Statistics and Probability Letters 235 (2026) 110747

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Poisson–Dirichlet approximation for the stationary distribution of the inclusion process

Han L. Gan

University of Waikato, New Zealand

ARTICLE INFO

Keywords:
Inclusion process
Poisson–Dirichlet distribution
Stein's method

ABSTRACT

We consider the approximation of the stationary distribution of the finite inclusion process with the Poisson–Dirichlet distribution. Using Stein's method, we derive an explicit bound for the approximation error, which is of order $1/N$ in the thermodynamic limit. The results are achieved from a minor modification to Stein's method for Poisson–Dirichlet distribution approximation developed in Gan and Ross, (2021). The derivatives used on test functions in Gan and Ross, (2021) were directional type derivatives specifically chosen for their measure preserving properties. Depending upon the application, these derivatives can prove cumbersome. In this note, we show that for certain test functions we can instead use more traditional derivatives, which simplifies the bounds for the Stein factors and is more amenable to the approximation of the inclusion process.

1. Introduction

The one parameter Poisson–Dirichlet distribution with parameter $\theta > 0$, PD(θ), is a probability measure on the infinite dimensional ordered simplex,

$$\nabla_\infty := \{(p_1, p_2, \dots) : p_1 \geq p_2 \geq \dots, \sum_{i=1}^{\infty} p_i = 1\}.$$

The distribution lacks an explicit form, but it can be represented in a variety of ways, such as a limit of Dirichlet distributions or via a stick breaking process, see Feng (2010) for more details and examples. As working with order statistics can be technically challenging, rather than using non-increasing ordering, Gan and Ross (2021) use a labelled version of the process with labels from a compact metric space E, and define the Dirichlet process in the following manner.

Definition 1.1 (*Dirichlet Process*). Let $\theta > 0$ and let π be a probability measure on a compact metric space E. Let $(P_1, P_2, \dots) \sim \mathrm{PD}(\theta)$ be independent of $\xi_1, \xi_2, \dots$, which are i.i.d. with distribution π. Define the probability measure on $M_1 := M_1(E)$, the space of probability measures on E, by

$$\mathrm{DP}(\theta, \pi) := \mathscr{L}\Big(\sum_{i=1}^{\infty} P_i \delta_{\xi_i}\Big). \tag{1.1}$$

In Gan and Ross (2021), Stein's method for Dirichlet process approximation and Poisson–Dirichlet approximation was developed and a general theorem for bounding the approximation error between a random measure of interest and a Dirichlet process was derived. Motivated by the approximation of the stationary distribution of the inclusion process, we make a minor refinement to the

E-mail address: han.gan@waikato.ac.nz.

https://doi.org/10.1016/j.spl.2026.110747
Received 17 December 2025; Received in revised form 22 March 2026; Accepted 23 March 2026

Available online 27 March 2026

framework of Gan and Ross (2021). This refinement makes the framework more amenable to the application for approximation of the stationary distribution of the inclusion process, and moreover may make Poisson–Dirichlet and Dirichlet process approximation using Stein's method more accessible and easier to use in further applications.

1.1. Application to the inclusion process

Originating in Giardinà et al. (2007, 2009), the discrete inclusion process is a stochastic interacting particle system with N particles on L sites, which we can index as $\{1,\dots,L\}$. We define the set of configurations corresponding to the proportions of the N particles on L sites as

$$\mathcal{G}_L = \left\{ \mu \in [0,1]^L : \sum_{i=1}^{L} \mu_i = 1 \right\},$$

where μ_i is the proportion of the N particles in site i. As the indices for the sites are arbitrary, we can embed the atomic masses of the process onto a lattice of width $1/L$ on $[0,1]$. If we set $x_i = i/L$, for any $\theta > 0$, the generator of this discrete inclusion process is as follows,

$$\mathcal{A}_1 f(\mu) = \sum_{i,j=1}^{L} N\mu_i \left(N\mu_j + \frac{\theta}{L} \right) \left[f\left(\mu - \frac{1}{N}\delta_{x_i} + \frac{1}{N}\delta_{x_j} \right) - f(\mu) \right]. \tag{1.2}$$

The process can be briefly described in the following manner. At unit rate, for each (ordered) pair of particles on sites (x_i, x_j) a particle is moved from site x_i to site x_j. In addition, for each particle in the system, at rate θ/L independently, the particle is removed and a new particle is added at one of the L sites chosen uniformly at random. Notably this model is also equivalent to a population genetics model, a type of Moran model with uniform mutation.

The limiting behaviour of this process has been extensively studied in a variety of different limiting regimes, see Chleboun et al. (2022), Jatuviriyapornchai et al. (2020), Bianchi et al. (2017) and Kim (2021) for example. Let W follow stationary distribution associated with $\mathcal{A}_1$. In this paper, our goal is to assess the distribution of the masses in W, and compare this to distribution of the masses of the stationary distribution of the Dirichlet process with $E = [0,1]$ and $\pi = \mathrm{U}[0,1]$. We now state the main approximation result. Note that the definition of the family of test functions $\mathcal{H}_2$ is formally defined in Section 2 and $\|\cdot\|_\infty$ denotes the supremum norm.

Theorem 1.2. *For any N, L and $\theta > 0$, let W be distributed as the stationary probability measure associated with $\mathcal{A}_1 f$ in (1.2). Set $E = [0,1]$, $\pi = \mathrm{U}[0,1]$ and let $Z \sim \mathrm{DP}(\theta,\pi)$. Then for any $h \in \mathcal{H}_2$ such that $h(\mu) = \langle \phi, \mu^k \rangle$,*

$$|\mathbb{E}h(W) - \mathbb{E}h(Z)| \le \left[\frac{k(k-1)}{2L(\theta+1)} + \frac{2k(k-1)\theta}{N(\theta+1)} + \frac{8k(k-1)(k-2)}{9N(\theta+2)} \right] \|\phi\|_\infty.$$

This bound is general, holds for any choice of N, L and θ, and will converge to zero as $N \to \infty$ when L scales with N in a positive manner. In Jatuviriyapornchai et al. (2020, Section 3), it is stated that if $d \to 0$ and $dL \to \alpha > 0$, then the limiting distribution is the Poisson–Dirichlet distribution. In the notation of this manuscript, $d = \theta/L$, and hence our bound is entirely consistent with existing results. Furthermore, in the so called thermodynamic limit, where $N, L \to \infty$ such that N/L converges to a positive constant θ, the bound is therefore of order $\frac{1}{N}$.

The remainder of the paper will consist of two sections. In Section 2 we will briefly outline Stein's method for Poisson–Dirichlet and Dirichlet process approximation and refine the bounds of Gan and Ross (2021) for the derivatives of the solutions to the Stein equation for test functions $h \in \mathcal{H}_2$. In the final section we prove Theorem 1.2.

2. Refining the Stein method for the Dirichlet process

Pioneered in Stein (1972) for Normal approximation, Stein's method is a powerful probabilistic tool used to explicitly bound the distributional distance between two distributions. Since the original seminal paper, it has been developed for a wide variety of distributions. For many examples and applications, see for example the surveys or monographs (Ross, 2011; Chatterjee, 2014; Barbour and Chen, 2005; Ley et al., 2017).

We first give a brief summary of the Stein's method approach used in Gan and Ross (2021). For any random probability measure W, and $Z \sim \mathrm{DP}(\theta,\pi)$, our goal is to bound quantities of the form $|\mathbb{E}h(W) - \mathbb{E}h(Z)|$, for all functions h in a sufficiently rich class of test functions. To achieve this, we begin by characterising $\mathrm{DP}(\theta,\pi)$ as the unique stationary distribution of a Fleming–Viot process (Fleming and Viot, 1979) with generator,

$$\mathcal{A}_2 f(\mu) = \frac{\theta}{2} \int_E \partial_x f(\mu)(\pi - \mu)(dx) + \frac{1}{2} \int_{E^2} \big(\mu(dx)\delta_x(dy) - \mu(dx)\mu(dy) \big) \partial_{xy} f(\mu). \tag{2.1}$$

The specific variant of Stein's method we use is often called the generator method, first introduced in Barbour (1988, 1990) and Götze (1991). The general idea of the approach is that as Z follows the stationary distribution of $\mathcal{A}_2$, then $\mathbb{E}\mathcal{A}_2 f(Z) = 0$ for all f. If a random measure W is close to Z in distribution, then one would expect $\mathbb{E}\mathcal{A}_2 f(W) \approx 0$. The next step is then to choose specific test functions f that quantify the distributional distance between W and Z. More precisely, for any function h from some family of functions $\mathcal{H}$, we solve for the function f_h that satisfies

$$\mathcal{A}_2 f_h(\mu) = h(\mu) - \mathbb{E}h(Z). \tag{2.2}$$

We then set $\mu = W$, which implies $|\mathbb{E}h(W) - \mathbb{E}h(Z)| = |\mathbb{E}\mathcal{A}_2 f_h(W)|$, and we then seek a bound for $|\mathbb{E}\mathcal{A}_2 f_h(W)|$.

Two different choices of families of test functions for $\mathcal{H}$ are used in Gan and Ross (2021). Let $\mathrm{C}(E^k)$ denote the set of continuous functions from $E^k \mapsto \mathbb{R}$. For $\mu \in M_1$ and $\phi \in \mathrm{C}(E)$, set $\langle \phi, \mu \rangle = \int_E \phi(x)\mu(dx)$ and $\mathrm{BC}^{2,1}(\mathbb{R}^k)$ to be the set of functions with two bounded and continuous derivatives with second derivative Lipschitz. The first family of test functions is defined as,

$$\mathcal{H}_1 = \{h(\mu) := H(\langle \phi_1, \mu\rangle, \dots, \langle \phi_k, \mu\rangle) : k \in \mathbb{N}, H \in \mathrm{BC}^{2,1}(\mathbb{R}^k), \phi_i \in \mathrm{C}(E), i = 1, \dots, k\}.$$

The second family of test functions is defined as,

$$\mathcal{H}_2 = \{h(\mu) := \langle \phi, \mu^k \rangle : k \in \mathbb{N}, \phi \in \mathrm{C}(E^k)\}.$$

To successfully apply Stein's method, bounds on f_h in (2.2) and its derivatives play a critical role. For each of the two families of test functions $\mathcal{H}_1$ and $\mathcal{H}_2$, two different approaches were used to bound the derivatives of f_h. For $\mathcal{H}_1$, a probabilistic coupling approach that utilises properties of the coalescent process was used and for $\mathcal{H}_2$ a different approach based upon a dual process was used. The differential operator for the derivatives used was defined by

$$\partial_x F(\mu) := \lim_{\varepsilon \to 0^+} \frac{F((1-\varepsilon)\mu + \varepsilon\delta_x) - F(\mu)}{\varepsilon}, \tag{2.3}$$

as opposed to a more standard definition of derivative of

$$\partial_x F(\mu) := \lim_{\varepsilon \to 0^+} \frac{F(\mu + \varepsilon\delta_x) - F(\mu)}{\varepsilon}. \tag{2.4}$$

Derivatives of the form of (2.3) were used as for the family $\mathcal{H}_1$, the probabilistic coupling arguments requires the measures to be evaluated by F to be probability measures. For consistency, the same form of derivative was then also used for the family $\mathcal{H}_2$ as well. It turns out that the proof method for $\mathcal{H}_2$ does not require the measures to be probability measures, and as such, the more commonly form for derivatives as in (2.4) (see Ethier and Kurtz (1993) and Ethier and Griffiths (1993) for example) can be used and this is advantageous for our application to approximate of the stationary distribution of the inclusion process.

Let W follow the stationary distribution of the discrete inclusion process associated with $\mathcal{A}_1 f$ in (1.2). Then noting that this implies $\mathbb{E}\mathcal{A}_1 f(W) = 0$, for any function $h \in \mathcal{H}_2$, our approach will be to derive a bound for

$$|\mathbb{E}h(W) - \mathbb{E}h(Z)| = |\mathbb{E}\mathcal{A}_2 f_h(W)| = |\mathbb{E}\mathcal{A}_2 f_h(W) - \mathbb{E}\mathcal{A}_1 f_h(W)|, \tag{2.5}$$

by exploiting properties of f_h and its derivatives. This method is known as the generator comparison method in Stein's method literature. To evaluate the right hand side of the above, the natural approach is to take the Taylor expansion of $f_h(\mu - \frac{1}{N}\delta_{i/L} + \frac{1}{N}\delta_{j/L}) - f_h(\mu)$ in $\mathcal{A}_1 f$ to match the derivatives in $\mathcal{A}_2 f$. At this point, derivatives of the form (2.4) are more useful as they naturally match the terms in the Taylor expansion, whereas derivatives of the form (2.3) do not. Hence we rederive the results of Gan and Ross (2021, Theorem 2.7) but for derivatives of the form (2.4). The proof is included in the appendix.

Theorem 2.1. *For $h \in \mathcal{H}_2$ such that $h(\mu) = \langle \phi, \mu^k \rangle$, the solution f_h given in (2.2) is well defined and*

$$\begin{aligned}
\sup_x \|\partial_x f_h\|_\infty &\le \frac{2k}{\theta}\|\phi\|_\infty,\\
\sup_{x,y} \|\partial_{xy} f_h\|_\infty &\le \frac{k(k-1)}{\theta+1}\|\phi\|_\infty,\\
\sup_{x,y,z} \|\partial_{xyz} f_h\|_\infty &\le \frac{2k(k-1)(k-2)}{3(\theta+2)}\|\phi\|_\infty.
\end{aligned}$$

3. Proof of the inclusion process approximation results

Proof of Theorem 1.2. Recall the generator we use for the Fleming–Viot process (2.1) with $E = [0,1]$ and π the uniform measure on E is

$$\mathcal{A}_2 f(\mu) = \theta \int_0^1 \partial_x f(\mu)(\pi - \mu)(dx) + \int_0^1 \int_0^1 \big(\mu(dx)\delta_x(dy) - \mu(dx)\mu(dy)\big)\partial_{xy} f(\mu), \tag{3.1}$$

and that the generator for the (discrete) inclusion process is

$$\mathcal{A}_1 f(\mu) = \sum_{i,j=1}^{L} N\mu_i \left(N\mu_j + \frac{\theta}{L}\right)\left[f\left(\mu - \frac{1}{N}\delta_{i/L} + \frac{1}{N}\delta_{j/L}\right) - f(\mu)\right].$$

Furthermore, recall from (2.5) that our goal is to bound

$$|\mathbb{E}h(W) - \mathbb{E}h(Z)| = |\mathbb{E}\mathcal{A}_2 f_h(W)| = |\mathbb{E}\mathcal{A}_2 f_h(W) - \mathbb{E}\mathcal{A}_1 f_h(W)|, \tag{3.2}$$

where W is a random measure following the stationary distribution associated with $\mathcal{A}_1$. For convenience we write $W = \sum_{i=1}^{L} W_i \delta_{x_i}$, where $x_i = i/L$. Our goal is to use the Taylor expansion $\mathcal{A}_1 f(W)$, match terms to $\mathcal{A}_2 f(W)$ and then carefully bound the remainder. We note the Taylor expansion,

$$f\left(W - \frac{1}{N}\delta_{x_i} + \frac{1}{N}\delta_{x_j}\right) - f(W) = -\frac{1}{N}\partial_{x_i} f(W) + \frac{1}{N}\partial_{x_j} f(W) + \frac{1}{2N^2}\partial^2_{x_i x_i} f(W) + \frac{1}{2N^2}\partial^2_{x_j x_j} f(W) - \frac{1}{N^2}\partial^2_{x_i x_j} f(W) + R(W).$$

Then given we can sum either over all i, j or $i \neq j$ when mathematically convenient as for $i = j$, $f\left(W - \frac{1}{N}\delta_{x_i} + \frac{1}{N}\delta_{x_j}\right) - f(W) = 0$,

$$\begin{aligned}
\mathcal{A}_1 f(W) &= \sum_{i\neq j}^{L} N^2 W_i W_j \Big[-\frac{1}{N}\partial_{x_i} f(W) + \frac{1}{N}\partial_{x_j} f(W) \\
&\quad + \frac{1}{2N^2}\partial^2_{x_i x_i} f(W) + \frac{1}{2N^2}\partial^2_{x_j x_j} f(W) - \frac{1}{N^2}\partial^2_{x_i x_j} f(W) + \varepsilon_3(W)\Big] \\
&\quad + \frac{\theta}{L}\sum_{i,j=1}^{L} N W_i \Big[-\frac{1}{N}\partial_{x_i} f(W) + \frac{1}{N}\partial_{x_j} f(W) + \varepsilon_2(W)\Big] \\
&= \sum_{i=1}^{L} W_i(1 - W_i)\partial^2_{x_i x_i} f(W) - \sum_{i\neq j}^{L} W_i W_j \partial^2_{x_i x_j} f(W) + N^2 \sum_{i\neq j}^{L} W_i W_j \varepsilon_3(W) \\
&\quad - \theta \sum_{i=1}^{L} W_i \partial_{x_i} f(W) + \frac{\theta}{L}\sum_{j=1}^{L} \partial_{x_j} f(W) + \frac{N\theta}{L}\sum_{i,j=1}^{L} W_i \varepsilon_2(W),
\end{aligned} \tag{3.3}$$

where $\varepsilon_2(W)$ and $\varepsilon_3(W)$ are higher order expansion terms that we currently suppress for readability but will explicitly bound later. Again using the representation $W = \sum_{i=1}^{L} W_i \delta x_i$, we can write

$$\begin{aligned}
\mathcal{A}_2 f(W) &= \theta \int_0^1 \partial_x f(W) dx - \theta \sum_{i=1}^{L} W_i \partial_{x_i} f(W) \\
&\quad + \sum_{i=1}^{L} W_i(1 - W_i)\partial^2_{x_i x_i} f(W) - \sum_{i\neq j}^{L} W_i W_j \partial^2_{x_i x_j} f(W).
\end{aligned} \tag{3.4}$$

Combining (3.2), (3.3) and (3.4) yields

$$|\mathbb{E}h(W) - \mathbb{E}h(Z)| \leq N\theta|\varepsilon_2(W)| + N^2|\varepsilon_3(W)| + \theta\left|\int_0^1 \partial_x f_h(W)dx - \sum_{j=1}^{L}\frac{1}{L}\partial_{x_j} f_h(W)\right|.$$

For the final term of the above, note that this is simply the Riemann sum approximation for the integral, and hence via the mean value theorem and Theorem 2.1,

$$\left|\int_0^1 \partial_x f_h(W)dx - \sum_{j=1}^{L}\frac{1}{L}\partial_{x_j} f_h(W)\right| \leq \frac{\sup_{x,y}\|\partial_{xy} f_h\|_\infty}{2L} \leq \frac{k(k-1)}{2L(\theta+1)}\|\phi\|_\infty.$$

The final result follows from bounding the higher order error terms as below.

$$\varepsilon_2(W) = \frac{1}{2N^2}\partial^2_{x_i x_i} f_h(W^*) + \frac{1}{2N^2}\partial^2_{x_j x_j} f_h(W^*) - \frac{1}{N^2}\partial^2_{x_i x_j} f_h(W^*),$$

where $W^* = (W + s_1(-\frac{1}{N}\delta_{x_i} + \frac{1}{N}\delta_{x_j}))$ for some $s_1 \in [0, 1]$. Hence,

$$|\varepsilon_2(W)| \leq \frac{2}{N^2}\sup_{x,y}\|\partial_{xy} f_h\|_\infty \leq \frac{2k(k-1)}{N^2(\theta+1)}\|\phi\|_\infty.$$

Similarly,

$$\varepsilon_3(W) = -\frac{1}{6N^3}\partial^3_{x_i x_i x_i} f_h(W^*) + \frac{1}{2N^3}\partial^3_{x_i x_i x_j} f_h(W^*) - \frac{1}{2N^3}\partial^3_{x_i x_i x_j} f_h(W^*) + \frac{1}{6N^3}\partial^3_{x_j x_j x_j} f_h(W^*),$$

where $W^* = (W + s_2(-\frac{1}{N}\delta_{x_i} + \frac{1}{N}\delta_{x_j}))$ for some $s_2 \in [0, 1]$. Hence

$$|\varepsilon_3(W)| \leq \frac{4}{3N^3}\sup_{x,y,z}\|\partial_{xyz} f_h\|_\infty \leq \frac{8k(k-1)(k-2)}{9N^3(\theta+2)}\|\phi\|_\infty. \quad \square$$

Appendix A. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.spl.2026.110747.

Data availability

No data was used for the research described in the article.

References

Barbour, A.D., 1988. Stein's method and Poisson process convergence. J. Appl. Probab. (Special Vol. 25A), 175–184, A celebration of applied probability.

Barbour, A., 1990. Stein's method for diffusion approximations. Probab. Theory Related Fields 84 (3), 297–322.

Barbour, A.D., Chen, L.H.Y. (Eds.), 2005. An Introduction to Stein's Method. In: Lecture Notes Series. Institute for Mathematical Sciences. National University of Singapore, vol. 4, Singapore University Press, Singapore; World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, Lectures from the Meeting on Stein's Method and Applications: a Program in Honor of Charles Stein held at the National University of Singapore, Singapore, July 28–August 31, 2003.

Bianchi, A., Dommers, S., Giardinà, C., 2017. Metastability in the reversible inclusion process. Electron. J. Probab. 22, Paper No. 70, 34.

Chatterjee, S., 2014. A short survey of stein's method. In: Proceedings of the International Congress of Mathematicians—Seoul 2014. Vol. IV. Kyung Moon Sa, Seoul, pp. 1–24.

Chleboun, P., Gabriel, S., Grosskinsky, S., 2022. Poisson-Dirichlet asymptotics in condensing particle systems. Electron. J. Probab. 27, Paper No. 155, 35.

Ethier, S.N., Griffiths, R.C., 1993. The transition function of a Fleming-Viot process. Ann. Probab. 21 (3), 1571–1590.

Ethier, S.N., Kurtz, T.G., 1993. Fleming-Viot processes in population genetics. SIAM J. Control Optim. 31 (2), 345–386.

Feng, S., 2010. The Poisson-Dirichlet Distribution and Related Topics. In: Probability and its Applications (New York), Springer, Models and asymptotic behaviors.

Fleming, W.H., Viot, M., 1979. Some measure-valued Markov processes in population genetics theory. Indiana Univ. Math. J. 28 (5), 817–843.

Gan, H.L., Ross, N., 2021. Stein's method for the Poisson-Dirichlet distribution and the Ewens sampling formula, with applications to Wright-Fisher models. Ann. Appl. Probab. 31 (2), 625–667.

Giardinà, C., Kurchan, J., Redig, F., 2007. Duality and exact correlations for a model of heat conduction. J. Math. Phys. 48 (3), 033301, 15.

Giardinà, C., Kurchan, J., Redig, F., Vafayi, K., 2009. Duality and hidden symmetries in interacting particle systems. J. Stat. Phys. 135 (1), 25–55.

Götze, F., 1991. On the rate of convergence in the multivariate CLT. Ann. Probab. 19 (2), 724–739.

Jatuviriyapornchai, W., Chleboun, P., Grosskinsky, S., 2020. Structure of the condensed phase in the inclusion process. J. Stat. Phys. 178 (3), 682–710.

Kim, S., 2021. Second time scale of the metastability of reversible inclusion processes. Probab. Theory Related Fields 180 (3–4), 1135–1187.

Ley, C., Reinert, G., Swan, Y., 2017. Stein's method for comparison of univariate distributions. Probab. Surv. 14, 1–52.

Ross, N., 2011. Fundamentals of Stein's method. Probab. Surv. 8, 210–293.

Stein, C., 1972. A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In: Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability Theory. Univ. California Press, Berkeley, Calif., pp. 583–602.