Annals of the Institute of Statistical Mathematics (2021) 73:1187–1202
https://doi.org/10.1007/s10463-020-00779-8

Improper versus finitely additive distributions as limits of countably additive probabilities

Pierre Druilhet[1] · **Erwan Saint Loubert Bié**[1]

Received: 27 August 2019 / Revised: 9 October 2020 / Accepted: 10 November 2020 /
Published online: 3 January 2021

Abstract
The Bayesian paradigm with proper priors can be extended either to improper distributions or to finitely additive probabilities (FAPs). Improper distributions and diffuse FAPs can be seen as limits of proper distribution sequences for specific convergence modes. In this paper, we compare these two kinds of limits. We show that improper distributions and FAPs represent two distinct features of the limit behavior of a sequence of proper distribution. More specifically, an improper distribution characterizes the behavior of the sequence inside the domain, whereas diffuse FAPs characterizes how the mass concentrates on the boundary of the domain. Therefore, a diffuse FAP cannot be seen as the counterpart of an improper distribution. As an illustration, we consider several approach to define uniform FAP distributions on natural numbers as an equivalent of improper flat prior. We also show that expected logarithmic convergence may depend on the chosen sequence of compact sets.

Keywords Bayesian statistics · Improper distribution · Finitely additive probability · Q-vague convergence · Uniform distribution · Expected logarithmic convergence. · Remote probability

1 Introduction

Improper priors and finitely additive probabilities (FAP) are the two main extensions of the standard Bayesian paradigm based on proper priors, i.e., countably additive probabilities (see Hartigan 1983, p. 15). Both extensions induce paradoxical phenomena such as strong inconsistency (Stone 1976; Dubins 1975) or marginalization paradoxes (Dawid et al. 1973) that do not occur with proper priors. To have a better understanding of these phenomena, some authors such as Stone (1982) or Kadane et al. (1986, p. 218), consider improper distributions and FAPs as limits of proper

✉ Pierre Druilhet
pierre.druilhet@uca.fr

[1] CNRS, LMBP, UMR 6620, Université Clermont Auvergne, Campus des Cézeaux, 3 Place Vasarely, TSA 60026, CS 60026, 63178 Aubiere Cedex, France

prior sequences w.r.t. to appropriate topologies. Heuristically, this approach seems to establish a link between improper distributions and FAP.

Seeing a FAP as a limit is a way to preserve the total mass equal to 1, while sacrificing the countable additivity. This point of view has been mainly supported by de Finetti (1972). On the other hand, improper distributions aim at preserving the countable additivity, while sacrificing a total mass equal to 1. Improper distributions appear naturally in the framework of conditional probability, see Rényi (1955) and more recently Taraldsen and Lindqvist (2010, 2016) and Lindqvist and Taraldsen (2018). Conditional probability spaces are also related to projective spaces of measures (Rényi 1970) which have a natural quotient space topology and a natural convergence mode, named q-vague convergence by Bioche and Druilhet (2016). Bayesian inference with improper posterior is justified by Taraldsen et al. (2019) from a theoretical point of view. Bord et al. (2018) consider the convergence of proper distribution sequences to an improper posterior for Bayesian estimation of abundance by removal sampling. Tufto et al. (2012) propose to adapt MCMC for the estimation of improper posteriors. In another approach, Akaike (1980) consider the convergence of posterior distributions w.r.t. an entropy criterion when the posterior distributions are proper.

In this paper, we mainly consider convergence of prior distributions to FAPs or to improper distributions, regardless to any statistical model. In Sect. 2, we define the notion of limits in the settings of improper distributions and of FAPs. We show that improper distributions and FAPs represent two distinct characteristics of a sequence of proper distributions. Therefore, they cannot be connected by the mean of proper distribution sequences. In Sect. 3, we revisit the notion of uniform distribution on integers in the light of our results. In Sect. 4, we illustrate with some examples the fundamental difference between convergence to an improper prior and to a FAP. In Sect. 5, we consider expected logarithmic convergence, defined by Berger et al. (2009) to approximate an improper distribution by a sequence of truncated proper priors. We apply some of the methods used in Sect. 2 to propose an example where this convergence mode depends on the chosen sequence of compact sets.

2 Convergence of probability sequences

We denote by $\mathcal{C}_b$ the set of continuous real-valued bounded functions on a space Θ and by $\mathcal{C}_K$ the set of continuous real-valued functions with compact support. For a σ-finite measure π, we denote $\pi(f) = \int f(\theta)\, \mathrm{d}\pi(\theta)$. Let $\{\pi_n\}_{n\in\mathbb{N}}$ be a sequence of proper distributions. The usual converge mode of $\{\pi_n\}_{n\in\mathbb{N}}$ to a proper prior π is the narrow convergence, also called weak convergence or convergence in law, defined by:

$$\pi_n \xrightarrow[n\to+\infty]{\text{narrowly}} \pi \iff \pi_n(f) \xrightarrow[n\to+\infty]{} \pi(f) \quad \forall f \in \mathcal{C}_b. \tag{1}$$

When it exists, the narrow limit of $\{\pi_n\}_n$ is necessarily unique. In this section, we consider two alternative convergence modes when there is no narrow limit, and especially when the total mass tends to concentrate around the boundary on the domain, more precisely when $\lim_n \pi_n(f) = 0$ for all f in $\mathcal{C}_K$. The idea is to consider a proper prior either as a special case of FAP or as a special case of a Radon measure, and for each case, to define a convergence mode in a formalized way.

In the following, Θ is a locally compact separable metric space. This is the case, for example, for usual topological finite-dimensional vector spaces or for denumerable sets with the discrete topology. In the latter case, any function is continuous and a compact set is a finite set.

2.1 Convergence to an improper distribution

To extend the notion of narrow limits, we consider here proper distributions within the set of projective space of positive Radon measures as follows: we denote by $\mathcal{R}$ the set of non-null Radon measures, that is regular countably additive measures with finite mass on each compact set. Note that, in the discrete case, any σ-finite measure is a Radon measure.

We define an improper distribution as an unbounded Radon measure which appears in parametric Bayesian statistics (see, e.g., Jeffreys 1970). The *projective space* $\overline{\mathcal{R}}$ associated with $\mathcal{R}$ is the quotient space for the equivalence relation $\sim$ defined by $\pi_1 \sim \pi_2$ iff $\pi_2 = \alpha\, \pi_1$ for some positive scalar factor α. To each Radon measure π is associated a unique equivalence class $\overline{\pi} = \{\pi' = \alpha\, \pi\, ; \alpha > 0\}$. Therefore, a projective space is a space where objects are defined up to a positive scalar factor. It is natural in Bayesian statistics to consider such projective space since two equivalent priors give the same posterior. The projective space $\overline{\mathcal{R}}$ is also naturally linked with conditional probability spaces (Rényi 1955). All the results presented below on the convergence mode w.r.t. to the projective space $\overline{\mathcal{R}}$ can be found in Bioche and Druilhet (2016). The usual topology on $\mathcal{R}$ is the vague topology defined by

$$\pi_n \xrightarrow[n\to+\infty]{\text{q - vaguely}} \pi \iff \pi_n(f) \xrightarrow[n\to+\infty]{\text{vaguely}} \pi(f) \quad \forall f \in \mathcal{C}_K. \tag{2}$$

From the related quotient topology, we can derive a convergence mode, called q-vague convergence: a sequence $\{\pi_n\}_n$ in $\mathcal{R}$ converges $q-vaguely$ to a (non-null) improper distribution π in $\mathcal{R}$ if $\overline{\pi}_n$ converges to $\overline{\pi}$ w.r.t. the quotient topology where $\overline{\pi}_n$ and $\overline{\pi}$ are the equivalence classes associated with π_n and π. The limit $\overline{\pi}$ is unique whereas π is unique only up to a positive scalar factor. It is not always tractable to check a convergence in the quotient space. However, there is an equivalent definition in the initial space $\mathcal{R}$: $\{\pi_n\}_n$ converges $q-vaguely$ to π if there exists some scalar factors α_n such that $\{\alpha_n\, \pi_n\}_n$ converges vaguely to π:

$$\pi_n \xrightarrow[n\to+\infty]{q-vaguely} \pi \iff a_n\pi_n \xrightarrow[n\to+\infty]{vaguely} \pi \quad \text{for some } a_1, a_2, \cdots > 0\,. \tag{3}$$

The q-vague convergence can be considered as an extension of the narrow convergence in the sense that if $\{\pi_n\}_n$ and π are proper distributions and $\{\pi_n\}_n$ converges narrowly to π then $\{\pi_n\}_n$ converges q-vaguely to π. Note that the converse part holds if and only if $\{\pi_n\}_n$ is tight (see Bioche and Druilhet 2016, Proposition 2.8).

When a sequence $\{\pi_n\}_n$ of proper distributions converges q-vaguely to an improper distribution, then $\lim_n \pi_n(K) = 0$ for any compact K (Bioche and Druilhet 2016, Proposition 2.11). The following lemma gives an apparently stronger, but in fact equivalent, result. It will be useful to establish our main result and to construct examples in Sects. 4.3 and 5.

Lemma 1 *Let $\{\pi_n\}_n$ be a sequence of proper distributions such that $\lim_n \pi_n(K) = 0$ for any compact K. Then there exists a non-decreasing sequence of compact sets K_n such that $\cup_n K_n = \Theta$ and $\lim_n \pi_n(K_n) = 0$. Moreover, K_n may be chosen such that, for any compact K, there exists an integer N such that $K \subset K_N$.*

Proof Let $\widetilde{K}_m$, $m \geq 1$, be an increasing sequence of compact sets with $\cup_m \widetilde{K}_m = \Theta$. For each m, $\lim_n \pi_n(\widetilde{K}_m) = 0$, so there exists an integer N_m such that $N_m > N_{m-1}$ and $\pi_n(\widetilde{K}_m) \leq 1/m$ for $n > N_m$. Consider now such a sequence of integers N_m, $m \geq 1$. For any n, there exists a unique integer m such that $N_m \leq n < N_{m+1}$. We define K_n by $K_n = \widetilde{K}_m$. So, $\pi_n(K_n) = \pi_n(\widetilde{K}_m) \leq 1/m$. Since m increases with n, $\lim_n \pi_n(K_n) = 0$. Furthermore, the sequence $\widetilde{K}_m$ can be chosen such that, for any compact K, K is a subset of all but finitely many $\widetilde{K}_m$, (see e.g. Bauer 2001, Lemma 29.8). By construction, the same property holds for the sequence K_n. $\square$

Note that, $\lim_n \pi_n(K) = 0$ for any compact set K does not imply that $\{\pi_n\}_n$ converge q-vaguely, as shown in Sect. 4 on some examples.

2.2 Convergence to a FAP

Here, we consider proper distributions as special cases of FAPs. Denote by $\mathcal{F}_b$ the set of bounded real-valued measurable functions on Θ. A FAP π is a linear functional on $\mathcal{F}_b$ which is positive, i.e., $\pi(f) \geq 0$ if $f \geq 0$, and which satisfies $\pi(1) = 1$. Therefore, the set of FAPs is included in the topological dual of $\mathcal{F}_b$ equipped with the sup-norm. For any measurable set $E \subset \Theta$, we define $\pi(E) = \pi(\mathbb{1}_E)$, where $\mathbb{1}_E(x) = 1$ if $x \in E$ and 0 otherwise. We also denote $\int f(\theta)\, \mathrm{d}\pi(\theta) = \pi(f)$.

For most authors (see, e.g., Heath and Sudderth 1978), a FAP is a linear functional on the set of bounded real-valued functions. Here, we do impose a measurability condition, since we require proper distributions to be special cases of FAPs. In the case where Θ is a denumerable set equipped with the usual discrete topology, any function or set is measurable, and so, both definitions of FAPs are equivalent.

Let $\{\pi_n\}_n$ be a sequence of FAPs. The usual convergence mode for FAPs is associated with the weak* topology: a sequence $\{\pi_n\}_n$ converges to π if $\lim_n \pi_n(f) = \pi(f)$ for any $f \in \mathcal{F}_b$.

When $\{\pi_n\}_n$ does not converge, we may consider limit points, as proposed by Stone (1982) for denumerable sets and extended here to more general sets. The existence of limit points relies on the Banach-Alaoglu-Bourbaki theorem (see, e.g., Rudin 1991), since a FAP belongs to the unit ball in the dual of $\mathcal{F}_b$, which is compact for the weak*-topology. Hence, for any sequence $\{\pi_n\}_n$ of FAPs, there exists at least one limit point π which is defined as a FAP limit. We recall that π is a limit point of $\{\pi_n\}_n$ for the weak*-topology if and only if for any integer p, any $f_1, \dots f_p$ in $\mathcal{F}_b$ and any $\varepsilon > 0$, there exists an infinite number of n such that $\left|\pi_n(f_i) - \pi(f_i)\right| \leq \varepsilon$, $i = 1, \dots, p$. Since $\mathcal{F}_b$ is not in general first-countable, there does not necessarily exist a subsequence $\{\pi_{n_k}\}_k$ that converges to π. We can only say that, for any $f_1, \dots, f_p$ in $\mathcal{F}_b$, there exists a subsequence $\{\pi_{n_k}\}_k$ such that $(\pi_{n_k}(f_1), \dots, \pi_{n_k}(f_p))$ converges to $(\pi(f_1), \dots, \pi(f_p))$.

When π_n and π are proper distributions, then $\{\pi_n\}_n$ converging narrowly to π does not imply that π is a FAP limit point of $\{\pi_n\}_n$. Therefore, unlike q-vague limits, FAP limit points cannot be considered as an extension of the narrow convergence.

For example, consider the proper distributions $\pi_n = \delta_{\sqrt{2}/n}$, where δ is the Dirac measure. The sequence $\{\pi_n\}_n$ converges narrowly to $\pi = \delta_0$ but π is not a FAP limit point of $\{\pi_n\}_n$. To show this, consider $f(\theta) = \mathbb{1}_{\mathbb{Q}}(\theta) \in \mathcal{F}_b$, with $\mathbb{Q}$ the set of rational numbers, we have $\lim_n \pi_n(f) = 0 \neq \pi(f) = 1$. We can only say that any FAP limit point of the sequence $\{\pi_n\}_n$ will coincide with $\pi = \delta_0$ on the set $\mathcal{C}_b$.

To consider a FAP limit as an extension of the narrow convergence, we should have defined FAPs on the space $\mathcal{C}_b$ rather than $\mathcal{F}_b$. However, with this choice, $\pi(E)$ is not well defined for all measurable sets E.

In the special case where Θ is a denumerable set, any real-valued function on Θ is continuous. So, if a sequence of proper distributions $\{\pi_n\}_n$ converges narrowly to a proper distribution π, then π is a FAP limit point.

Another way to extend the notion of limit can be obtained by using the Hahn-Banach theorem as follows (see Huisman 2016): let $\mathcal{S}_c$ be the set of $f \in \mathcal{F}_b$ such that $\lim_n \pi_n(f)$ exists. A FAP π is said to be an extended FAP limit of $\{\pi_n\}_n$ if $\lim_n \pi_n(f) = \pi(f)$ for any $f \in \mathcal{S}_c$, and if $\pi(f) \leq \limsup_n \pi_n(f)$. The existence of a FAP π satisfying this requirement is guaranteed by the Hahn-Banach theorem (see Rudin 1991): define the linear function $\boldsymbol{\Phi}$ on $\mathcal{S}_c$ by $\boldsymbol{\Phi}(f) = \lim_n \pi_n(f)$ and the sub-linear functional $p(f) = \limsup_n \pi_n(f)$. Then, there exists a linear functional π on $\mathcal{F}_b$ that coincides with $\boldsymbol{\Phi}$ on $\mathcal{S}_c$ and that satisfies $\pi(f) \leq p(f)$ on $\mathcal{F}_b$. The condition $\pi(f) \leq p(f)$ implies that π is a FAP. Conversely, an extended FAP limit necessarily satisfies $\pi(f) \leq p(f)$. Replacing f by $-f$ gives $\pi(f) \geq \liminf_n \pi_n(f)$. Therefore, an extended FAP limit can be characterized by the following lemma:

Lemma 2 *A FAP π is an extended FAP-limit of the sequence $\{\pi_n\}_n$ if and only if* for any $f \in \mathcal{F}_b$

$$\liminf_n \pi_n(f) \leq \pi(f) \leq \limsup_n \pi_n(f) \tag{4}$$

or equivalently if and only if for any measurable set E

$$\liminf_n \pi_n(E) \leq \pi(E) \leq \limsup_n \pi_n(E). \tag{5}$$

Note that the sequence $\{\pi_n\}_n$ converges to π for the weak* topology if and only if $\mathcal{S}_c = \mathcal{F}_b$. In general, an extended FAP limit is not unique and its existence relies on the axiom of choice.

The set of limit points of $\{\pi_n\}_n$ is included in the set of extended FAP limits. The converse inclusion is false in general. Inequalities (4) or (5) hold for limit points but are not sufficient to characterize them. It is easy to see that the closed convex hull of the set of limit points is included in the set of extended FAP limits. We conjecture that, conversely, the set of extended FAP limits defined by (4) is the closed convex hull of the set of limit points. As a simple example, consider the sequence $\{\pi_n\}_n$ with $\pi_{2n} = \delta_0$ and $\pi_{2n+1} = \delta_1$. There are only two limit points δ_0 and δ_1, whereas any $\pi = \alpha\delta_0 + (1-\alpha)\delta_1$, $0 \le \alpha \le 1$ is a extended FAP limit. In Sect. 4.1, we illustrate the difference between these two constructions of limits with another example.

Even if the notion of FAP limit points is more restrictive than the notion of extended FAP limit, the main results, especially Theorem 1, Corollary 1, Proposition 2, Lemmas 3 and 4 hold for both of them. In the following, we consider only FAP limit points.

2.3 FAP limit points versus q-vague convergence

The fact that a sequence of proper distributions has both improper and FAP limit points may suggest a connection between the two notions as proposed heuristically by several authors, such as Levi (1980), Stone (1982) and Kadane et al. (1986). The following results show that this is not the case. Roughly speaking, it is shown that any FAP which is a limit point of some proper distribution sequence can be connected to any improper prior by this mean.

Theorem 1 *Let $\{\pi_n\}_n$ be a sequence of proper distributions such that $\lim_n \pi_n(K) = 0$ for any compact set K. Then, for any improper distribution π, it can be constructed a sequence $\{\widetilde{\pi}_n\}_n$ which converges q-vaguely to π and which has the same set of FAP limit points as $\{\pi_n\}_n$.*

Proof For any FAP or any proper or improper distribution μ, we define the distribution $(\mathbb{1}_A\,\mu)$ by $(\mathbb{1}_A\,\mu)(f) = \mu(\mathbb{1}_A f)$ where A is any measurable set. From Lemma 1, it can be constructed an exhaustive increasing sequence K_n of compact sets such that $\lim_n \pi_n(K_n) = 0$. Put $\gamma_n = \pi_n(K_n)$ and define the sequence of proper distributions $\widetilde{\pi}_n = \gamma_n \frac{1}{\pi(K_n)} \mathbb{1}_{K_n}\pi + (1-\gamma_n)\frac{1}{\pi_n(K_n^c)}\mathbb{1}_{K_n^c}\,\pi_n$, with K^c the complement of K. By Lemmas 4 and 3 in "Appendix 1", $\widetilde{\pi}_n$ has the same FAP limit points as $\{\pi_n\}$. By Lemma 5, $\widetilde{\pi}_n$ converges q-vaguely to π. □

Corollary 1 *Let $\{\pi_n\}_n$ be a sequence of proper distributions that converges q-vaguely to an improper distribution $\pi^{(1)}$. Then, for any other improper distribution $\pi^{(2)}$, it can be constructed a sequence $\{\widetilde{\pi}_n\}_n$ that converges q-vaguely to $\pi^{(2)}$ and that has the same FAP limit points as $\{\pi_n\}_n$.*

We have shown that no direct link between an improper limit and a FAP limit point can be established. One can only say that if a sequence of proper distributions converges to an improper distribution, its FAP limit point π are *diffuse*, i.e. $\pi(K) = 0$ for any compact set K.

3 Uniform distribution on integers

In this section, we compare different notions of uniform distributions on the set $\mathbb{N}$ of integers, by using several considerations such as limit of proper uniform distributions.

We also illustrate the fact that FAP uniform distributions are not well-defined objects (de Finetti 1972, pp.122, 224). Contrary to uniform improper distributions, FAP limit points of uniform distributions on an exhaustive sequence of compact sets are highly dependent on the choice of that sequence.

3.1 Uniform improper distribution

There are several equivalent ways to define a uniform improper prior on integers. These definitions lead to a unique, up to a scalar factor, distribution. The uniform distribution can be defined directly as a flat distribution, i.e., $\pi(k) \propto 1$ for any integer k. It is the unique (up to a scalar factor) measure that is shift invariant, i.e., such that $\pi(k + A) = \pi(A)$ for any integer k and any set of integers A. The uniform distribution is also the q-vague limit of the sequence of uniform proper distributions on $K_n = \{0, 1, \dots, n\}$. More generally and equivalently, the uniform distribution is the q-vague limit of any sequence of proper uniform priors on an exhaustive increasing sequence $\{K_n\}_n$ of finite subsets of integers.

3.2 Uniform finitely additive probability

The notion of uniform finitely additive probabilities is more complex. Contrary to the improper case, there is no explicit definition since $\pi(k) = 0$ for any integer k. We present here several non equivalent approaches to define a uniform FAP. The first two ones can be found in Kadane and O'Hagan (1995) and Schirokauer and Kadane (2007).

3.2.1 Shift invariant (SI) uniform distribution

As for the improper case, a uniform FAP π can be defined as being any shift invariant FAP, i.e., a FAP satisfying $\pi(A) = \pi(A + k)$ for any subset of integers A and any integer k. Such a distribution will be called SI-uniform. In that case, one necessarily has: $\pi(k_1 + k_2 \times \mathbb{N}) = k_2^{-1}$, for any $(k_1, k_2) \in \mathbb{N} \times \mathbb{N}^*$. In Kadane and O'Hagan (1995), the authors investigate the properties of FAPs satisfying only $\pi(k_1 + k_2 \times \mathbb{N}) = k_2^{-1}$, where the sets $k_1 + k_2 \times \mathbb{N}$ are called *residue classes*.

3.2.2 Limiting relative frequency (LRF) uniform distributions

Kadane and O'Hagan (1995) consider a stronger condition to define uniformity. For a subset A, define its Limiting Relative Frequency LRF(A) by

$$\mathrm{LRF}(A) = \lim_{N\to\infty} \frac{\#\{k \leq N,\ \text{s.t. } k \in A\}}{N+1},$$

when this limit exists. A FAP π on $\mathbb{N}$ is said to be *LRF uniform* if $\pi(A) = p$ when $\mathrm{LRF}(A) = p$.

Let π_n be the uniform proper distribution on $K_n = \{0, 1, \dots, n\}$, then $\mathrm{LRF}(A) = \lim_{n\to\infty} \pi_n(A)$. Therefore, any FAP limit point of π_n is LRF uniform. In fact, a FAP π is LRF uniform if and only if it is an extended FAP limit of $\{\pi_n\}_n$.

It is worth noting that, unlike the q-vague limit,

FAP limit points are highly dependent on the choice of the increasing exhaustive sequence of finite sets K_n. Changing the sequence $\{K_n\}_n$ changes the notion of uniformity. For example, if $\widetilde{\pi}_n$ is the uniform distribution on $K_n = \{2k\,;\, 0 \leq k \leq n^2\} \cup \{2k+1;\, 0 \leq k \leq n\}$, then $\lim_n \widetilde{\pi}_n(2\mathbb{N}) = 1$, whereas $\lim_n \pi_n(2\mathbb{N}) = 1/2$.

3.2.3 Bernoulli Scheme (BS) uniform distribution

We propose here another notion of uniformity that is not dependent of the choice a particular increasing sequence of finite sets K_n as for the LRF uniformity. Consider a Bernoulli Scheme, that is, a sequence $\{X_k\}_{k\in\mathbb{N}}$ of i.i.d. Bernoulli distributed random variables with mean $p \in [0, 1]$. Define the random set $A(X) = \{k \in \mathbb{N},\ \text{s.t. } X_k = 1\}$. A FAP π is said to be *BS-uniform* if, for any $p \in [0;1]$, $\pi(A(X)) = p$, almost surely. By the strong law of large numbers, $\mathrm{LRF}(A(X)) = p$, almost surely.

Proposition 1 *Let $\{K_n\}$ be an increasing sequence of finite subsets of $\mathbb{N}$, with $\cup_{n\in\mathbb{N}} K_n$ being infinite. Then, any FAP which is a limit point of the sequence π_n of uniform distributions on K_n is BS-uniform.*

When $\cup_{n\in\mathbb{N}} K_n = \mathbb{N}$, this proposition shows that any FAP limit point of uniform distribution sequences is BS-uniform. In particular, a LRF uniform FAP is also BS uniform. However, if, for example, K_n is the set of even numbers less or equal to n, then any FAP limit point π of the sequence of uniform distributions on K_n is BS-uniform. However, π is not uniform on $\mathbb{N}$ but uniform on $2\mathbb{N}$. Therefore, BS uniformity looks much more like a necessary condition for a FAP to be uniform, than like a complete definition.

4 Comparison of convergence modes on examples

We consider here some examples that illustrate the difference between convergence of proper distributions to an improper distribution and to a FAP.

4.1 FAP limit points on $\mathbb{N}$

For a sequence $\{\pi_n\}_n$ of proper distributions on $\mathbb{N}$, it is known that there does not necessarily exist a q-vague limit, but if it exists, it is unique up to a scalar factor, i.e., it is unique in the projective space of Radon measures. At the opposite, we have seen that a FAP limit point always exists but is not necessarily unique.

We illustrate the non-uniqueness of FAP limit point with an extreme case. Consider the sequence of proper distributions $\pi_n = \delta_n$, where δ_n is the Dirac measure on n. This sequence has no q-vague limit since $\pi_n(k) = 0$ for $n > k$.

Let π be a FAP limit points. For any subset A, there exists a subsequence $\{\pi_{n_k}\}$ such that $\pi_{n_k}(A)$ convergences to $\pi(A)$. So, $\pi(A) \in \{0, 1\}$. Therefore π is any *remote* FAP, that is a diffuse FAP such that $\pi(A) \in \{0, 1\}$, as defined by Dubins (1975, p. 92). This also proves the existence of remote FAPs. Note that a remote FAP is neither BS uniform nor SI and therefore cannot be LRF uniform. As a remark, the extended FAP limits of π_n are all the diffuse FAPs.

4.2 Convergence of sequence of Poisson distributions

We consider here the sequence $\{\pi_n\}_n$ of Poisson distributions with mean n. Although $\lim_n \pi_n(K) = 0$ for any finite set K, this sequence of proper distributions does not converge q-vaguely to any improper distribution (Bioche and Druilhet 2016, § 5.2). As a remark, let $\widetilde{\pi}_n$ be the shifted measures defined on the set of positive and integers integers $\mathbb{Z}$ by $\widetilde{\pi}_n(B) = \pi_n(B + n)$, where π_n can be seen as a measure on the set $\mathbb{Z}$, with $\pi_n(k) = 0$ for $k < 0$. Then, using the approximation of the Poisson distribution by an normal distribution, it can be shown that the sequence $\widetilde{\pi}_n$ converges q-vaguely to the improper uniform measure on the set $\mathbb{Z}$.

We consider now the FAP limit point of the sequence $\{\pi_n\}_n$. The next result shows that these limits have some properties of uniformity described in Sect. 3 but not all of them. The proof is given in "Appendix 2".

Proposition 2 *Any FAP π which is a limit point of the sequence $\{\pi_n\}_n$ of Poisson distribution with mean n is SI-uniform and BS-uniform but not necessarily LRF-uniform.*

Therefore, the FAP limit points of the Poisson distribution sequence are examples of SI- and BS-uniform distributions that are not LRF uniform. Kadane and Jin (2014) give another example of SI but not LRF uniform FAPs using paths of random walks. Even if they consider FAPs on a subset of bounded functions, it can be extended to $\mathcal{F}_b$ by using the Hahn-Banach theorem similarly to Sect. 2.2.

4.3 FAP versus q-vague convergence of uniform proper distributions

To illustrate the fact that any FAP limit point can be related with any improper distribution, consider again the sequence $\{\pi_n\}_n$ of Poisson distributions with mean n and let π_0 be any improper distribution on the integers. Since $\lim_n \pi_n(K) = 0$ for any finite set, from Lemma 1, we can construct an exhaustive sequence of finite set K_n such that $\lim_n \pi_n(K_n) = 0$. Put $K_n = \{k \in \mathbb{N},\ k \leq n/2\}$, which satisfies this condition. Define the sequence of proper distributions $\widetilde{\pi}_n$ by:

$$\widetilde{\pi}_n(A) = \pi_n(K_n)\frac{\pi_0(A \cap K_n)}{\pi_0(K_n)} + (1 - \pi_n(K_n))\frac{\pi_n(A \cap K_n^c)}{\pi_n(K_n^c)} \tag{6}$$

for any set A. From Theorem 1, $\{\widetilde{\pi}_n\}_n$ converges q-vaguely to π_0 and has the same FAP limit points as $\{\pi_n\}_n$.

As another example, let $\{\pi_n\}_n$ be the sequence of uniform distributions on $\{0, 1, \dots, n\}$ and choose $K_n = \{k \in \mathbb{N},\ k \leq \sqrt{n}\}$. Then, $\lim_n \pi_n(K_n) = 0$. Therefore, for any improper distribution π_0 on the set of integers, the sequence constructed as in (6) has the same FAP limit points as those of the sequence of uniform distributions $\{\pi_n\}_n$ and converges q-vaguely to π_0. This shows again the difficulty to connect the uniform improper distribution and uniform FAPs by limits of proper distributions.

4.4 Convergence of beta distributions

In this section, we consider the limit of the sequence of Beta distribution $\pi_n = \text{Beta}(a_n, b_n)$ defined on $\Theta =]0, 1[$ when a_n and b_n go to 0. We will see that contrary to the improper limit, the FAP limit points depend on the way a_n and b_n go to 0. This illustrates again the difference between the two kinds of limits.

The density of a beta distribution Beta(a, b) is given by

$$\pi_{a,b}(x) = \frac{1}{\beta(a, b)}\, x^{a-1}(1 - x)^{b-1} \text{ for } x \in]0;1[$$

where $\beta(a, b)$ is the beta function.

From Bioche and Druilhet (2016), the unique (up to a scalar factor) q-vague limit of Beta(a_n, b_n) when a_n and b_n go to 0 is the Haldane improper distribution:

$$\pi_H(x) = \frac{1}{x(1 - x)} \text{ for } x \in]0;1[.$$

The q-vague limit gives no information on the relative concentration of the mass around 0 and 1: for $0 < u < v < 1$, $\pi_H(]0, u]) = \pi_H([v, 1[) = +\infty$. To explore this concentration, we temporarily replace the space Θ by $\overline{\Theta} = [0, 1]$. This has no consequence on the Beta distributions but the Haldane distribution is no longer a q-vague limit of the sequence. Put $c_n = a_n/b_n$ and assume that $\{c_n\}_n$ converges to some

$c \in [0,1]$. The sequence $\{\pi_n\}_n$ converges narrowly, and hence q-vaguely, to the proper distribution $\widetilde{\pi} = \frac{1}{1+c}\delta_0 + \frac{c}{1+c}\delta_1$. Contrary to the Haldane prior, $\widetilde{\pi}$ shows how the mass concentrates on the boundary of the domain, but gives no information on the behavior of the sequence inside the domain. Note that the Haldane distribution is not a Radon measure on $\overline{\Theta}$ since $\pi_H([0,1]) = +\infty$ where $[0, 1]$ is a compact set. Therefore π_H cannot be a candidate for the q-vague limit on $\overline{\Theta}$.

We now consider the FAP limit points on $\Theta =]0,1[$ of π_n, and we show that they give an information similar to that given by $\widetilde{\pi}$ on the way the mass concentrate on the boundary of the domain. Again, we assume that $c_n = a_n/b_n$ converges to some $c \in [0,1]$. Easy calculations show that for any $0 < \varepsilon < 1$ $\lim_n \pi_n(]0,\varepsilon[) = \frac{1}{1+c}$ and $\lim_n \pi_n(]1-\varepsilon,1[) = \frac{c}{1+c}$. Therefore, for any FAP limit point π and for any $\varepsilon \in]0,1[$, we have $\pi(]0,\varepsilon[) = \frac{1}{1+c}$ and $\pi(]1-\varepsilon,1[) = \frac{c}{1+c}$, with $\pi([u,v]) = 0$ for $0 < u < v < 1$.

5 Expected logarithmic convergence

In Bayesian statistics, consider a statistical model $p(x|\theta)$ and an improper prior $\pi(\theta)$ on Θ. Define the truncated proper prior $\pi_n(\theta) \propto \pi(\theta)\ \mathbb{1}_{\theta \in K_n}$, for some exhaustive increasing sequence of compact sets $\{K_n\}_n$. From Berger et al. (2009), a sequence of posteriors distributions $\pi_n(\theta|x)$ is said to be expected logarithmically convergent to $\pi(\theta|x)$ if

$$\lim_{i\to\infty} \int_{\mathcal{X}} p_i(x)\kappa(\pi(\cdot|x), \pi_i(\cdot|x))\mathrm{d}x = 0\,,$$

where $p_i(x) = \int_\Theta p(x|\theta)\pi_i(\theta)\mathrm{d}\theta$, and $\kappa(m_1, m_2)$ denotes the Kullback-Leibler distance between probability measures m_1 and m_2. The prior π is said to be permissible w.r.t. $p(x|\theta)$ if $\pi(\theta|x)$ is proper and if there exists some exhaustive sequence of compact sets $\{K_n\}_n$ such that $\pi_n(\theta|x)$ is expected logarithmically convergent to $\pi(\theta|x)$. Note that $\pi_n(\theta|x)$ converges q-vaguely to $\pi(\theta|x)$, provided that $p(x|\theta)$ is continuous w.r.t. θ for any x (Bioche and Druilhet 2016, Proposition 3.1).

An open problem is to know whether this property is always independent from the choice of the sequence of compact sets $\{K_n\}_n$. We present here a situation where it is not.

The construction of this counter-example relies on the fact that the tail behavior of a sequence of distributions is not directly related to its q-vague convergence as explained in Sect. 2.3.

Consider the following model: for any integers x and θ (included negative integers) define:

$$p(x|\theta) = \begin{cases} \frac{1}{3} & \text{if } \theta \geq 1,\ x \in \{\left[\frac{\theta}{2}\right],\ 2\theta,\ 2\theta+1\} \\ 1 & \text{if } \theta \leq 0, x = \theta \\ 0 & \text{otherwise} \end{cases}$$

where $[l]$ the integer part of l, with the particular case $[1/2]=1$. For $\theta \leq 0$, $x \leq 0$, we have a deterministic model. Remark also that for $x, \theta \geq 1$, we have the equivalence:

$$\left(x \in \left\{\left[\frac{\theta}{2}\right], 2\theta, 2\theta+1\right\}\right) \Leftrightarrow \left(\theta \in \left\{\left[\frac{x}{2}\right], 2x, 2x+1\right\}\right).$$

Consider the flat prior $\pi(\theta) \propto 1$. If we consider only $x, \theta \geq 1$, this model corresponds to a model proposed by Fraser et al. (1985) and used by Berger et al. (2009) to illustrate their approach.

Let $K_n = \{-a_n \leq \theta \leq b_n\}$ be an exhaustive sequence of compact sets, with $a_n, b_n \to +\infty$. Denote $I_n = \int_{\mathcal{X}} p_n(x)\kappa(\pi(\cdot|x), \pi_n(\cdot|x))\mathrm{d}x$. We have:

$$I_n = \frac{1}{a_n+b_n+1}\left(\ln(3)\frac{\left(2b_n+1-\left[\frac{b_n}{2}\right]\right)}{3} + \frac{\ln(3/2)}{3}\omega(b_n)\right)$$

where $\omega(b_n)=1$ if b_n is even, and 0 if b_n is odd. Therefore, as a_n and b_n tend to infinity, $I_n \sim \frac{\ln(3)}{2}\frac{b_n}{a_n+b_n+1}$.

When b_n/a_n tends to 0, $\lim_{n\to\infty} I_n = 0$, which gives an expected logarithmically convergent sequence of posteriors. However, taking $a_n = b_n$ leads to $\lim_{n\to\infty} I_n = \frac{\ln(3)}{4}$, and the sequence of corresponding posteriors is not expected logarithmically convergent.

This example shows that at least for some statistical models $p(x|\theta)$ and improper prior $\pi(\theta)$, the notion of expected logarithmic convergence may depend on the choice of the sequence of compact sets. It could be interesting to characterizes situations where it does not. This is left for future works.

6 Conclusion and perspectives

In this paper, we have shown that the characteristics of a sequence of proper distributions given by its FAP or improper limits are quite different. As a consequence, there is no clear link between improper distributions and FAPs: a diffuse FAP cannot be considered as the counterpart of some improper distribution.

In Bayesian statistics, improper distributions are commonly used in practice, even if some paradoxes may occur. They are easy to interpret either through their densities or as conditional probabilities (Taraldsen and Lindqvist 2016). At the opposite, diffuse FAPs are never used in practice, mainly because their constructions are always implicit and because diffuse FAPs give information only on the boundary of the parameter space, which is difficult to construe.

However, FAPs may provide a better understanding of some limit behavior that are not captured by improper distributions. The fact that our main results rely on explicit constructions of proper prior sequences may be useful to provide counterexamples. For example, in Sect. 5, we have shown that the notion of expected logarithmic convergence may depend on the sequence of compact sets. We hope to use our results in future works to have a better understanding of some paradoxical phenomena in Bayesian statistics, such as strong inconsistency or the marginalization paradox.

Appendix 1

We establish some lemmas useful to prove Theorem 1. The first one is straightforward.

Lemma 3 *Let* $\{\pi_n^{(1)}\}_n$ *and* $\{\pi_n^{(2)}\}_n$ *be two sequences of proper distributions and* $0 \leq \gamma_n \leq 1$ *be a sequence of scalars that converges to 0. Then, the sequence defined by* $\widetilde{\pi}_n = \gamma_n \pi_n^{(1)} + (1-\gamma_n)\pi_n^{(2)}$ *has the same FAP limit points as* $\{\pi_n^{(2)}\}_n$.

Proof For any $f_1, \dots, f_p \in \mathcal{F}_b$, then $(\pi_{n_k}^{(2)}(f_1), \dots, \pi_{n_k}^{(2)}(f_p))$ converges to $(\pi(f_1), \dots, \pi(f_p))$ iff $(\widetilde{\pi}_{n_k}(f_1), \dots, \widetilde{\pi}_{n_k}(f_p))$ converges to $(\pi(f_1), \dots, \pi(f_p))$. The result follows. □

Lemma 4 *Let* $\{\pi_n\}_n$ *be a sequence of proper priors and* K_n *be a non-decreasing sequence of compact sets such that* $\lim_n \pi_n(K_n) = 0$*, then the sequence defined by* $\widetilde{\pi}_n = \frac{1}{\pi_n(K_n^c)} \mathbb{1}_{K_n^c} \pi_n$ *has the same FAP limit points as* $\{\pi_n\}_n$.

Proof First, note that $\{\pi_n\}_n$ is not defined when $\pi_n(K_n) = 1$, but this cannot occur more than a finite number of times. For any $f \in \mathcal{F}_b$, $\pi_n(f) = \mathbb{1}_{K_n}\pi_n(f) + \mathbb{1}_{K_n^c}\pi_n(f) = \pi_n(\mathbb{1}_{K_n} f) + \pi_n(K_n^c)\widetilde{\pi}_n(f)$. Since f is bounded, $\lim_n \pi_n(\mathbb{1}_{K_n} f) = 0$. Moreover, $\lim_n (K_n^c) = 1$. Thus, for any $f_1, \dots, f_p \in \mathcal{F}_b$, $(\pi_{n_k}(f_1), \dots, \pi_{n_k}(f_p))$ converges to $(\pi(f_1), \dots, \pi(f_p))$ iff $(\widetilde{\pi}_{n_k}(f_1), \dots, \widetilde{\pi}_{n_k}(f_p))$ converges to $(\pi(f_1), \dots, \pi(f_p))$. □

At the opposite of Lemma 4, the following lemma shows that if we consider the restriction of a sequence $\{\pi_n\}_n$ of a proper or improper distribution on a exhaustive increasing sequence $\{K_n\}_n$ of compact sets, we preserve the q-vague limits.

Lemma 5 *Let* K_n *be a non-decreasing sequence of compact sets such that* $\cup_n K_n = \Theta$ *and such that, for any compact K, there exists N such that* $K \subset K_N$*. A sequence* $\{\pi_n\}_n$ *of Radon measures converges q-vaguely to the Radon measure* π *if and only if* $\widetilde{\pi}_n = \frac{1}{\pi_n(K_n)} \mathbb{1}_{K_n} \pi_n$ *converges q-vaguely to* π.

Proof Assume that π_n converges q-vaguely to π, then there exists some positive scalars $\{a_n\}_n$ such that for any f in $\mathcal{C}_K$, $\lim_n a_n \pi_n(f) = \pi(f)$. Put $\widetilde{a}_n = a_n\, \pi_n(K_n)$ and denote by K_f a compact set that includes the support of f. Then, there exists an integer N such that $K_f \subset K_n$ for $n > N$. Therefore, for $n > N$, $\widetilde{a}_n \widetilde{\pi}_n(f) = a_n \pi_n(f)$. The result and its reciprocal follow. □

Appendix 2

We prove here Proposition 2 of Sect. 4.2.

In order to show that π is SI-uniform, we consider π_n as a distribution on the set of positive and negative integers, extending it by 0 on the non-positive integers. Define by $\pi_n^{(k)}$ the shifted distribution: $\pi_n^{(k)}(A) = \pi_n(A+k)$, for any subset A of the set of integers. One knows that $\|\pi_n^{(k)} - \pi_n\|_{TV} \leq \frac{k}{\sqrt{2\pi n}}$, where $\|\cdot\|_{TV}$ is the total variation norm. Therefore, for any subset of $\mathbb{N}$, $\lim_{n\to\infty} |\pi_n(A+k) - \pi_n(A)| = 0$. Letting n go to infinity, we deduce that, for any FAP limit point π of π_n, and any integer k: $\pi(A+k) = \pi(A)$.

The fact that π is BS-uniform comes from an easy adaptation of the Hoeffding inequality in that context. Let $(X_k)_{k\in\mathbb{N}}$ be a Bernoulli scheme, of parameter p, and denote by $\mathbb{P}$ the associated probability. Hoeffding inequality gives, that, for any n:

$$\mathbb{P}\left\{ \left| \sum_{k=0}^{\infty} e^{-n}\frac{k^n}{n!}(X_k(\omega) - p) \right| \geq t \right\} \leq 2e^{-2c\sqrt{2\pi n}\,t^2},$$

for some positive constant c. The expected conclusion is then obtained thanks to the Borel-Cantelli lemma.

The fact that some of the limit points π of $\{\pi_n\}_n$ are not LRF uniform is a direct consequence of the following lemma.

Lemma 6 *For any $0 \leq p, p' \leq 1$, there exists a set A and some FAP limit points π of $\{\pi_n\}_n$ such that $LRF(A) = p$ and $\pi(A) = p'$.*

Proof First note that, for any set A', $LRF(A') = p$ if, and only if, $\sharp\{k \leq n,\ k \in A'\} = pn + o(n)$. Therefore, for any set A with $LRF(A) = p$ and for any set B such that $\sharp\{k \leq n,\ k \in B\} = o(N)$, one has both $LRF(A \cup B) = p$ and $LRF(A \backslash B) = p$. Take now for set B the following:

$$B = \bigcup_{k\in\mathbb{N}} \left\{ u \in \mathbb{N}\colon\ 4^k - 2^k k \leq u \leq 4^k + 2^k k \right\}.$$

For that B, one has:

$$\limsup_{n\to\infty} \frac{\sharp\{k \leq n,\ k \in B\}}{n+1} = \lim_{k\to\infty} \frac{\sum_{i=0}^{k} 2^{i+1} i}{4^k + 2^k k} \leq \lim_{k\to\infty} \frac{(k+1)2^{k+2}}{4^k} = 0,$$

and thus $LRF(B) = 0$. However, $\pi_{4^k}(B)$ converges to 1. Indeed, if U_k is some random variable with law π_{4^k}, one has:

$$\pi_{4^k}\left(\left\{ u \in \mathbb{N}\colon\ 4^k - 2^k k \leq u \leq 4^k + 2^k k \right\} \right) = \mathbb{P}\left(\frac{U_k - 4^k}{\sqrt{4^k}} \in \left[-k; k \right] \right).$$

The right-hand side term above converges to 1 thanks to the central limit theorem. Hence $LRF(A \cup B) = LRF(A \backslash B) = p$ while $\pi_{4^k}(A \cup B)$ converges to 1, and $\pi_{4^k}(A \backslash B)$

converges to 0. Now, for any $p' \in [0;1]$, choose two numbers $a < b$, so that $p' = \int_a^b \frac{e^{-u^2/2}}{\sqrt{2\pi}} \mathrm{d}u$. Take the set B' to be:

$$B' = \bigcup_{k \in \mathbb{N}} \left\{ u \in \mathbb{N}\colon\ 4^k + 2^k \max(-k, a) \le u \le 4^k + 2^k \min(k, b) \right\},$$

then $LRF(B') = 0$ again and $\pi_{4^k}(B')$ converges to p' , still thanks to the central limit theorem. Let $A = (A' \backslash B) \cup B'$. Then $LRF(A) = p$ and $\lim_{k\to\infty} \pi_{4^k}(A) = p'$. Now, any FAP limit point π of subsequence $\{\pi_{4^k}\}_k$ is also a FAP limit point of $\{\pi_k\}_k$. Hence, π is SI-uniform and BS-uniform, but one has $\pi(A) = p'$ and $LRF(A) = p$. □

References

Akaike, H. (1980). The interpretation of improper prior distributions as limits of data-dependent proper prior distributions. *Journal of the Royal Statistical Society, Series B*, *42*(1), 46–52.

Bauer, H. (2001). *Measure and integration theory. De Gruyter Studies in Mathematics*, Vol. 26. Berlin: Walter de Gruyter & Co. **(Translated from the German by Robert B. Burckel)**.

Berger, J. O., Bernardo, J. M., Sun, D. (2009). The formal definition of reference priors. *Annals of Statistics*, *37*(2), 905–938.

Bioche, C., Druilhet, P. (2016). Approximation of improper priors. *Bernoulli*, *22*(3), 1709–1728.

Bord, S., Bioche, C., Druilhet, P. (2018). A cautionary note on Bayesian estimation of population size by removal sampling with diffuse priors. *Biometrical Journal*, *60*(3), 450–462.

Dawid, A. P., Stone, M., Zidek, J. V. (1973). Marginalization paradoxes in Bayesian and structural inference. *Journal of the Royal Statistical Society, Series B*, *35*, 189–233.

de Finetti, B. (1972). *Probability, induction and statistics. The art of guessing. Wiley Series in Probability and Mathematical Statistics*. London: Wiley.

Dubins, L. E. (1975). Finitely additive conditional probabilities, conglomerability and disintegrations. *Annals of Probability*, *3*, 89–99.

Fraser, D. A. S., Monette, G., Ng, K. W. (1985). Marginalization, likelihood, and structural models. In P. R. Krishnaiah (Ed.), *Multivariate Analysis*, pp. 209–217. Amsterdam: North Holland. **(MR0822296)**.

Hartigan, J. A. (1983). *Bayes theory*. New York: Springer-Verlag.

Heath, D., Sudderth, W. (1978). On finitely additive priors, coherence, and extended admissibility. *Annals of Statistics*, *6*(2), 333–345.

Huisman, L. (2016). Infinitesimal distributions, improper priors and Bayesian inference. *Sankhya A*, *78*(2), 324–346.

Jeffreys, H. (1970). *Theory of probability. Oxford Science Publications*3rd ed. New York: The Clarendon Press, Oxford University Press.

Kadane, J. B., Jin, J. (2014). Uniform distributions on the integers: A connection to the Bernouilli random walk. *Econometric Reviews*, *33*(1–4), 372–378.

Kadane, J. B., O'Hagan, A. (1995). Using finitely additive probability: Uniform distributions on the natural numbers. *Journal of the American Statistical Association*, *90*(430), 626–631.

Kadane, J. B., Schervish, M. J., Seidenfeld, T. (1986). Statistical implications of finitely additive probability. *Bayesian inference and decision techniques. Studies Bayesian Econometrics and Statistics*, Vol. 6, pp. 59–76. Amsterdam: North-Holland.

Levi, I. (1980). *The enterprise of knowledge, No. 1–4*, Vol. 33, pp. 372–378. Cambridge, MA: The MIT Press.

Lindqvist, B. H., Taraldsen, G. (2018). On the proper treatment of improper distributions. *Journal of Statistical Planning and Inference*, *195*, 93–104.

Rényi, A. (1955). On a new axiomatic theory of probability. *Acta Mathematica Adacamiae Scientiarum Hungaricae*, *6*, 285–335.

Rényi, A. (1970). *Foundations of probability*. San Francisco, CA: Holden-Day Inc.

Rudin, W. (1991). *Functional analysis. International Series in Pure and Applied Mathematics*, 2nd ed. New York: McGraw-Hill Inc.

Schirokauer, O., Kadane, J. B. (2007). Uniform distributions on the natural numbers. *Journal of Theoretical Probability*, *20*(3), 429–441.

Stone, M. (1976). Strong inconsistency from uniform priors. *Journal of the American Statistical Association*, *71*(353), 114–125.

Stone, M. (1982). Review and analysis of some inconsistencies related to improper priors and finite additivity. *Logic, methodology and philosophy of science, VI (Hannover, 1979). Studies in Logic and the Foundations of Mathematics*, Vol. 10, pp. 413–426. Amsterdam: North-Holland.

Taraldsen, G., Lindqvist, B. H. (2010). Improper priors are not improper. *American Statistician*, *64*(2), 154–158.

Taraldsen, G., Lindqvist, B. H. (2016). Conditional probability and improper priors. *Communication in Statistics, Theory and Methods*, *45*(17), 5007–5016.

Taraldsen, G., Turfo, J., Lindqvist, B. H. (2019). Statistics with improper posteriors. Preprint, arXiv:1812.01314.

Tufto, J., Lande, R., Ringsby, T.-H., Engen, S., Saether, B.-E., Walla, T. R., et al. (2012). Estimating Brownian motion dispersal rate, longevity and population density from spatially explicit mark-recapture data on tropical butterflies. *Journal of Animal Ecology*, *4*(81), 756–769.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.