Available online at www.sciencedirect.com

ScienceDirect

Stochastic Processes and their Applications 152 (2022) 149–176

stochastic processes and their applications

www.elsevier.com/locate/spa

Shuffling cards by spatial motion[☆]

Persi Diaconis[a], Soumik Pal[b,*]

[a] Departments of Mathematics and Statistics, Stanford University, Stanford, CA, United States of America

[b] Department of Mathematics, University of Washington, Seattle, WA, United States of America

Received 10 June 2021; received in revised form 1 April 2022; accepted 28 June 2022
Available online 4 July 2022

Abstract

We propose a model of card shuffling where a pack of cards, spread as points on a square table, are repeatedly gathered locally at random spots and then spread towards a random direction. A shuffling of the cards is then obtained by arranging the cards by their increasing x-coordinate values. When there are m cards on the table we show that this random ordering gets mixed in time $O(\log m)$. Explicit constants are evaluated in a diffusion limit when the position of m cards evolves as an interesting $2m$-dimensional non-reversible reflected jump diffusion in time. Our main technique involves the use of multidimensional Skorokhod maps for double reflections in $[0, 1]^2$ in taking the discrete to continuous limit. The limiting computations are then based on the planar Brownian motion and properties of Bessel processes.

MSC: 60J10; 60J60

Keywords: Markov chains mixing time; Card shuffling; Reflected diffusions; Skorokhod maps; Planar Brownian motion; Stochastic flow of kernels

1. Introduction

1.1. The gather-and-spread model of spatial shuffling

Let $D = [0, 1]^2$ represent a square table. Imagine m labeled cards spread on this table. We will ignore the dimensions of the cards and represent them as particles with spatial positions in $[0, 1]^2$. Suppose at each discrete time step an agent selects a spot uniformly at random in D.

☆ This research is partially supported by National Science Foundation, United States of America grants DMS-1208775, DMS-1308340 and DMS-1612483.

* Corresponding author.
E-mail addresses: diaconis@math.stanford.edu (P. Diaconis), soumik@uw.edu (S. Pal).

https://doi.org/10.1016/j.spa.2022.06.023

Consider all cards whose current position lies in a disc of radius $\delta > 0$ centered at that point. She gathers all such cards to the center of the disc in a single heap, randomly selects a direction, tosses an independent coin for each card in that heap, and, for those cards whose coins turn up heads, pushes the cards in that direction for a fixed, bounded distance while keeping them within the boundaries of the table. Other cards, including those whose coins turn up tails, are not moved. She does this independently at each time step. After T steps the cards are projected on the x-axis and arranged in a line in the increasing order of the x-coordinate values. We are interested in the resulting random permutation of the set $[m] := \{1, 2, \ldots, m\}$, especially in estimating T that guarantees that this terminal random permutation is approximately uniform in total variation distance, irrespective of the initial positions of the cards. The model is designed to mimic a popular way to mix cards (called smooshing) by gathering cards locally using both palms and then spreading the cards by dragging them under the palm.

We now give a more formal definition. Let U be the closed disc of radius δ centered at the origin. For any $z \in \mathbb{R}^2$, the set $z + U$ will denote the disc centered at z. Colloquially we will refer to U as the "palm" of the agent put at the point z.

We now describe the "gather" operation. Consider a point $z_0 = (x_0, y_0) \in D$. Let $G^{z_0} : D \to D$ denote the map

$$G^{z_0}(z) := z_0 1\{z \in z_0 + U\} + z 1\{z \notin z_0 + U\}, \qquad z = (x, y) \in D.$$

That is, points under the palm are gathered to the center. If z is close to the boundary of D, there are fewer points to which it can be gathered. For tractability of our stochastic processes we will require some spatial homogeneity. This inspires the following extended definition.

Let $\overline{D}$ denote the Minkowski sum of the two sets D and U. That is $\overline{D} = \cup_{z\in D}\{z + U\}$. For $x \in \mathbb{R}$, let

$$\widetilde{x} := \max(\min(x, 1), 0) = \begin{cases} 0, & \text{if } x < 0. \\ x, & \text{if } 0 \le x \le 1. \\ 1, & \text{if } x > 1. \end{cases}$$

For $z_0 = (x_0, y_0) \in \overline{D}\backslash D$, define $\widetilde{z}_0 := (\widetilde{x}_0, \widetilde{y}_0)$ and

$$G^{z_0}(z) = \widetilde{z}_0 1\{z \in z_0 + U\} + z 1\{z \notin z_0 + U\}, \; z \in D.$$

That is, points under the palm are gathered to a boundary point in case the center is outside D. See Fig. 1 where the point $\mathbf{z}$ is outside the unit square and the corresponding $\tilde{\mathbf{z}}$ is on the boundary. Hence if the palm is placed such that the center is on $\mathbf{z}$, all cards under the palm will be gathered at $\tilde{\mathbf{z}}$.

We now define the "spread" operation. Fix $s_0 > 0$, a $\theta \in [0, 2\pi]$, and a point $z_0 = (x_0, y_0) \in \overline{D}$. For $z = (x, y) \in D$ such that $z \in z_0 + U$, let $x_1 = x + s_0 \cos(\theta)$ and $y_1 = y + s_0 \sin(\theta)$. Define the map $f_{s_0}^{z_0,\theta} : D \to D$, by

$$f_{s_0}^{z_0,\theta}(z) = \begin{cases} (\widetilde{x}_1, \widetilde{y}_1), & \text{if } z \in \{z_0 + U\} \cap D. \\ z, & \text{otherwise.} \end{cases} \tag{1}$$

Thus, for $z \in \{z_0 + U\} \cap [0, 1]^2$ ("cards under the palm") we move z linearly in the direction θ for distance s_0 until we hit the boundary of the table and stop moving the coordinate that is at the boundary. Nothing else is touched. Again, see Fig. 1 where the point (x, y), under the palm, is dragged until the x-coordinate hits one and does not increase any more. The y-coordinate, however, continues to decrease. (x_1, y_1) represents the position of the particle had there been no boundary. The actual position is given by the coordinates $(\widetilde{x}_1, \widetilde{y}_1)$, where $\widetilde{x}_1 = 1$ since $x_1 > 1$.

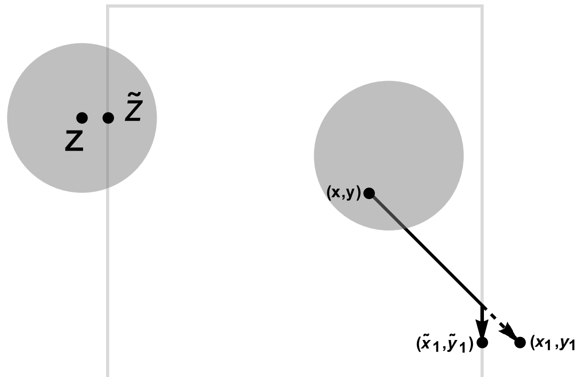

Fig. 1. The effect of the boundary. On the left, the palm is centered at z which is outside the table. All cards in the shaded area inside the square will then be gathered at $\tilde{z}$. On the right we see spread (without gather). (x, y) is a point under the palm which is not gathered at the center. The palm drags the card at (x, y) along a straight line until the x-coordinate hits the boundary. The x-coordinate freezes while the y-coordinate continues to decrease.

If two or more cards occupy the same position (say due to gathering) we need additional randomization to break the ties. Fix $0 < p < 1$. Every time a card is about to be spread, it tosses an independent coin with probability p of turning up heads. If it turns up heads, the card follows the palm in the chosen direction. Otherwise it does not move. We define this process formally below.

Fix $\lambda > 0$. For mathematical convenience we consider continuous time and model the random selection of spots by the agent as a Poisson point process (PPP) on $(0, \infty) \times \overline{D}$ of constant rate λ with respect to the product Lebesgue measure. Since $\overline{D}$ is bounded, it is possible to enumerate the atoms of this point process in a sequence $\{(t_i, w_i),\ i \in \mathbb{N}\}$ such that $t_1 < t_2 < t_3 < \cdots$, and each $w_i \in \overline{D}$. One can obtain a discrete time model by discarding t_is and considering the sequence $(w_i,\ i \in \mathbb{N})$ of i.i.d. uniformly chosen points in $\overline{D}$ at discrete time points $i \in \mathbb{N}$.

Let ν_0 be any probability distribution on $[0, 2\pi]$ that satisfies the following unbiasedness assumption. Here and throughout, $\nu(f(\cdot))$ for a probability measure ν and a function f, suitably measurable, will denote the expectation of f under ν.

Assumption 1. Assume that $\nu_0(\cos(\cdot)) = \nu_0(\sin(\cdot)) = \nu_0(\sin(\cdot)\cos(\cdot)) = 0$ and that $\nu_0\left(\cos^2(\cdot)\right) = \nu_0\left(\sin^2(\cdot)\right) = \sigma^2$, for some $\sigma > 0$.

Examples of ν_0 include the uniform distribution over $[0, 2\pi]$ and the discrete uniform distribution over the set $\{0, \pi/2, \pi, 3\pi/2\}$ with $\sigma^2 = 1/2$ in both cases.

Generate an i.i.d. sequence $(\theta_i,\ i \in \mathbb{N})$ sampled from ν_0 and consider the sequence of functions $\left(f_{s_0}^{w_i,\theta_i},\ i \in \mathbb{N}\right)$. Fix $m \in \mathbb{N}$. Let $0 < p < 1$ be fixed as before.

Definition 1 (*m Point Motion Under Gather-and-Spread*)**.** Let $Z_j(0) = z_j \in D$, $j \in [m]$, denote the initial positions of m cards. Generate an i.i.d. array $\left(H_j(i),\ j \in [m],\ i \in \mathbb{N}\right)$ of Bernoulli(p) random variables. Define sequentially, for $i \in \mathbb{N}$, $j \in [m]$, starting with $t_0 = 0$,

$$\begin{aligned} Z_j^0(t_i) &:= G^{w_i}\left(Z_j(t_{i-1})\right), \quad (\textbf{gather}) \textit{ and} \\ Z_j(t_i) &:= H_j(i) f_{s_0}^{w_i,\theta_i}\left(Z_j^0(t_i)\right) + \left(1 - H_j(i)\right) Z_j^0(t_i), \quad (\textbf{spread}). \end{aligned} \tag{2}$$

Extend the sequence $\left(Z_j(t_i),\ j \in [m],\ i = 0, 1, 2, \ldots\right)$ to all $t \in [0, \infty)$ by defining

$$Z_j(t) = Z_j(t_{i-1}), \quad t_{i-1} \le t < t_i, \quad i \in \mathbb{N},\ j \in [m]. \tag{3}$$

The resulting continuous time Markov chain $\left(Z_j(t),\ j \in [m], t \ge 0\right)$ on the state space D^m will be called the *m point motion under the gather-and-spread model*.

Remark 1. As before we choose to work with $\overline{D}$ instead of D for technical reasons. For a PPP on $(0, \infty) \times D$ of constant rate λ, the cards near the boundary move less frequently than those near the center of the table, affecting spatial homogeneity. Intuitively, the two models are not too different if U is small compared to D.

Many natural variations of this model can be analyzed by the methods of this paper. For example, instead of gathering all the cards under the palm at the center one can choose a new independent, uniformly at random, position under the palm for each card. This is an example of *local mixing*. It is also possible to change the spread by selecting a probability distribution on $[0, \infty)$ and deciding the spread s_0 of each gathered card by sampling independently from it. For all such models the analysis in the paper remains similar.

We now describe what we mean by shuffling using this spatial motion. Let $\mathcal{S}_m$ denote the group of permutations of m labels. Consider an m-tuple of real numbers $(x_1, \ldots, x_m)$. Define the rank-to-index permutation corresponding to this set in the following way. If every coordinate is distinct, then one can arrange the coordinates in increasing order $x_{(1)} < x_{(2)} < \cdots < x_{(m)}$, for a unique element $\gamma \in \mathcal{S}_m$ such that $x_{\gamma_i} = x_{(i)}$. Say that the *rank* of x_{γ_i} is i while the index of x_{γ_i} is γ_i. When all coordinates are not distinct, the rank-to-index will refer to a random variable taking values in $\mathcal{S}_m$ which is obtained by "resolving the ties at random". To do this rigorously, generate i.i.d. uniform $[0, 1]$ random variables $\{U_1, \ldots, U_m\}$. Consider the set of pairs $\{(x_1, U_1), \ldots, (x_m, U_m)\}$. Rank the above sequence in the increasing dictionary order. That is $(x_i, U_i) < \left(x_j, U_j\right)$ if, either $x_i < x_j$ or $\{x_i = x_j\} \cap \{U_i < U_j\}$. It is clear that, almost surely, this gives us a totally ordered sequence with no equalities. As before, let $\gamma \in \mathcal{S}_m$ be the unique element such that $\left(x_{\gamma_i}, U_{\gamma_i}\right)$ is the ith smallest element in the above ordering. Then γ will be called the rank-to-index permutation corresponding to $(x_1, \ldots, x_m)$. This also works when one replaces $(x_1, \ldots, x_m)$ by a random vector $(X_1, \ldots, X_m)$ by first obtaining the rank-to-index permutation conditioned on a realization of $(X_1, \ldots, X_m)$ and then integrating with respect to the law of the conditioned random vector. For example, if $m = 4$ and $x_1 = 1, x_2 = 2, x_3 = 1, x_4 = 3$ then resolve the tie between x_1 and x_3 by generating i.i.d. Uni$(0, 1)$ random variables U_1, U_3. Depending on whether $U_1 < U_3$ or $U_1 > U_3$, the rank-to-index permutation would be either of the following (equally likely)

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 2 & 4 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 2 & 4 \end{pmatrix}.$$

Consider the Markov chain $Z_j(\cdot) = \left(X_j(\cdot), Y_j(\cdot)\right),\ j \in [m]$, from Definition 1. For any time $t \ge 0$, let $\gamma(t)$ denote the rank-to-index permutation corresponding to the set of x-coordinates of the m points, $\{X_1(t), \ldots, X_m(t)\}$. Let $\|\gamma(t) - \mathrm{Uni}\|_{\mathrm{TV}}$ denote the total variation distance between the law of $\gamma(t)$ and the uniform distribution on $\mathcal{S}_m$. By "mixing time of shuffling" we refer to the first time $t_{\mathrm{mix}}(\epsilon)$ when this total variation distance is smaller than a given $\epsilon > 0$, say $1/4$, irrespective of the vector of initial positions $\left(Z_j(0) = z_j,\ j \in [m]\right)$.

In Section 2.2 we provide an $O(\log m)$ mixing time bound for this Markov chain by developing a general scheme for all such problems. However, precise calculations of constants are not easy to derive for the discrete model. This difficulty is partly due to the existence of the boundary of D. Of course, projection of Z_j's on the x-axis is arbitrarily chosen. By symmetry, the same bound holds for projection on the y-axis. Whether the problem of the boundary gets simplified by a more judicious choice of curve for projection is an interesting problem that is not answered here.

In Section 3 the problem is simplified under a jump-diffusion limit as follows. First we will consider the parameter $\lambda \to \infty$ and $s_0 = 1/\sqrt{\lambda} \to 0$ while keeping D, U, and ν_0 fixed. This means that we will make a lot of short spread moves. Furthermore, we will make gatherings rare by defining a *lazy gathering model*. For each t_i of the PPP in Definition 1 we will toss an independent coin with a probability of heads given by $1/\lambda$. If the coin turns heads, we perform both gather and spread steps in (2); otherwise we skip the gather step in (2) and only do the spread (with notations from (2)),

$$Z_j(t_i) := H_j(i) f_{s_0}^{w_i,\theta_i}\left(Z_j(t_{i-1})\right) + \left(1 - H_j(i)\right) Z_j(t_{i-1}), \quad i \in \mathbb{N},\ j \in [m]. \tag{4}$$

Informally, per unit amount of time, we spread cards about λ many times, each time by distance $O(1/\sqrt{\lambda})$, before gathering once. In fact, we gather at the jumps of a Poisson process of rate one. In between these jumps the Markov chain of m point motion converges in law to a $2m$ dimensional diffusion with state space D^m and reflected at the boundary. Thus the process evolves as a reflected diffusion that jumps according to a kernel at the points given by a Poisson process of rate one. This jump-diffusion is non-reversible and has reflections at the boundary of the non-smooth domain D. Hence information regarding its stationary distribution and rate of convergence cannot be inferred by standard methods. Nevertheless, we can bound the mixing time of shuffling thanks to our coupling scheme.

Theorem 1. *Under Assumption 1, for all $m \geq 2$, the m-point motion under the lazy gathering model converges weakly in the Skorokhod space to a $2m$-dimensional reflected jump-diffusion model with state space $[0, 1]^{2m}$ satisfying a stochastic differential equation described below in (12).*

The mixing time of shuffling m cards in this jump-diffusion model, $t_{\text{mix}}(\epsilon)$, is bounded above by $\frac{1}{\mathfrak{p}} \log(m/\epsilon)$, where $0 < \mathfrak{p} < 1$ is given by

$$\mathfrak{p} := \frac{\delta^2}{2p\pi\sigma^2(1+2\delta)^2} K_0\left(\frac{\sqrt{2}+2\delta}{\sigma\delta\sqrt{\pi p(1-p)}}\right). \tag{5}$$

Here K_0 is the modified Bessel function of the second kind.

Details about the modified Bessel function K_0 can be found in [1, Section 9.6]. Notably, for a real argument z, the following asymptotics hold: $K_0(z) \sim -\log(z)$ as $z \to 0$, and $K_0(z) \sim \sqrt{\pi/(2z)}\exp(-z)$, as $z \to \infty$.

The constant in (5) is an increasing function of δ, for a fixed choice of p and σ^2. However, it is rather small and the bound is far from optimal. For example, suppose the palm is large enough to cover the entire table if placed at the center. That is, $2\delta > \sqrt{2}$. Then, after an exponential amount of time, the palm will gather all points under it which are then automatically uniformly shuffled by our tie-breaking rule. But, of course, this bound will not work for moderate to small δ. On the other hand, for a choice of a moderate $\delta = 0.3$ we get a minuscule $\mathfrak{p} \approx 1.88 \times 10^{-7}$! Theorem 1 should be interpreted as simply an upper bound that is logarithmic in the number of cards with a large, but known, constant.

As mentioned before, gathering is a local mixing strategy. For example, imagine a viscous fluid (say, cake batter). We take a beater, randomly choose spots in the batter, and vigorously mix the location. If there are particles on the batter in that specific location, they will be so vigorously mixed as to be *exchangeable* in their future evolution. Our results are valid for all such local mixing procedures. The spread, on the other hand is an "advection–diffusion" where we imagine a rod dipped in the fluid being dragged in a direction and creating a shear in its wake. Now, for best mixing practices in the non-stochastic setting it is intuitive to desire a chaotic system. This is usually achieved by repeating two perpendicular directions of shear with self-crossing trajectories (see Aref [3]) such as a repeated figure eight movement through the fluid. This, along with diffusion in the fluid, cause mixing. In this sense, our gather-and-spread moves have been designed to study the effect of local mixing and a stochastic advection on particles in an underlying fluid.

The analogy, however, breaks down in the meaning of the word "mixing", which is used in a different sense in fluid mixing. In that context, the points are unlabeled and we are interested in the difference between the empirical distribution of the points from the uniform distribution. This is not the case here. In fact, as shown in Fig. 2, there will always be clumps of points. In fact, it is not hard to see from our diffusion analysis that even for a single point, the uniform distribution is not the stationary distribution since the corners will have slightly more mass than the rest. Nevertheless we find this analogy motivating to further study both subjects.

1.2. Review of literature

The mathematical study of shuffling cards has a long history going back to Poincaré [36]. We recommend the book [31] for an introduction and [39,40] for a comprehensive overview. Spatial mixing is one of three standard ways in which playing cards are shuffled, and the least investigated so far. The other two, riffle shuffle and the overhand shuffle have had substantial theoretical development and sharp rates of convergence under a variety of models are available. For riffle shuffling, the original work of Bayer and Diaconis [6] shows that n cards repeatedly mixed show a sharp cutoff in total variation convergence to the uniform distribution after $(3/2)\log_2(n) + c$ shuffles. This is accompanied by explicit determination of the eigenvalues and eigenvectors [14] and similarly sharp rates for functions of the current arrangement [4]. The attendant connection with random walk on the chambers of the braid arrangement [9,12,34] blends this shuffling scheme with a healthy part of geometry and algebra. For overhand shuffling, work of Pemantle [34] and Jonasson [25] shows that order $n^2 \log n$ overhand shuffles are necessary and sufficient for mixing (cutoff is an open problem at this time).

There are other well-studied Markov chains for card-shuffling such as the *random adjacent transposition* chains that come with a spatial flavor. Picture n labeled cards in a line, originally in order $1, 2, \dots, n$. At each step an adjacent pair of cards is chosen at random and the two cards are transposed. Results of Diaconis and Saloff-Coste [15] followed by Wilson [45] show that order $n^3 \log n$ steps are necessary and sufficient for convergence. Recently Lacoin [28] sharpened this to show that there is a total variation cut-off at $n^3 \log n/(2\pi^2)$. One can extend the analysis to higher dimensions. For example, in two dimensions cards can be arranged on the vertices of a $\sqrt{n} \times \sqrt{n}$ grid. At each step an edge is chosen at random and the two cards at the vertices of this edge are transposed. This takes order $n^2 \log n$ to mix [15]. These problems have become of recent interest as the 'interchange process' because of their connections to suggestions of Dirac and Feynman in quantum mechanics. See [32] for a tutorial and articles by Alon and Kozma [2] and Berestycki and Kozma [7] for interesting results. A related 'mean-field' walk is the 'random-to-random' walk. A randomly chosen card is removed and reinserted in a random position. In [15] order $n \log n$ steps are shown to be necessary and sufficient. In

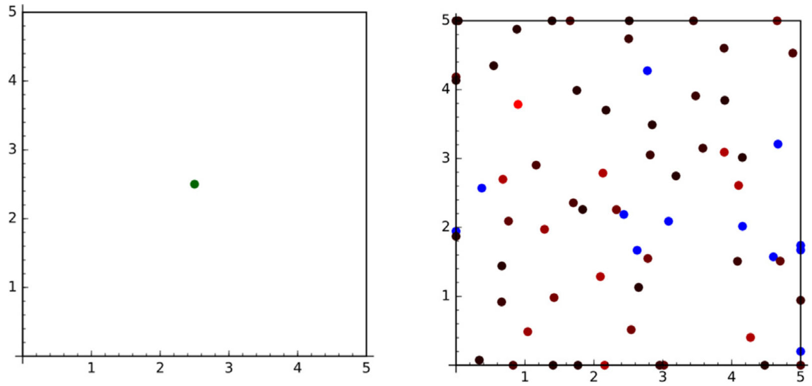

Fig. 2. Configuration of 250 cards after 3000 steps of gather-and-spread. Table is $[0, 5]^2$, $\delta = 0.5$, $s_0 = 1$, $p = 0.5$. Initially all cards are at the center shown on the left. At the terminal step, shown on the right, there are 68 clusters of cards which are colored by the number of points in the cluster relative to the maximum. Singleton clusters are colored blue while the largest clusters are in bright red. There are 23 clusters on the boundary.
Source: Figure produced by Yuqi Huang.

a tour-de-force [16] Dieker and Saliola determine all the eigenvalues and eigenvectors of this chain. A sharp cut-off has been recently established by Bernstein and Nestoridi in [8]. A more overtly spatial walk is studied by Pemantle [35]. He considers n^2 cards at the vertices of an $n \times n$ array. At each step an element x of the array is chosen uniformly at random. Then with probability $1/2$ the rectangle of cards above and to the left of x is rotated 180° degrees, and with probability $1/2$ the rectangle of cards below and to the right of x is rotated 180° degrees. While this is not a particularly natural model, it does have fascinating mixing properties. Pemantle shows that order $\Theta(n^2)$ steps are necessary and sufficient to mix all n^2 cards. However, for a fixed set of k cards, $c(k)n$ steps suffice. Here, $c(k)$ is of order $k^3 (\log k)^2$.

In contrast, there has been virtually *no* analysis of true spatial mixing of cards. This is despite this method being widely used in casino games such as Baccarat and in California Poker. The problems involved need a different kind of analysis. Spatial mixing is inherently a continuous process and fluid type models are needed. Although mixing of fluids has been studied in deterministic contexts [18–20,33,43], the literature does not cover random fluid mixing. The present paper presents a completely new model and analysis approach which gives a reasonable indication of the time to mixing. The model and analysis can be varied to handle natural variations and applications to mixing of things like dominoes and mahjong tiles. The mathematical details involve a substantial excursion into stochastic calculus and a novel jump diffusion approximation which we hope will be of independent interest. Note that a distinction between the above literature and our model is that the stochastic process of permutations given by the x-coordinates of the cards in our model is not a Markov process by itself, but a function of an underlying Markov process given by the spatial positions of the cards.

We conclude this review by reporting that we have also undertaken both simulations and a study of real world smooshing. Simulations on the gather-and-spread and related models were done by students at the University of Washington. The full report can be found at [17], from which Fig. 2 is taken. At Stanford, a group of students smooshed for various times (60 s, 30 s, 15 s) with 52 cards and 100 repetitions for each time (so 100 permutations for each of three times). A collection of ad-hoc test statistics were studied: position of original top (bottom) card, number of originally adjacent cards remaining adjacent, distance to the starting order in various metrics, length of the longest increasing subsequence, etc. The results suggest that randomness sets in after 30 s or so while 15 s was far from random. Since this kind of shuffling is used in both poker tournaments and Monte Carlo (see Diaconis et al. [13] for more on this), further study is of interest.

Outline of the argument

Our argument has three main steps. The first, in Section 2.2, is a coupling argument that gives $O(\log m)$ bound for the mixing time. The second derives a scaling limit of the gather-and-spread model to a jump diffusion. This is achieved in Section 3 via a method that originated from queueing theory. The idea is to the view the "constrained" discrete process as the result of applying a deterministic Skorokhod map to a discrete "non-constrained" martingale and then take weak limits of the martingale sequence. The martingale sequence converges to Brownian motion or a related process and the limit of the constrained process is then given by the Skorokhod map applied to the limiting Brownian motion. The third step bounds the constant in the $O(\log m)$ bound by using stochastic calculus. This is done in Section 4. The idea is that the specific properties of the reflected diffusion reduces the problem to a calculation involving the paths of a time-changed two-dimensional Bessel process (see (27)). Since the two-dimensional Bessel process is also the norm of a planar Brownian motion, the relevant quantities can all be explicitly calculated or bounded above. This gives Theorem 1.

2. Dimension consistency

2.1. A toy one-dimensional model

Before we employ our coupling scheme in the general setting, it is instructive to use the same strategy in a simpler model. Fix arbitrary positive integers m, N. Consider m labeled particles (representing cards) on a line. The position of each card can be one of the N positive integers $[N] := \{1, 2, \ldots, N\}$.

At time zero, the position of each particle is fixed, say, $X_i(0)$, $i \in [m]$. Time is discrete: $t = 0, 1, 2, 3, \ldots$. At each time t, pick a uniform random site, i.e., a random integer k from the set $[N]$. Toss a fair coin to decide left or right. For each card that is currently at site k (there may not be any), toss an independent coin with probability p of turning up heads.

- If we decided left, move all cards whose coins turn heads to the left by one, if possible. That is, if $X_i(t) = k$ and the coin for i turned heads, then $X_i(t+1) = (k-1)$, unless $X_i(t) = 1$, in which case $X_i(t+1) = 1$.
- If we decided right, move all cards whose coins turn heads to the right by one, if possible. That is, if $X_i(t) = k$ and the coin for i turned heads, then $X_i(t+1) = X_i(t) + 1$, unless $X_i(t) = N$, in which case $X_i(t+1) = N$.
- For all other cards $X_j(t+1) = X_j(t)$.

That is, imagine cards in boxes on a line, where $X_i(t)$ is the box of the ith card at time t, and cards pile on top of existing cards when they jump. But these details are not important mathematically. Repeat the above procedure by sampling a random site and the coin tosses at every time independent of the past. This gives us a stochastic process $X(t) := (X_1(t), X_2(t), \ldots, X_m(t))$, $t = 0, 1, 2, \ldots$, which is the m point motion on the line. At the end of time T, if we gather cards from left to right (breaking ties in any order), we get a shuffling of m cards. We will estimate the mixing time of this shuffle using coupling.

Consider another pack of cards whose positions will be denoted by $X'_i(\cdot)$, $i = 1, 2, \ldots, m$. Critically, assume that the joint distribution of the random vector $X'(0)$ is exchangeable. This is true, for example, if $X'(0)$ is distributed according to the stationary distribution for the Markov chain of the motion of m labeled particles described above. However, we do not need to assume that $X'(0)$ is distributed according to a stationary distribution for the following argument to hold. This is also true, for example, if the indices of the process $X'(0)$ are assigned, uniformly at random, independent of their starting positions. What is important is that, at any future point in time t, its rank-to-index permutation process $\gamma'(t)$ remains uniformly distributed over $\mathcal{S}_m$. This is guaranteed when the initial positions of X' are exchangeable since the motion of the m coordinates of X' is symmetric.

Let $Q_m(x)$ denote the law of the process X starting from $X(0) = x$ and let $\mathbb{Q}_m$ denote the law of the process X' (under randomized indices). Now consider $2m$ cards whose positions will be denoted by $\left(X_1, \ldots, X_m, X'_1, \ldots, X'_m\right)$. All $2m$ cards will now move according to the same choices of site selection and the direction to move. That is, both X and X' run using the same "noise", i.e., by the same realizations of uniform pick of random sites as well as the coin tosses to decide left or right.

Let τ_i denote the first time t such that $X_i(t) = X'_i(t)$. Couple the two processes for all subsequent times by defining $X_i(t) = X'_i(t)$ for all $t \geq \tau_i$. Note that, for the individual coin tosses for the cards in order to decide if it should move or not, we use independent tosses until any τ_i, after which X_i and X'_i are identified and will use the same coin toss.

The crucial observation is that the marginal distribution of the m-dimensional process $(X_1, \ldots, X_m)$ is still $Q_m(X(0))$, as if the other m coordinates $(X'_1, \ldots, X'_m)$ do not exist. Similarly, the law of the other m-dimensional process $(X'_1, \ldots, X'_m)$ is $\mathbb{Q}_m$. We will call this property *dimension consistency* and will be shared by all the models in this paper. This property is a consequence of the fact that our models are point motions of stochastic flows of kernels that arose from the article by Harris [21]. See [5,10,23,29,30] for subsequent developments in this theory. Hence, if we let $\tau^* = \max_i \tau_i$, then it follows that the total variation distance between $X(t)$ and $X'(t)$ is bounded above by $P(\tau^* > t)$. Thus, it suffices to estimate the tails of τ^*.

Consider again a system of $2m$ cards where two of them, say X_1 and X'_1, are initially at position x and y, respectively. Let τ_{xy} denote the coupling time of these two cards. It follows from dimension consistency that the law of τ_{xy}, under any initial positions of the $2m$ cards such that $X_1(0) = x$, $X'_1(0) = y$, is equal to its law under $Q_2(x, y)$.

Suppose that there is a random variable ζ such that every τ_{xy} is stochastically dominated by ζ, irrespective of x and y. In the above example, suppose $1 \leq X_1(0) = x \leq X'_1(0) = y$ without loss of generality. Both X_1 and X'_1 are lazy reflected random walks that jump at each turn with probability p/N (probability $1/N$ for its site to be chosen and p for its coin to turn up heads). Thus, irrespective of x, y, τ_{xy} is bounded by ζ^* which is the hitting time of 1 for a lazy reflected random walk starting at N. The distribution of ζ^* can be found explicitly, but what is more important for our purpose is that there is another (possibly a different choice) random variable ζ such that every τ_{xy} is stochastically dominated by ζ, and ζ has geometric tails. To see this note by the Markov property,

$$
\begin{aligned}
&Q_2(x, y)\left(\tau_{xy} > t + s \mid \tau_{xy} > t,\ X_1(t) = x', X'_1(t) = y'\right) \\
&\qquad = Q_2(x', y')\left(\tau_{x'y'} > s\right) \leq P(\zeta^* > s).
\end{aligned}
$$

Thus, $Q_2(x, y)\left(\tau_{xy} > t + s \mid \tau_{xy} > t\right) \le P(\zeta^* > s)$, and, hence

$$Q_2(x, y)\left(\tau_{xy} > t + s\right) \le Q_2(x, y)\left(\tau_{xy} > t\right) P(\zeta^* > s) \le P(\zeta^* > t)P(\zeta^* > s).$$

Choose an s such that $P(\zeta^* > s) \le 1/2$. Then, by iterating the above argument

$$Q_2(x, y)\left(\tau_{xy} > t + ks\right) \le Q_2(x, y)\left(\tau_{xy} > t\right)\left(P(\zeta^* > s)\right)^k \le \frac{1}{2^k}.$$

Thus, there exists a constant C, depending on s (say $C = 2^s$), such that

$$Q_2(x, y)(\tau_{xy} > t) \le C2^{-t}, \quad \text{for all } t > 0,$$

where we have chosen the parameter t to be continuous for convenience.

Now return to the scenario of $2m$ cards, m of which have randomized indices. Recall that $\tau^* = \max_i \tau_i$. By a union bound, for any $t > 1$,

$$P\left(\tau^* > t \log_2 m\right) \le P\left(\cup_{i=1}^m \{\tau_i > t \log_2 m\}\right) \le Cm2^{-t\log_2 m} \le Cm^{-t+1}.$$

This gives an upper bound on the total variation distance between random vectors $X(t)$ and $X'(t)$.

Consider from Section 1 the rank-to-index permutation process $\gamma(t)$ corresponding to $X(t)$, and $\gamma'(t)$ corresponding to $X'(t)$. Recall that $\gamma'(t)$, for each t, is distributed uniformly over $\mathcal{S}_m$. It follows (e.g. by choosing the same set of uniform random variables to break ranks) that the total variation distance between $\gamma(t)$ and $\gamma'(t)$ cannot be larger than that of $X(t)$ and $X'(t)$. Thus $\|\gamma(t) - \mathrm{Uni}\|_{\mathrm{TV}} \le Cm^{-t+1}$. It is easy to see that the right side of the above inequality is ϵ if we choose $t = O(\log m)$. Thus the mixing time of shuffling $t_{\mathrm{mix}} = O(\log m)$.

We wish to repeat the fact that it is unimportant for this argument to assume that X' is distributed according to the stationary distribution of the m-point motion. This is because we are only interested in the mixing time of shuffling which depends on the rank-to-index permutations and not on the spatial locations. However, this saves us the trouble of proving the existence of a stationary distribution for more complex models.

2.2. *A general bound for a triangular array of multidimensional models*

Although the gather-and-spread model is a two-dimensional model, all such fluid-dynamical models, irrespective of dimension, can be studied by a key consistency property that is described below.

Throughout this paper, RCLL will refer to functions from $[0, \infty) \to \mathbb{R}^d$ that are right continuous on $[0, \infty)$ and admit left limit at every point in $(0, \infty)$. For every $m \in \mathbb{N}$, suppose that on a right-continuous and complete filtered probability space we have a strong Markov process $Z^{(m)}(t) = \left(Z_1^{(m)}(t), \ldots, Z_m^{(m)}(t)\right)$, $t \ge 0$, where each component process has state space $\mathbb{R}^d$, for some $d \ge 1$, and RCLL paths. In particular, we allow discrete time Markov processes by extending them to continuous time by piecewise constant interpolation. Let $Q_m(z_1, \ldots, z_m)$ denote the process law starting from the initial vector $(z_1, \ldots, z_m)$. At time t extract the "x-coordinates" (which can be any of the d coordinates) $X_1^{(m)}(t), \ldots, X_m^{(m)}(t)$ and consider the rank-to-index permutation $\gamma^{(m)}(t)$. We are interested in the mixing time of shuffling for this permutation.

Assumption 2 (*Dimension Consistency*)**.** We say that the family of models Q_m, $m \in \mathbb{N}$, satisfies the dimension consistency property if the following holds true. For any $m \in \mathbb{N}$, let

$Z^{(m)}$ be distributed according to $Q_m(z_1, \ldots, z_m)$. Then, the marginal distribution of the process $\left(Z_i^{(m)}, Z_j^{(m)}\right)$, for any $i \neq j$, is $Q_2(z_i, z_j)$.

Note that the m point motion under the gather-and-spread model in Definition 1 satisfy this property. In fact, recall the comment on the structural similarities between a fluid mechanics model and our card shuffling model as remarked in Section 1.2. Assumption 2 will be valid for all such models.

For $x, y \in \mathbb{R}^d$, let $\overline{\tau}_{x,y}^{(m)}$ denote the stopping time

$$\overline{\tau}_{x,y}^{(m)} = \inf\left\{t \geq 0 :\ Z_i^{(m)}(t) = Z_j^{(m)}(t)\right\}, \quad Z_i^{(m)}(0) = x,\ Z_j^{(m)}(0) = y. \tag{6}$$

Note that $\overline{\tau}_{x,y}^{(m)}$ is indeed a stopping time due to our assumptions on the filtration. By Assumption 2, its distribution under $Q_m(z_1, \ldots, z_m)$ does not depend on m or z_k, $k \notin \{i, j\}$, and we will drop the m from its notation and denote it by $\overline{\tau}_{x,y}$.

Assumption 3. There is a positive random variable ζ^* such that $\overline{\tau}_{x,y}$ is stochastically dominated by ζ^* irrespective of x, y.

Theorem 2. *Suppose Assumptions 2 and 3 hold. Then, there exists a positive constant C such that for all $x, y \in \mathbb{R}^d$,*

$$P\left(\overline{\tau}_{x,y} > t\right) \leq C2^{-t}, \quad \text{for all } t \geq 0. \tag{7}$$

Moreover, if there is an $\alpha > 0$ such that

$$\liminf_{u\to 0+} \frac{1}{u} P\left(\zeta^* \leq u\right) \geq \alpha, \tag{8}$$

e.g. when ζ^* has a continuous density at zero that is positive, then every $\overline{\tau}_{x,y}$ is stochastically dominated by an exponential random variable with rate α. The mixing time for shuffling is $O(\log m)$ as $m \to \infty$, in particular, when (8) holds, $t_{\mathrm{mix}}(\epsilon) \leq \frac{1}{\alpha}\log(m/\epsilon)$.

Proof. Fix m. Let $G(t) = P(\zeta^* > t) = 1 - P(\zeta^* \leq t)$. As in Section 2, proceed by applying the strong Markov property: for any $t, s > 0$,

$$P\left(\overline{\tau}_{x,y} > t + s\right) \leq P\left(\overline{\tau}_{x,y} > t\right) G(s) \leq G(t)G(s).$$

Exactly as before (7) follows. For the stronger conclusion (8), note that for a rational $t = k/n$, iterating the above gives us

$$P\left(\overline{\tau}_{x,y} > k/n\right) \leq (G(1/n))^k = \left(1 - P(\zeta^* \leq 1/n)\right)^{nt}.$$

For any $0 < \delta < \alpha$, if n is taken large enough, by (8), $P(\zeta^* \leq 1/n) \geq (\alpha - \delta)/n$. Thus

$$P\left(\overline{\tau}_{x,y} > t\right) \leq \left(1 - \frac{\alpha - \delta}{n}\right)^{nt} \to e^{-(\alpha-\delta)t},$$

as $n \to \infty$. Now take $\delta \to 0+$ to get $P\left(\overline{\tau}_{x,y} > t\right) \leq \exp\left(-\alpha t\right)$ for all rational $t > 0$. For irrational t, the inequality follows by right-continuity of the distribution function of $\overline{\tau}_{x,y}$. This proves that $\overline{\tau}_{x,y}$ is dominated by an exponential(α) random variable.

The claim about mixing time of shuffling being $O(\log m)$ follows as before. Let us argue the special case when (8) holds. In that, by stochastic domination by an exponential(α) variable, we have $Q_2(x, y)\left(\tau_{xy} > t\right) \leq e^{-\alpha t}$.

Hence, as in the case of the toy one-dimensional model described in the last subsection, start now with $2m$ cards (X, X') where the indices of X' are assigned uniformly at random, independent of the rest. Let τ_i be the coupling time of cards X_i and X_i', and, let $\tau^* = \max_{i\in[m]} \tau_i$. One gets from the union bound

$$\left\| X(t) - X'(t) \right\|_{\mathrm{TV}} \leq P\left(\tau^* > t\right) \leq m e^{-\alpha t}.$$

By equating the RHS to be ϵ and considering the pushforward to the rank-to-index permutations of both processes, as before, we get $t_{\mathrm{mix}}(\epsilon) \leq \frac{1}{\alpha} \log(m/\epsilon)$. $\square$

The mixing time of shuffling for the m point motion under the gather-and-spread model would be $O(\log m)$ if a ζ^* exists. The difficulty in estimating the exact constant α is the boundary of the square table. The longer the cards stay at the boundary, the slower is the mixing. To do a finer analysis and working towards an invariance principle, we take a diffusion limit of the m-point motion. The jumps in the gathering are made rare and the spreads are made small and frequent. As we show in the next section, the sequence of processes converges to a jump-diffusion limit. This jump-diffusion spends a negligible amount of time at the boundary (a Lebesgue null set), and allows us to do a more precise estimate of the constants.

3. The jump-diffusion limit of the gather-and-spread model

It is hard to get a precise estimate on the missing constant in the $O(\log m)$ mixing time bound above. In this section we treat a continuum jump-diffusion limit on the m point motion of the lazy gather-and-spread model (4) to estimate that constant by explicit computations facilitated by stochastic calculus.

Consider a sequence of discrete gather-and-spread models with a corresponding sequence of parameters $\lambda^{(n)} = n$ and $s_0^{(n)} = 1/\sqrt{n}$, as $n \in \mathbb{N}$ tends to infinity while keeping every other parameter (U, ν_0, $0 < p < 1$, initial values etc.) fixed.

Recall the modification to Definition 1 as given in (4). Let $(t_i, \theta_i, w_i,\ \ i \in \mathbb{N})$ and $\left(H_j(i),\ j \in [m],\ i \in \mathbb{N}\right)$ be as in Definition 1. Starting with $t_0 = 0$, at each t_i, toss a coin with probability of heads given by $1/\lambda = 1/n$. If the coin turns heads, define $Z_j^{(n)}(t_i)$ for every $j \in [m]$ according to (2). Otherwise, define

$$Z_j^{(n)}(t_i) := H_j(i) f_{1/\sqrt{n}}^{w_i,\theta_i}\left(Z_j^{(n)}(t_{i-1})\right) + \left(1 - H_j(i)\right) Z_j^{(n)}(t_{i-1}),\ i \in \mathbb{N},\ j \in [m]. \tag{9}$$

Extending to all time periods as in (3) gives us a sequence of m point motions

$$Z^{(n)} := \left(Z_j^{(n)}(t) = \left(X_j^{(n)}(t), Y_j^{(n)}(t)\right),\ j \in [m],\ t \geq 0\right),$$

We are interested in the limit of this process as n tends to infinity.

Let $\varrho^{(n)}$ denote the first time we gather, i.e., the first time a coin with probability $1/n$ turns up heads. By the Poisson thinning property, this is distributed as a rate one exponential random variable (irrespective of n) and is independent of the process $Z^{(n)}(t)$, $0 \leq t < \varrho^{(n)}$.

Consider the joint law of $\varrho^{(n)}$ and $Z^{(n)}(t)$, $0 \leq t < \varrho^{(n)}$. As $n \to \infty$, we will show that the process $Z^{(n)}$ converges in law in the usual Skorokhod space to a continuous diffusion which is strong Markov, stopped at an independent exponential one time. This is enough for our purpose since by restarting the diffusion from a different initial condition (as dictated by the gather) at this random time gives us a limiting jump-diffusion.

In order to describe the limiting (unstopped) diffusion we abuse our notations and assume that $Z^{(n)}$ is updated at every time t_i by (9) (without the gather) and take a diffusion limit. By an abuse of notation we continue to refer to this process (without the gather) in this section by $Z^{(n)}$ while keeping in mind that the process is observed only till an independent exponential time.

Notice the following properties of the trajectories of cards under the spread moves (and no gather). Because of Assumption 1, the increments of every card has mean zero, if it stays within the interior of the table after the spread. Also, if two cards are under the palm and they both decide to move with the palm, then their increments are positively correlated. However, if they are not under the palm, then the increment in one is independent of the other. This is captured by the fact that the limiting diffusion has zero drift (in the interior of the square) and a diffusion matrix given below.

Define the 2×2 positive-definite matrix

$$\Sigma = \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix} = \begin{bmatrix} \nu_0 \left(\sin^2(\theta)\right) & \nu_0 \left(\sin(\theta)\cos(\theta)\right) \\ \nu_0 \left(\sin(\theta)\cos(\theta)\right) & \nu_0 \left(\cos^2(\theta)\right) \end{bmatrix}.$$

Note Area $(U) = \pi\delta^2$. Let F be the $m \times m$ symmetric, positive definite matrix given by

$$F_{ij}(z_1, \ldots, z_m) = \begin{cases} p^2 \text{Area}\left(\{z_i + U\} \cap \{z_j + U\}\right), & i, j \in [m],\ i \neq j. \\ p\text{Area}\left(\{z_i + U\}\right) = p\pi\delta^2, & i = j \in [m]. \end{cases} \tag{10}$$

We skip the proof of the following elementary fact.

Lemma 3. *Let z_1, z_2 be two arbitrary points on the plane. Then*

$$\text{Area}\left(\{z_1 + U\} \cap \{z_2 + U\}\right) = \varphi\left(\|z_1 - z_2\|\right)$$

where $\varphi : [0, \infty) \to [0, \infty)$ is given by

$$\varphi(r) := \begin{cases} 2\delta^2 \arccos\left(\frac{r}{2\delta}\right) - \frac{r}{2}\sqrt{4\delta^2 - r^2}, & \text{for } r \leq 2\delta, \\ 0, & \text{otherwise}. \end{cases}$$

In particular, φ is a decreasing convex function on $(0, \infty)$.

Let $B(z_1, \ldots, z_m)$ be the $2m \times 2m$ matrix Kronecker product $F \otimes \Sigma$. See [22, Definition 4.2.1] for the definition. Here and throughout we label the rows (and columns) of the diffusion matrix by assigning the $(2i-1)$th row to x_i and $2i$th row to y_i, successively for $i = 1, 2, \ldots, m$. In particular, the block of B corresponding to $z_i = (x_i, y_i)$ is a 2×2 matrix given by $p\pi\delta^2\Sigma$; the block corresponding to $(z_i = (x_i, y_i), z_j = (x_j, y_j))$th is a 4×4 matrix given by

$$\begin{bmatrix} p\pi\delta^2\sigma^2 & 0 & F_{ij}\sigma^2 & 0 \\ 0 & p\pi\delta^2\sigma^2 & 0 & F_{ij}\sigma^2 \\ F_{ij}\sigma^2 & 0 & p\pi\delta^2\sigma^2 & 0 \\ 0 & F_{ij}\sigma^2 & 0 & p\pi\delta^2\sigma^2 \end{bmatrix}. \tag{11}$$

By [22, Theorem 4.2.12], the eigenvalues of B are pairwise products of those of F and Σ. Hence, $B(z_1, \ldots, z_m)$, which is symmetric, is also nonnegative definite for any $(z_1, \ldots, z_m)$. We show later in Lemma 8 that this matrix is uniformly positive definite. Let $A(z_1, \ldots, z_m)$ denote the unique positive definite square-root of $B(z_1, \ldots, z_m)$. Let $A_{X_j}(z_1, \ldots, z_m)$ and $A_{Y_j}(z_1, \ldots, z_m)$ denote the row of $A(z_1, \ldots, z_m)$ corresponding to coordinate x_j and y_j, respectively, for $j \in [m]$. Thus, according to our convention, A_{X_j} is the $(2j-1)$th row and A_{Y_j} is the $(2j)$th row of A.

Definition 2. Fix points $z_j = (x_j, y_j) \in [0,1]^2$, $j \in [m]$. Let $\mathcal{R}_m(z_1, \ldots, z_m)$ denote the law of a time-homogeneous diffusion in $[0,1]^{2m}$ with zero drift, diffusion matrix B, normal reflection at the boundary, and initial conditions $Z_j(0) = z_j = (x_j, y_j)$, $j \in [m]$. The multidimensional vector-valued process $\left(Z_j(\cdot) = (X_j(\cdot), Y_j(\cdot))\right)$, $j \in [m]$, satisfies the stochastic differential equation (SDE):

$$
\begin{aligned}
X_j(t) &= x_j + \int_0^t A_{X_j}(Z_1(s), \ldots, Z_m(s)) \cdot d\beta(s) + L_j^{X,0}(t) - L_j^{X,1}(t),\\
Y_j(t) &= y_j + \int_0^t A_{Y_j}(Z_1(s), \ldots, Z_m(s)) \cdot d\beta(s) + L_j^{Y,0}(t) - L_j^{Y,1}(t).
\end{aligned}
\tag{12}
$$

Here $\beta = (\beta_1, \ldots, \beta_{2m})$ is a $2m$ dimensional Brownian motion. The symbol $\int A_\cdot \cdot d\beta$ refers to the multidimensional Itô stochastic integral with respect to the Brownian motion β. The processes $L_j^{X,0}(\cdot)$ and $L_j^{X,1}(\cdot)$ are the accumulated local times for the process X_j at zero and one, respectively. The processes $L_j^{Y,0}(\cdot)$ and $L_j^{Y,1}(\cdot)$ are similarly defined. See [26, Section 3.7] for the normalization factor of local times.

We show in Theorem 7 that, for $0 < p < 1$, the SDE (12) has a pathwise unique strong solution which is strong Markov. In particular, there is uniqueness in law and every solution is strong. The following is our main convergence result. Let $\mathcal{D}^{(2m)}[0,\infty)$ be the usual Skorokhod space of RCLL paths from $[0,\infty)$ to $\mathbb{R}^{2m}$. Unless otherwise mentioned, we work with the stronger locally uniform topology on this space. See [11, Section 15 and 16] for details on the Skorokhod space and the locally uniform and other topologies on it. This is for convenience. Since our limiting processes are continuous almost surely, the convergence with respect to the usual Skorokhod topology is equivalent to convergence in the locally uniform topology.

Theorem 4 (*Without Gather*)**.** *Fix U, $p \in (0,1)$, and ν_0. Fix an arbitrary set $\{z_j,\ j \in [m]\}$ in $[0,1]^2$. Let $\left(Z_j^{(n)}(\cdot),\ j \in [m]\right)$ denote the m-point motion given in (9) when $\lambda^{(n)} = n$ and $s_0^{(n)} = 1/\sqrt{n}$ starting with $Z_j^{(n)}(0) = z_j$ for $j \in [m]$. Then, as n tends to infinity, the sequence $\left(Z_j^{(n)}(t),\ t \geq 0,\ j \in [m]\right)$, $n \in \mathbb{N}$, converges in law in $\mathcal{D}^{(2m)}[0,\infty)$ in the locally uniform topology to $\left(Z_j(t),\ t \geq 0,\ j \in [m]\right)$ that is a solution of (12).*

That the limiting process should have zero drift and diffusion coefficients given by the matrix B is easy to guess by computing the mean and the covariance of the increments of the discrete model. The appearance of local time is the consequence of the boundary behavior of our model and this is where it is critical that we use the $x \mapsto \widetilde{x}$ function in definition (1).

Theorem 4 is proved in several steps below. We start with $m = 1$.

Lemma 5. *Let $Z_1^{(n)}(t) = \left(X_1^{(n)}(t), Y_1^{(n)}(t)\right)$, $t \geq 0$, denote the one point motion given in (9) (without gather) with $\lambda^{(n)} = n$ and $s_0^{(n)} = 1/\sqrt{n}$ and given initial condition $Z_1^{(n)}(0) = (x, y) \in [0,1]^2$. Then, as $n \to \infty$, $\left(\left(X_1^{(n)}(t), Y_1^{(n)}(t)\right), t \geq 0\right)$ converges in law to a pair $((X_1(t), Y_1(t)),\ t \geq 0)$ of independent doubly reflected Brownian motion (RBM) in the interval $[0,1]$ with zero drift and constant diffusion coefficient $p\pi\delta^2\sigma^2$, starting at (x, y). In other words, (X_1, Y_1) satisfies the SDE*

$$
\begin{aligned}
X_1(t) &= x_1 + \sigma\delta\sqrt{p\pi}\, W_1(t) + L_1^{X,0}(t) - L_1^{X,1}(t),\\
Y_1(t) &= y_1 + \sigma\delta\sqrt{p\pi}\, W_2(t) + L_1^{Y,0}(t) - L_1^{Y,1}(t),
\end{aligned}
\tag{13}
$$

where (W_1, W_2) is a pair of independent standard one-dimensional Brownian motions, $L_1^{X,0}(t)$, $L_1^{X,1}(t)$ are the accumulated local times at zero and one, respectively, till time t for the semimartingale X_1, and $L_1^{Y,0}(t)$, $L_1^{Y,1}(t)$ are similarly defined.

The existence, uniqueness of reflected process and corresponding discrete to continuum convergence problems are best handled by the established tools of Skorokhod problems and Skorokhod maps. Skorokhod maps transform an "un-reflected" (or unconstrained) process to a "reflected" (or constrained) one inside a domain via a deterministic transform on the RCLL path-space. The existence of such a deterministic transform is called a Skorokhod problem. Once it is established that such a map exists, i.e. the Skorokhod problem has a solution, then existence, uniqueness, convergence problems can be handled "pre-reflection" by the usual martingale methods, and let the deterministic transform take care of the rest. See the exposition in [27] where the reader can find more details.

To simplify the notation denote $\mathcal{D}^{(1)}[0,\infty)$, the Skorokhod space of RCLL functions from $[0,\infty)$ to $\mathbb{R}$, by $\mathcal{D}[0,\infty)$. Let $\mathcal{BV}[0,\infty)$ and $\mathcal{I}[0,\infty)$ denote the subsets of $\mathcal{D}[0,\infty)$ comprised of functions of bounded variations and nondecreasing functions, respectively.

Definition 3 (*Skorokhod Map on* $[0,1]$)**.** Given $\psi \in \mathcal{D}[0,\infty)$, there exists a unique pair of functions $\left(\bar{\phi}, \bar{\eta}\right) \in \mathcal{D}[0,\infty) \times \mathcal{BV}[0,\infty)$ that satisfy the following two properties:

(i) For every $t \geq 0$, $\bar{\phi}(t) = \psi(t) + \bar{\eta}(t) \in [0,1]$.

(ii) $\bar{\eta}(0-) = 0$ and $\bar{\eta}$ has the decomposition $\bar{\eta} = \bar{\eta}_l - \bar{\eta}_u$ as the difference of functions $\bar{\eta}_l, \bar{\eta}_u \in \mathcal{I}[0,\infty)$ satisfying the so-called complementarity conditions:

$$\int_0^\infty 1\left\{\bar{\phi}(s) > 0\right\} d\bar{\eta}_l(s) = 0 \quad \text{and} \quad \int_0^\infty 1\left\{\bar{\phi}(s) < 1\right\} d\bar{\eta}_u(s) = 0. \tag{14}$$

Here $\bar{\eta}(0-) = 0$ means that, if $\bar{\eta}(0) > 0$, then $d\bar{\eta}$ has an atom at zero. We refer to the map $\Gamma_{0,1} : \mathcal{D}[0,\infty) \to \mathcal{D}[0,\infty)$ that takes ψ to $\bar{\phi}$ as the Skorokhod map on $[0,1]$. The pair $\left(\bar{\phi}, \bar{\eta}\right)$ is said to solve the Skorokhod problem on $[0,1]$ with input ψ.

The existence and uniqueness of Skorokhod map over general domains is a classical topic. See, for example, Tanaka [44]. Let $x^+ = \max(x,0)$ for $x \in \mathbb{R}$. On $[0,1]$ the map has an explicit solution. In Theorem 1.4 of [27] it is shown that $\Gamma_{0,1} = \Lambda_1 \circ \Gamma_0$, where

$$\begin{aligned} &\Gamma_0(\psi)(t) = \psi(t) + \sup_{0\le s\le t}\,[-\psi(s)]^+ \quad \text{and} \\ &\Lambda_1(\phi)(t) = \phi(t) - \sup_{0\le s\le t}\left[(\phi(s)-1)^+ \wedge \inf_{s\le u\le t}\phi(u)\right]. \end{aligned} \tag{15}$$

In particular, both Γ_0 and $\Gamma_{0,1}$ are Lipschitz with respect to the (locally) uniform and the Skorokhod J_1 metric on $\mathcal{D}[0,\infty)$.

Proof of Lemma 5. Fix $n \in \mathbb{N}$. Let χ denote a PPP on $(0,\infty)\times\overline{D}$ with rate $\lambda^{(n)} = n$. Evaluate the atoms of the PPP as a sequence $\{(t_i, w_i),\ i \in \mathbb{N}\}$ where t_i is increasing with i. Recall the i.i.d. sequence $(\theta_i,\ i \in \mathbb{N})$ sampled from ν_0 and an independent i.i.d. sequence of Bernoulli(p) random variables $(H_i,\ i \in \mathbb{N})$ from Definition 1 (where we have substituted the notation H_i for $H_1(i)$). Define $Z_1^{(n)}(\cdot)$ as in (9).

For $t \geq 0$, let

$$N(t) := \int_{[0,t]\times\overline{D}} 1\left\{Z_1^{(n)}(s-) \in w + U\right\} d\chi(s,w) = \sum_{j:\, t_j \leq t} 1\left\{\left\| Z_1^{(n)}(t_j-) - w_j \right\| \leq \delta\right\}.$$

Then $N(t)$ counts the number of times the point $Z_1^{(n)}(\cdot-)$ is "under the palm" during time interval $[0,t]$. By the symmetry of the norm $\|\cdot\|$ and the spatial homogeneity of the PPP, if the current position of the card is $z_1^{(n)}$, the first time when it is under the palm is an exponential random variable with rate $n\mathrm{Area}\left(z_1^{(n)} + U\right) = n\pi\delta^2$, independent of the past. Thus, $(N(t),\ t \geq 0)$ is a Poisson process with rate $n\pi\delta^2$.

Mark each jump time t_i of N with the corresponding θ_i and H_i. For $j \in \mathbb{N}$, define

$$M^{X,n}(t_j) = \frac{1}{\sqrt{n}} \sum_{i=1}^{j} H_i \cos(\theta_i). \tag{16}$$

Extend to other values of t by defining that if $t_{j-1} \leq t < t_j$ for some $j \in \mathbb{N}$, then

$$M^{X,n}(t) = M^{X,n}(t_{j-1}). \tag{17}$$

Note that $M^{X,n}$ is a martingale since $\nu_0(\cos(\theta)) = 0$ and H_i is independent of θ_i.

Now suppose that $\cos(\theta_j) \leq 0$. Then the difference $X_1^{(n)}(t_j) - X_1^{(n)}(t_{j-1})$ is given by

$$\begin{cases} \dfrac{1}{\sqrt{n}} H_j \cos(\theta_j), & \text{if } X_1^{(n)}(t_{j-1}) + \dfrac{1}{\sqrt{n}} H_j \cos(\theta_j) > 0, \\ -X_1^{(n)}(t_{j-1}), & \text{if } X_1^{(n)}(t_{j-1}) + \dfrac{1}{\sqrt{n}} H_j \cos(\theta_j) \leq 0. \end{cases}$$

We can express this differently as

$$X_1^{(n)}(t_j) - X_1^{(n)}(t_{j-1}) = \frac{1}{\sqrt{n}} H_j \cos(\theta_j) + \left(X_1^{(n)}(t_{j-1}) + \frac{1}{\sqrt{n}} H_j \cos(\theta_j) \right)^-,$$

where $x^- := \max(-x, 0) \geq 0$.

Similarly, when $\cos(\theta_j) > 0$, $X_1^{(n)}(t_j) - X_1^{(n)}(t_{j-1})$ is given by

$$\begin{cases} \dfrac{1}{\sqrt{n}} H_j \cos(\theta_j), & \text{if } X_1^{(n)}(t_{j-1}) + \dfrac{1}{\sqrt{n}} H_j \cos(\theta_j) < 1, \\ 1 - X_1^{(n)}(t_{j-1}), & \text{if } X_1^{(n)}(t_{j-1}) + \dfrac{1}{\sqrt{n}} H_j \cos(\theta_j) \geq 1. \end{cases}$$

Hence,

$$X_1^{(n)}(t_j) - X_1^{(n)}(t_{j-1}) = \frac{1}{\sqrt{n}} H_j \cos(\theta_j) - \left(1 - X_1^{(n)}(t_{j-1}) - \frac{1}{\sqrt{n}} H_j \cos(\theta_j) \right)^-.$$

Combining the two cases note that we can always write

$$X_1^{(n)}(t_j) - X_1^{(n)}(t_{j-1}) = \frac{1}{\sqrt{n}} H_j \cos(\theta_j) + \left(X_1^{(n)}(t_{j-1}) + \frac{1}{\sqrt{n}} H_j \cos(\theta_j) \right)^- - \left(1 - X_1^{(n)}(t_{j-1}) - \frac{1}{\sqrt{n}} H_j \cos(\theta_j) \right)^-.$$

Define the following pair of increasing functions, both starting at zero:

$$I^{X,0,n}(t_j) = I^{X,0,n}(t_{j-1}) + \left(X_1^{(n)}(t_{j-1}) + \frac{1}{\sqrt{n}} H_j \cos(\theta_j) \right)^-$$
$$I^{X,1,n}(t_j) = I^{X,1,n}(t_{j-1}) + \left(1 - X_1^{(n)}(t_{j-1}) - \frac{1}{\sqrt{n}} H_j \cos(\theta_j) \right)^- . \tag{18}$$

Extend them to all values of $t \in [0, \infty)$ by defining $I^{X,0,n}(t) = I^{X,0,n}(t_{j-1})$ and $I^{X,1,n}(t) = I^{X,1,n}(t_{j-1})$, for all $t \in [t_{j-1}, t_j)$. Note that the jumps of the process $I^{X,0,n}$ occur at those t_j such that $X_1^{(n)}(t_j) = 0$ and $\cos(\theta_j) < 0$, $H_j = 1$ while the jumps of $I^{X,1,n}$ occur at those t_j such that $X_1^{(n)}(t_j) = 1$ and $\cos(\theta_j) > 0$, $H_j = 1$.

Recall the martingale $M^{X,n}$ from (16). From here it is not hard to see that $X_1^{(n)}(\cdot)$ is the solution of the following system of pathwise Eqs. (18) and

$$X_1^{(n)}(t) = x_1 + M^{X,n}(t) + \int_0^t 1\left\{X_1^{(n)}(s) = 0\right\} dI^{X,0,n}(s) - \int_0^t 1\left\{X_1^{(n)}(s) = 1\right\} dI^{X,1,n}(s) .$$

This is an expression that satisfies the Skorokhod problem decomposition given in Definition 3. The process $X_1^{(n)}$ is constrained to stay in $[0, 1]$, $M^{X,n}$ is RCLL, while

$$\bar{\eta}_l := \int_0^t 1\left\{X_1^{(n)}(s) = 0\right\} dI^{X,0,n}(s) , \quad \bar{\eta}_u := \int_0^t 1\left\{X_1^{(n)}(s) = 1\right\} dI^{X,1,n}(s)$$

are increasing and obviously satisfy the complementarity conditions (14). Hence

$$\left(X_1^{(n)}, I^{X,0,n} - I^{X,1,n} \right)$$

is the unique solution of the Skorokhod problem on $[0, 1]$ with input $x + M^{X,n}(\cdot)$.

Now take limits as n tends to infinity. It follows from Donsker's invariance principle that the continuous time martingale $M^{X,n}$ converges to $\sigma\delta\sqrt{p\pi} W_1$, where W_1 is a standard Brownian motion. This is because $M^{X,n}$ is a continuous time centered random walk that jumps at rate $n\pi\delta^2$ and the variance of its increments is $p\sigma^2/n$. By the Lipschitz continuity of the deterministic Skorokhod map, it immediately follows that the vector of processes

$$\left(M^{X,n}, X_1^{(n)}, I^{X,0,n}, I^{X,1,n} \right)$$

jointly converges in law to the vector of $\sigma\delta\sqrt{p\pi} W_1$ and the corresponding terms in the solution of the Skorokhod problem in $[0, 1]$ with input $x_1 + \sigma\delta\sqrt{p\pi} W_1$.

Let us now identify the limit as reflecting Brownian motion in the interval $[0, 1]$ with constant diffusion coefficient $\sigma^2 p\pi\delta^2$. The limit, say X_1, satisfies the SDE given by the Skorokhod equation:

$$X_1(t) = x_1 + \sigma\delta\sqrt{p\pi} W_1(t) + L^{X,0}(t) - L^{X,1}(t),$$

where $L^{X,1}$ and $L^{X,0}$ are outputs from the Skorokhod problem with input $x + \sigma\delta\sqrt{p\pi} W_1(t)$. To identify $L^{X,1}$ and $L^{X,0}$ with the local time of the process X_1 at the boundary zero and one, respectively, we apply the Tanaka ([26, page 220]) formula to the semimartingale X_1 for the functions $x \mapsto x^+$ and $x \mapsto (1 - x)^+$.

For the y-coordinate process repeat the above argument except that $\cos(\theta_j)$ will be replaced by $\sin(\theta_j)$. That is, define

$$M^{Y,n}(t_j) = \frac{1}{\sqrt{n}} \sum_{i=1}^{j} H_i \sin(\theta_i). \tag{19}$$

Extend to other values of t by keeping the process constant in each interval $[t_{j-1}, t_j)$, $j \in \mathbb{N}$. $M^{Y,n}$ is also a martingale since $\nu_0(\sin(\theta)) = 0$.

The naturally defined corresponding processes for the y-coordinate

$$\left(M^{Y,n}, Y_1^{(n)}, I^{Y,0,n}, I^{Y,1,n}\right)$$

jointly converges in law to the vector of a Brownian motion $\sigma\delta\sqrt{p\pi}W_2$ and the corresponding terms in the solution of the Skorokhod problem in [0, 1] with input $y + \sigma\delta\sqrt{p\pi}W_2$. We need to argue joint convergence of the vector

$$\left(M^{X,n}, X_1^{(n)}, I^{X,0,n}, I^{X,1,n}, M^{Y,n}, Y_1^{(n)}, I^{Y,0,n}, I^{Y,1,n}\right).$$

However, this will follow from the joint convergence of the pair $\left(M^{X,n}, M^{Y,n}\right)$ since everything else is a deterministic Lipschitz function applied to this pair of processes.

We first claim that $\left(M^{X,n}(t)M^{Y,n}(t),\ t \geq 0\right)$ is also a martingale. Since each process individually is a process of identically distributed independent increments, it suffices to check that the increments are uncorrelated. However, that is guaranteed by Assumption 1 that $\nu_0\left(\cos(\theta)\sin(\theta)\right) = 0$.

Now, by marginal convergence, it follows that the sequence of laws of the pair of processes $\left((p\pi\sigma^2\delta^2)^{-1/2}M^{X,n}, (p\pi\sigma^2\delta^2)^{-1/2}M^{Y,n}\right)$ in $\mathcal{D}^{(2)}[0,\infty)$ is tight in the locally uniform metric and that any limiting processes W_1, W_2 are marginally Brownian motions that additionally satisfy W_1W_2 is a local martingale. It follows by Knight's theorem (see [26, page 179]) that W_1, W_2 must be a pair of independent Brownian motions. Since X_1 and Y_1 are outputs of the deterministic Skorokhod map applied to W_1 and W_2, they too are independent. This completes the proof. □

Proof of Theorem 4. This proof is a generalization of the proof of Lemma 5. As in that proof, for every $n \in \mathbb{N}$, $j \in [m]$, and $i \in \{0, 1\}$, define the quantities

$$X_j^{(n)}, Y_j^{(n)}, M_j^{X,n}, M_j^{Y,n}, I_j^{X,i,n}, I_j^{Y,i,n}.$$

Then, for each $j \in [m]$, the vector $\left(X_j^{(n)}, Y_j^{(n)}\right)$ can be expressed as the solution of a system of Skorokhod equations in [0, 1] with given inputs

$$x_j + M_j^{X,n} \quad \text{and} \quad y_j + M_j^{Y,n}(\cdot), \quad \text{respectively.} \tag{20}$$

The strategy is now the following. Consider the vector of $4m$ many processes obtained by concatenating $\left(X_j^{(n)}, Y_j^{(n)},\ j \in [m]\right)$ with the $2m$ many inputs in (20). Each coordinate process is tight by Lemma 5 and has an almost sure continuous limit. Hence the joint law of these $4m$ processes is tight in $\mathcal{D}^{(4m)}[0,\infty)$, with the locally uniform metric, and any weak limit is a probability measure on $\mathcal{C}^{(4m)}[0,\infty)$. The latter is the space of all continuous functions from $[0,\infty)$ to $\mathbb{R}^{4m}$ equipped with the locally uniform metric.

Let the $4m$ dimensional vector

$$\left(X_j, Y_j, x_j + M_j^X, y_j + M_j^Y,\ \ j \in [m]\right) \tag{21}$$

denote a process whose law is any weak limit of the sequence of processes

$$\left(X_j^{(n)}, Y_j^{(n)}, x_j + M_j^{X,n}, y_j + M_j^{Y,n},\ \ j \in [m]\right), \quad n \in \mathbb{N}. \tag{22}$$

To prove the existence of the limiting SDE representation, it is therefore enough to argue that the vector of martingales $\left(M_j^X, M_j^Y,\ \ j \in [m]\right)$ has a stochastic integral representation as the local martingale component in (12). Once this is achieved, using the uniqueness in law of a process satisfying SDE (12) proved in Theorem 7, every weak limit must be the same and given by the solution of (12).

To carry this out carefully, start be expressing the processes $M_j^{X,n}, M_j^{Y,n}$, $j \in [m]$, as martingales with respect to natural filtrations. Fix $n \in \mathbb{N}$. Recall the PPP χ on $(0,\infty) \times \overline{D}$ from the beginning of the proof of Lemma 5. Extend the PPP by decorating each atom of χ by an independent vector of length $(m+1)$, $(\theta, H_1, \ldots, H_m)$ where we sample $\theta \sim \nu_0$, and $\left(H_j,\ j \in [m]\right)$ are i.i.d. Bernoulli(p) picks, independent of θ. This produces a PPP $\overline{\chi}$ on $[0,\infty) \times \overline{D} \times [0, 2\pi] \times \{0,1\}^m$. Choose a suitable probability space $(\Omega, \mathcal{F}_\infty, \mathcal{P})$ that supports $\overline{\chi}$. Let $(\mathcal{F}_t,\ t \geq 0)$ be the natural right continuous filtration generated by the process $\left(\overline{\chi}_t,\ t \geq 0\right)$ where $\overline{\chi}_t$ is the restriction of $\overline{\chi}$ to $[0,t] \times \overline{D} \times [0, 2\pi] \times \{0,1\}^m$. Note that, as opposed to Definition 1, in this proof we attach the random angle and Bernoulli variables whether or not there are cards "under the palm". They simply do not influence the motion of the cards unless the cards are under the palm. Enumerate the countably many atoms of $\overline{\chi}$ by $((t_i, w_i, \theta_i, H_1(i), \ldots, H_m(i)),\ i \in \mathbb{N})$ where $t_1 < t_2 < \cdots$. Then, on our sample space above we have the following expressions:

$$M_j^{X,n}(t) = \frac{1}{\sqrt{n}} \sum_{i: t_i \leq t} H_j(i) \cos(\theta_i) 1\left\{ \left\| Z_j^{(n)}(t_i-) - w_i \right\| \leq \delta \right\},$$

$$M_j^{Y,n}(t) = \frac{1}{\sqrt{n}} \sum_{i: t_i \leq t} H_j(i) \sin(\theta_i) 1\left\{ \left\| Z_j^{(n)}(t_i-) - w_i \right\| \leq \delta \right\}.$$

Here, as before, $Z_j^{(n)} = \left(X_j^{(n)}, Y_j^{(n)}\right)$.

Let $Z^{(n)}$ denote the $2m$ dimensional vector of $\left(Z_j^{(n)},\ j \in [m]\right)$. Recall the $2m \times 2m$ dimensional matrix $B(z_1, \ldots, z_m)$ from (12). For $z = (z_1 = (x_1, y_1), \ldots, z_m = (x_m, y_m)) \in \mathbb{R}^{2m}$, label the elements of B by $B_{x_j, x_k}(z)$, $B_{x_j, y_k}(z)$, or $B_{y_j, y_k}(z)$, for $j, k \in [m]$, by a natural correspondence.

Lemma 6. *In the filtered probability space described above each $M_j^{X,n}$ and $M_j^{Y,n}$ is an $(\mathcal{F}_t)$ martingale. Moreover, for all $(j,k) \in [m]^2$, the following processes are also $(\mathcal{F}_t)$ martingales:*

$$\begin{aligned}
\xi_{j,k}^{X,X,n}(t) &:= M_j^{X,n}(t) M_k^{X,n}(t) - \int_0^t B_{x_j, x_k}\left(Z^{(n)}(s)\right) ds. \\
\xi_{j,k}^{X,Y,n}(t) &:= M_j^{X,n}(t) M_k^{Y,n}(t) - \int_0^t B_{x_j, y_k}\left(Z^{(n)}(s)\right) ds. \\
\xi_{j,k}^{Y,Y,n}(t) &:= M_j^{Y,n}(t) M_k^{Y,n}(t) - \int_0^t B_{y_j, y_k}\left(Z^{(n)}(s)\right) ds.
\end{aligned} \tag{23}$$

Proof of Lemma 6. We start by arguing that $M_j^{X,n}$ and $M_j^{Y,n}$ are martingales. For every $\omega \in \Omega$, the processes $M_j^{X,n}$ and $M_j^{Y,n}$ are stochastic integrals of predictable integrands: for $\theta \in [0, 2\pi]$ and $h_j \in \{0, 1\}$, $j \in [m]$,

$$g_j^{X,n}(t,\omega) = \frac{1}{\sqrt{n}} h_j \cos(\theta) 1\left\{\left\| Z_j^{(n)}(t-) - w\right\| \le \delta\right\} \quad \text{and}$$
$$g_j^{Y,n}(t,\omega) = \frac{1}{\sqrt{n}} h_j \sin(\theta) 1\left\{\left\| Z_j^{(n)}(t-) - w\right\| \le \delta\right\}, \quad \omega \in \Omega,$$

with respect to the Poisson random measure $\overline{\chi}$ (see [24, Chapter II, Section 1]). Then, the claim follows from [24, Chapter II, Lemma 1.21], since, the predictable compensator of the processes are given by (respectively)

$$\frac{1}{\sqrt{n}} \int_0^t np\nu_0\left(\cos(\theta)\right) \mathrm{Area}\left(Z_j^{(n)}(s-) + U\right) ds \quad \text{and}$$
$$\frac{1}{\sqrt{n}} \int_0^t np\nu_0\left(\sin(\theta)\right) \mathrm{Area}\left(Z_j^{(n)}(s-) + U\right) ds.$$

Both expressions above are zero since $\nu_0(\cos\theta) = 0 = \nu_0(\sin\theta)$.

For the reader who might be uncomfortable with the stochastic calculus for Poisson processes, simply replace the Poisson process by a discrete time process with independent increments to derive the above conclusion "by hand". This is true for the argument below as well.

For the processes listed in (23), let us argue the martingale property of the first process in the display and leave the rest of the similar arguments for the reader. Consider the process $\xi_{j,k}^{X,X,n}$. Since $M_j^{X,n}(t)$ and $M_k^{X,n}(t)$ are both martingales, we simply need to argue that the predictable compensator for the product of the two process at time t is exactly $\int_0^t B_{x_j,x_k}\left(Z^{(n)}(s)\right) ds$. However, since $M_j^{X,n}(t)$ and $M_k^{X,n}(t)$ are both stochastic integrals of predictable integrands with respect to a Poisson random measure, the predictable compensator up to time t is given by the integral of the product of the integrands with respect to the intensity measure:

$$\int_0^t \left(p1\{j=k\} + p^2 1\{j \ne k\}\right) \sigma^2 \mathrm{Area}\left(\left\{Z_j^{(n)}(s-) + U\right\} \cap \left\{Z_k^{(n)}(s-) + U\right\}\right) ds.$$

The above is, of course, exactly equal to $\int_0^t B_{x_j,x_k}\left(Z^{(n)}(s)\right) ds$. $\square$

Returning to the proof of Theorem 4, recall that $\mathcal{C}^{(4m)}[0,\infty)$, the space of continuous functions from $[0,\infty)$ to $\mathbb{R}^{4m}$. Endow the space with a right-continuous natural filtration. We will use this as our sample space. Consider this sample space along with a probability measure that is any weak limit obtained from the joint weak convergence of the vector of processes in (22) to the processes in (21). Since the limit is continuous, we may assume that the weak convergence holds in the locally uniform norm on $\mathcal{C}^{(4m)}[0,\infty)$.

Since each $M_j^{X,n}$ is a stochastic integral of a mean-zero bounded predictable integrand with respect to a Poisson random measure, it is a martingale and any weak limit is necessarily also a continuous local martingale. We have already argued continuity above and the martingale property follows from weak convergence and boundedness of integrands. Furthermore, from (23) each of the following is a continuous stochastic process and a weak limit of a sequence of martingales.

$$\xi_{j,k}^{X,X}(t) := M_j^X(t)M_k^X(t) - \int_0^t B_{x_j,x_k}(Z(s))\,ds.$$
$$\xi_{j,k}^{X,Y}(t) := M_j^X(t)M_k^Y(t) - \int_0^t B_{x_j,y_k}(Z(s))\,ds.$$
$$\xi_{j,k}^{Y,Y}(t) := M_j^Y(t)M_k^Y(t) - \int_0^t B_{y_j,y_k}(Z(s))\,ds.$$

Here $Z_j(\cdot) = \big(X_j(\cdot), Y_j(\cdot)\big)$ and $Z(\cdot) = (Z_1(\cdot), \ldots, Z_m(\cdot))$. By a usual localization argument and the continuity of $B_{\cdot}$ they can all be seen as continuous local martingales. To wit, consider the convergence of $\xi_{j,k}^{X,X,n}$ to $\xi_{j,k}^{X,X}$, and consider a sequence of stopping times $\mathcal{T}_\ell$ that stops the process when its absolute value exceeds ℓ. Weak convergence and boundedness shows that each stopped process $\xi_{j,k}^{X,X}(\min(\cdot, \mathcal{T}_\ell))$ is a martingale. Take $\ell \to \infty$ to establish the local martingale property.

We now use [26, Chapter 3, Theorem 4.2] on the representation of continuous local martingales as stochastic integrals. According to this result, on a possibly extended probability space, one can find a $2m$ dimensional Brownian motion $(\beta_1, \ldots, \beta_{2m})$ such that for each $j \in [m]$ we have

$$M_j^X(t) = \int_0^t A_{X_j}(Z_1(s), \ldots, Z_m(s)) \cdot d\beta(s) \quad \text{and}$$
$$M_j^Y(t) = \int_0^t A_{Y_j}(Z_1(s), \ldots, Z_m(s)) \cdot d\beta(s).$$

This settles the local martingale component in the SDE representation (12). That the finite variation components are given by local times follow from Lemma 5. Finally, uniqueness in law from Theorem 7 completes the proof. □

Theorem 7. *Fix arbitrary initial points $z_1, \ldots, z_m$ in $[0, 1]^2$. Under Assumption 1 and when $p \in (0, 1)$, for any $m \geq 1$, there is a pathwise unique strong solution to the stochastic differential Eq. (12), starting at $(z_1, \ldots, z_m)$, under which the process is strong Markov. In particular, the law of such a solution is unique.*

The proof requires the following lemma.

Lemma 8. *The diffusion matrix $B(z_1, \ldots, z_m)$ is uniformly elliptic over $[0, 1]^{2m}$.*

Proof. Since $B = F \otimes \Sigma$ is the Kronecker product of F and Σ, the $2m$ eigenvalues of B are the pairwise product of the m eigenvalues of F and the two eigenvalues of Σ. Since the eigenvalues of Σ are both σ^2, they are both positive by Assumption 1. Therefore, to prove the lemma, it suffices to show uniform ellipticity for the matrix $F(z_1, \ldots, z_m)$.

Consider any $\xi = (\xi_1, \ldots, \xi_m) \in \mathbb{R}^m$ and note that

$$
\begin{aligned}
&\xi' F(z_1, \ldots, z_m)\xi \\
&\quad = \sum_{i\in[m]} \sum_{j\in[m]} p^2 \xi_i \xi_j \mathrm{Area}\left(\{z_i + U\} \cap \{z_j + U\}\right) + (p - p^2)\pi\delta^2 \sum_{i\in[m]} \xi_i^2 \\
&\quad = p^2 \int_{\mathbb{R}^2} \left[\sum_{i=1}^m \sum_{j=1}^m \xi_i \xi_j 1\{v \in z_i + U\} 1\{v \in z_j + U\} \right] dv + p(1-p)\pi\delta^2 \|\xi\|^2 \\
&\quad = p^2 \int_{\mathbb{R}^2} \left(\sum_{j=1}^m \xi_j 1\{v \in z_j + U\} \right)^2 dv + p(1-p)\pi\delta^2 \|\xi\|^2 \ge p(1-p)\pi\delta^2 \|\xi\|^2 .
\end{aligned}
$$

Since $p(1-p) > 0$ this proves uniform ellipticity. □

Proof of Theorem 7. We verify the assumptions of [37, Theorem 4.3] which has been proved for the so-called Extended Skorokhod Problem (ESP). In particular, it holds for the case of Skorokhod problems.

Our Skorokhod map is coordinatewise given by (15). Therefore, it is well-defined and Lipschitz. Therefore, it suffices to check Assumption 4.1 (1) in [37]. Since the drift is zero, we need to only check that the map $(z_1, \ldots, z_m) \mapsto A(z_1, \ldots, z_m)$, as a function on $[0, 1]^{2m}$, is Lipschitz. By [41, Lemma 21.10] and the uniform ellipticity condition from Lemma 8 it suffices to check that the map $(z_1, \ldots, z_m) \mapsto B(z_1, \ldots, z_m)$ is Lipschitz. This, in turn, follows from checking via Lemma 3 that the map $(z_1, \ldots, z_m) \mapsto F_{ij}(z_1, \ldots, z_m)$ for each (i, j) pair is Lipschitz which follows from the convexity of the function φ. □

4. Estimates on mixing time of shuffling for the jump diffusion

We now define the limiting lazy gather-and-spread model. Let us recall the diffusion model from Section 3. Consider a suitable probability space $\left(\Omega, (\mathcal{F}_t)_{t\ge 0}, P\right)$ with the usual filtration that supports $2m$ many standard linear Brownian motions $\left(\overline{\beta}_1, \ldots, \overline{\beta}_{2m}\right)$ and an independent PPP on $(0,\infty)\times\overline{D}$ with rate given by the product Lebesgue measure on $(0,\infty)$ and the uniform probability distribution on $\overline{D}$. That is the atoms of the PPP can be arranged as (t_i, w_i), $i \in \mathbb{N}$, where $0 < t_1 < t_2 < \cdots$ are the jumps of a Poisson process of rate one and the sequence $(w_i,\ i \in \mathbb{N})$ is i.i.d., sampled uniformly from $\overline{D}$, independently of $(t_i,\ i \in \mathbb{N})$.

Suppose the initial values $z_1 = (x_1, y_1), \ldots, z_m = (x_m, y_m)$. Let $t_0 = 0$ and define $Z(t) = (Z_1(t), \ldots, Z_m(t))$, $t \in [0, t_1)$, as the solution of SDE (12) with initial conditions $X_j(0) = x_j$ and $Y_j(0) = y_j$, $j \in [m]$, and the Brownian motions given by $\overline{\beta}_j$, $j \in [2m]$. Then, inductively, for $i = 1, 2, \ldots$, on $[t_i, t_{i+1})$, condition on $\mathcal{F}_{t_i}$, define initial conditions

$$
z_j(t_i) := G^{w_i}\left(Z_j(t_i-)\right), \quad j \in [m],
$$

and let $Z(t+t_i)$, $t \in [0, t_{i+1}-t_i)$ be the solution of SDE (12) with initial points $z_j(t_i)$, $j \in [m]$, and the Brownian motions given by $\beta_j^{(i)}(t) = \overline{\beta}_j(t_i+t) - \overline{\beta}_j(t_i)$. This gives us a jump-diffusion process $Z(t) = (Z_1(t), \ldots, Z_m(t))$, $t \in [0,\infty)$ with RCLL paths adapted to $\left(\Omega, (\mathcal{F}_t)_{t\ge 0}, P\right)$. The process is clearly strong Markov. Let $\mathcal{P}_m(z_1, \ldots, z_m)$ denote the law of the jump-diffusion process described above starting from the initial points $(z_1, \ldots, z_m) \in D^m$.

Lemma 9. *Assumption 2 in Section 2.2 holds for the jump-diffusion process Z.*

Proof. Assumption 2 is clearly true for the discrete gather-and-spread model without gathering as in the beginning of Section 3. Hence, by taking a weak limit, it is true for the diffusion satisfying (12). Stopping the diffusion at an independent exponential time and gathering at an independently chosen point in D preserves the property. Now, by iterating the argument, the statement is true for all intervals $[t_i, t_{i+1})$, $i \in \mathbb{N}$, and, therefore, over the entire $[0, \infty)$. □

Proof of Theorem 1. In order to employ Theorem 2, it suffices to find a ζ^* that stochastically dominates $\overline{\tau}_{xy} = \overline{\tau}^m_{xy}$ from (6) for any $x \neq y$. Without loss of generality, take $m = 2$.

The strategy is the following. Consider $\|Z_1(t) - Z_2(t)\|$ during $[0, t_1)$. Ignore the possibility that this norm hits zero during this interval. Consider

$$\{Z_1(t_1-) + U\} \cap \{Z_2(t_1-) + U\}\,.$$

The area of this set is given by Lemma 3. $Z_1(t_1) = Z_2(t_1)$ if the random point w_1 lies in this set (since the gathering will place both cards at w_1). In this event we get $\overline{\tau}_{12} \leq t_1$, otherwise we restart. Hence, at each t_i we toss a coin that indicates if $w_i \in \{Z_1(t_i-) + U\} \cap \{Z_2(t_i-) + U\}$. Suppose we bound this probability from below by $\mathfrak{p} > 0$ (say), irrespective of z_1, z_2. Then, by the strong Markov property, each such coin toss is independent, and $\overline{\tau}_{12}$ is stochastically dominated by t_J, where J is a geometric random variable with rate $\mathfrak{p}$. t_J is the sum of a random number of exponential one random variables. Such a random variable has a density everywhere on $[0, \infty)$, in particular, condition (8) is satisfied where α is simply the density at zero. Now, given $J = j$, t_J is a gamma random variable with mean j and scale 1. For all $j \geq 2$, this random variable has density zero at the origin. Given $J = 1$, t_J is exponential with rate one, which has a density one at the origin. Thus, the density at the origin of t_J is $1 \cdot P(J = 1) = \mathfrak{p}$. Thus, we can take $\alpha = \mathfrak{p}$ in (8). Hence, by Theorem 2, $t_{\mathrm{mix}}(\epsilon) = \frac{1}{\mathfrak{p}} \log(m/\epsilon)$.

What remains is to find such a $\mathfrak{p}$. This is done in the rest of this article. □

Consider any of the intervals $[t_i, t_{i+1})$, condition on $\mathcal{F}_{t_i}$ and shift time $t \mapsto t - t_i$, for $t \in [t_i, t_{i+1})$. During this interval the jump-diffusion Z is simply a diffusion stopped at an independent rate one exponential time $\varrho := t_{i+1} - t_i$. Thus, by the strong Markov property, we can assume that $i = 0 = t_i$ and $\varrho := t_1$ is an exponential one random variable, independent of the diffusion $Z(t)$, $t \in [0, \varrho)$.

Express this stopped diffusion $Z(\cdot) = (Z_1(\cdot), Z_2(\cdot))$, where $Z_1 = (X_1, Y_1)$, $Z_2 = (X_2, Y_2)$, which is a solution of (12), in the following way:

$$\begin{aligned}
X_1(t) &= W_1(t) + L_1^{X,0}(t) - L_1^{X,1}(t),\\
Y_1(t) &= B_1(t) + L_1^{Y,0}(t) - L_1^{Y,1}(t),\\
X_2(t) &= W_2(t) + L_2^{X,0}(t) - L_2^{X,1}(t),\\
Y_2(t) &= B_2(t) + L_2^{Y,0}(t) - L_2^{Y,1}(t).
\end{aligned}$$

Here (W_1, B_1, W_2, B_2) is a four-dimensional continuous semimartingale process such that each coordinate process is marginally distributed as a Brownian motion with constant diffusion coefficients $p\pi\delta^2\sigma^2$ and initial values x_1, y_1, x_2, y_2, respectively. But they are not all independent. This is because, in the discrete model, the cards that are under the palm tend to move together during the spread moves leading to a positive correlation between their increments, as compared to zero correlation when not under the palm.

Consider the two processes

$$\beta_1^0(t) := W_1(t) - W_2(t), \quad \beta_2^0(t) := B_1(t) - B_2(t).$$

Let $[X, Y](t)$ denote the mutual covariation between two continuous semimartingales X and Y over the time interval $[0, t]$. By (11),

$$[\beta_1^0, \beta_2^0](t) = 0, \quad [\beta_1^0, \beta_1^0](t) = [\beta_2^0, \beta_2^0](t) = 2\sigma^2 \int_0^t \left(p\pi\delta^2 - F_{12}\left(Z_1(s), Z_2(s)\right)\right) ds.$$

Let

$$\Gamma(t) = 2\sigma^2 \int_0^t \left(p\pi\delta^2 - F_{12}\left(Z_1(s), Z_2(s)\right)\right) ds.$$

Since $0 \le F_{12}(z_1, z_2) \le p^2 \text{Area}(U) = p^2\pi\delta^2$ for all $(z_1, z_2) \in D^2$, then

$$2\sigma^2 p\pi\delta^2 \ge \Gamma'(t) \ge 2\sigma^2 p(1-p)\pi\delta^2 > 0, \quad \text{and} \quad \Gamma(t) \ge 2\sigma^2 p(1-p)\pi\delta^2 t. \tag{24}$$

Hence Γ is strictly increasing and $\lim_{t\to\infty} \Gamma(t) = \infty$ almost surely. By Knight's Theorem (see [26, page 179]), on the same probability space, there exists a pair of independent standard Brownian motions β_1, β_2 such that

$$\beta_1^0(t) = \beta_1\left(\Gamma(t)\right), \quad \beta_2^0(t) = \beta_2\left(\Gamma(t)\right).$$

Let $R(t) = \sqrt{\beta_1^2(t) + \beta_2^2(t)}$. Then R is a 2-dimensional Bessel process starting from $\|z_1 - z_2\|$. Now, X_1 and X_2 are doubly reflected Brownian motions with continuous noises W_1 and W_2, respectively (as in Definition 3). The maps $W_1 \mapsto X_1$ and $W_2 \mapsto X_2$ are Lipschitz in the locally uniform metric on $\mathcal{C}[0, \infty)$. See [38, Remark 4.2 (ii)]. Therefore $(X_1(t) - X_2(t))^2 \le (W_1(t) - W_2(t))^2$. Similarly $(Y_1(t) - Y_2(t))^2 \le (B_1(t) - B_2(t))^2$. Hence

$$\|Z_1(t) - Z_2(t)\| \le \sqrt{\left(\beta_1^0(t)\right)^2 + \left(\beta_2^0(t)\right)^2} = R\left(\Gamma(t)\right), \quad 0 \le t < \varrho,$$

where ϱ is the independent exponential one random variable.

At ϱ pick w uniformly from $\overline{D}$ independently of the process $Z(t)$, $t \in [0, \varrho)$, and ϱ itself. Recall that U is the closed disc of radius δ around the origin. Consider the Bernoulli random variable

$$\chi_1 := 1\left\{w \in \{Z_1(\varrho-) + U\} \cap \{Z_2(\varrho-) + U\}\right\}. \tag{25}$$

Given $Z(t)$, $t \in [0, \varrho)$, the probability that $\{\chi_1 = 1\}$ is given by

$$\frac{1}{\text{Area}\left(\overline{D}\right)} \text{Area}\left(\{Z_1(\varrho-) + U\} \cap \{Z_2(\varrho-) + U\}\right) = \frac{1}{\text{Area}\left(\overline{D}\right)} \varphi\left(\|Z_1(\varrho-) - Z_2(\varrho-)\|\right),$$

by Lemma 3. Since φ, defined in Lemma 3, is decreasing the above expressions are bounded below by

$$\frac{1}{\text{Area}\left(\overline{D}\right)} \varphi\left(R(\Gamma(\varrho))\right). \tag{26}$$

Hence, integrating with respect to the law of $Z(t)$, $t \in [0, \varrho)$, and ϱ, we get

$$\begin{aligned}
\text{Area}\left(\overline{D}\right) \mathrm{E}\left(\chi_1\right) &\ge \mathrm{E}\left[\int_0^\infty e^{-t} \varphi\left(R(\Gamma(t))\right) dt\right] \\
&= \mathrm{E}\left[\int_0^\infty e^{-\Gamma^{-1}(u)} \varphi\left(R(u)\right) \left(\Gamma^{-1}(u)\right)' du\right], \quad u = \Gamma(t).
\end{aligned}$$

From the bounds in (24) we get

$$\frac{1}{2\sigma^2 p(1-p)\pi\delta^2} \ge \left(\Gamma^{-1}(u)\right)' \ge \frac{1}{2\sigma^2 p\pi\delta^2}, \quad \text{and} \quad \Gamma^{-1}(u) \le \frac{u}{2\sigma^2 p(1-p)\pi\delta^2}.$$

Hence

$$
\begin{aligned}
\operatorname{Area}\left(\overline{D}\right) \mathrm{E}\left(\chi_1\right) &\geq \frac{1}{2\sigma^2 p\pi\delta^2} \mathrm{E}\left[\int_0^\infty e^{-cu}\varphi(R(u))du\right], \quad c = \frac{1}{2\sigma^2 p(1-p)\pi\delta^2},\\
&= \frac{1}{2\sigma^2 p\pi\delta^2} \int_0^\infty e^{-cu} \mathrm{E}\left[\varphi(R(u))\right] du\\
&= \frac{1}{2\sigma^2 p\pi\delta^2} \int_0^\infty e^{-cu} Q^2_{\|z_1 - z_2\|}\left[\varphi(R(u))\right] du,
\end{aligned}
$$

where Q^2_x is the law of a two dimensional Bessel process starting from x.

Now $\|z_1 - z_2\| \leq \sqrt{2}$, the diameter of D. It follows from additivity of squared Bessel processes (see [42]) that the law of $R(u)$, under $Q^2_{\|z_1-z_2\|}$, is stochastically dominated by the law of $R(u)$, under $Q^2_{\sqrt{2}}$. Using the fact that φ is decreasing we get

$$
Q^2_{\|z_1-z_2\|}\left[\varphi\left(R(u)\right)\right] \geq Q^2_{\sqrt{2}}\left[\varphi\left(R(u)\right)\right].
$$

Combining all the bounds and noting that $\operatorname{Area}\left(\overline{D}\right) \leq (1+2\delta)^2$, we get

$$
\mathrm{E}\left(\chi_1\right) \geq \frac{1}{2\sigma^2 p\pi\delta^2(1+2\delta)^2} \int_0^\infty \exp\left(-\frac{u}{2\sigma^2 p(1-p)\pi\delta^2}\right) Q^2_{\sqrt{2}}\left[\varphi\left(R(u)\right)\right] du. \tag{27}
$$

Notice that this is a lower bound that is independent of the starting position of the diffusion.

To estimate the last expression we express it back in terms of planar Brownian motion. Let $v = (v_1, v_2) \in \mathbb{R}^2$ be arbitrary and, as before, let $c = \left(2\sigma^2 p(1-p)\pi\delta^2\right)^{-1}$. Let $V = (V_1, V_2)$ be a planar Brownian motion, starting from $v = (v_1, v_2)$. Then

$$
\int_0^\infty \exp\left(-\frac{u}{2\sigma^2 p(1-p)\pi\delta^2}\right) Q^2_{\|v\|}\left[\varphi\left(R(u)\right)\right] du = \int_0^\infty e^{-cu} \mathrm{E}_v\left[\varphi\left(\|V\|(u)\right)\right] du.
$$

The last expression is the c resolvent (sometimes called the c potential) operator for the generator of planar Brownian motion and is known explicitly. See [41, page 93]:

$$
\begin{aligned}
\int_0^\infty e^{-cu} \mathrm{E}_v\left[\varphi\left(\|V\|(u)\right)\right] du. &= \frac{1}{\pi} \int_{\mathbb{R}^2} K_0\left(\sqrt{2c}\,\|y\|\right) \varphi(\|v - y\|) dy\\
&= \frac{1}{\pi} \int_{z:\|z\|<2\delta} K_0\left(\sqrt{2c}\,\|v - z\|\right) \varphi(\|z\|) dz,
\end{aligned}
$$

where K_0 is the modified Bessel function of the second kind. The second equality above is due to the fact that $\varphi(r) = 0$ for $r \geq 2\delta$.

Now K_0 is a decreasing function of its argument (see [1, page 374]). Thus

$$
\frac{1}{\pi} \int_{z:\|z\|<2\delta} K_0\left(\sqrt{2c}\,\|v - z\|\right) \varphi(\|z\|) dz \geq K_0\left(\sqrt{2c}(\|v\| + 2\delta)\right) \frac{1}{\pi} \int_{z:\|z\|<2\delta} \varphi(\|z\|) dz.
$$

We now put $\|v\| = \sqrt{2}$ and evaluate the following integral:

$$
\begin{aligned}
&\frac{1}{\pi} \int_{z:\|z\|<2\delta} \varphi(\|z\|) dz = 2\int_0^{2\delta} r\varphi(r) dr\\
&= 4\delta^2 \int_0^{2\delta} r \arccos\left(\frac{r}{2\delta}\right) dr - \int_0^{2\delta} r^2\sqrt{4\delta^2 - r^2} dr\\
&= 16\delta^4 \left[\int_0^1 s \arccos\left(s\right) ds - \int_0^1 s^2\sqrt{1 - s^2} ds\right], \; s = \frac{r}{2\delta},\\
&= 16\delta^4 \left[\frac{\pi}{8} - \frac{\pi}{16}\right] = \pi\delta^4.
\end{aligned}
$$

Combining all the bounds,

$$\mathrm{E}\left(\chi_1\right) \geq \mathfrak{p} := \frac{\delta^2}{2p\pi\sigma^2(1+2\delta)^2} K_0\left(\frac{\sqrt{2}+2\delta}{\sigma\delta\sqrt{\pi p(1-p)}}\right). \tag{28}$$

Therefore, at $\varrho = t_1$, the probability that we gather cards 1 and 2 is at least $\mathfrak{p}$, irrespective of the starting positions of the two cards. This gives an upper bound on $\mathfrak{p}$ and completes the proof of Theorem 1.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

It is our pleasure to thank Lauren Bandklayder, Marc Coram, Emanuelle Gouillart, Kimberly Kinateder, Mark Perlman, and Graham White for their help with this project over the years. Many thanks to Keith Fife, Yuqi Huang, and Max Goering at the University of Washington for their help with the simulations. The paper has been improved significantly by multiple suggestions of our anonymous referees. We are grateful for their effort. The first author was supported in part by NSF grant DMS-1208775. The second author was supported in part by NSF grants DMS-1308340, DMS-1612483, and DMS-2052239.

References

[1] Milton Abramowitz, Irene A. Stegun, Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, in: National Bureau of Standards Applied Mathematics Series, vol. 55, For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C., 1964, p. xiv+1046.

[2] Gil Alon, Gady Kozma, The probability of long cycles in interchange processes, Duke Math. J. 162 (9) (2013) 1567–1585, http://dx.doi.org/10.1215/00127094-2266018.

[3] H. Aref, Stirring by chaotic advection, J. Fluid Mech. 143 (1984) 1–21.

[4] Sami Assaf, Persi Diaconis, K. Soundararajan, A rule of thumb for riffle shuffling, Ann. Appl. Probab. 21 (3) (2011) 843–875, http://dx.doi.org/10.1214/10-AAP701.

[5] P. Baxendale, T.E. Harris, Isotropic stochastic flows, Ann. Probab. 14 (1986) 1155–1179.

[6] Dave Bayer, Persi Diaconis, Trailing the dovetail shuffle to its lair, Ann. Appl. Probab. 2 (2) (1992) 294–313, http://dx.doi.org/10.1214/aoap/1177005705.

[7] N. Berestycki, G. Kozma, Cycle structure of the interchange process and representation theory, Bull. Soc. Math. France 143.2 (2015) 265–280.

[8] M. Bernstein, E. Nestoridi, Cutoff for random to random card shuffle, Ann. Probab. 47 (5) (2017).

[9] Pat Bidigare, Phil Hanlon, Dan Rockmore, A combinatorial description of the spectrum for the tsetlin library and its generalization to hyperplane arrangements, Duke Math. J. 99 (1) (1999) 135–174, http://dx.doi.org/10.1215/S0012-7094-99-09906-4.

[10] S. Billey, K. Burdzy, S. Pal, B.E. Sagan, On meteors, earthworms, and WIMPs, Ann. Appl. Probab. 25 (4) (2015) 1729–1779.

[11] Patrick Billingsley, Convergence of Probability Measures, Second Edition, John Wiley & Sons, Inc., ISBN: 9780470316962, 1999, http://dx.doi.org/10.1002/9780470316962.ch3.

[12] Kenneth S. Brown, Persi Diaconis, Random walks and hyperplane arrangements, Ann. Probab. 26 (4) (1998) 1813–1854, http://dx.doi.org/10.1214/aop/1022855884.

[13] Persi Diaconis, Steven N. Evans, Ron Graham, Unseparated pairs and fixed points in random permutations, Adv. Appl. Math. (ISSN: 0196-8858) 61 (2014) 102–124, http://dx.doi.org/10.1016/j.aam.2014.05.006.

[14] P. Diaconis, C.Y.A. Pang, A. Ram, Hopf algebras and Markov chains; two examples and a theory, J. Algebraic Combin. 39 (3) (2014) 527–585.

[15] Persi Diaconis, Laurent Saloff-Coste, Comparison techniques for random walk on finite groups, Ann. Probab. 21 (4) (1993) 2131–2156, http://dx.doi.org/10.1214/aop/1176989013.
[16] A.B. Dieker, F.V. Saliola, Spectra analysis of random-to-random Markov chains, Adv. Math. 323 (2018) 427–485.
[17] Keith Fife, Yuqi Huang, Randomly mixing fluids, 2018, Final report. Washington Experimental Math Lab. Available at.
[18] Emmanuelle Gouillart, Chaotic mixing by rod-stirring devices in open and closed flows, (Ph.D. thesis), Université Pierre et Marie Curie - Paris VI, 2007.
[19] E. Gouillart, O. Dauchot, B. Dubrulle, S. Roux, J.-L. Thiffeault, Slow decay of concentration variance due to no-slip walls in chaotic mixing, Phys. Rev. E 78 (2008) 026211, http://dx.doi.org/10.1103/PhysRevE.78.026211.
[20] E. Gouillart, N. Kuncio, O. Dauchot, B. Dubrulle, S. Roux, J.-L. Thiffeault, Walls inhibit chaotic mixing, Phys. Rev. Lett. 99 (2007) 114501, http://dx.doi.org/10.1103/PhysRevLett.99.114501.
[21] T.E. Harris, Brownian motions on the homeomorphisms of the plane, Ann. Probab. 9 (2) (1981) 232–254.
[22] R.A. Horn, C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, ISBN: 9780521467131, 1994.
[23] Chris Howitt, Jon Warren, Consistent families of Brownian motions and stochastic flows of kernels, Ann. Probab. 37 (4) (2009) 1237–1272.
[24] J. Jacod, A.N. Shiryaev, Limit Theorems for Stochastic Processes, Second Edition, in: A Series of Comprehensive Studies in Mathematics, vol. 288, Springer, 2002.
[25] Johan Jonasson, The overhand shuffle mixes in $\Theta(n^2 \log n)$ steps, Ann. Appl. Probab. 16 (1) (2006) 231–243, http://dx.doi.org/10.1214/105051605000000692.
[26] Ioannis Karatzas, Steven E. Shreve, Brownian Motion and Stochastic Calculus, Second, in: Graduate Texts in Mathematics, vol. 113, Springer-Verlag, New York, ISBN: 0-387-97655-8, 1991, p. xxiv+470, http://dx.doi.org/10.1007/978-1-4612-0949-2.
[27] Lukasz Kruk, John Lehoczky, Kavita Ramanan, Steven Shreve, An explicit formula for the skorokhod map on [0, a], Ann. Probab. 35 (5) (2007) 1740–1768, http://dx.doi.org/10.1214/009117906000000890.
[28] Hubert Lacoin, Mixing time and cutoff for the adjacent transposition shuffle and the simple exclusion, Ann. Probab. 44 (2) (2016) 1426–1487, http://dx.doi.org/10.1214/15-AOP1004.
[29] Y. Le Jan, On isotropic Brownian motions, Z. Wahrscheinlichkeitstheor. Verwandte Geb. 70 (1985) 609–620.
[30] Y. Le Jan, O. Raimond, Flows, coalescence and noise, Ann. Probab. 32 (2004) 1247–1315.
[31] D.A. Levin, Y. Peres, Markov Chains and Mixing Times. Second Edition., American Mathematical Society, ISBN: 9781470429621, 2017.
[32] L.L. Ng, Heisenberg Model, Bethe Ansatz, and Random Walks, Harvard University, 1996, Bachelor's thesis.
[33] E.L. Paul, V. Atiemo-Obeng, S.M. Kresta, Handbook of Industrial Mixing: Science and Practice, in: Handbook of Industrial Mixing: Science and Practice, (v. 1) John Wiley & Sons, ISBN: 9780471269199, 2003.
[34] Robin Pemantle, Randomization time for the overhand shuffle, J. Theoret. Probab. 2 (1) (1989) 37–49.
[35] Robin Pemantle, A shuffle that mixes sets of any fixed size much faster than it mixes the whole deck, Random Struct. Algorithms (ISSN: 1098-2418) 5 (5) (1994) 609–626, http://dx.doi.org/10.1002/rsa.3240050502.
[36] Henri Poincaré, Calcul des Probabilités, Gauthier-Villars, Paris, 1912.
[37] K. Ramanan, Reflected diffusions defined via the extended skorokhod map, EJP 11 (36) (2006) 934–992.
[38] Yasumasa Saisho, Stochastic differential equations for multi-dimensional domain with reflecting boundary, Probab. Theory Related Fields (ISSN: 1432-2064) 74 (3) (1987) 455–477, http://dx.doi.org/10.1007/BF00699100.
[39] Laurent Saloff-Coste, Probability on groups: random walks and invariant diffusions., Notices Amer. Math. Soc. 48 (2001) 968D977.
[40] Laurent Saloff-Coste, Random walks on finite groups, in: Harry Kesten (Ed.), Probability on Discrete Structures, Springer Berlin Heidelberg, Berlin, Heidelberg, ISBN: 978-3-662-09444-0, 2004, pp. 263–346, http://dx.doi.org/10.1007/978-3-662-09444-0_5.
[41] R.L. Schilling, L. Partzsch, B. Böttcher, Brownian Motion: An Introduction To Stochastic Processes, in: De Gruyter Textbook, De Gruyter, ISBN: 9783110307306, 2014.
[42] Tokuzo Shiga, Shinzo Watanabe, Bessel diffusions as a one-parameter family of diffusion processes, Z. Wahrscheinlichkeitstheor. Verw. Geb. 27 (1) (1973) 37–46.
[43] R. Sturman, J.M. Ottino, S. Wiggins, The Mathematical Foundations of Mixing: The Linked Twist Map As a Paradigm in Applications: Micro To Macro, Fluids To Solids, in: Cambridge Monographs on Applied and Computational Mathematics, Cambridge University Press, ISBN: 9781139459204, 2006.

[44] H. Tanaka, Stochastic differential equations with reflecting boundary conditions in convex regions, Hiroshima Math. J. 9 (1979) 163–177.

[45] David Bruce Wilson, Mixing times of lozenge tiling and card shuffling Markov chains, Ann. Appl. Probab. 14 (1) (2004) 274–325, http://dx.doi.org/10.1214/aoap/1075828054.