A Combinatorial Analysis of Interacting Diffusions

Sourav Chatterjee · Soumik Pal

Abstract We attempt to unify the analysis of several families of naturally occurring multidimensional stochastic processes by studying the underlying combinatorics involved. At equilibrium, the behavior of these processes is determined by the properties of a randomly chosen point of a corresponding polyhedron. How such a randomly chosen point behaves is a difficult question which is intertwined with the geometry and the symmetry of the polyhedron. The simplest of all cases is the simplex where a complete probabilistic study is known. A possible general strategy is through triangulation of the polyhedron where we decompose it as a union of simplices with non-intersecting interiors. In particular we study the case when the polyhedron is a simplicial polytope, since they correspond to the natural examples of stochastic processes. This is the case when the polytope is invariant under a Coxeter group action, which leads to a simple and explicit description of the equilibrium behavior of the stochastic processes in terms of independent and identically distributed Exponential random variables. Another class of examples is furnished by processes indexed by weighted graphs, all of which generate simplicial polytopes with $n!$ faces. We show that the proportion of volume contained in each component simplex corresponds to a probability distribution on the group of permutations, some of which have surprising connections with the classical urn models.

S. Chatterjee
University of California at Berkeley, 367 Evans Hall # 3860, Berkeley, CA 94720-3860, USA
e-mail: sourav@stat.berkeley.edu

S. Pal (✉)
University of Washington, C-547 Padelford Hall, Seattle, WA 98195, USA
e-mail: soumik@math.washington.edu

1 Introduction

Consider a sequence of n-dimensional cones $C_1, C_2, \ldots, C_r$ whose interiors are disjoint and the closure of their union is the entire space $\mathbb{R}^n$. Ignoring their mutual intersections (a set of measure zero) they can be thought of as a partition of $\mathbb{R}^n$. Let $b : \mathbb{R}^n \to \mathbb{R}^n$ be a function that is constant over each C_i, and consider the corresponding stochastic differential equation (SDE):

$$dX_t = b(X_t)\, dt + dW_t, \quad X_t \in \mathbb{R}^n, \tag{1}$$

where $W_t = (W_t(1), W_t(2), \ldots, W_t(n))$ is an n-dimensional Brownian motion.

In physical terms, the above equation models the n-dimensional position vector of a *particle* which moves randomly in a continuous motion under the influence of a potential. The SDE is an example of the well-known Langevin equation in statistical physics.

Define the function $k(x) = -\langle x, b(x) \rangle$. Since b is constant over cones, it follows that k is a positively homogeneous function. For example, k could be the Minkowski functional (i.e. the gauge function) of a set containing the origin. Conversely, if k is continuous and positive everywhere except at the origin, k must be a Minkowski functional of such a set. That is to say, if we define the *unit ball* and the *surface* given by k respectively as

$$C = \big\{x \,:\, k(x) \le 1\big\}, \qquad \mathcal{S} = \big\{x \,:\, k(x) = 1\big\}, \tag{2}$$

it is not difficult to see that k satisfies the relation $k(x) = \inf\{\alpha > 0 : x \in \alpha C\}$.

For instance, consider the following two examples of stochastic processes. The first one, called the bang-bang process is classical and is particularly important in stochastic control theory. See the articles by Karatzas and Shreve [16], Shreve [19], or the chapter by Warnecke [22]. This is a one-dimensional diffusion which solves the following SDE with a single real parameter α:

$$dX_t = -\alpha \, \text{sign}(X_t)\, dt + dW_t.$$

Here the function $k(x) = \alpha \, |x|$ which is a Minkowski functional if $\alpha > 0$.

The other example is the class of Brownian motions with rank-based interactions. This is a family of n-dimensional diffusion processes parametrized by a single vector δ in $\mathbb{R}^n$. These diffusions have an identity diffusion matrix and a drift that depends on the order in which the coordinates can be arranged in increasing values. If we think of each coordinate of the diffusion as recording the position of a particle moving on

a line, then at any instant of time the particle with the ith smallest position gets an instantaneous drift δ_i. The formal SDE for the diffusion can be described by

$$dX_t(i) = \sum_{j=1}^{n} \delta_j 1\{X_t(i) = X_t[j]\}\, dt + dW_t(i), \quad i = 1, 2, \ldots, n, \tag{3}$$

where $X_t[1] \leq X_t[2] \leq \cdots \leq X_t[n]$ are the coordinates arranged in increasing order. The $W_t(i)$'s are assumed to be independent Brownian motions for some suitable underlying filtration. The function k for the rank-based processes is given by $k(x) = -\sum_{i=1}^{n} \delta_i x[i]$, where $x[i]$ denotes the ith smallest coordinate of x.

The rank-based interacting Brownian motions and closely related models have appeared in several veins of the literature. Extensive reviews can be found in the articles by Pal and Pitman [17] and Chatterjee and Pal [3]. Some of the recent work include the articles by Jourdain and Malrieu [14] and Banner, Fernholz, and Karatzas [2]. Also see the related discrete time models by Ruzmaikina and Aizenman [18], and Arguin [1].

We are interested in the question: how does the diffusion solution of (1) behave in equilibrium? This is actually a two-part question: the first concerns the existence of the limit, the second is a description of the limiting probability measure on $\mathbb{R}^n$. For Langevin equations, the answer to the second question is classical. As we show below, the limiting probability measure, if it exists, always has a density $\exp(-2k)$ with respect to the Lebesgue measure (normalized to have a total integral of one).

But this is only the first step in understanding properties of the processes. Probability measures on high-dimensional spaces, specially non-product measures, can be notoriously hard to analyze. Since much of probability theory, in some way or the other, utilizes independence, it would be helpful to relate these measures with independent random variables.

To show concrete examples, consider the bang-bang process. It can be shown that the solution process is recurrent if and only if α is positive. In that case, the process has a unique reversible equilibrium distribution under which $|X_t|$ is distributed as Exponential with rate 2α. The recurrence and equilibrium properties of the rank-based Brownian motions can also be precisely determined. The following result is from [17].

Theorem 1 (Theorem 4 in [17]) *Consider a system of particles obeying the s.d.e.* (3). *For $1 \leq k \leq n$ let*

$$\alpha_k := \sum_{i=1}^{k} (\delta_i - \bar{\delta}), \quad \bar{\delta} := \frac{1}{n} \sum_{i=1}^{n} \delta_i. \tag{4}$$

For each fixed initial distribution the particles, the collection of laws of $X_t[n] - X_t[1]$ for $t \geq 0$ is tight if and only if

$$\alpha_k > 0 \quad \text{for all } 1 \leq k \leq N - 1, \tag{5}$$

in which case the following result holds:

The distribution of the spacings system $(X_t[j+1] - X_t[j], \; 1 \le j \le n-1)$ at time t converges in total variation as t tends to infinity to a unique stationary distribution for the spacings system, which is that of independent Exponential variables Y_k with rates $2\alpha_k$, $1 \le k \le n-1$. Moreover, the spacings system is reversible at equilibrium.

The independence of the spacings under the invariant distribution is somewhat puzzling since, due to the interaction, there is no independence between the spacing processes. The proof in [17], which does not shed light on this phenomenon, invokes Williams's results on reflected Brownian motions [23]. Notice the similarities between the two examples. Both involve drift functions that are piecewise constant; in fact the drift is a single vector under the action of a group (sign flips for the former and permutations for the latter). Moreover, in both cases the invariant distribution involves independent Exponentials which provide a friendly description of an otherwise abstract probability measure.

We provide alternate proofs of these results as an application of a general theory outlined below. Consider the general equation (1). Our first step is to reduce the problem of describing the stationary measure to the study of the combinatorics of the unit ball C. Now, by our assumptions about the structure of b, the unit ball is a polyhedron. Assume that it is a polytope. Because the density $\exp(-2k)$ is a function of k, a skew-product decomposition (proved later in the text) relates this probability measure to the uniform distribution on the polyhedron. Let us now focus on the uniform distribution on n-dimensional polytopes. If the polytope is the unit simplex $\{x = (x_1, \ldots, x_n) : x_i \ge 0, \; \sum_i x_i \le 1\}$, the uniform distribution can be effectively generated by dividing independent and identically distributed Exponential(1) random variables by their total sum. That is to say, let $X_1, \ldots, X_n$ be n independent and identically distributed random variables with an Exponential law with rate one, i.e., $P(X_i > t) = e^{-t}$, $t \ge 0$. Then the vector

$$(X_1/S, \ldots, X_n/S), \quad S = X_1 + \ldots X_n$$

is distributed uniformly in the set $\{x : x_i \ge 0, \; \sum_i x_i = 1\}$. Dropping any one coordinate gives us a vector chosen uniformly from the unit simplex of dimension $(n-1)$.

Now suppose we decompose the polytope as a union of K simplices with disjoint interiors. It is a well-known theorem that such a triangulation is possible for every polytope. Here, simplex means any linear transformation of the unit simplex. See, for example, the chapter by Lee in [10, Chap. 17]. Every component simplex, say S_i, in the triangulation is a linear map A_i of the unit simplex. Thus, an algorithm to generate a point uniformly from the polytope would be to generate a point X uniformly from the unit simplex, choose I between $\{1, 2, \ldots, K\}$ with probability

$$P(I = i) = \frac{\text{Vol}(S_i)}{\sum_{j=1}^{K} \text{Vol}(S_j)}, \quad i = 1, 2, \ldots, K, \tag{6}$$

and let $Y = A_I X$. Such an Y is clearly uniformly chosen from the polytope.

Here is a well-known example. Consider the unit cube in n-dimension. A point chosen randomly from this object has all coordinates uniformly distributed in the

interval $(0, 1)$ and are independent of one another. On the other hand, the unit cube has a natural triangulation consisting of the simplices

$$A_\pi := \left\{ x \in \mathbb{R}^n : 0 \leq x_{\pi(1)} \leq x_{\pi(2)} \leq \cdots \leq x_{\pi(n)} \leq 1 \right\}, \quad \pi \in \sigma_n,$$

for every π which is a permutation of n labels. Each of these A_π's have the same volume and a natural map to the unit simplex (the point $(x_1, \ldots, x_n)$ goes to $(x_{\pi(1)}, x_{\pi(2)} - x_{\pi(1)}, \ldots, x_{\pi(n)} - x_{\pi(n-1)})$). Thus, we readily arrive at the following fact. Pick n independent Uniform$(0, 1)$ random variables $U_1, U_2, \ldots, U_n$, and arrange them in increasing order $U_{(1)} \leq U_{(2)} \leq \cdots \leq U_{(n)}$. Then the vector formed by $U_{(1)}$ and the spacings $U_{(i+1)} - U_{(i)}$, $1 \leq i \leq n - 1$, between the order statistics is a random point from the n-simplex. Conversely, it can be stated as: pick a point Y uniformly from the simplex, and a random permutation π chosen uniformly from S_n. Then the vector $(U_{\pi(1)}, \ldots, U_{\pi(n)})$, where

$$U_i = \sum_{j=1}^{i} Y_j, \quad i = 1, 2, \ldots, n,$$

are independent Uniform$(0, 1)$ random variables.

A particularly explicit triangulation exists when the n-dimensional polytope is simplicial (i.e., every $(n - 1)$-dimensional extremal face is a simplex). One can then simply connect the origin (which is in the interior) with each of the extremal faces to generate a triangulation. Simplicial polytopes can sometimes be identified through their group of symmetries and have a well-explored relationship with a class of groups known as the Coxeter groups. Euclidean Coxeter groups are groups of orthogonal matrices generated purely by reflections. They find applications in several ares of mathematics and rightfully has a huge literature of its own. We direct the reader to the following excellent introductory books by Humphreys [13], Kane [15], and, Grove & Benson [12]. Also see the classics by Coxeter, [4] and [5]. The finite Coxeter groups include symmetries of regular polytopes and the Weyl groups of simple Lie algebras. They are usually defined formally as a set of generators and relations among them. The definitions and properties of finite reflection groups and Coxeter groups have been described in Sect. 3.1.

The polytopes related to Coxeter groups are particularly nice since not only they are simplicial, but they also have a group of symmetry given by the Coxeter group. Then the random variable I in (6) is uniformly distributed among $\{1, 2, \ldots, K\}$. This is exactly the case for the bang-bang or the rank-based processes. Generalizing these examples we prove the following result in this article.

Proposition 2 *Consider the SDE*

$$dX_t = b(X_t) \, dt + dW_t, \quad \text{where } W_t = \left(W_t(1), W_t(2), \ldots, W_t(n) \right) \tag{7}$$

is an n-dimensional Brownian motion. Assume that the drift function $b : \mathbb{R}^n \to \mathbb{R}^n$ is piecewise constant and satisfies

$$b(\alpha x) = b(x), \quad \forall \alpha > 0 \quad \text{and} \quad b(Ax) = Ab(x), \quad \forall A \in G, \tag{8}$$

where G is a finite group of orthogonal matrices with no non-trivial invariant subspaces. Then the following conclusions hold.

(i) *Let $k(x) = -\langle x, b(x) \rangle$. A sufficient condition for X to be recurrent is that, for some non-zero vector $\lambda \in \mathbb{R}^n$, we have*

$$k(x) = \max_{A \in G} \langle A\lambda, x \rangle.$$

   In that case X has a unique, reversible invariant distribution μ on $\mathbb{R}^n$. The marginal law of X_t converges in total variation to μ as t tends to infinity.

(ii) *Moreover, if G is Coxeter, there are n linearly independent vectors $\eta_1, \eta_2, \ldots, \eta_n$ and n constants $\alpha_1, \alpha_2, \ldots, \alpha_n$ such that under μ, the random variables*

$$Y_i = \alpha_i \langle A\eta_i, X \rangle, \quad i = 1, 2, \ldots, n, \tag{9}$$

   are iid exponential random variables with rate two, where A is a matrix-valued function of X such that $k(X) \equiv \langle A\lambda, X \rangle$. Alternately, this can be expressed by saying that each $\langle A\eta_i, X \rangle$ is distributed as Exponential with rate $2\alpha_i$.

(iii) *Additionally, if the stabilizer subgroup of λ in G is trivial, then the vectors $\eta_1, \eta_2, \ldots, \eta_n$ are determined as the generators of the conic hull of the set of vectors $\{\lambda - A\lambda, \ A \in G\}$. That is to say, every vector in the set $\{\lambda - A\lambda, \ A \in G\}$ can be represented as a linear combination of the subset $\{\eta_1, \eta_2, \ldots, \eta_n\}$ with non-negative coefficients.*

   The constants $\alpha_1, \ldots, \alpha_n$ are the unique positive coefficients of $\eta_1, \ldots, \eta_n$ in the expansion

$$\lambda = \sum_{i=1}^{n} \alpha_i \eta_i.$$

Note that we really do not need to know the details of the group structure to apply the previous result, except for the information that G is Coxeter. As we show in the examples in Sect. 3.2, necessary and sufficient conditions can be obtained if we have a better knowledge of the group structure. Finally, let us mention that a list of Coxeter groups up to isomorphisms is available and can be found in any standard textbook, say [12].

In Sect. 3.2, we describe several families of interacting diffusions that can be analyzed by the previous theorem. They all appear as solutions to stochastic differential equations of the type (7) with a piecewise constant drift function satisfying condition (8), but involving different families of orthogonal groups.

When G is the group of permutation matrices, we get back rank-based interactions. Using Proposition 2, we provide an alternative proof (not involving reflected Brownian motions) of the Pal–Pitman result.

Another class of examples are named *sign-rank-based interactions*. Here the drift vector not only changes when the coordinate values get permuted, but also, when the signs of the coordinates change. The relevant group is the one which generated by all the permutation matrices and all the diagonal matrices whose diagonal elements are either plus or minus one. In one dimension, this boils down to the simple Bang-bang process.

A final class of examples considered are similar to the sign-rank-based processes, but with more constraints. Here, too, the drift vector changes when we permute coordinates. It also changes when we change signs of coordinates, but only when done in pairs. The group behind the curtain is generated by permutation matrices and diagonal matrices whose diagonal elements are ± 1 with the additional constraint that only even number of -1's are allowed.

Readers acquainted with the theory of Coxeter groups will recognize that the previous three examples correspond to the three well-known families of Coxeter groups, denoted by $\mathcal{A}_n$, $\mathcal{B}_n$, and $\mathcal{D}_n$ for each $n \in \mathbb{N}$. In each case Proposition 2 allows us to formulate a simple sufficient condition for checking the existence of a unique invariant probability distribution and provides a complete description of the distribution in terms of independent Exponentials.

The second class of stochastic processes we consider is not regular (i.e., no obvious group of symmetries). The interaction is parametrized by all graphs with n vertices and possible edge weights. Thus, the probabilities in (6) are not easy to determine. In fact, these probabilities correspond to probabilities of various orderings of particles, increasingly arranged, under the invariant distribution. We take up a few examples and show surprising connections with existing probability models on permutations.

Let $\mathcal{G}$ be a graph on n vertices where the vertices are labeled by $\{1, 2, \dots, n\}$. The edge between i and j have an associated edge weight β_{ij}, which is zero if there is no edge between the two vertices. Consider the SDE on $\mathbb{R}^n$ given by

$$dX_t(i) = \sum_{j=1}^{n} \beta_{ij} \operatorname{sign}\big(X_t(j) - X_t(i)\big) dt + dW_t(i), \quad i = 1, 2, \dots, n,$$

where, W is again an n-dimensional Brownian motion.

When all the edge weights are non-negative, the model can be described by imagining n particles on the line whose positions are recorded at time t by the n-dimensional vector $(X_t(1), \dots, X_t(n))$. These particles are indexed by the vertices of the graphs, and gets attracted towards one another while performing random motion. The constants β_{ij} measure the strength of their attraction. Thus, it is intuitive since all the n particles attract each other, the *cloud* of particles stay together at all times. This has been made precise in the text by showing the following. Define

$$\bar{X}_t = \frac{1}{n} \sum_{i=1}^{n} X_t(i),$$

the centered vector $(X(1) - \bar{X}, X(2) - \bar{X}, \dots, X(n) - \bar{X})$ has a stationary distribution whenever the graph is connected and the edge weights are non-negative.

We focus on the case when we assume that each particle has a mass m_i, $i = 1, 2, \dots, n$, and that $\beta_{ij} = m_i m_j$, in the sense that the strength of the mutual attraction is proportional to the product of their masses. Due to this *gravitational* intuition, under the stationary distribution, we should expect heavier particles to stay at the middle of the pile, while the lighter ones should be at the edge.

In general, this is difficult to formulate and prove. However, in one particular case, this becomes apparent. For any $\alpha > 0$, consider n particles with the interaction described in the previous paragraph, where the mass of the first particle is α and the rest of the masses are 1. Similar models of 'heavy particles' (when $\alpha > 1$) have been considered widely in the mathematical physics literature. See the articles by Szász & Tóth [21], and by Sinai & Soloveychik [20].

We prove the following surprising connection with Polya's urn scheme. Let $\sigma(1)$ denote the rank of the first particle under the invariant distribution of the centered vector $(X(1) - \bar{X}, X(2) - \bar{X}, \ldots, X(n) - \bar{X})$. We prove the following.

Proposition 3 *Consider a Polya's urn scheme which initially has α red balls and α black balls. At every step one picks up a ball at random, returns the ball to the urn and adds an extra ball of the same color. Then, the distribution of $\sigma(1) - 1$ is the same as the number of red balls picked when we run the urn scheme described above for $n - 1$ steps.*

In particular, the sequence of random variables $\sigma(1)/n$ converges weakly to the Beta(α, α) distribution as n tends to infinity.

As a concluding remark we would like to say that this article is describing one facade of the relationship between triangulation and *factorization* of uniform distributions over polytopes in terms of independent exponentials. We believe that the central idea has a potential beyond the examples described here and can be fruitful in understanding both combinatorics and probabilistic structures on polytopes.

1.1 Outline of the Paper

In the next section we describe the set-up of the paper and prove general results about recurrence of interacting diffusions and their invariant distributions when they exist. Section 3 describes the combinatorics involved in the invariant distributions. Section 3.1 describes the connection with Coxeter groups followed by several examples in Sect. 3.2. The following Sect. 3.3 proves results about interactions parametrized by graphs.

2 Diffusions with Piecewise Constant Drift

Let us return to (1). The existence and uniqueness in law of the solution of the equation is immediate by an application of Girsanov's theorem [7, p. 202]. Define, as before, the function $k(x) = -\langle x, b(x) \rangle$. Since b is constant over cones, it follows that k is a positively homogeneous function. If k is continuous, by virtue of being piecewise linear, it follows easily that b is the negative of the gradient of k in the sense of distributions. The following proposition is well known.

Proposition 4 *Consider the stochastic differential equation* (1). *Let $k : \mathbb{R}^n \to \mathbb{R}$ be a continuous function such that b represents $-\nabla k$ in the sense of distribution. Assume that $\exp(-2k(x))$ is integrable. Then the probability distribution given by the unnormalized density $\exp(-2k(x))dx$ provides a reversible, invariant probability distribution μ for the process in* (1).

Proof Let μ be the measure on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ defined by $\mu(dx) = e^{-2k(x)}\, dx$. Consider the Sobolev space, $H^{1,2}$, of all measurable functions $f : \mathbb{R}^n \to \mathbb{R}$, such that f and all of its partial derivatives $\partial f / \partial x_i$, $i = 1, 2, \ldots, n$ (in the sense of distributions) are in $\mathbb{L}^2(\mu)$. Then we can define the following symmetric bilinear form, on the domain $H^{1,2}$, given by

$$\mathcal{E}(f, g) = \int_{\mathbb{R}^n} \langle \nabla f, \nabla g \rangle e^{-2k(x)}\, dx.$$

Since $e^{-2k(x)}$ is never zero, it follows that $H^{1,2}$ is a Hilbert space. Thus, $\mathcal{E}$ is closed, since it is defined everywhere on the Hilbert space $H^{1,2}$. It is also known to be $\mathcal{E}$ is Markovian (see, e.g., [9, Example 1.2.1]). Thus, it is clear that this is a Dirichlet form in $\mathbb{L}^2(\mathbb{R}^n, \mu)$.

By Theorem 1.3.1 in [9], we claim the existence of a unique non-positive definite self-adjoint operator $\mathcal{L}'$ on $H^{1,2}$ such that

$$\mathcal{E}(f, g) = \left\langle \sqrt{-\mathcal{L}'} f, \sqrt{-\mathcal{L}'} g \right\rangle_\mu, \quad \forall f, g \in H^{1,2}.$$

Here $\langle \cdot \rangle_\mu$ refers to the usual inner product in $\mathbb{L}^2(\mu)$. Or, in other words, (Corollary 1.3.1 of [9]) there is a unique self-adjoint operator $\mathcal{L}'$ on a domain $\mathcal{D}(\mathcal{L}') \subseteq H^{1,2}$ such that

$$\mathcal{E}(f, g) = \left\langle -\mathcal{L}' f, g \right\rangle_\mu, \quad \forall f \in \mathcal{D}(\mathcal{L}'), \forall g \in H^{1,2}. \tag{10}$$

We now show that (10) is satisfied by a multiple of the generator of the Markov process in (1). The generator, $\mathcal{L}$, is given by

$$\mathcal{L} f = \langle b, \nabla f \rangle + \frac{1}{2} \Delta f.$$

By our assumption $b = -\nabla k$ takes finitely many values. Thus, we can define $\mathcal{L}$ on the domain

$$H^{2,2} = \left\{ f \in \mathbb{L}^2(\mu) \,\bigg|\, \frac{\partial f}{\partial x_i} \in \mathbb{L}^2(\mu), \text{ and } \frac{\partial^2 f}{\partial x_i^2} \in \mathbb{L}^2(\mu) \, \forall i = 1, 2 \ldots, n \right\}.$$

It is clear that the domain of $\mathcal{L}$ above is a subset of $H^{1,2}$.

We claim that $2\mathcal{L}$ satisfies (10). In that direction, consider any $f \in H^{2,2}$ and any $g \in H^{1,2}$, we have

$$\int_{\mathbb{R}^n} \mathcal{L}(f) g e^{-2k(x)}\, dx = \int_{\mathbb{R}^n} \left(\langle b, \nabla f \rangle + \frac{1}{2} \Delta f \right) g e^{-2k(x)}\, dx$$

$$= \int_{\mathbb{R}^n} \langle b, \nabla f \rangle g e^{-2k(x)}\, dx + \int_{\mathbb{R}^n} \frac{1}{2} \Delta f g e^{-2k(x)}\, dx$$

$$= -\int_{\mathbb{R}^n} \langle \nabla k, \nabla f \rangle g e^{-2k(x)}\, dx - \frac{1}{2} \int_{\mathbb{R}^n} \left\langle \nabla f, \nabla \left(g e^{-2k(x)} \right) \right\rangle dx$$

$$= -\frac{1}{2} \int_{\mathbb{R}^n} \langle \nabla f, \nabla g \rangle e^{-2k(x)}\, dx = -\frac{1}{2} \mathcal{E}(f, g). \tag{11}$$

Note that the boundary terms are zero in the integration by parts above since both $\partial f / \partial x_i$ and g are in $\mathbb{L}^2(\mu)$, and thus

$$\left. \frac{\partial f}{\partial x_i} g e^{-2k(x)} \right|_{-\infty}^{\infty} = 0, \quad \forall i = 1, 2, \ldots, n.$$

We can rewrite (11) as

$$\mathcal{E}(f, g) = \langle -2\mathcal{L}f, g \rangle, \quad \forall f \in H^{2,2}, \forall g \in H^{1,2}$$

which, compared with (10), proves that $2\mathcal{L}$ to be the unique operator associated with the Dirichlet form $\mathcal{E}$. Further, from self-adjointness of $\mathcal{L}$, we infer

$$\langle \mathcal{L}f, g \rangle_\mu = \langle f, \mathcal{L}g \rangle_\mu, \quad \forall f, g \in H^{2,2}, \tag{12}$$

where $\langle \cdot, \cdot \rangle_\mu$ refers to the usual $\mathbb{L}^2$ inner product. We can take $g \equiv 1$ to get that μ is an invariant measure for the process X_t. This proves the claim. $\square$

When is the function $\exp(-2k(x))$ integrable? This question is critical to both the recurrence property of the diffusion process as well as the existence of a unique long term stationary distribution. Its answer, however, is geometric in nature. It is intuitive that k needs to be non-negative.

Definition 5 A continuous, non-negative, positively homogeneous function $k : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is said to be *irreducible* if it satisfies $k(x) = 0$ if and only if $x = 0$.

Lemma 6 *If k is continuous and irreducible either $k(x) \geq 0$, $\forall x \in \mathbb{R}^n$, or $k(x) \leq 0$, $\forall x \in \mathbb{R}^n$. Moreover, if $k(x) > 0$, $\forall x \in \mathbb{R}^n$, $x \neq 0$, then $\mathrm{Vol}(C) < \infty$.*

Proof Suppose that there are points x_0 and x_1 such that $k(x_0) > 0$ and $k(x_1) < 0$. We can choose a continuous curve γ_t, $t \in [0, 1]$, in $\mathbb{R}^n$ such that $\gamma_0 = x_1$ and $\gamma_1 = x_2$ and $0 \notin \gamma[0, 1]$. Since k is continuous, by the intermediate value theorem, there exists a $t^* \in (0, 1)$ such that $k(\gamma_{t^*}) = 0$ but $\gamma_{t^*} \neq 0$. But this is impossible if k is irreducible, and we have proved the first assertion of the lemma.

For the second assertion, we need to show that C is bounded. Suppose, on the contrary, we can find a sequence $\{x_n\} \subseteq C$ such that $\lim_{n \to \infty} \|x_n\| = \infty$. One can assume that $\|x_n\| \geq 1$, for all $n \in \mathbb{N}$. Then the points $y_n = x_n / \|x_n\|$ satisfy

$$k(y_n) = k(x_n) / \|x_n\| \leq 1, \quad \forall n \in \mathbb{N}.$$

Thus, $y_n \in C$, for all $n = 1, 2, \ldots$. However, there exists a subsequence of y_n, say $\{y_{n_m}\}$ such that $\lim_{m \to \infty} y_{n_m} = z$, for some z with $\|z\| = 1$. Hence, by continuity of k, we infer

$$k(z) = \lim_{m \to \infty} k(y_{n_m}) = \lim_{m \to \infty} \frac{k(x_{n_m})}{\|x_{n_m}\|} = 0.$$

The final equality is due to the fact that $0 \leq k(x_n) \leq 1$ for all n and $\lim \|x_n\| = \infty$. Since $z \neq 0$, this contradicts our assumption that k is irreducible. Hence we are done. $\square$

We now show that the process in (1) is Harris recurrent if k is non-negative and irreducible. It then follows (see [7, Sect. 7.5]) that it has a unique invariant measure μ described above in Proposition 4. Moreover, if $P_t(x)$ is the marginal distribution of X_t when $X_0 = x$, then $\lim_{t\to\infty} \|P_t(x) - \mu\|_{\text{TV}} = 0$. Here $\|\cdot\|_{\text{TV}}$ refers to the total variation norm on measures.

The following claim settles the argument.

Lemma 7 *Consider the notations and assumptions in Proposition* 4. *Suppose that the function k is a non-negative, irreducible, positively homogeneous function. Then the process X_t is recurrent.*

Proof We will use Corollary 7.5.4 in [7]. We need to consider the quantity

$$d(x) = n + 2\langle x, b(x)\rangle, \quad x \in \mathbb{R}^n.$$

By our definition we have $\langle x, b(x)\rangle = -k(x)$. Thus, $d(x) = n - 2k(x)$.

Now, since k is non-negative and irreducible, it is growing to infinity *uniformly* in all directions radially outward from zero. The way to see this is to note

$$k(x) = \|x\| k(x/\|x\|) \geq \|x\| \inf_{\|y\|=1} k(y) = c_1 \|x\|.$$

The constant $c_1 = \inf_{\|y\|=1} k(y)$ is positive since k is a strictly positive continuous function on the compact set $\{y : \|y\| = 1\}$.

Now, if we fix an $\epsilon > 0$, there exists $R > 0$ such that $d(x) < -\epsilon$ for all x with $\|x\| > R$. Let T_R be the first hitting time of the compact set $B_R = \{x \in \mathbb{R}^n : \|x\| \leq R\}$. By Corollary 7.5.4 in [7], we immediately obtain $E^x(T_R) \leq \|x\|^2/\epsilon$. Thus B_R gets visited infinitely often and hence the process X_t is recurrent. $\qquad\square$

The integrability of the function $\exp(-2k(x))$ requires precisely the same condition as in the last lemma.

Lemma 8 *Suppose that $k : \mathbb{R}^n \to \mathbb{R}$ is a non-negative, irreducible, continuous, positively homogeneous function. Then the $\exp(-2k(x))$ is an integrable function.*

To prove the previous lemma we need the following *polar decomposition* formula. For any non-negative, irreducible k, define the surface projection function $\Theta : \mathbb{R}^n \to \mathcal{S}$ by

$$\Theta(x) = \frac{x}{k(x)}, \quad x \in \mathbb{R}^n \backslash \{0\}, \text{ and } \Theta(0) = 0. \tag{13}$$

Clearly the range of Θ is the surface $\mathcal{S}$ defined in (2).

We take the following two measurable spaces

$$M_1 = \left(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n)\right) \quad \text{and} \quad M_2 = \left(\mathbb{R}^+ \times \mathcal{S}, \ \mathcal{B}(\mathbb{R}^+) \otimes \mathcal{B}(\mathcal{S})\right).$$

One can construct a measurable map $T : M_1 \to M_2$ given by $T(x) = (k(x), \Theta(x))$, $\forall x \in \mathbb{R}^n$. It clearly follows from the definition that T is a one-to-one map. We prove the following slightly general result for the sake of completeness.

Lemma 9 *For any non-negative, irreducible, positively homogeneous function* $k :$ $\mathbb{R}^n \to \mathbb{R}^+ \cup \{\infty\}$, *and any integrable* $f : \mathbb{R}^n \to \mathbb{R}$, *we have*

$$\int_{\{x:k(x)<\infty\}} f(x)\,dx = n\operatorname{Vol}(C) \int_0^\infty r^{n-1} \int_{\mathcal{S}} f(r \cdot z)\,d\mu(z)\,dr, \qquad (14)$$

where dx *refers to the* n-*dimensional Lebesgue measure and* μ *is the cone measure on* $\mathcal{S}$ *defined as*

$$\mu(E) = \frac{1}{\operatorname{Vol}(C)} \operatorname{Vol}\big(T^{-1}([0,1] \times E)\big), \quad \forall E \in \mathcal{B}(\mathcal{S}). \qquad (15)$$

Proof We first prove (14) for functions f equal to indicators of sets $A = T^{-1}([0,b] \times E)$, where $E \in \mathcal{B}(\mathcal{S})$ and $b \geq 0$. We have

$$\operatorname{Vol}(A) = \int 1_A\,dx = \int 1\{k(x) \leq b\} 1\{\Theta(x) \in E\}\,dx.$$

If we make let $y = b^{-1}x$, then, by positive homogeneity of k, one can write the last equation as

$$\operatorname{Vol}(A) = b^n \int 1\{k(y) \leq 1\} 1\{\Theta(y) \in E\}\,dy = b^n \operatorname{Vol}\big(T^{-1}([0,1] \times E)\big)$$

$$= b^n \operatorname{Vol}(C) \cdot \mu(E) = \operatorname{Vol}(C)\,n \int_0^b r^{n-1} \int_E d\mu\,dr,$$

which proves (14) for this particular case. The rest of the argument follows from standard measure theoretic approximation results.   □

Proof of Lemma 8 Since $2k$ is another non-negative, irreducible, positively homogeneous function, it suffices to show that $\exp(-k(x))$ is integrable.

By Lemma 6, the set $C = \{x \in \mathbb{R}^n : k(x) \leq 1\}$ has a finite volume, and hence the cone measure on $\mathcal{S} = \partial C$ is well defined. From the change of variable formula in Lemma 9 we then obtain

$$\int_{\mathbb{R}^n} e^{-k(x)}\,dx = n\operatorname{Vol}(C) \int_0^\infty r^{n-1} e^{-r}\,dr$$

$$= n\operatorname{Vol}(C) \int_0^\infty s^{n-1} e^{-s}\,ds = \operatorname{Vol}(C)\,n\Gamma(n)$$

$$= \operatorname{Vol}(C)\,\Gamma(n+1) < \infty.$$

□

Note that it follows from Lemma 9 that under the probability measure with the unnormalized density $e^{-k(x)}$, the random variable $k(X)$ is always distributed as Gamma(n) irrespective of k and independently of the vector $\Theta(X)$. Similarly, under the uniform measure on C, the random variable $k(X)$ is always distributed as

Beta$(n, 1)$ independently of $\Theta(X)$. Under both these measures, $\Theta(X)$ has the same law. This provides a link between the two probability measures which is important in their understanding.

What can be recovered when k is not irreducible? In general little, except when k is the gauge function of a lower-dimensional polytope. The following proposition generalizes Proposition 4.

Proposition 10 *Let $k : \mathbb{R}^n \to \mathbb{R}$ be a continuous function whose derivative in the sense of distribution is represented by a bounded function. Suppose there exists a subspace $H \subseteq \mathbb{R}^n$ such that if $y \in H$ and $z \in H^\perp$, then*

$$k(y + z) = k_1(y) + k_2(z).$$

Additionally, assume that k_1 is a non-negative, irreducible, positive homogeneous function on H. Consider the solution to (1) when $b(x)$ represents $-\nabla k(x)$, where ∇ is in the sense of distributions. Assume that a solution to (1) with this drift exists. Let $A : \mathbb{R}^n \to H$ be a projection matrix onto the subspace H. Then

(1) *The process $Y_t = AX_t$ has a unique reversible stationary probability distribution μ.*
(2) *Suppose k' is any other function defined as*

$$k'(x) = k_1(Ax) + k'_2(x - Ax), \quad \forall x \in \mathbb{R}^n,$$

for some non-negative function k'_2 such that $\exp(-2k')$ is integrable. Then μ is the law of the random vector $Y = AX$, where X is a random vector with density proportional to $\exp(-2k')$.

Proof Define $k'_2 : \mathbb{R}^n \to \mathbb{R}$ to be

$$k'_2(x) = \sup_{1 \le j \le n-d} \left| \langle x, \delta_j \rangle \right|, \tag{16}$$

where δ_i, $i = 1, \ldots, n - d$ is an orthogonal basis of $H^\perp$. Modify the function k by defining

$$k'(x) = k_1(y) + k'_2(z), \quad \forall y \in H, \ z \in H^\perp, \ x = y + z.$$

• We claim that k' is a non-negative, irreducible, positively homogeneous function on $\mathbb{R}^n$. Let us verify the condition in Definition 5. Since k_1 and k'_2 are both positive, for $x = y + z$, $y \in H$, $z \in H^\perp$, we get

$$\{h'(x) = 0\} \quad \Leftrightarrow \quad \{k'(x) = 0\} \quad \Leftrightarrow \quad \{k_1(y) = 0\} \quad \text{and} \quad \{k'_2(z) = 0\}.$$

But, by assumption for k_1 and by construction for k'_2, we get that $k'(x) = 0$ if and only if $y = 0$ and $z = 0$, that is, if and only if $x = 0$. This proves our claim.

Consider the solutions of the following two stochastic differential equations

$$dX_t = -\nabla k(X'_t)\, dt + dW_t, \tag{17}$$
$$dX'_t = -\nabla k'(X'_t)\, dt + dW_t, \tag{18}$$

where the drifts are function representing the derivatives in the sense of distributions. That X'_t exists (in the weak sense) is clear.

● We claim that X' has a unique reversible invariant distribution μ' with unnormalized density $\exp(-2k'(x))\, dx$. From Lemma 8, we deduce that $\exp(-2k')$ is integrable. Hence we can suitably normalize and get a probability measure whose density is proportional to $\exp(-2k'(x))$. Now, we can apply Proposition 4 to claim that the unique reversible invariant distribution of the process X' exists and is given by the unnormalized density $\exp(-2k'(x))$.

Now, consider a projection matrix A with range space H. Thus, $A^\perp = I - A$ is a projection onto $H^\perp$. Now, by assumption, the function k splits additively as

$$k(x) = k_1(Ax) + k_2(A^\perp x).$$

Taking gradients on both sides, one obtains

$$\Rightarrow \quad \nabla k(x) = A'\nabla k_1(Ax) + (A^\perp)'\nabla k_2(A^\perp x).$$

Here A' and $(A^\perp)'$ refers to taking adjoints. But A being a projection is self-adjoint and satisfies $A^2 = A$. Thus, it follows that

$$A\nabla k(x) = A^2\nabla k_1(Ax) + A(I - A)\nabla k_2(A^\perp x) = A\nabla k_1(Ax). \tag{19}$$

Exactly in the same way we also have

$$A\nabla k'(x) = AA'\nabla k_1(Ax) + A(A^\perp)'\nabla k'_2(A^\perp x) = Ak_1(Ax). \tag{20}$$

Now, consider the processes $Y = AX$ and $Y' = AX'$ which clearly satisfy the following differential equations

$$dY_t = -A\nabla k(X_t)\, dt + A\, dW_t, \quad \text{and}$$
$$dY'_t = -A\nabla k'(X'_t)\, dt + A\, dW_t. \tag{21}$$

Using (19), and (20) we can rewrite the above equations as

$$dY_t = -A\nabla k_1(Y) + A\, dW_t, \quad \text{and}$$
$$dY'_t = -A\nabla k_1(Y') + A\, dW_t.$$

Clearly the laws of the process Y and Y' are identical due to the uniqueness in law of the weak solutions of their stochastic differential equation. To get past the arbitrariness of the hyperplane H, one can simply observe that if dimension$(H) = d \leq n$,

there exists a $(d \times n)$ matrix D which is a bijection between H and $\mathbb{R}^d$. The laws of DY and DY' are identical by standard theory of SDE. Now one simply inverts D onto H to obtain our conclusion.

Now, since A is a linear map, the process Y' has a unique invariant distribution ν induced by the invariant distribution of X', and given by

$$\nu(B) = \mu'(Ax \in B), \quad \forall B \in \mathcal{B}(H), \tag{22}$$

where μ' is the unique invariant distribution for the process X'. Thus, by the equality in law of the processes Y and Y' it follows that the process $Y_t = AX_t$ has a unique invariant distribution given by ν according to the recipe above.

It is easy to give an intrinsic description of ν from (22). If X' is distributed according to μ', then for any $B \in \mathcal{B}(H)$, we have

$$\nu(B) = \mu'\big(AX' \in B\big) = \int_{Ax \in B} \exp\big(-2k_1(Ax) - 2k_2'(x - Ax)\big)\, dx$$

$$= \int_{y \in B} \exp\big(-2k_1(y)\big)\, dy \int_{H^\perp} \exp\big(-2k_2'(z)\big)\, dz.$$

Clearly, ν has a density proportional to $\exp(-2k_1(y))$ with respect to the Lebesgue measure restricted to the hyperplane H. □

3 Simplicial Cones and Exponential Distributions

We have seen in the last section that the invariant probability distributions for the SDEs described in (1) have unnormalized densities with respect to the Lebesgue measure given by $\exp(-2k(x))$, where k is a non-negative, irreducible, positively homogeneous function. Even with such an explicit description of the density function it can be very hard to compute any means, variances, or one-dimensional marginal distributions. Our objective in this section is to link them to the combinatorial structures of the unit ball generated by the positively homogeneous function $2k$, particularly when these are star-shaped or convex polytopes. In the special case of simplicial polytopes this allows us to furnish a complete description of the invariant measure in terms of independent Exponential random variables.

The main geometric idea is the following. Consider, as before, a drift function b which is constant over cones $C_1, C_2, \ldots, C_r$. We assume that b is the negative gradient in the sense of distributions of the function $k(x) = -\langle x, b(x) \rangle$. Thus, $2k(x)$ is a linear function inside each cone C_i.

Suppose now that C_i is simplicial. Simplicial cones are those that can be transformed to the positive quadrant by applying a non-singular linear transformation. That is to say, for each C_i there exists n linearly independent vectors in $\mathbb{R}^n$, denoted by $\beta_1, \beta_2, \ldots, \beta_n$, such that $C_i = \{x \in \mathbb{R}^n : x = \sum_i a_i \beta_i, \ a_i \geq 0 \ \forall i\}$. It is now not hard to see that in that case the probability measure given by $\exp(-2k(x))$ restricted to the cone C_i must be a linear transformation of independent Exponential distributions. The difficulty in the execution of the previous argument is to identify from the function b that a simplicial polytope is lurking behind the scenes and to compute the

necessary linear transformations. In the following two subsections we consider two general classes of examples where the argument can be fully carried out. One, the regular case, is where each C_i can be mapped to any other by a group of orthogonal transformations. This leads us to a connection with finite irreducible Coxeter groups. The other, which is not regular, deals with graphs and the combinatorics hidden in their structure.

3.1 The Regular Case: Groups of Orthogonal Transformations

Consider a finite-dimensional Euclidean vector space V. A linear transformation from V to itself is called orthogonal if the corresponding matrix is orthogonal. In that case the determinant of the transformation is ± 1. An important class of orthogonal matrices is given by reflections. A reflection along a unit vector $u \in V$ corresponds to the matrix $I - 2uu'$ (all vectors are columns and u' denotes the transpose of u). Geometrically it produces the mirror image of any vector with respect to the hyperplane orthogonal to u. A reflection group is a group of matrices with a generator each element of which is a reflection matrix.

Let G be a group of orthogonal matrices. G is called irreducible if there is no nontrivial subspace W of V which is stable under the action of G, i.e., $\rho(s)W \subseteq W$, for all $s \in G$. As a recurring example, considering the symmetric group of permutations on n elements. It has a natural representation as permutation matrices on $\mathbb{R}^n$. This is not irreducible since the one-dimensional subspace W_1 spanned by the vector of all ones remains invariant under the action of the group. However, the action restricted to W_1 and $W_2 = W_1^\perp$ is irreducible.

For a finite group of orthogonal linear transformations G we now define what is known as a *fundamental region*. Please see Chap. 3 in [12] for more details.

Fundamental Region *A subset $F \subseteq V$ is known as a fundamental region for a group of orthogonal transformations G if*

(1) *F is open.*
(2) *$F \cap AF = \emptyset$ if $A \neq I$, where I is the identity matrix and $A \in G$.*
(3) *$V = \bigcup_{A \in G} \overline{AF}$, where $\overline{B}$ denotes the topological closure of a subset B.*

For the representation of the symmetric group as permutation matrices such a region is provided by the cone $\{x \in \mathbb{R}^n : x_1 < x_2 < \cdots < x_n\}$. In general, fundamental regions are not unique.

Henceforth we will work with $V = \mathbb{R}^n$ even though proper subspaces of $\mathbb{R}^n$ provide another rich class of examples.

The groups we will be interested in, which includes the permutations as a special case, are generated purely by reflections. When irreducible, these groups are known as Coxeter groups and we review their basic structure below.

A reflection along a vector r is uniquely characterized by the fact that it keeps every vector orthogonal to r unchanged and flips the sign of every multiple of r. A Coxeter group is a finite irreducible group of orthogonal transformations generated by finitely many reflections. A frequent class of examples are the Dihedral groups which are the symmetry groups of regular polygons. Although, these groups contain

both rotations (think of a unit square being rotated by $\pi/2$) and reflections (the square getting reflected along the mid-axis), one can show that these groups can be generated purely by the reflection elements (reflecting the square twice along different axes amounts to a rotation).

Closely associated with the Coxeter groups is the idea of *root systems*. If G contains a reflection along r, then both r and $-r$ are known as *roots* of G. Let Δ be the set of all the roots of G, usually referred to as the *root system* of the Coxeter group. Recall the definition of a fundamental region. We are going to construct a fundamental region for G which is a simplicial cone.

Let G be a Coxeter group acting on $\mathbb{R}^n$. Thus, in particular, it is irreducible.

Choose any vector $u \in \mathbb{R}^n$ such that $\langle u, r \rangle \neq 0$ for any root r of G. Then the root system is partitioned into two subsets

$$\Delta_u^+ = \left\{ r \in \Delta : \langle r, u \rangle > 0 \right\}, \qquad \Delta_u^- = \left\{ r \in \Delta : \langle r, u \rangle < 0 \right\}.$$

Theorem 4.1.7 in [12] (and several lemmas preceding it) proves the following result.

Theorem 11 *There is a unique collection of n many vectors Π_u in Δ_u^+ such that every vector in Δ_u^+ can be written as a linear combination of elements in Π_u with non-negative coefficients. This collection, known as the u-base or fundamental roots, and denoted by Π_u, is linearly independent and forms a basis of $\mathbb{R}^n$.*

A u-base provides a fundamental region for a Coxeter group G, by defining

$$F_u := \left\{ x \in \mathbb{R}^n : \langle x, r \rangle > 0 \quad \forall r \in \Pi_u \right\}. \tag{23}$$

It has been proved in Theorem 4.2.4 in [12] that this is a fundamental region. It is also a simplicial cone since the vectors in Π_u are linearly independent.

Let us now return to the framework in Proposition 4. We start with a drift function $b : \mathbb{R}^n \to \mathbb{R}$ that is scale invariant, i.e., $b(\alpha x) = b(x)$ for all $\alpha > 0$. We would like to analyze the probability measure given by normalizing $\exp(-2k(x))$ where $k(x) = -\langle x, b(x) \rangle$. Suppose the drift function b takes finitely many values on $\mathbb{R}^n$ and satisfies the property that $b(Ax) = Ab(x)$ for all $A \in G$ for some Coxeter group G. Can we claim that there is a finite sequence of non-overlapping simplicial cones whose union is the whole space and such that b takes a constant value inside each cone?

The answer is no in general. However, there is a simple sufficient condition which indeed guarantees an affirmative answer to the question. Let, as before, $k(x) = -\langle x, b(x) \rangle$. If k is non-negative, then clearly, by positive homogeneity, k is the Minkowski functional of a star-shaped body containing the origin. The question in the previous paragraph is equivalent to asking whether this star-shaped body can be triangulated as a disjoint union of simplices each of which contains the origin as a extreme point. We are going to show below that the answer to the question is yes, if $k(x) = \max_{A \in G} \langle \lambda, Ax \rangle$ for some vector λ. Seen through a geometric lens, this is equivalent to the statement that the star-shaped body generated by k is a convex polytope which then turns out to be simplicial.

Lemma 12 *Let G be a finite irreducible group of orthogonal matrices on $\mathbb{R}^n$. For any $\lambda \in \mathbb{R}^n$, $\lambda \neq 0$, the function*

$$k(x) = \max_{A \in G} \langle \lambda, Ax \rangle, \quad x \in \mathbb{R}^n, \tag{24}$$

is a non-negative positively homogeneous function on $\mathbb{R}^n$, which is irreducible. Moreover, k is invariant under the action of the group. That is $k(x) = k(Ax)$, for all $A \in G$, and for all $x \in \mathbb{R}^n$.

Proof It is trivial to see that k is positively homogeneous. To show that it is non-negative we use the fact (see [6]) that for any non-trivial irreducible group G of orthogonal matrices, the sum $\sum_{A \in G} A$ is the zero matrix. Since G is irreducible, it follows that

$$\sum_{A \in G} \langle \lambda, Ax \rangle = \left\langle \lambda, \sum_A Ax \right\rangle = 0.$$

But this implies that the maximum must be non-negative. Thus k is non-negative.

To prove that k must be strictly positive for all non-zero vectors, note that for the previous argument, $k(x) = 0$ for some x would imply that

$$\langle \lambda, Ax \rangle = 0, \quad \forall A \in G.$$

Define the subspace

$$V_\lambda = \left\{ x \in \mathbb{R}^n : \langle \lambda, Ax \rangle = 0 \quad \forall A \in G \right\}. \tag{25}$$

Next, note that V_λ is stable under the action of G. To see this take any $B \in G$ and any $x \in V_\lambda$, then clearly $Bx \in V_\lambda$. Thus $AV_\lambda \subseteq V_\lambda$ for all $A \in G$. But, since G is irreducible, V_λ must be either zero or the entire space. If V_λ is the entire subspace, then by putting A to be the identity in the definition (25), we get $\langle \lambda, x \rangle = 0 \; \forall x \in V$. This shows that λ must be zero, which we have ruled out in our assumption. The invariance of k under the action of G is clear by the homomorphism property. $\square$

The last lemma proves that k is the gauge function of a convex polytope containing the origin. The following lemma shows that the number of extremal faces of the polytope is given by the size of the orbit of λ.

Lemma 13 *Consider k as in Lemma 12. Given any λ, there exists $x \neq 0$ such that $k(x) = \langle \lambda, x \rangle$. Moreover, $k(Ax) = \langle A\lambda, Ax \rangle$ for all $A \in G$.*

Proof To prove the first assertion, suppose that for all x, we have

$$\langle \lambda, x \rangle < k(x) = \max_{A \in G} \langle A\lambda, x \rangle.$$

Then, for any such x, for any $B \in G$, we also have

$$\langle B\lambda, x \rangle = \langle \lambda, B'x \rangle < k(B'x) = \max_{A \in G} \langle A\lambda, B'x \rangle = \max_{A \in G} \langle BA\lambda, x \rangle = k(x).$$

But that would imply $\max_{A \in G} \langle A\lambda, x \rangle < k(x)$ which is clearly a contradiction.

For the second assertion, consider x, λ such that $k(x) = \langle x, \lambda \rangle$. Now

$$k(Bx) = \max_{A \in G} \langle A\lambda, Bx \rangle = \max_{A \in G} \langle B'A\lambda, x \rangle = \max_{A \in G} \langle (B^{-1}A)\lambda, x \rangle.$$

The right hand side is maximized when $B^{-1}A = I$ which proves the lemma. □

We now show that for Coxeter groups that the polytope generated by k is simplicial. That is to say, all its extremal facets are simplices.

Lemma 14 *Consider any n-dimensional irreducible group of orthogonal matrices G. Let $\lambda \in \mathbb{R}^n$ be such that $A\lambda \neq \lambda$ for all $\lambda \neq e$. In other words, λ has no non-trivial stabilizer.*

Consider the region $\mathcal{C} = \{x \in \mathbb{R}^n : k(x) = \langle \lambda, x \rangle\}$. Then the interior of $\mathcal{C}$, given by $\mathcal{C}_0 = \{\langle \lambda, x \rangle > \langle A\lambda, x \rangle, \forall A \in G\}$, provides a fundamental region for the group. Additionally, if G is a Coxeter group, $\mathcal{C}$ is an n-dimensional closed simplicial cone.

Proof Note that $\mathcal{C}$ is the region $\{x \in \mathbb{R}^n : \langle \lambda, x \rangle \geq \langle A\lambda, x \rangle \ \forall A \in G\}$. We first show that $\mathcal{C}$ is a n-dimensional convex cone. Label the non-identity elements of the group G by $A_1, A_2, \ldots, A_N$ where $N + 1 = |G|$.

Consider the matrix Q defined by

$$Q = \begin{bmatrix} \lambda - A_1\lambda \\ \lambda - A_2\lambda \\ \vdots \\ \lambda - A_N\lambda \end{bmatrix}, \tag{26}$$

where all vectors are row vectors. This matrix Q when applied to vectors of $\mathcal{C}$ produces non-negative entries. The dimension of Q is $N \times n$. We first show that the rank of Q is n. Note that trivially the rank cannot be more than n. We show that the dimension of the kernel is zero which proves that the rank must be exactly n.

Let $\mathcal{K}$ denote the kernel, $\{x : Qx = 0\}$. Then we claim that $\mathcal{K}$ is invariant under the action of the group. This is because, $x \in \mathcal{K}$ iff $\langle \lambda, x \rangle = \langle A\lambda, x \rangle$ for all $A \in G$. But, for any $B, A \in G$, we also have

$$\langle A\lambda, Bx \rangle = \langle B^{-1}A\lambda, x \rangle = \langle \lambda, x \rangle, \quad \text{since } x \in \mathcal{K},$$
$$= \langle B^{-1}\lambda, x \rangle = \langle B'\lambda, x \rangle = \langle \lambda, Bx \rangle.$$

Thus $Bx \in \mathcal{K}$. This proves that $\mathcal{K}$ stable under the action of the group. But since the representation is irreducible, this implies that $\mathcal{K}$ is either zero or the full space. But, it is easy to see that if $\mathcal{K}$ is the full space, then λ must be zero. This proves that $\mathcal{C}$ is n-dimensional. That it is a convex cone is obvious.

Since the dimension of $\mathcal{C}$ is n and the stabilizer of λ is the identity, the interior of the cone is given by $\mathcal{C}_0$. We now show that $\mathcal{C}_0$ is a fundamental region for G by

verifying the definition. $\mathcal{C}_0$ is open by definition. For any $A \in G$, $A \neq I$, note that $A\mathcal{C}_0$ is the following set

$$\{Ax, \ x \in \mathcal{C}_0\} = \left\{y : \langle \lambda, A^{-1}y \rangle > \langle B\lambda, A^{-1}y \rangle, \ \forall B \in G \right\}$$
$$= \left\{y : \langle A\lambda, y \rangle > \langle B\lambda, y \rangle, \ \forall B \in G \right\}.$$

Thus $x \in \mathcal{C}_0 \cap A\mathcal{C}_0$ implies $\langle \lambda, x \rangle > \langle A\lambda, x \rangle > \langle \lambda, x \rangle$ which is impossible. Thus the intersection must be empty. It is also trivial to see that $\bigcup_{A \in G} A\mathcal{C} = \mathbb{R}^n$. This shows that $\mathcal{C}_0$ is a fundamental region.

For Coxeter groups we now show that $\mathcal{C}_0$ is the same region as F_λ defined in (23). Notice first that if A is a reflection along a vector r for some $A \in G$, then

$$\langle \lambda - A\lambda, x \rangle = 2\frac{\langle r, \lambda \rangle \langle r, x \rangle}{\|r\|^2}. \tag{27}$$

Now, suppose $x \in \mathcal{C}_0$. Then $\langle \lambda - A\lambda, x \rangle > 0$ for all non-identity $A \in G$, in particular, for all A which corresponds to reflections along the roots. Thus, for any root $r \in \Delta_\lambda^+$, from the above equality we get that $\langle r, x \rangle > 0$. From the definition of F_λ, it is now obvious that $x \in F_\lambda$. Thus we have shown that $\mathcal{C}_0 \subseteq F_\lambda$.

For the reverse equality, note that if $\mathcal{C}_0$ is a proper subset of F_λ, then for every $A \in G$, the set $A\mathcal{C}_0$ is a proper subset of AF_λ. But, each AF_λ is disjoint and the union of the closures of $A\mathcal{C}_0$ is the entire $\mathbb{R}^n$. This is clearly impossible. Thus, we have shown that equality holds among the two fundamental regions $\mathcal{C}_0$ and F_λ. Since F_λ is a simplicial cone, so is $\mathcal{C}_0$. Thus $\mathcal{C}$ is a closed simplicial cone. $\square$

The connection between simplicial cones and Exponential distributions is made precise in the next lemma.

Lemma 15 *Consider a sequence of simplicial cones $C_1, C_2, \ldots, C_r$ which are open, disjoint, and the closure of their union is the whole space. Let k be a non-negative, irreducible, positively homogeneous function such that k is linear on each C_i. That is, $k(x) = \langle \lambda_i, x \rangle$, for all $x \in C_i$, for some sequence of vectors $\lambda_1, \lambda_2, \ldots, \lambda_r$ which may not be all distinct.*

Let X be a random variable whose density with respect to the Lebesgue measure on $\mathbb{R}^n$ is proportional to $e^{-2k(x)}$. Let B_i, $i = 1, 2, \ldots, r$, be any set of invertible matrices such that for each i, the matrix B_i maps the cone C_i onto the n-dimensional quadrant. Also, let $\alpha_1(i), \ldots, \alpha_n(i)$ be the coefficients in the unique representation $\lambda_i' = \alpha'(i)B_i$. Then, the random vector $Y = (Y_1, Y_2, \ldots, Y_n)$, where

$$Y_j := \alpha_j(i)\langle (B_i)_{j*}, X \rangle, \quad \text{if } X \in C_i,$$

is a vector of iid Exponential(2) random variables. Here $(B_i)_{j*}$ denotes the jth row of the matrix B_i.

Proof Since the cones C_i's are simplicial, the existence of the matrices B_i follows from the definition. Moreover, it follows from the definition of $\alpha(i)$ that if $x \in C_i$ and

$z = B_i x$, then $k(x) = \langle \lambda_i, x \rangle = \sum_j \alpha_j(i) z_j$ irrespective of i. Let B_i^* be the matrix given by

$$\left(B_i^* \right)_{j*} = \alpha_j(i)(B_i)_{j*}, \quad j = 1, 2, \ldots, d.$$

Since the transformations are piecewise linear, it follows that

$$Y := \sum_j B_j^* X 1(X \in C_j)$$

has a density proportional to $\exp\{-2\sum_j y_j\}$ over the quadrant $\{y : y_1 \geq 0, y_2 \geq 0, \ldots, y_d \geq 0\}$. This immediately identifies itself as the joint density of iid exponentials with rate two. This proves part (2). □

Proposition 16 *Let G denote a Coxeter group acting on $\mathbb{R}^n$. For $\lambda \in \mathbb{R}^n$, $\lambda \neq 0$, let $k(x) = \max_{A \in G} \langle A\lambda, x \rangle$. Let ν be the probability measure with unnormalized density $\exp\{-2k(x)\}$ on $\mathbb{R}^n$. Then the following statements hold true.*

(1) *When the stabilizer of λ is trivial, the conic hull of the finite set $\{\lambda - A\lambda, \, A \in G\}$ contains n linearly independent generating vectors $\{\eta_1, \eta_2, \ldots, \eta_n\}$. That is, every other vector in the set can be expressed as a linear combination of the generators with non-negative coefficients.*

(2) *Let X denote a random variable with distribution ν. Also let α_i denote the unique positive coefficient of η_i in the expansion*

$$\lambda = \sum_{i=1}^n \alpha_i \eta_i.$$

For $i = 1, 2, \ldots, n$, define the change of variable

$$Y_i = \alpha_i \langle A\eta_i, X \rangle, \quad \text{when } k(X) = \langle A\lambda, X \rangle.$$

Then, the vector $Y = (Y_1, Y_2, \ldots, Y_n)$ are iid Exponential(2) random variables.

Proof To prove the first assertion, we use Lemma 14. Assume first that the stabilizer of λ is trivial. Then, Lemma 14 tells us that the cone

$$C = \left\{ x : \langle \lambda - A\lambda, x \rangle \geq 0 \right\} \tag{28}$$

is an n-dimensional simplicial cone. Hence there exists exactly n many linearly independent generators among the set $\{\lambda - A\lambda, \, A \in G\}$ such that every other vector is a linear combination with non-negative coefficients. The rest of the result follows directly from Lemma 15. Notice that the coefficients of λ in the expansion $\lambda = \sum_{i=1}^n \alpha_i \eta_i$ are positive by Farkas lemma. This is because, for any non-zero $y \in C$, the inner product $\langle \lambda, y \rangle = k(y) > 0$ by irreducibility and non-negativity of k.

When λ has a non-trivial stabilizer, the cone in (28) is a union of several simplicial cones. The simplest way to see this is to take a sequence λ_l which have no non-trivial stabilizers and which converges to λ. The component cones are then given by the limits of the sequence of simplicial cones generated by them. In any case Lemma 15

still holds, however, the vectors $\{\eta_1, \ldots, \eta_n\}$ have to be determined by the limiting procedure.   □

Combining the results above, the proof of Proposition 2 in the Introduction now follows easily.

3.2 Examples

Let us consider some examples of consequences of Proposition 2.

Example 1 (Rank-based interactions) Brownian motions with rank-based interactions have been considered in (3). Clearly the drift function b is constant over finitely many cones determined by the permutation generated by the ordered coordinates. Let $x[1] \leq x[2] \leq \cdots \leq x[n]$ denote the coordinates of an n-dimensional vector arranged in increasing order. It is easy to see that

$$k(x) = -\sum_{i=1}^{n} \delta_i x[i],$$

is a positively homogeneous function which is not irreducible since it takes a constant value over the linear span of the vector of all ones. However, if we let H be the subspace orthogonal to 1, then k splits additively as

$$k(x) = -\sum_{i=1}^{n} \delta_i \big(x[i] - \bar{x}\big) + n\bar{\delta}\bar{x}.$$

Let $k_1 : H \to \mathbb{R}$ denote the restriction of k to H, then it is clear that $k_1(x) = k_1(A_\sigma x)$ for any permutation matrix A_σ. Now, the group of permutation matrices acting on H is well known to be irreducible and generated by reflections along $e_{i+1} - e_i$ for $i = 1, 2, \ldots, n-1$. This is just a restatement of the fact that every permutation can be written as a product of transposes. Thus, it is a Coxeter group, often denoted by $\mathcal{A}_{n-1}$. Thus the conclusions of Proposition 12 applies and k_1 is irreducible if

$$k_1(x) = \max_{A_\sigma} \langle -A_\sigma \delta, x \rangle. \tag{29}$$

This condition is equivalent to the condition $\delta_1 \geq \delta_2 \geq \cdots \geq \delta_n$.

Hence, from Proposition 10 it also follows that the projection $P_H X_t$ of the diffusion on to H has an invariant distribution whose density with respect to Lebesgue measure on H is proportional to $\exp(-2k_1(x))$.

We now apply Proposition 16. One can see that the set of vectors $-\delta + A_\sigma \delta$, as σ ranges over permutations, contains positive multiples of vectors $e_{i+1} - e_i$, since they correspond to the transposition of i and $i + 1$. These $n - 1$ linearly independent vectors are the conic extremes of the set. Thus, by Proposition 16, the spacings $X[i + 1] - X[i]$ are independent Exponential random variables under the invariant distribution. The correct rates can be easily verified.

In this example it is easy to see the shortfall of the sufficient condition (29). The drift function is constant over a fundamental region $F = \{x \in H : x_1 < x_2 < \cdots < x_n\}$ which is clearly simplicial. Now $k(x)$ is irreducible if and only if the unit ball generated by k is compact. By symmetry, we can restrict our attention to F. Since F is simplicial we can apply a suitable linear transformation to map it to the positive quadrant. Thus, it can be easily verified that if $\bar{\delta}$ denoted the average of the coordinates of δ, the intersection of the unit ball with F is compact if and only if

$$\alpha_k := \sum_{i=1}^{k} \left(\delta_i - \bar{\delta}\right) > 0, \quad \forall i = 1, 2, \ldots, n-1.$$

This is precisely the condition derived by Pal & Pitman in [17] using the theory of reflected Brownian motions and is weaker than the sufficient condition that the coordinates of δ decreases with increasing values of the coordinates of x.

Example 2 (Sign-rank-based interactions) An example of interactions similar to rank-based can be generated by allowing both the rank and signs coordinates to determine the drift. As before, we start with the n-dimensional SDE:

$$dX_t = b(X_t)\,dt + dW_t, \tag{30}$$

where W_t is an n-dimensional Brownian motion. Suppose that the drift function takes finitely many values, is scale invariant, and $b(Ax) = Ab(x)$ whenever A is either a permutation matrix or a diagonal matrix with each diagonal entry being plus or minus one. Thus, not only that the values of the drift get permuted whenever the coordinates get permuted, but also the sign of the drift changes with the sign of the corresponding coordinate.

The group generated by the collection of permutation matrices and the diagonal matrices of sign flips is a Coxeter group denoted by $\mathcal{B}_n$. Please see pp. 66–71 of [12] for more details. We can safely apply Proposition 2. Thus, the n-dimensional process under such a sign-rank-based interaction is recurrent if there is a vector $\lambda \in \mathbb{R}^n$, $\lambda \neq 0$, such that

$$k(x) = -\langle x, b(x) \rangle = \max_{A \in \mathcal{B}_n} \langle A\lambda, x \rangle.$$

If we restrict the above condition to the cone $\{x : 0 < x_1 < x_2 < \cdots < x_n\}$ we see that the vector of drifts $b(x) = -\lambda$ where λ satisfies that $0 \leq \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$.

When this does hold true, X has an unique long term stationary distribution. To find the decomposition of this probability distribution in terms of independent Exponentials we consider a δ all of whose coordinates are non-zero and distinct. That is, it has a trivial stabilizer subgroup in $\mathcal{B}_n$. Consider the conic hull generated by the set of vectors $\{\lambda - A\lambda, \ A \in \mathcal{B}_n\}$. As in the case of rank-based interactions one can see that the generators of the conic hull are positive multiples of the vectors e_1 and $\{e_{i+1} - e_i, \ i = 1, 2, \ldots, n-1\}$.

Now we apply the final conclusion of Proposition 2. To get the vector of Exponentials under the invariant distribution, note that

$$k(x) = \sum_{i=1}^{n} \lambda_i \, |x| \, [i],$$

where $|x|[1] \le |x|[2] \le \cdots \le |x|[n]$ are the ordered values of the absolute values of the coordinates $(|x_1|, |x_2|, \ldots, |x_n|)$.

Thus, from Proposition 2 it follows that the random vector $(|X|[1], |X|[i+1] - |X|[i], \, i = 1, 2, \ldots, n-1)$ are distributed as independent Exponentials.

To compute the rates of these Exponentially distributed random variables, one needs to compute the coefficient of λ with respect to the basis vector e_1 and $\{e_{i+1} - e_i, \, i = 1, 2, \ldots n-1\}$. A simple computation leads us to the conclusion that the corresponding vector of rates of these Exponentials are given by

$$\left(2 \sum_{s=j}^{n} \lambda_s, \quad j = 1, 2, \ldots, n \right).$$

Example 3 (Constrained sign-rank-based interactions) This is an interesting class of constrained sign-rank-based interactions where not all sign changes of coordinates affect the drift vector. Consider again the stochastic differential equation (30). Suppose that the drift function b takes finitely many values, is scale invariant, and $b(Ax) = Ab(x)$ for all permutation matrices A and all diagonal matrices with each diagonal entry being positive or negative one *with the constraint* that there are even number of negative ones. Thus, the sign of the drift vector changes when either the ordering of coordinates change or when pairs of coordinates have flipped their signs.

The groups generated by the permutation matrices and the diagonal matrices with even number of sign flips is clearly a subgroup of the $\mathcal{B}_n$. They form, in fact, a family of Coxeter groups, usually denoted by $\mathcal{D}_n$ where n denotes the dimension of the underlying space. We again refer the reader to pp. 66–71 of [12] for more details.

We apply Proposition 2. Thus, the n-dimensional process under such a constrained sign-rank-based interaction is recurrent if there is a vector $\lambda \in \mathbb{R}^n$, $\lambda \ne 0$, such that

$$k(x) = -\langle x, b(x) \rangle = \max_{A \in \mathcal{D}_n} \langle A\lambda, x \rangle.$$

The above condition is more difficult to analyze than the previous examples. One can show using known results about the fundamental root systems of $\mathcal{D}_n$ (p. 71 in [12]) that the drift is determined by the fact that over the cone

$$\{x : 0 < x_1 + x_2, \quad x_1 < x_2 < \cdots < x_n\}, \tag{31}$$

the drift is a constant $b(x) = -\lambda$ where λ satisfies that $0 \le \lambda_1 + \lambda_2$ and $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$. This cone is actually a fundamental region for the group $\mathcal{D}_n$. Thus, the drift vector is now determined over entire $\mathbb{R}^n$ by the property $b(Ax) = Ab(x)$ for all $A \in \mathcal{D}_n$.

Under this condition the process has a long term stationary distribution. To find what functions turn out to be independent Exponentials, we need to understand, for a given x, what unique $A \in \mathcal{D}_n$ will produce $k(x) = \langle \lambda, Ax \rangle$. Clearly, this will happen if A is chosen such that Ax belongs to the cone (31). There are two cases to consider. One, when the number of coordinates of x that are negative is even. In this case, one simply flips the signs of these coordinates, and then ranks the absolute values to get a vector in (31). Both these actions are permissible since they correspond to multiplication by matrices in $\mathcal{D}_n$. The other case is when x has an odd number of negative coordinates. First, one has to flip the sign of all the negative coordinates except the least negative one and then rank all the coordinates. In this ordering, either the absolute value of the second least negative coordinate is less than the least positive coordinate in which case, the resulting vector is in (31). Or, it is not, in which case we need to compare the least negative coordinate with the least positive coordinate. Their sum is either positive or negative, and we make the appropriate (zero or two) sign flips to get the right transformation. Let $H(x)$ be the resulting vector produced by the above procedure.

Under this stationary distribution, the vector of random vector $H(X_1, X_2, \ldots, X_n)$ is distributed as n independent Exponentials. Furthermore, as in the case of sign-rank interactions, one can work out the linear algebra to compute that the corresponding vector of rates of these Exponentials are given by

$$\left(\sum_{s=1}^{n} \lambda_s, \quad -\lambda_1 + \sum_{s=2}^{n} \lambda_s, \quad 2 \sum_{s=j}^{n} \lambda_s, \quad j = 3, 4, \ldots, n \right).$$

3.3 An Example of Irregular Interaction

In this subsection we consider an example of interacting Brownian motions with drift functions that are still piecewise constant on cones, but are not consistent with any group action.

Consider a graph $\mathcal{G}$ on n vertices where the vertices are labeled by $\{1, 2, \ldots, n\}$. The edge between i and j have an associated edge weight β_{ij}, which is zero if there is no edge between the two vertices.

Consider the SDE on $\mathbb{R}^n$ given by

$$dX_t(i) = \sum_{j=1}^{n} \beta_{ij} \text{sign}\big(X_t(j) - X_t(i) \big) dt + dW_t(i), \qquad (32)$$

where, as before, $W = (W(1), W(2), \ldots, W(n))$ is an n-dimensional Brownian motion.

When all the edge weights are non-negative, the model can be described by saying that the Brownian motions, which are indexed by the vertices of the graphs, get attracted towards one another. The constants β_{ij} measure the strength of their attraction.

The appropriately defined drift function $b(x)$ is piecewise constant on the family of cones $C_\pi := \{x : x_{\pi(1)} \le x_{\pi(2)} \le \cdots \le x_{\pi(n)}\}$, where π ranges over all permutations

of n labels. However, it might not satisfy the condition that $b(A_\pi x) = A_\pi b(x)$ where A_π is the permutation matrix corresponding to π.

Note that the drift function b is the negative of the gradient of the positively homogeneous function

$$k(x) = \sum_{i<j} \beta_{ij} \, |x_i - x_j|, \quad \forall x \in \mathbb{R}^n.$$

It can be easily verified (and intuitive) that if β_{ij}'s are non-negative and $\mathcal{G}$ is connected, the function k is irreducible when restricted to the subspace H orthogonal to the vector of all ones (which we denote by 1). Then the conclusions of Proposition 10 applies. In particular, if we define

$$k'(x) = k(x) + \big|\langle x, 1\rangle\big|$$

then k' is integrable and both the probability measures induced by functions $\exp(-2k)$ and $\exp(-2k')$ on H must be the same.

Assume $\beta_{ij} \geq 0$ and $\mathcal{G}$ is connected. For convenience absorb the factor of two in $\exp(-2k)$ in the definition of β. Let μ_n be the probability measure whose unnormalized density is given by

$$\exp\left(-\sum_{i<j} \beta_{ij}|x_i - x_j| - \left|\sum_i x_i\right|\right). \tag{33}$$

What properties of the probability measure can we explicitly describe? Clearly, any deep inspection of such a general family is extremely difficult. We will improve our chances if we restrict the edge weights to the following class. Consider n positive constants $m_1, m_2, \ldots, m_n$. Let $\beta_{ij} = m_i m_j$ for all pairs i, j. In particular the graph is complete. One can think of m_i as the *mass* of the ith particle, and hence the strength of attraction between particles i and j is proportional to the product of their masses. In fact, the case when all the m_i's are equal to one has been dealt with in [3] where the model was named the one-dimensional gravity model.

Now suppose $X = (X_1, \ldots, X_n)$ follows the p.d.f.

$$C_n \exp\left(-\sum_{i<j} m_i m_j |x_i - x_j| - \left|\sum_i x_i\right|\right),$$

where C_n is the normalizing constant. Let $M = \sum_{i=1}^n m_i$ be the total mass of the system. For each i, let $Y_i = X[i]$ (the ith smallest coordinate) and $\Pi(i)$ be the (random) index j such that $Y_i = X_j$. The joint p.m.f. of (Π, Y) at a point (π, y) is

$$C_n \exp\left(-\sum_{i<j} m_{\pi_i} m_{\pi_j} (y_j - y_i) - \left|\sum_i y_i\right|\right),$$

where $\pi = (\pi_1, \ldots, \pi_n)$ is a permutation of $\{1, \ldots, n\}$ and $y_1 < \cdots < y_n \in \mathbb{R}$. Now let $\Delta_i = Y_{i+1} - Y_i$, $i = 1, \ldots, n-1$. For each i, and each $\pi \in S_n$, let

$$F_i(\pi) = \frac{\sum_{j=1}^i m_{\pi_j}}{M}.$$

Then $M^{-1} \sum_{j=i+1}^n m_{\pi_j} = 1 - F_i(\pi)$, and $F_n(\pi) \equiv 1$. The joint density of (Π, Δ, Y_1) at a point (π, δ, y_1) is

$$C_n \exp\left(-\sum_{i<j} m_{\pi_i} m_{\pi_j} \left(\sum_{k=i}^{j-1} \delta_k\right) - \left|\sum_{i=2}^n \sum_{j=1}^{i-1} \delta_j + n y_1\right|\right)$$

$$= C_n \exp\left(-M^2 \sum_{i=1}^{n-1} F_i(\pi)(1 - F_i(\pi))\delta_i - \left|\sum_{i=2}^n \sum_{j=1}^{i-1} \delta_j + n y_1\right|\right).$$

Now we can easily integrate out Y_1 to get the joint density of (Π, Δ):

$$C \exp\left(-M^2 \sum_{i=1}^{n-1} F_i(\pi)\left(1 - F_i(\pi)\right)\delta_i\right),$$

where C is now a different constant. Thus, conditional on $\Pi = \pi$, $Y_1, \ldots, Y_n$ are independent, with $Y_i \sim \mathrm{Exp}(M^2 F_i(\pi)(1 - F_i(\pi)))$. It is easy to see from this observation that the marginal p.m.f. of Π must be

$$C(m) \prod_{i=1}^{n-1} \frac{1}{F_i(\pi)(1 - F_i(\pi))}, \tag{34}$$

where $\pi = (\pi_1, \ldots, \pi_n)$ is any permutation of $\{1, \ldots, n\}$ and $C(m)$ is the normalizing constant that depends on the values of $m_1, \ldots, m_n$.

If $m_1 = m_2 = \cdots = m_n$, then this is the uniform distribution on S_n. Otherwise, it is a non-uniform distribution on the set of permutations. It is not immediate what sort of distributions on the space of permutations the probability mass function (34) induces. To the best of our knowledge this is not a member of any of the existing families of distributions on the group of permutations, although it is closely connected to what is known as the Luce probability model on permutations. Please see [6, p. 174] (and the references within) for the details. We describe below the following reformulation of Luce model by Gordon [11]. Let $U_1, U_2, \ldots, U_n$ be independent Exponentials with rates $w_1, w_2, \ldots, w_n$. Then the probability on the permutations induced by the order statistics of $U_1, U_2, \ldots, U_n$ belongs to the Luce model. Exploiting the memoryless property of the Exponentials, one can write down this probability measure explicitly: for any permutation π, the probability of π is

$$P\left(U_{\pi_1} \leq \cdots \leq U_{\pi_n}\right)$$

$$= \left(\frac{w_{\pi_1}}{w_{\pi_1} + \cdots + w_{\pi_n}}\right)\left(\frac{w_{\pi_2}}{w_{\pi_2} + \cdots + w_{\pi_n}}\right) \cdots \left(\frac{w_{\pi_{n-1}}}{w_{\pi_{n-1}} + w_{\pi_n}}\right).$$

The Luce model prefers orderings where the larger the w_i is, the smaller is the rank of the corresponding U_i. Let us call the Luce model corresponding to the choice of weights $m_1, m_2, \ldots, m_n$ by $q(\cdot)$. Then, it is not difficult to verify that the probability mass function given in (34) is proportional to $q(\pi)q(\hat{\pi})$, where $\hat{\pi}$ is the permutation satisfying $\pi_i + \hat{\pi}_i = n + 1$ for every $i = 1, \ldots, n$.

Clearly the p.m.f. is large when $F_i(\pi)$ is close to zero or one for most values of i. The intuition from gravity predicts that heavier particles should be close and should avoid being too high or too low in rank. We now show this to be true in a particularly simple case when there is a single distinguished particle. Suppose that $m_1 = \alpha$ and $m_2 = m_3 = \cdots = m_n = 1$. We consider the joint distribution as before

$$\frac{d\mu}{dx} = C_n \exp\left(-\sum_{i<j} m_i m_j |x_i - x_j| + \left|\sum x_i\right|\right).$$

We are interested in the derived joint distribution of the ranks of each particle given by the general expression in the previous section.

Let us compute the distribution of the rank of the first particle which is distinguished from the others due to a different mass.

$$P(\text{rank of the first particle} = j) = P\big(\Pi(j) = 1\big)$$

$$= \sum_{\sigma:\sigma(j)=1} C \prod_{i=1}^{n-1} \frac{1}{F_i(\sigma)(1 - F_i(\sigma))}, \tag{35}$$

where C is a constant depending on n and α. In the following text, we will freely use C as the normalizing constant keeping in mind that the constants might be different from one another although they only depend on n and α.

Now, there are $(n - 1)!$ many permutations σ such that $\sigma(j) = 1$. For any of them

$$F_i(\sigma) = \begin{cases} \frac{i}{\alpha+n-1} & \text{if } i < j, \\ \frac{\alpha+i-1}{\alpha+n-1} & \text{otherwise.} \end{cases}$$

And thus we can rewrite (35) as

$$P(\Pi(j) = 1) = C \frac{1}{\prod_{i=1}^{j-1} \frac{i}{\alpha+n-1}(1 - \frac{i}{\alpha+n-1}) \prod_{i=j}^{n-1} \frac{\alpha+i-1}{\alpha+n-1}(1 - \frac{\alpha+i-1}{\alpha+n-1})}$$

$$= C \frac{1}{(j-1)!(n-j)! \prod_{i=1}^{j-1}(\alpha+n-i-1) \prod_{i=j}^{n-1}(\alpha+i-1)}$$

$$= C \binom{n-1}{j-1} \frac{\alpha(\alpha+1)\cdots(\alpha+n-j-1)\,\alpha(\alpha+1)\cdots(\alpha+j-2)}{\alpha(\alpha+1)\cdots(\alpha+n-2)\,\alpha(\alpha+1)\cdots(\alpha+n-2)}$$

$$= C \binom{n-1}{j-1} \alpha(\alpha+1)\cdots(\alpha+n-j-1)\,\alpha(\alpha+1)\cdots(\alpha+j-2).$$

$$\tag{36}$$

We can immediately infer from the previous expression the following fact:

$$\frac{P(\text{rank of the first particle} = j + 1)}{P(\text{rank of the first particle} = j)} = \frac{P(\Pi(j+1) = 1)}{P(\Pi(j) = 1)}$$

$$= \frac{(n-j)(\alpha + j - 1)}{j(\alpha + n - j - 1)}.$$

In other words

$$P(\text{rank of the first particle} = j + 1) > P(\text{rank of the first particle} = j)$$

$$\text{iff} \quad (n-j)(\alpha + j - 1) > j(\alpha + n - j - 1),$$

$$\text{iff} \quad n(\alpha - 1) + (n - \alpha + 1)j - j^2 > (\alpha + n - 1)j - j^2,$$

$$\text{iff} \quad 2(\alpha - 1)j < n(\alpha - 1).$$

Thus, if α is more than 1, the probability of the rank being j increases till $j = \lceil n/2 \rceil$, and then strictly decreases. Clearly, the most likely position for the heavier particle is going to be the median. On the other hand, if $\alpha < 1$, just the opposite happens, and we are likely to see the lighter particle either at the top or trailing behind.

The probability computed in (36), although seemingly unfriendly, is a very familiar object. Consider a Polya's urn scheme which has α red balls and α black balls. We play a game where at each step we pick a ball at random and replace it in the urn with a ball of the same color. It is well known (see Feller [8]) that if we play this game for $n - 1$ steps the probability we pick exactly $j - 1$ red balls is given by

$$\binom{n-1}{j-1} \frac{\alpha(\alpha + 1) \cdots (\alpha + n - j - 1) \, \alpha(\alpha + 1) \cdots (\alpha + j - 2)}{2\alpha(2\alpha + 1)(2\alpha + 2) \cdots (2\alpha + n - 2)}.$$

If we compare the previous expression with (36), the differences are merely in the expression of the normalizing constants. Thus, if $\sigma(1) = \Pi^{-1}(1)$ is the rank of the first particle, it is clear that $\sigma(1) - 1$ has the same distribution as the number of red balls picked in a Polya's urn scheme run for $n - 1$ steps.

Proposition 17 *For any $\alpha > 0$, consider the SDE* (32) *with a distribution of mass such that the mass of the first particle being α and the rest being of mass 1. Then the sequence of random variables $\sigma(1)/n$ converges weakly to the Beta(α, α).*

Proof The proof follows from known results about Polya's urn. See Feller [8]. □

References

1. Arguin, L.-P.: A dynamical characterization of Poisson-Dirichlet distributions. Electron. Commun. Probab. **12**, 283–290 (2007)

2. Banner, A., Fernholz, R., Karatzas, I.: Atlas models of equity markets. Ann. Appl. Probab. **15**(4), 2296–2330 (2005)
3. Chatterjee, S., Pal, S.: A phase transition behavior for Brownian motions interacting through their ranks. Probab. Theory Rel. Fields. (2008, to appear)
4. Coxeter, H.S.M.: Discrete groups generated by reflections. Annal. Math. **35**, 588–621 (1934)
5. Coxeter, H.S.M.: Regular Polytopes, 2nd edn. Macmillan & Co., New York (1963)
6. Diaconis, P.: Group Representations in Probability and Statistics. IMS Lecture Notes—Monograph Series, vol. 11 (1988)
7. Durrett, R.: Stochastic Calculus—A Practical Introduction. Probability and Stochastics Series. CRC Press, Boca Raton (1996)
8. Feller, W.: An Introduction to Probability Theory and Its Applications, vol. 1, 3rd edn. Wiley, New York (1968)
9. Fukushima, M., Oshima, Y., Takeda, M.: Dirichlet Forms and Symmetric Markov Processes. Studies in Mathematics, vol. 19. de Gruyter, Berlin (1994)
10. Goodman, J.E., O'Rourke, J. (eds.): Handbook of Discrete and Computational Geometry, 2nd edn. CRC Press, Boca Raton (2004)
11. Gordon, L.: Successive sampling in large finite populations. Ann. Stat. **11**, 702–706 (1983)
12. Grove, L.C., Benson, C.T.: Finite Reflection Groups, 2nd edn. Graduate Texts in Mathematics, vol. 99. Springer, Berlin (1985)
13. Humphreys, J.E.: Reflection Groups and Coxeter Groups. Cambridge Studies in Advanced Mathematics, vol. 29. Cambridge University Press, Cambridge (1990)
14. Jourdain, B., Malrieu, F.: Propagation of chaos and Poincaré inequalities for a system of particles interacting through their cdf. Ann. Appl. Probab. **18**(5), 1706–1736 (2008)
15. Kane, R.: Reflection Groups and Invariant Theory. CMS Books in Mathematics, Canadian Mathematical Society. Springer, Berlin (2001)
16. Karatzas, I., Shreve, S.: Trivariate density of Brownian motion, its local and occupation times, with application to stochastic control. Ann. Probab. **12**(3), 819–828 (1984)
17. Pal, S., Pitman, J.: One-dimensional Brownian particle systems with rank dependent drifts. Ann. Appl. Probab. **18**(6), 2179–2207 (2008)
18. Ruzmaikina, A., Aizenman, M.: Characterization of invariant measures at the leading edge for competing particle systems. Ann. Probab. **33**(1), 82–113 (2005)
19. Shreve, S.: Reflected Brownian motion in the 'bang-bang' control of Brownian drift. SIAM J. Control Optim. **19**, 469–478 (1981)
20. Sinai, Ya.G., Soloveychik, M.R.: One-dimensional classical massive particle in the ideal gas. Commun. Math. Phys. **104**, 423–443 (1986)
21. Szász, D., Tóth, B.: Bounds for the limiting variance of the "heavy particle" in R^1. Commun. Math. Phys. **104**, 445–455 (1986)
22. Warnecke, G.: On singular and bang-bang processes in optimal control. In: System Modelling and Optimization. Lecture Notes in Control and Information Sciences, vol. 59, pp. 345–353. Springer, Berlin (1984)
23. Williams, R.J.: Reflected Brownian motion with skew symmetric data in a polyhedral domain. Probab. Theory Related Fields **75**(4), 459–485 (1987)