt_n^N}) \to (\nu_{t_1}, \dots, \nu_{t_n}) \quad \text{in } (\mathcal{P}_{\mathrm{W}_a}(\mathbb{K}))^n \text{ in distribution as } N \to \infty. \tag{5.7}$$

We observe that the only difference with Theorem 1.4 is the condition (5.6), which is necessitated by the fact that the domain is no longer compact. Indeed, when we considered reflected diffusions with soft killing, the domain $\bar{D}$ was compact, with the principal eigenfunction ϕ being bounded away from 0. However, in the case of hard killing, the domain D is noncompact, with ϕ vanishing at the boundary. As a consequence of this, we must establish controls on the mass near the boundary. In order to obtain a hydrodynamic limit theorem over a fixed time horizon for the Fleming–Viot particle system with hard killing, it is important to obtain such controls over a fixed time horizon, as have been established in [16, 46, 47]. Since Theorem 1.4 is a statement about the Fleming–Viot particle system over an $\mathcal{O}(N)$ time horizon, however, we require controls on the mass near the boundary over an $\mathcal{O}(N)$ time horizon. Such controls have not previously been established and represent the principle obstacle to extending Theorem 1.4 to include hard killing. We will obtain such controls in Section 5.2 in the case of Brownian dynamics with hard killing, allowing us to prove Theorem 5.2.

The rest of this Section is devoted to the proof of Theorem 5.2. We will outline in Section 5.1 the notation we will use for the proof of Theorem 5.2 and, in particular, where it differs from the notation outlined in Section 2 for the soft killing case. We will then obtain controls on the mass near the boundary ∂D in Section 5.2. The rest of the proof follows the same outline as the proof of Theorem 1.4. In Section 5.3 we will then perform calculations for $\mathcal{Y}_t^N$, which are analogous to those of Section 3. Finally, we conclude the proof of Theorem 5.2 in Section 5.4, analogously to Section 4. We collect the proofs of technical lemmas needed for the proof of Theorem 5.2 in Appendix E. We will not repeat calculations, which are identical to those found in the proof of Theorem 1.4, pointing out only where the proof differs.

5.1. *Notation for the proof of Theorem* 5.2. Recalling that L is the (half) Dirichlet Laplacian (there is no analogue of L_0 here), we define on the domain $\mathcal{D}(\Gamma) := \{f : f, f^2 \in \mathcal{D}(L)\} \subseteq C_0(D)$ the Carre du champs operator

$$\Gamma(f) := L(f^2) - 2fL(f) = |\nabla f|^2. \tag{5.8}$$

We note, in particular, that $\nabla\phi$ is globally bounded by [48], Theorem 1.1, so that $\phi \in \mathcal{D}(\Gamma)$.

In the soft killing case, the definitions of $P_t^{N,\mathcal{E}}$, Q_t^N, $Y_t^{N,\mathcal{E}}$ and $m_t^{N,\mathcal{E}}$ for $\mathcal{E} \in \mathcal{B}(\mathbb{K})$ are given in Section 2. We adopt the same definitions here, except that ϕ is now the principal eigenfunction of the half Dirichlet Laplacian. Moreover, we define F, H and $\vec{R}_t^N$, as in (3.4), (3.5) and (3.7), respectively, so that

$$F(\vec{r}) = \frac{p}{q}, \qquad H(\vec{r}) = \begin{pmatrix} 0 & -\frac{1}{q^2} \\ -\frac{1}{q^2} & 2\frac{p}{q^3} \end{pmatrix} \quad \text{for } \vec{r} = \begin{pmatrix} p \\ q \end{pmatrix} \in \mathbb{R}^2_{>0}, \qquad \vec{R}_t^{N,\mathcal{E}} := \begin{pmatrix} P_t^{N,\mathcal{E}} \\ Q_t^N \end{pmatrix}.$$

We adopt the notation described in Section 2.1 so that V_t, $\mathcal{O}_t^{\mathrm{MG}}$, $\mathcal{O}_t^{\mathrm{FV}}$, $\mathcal{O}_t^{\Delta}$, $\mathcal{O}_t^{\mathrm{Lip}}$ and $\mathcal{O}_t^{\mathrm{Cts}}$ are defined as in Section 2.1 in particular. We further define the following:

1. $X_t^N = \mathcal{O}_t^{\mathrm{MG}}(Y^N, J^N)$ if for all $N \geq N_0$ (for some $N_0 < \infty$) and for some $C < \infty$, X_t^N is a martingale whose quadratic variation is such that

$$[X^N]_t - \int_0^t CY_s^N \, dJ_s^N \quad \text{is a supermartingale.} \tag{5.9}$$

2. $X_t^N = \mathcal{O}_t^{\mathrm{FV}}(Y^N, J^N)$ if for all $N \geq N_0$ (for some $N_0 < \infty$) and for some $C < \infty$, X_t^N is a finite variation process whose total variation, $V_t(X^N)$, is such that

$$V_t(X^N) - \int_0^t CY_s^N \, dJ_s^N \quad \text{is a supermartingale.} \tag{5.10}$$

5.2. *Control on the mass near the boundary.*

LEMMA 5.3. *We fix $T < \infty$. We consider a sequence of Fleming–Viot particle systems $(\vec{X}_t^N)_{t \geq 0}$ satisfying* (5.6). *Then for all $\epsilon > 0$, there exists $\delta = \delta(\epsilon) > 0$ such that*

$$\liminf_{N \to \infty} \mathbb{P}_{\vec{X}_0^N}\big(m_t^N\big(B(\partial D, \delta)\big) \leq \epsilon \quad \textit{for all } 0 \leq t \leq NT\big) > 1 - \epsilon. \tag{5.11}$$

We define for all $\epsilon > 0$ the stopping time

$$\tau_\epsilon^N := \inf\{t > 0 : m_t^N\big(B\big(\partial D, \delta(\epsilon)\big)\big) > 2\epsilon\}. \tag{5.12}$$

We observe, in particular, that Q_t^N is uniformly bounded from below by a strictly positive constant dependent only upon ϵ, for all $0 \leq t \leq \tau_\epsilon^N$. Moreover, it follows from (5.11) that

$$\liminf_{N \to \infty} \mathbb{P}\big(\tau_\epsilon^N > NT\big) \geq 1 - \epsilon. \tag{5.13}$$

PROOF OF LEMMA 5.3. Controls on the mass of particles near the boundary over a fixed time horizon were established by Burdzy, Holyst and March in [16]. These involve a coupling between the particles and a family of Bessel processes. Similar couplings were established by Villemonais in [47], Section 3, and by Nolen and the present author in [46], Section 4. We utilise the coupling obtained in [46], Section 4, since this coupled family of processes is jointly independent, allowing us to apply Cramér's theorem.

Since ∂D is C^∞ and D is bounded, D satisfies an interior ball condition for some radius $r > 0$. For this $r > 0$, [46], Proposition 4.2, provides a family of independent $[0, r]$-valued continuous processes $(\eta_t^{N,1})_{t \geq 0}, \dots, (\eta_t^{N,N})_{t \geq 0}$, with $(\eta_t^{N,i})_{t \geq 0}$ having the same distribution for all i and all N, such that

$$d\big(B_t^{N,i}, \partial D\big) \geq r - \eta_t^{N,i} \quad \text{for } 0 \leq t < \infty. \tag{5.14}$$

We define $M_N := \lfloor \frac{N}{k} \rfloor$. We then define for $1 \leq m \leq M_N$, $1 \leq k \leq N < \infty$, $0 < T_0 < T_1 < \infty$ and $\delta_0 > 0$ the events

$$A_N^{k,T_0,T_1,\delta_0}(m) := \Big\{ \sup_{T_0 \leq t \leq T_1} \big|\{(m-1)k + 1 \leq i \leq mk : \eta_t^{N,i} \geq r - \delta\}\big| \geq 2 \Big\}. \tag{5.15}$$

We observe that for fixed k, T_0, T_1, δ, the events $A_N^{k,T_0,T_1,\delta}(m)$ for $m \leq M_N$ and $N \geq k$ have the same probability. Moreover, it follows from [46], Lemma 4.4, that, for fixed k, T_0, T_1,

$$\mathbb{P}\big(A_N^{k,T_0,T_1,\delta}(1)\big) \to 0 \quad \text{as } \delta \to 0. \tag{5.16}$$

Moreover, it follows from (5.14) that

$$\sup_{T_0 \leq t \leq T_1} m_t^N\big(B(\partial D, \delta)\big) \leq \frac{k}{N} \sum_{1 \leq m \leq M_N} \mathbb{1}\big(A_N^{k,T_0,T_1,\delta}(m)\big) + \frac{1}{k} + \frac{k}{N}. \tag{5.17}$$

We may, therefore, apply Cramér's theorem to conclude that, for all $\epsilon > 0$ and $0 < T_0 \leq T_1 < \infty$, there exists $\delta_0 > 0$ and $c_0 > 0$ such that

$$\mathbb{P}_{\vec{X}_0^N}\Big(\sup_{T_0 \leq t \leq T_1} m_t^N\big(B(\partial D, \delta)\big) \geq \epsilon\Big) \leq e^{-c_0 N} \quad \text{for all } N \text{ large enough.} \tag{5.18}$$

It follows from a union bound that, for all $\epsilon, T_0 > 0$, there exists $\delta_0 > 0$ such that

$$\liminf_{N\to\infty} \mathbb{P}_{\vec{X}_0^N}\big[m_t^N\big(B(\partial D, \delta_0)\big] \leq \epsilon \quad \text{for all } t_0 \leq t \leq NT\big) = 1. \tag{5.19}$$

All that remains is to deal with the initial time $[0, T_0]$, which can be addressed with a crude bound. We observe that, for any $T_0, \delta_1 > 0$, in order for a given particle to enter $B(\partial D, \delta_1)$, it either has to start within $B(\partial D, 2\delta_1)$ or else travel at least a distance δ_1 in time T_0 (note that killing only occurs at the boundary). The former possibility can be controlled by (5.6), the latter by controlling the distance travelled by Brownian motion in time T_0. We obtain that, for all $\epsilon > 0$, there exists $T_0 > 0$ and $\delta_1 > 0$ such that

$$\liminf_{N\to\infty} \mathbb{P}\big(m_t^N\big(B(\partial D, 2\delta_1)\big) \leq \epsilon \quad \text{for all } 0 \leq t \leq T_0\big) > 1 - \epsilon. \tag{5.20}$$

Therefore, for given $\epsilon > 0$, we choose $\delta_1, T_0 > 0$ for which we have (5.20). For this same $T_0, \epsilon > 0$, we then obtain $\delta_0 > 0$ such that we have (5.19). Taking $\delta := \delta_0 \wedge \delta_1$, we obtain (5.11). □

5.3. *Analogue of the calculations of Section 3.* We obtain the following analogue of Theorem 3.2.

THEOREM 5.4. *We fix arbitrary $\epsilon > 0$ and localise up to the stopping time τ_ϵ^N defined in (5.12). None of the following statements should be understood to be uniform in ϵ but rather should be understood as statements for arbitrary fixed $\epsilon > 0$. We have the following, uniformly over all choices of $\mathcal{E}, \mathcal{F} \in \mathcal{B}(\mathbb{K})$:*

1. *The covariation $[Y^{N,\mathcal{E}}, Y^{N,\mathcal{F}}]_{t\wedge\tau_\epsilon^N}$ is such that for disjoint $\mathcal{E}, \mathcal{F} \in \mathcal{B}(\mathbb{K})$ we have*

$$[Y^{N,\mathcal{E}}, Y^{N,\mathcal{F}}]_{t\wedge\tau_\epsilon^N} = \mathcal{O}_t^{\mathrm{FV}}\Big(\frac{Y^{N,\mathcal{E}} Y^{N,\mathcal{F}}}{N}, J^N\Big).$$

2. *There exists martingales $\mathcal{K}_t^{N,\mathcal{E}}$ for $\mathcal{E} \in \mathcal{B}(\mathbb{K})$ such that $Y_t^{N,\mathcal{E}}$ satisfies*

$$\begin{aligned} Y_{t\wedge\tau_\epsilon^N}^{N,\mathcal{E}} &= Y_0^{N,\mathcal{E}} + \mathcal{K}_{t\wedge\tau_\epsilon^N}^{N,\mathcal{E}} + \mathcal{O}_t^{\mathrm{MG}}\Big(\frac{Y^{N,\mathcal{E}}}{N^3}, J^N\Big) \\ &\quad + \frac{1}{N}\int_0^{t\wedge\tau_\epsilon^N} \frac{1}{(Q_{s-}^N)^2}\langle Y_{s-}^{N,\mathcal{E}} m_{s-}^N - m_{s-}^{N,\mathcal{E}}, \Gamma(\phi)\rangle\, ds \\ &\quad + \mathcal{O}_t^{\mathrm{FV}}\Big(\frac{Y^{N,\mathcal{E}}}{N^2}, J^N\Big) \cap \mathcal{O}_t^{\Delta}\Big(\frac{1}{N^3}\Big) \\ &\quad + \frac{1}{N-1}\int_0^{t\wedge\tau_\epsilon^N} \frac{1}{(Q_{s-}^N)^2}\langle Y_{s-}^{N,\mathcal{E}} m_{s-}^N - m_{s-}^{N,\mathcal{E}}, \phi^2\rangle\, dJ_s^N \end{aligned} \tag{5.21}$$

for $\mathcal{E} \in \mathcal{B}(\mathbb{K})$, *and such that*

$$
\begin{aligned}
[\mathcal{K}^{N,\mathcal{E}}, \mathcal{K}^{N,\mathcal{F}}]_{t\wedge\tau_\epsilon^N} &= \frac{1}{N-1}\int_0^{t\wedge\tau_\epsilon^N} \frac{1}{(Q_{s-}^N)^2}\langle m_{s-}^{N,\mathcal{E}\cap F} - Y_{s-}^{N,\mathcal{E}} m_{s-}^{N,\mathcal{F}} \\
&\quad - Y_{s-}^{N,\mathcal{E}} m_{s-}^{N,\mathcal{F}} + Y_{s-}^{N,\mathcal{E}} Y_{s-}^{N,\mathcal{F}} m_s^N, \phi^2\rangle \, dJ_s^N \\
&\quad + \frac{1}{N}\int_0^{t\wedge\tau_\epsilon^N} \frac{1}{(Q_{s-}^N)^2}\langle m_{s-}^{N,\mathcal{E}\cap F} - Y_{s-}^{N,\mathcal{E}} m_{s-}^{N,\mathcal{F}} - Y_{s-}^{N,\mathcal{E}} m_{s-}^{N,\mathcal{F}} \\
&\quad + Y_{s-}^{N,\mathcal{E}} Y_{s-}^{N,\mathcal{F}} m_s^N, \Gamma(\phi)\rangle \, ds.
\end{aligned} \tag{5.22}
$$

3. *Furthermore*, $Y_t^{N,\mathcal{E}}$ *satisfies*

$$
\begin{aligned}
Y_{t\wedge\tau_\epsilon^N}^{N,\mathcal{E}} &= \Big[\mathcal{O}_t^{\mathrm{FV}}\Big(\frac{Y^{N,\mathcal{E}}}{N}\Big) + \mathcal{O}_t^{\mathrm{FV}}\Big(\frac{Y^{N,\mathcal{E}}}{N}, J^N\Big) \\
&\quad + \mathcal{O}_t^{\mathrm{MG}}\Big(\frac{Y^{N,\mathcal{E}}}{N}\Big) + \mathcal{O}_t^{\mathrm{MG}}\Big(\frac{Y^{N,\mathcal{E}}}{N}, J^N\Big)\Big] \cap \mathcal{O}_t^{\Delta}\Big(\frac{1}{N}\Big).
\end{aligned} \tag{5.23}
$$

4. *Parts* 1–3 *remain true if* $\mathcal{E}$ *and* $\mathcal{F}$ *are replaced with a sequence of* σ_0*-measurable random sets* $\mathcal{E}^N$ *and* $\mathcal{F}^N$.

PROOF OF THEOREM 5.4. It is useful (and simplifies our calculations) to note that since ϕ vanishes on the boundary and killing only occurs on the boundary, we necessarily have that $\phi(B_{t-}^i) = 0$ if B^i is killed at time t.

It is straightforward to obtain the following analogue of Proposition 3.3 by examining the martingale

$$
P_t^{N,\mathcal{E}} - P_0^{N,\mathcal{E}} - \int_0^t P_{s-}^{N,\mathcal{E}}\Big(-\lambda\, ds + \frac{N}{N-1}\, dJ_s^N\Big).
$$

PROPOSITION 5.5. *We have for all* $\mathcal{E} \in \mathcal{B}(\mathbb{K})$ *that*

$$
dP_t^{N,\mathcal{E}} = P_{t-}^{N,\mathcal{E}}\Big(-\lambda\, dt + \frac{N}{N-1}\, dJ_t^N\Big) + dM_t^{N,\mathcal{E}}, \tag{5.24}
$$

whereby $M^{N,\mathcal{E}}$ *are martingales which satisfy for all* $\mathcal{E}, \mathcal{F} \in \mathcal{B}(\mathbb{K})$

$$
\begin{aligned}
&[M^{N,\mathcal{E}}, M^{N,\mathcal{F}}]_t \\
&\quad = \frac{1}{N}\int_0^t \langle m_s^{N,\mathcal{E}\cap\mathcal{F}}, \Gamma_0(\phi)\rangle\, ds \\
&\qquad + \frac{1}{N}\int_0^t \frac{N}{N-1}\langle m_{s-}^{N,\mathcal{E}\cap\mathcal{F}}, \phi^2\rangle - \Big(\frac{N}{N-1}\langle m_{s-}^{N,\mathcal{E}}, \phi\rangle\Big)\Big(\frac{N}{N-1}\langle m_{s-}^{N,\mathcal{F}}, \phi\rangle\Big)\, dJ_s^N.
\end{aligned} \tag{5.25}
$$

Moreover, it is apparent that $\sum_{s\le t} \Delta M_t^{N,\mathcal{E}} = \mathcal{O}_t^{\mathrm{MG}}(\frac{P^{N,\mathcal{E}}}{N}, J_t^N)$ *for* $\mathcal{E} \in \mathcal{B}(\mathbb{K})$. *We write* M_t^N *for* $M_t^{N,\mathbb{K}}$.

Using that Q_t^N is bounded from below away from 0 for $t \le \tau_\epsilon^N$, uniformly in N, we obtain Part 1 of Theorem 5.4 in precisely the same manner that we obtained Part 1 of Theorem 3.2 in Section 3.1.2.

Replacing the boundedness of the jump rate with a bound in terms of the number of jumps, we obtain the following analogue of Lemma 3.4.

LEMMA 5.6 (Itô's lemma). *We have*

$$Y_t^{N,\mathcal{E}} = Y_0^{N,\mathcal{E}} + \int_0^t \nabla F(R_{s-}^{N,\mathcal{E}}) \cdot dR_s^{N,\mathcal{E}} + \frac{1}{2} dR_s^{N,\mathcal{E}} \cdot H(R_{s-}^{N,\mathcal{E}})\, dR_s^{N,\mathcal{E}} + \mathcal{O}_t^{\mathrm{FV}}\Big(\frac{Y^{N,\mathcal{E}}}{N^2}, J^N\Big) \cap \mathcal{O}_t^{\Delta}\Big(\frac{1}{N^3}\Big). \tag{5.26}$$

We then obtain parts 2–4 of Theorem 5.4, as in the proof of Theorem 3.2. □

5.4. *Proof of Theorem* 5.2. We first prove the following analogue of proposition 4.1.

PROPOSITION 5.7. *For all $\mathcal{E} \in \mathcal{B}(\mathbb{K})$ and $f \in C_b(\bar{D})$, we have that*

$$\mathbb{1}(\tau_\epsilon^N > t) \sup_{t \le s_1, s_2 \le t+1} \big|(m_{s_1}^{N,\mathcal{E}} - Y_{s_2}^{N,\mathcal{E}}\pi)(f)\big| \to 0 \quad \textit{in probability as } t \wedge N \to \infty. \tag{5.27}$$

In particular, taking $f = 1$, we have for any $\mathcal{E} \in \mathcal{B}(\mathbb{K})$ that

$$\mathbb{1}(\tau_\epsilon^N > t) \sup_{t \le s_1, s_2 \le t+1} \big|\chi_{s_1}^N(\mathcal{E}) - \mathcal{Y}_{s_2}^N(\mathcal{E}))\big| \to 0 \quad \textit{in probability as } t \wedge N \to \infty. \tag{5.28}$$

We note that $\tau_\epsilon^N > 0$ implies a uniform positive lower bound on $\mathbb{P}_{m_0^N}(\tau_\partial > T)$, where we think of the empirical measure m_0^N as the initial condition of a single killed Brownian motion, killed at time τ_∂. Using this fact and using Theorem 5.4, Theorem E.3 and Proposition E.4 in place of Theorem 5.4, Theorem 1.8 and Proposition 1.9, respectively, the proof of Proposition 5.7 is identical to the proof of Proposition 4.1 found in Section 4.1.

The characterisation of $\mathcal{Y}_t^N$ in the setting of soft killing, given in Section 3, does not involve dJ_t^N terms. This is a consequence of the fact that the jumps occur at a (position dependent) Poisson rate. On the other hand and as a consequence of the hard catalyst killing, the characterisation of $\mathcal{Y}_t^N$, given in Section 5.3, does involve such terms. Consequentially, we shall require the following lemma.

LEMMA 5.8. *We consider a sequence of uniformly bounded processes $(Z_t^N)_{t\ge 0}$ such that $\sup_{t\le s\le t+1}|Z_s^N - Z_t^N| \to 0$ in probability as $t \wedge N \to \infty$. Then after rescaling time by $t \mapsto Nt$, we have that*

$$\int_0^T Z_{Ns}^N \Big(\frac{1}{N} dJ_{Ns}^N - \lambda\, ds\Big) \to 0 \quad \textit{in probability as } N \to \infty. \tag{5.29}$$

In particular, we have the convergence

$$\Big(\frac{1}{N} J_{Nt}^N\Big)_{0\le t\le T} \to (\lambda t)_{0\le t\le T} \quad \textit{uniformly in probability as } N \to \infty. \tag{5.30}$$

PROOF OF LEMMA 5.8. We fix $\epsilon > 0$. We first take $\mathcal{E} = \mathbb{K}$ in (5.27) and apply Proposition C.1 to see that $m_t^N \mathbb{1}(\tau_\epsilon^N > t) + \pi \mathbb{1}(\tau_\epsilon^N \le t)$ converges to π in W in probability. We then obtain from Theorem E.3 that

$$\mathbb{1}(\tau_\epsilon^N > t)(J_{t+1}^N - J_t^N - \lambda) \to 0 \quad \text{in probability as } N \wedge t \to \infty. \tag{5.31}$$

It follows from the proof of [46], Proposition 4.10, that there exists $M < \infty$ and $p \in (0, 1)$, dependent upon neither N nor t, such that the number of jumps of any particle between the times $t \wedge \tau_\epsilon^N$ and $(t+1) \wedge \tau_\epsilon^N$ is stochastically dominated by the sum of M independent geometric random variables. It follows that $\{J_{(t+1)\wedge\tau_\epsilon^N}^N - J_{t\wedge\tau_\epsilon^N}^N : t \ge 0, N \ge 2\}$ is uniformly bounded in $L^2(\mathbb{P})$, hence uniformly integrable.

We now calculate

$$\begin{aligned}\int_0^T \int_{Nt}^{Nt+1} Z_s^N \, dJ_s^N \, dt &= \int_0^T \int_0^{NT+1} \mathbb{1}(Nt \le s \le Nt+1) Z_s^N \, dJ_s^N \, dt \\ &= \int_0^{NT+1} Z_s^N \int_{\frac{s-1}{N} \vee 0}^{\frac{s}{N} \wedge T} dt \, dJ_s^N \\ &= \int_0^{NT+1} \min\left(\frac{1}{N}, \frac{s}{N}, \frac{T-s-1}{N}\right) Z_s^N \, dJ_s^N.\end{aligned}$$

We see that

$$\begin{aligned}&\mathbb{1}(\tau_\epsilon^N > NT) \left| \frac{1}{N} \int_0^T Z_{Ns}^N \, dJ_{Ns}^N - \int_0^T \int_{Nt}^{Nt+1} Z_s^N \, dJ_s^N \, dt \right| \\ &\quad \le \frac{C}{N} \mathbb{1}(\tau_\epsilon^N > NT)(J_1 - J_0 + J_{NT+1} - J_{NT}) \xrightarrow{p} 0 \quad \text{as } N \to \infty.\end{aligned}$$

It follows from the above that

$$\begin{aligned}&\mathbb{E}\left[\left|\int_0^T \mathbb{1}(\tau_\epsilon^N > NT) \left(\int_{Nt}^{Nt+1} Z_s^N (dJ_s^N - \lambda \, ds)\right) dt\right|\right] \\ &\quad \le \int_0^T \mathbb{E}\left[\left|\mathbb{1}(\tau_\epsilon^N > NT) \int_{Nt}^{Nt+1} Z_s^N (dJ_s^N - \lambda \, ds)\right|\right] dt \to 0 \quad \text{as } N \to \infty.\end{aligned}$$

Since $\epsilon > 0$ is arbitrary, Lemma 5.8 follows from (5.13). □

We then obtain the following analogue of Proposition 4.2.

PROPOSITION 5.9. *We take some deterministic initial profile $\nu^0 \in \mathcal{P}(\mathbb{K})$ and define $(\nu_t)_{0 \le t < \infty}$ to be a Wright–Fisher process of rate Θ and initial condition $\nu_0 := \nu^0$. We then consider a sequence of Fleming–Viot multicolor processes $(\vec{B}_t^N, \vec{\eta}_t^N)_{0 \le t < \infty}$. We assume that $\mathcal{Y}_0^N \to \nu^0$ in W_a in probability.*

We fix $T < \infty$ and rescale time by $t \mapsto Nt$. We then have the convergence

$$(\mathcal{Y}_{Nt}^N)_{0 \le t \le T} \to (\nu_t)_{0 \le t \le T} \quad \text{in } D([0, T]; \mathcal{P}_{\mathrm{W}}(\mathbb{K})) \text{ in distribution as } N \to \infty. \tag{5.32}$$

PROOF OF PROPOSITION 5.9. The proof of Proposition 5.9 follows in the same two steps as that of the proof of Proposition 4.2. In both steps we fix $\epsilon > 0$ and localise up to time τ_ϵ^N. We then repeat the proof found in Section 4.2, with Theorem 3.2 and Proposition 4.1 replaced by Theorem 5.4 and Proposition 5.7, respectively, and using Lemma 5.8 in the obvious manner. We then conclude each step by observing that $\epsilon > 0$ is arbitrary and applying (5.13). □

We continue as in the proof of Theorem 1.4. We recall from Theorem D.2 that $(\nu_t)_{0 \le t \le T} \in C([0, T]; \mathcal{P}_{\mathrm{W}}(\mathbb{K}))$ almost surely. We take a sequence $(\vec{t}^N)_{2 \le N < \infty} = ((t_1^N, \dots, t_n^N))_{2 \le t \le N}$ converging to $\vec{t} = (t^1, \dots, t^n)$, as in the statement of Theorem 5.2. It then follows that

$$(\mathcal{Y}_{Nt_1^N}^N, \dots, \mathcal{Y}_{Nt_n^N}^N) \to (\nu_{t_1}, \dots, \nu_{t_n}) \quad \text{in } (\mathcal{P}_{\mathrm{W}}(\mathbb{K}))^n \text{ in distribution as } N \to \infty.$$

Recalling the positivity and boundedness of ϕ from Theorem E.1 and the definition (5.12) of τ_ϵ^N, we observe that, for every $\epsilon > 0$, there exists a uniform constant $C_\epsilon < \infty$ and random measures $\delta_t^{N,\epsilon}$ for $0 \le t < \tau_\epsilon^N$ and $N \in \mathbb{N}$ such that

$$\chi_t^N \le C_\epsilon \mathcal{Y}_t^N + \delta_t^{N,\epsilon} \quad \text{and} \quad \delta_t^{N,\epsilon}(\mathbb{K}) \le \epsilon \quad \text{for all } 0 \le t < \tau_\epsilon^N, N \ge 2. \tag{5.33}$$

Note that the term $\delta_t^{N,\epsilon}$ does not appear in the soft killing case, (4.4). It arises here from the fact that ϕ is no longer bounded from below but instead vanishes at the boundary.

We now fix $1 \leq k \leq n$. Since $(\mathcal{Y}^N_{Nt_k^N})_{N\geq 1}$ is a tight sequence of random measures, it follows from (5.33) that $(\chi^N_{Nt_k^N})_{N\geq 1}$ must also be a tight sequence of random measures. It, therefore, follows from (5.28) and Lemma C.2 that $\mathrm{W}(\mathcal{Y}^N_{Nt_k^N}, \chi^N_{Nt_k^N}) \to 0$ as $N \to \infty$. Thus, we have established that

$$(\chi^N_{Nt_1^N}, \dots, \chi^N_{Nt_n^N}) \to (\nu_{t_1}, \dots, \nu_{t_n}) \quad \text{in } (\mathcal{P}_{\mathrm{W}}(\mathbb{K}))^n \text{ in distribution as } N \to \infty.$$

We have left only to strengthen the notion of convergence to convergence in the weak atomic metric. This is accomplished with the following analogue of Proposition 4.3.

PROPOSITION 5.10. *We recall that* $\Psi(u) := (1-u) \vee 0$ *is the function used to define the* W_a *metric in Appendix* C.2. *We fix* $\epsilon > 0$, *thereby defining the stopping defined* τ_ϵ^N *by* (5.12). *For all* $\delta > 0$, *there exists* $\epsilon' > 0$ *such that*

$$\liminf_{N\to\infty} \mathbb{P}\left(\tau_\epsilon^N > NT, \sup_{0\leq t\leq T} \sum_{\substack{k,\ell\in\mathbb{K}\\ k\neq\ell}} \chi^N_{Nt}(\{k\})\chi^N_{Nt}(\{\ell\})\Psi\left(\frac{d(k,\ell)}{\epsilon'}\right) \leq \delta + \epsilon\right) \geq 1-\delta.$$

Note that the above sum is well defined, as the terms are nonzero only for $k, \ell \in \operatorname{supp}(\chi_0^N)$.

PROOF OF PROPOSITION 5.10. We follow the same proof strategy as the proof of Proposition 4.3, replacing Theorem 3.2 with Theorem 5.4, applying Lemma 5.8 in the obvious manner and replacing the supermartingale (4.13) with

$$e^{-C(t+\frac{J^N_{Nt\wedge\tau_\epsilon^N}}{N})} \sum_{\substack{k,\ell\in\mathbb{K}\\ k\neq\ell}} Y^{N,\{k\}}_{Nt\wedge\tau_\epsilon^N} Y^{N,\{\ell\}}_{Nt\wedge\tau_\epsilon^N} \Psi\left(\frac{d(k,\ell)}{\epsilon}\right)$$

for some sufficiently large constant $C < \infty$. □

Having established Proposition 5.10, we may then apply Lemma C.5 along with (5.13) to conclude that $\{\mathcal{L}(\chi^N_{Nt_k^N})\}$ is tight in $\mathcal{P}(\mathcal{P}_{\mathrm{W}_a}(\mathbb{K}))$ for all $1 \leq k \leq n$ so that we have Theorem 1.4.

APPENDIX A: REFLECTED DIFFUSIONS WITH SOFT KILLING

A.1. Definition. We consider a normally reflected diffusion $(X_t^0)_{0\leq t<\infty}$ in the domain $\bar{D}$ corresponding to a solution of the Skorokhod problem. In particular, for any filtered probability space on which is defined the m-dimensional Brownian motion W_t and initial condition $x \in \bar{D}$, there exists by [33], Theorem 3.1, a pathwise unique strong solution of the Skorokhod problem

$$\begin{aligned} &X_t^0 = x + \int_0^t b(X_s^0)\,ds + \int_0^t \sigma(X_s^0)\,dW_s + \int_0^t \vec{n}(X_s^0)\,d\xi_s \in \bar{D}, \\ &\qquad 0 \leq t < \infty, \\ &\int_0^\infty \mathbb{1}_D(X_t^0)\,d\xi_t = 0, \end{aligned} \tag{A.1}$$

where W_s is a Brownian motion and the local time ξ_t is a nondecreasing process with $\xi_0 = 0$.

This corresponds to a solution of the submartingale problem, introduced by Stroock and Varadhan [41], and is a Feller process [41], Theorem 5.8, Remark 2 (and hence strong Markov). It is then straightforward (using a separate probability space on which is defined an exponential random variable) to construct an enlarged filtered probability space on which (X^0, W, ξ) is a solution of the Skorokhod problem and on which there is a stopping time τ_∂ corresponding to the ringing time of a Poisson clock with position dependent rate $\kappa(X_t^0)$ from which is constructed the killed process $(X_t)_{0\le t<\tau_\partial}$. This killed process is a solution to

$$\mathbb{1}(t \geq \tau_\partial) - \int_0^{t\wedge\tau_\partial} \kappa(X_s^0)\,ds \quad \text{is a martingale, } X_t := \begin{cases} X_t^0, & t < \tau_\partial, \\ 0, & t \geq \tau_\partial, \end{cases} \tag{A.2}$$

where W_s is an m-dimensional Brownian motion and the local time ξ_t is a nondecreasing process with $\xi_0 = 0$. Since X_t^0 is Feller, the process X_t is, therefore, also Feller (and hence strong Markov).

We write $L^0/L = L^0 - \kappa$ for the infinitesimal generators of X^0 and X, respectively, having the same domains $\mathcal{D}(L^0) = \mathcal{D}(L)$. We further define the Carre du Champs operator Γ_0 on the algebra $\mathcal{A}$,

$$\begin{aligned} &\Gamma_0(f,g) := L_0(fg) - fL_0(f) - gL_0(f), \qquad \Gamma_0(f) := \Gamma_0(f,f), \\ &f,g \in \mathcal{A} := \{f \in C^2(\bar{D}) : \vec{n}\cdot\nabla f \equiv 0 \text{ on } \partial D\}, \end{aligned} \tag{A.3}$$

so that, for $f \in \mathcal{A}$, we have

$$[f(X^0)]_t = \int_0^t \Gamma_0(f)(X_s^0)\,ds. \tag{A.4}$$

A.2. Convergence to a unique quasi-stationary distribution.

THEOREM A.1. *There exists a unique quasi-stationary distribution* (*QSD*) $\pi \in \mathcal{P}(\bar{D})$ *for* X_t. *Moreover, there exist constants* $C < \infty$ *and* $k > 0$ *such that*

$$\|\mathcal{L}_\mu(X_t|\tau_\partial > t) - \pi\|_{\mathrm{TV}} \leq Ce^{-kt} \quad \text{for all } \mu \in \mathcal{P}(\bar{D}) \text{ and } t \geq 0. \tag{A.5}$$

Furthermore, π *is a left eigenmeasure of* L *with eigenvalue* $-\lambda < 0$,

$$\langle \pi, Lf\rangle = -\lambda\langle \pi, f\rangle, \qquad f \in \mathcal{D}(L), \tag{A.6}$$

and with corresponding positive right eigenfunction $\phi \in \mathcal{A} \cap C^2(\bar{D}; \mathbb{R}_{>0})$. *This right eigenfunction is both the unique nonnegative right eigenfunction and the unique right eigenfunction of eigenvalue* $-\lambda$, *up to rescaling.*

PROOF OF THEOREM A.1. Our strategy is to check [19], Assumption (A), starting with [19], Assumption (A1).

We fix arbitrary $t_0 > 0$. It follows from the boundedness of the killing rate κ and the parabolic Harnack inequality that there exists $c_0 > 0$ and $\nu \in \mathcal{P}(\bar{D})$ such that

$$\mathcal{L}_x(X_{t_0}|\tau_\partial > t_0) \geq c_0\nu \quad \text{for all } x \in \bar{D}.$$

Thus, [19], Assumption (A1), is satisfied. We now turn to checking [19], Assumption (A2).

In [40], page 6, they use the Krein–Rutman theorem to prove that there exists $\phi \in \mathcal{A} \cap C^2(\bar{D}; \mathbb{R}_{>0})$ and $\lambda \in \mathbb{R}$ such that

$$L\phi + \lambda\phi = 0 \quad \text{on } \bar{D}, \qquad \phi > 0 \quad \text{on } \bar{D}, \qquad \nabla\phi\cdot\vec{n} = 0 \quad \text{on } \partial D.$$

We see that $e^{\lambda t}\phi(X_t)\mathbb{1}(t < \tau_\partial)$ is a martingale so that

$$\langle P_t(\mu,\cdot),\phi\rangle = e^{-\lambda t}\langle \mu,\phi\rangle \quad \text{for all } \mu \in \mathcal{P}(\bar{D}). \tag{A.7}$$

Therefore, by (A.7) we have for all $\mu \in \mathcal{P}(\bar{D})$ that

$$\text{(A.8)} \qquad \mathbb{P}_\mu(t < \tau_\partial) \geq \frac{\langle P_t(\mu, \cdot), \phi\rangle}{\sup_{x' \in \bar{D}} \phi(x')} = \frac{e^{-\lambda t}\langle \mu, \phi\rangle}{\sup_{x' \in \bar{D}} \phi(x')} \geq \frac{\inf_{x' \in \bar{D}} \phi(x')}{\sup_{x' \in \bar{D}} \phi(x')} e^{-\lambda t}.$$

Similarly, (A.7) gives that, for all $x \in \bar{D}$, we have

$$\text{(A.9)} \qquad \mathbb{P}_x(t < \tau_\partial) \leq \frac{\langle P_t(x, \cdot), \phi\rangle}{\inf_{x' \in \bar{D}} \phi(x')} \leq \frac{\sup_{x' \in \bar{D}} \phi(x')}{\inf_{x' \in \bar{D}} \phi(x')} e^{-\lambda t}.$$

Therefore, we have

$$\mathbb{P}_x(t < \tau_\partial) \leq \left(\frac{\sup_{x' \in \bar{D}} \phi(x')}{\inf_{x' \in \bar{D}} \phi(x')}\right)^2 \mathbb{P}_\mu(t < \tau_\partial), \quad \text{for all } t \geq 0 \text{ and } x \in \bar{D}.$$

Thus, we have verified [19], Assumption (A), so that [19], Theorem 1.1, implies the existence of a quasi-stationary distribution $\pi \in \mathcal{P}(\bar{D})$ satisfying (A.5), which must be the unique quasi-stationary distribution. Moreover, the uniqueness of ϕ up to renormalisation, both as a nonnegative right eigenfunction and eigenfunction of eigenvalue $-\lambda$, is given by [19], Corollary 2.4. Finally, the QSD π corresponds to the left eigenmeasure of L for some eigenvalue $-\lambda' < 0$ by [38], Proposition 4. This eigenvalue must be equal to $-\lambda$, since $-\lambda'\langle \pi, \phi\rangle = \langle \pi, L\phi\rangle = -\lambda\langle \pi, \phi\rangle$, so that we have (A.6). □

APPENDIX B: PROOF OF PROPOSITION 1.9

We recall that ϕ is normalised so that $\langle \pi, \phi\rangle = 1$. We take $(x^i, \eta^i)_{1 \leq i \leq n} \in (\bar{D} \times \mathbb{K})^n$ and calculate

$$\text{(B.1)} \qquad \begin{aligned} \mathcal{L}_{\frac{1}{n}\sum_{i=1}^n \delta_{(x^i,\eta^i)}}\big((X_s, \eta_s) | \tau_\partial > s\big) &= \frac{\frac{1}{n}\sum_{i=1}^n \mathbb{P}_{x^i}(\tau_\partial > s)\mathcal{L}_{x^i}(X_s | \tau_\partial > s) \otimes \delta_{\eta^i}}{\frac{1}{n}\sum_{i=1}^n \mathbb{P}_{x^i}(\tau_\partial > s)} \\ &= \frac{\frac{1}{n}\sum_{i=1}^n e^{\lambda s}\mathbb{P}_{x^i}(\tau_\partial > s)\mathcal{L}_{x^i}(X_s | \tau_\partial > s) \otimes \delta_{\eta^i}}{\frac{1}{n}\sum_{i=1}^n e^{\lambda s}\mathbb{P}_{x^i}(\tau_\partial > s)}. \end{aligned}$$

It then follows from the triangle inequality that

$$\begin{aligned} &\left\| \frac{1}{n}\sum_{i=1}^n e^{\lambda s}\mathbb{P}_{x^i}(\tau_\partial > s)\mathcal{L}_{x^i}\big(X_s | \tau_\partial > s\big) \otimes \delta_{\eta^i} - \frac{1}{n}\sum_{i=1}^n \phi(x^i)\pi \otimes \delta_{\eta^i} \right\|_{\mathrm{TV}} \\ &\quad \leq \left\| \frac{1}{n}\sum_{i=1}^n \big[e^{\lambda s}\mathbb{P}_{x^i}(\tau_\partial > s) - \phi(x^i)\big]\mathcal{L}_{x^i}\big(X_s | \tau_\partial > s\big) \otimes \delta_{\eta^i} \right\|_{\mathrm{TV}} \\ &\qquad + \left\| \frac{1}{n}\sum_{i=1}^n \phi(x^i)\big[\mathcal{L}_{x^i}\big(X_s | \tau_\partial > s\big) \otimes \delta_{\eta^i} - \pi \otimes \delta_{\eta^i}\big] \right\|_{\mathrm{TV}} \\ &\quad \leq \frac{1}{n}\sum_{i=1}^n \big|e^{\lambda s}\mathbb{P}_{x^i}(\tau_\partial > s) - \phi(x^i)\big| + \frac{1}{n}\sum_{i=1}^n \phi(x^i)\big\|\mathcal{L}_{x^i}\big(X_s | \tau_\partial > s\big) \otimes \delta_{\eta^i} - \pi \otimes \delta_{\eta^i}\big\|_{\mathrm{TV}}. \end{aligned}$$

We can apply [20], Theorem 2.1, by Theorem A.1. It follows from Theorem A.1 and [20], Theorem 2.1, that there exists $\epsilon_t \to 0$ such that, for any $n < \infty$ and $(x_i, \eta_i)_{1 \leq i \leq n} \in (D \times \mathbb{K})^n$, we have that

$$\left\| \frac{1}{n}\sum_{i=1}^n e^{\lambda s}\mathbb{P}_{x^i}(\tau_\partial > s)\mathcal{L}_{x^i}\big(X_s | \tau_\partial > s\big) \otimes \delta_{\eta^i} - \frac{1}{n}\sum_{i=1}^n \phi(x^i)\pi \otimes \delta_{\eta^i} \right\|_{\mathrm{TV}} \leq \epsilon_t \frac{1}{n}\sum_{i=1}^n \phi(x^i).$$

We apply this to both the numerator and denominator of the right-hand side of (B.1) to obtain (1.14).

APPENDIX C: SPACES OF MEASURES

For a given topological space $\mathbf{S}$, we write $\mathcal{B}(\mathbf{S})$ for the Borel σ-algebra on $\mathbf{S}$ and write $\mathcal{P}(\mathbf{S})$ for the space of probability measures on $\mathcal{B}(\mathbf{S})$, equipped with the topology of weak convergence of measures. We write $\mathcal{M}(\mathbf{S})$ for the space of all bounded Borel measurable functions on $\mathbf{S}$.

C.1. The Wasserstein metric. For general separable metric spaces $(\mathbf{S}, d)$, we let W denote the Wasserstein-1 metric on $\mathcal{P}(\mathbf{S})$, generated by the metric $d \wedge 1$, which metrises $\mathcal{P}(\mathbf{S})$ [27], Theorem 6. We write $\mathcal{P}_{\mathrm{W}}(\mathbf{S})$ for the metric space $(\mathcal{P}(\mathbf{S}), \mathrm{W})$. The following, therefore, follows from the Skorokhod representation theorem.

PROPOSITION C.1. *Let $(\mathbf{S}, d)$ be a separable metric space. Let $(\mu_n)_{n\geq 1}$ be a sequence of $\mathcal{P}(\mathbf{S})$-valued random measures, and $\mu \in \mathcal{P}(\mathbf{S})$ a deterministic measure. Then the following are equivalent*:

1. $\mathrm{W}(\mu_n, \mu) \xrightarrow{p} 0$ *as* $n \to \infty$;
2. $\mu_n(f) \xrightarrow{p} \mu(f)$ *as* $n \to \infty$ *for all* $f \in C_b(\mathbf{S})$.

We similarly obtain the following lemma.

LEMMA C.2. *Let $(\mathbf{S}, d)$ be a separable metric space. Let $(\mu_n^{(i)})_{n\geq 1}$ for $i = 1, 2$ be tight sequences of $\mathcal{P}(\mathbf{S})$-valued random measures, with $\mu_n^{(1)}$ and $\mu_n^{(2)}$ defined on the same probability space, for all n. We suppose that*

$$\left|\mu^{(1)}(A) - \mu_n^{(2)}(A)\right| \xrightarrow{p} 0 \quad \text{as } n \to \infty \text{ for all } A \in \mathcal{B}(\mathbf{S}). \tag{C.1}$$

Then $\mathrm{W}(\mu_n^{(1)}, \mu_n^{(2)}) \xrightarrow{p} 0$ *as* $n \to \infty$.

PROOF OF LEMMA C.2. Since $((\mu_n^{(1)}, \mu_n^{(2)}) : 1 \leq n < \infty)$ is a tight sequence of random measures, we may consider arbitrary subequential limits to which we apply Skorokhod's representation theorem. Then on this new probability space and along the subsequential limit, we have $(\mu_{n_k}^{(1)}, \mu_{n_k}^{(2)}) \to (\mu^{(1)}, \mu^{(2)})$ in $\mathcal{P}(\mathbf{S}) \times \mathcal{P}(\mathbf{S})$ as $k \to \infty$. We now use (C.1) to conclude that $\mu^{(1)} = \mu^{(2)}$ almost surely from which we conclude Lemma C.2. $\square$

C.2. The weak atomic metric. Convergence in our scaling limit is given in terms of the weak atomic metric, introduced by Ethier and Kurtz in [24]. We shall define the weak atomic metric on the color space $(\mathbb{K}, d)$ (which we recall is assumed to be a complete, separable metric space). We write $\mathcal{P}_{\mathrm{W}_a}(\mathbb{K})$ for $\mathcal{P}(\mathbb{K})$ equipped with the metric W_a.

In [24] Ethier and Kurtz defined the weak atomic metric on the space of all finite, positive, Borel measures, whereas we restrict our attention to probability measures on $\mathbb{K}$. We fix $\Psi(u) = (1-u) \vee 0$ and define the weak atomic metric to be

$$\begin{aligned} \mathrm{W}_a(\mu, \nu) := \mathrm{W}(\mu, \nu) + \sup_{0<\epsilon\leq 1} \bigg| &\int_{\mathbb{K}} \int_{\mathbb{K}} \Psi\left(\frac{d(x,y)}{\epsilon}\right) \mu(dx)\mu(dy) \\ &- \int_{\mathbb{K}} \int_{\mathbb{K}} \Psi\left(\frac{d(x,y)}{\epsilon}\right) \nu(dx)\nu(dy) \bigg|. \end{aligned} \tag{C.2}$$

In [24] they used the Lévy–Prokhorov metric, instead of the W-metric, and let Ψ be an arbitrary continuous, nondecreasing function such that $\Psi(0) = 1$ and $\Psi(1) = 0$. We make the above choices for simplicity (note that W is equivalent to the Lévy–Prokhorov metric [27], Theorem 2). Convergence in the weak atomic metric is equivalent to weak convergence of measures and convergence of the location and sizes of the atoms.

LEMMA C.3 (Lemma 2.5, [24]). *Consider a sequence of probability measures* $(\mu_n)_{n=1}^{\infty}$ *and a probability measure* μ, *all in* $\mathcal{P}(\mathbb{K})$. *The following are equivalent*:

1. $\mathrm{W}_a(\mu_n, \mu) \to 0$ *as* $n \to \infty$.
2. *We have both of the following*:
   (a) $\mathrm{W}(\mu_n, \mu) \to 0$ *as* $n \to \infty$.
   (b) *There exists an ordering of the atoms* $\{\alpha_i \delta_{x_i}\}$ *of* μ *and the atoms* $\{\alpha_i^n \delta_{x_i^n}\}$ *of* μ_n *so that* $\alpha_1 \geq \alpha_2 \geq \dots$ *and* $\lim_{n\to\infty}(\alpha_i^n, x_i^n) = (\alpha_i, x_i)$ *for all* i.

REMARK C.4. Note that (2a) is equivalent to $\mu_n \to \mu$ weakly by Proposition C.1.

Thus, measures are close in the weak atomic metric if and only if they are both close in the Wasserstein-1 metric W and have similar atoms. For instance, $\frac{1}{2}\mathrm{Leb}_{[0,1]} + \frac{1}{2}\delta_{\frac{1}{2}}$ is close in the weak atomic metric to $(\frac{1}{2} - \epsilon)\mathrm{Leb}_{[0,1]} + (\frac{1}{2} + \epsilon)\delta_{\frac{1}{2}+\epsilon}$ (for small $\epsilon > 0$) but not to $\frac{1}{2}\mathrm{Leb}_{[0,1]} + (\frac{1}{4}\delta_{\frac{1}{2}-\epsilon} + \frac{1}{4}\delta_{\frac{1}{2}+\epsilon})$ nor to $\frac{1}{3}\mathrm{Leb}_{[0,1]} + \frac{2}{3}\delta_{\frac{1}{2}}$.

We note by [24], page 5, that $\mathcal{B}(\mathcal{P}(\mathbb{K})) = \mathcal{B}(\mathcal{P}_{\mathrm{W}_a}(\mathbb{K}))$ so that probability measures in $\mathcal{P}(\mathcal{P}_{\mathrm{W}_a}(\mathbb{K}))$ are probability measures in $\mathcal{P}(\mathcal{P}(\mathbb{K}))$ and vice-versa. It will be useful to be able to characterise tightness in both $\mathcal{P}(\mathcal{P}_{\mathrm{W}_a}(\mathbb{K}))$ and $\mathcal{P}(D([0,T]; \mathcal{P}_{\mathrm{W}_a}(\mathbb{K})))$.

Ethier and Kurtz established in [24], Lemma 2.9, the following tightness criterion.

LEMMA C.5 (Lemma 2.9, [24]). *Consider a sequence of measures* $(\mu_n)_{n=1}^{\infty}$ *in* $\mathcal{P}(\mathcal{P}(\mathbb{K}))$. *Then the following are equivalent*:

1. $(\mu_n)_{n=1}^{\infty}$ *is tight in* $\mathcal{P}(\mathcal{P}_{\mathrm{W}_a}(\mathbb{K}))$.
2. $(\mu_n)_{n=1}^{\infty}$ *is tight in* $\mathcal{P}(\mathcal{P}_{\mathrm{W}}(\mathbb{K}))$, *and we also have*

$$\sup_n \mathbb{E}\left[\int_{\mathbb{K}}\int_{\mathbb{K}} \Psi\left(\frac{d(x,y)}{\epsilon}\right)\mathbb{1}(x \neq y)\mu_n(dx)\mu_n(dy)\right] \to 0 \quad \text{as } \epsilon \to 0. \tag{C.3}$$

Note that the above statement is slightly different from the statement of [24], Lemma 2.9. It is straightforward to see that the two statements are equivalent for families of probability measures; for our purposes this lemma statement will be easier to use.

APPENDIX D: THE WRIGHT–FISHER PROCESS

The Wright–Fisher process is defined as a solution of a martingale problem. There are various possible formulations of this martingale problem, which can be found in [23], Section 3. The formulation we employ is given by [23], (3.20) and (3.21), and is defined as follows.

DEFINITION D.1 (Wright–Fisher process). A Wright–Fisher process on $\mathcal{P}(\mathbb{K})$ of rate $\theta > 0$ with initial condition $\nu^0 \in \mathcal{P}(\mathbb{K})$ is a continuous $\mathcal{P}(\mathbb{K})$-valued process $(\nu_t)_{0\leq t<\infty}$ such that $\nu_0 := \nu^0$ and which is a solution of the following martingale problem.

We define, for all $n \geq 2$, the maps

$$\Phi_{ij}^{(n)} : \mathcal{B}_b(\mathbb{K}^n) \to \mathcal{B}_b(\mathbb{K}^{n-1}),$$

$$f \mapsto (\Phi_{ij}^{(n)} f : (x_1, \dots, x_{n-1}) \mapsto f(x_1, \dots, x_{j-1}, x_i, x_j, \dots, x_{n-1}).$$

We further define, for all $n \geq 1$ and $f \in \mathcal{B}_b(\mathbb{K}^n)$, the map $\varphi_f \in \mathcal{B}_b(\mathcal{P}(\mathbb{K}))$ by

$$\varphi_f(\nu) := \nu^{\otimes n}(f). \tag{D.1}$$

We then define the generator

(D.2) $$(\mathcal{L}\varphi_f)(\nu) := \theta \sum_{1\le i<j\le n} \big[\nu^{\otimes(n-1)}\big(\Phi_{ij}^{(n)} f\big) - \nu^{\otimes n}(f)\big], \mathcal{D}(\mathcal{L}) = \{\varphi_f \text{ given by (D.1)}\}.$$

The martingale problem defining the Wright–Fisher process of rate $\theta > 0$ is then the condition that, for all φ_f given by (D.1),

(D.3) $$\varphi_f(\nu_t) - \varphi_f(\nu_0) - \int_0^t (\mathcal{L}\varphi_f)(\nu_s)\,ds \quad \text{is a martingale.}$$

THEOREM D.2 ([23]). *We fix $\nu^0 \in \mathcal{P}(\mathbb{K})$ and $\theta > 0$. There exists a unique in law Wright–Fisher process on $\mathcal{P}(\mathbb{K})$ with initial condition $\nu_0 = \nu^0$. Moreover, the sample paths are continuous in the weak atomic metric,*

(D.4) $$(\nu_t)_{0\le t<\infty} \in C\big([0,\infty); \mathcal{P}_{\mathrm{W}_a}(\mathbb{K})\big) \subseteq C\big([0,\infty); \mathcal{P}(\mathbb{K})\big), \quad \textit{almost surely.}$$

Existence and uniqueness of the Wright–Fisher process in $C([0,\infty); \mathcal{P}(\mathbb{K}))$ is given by [23], Theorem 7.1. Continuity of sample paths in the weak atomic metric is given by [23], Corollary 7.4.

We now provide a proof of Proposition 1.3.

PROOF OF PROPOSITION 1.3. The martingales in (D.3) are continuous by [23], Proposition 7.3.

We fix some choice of disjoint measurable sets $A_1, \dots, A_n$ with $\dot\cup_{j=1}^n A_j = \mathbb{K}$. For arbitrary $1 \le i, j \le n$, we take the test functions φ_f, given by (D.1), with the choices of $f(x_1) = \mathbb{1}(x_1 \in A_i)$ and $f(x_1, x_2) = \mathbb{1}(x_1 \in A_i, x_2 \in A_j)$. It follows that

$$\nu_t(A_i) \quad \text{and} \quad \nu_t(A_i)\nu_t(A_j) - \theta \int_0^t \big[\nu_t(A_i)\mathbb{1}(i=j) - \nu_t(A_i)\nu_t(A_j)\big]\,dt$$

are continuous martingales for all $1 \le i, j \le n$. Proposition 1.3 follows. □

APPENDIX E: BROWNIAN MOTION WITH HARD KILLING AT THE BOUNDARY

In this appendix, $(B_t)_{0\le t<\tau_\partial}$ is Brownian motion in an open, connected, bounded domain D, killed instantaneously at the boundary. The Fleming–Viot particle system $(\vec{B}_t^N)_{t\ge 0}$ and Fleming–Viot multicolor process $(\vec{B}_t^N, \vec{\eta}_t^N)_{t\ge 0}$ are driven by this Brownian motion with hard killing (as in Definition 5.1).

We have the following analogue of Theorem A.1.

THEOREM E.1. *There exists a unique quasi-stationary distribution (QSD) $\pi \in \mathcal{P}(D)$ for $(B_t)_{0\le t<\tau_\partial}$. Moreover, there exist constants $C < \infty$ and $k > 0$ such that*

(E.1) $$\big\|\mathcal{L}_\mu(B_t | \tau_\partial > t) - \pi\big\|_{\mathrm{TV}} \le Ce^{-kt} \quad \textit{for all } \mu \in \mathcal{P}(D) \textit{ and } t \ge 0.$$

Furthermore, π is a left eigenmeasure of L with eigenvalue $-\lambda < 0$,

(E.2) $$\langle \pi, Lf\rangle = -\lambda\langle \pi, f\rangle, \quad f \in \mathcal{D}(L).$$

Moreover, L has a positive right eigenfunction belonging to the domain of the Carre du Champs operator described in (5.8), $\phi \in C_0(D; \mathbb{R}_{>0}) \cap C^\infty(D) \cap \mathcal{D}(\Gamma)$. This right eigenfunction is both the unique nonnegative right eigenfunction and the unique right eigenfunction of eigenvalue $-\lambda$, up to rescaling.

We have the following analogue of Theorem 1.8.

DEFINITION E.2. We define a $D \times \mathbb{K}$-valued killed Markov process, denoted by $((B_t, \eta_t))_{0 \le t < \tau_\partial}$, as follows. The process evolves in the first variable as a Brownian motion B_t killed instantaneously upon contact with the boundary. The killing time τ_∂ is then given by $\tau_\partial := \inf\{t > 0 : B_{t-} \in \partial D\}$. In the second variable η_t is a constant element of $\mathbb{K}$, up to the killing time τ_∂, so that $\eta_t = \eta_0$ for all $0 \le t < \tau_\partial$. After the killing time, the process is sent to a fixed cemetery state.

We recall that the stopping time τ_ϵ^N for $\epsilon > 0$ was defined in (5.12) by

$$\{t > 0 : m_t^N(B(\partial D, \delta(\epsilon))) > \epsilon\},$$

whereby $\delta = \delta(\epsilon) > 0$ is given by Lemma 5.3.

THEOREM E.3. *We consider the Fleming–Viot multicolor process $(\vec{B}_t^N, \vec{\eta}_t^N)_{t \ge 0}$ for $N \ge 2$. Then there exists constants $C_{\epsilon,T,N}$ for $\epsilon > 0$, $0 \le T < \infty$ and $N \ge 2$ such that $C_{\epsilon,T,N} \to 0$, as $N \to \infty$, and such that for any initial condition $(\vec{B}_0^N, \vec{\eta}_0^N)$ for which $\tau_\epsilon^N > 0$, and any $f \in \mathcal{B}_b(\bar{D} \times \mathbb{K}; \mathbb{R})$, we have that*

$$\text{(E.3)} \quad \begin{aligned} &\mathbb{E}_{(\vec{B}_0^N, \vec{\eta}_0^N)}\left[\sup_{t \le T}\left|\left(\frac{1}{N}\sum_{i=1}^N \delta_{(B_t^{N,i}, \eta_t^{N,i})} - \mathcal{L}_{\frac{1}{N}\sum_{i=1}^N \delta_{(B_0^{N,i}, \eta_0^{N,i})}}((B_t, \eta_t))\right)(f)\right|\right] \\ &\quad \le C_{\epsilon,T,N}\|f\|_\infty, \end{aligned}$$

$$\text{(E.4)} \quad \mathbb{E}_{(\vec{B}_0^N, \vec{\eta}_0^N)}\left[\sup_{t \le T}\left|J_t^N - \ln \mathbb{P}_{\frac{1}{N}\sum_{i=1}^N \delta_{(B_0^{N,i}, \eta_0^{N,i})}}(\tau_\partial > t)\right| \wedge 1\right] \le C_{\epsilon,T,N}.$$

Finally, we have the following large-time limit for $((B_t, \eta_t))_{0 \le t < \tau_\partial}$ by the same proof as the proof of Proposition 1.9.

PROPOSITION E.4. *For arbitrary sequences $(x^i, \eta^i)_{1 \le i \le n}$ in $D \times \mathbb{K}$, we consider the process $(B_t, \eta_t)_{0 \le t < \tau_\partial}$ with initial distribution given by the empirical measure $\frac{1}{n}\sum_{i=1}^n \delta_{(x^i, \eta^i)}$. Then there exists $c_t \to 0$ as $t \to \infty$ such that, for all sequences $(x^i, \eta^i)_{1 \le i \le n}$ in $\bar{D} \times \mathbb{K}$ and all $n \in \mathbb{N}$, we have*

$$\text{(E.5)} \quad \left\|\mathcal{L}_{\frac{1}{n}\sum_{i=1}^n \delta_{(x^i, \eta^i)}}((B_t, \eta_t) | \tau_\partial > t) - \frac{\sum_{i=1}^n \phi(x^i)\pi \otimes \delta_{\eta^i}}{\sum_{i=1}^n \phi(x^i)}\right\|_{\text{TV}} \le c_t, \quad 0 \le t < \infty.$$

PROOF OF THEOREM E.1, THEOREM E.3 AND PROPOSITION E.4. It is easy to check that $(B_t)_{0 \le t < \tau_\partial}$ satisfies [19], Assumption (A); a proof in the Hörmander setting is given by the present author in [44], Theorem 7.4 (see the proof of [44], Proposition 7.12). The fact that ϕ belongs to the domain of Γ follows from [48], Theorem 1.1. Otherwise, the proofs of Theorem E.1, Theorem E.3 and Proposition E.4 are identical to those of Theorem A.1, Theorem 1.8 and Proposition 1.9, respectively. □

Funding. This work was partially funded by grant 200020 196999 from the Swiss National Foundation and by the EPSRC MathRad programme grant EP/W026899/.

REFERENCES

[1] ALDOUS, D. (1978). Stopping times and tightness. *Ann. Probab.* **6** 335–340. MR0474446 https://doi.org/10.1214/aop/1176995579

[2] ASSELAH, A., FERRARI, P. A. and GROISMAN, P. (2011). Quasistationary distributions and Fleming–Viot processes in finite spaces. *J. Appl. Probab.* **48** 322–332. MR2840302 https://doi.org/10.1239/jap/1308662630

[3] ASSELAH, A., FERRARI, P. A., GROISMAN, P. and JONCKHEERE, M. (2016). Fleming–Viot selects the minimal quasi-stationary distribution: The Galton–Watson case. *Ann. Inst. Henri Poincaré Probab. Stat.* **52** 647–668. MR3498004 https://doi.org/10.1214/14-AIHP635

[4] BALL, K., KURTZ, T. G., POPOVIC, L. and REMPALA, G. (2006). Asymptotic analysis of multiscale approximations to reaction networks. *Ann. Appl. Probab.* **16** 1925–1961. MR2288709 https://doi.org/10.1214/105051606000000420

[5] BERESTYCKI, J., BERESTYCKI, N. and SCHWEINSBERG, J. (2013). The genealogy of branching Brownian motion with absorption. *Ann. Probab.* **41** 527–618. MR3077519 https://doi.org/10.1214/11-AOP728

[6] BERESTYCKI, J., BRUNET, É., NOLEN, J. and PENINGTON, S. (2022). Brownian bees in the infinite swarm limit. *Ann. Probab.* **50** 2133–2177. MR4499276 https://doi.org/10.1214/22-aop1578

[7] BIENIEK, M. and BURDZY, K. (2018). The distribution of the spine of a Fleming–Viot type process. *Stochastic Process. Appl.* **128** 3751–3777. MR3860009 https://doi.org/10.1016/j.spa.2017.12.003

[8] BIENIEK, M., BURDZY, K. and FINCH, S. (2012). Non-extinction of a Fleming–Viot particle model. *Probab. Theory Related Fields* **153** 293–332. MR2925576 https://doi.org/10.1007/s00440-011-0372-5

[9] BROWN, S., JENKINS, P. A., JOHANSEN, A. M. and KOSKELA, J. (2021). Simple conditions for convergence of sequential Monte Carlo genealogies with applications. *Electron. J. Probab.* **26** Paper No. 1. MR4216514 https://doi.org/10.1214/20-ejp561

[10] BRUNET, É. and DERRIDA, B. (1997). Shift in the velocity of a front due to a cutoff. *Phys. Rev. E* (3) **56** 2597–2604. MR1473413 https://doi.org/10.1103/PhysRevE.56.2597

[11] BRUNET, É. and DERRIDA, B. (2001). Effect of microscopic noise on front propagation. *J. Stat. Phys.* **103** 269–282. MR1828730 https://doi.org/10.1023/A:1004875804376

[12] BRUNET, É., DERRIDA, B., MUELLER, A. H. and MUNIER, S. (2006). Noisy traveling waves: Effect of selection on genealogies. *Europhys. Lett.* **76** 1–7. MR2299937 https://doi.org/10.1209/epl/i2006-10224-4

[13] BRUNET, É., DERRIDA, B., MUELLER, A. H. and MUNIER, S. (2007). Effect of selection on ancestry: An exactly soluble case and its phenomenological generalization. *Phys. Rev. E* (3) **76** 041104. MR2365627 https://doi.org/10.1103/PhysRevE.76.041104

[14] BURDZY, K. List of open problems.

[15] BURDZY, K. and ENGLÄNDER, J. (2021). The spine of the Fleming–Viot process driven by Brownian motion. Available at arXiv:2112.01720.

[16] BURDZY, K., HOŁYST, R. and MARCH, P. (2000). A Fleming–Viot particle representation of the Dirichlet Laplacian. *Comm. Math. Phys.* **214** 679–703. MR1800866 https://doi.org/10.1007/s002200000294

[17] BURDZY, K., KOŁODZIEJEK, B. and TADIĆ, T. (2019). Inverse exponential decay: Stochastic fixed point equation and ARMA models. *Bernoulli* **25** 3939–3977. MR4010978 https://doi.org/10.3150/19-bej1116

[18] BURDZY, K., KOŁODZIEJEK, B. and TADIĆ, T. (2022). Stochastic fixed-point equation and local dependence measure. *Ann. Appl. Probab.* **32** 2811–2840. MR4474520 https://doi.org/10.1214/21-aap1749

[19] CHAMPAGNAT, N. and VILLEMONAIS, D. (2016). Exponential convergence to quasi-stationary distribution and Q-process. *Probab. Theory Related Fields* **164** 243–283. MR3449390 https://doi.org/10.1007/s00440-014-0611-7

[20] CHAMPAGNAT, N. and VILLEMONAIS, D. (2017). Uniform convergence to the Q-process. *Electron. Commun. Probab.* **22** Paper No. 33. MR3663104 https://doi.org/10.1214/17-ECP63

[21] DE MULATIER, C., DUMONTEIL, E., ROSSO, A. and ZOIA, A. (2015). The critical catastrophe revisited. *J. Stat. Mech. Theory Exp.* **8** P08021. MR3400257 https://doi.org/10.1088/1742-5468/2015/08/p08021

[22] DURRETT, R. and REMENIK, D. (2011). Brunet–Derrida particle systems, free boundary problems and Wiener–Hopf equations. *Ann. Probab.* **39** 2043–2078. MR2932664 https://doi.org/10.1214/10-AOP601

[23] ETHIER, S. N. and KURTZ, T. G. (1993). Fleming–Viot processes in population genetics. *SIAM J. Control Optim.* **31** 345–386. MR1205982 https://doi.org/10.1137/0331019

[24] ETHIER, S. N. and KURTZ, T. G. (1994). Convergence to Fleming–Viot processes in the weak atomic topology. *Stochastic Process. Appl.* **54** 1–27. MR1302692 https://doi.org/10.1016/0304-4149(94)00006-9

[25] FLEMING, W. H. and VIOT, M. (1979). Some measure-valued Markov processes in population genetics theory. *Indiana Univ. Math. J.* **28** 817–843. MR0542340 https://doi.org/10.1512/iumj.1979.28.28058

[26] FRANKHAM, R. (1995). Effective population size/adult population size ratios in wildlife: A review. *Genet. Res.* **66** 95–107.

[27] GIBBS, A. L. and SU, F. E. (2002). On choosing and bounding probability metrics. *Int. Stat. Rev.* **70** 419–435.

[28] GRIGORESCU, I. (2007). Large deviations for a catalytic Fleming–Viot branching system. *Comm. Pure Appl. Math.* **60** 1056–1080. MR2319055 https://doi.org/10.1002/cpa.20174

[29] GRIGORESCU, I. and KANG, M. (2012). Immortal particle for a catalytic branching process. *Probab. Theory Related Fields* **153** 333–361. MR2925577 https://doi.org/10.1007/s00440-011-0347-6

[30] KATZENBERGER, G. S. (1991). Solutions of a stochastic differential equation forced onto a manifold by a large drift. *Ann. Probab.* **19** 1587–1628. MR1127717

[31] KURTZ, T. G. (1992). Averaging for martingale problems and stochastic approximation. In *Applied Stochastic Analysis* (*New Brunswick*, *NJ*, 1991) (I. Karatzas and D. Ocone, eds.). *Lect. Notes Control Inf. Sci.* **177** 186–209. Springer, Berlin. MR1169928 https://doi.org/10.1007/BFb0007058

[32] LABBÉ, C. (2013). Flots stochastiques et représentation lookdown. Ph.D. thesis, Université Pierre-et-Marie-Curie.

[33] LIONS, P.-L. and SZNITMAN, A.-S. (1984). Stochastic differential equations with reflecting boundary conditions. *Comm. Pure Appl. Math.* **37** 511–537. MR0745330 https://doi.org/10.1002/cpa.3160370408

[34] LÖBUS, J.-U. (2009). A stationary Fleming–Viot type Brownian particle system. *Math. Z.* **263** 541–581. MR2545857 https://doi.org/10.1007/s00209-008-0430-6

[35] MAILLARD, P. (2016). Speed and fluctuations of N-particle branching Brownian motion with spatial selection. *Probab. Theory Related Fields* **166** 1061–1173. MR3568046 https://doi.org/10.1007/s00440-016-0701-9

[36] MALLEIN, B. (2017). Branching random walk with selection at critical rate. *Bernoulli* **23** 1784–1821. MR3624878 https://doi.org/10.3150/15-BEJ796

[37] MÉLÉARD, S. and TRAN, V. C. (2012). Slow and fast scales for superprocess limits of age-structured populations. *Stochastic Process. Appl.* **122** 250–276. MR2860449 https://doi.org/10.1016/j.spa.2011.08.007

[38] MÉLÉARD, S. and VILLEMONAIS, D. (2012). Quasi-stationary distributions and population processes. *Probab. Surv.* **9** 340–410. MR2994898 https://doi.org/10.1214/11-PS191

[39] PENINGTON, S., ROBERTS, M. I. and TALYIGÁS, Z. (2022). Genealogy and spatial distribution of the N-particle branching random walk with polynomial tails. *Electron. J. Probab.* **27** Paper No. 93. MR4456776 https://doi.org/10.1214/22-ejp806

[40] SCHWAB, C. (2005). Krein–Rutman theorem and the principal eigenvalue. In *Numerical Methods for Elliptic and Parabolic PDEs* (*Lecture Notes*).

[41] STROOCK, D. W. and VARADHAN, S. R. S. (1971). Diffusion processes with boundary conditions. *Comm. Pure Appl. Math.* **24** 147–225. MR0277037 https://doi.org/10.1002/cpa.3160240206

[42] TOUGH, O. (2021). Scaling limit of the Fleming–Viot multicolor process. Available at arXiv:2110.05049.

[43] TOUGH, O. (2021). Asymptotic behaviour of the Fleming–Viot process. Ph.D. thesis, Duke Univ.

[44] TOUGH, O. (2022). L^∞-convergence to a quasi-stationary distribution. Available at arXiv:2210.13581.

[45] TOUGH, O. (2023). Selection principle for the Fleming–Viot process with drift -1. Available at arXiv:2306.03585.

[46] TOUGH, O. and NOLEN, J. (2022). The Fleming–Viot process with McKean–Vlasov dynamics. *Electron. J. Probab.* **27** Paper No. 101. MR4460269 https://doi.org/10.1214/22-ejp820

[47] VILLEMONAIS, D. (2014). General approximation method for the distribution of Markov processes conditioned not to be killed. *ESAIM Probab. Stat.* **18** 441–467. MR3333998 https://doi.org/10.1051/ps/2013045

[48] XU, X. (2009). Gradient estimates for the eigenfunctions on compact manifolds with boundary and Hörmander multiplier theorem. *Forum Math.* **21** 455–476. MR2526794 https://doi.org/10.1515/FORUM.2009.021