The Annals of Applied Probability
2012, Vol. 22, No. 2, 827–859
DOI: 10.1214/11-AAP788

STOCHASTIC MODEL FOR CELL POLARITY

BY ANKIT GUPTA[1]

University of Wisconsin, Madison

Cell polarity refers to the spatial asymmetry of molecules on the cell membrane. Altschuler, Angenent, Wang and Wu have proposed a stochastic model for studying the emergence of polarity in the presence of feedback between molecules. We analyze their model further by representing it as a model of an evolving population with interacting individuals. Under a suitable scaling of parameters, we show that in the infinite population limit we get a Fleming–Viot process. Using well-known results for such processes, we establish that cell polarity is exhibited by the model and also study its dependence on the biological parameters of the model.

1. Introduction. The phenomenon of polarity is ubiquitous in living organisms. It is known to occur at many levels: from cellular to organismic. Polarity is what causes one part of a biological system to be different from another. Understanding how polarity is established and maintained is a matter of fundamental concern for biologists.

In this paper we are interested in polarity at the level of individual cells. Consider a spherical cell consisting of the cytosol and the membrane. Suppose that it contains numerous molecules that may either reside in the cytosol or on the membrane. The phenomenon of cell polarity refers to an identifiable form of spatial asymmetry of molecules on the membrane. Biologists generally consider a cell to be in a *polarized* state when most of the membrane molecules appear to be concentrated around a single site or located in a single hemisphere on the membrane. It is known that many types of cells exhibit this phenomenon. The most common example is the yeast cell (see [22, 30, 34, 35]), but there are many others (see [9, 10]). As noted in [9], cell polarity is vital in the creation of functionally specialized regions on the membrane, which can then facilitate cellular processes such as localized membrane growth, activation of immune response, directional cell migration and vectorial transport of molecules across cell layers.

Due to its importance, many attempts have been made to investigate the mechanisms responsible for cell polarity. Drubin and Nelson [9] mention that the existence of cell polarity involves positive feedback from the signaling molecules on

Received July 2010; revised May 2011.
[1]Supported in part by NSF Grants DMS-05-53687 and DMS-08-05793.
MSC2010 subject classifications. 60G57, 60J68, 92C37, 92C42.
Key words and phrases. Fleming–Viot, cell polarity, spatial clustering, Donnelly–Kurtz, particle representation.

the membrane. This feedback enables the signaling molecules to perform localized recruitment, thereby causing concentration of molecules in a specified region on the membrane. Examples of such signaling molecules include Cdc 42 in budding yeast (see [3]), mPar3/mPar6 in neurons (see [38]), Rac in kidney cells (see [18]) and human chemotaxing neutrophils (see [44]). Even though the feedback mechanism may bring the molecules together, it is unclear if it can generate cell polarity alone. This is because the molecules on the membrane are constantly diffusing and, hence, any clusters that form may disappear quickly with time. Biologists have proposed that additional mechanisms like directed transport and coupled inhibitors are required to counter the spatial diffusion and generate spatial asymmetry (see [19, 22, 31, 41, 43]). However, these additional mechanisms are not always found in cells that exhibit polarity. Hence, the question arises whether feedback alone can cause polarization.

Altschuler, Angenent, Wang and Wu [1] show that indeed feedback alone can generate cell polarity when the number of molecules is small. They prove this result via a simple mathematical model derived by abstracting the feedback circuits found in cells. In their model, the feedback mechanism is given by the following: a molecule on the membrane may pull a molecule from the cytosol to its location on the membrane. In a stochastic setting they show that their model exhibits recurring cell polarity. However, the frequency of polarity is inversely proportional to the number of molecules in the cell. This suggests that no polarity can persist in the infinite population limit without any additional mechanisms to reinforce asymmetry.

In this paper we will scale some parameters of the model in [1] and study the resulting model. The main result of our paper is that if we let the feedback strength of each membrane bound molecule increase linearly with the population size, then we do get recurring cell polarity in the infinite population limit. Hence, under our scaling, the model suggests that feedback alone can generate cell polarity in the infinite population limit without any additional mechanisms. Our approach is to express the dynamics of cell molecules as a measure-valued Markov process and then prove that in the limit, the dynamics of molecules on the membrane can be described by a Fleming–Viot process. This process was introduced by Fleming and Viot [17] in 1979 and it has been very well studied since then. An excellent survey of Fleming–Viot processes is given by Ethier and Kurtz [12]. Using the results already known for such processes, we will first show that the limiting process is ergodic and hence has a unique stationary distribution. We will then illustrate that at stationarity the membrane molecules are arranged into *clans* of various sizes and molecules in a clan are spatially clustered. Moreover, the distribution of clan sizes and the expected spatial spread of the clans can be readily computed in terms of the biological parameters of the model. Our results will allow us to deduce that there are times when most of the molecules are part of a single clan and lie in a

single hemisphere on the membrane, thereby causing a cell to polarize. We now describe the model given in [1].

DESCRIPTION 1.1. There are N molecules in the cell (cytosol and membrane). The cell itself is a sphere of radius R. The following four events can change the molecular configuration in the cell:

- *Spontaneous membrane association*: A molecule in the cytosol moves to a random location on the membrane at rate k_{on}.
- *Spontaneous membrane dissociation*: A molecule on the membrane moves back into the cytosol at rate k_{off}.
- *Membrane association through recruitment* (*feedback mechanism*): A molecule on the membrane recruits another molecule from the cytosol at rate $k_{\text{fb}} \times$ (*fraction of molecules in the cytosol*).

  At the time of recruitment, the *recruited* particle moves to the location of the *recruiting* particle.
- *Membrane diffusion*: Each molecule on the membrane does Brownian motion with speed D.

The parameters of the model N, D, R, k_{on}, k_{fb} and k_{off} have clear biological interpretations. As mentioned in [1], k_{fb} and k_{off} are comparable and throughout this paper we will assume the following:

ASSUMPTION 1.2.

$$k_{\text{fb}} > k_{\text{off}} > 0.$$

In this paper we scale up k_{fb} and k_{off} by the population size N and leave k_{on} the same. We show that under this scaling the model becomes mathematically tractable as $N \to \infty$. In Section 3 we will discuss the choice of this scaling and the necessity of Assumption 1.2.

Since we will be relating this model to a well-known model in population genetics, it is convenient to think of cell molecules as individuals in an evolving population. Consider the membrane molecules as being *alive* and the cytosol molecules as being *dead*. Each membrane molecule has two attributes: location and clan indicator. When a membrane molecule recruits another molecule from the cytosol, this new molecule gets initially assigned the same location and clan indicator as the recruiting molecule. The location of this new molecule will change subsequently, as it does its own Brownian motion but its clan indicator remains the same. We can think of membrane recruitment as a *birth* process in which the recruiting membrane molecule (the *parent*) passes its characteristics to the recruited molecule (the *offspring*). The membrane molecules that have the same clan indicator are said to be in the same *clan*, which implies that they have a common ancestor. When a molecule spontaneously associates itself to the membrane, we assign it a new clan

indicator and a randomly chosen location on the membrane. Therefore, we can think of spontaneous association as *immigration* in which the individuals bring new genetic traits into the population. When a membrane molecule spontaneously dissociates from the membrane and goes into the cytosol, it loses both its attributes. So we can think of spontaneous dissociation as *death*. Note that a molecule that dies can get reincarnated.

At any time, the membrane molecules can be classified into clans based on their ancestry. Since the molecules in a clan have a common ancestor, if the diffusion constant D is small, we can expect them to be clustered on the membrane. However, the molecular diffusion may cause a clan to spread apart with time. Surprisingly, this does not happen in our model. We mentioned before that in the infinite population limit, the cell dynamics is ergodic and reaches a stationary state at which the spatial spread of the clans does not change with time. This is due to the extremely fast nature of the birth and death mechanisms in our model which causes most of the molecules in a clan to be *newly born*. Hence, they have been unable to move away from their common ancestor by too much. We will show that in the limit there are infinitely many clans present in the population at stationarity, but there are only a few *large* clans. These two results together imply that spatial asymmetry is present and persistent. We will then argue that there will be times when most of the population will be part of one large clan and also appear to concentrate around a single point. Consequently, the cell is polarized at these times. This shows that unlike the original model, cell polarity is present in our rescaled model as the population size goes to infinity. For a detailed study of the model considered here, we refer the readers to [21].

This paper is organized as follows. In Section 2 we give the main results of our paper. In Section 3 we interpret these results in the context of biology and compare our results with the results provided in [1] for the original model. We also state some interesting research questions that we were unable to answer in this paper. Finally, in Section 4 we provide the proofs of our results.

Notation. We now introduce some notation that will be used throughout the paper. Let (S, d) be a compact metric space. Then by $B(S)(C(S))$ we refer to the set of all bounded (continuous) real-valued Borel measurable functions. Since (S, d) is compact, $C(S) \subset B(S)$. Both $B(S)$ and $C(S)$ are Banach spaces under the sup norm $\|f\| = \sup_{x\in S} |f(x)|$. For any differentiable manifold M and $k \geq 1$, let $C^k(M)$ be the space of functions which are k-times continuously differentiable. Let $\mathcal{B}(S)$ be the Borel sigma field on S. The space of all positive Borel measures with total measure bounded above by 1 is denoted by $\mathcal{M}_1(S)$ and $\mathcal{P}(S)$ is the space of all Borel probability measures. Since (S, d) is compact, Prohorov's theorem implies that both $\mathcal{P}(S)$ and $\mathcal{M}_1(S)$ are compact under the topology of weak convergence. For any $\mu \in \mathcal{M}_1(S)$ and $f : S \to \mathbb{R}$ let

$$\langle f, \mu \rangle = \int_S f(s)\mu(ds).$$

If $\mu \in \mathcal{M}_1(S)$, then for any positive integer m, $\mu^m \in \mathcal{M}_1(S^m)$ refers to the m-fold product of μ. If μ is an atomic measure of the form $a_n \sum_{i=1}^n \delta_{x_i}$ for some $a_n > 0$, then $\mu^{(m)}$ is the *symmetric* m-fold product of μ defined by

$$\mu^{(m)} = \frac{1}{n(n-1)\cdots(n-m+1)} \sum_{1 \le i_1 \ne i_2 \ne \cdots \ne i_m \le n} \delta_{(x_{i_1}, x_{i_2}, \ldots, x_{i_m})}, \tag{1.1}$$

where the sum is over all distinct m-tuples of $\{1, 2, \ldots, n\}$. If $m > n$, then the sum above is empty and $\mu^{(m)}$ is taken to be 0. Observe that $\mu^{(m)}$ does not depend on a_n and for $n > m$ it is a probability measure over S^m. Also note that if μ is a probability measure (i.e., $a_n = 1/n$), then for large n, $\mu^{(m)}$ is approximately equal to μ^m.

The space of cadlag functions (i.e., right continuous functions with left limits) from $[0, \infty)$ to S is called $D_S[0, \infty)$ and it is endowed with the Skorohod topology (for details see Chapter 3, Ethier and Kurtz [11]). The space of continuous functions from $[0, \infty)$ to S is called $C_S[0, \infty)$ and it is endowed with the topology of uniform convergence over compact sets.

For any operator $A \subset B(S) \times B(S)$, let $\mathcal{D}(A)$ and $\mathcal{R}(A)$ designate the domain and range of A. The notion of the *martingale problem* associated to an operator A is introduced and developed in Chapter 4, Ethier and Kurtz [11]. In this paper, by a solution of the martingale problem for A, we mean a measurable stochastic process X with paths in $D_S[0, \infty)$ such that for any $f \in \mathcal{D}(A)$,

$$f(X(t)) - \int_0^t Af(X(s))\, ds$$

is a martingale with respect to the filtration generated by X. For a given initial distribution $\pi \in \mathcal{P}(S)$, a solution X of the martingale problem for A is a solution of the martingale problem for (A, π) if $\pi = PX(0)^{-1}$. If such a solution exists uniquely for all $\pi \in \mathcal{P}(S)$, then we say that the martingale problem for A is well posed.

2. The main results. Our first task in this section is to represent the dynamics of cell molecules as a measure-valued Markov process. Suppose there are N molecules in the cell (cytosol and membrane). The cell membrane will be denoted by E and it is a sphere of radius R in $\mathbb{R}^3$. As we mentioned before, each membrane molecule has two attributes: location and clan indicator. The locations are elements in E, while the clan indicators will be chosen as elements in the unit interval $[0, 1]$. Hence, $E \times [0, 1]$ is the type space for the molecules. A molecule of type $x = (y, z) \in E \times [0, 1]$ is located at y on the membrane and has z as its clan indicator. Note that a membrane molecule will change its type only due to Brownian motion. Therefore, during its stay on the membrane, only its location (first coordinate) changes while its clan indicator (second coordinate) remains fixed.

If there are N molecules in the cell, then we assign mass $1/N$ to each molecule. The membrane population at time t can be represented by an atomic measure as follows:

$$\mu^N(t) = \frac{1}{N} \sum_{i=1}^{n^N(t)} \delta_{x_i(t)}, \tag{2.1}$$

where $n^N(t) = N\langle 1, \mu^N(t)\rangle$ is the number of molecules on the membrane at time t and $x_1(t), \ldots, x_{n^N(t)}$ are their types. Viewed as a process, μ^N is Markov and its state space is given by

$$\mathcal{M}_a^N(E \times [0,1]) = \left\{ \frac{1}{N} \sum_{i=1}^{n} \delta_{x_i} : 0 \le n \le N \text{ and } x_1, \ldots, x_n \in E \times [0,1] \right\}.$$

For any $\mu \in \mathcal{M}_a^N(E \times [0,1])$, the total mass $\langle 1, \mu\rangle \le 1$ and, hence, $\mathcal{M}_a^N(E \times [0,1]) \subset \mathcal{M}_1(E \times [0,1])$. If we endow $\mathcal{M}_a^N(E \times [0,1])$ with the topology of weak convergence, then it is a compact space.

The generator of any Markov process is an operator which captures the rate of change of the distribution of the process. For a detailed discussion on generators, see Chapter 4 in Ethier and Kurtz [11]. For a speed D Brownian motion on the membrane E, the generator is given by $\frac{D}{2}\Delta$, where Δ is the Laplace–Beltrami operator on E. Note that $C^2(E) \subset \mathcal{D}(\Delta)$, where $C^2(E)$ is the space of twice continuously differentiable functions on E. Next we define the classes of functions that we will use in this paper.

DEFINITION 2.1.

$$\mathcal{C} = \big\{ f \in C\big((E \times [0,1])^m\big) \text{ such that } f(\cdot, z) \in C^2(E^m) \text{ for all } z \in [0,1]^m, \text{ and } \nabla f(x, \cdot), \Delta f(x, \cdot) \in C([0,1]^m) \text{ for all } x \in E^m \text{ and } m \ge 1 \big\}.$$

DEFINITION 2.2.

$$\bar{\mathcal{C}} = \big\{ F(\mu) = \big\langle f, \mu^{(m)} \big\rangle \text{ such that } f \in \mathcal{C} \text{ and } m \ge 1 \big\}.$$

We now specify the generator $\mathbb{A}^N$ for the process μ^N, which captures the dynamics of the cell population. For any $f \in \mathcal{C}$ let $\Delta_i f$ denote the action of the Laplace–Beltrami operator on f by considering it as a function of its ith coordinate. Let the domain of the operator $\mathbb{A}^N$ be $\mathcal{D}(\mathbb{A}^N) = \bar{\mathcal{C}}$ and for $F \in \bar{\mathcal{C}}$ of the form

$\langle f, \mu^{(m)} \rangle$, define

$$
\begin{aligned}
\mathbb{A}^N F(\mu) = {} & \frac{D}{2} \sum_{i=1}^{m} \langle \Delta_i f, \mu^{(m)} \rangle \\
& + k_{\mathrm{on}} N (1-h) \int_E \int_{[0,1]} \left(F\left(\mu + \frac{1}{N}\delta_{(y,z)}\right) - F(\mu) \right) \vartheta(dy)\, dz \\
& + k_{\mathrm{off}} N^2 \int_{E\times[0,1]} \left(F\left(\mu - \frac{1}{N}\delta_x\right) - F(\mu) \right) \mu(dx) \\
& + k_{\mathrm{fb}} N^2 (1-h) \int_{E\times[0,1]} \left(F\left(\mu + \frac{1}{N}\delta_x\right) - F(\mu) \right) \mu(dx),
\end{aligned}
\tag{2.2}
$$

where $h = \langle 1, \mu \rangle$ and $\vartheta(\cdot)$ is the surface area measure on the sphere E normalized to have total area as 1. Terms in the operator above correspond to the surface diffusion of the membrane molecules, spontaneous association, spontaneous dissociation and membrane recruitment, in that order. The martingale problem for $\mathbb{A}^N$ is well posed and this can be seen by viewing the operator $\mathbb{A}^N$ as a bounded perturbation of the diffusion operator (given by the first term of $\mathbb{A}^N$). It is easy to argue that the martingale problem for the diffusion operator is well posed and the solution for any initial distribution is just the empirical measure process of a system of particles doing independent speed D Brownian motion on E. Proposition 10.2 and Theorem 10.3 in Chapter 4 of Ethier and Kurtz [11] imply the well-posedness of the martingale problem for $\mathbb{A}^N$.

It will soon become evident that the initial distribution of cell molecules is not important for the discussion in this paper. For definiteness we will assume that the membrane is initially empty. Let $\bar{\pi}_0 \in \mathcal{P}(\mathcal{M}_a^N(E \times [0,1]))$ be the distribution that puts all the mass at the 0 measure. From now on μ^N will be the unique Markovian solution to the martingale problem corresponding to $(\mathbb{A}^N, \bar{\pi}_0)$.

We define another process h^N by

$$
h^N(t) = \langle 1, \mu^N(t) \rangle = \frac{n^N(t)}{N}, \qquad t \geq 0. \tag{2.3}
$$

At any time t, $h^N(t)$ is the fraction (or the total mass) of cell molecules that are on the membrane. We will refer to h^N as the *fraction process*. Observe that $h^N(0) = 0$.

We are interested in showing the convergence of the sequence of processes $\{\mu^N\}$ as $N \to \infty$. Note that the last two terms in $\mathbb{A}^N$ [see (2.2)] do not appear to converge independently. This is because terms like

$$
\int_{E\times[0,1]} \left(F\left(\mu \pm \frac{1}{N}\delta_x\right) - F(\mu) \right) \mu(dx)
$$

will typically be of order $1/N$ and we are multiplying them by N^2 outside. However, convergence does happen because these two terms combine to give a second-order term. Instead of directly dealing with the sequence of generators $\{\mathbb{A}^N\}$, we

will prove the convergence result in Section 4 by using the particle construction introduced by Donnelly and Kurtz in [8]. This construction provides a more probabilistic way of passing to the limit.

Define an operator $\mathbb{A}$ as follows. For any $F \in C(\mathcal{M}_1(E \times [0,1]))$ of the form $F(\mu) = \prod_{i=1}^m \langle f_i, \mu\rangle$, where $f_i \in \mathcal{C} \cap C(E \times [0,1])$ for all $i = 1, 2, \ldots, m$ and $m \geq 1$, define

$$\begin{aligned}\mathbb{A}F(\mu) &= \frac{D}{2}\sum_{i=1}^m \langle \Delta f_i, \mu\rangle \prod_{j\neq i}\langle f_j, \mu\rangle \\ &\quad + k_{\mathrm{on}}\frac{(1-h_{\mathrm{eq}})}{h_{\mathrm{eq}}} \\ &\quad \times \sum_{i=1}^m \int_E \int_{[0,1]} \bigl(f_i(y,z)\vartheta(dy)\,dz - f_i(x)\mu(dx)\bigr)\prod_{j\neq i}\langle f_j, \mu\rangle \\ &\quad + \frac{k_{\mathrm{fb}}(1-h_{\mathrm{eq}})}{h_{\mathrm{eq}}} \sum_{1\leq i\neq j\leq m} \bigl(\langle f_i f_j, \mu\rangle - \langle f_i, \mu\rangle\langle f_j, \mu\rangle\bigr) \prod_{k\neq i,j}\langle f_k, \mu\rangle.\end{aligned} \tag{2.4}$$

The operator $\mathbb{A}$ is the generator of a Fleming–Viot process. The martingale problem corresponding to it is well posed and each solution has paths in $C_{\mathcal{P}(E\times[0,1])}[0,\infty)$ by Theorem 3.2, Ethier and Kurtz [12]. We are now ready to state the convergence result. Throughout this paper $\Rightarrow$ will denote convergence in distribution.

THEOREM 2.3. *There exists a stopping time τ_N (with respect to filtration generated by μ^N) satisfying $\tau_N \to 0$ a.s. as $N \to \infty$, such that if we define processes $\hat{h}^N$ and $\hat{\mu}^N$ as*

$$\hat{h}^N(t) = h^N(t+\tau_N), \qquad t \geq 0, \tag{2.5}$$

and

$$\hat{\mu}^N(t) = \mu^N(t+\tau_N), \qquad t \geq 0, \tag{2.6}$$

then the following is true.

(A) *For any $T > 0$,*

$$\sup_{0\leq t\leq T} |\hat{h}^N(t) - h_{\mathrm{eq}}| \Rightarrow 0 \qquad \text{as } N \to \infty,$$

where $h_{\mathrm{eq}} = 1 - \frac{k_{\mathrm{off}}}{k_{\mathrm{fb}}}$.

(B) *Suppose that the sequence of random variables $\{\hat{\mu}^N(0)\}$ converges in distribution to $\mu(0)$ and let $\pi_0 \in \mathcal{P}(\mathcal{P}(E \times [0,1]))$ be the distribution of $\mu(0)/h_{\mathrm{eq}}$. Then $\hat{\mu}^N \Rightarrow \mu$ in $D_{\mathcal{M}_1(E\times[0,1])}[0,\infty)$ as $N \to \infty$, where $\mu = h_{\mathrm{eq}}\nu$ and ν is the Fleming–Viot process with type space $E \times [0,1]$, generator $\mathbb{A}$ and initial distribution π_0.*

REMARK 2.4. Note that the state space of the processes $\hat{\mu}^N$ is $\mathcal{M}_1(E \times [0,1])$, which is compact and so $\mathcal{P}(\mathcal{M}_1(E \times [0,1]))$ is also compact by Prohorov's theorem. Hence, the distributions of $\hat{\mu}^N(0)$ will certainly converge along a subsequence and the assertion of the theorem above will hold for this subsequence. In fact, the distributions of $\hat{\mu}^N(0)$ converge along the entire sequence (see Remark 4.4).

A heuristic explanation for the above result is as follows. As N gets larger, the extremely fast nature of the birth and death mechanisms forces the fraction process to immediately settle to an equilibrium value h_{eq} [given by part (A) of the above theorem]. Note that $k_{\mathrm{fb}} h_{\mathrm{eq}}(1 - h_{\mathrm{eq}}) = k_{\mathrm{off}} h_{\mathrm{eq}}$ and so at this equilibrium value, the net influx of population onto the membrane due to birth matches the net efflux of population from the membrane due to death. Since the total mass on the membrane is not allowed to deviate from this equilibrium, any addition of new mass due to immigration must be concurrently offset by an equal reduction in existing mass due to death. Hence, in the limit, the net demographic effect of immigration is the same as that of mutation and, therefore, we see a mutation-like term in the limiting generator $\mathbb{A}$ [see the second term in (2.4)]. Similarly, the addition of new mass on the membrane due to birth must be accompanied by the reduction of equal mass due to death. This gives rise to the second-order sampling term in $\mathbb{A}$ [see the third term in (2.4)]. These ideas are made rigorous in the proof of Theorem 2.3 given in Section 4.

From now on ν will denote the Fleming–Viot process given in the statement of Theorem 2.3. We next claim that ν has a unique stationary distribution and it is also *strongly ergodic* in the sense that its transition function converges asymptotically to the stationary distribution. In fact, this convergence is exponentially fast. Let S be any metric space and let $\mathcal{B}(S)$ be the Borel sigma field on S. Define the *total variation* metric over the space of probability measures $\mathcal{P}(S)$ by

$$\|\nu_1 - \nu_2\|_{\mathrm{var}} = \sup_{\Gamma \in \mathcal{B}(S)} \|\nu_1(\Gamma) - \nu_2(\Gamma)\|.$$

PROPOSITION 2.5. (A) *The process ν is strongly ergodic and it has a unique stationary distribution $\Pi \in \mathcal{P}(\mathcal{P}(E \times [0,1]))$.*

(B) *The transition function of ν converges to the stationary distribution exponentially fast. There exists a constant $C > 0$ such that*

$$\big\| P\big(\nu(t) \in \cdot\big) - \Pi(\cdot) \big\|_{\mathrm{var}} \le C \exp\left(-\left(\frac{k_{\mathrm{on}}(1 - h_{\mathrm{eq}})}{2h_{\mathrm{eq}}} \right) t \right).$$

PROOF. Both parts follow from Theorem 5.1 and Corollary 8.4 in Ethier and Kurtz [12]. □

Define a $\mathcal{P}([0,1])$-valued process ν_c by

$$\nu_c(t, A) = \nu(t, E \times A), \qquad A \in \mathcal{B}([0,1]) \text{ and } t \ge 0. \tag{2.7}$$

We will refer to ν_c as the *clan process*, as it will help us in the determination of clan sizes. We shall discuss this further in Section 3. As a consequence of Theorem 2.3, we get the following corollary.

COROLLARY 2.6. *The process ν_c is the Fleming–Viot process with type space* $[0, 1]$ *and generator $\mathbb{A}_c$ given by*

$$\mathbb{A}_c F(\mu) = k_{\text{on}} \frac{(1 - h_{\text{eq}})}{h_{\text{eq}}} \sum_{i=1}^{m} \int_{[0,1]} \big(f_i(z)\,dz - f_i(x)\mu(dx)\big) \prod_{j \neq i} \langle f_j, \mu \rangle + \frac{k_{\text{fb}}(1 - h_{\text{eq}})}{h_{\text{eq}}} \sum_{1 \le i \neq j \le m} \big(\langle f_i f_j, \mu \rangle - \langle f_i, \mu \rangle \langle f_j, \mu \rangle\big) \prod_{k \neq i,j} \langle f_k, \mu \rangle,$$

where $F(\mu) = \prod_{i=1}^{m} \langle f_i, \mu \rangle$ and $f_i \in C([0, 1])$ for $i = 1, 2, \ldots, m$.

PROOF. The proof is immediate from the definition of ν_c and the descriptions of the generators $\mathbb{A}$ and $\mathbb{A}_c$. □

Since the molecules are constantly diffusing on the membrane, we would expect each clan to spread out more and more with time. However, we will show that this does not happen in our model. We would like to measure the average spatial spread of the molecules that belong to the same clan. One way to measure it would be to randomly sample two molecules from the membrane population at any time t and compute their expected distance squared, given that they are in the same clan. We call this quantity $S_p(t)$. For $i = 1, 2$ let $X_i(t) = (Y_i(t), C_i(t)) \in E \times [0, 1]$ be the sampled molecules. Then given $\nu(t)$, $X_1(t)$ and $X_2(t)$ are i.i.d. with common distribution $\nu(t)$. Therefore,

$$\begin{aligned} S_p(t) &= E\big(\|Y_1(t) - Y_2(t)\|^2 | C_1(t) = C_2(t)\big) \\ &= \frac{E(\|Y_1(t) - Y_2(t)\|^2 1_{\{C_1(t)=C_2(t)\}})}{P(C_1(t) = C_2(t))} \\ &= \frac{E(E(\|Y_1(t) - Y_2(t)\|^2 1_{\{C_1(t)=C_2(t)\}} | \nu(t)))}{E(E(C_1(t) = C_2(t) | \nu(t)))} \\ &= \frac{E(\int_E \int_{[0,1]} \|y_1 - y_2\|^2 1_{\{c_1=c_2\}} \nu(t, dy_1, dc_1)\nu(t, dy_2, dc_2))}{E(\int_E \int_{[0,1]} 1_{\{c_1=c_2\}} \nu(t, dy_1, dc_1)\nu(t, dy_2, dc_2))}. \end{aligned}$$

From Proposition 2.5 we know that the process ν has a unique stationary distribution $\Pi \in \mathcal{P}(\mathcal{P}(E \times [0, 1]))$. At stationarity, $S_p(t)$ does not depend on t and can be

written as

$$S_p = \int_{\mathcal{P}(E\times[0,1])} \left(\int_E \int_{[0,1]} \|y_1 - y_2\|^2 1_{\{c_1=c_2\}} \mu(dy_1, dc_1)\mu(dy_2, dc_2) \right) \times \Pi(d\mu) \tag{2.8}$$
$$\times \left(\int_{\mathcal{P}(E\times[0,1])} \left(\int_E \int_{[0,1]} 1_{\{c_1=c_2\}} \mu(dy_1, dc_1)\mu(dy_2, dc_2) \right) \Pi(d\mu) \right)^{-1}.$$

The theorem below gives a precise formula for S_p. It will be proved in Section 4.

THEOREM 2.7. *Let* $\alpha = \frac{1-h_{\text{eq}}}{h_{\text{eq}}} = \frac{k_{\text{off}}}{k_{\text{fb}}-k_{\text{off}}}$. *Then*

$$S_p = \frac{2D}{((k_{\text{on}} + k_{\text{fb}})\alpha + D/R^2)}.$$

In the next section we connect all the results mentioned in this section and present the complete picture in our biological setting.

3. The biological interpretation. In this paper our main objective is to show that if we take the model for cell polarity given by Altschuler, Angenent, Wang and Wu [1] (see Description 1.1) and scale the parameters k_{fb} and k_{off} by the population size N, then, unlike the original model, we get cell polarity in the infinite population limit. In this section we describe how the results mentioned in the last section help us in making this conclusion. These results will also give us an insight into the influence of various biological parameters on polarity.

The main convergence result, Theorem 2.3, shows that as $N \to \infty$ the fraction of the molecules on the membrane at any time is equal to h_{eq} and the dynamics of cell molecules is given by a measure-valued process μ where $\mu = h_{\text{eq}}\nu$ with ν being a Fleming–Viot process. The process ν has a unique stationary distribution and its transition function converges exponentially to this stationary distribution (see Proposition 2.5).

At any time, the molecules on the membrane can be divided into clans based on their ancestral relationships. We now determine the distribution of the clan sizes. Let ν_c be the process given by (2.7). From Corollary 2.6 we know that it is a Fleming–Viot process with type space [0, 1] and generator $\mathbb{A}_c$. Such a Fleming–Viot process arises as a reformulation of the infinitely-many-neutral-alleles model due to Kimura and Crow [27] (see [11] for more details). By Theorem 7.2 in Ethier and Kurtz [12], at any time t, the random probability measure $\nu_c(t)$ is purely atomic. This means that $\nu_c(t)$ is of the form $\sum_{i=0}^{\infty} p_i \delta_{x_i}$, where p_i is the size of the atom corresponding to the point mass at x_i. At any $t \geq 0$ and any clan indicator $z \in [0, 1]$, the size of the clan at time t corresponding to z is just $\mu(t, E \times \{z\})$. The sum of all the clan sizes is quite clearly h_{eq}. If we normalize each clan size by dividing it by h_{eq}, then the normalized size of the clan at time t corresponding

to z is just $\nu(t, E \times \{z\}) = \nu_c(t, \{z\})$. In other words, the normalized clan sizes at time t are nothing but the sizes of the atoms in $\nu_c(t)$. From now on by *clan size* we always mean the *normalized clan size*.

The assertions of Proposition 2.5 are true for ν_c as well. Let Λ_∞ be the infinite simplex given by

$$\Lambda_\infty = \left\{(x_1, x_2, \ldots) : \sum_{i=1}^{\infty} x_i = 1 \text{ and } 0 < x_i < 1, i = 1, 2, \ldots\right\}.$$

GEM(θ) distribution is a distribution over the infinite simplex Λ_∞ that depends on a parameter $\theta \in [0, \infty)$. This distribution is named after three population geneticists McCloskey, Engen and Griffiths (see Johnson, Kotz and Balakrishnan [24] and Pitman and Yor [33]). It is defined as below.

DEFINITION 3.1 [GEM(θ) distribution]. Let $\{W_n : n = 1, 2, \ldots\}$ be a sequence of i.i.d. Beta$(1, \theta)$ random variables [i.e., each W_i has density $\theta(1-x)^{\theta-1}$ for $0 < x < 1$]. Define $P_1 = W_1$ and $P_n = (1 - W_1)(1 - W_2) \cdots (1 - W_{n-1})W_n$ for $n \geq 1$. Then the sequence $\{P_n : n = 1, 2, \ldots\}$ is said to have the GEM(θ) distribution.

If we define

$$\theta = \frac{k_{\text{on}}}{k_{\text{fb}}}, \tag{3.1}$$

then at stationarity the sizes of the atoms in $\nu_c(t)$ are distributed according to the GEM(θ) distribution. This is a direct consequence of Theorem 4.6 in Chapter 10, Ethier and Kurtz [11]. This result shows that at stationarity there are infinitely many clans on the membrane and their sizes follow the GEM(θ) distribution. If we arrange these sizes in descending order, then the resulting random infinite vector has the Poisson–Dirichlet distribution with the same parameter θ (see Chapter 2 in [16]). The Poisson–Dirichlet distribution was introduced by Kingman [28] in 1975 and many of its properties are well known. This characterization of clan sizes at stationarity makes it possible to compute the distribution and moments of the largest clan size, second largest clan size, third largest clan size and so on (see Griffiths [20]). The joint distribution of the first few largest clans can also be obtained (see Watterson [42]). If we sample n molecules from the membrane at stationarity from $\nu(t)$, then the distributional properties of the clans represented by this sample can be studied via the Ewen's Sampling Formula (see [14]). All these results indicate that the clan sizes at stationarity are far from uniform and there are a few *large* clans and many *small* clans. Most of the molecules are contained in these few large clans. In fact, if we sample n membrane molecules at stationarity, then they would belong to roughly $\theta \log n$ distinct clans asymptotically (see Theorem 2.11 in [16]).

The quantity S_p [given by (2.8)] measures the average spatial spread of the clans and its value at stationarity is given by Theorem 2.7. At stationarity there are only a few large clans and if S_p is small relative to the cell size, then the spatial spread of these large clans is also small. Therefore, the distribution of molecules at stationarity is highly asymmetrical at all times.

We now discuss the emergence of cell polarity. First we need to define it mathematically.

DEFINITION 3.2 (ε-polarity). For any $0 < \varepsilon \ll 1$, we say that the cell is ε-polarized if at least $(1 - \varepsilon)$ fraction of the membrane population belongs to a single clan and also resides in a single hemisphere on the membrane.

The above definition is motivated by the biological literature (see [1, 4, 5]). Note that the molecules in a clan will generally appear to cluster around the location of their most recent common ancestor (see [7]). Therefore, as in [1], if diffusion is small, having one predominant clan on the membrane is a good indication that a single site of polarity has formed.

At stationarity, the probability that the cell is ε-polarized at any time t can be expressed as

$$p_\varepsilon = \Pi\big(\{\beta \in \mathcal{P}(E \times [0,1]) \text{: there exists a } z \in [0,1] \text{ and a hemisphere } H \subset E \text{ such that } \beta(H \times \{z\}) \geq 1 - \varepsilon\}\big), \tag{3.2}$$

where $\Pi \in \mathcal{P}(\mathcal{P}(E \times [0,1]))$ is the stationary distribution of the process ν. We mentioned before that at stationarity, the vector of clan sizes in descending order follows the Poisson Dirichlet distribution with parameter θ. Let V_1 be the size of the largest clan. For any $\varepsilon > 0$, Theorem 2.5 in [16] implies that

$$q_\varepsilon := P\big(V_1 > \sqrt{1-\varepsilon}\big) > 0. \tag{3.3}$$

Suppose that we are at stationarity. Let r_ε denote the probability that $(\sqrt{1-\varepsilon})$-fraction of the molecules in the largest clan are situated in a single hemisphere on the membrane given that the size of the largest clan is at least $\sqrt{1-\varepsilon}$. Observe that S_p is like a weighted average of the spatial spreads of the clans, where the weight of each clan is proportional to its size. Therefore, if almost all the molecules are in the largest clan, then S_p is approximately the spatial spread of the largest clan. Hence, if S_p is small in comparison to the cell size, we can reasonably expect r_ε to be positive. Observe that $p_\varepsilon \geq q_\varepsilon r_\varepsilon$ and so p_ε is also positive for a small positive ε. Since the process ν is ergodic, Birkhoff's ergodic theorem (see Theorem 10.6 and Corollary 10.9 in [25]) implies that the cell will definitely reach the ε-polarized state and, in fact, spend p_ε proportion of its time there. Thus, the cell gets ε-polarized infinitely often and we have recurring cell polarity.

Before we proceed we need to define some new quantities. Let

$$\bar{S}_p = \frac{S_p}{R^2}, \tag{3.4}$$

$$\chi = \frac{D}{R^2} \tag{3.5}$$

and

$$\gamma = k_{\mathrm{fb}}\left(\frac{1-h_{\mathrm{eq}}}{h_{\mathrm{eq}}}\right) = \left(\frac{k_{\mathrm{fb}}k_{\mathrm{off}}}{k_{\mathrm{fb}}-k_{\mathrm{off}}}\right). \tag{3.6}$$

We can interpret $\bar{S}_p$ as the average spatial spread of the clans relative to the cell size, while χ can be seen as the speed of diffusion relative to the cell size. Note that the ratio $(1-h_{\mathrm{eq}})/h_{\mathrm{eq}}$ is the molecular mass available in the cytosol for recruitment per membrane molecule. The parameter γ is just the feedback rate tempered by this availability ratio. It can be interpreted as the effective feedback strength. We can recast the result of Theorem 2.7 as

$$\bar{S}_p = \frac{2\chi}{((1+\theta)\gamma+\chi)}. \tag{3.7}$$

Recall that the biological parameters in our model are D, R, k_{on}, k_{fb} and k_{off}. We now examine their impact on cell polarity. Instead of working with the original parameters, we will work with θ, χ and γ. From the above discussion it is clear that the formation of polarity will be facilitated if the probability q_ε [given by (3.3)] is high while the quantity $\bar{S}_p$ is low. As noted earlier, the parameter θ controls the distribution of molecules into the infinitely many clans present at stationarity. From the properties of Poisson Dirichlet distributions we know that the probability q_ε decreases as θ increases and vice-versa (see [6] and Chapter 2 in [16]). In fact, it can be shown that this probability is nearly 1 if $\theta \approx 0$ (see [15]). Hence, polarity will establish more easily if θ is small. Recall that the process ν_c [given by (2.7)] is the Fleming–Viot process corresponding to the infinitely-many-neutral-alleles model. The sample paths of this process take values over the space of atomic measures over $[0,1]$. Using Dirichlet forms, Schmuland [36] has shown that with probability 1 there will exist times at which this process will hit the state of having a single atom if and only if $\theta < 1$. Therefore, $\theta < 1$ assures that there will exist times when there is only one clan present. At these times the chances of observing polarity will be nearly 1 if $\bar{S}_p$ is sufficiently small. The formula (3.7) makes it clear that the quantity $\bar{S}_p$ gets smaller as the relative diffusion speed (χ) goes down or the effective feedback strength (γ) goes up.

Recall that the likelihood of finding a cell in the ε-polarized state at any time at stationarity is given by p_ε [given by (3.2)]. The observations made in the preceding paragraph show that p_ε increases with γ but decreases with θ and χ. Unfortunately, we do not have a precise formula for p_ε at the moment. Such a formula would be really useful in determining the chances of observing polarity in a cell

with any given set of parameters. It will also give us a clear idea of the time spent by the cell in the polarized state.

We would now like to compare our results to the results presented in [1] for the original model. To avoid confusion, we will denote the association, dissociation and recruitment rates in the original model as k^o_{on}, k^o_{off} and k^o_{fb}, respectively. Note that under our scaling $k^o_{\text{on}} = k_{\text{on}}$ while $k^o_{\text{off}} = Nk_{\text{off}}$ and $k^o_{\text{fb}} = Nk_{\text{fb}}$, where N is the total population size. The analysis in [1] assumes that $k^o_{\text{fb}} > k^o_{\text{off}}$ and k^o_{on} is much smaller in comparison to k^o_{off} or k^o_{fb}. Observe that spontaneous membrane association tends to spatially homogenize the molecules on the membrane and so if k^o_{on} is not small in comparison, we cannot hope to see cell polarity. Under the above assumptions it is shown in [1] that the fraction of molecules on the membrane approaches the equilibrium value

$$h_{\text{eq}} = 1 - \frac{k^o_{\text{off}}}{k^o_{\text{fb}}} + O\left(\frac{k^o_{\text{on}}}{k^o_{\text{fb}}}\right),$$

at an exponential rate with half-time of $(h_{\text{eq}}k^o_{\text{fb}})^{-1}$. In our scaling, $k^o_{\text{on}} = O(1)$, while k^o_{fb} and k^o_{off} are $O(N)$. Therefore, it is not surprising that as $N \to \infty$ the fraction of molecules on the membrane reaches the equilibrium value

$$h_{\text{eq}} = 1 - \frac{k_{\text{off}}}{k_{\text{fb}}}$$

almost instantly [see part (A) of Theorem 2.3]. Since k_{on} is small, the bulk of the population at equilibrium must come through membrane recruitment. It is mentioned in [1] that if $k^o_{\text{fb}} \le k^o_{\text{off}}$, the membrane will be nearly empty at equilibrium and so clusters cannot form. For the same reason Assumption 1.2 is required in this paper.

As we have discussed above, the emergence of cell polarity crucially depends on the likelihood of having just one large clan on the membrane. It is shown in [1] that for a finite population size N, the number of clans on the membrane will reduce to just 1 at certain times, giving rise to polarity (if D is small), if the spontaneous association events are rare (k^o_{on} is small). However, the frequency at which these times arrive is proportional to $1/N$ and, hence, there is no recurring polarity in the infinite population limit unlike the rescaled model that we consider.

It is observed in [1] that the clustering behavior for the original model is entirely determined by a simple relationship between the ratio $\frac{k^o_{\text{on}}}{k^o_{\text{fb}}}$ and the population size N. Their analysis shows that if $\frac{k^o_{\text{on}}}{k^o_{\text{fb}}} \ll N^{-2}$, then certainly one cluster will form and if $\frac{k^o_{\text{on}}}{k^o_{\text{fb}}} \gg (N^{-1}\log N)^{1/2}$, then no clusters will form. Using numerical simulations, they observe that the transition occurs when $\frac{k^o_{\text{on}}}{k^o_{\text{fb}}} \approx N^{-1}$. This motivated us to scale k^o_{fb} by a factor of N and analyze the resulting model. We also had to scale k^o_{off} by N because otherwise the entire population will soon be on the membrane (as h_{eq} will then be 1), depleting the cytosol and preventing further

membrane recruitment. In such a scenario, the feedback mechanism will be unable to counter the surface diffusion and, hence, there will not be any lasting cell polarity. Note that having $k_{\mathrm{fb}}^{o} = Nk_{\mathrm{fb}}$ is the same as changing the feedback rate in Description 1.1 to $k_{\mathrm{fb}} \times$ (*number of molecules in the cytosol*). This is the same as saying that each membrane molecule recruits each cytosol molecule at rate k_{fb}. Such a definition may be more natural for the feedback circuits found in certain cells. Our results provide an explanation for the existence of cell polarity in such cells if the population size is large.

There are many biologically appealing questions about the model that we have been unable to answer in this paper. As we mentioned above, it would be useful be have precise estimates for p_{ε}. It would also be interesting to compute the time it takes to hit the ε-polarized state and the time the cell stays polarized after that. These results would give us a better idea about the the onset and maintenance of polarity. The role of various model parameters will emerge clearly as well.

The model we study does have the drawback of being simplistic, as all the molecules in the cell are identical. Most cells that exhibit polarity have molecules of many different types recruiting each other at various type-dependent rates (see [3, 9, 40]). We would like to know if a multi-type generalization of our polarity model would also lead to tractable measure-valued dynamics in the infinite population limit. We hope to answer this question elsewhere in the very near future.

In this paper we have only looked at *single-site* polarity. Many cells exhibit *anterior–posterior* polarity (see [13, 32, 37]). In such cells there are usually two types of molecules and they segregate themselves in such a way that one type of molecule forms the *front* and the other type of molecule forms the *rear*. Such an arrangement is vital for cell division and locomotion. It has been suggested that this phenomenon is caused when molecules not only recruit the molecules of their own type but also locally inhibit the recruitment of the other type. It may be possible to extend the model considered here to account for anterior–posterior polarity as well.

The story of cell polarity is far from over and we hope that more work will be done to mathematically understand this biologically vital phenomenon and answer the challenging questions it poses.

4. Proofs. In this section we prove the main results of our paper: Theorems 2.3 and 2.7.

4.1. *Proof of part* (A) *of Theorem* 2.3. Recall that $n^N(t) = N\langle 1, \mu^N(t)\rangle$ denotes the number of molecules on the membrane at time t. Since μ^N has generator $\mathbb{A}^N$, we can write the generator K^N for the $\mathbb{N}_0$-valued process n^N as the following. For $f \in C(\mathbb{R})$ let

$$K^N f(n) = k_{\mathrm{on}}(N-n)\big(f(n+1) - f(n)\big) + Nk_{\mathrm{off}}n\big(f(n-1) - f(n)\big) + k_{\mathrm{fb}}n(N-n)\big(f(n+1) - f(n)\big).$$

We start with nothing on the membrane and, hence, $n^N(0) = 0$. The form of the generator K^N allows us to write the equation for n^N as

$$n^N(t) = Y_1\Big(k_{\text{on}} \int_0^t \big(N - n^N(s)\big)\, ds\Big) - Y_2\Big(N k_{\text{off}} \int_0^t n^N(s)\, ds\Big) + Y_3\Big(N k_{\text{fb}} \int_0^t n^N(s)\Big(1 - \frac{n^N(s)}{N}\Big)\, ds\Big). \tag{4.1}$$

We would like to estimate the first time n^N reaches a positive fraction of the population size N. Pick an $\varepsilon > 0$ such that $k_{\text{fb}}(1-\varepsilon) > k_{\text{off}}$ and define

$$\rho_\varepsilon^N = \inf\{t \geq 0 : n^N(t) \geq N\varepsilon\}. \tag{4.2}$$

LEMMA 4.1. *Let* $\lambda = k_{\text{fb}}(1-\varepsilon) - k_{\text{off}}$. *Then*

$$\lim_{N\to\infty} P\Big(\rho_\varepsilon^N \leq \frac{2\log N}{\lambda N}\Big) = 1.$$

Moreover, $\rho_\varepsilon^N \to 0$ *a.s. as* $N \to \infty$.

PROOF. We first slow the time by a factor of N. Let $\tilde{n}^N(t) = n^N(t/N)$. Since n^N satisfies equation (4.1), $\tilde{n}^N$ satisfies

$$\tilde{n}^N(t) = Y_1\Big(k_{\text{on}} \int_0^t \Big(1 - \frac{\tilde{n}^N(s)}{N}\Big)\, ds\Big) - Y_2\Big(k_{\text{off}} \int_0^t \tilde{n}^N(s)\, ds\Big) + Y_3\Big(k_{\text{fb}} \int_0^t \tilde{n}^N(s)\Big(1 - \frac{\tilde{n}^N(s)}{N}\Big)\, ds\Big). \tag{4.3}$$

Define

$$\tilde{\rho}_\varepsilon^N = \inf\{t \geq 0 : \tilde{n}^N(t) \geq N\varepsilon\} = N\rho_\varepsilon^N. \tag{4.4}$$

To prove the first claim of the lemma, we only need to show that

$$\lim_{N\to\infty} P\Big(\tilde{\rho}_\varepsilon^N \leq \frac{2\log N}{\lambda}\Big) = 1. \tag{4.5}$$

For $0 \leq t < \tilde{\rho}_\varepsilon^N$, $\frac{\tilde{n}^N(t)}{N} \leq \varepsilon$. Define another process Z by the equation

$$Z(t) = Y_1\big(k_{\text{on}}(1-\varepsilon)t\big) - Y_2\Big(k_{\text{off}} \int_0^t Z(s)\, ds\Big) + Y_3\Big(k_{\text{fb}}(1-\varepsilon) \int_0^t Z(s)\, ds\Big). \tag{4.6}$$

Note that Z is independent of N and ε is chosen so that $k_{\text{fb}}(1-\varepsilon) > k_{\text{off}}$. The form of the equation for Z shows that Z is a supercritical branching process with immigration. For $0 \leq t < \tilde{\rho}_\varepsilon^N$ we clearly have $Z(t) \leq \tilde{n}^N(t) < \varepsilon N$. Define

$$\bar{\rho}_\varepsilon^N = \inf\{t \geq 0 : Z(t) \geq N\varepsilon\}. \tag{4.7}$$

It is easy to see that $\tilde{\rho}_\varepsilon^N \le \bar{\rho}_\varepsilon^N$. We will find a probabilistic upper bound on $\bar{\rho}_\varepsilon^N$ which will show (4.5) and hence prove the first claim of the lemma.

A supercritical branching process with immigration can be written as a superposition of independent supercritical branching processes starting with an initial population of 1 at various times. This fact along with Theorems 1 and 2 in Chapter 3, Section 7, in Athreya and Ney [2] show that there exists a random variable W such that $W > 0$ a.s. and

$$\lim_{t\to\infty} e^{-\lambda t} Z(t) = W \qquad \text{a.s.}$$

Therefore,

$$\lim_{N\to\infty} e^{-\lambda \bar{\rho}_\varepsilon^N} Z(\bar{\rho}_\varepsilon^N) = W \qquad \text{a.s.},$$

which implies that

$$\lim_{N\to\infty} \log(e^{-\lambda \bar{\rho}_\varepsilon^N} Z(\bar{\rho}_\varepsilon^N)) = \lim_{N\to\infty} \bigl(-\lambda \bar{\rho}_\varepsilon^N + \log Z(\bar{\rho}_\varepsilon^N)\bigr) = \log W \qquad \text{a.s.}$$

Observe that $N\varepsilon \le Z(\bar{\rho}_\varepsilon^N) \le N\varepsilon + 1$. From above we get

$$\lim_{N\to\infty} \frac{\bar{\rho}_\varepsilon^N}{\log N} = \frac{1}{\lambda} \qquad \text{a.s.}$$

Since $N\rho_\varepsilon^N = \tilde{\rho}_\varepsilon^N \le \bar{\rho}_\varepsilon^N$ a.s., the above limit implies (4.5) and also shows that $\rho_\varepsilon^N \to 0$ a.s. as $N \to \infty$. This completes the proof of the lemma. □

Recall the definition of equilibrium fraction h_{eq} from the statement of Theorem 2.3. Fix ε to be $\frac{h_{\text{eq}}}{2} = \frac{1}{2}(1 - \frac{k_{\text{off}}}{k_{\text{fb}}})$ and let ρ^N be ρ_ε^N for this particular choice of ε. By Lemma 4.1 we obtain

$$\lim_{N\to\infty} P\left(\rho^N \le \frac{4\log N}{(k_{\text{fb}} - k_{\text{off}})N}\right) = 1. \tag{4.8}$$

Recall that the process h^N is given by (2.3). Using (4.1), we can write an equation for h^N as

$$\begin{aligned} h^N(t) = {} & \frac{1}{N} Y_1\left(Nk_{\text{on}} \int_0^t \bigl(1 - h^N(s)\bigr)\,ds\right) - \frac{1}{N} Y_2\left(N^2 k_{\text{off}} \int_0^t h^N(s)\,ds\right) \\ & + \frac{1}{N} Y_3\left(N^2 k_{\text{fb}} \int_0^t h^N(s)\bigl(1 - h^N(s)\bigr)\,ds\right). \end{aligned} \tag{4.9}$$

Let $\bar{h}^N$ be the process given by

$$\bar{h}^N(t) = h^N(t + \rho^N). \tag{4.10}$$

Note that

$$\bar{h}^N(0) = h^N(\rho^N) = \frac{\lceil Nh_{\text{eq}}/2 \rceil}{N}. \tag{4.11}$$

For $i = 1, 2, 3$ let

$$\bar{Y}_i(t) = Y_i(t + \delta_i^N) - Y_i(\delta_i^N),$$

where

$$\delta_1^N = Nk_{\text{on}} \int_0^{\rho^N} \bigl(1 - h^N(s)\bigr)\, ds,$$

$$\delta_2^N = N^2 k_{\text{off}} \int_0^{\rho^N} h^N(s)\, ds$$

and

$$\delta_3^N = N^2 k_{\text{fb}} \int_0^{\rho^N} h^N(s)\bigl(1 - h^N(s)\bigr)\, ds.$$

From the strong Markov property of the Poisson process we can conclude that $\bar{Y}_1$, $\bar{Y}_2$ and $\bar{Y}_3$ are independent unit Poisson processes. We can write the equation for process $\bar{h}^N$ as

$$\begin{aligned}
\bar{h}^N(t) &= \bar{h}^N(0) + h^N(t + \rho^N) - h^N(\rho^N) \\
&= \bar{h}^N(0) + \frac{1}{N}\bar{Y}_1\biggl(k_{\text{on}} N \int_0^t \bigl(1 - \bar{h}^N(s)\bigr)\, ds\biggr) \\
&\quad - \frac{1}{N}\bar{Y}_2\biggl(N^2 k_{\text{off}} \int_0^t \bar{h}^N(s)\, ds\biggr) \\
&\quad + \frac{1}{N}\bar{Y}_3\biggl(N^2 k_{\text{fb}} \int_0^t \bar{h}^N(s)\bigl(1 - \bar{h}^N(s)\bigr)\, ds\biggr).
\end{aligned} \tag{4.12}$$

Let $\bar{Y}_i^c$ be the centered version of $\bar{Y}_i$ [i.e., $\bar{Y}_i^c(u) = \bar{Y}_i(u) - u$ for $u \geq 0$]. Define

$$\begin{aligned}
M_N(t) &= \frac{1}{N}\bar{Y}_1^c\biggl(k_{\text{on}} N \int_0^t \bigl(1 - \bar{h}^N(s)\bigr)\, ds\biggr) - \frac{1}{N}\bar{Y}_2^c\biggl(N^2 k_{\text{off}} \int_0^t \bar{h}^N(s)\, ds\biggr) \\
&\quad + \frac{1}{N}\bar{Y}_3^c\biggl(N^2 k_{\text{fb}} \int_0^t \bar{h}^N(s)\bigl(1 - \bar{h}^N(s)\bigr)\, ds\biggr),
\end{aligned} \tag{4.13}$$

which is a martingale with quadratic variation given by

$$\begin{aligned}
[M_N]_t &= \frac{1}{N^2}\bar{Y}_1\biggl(k_{\text{on}} N \int_0^t \bigl(1 - \bar{h}^N(s)\bigr)\, ds\biggr) + \frac{1}{N^2}\bar{Y}_2\biggl(N^2 k_{\text{off}} \int_0^t \bar{h}^N(s)\, ds\biggr) \\
&\quad + \frac{1}{N^2}\bar{Y}_3\biggl(N^2 k_{\text{fb}} \int_0^t \bar{h}^N(s)\bigl(1 - \bar{h}^N(s)\bigr)\, ds\biggr).
\end{aligned} \tag{4.14}$$

Since $0 \leq \bar{h}^N \leq 1$, we have

$$E([M_N]_t) \leq k_{\text{on}} \frac{t}{N} + k_{\text{off}} t + k_{\text{fb}} t. \tag{4.15}$$

By centering the Poissons in equation (4.12), we can write

$$\begin{aligned}\bar{h}^N(t) &= \bar{h}^N(0) + k_{\text{on}} \int_0^t \bigl(1 - \bar{h}^N(s)\bigr)\,ds - N k_{\text{off}} \int_0^t \bar{h}^N(s)\,ds \\ &\quad + N k_{\text{fb}} \int_0^t \bar{h}^N(s)\bigl(1 - \bar{h}^N(s)\bigr)\,ds + M_N(t).\end{aligned} \tag{4.16}$$

Let $F(h) = k_{\text{fb}} h(1-h) - k_{\text{off}} h$ and

$$Z_N(t) = \int_0^t k_{\text{on}}\bigl(1 - \bar{h}^N(s)\bigr)\,ds + M_N(t).$$

From (4.15) and Corollary 2.3.3 in [23] we can conclude that $\{Z_N\}$ is a sequence of semimartingales that is relatively compact in $D_{\mathbb{R}}[0,\infty)$. The jumps in Z_N are of size $1/N$ and, hence, if Z is a limit point of this sequence, then Z must be a continuous process a.s. equation (4.16) can be written as

$$\bar{h}^N(t) = \bar{h}^N(0) + Z_N(t) + N \int_0^t F(\bar{h}^N(s))\,ds. \tag{4.17}$$

Define another process α^N by

$$\alpha^N(t) = \bar{h}^N(t) - h_{\text{eq}}.$$

Observe that

$$\begin{aligned}F(\bar{h}^N(t)) &= k_{\text{fb}}\bigl(\alpha^N(t) + h_{\text{eq}}\bigr)\bigl(1 - h_{\text{eq}} - \alpha^N(t)\bigr) - k_{\text{off}}\bigl(\alpha^N(t) + h_{\text{eq}}\bigr) \\ &= \alpha^N(t)\bigl(k_{\text{fb}}(1 - h_{\text{eq}}) - k_{\text{off}}\bigr) + k_{\text{fb}} h_{\text{eq}}(1 - h_{\text{eq}}) \\ &\quad - k_{\text{off}} h_{\text{eq}} - k_{\text{fb}} \alpha^N(t)\bigl(\alpha^N(t) + h_{\text{eq}}\bigr) \\ &= -k_{\text{fb}} \alpha^N(t) \bar{h}^N(t) \qquad \left(\text{using } h_{\text{eq}} = 1 - \frac{k_{\text{off}}}{k_{\text{fb}}}\right).\end{aligned}$$

From (4.17) we get

$$\alpha^N(t) = \alpha^N(0) - N k_{\text{fb}} \int_0^t \bar{h}^N(s)\alpha^N(s)\,ds + Z_N(t),$$

which can be written in differential form as

$$d\alpha^N(t) + N k_{\text{fb}} \bar{h}^N(t)\alpha^N(t)\,dt = dZ_N(t).$$

Let $\beta_N(t) = N k_{\text{fb}} \int_0^t \bar{h}^N(s)\,ds$. Then

$$d\bigl(\alpha^N(t) e^{\beta_N(t)}\bigr) = e^{\beta_N(t)}\,dZ_N(t).$$

Integrating from 0 to t, we get

$$\alpha^N(t) e^{\beta_N(t)} - \alpha^N(0) = \int_0^t e^{\beta_N(s)}\,dZ_N(s).$$

Therefore,

$$\alpha^N(t) - \alpha^N(0)e^{-\beta_N(t)} = e^{-\beta_N(t)} \int_0^t e^{\beta_N(s)}\, dZ_N(s). \tag{4.18}$$

Let

$$\bar{\sigma}^N = \inf\left\{ t \geq 0 : \bar{h}^N(t) \leq \frac{h_{\mathrm{eq}}}{4} \right\}.$$

For $0 \leq t \leq \bar{\sigma}^N$

$$\bar{h}^N(t) \geq \frac{h_{\mathrm{eq}}}{4} - \frac{1}{N}$$

and, hence, for any small positive number ε,

$$\inf_{0 \leq t \leq (T \wedge \bar{\sigma}^N - \varepsilon)} |\beta_N(t+\varepsilon) - \beta_N(t)| \Rightarrow \infty$$

as $N \to \infty$. Fix any $T > 0$. By Lemma 5.2 in [26],

$$\sup_{0 \leq t \leq T \wedge \bar{\sigma}^N} e^{-\beta_N(t)} \left| \int_0^t e^{\beta_N(s)}\, dZ_N(s) \right| \Rightarrow 0.$$

Hence, (4.18) gives

$$\sup_{0 \leq t \leq T \wedge \bar{\sigma}^N} |\alpha^N(t) - \alpha^N(0)e^{-\beta_N(t)}| \Rightarrow 0. \tag{4.19}$$

If $\bar{\sigma}^N < T$, then from the definitions of $\bar{\sigma}^N$, α^N and (4.11) we get

$$\begin{aligned} \sup_{0 \leq t \leq T \wedge \bar{\sigma}^N} |\alpha^N(t) - \alpha^N(0)e^{-\beta_N(t)}| &\geq |\alpha^N(\bar{\sigma}^N) - \alpha^N(0)e^{-\beta_N(\bar{\sigma}^N)}| \\ &\geq \left| -\frac{3h_{\mathrm{eq}}}{4} + \frac{h_{\mathrm{eq}}}{2} e^{-\beta_N(\bar{\sigma}^N)} \right| - \frac{2}{N} \\ &\geq \frac{h_{\mathrm{eq}}}{4} - \frac{2}{N}. \end{aligned}$$

This calculation shows that

$$P(\bar{\sigma}^N < T) \leq P\left(\sup_{0 \leq t \leq T \wedge \bar{\sigma}^N} |\alpha^N(t) - \alpha^N(0)e^{-\beta_N(t)}| \geq \frac{h_{\mathrm{eq}}}{4} - \frac{2}{N} \right).$$

Therefore, from (4.19), $P(\bar{\sigma}^N < T) \to 0$ for any $T > 0$ and, hence, $\bar{\sigma}^N \to \infty$ in probability.

Let the stopping time τ_N be given by

$$\tau_N = \rho^N + \frac{\log N}{N}. \tag{4.20}$$

From Lemma 4.1 $\rho^N \to 0$ a.s. and, hence, $\tau_N \to 0$ a.s. Let the process $\hat{h}^N$ be defined by

$$\hat{h}^N(t) = h^N(t+\tau_N) = \bar{h}^N\left(t + \frac{\log N}{N}\right), \qquad t \geq 0. \tag{4.21}$$

Note that

$$\begin{aligned}
&\sup_{0\leq t\leq((T\wedge\bar{\sigma}^N)-(\log N)/N)} |\hat{h}^N(t) - h_{\mathrm{eq}}| \\
&\quad = \sup_{0\leq t\leq((T\wedge\bar{\sigma}^N)-(\log N)/N)} \left|\alpha^N\left(t + \frac{\log N}{N}\right)\right| \\
&\quad \leq \sup_{0\leq t\leq((T\wedge\bar{\sigma}^N)-(\log N)/N)} \left|\alpha^N\left(t + \frac{\log N}{N}\right) - \alpha^N(0)e^{-\beta_N(t+(\log N)/N)}\right| \\
&\qquad + \sup_{0\leq t\leq((T\wedge\bar{\sigma}^N)-(\log N)/N)} |\alpha^N(0)e^{-\beta_N(t+(\log N)/N)}|.
\end{aligned}$$

The first term converges to 0 in probability due to (4.19). Observe that for $0 \leq t \leq ((T\wedge\bar{\sigma}^N) - \frac{\log N}{N})$,

$$\beta_N\left(t + \frac{\log N}{N}\right) \geq Nk_{\mathrm{fb}} \int_0^{t+(\log N)/N} \bar{h}^N(s)\,ds = k_{\mathrm{fb}}\left(\frac{h_{\mathrm{eq}}}{4} - \frac{1}{N}\right)\log N.$$

Thus, the second term above converges to 0 as $N \to \infty$ a.s. Since $\log N/N \to 0$ and $\bar{\sigma}^N \to \infty$, we get

$$\sup_{0\leq t\leq T} |\hat{h}^N(t) - h_{\mathrm{eq}}| \Rightarrow 0.$$

This proves part (A) of Theorem 2.3.

4.2. *Proof of part* (B) *of Theorem* 2.3. To prove part (B) of the theorem, we will use the particle construction introduced by Donnelly and Kurtz in [8]. In this construction the molecules are arranged in levels which are indexed by positive integers. The arrangement is such that for any positive integer n, the process determined by the first n levels is embedded in the process determined by the first $(n+1)$ levels. This allows us to pass to the projective limit. Another advantage of this construction is that it makes the ancestral relationships between molecules explicit. For any set of molecules we can trace back their genealogical tree to obtain results about the measure-valued process.

We first motivate the particle construction. Suppose the total population size is N and at any time t there are $n^N(t)$ molecules on the membrane. The process n^N follows equation (4.1) and suppose its evolution is known. Each molecule has a type in $E \times [0,1]$ as before. We can represent the population on the membrane at time t by a vector $(Y_1^N(t), Y_2^N(t), \ldots, Y_{n^N(t)}^N)$. Since the labeling of the

molecules is arbitrary, it contains exactly the same information as the measure $\tilde{Z}(t) = \sum_{i=1}^{n} \delta_{Y_i(t)}$. We can choose any labeling we find convenient. So we look into the future and order the individuals according to the time of survival of their lines of descent. In this new ordering we arrange the molecules into *levels*, which are taken to be positive integers. At any time t, if there are $n^N(t)$ molecules, we will represent the population as the vector $(X_1^N(t), X_2^N(t), \ldots, X_{n^N(t)}^N)$. We will refer to X_i^N as the ith level process, where $X_i^N(t) \in E \times [0,1]$ is the molecule type at level i at time t. Molecules are allowed to change levels with time. If a death happens at time t, then $n^N(t) = n^N(t-) - 1$ and we just remove the molecule at the highest index $n^N(t)$. If an immigration happens at time t, then $n^N(t) = n^N(t-)+1$ and we uniformly select a level from the first $n^N(t-)+1$ levels and insert the immigrant molecule there. If a birth event happens at time t, then $n^N(t) = n^N(t-)+1$ and we do the following. We first uniformly select two levels i and j from the first $n^N(t-)+1$ levels. Suppose i is the smaller of the two levels. Then we shall refer to the molecule $X_i^N(t-)$ as the parent and insert a copy of it at level j. The molecules $\{X_k^N(t-) : k = j, j+1, \ldots\}$ are shifted up by one level. So at time t, the offspring molecule $X_j^N(t)$ is a copy of $X_i^N(t-)$, while $X_k^N(t) = X_k^N(t-)$ for $k < j$ and $X_k^N(t) = X_{k-1}^N(t-)$ for $k > j$. In between all these events molecules are doing speed D Brownian motion on E and changing their location.

What we have described above is a Markov process X^N with state space

$$S^N = \bigcup_{n=0}^{N} (E \times [0,1])^n.$$

We adopt the convention that $(E \times [0,1])^0 = \{\triangle\}$. For $x \in S^N$, if $x \in (E \times [0,1])^n$, then let $|x| = n$ for any $n = 0, 1, \ldots, N$. If at time t, $X^N(t) = x \in S^N$ and $|x| = n$, then it means that there are n molecules on the membrane with the type vector $x = (x_1, x_2, \ldots, x_n) \in (E \times [0,1])^n$.

If $|x| = n \geq m$ and $x = (x_1, x_2, \ldots, x_n)$, then let $x^{|m} = (x_1, x_2, \ldots, x_m)$. Any $f \in B((E \times [0,1])^m)$ can be regarded as a function over S^N by defining it as $f(x) = 0$ if $|x| < m$ and $f(x) = f(x^{|m})$ if $|x| \geq m$. We now specify the generator A^N of the Markov process X^N by its action on functions in its domain $\mathcal{D}(A^N) = \mathcal{C}$ (see Definition 2.1) as

$$\begin{aligned} A^N f(x) = {} & \frac{D}{2} \sum_{i=1}^{n} \Delta_i f(x) + nNk_{\text{off}}\big(f(d_n(x)) - f(x)\big) \\ & + k_{\text{on}}\left(\frac{N-n}{n+1}\right) \sum_{i=1}^{n+1} \int_E \int_0^1 \big(f(\theta_i(x,(y,r))) - f(x)\big)\vartheta(dy)\,dr \\ & + 2k_{\text{fb}}\left(\frac{N-n}{n+1}\right) \sum_{1 \leq i < j \leq (n+1)} \big(f(\theta_{ij}(x)) - f(x)\big), \end{aligned} \tag{4.22}$$

where $n = |x|$ and if $x = (x_1, x_2, \ldots, x_n)$, then $d_i(x) = (x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$ (remove the ith coordinate), $\theta_{ij}(x) = (x_1, \ldots, x_{j-1}, x_i, x_j, \ldots, x_n)$ (insert a copy of x_i at the jth place) and $\theta_i(x, (y, r)) = (x_1, \ldots, x_{i-1}, (y, r), x_i, \ldots, x_n)$ [insert (y, r) at the ith place].

Viewing the operator A^N as a bounded perturbation of the diffusion operator [given by the first term on the right of (4.22)], we can argue that the martingale problem for A^N is well posed in the same way we argued for $\mathbb{A}^N$ in Section 2. We now relate any solution of the martingale problem for A^N to a solution of the martingale problem for $\mathbb{A}^N$ [see (2.2)] by using the Markov mapping theorem (see Theorem 2.7 in Kurtz [29]). Let

$$S_0^N = \mathcal{M}_a^N(E \times [0,1]) = \left\{ \frac{1}{N} \sum_{i=1}^{n} \delta_{x_i} : 0 \le n \le N \text{ and } x_1, \ldots, x_n \in E \times [0,1] \right\}$$

and

$$S^N = \bigcup_{n=0}^{N} (E \times [0,1])^n$$

as before. Define $\gamma : S^N \to S_0^N$ by

$$\gamma(x) = \frac{1}{N} \sum_{i=1}^{n} \delta_{x_i} \qquad \text{if } x = (x_1, x_2, \ldots, x_n).$$

Define the transition function $\alpha : S_0^N \to \mathcal{P}(S^N)$ by

$$\alpha\left(\frac{1}{N} \sum_{i=1}^{n} \delta_{x_i}, dz \right) = \frac{1}{n!} \sum_{\sigma \in \Sigma_n} \delta_{(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(n)})}\, dz,$$

where Σ_n is the set of all permutations on $\{1, 2, \ldots, n\}$.

LEMMA 4.2. *Let $\pi_0^N \in \mathcal{P}(S_0^N)$ and define $\pi^N = \int_{S_0^N} \alpha(y, \cdot)\pi_0^N(dy)$. If ν^N is the solution of the martingale problem for $(\mathbb{A}^N, \pi_0^N)$ and X^N is the solution of the martingale problem for (A^N, π^N), then $\gamma(X^N)$ and ν^N have the same distribution in $D_{\mathcal{M}_a^N(E\times[0,1])}[0, \infty)$. Furthermore, for any $t \ge 0$ the distribution of $X^N(t) = (X_1^N(t), X_2^N(t), \ldots)$ is exchangeable.*

REMARK 4.3. The length of the vector $X^N(t)$ is $n^N(t)$, which is a random variable. When we say that the distribution of $X^N(t) = (X_1^N(t), X_2^N(t), \ldots)$ is exchangeable we mean that given $n^N(t) = n$, the distribution of $(X_1^N(t), X_2^N(t), \ldots, X_n^N(t))$ is exchangeable.

PROOF OF LEMMA 4.2. The definition of α ensures that $\alpha(\mu, \gamma^{-1}(\mu)) = 1$ for all $\mu \in S_0^N$. If $f \in \mathcal{C} \cap C((E \times [0,1])^m)$ and $\mu = \frac{1}{N}\sum_{i=1}^n \delta_{x_i}$, then let

$$F(\mu) = \int_{S^N} f(z)\alpha(\mu, dz) = \frac{1}{n!}\sum_{\sigma\in\Sigma_n} f\big(x_{\sigma(1)}, \ldots, x_{\sigma(n)}\big) = \langle f, \mu^{(m)}\rangle. \tag{4.23}$$

Hence, $F \in \bar{\mathcal{C}} = \mathcal{D}(\mathbb{A})$. Now we show that for such a function F

$$\mathbb{A}^N F(\cdot) = \int_{S^N} A^N f(z)\alpha(\cdot, dz). \tag{4.24}$$

On writing down the expressions for $\mathbb{A}^N$ and A^N using (2.2) and (4.22), we observe that there are four terms on each side of (4.24). We will show that the equality holds term by term. It is easy to see that the first term corresponding to the Brownian diffusion of membrane molecules is equal on both sides. We check the equality for the next three terms below.

For $x = (x_1, \ldots, x_n)$ let $\sigma(x) = (x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(n)})$ for any $\sigma \in \Sigma_n$. Let $\mu = \frac{1}{N}\sum_{i=1}^n \delta_{x_i}$. Then

$$\begin{aligned} F\Big(\mu + \frac{1}{N}\delta_{(y,r)}\Big) &= \frac{1}{(n+1)!}\sum_{\sigma\in\Sigma_{n+1}} f(\sigma(\theta_{n+1}(x, (y, r)))) \\ &= \frac{1}{n+1}\sum_{i=1}^{n+1}\frac{1}{n!}\sum_{\sigma\in\Sigma_n} f(\theta_i(\sigma(x), (y, r))). \end{aligned} \tag{4.25}$$

Similarly,

$$\begin{aligned} N\int_{E\times[0,1]} F\Big(\mu + \frac{1}{N}\delta_x\Big)\mu(dx) &= \sum_{i=1}^n \frac{1}{(n+1)!}\sum_{\sigma\in\Sigma_{(n+1)}} f(\sigma(\theta_{in^+}(x))) \\ &= \frac{1}{(n+1)!}\sum_{i=1}^n\sum_{j\neq i}\sum_{\sigma\in\Sigma_n} f(\theta_{ij}(\sigma(x))) \\ &= \frac{2}{n+1}\sum_{1\le i<j\le(n+1)}\frac{1}{n!}\sum_{\sigma\in\Sigma_n} f(\theta_{ij}(\sigma(x))), \end{aligned} \tag{4.26}$$

where $n^+ = (n+1)$ in the first equation above. Finally,

$$\begin{aligned} N\int_{E\times[0,1]} F\Big(\mu - \frac{1}{N}\delta_x\Big)\mu(dx) &= \sum_{i=1}^n \frac{1}{(n-1)!}\sum_{\sigma\in\Sigma_{n-1}} f(\sigma(d_i(x))) \\ &= \frac{1}{(n-1)!}\sum_{\sigma\in\Sigma_n} f(\sigma(d_n(x))) \\ &= n\frac{1}{n!}\sum_{\sigma\in\Sigma_n} f(\sigma(d_n(x))). \end{aligned} \tag{4.27}$$

Equations (4.25), (4.26) and (4.27) show that the relation (4.24) holds and so the Markov mapping theorem is applicable. Therefore, we can conclude that $\gamma(X^N)$ and ν^N have the same distribution in $D_{\mathcal{M}_a^N(E\times[0,1])}[0,\infty)$. From Corollary 3.5 in Kurtz [29] we obtain that if $n^N(t)=n$, then

$$E(f(X_1^N(t),X_2^N(t),\ldots,X_n^N(t))|\mathcal{F}_t)=\int_{S^N} f(z)\alpha(\gamma(X^N(t)),dz),$$

where $\{\mathcal{F}_t\}$ is the filtration generated by the process $\gamma(X^N(\cdot))$. Since α is symmetric, the distribution of $(X_1^N(t),X_2^N(t),\ldots)$ is exchangeable. $\square$

Recall from Section 2 that $\bar{\pi}_0\in\mathcal{P}(\mathcal{M}_a^N(E\times[0,1]))$ is the distribution that puts all the mass at the 0 measure and μ^N [given by (2.1)] is the unique solution to the martingale problem corresponding to $(\mathbb{A}^N,\bar{\pi}_0)$. For τ_N given by (4.20), define the process $\hat{\mu}^N$ by

$$\hat{\mu}^N(t)=\mu^N(t+\tau_N),\qquad t\geq 0. \tag{4.28}$$

Also let

$$\hat{n}^N(t)=N\hat{h}^N(t)=n^N(t+\tau_N),\qquad t\geq 0. \tag{4.29}$$

Let $\hat{\pi}_0^N\in\mathcal{P}(\mathcal{M}_a^N(E\times[0,1]))$ be the distribution of $\mu^N(\tau_N)=\hat{\mu}^N(0)$ and define $\pi^N\in\mathcal{P}(S^N)$ by $\pi^N=\int_{S_0^N}\alpha(y,\cdot)\hat{\pi}_0^N(dy)$. Let X^N be the unique solution to the martingale problem for (A^N,π^N). Note that X^N lives in the space $S^N=\bigcup_{n=0}^N(E\times[0,1])^n$ and for any $t\geq 0$, $|X^N(t)|=\hat{n}^N(t)$. The process $\hat{h}^N$ converges to h_{eq} uniformly over compact time intervals [from part (A) of Theorem 2.3]. Hence, $\hat{n}^N$ converges to ∞ uniformly over compact time intervals as well.

For part (B) of Theorem 2.3 we assume that the sequence of random variables $\{\hat{\mu}^N(0)\}$ converges in distribution to $\mu(0)$ as $N\to\infty$. Let $\hat{\pi}_0\in\mathcal{P}(\mathcal{M}_1(E\times[0,1]))$ be the distribution of $\mu(0)$ and $\pi_0\in\mathcal{P}(\mathcal{P}(E\times[0,1]))$ be the distribution of $\mu(0)/h_{\mathrm{eq}}$. Our assumption implies that $\hat{\pi}_0^N$ converges weakly to $\hat{\pi}_0$. Due to part (A) of Theorem 2.3, this is equivalent to saying that the distributions of $\mu^N(\tau_N)/h^N(\tau_N)$ converge weakly to π_0.

Now sample a probability measure μ from π_0 and let $(Y_1,Y_2,\ldots)$ be an infinite sequence of exchangeable random variables with de Finetti measure μ. Let $\pi\in\mathcal{P}((E\times[0,1])^\infty)$ be the corresponding distribution of $(Y_1,Y_2,\ldots)$. Since $\hat{\pi}_0^N\Rightarrow\hat{\pi}_0$, we also have $\pi^N\Rightarrow\pi$.

From now on consider X^N as a process over $(E\times[0,1])^\infty$ in which the components greater than N do not vary. The space $(E\times[0,1])^\infty$ is given the usual product topology.

We can regard any function $f \in \mathcal{C} \cap B((E \times [0,1])^m)$ as a function over $(E \times [0,1])^\infty$ by defining it as $f(x) = f(x^{|m}) = f(x_1, \ldots, x_m)$ for any $x \in (E \times [0,1])^\infty$. By the definition of X^N, for any $f \in \mathcal{C} \cap B((E \times [0,1])^m)$,

$$M^N_{X,f}(t) = f(X^N(t)) - \int_0^t A^N f(X^N(s))\, ds \tag{4.30}$$

is a martingale. Define another process

$$\hat{Z}_N(t) = \gamma(X^N(t)) = \frac{1}{N} \sum_{i=1}^{\hat{n}^N(t)} \delta_{X_i^N(t)}, \qquad t \geq 0. \tag{4.31}$$

From Lemma 4.2, the process $\hat{Z}_N$ has the same distribution as the process $\hat{\mu}^N$. Hence, if $F \in \bar{\mathcal{C}}$ is given by $F(\mu) = \langle f, \mu^{(m)} \rangle$, then

$$M^N_{\hat{Z},F} = F(\hat{Z}^N(t)) - \int_0^t \mathbb{A}^N F(\hat{Z}^N(s))\, ds \tag{4.32}$$

is also a martingale. If $|X^N(t)| = \hat{n}^N(t) > m$, then

$$\begin{aligned} A^N f(X^N(t)) &= \frac{D}{2} \sum_{i=1}^m \Delta_i f(x) \\ &\quad + 2k_{\mathrm{fb}} \left(\frac{N - \hat{n}^N(t)}{\hat{n}^N(t) + 1} \right) \sum_{1 \leq i < j \leq m} \big(f(\theta_{ij}(X^N(t))) - f(x) \big) \\ &\quad + k_{\mathrm{on}} \left(\frac{N - \hat{n}^N(t)}{\hat{n}^N(t) + 1} \right) \sum_{i=1}^m \int_E \int_0^1 \big(f(\theta_i(X^N(t), (y,r))) - f(x) \big) \vartheta(dy)\, dr. \end{aligned}$$

The *death* term drops out because only the molecule at the highest level is allowed to die and f depends on only the first m levels. From above it can be easily seen that for any positive integer m and $f \in \mathcal{C} \cap B((E \times [0,1])^m)$, the supremum of the process $A^N f(X^N(\cdot))$ over compact time intervals stays bounded as $N \to \infty$. Using this fact along with (4.30), (4.32) and (4.24), it is easy to argue that the sequence of processes $\{(X^N, \hat{Z}_N)\}$ is relatively compact in $D_{(E \times [0,1])^\infty \times \mathcal{M}_1(E \times [0,1])}[0, \infty)$ (see Corollary 9.3 and Theorem 9.4 in Ethier and Kurtz [11]). Suppose $(X, \hat{Z})$ is any limit point and $(X^N, \hat{Z}_N) \Rightarrow (X, \hat{Z})$ along the subsequence k_N. By the continuous mapping theorem and the boundedness of $f \in \mathcal{C} \cap B((E \times [0,1])^m)$, the sequence of martingales $\{M^N_{X,f}(t)\}$ converges along the subsequence k_N to

$$M_{X,f}(t) = f(X(t)) - \int_0^t A_m f(X(s))\, ds, \tag{4.33}$$

which is a martingale with respect to the filtration generated by X. The operator A_m is given by

$$
\begin{aligned}
A_m f(x) = {} & \frac{D}{2}\sum_{i=1}^{m}\Delta_i f(x) + 2k_{\mathrm{fb}}\biggl(\frac{1-h_{\mathrm{eq}}}{h_{\mathrm{eq}}}\biggr)\sum_{1\le i<j\le m}\bigl(f(\theta_{ij}(x)) - f(x)\bigr) \\
& + k_{\mathrm{on}}\biggl(\frac{1-h_{\mathrm{eq}}}{h_{\mathrm{eq}}}\biggr)\sum_{i=1}^{m}\int_E\int_0^1 \bigl(f(\theta_i(x,(y,r))) - f(x)\bigr)\vartheta(dy)\,dr
\end{aligned}
\tag{4.34}
$$

for any $f \in \mathcal{D}(A_m) = \mathcal{C} \cap B((E \times [0,1])^m)$. The operator A_m is the generator for the process determined by the first m levels of the limiting process X. We can easily check that the martingale problem for A_m is well posed due to the same reasons that were given for A^N. Taking $\mathcal{D}(A) = \bigcup_{m=1}^{\infty}\mathcal{D}(A_m)$ and defining $Af = A_m f$ if $f \in \mathcal{D}(A_m)$, we see that the martingale problem for A is well posed. The distribution of $X^N(0)$ (denoted by π^N) converges to π. From (4.33) we know that for any positive integer m, the process followed by the first m levels of X has generator A_m. Hence, X is the unique solution to the martingale problem corresponding to (A, π).

Let

$$\gamma_N = \inf\{t \ge 0 : \hat{n}^N(t) = 0\}$$

and for any $0 \le t < \gamma_N$ define

$$Z_N(t) = \frac{1}{\hat{n}^N(t)}\sum_{k=1}^{\hat{n}^N(t)}\delta_{X_k^N(t)}. \tag{4.35}$$

Observe that

$$\hat{Z}_N(t) = \frac{1}{N}\sum_{k=1}^{\hat{n}^N(t)}\delta_{X_k^N(t)} = \biggl(\frac{\hat{n}^N(t)}{N}\biggr)\biggl(\frac{1}{\hat{n}^N(t)}\sum_{k=1}^{\hat{n}^N(t)}\delta_{X_k^N(t)}\biggr) = \hat{h}^N(t)Z_N(t).$$

The process $\hat{h}^N$ converges to the constant process h_{eq}. Therefore, $\gamma_N \to \infty$ in probability and for any $t \ge 0$, $\hat{n}^N(t) \to \infty$ in probability. Define Z to be the process

$$Z(t) = \frac{\hat{Z}(t)}{h_{\mathrm{eq}}}, \qquad t \ge 0.$$

Then $Z(0)$ has distribution π_0 and since $(X^N, \hat{Z}_N) \Rightarrow (X, \hat{Z})$ along the subsequence k_N, we must have that $(X^N, Z_N) \Rightarrow (X, Z)$ along the same subsequence. Notice that for any t, $Z_N(t) \Rightarrow Z(t)$ implies that $Z_N^{(m)}(t) \Rightarrow Z^m(t)$. From the exchangeability of $X^N(t)$ we get that for any $f \in \mathcal{C} \cap B((E \times [0,1])^m)$,

$$E(f(X_1^N(t), \ldots, X_m^N(t))) = E\bigl(\langle f, \hat{Z}_N^{(m)}(t)\rangle\bigr) = E\bigl(\langle f, Z_N^{(m)}(t)\rangle\bigr)$$

for any $0 \le t < \gamma_N$. Passing to the limit along the subsequence k_N, we obtain

$$E(f(X_1(t), \dots, X_m(t))) = E(\langle f, Z^m(t)\rangle)$$

for any $t \ge 0$. It shows that conditional on $Z(t)$, $X_1(t)$, $X_2(t)$, ... are i.i.d. random variables with distribution $Z(t)$. Hence, for any $t \ge 0$, $X(t)$ is exchangeable with de Finetti measure $Z(t)$. Therefore,

$$Z(t) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \delta_{X_i(t)} \qquad \text{a.s.} \tag{4.36}$$

and if $\{\mathcal{F}_t^Z\}$ is the filtration generated by the process Z, then

$$E(f(X_1(t), \dots, X_m(t))|\mathcal{F}_t^Z) = \langle f, Z^m(t)\rangle.$$

Conditioning (4.33) with respect to $\{\mathcal{F}_t^Z\}$, we obtain that

$$\langle f, Z^m(t)\rangle - \int_0^t \langle A_m f, Z^m(s)\rangle \, ds$$

is a martingale. If we define an operator $\mathbb{A}$ by

$$\mathbb{A}F(\mu) = \langle A_m f, \mu^m\rangle$$

for $F(\mu) = \langle f, \mu^m\rangle$, then this definition agrees with the definition of the Fleming–Viot generator $\mathbb{A}$ given in Section 2 by (2.4). The martingale problem for $\mathbb{A}$ is well posed and, hence, Z is the unique solution to the martingale problem for $(\mathbb{A}, \pi_0)$.

From the discussion above it is clear that $(X^N, \hat{Z}_N) \Rightarrow (X, \hat{Z})$, where $\hat{Z} = h_{\text{eq}}Z$ and Z is a Fleming–Viot process with generator $\mathbb{A}$. Since the process $\hat{\mu}^N$ has the same distribution as the process $\hat{Z}_N$, we also have $\hat{\mu}^N \Rightarrow h_{\text{eq}}Z$ in $D_{\mathcal{M}_1(E\times[0,1])}[0,\infty)$. This proves part (B) of Theorem 2.3 with the process ν being the same as the process Z.

REMARK 4.4. The approach of particle construction that we used to prove part (B) of Theorem 2.3 can also be used to show that the distributions of $\hat{\mu}^N(0) = \mu^N(\tau_N)$ converge along the entire sequence. Since $\tau_N \to 0$ a.s., we first do a random time change γ^N (which is a bijection from $[0,\infty)$ to $[0,\infty)$ depending on the population size n^N) such that $\gamma^N(\tau_N) \to \rho$ a.s., where ρ is a positive random variable. We then use the particle construction $\hat{X}^N(t) = (\hat{X}_1^N(t), \hat{X}_2^N(t), \dots)$ similar to the one used here, except that the birth, death and immigration rates are altered according to the random time change. Next we show that the process $\hat{X}^N$ converges on the random time interval $[0, \gamma^N(\tau_N)]$ as $N \to \infty$. With a bit more work it is possible to conclude from this convergence that the distributions of $\mu^N(\tau_N)$ converge as well.

4.3. *Proof of Theorem 2.7.* The membrane molecules are doing speed D Brownian motion on the sphere of radius R, which we call E. Suppose the sphere E is embedded in $\mathbb{R}^3$ with its center at the origin. Let $B = (B_1, B_2, B_3)^T$ be a Brownian motion on E with speed D and let $W = (W_1, W_2, W_3)^T$ be a standard Brownian motion in $\mathbb{R}^3$. Henceforth, let $\langle \cdot, \cdot \rangle$ denote the standard inner product in $\mathbb{R}^3$ and let $\| \cdot \|$ denote the corresponding Euclidean norm. From Stroock [39] it follows that we can express B as the solution of Itô's equation

$$dB = \sqrt{D}\left(I - \frac{BB^T}{R^2}\right) dW - D\frac{B}{R^2}\, dt. \tag{4.37}$$

From above, it is immediate that for any $t \geq 0$,

$$E(B_i(t)) = B_i(0)e^{-2Dt/R^2} \qquad \text{for } i = 1, 2, 3. \tag{4.38}$$

LEMMA 4.5. *Let B and $\bar{B}$ be two independent speed D Brownian motions on the sphere E. Then for any $t > 0$,*

$$E\big(\|B(t) - \bar{B}(t)\|^2\big) = 2R^2\left(1 - \frac{\langle B(0), \bar{B}(0)\rangle}{R^2} e^{-2Dt/R^2}\right).$$

PROOF. This result is a consequence of the simple calculation below:

$$\begin{aligned}
&E\big(\|B(t) - \bar{B}(t)\|^2\big) \\
&\quad = E\big((B_1(t) - \bar{B}_1(t))^2 + (B_2(t) - \bar{B}_2(t))^2 + (B_3(t) - \bar{B}_3(t))^2\big) \\
&\quad = E\big(B_1^2(t) + B_2^2(t) + B_3(t)^2 + \bar{B}_1^2(t) + \bar{B}_2^2(t) + \bar{B}_3(t)^2 \\
&\qquad\qquad - 2B_1(t)\bar{B}_1(t) - 2B_2(t)\bar{B}_2(t) - 2B_3(t)\bar{B}_3(t)\big) \\
&\quad = 2R^2 - 2E(B_1(t))E(\bar{B}_1(t)) - 2E(B_2(t))E(\bar{B}_2(t)) - 2E(B_3(t))E(\bar{B}_3(t)) \\
&\quad = 2R^2\left(1 - \frac{\langle B(0), \bar{B}(0)\rangle}{R^2} e^{-2Dt/R^2}\right) \qquad \text{[using (4.38)]}.
\end{aligned}$$

□

Recall the definition of S_p given by (2.8) and the definition of the process X. We assume that we are at stationarity and, hence, we can also assume that X is defined for all $t \in (-\infty, \infty)$. At any fixed time t the sequence $X(t) = (X_1(t), X_2(t), \ldots)$ is exchangeable and its de Finetti measure $Z(t)$ has the same distribution as $\nu(t)$. Thus, the distribution of two molecules sampled from $\nu(t)$ is the same as the distribution of the first 2 levels $X_1(t)$ and $X_2(t)$. For $i = 1, 2$ let $X_i(t) = (Y_i(t), C_i(t))$, where $Y_i(t) \in E$ and $C_i(t) \in [0, 1]$. From the calculation in Section 2 we can write

$$S_p = E\big(\|Y_1(t) - Y_2(t)\|^2 | C_1(t) = C_2(t)\big). \tag{4.39}$$

The process determined by the first two levels of X evolves according to the generator A_2 given by (4.34) with $m = 2$. From the definition of A_2 it is clear that

level 2 *looks down* to level 1 at rate $2k_{\mathrm{fb}}\alpha$, where $\alpha = (1 - h_{\mathrm{eq}})/h_{\mathrm{eq}}$. Moreover, at both the levels there is an *immigration* event at rate $k_{\mathrm{on}}\alpha$, in which a molecule with a uniformly chosen type in $E \times [0, 1]$ is inserted at that level. In between these lookdowns and immigrations, the molecules at levels 1 and 2 are diffusing on the membrane according to independent speed D Brownian motions.

The quantity S_p can be calculated by tracing back the history from time t. Let τ_{12} be the last lookdown time between the first two levels and τ_i be the last immigration time at level i for $i = 1, 2$. The random variables τ_{12}, τ_1 and τ_2 are independent and exponentially distributed with rates $2k_{\mathrm{fb}}\alpha$, $k_{\mathrm{on}}\alpha$ and $k_{\mathrm{on}}\alpha$, respectively. Let τ be the minimum of τ_1, τ_2 and τ_{12} and so it is an exponential random variable with rate $2(k_{\mathrm{on}} + k_{\mathrm{fb}})\alpha$.

The molecules at levels 1 and 2 will be in the same clan provided $\tau = \tau_{12}$. Molecules at levels 1 and 2 were at the same place at time $t - \tau$ and have been doing independent speed D Brownian motions on the sphere E since then. Using Lemma 4.5, we get

$$\begin{aligned} &E\big(\|Y_1(t) - Y_2(t)\|^2 | C_1(t) = C_2(t)\big) \\ &\quad = 4R^2(k_{\mathrm{on}} + k_{\mathrm{fb}})\alpha \int_0^\infty (1 - e^{-2Ds/R^2}) e^{-2(k_{\mathrm{on}}+k_{\mathrm{fb}})\alpha s}\, ds \\ &\quad = \frac{2D}{((k_{\mathrm{on}} + k_{\mathrm{fb}})\alpha + D/R^2)}. \end{aligned}$$

This proves Theorem 2.7.

Acknowledgments. I wish to sincerely thank my adviser, Professor Thomas G. Kurtz, for his constant support and guidance. A very special thanks to Professor Sigurd Angenent for introducing me to this problem and asking many interesting questions. I also wish to thank Professor Steve Altschuler and Professor Lani Wu for inviting me to their lab at the University of Texas, Southwestern and giving me the opportunity to better understand the biological aspects of this problem.

REFERENCES

[1] ALTSCHULER, S. J., ANGENENT, S. B., WANG, Y. and WU, L. F. (2008). On the spontaneous emergence of cell polarity. *Nature* **454** 886–889.

[2] ATHREYA, K. B. and NEY, P. E. (1972). *Branching Processes*. *Die Grundlehren der Mathematischen Wissenschaften* **196**. Springer, New York. MR0373040

[3] BUTTY, A. C., PERRINJAQUET, N., PETIT, A., JAQUENOUD, M., SEGALL, J. E., HOFMANN, K., ZWAHLEN, C. and PETER, M. (2002). A positive feedback loop stabilizes the guanine-nucleotide exchange factor Cdc24 at sites of polarization. *EMBO J.* **21** 1565–1576.

[4] COVE, D. J. (2000). The generation and modification of cell polarity. *Journal of Experimental Botany* **51** 831–838.

[5] COVE, D. J., HOPE, I. A. and QUATRANO, R. S. (1999). Polarity in biological systems. In *Development*, *Genetics*, *Epigenetics and Environmental Regulation* (V. E. A. Russo, D. J. Cove, L. G. Edgar, R. Jaenisch and F. Salamini, eds.) 507–524. Springer, Berlin.

[6] DAWSON, D. A. and FENG, S. (2006). Asymptotic behavior of the Poisson–Dirichlet distribution for large mutation rate. *Ann. Appl. Probab.* **16** 562–582. MR2244425
[7] DAWSON, D. A. and HOCHBERG, K. J. (1982). Wandering random measures in the Fleming–Viot model. *Ann. Probab.* **10** 554–580. MR0659528
[8] DONNELLY, P. and KURTZ, T. G. (1999). Particle representations for measure-valued population models. *Ann. Probab.* **27** 166–205. MR1681126
[9] DRUBIN, D. G. and NELSON, W. J. (1996). Origins of cell polarity. *Cell* **84** 335–344.
[10] EBERSBACH, G. and JACOBS-WAGNER, C. (2007). Exploration into the spatial and temporal mechanisms of bacterial polarity. *Trends Microbiology* **15** 101–108.
[11] ETHIER, S. N. and KURTZ, T. G. (1986). *Markov Processes*: *Characterization and Convergence*. Wiley, New York. MR0838085
[12] ETHIER, S. N. and KURTZ, T. G. (1993). Fleming–Viot processes in population genetics. *SIAM J. Control Optim.* **31** 345–386. MR1205982
[13] EVANS, J. P., FOSTER, J. A., MCAVEY, B. A., GERTON, G. L., KOPF, G. S. and SCHULTZ, R. M. (2000). Effects of perturbation of cell polarity on molecular markers of sperm–egg binding sites on mouse eggs. *Biology of Reproduction* **62** 76–84.
[14] EWENS, W. J. (1972). The sampling theory of selectively neutral alleles. *Theoret. Population Biology* **3** 87–112. MR0325177
[15] FENG, S. (2009). Poisson–Dirichlet distribution with small mutation rate. *Stochastic Process. Appl.* **119** 2082–2094. MR2519357
[16] FENG, S. (2010). *The Poisson–Dirichlet Distribution and Related Topics*: *Models and Asymptotic Behaviors*. Springer, Heidelberg. MR2663265
[17] FLEMING, W. H. and VIOT, M. (1979). Some measure-valued Markov processes in population genetics theory. *Indiana Univ. Math. J.* **28** 817–843. MR0542340
[18] GASSAMA-DIAGNE, A. E. A. (2006). Phosphatidylinositol-3,4,5-trisphosphate regulates the formation of the basolateral plasma membrane in epithelial cells. *Nature Cell Biol.* **8** 963–970.
[19] GIERER, A. and MEINHARDT, H. (1972). A theory of biological pattern formation. *Kybernetik* **12** 30–39.
[20] GRIFFITHS, R. C. (1979). On the distribution of allele frequencies in a diffusion model. *Theoret. Population Biol.* **15** 140–158. MR0528914
[21] GUPTA, A. (2010). Stochastic models for cell polarity. Ph.D. thesis, Univ. Wisconsin, Madison.
[22] IRAZOQUI, J. E., GLADFELTER, A. S. and LEW, D. J. (2003). Scaffold-mediated symmetry breaking by Cdc42p. *Nature Cell Biol.* **5** 1062–1070.
[23] JOFFE, A. and MÉTIVIER, M. (1986). Weak convergence of sequences of semimartingales with applications to multitype branching processes. *Adv. in Appl. Probab.* **18** 20–65. MR0827331
[24] JOHNSON, N. L., KOTZ, S. and BALAKRISHNAN, N. (1997). *Discrete Multivariate Distributions*. Wiley, New York. MR1429617
[25] KALLENBERG, O. (2002). *Foundations of Modern Probability*, 2nd ed. Springer, New York. MR1876169
[26] KATZENBERGER, G. S. (1991). Solutions of a stochastic differential equation forced onto a manifold by a large drift. *Ann. Probab.* **19** 1587–1628. MR1127717
[27] KIMURA, M. and CROW, J. F. (1964). The number of alleles that can be maintained in a finite population. *Genetics* **49** 725–738.
[28] KINGMAN, J. F. C., TAYLOR, S. J., HAWKES, A. G., WALKER, A. M., COX, D. R., SMITH, A. F. M., HILL, B. M., BURVILLE, P. J. and LEONARD, T. (1975). Random discrete distribution. *J. Roy. Statist. Soc. Ser. B* **37** 1–22. MR0368264
[29] KURTZ, T. G. (1998). Martingale problems for conditional distributions of Markov processes. *Electron. J. Probab.* **3** 29 pp. (electronic). MR1637085

[30] MARCO, E., WEDLICH-SOLDNER, R., LI, R., ALTSCHULER, S. J. and WU, L. F. (2007). Endocytosis optimizes the dynamic localization of membrane proteins that regulate cortical polarity. *Cell* **129** 411–422.

[31] OZBUDAK, E. M., BECSKEI, A. and VAN OUDENAARDEN, A. (2005). A system of counteracting feedback loops regulates Cdc42p activity during spontaneous cell polarization. *Dev. Cell* **9** 565–571.

[32] PELLETTIERI, J. and SEYDOUX, G. (2002). Anterior–posterior polarity in C. elegans and drosophila–PARallels and differences. *Science* **298** 1946–1950.

[33] PITMAN, J. and YOR, M. (1997). The two-parameter Poisson–Dirichlet distribution derived from a stable subordinator. *Ann. Probab.* **25** 855–900. MR1434129

[34] PRUYNE, D. and BRETSCHER, A. (2000). Polarization of cell growth in yeast. *Journal of Cell Science* **113** 571–585.

[35] PRUYNE, D. and BRETSCHER, A. (2000). Polarization of cell growth in yeast I. Establishment and maintenance of polarity states. *Journal of Cell Science* **113** 365–375.

[36] SCHMULAND, B. (1991). A result on the infinitely many neutral alleles diffusion model. *J. Appl. Probab.* **28** 253–267. MR1104564

[37] SHAPIRO, L., MCADAMS, H. H. and LOSICK, R. (2002). Generating and exploiting polarity in bacteria. *Science* **298** 1942–1946.

[38] SHI, S. H., JAN, L. Y. and JAN, Y. N. (2003). Hippocampal neuronal polarity specified by spatially localized mPar3/mPar6 and PI 3-kinase activity. *Cell* **112** 63–75.

[39] STROOCK, D. W. (1971). On the growth of stochastic integrals. *Z. Wahrsch. Verw. Gebiete* **18** 340–344. MR0287622

[40] TAKAKU, T., OGURA, K., KUMETA, H., YOSHIDA, N. and INAGAKI, F. (2010). Solution structure of a novel Cdc42 binding module of Bem1 and its interaction with Ste20 and Cdc42. *Journal of Biological Chemistry* **285** 19346–19353.

[41] TURING, A. M. (1952). The chemical basis of morphogenesis. *Phil. Trans. R. Soc. B* **237** 37–72.

[42] WATTERSON, G. A. (1976). The stationary distribution of the infinitely-many neutral alleles diffusion model. *J. Appl. Probab.* **13** 639–651. MR0504014

[43] WEDLICH-SOLDNER, R., WAI, S. C., SCHMIDT, T. and LI, R. (2004). Robust cell polarity is a dynamic state established by coupling transport and GTPase signaling. *J. Cell Biol.* **166** 889–900.

[44] WEINER, O. D. E. A. (2002). A PtdInsP(3)- and Rho GTPase-mediated positive feedback loop regulates neutrophil polarity. *Nature Cell Biol.* **4** 509–513.

DEPARTMENT OF MATHEMATICS
480 LINCOLN DRIVE
MADISON, WISCONSIN 53706
USA
E-MAIL: gupta@math.wisc.edu