Electronic Journal of Statistics
Vol. 5 (2011) 1354–1373
ISSN: 1935-7524
DOI: 10.1214/11-EJS641

Projective limit random probabilities on Polish spaces

Peter Orbanz*

Computational and Biological Learning Laboratory
University of Cambridge
e-mail: p.orbanz@eng.cam.ac.uk

Abstract: A pivotal problem in Bayesian nonparametrics is the construction of prior distributions on the space $\mathbf{M}(V)$ of probability measures on a given domain V. In principle, such distributions on the infinite-dimensional space $\mathbf{M}(V)$ can be constructed from their finite-dimensional marginals—the most prominent example being the construction of the Dirichlet process from finite-dimensional Dirichlet distributions. This approach is both intuitive and applicable to the construction of arbitrary distributions on $\mathbf{M}(V)$, but also hamstrung by a number of technical difficulties. We show how these difficulties can be resolved if the domain V is a Polish topological space, and give a representation theorem directly applicable to the construction of any probability distribution on $\mathbf{M}(V)$ whose first moment measure is well-defined. The proof draws on a projective limit theorem of Bochner, and on properties of set functions on Polish spaces to establish countable additivity of the resulting random probabilities.

AMS 2000 subject classifications: Primary 62C10; secondary 60G57.
Keywords and phrases: Bayesian nonparametrics, Dirichlet processes, random probability measures.

Received January 2011.

1. Introduction

A variety of ways exists to construct the Dirichlet process. For this particular case of a random probability measure, the spectrum of construction approaches ranges from the projective limit construction from finite-dimensional Dirichlet distributions proposed by Ferguson [8] to the stick-breaking construction of Sethuraman [25]; see e.g. the survey by Walker *et al.* [27] for an overview. Most of these constructions are bespoke representations more or less specific to the Dirichlet. An exception is the projective limit representation, which can represent any probability distribution on the space of probability measures. However, several authors [e.g. 12, 13] have noted technical problems arising for this construction. The key role of the Dirichlet process, and the proven utility of its representation by stick-breaking or by Poisson processes, may account for the slightly surprising fact that these problems have not yet been addressed in the literature.

*Research supported by EPSRC grant EP/F028628/1.

The purpose of this paper is to provide a projective limit result directly applicable to the construction of any probability distribution on $\mathbf{M}(V)$. We do so by first modifying and then proving a construction idea put forth by Ferguson [8]. Intuitively speaking, our main result (Theorem 1.1) allows us to construct distributions on $\mathbf{M}(V)$ by substituting the Dirichlet distributions used in the derivation of the Dirichlet process by other families of distributions, and by verifying that these families satisfy the two necessary and sufficient conditions of the theorem. Stick-breaking, urn schemes [3] and other specialized representations of the Dirichlet process all rely on the latter's particular discreteness and spatial decorrelation properties. Our approach may facilitate the derivation of models for which no such representations can be expected to exist, for example, of smooth random measures. For Bayesian nonparametrics, the result provides what currently seems to be the only available tool to construct an arbitrary prior distribution on the set $\mathbf{M}(V)$. It also makes Bayesian methods based on random measures more readily comparable to other types of nonparametric priors constructed in a similar fashion, notably to Gaussian processes [2, 28].

The technical difficulties arising for the construction proposed in [8] can be summarized as three separate problems, which Appendix A reviews in detail. In short:

i **Product spaces.** The product space setting of the standard Kolmogorov extension theorem is not well-adapted to the problem of constructing random probability measures.
ii **Measurability problems.** A straightforward formalization of the construction in terms of an extension or projective limit theorem results in a space whose dimensions are labeled by the Borel sets of V, and is hence of uncountable dimension. As a consequence, the constructed measure cannot resolve most events of interest. In particular, singletons, and hence the event that the random measure assumes a specific measure as its value, are not measurable [13, Sec. 2.3.2].
iii **σ-additivity.** The constructed measure is supported on finitely additive probabilities (charges), rather than σ-additive probabilities (measures); see Ghosal [12, Sec. 2.2]. Further conditions are necessary to obtain a measure on probability measures.

To make the projective limit construction feasible, we have to impose some topological requirements on the domain V of the random measure. Specifically, we require that V is a Polish space, i.e. a topological space which is complete, separable and metrizable [17]. This setting is sufficiently general to accommodate any applications in Bayesian nonparametrics—Bayesian methods do not solicit the generality of arbitrary measurable spaces, since no useful notion of conditional probability can be defined without a modicum of topological structure. Polish spaces are in many regards the natural habitat of Bayesian statistics, whether parametric or nonparametric, since they guarantee both the existence of regular conditional probabilities and the validity of de Finetti's theorem [16, Theorem 11.10]. The restriction to Polish spaces is hence unlikely to incur any loss of generality. We address problem (i) by means of a generalization of Kolmogorov's extension theorem, due to Bochner [4]; problem (ii) by means of the fact that the Borel σ-algebra of a Polish space V is generated by a countable subsystem of sets, which allows us to substitute the uncountable-dimensional projective limit space by a countable-dimensional surrogate; and problem (iii) using a result of Harris [14] on σ-additivity of set functions on Polish spaces.

The remainder of the article is structured as follows: The main result is stated in Sec. 1.1, which is meant to provide all information required to apply the theorem, without going into the details of the proof. Related work is summarized in Sec. 1.3. A brief overview of projective limit constructions is given in Sec. 2, to the extent relevant to the proof. Secs. 3 and 4 contain the actual proof of Theorem 1.1: The projective limit construction of random set functions is described in Sec. 3. A necessary and sufficient condition for these random set functions to be σ-additive is given in Sec. 4. Appendix A reviews problems (i)-(iii) above in more detail.

1.1. Main result

To state our main theorem, we must introduce some notation, and specify the relevant notion of a marginal distribution in the present context. Let $\mathbf{M}(V)$ be the set of Borel probability measures over a Polish topological space $(V, \mathcal{T}_{\mathrm{V}})$; recall that the space is Polish if $\mathcal{T}_{\mathrm{V}}$ is a metrizable topology under which V is complete and separable [1, 17]. Throughout, the underlying model of randomness is an abstract probability space $(\Omega, \mathcal{A}, \mathbb{P})$. A random variable $X\colon \Omega \to \mathbf{M}(V)$, with the image measure $P := X\mathbb{P}$ as its distribution, is called a *random probability measure* on V. Our main result, Theorem 1.1, is a general representation result for the distribution P of such a random measure. To define measures on the space $\mathbf{M}(V)$, we endow it with the weak* topology $\mathcal{T}_{w^*}$ (which in the context of probability is often called the topology of weak convergence) and with the corresponding Borel σ-algebra $\mathcal{B}_{w^*} := \sigma(\mathcal{T}_{w^*})$. Since V is Polish, the topological space $(\mathbf{M}(V), \mathcal{T}_{w^*})$ is Polish as well [17, Theorem 17.23].

Let $I = (A_1, \ldots, A_n)$ be a *measurable partition* of V, i.e. a partition of V into a finite number of measurable, disjoint sets. Denote the set of all such partitions $\mathcal{H}(\mathcal{B}_{\mathrm{V}})$. Any probability measure $x \in \mathbf{M}(V)$ can be evaluated on a partition I to produce a vector $x_{\mathrm{I}} := (x(A_1), \ldots, x(A_n))$, and we write $\phi_{\mathrm{I}} : x \mapsto x_{\mathrm{I}}$ for the evaluation functional so defined. Clearly, x_{I} represents a probability measure on the finite σ-algebra $\sigma(I)$ generated by the partition. Let $\triangle_{\mathrm{I}}$ be the set of all measures $x_{\mathrm{I}} = \phi_{\mathrm{I}}(x)$ obtained in this manner, where x runs through all measures in $\mathbf{M}(V)$. This set, $\triangle_{\mathrm{I}} = \phi_{\mathrm{I}}\mathbf{M}(V)$, is precisely the unit simplex in the n-dimensional Euclidean space $\mathbb{R}^{\mathrm{I}}$,

$$\triangle_{\mathrm{I}} = \left\{ x_{\mathrm{I}} \in \mathbb{R}^{\mathrm{I}} \middle| x_{\mathrm{I}}(A_i) \geq 0 \text{ and } \sum_{A_i \in I} x_{\mathrm{I}}(A_i) = 1 \right\} . \tag{1.1}$$

Let $J = (B_1, \ldots, B_m)$ and $I = (A_1, \ldots, A_n)$ be partitions such that I is a coarsening of J, that is, for each $A_i \in I$, there is a set $\mathcal{J}_i \subset \{1, \ldots, m\}$ of

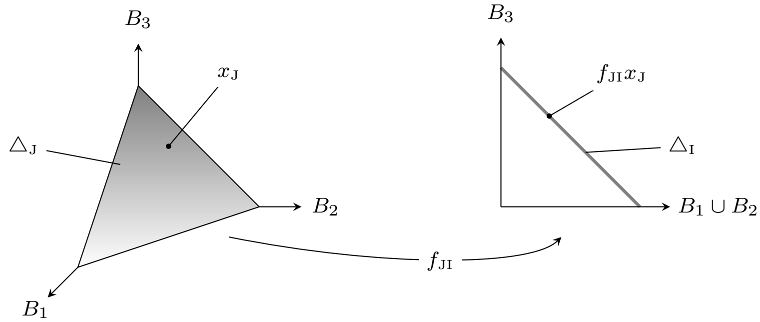


FIG 1. *Left: The simplex $\triangle_J \subset \mathbb{R}^J$ for a partition $J = (B_1, B_2, B_3)$. Right: A new simplex $\triangle_I = f_{JI}\triangle_J$ is obtained by merging the sets B_1 and B_2, producing the partition $I = (B_1 \cup B_2, B_3)$. The mapping f_{JI} is given by $f_{JI}x_J = (x_J(B_1) + x_J(B_2), x_J(B_3))$. Its image $\triangle_I$ is a subset of the product space $\mathbb{R}^I$, which shares only a single axis, B_3, with the space $\mathbb{R}^J$.*

indices such that $A_i = \cup_{j \in \mathcal{J}_i} B_j$. The sets $\mathcal{J}_i$ form a partition of the index set $\{1, \ldots, m\}$. If I is a coarsening of J, we write $I \preceq J$.

Let $x, x' \in \mathbf{M}(V)$. If $I \preceq J$, then $\phi_J x = \phi_J x'$ implies $\phi_I x = \phi_I x'$. In other words, $\phi_I x$ is completely determined by $\phi_J x$, and invariant under any changes to x which do not affect $\phi_J x$. Therefore, the implicit definition $f_{JI}(\phi_J(x)) := \phi_I(x)$ determines a well-defined mapping $f_{JI} : \triangle_J \to \triangle_I$. With notation for J and I as above, f_{JI} can equivalently be defined as

$$(f_{JI}x_J)(A_i) = \sum_{j \in \mathcal{J}_i} x_J(B_j) \; . \tag{1.2}$$

Figure 1 illustrates the mapping f_{JI} and the simplices $\triangle_J$ and $\triangle_I$. The image $f_{JI}x_J \in \triangle_I$ constitutes a probability distribution on the events in I. The following intuition is often helpful: The space $\mathbf{M}(V)$ is convex, with the Dirac measures on V as its extreme points, and we can roughly think of $\mathbf{M}(V)$ as the infinite-dimensional analogue of the simplices $\triangle_I$. Similarly, we can regard the evaluations maps $\phi_I : \mathbf{M}(V) \to \triangle_I$ as analogues of the maps $f_{JI} : \triangle_J \to \triangle_I$. Even though both $\mathbf{M}(V)$ and the spaces $\triangle_I$ are Polish, however, we have to keep in mind that the weak* topology on $\mathbf{M}(V)$ is, in many regards, quite different from the topology which $\triangle_I$ inherits from Euclidean space. For further properties of the space $\mathbf{M}(V)$, we refer to the excellent exposition given by Aliprantis and Border [1, Chapter 15].

Suppose that P is a probability measure on $\mathbf{M}(V)$. Denote by $\phi_I P$ the image measure of P under ϕ_I, i.e. the measure on $\triangle_I$ defined by $(\phi_I P)(A_I) := P(\phi_I^{-1} A_I)$ for all $A_I \in \mathcal{B}(\triangle_I)$. We refer to $\phi_I P$ as the *marginal* of P on $\triangle_I$. Similarly, if P_J is a measure on $\triangle_J$, then for any $I \preceq J$, the image measure $f_{JI}P_J$ is called the marginal of P_J on $\triangle_I$. The following theorem, our main result, states that a measure P on $\mathbf{M}(V)$ can be constructed from a suitable family of marginals P_I on the simplices $\triangle_I$. The notation $\mathbb{E}_Q[\,.\,]$ refers to expectation with respect to the law Q.

Theorem 1.1. *Let V be a Polish space with Borel sets $\mathcal{B}_{\mathrm{V}}$. Let $\mathbf{M}(V)$ be the set of probability measures on $(V, \mathcal{B}_{\mathrm{V}})$, and $\mathcal{B}_{w^*}$ the Borel σ-algebra generated by the weak* topology on $\mathbf{M}(V)$. Let $\langle P_{\mathrm{I}} \rangle_{\mathcal{H}(\mathcal{B}_{\mathrm{V}})} := \{P_{\mathrm{I}} | I \in \mathcal{H}(\mathcal{B}_{\mathrm{V}})\}$ be a family of probability measures on the finite-dimensional simplices $\triangle_{\mathrm{I}}$. The following statements are equivalent:*
(1) The family $\langle P_{\mathrm{I}} \rangle_{\mathcal{H}(\mathcal{B}_{\mathrm{V}})}$ is projective,

$$P_{\mathrm{I}} = f_{\mathrm{JI}} P_{\mathrm{J}} \qquad \text{whenever } I \preceq J \tag{1.3}$$

and satisfies

$$\mathbb{E}_{P_{\mathrm{I}}}[X_{\mathrm{I}}] = \phi_{\mathrm{I}} G_0 \qquad \text{for all } I \in \mathcal{H}(\mathcal{B}_{\mathrm{V}}) \,. \tag{1.4}$$

(2) There exists a unique probability measure P on $(\mathbf{M}(V), \mathcal{B}_{w^*})$ satisfying

$$P_{\mathrm{I}} = \phi_{\mathrm{I}} P \qquad \text{for all } I \in \mathcal{H}(\mathcal{B}_{\mathrm{V}}) \tag{1.5}$$

and

$$\mathbb{E}_P[X] = G_0 \qquad \text{for some } G_0 \in \mathbf{M}(V) \,. \tag{1.6}$$

If either statement holds, P is a Radon measure.

Remark 1.2. Theorem 1.1 is applicable to the construction of any random probability measure X on V whose first moment $\mathbb{E}_P[X]$ exists. In particular, the random measure X need not be discrete. See Sec. 1.2 for examples.

The two conditions of Theorem 1.1 serve two separate purposes: Condition (1.3) guarantees that the family $\langle P_{\mathrm{I}} \rangle_{\mathcal{H}(\mathcal{B}_{\mathrm{V}})}$ defines a unique probability measure $P_{\mathcal{H}(\mathcal{B}_{\mathrm{V}})}$. The support of this measure is not actually $\mathbf{M}(V)$, but a larger set—specifically, the set $\mathbf{C}(\mathcal{Q})$ of *finitely* additive probability measures (charges) defined on a certain subsystem $Q \subset \mathcal{B}_{\mathrm{V}}$, which we will make precise in Sec. 3. The set $\mathbf{C}(\mathcal{Q})$ contains the set $\mathbf{M}(\mathcal{Q})$ of σ-additive probability measures on Q as a measurable subset, and $\mathbf{M}(\mathcal{Q})$ is in turn isomorphic to $\mathbf{M}(V)$, by Carathéodory's extension theorem [16, Theorem 2.5]. To obtain the distribution of a random measure, we need to ensure that $P_{\mathcal{H}(\mathcal{B}_{\mathrm{V}})}$ concentrates on the subset $\mathbf{M}(\mathcal{Q}) \cong \mathbf{M}(V)$, or in other words, that draws from $P_{\mathcal{H}(\mathcal{B}_{\mathrm{V}})}$ are σ-additive almost surely. Condition (1.4) is sufficient—and in fact necessary—for $P_{\mathcal{H}(\mathcal{B}_{\mathrm{V}})}$ to concentrate on $\mathbf{M}(V)$, and therefore for a random variable $X_{\mathcal{H}(\mathcal{B}_{\mathrm{V}})}$ with distribution $P_{\mathcal{H}(\mathcal{B}_{\mathrm{V}})}$ to constitute a random measure. If (1.4) is satisfied, the measure constructed on $\mathbf{C}(\mathcal{Q})$ can be restricted to a measure on $\mathbf{M}(V)$, resulting in the measure P described by Theorem 1.1. Sec. 3 provides more details.

The technical restriction that V be Polish is a mild one for all practical purposes, a fact best illustrated by some concrete examples of Polish spaces: The real line is Polish, and so are $\mathbb{R}^n$ and $\mathbb{C}^n$; any finite space; all separable Banach spaces (since Banach spaces are complete metric spaces), in particular $\mathcal{L}_2$ and any other separable Hilbert space; the space $\mathbf{M}(V)$ of probability measures over a Polish domain V, in the weak* topology [1, Chapter 15]; the spaces $\mathcal{C}([0,1], \mathbb{R})$ and $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$ of continuous functions, in the topology of compact convergence [2, §38]; and the Skorohod space $\mathcal{D}(\mathbb{R}_+, \mathbb{R})$ of càdlàg functions [24, Chapter VI]. Any countable product of Polish spaces is Polish, in particular $\mathbb{R}^{\mathbb{N}}$, $\mathbb{C}^{\mathbb{N}}$, and the Hilbert cube $[0,1]^{\mathbb{N}}$. A subset of a given Polish space is Polish in the relative topology if and only if it is a G_δ set [17, Theorem 3.11]. A borderline example are the spaces $\mathcal{C}(T,E)$ of continuous functions with Polish range E. This space is Polish if $T=\mathbb{R}_{\geq 0}$ or if T is compact and Polish, but not e.g. for $T=\mathbb{R}$ [10, §454]. In Bayesian nonparametrics, this distinction may be relevant in the context of the "dependent Dirichlet process" model of MacEachern [21], which involves Dirichlet processes on spaces of continuous functions. For more background on Polish spaces, see [1, 10, 17].

1.2. Examples

Theorem 1.1 yields straightforward constructions for several models studied in the literature, and we consider three specific examples to illustrate the result. First, by choosing the finite-dimensional marginals P_{I} in Theorem 1.1 as a suitable family of Dirichlet distributions, we obtain a construction of the Dirichlet process in the spirit of Ferguson [8].

Corollary 1.3 (Dirichlet Process)**.** *Let V be a Polish space, G_0 a probability measure on $\mathcal{B}_{\mathrm{V}}$, and let $\alpha \in \mathbb{R}_{>0}$. For each $I \in \mathcal{H}(\mathcal{B}_{\mathrm{V}})$, define P_{I} as the Dirichlet distribution on $\triangle_{\mathrm{I}} \subset \mathbb{R}^{\mathrm{I}}$, with concentration α and expectation $\phi_{\mathrm{I}} G_0 \in \triangle_{\mathrm{I}}$. Then there is a uniquely determined probability measure P on $\mathbf{M}(V)$ with expectation G_0 and the distributions P_{I} as its marginals, that is, $\phi_{\mathrm{I}} P = P_{\mathrm{I}}$ for all $I \in \mathcal{H}(\mathcal{B}_{\mathrm{V}})$.*

A similar construction yields the *normalized inverse Gaussian process* of Lijoi *et al.* [20]. The inverse Gaussian distribution on $\mathbb{R}_{\geq 0}$ is given by the density $p_{\mathrm{IG}}(z|\alpha,\gamma) = \frac{\alpha}{\sqrt{2\pi}} x^{-3/2} \exp(-\frac{1}{2}(\frac{\alpha^2}{x} + \gamma^2 x) + \gamma\alpha)$ with respect to Lebesgue measure. Lijoi *et al.* [20] define a normalized inverse Gaussian distribution $\mathrm{NIG}(\alpha_1,\ldots,\alpha_n)$ on the simplex $\triangle_n \subset \mathbb{R}^n$ as the distribution of the vector $w = (\frac{z_1}{\sum_i z_i},\ldots,\frac{z_n}{\sum_i z_i})$, where z_i is distributed according to $p_{\mathrm{IG}}(z_i|\alpha_i,\gamma=1)$. The density of w can be derived explicitly [20, Equation (4)]. Applicability of Theorem 1.1 is a direct consequence of the results of Lijoi *et al.* [20], which imply conditions (1.3) [20, (C3)] and (1.4) [20, Proposition 2].

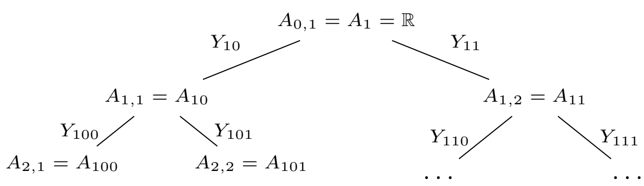
Corollary 1.4 (Normalized Inverse Gaussian Process)**.** *Let $\alpha \in \mathbb{R}_+$ and $G_0 \in \mathbf{M}(V)$. For any partition $I=(A_1,\ldots,A_n)$ in $\mathcal{H}(\mathcal{B}_{\mathrm{V}})$, choose the measure P_{I} as the normalized inverse Gaussian distribution $\mathrm{NIG}(\alpha G_0(A_1),\ldots,\alpha G_0(A_n))$. There is a uniquely determined probability measure P on $\mathbf{M}(V)$ with expectation G_0 and $\phi_{\mathrm{I}} P = P_{\mathrm{I}}$ for all $I \in \mathcal{H}(\mathcal{B}_{\mathrm{V}})$.*

Although both the Dirichlet process and the normalized inverse Gaussian process are discrete almost surely, Theorem 1.1 is applicable to the construction of continuous random measures. The *Pólya tree* random measures introduced by Ferguson [9] provide a convenient example. They can be obtained as projective limits as follows: Choose $V=\mathbb{R}$ and let $G_0 \in \mathbf{M}(\mathbb{R})$ be a probability measure with cumulative distribution function g_0. For each n, let I_n be the partition of $\mathbb{R}$ into intervals $[g_0^{-1}(\frac{k-1}{2^n}), g_0^{-1}(\frac{k}{2^n}))$, where $k=1,\ldots,2^n$. All sets in I_n have identical probability $1/2^n$ under G_0. Since each partition I_n is obtained from

I_{n-1} by splitting each set in I_{n-1} at a single point, the sequence (I_n) satisfies $I_1 \preceq I_2 \preceq \ldots$. It can be represented as a binary tree whose nth level corresponds to I_n, each node representing one constituent set. There are two natural ways of indexing sets in the partitions: One is to write $A_{n,k}$ for the kth set in I_n, i.e. n indexes tree levels and k enumerates sets within each level. The other is to index sets as $A_{m_1,\ldots,m_n}$ by a binary sequence encoding the unique path from the root node $\mathbb{R}$ and the set in question, where $m_i = 1$ indicates passing to a right child node. Let $[m]_2$ denote the binary representation of an arbitrary positive integer m. Then

$$A_{n,k} = \Big[g_0^{-1}\Big(\frac{k-1}{2^n}\Big), g_0^{-1}\Big(\frac{k}{2^n}\Big)\Big) = A_{[2^n+(k-1)]_2} \quad \text{and} \quad I_n = (A_{n,1}, \ldots, A_{n,2^n}) \,.$$

It is useful to use both index conventions interchangeably. With each node $A_{m_1\cdots m_n}$, we associate a pair $(Y_{m_1\cdots m_n 0}, Y_{m_1\cdots m_n 1}) \sim \text{Beta}(\alpha_{m_1\cdots m_n 0}, \alpha_{m_1\cdots m_n 1})$ of beta random variables:



To apply Theorem 1.1, define probability measures $P_{\mathrm{I_n}}$ on the simplices $\triangle_{\mathrm{I_n}}$ as follows: Suppose a particle slides down the tree, moving along each edge with the associated probability $Y_{m_1\cdots m_n}$. The probability of reaching the set $A_{n,k}$ is a random variable $X_{n,k}$, defined recursively in terms of the beta variables as $X_{m_1\cdots m_n m_{n+1}} := X_{m_1\cdots m_n} Y_{m_1\cdots m_n m_{n+1}}$. Choose $P_{\mathrm{I_n}}$ as the distribution of $X_{\mathrm{I_n}} = (X_{n,1}, \ldots, X_{n,2^n})$. Applicability of Theorem 1.1 follows from two results of Ferguson [9]: (a) The partitions I_n generate the Borel sets $\mathcal{B}(\mathbb{R})$ and (b) each random measure $X_{\mathrm{I_n}} \in \triangle_{\mathrm{I_n}}$ has expectation $\mathbb{E}[X_{\mathrm{I_n}}] = (G_0(A_{n,1}), \ldots, G_0(A_{n,2^n}))$. Property (a) implies that the sequence $P_{\mathrm{I_n}}$ induces a complete family $\langle P_{\mathrm{I}} \rangle$ of probability measures on all simplices $\triangle_{\mathrm{I}}$, $I \in \mathcal{H}(\mathcal{B}(\mathbb{R}))$. By construction, $\langle P_{\mathrm{I}} \rangle$ satisfies (1.3). According to (b), (1.4) holds. Theorem 1.1 and the well-known continuity properties of Pólya trees [19, Theorem 3] yield:

Corollary 1.5 (Pólya tree)**.** *Let $\langle P_{\mathrm{I}} \rangle$ be a family of measures defined as above. There is a unique probability measure P on $\mathbf{M}(\mathbb{R})$ satisfying $\phi_{\mathrm{I}} P = P_{\mathrm{I}}$. The distribution P is a Pólya tree in the sense of Ferguson [9], with parameters G_0 and $(\alpha_{[n]_2})_{n \in \mathbb{N}}$. The random probability measure X on $\mathbb{R}$ with distribution P has expected measure $\mathbb{E}_P[X] = G_0$. If $\alpha_{n,k} = cn^2$ for some $c > 0$, then X is absolutely continuous with respect to Lebesgue measure on $\mathbb{R}$ almost surely.*

1.3. Related work

Theorem 1.1 was effectively conjectured by Ferguson [8]. Although he only considered the special case of the Dirichlet process, and despite the technical difficulties already mentioned, he recognized both the usefulness of indexing spaces by measurable partitions (a key ingredient of the construction in Sec. 3), and the connection between σ-additivity of random draws from the Dirichlet process and σ-additivity of its parameter measure [cf. 8, Proposition 2]. Authors who have recognized problems to the effect that such a construction is not feasible on an arbitrary measurable space V include Ghosh and Ramamoorthi [13] and Ghosal [12]; both references also provide excellent surveys of the different construction approaches available for the Dirichlet process. Ghosal [12] additionally points out, in the context of problem (ii), that a countable generator may be substituted for $\mathcal{B}_{\mathrm{V}}$, provided the underlying space is separable and metrizable.

To resolve the σ-additivity problem (iii), we appeal to a result of Harris [14], which reduces the conditions for σ-additivity of random set functions to their behavior on a countable number of sequences. This result is well-known in the theory of point processes and random measures [7, 15]. Although Sethuraman was aware of Harris' work and referenced it in his well-known article [25], it has to our knowledge never been followed up on in the nonparametric Bayesian literature.

For the specific problem of defining the Dirichlet process, it is possible to forego the projective limit construction altogether and invoke approaches specifically tailored to the properties of the Dirichlet [12, 13, 27]. On the real line, both the Dirichlet process and the closely related Poisson-Dirichlet distribution of Kingman [18] arise in a variety of contexts throughout mathematics, each of which can be regarded as a possible means of definition [e.g. 23, 26]. On arbitrary Polish spaces, the Dirichlet process can be derived implicitly as de Finetti mixing measure of an urn scheme [3], or as special case of a Pólya tree [9].

Sethuraman's stick-breaking scheme [25] is remarkable not only for its simplicity. In contrast to all other constructions listed above, it does not require V to be Polish, but is applicable on an arbitrary measurable space with measurable singletons. The stick-breaking and projective limit representations of the Dirichlet process trade off two different types of generality: Stick-breaking imposes less restrictions on the choice of V, but is not applicable to represent other types of distributions on $\mathbf{M}(V)$. The projective limit approach requires more structure on V, but can represent any probability measure on $\mathbf{M}(V)$. The trade-off is reminiscent of similar phenomena encountered throughout stochastic process theory. For example, probability measures on infinite-dimensional product spaces can be constructed by means of Kolmogorov's extension theorem. If the measure to be constructed is factorial over the product, the component spaces of the product may be chosen as arbitrary measurable spaces [2, Theorem 9.2]. To model stochastic dependence across different subspaces, however, a minimum of topological structure is indispensable, and Kolmogorov's theorem hence requires the component spaces to be Polish [16, Theorem 6.16]. The Dirichlet process, as a purely atomic random measure whose different atoms are stochastically dependent only through the global normalization constraint, can be regarded as the closest analogue of a factorial measure on the space $\mathbf{M}(V)$. In analogy to a factorial measure, it can be constructed on very general spaces, whereas the projective limit approach, which can represent arbitrary correlation structure, requires stronger topological properties.

2. Background: Projective limits

A projective limit is constructed from a family of mathematical structures, indexed by the elements of an index set D [5, 22]. For our purposes, the structures in question will be topological measurable spaces $(\mathcal{X}_{\mathrm{I}}, \mathcal{B}_{\mathrm{I}})$, with $I \in D$. The projective limit defined by this family is again a measurable space, denoted $(\mathcal{X}_{\mathrm{D}}, \mathcal{B}_{\mathrm{D}})$. This projective limit space is the smallest space containing all spaces $(\mathcal{X}_{\mathrm{I}}, \mathcal{B}_{\mathrm{I}})$ as its substructures, in a sense to be made precise shortly. To obtain a meaningful notion of a limit, the index set D need not be totally ordered, but it must be possible to form infinite sequences of suitably chosen elements. The set is therefore required to be *directed*: There is a partial order relation $\preceq$ on D and, whenever $I, J \in D$, there exists $K \in D$ such that $I \preceq K$ and $J \preceq K$. A simple example of a directed set is the set $D := \mathcal{F}(L)$ of all finite subsets of an infinite set L, where D is partially ordered by inclusion.

The component spaces $\mathcal{X}_{\mathrm{I}}$ used to define the projective limit need to "fit in" with each other in a suitable manner. This idea is formalized by defining a family of mappings f_{JI} between the spaces which are regular with respect to the structure posited on the point sets $\mathcal{X}_{\mathrm{I}}$. For measurable spaces, the adequate notion of regularity is measurability. Since we assume each σ-algebra $\mathcal{B}_{\mathrm{I}}$ to be generated by an underlying topology $\mathcal{T}_{\mathrm{I}}$, we slightly strengthen this requirement to continuity.

Definition 2.1 (Projective limit set)**.** Let D be a directed set and $(\mathcal{X}_{\mathrm{I}}, \mathcal{T}_{\mathrm{I}})$, with $I \in D$, a family of topological spaces. For any pair $I \preceq J \in D$, let $f_{\mathrm{JI}} : \mathcal{X}_{\mathrm{J}} \to \mathcal{X}_{\mathrm{I}}$ be a function such that

1. f_{JI} is $\mathcal{T}_{\mathrm{J}}$-$\mathcal{T}_{\mathrm{I}}$-continuous.
2. $f_{\mathrm{II}} = \mathrm{Id}_{\mathcal{X}_{\mathrm{I}}}$.
3. $f_{\mathrm{KJ}} \circ f_{\mathrm{JI}} = f_{\mathrm{KI}}$ whenever $I \preceq J \preceq K$.

The functions f_{JI} are called *generalized projections*. The family $\{\mathcal{X}_{\mathrm{I}}, \mathcal{T}_{\mathrm{I}}, f_{\mathrm{JI}} | I \preceq J \in D\}$, which we denote $\langle \mathcal{X}_{\mathrm{I}}, \mathcal{T}_{\mathrm{I}}, f_{\mathrm{JI}} \rangle_{\mathrm{D}}$, is called a *projective system* of topological spaces. Define a set $\mathcal{X}_{\mathrm{D}}$ as follows: For each collection $\{x_{\mathrm{I}} \in \mathcal{X}_{\mathrm{I}} | I \in D\}$ of points satisfying

$$x_{\mathrm{I}} = f_{\mathrm{JI}} x_{\mathrm{J}} \qquad \text{whenever } I \preceq J \ , \tag{2.1}$$

identify the set $\{x_{\mathrm{I}} \in \mathcal{X}_{\mathrm{I}} | I \in D\}$ with a point x_{D}, and let $\mathcal{X}_{\mathrm{D}}$ be the collection of all such points. The set $\mathcal{X}_{\mathrm{D}}$ is called the *projective limit set* of $\langle \mathcal{X}_{\mathrm{I}}, f_{\mathrm{JI}} \rangle_{\mathrm{D}}$.

Denote the Borel σ-algebras on the topological spaces $\mathcal{X}_{\mathrm{I}}$ by $\mathcal{B}_{\mathrm{I}} := \sigma(\mathcal{T}_{\mathrm{I}})$. For each $I \in D$, the map defined as $f_{\mathrm{I}} \colon x_{\mathrm{D}} \mapsto x_{\mathrm{I}}$ is a well-defined function $f_{\mathrm{I}} \colon \mathcal{X}_{\mathrm{D}} \to \mathcal{X}_{\mathrm{I}}$. These functions are called *canonical mappings*. They define a topology $\mathcal{T}_{\mathrm{D}}$ and a σ-algebra on the projective limit space $\mathcal{X}_{\mathrm{D}}$, as the smallest topology (resp. σ-algebra) which makes all canonical mappings f_{I} continuous (resp. measurable). In particular,

$$\mathcal{B}_{\mathrm{D}} := \sigma(f_{\mathrm{I}} | I \in D) = \sigma(\cup_{I \in D} f_{\mathrm{I}}^{-1} \mathcal{B}_{\mathrm{I}}) = \sigma(\mathcal{T}_{\mathrm{D}}) \ . \tag{2.2}$$

In analogy to the set $\mathcal{X}_{\text{D}}$, the topological space $(\mathcal{X}_{\text{D}}, \mathcal{T}_{\text{D}})$ is called the projective limit of $\langle \mathcal{X}_{\text{I}}, \mathcal{T}_{\text{I}}, f_{\text{JI}} \rangle_{\text{D}}$, and the measurable space $(\mathcal{X}_{\text{D}}, \mathcal{B}_{\text{D}})$ the projective limit of $\langle \mathcal{X}_{\text{I}}, \mathcal{B}_{\text{I}}, f_{\text{JI}} \rangle_{\text{D}}$.

A measure P_{D} on the projective limit $(\mathcal{X}_{\text{D}}, \mathcal{B}_{\text{D}})$ can be constructed by defining a measure P_{I} on each space $(\mathcal{X}_{\text{I}}, \mathcal{B}_{\text{I}})$. By simultanously applying the projective limit to the projective system $\langle \mathcal{X}_{\text{I}}, \mathcal{B}_{\text{I}}, f_{\text{JI}} \rangle_{\text{D}}$ and to the measures P_{I}, the family $\langle P_{\text{I}} \rangle_{\text{D}}$ is assembled into the measure P_{D}. The only requirement is that the measures P_{I} satisfy a condition analogous to the one imposed on points by (2.1). More precisely, P_{I} has to coincide with the image measure of P_{J} under f_{JI},

$$P_{\text{I}} = f_{\text{JI}} P_{\text{J}} = P_{\text{J}} \circ f_{\text{JI}}^{-1} \qquad \text{whenever } I \preceq J\ . \tag{2.3}$$

A family of measures $\langle P_{\text{I}} \rangle_{\text{D}}$ satisfying (2.3) is called a *projective family*. The existence and uniqueness of P_{D} on $(\mathcal{X}_{\text{D}}, \mathcal{B}_{\text{D}})$ is guaranteed by the following result [6, IX.4.3, Theorem 2].

Theorem 2.2 (Bochner)**.** *Let $\langle \mathcal{X}_{\text{I}}, \mathcal{B}_{\text{I}}, f_{\text{JI}} \rangle_{\text{D}}$ be a projective system of measurable spaces with countable, directed index set D, and $\langle P_{\text{I}} \rangle_{\text{D}}$ a projective family of probability measures on these spaces. Then there exists a uniquely defined measure P_{D} on the projective limit space $(\mathcal{X}_{\text{D}}, \mathcal{B}_{\text{D}})$ such that*

$$P_{\text{I}} = f_{\text{I}} P_{\text{D}} = P_{\text{D}} \circ f_{\text{I}}^{-1} \qquad \textit{for all } I \in D\ . \tag{2.4}$$

We refer to the measures in the family $\langle P_{\text{I}} \rangle_{\text{D}}$ as the *marginals* of the stochastic process P_{D}. Since the marginals completely determine P_{D}, some authors refer to $\langle P_{\text{I}} \rangle_{\text{D}}$ as the *weak distribution* of the process, or as a *promeasure* [6].

Theorem 2.2 was introduced by Bochner [4, Theorem 5.1.1], for a possibly uncountable index set D. The uncountable case requires an additional condition known as *sequential maximality*, which ensures the projective limit space is non-empty. For our purposes, however, countability of the index set is essential: Measurability problems (problem (ii) in Sec. 1) arise whenever D is uncountable, and are not resolved by sequential maximality.

The most common example of a projective limit theorem in probability theory is Kolmogorov's extension theorem [16, Theorem 6.16], which can be regarded as the special case of Bochner's theorem obtained for product spaces: Let D be the set of all finite subsets of an infinite set L, partially ordered by inclusion. Choose any Polish measurable space $(\mathcal{X}_0, \mathcal{B}_0)$, and set $\mathcal{X}_{\text{I}} := \prod_{i \in I} \mathcal{X}_0$. The resulting projective limit space is the infinite product $\mathcal{X}_{\text{D}} = \prod_{i \in L} \mathcal{X}_0$, and $\mathcal{B}_{\text{D}}$ coincides with the Borel σ-algebra generated by the product topology. For product spaces, the sequential maximality condition mentioned above holds automatically, so L may be either countable or uncountable. Once again, though, the measurability problem (ii) arises unless L is countable. The product space form of the theorem is typically used in the construction of Gaussian process distributions on random functions [2]. For random measures, a more adequate projective system is constructed in following section.

3. Projective limits of probability simplices

This section constitutes the first part of the proof of Theorem 1.1: The construction of a projective limit space $\mathcal{X}_{\mathrm{D}}$ from simplices $\triangle_{\mathrm{I}}$, and the analysis of its properties. The space $\mathcal{X}_{\mathrm{D}}$ turns out to consist of set functions which are not necessarily σ-additive, and the remaining part of the proof in Sec. 4 will be the derivation of a criterion for σ-additivity.

The distinction between finitely additive and σ-additive set functions will be crucial to the ensuing discussion. We consider two types of set systems $\mathcal{Q}$ on the space V: *Algebras*, which contain both $\emptyset$ and V, and are closed under complements and finite unions, and *σ-algebras*, which are algebras and additionally closed under countable unions. A non-negative set function μ on either an algebra or σ-algebra $\mathcal{Q}$ is called a *charge* if it satisfies $\mu(\emptyset) = 0$ and is finitely additive. If a charge is normalized, i.e. if $\mu(V) = 1$, it is called a *probability charge*. A charge is a measure if and only if it is σ-additive. If $\mathcal{Q}$ is an algebra, and not closed under countable unions, the definition of σ-additivity only requires μ to be additive along those countable sequences of sets $A_n \in \mathcal{Q}$ whose union is in $\mathcal{Q}$.

3.1. Definition of the projective system

For the choice of components in a projective system, it can be helpful to regard the elements x_{D} of the projective limit space $\mathcal{X}_{\mathrm{D}}$ as mappings, from a domain defined by the index set D to a range defined by the spaces $\mathcal{X}_{\mathrm{I}}$. The simplest example is once again the product space $\mathcal{X}_{\mathrm{D}} = \mathcal{X}_0^{\mathrm{L}}$ in Kolmogorov's theorem, for which each $x_{\mathrm{D}} \in \mathcal{X}_{\mathrm{D}}$ can be interpreted as a function $x_{\mathrm{D}} : L \to \mathcal{X}_0$. Probability measures on $(V, \mathcal{B}_{\mathrm{V}})$ are in particular set functions $\mathcal{B}_{\mathrm{V}} \to [0, 1]$, so it is natural to construct D from the sets in $\mathcal{B}_{\mathrm{V}}$. It is not necessary to include all measurable sets: If $\mathcal{Q}$ is an algebra that generates $\mathcal{B}_{\mathrm{V}}$, any probability measure on $\mathcal{Q}$ has, by Carathéodory's theorem [16, Theorem 2.5], a unique extension to a probability measure on $\mathcal{B}_{\mathrm{V}}$. In other words, the space $\mathbf{M}(\mathcal{Q})$ of probability measures on $\mathcal{Q}$ is isomorphic to $\mathbf{M}(V)$, and $\mathcal{Q}$ can be substituted for $\mathcal{B}_{\mathrm{V}}$ in the projective limit construction.

Desiderata for the projective limit are: (1) The projective limit space $\mathcal{X}_{\mathrm{D}}$ should contain all measures on $\mathcal{Q}$ (and hence on $\mathcal{B}_{\mathrm{V}}$). (2) $\mathcal{Q}$ should be countable, to address the measurability problem (ii) in Sec. 1. (3) The marginal spaces $\mathcal{X}_{\mathrm{I}}$ should consist of the finite-dimensional analogues of measures on $\mathcal{Q}$, and hence of measures on finite subsets of events in $\mathcal{Q}$. (4) The definition of the system should facilitate a proof of σ-additivity. In this section, we will recapitulate the projective limit specified in Sec. 1.1 and show it indeed satisfies (1)-(3); that (4) is satisfied as well will be shown in Sec. 4.

Choice of $\mathcal{Q}$. We start with the prototypical choice of basis for any Polish topology: Let $W \subset V$ be a countable, dense subset of V. Fix a metric $d : V \times V \to \mathbb{R}_+$ which generates the topology $\mathcal{T}_{\mathrm{V}}$, and denote by $B(v, r)$ the open d-ball of radius

r around v. Denote the system of open balls with rational radii and centers in W by

$$\mathcal{U} := \{B(v,r) | v \in W, r \in \mathbb{Q}_+\} \cup \{\emptyset\} . \tag{3.1}$$

Since V is separable and metrizable, $\mathcal{U}$ forms a countable basis of the topology $\mathcal{T}_{\text{V}}$ [1, Lemma 3.4]. Let $\mathcal{Q}(\mathcal{U})$ be the algebra generated by $\mathcal{U}$. Then $\mathcal{U} \subset \mathcal{Q}(\mathcal{U}) \subset \mathcal{B}_{\text{V}}$. In particular, $\mathcal{Q}(\mathcal{U})$ is a countable generator of $\mathcal{B}_{\text{V}}$.

Index set. As the index set D, we do not choose $\mathcal{Q}(\mathcal{U})$ itself, but rather the set of all finite partitions of V consisting of disjoint sets $A_i \in \mathcal{Q}(\mathcal{U})$,

$$D := \mathcal{H}(\mathcal{Q}) = \left\{ (A_1, \dots, A_n) \,\middle|\, n \in \mathbb{N},\ A_i \in \mathcal{Q}(\mathcal{U}),\ \dot{\cup} A_i = V \right\} . \tag{3.2}$$

Each element $I \in D$ is a finite partition, and the set of probability measures on the events in this partition is precisely the simplex $\triangle_{\text{I}}$. To define a partial order on D, let $I = (A_1, \dots, A_m)$ and $J = (B_1, \dots, B_n)$ be any two partitions in D, and denote their intersection (common refinement) by $I \cap J := (A_i \cap B_j)_{i,j}$. Since $\mathcal{Q}(\mathcal{U})$ forms an algebra, $I \cap J$ is again an element of D. Now define a partial order relation $\preceq$ as

$$I \preceq J \quad :\Leftrightarrow \quad I \cap J = J , \tag{3.3}$$

that is, $I \preceq J$ iff J is a refinement of I. The set $(D, \preceq)$ is a valid index set for a projective limit system, because it is directed: $K := I \cap J$ always satisfies $I \preceq K$ and $J \preceq K$.

Projection functions. What remains to be done is to specify the functions f_{JI}. Consider a partition $J = (A_1, \dots, A_n)$, and any $x_{\text{J}} \in \triangle_{\text{J}}$. Each entry $x_{\text{J}}(A_j)$ assigns a number (a probability) to the event A_j, and we define f_{JI} accordingly to preserve this property. To this end, let $J = (B_1, \dots, B_n)$ be a partition in D, and let $I = (A_1, \dots, A_m)$ be a coarsening of J (that is, $I \preceq J$). For each A_i, let $\mathcal{J}_i \subset \{1, \dots, n\}$ be the subset of indices for which $A_i = \cup_{j \in \mathcal{J}_i} B_j$. Then define f_{JI} as

$$(f_{\text{JI}} x_{\text{J}})(A_i) := \sum_{j \in \mathcal{J}_i} x_{\text{J}}(B_j) . \tag{3.4}$$

We choose $\mathcal{X}_{\text{I}} := \triangle_{\text{I}}$ as defined in (1.1), and endow $\triangle_{\text{I}}$ with the relative topology $\mathcal{T}_{\text{I}} := \mathcal{T}(\mathbb{R}^{\text{I}}) \cap \triangle_{\text{I}}$ and the corresponding Borel sets $\mathcal{B}_{\text{I}} := \mathcal{B}(\mathcal{T}_{\text{I}}) = \mathcal{B}(\mathbb{R}^{\text{I}}) \cap \triangle_{\text{I}}$. The relative topology makes additions on $\triangle_{\text{I}}$, and hence the mappings f_{JI}, continuous. Each f_{II} is the identity on $\triangle_{\text{I}}$, and $f_{\text{KI}} = f_{\text{KJ}} \circ f_{\text{JI}}$. For any pair $I \preceq J \in D$, $\triangle_{\text{I}} = f_{\text{JI}} \triangle_{\text{J}}$ and conversely, $\triangle_{\text{J}} = f_{\text{JI}}^{-1} \triangle_{\text{I}}$. Therefore, $\left\langle \triangle_{\text{I}}, \mathcal{B}_{\text{I}}, f_{\text{JI}} \right\rangle_{\text{D}}$ is a projective system.

3.2. Structure of the projective limit space

Let $(\mathcal{X}_{\text{D}}, \mathcal{B}_{\text{D}})$ be the projective limit of $\left\langle \triangle_{\text{I}}, \mathcal{B}_{\text{I}}, f_{\text{JI}} \right\rangle_{\text{D}}$. We observe immediately that $\mathcal{X}_{\text{D}}$ contains $\mathbf{M}(\mathcal{Q})$: If x is a probability measure on $\mathcal{Q}(\mathcal{U})$, let $x_{\text{I}} := f_{\text{I}} x$

for each partition $I \in D$. The collection $\{x_{\mathrm{I}} | I \in D\}$ satisfies (2.1), and hence constitutes a point in $\mathcal{X}_{\mathrm{D}}$. The following result provides more details about the constructed measurable space $(\mathcal{X}_{\mathrm{D}}, \mathcal{B}_{\mathrm{D}})$, which turns out to be the space $\mathbf{C}(\mathcal{Q})$ of all probabiliy charges defined on $\mathcal{Q}(\mathcal{U})$. By $\mathcal{B}_{w^*}$, we again denote the Borel σ-algebra on $\mathbf{M}(V)$ generated by the weak* topology.

Proposition 3.1. *Let V be a Polish space, and $(\mathcal{X}_{\mathrm{D}}, \mathcal{B}_{\mathrm{D}})$ the projective limit of the projective system $\langle \triangle_{\mathrm{I}}, \mathcal{B}_{\mathrm{I}}, f_{\mathrm{JI}} \rangle_{\mathrm{D}}$ defined in Sec. 3.1. Denote by $\psi : \mathbf{M}(V) \to \mathbf{M}(\mathcal{Q})$ the restriction mapping which takes each measure x on $\mathcal{B}_{\mathrm{V}}$ to its restriction $x_{\mathrm{D}} = x|_{\mathcal{Q}}$ on $\mathcal{Q} \subset \mathcal{B}_{\mathrm{V}}$. Then the following hold:*

(i) $\mathcal{X}_{\mathrm{D}} = \mathbf{C}(\mathcal{Q})$*, the space of probability charges on* $\mathcal{Q}(\mathcal{U})$*.*
(ii) $\mathbf{M}(\mathcal{Q})$ *is a measurable subset of* $\mathbf{C}(\mathcal{Q})$*.*
(iii) ψ *is a Borel isomorphism of* $(\mathbf{M}(V), \mathcal{B}_{w^*})$ *and* $(\mathbf{M}(\mathcal{Q}), \mathcal{B}_{\mathrm{D}} \cap \mathbf{M}(\mathcal{Q}))$*.*

Part (ii) implies that a projective limit measure P_{D} constructed on $\mathbf{C}(\mathcal{Q})$ by means of Theorem 2.2 can be restricted to a measure on $\mathbf{M}(\mathcal{Q})$ without further complications, in particular without appealing to outer measures. According to (iii), there is a measure P on $\mathbf{M}(V)$ which can be regarded as equivalent to P_{D}, namely the image measure $P := \psi^{-1} P_{\mathrm{D}}$ under the inverse of the restriction map ψ. This is of course the measure P described in Theorem 1.1, though some details still remain to be established later on. Since ψ is a Borel isomorphism, P constitutes a measure with respect to the "natural" topology on $\mathbf{M}(V)$.

Proof. Part (i). Let $x_{\mathrm{D}} \in \mathcal{X}_{\mathrm{D}}$. The trivial partition $I_0 := (V)$ is in D, which implies $x_{\mathrm{D}}(V) = f_{I_0} x_{\mathrm{D}} = 1$ and $x_{\mathrm{D}}(\emptyset) = 0$. To show finite additivity, let $A_1, A_2 \in \mathcal{Q}(\mathcal{U})$ be disjoint sets and choose a partition $J \in D$ such that $A_1, A_2 \in J$. Let $I \preceq J$ be the coarsening of J obtained by joining the two sets. As the elements of each space $\triangle_{\mathrm{I}}$ are finitely additive,

$$x_{\mathrm{D}}(A_1) + x_{\mathrm{D}}(A_2) = (f_{\mathrm{J}} x_{\mathrm{D}})(A_1) + (f_{\mathrm{J}} x_{\mathrm{D}})(A_2) \overset{(3.4)}{=} (f_{\mathrm{I}} x_{\mathrm{D}})(A_1 \cup A_2) = x_{\mathrm{D}}(A_1 \cup A_2)\ .$$

Hence, x_{D} is a charge. Conversely, assume that x_{D} is a probability charge on $\mathcal{Q}(\mathcal{U})$. The evaluation $f_{\mathrm{I}} x_{\mathrm{D}}$ of x_{D} on a partition $I \in D$ defines a probability measure on the finite σ-algebra $\sigma(I)$, and thus $f_{\mathrm{I}} x_{\mathrm{D}} \in \triangle_{\mathrm{I}}$. Since additionally $f_{\mathrm{JI}}(f_{\mathrm{J}} x_{\mathrm{D}}) = f_{\mathrm{I}} x_{\mathrm{D}}$, the set $\langle f_{\mathrm{I}} x_{\mathrm{D}} \rangle_{\mathrm{D}}$ forms a collection of points $f_{\mathrm{I}} x_{\mathrm{D}} \in \triangle_{\mathrm{I}}$ satisfying (2.1), and hence $x_{\mathrm{D}} \in \mathcal{X}_{\mathrm{D}}$.
Part (ii). Regard the restriction map ψ as a mapping into $\mathbf{C}(\mathcal{Q})$, with image $\mathbf{M}(\mathcal{Q})$. By Caratheodory's extension theorem, ψ is injective [16, Theorem 2.5]. If an injective mapping between Polish spaces is measurable, its inverse is measurable as well [16, Theorem A1.3]. Thus, if we can show ψ to be measurable, $\mathbf{M}(\mathcal{Q}) = \psi(\mathbf{M}(V))$ is a measurable set.

First observe that ψ relates the evaluation functionals $f_{\mathrm{I}} : \mathbf{C}(\mathcal{Q}) \to \triangle_{\mathrm{I}}$ on probability charges to the evaluation functionals $\phi_{\mathrm{I}} : \mathbf{M}(V) \to \triangle_{\mathrm{I}}$ on probability measures via the equations

$$\phi_{\mathrm{I}} = f_{\mathrm{I}} \circ \psi \qquad \text{for all } I \in D\ . \tag{3.5}$$

We will show that the mappings ϕ_{I} generate the σ-algebra $\mathcal{B}_{w^*}$ on $\mathbf{M}(V)$. Since the canonical mappings f_{I} generate $\mathcal{B}_{\text{D}}$ on $\mathbf{C}(\mathcal{Q})$ by definition, (3.5) then implies $\mathcal{B}_{w^*}$-$\mathcal{B}_{\text{D}}$-measurability of ψ:

Let $\phi_A : \mathbf{M}(V) \to [0,1]$ be the evaluation functional $x \mapsto x(A)$. Since $\mathbf{M}(V)$ is separable, the Borel sets of the weak* topology coincide with those generated by the maps ϕ_A [11, Theorem 2.3], thus $\mathcal{B}_{w^*} = \sigma(\phi_A | A \in \mathcal{B}_{\text{V}})$. Each mapping ϕ_A can be identified with ϕ_{I} for $I = (A, A^c)$, because $\phi_{(A,A^c)}(x) = (x(A), 1 - x(A))$. Hence equivalently, $\mathcal{B}_{w^*} = \sigma(\phi_{(A,A^c)} | A \in \mathcal{B}_{\text{V}})$, and with (3.5),

$$\mathcal{B}_{w^*} = \psi^{-1}\sigma(f_{(A,A^c)} | A \in \mathcal{B}_{\text{V}}) \; . \tag{3.6}$$

Clearly, the maps $f_{(A,A^c)}$ for $A \in \mathcal{Q}$ are sufficient to express all information expressible by the larger family of maps f_{I}, $I \in D$, and thus generate the projective limit σ-algebra,

$$\sigma(f_{(A,A^c)} \,|\, A \in \mathcal{Q}) = \mathcal{B}_{\text{D}} \; . \tag{3.7}$$

In summary, ψ is $\mathcal{B}_{w^*}$-$\mathcal{B}_{\text{D}}$-measurable, and we deduce $\mathbf{M}(\mathcal{Q}) = \psi(\mathbf{M}(V)) \in \mathcal{B}_{\text{D}}$.
Part (iii). As shown above, ψ is injective and measurable, and regarded as a mapping onto its image $\mathbf{M}(\mathcal{Q})$, it is trivially surjective. What remains to be shown is measurability of the inverse. By part *(ii)*, the image $\psi(\mathbf{M}(V)) = \mathbf{M}(\mathcal{Q})$ is a Borel subset of $\mathbf{C}(\mathcal{Q})$. As a countable projective limit of Polish spaces, $(\mathbf{C}(\mathcal{Q}), \mathcal{T}_{\text{D}})$ is Polish [5, Chapter IX]. Since $\mathbf{M}(V)$ is Polish, $(\mathbf{M}(V), \mathcal{B}_{w^*})$ is a standard Borel space, i.e. a Borel space generated by a Polish topology. The space $(\mathbf{M}(\mathcal{Q}), \mathcal{B}_{\text{D}} \cap \mathbf{M}(\mathcal{Q}))$ is standard Borel as well, since $\mathbf{M}(\mathcal{Q})$ is a Borel subset of a Polish space [16, Theorem A1.2]. As noted above, measurable bijections between standard Borel spaces are automatically bimeasurable [16, Theorem A1.3], which shows ψ to be a Borel isomorphism. □

4. σ-additivity of random charges

The previous section provides the means to construct the distribution P_{D} of a random charge $X_{\text{D}} : \Omega \to \mathbf{C}(\mathcal{Q})$ as a projective limit measure. To obtain random measures rather than random charges in this manner, we need to additionally ensure that P_{D} concentrates on the measurable subspace $\mathbf{M}(V)$, or in other words, that X_{D} is σ-additive $\mathbb{P}$-almost surely.

Consider a projective limit random charge X_{D}, distributed according to a projective limit measure P_{D} on $\mathbf{C}(\mathcal{Q})$. The following proposition gives a necessary and sufficient condition for almost sure σ-additivity of X_{D}, formulated in terms of its expectation $\mathbb{E}_{P_{\text{D}}}[X_{\text{D}}]$. It also shows that the expected values of P_{D} and the projective family $\langle P_{\text{I}} \rangle_{\text{D}}$ are themselves projective, in the sense that $f_{\text{I}}\mathbb{E}_{P_{\text{D}}}[X_{\text{D}}] = \mathbb{E}_{P_{\text{I}}}[X_{\text{I}}]$, and accordingly $f_{\text{JI}}\mathbb{E}_{P_{\text{J}}}[X_{\text{J}}] = \mathbb{E}_{P_{\text{I}}}[X_{\text{I}}]$ for any pair $I \preceq J$. The latter makes the criterion directly applicable to construction problems: If we initiate the construction by choosing an expected measure $G_0 \in \mathbf{M}(V)$ for the prospective measure P_{D}, and then choose the projective family such that $\mathbb{E}_{P_{\text{I}}}[X_{\text{I}}] = f_{\text{I}} G_0$, random draws from P_{D} will take values in $\mathbf{M}(V)$ almost surely.

Proposition 4.1. *Let $(\mathcal{X}_{\mathrm{D}}, \mathcal{B}_{\mathrm{D}})$ be the projective limit of finite-dimensional probability simplices defined in Proposition 3.1, and let $\langle P_{\mathrm{I}} \rangle_{\mathrm{D}}$ be a projective family of probability measures on the spaces $(\triangle_{\mathrm{I}}, \mathcal{B}_{\mathrm{I}})$. Denote by P_{D} the projective limit measure, and by $G_0 := \mathbb{E}_{P_{\mathrm{D}}}[X_{\mathrm{D}}]$ its expectation. Then:*

(i) The expectation G_0 is an element of $\mathcal{X}_{\mathrm{D}}$ and

$$f_{\mathrm{I}} G_0 = \mathbb{E}_{P_{\mathrm{I}}}[X_{\mathrm{I}}] \qquad \text{for any } I \in D\,. \tag{4.1}$$

(ii) X_{D} is σ-additive $\mathbb{P}$-almost surely if and only if G_0 is σ-additive.

The proof requires a criterion for σ-additivity of probability charges expressible in terms of a countable number of conditions. Assuming that G_0 is σ-additive, we will deduce from the projective limit construction that, if a fixed sequence of sets is given, the random content X_{D} is countably additive along this sequence with probability one. This only implies almost sure σ-additivity of X_{D} on $\mathcal{Q}(\mathcal{U})$ if the condition for σ-additivity can be reduced to a countable subset of sequences in $\mathcal{Q}(\mathcal{U})$ (cf. Appendix A.3). Such a reduction was derived by Harris [14, Lemma 6.1]. For our particular choice of $\mathcal{Q}(\mathcal{U})$, his result can be stated as follows:

Lemma 4.2 (Harris)**.** *Let V be any Polish space and $\mathcal{Q}(\mathcal{U})$ the countable algebra generated by the open balls (3.1). Then the set of all sequences of elements of $\mathcal{Q}(\mathcal{U})$ contains a countable subset of sequences $(A_n^m)_n$, where $A_n^m \searrow \emptyset$ for all $m \in \mathbb{N}$, such that any probability charge μ on $\mathcal{Q}(\mathcal{U})$ is σ-additive if and only if it satisfies*

$$\lim_{n\to\infty} \mu(A_n^m) = 0 \qquad \text{for all } m \in \mathbb{N}\,. \tag{4.2}$$

Proof of Proposition 4.1. Part (i). The expectation $\mathbb{E}_{P_{\mathrm{D}}}[X_{\mathrm{D}}]$ is finitely additive: For any finite number of disjoint sets $A_i \in \mathcal{B}_{\mathrm{D}}$,

$$\sum_{i=1}^{n} \mathbb{E}_{P_{\mathrm{D}}}[X_{\mathrm{D}}](A_i) = \int_{\mathbf{C}(\mathcal{Q})} \sum_{i=1}^{n} x_{\mathrm{D}}(A_i) P_{\mathrm{D}}(dx_{\mathrm{D}}) = \mathbb{E}_{P_{\mathrm{D}}}[X_{\mathrm{D}}](\cup_i A_i)\,. \tag{4.3}$$

Since clearly also $\mathbb{E}_{P_{\mathrm{D}}}[X_{\mathrm{D}}](\emptyset) = 0$ and $\mathbb{E}_{P_{\mathrm{D}}}[X_{\mathrm{D}}](V) = 1$, the expectation is an element of $\mathcal{X}_{\mathrm{D}}$. To verify (4.1), note the mappings $f_{\mathrm{JI}} : \triangle_{\mathrm{J}} \to \triangle_{\mathrm{I}}$ are affine, and hence

$$\begin{aligned} f_{\mathrm{JI}} \mathbb{E}_{P_{\mathrm{J}}}[X_{\mathrm{J}}] \overset{f \text{affine}}{=} \mathbb{E}_{P_{\mathrm{J}}}[f_{\mathrm{JI}} X_{\mathrm{J}}] &= \int_{\triangle_{\mathrm{J}} = f_{\mathrm{JI}}^{-1} \triangle_{\mathrm{I}}} f_{\mathrm{JI}} x_{\mathrm{J}} P_{\mathrm{J}}(dx_{\mathrm{J}}) \\ &= \int_{\triangle_{\mathrm{I}}} x_{\mathrm{I}} (f_{\mathrm{JI}} P_{\mathrm{J}})(dx_{\mathrm{I}}) = \int_{\triangle_{\mathrm{I}}} x_{\mathrm{I}} P_{\mathrm{I}}(dx_{\mathrm{I}})\,. \end{aligned} \tag{4.4}$$

Therefore, the expectations of a projective family $\langle P_{\mathrm{I}} \rangle_{\mathrm{D}}$ satisfy $f_{\mathrm{JI}} \mathbb{E}_{P_{\mathrm{J}}}[X_{\mathrm{J}}] = \mathbb{E}_{P_{\mathrm{I}}}[X_{\mathrm{I}}]$. By the same device, $f_{\mathrm{I}} G_0 = f_{\mathrm{I}} \mathbb{E}_{P_{\mathrm{D}}}[X_{\mathrm{D}}] = \mathbb{E}_{P_{\mathrm{I}}}[X_{\mathrm{I}}]$ holds for the projective limit measure P_{D}.

Part (ii). First assume that G_0 is σ-additive. Let $(A_n^m)_n$ be any of the set sequences given by Lemma 4.2. As $n \to \infty$, the random sequence $(X_{\mathrm{D}}(A_n^m))$

converges to 0 almost surely: σ-Additivity of G_0 implies

$$\lim_{n\to\infty} \mathbb{E}_{P_{\mathrm{D}}}[X_{\mathrm{D}}](A_n^m) = \lim_{n\to\infty} G_0(A_n^m) = G_0(\emptyset) = \mathbb{E}_{P_{\mathrm{D}}}[X_{\mathrm{D}}](\emptyset) \,, \tag{4.5}$$

hence $X_{\mathrm{D}}(A_n^m) \xrightarrow{L_1} 0$. The sequence (A_n^m) is decreasing and the random variable X_{D} is charge-valued, which implies $X_{\mathrm{D}}(A_{n+1}^m) \leq X_{\mathrm{D}}(A_n^m)$ a.s. In particular, the sequence $(X_{\mathrm{D}}(A_n^m))$ forms a supermartingale when endowed with its canonical filtration. For supermartingales, convergence in the mean implies almost sure convergence [2, Theorem 19.3], and thus indeed $X_{\mathrm{D}}(A_n^m) \xrightarrow{\text{a.s.}} 0$.

Consequently, there is a $\mathbb{P}$-null subset N_m of the abstract probability space Ω such that

$$(X_{\mathrm{D}}(\omega))(A_n^m) \xrightarrow{n\to\infty} (X_{\mathrm{D}}(\omega))(\emptyset) \qquad \text{for } \omega \notin N_m \,. \tag{4.6}$$

The union $N := \cup_{m\in\mathbb{N}} N_m$ of these null sets, taken over all sequences (A_n^m) required by Lemma 4.2, is again a $\mathbb{P}$-null set. The charge $X_{\mathrm{D}}(\omega)$ satisfies (4.2) for all m whenever $\omega \notin N$. Therefore, X_{D} is σ-additive $\mathbb{P}$-a.s. by Lemma 4.2, and hence almost surely a probability measure.

Conversely, let X_{D} assume values in $\mathbf{M}(V) \cong \mathbf{M}(\mathcal{Q})$ almost surely. Since $A_n^m \searrow \emptyset$, the sequence of measurable functions $\omega \mapsto (X_{\mathrm{D}}(\omega))(A_n^m)$ converges to 0 almost everywhere. By hypothesis, $\mathbf{C}(\mathcal{Q}) \smallsetminus \mathbf{M}(\mathcal{Q})$ is a null set, hence

$$\lim_{n\to\infty} \mathbb{E}_{P_{\mathrm{D}}}[X_{\mathrm{D}}](A_n^m) = \lim_{n\to\infty} \int_{X_{\mathrm{D}}^{-1}\mathbf{M}(\mathcal{Q})} (X_{\mathrm{D}}(\omega))(A_n^m)\mathbb{P}(d\omega) = 0 \,, \tag{4.7}$$

where the second identity holds by dominated convergence [16, Theorem 1.21]. Since $\mathbb{E}_{P_{\mathrm{D}}}[X_{\mathrm{D}}]$ is a probability charge according to part (i) and satisfies (4.7), it satisfies the conditions of Lemma 4.2, and we conclude $\mathbb{E}_{P_{\mathrm{D}}}[X_{\mathrm{D}}] \in \mathbf{M}(\mathcal{Q})$. □

Theorem 1.1 is now finally obtained by deducing the properties of P from those of P_{D} as established by Proposition 4.1.

Proof of Theorem 1.1. First suppose that (1.3) and (1.4) hold. By Theorem 2.2, a unique projective limit measure P_{D} exists on $\mathcal{X}_{\mathrm{D}} = \mathbf{C}(\mathcal{Q})$, with $f_{\mathrm{I}}P_{\mathrm{D}} = P_{\mathrm{I}}$. Proposition 4.1$(ii)$ shows P_{D} is concentrated on the measurable subset $\mathbf{M}(\mathcal{Q})$. By Proposition 3.1(iii), it uniquely defines an equivalent measure $P := \psi^{-1}P_{\mathrm{D}}$ on $\mathbf{M}(V)$, which satisfies (1.5). As a probability measure on a Polish space, P is a Radon measure [6, IX.3.3, Proposition 3].

Conversely, assume that P is given. Then (1.3) follows from (1.5). The expectation $G_0 = \mathbb{E}_P[X]$ is in $\mathbf{M}(V)$ by Proposition 4.1(ii). Any measure on $\mathbf{M}(V)$ can be represented as a measure on $\mathcal{X}_{\mathrm{D}} = \mathbf{C}(\mathcal{Q})$, hence by Proposition 4.1(i), the expectation G_0 and the marginals $P_{\mathrm{I}} = f_{\mathrm{I}}P$ satisfy (4.1). Thus, (1.4) holds, and the proof is complete. □

Appendix A: Review of technical problems

This appendix provides a more detailed description of problems (i)–(iii) listed in Sec. 1. The discussion addresses readers of passing familiarity with measure-theoretic probability; to the probabilist, it will only state the obvious.

The approach proposed in [8] is, in summary, the following: A probability measure on $(V, \mathcal{B}_{\mathrm{V}})$ is a set function $\mathcal{B}_{\mathrm{V}} \to [0,1]$. The set $\mathbf{M}(V)$ of probability measures can be regarded as a subset of the space $[0,1]^{\mathcal{B}_{\mathrm{V}}}$ of all such functions. More precisely, the space chosen in [8] is $[0,1]^{\mathcal{H}(\mathcal{B}_{\mathrm{V}})}$, where $\mathcal{H}(\mathcal{B}_{\mathrm{V}})$ again denotes the set of all measurable, finite partitions of V. This space contains one axis for each partition, and hence is a larger space than $[0,1]^{\mathcal{B}_{\mathrm{V}}}$, but redundantly encodes the same information. The Kolmogorov extension theorem [16, Theorem 6.16] is then applied to a family of Dirichlet distributions defined on the finite-dimensional subspaces of the product space $[0,1]^{\mathcal{H}(\mathcal{B}_{\mathrm{V}})}$.

A.1. Product spaces

The Kolmogorov extension theorem used in the construction is not well-adapted to the problem of constructing measures on measures, because the setting assumed by the theorem is that of a product space: A finite-dimensional marginal of a measure P on $\mathbf{M}(V)$ is a measure P_{I} on the set of measures over a finite σ-algebra $\mathcal{C}$ of events. Any such σ-algebra can be generated by a partition I of events in $\mathcal{B}_{\mathrm{V}}$. The set consisting of the marginals on I of all measures $x \in \mathbf{M}(V)$ is necessarily isomorphic to the unit simplex in $|I|$-dimensional Euclidean space. Hence, the marginals of a measure P defined *on* $\mathbf{M}(V)$ always live on simplices of the form $\triangle_{\mathrm{I}}$ as described in Sec. 1.1. In other words, when we set up a projective limit construction for measures on $\mathbf{M}(V)$, the choice of possible finite-dimensional marginal spaces is limited—either the simplices are used directly, as in Sec. 1.1, or they are embedded into some other finite-dimensional space. If the projective limit result to be applied is the Kolmogorov extension theorem, the simplices must be embedded into Euclidean product spaces, as proposed in [8]. The problem here is that it is difficult to properly formalize marginalization to subspaces, as required by the theorem. For constructions on $[0,1]^{\mathcal{B}_{\mathrm{V}}}$, the problem can be illustrated by the example in Fig. 1: For $J = (B_1, B_2, B_3)$, the simplex $\triangle_{\mathrm{J}}$ is a subspace of $\mathbb{R}^{\mathrm{J}} \cong \mathbb{R}^3$. Marginalization corresponds to merging two events, such as B_1 and B_2 in the example. The resulting simplex $\triangle_{\mathrm{I}}$ for $I = (B_1 \cup B_2, B_3)$ is a subspace of $\mathbb{R}^{\mathrm{I}}$. However, $\mathbb{R}^{\mathrm{I}}$ is not a subspace of $\mathbb{R}^{\mathrm{J}}$, nor is $\triangle_{\mathrm{I}}$ a subspace of $\triangle_{\mathrm{J}}$. Hence, in the product space setting of the Kolmogorov theorem, the natural way to formalize a reduction in dimension for measures on a finite number of events does not correspond to a projection onto a subspace.

A.2. Measurability problems

A general property of projective limit constructions of stochastic processes is that the index set—intuitively, the set of axes labels of a product, or of dimensions in a more general setting—must be countable to obtain a useful probability measure. This is due to the fact that all projective limit theorems implicitly generate a σ-algebra on the infinite-dimensional space—the σ-algebra $\mathcal{B}_{\mathrm{D}}$ specified by (2.2)—based on the σ-algebras on the marginal spaces used in the construction. The constructed measure lives on this σ-algebra.

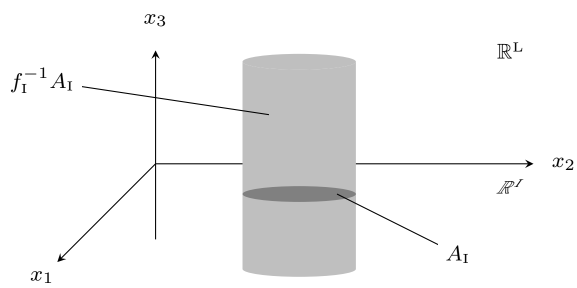


FIG 2. *Three-dimensional analogue of a cylinder set in the product space setting. An event $A_D \subset \mathbb{R}^D$ is independent of the random variable X_3 if it is the preimage $A_D = f_I^{-1}A_I$ of some event $A_I \subset \mathbb{R}^I$, that is, if the set A_D is of "axis parallel" shape in direction of X_3. The event A_I in the figure occurs if $(X_1, X_2) \in A_I$, or equivalently, if $(X_1, X_2, X_3) \in f_{JI}^{-1}A_I$.*

If the dimension is uncountable, the resolution of the σ-algebra is too coarse to resolve most events of interest. In particular, it does not contain singletons. The problem is most readily illustrated in the product space setting: Suppose the Kolmogorov theorem is used to define a measure P on an infinite-dimensional product space $\mathcal{X}_D := \mathbb{R}^L$, where L is some infinite set. The measure P is constructed from given measures P_I defined on the finite-dimensional sub-products $\mathbb{R}^I$, where $I \in D$ are finite subsets of L. The σ-algebra on $\mathbb{R}^L$ on which P_D is defined is generated as follows: Denote by f_I the product space projector $\mathbb{R}^L \to \mathbb{R}^I$. For any measurable set $A^I \in \mathbb{R}^I$, the preimage $f_I^{-1}A_I$ is a subset of $\mathbb{R}^L$, which is of "axis-parallel" shape in direction of all axis not contained in I. The finite-dimensional analogue of this situation is illustrated in Fig. 2, where A^I is assumed to be an elliptically shaped set in the plane $\mathbb{R}^I$, and the overall space $\mathbb{R}^L$ is depicted as three-dimensional. Preimages $f_I^{-1}A_I$ of measurable sets are, for obvious reasons, called *cylinder sets* in the probability literature. The σ-algebra defined by the Kolmogorov theorem is the smallest σ-algebra containing all cylinder sets $f_I^{-1}A_I$, for all measurable sets $A_I \in \mathbb{R}^I$ and all finite sub-products $\mathbb{R}^I$. Since σ-algebras are defined by closure under countable operations, the sets in this σ-algebra can be thought of as cylinder sets that are of axis-parallel shape along all but a countable number of dimensions. If the overall space is of countable dimension, any set of interest can be expressed in this form. If the dimension is uncountable, however, these events only specify the joint behavior of a countable subset of random variables—in Fig. 2, $\mathbb{R}^I$ would represent a subspace of countable dimension of the uncountable-dimensional space $\mathbb{R}^L$.

For example, consider the set $\mathbb{R}^L := \mathbb{R}^{\mathbb{R}}$, regarded as the set of all functions $x_D : \mathbb{R} \to \mathbb{R}$, which arises in the construction of Gaussian processes. Although the constructed measure P_D is a distribution on random functions x_D, this measure cannot assign a probability to events of the form $\{X_D = x_D\}$, i.e. to the event that the outcome of a random draw is a particular function x_D. The only measurable events are of the form $\{X_D(s_1) = t_1, X_D(s_2) = t_2, \dots\}$ and specify the value of the function at a countable subset of points $s_1, s_2, \dots \in \mathbb{R}$.

A.3. σ-additivity

The marginal distributions used in the construction specify the joint behavior of the constructed measure P_{D} on any finite subset of measurable sets. σ-additivity requires additivity along an infinite sequence, and cannot be deduced directly from additivity of the marginals. Suppose that some sequence $A_1, A_2, \ldots$ of measurable sets in V is given, and that x_{D} is a random set function drawn from P_{D}. Countable additivity of x_{D} along the sequence can be shown to hold almost surely (with respect to P_{D}) by means of a simple convergence argument [8, Proposition 2]. However, as a σ-algebra, $\mathcal{B}_{\mathrm{V}}$ is either finite or uncountable. Hence, if V is infinite, $\mathcal{B}_{\mathrm{V}}$ contains an uncountable number of such sequences. Even though x_{D} is additive along any given sequence with probability one, the null sets of exceptions aggregate into a non-null set over all sequences, and x_{D} is not σ-additive with probability one. Substituting a countable generator $\mathcal{Q}$ for $\mathcal{B}_{\mathrm{V}}$ does not resolve the problem, since the number of sequences in $\mathcal{Q}$ remains uncountable.

Acknowledgments

I would like to thank the associate editor and two referees for valuable suggestions, in particular for pointing out the example in Corollary 1.4. I am grateful to Daniel M. Roy for helpful comments and corrections.

References

[1] Aliprantis, C. D. and Border, K. C. (2006). *Infinite Dimensional Analysis*. Springer, 3rd edition. MR2378491
[2] Bauer, H. (1996). *Probability Theory*. W. de Gruyter. MR1385460
[3] Blackwell, D. and MacQueen, J. B. (1973). Ferguson distributions via Pólya urn schemes. *Ann. Statist.*, **1**, 353–355. MR0362614
[4] Bochner, S. (1955). *Harmonic Analysis and the Theory of Probability*. University of California Press. MR0072370
[5] Bourbaki, N. (1966). *Elements of Mathematics: General Topology*. Hermann (Paris) and Addison-Wesley.
[6] Bourbaki, N. (2004). *Elements of Mathematics: Integration*. Springer.
[7] Crauel, H. (2002). *Random probability measures on Polish spaces*. Taylor & Francis. MR1993844
[8] Ferguson, T. S. (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.*, **1**(2). MR0350949
[9] Ferguson, T. S. (1974). Prior distributions on spaces of probability measures. *Ann. Statist.*, **2**(4), 615–629. MR0438568
[10] Fremlin, D. H. (2000–2006). *Measure Theory*, volume I–IV. Torres Fremlin. MR2462372
[11] Gaudard, M. and Hadwin, D. (1989). Sigma-algebras on spaces of probability measures. *Scand. J. Stat.*, **16**, 169–165. MR1028976

[12] Ghosal, S. (2010). Dirichlet process, related priors and posterior asymptotics. In N. L. Hjort *et al.*, editors, *Bayesian Nonparametrics*. Cambridge University Press. MR2730660

[13] Ghosh, J. K. and Ramamoorthi, R. V. (2002). *Bayesian Nonparametrics*. Springer. MR1992245

[14] Harris, T. E. (1968). Counting measures, monotone random set functions. *Probab. Theory Related Fields*, **10**, 102–119. MR0235592

[15] Kallenberg, O. (1983). *Random Measures*. Academic Press. MR0818219

[16] Kallenberg, O. (2001). *Foundations of Modern Probability*. Springer, 2nd edition. MR1464694

[17] Kechris, A. S. (1995). *Classical Descriptive Set Theory*. Springer. MR1321597

[18] Kingman, J. F. C. (1975). Random discrete distributions. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, **37**, 1–22. MR0368264

[19] Lavine, M. (1992). Some aspects of Pólya tree distributions for statistical modelling. *Ann. Statist.*, **20**(3), 1222–1235. MR1186248

[20] Lijoi, A., Mena, R. H., and Prünster, I. (2005). Hierarchical mixture modeling with normalized inverse-Gaussian priors. *J. Amer. Statist. Assoc.*, **100**, 1278–1291. MR2236441

[21] MacEachern, S. N. (2000). Dependent Dirichlet processes. Technical report, Ohio State University.

[22] Mallory, D. J. and Sion, M. (1971). Limits of inverse systems of measures. *Ann. Inst. Fourier (Grenoble)*, **21**(1), 25–57. MR0284557

[23] Olshanski, G. (2003). An introduction to harmonic analysis on the infinite symmetric group. In *Asymptotic Combinatorics with Applications to Mathematical Physics*, volume 1815 of *Lecture Notes in Mathematics*, pages 127–160. Springer. MR2009838

[24] Pollard, D. (1984). *Convergence of Stochastic Processes*. MR0762984

[25] Sethuraman, J. (1994). A constructive definition of Dirichlet priors. *Statist. Sinica*, **4**, 639–650. MR1309433

[26] Talagrand, M. (2003). *Spin Glasses: A Challenge for Mathematicians*. Springer. MR1993891

[27] Walker, S. G., Damien, P., Laud, P. W., and Smith, A. F. M. (1999). Bayesian nonparametric inference for random distributions and related functions. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, **61**(3), 485–527. MR1707858

[28] Zhao, L. H. (2000). Bayesian aspects of some nonparametric problems. *Ann. Statist.*, **28**, 532–552. MR1790008