Predicting extreme events is important in many applications in risk analysis. Extreme-value theory suggests modelling extremes by max-stable distributions. The Bayesian approach provides a natural framework for statistical prediction. Although various Bayesian inferential procedures have been proposed in the literature of univariate extremes and some for multivariate extremes, the study of their asymptotic properties has been left largely untouched. In this paper we focus on a semiparametric Bayesian method for estimating max-stable distributions in arbitrary dimension. We establish consistency of the pertaining posterior distributions for fairly general, well-specified max-stable models, whose margins can be short-, light- or heavy-tailed. We then extend our consistency results to the case where data are samples of block maxima whose distribution is only approximately a max-stable one, which represents the most realistic inferential setting.

1. Introduction. Predicting the extremes of multiple variables is important in many applied fields for risk management. For instance, when designing bridges in civil engineering it is crucial to quantify what forces they must sustain in the future, e.g. the maximum wind speed, maximum river level, etc. [e.g. 12, Ch. 9.3]. In finance, the solvability of an investment is influenced by extreme fluctuations in the financial market affecting multiple assets, such as share prices, market indexes, currency values, etc. [e.g., 51]. Extreme-value theory encompasses several approaches for modelling multivariate extremes [e.g. 31]. In this paper we focus on the family of max-stable models, which arises as a class of asymptotic distributions for linearly normalised componentwise maxima of random vectors [31, Ch. 4]. Max-stable models have been successfully applied in several areas, e.g. in meteorological, environmental, medical and actuarial studies for analysing heavy rainfall, extreme temperatures, air pollution, clinical trials, insurance claims, etc. [e.g. 15, 25], in addition to those previously mentioned. In recent years, the popularity of some max-stable models is due to max-stable processes, which have been widely used in spatial applications [e.g. 3, 11, 20].

1.1. Bayesian approach for extremes. The Bayesian approach provides a natural framework for statistical prediction. The study of asymptotic properties like the consistency of the posterior distribution of the parameter of interest is informative for the robustness of the underlying Bayesian procedure [39]. In the last two decades the asymptotic theory of infinite-dimensional Bayesian statistics has been a very active research area. Applications to several interesting non- and semiparametric statistical problems have been proposed; among the most recent works, see [39, 47, 56, 66] and the references therein. To the best of our knowledge, to date there is no such a rigorous study concerning problems in extreme value analysis, both
in the univariate and multivariate context. This article fills such a gap, providing the first results on posterior consistency of non- and semiparametric Bayesian inference for multivariate max-stable models. The Bayesian literature for univariate extremes includes several methodological and applied contributions [e.g., see 5, 16, 25, 72], while only few works address the analysis of multivariate extremes [e.g., see 29, 70, 76]. There are two main reasons for the slow progress in the multivariate context, summarised below.

1.2. Max-stable models challenges. The first motivation is that multivariate max-stable distributions define a complex, semiparametric model class. A multivariate max-stable distribution is of the form $G_{\vartheta}(x|\theta) = C_{EV}(G_{\vartheta_1}(x_1), \ldots, G_{\vartheta_d}(x_d)|\theta)$, $x \in \mathbb{R}^d$, where $C_{EV}(\cdot|\theta)$ is the so-called extreme value copula, $\vartheta = (\vartheta_1, \ldots, \vartheta_d)$ is a finite-dimensional vector of marginal parameters and $\theta$ is an infinite-dimensional dependence parameter, both living in suitable parameter spaces (see Section 2.2 for details). In a commonly adopted representation of extreme value copulas, the dependence parameter $\theta$ is a probability measure subject to specific mean constraints, known as angular probability measure. A special transform (reparametrisation) of such a probability measure yields the well-known Pickands dependence function, complying in turn with specific shape restrictions. The copula $C_{EV}(\cdot|\theta)$ is also commonly parametrised by $\theta$ as the Pickands dependence function, since the latter is simple to interpret [e.g., 6]. Proper estimation of such infinite dimensional objects, accounting for their specificities, is not a simple task. Quite sophisticated non- and semiparametric estimation methods based on polynomials and splines have been proposed for inferring the dependence structure under both parametrisations [e.g., 17, 43, 49, 53, 70]. In particular, [54] propose a fully nonparametric Bayesian estimation method for bivariate max-stable distributions, where both dependence parametrisations are simultaneously dealt with by means of Bernstein polynomial representations.

The second motivation is that the expression of the density function $g_{\vartheta}(x|\theta)$ of $G_{\vartheta}(x|\theta)$ is complicated (see formulas 3 and 4) and therefore the corresponding likelihood function is computationally burdensome to calculate in practice [e.g., 28]. Accordingly, in high dimensions the statistical inference is often performed using a composite-likelihood approach [see 59] and the development of efficient full-likelihood estimation methods still represents an active research area [e.g., 28, 78, 44]. Notably, [29] have been able to derive a Bayesian inferential method based on the full-likelihood for fitting max-stable distributions to data of arbitrary dimensions (greater than two). They also establish asymptotic normality of the posterior distribution, yet under the rather restrictive assumptions that the extreme-value copula belongs to a known parametric model and the margins are all unit-Fr閏het.

1.3. Goals. The present paper provides several contributions. The first main statistical question that we address in this paper is how to construct a Bayesian inferential procedure that allows to consistently estimate the parameter set $(\vartheta, \theta)$ of a max-stable distribution $G_{\vartheta}(x|\theta)$ and its density function $g_{\vartheta}(x|\theta)$. These two missions are closely linked within a Bayesian approach, both from a mathematical and a computational viewpoint. We establish the almost sure posterior consistency of non- and semiparametric Bayesian inferential procedures for max-stable distributions in arbitrary dimensions $d \geq 2$, where the prior on the dependence structure is specified through a Bernstein polynomial representation of the angular probability measure. The extreme value copula is model free in the proposed framework, which is thus more widely applicable if compared to the model-specific approach in [29]. In the bivariate case, we also show that our almost sure consistency results can be extended to priors specified on the Pickands dependence function, i.e. to the case where the latter is considered as the parameter $\theta$ of the extreme value copula.

Our asymptotic results are initially derived assuming that the observable dataset is sampled from a max-stable distribution with known unit-Fr閏het margins, as in [29]. In this case, the
marginal parameters $\vartheta$ are known and it suffices to specify a prior distribution for $\theta$. In practice, max-stable distributions are typically used for modelling the so-called block maxima, i.e. a vector of maxima obtained componentwise on a series of multidimensional observations of a certain length (block), e.g. yearly maxima. In this case, the use of max-stable models is only asymptotically justified, under regularity conditions, for increasingly large block sizes. Each univariate sequence of maxima (suitably normalised) must approximatively follow one of the following three types of distributions: (reverse) Weibull (short-tailed), Gumbel (light-tailed) or Fréchet (heavy-tailed). Accordingly, as a first step we extend our posterior consistency results to well-specified max-stable statistical models whose margins are all short- or light- or heavy-tailed distributions. In this case, both the marginal and dependence parameters $\vartheta$ and $\theta$ are unknown and a prior has to be specified for both of them. Typically, the assumption of marginal distributions with such a homogenous tail behaviour entails no significant practical restriction. For example, in several environmental applications physical phenomena are well described by short- or light-tailed distributions, due to natural constraints [e.g. 12, 15]. Moreover, heavy-tailed distributions are found to represent quite well the tail structure of many actuarial and financial data examples [e.g. 5, 15, 61]. The unified framework for the three tail types, embedded in the generalised extreme-value family [e.g. 23, Ch. 1], is not discussed in the present article for non-trivial technical issues, as the technical drawback of dealing with a multivariate distribution support that depends on 3d marginal distributions parameters (among others).

In a second step, we study the case where data are in the form of block maxima, which only approximately follow a max-stable distribution, and a max-stable statistical model is thus misspecified. We therefore address a second main statistical question: how to adapt the paradigm of our bonafide Bayesian procedure when data are not exactly sampled from a distribution in the considered model class, in order to obtain consistent inference on the true block maxima distribution. We then provide conditions under which a quasi-Bayesian procedure is mathematically justified, guaranteeing the consistent estimation of the true data generating density. In particular, we consider a Bayesian approach where data-dependent priors are specified in an empirical Bayes fashion [e.g., 62]. Their use turns out to be essential to adapt classical asymptotic arguments on posterior consistency to the present nonstandard framework. Our asymptotic results are derived by leveraging on the recent theory of remote contiguity [47, 33], which draws a link between the limiting statistical model and the actual joint probability law of the data sample. The technical tools developed in this work can be also of independent interest, beyond the extreme values context, and adapted to other statistical methods affected by a model convergence bias.

In extreme value analysis a crucial objective, if not the most important one, is to predict future events that are possibly more extreme than those already observed. Density estimation within the framework of max-stable models is intimately connected with this goal. As a final step, we translate our consistency theorems into consistent probabilistic forecasting results for future high-dimensional observations. In particular, we show how to capitalise on them to deliver asymptotically accurate predictive regions.

1.4. Organization. The remainder of this paper is organised as follows. We start providing the necessary background (Section 2) through: the introduction of notation used throughout the paper (Section 2.1), a brief review of the theory on max-stable distributions (Sections 2.2, 2.3), a concise description of how their dependence structure can be represented via Bernstein polynomials (Section 2.4). The basic asymptotic theory is developed in Section 3. After a short introduction on the Bayesian paradigm for well-specified max-stable models (Section 3.1) consistency results are firstly established for the class of so-called simple max-stable distributions (Section 3.2). These are then extended to more general families of
max-stable distributions (Section 3.3). We refine our asymptotic theory to account for the more realistic sampling scheme, where the data are samples of block maxima which only approximately follow a max-stable distribution (Section 4). Implications of our consistency results for statistical prediction are illustrated in Section 5. The Supplementary Material document provided with this article offers additional theoretical findings, along with a series of auxiliary lemmas, technical details on the presented examples and all the proofs of our main results.

2. Background. In this section, we report general notation used throughout the paper and review basic results on max-stable models. The latter provide the mathematical and probabilistic background for our main theoretical findings.

2.1. Notation. Given $X \subset \mathbb{R}^d$, with $d \in \mathbb{N}_+ = \{1, 2, \ldots\}$, and $f : X \rightarrow \mathbb{R}$, let $\|f\|_{\infty} = \sup_{x \in X} |f(x)|$ and $\|f\|_1 = \int_X |f(x)| \, dx$. Let $I = (i_1, \ldots, i_k) \subset \{1, \ldots, d\}$ with $1 \leq k \leq d$ and $x_I := (x_{i_1}, i)$. For a differentiable function $f$ at the point $x_0 \in X$, we denote by $f'(x_0)$ its mixed derivative of order $k$ with respect to $x_I$. If $d = 1$ and $f$ is nondecreasing, we also denote by $f^{-1}(t) = \inf\{x \in X : f(x) \geq t\}$ the left generalised inverse, with $t \in \mathbb{R}$.

Let $F$ and $G$ be two probability measures (pm’s) on a generic measurable space $\mathbb{X} := (\mathcal{X}, \sigma(\mathcal{X}))$. When $F$ and $G$ are absolutely continuous with respect to a measure $\nu$ on $\mathbb{X}$, with density functions $f$ and $g$, $\mathcal{D}(f, g) = \int_{\mathbb{X}} (\log f - \log g) \, d\nu$ and $\mathcal{D}_H^2(f, g) = \int_{\mathbb{X}} (\sqrt{f} - \sqrt{g})^2 \, d\nu$ are the Kullback-Leibler divergence and the squared Hellinger distance, respectively. Furthermore, $\mathcal{D}_\infty(f, g) = \sup_{x \in X} |f(x) - g(x)|$. When $X$ is a separable metric space and $\sigma(X)$ is its Borel $\sigma$-algebra, $\mathcal{D}_T(F, G) = \sup_{B \in \sigma(X)} |F(B) - G(B)|$ is the total variation distance, while $\mathcal{D}_W(F, G)$ denotes a metric between pm’s that metrizes the topology of weak convergence [e.g., 39, p. 508]. When $X$ is a partially ordered subset of $\mathbb{R}^d$, for simplicity, given a pm $F$ we also denote with $F$ the corresponding (cumulative) distribution function. We then denote by $\mathcal{D}_{KS}(F, G) = \mathcal{D}_\infty(F, G)$ the Kolmogorov-Smirnov distance [e.g., 39, Ch 3.1.1]. Moreover, we denote by $F^{(n)}$ and $F^{(\infty)}$ the $n$-fold and infinite-fold product pm’s. Finally, $\delta_x(\cdot)$ is the Dirac-delta pm at $x \in X$, $\text{Leb}_X$ is the restriction of the Lebesgue measure on $X$.

2.2. Max-stable distributions. Let $Z = (Z_1, \ldots, Z_d)$ be a random vector (rv) with joint distribution $F$ and $Z_1, Z_2, \ldots$ be independent and identically distributed (iid) copies of it. Hereafter, operations between vectors are meant componentwise. We say that $F$ is in the max-domain of attraction of a max-stable distribution $G$, in symbols $F \in \mathcal{D}(G)$, if there exist norming sequences $a_m > 0 = (0, \ldots, 0)$ and $b_m \in \mathbb{R}^d$, $m = 1, 2, \ldots$, such that

$$\lim_{m \rightarrow \infty} F^m(a_m x + b_m) = G(x), \quad x \in \mathbb{R}^d,$$

where the margins of $G$ are nondegenerate. Essentially, the asymptotic distribution of the location-scale normalised componentwise maxima vector (obtained over an increasingly number of $Z$’s copies) must be max-stable, no matter what the form of $F$ is. Accordingly, the distribution $G$ satisfies the max-stability property: $G(x) = G^k(\alpha_k x + \beta_k)$ for some $\alpha_k > 0, \beta_k \in \mathbb{R}^d$ and all $k = 1, 2, \ldots$ [e.g., 31, p. 143]. Its form is

$$G(x) = C_{\text{EV}}(G_1(x_1), \ldots, G_d(x_d)), \quad x \in \mathbb{R}^d,$$

where $C_{\text{EV}}$ is an extreme-value copula, allowing the representation

$$C_{\text{EV}}(u) = \exp \left(-L((−\ln u_1), \ldots, (−\ln u_d))\right), \quad u \in (0, 1]^d,$$

where $L : [0, \infty)^d \rightarrow [0, \infty)$ is a homogeneous function of order 1, named stable-tail dependence function. We refer to Section 2.3 for additional details on the dependence structure.
For any $j \in \{1, \ldots, d\}$, the $j$-th margin of $G$ is of one of the following three types:

$$G_j(x) = \begin{cases} 
\exp(-x^{-\rho_j}), & x > 0, \rho_j > 0, \\
\exp(-\exp(-x)), & x \in \mathbb{R}, \\
\exp(-(-x)^{\omega_j}), & x < 0, \omega_j > 0,
\end{cases}$$

(2)

known as the $\rho_j$-Fréchet, the Gumbel and the (reverse) $\omega_j$-Weibull distribution function, respectively. We focus on classes of scale or location-scale multivariate max-stable distributions having margins all of the same type, denoted generically by $\{G_{\varrho}, \vartheta \in \Theta \}$.

**Definition 2.1.** We refer to the family of multivariate max-stable distributions $G_{\varrho}(x) = C_{\text{EV}}(G_{\varrho_1}(x_1), \ldots, G_{\varrho_d}(x_d))$ as:

(i) **multivariate $\rho$-Fréchet**, when $G_{\varrho_1}(x_j) = \exp(-(x_j/\sigma_j)^{-\rho_j})$, with $\varrho_j = (\rho_j, \sigma_j) \in \Theta_j = (0, \infty)^2$, $x_j > 0$, for all $j = 1, \ldots, d$;

(ii) **multivariate Gumbel**, when $G_{\varrho_1}(x_j) = \exp(-\exp(-(x_j - \mu_j)/\sigma_j))$, with $\varrho_j = (\sigma_j, \mu_j) \in \Theta_j = (0, \infty) \times \mathbb{R}$, $x_j \in \mathbb{R}$, for all $j = 1, \ldots, d$;

(iii) **multivariate $\omega$-Weibull**, when $G_{\varrho_1}(x_j) = \exp(-(-x_j - \mu_j)/\sigma_j)^{\omega_j})$, with $\varrho_j = (\omega_j, \sigma_j, \mu_j) \in \Theta_j = (0, \infty)^2 \times \mathbb{R}$, $x_j < \mu_j$, for all $j = 1, \ldots, d$.

In the three cases, the marginal parameters and their spaces are $\varrho = (\rho, \sigma) \in \Theta = (0, \infty)^{2d}$, $\varrho = (\sigma, \mu) \in \Theta = (0, \infty)^{d} \times \mathbb{R}^d$ and $\varrho = (\omega, \sigma, \mu) \in \Theta = (0, \infty)^{2d} \times \mathbb{R}^d$, respectively.

The density functions of the three classes are related as follows. Let $X = (X_1, \ldots, X_d)$ be a rv with distribution $G_{\varrho}$ and define $Y_j = U_{\varrho_1}(X_j)$, $j \in \{1, \ldots, d\}$, where

$$U_{\varrho_1}(x_j) = -1/\log G_{\varrho_1}(x_j),$$

with $x_j \in \text{supp}(G_{\varrho_1})$. Then, $Y = (Y_1, \ldots, Y_d)$ follows a max-stable distribution with common 1-Fréchet margins. We refer to the latter as multivariate 1-Fréchet-max-stable or simple max-stable distribution and denote it by $G_1$. Specifically, for $y > 0$, $G_1(y) = \exp(-V(y))$, where $V(y) = L(1/y)$ is known as the exponent function and is hereafter assumed to have mixed derivatives up to order $d$ at almost every $y > 0$. Thus, the multivariate simple max-stable density function is given by Faà di Bruno’s formula

$$g_1(y) = \sum_{P \in \mathcal{P}_d} G_1(y) \prod_{i=1}^m \{-V_i(y)\},$$

(3)

where $\mathcal{P}_d$ is the set of all the partitions $P = \{I_1, \ldots, I_m\}$ of $\{1, \ldots, d\}$, with $m = |P|$. As a result, the multivariate max-stable density of $G_\varrho$ is

$$g_\varrho(x) = U_{\varrho}'(x) g_1(U_{\varrho}(x)),$$

(4)

where $U_\varrho(x) = (U_{\varrho_1}(x_1), \ldots, U_{\varrho_d}(x_d))$ and $U_{\varrho}'(x) := \prod_{j=1}^d (\partial/\partial x_j) U_{\varrho_1}(x_j)$.

2.3. Extremal dependence. The extreme-value copula is fully characterised by the stable-tail dependence function, which in turn is by its restriction on $\mathcal{R} := \{t \in [0, 1]^{d-1} : \|t\|_1 \leq 1\}$, defined as $A(t) = L(1 - t_1 - \cdots - t_{d-1}, t_1, \ldots, t_{d-1})$ and named Pickands dependence function. Specifically,

$$A(t) = d \int_{\mathcal{S}} \max\{(1 - t_1 - \cdots - t_{d-1})w_1, \ldots, t_{d-1}w_d\}dH(w),$$

(5)

where $H$ is a pm on the $d$-dimensional unit simplex $\mathcal{S} := \{w \geq 0 : \|w\|_1 = 1\}$, named angular pm, and satisfies the mean constraints
(C1) \( \int_S w_j H(\mathrm{d}w) = 1/d, \forall j \in \{1, \ldots, d\} \).

As a result of (5), the Pickands dependence function satisfies the convexity and boundary constraints

\[
\begin{align*}
(C2) \quad A(at_1 + (1-a)t_2) &\leq aA(t_1) + (1-a)A(t_2), \quad a \in [0,1], \forall t_1, t_2 \in \mathcal{R}, \\
(C3) \quad 1/d &\leq \max (t_1, \ldots, t_{d-1}, 1 - t_1 - \cdots - t_{d-1}) \leq A(t) \leq 1, \forall t \in \mathcal{R}.
\end{align*}
\]

These are necessary and sufficient conditions to characterise the class of valid Pickands dependence functions in the case \( d = 2 \), while they are only necessary but not sufficient when \( d > 2 \), see e.g. [5, p. 257] for a counterexample.

In general, \( H \) can place mass on all the \( 2^d - 1 \) subspaces of \( S \) of the form \( S_I = \{ v \in \mathcal{S} : v_j > 0 \text{ if } j \in I; v_j = 0 \text{ if } j \not\in I \} \), with \( I \) a non-empty subset of \( \{1, \ldots, d\} \) [5, Ch. 7]. This allows to represent dependence intensity specific to subsets of variables (pairs, triplets, etc.) as well as the amount of dependence among all the \( d \) variables jointly. Notice that \( \mathcal{S}_{(j)} = \{ e_j \}, j = 1, \ldots, d \), where \( e_j \) is the \( j \)-th canonical basis vector, and \( \mathcal{S}_{\{1, \ldots, d\}} \) corresponds to the subset of points of \( \mathcal{S} \) having positive coordinates. In the sequel, we focus on the subset of all the possible angular pm’s given in Definition 2.2. Such a class makes statistical inference not too complicated and, at the same time, is sufficiently rich for applications.

**Definition 2.2.** Let \( \mathcal{H} \) denote the class of pm’s on the Borel sets of \( \mathcal{S} \) satisfying (C.1) and having null mass outside the subset \( \mathcal{S} := \mathcal{S}_{\{1, \ldots, d\}} \cup \{ e_1 \} \cup \cdots \cup \{ e_d \} \). For any \( H \in \mathcal{H} \) there are point masses \( p_j \in [0,1/d], j = 1, \ldots, d \), and a Lebesgue integrable function \( h : \mathcal{R} \mapsto [0, \infty), \mathcal{R} := \{ v \in (0,1)^{d-1} : \|v\|_1 < 1 \} \), named angular density, such that for all Borel subsets \( B \subset \mathcal{S} \) we have

\[
H(B) = \sum_{j=1}^d p_j \delta_{e_j}(B) + \int_{\pi_{\mathcal{R}}(B \cap \mathcal{S}_{\{1, \ldots, d\}})} h(v) \mathrm{d}v,
\]

with \( \pi_{\mathcal{R}} : \mathcal{S} \mapsto \mathcal{R} : (w_1, \ldots, w_{d-1}, w_d) \mapsto (w_1, \ldots, w_{d-1}) \).

H(\cdot), \; \tau \in \mathcal{R}, \text{ is the distribution function pertaining to the pm } H \circ \pi_{\mathcal{R}}, \text{ called angular distribution. Let } \mathcal{A} \text{ be the set of functions on } \mathcal{R} \text{ defined through the representation in (5), with } H \in \mathcal{H}. \text{ Hereafter, we refer to } \mathcal{H} \text{ and } \mathcal{A} \text{ as the spaces of valid angular pm’s and Pickands functions, respectively.}

The dependence level among the components of a max-stable rv can be described by means of a geometric interpretation of the angular pm. The more the mass of \( H \) concentrates around \((1/d, \ldots, 1/d)\) (the barycenter of the simplex) the more the variables are dependent on each other. On the contrary, the more the mass of \( H \) accumulates close to the vertices of the simplex, the less dependent the variables are. Alternatively, the dependence level can be described via the Pickands dependence function \( A \), as it satisfies the inequality in (C3), where the lower and upper bounds represent the cases of complete dependence and independence, respectively.

2.4. Polynomial representation of the extremal dependence. In recent years, different polynomial functions have been used to model the extremal dependence more flexibly than using specific parametric models. For example, polynomials in Bernstein form have been used to model the univariate angular distribution and Pickands dependence functions [e.g., 54, 41]. Piecewise polynomials as linear combinations of B-splines have been used to model the univariate Pickands dependence function in a regression setting [17]. Alternative polynomial representations have been considered, e.g., in [49] and [42]. In higher dimensions, the
Pickands dependence function is less tractable [53], while multivariate pm’s on the simplex as the angular pm can be conveniently modelled through density functions and point masses [e.g., 50, 4, 43]. Bernstein polynomials have shown to be very tractable from a theoretical and a computational viewpoint when performing Bayesian nonparametric inference [e.g. 38, 63, 54, 43] and allow for a representation of the multivariate angular density in terms of Dirichlet mixtures, one of the most popular model classes for angular pm’s [8, 70]. Hereafter, we focus on such an approach, aiming for a general and concise discussion. In Section B.1 of the supplement, we provide further details on the Bernstein polynomial characterisation of univariate angular distribution and Pickands dependence functions. These representations are useful to construct prior distributions yielding consistent Bayesian semiparametric procedures.

We briefly describe angular pm modelling via a mixture of polynomial densities and point masses. For an integer \( k > d \), let \( \Gamma_k \) be the set of multi-indices \( \alpha \in \{1, \ldots, k-d+1\}^d \) such that \( \alpha_1 + \cdots + \alpha_{d-1} \leq k-1 \) and \( \alpha_d = k - \alpha_1 - \cdots - \alpha_{d-1} \), whose cardinality is [e.g., 53]

\[
|\Gamma_k| = \binom{k-1}{d-1}.
\]

For each \( \alpha \in \Gamma_k \), the Bernstein polynomial basis function of index \( \alpha - 1 \) and degree \( k-d \)

\[
b_{\alpha-1}(t; k-d) = \frac{(k-d)!}{\prod_{j=1}^d (\alpha_j - 1)!} \prod_{j=1}^{d-1} t_j^{\alpha_j - 1} (1-t_1 - \cdots - t_{d-1})^{\alpha_d - 1}, \quad t \in \bar{R},
\]

can be rewritten as \( b_{\alpha-1}(t; k-d) = \text{Dir}(t; \alpha)(k-d)!/(k-1)! \) for all \( t \in \bar{R} \), where \( \text{Dir}(t; \alpha) \) denotes the Dirichlet probability density with parameters \( \alpha \) [e.g., 57]. Therefore, a \( (k-d) \)-th degree Bernstein polynomial representation of the angular density is given by

\[
h_{k-d}(t) = \sum_{\alpha \in \Gamma_k} \varphi_{\alpha} \text{Dir}(t; \alpha), \quad t \in \bar{R},
\]

where \( \varphi_{\alpha} \in [0, 1] \) for any \( \alpha \in \Gamma_k \). Set \( \kappa_j = k \epsilon_j, \ j \in \{1, \ldots, d\} \). According to [43], the following pm on the Borel subsets \( B \subset S \)

\[
H_k(B) := \sum_{j=1}^d \delta_{\epsilon_j}(B) \varphi_{\kappa_j} + \int_{\pi_B(B \cap S_{1, \ldots, d})} h_{k-d}(t) dt,
\]

is a valid angular pm if and only if the non-negative coefficients \( \varphi^{(k)} = (\varphi_{\kappa_1}, \ldots, \varphi_{\kappa_d}, \varphi_{\alpha}, \alpha \in \Gamma_k) \) satisfy the restrictions:

(R1) \( \sum_{\alpha \in \Gamma_k} \varphi_{\alpha} = 1 - \varphi_{\kappa_1} - \cdots - \varphi_{\kappa_d} \);

(R2) \( \sum_{j=1}^{d-1} \sum_{\alpha \in \Gamma_k : \alpha_j = 0} \varphi_{\alpha} = \frac{1}{d} - \varphi_{\kappa_j}, \forall j = 1, \ldots, d. \)

Let the classes of angular pm’s with density in Bernstein Polynomial (BP) form be defined by

\[
\mathcal{H}_k := \{ H_k \in \mathcal{H} : H(B) = \sum_{j=1}^d \delta_{\epsilon_j}(B) \varphi_{\kappa_j} + \int_{\pi_B(B \cap S_{1, \ldots, d})} h_{k-d}(t) dt, \ (\text{R1})-\text{(R2)} \text{ hold true}, \}
\]

for each integer \( k > d \). We have then the following approximation property, see Section D.1.2 of the supplement for a proof.

**Proposition 2.3.** For every \( H \in \mathcal{H} \) and any \( \epsilon > 0 \), there exists \( k > d \) and \( H_k \in \mathcal{H}_k \) such that \( \|h - h_{k-d}\|_1 < \epsilon. \)
As a result, $\bigcup_{k=d+1}^{\infty} H_k$ is a dense subset of $(\mathcal{H}, \mathcal{D}_T)$ and, therefore, of $(\mathcal{H}, \mathcal{D}_W)$ and $(\mathcal{H}, \mathcal{D}_{KS})$, where $\mathcal{D}_{KS}$ metrizes the space of angular pm’s through the uniform distance (over $\mathbb{R}$) between their angular distribution functions. This guarantees that a prior on the angular pm suitably specified via the Bernstein polynomial representation in (7) has full support. Proposition 2.3 further gives the mathematical ground for devising full support priors on $d$-vari ate max-stable densities and consistent Bayesian predictive methods.

### 3. Bayesian inference for max-stable models.

#### 3.1. Bayesian paradigm for max-stable models. In this section, we firstly consider a simple max-stable observational model and a Bayesian statistical setting that can be described in the following general terms. We assume that the observables are iid rv’s $Y_1, \ldots, Y_n$ in $(0, \infty)^d$ following a simple max-stable distribution $G_1(\cdot|\theta_0)$. In particular, $\theta_0$ is an unknown infinite-dimensional parameter corresponding either to the true Pickands dependence function $A_0$ or to the true angular pm $H_0$. Hereafter, the corresponding stable tail and exponent functions are denoted by $L(\cdot|\theta_0)$ and $V(\cdot|\theta_0)$, respectively. Observe that, due to the 1-to-1 relation between $A_0$ and $H_0$, these two yield equivalent representations of the dependence structure. We thus consider a statistical model of the form $\{G_1^{(n)}(\cdot|\theta) : \theta \in \Theta\}$, where $\Theta$ equals either $\mathcal{A}$ or $\mathcal{H}$. We endow $\Theta$ with a metric $\mathcal{D}$ that makes it separable and such that, equipping $\Theta$ with the associated Borel $\sigma$-field $\mathcal{B}_\Theta$, $(\Theta, \mathcal{B}_\Theta)$ is a standard Borel measurable space [e.g., 73, p. 96]. Thus, for every Borel set $B$, $\Pi_\Theta(B) = \mathbb{P}(\theta \in B)$ is a prior distribution. By Theorem V.58 in [24], there exists a version of the probability density $g_1(y|\theta)$ (see (3)), which is jointly measurable in the parameter and the observation for almost every $y \in (0, \infty)^d$. Thus, a version of the posterior distribution is given by Bayes’ theorem

$$
\Pi_n(B) := \Pi_\Theta(B|Y_1, \ldots, Y_n) = \frac{\int_B \prod_{i=1}^{n} g_1(Y_i|\theta) d\Pi_\Theta(\theta)}{\int_{\Theta} \prod_{i=1}^{n} g_1(Y_i|\theta) d\Pi_\Theta(\theta)}.
$$

In particular, such a version of the conditional distribution of $\theta$ given $(Y_1, \ldots, Y_n)$ is regular: i.e., a Markov kernel from the sample space of data into $(\Theta, \mathcal{B}_\Theta)$. Moreover, defining

$$
G_1 := \{g_1(\cdot|\theta) : \theta \in \Theta\},
$$

the map $\phi_\Theta : (\Theta, \mathcal{D}) \rightarrow (G_1, \mathcal{D}_H) : \theta \mapsto g_1(\cdot|\theta)$ is Borel. Therefore, the prior $\Pi_\Theta$ on the dependence parameter induces a prior $\Pi_{G_1} = \Pi_\Theta \circ \phi_\Theta^{-1}$ on the Borel sets of $(G_1, \mathcal{D}_H)$, whose posterior distribution is the random pm $\Pi_n = \Pi_{G_1} \circ \phi_{G_1}^{-1}$. In this setup, almost sure convergence of the posterior distributions $\Pi_n$ and $\tilde{\Pi}_n$ to the Dirac pm’s $\delta_{\theta_0}$ and $\delta_{g_1(\cdot|\theta_0)}$ is equivalent to the following fact [e.g., 39, Ch 6].

\begin{definition}
The posterior distributions $\Pi_n(\cdot)$ and $\tilde{\Pi}_n$ are almost surely (as) consistent at $\theta_0$ and $g_1(\cdot|\theta_0)$ if, for all the neighbourhoods $\theta_0 \in U \subset \Theta$ and $g_1(\cdot|\theta_0) \in U \subset G_1$,

$$
\lim_{n \to \infty} \Pi_n(U^c) \to 0, \quad \lim_{n \to \infty} \tilde{\Pi}_n(U^c) \to 0, \quad G_1^{(\infty)}(\cdot|\theta_0) - a.s.
$$

\end{definition}

These notions have analogous extensions to semiparametric statistical models $\{G_{\theta}(\cdot|\theta) : (\theta, \Theta) \in \Theta \times \Theta\}$, corresponding to the three types of max-stable distribution classes introduced in Section 2.2.
3.2. Simple max-stable distributions. In this subsection, we establish almost sure consistency of the posterior distributions \( \Pi_n \) and \( \tilde{\Pi}_n \), previously introduced for data following a simple max-stable model. This first step allows to derive most of the mathematical ground used to establish consistency also for the more general max-stable models, which are more realistic for applications, see Section 3.3. The consistency results presented in the sequel concern prior distributions \( \Pi_\Theta \) on the extremal dependence constructed via the representation in Section 2.4. Nevertheless, they are based on a fairly general theory of Borel pm’s on the space \( \mathcal{H} \) and can be adapted to alternative prior specifications. A similar approach tailored to Pickands dependence function in the case \( d = 2 \) is presented in the supplement. A plethora of extremal dependence functionals used by practitioners [e.g., 25, Ch 2.2] are defined for pairs of variables and are readily obtainable from the Pickands dependence function. Thus, posterior consistency for those can be straightforwardly deduced from our results in Sections B.2 and B.3 of the supplement.

Consistency of the posterior distribution \( \tilde{\Pi}_n \) on simple max-stable densities is obtained via an extended version of Schwartz’s theorem [39, Theorem 6.23]. The latter yields consistency via sequential partitioning of the density space into a set whose entropy grows at most linearly in the sample size and a set of exponentially decaying prior probability. A key requirement is the Kullback-Leibler property of prior distributions defined below.

**Definition 3.2.** We say that the prior \( \Pi_\Theta \) possess the Kullback-Leibler property at \( \theta_0 \) (equivalently, \( g_1(\cdot|\theta_0) \) belongs to the Kullback-Leibler support of \( \Pi_\Theta \)) if

\[
\Pi_\Theta(\theta \in \Theta : g_1(\cdot|\theta) \in K_\epsilon) = \Pi_{\Phi^1}(K_\epsilon) > 0,
\]

for all \( K_\epsilon := \{g \in \mathcal{G}_1 : \mathcal{K}(g_1(\cdot|\theta_0), g) < \epsilon\}, \epsilon > 0 \).

In Section 3.2.1 we first study the metrization of \( \mathcal{H} \) under the two classical distances \( D_W \) and \( D_{KS} \) between pm’s. Secondly, we give sufficient conditions under which a generic Borel prior on the angular pm possesses the Kullback-Leibler property. These technical preliminaries set the stage for then establishing Hellinger consistency of \( \tilde{\Pi}_n \) (Section 3.2.2). Consistency of \( \Pi_n \) is also deduced under one of the two metrics for angular pm’s, according to the dimensionality of the data.

3.2.1. Kullback-Leibler theory for priors on the extremal dependence. We focus on prior distributions specified on the angular pm. From now on, we thus stick to the parametrisation where \( \theta = H, \Theta = \mathcal{H} \), and we set \( \theta = h \). Specifically, we consider prior distributions \( \Pi_\Theta(B) = \mathbb{P}(\theta \in B) \) on the Borel sets \( B \) of \((\Theta, D_W)\).

**Proposition 3.3.** In arbitrary dimension \( d \geq 2 \), the space of Borel pm’s on \( \tilde{S} \), endowed with the topology of weak convergence of measures, is separable and completely metrisable. Moreover, \( \Theta \), endowed with the associated subspace topology, is a standard Borel space, whose Borel \( \sigma \)-field coincides with the one induced by \( D_{KS} \).

**Corollary 3.4.** In dimension \( d \geq 2 \), there exists a version of the simple max-stable density such that the map \((\Theta, y) \mapsto g_1(y|\theta)\) is jointly measurable and \( \phi_{\Theta} : (\Theta, D_W) \mapsto (\mathcal{G}_1, D_H) : \theta \mapsto g_1(\cdot|\theta) \) is a Borel map, where \( \mathcal{G}_1 \) is defined as in (10). Moreover, for all \( \epsilon > 0 \) and \( K_\epsilon \) as in Definition 3.2, \( \phi_{\Theta}^{-1}(K_\epsilon) \) is a Borel set of \((\Theta, D_W)\).

The proofs of the above results are provided in Sections D.3.1 and D.3.2 of the supplement. While it is already known that the class of all pm’s on \( S \), equipped with the Borel \( \sigma \)-algebra induced by \( D_W \), is a standard Borel space, this is not necessarily true for pm’s on
subspaces of $\mathcal{S}$ or different metrics. According to Proposition 3.3, the metrization of $\Theta$ via $\mathcal{D}_W$ generates a standard Borel space, therefore the existence of a regular version of the posterior distribution $\Pi_\Theta$ arising from $\Pi_\Theta$ is guaranteed. This is a necessary requirement in order to legitimately study its consistency [e.g., 39, Ch 1.3]. As a consequence of Corollary 3.4, $\Pi_\Theta$ also induces a prior $\Pi_{G_1}$ on the Borel sets of $(G_1, \mathcal{D}_H)$. To inspect its Kullback-Leibler support, we introduce the following notion.

**Definition 3.5.** Let $\theta_* \in \Theta$ and $\|\dot{\theta}_*\|_\infty < \infty$. A prior $\Pi_\Theta$ is said to possess the $\mathcal{D}_\infty$-property at $\theta_*$ if, for any $\epsilon > 0$, it has positive inner probability on \{ $\theta \in \Theta : \mathcal{D}_\infty(\dot{\theta}, \dot{\theta}_*) \leq \epsilon$ \}.

We focus on true angular pm’s which belong to a particular subclass of $\Theta$, specified in Definition 3.6. In simple terms, we require true angular pm’s whose angular density is uniformly continuous on $\mathring{\mathbb{R}}$, while allowing for nonnegative point masses on the vertices of $\mathcal{S}$ (see Definition 3.6(i)-(ia)). In the bivariate case, we however allow for angular densities that diverge at the boundary of $\mathring{\mathbb{R}}$ (Definition 3.6(i)-(ib)). Unbounded angular densities are less tractable in high dimensions, while nonessential for practical statistical purposes when point masses are admitted. Popular models such as the Asymmetric Logistic, Dirichlet and Symmetric Logistic and Husler-Reiss for suitable values of the parameters [e.g., 6] are comprehended by the family in Definition 3.6(i)-(ia). We establish that a sufficient condition for $\Pi_\Theta$ to possess the Kullback-Leibler property at true admissible pm’s is that it possesses the $\mathcal{D}_\infty$-property at a subfamily of those (Definition 3.6(ii)), see Section D.3.3 of the supplement for details.

**Definition 3.6.** Let $\Theta$ be as in Definition 2.2 and:

(i) Let $\Theta_0 \subset \Theta$ be the class of angular pm’s whose angular density functions $\dot{\theta}$ are continuous on $\mathring{\mathbb{R}}$ and satisfy one of the following:

   (ia) $\dot{\theta}$ is uniformly continuous on $\mathring{\mathbb{R}}$;
   (ib) (only if $d = 2$) $\inf_{t \in (0, 1)} \dot{\theta}(t) > 0$ and $\lim_{t \downarrow 0} \dot{\theta}(t) = \lim_{t \uparrow 1} \dot{\theta}(t) = +\infty$.

(ii) Let $\Theta' \subset \Theta_0$ be the set of angular pm’s such that $\theta(\{e_j\}) = p_j > 0$, for $j = 1, \ldots, d$, with angular density satisfying the property in (ia) and $\inf_{t \in \mathring{\mathbb{R}}} \dot{\theta}(t) > 0$.

**Theorem 3.7.** Let $\theta_0 \in \Theta_0$ be the true angular pm. Assume that for any $\theta \in \Theta'$, the prior $\Pi_\Theta$ posses the $\mathcal{D}_\infty$-property at $\theta$. Then, for all $\epsilon > 0$

$$\Pi_\Theta(\phi^{-1}_\Theta(K_\epsilon)) = \Pi_{G_1}(K_\epsilon) > 0.$$

In the general case $d \geq 2$, since we concentrate on uniformly continuous angular densities, which can be uniformly approximated by polynomials, priors constructed through suitable polynomial representations thus satisfy the $\mathcal{D}_\infty$-property of Theorem 3.7. The latter is leveraged in the proof of Theorem 3.9 to establish posterior consistency for the specific case of priors on the angular pm constructed via the Bernstein polynomial representation in Section 2.4. See also Example 3.10 for the outline of a prior specification complying with the assumptions of Theorem 3.7. A slightly more general version of Theorem 3.7 can be achieved by allowing for bounded and possibly discontinuous multivariate angular densities. Nevertheless, since $\mathcal{D}_H$-dense in $G_1$, the class of simple max-stable densities with angular pm in $\Theta_0$ seems sufficiently rich.
3.2.2. Posterior consistency. For each $\theta$ in the class $\Theta$ given in Definition 2.2, $g_1(\cdot|\theta)$ equals almost everywhere the expression in (3) and is explicitly linked to $\theta$ via the relation (11)

$$-V_1(y|\theta) = \begin{cases} 
 dp_j y_j^{-2} + d \left( \int_{[0,y_j^2]} \|z\|_1^{d-1} \hat{\theta} \circ \pi_R(z/\|z\|_1) \big|_{z_j = y_j} \, dz_j \right), & \text{if } |I_i| = \{j\}, \\
 d \left( \int_{[0,y_j^2]} \|z\|_1^{d-1} \hat{\theta} \circ \pi_R(z/\|z\|_1) \big|_{z_j = y_j} \, dz_j \right), & \text{otherwise}
\end{cases}$$

where $y_{I_i}$ and $y_{I_i^c}$ are the restrictions of $y$ to $I_i$ and $I_i^c = \{1, \ldots, d\} \setminus I_i$, and $\pi_R$ is the projection map in (2.2). See also [27, 28] and Section C.1 of the supplement for details. We consider priors on the Borel sets of $\Theta$ by projection map in (2.2). See also [27, 28] and Section C.1 of the supplement for details. We consider priors on the Borel sets of $(\Theta, \mathcal{D}_W)$ constructed via the representation in Section 2.4 as follows, see also Section C.4.2 of the supplement for additional set-theoretical details.

**Condition 3.8.** For $k = 1, 2, \ldots$, let $\Phi_k := \{\varphi^{(k)} : (R1)-(R2) hold true\}$. For some $k_* \in \mathbb{N}_+ \setminus \{1, \ldots, d\}$, let $\Pi$ be a pm defined on the disjoint union space $\bigcup_{k \geq k_*} \{\{k\} \times \Phi_k\}$ and constructed via direct sum of coefficient spaces

$$\{(\Phi_k, \Sigma_k, \lambda(k) \nu_k), k \geq k_*\},$$

such that:

(i) $\nu_k$ is a fully supported pm on $\Phi_k$, equipped with Borel $\sigma$-field $\Sigma_k$;

(ii) $\lambda$ is a positive probability mass function on $\{k_* + 1, \ldots\}$, satisfying for some $q > 0$,

$$\sum_{i \geq k} \lambda(i) \lesssim \exp(-qk^{d-1}), \quad k \to \infty.$$ 

We assume the prior $\Pi_0\Theta$ is induced by the pm $\Pi$, via the representation (8).

**Theorem 3.9.** Let $Y_1, \ldots, Y_n$ be iid rv’s with distribution $G_1(\cdot|\theta_0)$, where $\theta_0 \in \Theta_0$ and $\Theta_0$ is as in Definition 3.6(i). Assume $\Pi_0\Theta$ satisfies Condition 3.8. Then, $\Pi G_1$ has full Hellinger support and, $G_1^{(\infty)}(\cdot|\theta_0)$ as:

(a) $\lim_{n \to \infty} \Pi_n(U_{\mathcal{D}_H}^{\varphi}) = 0$, for every $\mathcal{D}_H$-neighbourhood $U$ of $g_1(\cdot|\theta_0)$;

(b) $\lim_{n \to \infty} \Pi_n(U_{\mathcal{D}_W}^{\varphi}) = 0$, for every $\mathcal{D}_W$-neighborhood (if $d \geq 2$) or $\mathcal{D}_{KS}$-neighborhood $U_1$ of $\theta_0$ (if $d = 2$).

The proof of Theorem 3.9 is provided in Section D.3.5 of the supplement. Consistency of the posterior of the angular pm is obtained with a stronger metric in the bivariate case. Hellinger consistency of $\Pi_n$ entails that $\Pi_n$ concentrates on sets of angular pm’s whose Pickands dependence functions lie in a neighbourhood of $A_0$ under a Sobolev-type metric, induced by the norm $\|\mathcal{A}\|_\infty + \sum_{1 \leq j \leq d-1} \|A_{(j)}\|_\infty$. The relations

$$A_{(j)}(t) = d \int_S \{w_{j+1} \mathbb{I}_{(w_{j+1}+t,1]}(w_{j+1}) - w_{j+1} \mathbb{I}_{(w_{j+1},1]}(w_{j+1})\} \, d\theta(w),$$

for $j = 1, \ldots, d-1$, where $w^*(t) = y^*(t)/\|y^*(t)\|_1$ with $t \in \mathcal{R}$ and $y^*(t) = (1/(1 - \|t\|_1), 1/t_{d-1}, \ldots, 1/t_{d-1})$, allow to turn such a concentration result into consistency on the space of valid angular pm’s $\Theta$ with metric $\mathcal{D}_{KS}$, only when $d = 2$. An exception is the case where the true angular pm $\theta_0$ has no point masses. If so, Polya’s theorem [e.g., 39, Proposition A.11] guarantees that every open $\mathcal{D}_{KS}$-neighbourhood of $\theta_0$ contains a $\mathcal{D}_W$-neighbourhood, thus consistency under the weaker metric extends to consistency under the stronger one. We next sketch a prior construction for $\Pi$ that satisfies Condition 3.8.
Example 3.10. A prior $\Pi$ that exploits the representation of the angular pm in (8) can be obtained by choosing, for each $k \geq d+1$, a truncated Dirichlet prior $\nu_k(\varphi_k)$ on $|\Gamma_k|+d$-dimensional weight coefficients, with truncation outside $\Phi_k$ [43, Section 3]. Prior specification can thus be completed by choosing $\lambda(k)$ as the probability mass function of a truncated discrete Weibull distribution, with shape parameter $d-1$.

3.3. $\rho$-Fréchet-, $\omega$-Weibull- and Gumbel-max-stable distributions. In this section we extend the consistency results of Section 3.2 to the more general family of statistical models $\{G_{\theta,\varphi}^{(n)} : (\theta, \varphi) \in \Theta \times \Theta\}$. All the proofs can be found in Section D.4 of the supplement. Notably, they give the mathematical ground for addressing the problem discussed in Sections 4.3–4.5, where a max-stable distribution is fitted to sample maxima.

3.3.1. General problem formulation. Uncertainty on the finite dimensional parameter $\vartheta$ is expressed by means of a Borel prior $\Pi_{\Theta \times \Theta}$ on the space $\Theta$, endowed with the $L_1$ metric. The joint prior distribution $\Pi_{\Theta \times \Theta}$ on the Borel subsets of $\Theta \times \Theta$ gives rise to the posterior distribution

$$
\Pi_n(B) := \Pi_{\Theta \times \Theta}(B|\{X_i\}_{i=1}^n) = \frac{\int_B \prod_{i=1}^n g_{\theta}(X_i|\theta)\Pi_{\Theta \times \Theta}(d\theta, d\vartheta)}{\int_{\Theta \times \Theta} \prod_{i=1}^n g_{\theta}(X_i|\theta)\Pi_{\Theta \times \Theta}(d\theta, d\vartheta)},
$$

where $X_1, \ldots, X_n$ are iid rv's with distribution $G_{\theta,\varphi}(\cdot|\theta_0)$. In particular, $\varphi_{\Theta \times \Theta} : (\theta, \varphi) \mapsto g_{\theta}(\cdot|\theta) = U_{\varphi}(\cdot|\theta_0)$, with $U_\varphi$ and $U'_{\varphi}$ as in (4), is a Borel measurable map between $\Theta \times \Theta$ and the class of max-stable densities $G_{\Theta} := \{g_{\theta}(\cdot|\theta) : (\theta, \varphi) \in (\Theta, \Theta)\}$ equipped with the metric $d_{\varphi}$. Thus, $\Pi_{\Theta \times \Theta}$ also induces a prior $\Pi_{\Theta \times \Theta} = \Pi_{\Theta \times \Theta} \circ \varphi_{\Theta \times \Theta}^{-1}$ on the density $g_{\theta}(\cdot|\theta)$. In the sequel, almost sure consistency of the posterior distributions $\tilde{\Pi}_n(\cdot) = \Pi_{G_{\theta,\varphi}}(\cdot|\{X_i\}_{i=1}^n)$ and $\Pi_n$ is obtained by adapting arguments from Schwartz' theorem to the semiparametric model under study. Within this approach, consistency is derived from the exponential decay to 0 of posterior distributions (Proposition 3.12, see also Remark D.7 in the supplement), which is obtained under the following general sufficient condition.

Condition 3.11. The observational model $\{G_{\theta}(\cdot|\theta) : \theta \in \Theta, \varphi \in \Theta\}$ belongs to one of the classes in Definition 2.1. Moreover:

(i) the Kullback-Leibler support of $\Pi_{\Theta \times \Theta}$ contains $g_{\theta,\varphi}(\cdot|\theta_0)$;

(ii) for every $\mathcal{H}$-ball $\mathcal{U}_\varepsilon := \{g \in \mathcal{G}_{\Theta} : \mathcal{H}(g, g_{\theta_0}, \cdot|\theta_0) \leq 4\varepsilon\}$, $\varepsilon > 0$, and every $\delta \in (0, \delta_s)$, for $\delta_s > 0$, there exist a sequence of measurable partitions $\{\mathcal{G}_{\theta,\varphi}^0, \mathcal{G}_{\theta,\varphi}^0\}$ of the submodel

$$
\mathcal{U}_\varepsilon \cap \{g_{\theta}(\cdot|\theta) : (\theta, \varphi) \in \Theta \times \Theta, \|\varphi - \varphi_0\|_1 \leq \delta\}
$$

and test functionals $\tau_n(\{X_i\}_{i=1}^n) = (s_n(\{X_i\}_{i=1}^n), t_{n,1}(\{X_{i,1}\}_{i=1}^n), \ldots, t_{n,d}(\{X_{i,d}\}_{i=1}^n))$ with values in $[0, 1]$, such that

$$
\Pi_{\mathcal{G}_{\theta,\varphi}^0}(\mathcal{G}_{\theta,\varphi}^0) \leq e^{-r n}, \quad n \to \infty,
$$

for some $r > 0$ and:

(iia) $s_n$ satisfies

$$
\int s_n(\{X_i\}_{i=1}^n) dG^{(n)}_{\theta_0}(\{X_i\}_{i=1}^n|\theta_0) \leq e^{-n\varepsilon^2},
$$

$$
sup_{G : g \in \mathcal{G}_{\theta,\varphi}^0} \int \{1 - s_n(\{X_i\}_{i=1}^n)\} dG^{(n)}(\{X_i\}_{i=1}^n) \leq e^{-2n\varepsilon^2};
$$
consistency of Bayesian inference for max-stable distributions

(iib) for each \( j = 1, \ldots, d \), \( t_{n,j} \) satisfies

\[
\sup_{\vartheta_j \in \text{supp}(\Pi_\vartheta_j)} \frac{1}{\|\vartheta_j - \vartheta_{0,j}\|_\infty > \delta/d} \int \{1 - t_{n,j}((x_{i,j})_{i=1}^n)\} dG_{\vartheta_j}((x_{i,j})_{i=1}^n) \lesssim e^{-nc_j(\delta)},
\]

as \( n \to \infty \), where \( c_j(\delta) \) is a positive constant and \( \Pi_\vartheta_j \) is the marginal prior on \( \vartheta_j \).

Condition 3.11(i) requires that the prior satisfies the Kullback-Leibler property (this is already studied in Section 3.2.1 for the multivariate simple max-stable distributions). Condition 3.11(ii) adapts Schwartz’ s testing condition to the three model classes in Definition 2.1. It requires the existence of exponentially powerful tests \( t_{n,j} \), \( j = 1, \ldots, d \), each of them depending on the \( j \)-th marginal sample, to test the null hypothesis on the \( j \)-th marginal parameter “\( \vartheta_j = \vartheta_{0,j} \)” against alternatives lying outside an \( L_\infty \)-neighbourhood (point (iib)). Intuitively, they guarantee that the posterior asymptotically concentrates its mass away from those max-stable densities having the \( j \)-th margin “dissimilar” from the true one. Exponentially powerful tests \( s_n \) depending on the full sample are also required, for testing the null hypothesis “\( g_{\vartheta}(|\theta|) = g_{\vartheta_0}(|\theta_0|) \)” against alternative densities which have univariate margins close to the true ones, but lying outside a \( D_H \)-neighbourhood (point (iia)). These ensure instead that the posterior asymptotically concentrates no mass on max-stable densities having univariate margins and dependence structure possibly “similar” the true one \( g_{\vartheta_0}(|\vartheta|) \), but still being topologically distinguishable from it.

**Proposition 3.12.** Let \( X_1, \ldots, X_n \) be iid rv’s with distribution \( G_{\vartheta_0}(\cdot|\theta_0) \) and assume that Condition 3.11 is satisfied. Then, for every \( D_H \)-neighbourhood \( U \) of \( g_{\vartheta_0}(\cdot|\theta_0) \), \( D_W \)-neighbourhood (if \( d \geq 2 \)) or \( D_{KS} \)-neighbourhood (if \( d = 2 \)) \( U_1 \) of \( \theta_0 \) and \( L_2 \)-neighbourhood \( U_2 \) of \( \vartheta_0 \), eventually \( G_{\vartheta_0}^{(\infty)}(\cdot|\theta_0) \) as

\[
\max \left\{ \Pi_n(U_1^C), \Pi_n((U_1 \times U_2)^C) \right\} \leq e^{nc_n \Xi_n((X_i)_{i=1}^n, \tau_n, \Pi_{\Theta \times \Theta})}
\]

for a constant \( c > 0 \) and a functional \( \Xi_n \) which depend on the neighbourhoods’ choice and comply with

\[
G_{\vartheta_0}^{(n)}(e^{2nc_n \Xi_n((X_i)_{i=1}^n, \tau_n, \Pi_{\Theta \times \Theta})} > \varepsilon|\theta_0) \lesssim e^{-c'n}, \quad \forall \varepsilon > 0,
\]

for some constant \( c' > 0 \).

Strong consistency for subclasses of the three max-stable models in Definition 2.1 can be simultaneously deduced from Proposition 3.12 (see Corollary D.8 in the supplement), yet under the rather abstract Condition 3.11. Our specific working assumptions on semiparametric prior specifications and true parameter values are compactly reported next.

**Condition 3.13.** (i) The observational model \( \{G_{\vartheta}(\cdot|\theta) : \theta \in \Theta, \vartheta \in \Theta \} \) belongs to one of the classes in Definition 2.1, \( \theta_0 \in \Theta_0 \) and the prior \( \Pi_{\Theta} \):

- (ia) has full support on \((0, \infty)^{2d}\) when \( G_{\vartheta}(\cdot|\theta) \) is \( \rho \)-Fréchet;
- (ib) has full support on \((0, \infty)^d \times \mathbb{R}^d\) when \( G_{\vartheta}(\cdot|\theta) \) is Gumbel;
- (ic) has marginal distributions \( \Pi_{\vartheta_j} \) on \( \vartheta_j \) supported on compact subsets \( K_j \) of \((1, \infty) \times (0, \infty) \times \mathbb{R} \), for \( j = 1, \ldots, d \), when \( G_{\vartheta_j}(\cdot|\theta) \) is \( \omega \)-Weibull; moreover, \( \Pi_{\vartheta_j}(U) > 0 \), for every neighbourhood \( U \) of the true parameter \( \vartheta_0 = (\omega_0, \sigma_0, \mu_0) \).

(ii) The prior distribution on \((\theta, \vartheta)\) is \( \Pi_{\Theta \times \Theta} = \Pi_\Theta \times \Pi_\Theta \).

(iii) \( \Pi_\Theta \) satisfies Condition 3.8.
In Section 3.3.2, we establish posterior consistency for the specific cases of $\rho$-Fréchet, $\omega$-Weibull and Gumbel multivariate max-stable models, respectively, by verifying for each model that Condition 3.13 entails Condition 3.11 and then applying Proposition 3.12.

3.3.2. Model specific results. We start extending consistency to the case of $\rho$-Fréchet max-stable models, with unknown shape and scale parameters $\rho = (\rho_1, \ldots, \rho_d)$ and $\sigma = (\sigma_1, \ldots, \sigma_d)$, respectively. Their density function is

$$g_{\rho, \sigma}(x|\theta) = \prod_{j=1}^{d} \frac{\rho_j \sigma_j}{\sigma_j} x_j^{\rho_j - 1} g_1 \left( \frac{x_1}{\sigma_1}, \ldots, \frac{x_d}{\sigma_d} \mid \theta \right), \quad x > 0,$$

with $g_1(\cdot|\theta)$ almost everywhere as in (3). In this case, a joint prior $\Pi_{\Theta \times \Theta}$ is assigned to $\theta$ and $(\rho, \sigma)$, inducing a prior $\Pi_{\Theta}$ on $\mathcal{G}_\Theta := \{g_{\rho, \sigma}(\cdot|\theta) : \theta \in \Theta, (\rho, \sigma) \in (0, \infty)^{2d}\}$.

**Theorem 3.14.** Let $X_1, \ldots, X_n$ be iid rv's with distribution $G_{\rho_0, \sigma_0}(\cdot|\theta_0)$. Then, under Conditions 3.13(ii)–(iii) and 3.13(ia), $G_{\rho_0, \sigma_0}(\cdot|\theta_0) - \mathbb{P}$ as

(a) $\lim_{n \to \infty} \bar{\Pi}_n(U^0) = 0$, for every $\mathcal{H}$-neighbourhood $\bar{U}$ of $g_{\rho_0, \sigma_0}(\cdot|\theta_0)$;

(b) $\lim_{n \to \infty} \Pi_n((U_1 \times U_2)^0) = 0$, for every $\mathcal{W}$-neighbourhood (if $d \geq 2$) or $\mathcal{K}_S$-neighborhood $U_1$ of $H_0$ (if $d=2$) and $L_1$-neighborhood $U_2$ of $(\rho_0, \sigma_0)$.

We next consider $\omega$-Weibull max-stable models, with unknown shape, scale and location parameters $\omega = (\omega_1, \ldots, \omega_d)$, $\sigma = (\sigma_1, \ldots, \sigma_d)$ and $\mu = (\mu_1, \ldots, \mu_d)$, respectively. Their density function is

$$g_{\omega, \sigma, \mu}(x|\theta) = \prod_{j=1}^{d} \frac{\omega_j - x_j}{\sigma_j} \left( \frac{\mu_j - x_j}{\sigma_j} \right)^{-\omega_j - 1} g_1 \left( \frac{\mu_1 - x_1}{\sigma_1}, \ldots, \frac{\mu_d - x_d}{\sigma_d} \mid \theta \right)$$

for $x < \mu$. For technical convenience, we restrict the parameter space to $\Theta = (1, \infty)^d \times (0, \infty)^d$ and consider a prior $\Pi_{\Theta}$ on $(\omega, \sigma, \mu)$ with compactly supported marginal priors on $(\omega_j, \sigma_j, \mu_j)$, $j = 1, \ldots, d$, yielding a prior $\Pi_{\mathcal{G}_\Theta}$ supported on a subset of the model class $\mathcal{G}_\Theta = \{g_{\omega, \sigma, \mu} : \theta \in \Theta, (\omega, \sigma, \mu) \in \Theta\}$.

**Theorem 3.15.** Let $X_1, \ldots, X_n$ be iid rv's with distribution $G_{\omega_0, \sigma_0, \mu}(\cdot|\theta_0)$. Then, under Conditions 3.13(ii)–(iii) and 3.13(ia), $G_{\omega_0, \sigma_0, \mu}(\cdot|\theta_0) - \mathbb{P}$ as

(a) $\lim_{n \to \infty} \bar{\Pi}_n(U^0) = 0$, for every $\mathcal{H}$-neighbourhood $\bar{U}$ of $g_{\omega_0, \sigma_0, \mu}(\cdot|\theta_0)$;

(b) $\lim_{n \to \infty} \Pi_n((U_1 \times U_2)^0) = 0$, for every $\mathcal{W}$-neighbourhood (if $d \geq 2$) or $\mathcal{K}_S$-neighborhood $U_1$ of $H_0$ (if $d=2$) and $L_1$-neighborhood $U_2$ of $(\omega_0, \sigma_0, \mu_0)$.

**Remark 3.16.** The restriction $\omega > 1$ imposed to shape parameters rules out some $\omega$-Weibull max-stable densities for which the construction of Kullback-Leibler neighbourhoods is intractable. We point out that such a restriction is a quite common practice in extreme-value applications [e.g., 58, page 725]. The three-parameter Weibull distribution is an irregular model and testing $(\omega_{0,j}, \sigma_{0,j}, \mu_{0,j})$, for $j \in \{1, \ldots, d\}$, against the complement of a large compact neighbourhood with exponentially bounded errors remains an open problem. Herein, this issue is circumvented using a prior $\Pi_{\Theta}$ with compactly supported margins. In practice, confining prior specification on shape, scale and location parameters to an arbitrarily large compact set is hardly a restriction, as in applied sciences the physically reasonable ranges for the parameters are often finite. See also [9, page 845] for similar considerations concerning maximum likelihood inference for the univariate generalised extreme-value distribution, where maximisation is restricted to compact parametric subspaces.
Finally, we consider the case where the data come from a Gumbel max-stable distribution, with unknown scale and location parameters $\sigma = (\sigma_1, \ldots, \sigma_d)$ and $\mu = (\mu_1, \ldots, \mu_d)$, respectively. The pertaining density function is

$$
g_{\sigma, \mu}(x|\theta) = \prod_{j=1}^{d} \frac{e^{(x_j-\mu_j)/\sigma_j}}{\sigma_j} g_1\left(\frac{e^{(x_1-\mu_1)/\sigma_1}, \ldots, e^{(x_d-\mu_d)/\sigma_d}}{\theta}\right), \quad x \in \mathbb{R}^d.
$$

A prior $\Pi_\theta$ is now specified on $(\sigma, \mu)$, yielding a prior $\Pi_{G_\theta}$ on the max-stable class $G_\theta = \{g_{\sigma, \mu}(|\theta) : \theta \in \Theta, (\sigma, \mu) \in (0, \infty)^d \times \mathbb{R}^d\}$.

**Theorem 3.17.** Let $X_1, \ldots, X_n$ be iid rv’s following $G_{\sigma_0, \mu_0}(\cdot|\theta_0)$. Then, under Conditions 3.13(ii)–(iii) and 3.13(iii) and 3.13(b), $G_{\sigma_0, \mu_0}^{(\infty)}(\cdot|\theta_0)$ — as

(a) $\lim_{n \to \infty} \Pi_n(U^B) = 0$, for every $D_H$-neighbourhood $U$ of $g_{\sigma_0, \mu_0}(\cdot|\theta_0)$;

(b) $\lim_{n \to \infty} \Pi_n((U_1 \times U_2)^B) = 0$, for every $D_W$-neighbourhood (if $d \geq 2$) or $D_{KS}$-neighborhood $U_1$ of $\theta_0$ (if $d=2$) and $L_1$-neighborhood $U_2$ of $(\sigma_0, \mu_0)$.

4. Empirical Bayes inference for misspecified max-stable models. Consistency of a Bayesian procedure requires the posterior distribution to allow increasingly accurate inferences on the true parameter under study, as the sample size grows larger. In infinite-dimensional settings, the related asymptotic theory is established mostly in the case of well specified statistical models, where the notion of posterior consistency is formalised as done, for example, in Definition 3.1 for simple max-stable data. In statistical applications, max-stable models are fit to data that are samples of maxima, whose distribution is therefore not exactly max-stable. Nevertheless, under regularity conditions (see Condition 4.1 below) the distribution which truly generates each observation can be approximated by a max-stable one. In this setting, the use of a max-stable family for statistical inference based on maxima computed over blocks of finite size leads to a misspecified model. However, for increasingly larger blocks the misspecification becomes less and less an issue. To the best of our knowledge, there is no study offering a rigorous mathematical justification to Bayesian inference on max-stable models when the data sample consists of maxima, following only approximately a max-stable distribution. In particular, there is no indication on genuinely Bayesian prior specifications that appropriately takes into account model misspecification by adapting to the size of the blocks. The goal of this section is to provide conditions under which a quasi-Bayesian approach produces posterior-like distributions which concentrate near the appropriate max-stable density and leads to Hellinger consistent estimation of the true probability density of maxima. We resort to the use of a data-dependent prior, capitalising on the existing literature on frequentist inference on max-stable distributions, and propose guidelines for an empirical Bayesian approach. We develop new asymptotic techniques stemming from remote contiguity [47], devised to bridge the limiting statistical model and the actual data generating distribution. Such techniques are of independent interest for further applications to misspecified models [48], also beyond the extreme values context.

4.1. Empirical Bayes analysis of maxima. The posterior distributions $\Pi_n$ and $\Pi_n$ considered so far are conditional pm’s depending on a dataset coming from a max-stable model with true unknown marginal and dependence parameters. We now assume that the data sample consists of $n$ rv’s of componentwise maxima

$$M_{m_n,i} = \max(Z_{(i-1)m_n+1}, \ldots, Z_{im_n}), \quad i = 1, \ldots, n,$$

obtained by dividing a sample $(Z_1, \ldots, Z_{nm_n})$ of $nm_n$ iid rv’s with distribution $F_0$ into blocks of size $m_n$. Consequently, $M_{m_n,1}, \ldots, M_{m_n,n}$ are iid rv’s with distribution $F_0^{m_n}$.
Moreover, we assume the block size increases with \( n \), i.e. \( m_n \to \infty \) as \( n \to \infty \), which avoids considerations about double limits. Finally, we assume that \( F_0 \) is in the variational max-domain of \( G_{\theta_0}(\cdot|\theta_0) \) that is

\[
\lim_{n \to \infty} \mathcal{D}_T \left( F_0^{m_n}(a_{m_n} \cdot + b_{m_n}), G_{\theta_0}(\cdot|\theta_0) \right) = 0,
\]

for suitable norming sequences \( a_{m_n} \) and \( b_{m_n} \) that leads to a zero-location and unit-scale limiting distribution. Precisely, we assume that the following conditions hold true: by Corollary 3.1 in [33], they ensure that the limiting relation in (17) is satisfied.

\begin{align*}
\text{CONDITION 4.1.} \quad & \text{Let } F_0 \in \mathcal{D}(G_{\theta_0}(\cdot|\theta_0)), \text{ where } F_0(z) = C_0(F_{0,1}(z_1), \ldots, F_{0,d}(z_d)) \text{ and } G_{\theta_0}(\cdot|\theta_0) \text{ belongs to one of the three model classes in Definition 2.1. Suppose that } \\
& \theta_0 \in \Theta_0 \text{ and:} \\
& (i) \ C_0 \text{ is a } d\text{-times continuously differentiable copula on } (0,1)^d \text{ satisfying} \\
& \lim_{n \to \infty} \frac{\partial |I|}{\partial x_i, i \in I} m_n \left( C_0 \left( 1 - \frac{x}{m_n} \right) - 1 \right) = -L_I(x|\theta_0), \quad x > 0,
\end{align*}

for all \( I \subset \{1, \ldots, d\} \); (ii) \( F_{0,j}, j \in \{1, \ldots, d\} \), are continuously differentiable and satisfy one of the following

\[
\begin{align*}
& \lim_{z \to \infty} \frac{z F_{0,j}'(z)}{1 - F_{0,j}(z)} = \rho_{0,j}, \\
& \lim_{z \to z_{0,j}} \frac{(z_{0,j} - z) F_{0,j}'(z)}{1 - F_{0,j}(z)} = -\omega_{0,j}, \\
& \lim_{z \to z_{0,j}} \frac{F_{0,j}'(z)}{(1 - F_{0,j}(z))^2} \int_{z_{0,j}}^{20,j} (1 - F_{0,j}(t)) \, dt = 1,
\end{align*}
\]

where \( z_{0,j} := \sup \{ z \in \mathbb{R} : F_{0,j}(z) < 1 \} \). When \( G_{\theta_0}(\cdot|\theta_0) \) is multivariate \( \rho\)-Fréchet we also assume \( \text{supp}(F_{0,j}) \subset (0, \infty), j = 1, \ldots, d \), without loss of generality.

The notion of variational max-domain involves a stronger form of convergence than that exploited by the traditional max-domain, given in the first formula of Section (2.2), which only requires

\[
\lim_{n \to \infty} m_n \left( C_0 \left( 1 - \frac{x}{m_n} \right) - 1 \right) = -L(x|\theta_0), \quad x > 0,
\]

along with mild conditions for the convergence of the marginal distributions [23, Ch. 1]. These are indeed weaker requirements than Condition 4.1(i) and 4.1(ii). Notably, our setup guarantees that the probability density of the rv’s

\[
\bar{M}_{m_n,i} := (M_{m_n,i} - b_{m_n})/a_{m_n}, \quad i = 1, \ldots, n,
\]

is in a Hellinger-neighborhood of \( g_{\theta_0}(\cdot|\theta_0) \), for all sufficiently large block sizes \( m_n \). Without loss of generality, in the sequel we consider the valid choices of \( a_{m_n} \) and \( b_{m_n} \) given by

\[
(a_{m_n,j}, b_{m_n,j}) = \begin{cases}
(F_{0,j}^- (1 - 1/m_n), 0) \\
(z_{0,j} - F_{0,j}^- (1 - 1/m_n), z_{0,j}) \\
(m_n \int_{z_{0,j}}^{20,j} (1 - F_{0,j}(z)) \, dz, F_{0,j}^- (1 - 1/m_n))
\end{cases}
\]

and \( j = 1, \ldots, d \), for the three limit classes \( \rho\)-Fréchet, \( \omega\)-Weibull, Gumbel, respectively. These are generally unknown, thus the sequence \((\bar{M}_{m_n,i})_{i=1}^n\) is not directly available for
approximate Bayesian inference on the limiting max-stable model. A common practice in extreme-value analysis is to fit a max-stable model directly to unnormalised maxima encompassing scale and location parameters, whose estimates ultimately absorb $a_{m_n}$ and $b_{m_n}$ [e.g., 26]. Following this approach, we consider the case where a misspecified semiparametric max-stable model as in Section 3.3 is fitted to sample maxima, but replace the prior $\Pi_\Theta$ on the finite dimensional model component with a data dependent prior sequence $\Psi_n$ of the following general form:

\[
\begin{align*}
&d\Psi_n(\theta) \propto \begin{cases} 
   & d\pi_{sh}(\rho) \times \prod_{j=1}^{d} \pi_{sc} \left( \frac{\sigma_j}{a_{m_n,j}} \right) \frac{d\sigma_j}{a_{m_n,j}}, \\
   & d\pi_{sh}(\omega) \times \prod_{j=1}^{d} \pi_{sc} \left( \frac{\sigma_{j-1}}{a_{m_n,j-1}} \right) \frac{d\sigma_j}{a_{m_n,j}}, \\
   & \prod_{j=1}^{d} \pi_{sc} \left( \frac{\sigma_j}{a_{m_n,j}} \right) \frac{d\sigma_j}{a_{m_n,j}} \times \prod_{j=1}^{d} \pi_{loc} \left( \frac{\mu_j-b_{m_n,j}}{a_{m_n,j}} \right) \frac{d\mu_j}{a_{m_n,j}},
\end{cases}
\end{align*}
\]

for the three limit classes $\rho$-Fréchet, $\omega$-Weibull, Gumbel respectively, where $\hat{a}_{m_n} = \left(\hat{a}_{m_n,1}, \ldots, \hat{a}_{m_n,d}\right)$ and $\hat{b}_{m_n} = \left(\hat{b}_{m_n,1}, \ldots, \hat{b}_{m_n,d}\right)$ are estimators of $a_{m_n}$ and $b_{m_n}$, $\Pi_0$ is a pm with full support on a suitable subset of $(0, \infty)^d$ and $\pi_{sc}$ and $\pi_{loc}$ are Lebesgue probability densities whose properties are made precise in the following shares. Since priors on scale and location parameters should now incorporate information on the norming sequences, which is typically not available a priori, a genuinely subjective specification is hardly viable. In such a case, a data driven prior selection, also known as empirical Bayes, is a popular approach in Bayesian analysis [e.g., 62, Sections 1 and 3]. The data-dependent prior $\Psi_n$, combined with the misspecified max-stable likelihood, results into a quasi-posterior distribution, defined via

\[
\Pi^{(o)}_n(B) := \frac{\int_B \prod_{i=1}^{n} g_{\theta}(M_{m_n,i}|\theta)d(\Pi_\Theta \times \Psi_n)(\theta, \theta)}{\int_{\Theta \times \Theta} \prod_{i=1}^{n} g_{\theta}(M_{m_n,i}|\theta)d(\Pi_\Theta \times \Psi_n)(\theta, \theta)},
\]

for all $\Pi_\Theta \times \Psi_n$-measurable sets $B$. The superscript $(o)$ denotes dependence on observables.

**Remark 4.2.** The one above is a valid random pm but not precisely a bona fide posterior distribution (then called quasi-posterior), since it is obtained by plugging a data-dependent prior into Bayes formula. Unlike a genuinely Bayesian posterior it does not allow then for the classical interpretation as a conditional distribution of the parameters, given the data. A proper posterior would be achieved recovering the statistics $\hat{a}_{m_n}$ and $\hat{b}_{m_n}$ on the basis of a different dataset than the one used for posterior computation. However, using a separate sample for estimating the marginal norming sequences would be infeasible in many applications, due to the scarcity of extreme data to also carry out inference on marginal tails. This justifies our choice in proposing an empirical Bayes prior distribution.

Empirical Bayes inference is carried out on the basis of quasi-posterior distributions. Consequently, we are interested in establishing asymptotic concentration properties of $\Pi^{(o)}_n$ over $\Theta \times \Theta$ and of its counterpart $\Pi^{(o)}_n$ over the corresponding class of max-stable densities $G_\Theta$. We point out that, under Condition 4.1, the true probability density of unnormalised maxima $f^{(o)}_{m_n}(x) := (\partial/\partial x) F^{(o)}_{m_n}(x)$ becomes topologically undistinguishable from the density of $G_{\theta_0}(\cdot - b_{m_n})/a_{m_n}|\theta_0)$ as $n \to \infty$. Thus, we ultimately aim at showing that $\Pi^{(o)}_n$ cumulates an increasingly large fraction of its total mass near the latter. Practical implications for prediction and quantile estimation are showcased in Section 5.

4.2. Reparametrisation, remote contiguity. To accomplish the objective above, we firstly provide an alternative representation of the quasi-posterior distributions under study. A
change of variables in the integrals in (21)

$$
(\theta, \vartheta) \mapsto \psi_n(\theta) = \left\{ \begin{array}{ll}
(\theta, \rho, \sigma/a_{m_n}), \\
(\theta, \omega, \sigma/a_{m_n}, \{\mu - b_{m_n}\}/a_{m_n}), \\
(\theta, \sigma/a_{m_n}, \{\mu - b_{m_n}\}/a_{m_n}), 
\end{array} \right.
$$

for the three limit classes $\rho$-Fréchet, $\omega$-Weibull, Gumbel respectively, corresponding to a change of parametrisation, yields the equality $\Pi_n = \Pi_n^{(o)} \circ \psi_n^{-1}$, where $\Pi_n$ is the quasi-posterior defined via

$$
\Pi_n(B) := \frac{\int_B \prod_{i=1}^{n} g_{\theta}(\mathbf{M}_{m_i}^n|\theta)d(\Pi_{\Theta} \times \Psi_n)(\theta, \vartheta)}{\int_{\theta \times \Theta} \prod_{i=1}^{n} g_{\theta}(\mathbf{M}_{m_i}^n|\theta)d(\Pi_{\Theta} \times \Psi_n)(\theta, \vartheta)},
$$

for every $\Pi_{\Theta} \times \Psi_n$-measurable set $B$, where $\Psi_n = \psi_n \circ \psi_n^{-1}$. The latter also induces a quasi-posterior $\tilde{\Pi}_n$ on a corresponding class of max-stable densities, which is in turn linked to $\Pi_n^{(o)}$ via the relation $\tilde{\Pi}_n = \tilde{\Pi}_n^{(o)} \circ \tilde{\psi}_n^{-1}$, where

$$
\tilde{\psi}_n(g_{\theta}(\theta)) = g_{\psi_n(\theta)}(\theta).
$$

Consequently, the required asymptotic analysis boils down to establishing consistency of $\Pi_n$ and $\tilde{\Pi}_n$ almost surely (or in probability) at $(\theta_0, \vartheta_0)$ and $g_{\theta_0}(\theta_0)$, respectively, with respect to $Q_n$, the joint pm of $(\mathbf{M}_{m_1}^n, \ldots, \mathbf{M}_{m_n}^n)$. We stress that herein $\Pi_n$ and $\tilde{\Pi}_n$ are mathematical devices introduced to enable the study of the asymptotic behaviour of $\Pi_n^{(o)}$ and $\tilde{\Pi}_n^{(o)}$, but are not practical for statistical inference, as they depend on unobservables. Under appropriate assumptions on the data dependent prior, $\Pi_n$ and $\tilde{\Pi}_n$ satisfy an inequality like (12). Therefore, consistency can be obtained by establishing a variant of (13) where the sample of normalised maxima $\mathbf{M}_{m_1}^n, \ldots, \mathbf{M}_{m_n}^n$ replaces the sample $X_1, \ldots, X_n$ and the joint distribution of the former, $Q_n$, replaces the joint distribution of the latter, $G^{(o)}_{\Theta_0}(\theta_0)$. To do this, we resort to a specific form of remote-contiguity.

Recently, [47] has introduced a generalised form of contiguity. The classical notion of contiguity can be successfully exploited to establish consistency in probability of the quasi-posterior distribution with parametric limiting models [e.g., 14], while is unsuitable to obtain almost sure consistency and less accessible for nonparametric models. See [47, Section 3] for a comprehensive account.

**Definition 4.3.** Consider two sequences $\nu_n, \tau_n > 0$ such that $\nu_n, \tau_n \to 0$ as $n \to \infty$. As $n \to \infty$, $Q_n$ is said to be:

i. contiguous with respect to $G^{(o)}_{\Theta_0}(\theta_0)$ if, for a sequence of measurable events $E_n$,

$$
G^{(o)}_{\Theta_0}(E_n|\theta_0) = o(1) \Rightarrow Q_n(E_n) = o(1);
$$

ii. $\nu_n$-remotely contiguous with respect to $G^{(o)}_{\Theta_0}(\theta_0)$, if $G^{(o)}_{\Theta_0}(E_n|\theta_0) = o(\nu_n) \Rightarrow Q_n(E_n) = o(1)$;

iii. $\nu_n$-to-$\tau_n$-remotely contiguous with respect to $G^{(o)}_{\Theta_0}(\theta_0)$, if $G^{(o)}_{\Theta_0}(E_n|\theta_0) = o(\nu_n) \Rightarrow Q_n(E_n) = o(\tau_n)$,

where “$\Rightarrow$” denotes the usual implication symbol.

Essentially, remote contiguity allows to assert that an event sequence $E_n$ becomes increasingly unlucky to take place under $Q_n$ provided it is so under the law $G^{(o)}_{\Theta_0}(\theta_0)$, while explicitly accounting for the probability decay rates. In particular, it allows to deduce “suitably fast” convergence in $Q_n$-probability to zero of nonnegative data functionals (e.g., stochastic upper bounds on quasi-posteriors) from “sufficiently quick” decay of their expectation under
we consider the scenario where a max-stable density as in (14) is fitted to student-distributions defined on the real line have symmetric tails [23, Ch. 1–2], such as Cauchy, allowing to lift the approximation in (17) to remote contiguity results for \( Q_n \).

**Condition 4.4.** Let \( f_m \) denote the density of \( F^{m,n}_0(\alpha_{m,n} \cdot + b_{m,n}) \), then assume that there exists \( n_0 \in \mathbb{N}_+ \) and \( J_0 \in (0, \infty) \) such that \( \sup_{n \geq n_0} \| f_m / g_{\theta_0}(\cdot | \theta_0) \|_\infty < J_0 \).

Conditions 4.1 and 4.4 imply that the Kullback-Leibler divergence from the true limiting max-stable density to the density of normalised maxima is asymptotically null and the positive Kullback-Leibler variations between the two up to the fourth order [e.g., 39, equation 4.5.].

**Proposition 4.5.** Let \( F_0 \) and \( G_{\theta_0}(\cdot | \theta_0) \) satisfy Conditions 4.1 and 4.4. Then \( Q_n \) is \( e^{-cn} \)-to-\( -n^{-1-c'} \)-remotely contiguous with respect to \( G_{\theta_0}(\cdot | \theta_0) \), \( \forall c > 0, \forall c' \in (0, 1) \).

4.3. Fréchet domain of attraction. We start by analysing consistency in the case where maxima are obtained from rv's whose distribution has heavy-tailed margins, which are therefore in the max-domain of attraction of a Fréchet distribution. We recall that in this case the upper end-point of each marginal distribution is infinity. Without loss of generality, we restrict our attention to non-negative random variables (Condition 4.1(ii)). Popular heavy-tailed distributions defined on the real line have symmetric tails [23, Ch. 1–2], such as Cauchy, Student-\( t \), etc., and can be casted in our framework by considering their folded version. Then, we consider the scenario where a max-stable density as in (14) is fitted to \( (M_{m,n,i})_{i=1}^n \) and a data-dependent prior on \((\rho, \sigma)\) is specified as in (19). Our results rely on the following conditions.

**Condition 4.6.** (i) \( \Pi_{sh} \) has full support on \((0, \infty)^d \).

(ii) The probability density function \( \pi_{sc} \) satisfies \( \{ x \in \mathbb{R} : \pi_{sc}(x) > 0 \} \subset (0, \infty) \); moreover, there exist \( \eta \in (0, 1) \) and a Lebesgue integrable continuous function \( u_{sc} : (0, \infty) \to (0, \infty) \) such that:

(iia) \( \pi_{sc} \) is continuous on \([1 \pm \eta]\) and \( \inf_{x \in [1 \pm \eta]} \pi_{sc}(x) > 0 \);

(iiib) \( \sup_{t \in (1 \pm \eta)} \pi_{sc}(x / t) \leq u_{sc}(x) \), for all \( x > 0 \).

(iii) \( \lim_{n \to \infty} \tilde{a}_{m,n,j} / a_{m,n,j} = 1, \quad j = 1, \ldots, d, \quad F^{(\infty)}_0 - \text{as.} \)

Condition 4.6(ii) is very mild. It is satisfied by regular density functions on the positive half-line (or suitable subsets), such as gamma, Pareto, half-Cauchy, etc. The proof of the following theorem is deferred to Section D.6.2 of the supplement.

**Theorem 4.7.** Let \( M_{m,n,1}, \ldots, M_{m,n,n} \) be iid according to \( F^{m,n}_0 \). Assume \( F_0 \) and \( G_{\rho_0}(\cdot | \theta_0) \) satisfy Conditions 4.1 and 4.4. Then, under Conditions 3.13(iii) and 4.6, \( \Pi_n \) and \( \Pi_{sh} \) satisfy the properties at points (a)-(b) of Theorem 3.14, \( F^{(\infty)}_0 - \text{as and:} \)

(a') \( \lim_{n \to \infty} \Pi_n^{(\infty)}(U_n^r) = 0 \), for every sequence of \( \mathcal{D}_H \)-balls \( U_n \) centred at \( g_{\rho_0,a_{m,n}}(\cdot | \theta_0) \) of positive fixed radius;
(b') \( \lim_{n \to \infty} \Pi_n^0((\mathcal{U}_1 \times \mathcal{U}_{2n})^c) = 0 \), for every \( \mathcal{D}_W \)-neighborhood (if \( d \geq 2 \)) or \( \mathcal{D}_{KS} \)-neighborhood \( \mathcal{U}_1 \) of \( \theta_0 \) (if \( d=2 \)) and every sequence of rectangles \( \mathcal{U}_{2,n} = (\rho_0 \pm 1\varepsilon) \times (a_m (1 - \epsilon), a_m (1 + \epsilon)), \varepsilon > 0 \).

**Remark 4.8.** The weak convergence result in the first formula of Section (2.2) is often used as a justification for the following somewhat informal approximation for \( i = 1, \ldots, n \), as \( n \to \infty \),

\[
\mathbb{P}(M_{m,i} \leq x) \approx G_{\rho_0,1}(x/a_m | \theta_0) = G_{\rho_0,a_m}(x | \theta_0),
\]

and, hence, for using a family of max-stable distributions as an observational model for (unnormalised) sample maxima [e.g., 15, p. 48]. From a Bayesian perspective, Theorem 4.7(a') provides a rigorous mathematical justification for such a practice under the strong domain of attraction \( (17) \), as it ensures that the pseudo-posterior distribution \( \Pi_n^0 \) asymptotically concentrates on a Hellinger-neighbourhood of \( g_{\rho_0,a_m}(\cdot | \theta_0) \), the density of \( G_{\rho_0,a_m}(\cdot | \theta_0) \). As for the finite dimensional model components, Theorem 4.7(b') establishes that the marginal pseudo-posterior of the shape and scale parameters concentrates on a set of \( (\rho, \sigma) \) such that \( |\rho_j - \rho_{0,j}| \) is small and \( \sigma_j/a_{m,j} \) is close to one for any \( j \in \{1, \ldots, d\} \), reminiscing the behaviour of the MLE in the frequentist approach [26, Theorem 2].

**Remark 4.9.** Under Condition 4.6, the data dependent prior \( \overline{\pi}_n \), obtained via the reparametrisation in (20), is positive on a neighbourhood of the true limiting parameter \( (\theta_0, \rho_0, 1) \). This property is crucial for obtaining consistency of the quasi-posteriors \( \hat{\pi}_n \), which may fail in absence of the empirical Bayes rescaling in (19) (first line), through estimators \( \hat{a}_{m,j} \) complying with Condition 4.6(iii). Indeed, the norming sequence \( a_{m,j}, j = 1, \ldots, d, \) in (18) diverges to infinity as \( n \to \infty \). Therefore, a prior density \( \pi_{sc} \) specified on each original scale parameter with no scale correction is after reparametrisation asymptotically null at the true value, since the integrability of \( \pi_{sc} \) requires \( \pi_{sc} (a_{m,j} \mid \theta_0) \to 0 \).

**Example 4.10.** Let \((Z_1, \ldots, Z_{nm})\) be an observable sample. For each \( j = 1, \ldots, d \), choose \( \hat{a}_{m,j} = \hat{F}_{n,j}^{-1}(1 - 1/m_n) \), where \( \hat{F}_{n,j}(x) = n^{-1} \sum_{i=1}^{nm} I(Z_{i,j} \leq x) \). Then, we have that for all \( \epsilon > 0 \), as \( n \to \infty \),

\[
\hat{F}_{n,j}^{nm}(|\hat{a}_{m,j} / a_{m,j} - 1| > \epsilon) \leq 4e^{-\tau_j \sqrt{n+1}}, \quad j = 1, \ldots, d,
\]

where \( \tau_j \equiv \tau_j(\epsilon) \) are positive constants. A proof of the above inequality is provided in Section E.1.1 of the supplementary material. Thus, by Borel-Cantelli lemma, \( \hat{a}_{m,j} \) complies with Condition 4.6(iii).

**Remark 4.11.** If the limiting relations in Condition 4.6(iii) hold true in probability rather than almost surely, the asymptotic results in Theorem 4.7 are obtained then in probability. Examples of estimators \( \hat{a}_{m,j} \) complying with such a weaker requirement can be obtained from those in [23, pp. 130–131] and [10].

Finally, we provide two examples of bivariate models complying with Conditions 4.1 and 4.4. Full derivations are provided in Sections E.1.2-E.1.3 of the supplement.

**Example 4.12.** Consider \( F_0(x) = 1 - 1/x_1^{\rho_{0,1}} - 1/x_2^{\rho_{0,2}} + 1/(x_1^{\rho_{0,1}} + x_2^{\rho_{0,2}} - 1) \), for \( x > 1 \), which is a slightly more general version of the distribution examined in [65, p. 289]. The distribution \( F_0(x) \) has Pareto margins, i.e. \( F_{0,j}(x_j) = 1 - 1/x_j^{\rho_{0,j}}, j = 1, 2 \), and belongs to the variational max-domain of

\[
G_{\rho_0,1}(x | H_0) = \exp \left( -x_1^{-\rho_{0,1}} - x_2^{-\rho_{0,2}} + (x_1^{\rho_{0,1}} + x_2^{\rho_{0,2}})^{-1} \right), \quad x > 0,
\]

whose angular distribution \( \theta_0 \) is the uniform on \([0, 1]\).
EXAMPLE 4.13. Consider
\[ F_0(x) = 1 - \left( x_1^{-3\rho_{0,1}} + x_2^{-3\rho_{0,2}} - (x_1^{\rho_{0,1}}x_2^{\rho_{0,2}})^{-3} \right)^{1/3}, \quad x > 1. \]
This distribution has Pareto margins and Joe/B5 copula [e.g., 45, p. 170]; it belongs to the variational max-domain of
\[ G_{\rho_{0,1}}(x|\theta_0) = \exp \left( - \left( x_1^{-3\rho_{0,1}} + x_2^{-3\rho_{0,1}} \right)^{1/3} \right), \quad x > 0, \]
whose extreme-value copula is a member of the so-called logistic family [e.g., 45, p. 172], with angular density \[ \hat{\theta}_0(t) = t(1-t)\{t^3 + (1-t)^3\}^{-5/3}, \] for \( t \in (0,1) \).

4.4. Gumbel domain of attraction. In this subsection we establish results similar to those in Theorem 4.7 for the case where the marginal distributions of \( F_0 \) belong to the Gumbel max-domain of attraction. In this setup the max-stable density in (16) provides a misspecified statistical model for \( M_{m_1,n}, \ldots, M_{m_n,n} \). We assume that a data-dependent prior is assigned to \( (\sigma, \mu) \), with prior densities \( \pi_{sc} \) and \( \pi_{loc} \) and estimators \( \hat{a}_{m_n} \) and \( \hat{b}_{m_n} \) complying with the following assumptions.

CONDITION 4.14. \( \pi_{sc} \) and \( \hat{a}_{m_n} \) satisfy Conditions 4.6(ii) and 4.6(iii), respectively, and:
(i) there exist \( \eta \in (0,1) \) and a Lebesgue integrable continuous function \( u_{loc} : \mathbb{R} \to (0, \infty) \) such that:
\( (ia) \) \( \pi_{loc} \) is continuous on \([-\eta, +\eta]\) and \( \inf_{x \in [-\eta, +\eta]} \pi_{loc}(x) > 0; \)
\( (ib) \) \( \sup_{t_1 \in (1\pm \eta), t_2 \in (-\eta, +\eta)} \pi_{loc}((x - t_2)/t_1) \leq u_{loc}(x), \) for all \( x > 0. \)
(ii) \( \lim_{n \to \infty} (\hat{b}_{m_{n,j}} - b_{m_{n,j}})/a_{m_{n,j}} = 0, \quad j = 1, \ldots, d, \quad F_{0}(\infty) \) as.

THEOREM 4.15. Let \( M_{m_1,n}, \ldots, M_{m_n,n} \) be iid according to \( F_{0,m_n} \), let \( F_0 \) and \( G_{1,0}(\cdot|\theta_0) \) satisfy Conditions 4.1 and 4.4. Then, under Conditions 3.13(iii) and 4.14, \( \hat{\Pi}_n \) and \( \Pi_n \) satisfy the properties at points (a)–(b) of Theorem 3.17 \( F_{0}(\infty) \) as and:
\( (a') \) \( \lim_{n \to \infty} \hat{\Pi}_n^{(\infty)}(U_n^c) = 0, \) for every sequence of \( \mathcal{D}_H \)-balls \( U_n \) centered at \( g_{m_n,b_{m_n}} (\cdot|\theta_0) \) of positive fixed radius;
\( (b') \) \( \lim_{n \to \infty} \Pi_n^{(\infty)}(U_1 \times U_2,2^c) = 0, \) for every \( \mathcal{D}_W \)-neighborhood (if \( d \geq 2 \)) or \( \mathcal{D}_{KS} \)-neighborhood \( U_1 \) of \( \theta_0 \) (if \( d=2 \)) and every sequence of rectangles \( U_{2,n} = (a_{m,n}(1-\epsilon), a_{m,n}(1+\epsilon)) \times (b_{m,n} \pm \epsilon a_{m,n}), \epsilon > 0. \)

The proof of Theorem 4.15 is similar to that of Theorem 4.7 and is therefore omitted. The main points of Remarks 4.8, 4.9 and 4.11 carry over to the present case and the asymptotic conditions imposed to \( \hat{a}_{m_n} \) and \( \hat{b}_{m_n} \) are satisfied by empirical estimators, as illustrated by the following example.

EXAMPLE 4.16. For \( j \in \{1, \ldots, d\} \), define \( \hat{F}_{nm_{n,j}} \) as in Example 4.10 and set
\[ \hat{b}_{m_{n,j}} = \hat{F}_{nm_{n,j}}^{-1} \left( 1 - \frac{1}{m_n} \right), \quad \hat{a}_{m_{n,j}} = m_n \int_{b_{m_{n,j}}}^{\infty} (1 - \hat{F}_{nm_{n,j}}(x))dx. \]
To satisfy Conditions 4.6(iii) and 4.14(ii) it is sufficient that the following requirements are met: there exists \( (\gamma_j, \alpha_j) \in (0, \infty)^2 \) for every \( j = 1, \ldots, d \) such that
\[ \int \exp(\gamma_j |z_j^{\alpha_j}) F_0,j(dz) < \infty \]
and for some $s \in (0, \min\{1, \alpha_1, \ldots, \alpha_d\})$, as $n \to \infty$,

\begin{equation}
\log n = \begin{cases}
o\left(na_{m_n,j}^2/m_n\right), & \text{if } \alpha_j \geq 1, \\
o\left(\min\{na_{m_n,j}^2/m_n, a_{m_n,j}^s n^s\}\right), & \text{if } \alpha_j < 1.
\end{cases}
\end{equation}

A proof is provided in Section E.2.1 of the supplement, where we exploit the concentration inequalities for the 1-Wasserstein distances

\[
\int |F_{0,j}(z) - \hat{F}_{n,m_{n,j}}(z)|\,dz, \quad j = 1, \ldots, d,
\]

recently derived by [34] under the exponential moment condition (23). The latter is satisfied by all the most common univariate distributions in the Gumbel max-domain of attraction, such as the exponential, Gaussian, Beta and log-normal distributions. Moreover, for each $j \in \{1, \ldots, d\}$, the map $m_n \mapsto a_{m_n,j}$ in the third line of (18) is slowly varying at infinity, then (24) is easily obtained, for example, by requiring that $m_n \sim n^t$ as $n \to \infty$, with $t \in (0, 1)$.

We complete this section with an example of a bivariate distribution belonging to the Gumbel max-domain of attraction that satisfies Conditions 4.1 and 4.4. This is formally verified in Section E.2.2 of the supplement.

**Example 4.17.** Consider the bivariate exponential distribution [e.g., 55], i.e.

\[F_0(x) = 1 - e^{-x_1} - e^{-x_2} + (e^{x_1} + e^{x_2} - 1)^{-1}, \quad x > 0,\]

with exponential margins $F_{0,j}(x_j) = 1 - e^{-x_j}$, $j = 1, 2$. This distribution belongs to the variational max-domain of

\[G_{1,0}(x|\theta_0) = \exp\left\{-e^{-x_1} - e^{-x_2} + (e^{x_1} + e^{x_2} - 1)^{-1}\right\}, \quad x \in \mathbb{R}^d.\]

Herein, the distributions $F_0$ and $G_{1,0}(x|\theta_0)$ have the same copulas of the distributions in Example 4.12.

4.5. **Reverse Weibull domain of attraction.** We finally study consistency in the case where a sample of maxima is obtained from random variables with distribution in the Weibull max-domain of attraction. We consider a misspecified statistical model given by the max-stable density (15), with $(\omega, \sigma, \mu)$ in the restricted space $\Theta = (1, \infty)^d \times (0, \infty)^d \times \mathbb{R}^d$. For technical convenience, we assume the following conditions on the support of the prior distributions in (19).

**Condition 4.18.** \(\pi_{sc}\) and \(\pi_{loc}\) satisfy Conditions 4.6(ii) and 4.14(i), respectively, and:

(i) there exist closed and bounded intervals $I_{sc} \subset (0, \infty)$ and $I_{loc} \subset \mathbb{R}$ such that \(\{x \in \mathbb{R} : \pi_{sc}(x) > 0\} = I_{sc}\) and \(\{x \in \mathbb{R} : \pi_{loc}(x) > 0\} = I_{loc}\);

(ii) there exists a closed and bounded rectangle $K_{sh} \subset (1, \infty)^d$ such that \(\text{supp}(\Pi_{sh}) = K_{sh}\)

and \(\Pi_{sh}\) assigns positive mass to every neighbourhood of $\omega_0$;

(iii) in $F_0^{(m_{ma})}$-probability as $n \to \infty$, $\tilde{a}_{m_n,j}/a_{m_n,j} \to 1$, and $(\tilde{b}_{m_n,j} - b_{m_n,j})/a_{m_n,j} \to 0$, for $j = 1, \ldots, d$.

**Theorem 4.19.** Let $M_{m_n,1}, \ldots, M_{m_n,d}$ be iid according to $F_0^{m_n}$. Let $F_0$ and $G_{\omega_0,1,0}(-|\theta_0)$ satisfy Conditions 4.1 and 4.4. Then, under Conditions 3.13(iii) and 4.18, \(\bar{\Pi}_n\) and \(\Pi_n\) satisfy the properties at points (a)–(b) of Theorem 3.15 in $F_0^{(m_{ma})}$-probability as $n \to \infty$ and:
(a') \( \lim_{n \to \infty} \bar{\Pi}^{(\omega)}(U_{n}) = 0 \), for every sequence of \( \mathcal{D}_{H} \)-balls \( U_{n} \) centred at \( g_{\omega,n}, a_{m,n}, b_{m,n} \) of positive fixed radius;
(b') \( \lim_{n \to \infty} \Pi^{(\omega)}((U_{1} \times U_{n}, 2)^{\omega}) = 0 \), for every \( \mathcal{D}_{W} \)-neighborhood (if \( d \geq 2 \)) or \( \mathcal{D}_{KS} \)-neighborhood \( U_{1} \) of \( \theta_{0} \) (if \( d = 2 \)) and every sequence of rectangles \( U_{2,n} = (a_{m,n}(1 - \epsilon), a_{m,n}(1 + \epsilon)) \times (b_{m,n} - \epsilon a_{m,n}, b_{m,n}) \), \( \epsilon > 0 \).

Theorem 4.19 provides convergence results which are valid in probability (see Section D.7.2 of the supplement for a proof). In order to obtain stronger forms of convergence as in Theorems 4.7 and 4.15, the moments of the positive part of the pseudo log likelihood ratios

\[
\log \{ g_{\omega,1,0}(M_{m,n,i}\mid \theta_{0})/g_{\omega,\sigma,\mu}(M_{m,n,i}\mid \theta) \},
\]

\( i = 1, \ldots, n \), have to be bounded uniformly over a \( (H, \omega, \sigma, \mu) \)-set of positive prior mass. For multivariate \( p \)-Fréchet and Gumbel models, the tractability of such quantities is guaranteed by the assumption \( \theta_{0} \in \Theta_{0} \), along with Condition 3.13 (iii). In the \( \omega \)-Weibull model, the latter only guarantee control over the first moments, while the higher order ones are less tractable unless further restrictions are imposed on the shape parameters. See Section D.4.4 of the supplement for details. Accordingly, we only focus on estimators \( \tilde{a}_{m,n}, \tilde{b}_{m,n} \) satisfying the convergence properties in Condition 4.18(iii) in probability, an example of which is provided next.

**Example 4.20.** Consider an iid observable sample \( Z_{1}, \ldots, Z_{nm,n} \) from \( F_{0} \). According to [23, Ch. 4.2 and Ch. 4.5], estimators \( \tilde{a}_{m,n}, \tilde{b}_{m,n} \) complying with Condition 4.18(iii) can be constructed by selecting

\[
\tilde{a}_{m,n,j} = Z_{n(m_{n}-1):nm_{n},j}\xi_{n,j}^{(1)}(1 - \gamma_{n,j})/(-\gamma_{n,j}), \quad j = 1, \ldots, d,
\]

and \( \tilde{b}_{m,n,j} = Z_{n(m_{n}-1):nm_{n},j} + \tilde{a}_{m,n,j} \), for \( j = 1, \ldots, d \), where \( Z_{s:mm,n,j} \) is the \( s \)-th order statistic of the marginal sample \( Z_{1,j}, \ldots, Z_{nm,n,j} \) and

\[
\text{\( \gamma_{n,j} = 2^{-1} \left( \xi_{n,j}^{(2)} - 2 \left( \xi_{n,j}^{(1)} \right)^{2} \right) \left( \xi_{n,j}^{(2)} - \left( \xi_{n,j}^{(1)} \right)^{2} \right)^{-1} \),}
\]

with

\[
\xi_{n,j}^{(l)} = \frac{1}{n} \sum_{i=1}^{n-1} \left( \log Z_{nm,n-i:mm,n,j} - \log Z_{n(m_{n}-1):mm,n,j} \right)^{l}, \quad l = 1, 2,
\]

and \( \tilde{\gamma}_{n,j} \) is an estimator satisfying \( \tilde{\gamma}_{n,j} = -1/\omega_{0,j} + o_{p}(1) \) [see 23, Ch. 3 for examples].

**5. Predictive consistency.** Predicting future observations is the central goal of several statistical applications concerning extreme value data. Quantifying the uncertainty associated with the forecast of future extremes is a crucial and difficult task, thus probabilistic forecasting of predictive density form [40] appears particularly attractive to extreme-value analysis. Given a sample of size \( n \), the Bayesian paradigm offers a direct approach to probabilistic forecasting of the \( (n + 1) \)-th observation via the posterior predictive density, defined pointwise as the mean of the posterior distribution of the data generating density. The true data generating density coincides with the true (unknown) predictive density of the next observation, when observations are i.i.d. Thus, Hellinger consistency of the posterior predictive density is a strong large-sample guarantee on the quality of prediction. Leveraging on the asymptotic theory of Sections 3.3, we establish such a property for posterior predictive densities stemming from well-specified max-stable models. We then extend consistency to quasi-posterior
predictive densities attained by the empirical Bayes analysis of sample block maxima (see Section 4). Proofs and technicalities are deferred to Sections D.8 and E.3 of the supplement.

For a sample of iid rv’s $X_1, \ldots, X_n$ following a multivariate max-stable distribution $G_{\theta_0}(\cdot|\theta_0)$, the posterior predictive density of the next observation $X_{n+1}$ is

$$\hat{g}_n(x) := \int_{\Theta} g(x|\theta) d\Pi_n(\theta) = \int_{\Theta} g(\theta) d\Pi_n(\theta|\theta_0),$$

where $x \in \mathbb{R}^d$ (note that the posterior distributions $\Pi_n$ and $\tilde{\Pi}_n$, arising from the prior distributions $\Pi_{\Theta \times \Theta}$ and $\Pi_{\Theta \times \Theta}$, are linked via $\tilde{\Pi}_n = \Pi_n \circ \varphi_{\Theta \times \Theta}^{-1}$, see Section 3.3). Its distribution $\hat{G}_n$ is the distribution of $X_{n+1}$ conditionally on $X_1, \ldots, X_n$, from a genuinely subjective viewpoint. It is then parameters free, as the latter are averaged out with respect to the posterior distribution. Thus, $\hat{g}_n$ represents an estimator of the true predictive density $g_{\theta_0}(\cdot|\theta_0)$ relative to the next observation $X_{n+1}$, from a frequentist perspective. Posterior predictive density consistency can be deduced from the asymptotic concentration properties established in Theorems 3.14, 3.15 and 3.17.

**Corollary 5.1.** Under the assumptions of Theorem 3.14 or 3.15 or 3.17,

$$\lim_{n \to \infty} \mathcal{D}_H(\hat{g}_n, g_{\theta_0}(\cdot|\theta_0)) = 0, \quad G_{\theta_0}^{(\infty)}(\cdot|\theta_0) - \text{as.}$$

The practically most relevant situation is when the observables are actually iid block maxima rv’s, whose underlying distribution is in the max-domain of attraction of $G_{\theta_0}(\cdot|\theta_0)$. In this case, prediction can be achieved by exploiting the empirical-Bayes approach outlined in Section 4.1, yielding the quasi-posterior predictive density

$$\hat{g}_n^{(o)}(x) := \int_{\Theta} g(x|\theta) d\Pi_n^{(o)}(\theta) = \int_{\Theta} g(\theta) d\Pi_n^{(o)}(\theta),$$

where $\Pi_n^{(o)}$ and $\tilde{\Pi}_n^{(o)}$ are quasi-posterior distributions (see Remark 4.2). Although the corresponding distribution cannot be subjectively interpreted as a conditional distribution, $\hat{g}_n^{(o)}$ is a Hellinger consistent estimator of the true predictive density $f_{m_n}^{(o)}$, under the assumptions of Theorem 4.7, 4.15 and 4.19.

**Corollary 5.2.** Under the assumptions of Theorem 4.7 or 4.15,

$$\lim_{n \to \infty} \mathcal{D}_H(\hat{g}_n^{(o)}, f_{m_n}^{(o)}) = 0, \quad F_{0}^{(\infty)} - \text{as.}$$

Under the assumptions of Theorem 4.19, $\mathcal{D}_H(\hat{g}_n^{(o)}, f_{m_n}^{(o)}) \to 0$ in $F_{0}^{(\infty)}$-probability as $n \to \infty$.

In the univariate case, predictive distributions can be used for extreme quantile and return level estimation (see [16] for an early discussion of predictive return levels). In the multivariate case, the identification of extreme subsets of the sample space, where future high dimensional observations ($d \geq 2$) occur with low probability $p \in (0, 1)$, is really a more challenging problem. These are generically referred to as extreme quantile regions. Since there is no univocal extension of real-valued quantiles to manifolds on $\mathbb{R}^d$, the latter being only a partially ordered space, several quantile regions definitions have been proposed in the literature. We next focus on a definition based on density sublevel sets, mutated from [30]. We then showcase a possible use of $\hat{g}_n^{(o)}$ for inferring quantile regions, pinpointing Hellinger consistency implications on the asymptotic accuracy of the resulting procedure.
EXAMPLE 5.3. Let $S_{m_n} := \text{supp}(F_{0m_n})$ and denote by $\hat{S}_{m_n}$ its interior. Assume Condition 4.1 holds true and

\begin{align}
\text{Leb}_{S_{m_n}}\left(\{x \in \hat{S}_{m_n} : f_{m_n}^{(o)}(x) = \alpha\}\right) &= 0, \quad \forall \alpha \in [0, \infty), \\
\text{Leb}_{S_{m_n}}(S_{m_n} \setminus \hat{S}_{m_n}) &= 0.
\end{align}

Then, an adaptation of Proposition 2.1 in [30] guarantees that for any $p \in (0, 1)$ there exists $\alpha_n \equiv \alpha_n(p)$ such that, defining

$$Q_n := \{x \in \hat{S}_{m_n} : f_{m_n}^{(o)}(x) \leq \alpha_n\},$$

the set $Q_n^c = S_{m_n} \setminus Q_n^c$ is the one with the minimum Leb$_{S_{m_n}}$-volume among those satisfying $F_{0m_n}(Q^n) = 1 - p$. Specifically, for any other Borel subset $B \subset S_{m_n}$ such that Leb$_{S_{m_n}}(B) = \text{Leb}_{S_{m_n}}(Q_n^c)$ and $F_{0m_n}(B) = 1 - p$, it must be that

$$\text{Leb}_{S_{m_n}}(B \triangle Q_n^c) = 0.$$

If for every $n$ the quasi-posterior predictive density $\hat{g}_n^{(o)}$ is continuous and satisfies properties analogous to (26)-(27) with probability 1, for a given $p \in (0, 1)$ there also exists $\hat{\alpha}_n$ such that

$$\hat{Q}_n := \{x \in \hat{S}_{m_n} : \hat{g}_n^{(o)}(x) \leq \hat{\alpha}_n\}$$

satisfies $\hat{G}_n^{(o)}(\hat{Q}_n) = \hat{\alpha}_n(\hat{S}_{m_n} \setminus \hat{Q}_n) = 1 - p$, where $\hat{S}_{m_n} := \text{supp}(\hat{G}_n^{(o)})$. Intuitively, $\hat{Q}_n$ represents an estimator of the $p$-quantile region $Q_n$, obtained by plugging $\hat{g}_n^{(o)}$ and $\hat{\alpha}_n$ in place of $f_{m_n}^{(o)}$ and $\alpha_n$ into (28). Almost sure Hellinger consistency of $\hat{g}_n^{(o)}$ as in (25) now implies that

$$\lim_{n \to \infty} F_{0m_n}(Q_n \triangle \hat{Q}_n) = 0, \quad F_{0}^{(\infty)} - \text{as.}$$

An analogous convergence is obtained in probability, if $\hat{g}_n^{(o)}$ is Hellinger consistent in probability. As a consequence, $\hat{Q}_n$ offers an accurate approximation of the quantile region corresponding to the block-maxima rv $M_{m_n}$, for an arbitrarily small $p$, provided that the block size $m_n$ and the number $n$ of its iid replicates are sufficiently large.

While fitting a multivariate max-stable model to componentwise maxima, sometimes practitioners tackle the simpler problem of estimating marginal quantiles. From a practical viewpoint, the reason for preferring joint modelling to individual inference lies in a more robust quantile estimation. E.g. in spatial applications, where each component of the rv of maxima pertains to a different location on a grid, using a model for the whole set of grid points yields estimates of locationwise quantiles which are less affected by local fluctuations [71]. Hellinger predictive consistency has implications also on such inferential task, as argued next.

EXAMPLE 5.4. Assume Condition 4.1 is satisfied. If (25) holds true, then $F_{0}^{(\infty)} - \text{as}$

$$\lim_{n \to \infty} \mathcal{D}_K(F_{0m_n}, G_n^{(o)}) = 0$$

and

$$\lim_{n \to \infty} \frac{(F_{0m_n})^{(1-p)}}{(G_n^{(o)})^{(1-p)}} = 1, \quad j = 1, \ldots, d, \quad \forall p \in (0, 1).$$

I.e., for $j = 1, \ldots, d$, the quasi-posterior predictive $(1 - p)$-quantile obtained inverting the $j$-th marginal cdf of $G_n^{(o)}$ is a consistent estimator of the $(1 - p)$-quantile of the $j$-th component of $M_{m_n}$. Noticeably, the latter also corresponds to the $(1 - p)^{1/m_n}$-quantile of $F_{0,j}$, representing a more extreme quantile for smaller probabilities $p$ and/or larger block sizes $m_n$. 
Acknowledgements. The authors are grateful to Michael Falk, Bas Kleijn and Anthony Davison for their valuable support and help. The authors are also grateful to two anonymous referees for their constructive suggestions, which have undoubtedly improved the presentation of this work. Simone Padoan is supported by the Bocconi Institute for Data Science and Analytics (BIDSA), Italy.

REFERENCES


DEPARTMENT OF DECISION SCIENCES, 
BOCCONI UNIVERSITY, 
VIA ROENTGEN 1, 20136, 
MILAN, ITALY. 
EMAIL: simone.padoan@unibocconi.it

CHAIR OF STATISTICS STAT, 
EPFL, MA B1 507, 
STATION 8, CH-1015, 
LAUSANNE, SWITZERLAND, 
EMAIL: stefano.rizzelli@epfl.ch
## SUPPLEMENTARY MATERIAL

### CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Notation</td>
<td>2</td>
</tr>
<tr>
<td>B</td>
<td>Bivariate simple max-stable distributions</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>B.1 Bernstein polynomial representation of the extremal dependence</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>B.2 Kullback-Leibler theory for priors on the extremal dependence</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>B.3 Posterior consistency</td>
<td>5</td>
</tr>
<tr>
<td>C</td>
<td>Multivariate simple max-stable distributions</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>C.1 Spectral representation</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>C.2 Metric properties</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>C.3 Positive Kullback-Leibler divergences</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>C.4 Construction of prior distributions</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>C.4.1 Prior distributions on univariate $A$</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>C.4.2 Prior distributions on multivariate $H$</td>
<td>14</td>
</tr>
<tr>
<td>D</td>
<td>Proofs</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>D.1 Proofs of the results in Section 2.4</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>D.1.1 Auxiliary results for the proof of Proposition 2.3</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>D.1.2 Proof of Proposition 2.3</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>D.2 Proofs of the results in Sections B.2–B.3</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>D.2.1 Proof of Proposition B.1</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>D.2.2 Proof of Corollary B.2</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>D.2.3 Proof of Theorem B.8</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>D.3 Proofs of the results in Section 3.2</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>D.3.1 Proof of Proposition 3.3</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>D.3.2 Proof of Corollary 3.4</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>D.3.3 Proof of Theorem 3.7</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>D.3.4 Auxiliary results for the proof of Theorem 3.9</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>D.3.5 Proof of Theorem 3.9</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>D.4 Proofs of the results in Section 3.3</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>D.4.1 Proof of Proposition 3.12</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>D.4.2 Auxiliary results for the proof of Theorem 3.14</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>D.4.3 Proof of Theorem 3.14</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>D.4.4 Auxiliary results for the proof of Theorem 3.15</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>D.4.5 Proof of Theorem 3.15</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td>D.4.6 Proof of Theorem 3.17</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td>D.5 Proofs of the results in Section 4.2</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>D.5.1 Proof of Proposition 4.5</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>D.6 Proofs of the results in Section 4.3</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>D.6.1 Auxiliary results for the proof of Theorem 4.7</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>D.6.2 Proof of Theorem 4.7</td>
<td>53</td>
</tr>
<tr>
<td></td>
<td>D.7 Proofs of the results in Section 4.5</td>
<td>53</td>
</tr>
<tr>
<td></td>
<td>D.7.1 Auxiliary results for the proof of Theorem 4.19</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>D.7.2 Proof of Theorem 4.19</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>D.8 Proofs of the results in Section 5</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>D.8.1 Proof of Corollary 5.1</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>D.8.2 Proof of Corollary 5.2</td>
<td>56</td>
</tr>
<tr>
<td>E</td>
<td>Examples and technicalities</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>E.1 Examples of Section 4.3</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>E.1.1 Technical derivations for Example 4.10</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>E.1.2 Technical derivations for Example 4.12</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>E.1.3 Technical derivations for Example 4.13</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td>E.2 Examples of Section 4.4</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>E.2.1 Technical derivations for Example 4.16</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>E.2.2 Technical derivations for Example 4.17</td>
<td>62</td>
</tr>
<tr>
<td></td>
<td>E.3 Examples of Section 5</td>
<td>62</td>
</tr>
<tr>
<td></td>
<td>E.3.1 Technical derivations for Example 5.3</td>
<td>62</td>
</tr>
<tr>
<td></td>
<td>E.3.2 Technical derivations for Example 5.4</td>
<td>65</td>
</tr>
</tbody>
</table>
This supplementary material document contains a discussion on additional results derived for bivariate simple max-stable distributions (see Section B), a collection of auxiliary results for multivariate simple max-stable models (see Section C), the proofs of all theoretical findings presented both in the main paper and in this manuscript (see Section D), some technical derivations for the examples proposed in the main text (Section E). In particular, the analysis in Sections 2.3–2.4 and 3.2 of the main paper is revisited in Section B, considering bivariate simple max-stable models and prior distributions specified on the univariate Pickands dependence function. A useful representation of $d$-variate simple max-stable models ($d \geq 2$) is recalled in Section C.1 and applied in Sections C.2–C.3, to derive properties of distances and divergences between simple max-stable densities. The latter may be of independent interest, beyond their use for establishing our consistency results. The construction of prior representations of the extremal dependence and the related Bayesian inference in the bivariate case. Given a

**APPENDIX A: NOTATION**

In this section we firstly provide some additional notation, used throughout the supplement along with those in Section 2.1 of the main article. Let $\mathcal{X} \subset \mathbb{R}$ and $f \in C^1(\mathcal{X})$ with an absolutely continuous first derivative. We denote the latter by $f’$ and its weak derivative by $f’’$ [e.g. 1, p. 22]. If $f$ is twice continuously differentiable, then $(\partial^2 f/\partial x^2)f(x) = f''(x)$ almost everywhere; therefore, in such a case, we consider the continuous representative given by the strong (i.e. the canonical) second derivative. We define the functionals

$$\|f\|_{1,\infty} := \|f\|_{\infty} + \|f’\|_{\infty}, \quad \|f\|_{2,\infty} := \|f\|_{1,\infty} + \|f’’\|_{\infty},$$

and the associated metrics $\mathcal{D}_{p,\infty}(f, g) := \|f - g\|_{p,\infty}$, $p = 1, 2$, with $g : \mathcal{X} \rightarrow \mathbb{R}$ sharing the same properties of $f$. If, instead, $f$ is only weakly differentiable, then $f’’$ denotes its weak derivative. For a generic space $\mathcal{X}$, endowed with a metric $\mathcal{D}$, $x_\ast \in \mathcal{X}$ and $\epsilon > 0$ we denote

$$B_{\epsilon,\mathcal{D}}(x_\ast) = \{x \in \mathcal{X} : \mathcal{D}(x, x_\ast) \leq \epsilon\};$$

for the sake of a lighter notation, when $\mathcal{D}$ is an $L_p$-distance, $p \in [1, \infty]$, we write $B_{\epsilon,p}(\cdot)$. With a little notational abuse, when $H_s \in \mathcal{H}$, with $\mathcal{H}$ given in Definition 2.2 of the main article, we also use the symbol

$$B_{\epsilon,\mathcal{H}}(H_s) := \{H \in \mathcal{H} : \|h - h_s\|_{\infty} \leq \epsilon\},$$

where $h$ and $h_s$ denote the angular densities of $H$ and $H_s$, respectively. Moreover, for any $B \subset \mathcal{X}$, we denote by $\mathcal{X}(\delta, B, \mathcal{D})$ the $\delta$-covering number of a set $B$ with respect to the metric $\mathcal{D}$ [e.g., 39, Appendix C]. Finally, for pairs of pm’s $F$ and $G$ with density functions $f$ and $g$ with respect to some dominating measure $\nu$ on $\mathcal{X}$, we denote the higher-order positive Kullback-Leibler divergences [e.g., 39, Appendix B] by

$$\mathcal{K}_+^{(l)}(f, g) = \int_{\mathcal{X}} \left[\log^+ \{f(x)/g(x)\}\right]^l f(x) d\nu(x),$$

where $l \in \mathbb{N}_+$ and, for all $y > 0$, $\log^+(y) = \max\{\log(y), 0\}$.

Secondly, we make a clarification on the notation used for max-stable observational models. For simplicity, in the remainder of this supplement we do not use the notation $G_\theta(\cdot | \theta)$, $\theta \in \Theta$, introduced in Section 3 of the main article for max-stable distributions, but directly switch to $G_\theta(\cdot | H)$, $H \in \mathcal{H}$, or $G_\theta(\cdot | A)$, $A \in A$. Accordingly, angular densities are hereafter always denoted by $h$ instead of $\theta$, occasionally used in the main paper. This is particularly the case whenever the results stated in the main paper are proved (i.e., in Section D). To denote the classes of angular pm’s given in Definition 3.6(i) and 3.6(ii) of the main paper, we use the symbols $\mathcal{H}_0$ and $\mathcal{H}_f$ in place of $\Theta_0$ and $\Theta^\prime$, respectively. Finally, whenever we refer to Definition 3.5 or Conditions 3.8, 3.11 and 3.13(ii)–(iii) in the main text, we replace the notation $\Pi_0$ (used for the prior distribution of the angular pm) with the more explicit symbol $\Pi_{H_0}$.

**APPENDIX B: BIVARIATE SIMPLE MAX-STABLE DISTRIBUTIONS**

This section complements Sections 2–3 of the main paper, by providing additional notions and results on representations of the extremal dependence and the related Bayesian inference in the bivariate case. Given a
Pickands dependence function $A \in \mathcal{A}$, when $d = 2$ the pertaining simple max-stable density function (equation (3) of the main article) simplifies to

$$g_1(y|A) = G_1(y|A) \left( \frac{(A(t) - t A'(t))(A(t) + (1-t) A'(t))}{(y_1 y_2)^2} + \frac{A''(t)}{(y_1 + y_2)^2} \right),$$

where $y > 0$ and $t \equiv t(y) = y_1/(y_1 + y_2)$. Moreover, conveniently setting $p_0 = H(\{e_2\})$, the pertaining angular cumulative distribution function (cdf) is of the form

$$H(t) = p_0 + \int_0^t h(v)dv + p_1 \delta_1([0, t]), \quad t \in [0, 1].$$

Consequently, $A$ and its derivatives are related to $H$ and $h$ through the explicit formulas

$$A(t) = 1 + 2 \int_0^t H(v)dv - t, \quad t \in [0, 1],$$

$$A'(t) = -1 + 2H(t), \quad t \in (0, 1),$$

$$A''(t) = 2h(t), \quad \text{for almost every } t \in (0, 1).$$

Furthermore, $A'(0) = 2p_0 - 1$ and $A'(1) = 1 - 2p_1$, where, with an abuse of notation, $A'$ also denotes the continuous extension of the first derivative of $A$ on $[0, 1]$. We stress that the notation $p_0$ should not be confused with a true parameter value, used to actually generate the data. In the sequel, the point masses pertaining to the true bivariate angular pm $H_0$ are denoted by $p_0, 0$ and $p_1, 1$, to avoid ambiguities.

The relations in (30)-(31) makes simultaneous polynomial modelling of the Pickands dependence and the angular distribution functions very tractable, as illustrated in the next subsection. In Section B.1 we review a Bernstein polynomial representation of the Pickands dependence and the angular distribution functions which, in the specific case $d = 2$, corresponds to the Bernstein polynomial representation of the angular density in Section 2.4 of the main article. Such construction can be used for prior specifications on the extremal dependence yielding consistent posterior distributions, as shown in Section B.3. The consistency results provided therein build on the general theory on Kullback-Leibler property presented in Section B.2. The latter rephrases the analysis in Section 3.2.1 of the main article from the perspective where uncertainty on the extremal dependence structure is dealt with by assigning a prior to the Pickands dependence function. We recall that the regularity conditions (C.2)-(C.3) in the main article univocally identify the class of valid Pickands dependence functions only when $d = 2$. Therefore, a direct specification of a prior with full support on the space of Pickands dependence functions is essentially tractable only in the bivariate case. From a practical viewpoint, the Pickands dependence function provides an easy-to-interpret description of the dependence structure of a bivariate max-stable rv and allows to readily retrieve some of the most commonly used extremal dependence measures. Examples are the extremal coefficient, the coefficient of upper tail dependence and Spearman’s rho for extreme-value copulas, given by $2A(1/2)$, $2 - 2A(1/2)$ and

$$12 \int_0^1 \frac{1}{(1 + A(t))^2} dt - 3,$$

respectively, see Ch. 8.2.7 in [5] for a comprehensive account. From a mathematical stance, we highlight that the case $d = 2$ is the only one where the map $(A, \mathcal{D}_{\infty}) \mapsto (H, \mathcal{D}_K) : A \mapsto H$ is homeomorphic (see also Proposition C.2). Such a relation between metric spaces is underpinned by equation (32) and the convexity of the Pickands dependence function. In inferential terms, this guarantees that $\mathcal{D}_{\infty}$-consistency at the true Pickands dependence function directly translates into $\mathcal{D}_K$-consistency at the true angular pm, while an equivalent implication fails in higher-dimensions. For technical convenience, the class $\mathcal{A}$ is hereafter endowed with the equivalent metric $\mathcal{D}_{1, \infty}$ and consistency results are provided with respect to the latter, paralleling the analysis in Section 3.2.2 of the main article.

### B.1. Bernstein polynomial representation of the extremal dependence.

According to [54], for $k \geq 2$, a $k - 1$-th degree Bernstein polynomial representation of the angular distribution function is given by

$$H_{k-1}(t) := \begin{cases} \sum_{j=0}^{k-1} \eta_j k^{-1} \text{Be}(t|j + 1, k - j), & \text{if } t \in [0, 1) \\ 1, & \text{if } t = 1 \end{cases}$$

where $\text{Be}(a|b)$ denotes the beta density function with shape parameters $a, b > 0$ and, for $0 \leq j \leq k - 1$, $k^{-1} \text{Be}(t|j + 1, k - j)$ is the $j$-th Bernstein basis function. If the polynomial coefficients $\eta_0, \ldots, \eta_{k-1}$ satisfy the restrictions:
$\beta$ defines a valid Pickands dependence function if its coefficients satisfy the restrictions:

- **(R3)** $0 \leq p_0 = \eta_0 \leq \eta_1 \leq \cdots \leq \eta_{k-1} = 1 - p_1 \leq 1$,
- **(R4)** $\eta_0 + \cdots + \eta_{k-1} = k/2$,

where $0 \leq p_0, p_1 \leq 1/2$, the function in (33) is a valid angular distribution function. Similarly, the Bernstein polynomial of degree $k$, for $k = 2, 3, \ldots$, given by

\begin{equation}
A_k(t) := \sum_{j=0}^{k} \beta_j (k+1)^{-1} \text{Be}(t|j+1, k-j-1), \quad t \in [0, 1],
\end{equation}

defines a valid Pickands dependence function if its coefficients satisfy the restrictions:

- **(R5)** $\beta_0 = \beta_k = 1 \geq \beta_j$, for all $j = 1, \ldots, k-1$;
- **(R6)** $\beta_1 = \frac{k-1+2p_0}{k}$ and $\beta_{k-1} = \frac{k-1+2p_1}{k};$
- **(R7)** $\beta_{j+2} - 2\beta_j + \beta_j \geq 0$, for $j = 0, \ldots, k-2$.

Thus, for $k = 2, 3, \ldots$, we define the classes of Pickands dependence functions and angular pm’s with distribution functions in *Bernstein polynomial (BP) form* via

\begin{align*}
A_k &= \{A_k \in A : A_k(t) = \sum_{j=0}^{k} \frac{\beta_j}{k+1} \text{Be}(t|j+1, k-j-1), \text{ (R5)-(R7) hold true} \}, \\
H_{k-1} &= \{H_{k-1} \in H : H_{k-1}(t) = \sum_{j=0}^{k-1} \frac{\eta_j}{k} \text{Be}(t|j+1, k-j), \text{ (R3)-(R4) hold true} \}.
\end{align*}

Notably, for each $A_k \in A_k$, it is possible to derive a polynomial $H_{k-1} \in H_{k-1}$ and vice versa, by means of precise relationships between the two polynomials’ coefficients, see [54, Proposition 3.2] for details. Moreover, by arguments in [54, Propositions 3.1-3.3], $\cup_{k=2}^{\infty} A_k$ and $\cup_{k=2}^{\infty} H_{k-1}$ are dense subsets of the spaces $(A, \mathcal{D}_1, \infty)$ and $(H, \mathcal{D}_K)$, respectively.

Concluding, we point out that, for any $t \in (0, 1)$, the angular density corresponding to the distribution in (33) is given by

\begin{equation}
h_{k-2}(t) = \sum_{j=1}^{k-1} (\eta_j - \eta_{j-1}) \text{Be}(t|j, k-j).
\end{equation}

Letting $\varphi_{\kappa_2} = \eta_0, \varphi_{\alpha} = \eta_{\alpha_1} - \eta_{\alpha_1-1}, \varphi_{\alpha} = (j, k-j), j = 1, \ldots, k-1, \text{ and } \varphi_{\alpha_1} = 1 - \eta_{k-1}$, we have that if $\eta_0, \ldots, \eta_{k-1}$ satisfy (R3)-(R4), then $(\varphi_{\kappa_1}, \varphi_{\alpha}, \alpha \in \Gamma_k)$ satisfy (R1)-(R2) and the Bernstein polynomial representation in Section 2.4 of the main article is retrieved.

**B.2. Kullback-Leibler theory for priors on the extremal dependence.** The Pickands dependence functions in $A$ are continuous and differentiable, with absolutely continuous first derivatives. Accordingly, we denote with $\Pi_A(B) = P(A \in B)$ a prior distribution on the Borel sets $B$ of $(A, \mathcal{D}_1, \infty)$. We have the following set theoretical results (the referenced definitions and equations are given in the main paper).

**PROPOSITION B.1.** Let $\mathcal{W}^{1,\infty}((0, 1))$ be the Sobolev space of bounded functions with bounded weak derivative on $(0, 1)$, endowed with $\mathcal{D}_1, \infty$. Then, in dimension $d = 2$, $A$ is a closed subset of $\mathcal{W}^{1,\infty}((0, 1))$. In particular, $(A, \mathcal{D}_1, \infty)$ is a Polish subspace.

**COROLLARY B.2.** In dimension $d = 2$, there exists a version of the simple max-stable density such that the map $(A, y) \mapsto g_1(y|A)$ is jointly measurable and $\phi_A : (A, \mathcal{D}_1, \infty) \mapsto (\mathcal{G}_1, \mathcal{D}_H) : A \mapsto g_1(|A|$ is a Borel map, where $\mathcal{G}_1$ is defined as in (10) with $\alpha = A$ and $\Theta = A$. Moreover, for all $\epsilon > 0$ and $K_\epsilon$ as in Definition 3.2, $\phi_A^{-1}(K_\epsilon)$ is a Borel set of $(A, \mathcal{D}_1, \infty)$.

By Corollary B.2, $\Pi_A$ induces a prior $\Pi_{\mathcal{G}_1}$ on the Borel sets of $(\mathcal{G}_1, \mathcal{D}_H)$. We next give conditions under which the true density is in the Kullback-Leibler support of the latter. In particular, we make use of the following notion.

**DEFINITION B.3.** Let $A_\epsilon \in A$ and $||A_\epsilon||_{2, \infty} < \infty$. The prior $\Pi_A$ is said to posses the $\mathcal{D}_2, \infty$ property at $A_\epsilon$ if it has positive inner probability on the sets $\{A \in A : \mathcal{D}_{2, \infty}(A, A_\epsilon) \leq \epsilon\}$, for all $\epsilon > 0$. 

To establish the Kullback-Leibler property at $A_0$ for a prior $\Pi_A$, we postulate twice continuous differentiability of the true Pickands dependence function on $(0,1)$ and impose mild restrictions on the behaviour of the second derivative near the boundary. The resulting class $A_0$ of admissible true Pickands dependence functions (Definition B.4(i)) is rich enough for applications, since it includes many well known parametric models, such as Symmetric Logistic, Asymmetric Logistic, Hüsler-Reiss, Pairwise-Beta, Dirichlet [see, e.g., 6]. We also point out that, by the fundamental theorem of calculus, $A_0$ coincides with the class of Pickands dependence functions whose angular pm’s lie in the class $\mathcal{H}_0$. As we claim next, it suffices to check that $\Pi_A$ possesses the $\mathcal{D}_{2,\infty}$-property at the Pickands dependence functions in a subclass (Definition B.4(ii)). Loosely speaking, the use of the stronger metric $\mathcal{D}_{2,\infty}$ is justified by the fact that the map $f \mapsto \mathcal{X}(g_1(\cdot|A), g_1(\cdot|f))$ on $(\mathcal{A}, \mathcal{D}_{1,\infty})$ is measurable for any fixed $A \in \mathcal{A}$ but not necessarily continuous.

**Definition B.4.** Let $\mathcal{A}$ be as in Definition 2.2 of the main paper and:

(i) Let $A_0 \subset \mathcal{A}$ be the class of twice continuously on $(0,1)$ (strongly) differentiable Pickands dependence functions $A$, whose second derivative satisfies one of the following:

(a) $\lim_{t \in (0,1)} A''(t) < +\infty$ and $\lim_{t \uparrow 1} A''(t) < +\infty$;

(b) $\inf_{t \in (0,1)} A''(t) > 0$ and $\lim_{t \downarrow 0} A''(t) = \lim_{t \uparrow 1} A''(t) = +\infty$.

(ii) Let $A' \subset A_0$ be a class of Pickands dependence functions whose extended first derivative satisfies $A'(0) > -1, A'(1) < 1$, and whose second derivative complies with $\inf_{t \in (0,1)} A''(t) > 0$, together with property (ia).

**Theorem B.5.** Let $A_0 \in A_0$ be the true Pickands dependence function. Assume that $\Pi_A$ possesses the $\mathcal{D}_{2,\infty}$-property at every $A \in A'$. Then, for any $\epsilon > 0$

$$\Pi_A(\phi_{\Delta}(K_\epsilon)) = \Pi_{G_1}(K_\epsilon) > 0.$$  

**Remark B.6.** As a map from $(\mathcal{A}, \mathcal{D}_{1,\infty})$ to $(\mathcal{H}, \mathcal{D}_{K,\Sigma})$, $A \mapsto H$ is 1-to-1 and continuous (thus, Borel). Hence, the prior $\Pi_A$ on the Pickands dependence function also induces a prior $\Pi_H$ on the Borel sets of $(\mathcal{H}, \mathcal{D}_{K,\Sigma})$. If $\Pi_A$ possesses the $\mathcal{D}_{2,\infty}$-property at every $A \in A'$, then $\Pi_H$ possesses the $\mathcal{D}_{\infty}$-property at every $H \in \mathcal{H}$ (Definition 3.5 of the main paper). Consequently, the result in Theorem B.5 readily follows from Theorem 3.7.

To establish the consistency results in Section B.3, Theorem B.5 is applied to the specific case of prior distributions constructed via the BP representation of the Pickands dependence functions (Section B.1). An explicit construction of such priors is detailed in Example B.9.

**B.3. Posterior consistency.** We consider priors on $A$ constructed through the following scheme, which relies on the representation in Section B.1, see also Section C.4.1 for technical details.

**Condition B.7.** Let $k_* \in \mathbb{N}_+$ and let the set sequence $(B_k)_{k=k_*}^{\infty}$ be defined via $B_k := \{\beta_k \in [0,1]^{k+1} : (R5)-(R7)$ hold true$.\}$ Assume $\Pi_A$ is the prior on the Borel sets of $(\mathcal{A}, \mathcal{D}_{1,\infty})$ induced by $\Pi$, the Borel pm on the disjoint union space $(\cup_{k=k_*}^{\infty} (\{k\} \times B_k), \Sigma)$ obtained via direct sum of the family $\{(B_k, \Sigma_k, \lambda(k)\nu_k), k \geq k_*\}$, where, for every $k \geq k_*$:

(i) $\nu_k$ is a pm with full support on $B_k$, equipped with its Borel $\sigma$-field $\Sigma_k$,

(ii) $\lambda(k) > 0$, with $\lambda(\cdot)$ denoting a probability mass function on $\{k_*, k_* + 1, \ldots\}$ such that

$$\sum_{i \geq k} \lambda(i) \lesssim e^{-qk}, \quad k \to \infty,$$

for some $q > 0$.

**Theorem B.8.** Let $Y_1, \ldots, Y_n$ be iid rv’s with distribution $G_1(\cdot|A_0)$, where $A_0 \in A_0$ \hspace{1pt} and $A_0$ is given in Definition B.4(i). Let $\Pi_A$ be a prior distribution on $A$ satisfying Condition B.7. Then, $G_1(\cdot|\theta_0) \sim\sim$

(a) $\lim_{n \to \infty} \Pi_n(\hat{U}^\theta) = 0$, for every $\mathcal{D}_{H}$-neighbourhood $\hat{U}$ of $g_1(\cdot|A_0)$;

(b) $\lim_{n \to \infty} \Pi_n(\hat{U}^\theta) = 0$, for every $\mathcal{D}_{1,\infty}$-neighbourhood $\hat{U}$ of $A_0$;

where $\Pi_n(\cdot) = \Pi_A(\cdot|Y_{1:n})$, $\hat{\Pi}_n(\cdot) = \Pi_{G_1}(\cdot|Y_{1:n})$ and $\Pi_{G_1} = \Pi_{G_1}$ is the prior on $g_1(\cdot|A)$ induced by $\Pi_A$. 


For the sake of completeness, a sketch of the proof is provided in Section D.2.3. However, note that the consistency results in Theorem B.8 also obtains as a byproduct of Theorem 3.9, once noticing that the prior $\Pi_A$ induces a prior $\Pi_H$ on the angular pm satisfying the assumptions therein, see Remark C.9 for technical aspects. Concluding, we provide an example of prior distribution $\Pi$ complying with Condition 3.8.

**Example B.9.** Exploiting the BP representation of the Pickands dependence function in (34), a prior $\Pi$ can be constructed by specifying priors $\nu_k$ for $k+1$-dimensional coefficient vectors $\mathbf{\beta}_k = (\beta_0, \ldots, \beta_k)$ via mixtures of the kernels
\[
dr_k(\mathbf{\beta}_k | p_1, p_0) = \delta_1(\beta_0) \delta(\beta_1) / k \prod_{j=2}^{k-2} \text{Unif}(\beta_j; a_{k,j}, b_{k,j}) \ (k-2) d\beta_j
\]
with respect to the mixing densities $\lambda_k(p_1, p_0) = \text{Unif}(p_1; a_{k,1}, b_{k,1}) \text{Unif}(p_0; 0, 1/2)$, where $\text{Unif}(a, b)$ is the uniform probability density over the interval $(a, b)$.

[34, Corollary 3.4]. Then, the prior probability mass function $\lambda$ for the polynomial degree $k$ can be chosen to be, e.g., a Poisson truncated outside $\{3, 4, \ldots\}$.

**Appendix C: Multivariate Simple Max-Stable Distributions**

In this section we provide some results on simple max-stable models and their dependence structure, that are useful for the proofs of the results in Sections 2–3 of the main article and B.2–B.3 of the present manuscript. In the sequel, we consider distributions of arbitrary dimension $d \geq 2$, unless otherwise specified.

**C.1. Spectral representation.** We start by recalling some known characterisations of simple max-stable models and setting up the corresponding notation. In the forthcoming subsections, these are exploited for technical derivations (e.g., Proposition C.1, Proposition C.8, Lemma D.3 and Corollary D.6). Any simple max-stable rv allows the following spectral representation [e.g., 31, Ch. 4–5]
\[
\mathbf{Y} = \max(\xi_i, i = 1, 2, \ldots),
\]
where $\xi_i, i = 1, 2, \ldots$ are the points of a Poisson process on $E := [0, \infty]^d \setminus \{0\}$ with mean measure $\Lambda(\cdot | H)$ satisfying
\[
\Lambda([0, y]^c | H) = \Lambda_1(y) = \int_{\mathcal{S}} \max_{1 \leq j \leq d} (w_j/y_j) dH(w).
\]
For each nonempty subset $I \subset \{1, \ldots, d\}$, define the subspace
\[
E_I := \{y \in E : y_j > 0, y_j = 0, \forall j \in I^c\}.
\]
By Definition 2.2 in the main paper, for any $H \in \mathcal{H}$, the restriction $\Lambda_{\{1, \ldots, d\} \setminus I}(\cdot | H)$ of $\Lambda(\cdot | H)$ on $E_{\{1, \ldots, d\} \setminus I}$ has Lebesgue density
\[
\lambda_{\{1, \ldots, d\} \setminus I}(z | H) = d||z||_1^{-d-1} h \circ \pi_{R}(z/||z||_1),
\]
while, for $j = 1, \ldots, d$, the restrictions $\Lambda_{\{j\}}(\cdot | H)$ on $E_{\{j\}}$ satisfy
\[
\Lambda_{\{j\}}(B | H) = dH(\{e_j\}) \int_{\pi_j(B \cap E_{\{j\}})} z_j^{-2} d\nu_j,
\]
for all Borel sets $B \subset E$, where, for $x \in E$, $\pi_j(x) = x_j$. Also note that $\Lambda(\cdot | H)$ has no mass on the subsets $E_I$ where $I$ is neither a singleton or the entire set $\{1, \ldots, d\}$. These properties yield the identity (11) in the main article.

For $j = 1, \ldots, d$, define the random index $\xi_{i,j}^*$ by $\xi_{i,j}^* = \max_{i=1,2,\ldots} \xi_i$. Then, the set $\{\xi_1^*, \ldots, \xi_d^*\}$ induces a random partition of $\{1, \ldots, d\}$. For a given angular pm $H \in \mathcal{H}$, we denote the joint density of a simple max-stable rv and the corresponding random partition by
\[
g_1(y, \mathcal{P} | H) = G_1(y | H) \prod_{i = 1}^m \{ -V_{I_i}(y | H) \}, \quad \mathcal{P} = \{I_1, \ldots, I_m\} \in \mathcal{P}_d,
\]
for almost every \( y \in \mathcal{Y} \), where \( \mathcal{Y} = (0, \infty)^d \) is the sample space for simple max-stable models, see e.g. [29, pp. 4816-4818] for additional details. In the sequel, we denote by \( G_1(\cdot, \cdot|H) \) the associated pm on the subsets of \( \mathcal{Y} \times \mathcal{P}_d \).

**C.2. Metric properties.** The results presented in Section 3.2 of the main article and B of the present manuscript involve several types of distance for simple max-stable distributions and the associated dependence functions. The following propositions describe the relationships existing among those. The first result we provide is valid in arbitrary dimension \( d \geq 2 \).

**Proposition C.1.** Let \( A_1 \) and \( A_2 \) be the Pickands dependence functions corresponding to the angular pm’s \( H_1, H_2 \in \mathcal{H} \), respectively. The following results hold true:

(i) \( e^{-1} \mathcal{D}_\infty(A_1, A_2) \leq \mathcal{D}_H(g_1(\cdot|H_1), g_1(\cdot|H_2)) \leq \| g_1(\cdot|H_1) - g_1(\cdot|H_2) \|_1^{1/2} \);

(ii) \( \mathcal{D}_\infty(A_1, A_2) \leq \min \{ 2\mathcal{D}_H(h_1, h_2)/\Gamma(d), 2d\|h_1 - h_2\|_1 \} \);

(iii) \( \forall \varepsilon > 0, \exists \eta > 0 \) such that, if \( \mathcal{D}_W(G_1(\cdot|H_1), G_1(\cdot|H_2)) < \eta \), then \( \mathcal{D}_H(H_1, H_2) < \varepsilon \);

where \( \Gamma(\cdot) \) is the gamma function.

**Proof.** The first half of the result at point (i) can be established by observing that for all \( t \in \mathcal{R} \), denoting \( y = (1/(1 - \|t\|_1), 1/t) \in \mathcal{Y} \), we have

\[
|A_1(t) - A_2(t)| \leq e|e^{-A_1(t)} - e^{-A_2(t)}| = e|G_1(y|H_1) - G_1(y|H_2)| \\
\leq e\mathcal{D}_H(G_1(\cdot|H_1), G_1(\cdot|H_2)) \\
\leq e\mathcal{D}_H(g_1(\cdot|H_1), g_1(\cdot|H_2)),
\]

see e.g. equation (B.1) and Lemma B.1(i) in [39] for the latter inequality. The second half directly follows from the inequalities

\[
\mathcal{D}_H(g_1(\cdot|H_1), g_1(\cdot|H_2)) \leq \| g_1(\cdot|H_1) - g_1(\cdot|H_2) \|_1^{1/2} \\
= (2\mathcal{D}_H(G_1(\cdot|H_1), G_1(\cdot|H_2)))^{1/2} \\
\leq (2\mathcal{D}_H(G_1(\cdot|H_1), g_1(\cdot|H_2)))^{1/2} \\
= \| g_1(\cdot|H_1) - g_1(\cdot|H_2) \|_1^{1/2},
\]

see also Lemma B.1(ii) in [39].

Next, observe that, using equation (5) and Definition 2.2, the term \( |A_1(t) - A_2(t)| \) can be bounded from above by

\[
d \left[ (1 - \|t\|_1)|p_{1,1} - p_{2,1}| + \sum_{j=2}^d t_{j-1}|p_{1,j} - p_{2,j}| \right] \\
+ d \int_{\mathcal{R}} \max \left\{ (1 - \|v\|_1)v_1, t_1v_2, \ldots, t_{d-2}v_{d-1}, t_{d-1}(1 - \|v\|_1) \right\} |h_1(v) - h_2(v)| dv,
\]

where, by the mean-constraints (C1),

\[
|p_{1,j} - p_{2,j}| \leq \int_{\mathcal{R}} v_j|h_1(v) - h_2(v)| dv, \quad j = 1, \ldots, d - 1, \\
|p_{1,d} - p_{2,d}| \leq \int_{\mathcal{R}} (1 - \|v\|_1)|h_1(v) - h_2(v)| dv.
\]

With few algebraic steps it can be now verified that the upper bound at point (ii) bounds from above the term in (37).

Finally, consider a sequence \( (H_k, k = 1, 2, \ldots) \in \mathcal{H} \) such that

\[
\lim_{k \to \infty} \mathcal{D}_W(G_1(\cdot|H_k), G_1(\cdot|H)) = 0.
\]

Then,

\[
\lim_{k \to \infty} \Lambda([0, y]_C|H_k) = \lim_{k \to \infty} -\log G_1(y|H_k) = -\log G_1(y|H) = \Lambda([0, y]_C|H),
\]
pointwise for every \( y \in \mathcal{G} \). This is sufficient to claim that \( \Lambda(\cdot|H_k) \) converges vaguely to \( \Lambda(\cdot|H) \). The identities \( \Lambda(B|H_k) = \Lambda_k(B), \Lambda(B|H) = \Lambda(B) \), for all Borel subset of \( \mathcal{S} \), together with Proposition 3.12(ii) in [65] now entail that \( \lim_{k \to \infty} \mathcal{D}_W(H_k, H) = 0 \). We can conclude that the map \( G_1(\cdot|H) \) is sequentially \( \mathcal{D}_W / \mathcal{D}_W \)-continuous and therefore \( \mathcal{D}_W / \mathcal{D}_W \)-continuous. The result at point (iii) follows.

In the bivariate case \( (d = 2) \), the simpler relations between the Pickands dependence and the angular distribution functions (equations (30)–(31)) underpin a stronger form of continuity for the map \( A \to H \), mapping a Pickands dependence function to its angular pm, than in the higher dimensional case \( (d > 2) \). Moreover, the more tractable form of the simple max-stable density (equation (29)) allows to establish a general Lipschitz continuity result for the map \( H \to g_1(\cdot|H) \), mapping an angular pm to the pertaining simple max-stable density.

**Proposition C.2.** In the specific case of \( d = 2 \), the following results hold true for any \( H_1, H_2 \in \mathcal{H} \), with corresponding Pickands \( A_1, A_2 \):

(i) \( \mathcal{D}_W^2(g_1(\cdot|H_1), g_1(\cdot|H_2)) \leq c\|h_1 - h_2\|_1 \), for a positive global constant \( c \);

(ii) \( \forall \epsilon > 0, \exists \eta > 0 \) such that, if \( \mathcal{D}_\infty(A_1, A_2) < \eta \), then \( \mathcal{D}_1(\cdot|H_1, A_2) \leq \epsilon \); in particular, \( \mathcal{D}_K(\cdot|H_1, H_2) \leq \epsilon \).

**Proof.** Recall the general expression of the simple-max stable density in (29) and observe that, by a change of variables, the term \( \|g_1(\cdot|H_1) - g_1(\cdot|H_2)\|_1 \) can be re-expressed as

\[
\int_0^1 \int_0^\infty \left| \frac{\phi_1(t) + r\chi_1(t)}{r^2t(1-t)^2} - \frac{\phi_2(t) + r\chi_2(t)}{r^2t(1-t)^2} \right| \exp \left( -\frac{A_1(t)}{rt(1-t)} \right) \, dr \, dt
\]

\[
\leq \int_0^1 \int_0^\infty \left| \frac{\phi_1(t) + r\chi_1(t)}{r^2t(1-t)^2} - \frac{\phi_2(t) + r\chi_2(t)}{r^2t(1-t)^2} \right| \exp \left( -\frac{A_1(t)}{rt(1-t)} \right) \, dr \, dt
\]

\[
+ \int_0^1 \int_0^\infty \left| \exp \left( -\frac{A_1(t)}{rt(1-t)} \right) - \exp \left( -\frac{A_2(t)}{rt(1-t)} \right) \right| \frac{\phi_2(t) + r\chi_2(t)}{r^2t(1-t)^2} \, dr \, dt.
\]

\[
=: I_1 + I_2,
\]

where, for \( i = 1, 2 \) and \( t \in (0, 1) \),

\[
\phi_i(t) = \left( A_i(t) - tA_i'(t) \right) \left( A_i(t) + (1-t)A_i'(t) \right), \quad \chi_i(t) = t^2(1-t)^2 A_i''(t).
\]

On one hand, Proposition C.1(iii), the identities in (31)–(32) and few algebraic manipulations yield

\[
I_1 \leq \int_0^1 \int_0^\infty \left( \frac{|\phi_1(t) - \phi_2(t)|}{r^2t(1-t)^2} + \frac{|\chi_1(t) - \chi_2(t)|}{r^2t(1-t)^2} \right) \exp \left( -\frac{A_1(t)}{rt(1-t)} \right) \, dr \, dt
\]

\[
\leq \int_0^1 |\phi_1(t) - \phi_2(t)| \frac{\Gamma(2)}{(A_1(t))^2} \, dt + \int_0^1 |A_1(t) - h_2(t)| \frac{\Gamma(1)(1-t)}{A_1(t)} \, dt
\]

\[
\leq c_1\|h_1 - h_2\|_1,
\]

for a positive global constant \( c_1 \). On the other, similar arguments together with the Lipschitz continuity of the exponential function on bounded sets give

\[
I_2 \leq \int_0^1 \int_0^\infty |A_1(t) - A_2(t)| \frac{e^{-\frac{1}{2}t^{-1}(1-t)^{-1}}}{r^2t(1-t)^2} \phi_2(t) + r\chi_2(t) \, dr \, dt
\]

\[
= \int_0^1 |A_1(t) - A_2(t)| \left( 8\phi_2(t)\Gamma(3) + 4\chi_2(t)\Gamma(2) \right) \, dt
\]

\[
\leq c_2\|h_1 - h_2\|_1,
\]

where \( c_2 \) is a positive global constant. The result at point (i) now follows from Lemma B.1(ii) in [39].

To establish the result at point (ii), start by considering a sequence \( (A_k, k = 1, 2, \ldots) \subset A \) such that \( \lim_{k \to \infty} \mathcal{D}_\infty(A_k, A_*) = 0 \). Recall that the first derivatives of \( A_\ast \) and \( A_k \) are uniformly continuous on \([0, 1]\). For, extend the first derivatives by considering the right and the left derivatives at 0 and 1, respectively. Thus, Theorem 25.7 in [68] entails that \( \lim_{k \to \infty} \mathcal{D}_\infty(A_k', A_*) = 0 \). We have then established that the map \( A \to A' \) is sequentially \( \mathcal{D}_\infty / \mathcal{D}_\infty \)-continuous, and thus continuous, at \( A_\ast \). The first half of the claim at point (ii) now follows. The second half is a direct consequence of the first one, together with the identity in (31). The proof is now complete.
C.3. Positive Kullback-Leibler divergences. In this section we introduce some technical results concerning the classical Kullback-Leibler and other related divergences for simple max-stable densities, which give the mathematical ground for proving the theorems of Section 3.2.1 of the main paper and Section B.2 of the present manuscript. They also represent a first step for establishing some of the lemmas in Sections D.4.2 and D.4.4 of the present manuscript, concerning semiparametric max-stable models. The actual interest on stronger notions of Kullback-Leibler divergences is motivated by the results in Section 4 of the main paper, whose proofs require control on higher-order moments of log likelihood ratios.

Before stating our first technical result, we introduce the following class of angular pm’s.

For any \( \{H, H_1, H_2\} \subset \mathcal{H} \) and \( l \in \mathbb{N}_+ \), define the functional

\[
\gamma_H^{(l)}(H_1; H_2) := \left( \int_Y \left[ \log^+ \left( \frac{g_1(y|H_1)}{g_1(y|H_2)} \right) \right]^l g_1(y|H)dy \right)^{1/l}.
\]

In particular, \( \gamma_H^{(l)}(H_1; H_2) \) equals the \( l \)-th root of the \( l \)-th order positive Kullback-Leibler divergence from \( g_1(\cdot|H_2) \) to \( g_1(\cdot|H_1) \) and it holds that

\[
\mathcal{K}(g_1(\cdot|H_1), g_1(\cdot|H_2)) \leq \gamma_H^{(1)}(H_1; H_2) = \mathcal{K}_1(g_1(\cdot|H_1), g_1(\cdot|H_2)).
\]

Before stating our first technical result, we introduce the following class of angular pm’s.

DEFINITION C.3. Let \( \mathcal{H}' \subset \mathcal{H} \) be the set of angular pm’s such that \( H(\{e_j\}) = p_j > 0 \), for \( j = 1, \ldots, d \), with angular density satisfying \( \|h\|_\infty < \infty \) and \( \inf_{t \in \mathcal{T}} h(t) > 0 \).

LEMMA C.4. Let \( H \in \mathcal{H}' \). Then, for every \( l \in \mathbb{N}_+ \) and \( \varepsilon > 0 \) there exists \( \delta > 0 \) such that, for all \( H_1 \in B_{\delta,\infty}(H) \) and \( H_2 \in \mathcal{H} \),

\[
\gamma_H^{(l)}(H; H_1) \leq \varepsilon + \log(1 + \varepsilon).
\]

In particular, when \( l = 1 \), it also holds that \( \mathcal{K}(g_1(\cdot|H), g_1(\cdot|H_1)) \leq \varepsilon + \log(1 + \varepsilon) \).

PROOF. Define \( \varepsilon' \) via \((1 + \varepsilon') = (1 + \varepsilon)^{1/d} \) and \( h_{\inf} := \inf_{v \in \mathcal{T}} h(v) \). Let

\[
0 < \delta < \min \left\{ \frac{\varepsilon}{2c(d\Gamma(l+1))(1+\varepsilon)^{1/l}}, \frac{\varepsilon'}{1+\varepsilon}, \frac{d}{1+\varepsilon} \min_{j=1,\ldots,d} p_j \right\},
\]

with \( c = 1/d \Gamma(d) \). Fix \( H_1 \in B_{\delta,\infty}(H) \). By Minkowski’s inequality, the term on the left hand-side of (38) is bounded from above by

\[
T_1 + T_2 \equiv \left( \int_Y \left| V(y|H_1) - V(y|H) \right|^l g_1(y|H_2)dy \right)^{1/l} + \left( \int_Y \left[ \log^+ \left( \frac{\sum_{P \in \mathcal{P}} \prod_{i=1}^m \{-V_{l_i}(y|H)\}}{\sum_{P \in \mathcal{P}} \prod_{i=1}^m \{-V_{l_i}(y|H_1)\}} \right) \right]^l g_1(y|H_2)dy \right)^{1/l}.
\]

Proposition C.1(ii), the bound in (39) and the fact that \( g_1(\cdot|H_2) \) has unit Fréchet margins allow to claim that

\[
T_1 \leq \left( \int_Y A_1 \left( \frac{1/y_1}{\|1/y_1\|} \right)^l - A \left( \frac{1/y_1}{\|1/y_1\|} \right)^l \right)^{1/l} {1/y_1} g_1(y|H_2)dy \leq (d\Gamma(l+1))^{1/l} \|\| (A, A_1) < \varepsilon,
\]

where \( I = \{2, \ldots, d\} \). Moreover, defining for all \( P \in \mathcal{P} \) the positive weights

\[
w_H(P) = \frac{\prod_{i=1}^m \{-V_{l_i}(y|H)\}}{\sum_{P' \in \mathcal{P}} \prod_{i=1}^m \{-V_{l_i}(y|H)\}},
\]
Jensen’s inequality yields that
\[
\log \left( \frac{\sum_{P \in \mathcal{P}} \prod_{i=1}^{m} (-V_{I_i}(y|H))}{\sum_{P \in \mathcal{P}} \prod_{i=1}^{m} (-V_{I_i}(y|H_1))} \right) = \max \left[ 0, -\log \left( \frac{\sum_{P \in \mathcal{P}} \prod_{i=1}^{m} (-V_{I_i}(y|H_1))}{\sum_{P \in \mathcal{P}} \prod_{i=1}^{m} (-V_{I_i}(y|H))} \right) \right]
\]
\[
= \max \left[ 0, -\log \left( \sum_{P \in \mathcal{P}} w_{H}(P) \prod_{i=1}^{m} \frac{(-V_{I_i}(y|H_1))}{(-V_{I_i}(y|H))} \right) \right]
\]
\[
\leq \max \left[ 0, -\sum_{P \in \mathcal{P}} w_{H}(P) \log \left( \frac{\prod_{i=1}^{m} (-V_{I_i}(y|H_1))}{\prod_{i=1}^{m} (-V_{I_i}(y|H))} \right) \right]
\]
(40)

Therefore, \( T_2 \leq \log(1+\epsilon) \), since for all \( P \in \mathcal{P}, i \in \{1, \ldots, m\} \) and (almost) every \( y \in \mathcal{Y} \) we have
\[
\left\{ \frac{-V_{I_i}(y|H)}{-V_{I_i}(y|H_1)} \right\} \leq 1 + \epsilon.
\]

The latter inequality easily follows from (11) and the subsequent facts:
(i) since \( \mathcal{D}_\infty(h, h_1) < \delta \), then \( h - h_1 < \delta = \frac{\delta}{h_{\inf}-\delta}(h_{\inf}-\delta) < \epsilon(h_{\inf}-\delta) < \epsilon h_1 \). As a result,
\[
\int_{0, \mathcal{Y}} \frac{dh \circ \pi_{\mathcal{R}}(z/\|z\|_1)}{\|z\|_{d+1}^i} \left| z_{i} = y_{i} \right| \leq \frac{1}{\epsilon} \int_{0, \mathcal{Y}} \frac{dh_1 \circ \pi_{\mathcal{R}}(z/\|z\|_1)}{\|z\|_{d+1}^i} \left| z_{i} = y_{i} \right| \; dz_{\mathcal{R}};
\]
(ii) by exploiting the mean constraints in (C.1) we deduce that for \( j = 1, \ldots, d - 1 \),
\[
\frac{(p_j - p_{1,j})/p_{1,j}}{p_{d} - p_{1,d}/p_{1,d}} = \frac{\int_0 \mathcal{R} v_{j}(h_1(v) - h(v))dv}{\int_0 \mathcal{R} (1 - \|v\|_1)(h_1(v) - h(v))dv} \leq \frac{\delta cd^{-1}}{\epsilon},
\]
and
\[
\frac{(p_j - p_{1,j})/p_{1,j}}{p_{d} - p_{1,d}/p_{1,d}} = \frac{\int_0 \mathcal{R} (1 - \|v\|_1)(h_1(v) - h(v))dv}{\int_0 \mathcal{R} (1 - \|v\|_1)dv} \leq \frac{\delta cd^{-1}}{\epsilon},
\]
therefore, for all \( j = 1, \ldots, d \), \( p_j \leq (1 + \epsilon)p_{j}^{'} \), wherefrom we conclude \( pjd_{I_i}^{-2} \leq (1 + \epsilon)p_{j}^{'}d_{I_i}^{-2} \), with \( I_i = \{j\} \).
The proof is now complete.

Remark C.5. Observe that the definition of \( \mathcal{H}'' \) is similar to that of \( \mathcal{H}' \). Though, the pm’s in \( \mathcal{H}'' \) are only required to have a bounded angular density, which may be discontinuous. As a consequence, \( \mathcal{H}' \subset \mathcal{H}'' \).

Lemma C.6. Let \( H \in \mathcal{H}_0 \) have angular density \( h \) which is uniformly continuous on \( \mathcal{R} \). Then, for all \( l \in \mathbb{N}^+ \) and \( \epsilon > 0 \) there exist \( H_\ast \in \mathcal{H}_l \) such that
\[
\mathcal{X}_{l}^{(l)}(g_1(\cdot|H), g_1(\cdot|H_\ast)) \leq \epsilon.
\]
In the particular case of \( l = 1 \), it also holds that \( \mathcal{X}(g_1(\cdot|H), g_1(\cdot|H_\ast)) \leq \epsilon \).

Proof. Consider the non-trivial case where \( H \notin \mathcal{H}' \) and fix \( \epsilon > 0 \). Without loss of generality, assume \( \epsilon < 1 \). For any \( H_1, H_2 \in \mathcal{H} \), Minkowski’s inequality entails that
\[
\left\{ \mathcal{X}_{l}^{(l)}(g_1(\cdot|H), g_1(\cdot|H_2)) \right\}^{1/l} = \mathcal{Y}_{H}^{(l)}(H; H_2)
\]
\[
\leq \mathcal{Y}_{H}^{(l)}(H_1; H) + \mathcal{Y}_{H}^{(l)}(H_1; H_2).
\]
Let \( \epsilon = 1/\Gamma(d) \) and fix an arbitrarily small \( \epsilon' > 0 \) such that
\[
\epsilon' + \log(1 + \epsilon') < \epsilon/2.
\]
Set $\epsilon'$ via $(1 + \epsilon') = (1 + \epsilon')^{1/d}$ and let $c_1, c_2$ be any two positive constants satisfying

$$c_1 < \min \left\{ \frac{\epsilon'}{\delta} \left\{ \frac{1}{d(1 + \epsilon')^{1/2e}} \right\}, \quad c_2 < \min \left\{ \frac{\epsilon'}{\delta} \left\{ \frac{1}{d(1 + \epsilon')^{1/2e}} \right\}, \frac{1}{1 + \epsilon' + c_1} \right\}. $$

Choose $H_1$ and $H_2$ in (41) as the angular pm’s corresponding to densities $h_1 = h/(1 + c_1)$ and $h_2 = h_1 + c_2$. respectively, and point masses

$$p_{i,j} = \frac{1}{d} - \int_{\mathcal{R}} v_j h_i(v) dv, \quad i = 1, 2, \quad j = 1, \ldots, d - 1,$n

$$p_{i,d} = \frac{1}{d} - \int_{\mathcal{R}} (1 - \|v\|_1) h_i(v) dv, \quad i = 1, 2.$n

Proceeding as in the proof of Lemma C.4 and using results similar to facts (i)-(ii) therein, we can show that both the terms on the right-hand side of (41) are smaller than $\epsilon' + \log(1 + \epsilon')$. Since $H_2 \in \mathcal{H}_l$, the result in the statement now follows selecting $H_* = H_2$.

In passing, note that, if $\inf_{w \in \mathcal{R}} h(w) > 0$, then $H_1 \in \mathcal{H}_l$ and, selecting $H_* = H_1$, the result follows directly from

$$\left\{ \mathcal{X}_+^{(l)} (g_1(\cdot|H), g_1(\cdot|H_1)) \right\}^{1/l} = \mathcal{Y}_H^{(l)} (H; H_1) < \epsilon' + \log(1 + \epsilon').$$

The proof is now complete. \qed

**Proposition C.7.** Let $H \in \mathcal{H}_{0+}$ have angular density $h$ which is uniformly continuous on $\mathcal{R}$. Then for any $l \in \mathbb{N}_+$ and $\epsilon > 0$ there exist $H_1 \in \mathcal{H}_l$ and $\delta > 0$ such that, for all $H_2 \in B_{\delta, \infty}(H_1)$,

$$\mathcal{X}_+^{(l)} (g_1(\cdot|H), g_1(\cdot|H_2)) \leq \epsilon.$$  

In the particular case of $l = 1$, it also holds that $\mathcal{X} (g_1(\cdot|H), g_1(\cdot|H_2)) \leq \epsilon$.

**Proof.** The result readily obtains by bounding from above $\{ \mathcal{X}_+^{(l)} (g_1(\cdot|H), g_1(\cdot|H_2)) \}$ as in (41). On one hand, by Lemma C.6, the first term on the right-hand side can be made arbitrarily small via an appropriate choice of $H_1 \in \mathcal{H}_l$. On the other hand, by Lemma C.4, also the second term can be made arbitrarily small by choosing $H_2 \in B_{\delta, \infty}(H_1)$ and $\delta > 0$ small enough. \qed

In the specific case of $d = 2$, we establish a result similar to the previous one for unbounded angular densities.

**Proposition C.8.** Let $d = 2$ and $H \in \mathcal{H}_{0+}$, with angular density $h$ satisfying

$$\inf_{t \in (0,1)} h(t) > 0, \quad \lim_{t \uparrow 1} h(t) = \lim_{t \downarrow 0} h(t) = +\infty.$$  

Then, for each $l \in \mathbb{N}_+$ and for all $\epsilon > 0$, there exist $H_1 \in \mathcal{H}_l$ and $\delta > 0$ such that the inequality (42) is satisfied by all $H_2 \in B_{\delta, \infty}(H_1)$. In the particular case of $l = 1$, it also holds that $\mathcal{X} (g_1(\cdot|H), g_1(\cdot|H_2)) \leq \epsilon$.

**Proof.** We start by constructing the angular pm $H_1$ as follows. For small constants $\epsilon_1, \epsilon_2$, there exist positive bounded functions $\gamma_{\epsilon_1}, \gamma_{\epsilon_2}$ which are continuous on $[0,1]$, satisfy $\gamma_{\epsilon_1} < h(t), \forall t \in (0, \epsilon_1), \gamma_{\epsilon_2} < h(t), \forall t \in (1 - \epsilon_2, 1)$, and are such that the function

$$h_1(t) := \begin{cases} 
\gamma_{\epsilon_1}(t), & t \in (0, \epsilon_1) \\
h(t), & t \in [\epsilon_1, 1 - \epsilon_2] \\
\gamma_{\epsilon_2}(t), & t \in (1 - \epsilon_2, 1) 
\end{cases}$$

can be continuously extended to $[0,1]$ and is bounded from below by $\inf_{t \in (0,1)} h(t) > 0$. Set

$$p_{1,1} = 1/2 - \int_0^1 t h_1(t) dt, \quad p_{1,2} = 1/2 - \int_0^1 (1 - t) h_1(t) dt,$$  

Then, the corresponding angular pm $H_1$ is an element of $\mathcal{H}_l$. Fix $l \in \mathbb{N}_+$. To establish the result in the statement, we now resort to the inequality (41), where, by Lemma C.4, the second term on the right-hand side can be made arbitrarily small by choosing $\delta$ small enough. As for the first term, i.e. $\mathcal{Y}_H^{(l)} (H; H_1)$, we next show that it can be made arbitrarily small by choosing $\epsilon_1, \epsilon_2$ sufficiently small.
Preliminary observe that, in the present bivariate case, \( V(y|H) = A(ty)(y_1^{-1} + y_2^{-2}) \) and

\[
V_{\{1\}}(y|H)V_{\{1\}}(y|H) = \frac{\left\{ A(ty) - tA'(ty) \right\}}{(y_1y_2)^2} \left\{ A(ty) + (1 - ty)A'(ty) \right\},
\]

where \( ty = y_1/\|y\|_1 \) and \( A \) is the Pickands dependence function pertaining to \( H \). Similar equalities hold true for \( H_1 \) and the pertaining Pickands dependence and exponent functions, \( A_1 \) and \( V(\cdot|H_1) \), respectively. This fact, Minkowski’s inequality and a few algebraic manipulations allow to deduce that

\[
\gamma_{H}^{(l)}(H;H_1) \leq R_1 + R_2 + R_3
\]

where

\[
R_1 = \mathcal{G}_\infty(A,A_1)\Gamma^{1/l}(1 + l),
\]

\[
R_2 = \left\{ \int_{y} \left[ \log^+ \left( \frac{A(ty) - tyA'(ty)}{A_1(ty) - tyA_1'(ty)} \right) \right]^{1/l} g_1(y|H)dy \right\} \frac{1}{l},
\]

\[
R_3 = \left\{ \int_{y} \left[ \log^+ \left( \frac{h(ty)}{h_1(ty)} \right) \right]^{1/l} g_1(y|H)dy \right\} \frac{1}{l}.
\]

By Proposition C.1(ii), we have that

\[
\frac{R_1}{\Gamma^{1/l}(1 + l)} \leq 4\|h - h_1\|_1 \leq 4 \left( \int_0^{\epsilon_1} + \int_1^{1-\epsilon_2} \right) h(t)dt,
\]

and the right-hand side can be made arbitrarily small by choosing \( \epsilon_1, \epsilon_2 \) small enough. Hence, we conclude that \( R_1 \) can be made arbitrarily small by choosing \( \epsilon_1, \epsilon_2 \) sufficiently small. Next, observe that for every \( t \in (0,1) \)

\[
0 < R_2^{(1)}(t) := \frac{A(t) - tA'(t)}{A_1(t) - tA_1'(t)} = \frac{1 - 2 \int_0^t \int_v^1 h(s)dsdv}{1 - 2 \int_0^t \int_v^1 h_1(s)dsdv} \leq 1,
\]

\[
0 < R_2^{(2)}(t) := \frac{A(t) + (1 - t)A'(t)}{A_1(t) + (1 - t)A_1'(t)} = \frac{1/2 - \int_0^{1-t} \int_v^{1-v} h(s)dsdv}{1/2 - \int_0^{1-t} \int_v^{1-v} h_1(s)dsdv} \leq 1,
\]

moreover

\[
\lim_{t \to 0} R_2^{(1)}(t)R_2^{(2)}(t) = p_2/p_{1,2} \leq 1, \quad \lim_{t \to 1} R_2^{(1)}(t)R_2^{(2)}(t) = p_1/p_{1,1} \leq 1.
\]

Therefore, \( R_2 = 0 \). Finally, due to the integrability of \( h \), there exists a constant \( \varsigma \in (0,1) \) such that

\[
\lim_{t \to 0} t^{\varsigma}h(t) = \lim_{t \to 1} (1 - t)^{\varsigma}h(t) = 0.
\]

Consequently, Minkowski’s inequality and a few algebraic manipulations allow to deduce that, for \( \epsilon_1, \epsilon_2 \) sufficiently small,

\[
R_3 \leq \frac{2\varsigma}{(1 - \varsigma)^{1+1/l}} \left( \Gamma(l + 1, -\log(\max\{\epsilon_1, \epsilon_2\})(1 - \varsigma)) \right)^{1/l}
\]

where \( a \) is a positive global constant and \( \Gamma(x,y) \) is the upper incomplete Gamma function. The term on the right-hand side can be made arbitrarily small by choosing \( \epsilon_1, \epsilon_2 \) small enough, whence the conclusion.

**C.4. Construction of prior distributions.** In this subsection we discuss more extensively some set theoretical aspects of the prior constructions in Section 3.2.2 of the main paper and Section B.3 of this manuscript.

**C.4.1. Prior distributions on univariate \( A \).** We first deal with the case where \( d = 2 \) and give additional details on the prior construction in Condition B.7. As in Sections B.2–B.3, we endow \( A \) with the Sobolev metric \( \mathcal{D}_{1,\infty} \) and denote by \( \mathcal{B}_A \) the induced Borel \( \sigma \)-algebra. In what follows, for a fixed positive
integer $k_s$, the subsets $\{A_k, k \geq k_s\}$ denote the classes of Pickands dependence functions in BP form, given in Section B.1. Note that, for each $k \geq k_s$, $A_k$ is $\mathcal{D}_{1,\infty}$-close, therefore $A_k \in \mathcal{B}_A$ and

$$A_* := \bigcup_{k \geq k_s} A_k \in \mathcal{B}_A.$$  

A pm on $(A, \mathcal{B}_A)$, say $\Pi_{A_*}$, can be obtained from a pm $\Pi_{A_*}$ on the subspace $\sigma$-algebra $\mathcal{B}_{A_*}$ pertaining to $A_*$ via

$$\Pi_{A_*}(B) := \Pi_{A_*}(B \cap A_*), \quad \forall B \in \mathcal{B}_{A_*}.$$  

Clearly, since $A_*$ is dense in $A$, if $\Pi_{A_*}$ has full support, the same is true for $\Pi_{A_*}$. This fact is exploited in Section B.3 to specify a prior distribution on the Pickands dependence function. Therein, a prior on $(A_*, \mathcal{B}_{A_*})$ is first induced by specifying a probability measure $\Pi$ on the Borel sets of the disjoint union topological space

$$\mathcal{B}_* := \bigcup_{k=k_s}^{\infty} \{k\} \times \mathcal{B}_k,$$

where the $\mathcal{B}_k$’s are Euclidean subspaces of suitable linear coefficient vectors (Condition B.7). Precisely, letting $\lambda(\cdot)$ be a positive probability mass function on $\{k_s, k_s + 1, \ldots\}$ and $\nu_k$, $k \geq k_s$, a sequence of pm’s on $\mathcal{B}_k$’s endowed with the Borel $\sigma$-algebras $\mathcal{B}_k$. Then $\Pi_k := \lambda(k)\nu_k(\cdot)$. Note that the $\sigma$-field of the measure space constructed via direct sum coincides with the Borel $\sigma$-algebra generated by the disjoint union topology, hereafter denoted by $\mathcal{B}_*$; see e.g. [35, p. 41] and [36, pp. 18, 84 and 828]. For $k \geq k_s$, define by

$$\phi^{(BP)}_{A_k}(\cdot) : B_k \mapsto \mathcal{A}_*: (\beta_0, \ldots, \beta_k) \mapsto \sum_{0 \leq j \leq k} \beta_j (k + 1)^{-1} \text{Be}(j + 1, k - j - 1)$$

the maps transforming linear coefficients into the corresponding Pickands dependence functions in BP form. Clearly, these are injective and continuous with respect to the Euclidean distance on $B_k$ and $\mathcal{D}_{1,\infty}$ on $A_*$. Thus, by the universal property of disjoint union topologies, there exists a unique continuous map

$$\phi^{(BP)}_{A_*} : \mathcal{B}_* \mapsto \mathcal{A}_*$$

such that, for all $k \geq k_s$,

$$\phi^{(BP)}_{A_k}(\beta) = \phi^{(BP)}_{A_*}((k, \beta))$$

whenever $\beta \in \mathcal{B}_k$. Finally, we have that $\Pi_{A_*} = \Pi \circ \{\phi^{(BP)}_{A_*}\}^{-1}$.

Next, we pinpoint the construction of a prior on the angular pm starting from the prior $\Pi$ on degree $k$ and coefficients $(\beta_0, \ldots, \beta_k)$ in the BP form of the Pickands dependence function. For $k \geq k_s$, let $\Phi_{k-1}$ be defined as in Condition 3.8. Then, define the maps

$$\phi^{(BP)}_{k} : B_k \mapsto \Phi_{k-1} : (\beta_0, \ldots, \beta_k) \mapsto (\varphi_{\kappa_1}, \varphi_{\kappa_2}, \varphi_{\alpha}, \alpha \in \Gamma_k) \equiv \varphi^{(k)}$$

transforming linear coefficients for the Pickands dependence function into the linear coefficients for the associated angular pm on the two dimensional simplex with density in BP form, where

$$\varphi_{\kappa_1} = \frac{k \beta_{k-1} - k + 1}{2}, \quad \varphi_{\kappa_2} = \frac{k \beta_1 - k + 1}{2},$$

$$\varphi_{\alpha} = \frac{k}{2}(\beta_{j+1} - 2 \beta_j + \beta_{j-1}), \quad \alpha = (j, k - j), \quad j = 1, \ldots, k - 1,$$

see Proposition 3.2 in [54] and equation (35) of this manuscript. Consider the disjoint union topological space

$$\Phi_* := \bigcup_{k \geq k_s} \{k\} \times \Phi_{k-1}$$

and endow it with the corresponding Borel $\sigma$-field $\mathcal{B}_{\Phi_*}$. The maps $\phi^{(BP)}_{k}$ are 1-to-1 and continuous; thus, by the universal property of disjoint union topologies, there exists a unique continuous map

$$\phi^{(BP)}_{*} : \mathcal{B}_* \mapsto \Phi_*$$

such that, for all $k \geq k_s$,

$$\phi^{(BP)}_{*}(\beta) = \phi^{(BP)}_{*}((k, \beta))$$
whenever $\beta \in B_k$. Hence, a prior $\Pi$ on $(B_k, \mathcal{B}_{B_k})$ induces a prior $\Pi' = \Pi \circ \{\phi_s^{(BP)}\}^{-1}$ on $(\Phi_s, \mathcal{B}_{\Phi_s})$. As done before for Pickands dependence functions, let $\mathcal{H}_s := \bigcup_{k \geq k_s} \mathcal{H}_{k-1}$ and, for $k \geq k_s$, define by

$$\phi^{(BP)}_{H_{k-1}} : \Phi_{k-1} \mapsto \mathcal{H}_s : \varphi \mapsto \left( \frac{2}{\pi} \sum_{j=1}^{k} \delta_{e_j} h_{k-2}(t) dt \right)$$

with $h_{k-2}(t) = \sum_{\alpha \in \Gamma_k} \varphi_{\alpha} \beta(t | \alpha)$ $\mathcal{H}_s$ such that, for all $k \geq k_s$, define by

$$\phi^{(BP)}_{H_{k-1}} \circ \varphi = \phi^{(BP)}_{H_{k-1}}((k, \varphi))$$

whenever $\varphi \in \Phi_{k-1}$. Thus, a prior is induced on $(\mathcal{H}_s, \mathcal{B}_{\mathcal{H}_s})$ via

$$\Pi_{\mathcal{H}_s}(B) = \Pi' \circ \{\phi^{(BP)}_{H_{k-1}}\}^{-1}(B), \quad \forall B \in \mathcal{B}_{\mathcal{H}_s},$$

where $\mathcal{B}_{\mathcal{H}_s}$ is the Borel $\sigma$-field generated by $\mathcal{B}_{\mathcal{H}_s}$. In turn, a prior on the Borel subsets $B$ of $(\mathcal{H}, \mathcal{B}_{\mathcal{H}_s})$ is defined via

$$\Pi_{\mathcal{H}_s}(B) = \Pi_{\mathcal{H}_s}(B \cap \mathcal{H}_s).$$

On the other hand, the map $\phi_{A/\mathcal{H}} : \mathcal{A} \mapsto \mathcal{H} : A \mapsto H$ is 1-to-1 and measurable with respect to $\mathcal{B}_{A}$ and the Borel $\sigma$-field induced by $\mathcal{B}_{\mathcal{H}_s}$ on $\mathcal{H}$, respectively. Consequently, the prior $\Pi_{\mathcal{A}}$ on $(\mathcal{A}, \mathcal{B}_{\mathcal{A}})$ induces a prior $\Pi_{\mathcal{A}} \circ \phi_{A/\mathcal{H}}^{-1}$ on the Borel subsets of $(\mathcal{H}, \mathcal{B}_{\mathcal{H}_s})$. We thus point out the following equivalence between pm’s.

**Remark C.9.** For each $B \in \mathcal{B}_{\mathcal{H}_s}$ we have the identity

$$\{\phi^{(BP)}_{A}\}^{-1} \circ \phi_{A/\mathcal{H}}^{-1}(B) = \{\phi^{(BP)}_{s}\}^{-1} \circ \{\phi^{(BP)}_{H_{k-1}}\}^{-1}(B),$$

where $\mathcal{H}_s$ immediately follows that

$$\Pi_{\mathcal{H}} = \Pi_{\mathcal{A}} \circ \phi_{A/\mathcal{H}}^{-1}.$$
to $d + 2$, $\Pi'$ is replaced by $\Pi$ and equation (43) is generalised to

$$\phi^{(BP)}_{\mathcal{H}_{k-1}}(\mathcal{H} : \Phi_{k-1} \mapsto \mathcal{H} : \phi^{(k)}(\cdot) \mapsto \sum_{j=1}^{d} \delta_{e_{j}}(\cdot) \phi_{\alpha_{j}} + \int_{\mathcal{R} \cap \mathcal{S}_{1,\ldots,d}} h_{k-d}(t) dt$$

with $h_{k-d}(t) = \sum_{\alpha \in \Gamma_{k}} \phi_{\alpha} \text{Dir}(t|\alpha)$, $t \in \mathcal{R}$; $\phi^{(k)}(\cdot) = (\phi_{\alpha_{1}}, \ldots, \phi_{\alpha_{d}}, \phi_{\alpha}, \alpha \in \Gamma_{k}$).

APPENDIX D: PROOFS

D.1. Proofs of the results in Section 2.4.

D.1.1. Auxiliary results for the proof of Proposition 2.3. The following approximation result is instrumental in proving Proposition 2.3. It is also exploited later on for inspecting the Kullback-Leibler support of prior distributions on max-stable densities (e.g., proof of Theorem 3.9; Lemma D.12 and Corollary D.13, auxiliary to the proof of Theorem 3.14).

**Lemma D.1.** For every $H \in \mathcal{H}'$ there exists a sequence $H_{k} \in \mathcal{H}_{k}$, with coefficients $\phi^{(k)}$, such that

$$\lim_{k \to \infty} \mathcal{D}_{\infty}(h, h_{k-d}) = 0.$$

**Proof.** For $\alpha \in \Gamma_{k}$ and $I = \{1, \ldots, d-1\}$, define

$$b_{\alpha-1}(t) = \frac{(k-d)!}{\prod_{j=1}^{d-1} (\alpha_{j} - 1)! (k-d - ||\alpha_{I} - 1||_{1})} \prod_{j=1}^{d-1} t_{j}^{\alpha_{j} - 1} (1 - ||t||_{1})^{k-d - ||\alpha_{I} - 1||_{1}}$$

$$= \frac{(k-d)!}{\prod_{j=1}^{d-1} (\alpha_{j} - 1)!} \prod_{j=1}^{d-1} t_{j}^{\alpha_{j}} (1 - ||t||_{1})^{\alpha_{d}-1}$$

$$= \frac{(k-d)!}{(k-1)!} \text{Dir}(t|\alpha).$$

By assumption, the angular density $h$ admits an extension $\tilde{h}$ which is bounded and continuous on $\mathcal{R}$. Then, it is well known that as $k \to \infty$

$$B_{k-d}(\tilde{h}; t) := \sum_{\alpha \in \Gamma_{k}} \frac{\alpha_{d} - 1}{k-d} b_{\alpha-1}(t) = \tilde{h}(t) + o(1),$$

where the error term is uniform over $\mathcal{R}$, see e.g. [52, p. 51]. Therefore, letting $c = 1/\Gamma(d)$, for every $\epsilon \in (0,1 \leq I \leq d)\cup$ there exists $k_{\epsilon}$ such that for all $k \geq k_{\epsilon}$

- $\mathcal{D}_{\infty}(B_{k-d}(\tilde{h}; \cdot), \tilde{h}) < \epsilon$ and $\int_{\mathcal{R}} B_{k-d}(\tilde{h}; t) dt < \int_{\mathcal{R}} h(t) dt + \epsilon < 1$;
- $\int_{\mathcal{R}} t_{j} B_{k-d}(\tilde{h}; t) dt < d^{-1} - p_{j} + c\alpha_{d} - 1 < d^{-1},$ for $j = 1, \ldots, d-1$ and $\int_{\mathcal{R}} (1 - ||t||_{1}) B_{k-d}(\tilde{h}; t) dt < d^{-1} - p_{d} + c\alpha_{d-1} - 1 < d^{-1}$.

As a consequence, setting for each $k \geq k_{\epsilon}$

$$\phi^{(k)}_{\alpha_{j}} = \frac{1}{d} - \int_{\mathcal{R}} t_{j} B_{k-d}(\tilde{h}; t) dt, \quad j = 1, \ldots, d-1,$$

$$\phi^{(k)}_{\alpha_{d}} = \frac{1}{d} - \int_{\mathcal{R}} (1 - ||t||_{1}) B_{k-d}(\tilde{h}; t) dt,$$

$$\phi^{(k)}_{\alpha} = \tilde{h} \left( \frac{\alpha_{d} - 1}{k-d} \right) \frac{(k-d)!}{(k-1)!}, \quad \alpha \in \Gamma_{k},$$

we obtain a sequence of valid angular pm’s $H_{k}$ of the form (8). In particular, the coefficients $\phi^{(k)}$ satisfy (R1)-(R2) in the main article, being the latter necessary and sufficient conditions to define valid angular pm’s via multivariate Bernstein polynomials. □
D.1.2. Proof of Proposition 2.3. Fix an arbitrary $\epsilon > 0$. For any positive constant $c < \epsilon/3$, we can define a valid spectral pm $H^* \in \mathcal{H}$ via

$$H^*(B) = \frac{d}{\sum_{j=1}^{d} p_j^* \delta_{e_j}(B)} + \int_{\pi_{R}(B\cap S_{1,\ldots,d})} h^*(v)dv,$$

for all Borel subsets $B$ of $\hat{S}$, where $h^* = h/(1 + c)$ and $p_j^* = 1/d - \int_{\hat{R}} t_j h^*(t)dt$, $j = 1, \ldots, d - 1$, $p_d^* = 1/d - \int_{\hat{R}} (1 - ||t||) h^*(t)dt$. Note that

$$\|h^* - h_k^*\|_1 \leq \epsilon/(1 + c) < \epsilon/3$$

and $\int_{\hat{R}} h^*(t)dt < 1$, $\min_{j=1,\ldots,d} p_j^* > 0$. There exists a nonnegative continuous function $h_K^*$ with compact support $K \subset \hat{R}$ such that

$$\|h^* - h_k^*\|_1 < \frac{1}{6} \min \left\{ \epsilon, \min_{j=1,\ldots,d} p_j^* \right\} =: c'$$

see e.g. [69, Ch. 3]. Consequently, choosing $c'' \in (0, c')$, setting $h^{**} = h_K^* + c''$, $p_j^{**} = 1/d - \int_{\hat{R}} t_j h^{**}(t)dt$, $j = 1, \ldots, d - 1$, $p_d^{**} = 1/d - \int_{\hat{R}} (1 - ||t||) h^{**}(t)dt$ and defining $H^{**}$ via

$$H^{**}(B) = \frac{d}{\sum_{j=1}^{d} p_j^{**} \delta_{e_j}(B)} + \int_{\pi_{R}(B\cap S_{1,\ldots,d})} h^{**}(v)dv,$$

for all Borel subsets $B$ of $\hat{S}$, we have $H^{**} \in \mathcal{H}'$ and

$$\|h^* - h^{**}\|_1 \leq \|h^* - h_k^*\|_1 + \|h_k^* - h^{**}\|_1 < \epsilon/3.$$

Finally, by Lemma D.1, there exists a sequence $H_0 \in \mathcal{H}_k$ such that, for $k$ sufficiently large, $\|h^{**} - h_{k-d}\|_1 < \epsilon/3$. The result now follows by triangular inequality.

D.2.2. Proof of the results in Sections B.2–B.3.

D.2.1. Proof of Proposition B.1. Let $(A_k)_{k=1}^{\infty} \subset A$ be a sequence which converges in $\mathcal{D}_{1,\infty}$-metric to some $A_* \in \mathcal{W}^{1,\infty}((0, 1))$. We now show it must be that $A_* \in A$, whence we conclude that $A$ is a closed subset of $\mathcal{W}^{1,\infty}((0, 1))$. By the uniform limit theorem, $A_*$ must be continuous. Consequently, it must also be convex [67, Theorem E, p. 17]. Obviously,

$$\max(t, 1-t) \leq \lim_{k \to \infty} A_k(t) \leq 1, \quad t \in [0, 1].$$

Thus, $A_*$ is a Pickands dependence function. Moreover, for any collection of nonoverlapping intervals $[a_i, b_i]$, $i = 1, \ldots, s$, with $s \in \mathbb{N}_+$, for any $\epsilon > 0$ and $k$ large enough we have

$$\sum_{i=1}^{s} |A_k'(b_i) - A_k'(a_i)| \leq \sum_{i=1}^{s} |A_k'(b_i) - A_k'(a_i)| + \epsilon/2.$$
D.2.3. Proof of Theorem B.8. To prove the first statement, we resort to Theorem 6.23 in [39] and verify that the conditions therein are satisfied. The second one obtains as a by-product.

Kalbuck-Leibler property. For all $A \in \mathcal{A}$ and $\delta > 0$ there exists a sequence $A_k \in \mathcal{A}_k$, with linear coefficients $\beta(k) \in \mathcal{B}_k$, such that, as $k \to \infty$,

$$B_k := \{A_k \in A_k : 2k^2 \|\beta(k) - \beta(k)\|_{\infty} < \delta\} \subset \{A_k \in A_k : \mathcal{D}_{2,\infty}(A_k, A_k) < \delta/2\} \subset \{A \in \mathcal{A} : \mathcal{D}_{2,\infty}(A, A) < \delta\}.$$ 

This is guaranteed by Theorem 6.3.2 in [19]. Furthermore, by Conditions B.7(i)--(ii), $\Pi_A B_k > 0$. We can now apply Theorem B.5 and conclude that $\Pi_A$ complies with (36), for all $\epsilon > 0$.

Metric entropy. Let $A_k$ and $\hat{A}_k$ be two Pickands dependence functions in BP form, with degree $k$. Let $H_{k-1}$ and $\tilde{H}_{k-1}$ be the corresponding angular distributions in BP form, obtained by applying Proposition 3.2(i) in [54].

The associated angular densities satisfy

$$\|h_{k-2} - \tilde{h}_{k-2}\|_1 = \int_0^1 \sum_{j=0}^{k-2} (\eta_{j+1} - \eta_j - \tilde{\eta}_{j+1} + \tilde{\eta}_j) \operatorname{Be}(\eta j + 1, k - j - 1) \, dv \leq \sum_{j=0}^{k-2} |\eta_{j+1} - \eta_j - \tilde{\eta}_{j+1} + \tilde{\eta}_j|.$$ 

By the above inequality, Proposition C.2(i) and Proposition C.2 in [39], we have that, for any set of the form $\mathcal{G}_1 = \{g_1(\cdot|A_k) : A_k \in A_k\}$,

$$\mathcal{N}(\epsilon, \mathcal{G}_1, \mathcal{D}_H) \leq \mathcal{N}(c' \epsilon^2, \{x \in \mathbb{R}^{k-2} : \|x\|_1 \leq 1\}, L^1) \leq (3/c' \epsilon^2)^{k-2},$$

for some $c' > 0$, where, without loss of generality, we assume $c' \epsilon^2 < 1$. Therefore, Conditions i. and ii. of Theorem 6.23 in [39] can be verified by arguments similar to those in Section D.3.5.

Conclusion. The first statement is now proven. As for the second statement, it follows from the first one together with Proposition C.1(i) and Proposition C.2(ii).

D.3. Proofs of the results in Section 3.2.

D.3.1. Proof of Proposition 3.3. Observe that, the union of the $d$ closed sets $\{e_j\}, j = 1, \ldots, d$, is a $G_\delta$ subset of $\mathcal{S}$. Moreover, $\pi_{\mathcal{R}} : \mathcal{S} \to \mathcal{R}$ is a continuous bijection and

$$\mathcal{S}_{\{1, \ldots, d\}} = \pi_{\mathcal{R}}^{-1}(\mathcal{R})$$

is the image of a $G_\delta$ subset of $\mathcal{R}$, thus $\mathcal{S}_{\{1, \ldots, d\}}$ is a $G_\delta$ subset of $\mathcal{S}$. As the union of two $G_\delta$ subsets,

$$\tilde{S} = \mathcal{S}_{\{1, \ldots, d\}} \cup \{e_1, \ldots, e_d\}$$

is also a $G_\delta$ subset of $\mathcal{S}$. As a result, $\tilde{S}$ is a Polish subspace of the simplex. Consequently, the class of Borel pm’s on the latter space, say $\mathcal{M}$, endowed with the weak topology, is metrizable by some metric $\mathcal{D}_W$ and Polish [39, Theorem A.3].

To establish the result in the second part of the statement, we use arguments similar to those of Theorem 4.1 in [37]. Let $\operatorname{Leb}_\mathcal{R}$ be the Lebesgue measure on $\mathcal{R}$ and define the class

$$\mathcal{I} := \{f \in L_1(\operatorname{Leb}_\mathcal{R}) : \int_\mathcal{R} v_j f(\nu) \operatorname{Leb}_\mathcal{R}(d\nu) \leq 1/d, j = 1, \ldots, d - 1, \int_\mathcal{R} (1 - \|v\|_1) f(\nu) \operatorname{Leb}_\mathcal{R}(d\nu) \leq 1/d, f \geq 0 \text{ a.e.}\}$$

and the map $\phi : \mathcal{I} \to \mathcal{M}$ via

$$[\phi(f)](B) = \int_{\pi_\mathcal{R}(B \cap \mathcal{S}_{\{1, \ldots, d\}})} f(\nu) \operatorname{Leb}_\mathcal{R}(d\nu) + \sum_{j=1}^{d-1} \left[1/d - \int_\mathcal{R} v_j f(\nu) \operatorname{Leb}_\mathcal{R}(d\nu) \delta e_j(B) \right] \delta e_j(B) + \left[1/d - \int_\mathcal{R} (1 - \|v\|_1) f(\nu) \operatorname{Leb}_\mathcal{R}(d\nu) \right] \delta e_d(B)$$

(48)
for any Borel subset $B$ of $\mathcal{S}$. It can be easily seen that $\mathcal{I}$ is the intersection of the $L_1$-closed sets \( \{ f \in L_1(\text{Leb}_R^+): f \geq 0 \text{ a.e.} \} \) and

$$\{ f \in L_1(\text{Leb}_R^+): \int_R v_j f(v) \text{Leb}_R^+(dv) \leq 1/d, \quad j = 1, \ldots, d-1, \}$$

$$\{ f \in L_1(\text{Leb}_R^+): \int_R (1 - \|v\|_1) f(v) \text{Leb}_R^+(dv) \leq 1/d, \}$$

thus $\mathcal{I}$ is closed. We deduce that $\mathcal{I}$ is a Borel subset of the Polish space $L_1(\text{Leb}_R^+)$. Consequently, to establish the second result, it is sufficient to show that $\phi$ is a Borel isomorphism from $\mathcal{I}$ onto $\mathcal{H}$.

Since $\mathcal{I}$, equipped with the subspace $L_1$-topology, is standard Borel (in fact, Polish) and $(\mathcal{M}, \mathcal{D}_W)$ is Polish, we can resort to Theorem 3.3.2 in [2] and prove that $\phi(\mathcal{I}) = \mathcal{H}$. From the definition in (48) it immediately follows that, for each $f \in \mathcal{I}$, we have $H(\cdot) = [\phi(f)](\cdot) \in \mathcal{H}$, thus $\phi(\mathcal{I}) \subseteq \mathcal{H}$. On the other hand, by Definition 2.2 in the main article, each $H \in \mathcal{H}$ admits the representation in (48), for some nonnegative Lebesgue-integrable function $f \in \mathcal{I}$, hence $\phi^{-1}(\mathcal{H}) \subseteq \mathcal{I}$, i.e. $\mathcal{H} \subseteq \phi(\mathcal{I})$. By the Radon-Nikodym theorem, the function $f$ satisfying such a representation is uniquely defined (up to a Leb$_R^+$-null set), thus $\phi$ is also $1 \to 1$. Finally, for any $f, h \in \mathcal{I}$ and any Borel subset $B \subseteq \mathcal{S}$, we have

$$[\phi(f)](B) - [\phi(h)](B) \leq (d + 1\|f - h\|_1,$$

i.e. the map $\phi: \mathcal{I} \to \mathcal{M}$ is Lipschitz continuous and, thus, a Borel map.

We have now established that $\mathcal{H}$ endowed with the subspace topology of weak convergence of pm's, metrized by $\mathcal{D}_W$, is standard Borel. Denote by $\mathcal{B}_W$ the pertaining Borel $\sigma$-field of subsets of $\mathcal{H}$. To prove the last claim in the statement, observe that $\mathcal{D}_K$ induces a stronger topology on $\mathcal{H}$ than $\mathcal{D}_W$. Then, the corresponding $\sigma$-field, say $\mathcal{B}_K$, contains $\mathcal{B}_W$. On the other hand, for any pair of angular pm's $H_1, H_2 \in \mathcal{H}$, we also have

$$\mathcal{D}_K(H_1, H_2) \leq (d + 1\|\phi^{-1}(H_1) - \phi^{-1}(H_2)\|_1 =: \mathcal{D}_\mathcal{I}(H_1, H_2)$$

and $\mathcal{I}$, endowed with the $L_1$ distance, is Polish (as a closed subset of a Polish space). Consequently, denoting by $\mathcal{B}_\mathcal{I}$ the Borel $\sigma$-algebra induced on $\mathcal{H}$ by the $\mathcal{D}_\mathcal{I}$-metric topology, we can deduce

$$\mathcal{B}_K \subseteq \mathcal{B}_\mathcal{I} \subseteq \mathcal{B}_W.$$  

It now follows that $\mathcal{B}_K = \mathcal{B}_W$, which completes the proof.

**D.3.2. Proof of Corollary 3.4.** Preliminary notice that the map $H \mapsto g_1(\cdot|H)$ 1-to-1. Indeed, each simple max-stable distribution can be seen as the distribution of the componentwise maxima over the points of a Poisson process, whose mean measure is univocally identified by an angular pm on the unit simplex $\mathcal{S}$; see Section C.1 and, e.g., Chapter 1.1 of [65]. Hence, the map linking an angular pm to the corresponding max-stable distribution is injective.

Next, denote by $\mathcal{B}_Y$ the Borel $\sigma$-algebra of $Y = (0, \infty)^d$ and by $\mathcal{B}_H$ the Borel $\sigma$-algebra induced on $\mathcal{H}$ by the metric $\mathcal{D}_W$. By Theorem V.58 in [24], there exists a positive $B_Y \times \mathcal{B}_H$-measurable function $(y, H) \mapsto g_1(y|H)$ such that, for all $H, g_1(\cdot|H)$ is a Lebesgue density of $G_1(\cdot|H)$. Then, arguments analogous to those in Appendix A of [63] yield that, for any fixed $H \in \mathcal{H}$ and $\epsilon > 0$,

$$\{ H \in \mathcal{H}: \mathcal{D}_H(g_1(\cdot|H_\epsilon), g_1(\cdot|H)) < \epsilon \} \in \mathcal{B}_H$$

and

$$\{ H \in \mathcal{H}: \mathcal{X}(g_1(\cdot|H_\epsilon), g_1(\cdot|H)) < \epsilon \} \in \mathcal{B}_H.$$  

Since $(G_1, \mathcal{D}_H)$ is separable, the associated Borel $\sigma$-algebra, $\mathcal{B}_{G_1}$, is generated by $\mathcal{D}_H$-open balls and we can conclude [e.g., 73, p. 86] that the map $H \mapsto g_1(\cdot|H)$ is $\mathcal{B}_H/\mathcal{B}_{G_1}$-measurable.

**D.3.3. Proof of Theorem 3.7.** It is sufficient to show that, for any $\epsilon > 0$, there exists a measurable subset of $\mathcal{H}$ with positive $\Pi_H$-mass, whose elements $H$ satisfy

$$\mathcal{X}(g_1(\cdot|H_\delta), g_1(\cdot|H)) < \epsilon.$$  

By Propositions C.7–C.8, the above inequality is satisfied for all $H \in B_{\delta,\infty}(H_\epsilon)$, for some $H_\epsilon \in \mathcal{H}$ and $\delta > 0$. Moreover, by assumption, there exists a measurable set $B_\delta \subseteq B_{\delta,\infty}(H_\epsilon)$ such that $\Pi_H(B_\delta) > 0$. The result now follows.
D.3.4. **Auxiliary results for the proof of Theorem 3.9.** We introduce a series of technical lemmas involving max-stable distributions with angular densities in BP form, culminating in Lemma D.5, used in the proof of Theorem 3.9. Then, we state and prove a corollary of Proposition 2.3, namely Corollary D.6, also exploited in the proof of Theorem 3.9.

In order to provide a concise account of the rather involved algebraic arguments used to establish Lemmas D.2–D.5, we make use of the following compact notation. For $I \subset \{1, \ldots, d\}$, denote by

\[ I^- := I \setminus \{\text{max}(i : i \in I)\} \]

the set obtained by removing from $I$ its largest element. As a convention, products over empty sets are meant as equal to 1, e.g., $x \in \mathbb{R}^d$ and $I = 0$ yield $\prod_{i \in I} x_i = 1$. For a function $f : \mathbb{R}^d \rightarrow \mathbb{R}$ and subsets $I_1, I_2 \subset \{1, \ldots, d\}$, $I_1 \cap I_2 = \emptyset$, we denote by $f(x)|_{x_{I_2} = y_{I_2}}$ the function $f$ evaluated at $x$ with fixed components $x_i = y_i$, for $i \in I_j$, $j = 1, 2$. Finally, for $v \in \mathbb{R}$, we define the function

\[ q(I, \alpha, v) := \begin{cases} 0, & \alpha = \kappa_j, j \in \{1, \ldots, d\}, I \neq \{j\}, \\ d(v_I - 2, d(1 - \|v\|_1)^{-2}, d(1 - \|v\|_1)^{-2}a_I \left(1 - \mathcal{I}_{v} \left(\alpha_I + 1, \|\alpha_{Jc}\|_1\right)\right) + d(1 - \|v\|_1)^{-2}\alpha_{Ic} \left(1 - \mathcal{I}_{v} \left(\alpha_I + 1, \|\alpha_{Jc}\|_1\right)\right) + \text{Dir}(v, \alpha) + d(\|\alpha_{Jc}\|_1), \\ \text{Dir} \left(\frac{v_I}{\|v_I\|_1}, \alpha_I\right) \left(1 - \mathcal{I}_{v} \left(\alpha_I + 1, \|\alpha_{Jc}\|_1\right)\right), & I = \{1, \ldots, d\}, \alpha \in \Gamma_k, \\ \text{Dir} \left(\frac{v_I}{\|v_I\|_1}, \alpha_I\right) \left(1 - \mathcal{I}_{v} \left(\alpha_I + 1, \|\alpha_{Jc}\|_1\right)\right), & d \notin I, 1 < |I| < d, \alpha \in \Gamma_k. \end{cases} \]

For any $a, b > 0$ and $x \in [0, 1]$, $\mathcal{I}_{x}(a,b)$ is the regularised incomplete Beta function, namely the cdf of a Beta distribution of parameters $a, b$ evaluated at $x$, while the (complete) Beta function is denoted by $\mathcal{B}(a,b)$.

In what follows, for a given integer $k > d$, the symbols $H_k$ and $\Gamma_k$ denotes the sets defined in Section 2.4 of the main article. For $\varphi \in \Phi_k$, where $\Phi_k$ is as in Condition 3.8 of the main article, we denote

\begin{align*}
&\varphi_{\beta} := (\varphi_{\kappa_1}, \ldots, \varphi_{\kappa_d}), \\
&\varphi_{\alpha} := (\varphi_{\alpha}, \alpha \in \Gamma_k),
\end{align*}

thus we have $\varphi = (\varphi_{\beta}, \varphi_{\alpha})$. Contrary to the notational convention adopted in Section 2.4 and 3.2.2 of the main article or Section C.4 of this manuscript, here the superscript “($k$)” is omitted. We make this omission since $k$ is clear from the context.

After these notational preliminaries, we now start by showing that, for an angular pm $H$ with angular density in BP form, the partial mixed derivatives $V_I(\cdot | H_k)$ can be expressed in terms of functions $q(I, \cdot, \cdot)$. Recall that $\mathcal{Y} = (0, \infty)^d$.

**Lemma D.2.** Let $k \geq d + 1$, $H_k \in \mathcal{H}_k$ and $I \subset \{1, \ldots, d\}$. Then, for every $y \in \mathcal{Y}$

\[ -V_I(y | H_k) = \|y\|_1^{-|I| - 1} \sum_{\alpha} \varphi_{\alpha} q \left( I, \alpha, \frac{y_1 \cdots y_{d-1}}{\|y\|_1} \right), \]

where $y_{1:d-1} \equiv (y_1, \ldots, y_{d-1})$ and $\alpha$ ranges over $\cup_{i=1}^{d} \{\kappa_i\} \cup \Gamma_k$.

**Proof.** We consider different types of $I$ separately.
**Case 1:** $I = \{j\}$, for some $j \in \{1, \ldots, d\}$. Some changes of variables allow to express $-V_I(y|H_k)$ as

$$
\frac{d\varphi_{kj}}{dz} + \int_{(y_j)} d||z||^{d-1} h_{k-d} \circ \pi_{\mathbb{R}} (z/||z||_1) |_{z_j=y_j} dz \\
= \frac{d\varphi_{kj}}{dz} \int_{y_j} d||z||^{d-1} h_{k-d} \circ \pi_{\mathbb{R}} (w) |_{w_j=y_j} dw |_{w_i=1}^{1-||w_i||_1} \\
= \frac{d\varphi_{kj}}{dz} \int_{y_j} d||z||^{d-1} h_{k-d} \circ \pi_{\mathbb{R}} (w) |_{w_j=t}^{1-||w_i||_1} dw |_{w_i=t} \\
$$

where $\mathbb{V}(\mathbb{Z}^d, x) = \{v \in (0,1)^d : \|v\|_1 \leq 1-x\}$, $w_{-l} = w \{1, \ldots, d\ \backslash \{l\}$ and $l = \max \{j : j \in I^c\}$. Recall that $h_{k-d}(t) = \sum_{\alpha \in \Gamma_k} \alpha_k \text{Dir}(t; \alpha)$, $t \in \mathbb{R}$. Then, the aggregation property of the Dirichlet distribution and some algebraic manipulations allow to rephrase the right-hand side as

$$
\frac{d\varphi_{kj}}{dy_j} \left( \varphi_{kj} + \sum_{\alpha \in \Gamma_k} \varphi_{\alpha} \int_{y_j/\|y\|_1}^1 \text{Dir} \left( t; \alpha \| \alpha \|_1 \right) dt \right) \\
= \frac{d\varphi_{kj}}{dy_j} \left( \varphi_{kj} + \sum_{\alpha \in \Gamma_k} \frac{\alpha_k}{\alpha_1} \frac{\alpha_{\alpha, 1} + 1, \| \alpha \|_1}{B(\alpha_1, \| \alpha \|_1)} \right) \\
= \frac{d\varphi_{kj}}{dy_j} \left( \varphi_{kj} + \sum_{\alpha \in \Gamma_k} \frac{\alpha_k}{\alpha_1} \frac{\alpha_{\alpha, 1} + 1, \| \alpha \|_1}{B(\alpha_1, \| \alpha \|_1)} \right) .
$$

The result now follows after some simple algebraic tweaking.

**Case 2:** $1 < |I| < d$. Arguments similar to the ones above allow to rephrase $-V_I(y|H_k)$ as follows

$$
\int_{(y_j)} d||z||^{d-1} h_{k-d} \circ \pi_{\mathbb{R}} (z/||z||_1) |_{z_j=y_j} dz \\
= \int_{y_j} d||z||^{d-1} h_{k-d} \circ \pi_{\mathbb{R}} (w) |_{w_j=y_j} dw |_{w_i=1}^{1-||w_i||_1} \\
= \int_{y_j} d||z||^{d-1} h_{k-d} \circ \pi_{\mathbb{R}} (w) |_{w_j=t}^{1-||w_i||_1} dw |_{w_i=t} \\
= \int_{y_j} d||z||^{d-1} \left[ \varphi_{\alpha} \int_{y_j/\|y\|_1}^1 \text{Dir} \left( t; \alpha \| \alpha \|_1 \right) dt \right] \\
= \int_{y_j} d||z||^{d-1} \left[ \varphi_{\alpha} \frac{\Pi_{i \in I} \left( \frac{\alpha_i}{\alpha_1} \right) \frac{\alpha_{\alpha, 1} + 1, \| \alpha \|_1}{B(\alpha_1, \| \alpha \|_1)} \right] \\
	imes (1-\mathbb{I}||y||_1/\|y\|_1, \| \alpha \|_1 + 1, \| \alpha \|_1) \\
= \int_{y_j} d||z||^{d-1} \left[ \varphi_{\alpha} \frac{\alpha_{\alpha, 1}}{\alpha_1} \text{Dir} \left( \frac{y_j}{\|y\|_1}; \alpha \| \alpha \|_1 \right) \right] \\
(1-\mathbb{I}||y||_1/\|y\|_1, \| \alpha \|_1 + 1, \| \alpha \|_1) ,
$$

where $I^c$ is defined as in (50). Once more, the result follows after some simple algebraic tweaking.

**Case 3:** $I = \{1, \ldots, d\}$. In this case, the result straightforwardly follows from (11). □

In the following three chained lemmas, we make use of some notions and notation introduced in Section C.1. In particular, in the next lemma, we derive the probability distribution of the random partition associated to the spectral representation of a max-stable distribution whose angular density is in BP form.
Then, for any \( \alpha \in \mathbb{R}^d \), \( \phi(\alpha) \in \Phi \) and, for \( \phi(\alpha) \in \Phi \) and, \( \tilde{\phi}(\alpha) \), the related Pickands dependence functions. Let \( \tilde{\phi}(\alpha) \) range over \( \bigcup_{i=1}^{d} \{ \kappa_i \} \cup \Gamma_k \) and
\[
\tilde{t}_v := \left( \frac{1}{v_2}, \ldots, \frac{1}{1 - \|v\|_1} \right), \quad r_v := \|v\|_1 + 1/\|v\|_1.
\]

PROOF. The change of variables
\[
y \mapsto (r, v) := \left( \|v\|_1, v_1/\|v\|_1, \ldots, v_{d-1}/\|v\|_1 \right),
\]

and few algebraic manipulations yield
\[
\int_{\mathcal{Y}} g_1(y, \mathcal{P}|H_k) dy = \int_{\mathcal{Y}} \prod_{j=1}^{m} (-V_{I_j}(y|H_k)) e^{-V(y|H_k)} dy
\]
\[
= \int_{\mathcal{Y}} \prod_{j=1}^{m} (-V_{I_j}(r_v, v(r(1-\|v\|_1)|H_k)) e^{-V(v,1-\|v\|_1|H_k)})/r_v dr_v dv
\]
\[
= \int_{\mathcal{Y}} \prod_{j=1}^{m} \sum_{\alpha(j)} \varphi_{\alpha(j)} q(I_j, \alpha(j), v) e^{-V(v,1-\|v\|_1|H_k)})/r_v dr_v dv
\]
\[
= \sum_{\alpha^{(1)}} \cdots \sum_{\alpha^{(m)}} \left( \prod_{j=1}^{m} \varphi_{\alpha(j)} \right) \int_{\mathcal{Y}} \frac{\Gamma(m) \prod_{j=1}^{m} q(I_j, \alpha(j), v)}{(A_k(t_v)r_v)^m} dv.
\]
The first equality of the statement follows from the fifth line and equation (53), with \( y \) replaced by \( (v, 1-\|v\|_1) \). The second equality is given in the sixth line.

LEMMA D.4. Let \( k \geq d + 1 \) and \( H_k \in \mathcal{H}_k \). Then, conditionally on \( H_k \), the probability of a partition \( \mathcal{P} = (I_1, \ldots, I_m) \) is
\[
\int_{\mathcal{Y}} g_1(y, \mathcal{P}|H_k) dy = \int_{\mathcal{Y}} \Gamma(m) \prod_{j=1}^{m} (-V_{I_j}(v, 1-\|v\|_1|H_k)) (A_k(t_v)r_v)^m dv
\]
\[
= \sum_{\alpha^{(1)}} \cdots \sum_{\alpha^{(m)}} \left( \prod_{j=1}^{m} \varphi_{\alpha(j)} \right) \int_{\mathcal{Y}} \frac{\Gamma(m) \prod_{j=1}^{m} q(I_j, \alpha(j), v)}{(A_k(t_v)r_v)^m} dv,
\]
where all the integer vectors \( \alpha^{(j)} \) range over \( \bigcup_{i=1}^{d} \{ \kappa_i \} \cup \Gamma_k \) and
\[
t_v := \left( \frac{1}{v_2}, \ldots, \frac{1}{1 - \|v\|_1} \right), \quad r_v := \|v\|_1 + 1/\|v\|_1.
\]

PROOF. By the second line of (54), we have that \( \sum_{\alpha^{(1)}} \varphi_{\alpha^{(1)}} s_{\alpha^{(1)}} \leq 1 \), where
\[
s_{\alpha^{(1)}} := \sum_{-l} \left( \prod_{s < l} \varphi_{\alpha^{(s)}} \right) \left( \prod_{b > l} \tilde{\varphi}_{\alpha^{(b)}} \right) \int_{\mathcal{Y}} \frac{\Gamma(m) \prod_{j=1}^{m} q(I_j, \alpha(j), v)}{(A_k(t_v)r_v)^m} dv,
\]
\( \alpha^{(1)} \in \bigcup_{i=1}^{d} \{ \kappa_i \} \cup \Gamma_k \).
Consequently, we have that $s_{\alpha^*} \leq d$, for all $\alpha^* \in \bigcup_{i=1}^{d} \{\kappa_i\} \cup \Gamma_k$. We easily prove this by contradiction. For an arbitrary choice of $\alpha^* \in \bigcup_{i=1}^{d} \{\kappa_i\} \cup \Gamma_k$, we might have chosen $\varphi_{\alpha^*}$ equal to the maximum value allowed by the mean constraint, i.e.

$$\varphi_{\alpha^*} = \frac{k}{d} \max_{1 \leq j \leq d} \varphi_{\alpha^*_j}.$$ 

If $s_{\alpha^*} > d > d \max_{1 \leq j \leq d} \varphi_{\alpha^*_j}/k$, we would have

$$1 \geq \sum_{\alpha^* \neq \alpha^*_j} \varphi_{\alpha^*_j}^{(s)} + \varphi_{\alpha^*}^{(s)} > \sum_{\alpha^* \neq \alpha^*_j} \varphi_{\alpha^*_j}^{(s)} + 1,$$

yielding a contradiction. Next, observe that each $s_{\alpha^*}$ can be written in the form

$$s_{\alpha^*} = \sum_{\alpha^*} \varphi_{\alpha^*} \tau_{\alpha^*}$$

and, once more, a contradiction argument allows to prove $\tau_{\alpha^*} \leq d^{2}$, for all $\alpha^* \in \bigcup_{i=1}^{d} \{\kappa_i\} \cup \Gamma_k$. Proceeding in this way, we can finally prove that

$$\int_{\mathcal{R}} \frac{\Gamma(m) \prod_{j=1}^{m} q(I_j, \alpha^*_j, v)}{(A_k(t_v) v)^m} \, dv \leq d^{m}, \quad \text{for all} \quad \alpha^*(1), \ldots, \alpha^*(d).$$

As a consequence, for all $l \in \{1, \ldots, d\}$ and $\alpha^*_l \in \bigcup_{i=1}^{d} \{\kappa_i\} \cup \Gamma_k$,

$$\sum_{-l} \left( \prod_{s < l} \varphi_{\alpha^*_s} \right) \left( \prod_{b > l} \varphi_{\alpha^*_b} \right) \int_{\mathcal{R}} \frac{\Gamma(m) \prod_{j=1}^{m} q(I_j, \alpha^*_j, v)}{(A_k(t_v) v)^m} \, dv \leq d^{m} \sum_{-l} \left( \prod_{s < l} \varphi_{\alpha^*_s} \right) \left( \prod_{b > l} \varphi_{\alpha^*_b} \right),$$

where the third line follows from the constraint (R1). The proof is now complete. \hfill \Box

**Lemma D.5.** Let $k \geq d + 1$ and $H_k, \tilde{H}_k \in H_k$, with coefficients $\varphi, \tilde{\varphi} \in \Phi_k$. Then, there exists a positive constant $c$ (depending on $d$) such that

$$\varphi_H^2(g_1(\cdot | H_k), g_1(\cdot | \tilde{H}_k)) \leq c \|\varphi_0 - \tilde{\varphi}_0\|_1,$$

where $\varphi_0$ and $\tilde{\varphi}_0$ are defined as in (52).

**Proof.** By Proposition C.1(i), the left-hand-side of (57) is bounded from above by

$$\|g_1(\cdot | H_k) - g_1(\cdot | \tilde{H}_k)\|_1 = \sum_{\mathcal{P} \in \mathcal{P}_d} \int_{\mathcal{Y}} \left| e^{-V(y|H_k)} \prod_{j=1}^{m} \{-V_{I_j}(y|H_k)\} - e^{-V(y|\tilde{H}_k)} \prod_{j=1}^{m} \{-V_{I_j}(y|\tilde{H}_k)\} \right| \, dy.$$ 

Thus, to establish (57), it is sufficient to show that for each $\mathcal{P} \in \mathcal{P}_d$

$$\int_{\mathcal{Y}} \left| e^{-V(y|H_k)} \prod_{j=1}^{m} \{-V_{I_j}(y|H_k)\} - e^{-V(y|\tilde{H}_k)} \prod_{j=1}^{m} \{-V_{I_j}(y|\tilde{H}_k)\} \right| \, dy \leq c_d \|\varphi_0 - \tilde{\varphi}_0\|_1,$$

where $c_d$ is a positive constant. The term on the left-hand-side can be bounded from above by

$$T_1 + T_2 := \int_{\mathcal{Y}} \left| e^{-V(y|H_k)} - e^{-V(y|\tilde{H}_k)} \prod_{j=1}^{m} \{-V_{I_j}(y|H_k)\} \right| \, dy$$

$$+ \int_{\mathcal{Y}} e^{-V(y|\tilde{H}_k)} \prod_{j=1}^{m} \{-V_{I_j}(y|H_k)\} - \prod_{j=1}^{m} \{-V_{I_j}(y|\tilde{H}_k)\} \right| \, dy.$$
We now derive upper bounds for $T_1$ and $T_2$, starting from $T_1$. Assume first that $\mathcal{P} \neq \{1, \ldots, d\}$. The change of variables in (56) together with a Lipschitz continuity argument, Proposition C.1(ii), Lemma D.2 and the first line of (54) lead to the following upper-bounds

\[
\varphi_\infty(A_k, \tilde{A}_k) \int_\mathcal{R} \int_0^\infty r^{d-1} \frac{V}{e} e^{-rV/d} \prod_{j=1}^m \{-V_{I_j}(rv, r(1-\|v\|_1)|H_k)\} dr dv \\
\leq 2d\|\varphi_0 - \tilde{\varphi}_0\| \int_\mathcal{R} \int_0^\infty r^{m-2} rVe^{-rV/d} \prod_{j=1}^m \{-V_{I_j}(v, 1-\|v\|_1)|H_k)\} dr dv \\
\leq 2d\|\varphi_0 - \tilde{\varphi}_0\|_1 mad^{m+1} \int_\mathcal{R} \Gamma(m) \prod_{j=1}^m \{-V_{I_j}(v, 1-\|v\|_1)|H_k)\} \frac{1}{(A_k(t_v)r_v)^m} dm \\
= 2d\|\varphi_0 - \tilde{\varphi}_0\|_1 mad^{m+1} \int_\mathcal{Y} g_1(y, \mathcal{P}|H_k) dy.
\]

whence we conclude $T_1 \leq 2d^{d+3}\|\varphi_0 - \tilde{\varphi}_0\|_1$. Furthermore, we have that

\[
T_2 \leq \int_\mathcal{Y} e^{-V(y|H_k)} \left| -V_{I_1}(y|H_k) \right| \prod_{j=2}^m \{-V_{I_j}(y|H_k)\} - \prod_{j=2}^m \{-V_{I_j}(y|\tilde{H}_k)\} dy \\
+ \int_\mathcal{Y} e^{-V(y|\tilde{H}_k)} \left| V_{I_1}(y|\tilde{H}_k) - V_{I_1}(y|H_k) \right| \prod_{j=2}^m \{-V_{I_j}(y|\tilde{H}_k)\} dy
\]

and recursively upper-bounding the first term on the right hand-side we finally obtain the bound from above

\[
\sum_{l=1}^m T_2^{(l)} := \sum_{l=1}^m \int_\mathcal{Y} e^{-V(y|H_k)} \left| V_{I_1}(y|H_k) - V_{I_1}(y|\tilde{H}_k) \right| \\
\times \left( \prod_{j=1}^{l-1} \{-V_{I_j}(y|H_k)\} \right) \left( \prod_{j=l+1}^m \{-V_{I_j}(y|\tilde{H}_k)\} \right) dy.
\]

For each $l \in \{1, \ldots, m\}$, the change of variables in (56) and the homogeneity of the functions $V_I(x|H_k)$, $V_I(x|\tilde{H}_k)$, $I \subseteq \{1, \ldots, d\}$, allow to re-express $T_2^{(l)}$ as

\[
\int_\mathcal{R} \int_0^\infty r^{d-1} e^{-V(v, 1-\|v\|_1|H_k)/r} \left| V_{I_1}(rv, r(1-\|v\|_1)|H_k) - V_{I_1}(rv, r(1-\|v\|_1)|\tilde{H}_k) \right| \\
\times \left( \prod_{j=1}^{l-1} \{-V_{I_j}(rv, r(1-\|v\|_1)|H_k)\} \right) \left( \prod_{j=l+1}^m \{-V_{I_j}(rv, r(1-\|v\|_1)|\tilde{H}_k)\} \right) dr dv \\
= \int_\mathcal{R} \int_0^\infty r^{m-1} e^{-V(v, 1-\|v\|_1|H_k)/r} \left| V_{I_1}(v, 1-\|v\|_1|H_k) - V_{I_1}(v, 1-\|v\|_1|\tilde{H}_k) \right| \\
\times \left( \prod_{j=1}^{l-1} \{-V_{I_j}(v, 1-\|v\|_1|H_k)\} \right) \left( \prod_{j=l+1}^m V_{I_j}(v, 1-\|v\|_1|\tilde{H}_k) \right) dr dv \\
= \int_\mathcal{R} \left| V_{I_1}(v, 1-\|v\|_1|H_k) - V_{I_1}(v, 1-\|v\|_1|\tilde{H}_k) \right| \frac{\Gamma(m)}{(A_k(t_v)r_v)^m} \\
\times \left( \prod_{j=1}^{l-1} \{-V_{I_j}(v, 1-\|v\|_1|H_k)\} \right) \left( \prod_{j=l+1}^m V_{I_j}(v, 1-\|v\|_1|\tilde{H}_k) \right) dv
\]

where, by Lemmas D.2-D.4, the term on the right-hand side is bounded from above by

\[
\sum_{\alpha^{(l)}} \left| \varphi_\alpha^{(l)} - \tilde{\varphi}_\alpha^{(l)} \right| \sum_{s \leq l} \left( \prod_{s \leq l} \varphi_\alpha^{(s)} \right) \left( \prod_{b > l} \tilde{\varphi}_\alpha^{(b)} \right) \int_\mathcal{R} \frac{\Gamma(m)}{(A_k(t_v)r_v)^m} q(I_j, \alpha^{(j)}, v) dv \\
\leq d^m \|\varphi_0 - \tilde{\varphi}_0\|_1.
\]
Therefore, \( T_2 \leq d^{d+1} \| \phi_p - \phi_q \|_1 \). The inequality (58) now follows for every \( p \neq \{1, \ldots, d\} \). Using similar arguments, the same upper bounds can be obtained for \( T_1 \) and \( T_2 \) in the simpler case of \( p = \{1, \ldots, d\} \). Hence, inequality (58) holds true also in this instance, completing the proof.

We conclude by showing that the class of simple max-stable densities with angular density in BP form is a "rich" subset of \( \mathcal{G}_1 \). Recall that \( \mathcal{G}_1 = \{g_1(\cdot|H) : H \in \mathcal{H}\} \).

**Corollary D.6.** The set \( \{g_1(\cdot|H) : H \in \cup_{k=d+1}^{\infty} \mathcal{H}_k\} \) is \( \mathcal{D}_H \)-dense in \( \mathcal{G}_1 \).

**Proof.** As a consequence of Proposition 2.3, for any \( H \in \mathcal{H} \) there exists a sequence \( H_k = \mathcal{H}_k \) such that \( \lim_{k \to \infty} \| h - h_{k-d} \|_1 \to 0 \). Then, there exists a subsequence \( h_{k_d} \) that converges to \( h \) pointwise almost-everywhere [75, Corollary 1.5.10]. We next show that also

\[
\mathcal{D}_T(G_1, G_1(\cdot|H)) \to 0
\]

as \( k \to \infty \), entailing that \( \mathcal{D}_H(g_1(\cdot|H_{k_d}), g_1(\cdot|H)) \to 0 \) and establishing the result.

In what follows, we use the notation introduced in Section C.1 of this manuscript. Define

\[
\Lambda_{k_d} := \Lambda(\cdot|H_{k_d}), \quad \Lambda := \Lambda(\cdot|H),
\]

then denote by \( N(k_d) \) and \( N \) two Poisson random measures on \( E \), with mean measures \( \Lambda_{k_d} \) and \( \Lambda \), respectively. For \( I \subset \{1, \ldots, d\} \), denote by \( \pi_I \) the projection map \( \pi_I(x) = (x_j)_{j \in I} \). Accordingly, denote by

\[
N^j(k_d) := N(k_d) \circ \pi^{-1}_j, \quad N_j := N \circ \pi^{-1}_j, \quad j = 1, \ldots, d,
\]

the marginal Poisson random measures. Moreover, let \( Y^{(k_d)} \) and \( Y \) be rv's distributed according to \( G_1(\cdot|H_{k_d}) \) and \( G_1(\cdot|H) \), respectively. For \( t > 0 \), define

\[
Y^{(k_d)}_t := (Y_1^{(k_d)} 1_{Y_1^{(k_d)} > t}, \ldots, Y_d^{(k_d)} 1_{Y_d^{(k_d)} > t})
\]

and \( Y_t \) in an analogous fashion. For a random element \( X \), let \( \mathcal{L}(X) \) denote the pertaining pm. Fix a small \( \epsilon > 0 \) and observe that by triangular inequality

\[
\mathcal{D}_T((G_1(\cdot|H_{k_d}), G_1(\cdot|H))) \leq \mathcal{D}_T(\mathcal{L}(Y^{(k_d)}), \mathcal{L}(Y^{(k_d)}_t)) + \mathcal{D}_T(\mathcal{L}(Y_t), \mathcal{L}(Y_t))
\]

(60)

Since \( G_1(\cdot|H_{k_d}) \) has unit Fréchet margins, it holds that

\[
\mathcal{D}_T(\mathcal{L}(Y^{(k_d)}), \mathcal{L}(Y^{(k_d)}_t)) \leq P(Y^{(k_d)} \neq Y^{(k_d)}_t) \leq d e^{-1/t}.
\]

Analogously, \( \mathcal{D}_T(\mathcal{L}(Y), \mathcal{L}(Y_t)) < d e^{-1/t} \). Thus, for \( t \) sufficiently small, the sum of the first two terms on the right-hand side of (60) is smaller than \( \epsilon/2 \). Next, for \( j = 1, \ldots, d \), denote

\[
N^j_{k_d} = N^j(k_d) (\cdot \cap (t, \infty)), \quad N_{j,t} = N_j (\cdot \cap (t, \infty)).
\]

For a Poisson random measure \( N^* = \sum_{t \geq 1} \delta_{X_t} \) on \( (t, \infty] \), define the random functional

\[
\phi(N^*) = \begin{cases} 
0, & \text{if } N^*((t, \infty)) = 0 \\
\max_{t \geq 1} X_t, & \text{otherwise}
\end{cases}
\]

Then, we have that \( Y^{(k_d)}_t \) and \( Y_t \) are equivalent in distribution to \( \Phi^{(k_d)}_t := \{\phi(N^{(k_d)}_{j,t})\}_{j=1}^d \) and \( \Phi_t := \{\phi(N_{j,t})\}_{j=1}^d \), respectively. Corollary 1.4.2 in [64] and arguments on pages 237–239 in [46], with a few adaptations, now yield

\[
\mathcal{D}_T(\mathcal{L}(Y^{(k_d)}_t), \mathcal{L}(Y_t)) = \mathcal{D}_T(\mathcal{L}(\Phi^{(k_d)}_t), \mathcal{L}(\Phi_t)) \leq c_d \sum_{I \subset \{1, \ldots, d\}} \mathcal{D}_T(\Lambda^{(I,t)}_{k_d} \circ \pi^{-1}_I, \Lambda^{(I,t)} \circ \pi^{-1}_I)
\]

(61)

\[
\leq c_d \sum_{I \subset \{1, \ldots, d\}} \mathcal{D}_T(\Lambda^{(I,t)}_{k_d}, \Lambda^{(I,t)})
\]
where $c_d$ is a positive constant depending only on $d$, $E_{I,t} = \{ y \in E : y_j > t, \forall j \in I \}$ and

$$\Lambda_{k_s}^{(I,t)} = \Lambda_{k_s}^{(I,t)} \cap E_{I,t}, \quad \Lambda^{(I,t)} = \Lambda \cap E_{I,t}.$$

Observe that each $E_{I,t}$ is relatively compact, thus $\Lambda_{k_s}^{(I,t)}$ and $\Lambda^{(I,t)}$ are finite measures. Moreover,

$$\Lambda_{k_s}^{(I,t)}(E_{I,t} \cap E_{1,\ldots,d}) = d \int_{S(1,\ldots,d)} \min(w_j / t) H_{k_s}(dw),$$

$$\Lambda^{(I,t)}(E_{I,t} \cap E_{1,\ldots,d}) = d \int_{S(1,\ldots,d)} \min(w_j / t) H(dw),$$

therefore it can be easily seen that $\Lambda_{k_s}^{(I,t)}(E_{I,t} \cap E_{1,\ldots,d}) \rightarrow \Lambda^{(I,t)}(E_{I,t} \cap E_{1,\ldots,d})$ as $k_s \rightarrow \infty$. Consequently, by Scheffé’s lemma, as $k_s \rightarrow \infty$,

$$|\mathcal{D}_r(\Lambda_{k_s}^{(I,t)}, \Lambda^{(I,t)})| \leq \frac{d^2}{r} \|h_{k_s} - h\|_1 + \int_{E_{I,t} \cap E_{1,\ldots,d}} d\|z\|_1^{d-1} |h_{k_s} \circ \pi_R(z \mid |z|_1) - h \circ \pi_R(z \mid |z|_1)|dz \rightarrow 0.$$ 

Due to (61), we now deduce that, for large enough $k_s$, also the third term on the right-hand side of (60) is smaller than $\epsilon/2$ and $\mathcal{D}_r((\mathcal{G}_1(\{ |H_k \}), \mathcal{G}_1(\{ |H \})) < \epsilon$. The result in (39) now follows and the proof is complete. \hfill \Box

\textbf{D.3.5. Proof of Theorem 3.9.} Firstly, we verify the conditions of Theorem 6.23 in [39] to establish the result at point (a). Then, the one at point (b) is deduced from the latter. Finally, the claim concerning the support of $\Pi_{I_1}$ is proved.

\textbf{Kulback-Leibler property.} By assumption, $H_0 \in \mathcal{H}_0$. Also, by Lemma D.1, for all $H \in \mathcal{H}'$ and $\delta > 0$, there exists a sequence $H_k \in \mathcal{H}_k$, with weights $\varphi^{(k)} \in \Phi_k$, such that, defining

$$B_k := \left\{ H_k \in \mathcal{H}_k : 2(k - 1)! \| \varphi^{(k)} - \varphi^{(k)} \|_\infty < \delta \right\},$$

as $k \rightarrow \infty$ we have

$$B_k \subset B_{\delta/2,\infty}(H_k) \subset B_{\delta,\infty}(H).$$

By Condition 3.8, $\Pi_{I_1}(B_k) > 0$, hence the prior $\Pi_{I_1}$ posses the $\mathcal{D}_\infty$-property at $H$ (Definition 3.5 of the main article). Therefore, by Theorem 3.7, we conclude that $\Pi_{G_{I_1}}(K_0) > 0$, for all $\epsilon > 0$, where $K_0$ is as in Definition 3.2 of the main paper. That is, $g_1(\{ |H_0 \})$ is in the Kulback-Leibler support of the prior $\Pi_{G_{I_1}}$.

\textbf{Metric entropy.} Observe that $\mathcal{D}_H$ is a metric that generates convex balls and define

$$\mathcal{G}_1^{(k)} := \{ g_1(\{ |H_k \}) : H_k \in \mathcal{H}_k; \mathcal{D}_H(g_1(\{ |H_k \}), g_1(\{ |H_0 \})) > 4\epsilon \},$$

$$\mathcal{G}_{1,n} := \bigcup_{k=d+1}^{\nu_n} \mathcal{G}_1^{(k)},$$

$$\mathcal{G}_{1,n,2} := \mathcal{G}_{1,n,1} \setminus \mathcal{G}_{1,n,1},$$

where $\nu_n$ is a sequence of positive integers and $\mathcal{G}_{1,n} := \{ g_1(\{ |H \}) : H \in \mathcal{H} \}$. Then, by Lemma D.5 and Proposition C.2 in [39] we have

$$N(2\epsilon, \mathcal{G}_{1,n,1}^{(k)} \cap \mathcal{D}_H) \leq N\left(\epsilon', \left\{ x \in \mathbb{R}^{[d]} : \|x\|_1 \leq 1 \right\}, L_1 \right) \leq \left(3/\epsilon'\right)^{k-1} \left(\epsilon'\right)^{d-1},$$

where $\epsilon'$ is a positive global constant and, without loss of generality, we assume $\epsilon' < 1$. As a consequence, we also have that

$$N(2\epsilon, \mathcal{G}_{1,n,1} \cap \mathcal{D}_H) \leq \sum_{k=d+1}^{\nu_n} \left(3/\epsilon'\right)^{k-1} \left(\epsilon'\right)^{d-1} \leq \nu_n \left(3/\epsilon'\right)^{d-1} \left(3/\epsilon'\right)^{d-1},$$

and choosing $\nu_n = \lfloor (n\epsilon^2)^{1/(d-1)} \rfloor$, with $\epsilon'' = (2 \log(3/\epsilon''))^{-1/(d-1)}$, for all large $n$ we have

$$\log N(2\epsilon, \mathcal{G}_{1,n,1} \cap \mathcal{D}_H) = \log \nu_n + \nu_n^{d-1} \log(3/\epsilon' \epsilon^2) \leq n\epsilon^2.$$
The first condition of Theorem 6.23 in [39] is therefore satisfied. The second condition therein is satisfied by assumption (ii) in Condition 3.8.

**Conclusion.** The result at point (a) in the statement of Theorem 3.9 follows from the above considerations. For the general case \( d \geq 2 \), the result at point (b) is a direct consequence of the one at point (a) and Proposition C.1(iii). For the specific case \( d = 2 \), the result follows from an application of Propositions C.1(i) and C.2(ii).

As for the claim on the full support of \( \Pi_{G_{1}} \), by Corollary D.6 we have that for any \( \epsilon > 0 \) and \( g_1(\cdot|H) \in G_{1} \) there exists \( g_1(\cdot|H) \in \{ g_1(\cdot|H) : H \in \bigcup_{i=d+1}^{\infty} \mathcal{H}_i \} \) such that \( \mathcal{D}_H(g_1(\cdot|H),g_1(\cdot|\tilde{H}_k)) < \epsilon/2 \). Thus, by Lemma D.5 and Condition 3.8, we have that

\[
\Pi_{G_{1}}(g \in G_1 : \mathcal{D}_H(g,g_1(\cdot|\tilde{H})) < \epsilon) \\
\geq \Pi_{G_{1}}(g \in G_1 : \mathcal{D}_H(g,g_1(\cdot|\tilde{H})) < \epsilon/2) \\
\geq \Pi_{G_{1}}(g \in \{ g_1(\cdot|H) : H \in \mathcal{H}_k \} : \mathcal{D}_H(g,g_1(\cdot|\tilde{H})) < \epsilon/2) \\
\geq \lambda(k)\nu_k(\varphi \in \Phi_k : \|\varphi - \tilde{\varphi}\|_1 < \epsilon/2c) \\
> 0.
\]

where \( \tilde{\varphi} \) is the vector of linear coefficients corresponding to \( \tilde{H}_{k-d} \) and \( c \) is a positive constants. In passing, observe that the set on the second line is \( \mathcal{D}(G_{1}) \)-measurable, since it is the intersection of the measurable set \( \{ g_1(\cdot|H) : H \in \mathcal{H}_k \} \) and the Hellinger ball of radius \( \epsilon/2 \) around \( g_1(\cdot|\tilde{H}_k) \). The proof is now complete.

**D.4. Proofs of the results in Section 3.3.**

**D.4.1. Proof of Proposition 3.12.** The proof is organised in two parts: set theoretical preliminaries and main body of the proof.

**Set theoretical preliminaries.** Denote

\[
\phi_{H \times \Theta} : (\mathcal{H} \times \Theta, \mathcal{D}_H \otimes \mathcal{D}_\Theta) \mapsto (\mathcal{G}_{\Theta}, \mathcal{G}_{\Theta}) : (H, \vartheta) \mapsto g_{\vartheta}(\cdot|H),
\]

where \( \mathcal{D}_H, \mathcal{D}_\Theta \) and \( \mathcal{G}_{\Theta} \) are the Borel \( \sigma \)-field induced on \( \mathcal{H}, \Theta \) and \( \mathcal{G}_{\Theta} \) by the metrics \( \mathcal{D}_W, L_1 \) and \( \mathcal{D}_H \), respectively. Let \( \mathcal{U}_L \) be any \( \mathcal{D}_H \)-neighbourhood of \( g_{\vartheta_0}(\cdot|H_0) \), \( \mathcal{U}_1 \) be any \( \mathcal{D}_W \)-neighborhood (if \( d > 2 \)) or \( \mathcal{D}_K \)-neighborhood (if \( d=2 \)) of \( H_0 \) and \( \mathcal{U}_2 \) be any \( L_1 \)-neighborhood of \( \vartheta_0 \). We next derive a preliminary upper bound for the term

\[
\max \left\{ \Pi_n(\mathcal{U}_L^G), \Pi_n((\mathcal{U}_1 \times \mathcal{U}_2)^G) \right\}
\]

using some simple set theoretical arguments.

Observe that, for some sufficiently small \( \epsilon', \delta > 0 \), \( B_{\epsilon', \varphi}(H_0) := \{ H \in \mathcal{H} : \mathcal{D}(H,H_0) < \epsilon' \} \subset \mathcal{U}_1 \), where \( \mathcal{D} = \mathcal{D}_K \) if \( d > 2 \) and \( \mathcal{D} = \mathcal{D}_W \) if \( d > 2 \) and \( B_{\delta, \varphi}(\vartheta_0) \subset \mathcal{U}_2 \). Therefore,

\[
\Pi_n((\mathcal{U}_1 \times \mathcal{U}_2)^G) \leq \Pi_n \left( B_{\epsilon', \varphi}(H_0) \times B_{\delta, \varphi}(\vartheta_0) \right) + \Pi_n \left( \mathcal{H} \times B_{\delta, \varphi}(\vartheta_0) \right).
\]

Moreover, notice that by Propositions C.1-C.2, for some \( \epsilon'' > 0 \),

\[
B_{\epsilon''} := \{ H \in \mathcal{H} : \mathcal{D}(g_{\vartheta}(\cdot|H),g_1(\cdot|H_0)) < 2\epsilon'' \} \subset B_{\epsilon', \varphi}(H_0).
\]

Furthermore, by continuity of the map \( \vartheta \mapsto g_{\vartheta}(\cdot|H_0) \) with respect to the \( L_1 \) and the Hellinger metrics on a neighbourhood of \( \vartheta_0 \), we can choose \( \delta \) small enough to guarantee that for all \( \vartheta \in B_{\delta, \varphi}(\vartheta_0) \)

\[
\mathcal{D}(g_{\vartheta}(\cdot|H_0),g_{\vartheta_0}(\cdot|H_0)) < \epsilon''.
\]

Hence, reverse triangular inequality entails that for all \( (H, \vartheta) \in B_{\epsilon''} \times B_{\delta, \varphi}(\vartheta_0) \)

\[
\mathcal{D}(g_{\vartheta}(\cdot|H),g_{\vartheta_0}(\cdot|H_0)) \geq \mathcal{D}(g_{\vartheta}(\cdot|H),g_{\vartheta}(\cdot|H_0)) - \mathcal{D}(g_{\vartheta}(\cdot|H_0),g_{\vartheta_0}(\cdot|H_0)) \\
= \mathcal{D}(g_1(\cdot|H),g_1(\cdot|H_0)) - \mathcal{D}(g_{\vartheta}(\cdot|H_0),g_{\vartheta_0}(\cdot|H_0)) \\
> \epsilon'',
\]

from which we conclude that

\[
B_{\epsilon', \varphi}(H_0) \times B_{\delta, \varphi}(\vartheta_0) \subset B_{\epsilon''} \times B_{\delta, \varphi}(\vartheta_0) \subset \phi_{\mathcal{H} \times \Theta}^{-1}(\mathcal{U}_L^G),
\]
where $\epsilon = \epsilon''/4$ and $\tilde{U}_c := \{ g \in G_{\Theta} : \mathcal{H}(g, g_{\theta_0}^{-} | \{H_0\}) \leq 4\epsilon \}$. As a result,

$$\Pi_n \left( B_{\epsilon, \delta}(H_0) \times B_{\delta, 1}(\theta_0) \right) \leq \tilde{\Pi}_n(\tilde{U}_c^c).$$

(66)

In particular, we can choose $\epsilon''$, and thus $\epsilon$, small enough to also guarantee that $\tilde{U}_c \subset \tilde{U}$. Therefore,

$$\Pi_n(\tilde{U}_c^c) \leq \Pi_n(\tilde{U}_c^c)$$

(67)

Moreover, denoting for $j$

$$\phi^{-1}_{H \times \Theta}(\tilde{U}_c^c) \cap \{ H \times B_{\delta, 1}(\theta_0) \} + \Pi_n \left( H \times B_{\delta/d, \infty}(\theta_0) \right).$$

(68)

By combining the inequalities in (65)-(67) we finally deduce that

$$\max \left\{ \Pi_n(\tilde{U}_c^c), \Pi_n((U_1 \times U_2)^c) \right\} \leq \Pi_n \left( \phi^{-1}_{H \times \Theta}(\tilde{U}_c^c) \cap \{ H \times B_{\delta, 1}(\theta_0) \} \right) + 2\Pi_n \left( H \times B_{\delta/d, \infty}(\theta_0) \right).$$

(69)

We now analyse the terms on the right-hand side of (68). The first term can be decomposed as

$$\Pi_n \left( \phi^{-1}_{H \times \Theta}(\tilde{U}_c^c) \cap \{ H \times B_{\delta, 1}(\theta_0) \} \right) = \tilde{\Pi}_n(G_{\Theta,n}) + \tilde{\Pi}_n(G_{\Theta,n}^c),$$

(70)

where $G_{\Theta,n}$ is any sequence of measurable subsets of $\tilde{U}_c^c \cap \phi_{H \times \Theta} (\{ H \times B_{\delta, 1}(\theta_0) \})$

and $G_{\Theta,n}^c$ is the relative complement of $G_{\Theta,n}$ in the set above. Let $X_{1:n} := (X_i)_{i=1}^n$ and denote by $R_n(X_{1:n}; \Pi_{H \times \Theta})$ the ratio between the marginal likelihood and the likelihood at $(H_0, \theta_0)$, i.e.

$$R_n(X_{1:n}; \Pi_{H \times \Theta}) := \int_{H \times \Theta} \prod_{i=1}^n \frac{g(X_i|H)}{g_0(X_i|H_0)} \, d\Pi_{H \times \Theta}(H, \theta)$$

$$= \int_{G_{\Theta,n}} \prod_{i=1}^n \frac{g(X_i)}{g_0(X_i|H_0)} \, d\Pi_{G_{\Theta}}(g).$$

and let $\tau_n = (s_n, t_{n,1}, \ldots, t_{n,d})$ be any collection of measurable functions $s_n : \mathbb{R}^{nd} \to [0, 1]$ and $t_{n,j} : \mathbb{R}^n \to [0, 1], j = 1, \ldots, d$. Then, on one hand, we have the inequality

$$\tilde{\Pi}_n(G_{\Theta,n}) \leq s_n(X_{1:n}) + (1 - s_n(X_{1:n}))\tilde{\Pi}_n(G_{\Theta,n})$$

(71)

$$= s_n(X_{1:n}) + \frac{1 - s_n(X_{1:n})}{R_n(X_{1:n}; \Pi_{H \times \Theta})} \int_{G_{\Theta,n}} \prod_{i=1}^n \frac{g(X_i)}{g_0(X_i|H_0)} \, d\Pi_{G_{\Theta}}(g).$$

on the other hand, we have the identity

$$\tilde{\Pi}_n(G_{\Theta,n}^c) = \frac{1}{R_n(X_{1:n}; \Pi_{H \times \Theta})} \int_{G_{\Theta,n}^c} \prod_{i=1}^n \frac{g(X_i)}{g_0(X_i|H_0)} \, d\Pi_{G_{\Theta}}(g).$$

Moreover, denoting for $j = 1, \ldots, d, X_{j,1:n} = (X_{j,i})_{i=1}^n$ and

$$D_{n,j} := H \times \{ \theta \in \Theta : ||\theta - \theta_{0,j}||_{\infty} > \delta/d \}.$$

(72)
we can observe that
\[ \Pi_n \left( \mathcal{H} \times B_{\mathcal{P}/d,\infty}(\vartheta_0) \right) \leq \sum_{j=1}^{d} \Pi_n(D_{n,j}) \]
\[ \leq \sum_{j=1}^{d} t_{n,j}(X_{j,1:n}) \leq \sum_{j=1}^{d} (1 - t_{n,j}(X_{j,1:n})) \Pi_n(D_{n,j}) \]
\[ = \sum_{j=1}^{d} t_{n,j}(X_{j,1:n}) \leq \sum_{j=1}^{d} \frac{1 - t_{n,j}(X_{j,1:n})}{\Pi_n(X_{1:n}; \Pi_{\mathcal{H} \times \Theta})} \int_{D_{n,j}}^{\sum_{i=1}^{n}} g_{\vartheta_0}(X_{i}|H_0) \Pi_{\mathcal{H} \times \Theta}(H, \vartheta) \, d\Pi_{\mathcal{H} \times \Theta}(H, \vartheta). \]

For all choices of measurable functions \( \tau = (s, t_1, \ldots, t_d) \), with \( s : \mathbb{R}^{nd} \rightarrow [0, 1] \) and \( t_j : \mathbb{R}^{n_j} \rightarrow [0, 1] \), \( j = 1, \ldots, d \), and Borel pm \( P \) on \( (\mathcal{H} \times \Theta, \mathcal{B}_\mathcal{H} \otimes \mathcal{B}_{\Theta}) \), define the functional
\[ \Xi_n(X_{1:n}, \tau, P) := \Xi_{n,1}(X_{1:n}, \tau) + \Xi_{n,2}(X_{1:n}, \tau, P), \]
where
\[ \Xi_{n,1}(X_{1:n}, \tau) := s_n(X_{1:n}) + 2 \sum_{j=1}^{d} t_{n,j}(X_{j,1:n}) \]
and
\[ \Xi_{n,2}(X_{1:n}, \tau, P) := (1 - s_n(X_{1:n})) \int_{\phi_{\mathcal{H} \times \Theta}}^{-1} n \prod_{i=1}^{n} g_{\vartheta_0}(X_{i}|H_0) \Pi_{\mathcal{H} \times \Theta}(H_0) \, dP(H, \vartheta) \]
\[ + \int_{\phi_{\mathcal{H} \times \Theta}}^{-1} n \prod_{i=1}^{n} g_{\vartheta_0}(X_{i}|H_0) \Pi_{\mathcal{H} \times \Theta}(H_0) \, dP(H, \vartheta) \]
\[ + 2 \sum_{j=1}^{d} (1 - t_{j}(X_{j,1:n})) \int_{D_{n,j}}^{\sum_{i=1}^{n}} g_{\vartheta_0}(X_{i}|H_0) \Pi_{\mathcal{H} \times \Theta}(H, \vartheta) \, d\Pi_{\mathcal{H} \times \Theta}(H, \vartheta). \]

Combining (68)-(73), we can now conclude that
\[ \max \left\{ \Pi_n(\tilde{U}^C), \Pi_n(\tilde{(H_1 \times H_2)}^C) \right\} \leq \Xi_{n,1}(X_{1:n}, \tau_n) + \frac{\Xi_{n,2}(X_{1:n}, \tau_n, \Pi_{\mathcal{H} \times \Theta})}{R_n(X_{1:n}; \Pi_{\mathcal{H} \times \Theta})}. \]

Moreover, by the Kullback-Leibler property, for any \( c > 0 \) we have that \( R_n(X_{1:n}; \Pi_{\mathcal{H} \times \Theta}) \geq e^{-nc} \) eventually almost surely as \( n \to \infty \), hence
\[ \Xi_{n,1}(X_{1:n}, \tau_n) + \frac{\Xi_{n,2}(X_{1:n}, \tau_n, \Pi_{\mathcal{H} \times \Theta})}{R_n(X_{1:n}; \Pi_{\mathcal{H} \times \Theta})} \leq e^{nc} \Xi_n(X_{1:n}, \tau_n, \Pi_{\mathcal{H} \times \Theta}). \]

Without loss of generality, we can assume that \( \delta < \delta_s \), with \( \delta_s \) as in Condition 3.11, and select \( \tilde{G}_{\Theta,n} \) and \( \tau_n \) satisfying the properties therein. By hypothesis, as \( n \to \infty \)
\[ \mathbb{E} \left\{ \Xi_{n,1}(X_{1:n}, \tau_n) \right\} \leq e^{-nc} + \sum_{j=1}^{d} e^{-nc(j)} \delta. \]

Moreover, an application of Fubini's theorem yields that as \( n \to \infty \)
\[ \mathbb{E} \left\{ \Xi_{n,2}(X_{1:n}, \tau_n, \Pi_{\mathcal{H} \times \Theta}) \right\} = \int_{\phi_{\mathcal{H} \times \Theta}}^{\sum_{i=1}^{n}} \left\{ 1 - s_n(x_{1:n}) \right\} dG^{(0)}(x_{1:n}) \, d\Pi_{\mathcal{H} \times \Theta}(g) + \Pi_{\mathcal{H} \times \Theta}(\tilde{G}_{\Theta,n}^{C}) \]
\[ + 2 \sum_{j=1}^{d} \int_{D_{n,j}}^{\sum_{i=1}^{n}} \left\{ 1 - t_{n,j}(x_{j,1:n}) \right\} dG^{(0)}_{\vartheta_0}(x_{j,1:n}) \, d\Pi_{\mathcal{H} \times \Theta}(H, \vartheta) \]
\[ \leq e^{-2nc} + e^{-rn} + \sum_{j=1}^{d} e^{-nc(j)} \delta, \]
where, for $g \in G_{\Theta,n}$ and $(H, \Theta) \in D_{n,j}$, $j = 1, \ldots, d$, $G_{\Theta}^{(n)}$ and $G_{\Theta,j}^{(n)}$ are the $n$-fold products of the pm $G$ with density $g$ and of the $j$-th marginal pm of $G_{\Theta}^{(n)}(\cdot | H)$, respectively. Consequently, choosing $c$ such that
\begin{equation}
2c < \frac{1}{2} \min \{ e^2, r, c_1(\delta), \ldots, c_d(\delta) \} =: c',
\end{equation}
an application of Markov’s inequality allows to deduce that for all $\varepsilon > 0$, as $n \to \infty$,
\begin{equation}
G_{\Theta_0} \left( e^{2cn} \Xi_n(X_{1:n}, \tau_n, \Pi_{H \times \Theta}) > \varepsilon \right) \leq e^{-nc'}.
\end{equation}
The result now follows from (77), (78) and (82).

\begin{remark}
The proof of Proposition 3.12 adapts arguments from the proofs of Theorems 6.17 and 6.23 in [39]. The key derivations yielding equations (12)-(13) of the main paper are discussed in detail, as they also play a crucial role in establishing the results of Section 4. In the present context of well-specified max-stable models, Proposition 3.12 allows to obtain a common form of consistency at an exponential rate [e.g., 13, Definition 3.1].
\end{remark}

\begin{corollary}
Under the assumptions of Proposition 3.12, for every $T_{H}$-neighborhood $U$ of $g_{\Theta_0}(|H_0)$, $T_{W}$-neighborhood (if $d \geq 2$) or $T_{KS}$-neighborhood (if $d=2$) $U_{1}$ of $H_{0}$ and $L_{1}$-neighborhood $U_{2}$ of $\Theta_{0}$ there exists a constant $c > 0$ such that as $n \to \infty$ eventually $G_{\Theta_0}^{(n)}(|H_0) \rightarrow$
\begin{equation}
\bar{\Pi}_n(U_{1} \times U_{2}) < e^{-nc}, \quad \Pi_n((U_{1} \times U_{2})_{n}) < e^{-nc}.
\end{equation}
\end{corollary}

\begin{proof}
Defining the events
\begin{align*}
E_{n,1} & := \left\{ \bar{\Pi}_n(U_{1}^{n}) > e^{-nc} \right\} \cup \left\{ \Pi_n((U_{1} \times U_{2})_{n}) > e^{-nc} \right\}, \\
E_{n,2} & := \left\{ \max \left[ \bar{\Pi}_n(U_{1}^{n}), \Pi_n((U_{1} \times U_{2})_{n}) \right] \leq e^{nc} \right\}, \\
E_{n,3} & := \left\{ \Xi_n((X_{i})_{i=1}^{n}, \tau_n, \Pi_{H \times \Theta}) > e^{-2nc} \right\}
\end{align*}
with $\Xi_n(\cdot, \cdot, \cdot)$ and $e$ as in Proposition 3.12, we have that
\begin{equation}
G_{\Theta_0}^{(n)} \left( \limsup_{n \to \infty} E_{n,1} \big| H_0 \right) \leq 1 - G_{\Theta_0}^{(n)} \left( \liminf_{n \to \infty} E_{n,2} \big| H_0 \right) + G_{\Theta_0}^{(n)} \left( \limsup_{n \to \infty} E_{n,3} \big| H_0 \right).
\end{equation}
On one hand, since (12) holds eventually $G_{\Theta_0}^{(n)}(|H_0) \rightarrow$ as for $n \to \infty$,
\begin{equation}
G_{\Theta_0}^{(n)} \left( \liminf_{n \to \infty} E_{n,2} \big| H_0 \right) = 1.
\end{equation}
On the other hand, by (13) and Borel-Cantelli lemma,
\begin{equation}
G_{\Theta_0}^{(n)} \left( \limsup_{n \to \infty} E_{n,3} \big| H_0 \right) = 0.
\end{equation}
The result now follows.
\end{proof}

\begin{remark}
We stress that, once $\varepsilon, \delta$ and an arbitrary measurable subset $G_{\Theta,n}$ are fixed, functionals $\Xi_{n,1}$ and $\Xi_{n,2}$ can be defined as in (76)-(75) and the inequalities in (71) and (73) are satisfied for all choices of measurable functions $\tau_n = (s_n, t_{n,1}, \ldots, t_{n,d})$, with
\begin{equation}
s_n : \mathbb{R}^{nd} \mapsto [0, 1], \quad t_{n,j} : \mathbb{R}^n \mapsto [0, 1], \quad j = 1, \ldots, d,
\end{equation}
even in the case where $\Pi_{H \times \Theta}$ is replaced by a possibly random pm $P_0$ on $(H \times \Theta, \mathcal{F}_H \otimes \mathcal{F}_\Theta)$ and $X_{1:n} = (X_1, \ldots, X_n)$ is a sample of possibly non max-stable rv’s such that, with probability 1,
\begin{equation}
\prod_{i=1}^{n} g_{\Theta_0}(X_i | H_0) > 0, \quad R_n(X_{1:n}; \Theta_0) > 0.
\end{equation}
\end{remark}
Therefore, replacing $\Pi_{\mathcal{H} \times \Theta}$ with $P_0$ on the right-hand side of (77), the validity of the latter bound extends to the case of more general random probabilities

$$\Pi_n((U_1 \times U_2)^C) := \frac{\int_{U_1 \times U_2} \prod_{i=1}^n g_i (X_i | H) d P_0 (H, \theta)}{\int_{\mathcal{H} \times \Theta} \prod_{i=1}^n g_i (X_i | H) d P_0 (H, \theta)}, \quad \tilde{\Pi}_n (U^C) := \Pi_n \circ \phi^{-1}_{\mathcal{H} \times \Theta} (U^C).$$

This fact is used in the proofs of Theorems 4.7, 4.15 and 4.19. We also point out that $\Xi_n$ tacitly depends on $\epsilon$ and $\delta$ via the sets $G_{n, \rho}$ (depending on $\epsilon$ and $\delta$ via the set in (70)) and $D_{n,1}, \ldots, D_{n,d}$ (depending on $\delta$). We recall that $\epsilon$ and $\delta$ are determined in the first place using some set theoretical arguments which move from the choices of $U_1, U_1$ and $U_2$. Thus, the functional $\Xi_n$ ultimately depends on $U_1, U_1$ and $U_2$. This consideration makes precise the remark following inequality (12) in the main paper.

**D.4.2. Auxiliary results for the proof of Theorem 3.14.** The results presented next provide the mathematical ground for showing that, for $\rho$-Fréchet statistical models, if Conditions 3.13(ii)–(iii) and 3.13(iv) in the main text hold true, Condition 3.11 is then satisfied. This conclusion is used to prove the posterior consistency results in Theorem 3.14. There are also implications for consistency in the case where data are componentwise maxima and a $\rho$-Fréchet statistical model is misspecified (Theorem 4.7), as highlighted below.

We start by stating a sequence of chained results culminating in Corollary D.13, which guarantees that the prior specification outlined in Theorem 3.14 satisfies the Kullback-Leibler property. We make use of the following notation. For any $y = (y_1, \ldots, y_d) > 0$ and $c = (c_1, \ldots, c_d) \in \mathbb{R}^d$, we denote

$$y^c = (y_1^{c_1}, \ldots, y_d^{c_d}).$$

For the sake of a lighter notation, all the intervals of the form $\bigtimes_{j=1}^d (x_j, y_j)$, for $x, y \in [-\infty, \infty]^d$, $x < y$, are denoted by $(x, y)$. We recall that, for $(\rho, \sigma) \in (0, \infty)^{2d}$ and $H \in \mathcal{H}$, the density $g_{\rho, \sigma}(\cdot | H)$ is given by

$$g_{\rho, \sigma}(y | H) = g_1 (\langle y/\sigma \rangle^\rho | H) \prod_{j=1}^d \rho_j y_j^{\rho_j - 1} \sigma_j^{-\rho_j}, \quad y \in (0, \infty).$$

For $l \in \mathbb{N}_+$, we define

$$\gamma (l, +) := \int_0^\infty x^l e^{-x} e^{-e^{-x}} dx, \quad \gamma (l, -) := \int_0^\infty x^l e^{-x} e^{e^{-x}} dx.$$

and

$$v_1 := \Gamma (2l + 1) \int_1^\infty \left( \log (s) \right)^l \text{Gamma} (s; 2l + 1, 1) ds,$$

where Gamma$(s; a, b), s > 0,$ denotes the Gamma probability density function with shape parameter $a$ and scale parameter $b$. Notice that if $X$ is a rv with standard Gumbel distribution and we denote $X^+ = \max (X, 0), X^- = \max (-X, 0)$, then $\gamma (l, +) = E[X^+]^l$ and $\gamma (l, -) = E[X^-]^l$. Of particular interest for later derivations are the cases $l = 1, \ldots, 4$, where

$$\begin{align*}
\gamma (1, +) &= \gamma - \text{Ei} (-1), \\
\gamma (1, -) &= \text{Ei} (-1), \\
\gamma (2, +) &= 2 \, s_2 F_3 (1, 1, 1; 2, 2, 2; -1), \\
\gamma (2, -) &= 2 G^{3, 0}_{2, 3} (1 \mid \begin{array}{c} 1,1,0,0,0 \end{array}), \\
\gamma (3, +) &= 6 \, s_4 F_4 (1, 1, 1, 1; 2, 2, 2, 2; -1), \\
\gamma (3, -) &= 6 G^{4, 0}_{3, 4} (1 \mid \begin{array}{c} 1,1,0,0,0,0 \end{array}), \\
\gamma (4, +) &= 24 \, s_4 F_5 (1, 1, 1, 1, 1; 2, 2, 2, 2, 2; -1), \\
\gamma (4, -) &= 24 G^{5, 0}_{4, 5} (1 \mid \begin{array}{c} 1,1,1,1,0,0,0,0,0 \end{array}),
\end{align*}$$

and

$$\begin{align*}
v_1 &= \frac{4}{e} - 2 \text{Ei} (-1), \\
v_2 &= 2 G^{4, 0}_{3, 4} (1 \mid \begin{array}{c} 1,1,1,0,0,0,0,5 \end{array}), \\
v_3 &= 6 G^{5, 0}_{4, 5} (1 \mid \begin{array}{c} 1,1,1,1,1,0,0,0,0,0 \end{array}), \\
v_4 &= 24 G^{6, 0}_{5, 6} (1 \mid \begin{array}{c} 1,1,1,1,1,1,0,0,0,0,0,9 \end{array}),
\end{align*}$$

with $\gamma$ denoting the Euler-Mascheroni constant, $\text{Ei}(z), z \in \mathbb{R} \setminus \{0\}$, denoting the exponential integral function, $G^{\nu}_{\mu, \lambda} (z | \nu_1, \ldots, \nu_\nu; \lambda_1, \ldots, \lambda_\lambda), z \in \mathbb{R} \setminus \{0\}$, denoting the Meijer $G$-function and $p_4 F_q (a_1, \ldots, a_p; b_1, \ldots, b_q; z), z > 0,$
denoting the generalised hypergeometric function. See Corollary D.13 and Lemma D.22. Finally, for \( l \in \mathbb{N}_+ \), \( H_*, H, \tilde{H} \in \mathcal{H} \) and \( \rho_*, \rho, \sigma_*, \sigma, \tilde{\sigma} \in (0, \infty) \), we define

\[
gamma_{H_*, \rho_*, \sigma_*}^{(l)}(H, \rho, \sigma; \tilde{H}, \tilde{\rho}, \tilde{\sigma}) := \left( \int_{(0, \infty)} \log^+ \left( \frac{g_{\rho_*, \sigma_*}(y|H)}{g_{\tilde{\rho}, \tilde{\sigma}}(y|\tilde{H})} \right) y^l \, \sigma \right)^{1/l}.
\]

Notice that the above functional is analogous to that introduced in Section C.3 in the context of simple max-stable distributions. In particular,

\[
gamma_{H, \rho, \sigma}^{(l)}(H, \rho, \sigma; \tilde{H}, \tilde{\rho}, \tilde{\sigma}) := \left( \chi_{+}^{(l)}(g_{\rho, \sigma}(\cdot|H), g_{\tilde{\rho}, \tilde{\sigma}}(\cdot|\tilde{H})) \right)^{1/l}
\]

where the term on the right-hand side denotes the \( l \)-th root of the \( l \)-th order positive Kullback-Leibler divergence from \( g_{\tilde{\rho}, \tilde{\sigma}}(\cdot|\tilde{H}) \) to \( g_{\rho, \sigma}(\cdot|H) \).

**Lemma D.10.** Let \( H \in \mathcal{H} \), with exponent function \( V(\cdot|H) \). Then, for any \( l \in \mathbb{N}_+ \) and \( \tilde{H} \in \mathcal{H} \), \( \rho \in (0, \infty) \), \( \tilde{\rho} \in B_{\delta_1, \infty}(\rho) \), \( \sigma \in (0, \infty) \), \( \tilde{\sigma} \in B_{\delta_2, \infty}(\sigma) \), with

\[
\delta_1 \in \left( 0, \min_{1 \leq j \leq d} \rho_j \varepsilon \right), \quad \delta_2 \in \left( 0, \min_{1 \leq j \leq d} \sigma_j \varepsilon \right), \quad \varepsilon \in (0, 1/2],
\]

we have that

\[
\left( \int_{(0, \infty)} \left| V(y^{\tilde{\rho}/\rho}(\sigma/\tilde{\sigma}) \tilde{\rho}|H) - V(y|H) \right| g_1(y|\tilde{H}) \, dy \right)^{1/l}
\]

is bounded from above by

\[
\frac{\{v_l + \Gamma(1 + l)\}^{1/l}}{\min_{1 \leq j \leq d} \rho_j} ||\rho - \tilde{\rho}||_1 + \{\Gamma(1 + 2l)\}^{1/l} \max_{1 \leq j \leq d} \left\{ \frac{3\rho_j}{\sigma_j} \left( \frac{3}{2} \right)^{\frac{3}{2}} \right\} ||\sigma - \tilde{\sigma}||_1.
\]

**Proof.** For \( x > 0 \), the function \( V(1/x|H) \) defines a D-norm, see [31]. Property (4.37) therein together with reverse triangle inequality yield

\[
\left| V(y^{\tilde{\rho}/\rho}(\sigma/\tilde{\sigma}) \tilde{\rho}|H) - V(y|H) \right| \leq \|y^{\tilde{\rho}/\rho}(\sigma/\tilde{\sigma}) - \tilde{\rho} - y^{-1} \|_1.
\]

The above inequality, the fact that \( g_1(\cdot|\tilde{H}) \) has unit Fréchet margins and a few manipulations lead to conclude that

\[
\left( \int_{(0, \infty)} \left| V(y^{\tilde{\rho}/\rho}(\sigma/\tilde{\sigma}) \tilde{\rho}|H) - V(y|H) \right| g_1(y|\tilde{H}) \, dy \right)^{1/l}
\]

\[
\leq \sum_{j=1}^{d} \left( \int_{0}^{\infty} \left| y_j^{\tilde{\rho}_j/\rho_j} - y_j^{-1} \right| y_j^{2-\varepsilon} e^{-y_j^{-1}} \, dy_j \right)^{1/l} + \sum_{j=1}^{d} \left( \int_{0}^{\infty} \left| \frac{\tilde{\sigma}_j}{\sigma_j} - 1 \right| y_j^{-1} e^{-y_j^{-1}} \, dy_j \right)^{1/l}
\]

\[
=: T_1 + T_2.
\]

On one hand, an application of the mean-value theorem and the inequality \( \log(x) \leq x, x \geq 1 \), yield

\[
T_1 \leq \sum_{j=1}^{d} \left| 1 - \tilde{\rho}_j/\rho_j \right| \left( \int_{0}^{1} \left| -\log \frac{y_j^{1-\varepsilon}}{y_j^{1-\varepsilon}} \right| y_j^{2-\varepsilon} e^{-y_j^{-1}} \, dy_j + \int_{1}^{\infty} \left| \log \frac{y_j}{y_j^{1-\varepsilon}} \right| y_j^{2-\varepsilon} e^{-y_j^{-1}} \, dy_j \right)^{1/l}
\]

\[
\leq \{v_l + \Gamma(1 + l)\}^{1/l} \sum_{j=1}^{d} \left| 1 - \tilde{\rho}_j/\rho_j \right| \left| 1 - \tilde{\rho}_j/\rho_j \right|^{1/l}.
\]
On the other hand, using sequentially the integral representation of the Gamma function, the mean-value theorem, the fact that \(|\log(x)| \leq \max(|x-1|, |x-1|/x), x > 0\), together with the bounds in (87), we obtain

\[
T_2 \leq \left\{ \Gamma(1 + l(1 + \varepsilon)) \right\}^{1/l} \sum_{j=1}^{d} \left| -\frac{\bar{\rho}_j}{\bar{\sigma}_j} \right| \\
\leq \left\{ \Gamma(1 + 2l) \right\}^{1/l} \sum_{j=1}^{d} \rho_j \left| \log(\bar{\sigma}_j/\bar{\sigma}_j) \right| \left\{ \left(\bar{\sigma}_j/\bar{\sigma}_j \right) \right\} \left\{ \left| \bar{\sigma}_j - \bar{\sigma}_j \right| \right\}.
\]

(88)

The result now follows. \(\square\)

**Lemma D.11.** For all \(k \geq d + 1, \ l \in \mathbb{N}_+, \ H \in \mathcal{H}, \ H_k \in \mathcal{H}_k, \ \rho, \tilde{\rho} \in (0, \infty)\) and \(\sigma, \tilde{\sigma} \in (0, \infty)\) it holds that

\[
\left( \int_{(0, \infty)} \left[ \max_{P \in \mathcal{P}_d, I \in \mathcal{P}} \log^+ \left\{ \frac{-V_I(y|H_k)}{-V_I(\tilde{y}^{\rho}/\rho|H_k)} \right\} \right]^l g_1(y|H)dy \right)^{1/l}
\]

is bounded from above by

\[
3k \left\{ \left( \gamma(+, l) + \gamma(l, -) \right) \sum_{j=1}^{d} \frac{|\rho_j - \tilde{\rho}_j|}{\rho_j} + \sum_{s=1}^{d} \frac{1 - \{ s/s \}}{\min\{ 1, (s/s) \}^{\rho_s}} \right\},
\]

where \(\gamma(+, l), \gamma(l, -)\) are given in (83).

**Proof.** By Minkowski’s inequality, the term in (89) is bounded from above by \(T_1 + T_2\), where

\[
T_1 := \left( \int_{(0, \infty)} \left[ \max_{P \in \mathcal{P}_d, I \in \mathcal{P}} \log^+ \left\{ \frac{-V_I(y|H_k)}{-V_I(\tilde{y}^{\rho}/\rho|H_k)} \right\} \right]^l g_1(y|H)dy \right)^{1/l},
\]

\[
T_2 := \left( \int_{(0, \infty)} \left[ \max_{P \in \mathcal{P}_d, I \in \mathcal{P}} \log^+ \left\{ \frac{-V_I(y|H_k)}{-V_I(\tilde{y}^{\rho}/\rho|H_k)} \right\} \right]^l g_1(y|H)dy \right)^{1/l}.
\]

The remainder of the proof consists of two parts: the derivation of upper bounds for the terms in curly brackets in the expression of \(T_1, T_2\) and the conclusion.

**Derivation of upper bounds.** To derive upper bounds for the terms of the form

\[
\log^+ \left\{ \frac{-V_I(y|H_k)}{-V_I(\tilde{y}^{\rho}/\rho|H_k)} \right\}, \log^+ \left\{ \frac{-V_I(y|H_k)}{-V_I(y^{\rho}/\rho|H_k)} \right\},
\]

we treat different types of \(I \subset \{1, \ldots, d\}\) separately. To improve the readability of some formulas, in the sequel we use the symbols \(\tilde{y} = \tilde{y}^{\rho}/\rho\) and \(\tilde{y} = \tilde{y}^{\sigma}/\sigma\). As usual, when \(x \in \mathbb{R}^d, x_I = (x_i, i \in I)\); in particular, if \(I = \{j\}\) for some \(j \in \{1, \ldots, d\}\), then \(x_I = x_j\).

**Case 1:** \(I = \{j\}, \) **for some** \(j \in \{1, \ldots, d\}. \) By Lemma D.2, we have that

\[
\log^+ \left\{ \frac{-V_I(y|H_k)}{-V_I(y^{\rho}/\rho|H_k)} \right\} \leq 2 \left| \tilde{\rho}_I - 1 \right| \left| \log y_I \right| + \log^+ \left\{ \frac{\varphi \kappa_j + \sum_{\alpha \in \Gamma_k} \varphi \alpha \frac{\alpha_I}{k} \left( 1 - \frac{I y_I/\|y_I\| \alpha_I}{\|y_I\| (\alpha_I + 1, \|y_I\|)} \right)}{\frac{\varphi \kappa_j + \sum_{\alpha \in \Gamma_k} \varphi \alpha \frac{\alpha_I}{k} \left( 1 - \frac{I \tilde{y}_I/\|	ilde{y}_I\| \alpha_I}{\|	ilde{y}_I\| (\alpha_I + 1, \|	ilde{y}_I\|)} \right)}{}}.\right.
\]
where, for all \( \alpha \in \Gamma_k \) and \( \beta \in (0, \infty) \),
\[
\ell_{\alpha, y}(\beta) := \log \left( \mathcal{I}_{1-x}(\|\alpha Fc\|_1, \alpha I + 1) \right)_{x=y/\rho}.
\]

By the multivariate mean-value theorem there exists \( c \in (0, 1) \) such that, setting \( \beta = c1 + (1-c)\hat{\rho}/\rho \),
\[
\ell_{\alpha, y}(1) - \ell_{\alpha, y}(\hat{\rho}/\rho) = \left( 1 - \frac{\hat{\beta}^I_{F/|y^\beta|}}{\|y^\beta\|_1} \right)^{\alpha I + 1} \frac{B}{B^I(1 - \frac{\hat{\beta}^I_{F/|y^\beta|}}{\|y^\beta\|_1}; \alpha Fc, 1, \alpha I + 1)} \left[ \log y_{\hat{s}} \right] \left( 1 - \frac{\hat{\rho}_s}{\rho_s} \right) 
\]
\[
\leq k \sum_{s=1}^d \frac{\|\rho_s - \hat{\rho}_s\|}{\rho_s} |\log y_{\hat{s}}|.
\]

where \( B(x; q_1, q_2) = B(q_1, q_2) I(x; q_1, q_2) \), \( x \in (0, 1) \), is the incomplete Beta function. Therefore, we conclude that
\[
\log^+ \left\{ \frac{-V_I(y|H_k)}{-V_I(y^\rho/p|H_k)} \right\} \leq (2 + k) \sum_{s=1}^d \frac{\|\rho_s - \hat{\rho}_s\|}{\rho_s} |\log y_{\hat{s}}|.
\]

By a similar reasoning and exploiting the inequality \( \log^+(x) \leq |x - 1| \), \( x > 0 \), we also obtain
\[
\log^+ \left\{ \frac{-V_I(y^\rho/p|H_k)}{-V_I(y^\rho/p|H_k)} \right\} \leq 2 \log^+ \left\{ (\sigma_I/\hat{\sigma}) \hat{\rho}_I \right\} + \log \left\{ \varphi_{\kappa_I} + \sum_{\alpha \in \Gamma_k} \varphi_{\alpha} \frac{\alpha}{\beta} \left( 1 - \frac{\hat{\beta}^I_{F/|y^\beta|}}{\|y^\beta\|_1} \right)^{\|\alpha Fc\|_1} \right\}
\]
\[
\leq 2 (2 + k) \sum_{s=1}^d \frac{|1 - \{\sigma_s/\hat{\sigma}_s\}\hat{\rho}_s|}{\min\{1, \{\sigma_s/\hat{\sigma}_s\}\hat{\rho}_s\}}.
\]

Case 2: \( 1 < |I| \leq d - 1 \). By arguments similar to those used in the previous case, we have that
\[
\log^+ \left\{ \frac{-V_I(y|H_k)}{-V_I(y^\rho/p|H_k)} \right\} \leq S_1 + S_2,
\]
where
\[
S_1 := (|I| + 1) \log^+ \frac{\|\hat{y}_I\|}{\|y_I\|_1} \leq d \sum_{s=1}^d |\log y_{\hat{s}}| \frac{|\hat{\rho}_s - \rho_s|}{\rho_s}
\]
and

\[ S_2 := \log^+ \left\{ \sum_{\alpha \in \Gamma_k} \varphi^{\alpha_k} \left\| \alpha_k \right\|_k \text{Dir} \left( \frac{y_I - \left\| y_I \right\|_1}{\left\| y_I \right\|_1} ; \alpha_I \right) \left( 1 - \frac{1}{\left\| y_I \right\|_1 / \left\| y_I \right\|_1} \left( \left\| \alpha_I \right\|_1 + 1, \left\| \alpha_I \right\|_1 \right) \right) \right\} \]

\[ \leq \max_{\alpha \in \Gamma_k} \left\{ \log^+ \left( \frac{\text{Dir} \left( \frac{y_I - \left\| y_I \right\|_1}{\left\| y_I \right\|_1} ; \alpha_I \right)}{\text{Dir} \left( \frac{\tilde{y}_I - \left\| \tilde{y}_I \right\|_1}{\left\| \tilde{y}_I \right\|_1} ; \alpha_I \right)} \right) \right\} + \max_{\alpha \in \Gamma_k} \left\{ \log^+ \left( \frac{1 - \frac{1}{\left\| y_I \right\|_1 / \left\| y_I \right\|_1} \left( \left\| \alpha_I \right\|_1 + 1, \left\| \alpha_I \right\|_1 \right)}{1 - \frac{1}{\left\| \tilde{y}_I \right\|_1 / \left\| \tilde{y}_I \right\|_1} \left( \left\| \alpha_I \right\|_1 + 1, \left\| \alpha_I \right\|_1 \right)} \right) \right\} \]

\[ \leq 2k \sum_{s=1}^d \left| \frac{\rho_s - \tilde{\rho}_s}{\rho_s} \right| \log |y_s|, \]

with \( I^- \) as in (50). Thus, we conclude that

\[ \log^+ \left\{ \frac{-V_I(y_I|H_K)}{-V_I(y^\rho/P|H_K)} \right\} \leq 3k \sum_{s=1}^d \frac{\left| \rho_s - \tilde{\rho}_s \right|}{\rho_s} \log |y_s|. \]  

Moreover, we have that

\[ \log^+ \left\{ \frac{-V_I(y^\rho/P|H_K)}{-V_I(y^\rho/P(\sigma/\tilde{\sigma})\tilde{P}|H_K)} \right\} \leq S_3 + S_4, \]

where

\[ S_3 := (|I| + 1) \log^+ \left( \frac{\tilde{y}_I}{\tilde{y}_I} \right) \leq d \log^+ \left\{ \max_{1 \leq s \leq d} (\sigma_s/\tilde{\sigma})\tilde{\rho}_s \right\} \leq d \sum_{s=1}^d \left| 1 - (\sigma_s/\tilde{\sigma})\tilde{\rho}_s \right| \]

and, using the inequality \( \log^+ (1/x) \leq |1 - x|/x, x > 0, \)

\[ S_4 := \log^+ \left\{ \sum_{\alpha \in \Gamma_k} \varphi^{\alpha_k} \left\| \alpha_k \right\|_k \text{Dir} \left( \frac{y_I - \left\| y_I \right\|_1}{\left\| y_I \right\|_1} ; \alpha_I \right) \left( 1 - \frac{1}{\left\| y_I \right\|_1 / \left\| y_I \right\|_1} \left( \left\| \alpha_I \right\|_1 + 1, \left\| \alpha_I \right\|_1 \right) \right) \right\} \]

\[ \leq \max_{\alpha \in \Gamma_k} \left\{ \log^+ \left( \frac{\text{Dir} \left( \frac{y_I - \left\| y_I \right\|_1}{\left\| y_I \right\|_1} ; \alpha_I \right)}{\text{Dir} \left( \frac{\tilde{y}_I - \left\| \tilde{y}_I \right\|_1}{\left\| \tilde{y}_I \right\|_1} ; \alpha_I \right)} \right) \right\} + \max_{\alpha \in \Gamma_k} \left\{ \log^+ \left( \frac{1 - \frac{1}{\left\| y_I \right\|_1 / \left\| y_I \right\|_1} \left( \left\| \alpha_I \right\|_1 + 1, \left\| \alpha_I \right\|_1 \right)}{1 - \frac{1}{\left\| \tilde{y}_I \right\|_1 / \left\| \tilde{y}_I \right\|_1} \left( \left\| \alpha_I \right\|_1 + 1, \left\| \alpha_I \right\|_1 \right)} \right) \right\} \]

\[ \leq \max_{\alpha \in \Gamma_k} \sum_{s=1}^d (\sigma_s - 1) \log^+ \left( \left( \sigma_s/\tilde{\sigma} \right)^{-\tilde{\rho}_s} \right) + k \sum_{s=1}^d \frac{ \left| 1 - (\sigma_s/\tilde{\sigma})\tilde{\rho}_s \right| }{ \min \left( 1, (\sigma_s/\tilde{\sigma})\tilde{\rho}_s \right) } \]

\[ \leq 2k \sum_{s=1}^d \frac{ \left| 1 - (\sigma_s/\tilde{\sigma})\tilde{\rho}_s \right| }{ \min \left( 1, (\sigma_s/\tilde{\sigma})\tilde{\rho}_s \right) }, \]

consequently

\[ \log^+ \left\{ \frac{-V_I(y^\rho/P|H_K)}{-V_I(y^\rho/P(\sigma/\tilde{\sigma})\tilde{P}|H_K)} \right\} \leq 3k \sum_{s=1}^d \frac{ \left| 1 - (\sigma_s/\tilde{\sigma})\tilde{\rho}_s \right| }{ \min \left( 1, (\sigma_s/\tilde{\sigma})\tilde{\rho}_s \right) }. \]
Case 3: \( \{1, \ldots, d\} \). Once again, Lemma D.2 and a few algebraic manipulations similar to those used above yield

\[
\begin{align*}
\log^+ & \left\{ \frac{-V_f(y|H_k)}{-V_f(y^{\rho/\rho}|H_k)} \right\} \\
= & \log^+ \left\{ \frac{\|y\|_{1,d-1}^{-1} \sum_{\alpha \in \Gamma_k} \varphi_\alpha \text{Dir}(y_{1:d-1}/\|y\|_{1}; \alpha)}{\|\tilde{y}\|_{1,d-1}^{-1} \sum_{\alpha \in \Gamma_k} \varphi_\alpha \text{Dir}(\tilde{y}_{1:d-1}/\|\tilde{y}\|_{1}; \alpha)} \right\} \\
\leq & (d + 1) \log^+ \frac{\|\tilde{y}\|_{1}}{\|y\|_{1}} + \max_{\alpha \in \Gamma_k} \log \text{Dir}(y_{1:d-1}/\|y\|_{1}; \alpha)) \\
\leq & (d + 1) \sum_{s=1}^{d} \frac{|\rho_s - \tilde{\rho}_s|}{\rho_s} |\log y_s| + k \sum_{s=1}^{d} \frac{|\rho_s - \tilde{\rho}_s|}{\rho_s} |\log y_s| \\
\leq & 2k \sum_{s=1}^{d} \frac{|\rho_s - \tilde{\rho}_s|}{\rho_s} |\log y_s|
\end{align*}
\]

(94)

and

\[
\begin{align*}
\log^+ & \left\{ \frac{-V_f(y^{\rho/\rho}|H_k)}{-V_f(y^{\rho/\rho}|\sigma|/\rho|H_k)} \right\} \\
= & \log^+ \left\{ \frac{\|\tilde{y}\|_{1,d-1}^{-1} \sum_{\alpha \in \Gamma_k} \varphi_\alpha \text{Dir}(\tilde{y}_{1,d-1}/\|\tilde{y}\|_{1}; \alpha)}{\|\tilde{y}\|_{1,d-1}^{-1} \sum_{\alpha \in \Gamma_k} \varphi_\alpha \text{Dir}(\tilde{y}_{1,d-1}/\|\tilde{y}\|_{1}; \alpha)} \right\} \\
\leq & (d + 1) \log^+ \frac{\|\tilde{y}\|_{1}}{\|y\|_{1}} + \max_{\alpha \in \Gamma_k} \log \text{Dir}(\tilde{y}_{1,d-1}/\|\tilde{y}\|_{1}; \alpha)) \\
\leq & (d + 1) \log^+ \left\{ \max_{1 \leq s \leq d} (\sigma_s/\tilde{\sigma}_s)^{\tilde{\rho}_s} \right\} + \max_{\alpha \in \Gamma_k} \sum_{s=1}^{d} (\alpha_s - 1) \log^+ \left\{ (\sigma_s/\tilde{\sigma}_s)^{\tilde{\rho}_s} \right\} \\
\leq & 2k \sum_{s=1}^{d} \frac{1 - (\sigma_s/\tilde{\sigma}_s)^{\tilde{\rho}_s}}{\min(1, (\sigma_s/\tilde{\sigma}_s)^{\tilde{\rho}_s})}.
\end{align*}
\]

(95)

Conclusion. Combining the bounds in (90), (92) and (94) and using Minkowski’s inequality together with the fact that \( g_1 (\cdot|H) \) has unit Fréchet margins, we obtain

\[
T_1 \leq \left( \int_{(0, \infty)} \left[ 3k \sum_{s=1}^{d} \frac{|\rho_s - \tilde{\rho}_s|}{\rho_s} |\log y_s| \right]^{l} g_1 (y|H) dy \right)^{1/l}
\]

(96)

\[
\leq 3k \sum_{s=1}^{d} \frac{|\rho_s - \tilde{\rho}_s|}{\rho_s} \left( \int_{(0, \infty)} |\log y_s|^{l} g_1 (y|H) dy \right)^{1/l}
\]

\[
= 3k \left( \gamma_{(+, l)} + \gamma_{(+, l)} \right) \sum_{s=1}^{d} \frac{|\rho_s - \tilde{\rho}_s|}{\rho_s}.
\]

Moreover, combining the bounds in (91), (93) and (95) and using Minkowski’s inequality together with the fact that \( g_1 (\cdot|H) \) has unit Fréchet margins, we conclude

\[
T_2 \leq \left( \int_{(0, \infty)} \left[ 3k \sum_{s=1}^{d} \frac{1 - (\sigma_s/\tilde{\sigma}_s)^{\tilde{\rho}_s}}{\min(1, (\sigma_s/\tilde{\sigma}_s)^{\tilde{\rho}_s})} \right]^{l} g_1 (y|H) dy \right)^{1/l}
\]

\[
\leq 3k \sum_{s=1}^{d} \frac{1 - (\sigma_s/\tilde{\sigma}_s)^{\tilde{\rho}_s}}{\min(1, (\sigma_s/\tilde{\sigma}_s)^{\tilde{\rho}_s})}.
\]
The result now follows.

**Lemma D.12.** Let $H_0 \in \mathcal{H}_0$ and $\rho_0 \in (0, \infty)$, $\sigma_0 \in (0, \infty)$. Then, for all $l \in \mathbb{N}_+$ and $\epsilon > 0$ there exist $H^* \in \mathcal{H}'$, $\delta_1, \delta_2, \delta_3 > 0$ such that

\[ \mathcal{X}_+^{(l)}(g_{\rho_0, \sigma_0}(|H_0|), g_{\rho, \sigma}(|H|)) < \epsilon, \]

for all $\rho \in B_{\delta_1,1}(\rho_0)$, $\sigma \in B_{\delta_2,1}(\sigma_0)$ and $H \in B_{\delta_3,\infty}(H^*)$. In the particular case of $l = 1$, it also holds that $\mathcal{X}(g_{\rho_0, \sigma_0}(|H_0|), g_{\rho, \sigma}(|H|)) < \epsilon$.

**Proof.** Observe that, for any $H \in \mathcal{H}$, $\rho \in (0, \infty)$, $\sigma \in (0, \infty)$, Minkowski’s inequality yields

\[ \left\{ \mathcal{X}_+^{(l)}(g_{\rho_0, \sigma_0}(|H_0|), g_{\rho, \sigma}(|H|)) \right\}^{1/l} = \mathcal{X}_+^{(l)}(g_{H_0, \rho_0, \sigma_0}(H_0, \rho_0, \sigma_0; H, \rho, \sigma)
\]

\[ \leq \mathcal{X}_+^{(l)}(g_{H_0, \rho_0, \sigma_0}(H_0, \rho_0, \sigma_0; H, \rho_0, \sigma_0))
\]

\[ + \mathcal{X}_+^{(l)}(g_{H_0, \rho_0, \sigma_0}(H, \rho_0, \sigma_0; H, \rho, \sigma)). \]

A change of variables allows to re-express the first term on the right-hand side of (98) as follows

\[ \left\{ \mathcal{X}_+^{(l)}(g_{\rho_0, \sigma_0}(|H_0|), g_{\rho, \sigma}(|H|)) \right\}^{1/l} = \left\{ \mathcal{X}_+^{(l)}(g_1(|H_0|), g_1(|H|)) \right\}^{1/l}. \]

Thus, by Propositions C.7 and C.8, such a term can be made arbitrarily small by choosing $H \in B_{\delta_3,\infty}(H^*)$, for suitable $H^* \in \mathcal{H}'$ and $\delta_3 > 0$. As for the second term on the right-hand side of (98), a change of variables and Minkowski’s inequality yield

\[ \mathcal{X}_+^{(l)}(H_0, \rho_0, \sigma_0; H, \rho, \sigma)
\]

\[ \leq \sum_{j=1}^{d} \left[ \int_{(0, \infty)} \left[ \log + \left\{ \frac{\rho_0, j}{\rho_j} \frac{\sigma^0_j}{\sigma_j} \right\} \frac{1 - \rho_j/\rho_0, j}{\rho_j} \right] \right] g_1(y|H_0) dy
\]

\[ + \left( \int_{(0, \infty)} \left[ \log + \left\{ \frac{\sum_{P \in \mathcal{P}_d} \prod_{j=1}^{n} (-V_i(y|H))}{\sum_{P \in \mathcal{P}_d} \prod_{j=1}^{n} (-V_i(y|H))} \right\} \right] g_1(y|H_0) dy \right)^{1/l}
\]

\[ =: T_1 + T_2 + T_3. \]

Applying once more Minkowski’s inequality and exploiting the fact that $g_1(|H_0|)$ has unit-Fréchet margins we obtain

\[ T_1 \leq \sum_{j=1}^{d} \left[ \log + \left\{ \frac{\rho_0, j}{\rho_j} \frac{\sigma^0_j}{\sigma_j} \right\} \right] + \left( 1 - \frac{\rho_j}{\rho_0, j} \right) \mathbb{I}(\rho_0, j > \rho_j) \gamma_{(l,+)} \]

\[ + \left( \frac{\rho_j}{\rho_0, j} - 1 \right) \mathbb{I}(\rho_0, j > \rho_j) \gamma_{(l,-)} \]

with $\gamma_{(l,+)}$, $\gamma_{(l,-)}$ as in (83). Consequently, $T_1$ can be made arbitrarily small by choosing $\delta_1, \delta_2$ small enough.

If $\delta_1, \delta_2$ also comply with (87) for some $\epsilon \in (0, 1/2)$, with $(\rho_0, \sigma_0)$ and $(\rho, \sigma)$ in place of $(\rho, \sigma)$ and $(\rho, \tilde{\sigma})$, respectively, the facts that $B_{\delta_1,1}(\rho_0) \subset B_{\delta_2,\infty}(\rho_0)$ and $B_{\delta_2,1}(\sigma_0) \subset B_{\delta_2,\infty}(\sigma_0)$ together with Lemma D.10 entail that

\[ T_2 \leq \sum_{j=1}^{d} \leq c_1 \delta_1 + c_2 \delta_2, \]
where $c_1, c_2$ are positive constant depending on $l, \rho_0, \sigma_0$. Furthermore, selecting $\delta_3$ such that
\[
0 < \delta_3 < \min \left\{ 1, \frac{\inf_{v \in \mathcal{R}} h^* (v)}{2}, \frac{\min_{1 \leq j \leq d} p_j^*} {\delta_3 cd - 1} \right\}^2
\]
where $c = 1/\Gamma (d)$ and $p_j^* = H^* (\{ e_j \})$, $j = 1, \ldots, d$, by Lemma D.1 there exists $k > d$ and $H_k \in \mathcal{H}_k \cap B_{\delta_3, \infty} (H^* \in \mathcal{H})$ with coefficients $\varphi_{\kappa_1}, \ldots, \varphi_{\kappa_d}, \varphi_{\alpha} \in \Gamma_k$, satisfying
\[
| h - h_{k-d} | \leq \frac{2\delta_3}{\inf_{v \in \mathcal{R}} h^* (v) - \delta_3} \left( \inf_{v \in \mathcal{R}} h^* (v) - \delta_3 \right) \leq \sqrt{\delta_3} \min (h, h_{k-d}),
\]
\[
\max \left( \frac{p_j - \varphi_{\kappa_j}}{\varphi_{\kappa_j}}, \frac{\varphi_{\kappa_j} - p_j}{\varphi_{\kappa_j}} \right) \leq \frac{2\delta_3}{\delta_3 cd - 1} \leq \sqrt{\delta_3}, \quad j = 1, \ldots, d,
\]
wherefrom we deduce that facts analogous to (i)-(ii) in the proof of Lemma C.4 hold true and we conclude that for all $z \in (0, \infty)$
\[
\max_{P \in \mathcal{P}_d} \max_{I_i \in \mathcal{P}} \left( \log^+ \left\{ \frac{-V_{I_1} (z | H_k)}{-V_{I_1} (z | H)} \right\}, \log^+ \left\{ \frac{-V_{I_1} (z | H_k)}{-V_{I_1} (z | H)} \right\} \right) \leq \log (1 + \sqrt{\delta_3}).
\]
Now, following steps like those in (40) and applying Minkowski’s inequality, we obtain
\[
T_3 \leq \left( \int_{(0, \infty)} \left[ \max_{P \in \mathcal{P}_d} \max_{I_i \in \mathcal{P}} \log^+ \left\{ \prod_{i=1}^{m} \frac{-V_{I_1} (y | H)}{-V_{I_1} (y^{\rho} / \rho_0, \sigma_0 / | \sigma | H)} \right\} \right]^l g_1 (y | H_0) dy \right)^{1/l}.
\]
\[
\leq \left( \int_{(0, \infty)} \left[ \max_{P \in \mathcal{P}_d} \max_{I_i \in \mathcal{P}} \log^+ \left\{ \frac{\rho}{\rho_0} \rho_0 \left( \frac{\sigma_0}{| \sigma |} \rho(H_0) \right) \right\} \right]^l g_1 (y | H_0) dy \right)^{1/l}.
\]
\[
\leq \left( \int_{(0, \infty)} \left[ \max_{P \in \mathcal{P}_d} \max_{I_i \in \mathcal{P}} \log^+ \left\{ \frac{\sigma_0 / | \sigma | H_0}{\sigma_0 / | \sigma | H} \right\} \right]^l g_1 (y | H_0) dy \right)^{1/l}.
\]
\[
+ \left( \int_{(0, \infty)} \left[ \max_{P \in \mathcal{P}_d} \max_{I_i \in \mathcal{P}} \log^+ \left\{ \frac{\rho}{\rho_0} \rho_0 \left( \frac{\sigma_0}{| \sigma |} \rho(H_0) \right) \right\} \right]^l g_1 (y | H_0) dy \right)^{1/l}.
\]
\[
= T_3^{(1)} + T_3^{(2)} + T_3^{(3)}.
\]
The term $T_3^{(1)}$ can be made arbitrarily small by resorting to Lemma D.11 and choosing $\delta_1, \delta_2$ sufficiently small. By the bound in (99), the same is true for $T_3^{(2)}$ and $T_3^{(3)}$. The proof is now complete. \hfill \Box

For given $H_0 \in \mathcal{H}_0, \rho_0 \in (0, \infty), \sigma_0 \in (0, \infty)$ and for all $\epsilon > 0$ define
\[
\mathcal{V}_\epsilon := \{ (H, \rho, \sigma) \in \mathcal{H} \times (0, \infty) \times (0, \infty) \in \mathcal{X} (g \rho_0, \sigma_0 (| H_0 |), g \rho, \sigma (| H |)) < \epsilon \}.
\]
moreover, for all $l \in \mathbb{N}_+$, define
\[
\mathcal{V}_\epsilon^{(l)} := \{ (H, \rho, \sigma) \in \mathcal{H} \times (0, \infty) \times (0, \infty) \in \mathcal{X} (g \rho_0, \sigma_0 (| H_0 |), g \rho, \sigma (| H |)) < \epsilon \}.
\]
We are now in the position to establish the following result, which is used in the proof of both Theorem 3.14 and Lemma D.22 (auxiliary to Theorem 4.7).

**Corollary D.13.** Under Condition 3.13(iia) and Conditions 3.13(ii)-(iii), for all $\epsilon > 0$ we have that
\[
\Pi_{H \times \Theta} (\mathcal{V}_\epsilon) \geq \Pi_{H \times \Theta} \left( \bigcap_{l=1}^{4} \mathcal{V}_\epsilon^{(l)} \right) > 0.
\]
Thus, the induced prior $\nu_{\Theta}$ on the Borel sets of $\Theta = \{ g_{\rho, \sigma}(|H) : H \in \mathcal{H}, (\rho, \sigma) \in (0, \infty) \times (0, \infty) \}$ possesses the Kullback-Leibler property.

**Proof.** Preliminarily observe that, since $(x, H, \rho, \sigma) \mapsto g_{\rho, \sigma}(x|H)$ is tacitly chosen as a Borel measurable map between $(0, \infty) \times \{ H \times (0, \infty) \times (0, \infty) \}$ and $[0, \infty]$, arguments similar to those in Appendix A of [63] guarantee that

$$\nu_{\Theta} \in \mathcal{B}_{\mathcal{H}} \otimes \mathcal{B}_{\Theta}, \quad \nu_{\Theta}^{(l)} \in \mathcal{B}_{\mathcal{H}} \otimes \mathcal{B}_{\Theta}, \quad l = 1, \ldots, 4,$$

where $\mathcal{B}_{\mathcal{H}}$ is the Borel $\sigma$-algebra of $(\mathcal{H}, \mathcal{B}_{\mathcal{H}})$ and $\mathcal{B}_{\Theta}$ is the Borel $\sigma$-algebra induced by the $L_1$-topology on $(0, \infty) \times (0, \infty)$. By Lemma D.12, whenever $H_0 \in \mathcal{H}_{\Theta}$, the higher-order positive divergences

$$\mathcal{D}_{\Theta}^{(l)}(g_{\rho_0, \sigma_0}(|H_0), g_{\rho, \sigma}(|H)), \quad l = 1, \ldots, 4,$$

can be made arbitrarily small over

$$B_{\delta_3, \infty}(H^*) \times B_{\delta_1, 1}(\rho_0) \times B_{\delta_2, 1}(\sigma_0),$$

for suitable $H^* \in \mathcal{H}^d$ and suitably small radii $\delta_1 \in (0, \epsilon), \delta_2 \in (0, \epsilon), \delta_3 > 0$. On one hand, Condition 3.13(iii) guarantees the existence of a $\mathcal{B}_{\mathcal{H}}$-measurable subset of $B_{\delta_3, \infty}(H^*)$ with positive $\Pi_{\mathcal{H}}$-probability, see the Kullback-Leibler property section in the proof of Theorem 3.9 for details (Section D.3.5 of this manuscript). On the other hand, $B_{\delta_1, 1}(\rho_0) \times B_{\delta_2, 1}(\sigma_0)$ is an open set with respect to the $L_1$-topology on

$$(0, \infty) \times (0, \infty) \equiv (0, \infty)^{2d},$$

hence it is $\mathcal{B}_{\Theta}$-measurable and, by Condition 3.13(iii), $\Pi_{\Theta}(B_{\delta_1, 1}(\rho_0) \times B_{\delta_2, 1}(\sigma_0)) > 0$. The result now follows.

The last two lemmas are exploited for proving the existence of test sequences that comply with Condition 3.11(ii) in the proof of Theorem 3.14. These are also useful for ensuring an exponential decay to zero for a stochastic upper-bound of the quasi posterior probabilities in Theorem 4.7(a')–(b'), as elucidated in Section D.6.2.

**Lemma D.14.** For all $k > d, H_k, \tilde{H}_k \in H_k$ with mixture weights $\varphi, \tilde{\varphi} \in \Phi_k, \rho_0 \in (0, \infty), \rho, \tilde{\rho} \in B_{\delta_0, 1}(\rho_0), \sigma_0 \in (0, \infty), \sigma, \tilde{\sigma} \in B_{\delta_0, 1}(\sigma_0)$, with

$$0 < \delta_0 < \frac{\min}{1 \leq j \leq d} \rho_0, \frac{\min}{1 \leq j \leq d} \sigma_0 =: \delta_*,$$

it holds that

$$\mathcal{D}_{\mathcal{H}}(g_{\rho, \sigma}(|H_k), g_{\tilde{\rho}, \tilde{\sigma}}(|\tilde{H}_k)) \leq \sqrt{c_0} ||\varphi_0 - \tilde{\varphi}_0||_1 + \sqrt{c_0} ||\rho - \tilde{\rho}||_1 + c_0 ||\sigma - \tilde{\sigma}||_1,$$

where $c_0$ is a positive constant depending on $d, \rho_0$ and $\sigma_0$, while $\varphi_0$ and $\tilde{\varphi}_0$ are defined as in (52).

**Proof.** Preliminary observe that $B_{\delta_0, 1}(\rho_0) \subset B_{\delta_0, \infty}(\rho_0)$ and $B_{\delta_0, 1}(\sigma_0) \subset B_{\delta_0, \infty}(\sigma_0)$. Therefore, $\tilde{\rho} \in B_{2\delta_0, \infty}(\rho)$, as well as $\tilde{\sigma} \in B_{2\delta_0, \infty}(\sigma)$. Moreover,

$$2\delta_0 < \min_{1 \leq j \leq d} \rho_0, \frac{\min}{1 \leq j \leq d} \rho_0 \cdot \frac{\delta_0}{2} < \frac{1}{2} \min_{1 \leq j \leq d} \rho_0,$$

and

$$2\delta_0 < \min_{1 \leq j \leq d} \sigma_0, \frac{\min}{1 \leq j \leq d} \rho_0 \cdot \frac{\delta_0}{2} < \frac{1}{2} \min_{1 \leq j \leq d} \sigma_0.$$

Therefore, $2\delta_0$ satisfies the condition in (87), with $\delta_1$ and $\delta_2$ replaced by $2\delta_0$ and $\varepsilon = 1/2$.

Next, observe that,

$$\mathcal{D}_{\mathcal{H}}(g_{\rho, \sigma}(|H_k), g_{\tilde{\rho}, \tilde{\sigma}}(|\tilde{H}_k)) = \mathcal{D}_{\mathcal{H}}(g_1(|H_k), g_{\tilde{\rho}, \tilde{\sigma}}(|\tilde{H}_k))$$

$$\leq \mathcal{D}_{\mathcal{H}}(g_1(|H_k), g_1(|\tilde{H}_k)) + \mathcal{D}_{\mathcal{H}}(g_1(|\tilde{H}_k), g_{\tilde{\rho}, \tilde{\sigma}}(|\tilde{H}_k)),$$
where $\hat{\rho} = \hat{\rho}/\rho$ and $\hat{\sigma} = \{\hat{\sigma}/\sigma\}^{1/\rho}$. By Lemma D.5, we have that the first term on the right-hand side is bounded from above by the square root of $\|\hat{\sigma}_0 - \hat{\sigma}_0\|_1$, for some $c > 0$ depending only on $d$. Moreover, by Lemma B.1(iv) in [39], the second term on the right hand side is bounded from above by the square root of 

$$X(g_1(\cdot|\tilde{H}_k), g_{\hat{\rho}, \hat{\sigma}}(\cdot|\tilde{H}_k)) = \sum_{j=1}^d \frac{\hat{\rho}}{\rho_j} \log_j \frac{1}{\hat{\rho}_j} + \sum_{j=1}^d (1 - \hat{\rho}_j) \int_{(0,\infty)} \log_j g_1(y|\tilde{H}_k) dy$$

$$+ \int_{(0,\infty)} \left[ V(\{y/\hat{\sigma}\}^\hat{\rho}|\tilde{H}_k) - \tilde{V}_k(y)g_1(y|\tilde{H}_k) \right] dy$$

$$+ \int_{(0,\infty)} \log \frac{\sum_{j=1}^d \left\{ -V_{\hat{\rho}}(\{y/\hat{\sigma}\}^\hat{\rho}|\tilde{H}_k) \right\} \#_{j=1}^d \left\{ -V_{\hat{\rho}}(\{y/\hat{\sigma}\}^\hat{\rho}|\tilde{H}_k) \right\} }{g_1(y|\tilde{H}_k) dy}$$

$$=: T_1 + T_2 + T_3 + T_4$$

A few algebraic manipulations lead to show that the first two terms satisfy

$$T_1 \leq \frac{5}{4 \min_{1 \leq j \leq d} \rho_{0,j}} \|\rho - \hat{\rho}\|_1 + \frac{75}{8 \min_{1 \leq j \leq d} \rho_{0,j}} \|\sigma - \hat{\sigma}\|_1$$

and

$$T_2 \leq \frac{5\gamma}{4 \min_{1 \leq j \leq d} \rho_{0,j}} \|\rho - \hat{\rho}\|_1.$$ 

By Lemma D.10 and the bounds in (102)-(103), we also have that

$$T_3 \leq \frac{v_1 + \Gamma(2)}{\min_{1 \leq j \leq d} \rho_j} \|\rho - \hat{\rho}\|_1 + \Gamma(3) \max_{1 \leq j \leq d} \left\{ \frac{3\rho_j}{\sigma_j} \left( \frac{3}{2} \right) \right\} \|\sigma - \hat{\sigma}\|_1$$

$$\leq \frac{5(v_1 + 1)}{4 \min_{1 \leq j \leq d} \rho_{0,j}} \|\rho - \hat{\rho}\|_1 + \Gamma(3) \max_{1 \leq j \leq d} \left\{ \frac{9\rho_{0,j}}{2\sigma_{0,j}} \left( \frac{3}{2} \right) \right\} \|\sigma - \hat{\sigma}\|_1,$$

with $v_1$ as in (86). Finally, a Jensen's inequality argument like that in (40), Lemma D.11, the bounds in (102)-(103) and a few algebraic steps analogous to those in (88) yield

$$T_4 \leq 3k \left\{ \left( \gamma(1,+) + \gamma(1,-) \right) \sum_{j=1}^d \frac{|\rho_j - \hat{\rho}_j|}{\rho_j} + d \frac{1 - (\sigma_s/\hat{\sigma})^{\hat{\rho}_s}}{\min \{1, (\sigma_s/\hat{\sigma})^{\hat{\rho}_s} \}} \right\}$$

$$\leq 15k \left[ \gamma(1,+,1) + \gamma(1, +) \right] \frac{\min_{1 \leq j \leq d} \rho_{0,j}}{\max_{1 \leq j \leq d} \rho_{0,j}} \max_{1 \leq j \leq d} (2/3)^{3\rho_{0,j}} \|\sigma - \hat{\sigma}\|_1,$$

where $\gamma(1,\bullet)$ is as in (85). The result now follows. $\square$

Before stating the last lemma, we recall that for $\rho, \sigma > 0, x > 0, G_{\rho,\sigma}(x) := e^{-(x/\sigma)^{-\rho}}$ denotes the (uni-variate) two-parameter Fréchet distribution function. We denote by $G_{\rho,\sigma}^{(n)}$ the pertaining $n$-fold pmf and, for all measurable functions $f : (0,\infty)^n \rightarrow \mathbb{R}$, we use the following short notation for expected values

$$G_{\rho,\sigma}^{(n)} f := \int_{(0,\infty)^n} f(x_1, \ldots, x_n) dG_{\rho,\sigma}^{(n)}(x_1, \ldots, x_n)$$

**Lemma D.15.** For each $\epsilon > 0$, there exist Borel-measurable functions $t_{n,j} : (0,\infty)^n \rightarrow \{0,1\}, j = 1, \ldots, d, n \in \mathbb{N}_+$, and positive constants $c_j(\epsilon), \epsilon \in (0,\infty)$, such that

$$G_{\rho_{0,j},\sigma_{0,j}}^{(n)} t_{n,j} \leq 2e^{-nc_j(\epsilon)}, \quad \sup_{(\rho_j,\sigma_j) \in B_{\rho,(0,\infty)}^{(n)}((\rho_{0,j},\sigma_{0,j}))} G_{\rho_j,\sigma_j}^{(n)}(1-t_{n,j}) \leq 2e^{-nc_j(\epsilon)}.$$ 

**Proof.** First, assume $(\rho, \sigma) \in B_{\rho,(0,\infty)}^{(n)}((\rho_{0,j},\sigma_{0,j}))$ and, in particular, $\sigma \in (\sigma_{0,j} - \epsilon, \sigma_{0,j} + \epsilon)$, then, it is not difficult to check that

$$\mathcal{D}_{KS}(G_{\rho,\sigma}, G_{\rho_{0,j},\sigma_{0,j}}) \geq \min \left\{ e^{-\left( \frac{\sigma_{0,j}}{\sigma_{0,j} + \epsilon} \right)^{\rho_0}} - e^{-1}, e^{-1} - e^{-\left( \frac{\sigma_{0,j}}{\sigma_{0,j} - \epsilon} \right)^{\rho_0}} \right\}.$$
Next, assume $(\rho, \sigma) \in B^{\epsilon}_{e, \infty}((\rho_{0,j}, \sigma_{0,j}))$ and $\sigma \in (\sigma_{0,j} - \epsilon, \sigma_{0,j} + \epsilon)$, thus it must be that $\rho \in (\rho_{0,j} - \epsilon, \rho_{0,j} + \epsilon)$. In this case, it holds that
\[
D_{KS}(G_{\rho, \sigma}, G_{\rho_{0,j}, \sigma_{0,j}}) \\
\geq \min \left\{ e^{-1/2} - e^{-2^{-(\rho_{0,j} - \epsilon)/\rho_{0,j}}}, e^{-1/2} - e^{-2^{-(\rho_{0,j} + \epsilon)/\rho_{0,j}}}, e^{-2^{-(\rho_{0,j} + \epsilon)/\rho_{0,j}} - 2}, e^{-2^{-(\rho_{0,j} - \epsilon)/\rho_{0,j}} - 2} \right\}.
\]
Therefore, we deduce
\[
\epsilon_j := \inf_{(\rho, \sigma) \in B^{\epsilon}_{e, \infty}((\rho_{0,j}, \sigma_{0,j}))} D_{KS}(G_{\rho, \sigma}, G_{\rho_{0,j}, \sigma_{0,j}}) > 0.
\]
Define the empirical distribution function map
\[
\hat{G}_{n,j} : (x_1, \ldots, x_n) \mapsto \left\{ \frac{1}{n} \sum_{i=1}^{n} I(x_i \leq x) \right\}, x \in (0, \infty)
\]
and the associated test functions
\[
t_{n,j} : (x_1, \ldots, x_n) \mapsto I \left( D_{KS}(\hat{G}_{n,j}, G_{\rho_{0,j}, \sigma_{0,j}}) > \epsilon_j/2 \right), \quad j = 1, \ldots, d.
\]
For each $j = 1, \ldots, d$, $(\rho_j, \sigma_j) \in B^{\epsilon}_{e, \infty}((\rho_{0,j}, \sigma_{0,j}))$, reverse triangular inequality and the first formula on page 253 of [79] yield
\[
G^{(n)}_{\rho_j, \sigma_j}(1 - t_{n,j}) = G^{(n)}_{\rho_j, \sigma_j} \{ D_{KS}(\hat{G}_{n,j}, G_{\rho_{0,j}, \sigma_{0,j}}) \leq \epsilon_j/2 \} \\
\leq G^{(n)}_{\rho_j, \sigma_j} \{ D_{KS}(G_{\rho, \sigma}, G_{\rho_{0,j}, \sigma_{0,j}}) \leq \epsilon_j/2 + D_{KS}(\hat{G}_{n,j}, G_{\rho, \sigma}) \} \\
\leq G^{(n)}_{\rho_j, \sigma_j} \{ \epsilon_j/2 \leq D_{KS}(\hat{G}_{n,j}, G_{\rho, \sigma}) \} \\
= G^{(n)}_{\rho_j, \sigma_j} \{ \sqrt{n} \epsilon_j/2 \leq \sqrt{n} D_{KS}(\hat{G}_{n,j}, G_{\rho, \sigma}) \} \\
\leq 2e^{-n\epsilon^2_j/2}.
\]
Similarly, we also obtain $G^{(n)}_{\rho_{0,j}, \sigma_{0,j}} t_{n,j} \leq 2e^{-n\epsilon^2_j/2}$. The result now follows.

**D.4.3. Proof of Theorem 3.14.** The proof is developed by verifying Condition 3.11 for the semiparametric $\rho$-Fréchet model under study and by applying Proposition 3.12. In passing, note that such derivations also play a crucial role in establishing the results of Theorem 4.7.

Preliminary observe that, by Corollary D.13, the Kullback-Leibler support of the prior $\Pi_{\Theta^{(0)}}$ given in Theorem 3.14 contains $\Pi_{\rho_0, \sigma_0}(|H_0)$. Hence, the first requirement of Condition 3.11 is met. Next, fixing $\epsilon > 0$ and $\delta < \delta_\epsilon$, with $\delta_\epsilon$ as in (101), define
\[
U_\delta := \{ (\rho, \sigma) \in (0, \infty)^{2d} : \| (\rho, \sigma) - (\rho_0, \sigma_0) \|_1 < \delta \},
\]
(104)
\[
\tilde{U}_\epsilon := \{ g \in \mathcal{G}_{\Theta} : D_{H}(g, \rho_0, \sigma_0, \cdot | H_0) \leq 4\epsilon \} \subset \tilde{U},
\]
and
\[
G^{(k)}_{\Theta, \epsilon, \delta} := \phi_{H \times \Theta} (H_k \times U_\delta) \cap \tilde{U}_\epsilon, \quad k = d + 1, \ldots, \nu_n,
\]
\[
G^{(k)}_{\Theta, n} := \cup_{k=d+1}^{\nu_n} G^{(k)}_{\Theta, \epsilon, \delta},
\]
where $\nu_n$ is a sequence of integers increasing with $n$. The latter is next selected by means of a metric entropy argument, which allows to derive a sequence of test functionals $s_n$ complying with Condition 3.11(iiia). By
Lemma D.14 and Proposition C.2 in [39], we have that, for $k = d + 1, \ldots, \nu_n$,
\[
\mathcal{N} \left( 2\epsilon, \mathcal{G}_{\theta, \epsilon, \delta}^{(k)}, \mathcal{D}_H \right) \leq \mathcal{N} \left( c_1 \epsilon^2, \{ x \in \mathbb{R}^{\left| \Gamma_k \right|} : \|x\|_1 \leq 1 \}, \mathcal{L}_1 \right)
\times \mathcal{N} \left( c_2 \epsilon^2/k, B_{\delta, 1}(\rho_0), L_1 \right) \mathcal{N} \left( c_3 \epsilon^2/k, B_{\delta, 1}(\sigma_0), L_1 \right)
\leq \left( \frac{c_4}{\epsilon^2} \right)^{\left( k - 1 \right) \frac{c_5 k}{\epsilon^2}} \mathcal{N} \left( c_2 \epsilon^2/k, B_{\delta, 1}(\rho_0), L_1 \right) \mathcal{N} \left( c_3 \epsilon^2/k, B_{\delta, 1}(\sigma_0), L_1 \right)
\leq \left( \frac{c_4}{\epsilon^2} \right)^{\left( k - 1 \right) \left( \frac{c_5 k}{\epsilon^2} \right)} 2^d
\]
where $c_1, c_2, c_3, c_4$ are positive constants depending only on $d, \rho_0, \sigma_0$, $c_5$ is given by
\[
c_5 = \max \left( \sup_{x \in B_{\delta, 1}(\rho_0)} \|x\|_1 3 c_2^{-1}, \sup_{x \in B_{\delta, 1}(\sigma_0)} \|x\|_1 3 c_3^{-1} \right)
\]
and, without loss of generality, we assume that $\epsilon$ satisfies $c_1 \epsilon^2 < 1$ and
\[
c_2 \epsilon^2/k < \sup_{x \in B_{\delta, 1}(\rho_0)} \|x\|_1, \quad c_3 \epsilon^2/k < \sup_{x \in B_{\delta, 1}(\sigma_0)} \|x\|_1.
\]
Therefore,
\[
\mathcal{N} \left( 2\epsilon, \mathcal{G}_{\theta, \epsilon, \delta}^{(k)}, \mathcal{D}_H \right) \leq \sum_{k=d+1}^{\nu_n} \left( \frac{c_4}{\epsilon^2} \right)^{\left( k - 1 \right) \left( \frac{c_5 k}{\epsilon^2} \right)} 2^d
\leq \nu_n \left( \frac{c_4}{\epsilon^2} \right)^{\nu_n - 1} \left( \frac{c_5 \nu_n}{\epsilon^2} \right)^2 2^d
\]
and, for $\nu_n$ sufficiently large,
\[
\log \mathcal{N} \left( 2\epsilon, \mathcal{G}_{\theta, \epsilon, \delta}^{(k)}, \mathcal{D}_H \right) \leq c_6 \log \nu_n + c_7 \nu_n^{-1} \leq c_8 \nu_n^{-1},
\]
for some positive $c_6, c_7, c_8$. Thus, for large enough $n$, choosing
\[
\nu_n = \left\lfloor \left( \frac{n \epsilon^2}{c_8} \right)^{1/(d-1)} \right\rfloor,
\]
it holds that $\log \mathcal{N} \left( 2\epsilon, \mathcal{G}_{\theta, \epsilon, \delta}^{(k)}, \mathcal{D}_H \right) \leq n \epsilon^2$. As a result, arguments in the proof of Theorem 6.23 in [39] guarantee the existence of a sequence of measurable test functions $s_n : (0, \infty)^n \to [0, 1]$ satisfying
\[
\int_{(0, \infty)^n} s_n(x_1, \ldots, x_n) \prod_{i=1}^n g_{\rho_0, \sigma_0}(x_i|\theta_0) \, dx_i \leq e^{n \epsilon^2} e^{-2 n \epsilon^2}
\]
\[
\sup_{G \in \mathcal{G}_{\theta, \epsilon, \delta}} \int_{(0, \infty)^n} \{ 1 - s_n(x_1, \ldots, x_n) \} \prod_{i=1}^n g(x_i) \, dx_i \leq e^{-2 n \epsilon^2}.
\]
Furthermore, for $\mathcal{G}_{\epsilon, \theta, \delta}$ denoting the relative complement of $\mathcal{G}_{\theta, \epsilon, \delta}$ in $\phi_{\mathcal{H} \times \Theta} (\mathcal{H} \times \mathcal{U}_{\delta}) \cap \mathcal{U}_{\delta}$ and $q$ as in Condition 3.8(ii), as $n \to \infty$ we have
\[
\Pi_{_{\mathcal{G}_{\theta}}} (\mathcal{G}_{\epsilon, \theta, \delta}) \leq \Pi_{_{\mathcal{H} \times \Theta}} \left( (\mathcal{H} \setminus \cup_{k=d+1}^{\nu_n} \mathcal{H}_k) \times (0, \infty)^{2d} \right)
\leq e^{-q \nu_n^{-1}} \sim e^{-n q / c_8}.
\]
Finally, the existence of tests complying with Condition 3.11(iiib) is guaranteed by Lemma D.15. Condition 3.11 is now verified and the final results follows from an application of Proposition 3.12 through Corollary D.8.

D.4.4. Auxiliary results for the proof of Theorem 3.15. Following the main lines of Section D.4.2, we now provide arguments for proving that if Conditions 3.13(ii)–(iii) and 3.13(iii) in the main text hold true, Condition 3.11 is then satisfied when a $\omega$-Weibull statistical model is adopted. This fact underpins the posterior consistency results in Theorem 3.15. Connections with the proof of Theorem 4.19 are also pointed out.
We recall that, for \( H \in \mathcal{H}, (\omega, \sigma, \mu) \in (0, \infty) \times (0, \infty) \times \mathbb{R}^d \) and \( x \in (-\infty, \mu) \), the \( \omega \)-Weibull max-stable density is given by

\[
g_{\omega, \sigma, \mu}(x|H) = \prod_{j=1}^{d} \omega_j \left( \frac{\mu_j - x_j}{\sigma_j} \right)^{-\omega_j - 1} g_{l} \left( \left( \frac{\mu_1 - x_1}{\sigma_1} \right)^{-\omega_1}, \ldots, \left( \frac{\mu_d - x_d}{\sigma_d} \right)^{-\omega_d} \right) |H|.
\]

Once more, we resort to the notational convention for vectors raised to powers and rectangles in \([-\infty, \infty]^d\) introduced before Lemma D.11. Similarly to Section D.4.2, we start by defining the functional

\[
y^{(l)}_{H, \omega, \sigma, \mu} (H, \omega, \sigma, \mu; \tilde{H}, \tilde{\omega}, \tilde{\sigma}, \tilde{\mu}) := \left( \int_{(\omega, \sigma, \mu)} \log^{+} \left( \frac{g_{\omega, \sigma, \mu}(x|H)}{g_{\tilde{\omega}, \tilde{\sigma}, \tilde{\mu}}(x|\tilde{H})} \right) \right)^{1/l} g_{\tilde{\omega}, \tilde{\sigma}, \tilde{\mu}}(x|\tilde{H}) dx \right)^{1/l}
\]

for \( l \in \mathbb{N}_{+}, H, \tilde{H} \in \mathcal{H} \) and \( \omega, \tilde{\omega} \in (0, \infty), \sigma, \tilde{\sigma} \in (0, \infty), \mu, \tilde{\mu} \in \mathbb{R}^d \). Notice that, if for some \( j \in \{1, \ldots, d\} \) it holds that \( \tilde{\mu}_j < \min(\mu_j, \mu_{\ast,j}) \), then the above functional equals \(+\infty\). More generally, the semi-parametric model

\[
\{g_{\omega, \sigma, \mu}(y|H) : H \in \mathcal{H}, (\omega, \sigma, \mu) \in \Theta\},
\]

with \( \Theta = (1, \infty)^d \times (0, \infty)^d \times \mathbb{R}^d \), is less regular than the other ones considered in Section 3.3 of the main article. Therefore, the analysis of positive Kullback-Leibler divergences is herein confined to the case of \( l = 1 \). Control over the first-order positive divergence is sufficient to establish the strong consistency results of Theorem 3.15, while yielding weaker results when consistency is extended to the case of data of the form of sample maxima.

See Sections 4.5 of the main paper and Section D.7 of the present manuscript for details.

**Lemma D.16.** Let \( H \in \mathcal{H} \), with exponent functions \( V(\cdot|H) \). Let \( \tilde{H} \in \mathcal{H} \), \( \tilde{\omega} \in (1, \infty) \), \( \omega \in B_{3,1}(\tilde{\omega}) \), \( \tilde{\sigma} \in (0, \infty) \), \( \sigma \in B_{2,1}(\tilde{\sigma}) \), \( \tilde{\mu} \in \mathbb{R}^d \), \( \mu \in B_{2,1}(\tilde{\mu}) \): \( \{\mu' > \tilde{\mu} : \|\mu' - \tilde{\mu}\|_1 < \delta_3\} \) and \( \varepsilon \in (0, 1/2) \), with

\[
(105) \quad \delta_1 \in \left( 0, \min_{1 \leq j \leq d} \frac{(\omega_j - 1)}{\varepsilon} \right) \land 1, \quad \delta_2 \in \left( 0, \min_{1 \leq j \leq d} \frac{\tilde{\sigma}_j}{\varepsilon} \right), \quad \delta_3 \in (0, \varepsilon).
\]

Then, defining \( b_j := b_j(\tilde{\omega}_j, \tilde{\sigma}_j) \), \( j = 1, \ldots, d \), as in (106)

\[
c(\tilde{\omega}, \tilde{\sigma}) := \frac{2}{\min_{1 \leq j \leq d} \tilde{\omega}_j} \max_{1 \leq j \leq d} \left[ 1 + 2 \omega_j + \frac{2 \omega_j}{2 \omega_j - 1} \frac{\Gamma(\omega_j + 1)}{\Gamma(2 \omega_j)} \right] \left( 1 - \frac{1}{2 \tilde{\sigma}_j} \right) \frac{\Gamma(2 \omega_j + 1)}{\Gamma(2 \omega_j)} \left( 1 - \frac{1}{2 \tilde{\sigma}_j} \right).
\]

the term

\[
\int_{(-\infty, \tilde{\mu})} V \left( \left( \frac{\mu - x}{\sigma} \right)^{-\omega} \right) H - V \left( \left( \frac{\tilde{\mu} - x}{\sigma} \right)^{-\omega} \right) H \right) + g_{\tilde{\omega}, \tilde{\sigma}, \tilde{\mu}}(x|\tilde{H}) dx
\]

is bounded from above by \( c(\tilde{\omega}, \tilde{\sigma}) \|\mu - \tilde{\mu}\|_1 \).

**Proof.** By arguments similar to those in the proof of Lemma D.10 we obtain

\[
\left[ V \left( \left( \frac{\mu - x}{\sigma} \right)^{-\omega} \right) H - V \left( \left( \frac{\tilde{\mu} - x}{\sigma} \right)^{-\omega} \right) H \right] \leq \| \left( \frac{\mu - x}{\sigma} \right)^{-\omega} - \left( \frac{\tilde{\mu} - x}{\sigma} \right)^{-\omega} \|_1,
\]

where, by the multivariate mean-value theorem and the bounds in (105), the term on the right-hand side is bounded from above by

\[
\sum_{j=1}^{d} \frac{\omega_j}{\sigma_j} |\mu_j - \tilde{\mu}_j| |\mu_j - x_j|^{\omega_j - 1} \leq \frac{2}{\min_{1 \leq j \leq d} \tilde{\omega}_j} \sum_{j=1}^{d} (\mu_j - \tilde{\mu}_j) |\mu_j - x_j|^{\omega_j - 1},
\]

with

\[
b_j = \inf_{\sigma_j \in \{\tilde{\sigma}_j/2, \tilde{\omega}_j/2\}} \omega_j \in \{\tilde{\omega}_j/2, \tilde{\omega}_j/2\}
\]

and

\[
(106) \quad b_j = \inf_{\sigma_j \in \{\tilde{\sigma}_j/2, \tilde{\omega}_j/2\}} \omega_j \in \{\tilde{\omega}_j/2, \tilde{\omega}_j/2\}.\]
Moreover, for $j = 1, \ldots, d$, it holds that

\[
\int_{(-\infty, \bar{\mu})} |\mu_j - x_j|^{\omega_j - 1} g_{\omega_j, \bar{\sigma}}(x|\bar{H})dx \\
\leq 1 + \int_{-\infty}^{\mu_j - \bar{\mu}} |\mu_j - x_j|^{\omega_j - 1} g_{\omega_j, \bar{\sigma}}(x_j)dx_j \\
\leq 1 + 2^{\omega_j - 1} \int_{-\infty}^{\mu_j - \bar{\mu}} (-x_j)^{\omega_j - 1} g_{\omega_j, \bar{\sigma}}(x_j)dx_j \\
\leq 1 + 2^{\omega_j} + 2^{\omega_j} \int_{-\infty}^{-1} (-x_j)^{\omega_j} g_{\omega_j, \bar{\sigma}}(x_j)dx_j
\]

where, for $x_j < 0$,

\[
g_{\omega_j, \bar{\sigma}}(x_j) := \exp\left(-(-x_j/\bar{\sigma}_j)^{\omega_j}\right)(-x_j/\bar{\sigma}_j)^{-\omega_j - 1} (\bar{\omega}_j/\bar{\sigma}_j)
\]

is the two parameter (reverse) Weibull probability density and

\[
\int_{-\infty}^{-1} (-x_j)^{\omega_j} g_{\omega_j, \bar{\sigma}}(x_j)dx_j \leq \bar{\omega}_j \int_1^\infty \sigma_j^{-\bar{\omega}_j} x^{2\bar{\omega}_j - 1} \exp\left(-x_j/\bar{\sigma}_j\right) dx_j \\
= \frac{\Gamma(2\bar{\omega}_j + 1)}{2\bar{\sigma}_j(1 - 2\bar{\omega}_j)} \left\{1 - \frac{\gamma(2\bar{\omega}_j, \bar{\sigma}_j^{-\bar{\omega}_j})}{\Gamma(2\bar{\omega}_j)}\right\}
\]

with $\Gamma(\cdot)$ and $\gamma(\cdot, \cdot)$ denoting the Gamma and the lower incomplete Gamma function. The result now follows. $\square$

**Lemma D.17.** For all $k \geq d + 1$, $\bar{H} \in \mathcal{H}$, $H_k \in \mathcal{H}_k$, $\omega, \bar{\omega} \in (1, \infty)$, $\sigma, \bar{\sigma} \in (0, \infty)$ and $\mu > \bar{\mu} \in \mathbb{R}^d$ it holds that

\[
\int_{(-\infty, \bar{\mu})} \left[\max_{P \in \mathcal{P}} \max_{I_j \in P} \log^+ \left\{\frac{-V_{I_j} \left(\frac{\mu - x}{\sigma}\right)^{-\omega} \left|H_k\right|}{-V_{I_j} \left(\mu - \bar{\mu} - x\right)^{-\omega} \left|H_k\right|}\right\}\right] g_{\omega, \bar{\sigma}}(x|\bar{H})dx
\]

is bounded from above by

\[
4k \|\mu - \bar{\mu}\|_1 \max_{1 \leq j \leq d} \frac{\omega_j}{\sigma_j} \Gamma \left(1 - \frac{1}{\omega_j}\right).
\]

**Proof.** A change of variables allow to re-express the term in (108) as

\[
\int_{(-\infty, 0)} \left[\max_{P \in \mathcal{P}} \max_{I_j \in P} \log^+ \left\{\frac{-V_{I_j} \left(\frac{\mu - x}{\sigma}\right)^{-\omega} \left|H_k\right|}{-V_{I_j} \left(\mu - \bar{\mu} - x\right)^{-\omega} \left|H_k\right|}\right\}\right] g_{\omega, \bar{\sigma}, 0}(x|\bar{H})dx.
\]

As done in the proof of Lemma D.11, by reasoning on a case by case basis via Lemma D.2 and the multivariate mean-value theorem, we end up with the general inequality

\[
\log^+ \left\{\frac{-V_{I_j} \left(\frac{-x}{\sigma}\right)^{-\omega} \left|H_k\right|}{-V_{I_j} \left(\mu - \bar{\mu} - x\right)^{-\omega} \left|H_k\right|}\right\} \leq 4k \sum_{j=1}^{d} \omega_j \frac{\mu_j - \bar{\mu}_j}{-x_j}
\]

for all $I_j \in P$, $P \in \mathcal{P}$. Consequently, the term in (108) is bounded from above by

\[
4k \sum_{j=1}^{d} \omega_j \int_{-\infty}^{0} \frac{\mu_j - \bar{\mu}_j}{-x_j} g_{\omega_j, \bar{\sigma}}(x_j)dx_j = 4k \sum_{j=1}^{d} \frac{\omega_j}{\sigma_j} \Gamma \left(1 - \frac{1}{\omega_j}\right) \left(\mu_j - \bar{\mu}_j\right),
\]

where $g_{\omega_j, \bar{\sigma}}(x_j)$ is the univariate probability density function defined in (107). The result now follows. $\square$
LEMMA D.18. Let $H_0 \in \mathcal{H}$ and $(\omega_0, \sigma_0, \mu_0) \in \Theta$. Then, for all $\epsilon > 0$ there exists $H^* \in \mathcal{H}'$ and $\delta_1, \delta_2, \delta_3, \delta_4 > 0$ such that

$$\mathcal{X}^1_+ (g_{\omega_0, \sigma_0, \mu_0}(\cdot|H_0), g_{\omega, \sigma, \mu}(\cdot|H)) < \epsilon$$

for all $\omega \in B_{\delta_1}(\omega_0), \sigma \in B_{\delta_2}(\sigma_0), \mu \in B_{\delta_3}(\mu_0)$ and $H \in B_{\delta_4}(H^*)$. Thus, in particular,

$$\mathcal{X}^1(g_{\omega_0, \sigma_0, \mu_0}(\cdot|H_0), g_{\omega, \sigma, \mu}(\cdot|H)) < \epsilon.$$

PROOF. Observe that

$$\mathcal{X}^1_+ (g_{\omega_0, \sigma_0, \mu_0}(\cdot|H_0), g_{\omega, \sigma, \mu}(\cdot|H)) = \gamma_{H_0, \omega_0, \sigma_0, \mu_0}^1 (H_0, \omega_0, \sigma_0, \mu_0; H, \omega, \sigma, \mu)$$

$$\leq \gamma_{H_0, \omega_0, \sigma_0, \mu_0}^1 (H_0, \omega_0, \sigma_0, \mu_0; H, \omega_0, \sigma_0, \mu_0)$$

$$\phantom{\leq} + \gamma_{H_0, \omega_0, \sigma_0, \mu_0}^1 (H, \omega_0, \sigma_0, \mu_0; H, \omega, \sigma, \mu)$$

$$\phantom{\leq} + \gamma_{H_0, \omega_0, \sigma_0, \mu_0}^1 (H, \omega, \sigma, \mu; H, \omega, \sigma, \mu)$$

$$=: S_1 + S_2 + S_3.$$

A change of variables yields $S_1 = \mathcal{X}^1(g_1(\cdot|H_0), g_1(\cdot|H))$, where the term on the right-hand side is defined as in Section C.3. Thus, by Proposition C.7-C.8, $S_1$ can be made strictly smaller than $\epsilon/3$ by choosing $H \in B_{\delta_4}(H^*)$, for a suitable $H^* \in \mathcal{H}'$ and a sufficiently small $\delta_4$. Similarly, we have

$$S_2 = \gamma_{H_0, \rho_0, \sigma_0}^1 (H, \rho_0, \sigma_0^{-1}; H, \rho, \sigma^{-1}),$$

where the term on the right-hand side is defined as in Section D.4.2, with $\rho_0 \equiv \omega_0$ and $\rho = \omega$. Thus, $S_2$ can be made strictly smaller than $\epsilon/3$ by reasoning as in the proof of Lemma D.12. We are now left with the analysis of $S_3$.

Using the inequality $\log^+ (xyz) \leq \log^+ (x) + \log^+ (y) + \log^+ (z)$, $x, y, z > 0$, we obtain

$$S_3 \leq \int_{(-\infty, \mu_0)} \left[ \mathcal{V} \left( \left\{ \frac{\mu - x}{\sigma} \right\}^{-\omega} H \right) - \mathcal{V} \left( \left\{ \frac{\mu_0 - x}{\sigma} \right\}^{-\omega} H \right) \right] g_{\omega_0, \sigma_0, \mu_0}(x|H_0) \, dx$$

$$\phantom{S_3 \leq} + \sum_{j=1}^d \int_{(-\infty, \mu_0)} \log^+ \left\{ \frac{\mu_j - x_j}{\mu_{0,j} - x_j} \right\} g_{\omega_0, \sigma_0, \mu_0}(x|H_0) \, dx$$

$$\phantom{S_3 \leq} + \int_{(-\infty, \mu_0)} \log^+ \left\{ \sum_{\mathcal{P} \in \mathcal{P}_d} \prod_{i=1}^m \left[ -V_{l_i} \left( \left\{ \frac{\mu_0 - x}{\sigma} \right\}^{-\omega} H \right) \right] \right\} g_{\omega_0, \sigma_0, \mu_0}(x|H_0) \, dx$$

$$=: T_1 + T_2 + T_3.$$

By Lemma D.16, the term $T_1$ can be made arbitrarily small by choosing $\delta_1, \delta_2, \delta_3$ small enough. Furthermore, using the inequality $\log^+(1 + z) \leq z$ for $z > 0$, we conclude that

$$T_2 \leq \sum_{j=1}^d \int_{(-\infty, \mu_0)} \frac{\mu_j - \mu_{0,j}}{\mu_{0,j} - x_j} g_{\omega_0, \sigma_0, \mu_0}(x|H_0) \, dx$$

$$= \sum_{j=1}^d \frac{\mu_j - \mu_{0,j}}{\sigma_{0,j}} \Gamma(1 - 1/\omega_{0,j})$$

where the term on the right-hand side can be made arbitrarily small by choosing $\delta_3$ small enough. Reasoning as in the proof of Lemma D.12 and using Lemma D.17, we also conclude that $T_3$ can be made small by choosing $\delta_1, \delta_2, \delta_3, \delta_4$ small enough. Hence, there exists suitable choices of the radii yielding $S_3 < \epsilon/3$. The proof is now complete.
For given $H_0 \in \mathcal{H}_0$, $\omega_0 \in (1, \infty)$, $\sigma_0 \in (0, \infty)$, $\mu_0 \in \mathbb{R}^d$ and for all $\epsilon > 0$, define
\begin{equation}
\mathcal{V}_\epsilon := \{(H, \omega, \sigma, \mu) \in \mathcal{H} \times \Theta : \mathcal{X}(g_{\omega_0, \sigma_0, \mu_0}(\cdot|H_0), g_{\omega, \sigma, \mu}(\cdot|H)) < \epsilon \}.
\end{equation}
\begin{equation}
\mathcal{V}^{(1)}_\epsilon := \{(H, \omega, \sigma, \mu) \in \mathcal{H} \times \Theta : (\omega, \sigma, \mu) \in B_{\epsilon, 1}(\omega_0) \times B_{\epsilon, 1}(\sigma_0) \times B_{\epsilon, 1}(\mu_0), \mathcal{X}^{(1)}(g_{\omega_0, \sigma_0, \mu_0}(\cdot|H_0), g_{\omega, \sigma, \mu}(\cdot|H)) < \epsilon \}.
\end{equation}
The following result is an immediate consequence of Lemma D.18. Its proof is analogous to that of Corollary D.19. Preliminarily observe that, by triangular inequality,

\textbf{Corollary D.19.} Under Condition 3.13(ii) and Conditions 3.13(ii)–(iii), for all $\epsilon > 0$ we have that
\begin{equation}
\Pi_{\mathcal{H} \times \Theta}(\mathcal{V}_\epsilon) \geq \Pi_{\mathcal{H} \times \Theta}(\mathcal{V}^{(1)}_\epsilon) > 0.
\end{equation}
Thus, the induced prior $\Pi_{\mathcal{G}_\Theta}$ on the Borel sets of $\mathcal{G}_\Theta = \{g_{\omega, \sigma, \mu}(\cdot|H) : H \in \mathcal{H}, (\omega, \sigma, \mu) \in \Theta \}$, where
\begin{equation}
\Theta = (1, \infty)^d \times (0, \infty)^d \times \mathbb{R}^d,
\end{equation}
possesses the Kullback-Leibler property.

The last two lemmas provide results for $\omega$-Weibull models which parallel those established in Lemmas D.14–D.15 for $\rho$-Fréchet distributions. We start by determining a convenient upper-bound for Hellinger distances.

\textbf{Lemma D.20.} For all $k \geq d + 1$, $H_k, \tilde{H}_k \in \mathcal{H}_k$ with mixture weights $\varphi, \hat{\varphi} \in \Phi_k$, $\omega_0 \in (1, \infty)$, $\omega, \tilde{\omega} \in B_{\delta_1, 1}(\omega_0)$, $\sigma \in (0, \infty)$, $\sigma, \tilde{\sigma} \in B_{\delta_2, 1}(\sigma_0)$, $\mu_0 \in \mathbb{R}^d$, $\mu, \tilde{\mu} \in B_{\delta_3, 1}(\mu_0)$ with
\begin{equation}
\delta_0 < \frac{1}{5} \min \left( \min_{1 \leq j \leq d} \omega_0, j, \min_{1 \leq j \leq d} \sigma_0^{-1} \right)
\end{equation}
\begin{equation}
\delta_1 < (1 - p) \min_{1 \leq j \leq d} \omega_0, j, \left( \min_{1 \leq j \leq d} \omega_0, j \right)^{-1}, \quad p < 1,
\end{equation}
\begin{equation}
\delta_2 < \frac{16}{25} \min_{1 \leq j \leq d} \sigma_0^2, \quad \delta_3 < 1/4,
\end{equation}
there exists a constant $c_0 > 0$ depending only on $\omega_0, \sigma_0, \mu_0$ such that
\begin{equation}
\mathcal{D}_H(g_{\omega, \sigma, \mu}(\cdot|H_k), g_{\omega, \tilde{\sigma}, \tilde{\mu}}(\cdot|\tilde{H}_k)) \leq \sqrt{c_0} \| \varphi - \hat{\varphi} \|_1 + \sqrt{c_0} \| \omega - \tilde{\omega} \|_1 + c_0 k \| \sigma - \tilde{\sigma} \|_1
\end{equation}
\begin{equation}
+ c_0 k \| \mu - \tilde{\mu} \|_1,
\end{equation}
where $\varphi_0$ and $\hat{\varphi}_0$ are defined as in (52).

\textbf{Proof.} Preliminarily observe that, by triangular inequality,
\begin{equation}
\mathcal{D}_H(g_{\omega, \sigma, \mu}(\cdot|H_k), g_{\omega, \tilde{\sigma}, \tilde{\mu}}(\cdot|\tilde{H}_k)) \leq \mathcal{D}_H(g_{\omega, \sigma, \mu}(\cdot|H_k), g_{\omega, \tilde{\sigma}, \tilde{\mu}}(\cdot|\tilde{H}_k))
\end{equation}
\begin{equation}
+ \mathcal{D}_H(g_{\omega, \tilde{\sigma}, \tilde{\mu}}(\cdot|H_k), g_{\omega, \sigma, \mu}(\cdot|H_k))
\end{equation}
\begin{equation}
+ \mathcal{D}_H(g_{\omega, \sigma, \mu}(\cdot|H_k), g_{\omega, \sigma, \mu}(\cdot|\tilde{H}_k))
\end{equation}
\begin{equation}
=: S_1 + S_2 + S_3.
\end{equation}
We have that
\begin{equation}
\| \sigma^{-1} - \tilde{\sigma}^{-1} \|_1 \leq \left( \min_{1 \leq j \leq d} \sigma_j \tilde{\sigma}_j \right)^{-1} \| \sigma - \tilde{\sigma} \|_1
\end{equation}
\begin{equation}
\leq \left( \frac{4}{5} \min_{1 \leq j \leq d} \sigma_0^{-1} \right)^{-2} \| \sigma - \tilde{\sigma} \|_1
\end{equation}
\begin{equation}
< \delta_0,
\end{equation}
therefore, denoting $\rho_0 = \omega_0$, $\rho = \omega$, $\tilde{\rho} = \tilde{\omega}$, it holds that $\rho, \tilde{\rho} \in B_{\delta_0, 1}(\rho_0)$ and $\sigma^{-1}, \tilde{\sigma}^{-1} \in B_{\delta_0, 1}(\sigma_0^{-1})$. Moreover, a change of variables yields
\begin{equation}
S_1 = \mathcal{D}_H(g_{\rho, \sigma^{-1}}(\cdot|H_k), g_{\tilde{\rho}, \tilde{\sigma}^{-1}}(\cdot|\tilde{H}_k)).
\end{equation}
Consequently, by Lemma D.14 and \( (115) \),

\[
S_1 \leq \sqrt{c_1 \| \varphi - \varphi_c \|_1^2 + \sqrt{c_1 k \| \omega - \omega_c \|_1^2 + c_2 k \| \sigma - \sigma_c \|_1^2}},
\]

where \( c_1 \) is a positive constant depending only on \( d, \omega_0 \) and \( \sigma_0 \). Furthermore, by Lemma B.1(iv) in \([39]\)

\[
S_2^2 \leq \mathcal{X}(g_{\omega, \sigma, \mu}(\cdot | H_k), g_{\omega, \sigma, \mu \vee \mu}(\cdot | H_k))
\]

\[
\leq \int_{(-\infty, \mu)} \left[ \Gamma \left( \left( \frac{\mu - x}{\sigma} \right)^{\omega} \right) - x \right] g_{\omega, \sigma, \mu}(x \mid H_k)dx + \sum_{j=1}^{d} \int_{(-\infty, \mu)} \log \left( \frac{\mu_j \vee \hat{\mu}_j - x_j}{\mu_j - x_j} \right) g_{\omega, \sigma, \mu}(x \mid H_k)dx
\]

\[
+ \int_{(-\infty, \mu)} \log \left( \frac{\sum_{P \in \mathcal{P}_d} \prod_{i=1}^{n} V_{I_i} \left( \left( \frac{\mu \vee \hat{\mu} - x}{\sigma} \right)^{-\omega} \right) H_k }{\sum_{P \in \mathcal{P}_d} \prod_{i=1}^{n} V_{I_i} \left( \left( \frac{\mu - x}{\sigma} \right)^{-\omega} \right) H_k } \right) g_{\omega, \sigma, \mu}(x \mid H_k)dx
\]

\[= T_1 + T_2 + T_3.\]

Since \( \mu \vee \hat{\mu} \in B_{2d, \infty}(\mu) \), by Lemma D.16 and the bounds in \((114)\)

\[
T_1 \leq c_2 \| \mu \vee \hat{\mu} - \mu \|_1 \leq c_2 \| \hat{\mu} - \mu \|_1,
\]

where \( c_2 > 0 \) only depends on \( \omega_0, \sigma_0 \). Moreover, following steps similar to those in \((111)\) and exploiting \((114)-(115)\), we obtain

\[
T_2 \leq \sum_{j=1}^{d} \frac{\mu_j \vee \hat{\mu}_j - \mu_j}{\sigma_j} \Gamma(1 - 1/\omega_j)
\]

\[
\leq \frac{5}{4} \max_{1 \leq j \leq d} \frac{\Gamma(1 - 1/(p_{\omega_{0,j}}))}{\sigma_{0,j}} \| \mu - \hat{\mu} \|_1.
\]

Finally, by Lemma D.17 and the bounds in \((114)\)

\[
T_3 \leq 4k \| \hat{\mu} \vee \mu - \mu \|_1 \max_{1 \leq j \leq d} \frac{\omega_{0,j}}{\sigma_{0,j}} \Gamma \left( 1 - \frac{1}{\omega_j} \right)
\]

\[
\leq 10k \| \hat{\mu} - \mu \|_1 \max_{1 \leq j \leq d} \frac{\omega_{0,j}}{\sigma_{0,j}} \Gamma \left( 1 - \frac{1}{p_{\omega_{0,j}}} \right).
\]

The term \( S_3 \) can be bounded from above analogously, whence the conclusion. \( \square \)

We now state the last technical result of this subsection, concerning tests on the marginal shape, scale and location parameters. We recall that, for all \( \omega, \sigma > 0, \mu \in \mathbb{R} \), the distribution function of a univariate location-scale (reverse) \( \omega \)-Weibull distribution is

\[
F_{\omega, \sigma, \mu}(x) = \exp \left( \frac{-(\mu - x)}{\sigma} \right), \quad x < \mu.
\]

As usual, \( F_{\omega, \sigma, \mu}^{(n)} \) denotes the pertaining \( n \)-fold pm. Moreover, we make use following short notation for measurable functions \( f \)

\[
F_{\omega, \sigma, \mu}^{(n)} f = \int_{(-\infty, \mu)^n} f(x_1, \ldots, x_n) d F_{\omega, \sigma, \mu}^{(n)}(x_1, \ldots, x_n).
\]

In what follows, as a convention, real valued terms of the form of supreme over an empty set are interpreted as equal to zero. The proof makes use of some basic properties of the \( 1 \)-Wasserstein distance for univariate distributions, we point the interested reader to Section 1.2.3 of \([60]\).
For \( j = 1, \ldots, d \), let \( K_j \subset (1, \infty) \times (0, \infty) \times \mathbb{R} \) be a compact neighbourhood of \((\omega_{0,j}, \sigma_{0,j}, \mu_{0,j})\). Then, for each \( \epsilon > 0 \) and \( j = 1, \ldots, d \), there exist Borel-measurable functions \( t_{n,j} : \mathbb{R}^n \rightarrow \{0, 1\} \), \( n \in \mathbb{N}_+ \), and a positive constant \( c_j(\epsilon) \) such that

\[
G_{\omega_0,j, \sigma_0,j, \mu_0,j}^{(n)} t_{n,j} \leq 2e^{-nc_j(\epsilon)},
\]

where \( \Theta_j^{(A)} = \{(\omega, \sigma, \mu) \in K_j : \| (\omega, \sigma, \mu) - (\omega_{0,j}, \sigma_{0,j}, \mu_{0,j}) \|_\infty > \epsilon \} \).

**Proof.** Fix \( j \in \{1, \ldots, d\} \). Observe that for any \( \epsilon > 0 \) there exist \( 0 < \epsilon_i < \min(\epsilon, 1) \), \( i = 1, 2, 3 \), such that

\[
(\omega_{0,j} \pm \epsilon_1) \times (\sigma_{0,j} \pm \epsilon_2) \times (\mu_{0,j} \pm \epsilon_3) \subset K_j,
\]

and \( \Theta_j^{(A)} \subset \bigcup_{i=1}^{3} \Theta_j^{(A)_i} \), where

\[
\Theta_j^{(A)_1} := \{(\omega, \sigma, \mu) \in K_j : |\mu - \mu_{0,j}| > \epsilon_3\},
\]

\[
\Theta_j^{(A)_2} := \{(\omega, \sigma, \mu) \in K_j : |\omega - \omega_{0,j}| > \epsilon_1, |\mu - \mu_{0,j}| < \epsilon_3\},
\]

\[
\Theta_j^{(A)_3} := \{(\omega, \sigma, \mu) \in K_j : |\sigma - \sigma_{0,j}| > \epsilon_2, |\mu - \mu_{0,j}| < \epsilon_3\}.
\]

We show that for \( i = 1, 2, 3 \) there exists \( \delta_i > 0 \) such that

\[
\sup_{(\omega, \sigma, \mu) \in \Theta_j^{(A)}} \mathcal{D}_{KS}(G_{\omega, \sigma, \mu}, G_{\omega_0,j, \sigma_0,j, \mu_0,j}) > \delta_i.
\]

This allows to conclude that there exists \( \delta > 0 \) such that

\[
\sup_{(\omega, \sigma, \mu) \in \Theta_j^{(A)}} \mathcal{D}_{KS}(G_{\omega, \sigma, \mu}, G_{\omega_0,j, \sigma_0,j, \mu_0,j}) > \delta,
\]

therefore a test functional \( t_{n,j} \) complying with (117) can be constructed via the empirical distribution function, as in the proof of Lemma D.15.

We start by remarking that, in view of the Heine-Borel theorem, there exist \( 1 < \omega < \overline{\omega} < \infty \), \( 0 < \underline{\sigma} < \sigma < \infty \) and \( -\infty < \mu < \overline{\mu} < \infty \) such that the projection of \( K_j \) into its \( \omega \)-, \( \sigma \)- or \( \mu \)-component is contained in the interval \((\omega, \overline{\omega})\), \((\underline{\sigma}, \sigma)\) and \((\mu, \overline{\mu})\), respectively. Without loss of generality, we assume \( \sigma < 1, \overline{\sigma} > 1 \). Consequently, if \( (\omega, \sigma, \mu) \in \Theta_j^{(A)} \) and \( \mu > \mu_{0,j} \), we have that

\[
\mathcal{D}_{KS}(G_{\omega, \sigma, \mu}, G_{\omega_0,j, \sigma_0,j, \mu_0,j}) > 1 - \exp\left\{ -\left(\frac{\mu - \mu_0}{\sigma}\right) \omega \right\} > 1 - \exp\left\{ -\left(\frac{\epsilon_3}{\sigma}\right) \overline{\omega} \right\} =: \delta_1^+.
\]

Analogously, if \( (\omega, \sigma, \mu) \in \Theta_j^{(A)_1} \) and \( \mu < \mu_{0,j} \), we have that

\[
\mathcal{D}_{KS}(G_{\omega, \sigma, \mu}, G_{\omega_0,j, \sigma_0,j, \mu_0,j}) > 1 - \exp\left\{ -\left(\frac{\mu_0 - \mu}{\sigma_0}\right) \omega_0 \right\} > 1 - \exp\left\{ -\left(\frac{\epsilon_3}{\sigma_0}\right) \omega_0 \right\} =: \delta_1^-,
\]

thus a valid choice of \( \delta_1 \) is \( \delta_1 = \min(\delta_1^+, \delta_1^-) \).
Next, assume that \((\omega, \sigma, \mu) \in \Theta^{(A)}_{j,2}\) and \(\omega > \omega_0\). Then, for \(z \to -\infty\), we have
\[
\frac{\exp\left\{-(\mu - z)\omega \sigma^{-\omega}\right\}}{\exp\left\{-(\mu_0 - z)\omega_0 \sigma_0^{-\omega_0}\right\}} \leq \frac{\exp\left\{-(\mu_0 - \epsilon_3 - z)\omega_0 + \epsilon_1 \sigma_0^{-\omega_0 - \epsilon_1}\right\}}{\exp\left\{-(\mu_0 - z)\omega_0 \sigma_0^{-\omega_0}\right\}} \to 0.
\]

Hence, for any \(s \in (0,1)\), choosing \(z < 0\) with \(|z|\) sufficiently large we obtain:
\[
\mathcal{D}_{KS}(G_{\omega,\mu,\sigma, G_{\omega_0,\sigma_0,\mu_0,0}}) > 1 - \frac{\exp\left\{-(\mu_0 - y)\omega_0 \sigma_0^{-\omega_0}\right\}}{\exp\left\{-(\mu - y)\omega \sigma^{-\omega}\right\}} G_{\omega_0,\sigma_0,\mu_0,0}(z) \\
> (1 - s) G_{\omega_0,\sigma_0,\mu_0,0}(z) \\
= : \delta^+_2.
\]

By a similar reasoning, choosing \(y < 0\) with \(|y|\) sufficiently large, for all \((\omega, \sigma, \mu) \in \Theta^{(A)}_{j,2}\) such that \(\omega < \omega_0\) we obtain
\[
\mathcal{D}_{KS}(G_{\omega,\sigma,\mu, G_{\omega_0,\sigma_0,\mu_0,0}}) > 1 - \frac{\exp\left\{-(\mu_0 - y)\omega_0 \sigma_0^{-\omega_0}\right\}}{\exp\left\{-(\mu - y)\omega \sigma^{-\omega}\right\}} G_{\omega,\sigma,\mu}(y) \\
> (1 - s) \exp\left\{-(\mu_0 + \epsilon_3 - y)\omega_0 - \epsilon_1 \sigma_0^{-\omega_0 + \epsilon_1}\right\} \\
=: \delta^-_2.
\]

A suitable choice of \(\delta_2\) is therefore \(\delta_2 = \min(\delta^+_2, \delta^-_2)\).

Finally, assume that \((\omega, \sigma, \mu) \in \Theta^{(A)}_{j,3}\) and denote by \(\mathcal{D}_{W,1}\) the 1-Wasserstein distance. We have
\[
\mathcal{D}_{W,1}(G_{\omega,\sigma,\mu, G_{\omega_0,\sigma_0,\mu_0,0}}) = \|G_{\omega,\sigma,\mu} - G_{\omega_0,\sigma_0,\mu_0,0}\|_1 \\
= \int_0^1 |\mu - \mu_0 - \sigma (\log(1/u))^{1/\omega} + \sigma_0 (\log(1/u))^{1/\omega_0}| du \\
\geq \int_0^1 |\sigma_0 (\log(1/u))^{1/\omega_0} - \sigma (\log(1/u))^{1/\omega}| du - \epsilon_3
\]
where the first term on the right-hand side equals
\[
\mathcal{D}_{W,1}(G_{\omega,\sigma,0, G_{\omega_0,\sigma_0,0,0}}) = \|G_{\omega,\sigma,0} - G_{\omega_0,\sigma_0,0,0}\|_1 \\
\geq |\sigma_0 \Gamma(1 + 1/\omega_0) - \sigma \Gamma(1 + 1/\omega)| \\
\geq \Gamma(1 + 1/\omega)|\sigma - \sigma_0| - \sigma_0 |\Gamma(1 + 1/\omega_0) - \Gamma(1 + 1/\omega)|.
\]

Therefore, using (118) along with the bound \(\inf_{x \in [1,2]} \Gamma(x) > 4/5\), we deduce that
\[
\mathcal{D}_{W,1}(G_{\omega,\sigma,\mu, G_{\omega_0,\sigma_0,\mu_0,0}}) > \frac{4}{15} \epsilon_2 =: \delta_{W,1}
\]
and a reasoning by contradiction now allows to conclude that there exists \(\delta_3 > 0\) satisfying (119) with \(i = 3\). Indeed, if this was not the case, we could select a negative \(t < \mu_0 - \epsilon_3\) with \(|t|\) sufficiently large to guarantee
\[
\int_{-\infty}^t \exp\left\{-\left(\frac{\mu - x}{\sigma}\right)^\omega\right\} dx + \int_{-\infty}^t \exp\left\{-\left(\frac{\mu_0 - x}{\sigma_0}\right)^\omega_0\right\} dx \leq 2\frac{\sigma}{\sigma_0^{\omega_0}} \exp\left\{-\left(\frac{\mu - x}{\sigma}\right)^\omega\right\} < \delta_{W,1}/2.
\]
for all \((\omega, \sigma, \mu) \in \Theta(A)\) and, for \(\delta_3\) satisfying \((\mu + \varepsilon_3 - t)\delta_3 < \delta_{W,1}/2\), there would exist some \((\omega, \sigma, \mu) \in \Theta(A)\) satisfying \(\mathcal{D}_{KS}(G_{\omega, \sigma, \mu}, G_{\omega, j, \sigma_{0,j}, \mu_{0,j}}) < \delta_3\), whence

\[
\mathcal{D}_{W,1}(G_{\omega, \sigma, \mu}, G_{\omega, j, \sigma_{0,j}, \mu_{0,j}}) = \|G_{\omega, \sigma, \mu} - G_{\omega, j, \sigma_{0,j}, \mu_{0,j}}\|_1 \\
\leq \int_{-\infty}^{t} \exp \left\{ - \left( \frac{\mu - x}{\sigma} \right)^{\omega_{0,j}} \right\} dx + \int_{-\infty}^{t} \exp \left\{ - \left( \frac{\mu_{0,j} - x}{\sigma_{0,j}} \right)^{\omega_{0,j}} \right\} dx \\
+ |\max(\mu, \mu_0) - t| \mathcal{D}_{KS}(G_{\omega, \sigma, \mu}, G_{\omega, j, \sigma_{0,j}, \mu_{0,j}}) \leq \delta_{W,1},
\]

yielding a contradiction. This completes the proof. \(\square\)

**D.4.5. Proof of Theorem 3.15.** The proof proceeds analogously to that of Theorem 3.14 (Section D.4.3 of the present manuscript), now using Corollary D.19 in place of Corollary D.13, Lemma D.20 in place of D.14 and Lemma D.21 in place of Lemma D.15. Full derivations are therefore omitted.

**D.4.6. Proof of Theorem 3.17.** We have the following facts:

**Fact 1.** For all \(l \in \mathbb{N}_+, H \in \mathcal{H}, \sigma > 0\) and \(\mu \in \mathbb{R}^d\), the \(l\)-th order positive Kullback-Leibler divergence from \(g_{\sigma, \mu}(\cdot|H)\) to the true data generating density \(g_{\sigma_0, \mu_0}(\cdot|H_0)\), i.e.

\[
\int_{\mathbb{R}^d} \log^+ \left\{ \frac{g_{\sigma_0, \mu_0}(x|H_0)}{g_{\sigma, \mu}(x|H)} \right\}^l g_{\sigma_0, \mu_0}(x|H_0) dx,
\]

equals the \(l\)-th power of

\[
g_{H_0,1, \sigma_0}^{(l)}(H_0, \rho_0, \sigma_0^*, H, \rho^*, \sigma^*),
\]

which is defined as in Section D.4.2, with

\[
\rho_0 = \left( \begin{array}{c} -1 \\ \rho_{0,1} \cdots, \rho_{0,d} \end{array} \right), \quad \sigma_0^* = \left( e^{\rho_{0,1}}, \ldots, e^{\rho_{0,d}} \right),
\]

\[
\rho^* = (\sigma_{0,1}^{-1}, \ldots, \sigma_{0,d}^{-1}), \quad \sigma^* = \left( e^{\rho_{1,1}}, \ldots, e^{\rho_{d,d}} \right).
\]

Analogously,

\[
\mathcal{X}(g_{\sigma_0, \mu_0}(\cdot|H_0), g_{\sigma, \mu}(\cdot|H)) = \mathcal{X}(g_{\rho_0, \sigma_0^*}(\cdot|H_0), g_{\rho^*, \sigma^*}(\cdot|H)).
\]

Clearly, for any \(\delta > 0\) there exists \(\varepsilon > 0\) such that

\[
\sigma \in B_{\varepsilon,1}(\sigma_0) \quad \text{and} \quad \mu \in B_{\varepsilon,1}(\mu_0) \implies \rho^* \in B_{\delta,1}(\rho_0^*) \quad \text{and} \quad \sigma^* \in B_{\delta,1}(\sigma_0^*).
\]

Therefore, results analogous to Lemma D.12 and Corollary D.13 hold true also in the present setting. In particular, for any \(\varepsilon > 0\), the prior \(\Pi_{H \times \Theta}\) resulting from Conditions 3.13(ii)–(iii) and 3.13(iii) possesses the Kullback-Leibler property.

**Fact 2.** For \(j = 1, \ldots, d, \sigma_j > 0\) and \(\mu_j \in \mathbb{R}\), recalling that \(G_{\mu_j, \sigma_j}(x) = \exp(-\exp((\mu_j - x)/\sigma_j)), x \in \mathbb{R}\), we have that

\[
\mathcal{D}_{KS}(G_{\sigma_{0,j}, \mu_{0,j}}, G_{\sigma, \mu}) = \mathcal{D}_{KS}\left(G_{\rho_{0,j}^*, \sigma_{0,j}^*}, G_{\rho_j^*, \sigma_j^*}\right)
\]

where the parameters \(\rho_{0,j}^*, \sigma_{0,j}^*, \rho_j^*, \sigma_j^*\) are defined as in (121) and, for \(x > 0\),

\[
G_{\rho_{0,j}^*, \sigma_{0,j}^*}(x) = \exp \left\{ - \left( \frac{x}{\sigma_{0,j}^*} \right)^{-\rho_{0,j}^*} \right\},
\]

\[
G_{\rho_j^*, \sigma_j^*}(x) = \exp \left\{ - \left( \frac{x}{\sigma_j^*} \right)^{-\rho_j^*} \right\}.
\]

Furthermore, for any \(\delta > 0\) there exists \(\varepsilon > 0\) such that

\[
(\sigma_j, \mu_j) \in B_{\varepsilon, \infty}(\sigma_{0,j}, \mu_{0,j}) \implies (\rho_j^*, \sigma_j^*) \in B_{\varepsilon, \infty}(\rho_{0,j}^*, \sigma_{0,j}^*).
\]
Consequently, arguments analogous to those in the proof of Lemma D.15 apply also in the present setting and a similar result obtains.

**Fact 3.** For all \( k \geq d + 1 \), \( H_k, \tilde{H}_k \in H_k \), \( \sigma, \tilde{\sigma} \in B_{\varepsilon,1}(\sigma_0) \) and \( \mu, \tilde{\mu} \in B_{\varepsilon,1}(\mu_0) \), with \( \varepsilon > 0 \), a change of variables yields

\[
\varphi_H(g_{\sigma,\mu}(\cdot|H_k), g_{\tilde{\sigma},\tilde{\mu}}(\cdot|\tilde{H}_k)) = \varphi_H(g_{\rho^*,\sigma^*}(\cdot|H_k), g_{\tilde{\rho}^*,\tilde{\sigma}^*}(\cdot|\tilde{H}_k)),
\]

where \( \rho^*, \sigma^* \) are given in (121) and \( \tilde{\rho}^*, \tilde{\sigma}^* \) are defined in an analogous fashion. Let \( \rho_0^*, \sigma_0^* \) be as in (121) and choose \( \varepsilon > 0 \) such that the implication in (122) is satisfied for a positive \( \delta \equiv \delta_0 > 0 \) complying with (101), where \( \rho_{0,j} \) and \( \sigma_{0,j} \) are now replaced by \( \rho_0^* \) and \( \sigma_0^* \), respectively, for \( j = 1, \ldots, d \). Hence, assuming without loss of generality that

\[
\varepsilon < \frac{1}{2} \min \left\{ \min_{1 \leq j \leq d} \sigma_{0,j}, \min_{1 \leq j \leq d} |\mu_{0,j}| \right\},
\]

by Lemma D.14, the term on the right-hand side of (123) is bounded from above by

\[
\sqrt{c_0} \| \varphi_0 - \varphi_\delta \|_1 + \sqrt{c_0 k} \| \rho^* - \tilde{\rho}^* \|_1 + c_0 k \| \sigma^* - \tilde{\sigma}^* \|_1
\]

\[
\leq \sqrt{c_0} \| \varphi_0 - \varphi_\delta \|_1 + \sqrt{c_0 k} \| \sigma - \tilde{\sigma} \|_1 + c_0 k \| \mu - \tilde{\mu} \|_1,
\]

where \( c_0 \) is a positive constant depending on \( d, \rho_0^* \) and \( \sigma_0^* \),

\[
c_0^* := c_0 \max \left\{ \frac{4}{\min_{1 \leq j \leq d} \sigma_{0,j}^2}, \frac{\max_{1 \leq j \leq d} e^{2 |\mu_{0,j}|}}{\varepsilon} \right\},
\]

while \( \| \varphi_0 - \varphi_\delta \|_1 \) is as in Lemma D.5.

Combining Facts 1-3, the results in the statement can be established by mirroring the proof of Theorem 3.14.

### D.5. Proofs of the results in Section 4.2.

#### D.5.1. Proof of Proposition 4.5.  

Preliminary observe that, under Condition 4.1, the true limiting max-stable density \( g_{\vartheta_0}(\cdot|H_0) \) is positive and continuous on:

* \( (0, \infty) \) in the multivariate \( \rho \)-Fréchet case, where \( \vartheta_0 = (\rho_0, 1) \), or
* \( (-\infty, 0) \) in the multivariate \( \omega \)-Weibull case, where \( \vartheta_0 = (\omega_0, 1, 0) \), or
* \( \mathbb{R}^d \), in the multivariate Gumbel case, where \( \vartheta_0 = (1, 0) \).

Moreover, for the choices of \( a_{m_n} \) and \( b_{m_n} \) given in Section 4.1 of the main article, we have that

\[
\{ x \in \mathbb{R}^d : f_{m_n}(x) > 0 \} \subset \{ x \in \mathbb{R}^d : g_{\vartheta_0}(x|H_0) > 0 \}.
\]

In this view, let \( E_n \) be a sequence of measurable events such that \( C_{\vartheta_0}^a(E_n|H_0) = o(e^{-nc}) \) as \( n \to \infty \), for some \( c > 0 \), and notice that for any \( \epsilon \in (0, c) \)

\[
Q_n(E_n) \leq e^{nc} C_{\vartheta_0}^a(E_n|H_0) + Q_n(E_n \cap \{ R_n > \epsilon \})
\]

(124)

where \( R_n \) denotes the sequence of rescaled of log likelihood ratios

\[
R_n := \frac{1}{n} \sum \log \left\{ \frac{f_{m_n}(\mathcal{M}_{m_n}\cdot)}{g_{\vartheta_0}(\mathcal{M}_{m_n}\cdot|H_0)} \right\}
\]

and \( \mathcal{M}_{m_n,1}, \ldots, \mathcal{M}_{m_n,d} \) are iid rv’s distributed according to \( F_{0}^{m_n}(a_{m_n} \cdot + b_{m_n}) \). By assumption, the first term on the right-hand side of (124) is of order \( O(e^{-c \epsilon n}) \). We next show that the second term is of order \( O(1/n^2) \), wherefrom the final result follows.

Let \( s_n^2 = \mathbb{E} \left[ \log^2 f_{m_n}(\cdot|H_0) \right] \) and observe that, under Condition 4.1, Corollary 3.1 in [33] together with Lemma B.1(ii) and equation (B.1) in [39] imply \( \varepsilon_n \to 0 \) as \( n \to \infty \). This fact, together with Condition 4.4 and Theorem 5 in [80] allow to deduce that, for all sufficiently large \( n \geq n_0 \)

\[
\mathbb{E} R_n = \mathcal{N}(f_{m_n}, g_{\vartheta_0}(\cdot|H_0)) \leq \left[ a + 8 \max \left\{ 1, \log \left( \frac{J_0}{\varepsilon_n} \right) \right\} \right] s_n^2 < \frac{\varepsilon}{2}
\]

(125)
where $a$ is a positive global constant. Moreover, arguments analogous to those in the proof of Lemma 4.1 in [33] allow to deduce that

\begin{equation}
\max(c_2, c_3) < 1 + c_4 < \infty, \tag{126}
\end{equation}

where, for $l = 2, 3, 4$,

$$q_l := \sup_{n \geq n_0} (-1)^l \int \log \left( \frac{f_{m_n}(x)}{g_{\Theta_0}(x|H_0)} \right) dx.$$

The bounds in (125)-(126), Markov’s inequality and a few simple manipulations now yield

$$Q_n(E_n \cap \{ R_n > \epsilon \}) \leq Q_n(\{ R_n - \mathbb{E} R_n > \epsilon / 2 \})$$

$$\leq Q_n(\{ |R_n - \mathbb{E} R_n| > \epsilon / 2 \})$$

$$\leq (2/\epsilon)^4 \left[ \frac{1 + c_4}{n^3} \left( 1 + 2\epsilon + 3\epsilon^2 / 2 \right) + \frac{3}{n^2} \right]$$

$$= O(1/n^2).$$

The proof is now complete.

**D.6. Proofs of the results in Section 4.3.**

**D.6.1. Auxiliary results for the proof of Theorem 4.7.**

**Lemma D.22.** Under the assumptions of Theorem 4.7, for any $\epsilon > 0$ eventually almost surely as $n \to \infty$

$$\int_{H \times \Theta} \prod_{i=1}^n \left\{ \frac{g_{\rho, \sigma}(M_{m_n} \cdot |H)}{g_{\theta_0}(M_{m_n} \cdot |H_0)} \right\} d(\Pi_H \times \Theta_n)(H, \rho, \sigma) \geq e^{-nc}.$$

**Proof.** Denote by $\Pi_{sc}^{(d)}$ any pm with positive Lebesgue density $\pi_{sc}^{(d)}$ on $(0, \infty)^d$ satisfying

$$\sup_{x \in (1 \pm \eta)^d} \pi_{sc}^{(d)}(x) < M,$$

where $M$ is defined via

$$M^{1/d} := \inf_{x \in (1 \pm \eta)^d} \pi_{sc}(x)$$

and is positive and bounded by Condition 4.6(ii). Define $\Pi_{\Theta} = \Pi_{sh} \times \Pi_{sc}^{(d)}$, where $\Theta = (0, \infty)^{2d}$. Then, denote $\Pi_{H \times \Theta} = \Pi_H \times \Pi_{sc}^{(d)}$ and observe it satisfies the hypotheses of Corollary D.13. Consequently, for all $\epsilon > 0$ we have that

$$\Pi_{H \times \Theta} \left( \cap_{l=1}^4 \mathcal{V}^{(l)}_{\epsilon} \right) > 0.$$

where $\mathcal{V}^{(l)}_{\epsilon}, l \in \mathbb{N}_+$, is defined as in (100), with $\sigma_0 = 1$. Letting $1 - \eta < a < 1, 1 < b < 1 + \eta$ and

$$\mathcal{V} = \{(H, \rho, \sigma) \in H \times (0, \infty)^{2d} : a < \sigma_j < b, j = 1, \ldots, d;$$

$$\mathcal{X}^{(l)}_{\epsilon} := \left\{ (g_{\rho, \sigma}(\cdot |H_0), g_{\rho, \sigma}(\cdot |H)) < \epsilon, l = 1, \ldots, 4 \right\},$$

we also have $\cap_{l=1}^4 \mathcal{V}^{(l)}_{\epsilon} \subset \mathcal{V}$ for a sufficiently small $\epsilon$, thus $\Pi_{H \times \Theta}(\mathcal{V}) > 0$. Then, define

$$\Pi_{H \times \Theta}(\cdot) := \Pi_{H \times \Theta}(\cdot \cap \mathcal{V}) / \Pi_{H \times \Theta}(\mathcal{V}).$$

Next, observe that by assumption eventually almost surely as $n \to \infty$

$$\frac{b}{1 + \eta} < \tilde{a}_{m_n, j} / a_{m_n, j} < \frac{a}{1 - \eta}, \quad j = 1, \ldots, d,$$
thus, by construction, \( \prod_{j=1}^{d} \pi_{sc}(x_j / \{ \hat{a}_{m_{n,j}}, a_{m_{n,j}} \}) \geq \zeta_{sc}^{(d)}(x) \) for all \( x \in (a, b)^d \) and

\[
\int_{\mathcal{H} \times \Theta} \prod_{i=1}^{n} \left\{ \frac{g_p, \sigma(\mathcal{M}_{m_{n,i}} \mid H)}{g_{\rho_0, 1}(\mathcal{M}_{m_{n,i}} \mid H_0)} \right\} d(\Pi_{\mathcal{H} \times \mathcal{V}})(H, \rho, \sigma)
\]

\[
\geq \int_{\mathcal{H} \times \Theta} \prod_{i=1}^{n} \left\{ \frac{g_p, \sigma(\mathcal{M}_{m_{n,i}} \mid H)}{g_{\rho_0, 1}(\mathcal{M}_{m_{n,i}} \mid H_0)} \right\} d(\Pi_{\mathcal{H} \times \mathcal{V}})(H, \rho, \sigma)
\]

\[
\geq \Pi_{\mathcal{H} \times \Theta}(\mathcal{V}) \exp \{-n \zeta_n\},
\]

where the last line follows from an application of Jensen’s inequality and

\[
I_n := \int_{\mathcal{V}} \frac{1}{n} \prod_{i=1}^{n} \log^+ \left\{ \frac{g_{\rho_0, 1}(\mathcal{M}_{m_{n,i}} \mid H_0)}{g_p, \sigma(\mathcal{M}_{m_{n,i}} \mid H)} \right\} d(\Pi_{\mathcal{H} \times \Theta})(H, \rho, \sigma).
\]

By Fubini’s theorem and Condition 4.4, for all \( n \geq n_0 \),

\[
E(I_n) = \int_{\mathcal{V}} \left[ \int \log^+ \left\{ \frac{g_{\rho_0, 1}(x \mid H_0)}{g_p, \sigma(x \mid H)} \right\} f_{m_n}(x) dx \right] d(P_{\mathcal{H} \times \Theta})(H, \rho, \sigma)
\]

\[
\leq J_0 \int_{\mathcal{V}} \mathcal{X}_+^{(1)}(g_{\rho_0, 1}(\cdot \mid H_0), g_p, \sigma(\cdot \mid H)) d(\Pi_{\mathcal{H} \times \Theta})(H, \rho, \sigma)
\]

\[
< J_0 \epsilon.
\]

Consequently, for all \( n \geq n_0 \), the term on the right hand side of (127) is bounded from below by

\[
e^{-nJ_0 \epsilon} \Pi_{\mathcal{H} \times \Theta}(\mathcal{V}) \exp \{-n \{I_n - E(I_n)\}\}.
\]

Moreover, denoting

\[
\zeta_n := E \left( \log^+ \left\{ \frac{g_{\rho_0, 1}(\mathcal{M}_{m_{n,1}} \mid H_0)}{g_p, \sigma(\mathcal{M}_{m_{n,1}} \mid H)} \right\} \right),
\]

\[
\zeta_n^{(l)} := E \left[ \left( \log^+ \left\{ \frac{g_{\rho_0, 1}(\mathcal{M}_{m_{n,1}} \mid H_0)}{g_p, \sigma(\mathcal{M}_{m_{n,1}} \mid H)} \right\} - \zeta_n \right)^l \right], \quad l \in \mathbb{N}_+,
\]

by Markov’s inequality and equation (6.2) in [7], for all \( \epsilon > 0 \)

\[
Q_n(|I_n - E(I_n)| > \epsilon) \leq \epsilon^4 \left\{ \frac{1}{n^3} \zeta_n^{(4)} + \frac{1}{n^2} \left( \zeta_n^{(2)} \right)^2 \right\},
\]

where, by Jensen’s inequality and Fubini’s theorem, for all \( n \geq n_0 \)

\[
\zeta_n^{(4)} \leq E \int_{\mathcal{Y}} \left[ \log^+ \left\{ \frac{g_{\rho_0, 1}(\mathcal{M}_{m_{n,1}} \mid H_0)}{g_p, \sigma(\mathcal{M}_{m_{n,1}} \mid H)} \right\} - \zeta_n \right]^4 d(\Pi_{\mathcal{H} \times \Theta})(H, \rho, \sigma)
\]

\[
= \int_{\mathcal{Y}} \left[ \log^+ \left\{ \frac{g_{\rho_0, 1}(\mathcal{M}_{m_{n,1}} \mid H_0)}{g_p, \sigma(\mathcal{M}_{m_{n,1}} \mid H)} \right\} - \zeta_n \right]^4 d(\Pi_{\mathcal{H} \times \Theta})(H, \rho, \sigma)
\]

\[
\leq \int_{\mathcal{Y}} \left[ J_0 \mathcal{X}_+^{(4)}(g_{\rho_0, 1}(\cdot \mid H_0), g_p, \sigma(\cdot \mid H)) + J_0^4 \left\{ \mathcal{X}_+^{(1)}(g_{\rho_0, 1}(\cdot \mid H_0), g_p, \sigma(\cdot \mid H)) \right\}^4
\]

\[
+ 6J_0^3 \left\{ \mathcal{X}_+^{(1)}(g_{\rho_0, 1}(\cdot \mid H_0), g_p, \sigma(\cdot \mid H)) \right\}^2 \mathcal{X}_+^{(2)}(g_{\rho_0, 1}(\cdot \mid H_0), g_p, \sigma(\cdot \mid H)) \right] d(\Pi_{\mathcal{H} \times \Theta})(H, \rho, \sigma)
\]

\[
\leq M'
\]

with \( M' \) a positive bounded constant, and

\[
\zeta_n^{(2)} \leq 1 + \zeta_n^{(4)} \leq 1 + M'.
\]
Thus the term on the right-hand side of (130) is of order \(O(1/n^2)\) and, by Borel-Cantelli lemma, we conclude that eventually almost surely

\[
I_n - E(I_n) \leq \varepsilon
\]

for all \(\varepsilon > 0\). As a result, eventually almost surely as \(n \to \infty\), the term in (129) is bounded from below by

\[
e^{-n(\varepsilon + J_0 \varepsilon)} \Pi_{H \times \Theta}(\mathcal{V}).
\]

Since \(\varepsilon\) and \(\varepsilon\) can be selected arbitrarily small, the final result now follows. \(\square\)

**D.6.2. Proof of Theorem 4.7.** Denote by \(\Pi_{uc}^{(d)}\) the pm on \((0, \infty)^d\) whose Lebesgue density equals \(\prod_{j=1}^d \{ u_{uc}(x_j)/u \}, \ x = (x_1, \ldots, x_d) > 0\), where

\[
u = \int_{(0, \infty)} u_{uc}(s) ds,
\]

and define \(\Pi_{\Theta} = \Pi_{uc} \times \Pi_{uc}^{(d)}, \) where \(\Theta = (0, \infty)^{2d}\). Then, denote \(\Pi_{H \times \Theta} = \Pi_H \times \Pi_{\Theta}\) and observe it satisfies the assumptions of Theorem 3.14. Consequently, the induced prior \(\Pi_{G_{\Theta}}\) on \(G_{\Theta} = \{ g_{\rho, \sigma}(\cdot | H) : (H, \rho, \sigma) \in (H \times \Theta) \}\) and the latter observational model jointly satisfy Condition 3.11.

Reasoning as in the proof of Proposition 3.12, we can conclude that there exist \(\varepsilon > 0\) and \(\delta > 0\) such that, for any sequence \(G_{\Theta, n}\) of measurable subsets of

\[
\mathcal{U}_n^c \cap \phi_{H \times \Theta} \left( \mathcal{H} \times B_{\delta, 1}((\rho_0, 1)) \right)
\]

and any collection \(\tau_n = (s_n, t_{n, 1}, \ldots, t_{n, d})\) of measurable functions \(s_n : (0, \infty)^{d n} \to [0, 1], t_{n, j} : (0, \infty)^n \to [0, 1], j = 1, \ldots, d, \) we have

\[
\max \left\{ \bar{\Pi}_{n}(\mathcal{U}_n^c), \Pi_n((\mathcal{U}_{1} \times \mathcal{U}_2)^c) \right\} \leq \Xi_{n, 1}(M_{m, 1:n, \tau_n}) + \Xi_{n, 2}(M_{m, 1:n, \tau_n, \Pi_H \times \Psi_n})
\]

where \(\mathcal{U}_n = \{ g \in G_{\Theta} : \mathcal{D}_H(g, g_{\rho_0, 1}(\cdot | H_0)) \leq 4\epsilon \}, \ \Psi_n = \Psi_n \circ \psi_n^{-1}, \) with \(\psi_n\) as in the first line of (20).

\[
R_n(M_{m, 1:n, \tau_n, \Pi_H \times \Psi_n}) = \int_{H \times \Theta} \prod_{i=1}^n \left( g_{\rho, \sigma}(M_{m, 1:n, \tau_n, \Pi_H \times \Psi_n}(H_0) \right) d(\Pi_{H \times \Psi_n})(H, \rho, \sigma),
\]

\(M_{m, 1:n} = (M_{m, i, j})_{i=1}^n\) and the functionals \(\Xi_{n, i}, i = 1, 2, \) are defined as in (75) - (76) (see Remark D.9 for further details). By assumption, eventually almost surely as \(n \to \infty\)

\[
1 - \eta < \tilde{\sigma}_{m, j}/\alpha_{m, j} < 1 + \eta, \quad j = 1, \ldots, d,
\]

thus Condition 4.6(iib) guarantees that \(\tau_{uc}(x/\tilde{\sigma}_{n, j}/\alpha_{n, j}) \leq \tau_{uc}(x), \) for all \(x > 0,\) and

\[
\Xi_{n, 2}(M_{m, 1:n, \tau_n, \Pi_H \times \Psi_n}) \leq \frac{d}{1 - \eta} \Xi_{n, 2}(M_{m, 1:n, \tau_n, \Pi_H \times \Theta}).
\]

Moreover, by Lemma D.22, for any \(c > 0\), eventually almost surely as \(n \to \infty\), \(R_n(M_{m, 1:n, \tau_n, \Pi_H \times \Psi_n}) \geq e^{-nc}.\) Therefore, there exists a constant \(M > 0\) such that eventually almost surely as \(n \to \infty\) the term on the right-hand side of (131) is bounded from above by

\[
Me^{cn} \Xi_n(M_{m, 1:n, \tau_n, \Pi_H \times \Theta}),
\]

where \(\Xi_n\) is defined as in (74). Without loss of generality, we can assume that \(\delta < \delta^*\), with \(\delta^*\) as in Condition 3.11, and select \(G_{\Theta, n}\) and \(\tau_n\) satisfying the properties therein. Hence, selecting \(c\) as in (81), the exponential tail bound in (82) is satisfied for a \(n\)-dimensional iid sample from \(G_{\rho_0, 1}(\cdot | H_0)\). Applying Proposition 4.5 we can thus conclude that for all \(\varepsilon > 0\) as \(n \to \infty\)

\[
Q_n \left( e^{2cn} \Xi_n(M_{m, 1:n, \tau_n, \Pi_H \times \Theta}) \geq \varepsilon \right) \leq n^{-1-c'},
\]

for some \(c' \in (0, 1)\). The first part of the statement now follows from Borel-Cantelli lemma, by adapting arguments in the proof of Corollary D.8. By the definitions of \(\Pi_n, \Pi_{\Theta, n}\) and \(\Pi_{\Theta, n}^{(o)}\), the results at points \((d')-(b')\) are immediate consequences.

**D.7. Proofs of the results in Section 4.5.**
D.7.1. Auxiliary results for the proof of Theorem 4.19.

**Lemma D.23.** Under the assumptions of Theorem 4.19, for any \( \epsilon \in (0, 1) \) there exists a constant \( k \equiv k(\epsilon) > 0 \) such that

\[
\lim_{n \to \infty} \inf F_{i(n,m)}^{0} \left( \frac{\sqrt{\text{det}(\mathcal{M}_{n,m}^{i}))}}{g_{0,1} \mathcal{M}_{n,m}^{i}(H)} \right) \geq \epsilon \left( 1 - nJ_0 \sqrt{\epsilon} \right)
\]

where \( J_0 \) is as in Condition 4.4.

**Proof.** We follow the main lines of the proof of Lemma D.22, with a few adaptations. Let \( \Pi_{\text{sc}}^{(d)} \) and \( \Pi_{\text{loc}}^{(d)} \) be pm with positive Lebesgue density \( \pi^{(d)}_{\text{sc}} \) on \( \mathcal{I}_{\text{sc}}^{d} \) and \( \pi^{(d)}_{\text{loc}} \) on \( \mathcal{I}_{\text{loc}}^{d} \), respectively, and satisfying

\[
\sup_{x \in (1 \pm \eta)} \pi^{(d)}_{\text{sc}}(x) < M_{\text{sc}}, \quad \sup_{x \in (-\eta, \eta)} \pi^{(d)}_{\text{loc}}(x) < M_{\text{loc}}.
\]

where \( M_{\text{sc}} \) and \( M_{\text{loc}} \) are defined via

\[
M_{\text{sc}}^{1/d} := \inf_{x \in (1 \pm \eta)} \pi_{\text{sc}}(x), \quad M_{\text{loc}}^{1/d} := \inf_{x \in (-\eta, \eta)} \pi_{\text{loc}}(x),
\]

and are positive and bounded by Conditions 4.6(ii) and 4.14(ii). Define

\[
\Pi_{\Theta} = \Pi_{\text{sh}} \times \Pi_{\text{sc}}^{(d)} \times \Pi_{\text{loc}}^{(d)}.
\]

By construction, \( \Pi_{\Theta} \) satisfies Condition 3.8(ic). Then, denote \( \Pi_{\mathcal{H} \times \Theta} = \Pi_{\mathcal{H}} \times \Pi_{\Theta} \) and observe that the other assumptions of Corollary D.19 are satisfied by hypothesis. We can thus conclude that for all \( \epsilon > 0 \)

\[
\Pi_{\mathcal{H} \times \Theta}(\mathcal{V}^{(1)}_{\epsilon}) > 0,
\]

where \( \mathcal{V}^{(1)}_{\epsilon} \) is defined as in (113), with \( \sigma_0 = 1 \) and \( \mu_0 = 0 \).

Next, observe that by Condition 4.18(ii) there exist neighborhoods \( \mathcal{U}^{1} \subset I_{\text{sc}}^{d} \) of \( 1 \) and \( \mathcal{U}^{0} \subset I_{\text{loc}}^{d} \) of \( 0 \) such that, with \( F_{0}^{(n,m)} \)-probability tending to 1, as \( n \to \infty \)

\[
\prod_{j=1}^{d} \pi_{\text{sc}}(x_{j}/\{\tilde{a}_{m,n,j}, \hat{a}_{m,n,j}\}) > \pi^{(d)}_{\text{sc}}(x), \quad \forall x \in \mathcal{U}^{1},
\]

\[
\prod_{j=1}^{d} \pi_{\text{loc}}(x_{j}/\{\hat{b}_{m,n,j}, \hat{a}_{m,n,j}\}) > \pi^{(d)}_{\text{loc}}(x), \quad \forall x \in \mathcal{U}^{0}.
\]

For a sufficiently small \( \epsilon \), defining

\[
\mathcal{V} = \{(H, \omega, \sigma, \mu) \in \mathcal{H} \times \Theta : (\omega, \sigma, \mu) \in K_{\text{sh}} \times \mathcal{U}^{1} \times \mathcal{U}^{0},
\]

we also have \( \mathcal{V}^{(1)}_{\epsilon} \subset \mathcal{V} \) and thus \( \Pi_{\mathcal{H} \times \Theta}(\mathcal{V}) > 0 \). Accordingly, let \( \Pi_{\mathcal{H} \times \Theta}(\cdot) := \Pi_{\mathcal{H} \times \Theta}(\cdot \cap \mathcal{V})/\Pi_{\mathcal{H} \times \Theta}(\mathcal{V}) \) and observe that, analogously to (127),

\[
\int_{\mathcal{H} \times \Theta} \prod_{i=1}^{n} \frac{g_{\omega, \sigma, \mu}(\mathcal{M}_{n,m}^{i}(H))}{g_{0,1,0}(\mathcal{M}_{n,m}^{i}(H))} d(\Pi_{\mathcal{H} \times \Theta}(H, \omega, \sigma, \mu) \geq \Pi_{\mathcal{H} \times \Theta}(\mathcal{V}) \exp \{-nI_n\},
\]

where

\[
I_n := \int_{\mathcal{V}} \frac{1}{n} \sum_{i=1}^{n} \log \left\{ \frac{g_{\omega, \sigma, \mu}(\mathcal{M}_{n,m}^{i}(H))}{g_{0,1,0}(\mathcal{M}_{n,m}^{i}(H))} \right\} d\Pi_{\mathcal{H} \times \Theta}(H, \omega, \sigma, \mu).
\]

Using Condition 4.4 and reasoning as in (128), we obtain once more the inequality \( \mathbb{E}(I_n) \leq J_0 \epsilon \), for all \( n \geq n_0 \). Hence, by Markov’s inequality, on a set of \( Q_n \)-probability larger than \( 1 - \sqrt{\epsilon} \), the term on the right-hand side of (133) is bounded from below by \( \Pi_{\mathcal{H} \times \Theta}(\mathcal{V}) \exp \{-nJ_0 \sqrt{\epsilon}\} \), for all \( n \geq n_0 \). Since \( \epsilon \) can be chosen arbitrarily small, the result now follows. \( \square \)
D.7.2. **Proof of Theorem 4.19.** The proof follows the main lines of the proof of Theorem 4.7 (Section D.6.2 of this manuscript), with a few adaptations. We only sketch the main changes.

By Conditions 4.18(i–iii), the data-dependent prior \( \Psi_n \) obtained after the reparametrisation in the second line of (20) have marginal densities on scale and location components satisfying
\[
\left\{ x \in \mathbb{R} : \pi_{\text{sc}}(x_j / \{ \tilde{a}_{mn,j} / a_{mn,j} \}) \right\} \left\{ \tilde{a}_{mn,j} / a_{mn,j} \right\}^{-1} > 0 \subset K_{\text{sc}}
\]
\[
\left\{ x \in \mathbb{R} : \pi_{\text{loc}} \left( x_j - \frac{\tilde{b}_{mn,j} - b_{mn,j}}{a_{mn,j}} \right) \right\} \left\{ \tilde{a}_{mn,j} / a_{mn,j} \right\}^{-1} > 0 \subset K_{\text{loc}}
\]
with \( F_0^{(\text{nom})} \)-probability tending to 1 as \( n \to \infty \), for a compact subinterval of \((0, \infty)\), \( K_{\text{sc}} \), having 1 as an interior point, and a compact interval in \( \mathbb{R} \), \( K_{\text{loc}} \), having 0 as an interior point. Then, by Conditions 4.6(ii–b) and 4.14(ii) there exist positive constants \( c_{\text{sc}} \) and \( c_{\text{loc}} \) such that
\[
\pi_{\text{sc}}(x_j / \{ \tilde{a}_{mn,j} / a_{mn,j} \}) \left\{ \tilde{a}_{mn,j} / a_{mn,j} \right\}^{-1} \leq c_{\text{sc}} \pi_{\text{sc}}(x), \quad \forall x \in K_{\text{sc}}
\]
\[
\pi_{\text{loc}} \left( x_j - \frac{\tilde{b}_{mn,j} - b_{mn,j}}{a_{mn,j}} \right) \left\{ \tilde{a}_{mn,j} / a_{mn,j} \right\}^{-1} \leq c_{\text{loc}} \pi_{\text{loc}}(x), \quad \forall x \in K_{\text{loc}}
\]
with \( F_0^{(\text{nom})} \)-probability tending to 1 as \( n \to \infty \), where
\[
\pi_{\text{sc}}(x) := u_{\text{sc}}(x) / \int_{K_{\text{sc}}} u_{\text{sc}}(s) \, ds, \quad x \in K_{\text{sc}}
\]
and “•” stands either for “sc” or “loc”. Denote by \( \Pi^{(d)}_{\Theta} \) the pm on \( K_{\text{sc}}^d \) whose Lebesgue density equals
\[
\prod_{j=1}^d \pi_{\text{sc}}(x_j)
\]
and define \( \Pi_{\Theta} = \Pi_{\text{sh}} \times \Pi_{\text{sc}}^{(d)} \times \Pi_{\text{loc}}^{(d)} \), where \( \Theta = (0, \infty)^{2d} \times \mathbb{R}^d \). Notice that, by construction, \( \Pi_{\Theta} \) satisfies Condition 3.8(ie) and, by assumption, \( \Pi_{\mathcal{H}} \) satisfies Condition 3.13(iii). Therefore, the arguments in the proof of Theorem 3.15 apply to \( \Pi_{\mathcal{H} \times \Theta} := \Pi_{\mathcal{H}} \times \Pi_{\Theta} \). Next, observe that it is possible to establish an inequality analogous to (131), with
\[
R_n(\overline{M}_{m_0,1:n}, \Pi_{\mathcal{H}} \times \Psi_n) := \int_{\mathcal{H} \times \Theta} \left\{ \frac{g_{\omega, \sigma, \mu}(\overline{M}_{m_0,1:n}[H])}{g_{\omega_0, 1, 0}(\overline{M}_{m_0,1:n}[H_0])} \right\} d(\Pi_{\mathcal{H}} \times \Psi_n)(H, \omega, \sigma, \mu).
\]
Lemma D.23 and the facts listed here above entail that with probability at least \( 1 - 2\sqrt{3} \) as \( n \to \infty \)
\[
(134) \quad \max \left\{ \widehat{\Pi}_{\text{sh}}(\overline{U}_{\Theta}), \Pi_{\text{sc}}(\{ \overline{U}_1 \times U_2 \}) \right\} \leq Me^{n J_0 \sqrt{\varepsilon}} \Xi_n(\overline{M}_{m_0,1:n}, \tau_n, \Pi_{\mathcal{H} \times \Theta}),
\]
where \( M \equiv M(\varepsilon) \) is a positive constant, \( J_0 \) is as in Condition 4.4 and \( \varepsilon \in (0, 1/4) \), while \( \Xi_n(\cdot, \tau_n, \Pi_{\mathcal{H} \times \Theta}) \) complies with (133). Therefore, for \( \varepsilon \) sufficiently small, Proposition 4.5 guarantees that the term on the right hand-side of (134) converges to zero in probability, allowing to conclude that
\[
\limsup_{n \to \infty} F_{0}^{(\text{nom})} \left( \frac{\widehat{\Pi}_{\text{sh}}(\overline{U}_{\Theta}), \Pi_{\text{sc}}(\{ \overline{U}_1 \times U_2 \})}{\varepsilon} \right) < 3\sqrt{3}, \quad \forall \varepsilon > 0.
\]
Since \( \varepsilon \) can be chosen arbitrarily small, the result follows.

D.8. **Proofs of the results in Section 5.**

D.8.1. **Proof of Corollary 5.1.** Under the assumptions of Theorem 3.14 or 3.15 or 3.17, the true max-stable data generating density \( g_{\vartheta_0}(\cdot | H_0) \) is of the form \( g_{\omega_0, \sigma_0}(\cdot | H_0) \) or \( g_{\omega_0, \sigma_0, \mu_0}(\cdot | H_0) \) or \( g_{\sigma_0, \mu_0}(\cdot | H_0) \), respectively. Accordingly, the set \( G_{\Theta} \) is \( \{ g_{\vartheta, \sigma}(\cdot | H) : H \in \mathcal{H}, (\vartheta, \sigma) \in (0, \infty)^{d} \times (1, \infty)^{d} \times \mathbb{R}^{d} \} \) or \( \{ g_{\omega, \sigma, \mu}(\cdot | H) : H \in \mathcal{H}, (\vartheta, \sigma, \mu) \in (0, \infty)^{d} \times (1, \infty)^{d} \times \mathbb{R}^{d} \} \) or \( \{ g_{\sigma, \mu}(\cdot | H) : H \in \mathcal{H}, (\vartheta, \sigma, \mu) \in (0, \infty)^{d} \times \mathbb{R}^{d} \} \), respectively. In the three cases, the map
\[
g \mapsto g_{\vartheta_0}(g, g_{\vartheta_0}(\cdot | H_0))
\]
from \( G_{\Theta} \) to \((0, \infty)^{d}) \) is convex. Thus, the result follows directly from that at point (a) of Theorem 3.14 or 3.15 or 3.17, respectively, together with arguments in the proof of Theorem 6.8 in [39].
D.8.2. Proof of Corollary 5.2. Preliminarily, recall that by point \((a')\) of Theorems 4.7, 4.15 and 4.19, for all \(\epsilon > 0\), as \(n \to \infty\) the quasi-posterior satisfies

\[
\tilde{\Pi}_n^{(\epsilon)} \left( \left\{ g \in \mathcal{G} : \mathcal{D}_H \left( g \cdot g_{\theta_0} \left( \cdot - b_{m_n} \right) /a_{m_n} | H_0 \right) \prod_{j=1}^{d} a_{m_n,j}^{-1} \right\} > \epsilon \right) \to 0
\]

almost surely (Theorems 4.7 and 4.15) or in probability (Theorem 4.19) as \(n \to \infty\). Here above, the density \(g_{\theta_0} \left( \cdot - b_{m_n} \right) /a_{m_n} | H_0 \right) \prod_{j=1}^{d} a_{m_n,j}^{-1}\) equals \(g_{\phi_0} a_{m_n} | H_0 \right) \prod_{j=1}^{d} a_{m_n,j}^{-1}\) or \(g_{\phi_0} a_{m_n} b_{m_n} | H_0 \right) \prod_{j=1}^{d} a_{m_n,j}^{-1}\) under the assumptions of Theorems 4.7 or 4.15 or 4.19, respectively. In each of the three cases, the norming sequences \(a_{m_n}\) and \(b_{m_n}\) are as obtained from the pairs in (18). Thus, reasoning as in the proof of Corollary 5.1 and adapting arguments from the proof of Theorem 6.8 in [39], we conclude that as \(n \to \infty\)

\[
\mathcal{D}_H \left( g_{\theta_n} \cdot g_{\theta_0} \left( \cdot - b_{m_n} \right) /a_{m_n} | H_0 \right) \prod_{j=1}^{d} a_{m_n,j}^{-1} \to 0
\]

almost surely under the assumptions of Theorems 4.7 and 4.15, or in probability under the assumptions of 4.19. On the other hand, Condition 4.1 guarantees that as \(n \to \infty\)

\[
\mathcal{D}_H \left( f_{m_n} \cdot g_{\theta_0} | H_0 \right) \prod_{j=1}^{d} a_{m_n,j}^{-1} \to 0,
\]

where \(f_{m_n} = f_{m_n}^{(\epsilon)} \left( a_{m_n} \cdot b_{m_n} \right) \prod_{j=1}^{d} a_{m_n,j}^{-1}\). The result now follows by triangular inequality.

APPENDIX E: EXAMPLES AND TECHNICALITIES

E.1. Examples of Section 4.3.

E.1.1. Technical derivations for Example 4.10. Herein, we prove the inequality in Example 4.10. Recall that \(Z_{1,1}, \ldots, Z_{nm_n}\) are iid rv’s with absolutely continuous distributions \(F_0\) and copula \(C_0\), thus we have the representation

\[
Z_{i,j} = F_{0,j}^{-1} \left( U_{i,j} \right), \quad j = 1, \ldots, d, \quad i = 1, \ldots nm_n,
\]

where \(U_i = (U_{i,1}, \ldots, U_{i,d})\) are iid rv’s with distribution \(C_0\). In the sequel, we also denote

\[
\check{U}_{i,j} = 1 - U_{i,j}, \quad j = 1, \ldots, d, \quad i = 1, \ldots nm_n.
\]

Moreover, denoting by \(U_{k;nm_n,j}\) the \(k\)-th order statistic of the marginal sample \(U_{1,j}, \ldots, U_{nm_n,j}\), observe that for \(j = 1, \ldots, d\),

\[
\hat{a}_{m_n,j} := F_{nm_n,j}^{-1} \left( 1 - 1/m_n \right) = F_{0,j}^{-1} \left( U_{n(m_n-1);nm_n,j} \right).
\]

We next show that, for all \(\epsilon > 0\), as \(n \to \infty\), the estimator \(\hat{a}_{m_n,j}\) of \(F_{0,j}^{-1} \left( 1 - 1/m_n \right)\) satisfies

\[
F_{0}^{(\text{nom})} \left( |\hat{a}_{m_n,j}/a_{m_n,j} - 1| > \epsilon \right) \leq 4\epsilon^{-1} \sqrt{n+1},
\]

where \(\tau_j \equiv \tau_j(\epsilon)\) is a positive constant.

Preliminary observe that

\[
F_{0}^{(\text{nom})} \left( |\hat{a}_{m_n,j}/a_{m_n,j} - 1| > \epsilon \right) = F_{0}^{(\text{nom})} \left( |\hat{a}_{m_n,j}/a_{m_n,j} - 1| < \epsilon \right) + F_{0}^{(\text{nom})} \left( |\hat{a}_{m_n,j}/a_{m_n,j} - 1| < \epsilon \right)
\]

\[
= T_{n,j}^{(1)} + T_{n,j}^{(2)}.
\]

As for \(T_{n,j}^{(1)}\), since \(F_{0,j} \in \mathcal{D}(G_{\rho_0,j})\), as \(n \to \infty\) we have that

\[
T_{n,j}^{(1)} = C \left( 1 - U_{n(m_n-1);nm_n,j} < 1 - F_{0,j} (a_{m_n,j} (1 + \epsilon)) \right) \leq C \left( U_{n+1;nm_n,j} < m_n \{ 1 - F_{0,j} (a_{m_n,j} (1 + \epsilon)) \} / m_n \right)
\]

\[
\leq C \left( U_{n+1;nm_n,j} < (1 + \epsilon)' (1 + \epsilon)^{-\rho_0,j / m_n} \right).
\]
for an arbitrarily small $\epsilon' > 0$ such that
$$v_j := 1 - (1 + \epsilon')(1 + \epsilon)^{-\rho_{0,j}} > 0.$$ 
Thus, the term on the right-hand side of the above inequality is bounded from above by
$$C_0^{\text{norm}}(\bar{U}_{n+1}, m_{n,j}, \bar{C} < (1 + \epsilon')(1 + \epsilon)^{-\rho_{0,j}} (n + 1)/(nm_n))
= C_0^{\text{norm}}(\bar{U}_{n+1}, m_{n,j} - (n + 1)/(nm_n) < -v_j(n + 1)/(nm_n))
\leq C_0^{\text{norm}}(\bar{U}_{n+1}, m_{n,j} - (n + 1)/(nm_n) > v_j(n + 1)/(nm_n)).$$
By Fact 3 in [18] we now conclude that, as $n \to \infty$,
$$\tau_{n,j}^{(1)} \leq 2\exp(-\tau_{j,1} \sqrt{n + 1}),$$
where $\tau_{j,1} = v_j/10$. By a similar reasoning, for some $\tau_{j,2} \equiv \tau_{j,2}(\epsilon) > 0$, as $n \to \infty$,
$$\tau_{n,j}^{(2)} \leq 2\exp(-\tau_{j,2} \sqrt{n + 1}).$$
The inequality in (135) now follows by setting $\tau_j = \min(\tau_{j,1}, \tau_{j,2}).$

### E.1.2. Technical derivations for Example 4.12

Note that, for $x \in (1, \infty)^2$, $F_0(x)$ allows the representation
$$F_0(x) = C_0\left(1 - 1/x_1^{\rho_{0,1}}, 1 - 1/x_2^{\rho_{0,2}}\right),$$
where
$$C_0(u) = 1 - (1 - u_1) - (1 - u_2) + \left(\frac{1}{1 - u_1} + \frac{1}{1 - u_2}\right)^{-1}.\tag{136}$$
As established in [33, Example 2.2], $C_0$ satisfies Condition 4.1(i). The verification of Condition 4.1(ii) is immediate for Pareto margins $F_{0,j}(x_j) = 1 - 1/x_j^{\rho_{0,j}}$, $x_j > 1$, $j = 1, 2$. Moreover, it is already known [e.g. 65, Example 5.16 and p. 289] that $F_0 \in \mathcal{D}(G_{\rho_{0},1}(\cdot|H_0))$. We recall that herein, for $x \in (0, \infty)^2$,
$$G_{\rho_{0},1}(x|H_0) = \exp\left\{-x_1^{-\rho_{0,1}} - x_2^{-\rho_{0,2}} + \left(x_1^{\rho_{0,1}} + x_2^{\rho_{0,2}}\right)^{-1}\right\},$$
where $x^{\rho} = (x_1^{\rho_{0,1}}, x_2^{\rho_{0,2}})$ and $V(\cdot) \equiv V(\cdot|H_0)$. A valid choice of the norming sequences, asymptotically equivalent to that in (18), is
$$a_{mn} = \left((mn - 1)^{1/\rho_{0,1}}, (mn - 1)^{1/\rho_{0,2}}\right), \quad b_{mn} = 0.$$ 
Thus, the cdf associated to the rv’s $M_{mn,i}/a_{mn}$ is given by
$$F_{0,mn}(a_{mn}, y) = \left\{1 - \frac{V(mn)}{m_n - 1}\right\}^{m_n} \tag{137}$$
where $y \in x_{j=1}^2(1/(mn - 1)^{1/\rho_{0,j}}, \infty)$ and, for all $z \in x_{j=1}^2(1/(mn - 1), \infty)$
$$V(mn)(z) = \frac{1}{z_1} + \frac{1}{z_2} + \left(z_1 + z_2 - \frac{1}{mn - 1}\right)^{-1}.\tag{138}$$
In the present example, for $y \in x_{j=1}^2(1/(mn - 1)^{1/\rho_{0,j}}, \infty)$,
$$f_{mn}(y) = \frac{m_n}{m_{n-1}} \left\{1 - \frac{V(mn)}{m_n - 1}\right\}^{m_n-1} \tag{139}$$
\[\times \left\{V(mn)(y^{\rho_0})V(mn)(y^{\rho_0}) - V(mn)(y^{\rho_0})\right\}^2 \prod_{j=1}^2 \rho_{0,j}^{y^{\rho_{0,j}-1}},\]
with
\[-V_{\{m\}}^{(m_n)}(z) = \frac{1}{z_j^2} \left\{ 1 - \frac{z_j^2}{(z_1 + z_2 - 1/(m\_n - 1))^2} \right\}, \quad j = 1, 2,\]
\[-V_{\{1,2\}}^{(m_n)}(z) = 2z_1z_2 - 1/(m\_n - 1)^{-3},\]
for \(z \in \mathbb{R}_+^{2} \times (1/(m\_n - 1), \infty)\), while, for \(y \in (0, \infty)^2\),
\[(141) \quad g_{\rho_0,1}(y|H_0) = \exp \left\{ -V(x^{(\rho_0)}) \right\} \left\{ V_1(y^{(\rho_0)})V_2(y^{(\rho_0)}) - V_{\{1,2\}}^{(\rho_0)}(y^{(\rho_0)}) \right\} \prod_{j=1}^2 \rho_{0,j} y_j^{\rho_{0,j} - 1},\]
with
\[(142) \quad V_{\{j\}}(y) = \frac{1}{y_j^2} \left\{ 1 - \frac{y_j^2}{(y_1 + y_2)^2} \right\}, \quad j = 1, 2, \quad V_{\{1,2\}}(y) = 2(y_1 + y_2)^{-3}.\]
Therefore, for \(y \in \mathbb{R}_+^{2} \times (1/(m\_n - 1)^{1/\rho_{0,j}}, \infty)\)
\[(143) \quad \frac{f_{m_n}(y)}{g_{\rho_0,1}(y|H_0)} = \frac{m_n - 1}{m_n - 1} \left\{ 1 - \frac{V_{\{m\}}^{(m_n)}(y^{(\rho_0)})}{m_n - 1} \right\}^{m_n - 1} \exp \left\{ V(y^{(\rho_0)}) \right\}
\times \left\{ \frac{V_1(y^{(\rho_0)})V_2(y^{(\rho_0)}) - V_{\{1,2\}}^{(\rho_0)}(y^{(\rho_0)})}{V_1(y^{(\rho_0)})V_2(y^{(\rho_0)})} \frac{V_{\{1\}}^{(\rho_0)}(y^{(\rho_0)})V_{\{2\}}^{(\rho_0)}(y^{(\rho_0)})}{V_{\{1\}}^{(\rho_0)}(y^{(\rho_0)})V_{\{2\}}^{(\rho_0)}(y^{(\rho_0)})} \right\}
= \frac{m_n - 1}{m_n - 1} I_n^{(1)}(y)I_n^{(2)}(y) + I_n^{(3)}(y).\]
Since for \(y_j > 1/(m\_n - 1)^{1/\rho_{0,j}}\) we have \(u_j := 1/(y_j^{\rho_{0,j}}(m\_n - 1)) \leq 1, j = 1, 2,\) then
\[I_n^{(1)}(y) = \left\{ 1 - \frac{V_{\{m\}}^{(m_n)}(y^{(\rho_0)})}{m_n - 1} \right\}^{m_n - 1} \exp \left\{ \frac{1}{m_n - 1} V(y^{(\rho_0)}) \right\}
\times \left\{ 1 + \left( \frac{1}{u_1} \right)^{-1} \right\} \left\{ 1 + \left( \frac{1}{u_2} \right)^{-1} \right\} \exp \left\{ u_1 + u_2 - \left( \frac{1}{u_1} + \frac{1}{u_2} \right)^{-1} \right\} \leq 1.\]
Moreover, for \(j = 1, 2,\) we have
\[-V_{\{j\}}^{(\rho_0)}(y^{(\rho_0)}) \quad \text{and} \quad -V_{\{1,2\}}^{(\rho_0)}(y^{(\rho_0)}) \leq 4\]
and
\[I_n^{(2)}(y) + I_n^{(3)}(y) \leq 16\] and we deduce that for any \(\tau > 0\) as \(n \rightarrow \infty\)
\[(144) \quad \|f_{m_n}/g_{\rho_0,1}(:|H_0)||_{\infty} \leq (1 + \tau)16,\]
wherefrom the final result follows.

E.1.3. Technical derivations for Example 4.13. Note that also in this case, for \( x \in (1, \infty)^2 \), 
\( F_0(x) \) allows the representation 
\[
F_0(x) = C_0 \left( 1 - 1/x_1^{\rho_0,1}, 1 - 1/x_2^{\rho_0,2} \right),
\]
where \( C_0 \) is now the Joe-B5 copula with dependence parameter 3, i.e., for \( u \in [0, 1]^2 \).
\[
C_0(u) = 1 - \left[ (1 - (1 - u_1)^3)(1 - (1 - u_2)^3) \right]^{1/3}.
\]
On the other hand, for all \( x \in (0, \infty)^2 \), in this example \( G_{\rho_0,1}(x|H_0) \) allows the representation 
\[
G_{\rho_0,1}(x|H_0) = C_{EV} \left( e^{1/x_1\rho_0,1}, e^{1/x_1\rho_0,1}; x_1 \right),
\]
where \( C_{EV}(x|H_0) \) is the logistic extreme-value copula with dependence parameter 3 
\[
C_{EV}(u|H_0) = \exp \left\{ -\left( -\log u_1 \right)^3 + (-\log u_3)^3 \right\}^{1/3}
\]
\[
=: \exp \left\{ -L(-\log u_1, -\log u_2) \right\},
\]
with \( L(\cdot) \equiv L(\cdot|H_0) \). Standard multivariate calculus allows to show that 
\[
\lim_{m_n \to \infty} C_{m_n,0}(1/\log u) = C_{EV}(u|H_0), \quad \forall u \in [0, 1],
\]
and that \( C_0 \) and \( L(\cdot|H_0) \) satisfy Condition 4.1(i). In particular, the above equation and the fact that \( F_{0,j} \in \mathcal{D}(G_{\rho_0,j}), j = 1, \ldots, d \), allow to conclude that \( F_0 \in \mathcal{D}(G_{\rho_0,1}(1/H_0)) \). Also observe that, since \( F_{0,j} \) are one-parameter Pareto, Condition 4.1(ii) is still satisfied. To establish the property in Condition 4.4, it is possible to follow the lines of Section E.1.2. Thus, we herein highlight only the main changes.

Observe that, for all \( y \in \times_{j=1}^3 (1/(m_n - 1)^{1/\rho_0,j}, \infty) \), \( F_{m_n}^0(a_{m_n}, y) \) and its density \( f_{m_n}(y) \) are still of the form in (138) and (140), with 
\[
V^{(m_n)}(z) = \left\{ \frac{1}{z_1^3} + \frac{1}{z_2^3} - \left( \frac{1}{m_n - 1} \frac{1}{z_1z_2} \right)^3 \right\}^{1/3}
\]
for all \( z \in \times_{j=1}^3 (1/(m_n - 1), \infty) \), whose derivatives are
\[
-V^{(m_n)}_{(j)}(z) = \left\{ \frac{1}{z_1^3} + \frac{1}{z_2^3} - \left( \frac{1}{m_n - 1} \frac{1}{z_1z_2} \right)^3 \right\}^{-2/3} \frac{1}{z_j^3} \left\{ 1 - \left( \frac{1}{m_n - 1} \frac{1}{z_{-j}} \right)^3 \right\}, \quad j = 1, 2,
\]
\[
-V^{(m_n)}_{(1,2)}(z) = 2 \left\{ \frac{1}{z_1^3} + \frac{1}{z_2^3} - \left( \frac{1}{m_n - 1} \frac{1}{z_1z_2} \right)^3 \right\}^{-5/3} \prod_{j=1}^2 \frac{1}{z_j^3} \left\{ 1 - \left( \frac{1}{m_n - 1} \frac{1}{z_{-j}} \right)^3 \right\}
\]
\[
+ 3 \left\{ \frac{1}{z_1^3} + \frac{1}{z_2^3} - \left( \frac{1}{m_n - 1} \frac{1}{z_1z_2} \right)^3 \right\}^{-2/3} \prod_{j=1}^2 \frac{1}{z_j^3},
\]
with \( z_{-j} = z_2 \) if \( j = 1 \) and \( z_{-j} = z_1 \) if \( j = 2 \). Whereas, for \( y \in (0, \infty)^2 \), \( g_{\rho_0,1}(y|H_0) \) and \( G_{\rho_0,1}(y|H_0) \) are still of the form in (141) and in the second line of (137), respectively, but now \( V(y) \equiv V(y|H_0) = (1/y_1^3 + 1/y_2^3)^{1/3} \) and 
\[
-V_{(j)}(y) = \frac{1}{y_j^3} \left( \frac{1}{y_1^3} + \frac{1}{y_2^3} \right)^{-2/3}, \quad j = 1, 2,
\]
\[
-V_{(1,2)}(y) = 2 \frac{1}{y_1^3} \frac{1}{y_2^3} \left( \frac{1}{y_1^3} + \frac{1}{y_2^3} \right)^{-5/3}.
\]
Notice that the equality in (143) is still valid. Since for \( y_j > 1/(mn - 1)^{1/3} \) we have \( w_j := 1/(y_j^{1/3}(mn - 1)) \leq 1, j = 1, 2, \) then

\[
s(u_1, u_2) := \left\{ 1 - (u_1 + u_2 - u_1 u_2)^{1/3} \right\} \exp\left\{ (u_1 + u_2)^{1/3} \right\} \leq 1,
\]

from which we deduce that

\[
\left\{ 1 - \frac{V(m_n)(y^{p_0})}{mn - 1} \right\}^{m_n-1} \exp\left\{ V(y^{p_0}) \right\} = \{s(u_1, u_2)\}^{m_n-1} \leq 1.
\]

Moreover, we have that

\[ 1 \leq q(u_1, u_2) := \frac{u_1 + u_2}{u_1 + u_2 - u_1 u_2} \leq 2, \]

which implies that

\[
\frac{V(m_n)(y^{p_0})V(m_n)(y^{p_0})}{V(1)(y^{p_0})V(2)(y^{p_0})} = \{q(u_1, u_2)\}^{4/3} \prod_{j=1}^{2} (1 - u_j)^2 \leq 2^{4/3}
\]

and

\[
-\frac{V\{1\}2(y^{p_0})}{V\{1\}2(y^{p_0})} = \{q(u_1, u_2)\}^{5/3} \prod_{j=1}^{2} (1 - u_j) + 3\{q(u_1, u_2)\}^{5/3}(u_1 + u_2 - u_1 u_2) \leq 2^{10/3}.
\]

Consequently, for all \( \tau > 0 \) and \( n \) sufficiently large,

\[
\|f_{m_n} / g_{p_0, 1} (\cdot | H_0)\|_\infty \leq (1 + \tau)2^{10/3},
\]

whence the final result.

**E.2. Examples of Section 4.4.**

E.2.1. **Technical derivations for Example 4.16.** We provide a proof of the claim in Example 4.16. Without loss of generality, assume \( \epsilon < 1 \). Observe that, for \( j = 1, \ldots, d, \)

\[
F_0^{(\text{rms})}(\{\hat{b}_{m,j} - b_{m,j} / m_{m,j} > \epsilon\}) = C_0^{(\text{rms})}(1 - U_{n(m_n - 1); mn_m,j} < 1 - F_0,j(a_{m,j} + b_{m,j}))
\]

\[
+ C_0^{(\text{rms})}(1 - U_{n(m_n - 1); mn_m,j} > 1 - F_0,j(-a_{m,j} + b_{m,j})),
\]

where \( U_i = (U_{i,1}, \ldots, U_{i,d}), i = 1, \ldots, mn_m, \) are iid rv’s with cdf \( C_0(u), \) the copula function of \( F_0, \) and \( U_k:mn_{m,j} \) denotes the \( k \)-th order statistic of the marginal sample \( U_1,j, \ldots, U_{mn_m,j}. \) Hence, a few adaptations to the arguments in Section E.1.1 yield that, for \( j = 1, \ldots, d, \) as \( n \to \infty, \)

\[
F_0^{(\text{rms})}(\{\hat{b}_{m,j} - b_{m,j} / m_{m,j} > \epsilon\}) \leq 2 \exp\{-\tau \sqrt{n + 1}\},
\]

where \( \tau = \tau(\epsilon) \) is a positive constant. Therefore, by Borel-Cantelli lemma, almost surely as \( n \to \infty \)

\[
(\hat{b}_{m,j} - b_{m,j}) / m_{m,j} \to 0, \quad j = 1, \ldots, d.
\]

Furthermore, we have that, for \( j = 1, \ldots, d, \)

\[
F_0^{(\text{rms})}(\{\hat{a}_{m,j} / m_{m,j} - 1 > \epsilon\})
\]

\[
= F_0^{(\text{rms})}\left(\int_{b_{m,j}}^{\hat{a}_{m,j}} (1 - \hat{F}_{mn,m,j}(z))dz - \int_{b_{m,j}}^{a_{m,j}} (1 - F_0,j(z))dz > \epsilon \frac{a_{m,j}}{mn_m}\right)
\]

\[
+ F_0^{(\text{rms})}\left(\int_{b_{m,j}}^{\hat{a}_{m,j}} (1 - \hat{F}_{mn,m,j}(z))dz - \int_{b_{m,j}}^{a_{m,j}} (1 - F_0,j(z))dz < -\epsilon \frac{a_{m,j}}{mn_m}\right)
\]

\[
= T_{1,n}^{(j)} + T_{2,n}^{(j)}.
\]
An inequality analogous to that in (145), for some \( \epsilon' < \epsilon/2 \) and \( \tau' = \tau(\epsilon') > 0 \), together with a few simple manipulations, yield that, for \( j = 1, \ldots, d \), as \( n \to \infty \),

\[
F_{0}^{(nmn)} \left( \int_{b_{mn,j}}^{b_{nmn,j}} (1 - \hat{F}_{nmn,j}(z))dz > \frac{\epsilon a_{mn,j}}{2m_{n}} \right) \leq 2 \exp\{-\tau' \sqrt{n+1}\}.
\]

Therefore, as \( n \to \infty \),

\[
T_{1,n}^{(j)} \leq F_{0}^{(nmn)} \left( \int_{b_{mn,j}}^{b_{nmn,j}} (1 - \hat{F}_{nmn,j}(z))dz + \int_{b_{mn,j}}^{z_{0,j}} (F_{0,j}(z) - \hat{F}_{nmn,j}(z))dz > \frac{\epsilon a_{mn,j}}{2m_{n}} \right)
\]

\[
\leq F_{0}^{(nmn)} \left( \mathcal{D}_{W,1}(F_{0,j}, \hat{F}_{nmn,j}) > \frac{\epsilon a_{mn,j}}{2m_{n}} \right) + \exp\{-\tau' \sqrt{n+1}\}
\]

where \( \mathcal{D}_{W,1} \) is the 1-Wasserstein distance. Since \((\epsilon/2)a_{mn,j}/m_{n} < 1\) as \( n \to \infty \), by Theorem 2 in [34] it holds that

\[
F_{0}^{(nmn)} \left( \mathcal{D}_{W,1}(F_{0,j}, \hat{F}_{nmn,j}) > \frac{\epsilon a_{mn,j}}{2m_{n}} \right) \leq k_{j} \exp \left\{ -k_{j}' \frac{\epsilon^2 a_{mn,j}^2}{4m_{n}^2} n \right\}
+ k_{j} \exp \left\{ -k_{j}' \frac{\epsilon a_{mn,j}}{2m_{n}} \right\} \alpha_{j} - k_{j}'' \right\} 1(\alpha_{j} < 1),
\]

for some positive constants \( k_{j}, k_{j}' \) and any \( k_{j}'' \in (0, \alpha_{j}) \). By the assumption in (24), if \( \alpha_{j} \geq 1 \), the expression on the right-hand-side of the above display simplifies to

\[
k_{j} \exp \left\{ -k_{j}' \frac{\epsilon^2 a_{mn,j}^2}{4m_{n}^2} n \right\} = k_{j} n^{1-\epsilon'} \exp \left\{ -n \frac{a_{mn,j}^2}{m_{n}} \left( k_{j}' \frac{\epsilon^2}{4} - (1 + \epsilon') \frac{\log n}{n a_{mn,j}^2} \right) \right\}
\]

\[
\leq k_{j} n^{1-\epsilon'}
\]

for any \( \epsilon' > 0 \) and \( n \) large enough, while, if \( \alpha_{j} < 1 \) and \( k_{j}'' = \alpha_{j} - s \), it boils down to

\[
k_{j} \exp \left\{ -k_{j}' \frac{\epsilon^2 a_{mn,j}^2}{4m_{n}^2} n \right\} + k_{j} \exp \left\{ -k_{j}' \frac{\epsilon a_{mn,j}}{2m_{n}} \right\}
\]

\[
\leq 2k_{j} \exp \left\{ -k_{j}' \frac{\epsilon^2}{4} \min\{n a_{mn,j}^2/m_{n}, a_{mn,j}^2 n^s \} \right\}
\]

\[
= 2k_{j} n^{1-\epsilon'} \exp \left\{ -\min \left[ a_{mn,j}^2 n \left( k_{j}' \frac{\epsilon^2}{4} - (1 + \epsilon') \frac{\log n}{n a_{mn,j}^2} \right), a_{mn,j}^2 n^s \left( k_{j}' \frac{\epsilon^2}{2s} - (1 + \epsilon') \frac{\log n}{n a_{mn,j}^2 n^s} \right) \right] \right\}
\]

\[
\leq 2k_{j} n^{1-\epsilon'}.
\]

Consequently, letting \( k = \max_{j} k_{j} \), as \( n \to \infty \),

\[
\max_{1 \leq j \leq d} T_{1,n}^{(j)} \leq 2kn^{1-\epsilon'} + \exp\{-\tau' \sqrt{n+1}\} = O(n^{1-\epsilon'}).
\]

A similar reasoning leads to conclude that, as \( n \to \infty \),

\[
\max_{1 \leq j \leq d} T_{2,n}^{(j)} = O(n^{1-\epsilon'}).
\]

In particular, since \( \epsilon' > 0 \) is arbitrary, in the two displays above we can replace \( O \) with \( o \). By Borel-Cantelli lemma we can now conclude that with probability 1

\[
\tilde{a}_{mn,j} / a_{mn,j} \to 1, \quad j = 1, \ldots, d,
\]

as \( n \to \infty \), which completes the proof.
E.2.2. Technical derivations for Example 4.17. Note that, for $x \in (0, \infty)^2$, $F_0(x)$ allows the representation

$$F_0(x) = C_0 \left(1 - e^{-x_1}, 1 - e^{-x_2}\right),$$

where $C_0$ is as in (136) and satisfies Condition 4.1(i) (see Section E.1.2). The marginal distributions $F_{0,j}$, $j = 1, 2$, are exponential, thus satisfy Condition 4.1(ii) [e.g., 32, p. 1311]. Moreover, $F_0 \in \mathcal{D}(G_{1,0}(\cdot|H_0))$, where, for $x \in \mathbb{R}^d$,

$$G_{1,0}(x|H_0) = \exp\left\{-e^{-x_1} - e^{-x_2} + (e^{x_1} + e^{x_2})^{-1}\right\},$$

[e.g. 65, Example 5.16]. Valid norming sequences, asymptotically equivalent to those in (18), are

$$a_{m_n} = (1, 1), \quad b_{m_n} = (\log(m_n - 1), \log(m_n - 1)).$$

Thus, the probability density pertaining to $F_{m_n}^*(a_{m_n}x + b_{m_n})$, $x \in (-\log(m_n - 1), \infty)^2$, is given by

$$f_{m_n}(x) = f_{m_n}^*(e^{x_1}, e^{x_2})e^{x_1}e^{x_2},$$

where $f_{m_n}^*$ is a probability density defined as in (140), with $\rho_0 = 1$ and $V(m_n)$ as in (139). Furthermore, for $x \in \mathbb{R}^2$,

$$g_{1,0}(x|H_0) = g_1(e^{x_1}, e^{x_2}|H_0)e^{x_1}e^{x_2},$$

where the simple max-stable density $g_1(\cdot|H_0)$ equals the density in (141), with $\rho_0 = 1$ and $V(\cdot|H_0)$ as in (137). Hence,

$$\|f_{m_n}/g_{1,0}(\cdot|H_0)\|_\infty \leq \|f_{m_n}^*/g_1(\cdot|H_0)\|_\infty$$

and, by (144), we can conclude that the requirement in Condition 4.4 is satisfied in the present example.

E.3. Examples of Section 5.

E.3.1. Technical derivations for Example 5.3. In what follows, it is implicitly assumed that all the regularity conditions introduced in Example 5.3 are satisfied. To provide a simplified and concise account, we deal only with the limiting Fréchet case, assuming almost sure Hellinger consistency of $\hat{\rho}_n$. Derivations in the other cases obtain by a few adaptations. We start by providing the necessary notation, then we give three technical lemmas, finally we prove the convergence result

$$\lim_{n \to \infty} F_0^{m_n}(Q_n \triangle \hat{Q}_n) = 0, \quad F_0^{(\infty)} - \text{as.}$$

We firstly point out that Condition 4.1 and (26)-(27) guarantee that for the chosen $p \in (0, 1)$ there exists $\alpha > 0$ such that, defining $S := \text{supp}(G_{\rho_0,1}(\cdot|H_0))$ and

$$Q := \{x \in \hat{S} : g_{\rho_0,1}(x|H_0) \leq \alpha\},$$

we have $G_{\rho_0,1}(Q|H_0) = p$. Next, we denote $\varpi_{m_n} : (0, \infty)^d \mapsto (0, \infty)^d : x \mapsto x/a_{m_n}$ and define $S_{m_n}' := \text{supp}(F_{0}^{m_n} \circ \varpi_{m_n}^{-1})$, $\hat{S}_{m_n}' := \text{supp}(\hat{G}_{0}^{(\infty)} \circ \varpi_{m_n}^{-1})$. Then, for $\alpha_n := \alpha_n \prod_{j=1}^{d} a_{m_n,j}$ and $\hat{\alpha}_n := \hat{\alpha}_n \prod_{j=1}^{d} a_{m_n,j}$, we introduce the sublevel sets

$$Q_n' := \{x \in \hat{S}_{m_n}' : f_{m_n}(x) \leq \alpha_n\},$$

$$Q_n' := \{x \in \hat{S}_{m_n}' : \hat{g}_{n}^{(\infty)}(a_{m_n}x) \prod_{j=1}^{d} a_{m_n,j} \leq \hat{\alpha}_n\}.$$

Observe that $f_{m_n}$ and $\hat{g}_{n}^{(\infty)}(a_{m_n} \cdot) \prod_{j=1}^{d} a_{m_n,j}$ are the densities of $F_{0}^{m_n} \circ \varpi_{m_n}^{-1} \equiv F_{0}^{m_n}(a_{m_n} \cdot)$ and $\hat{G}_{0}^{(\infty)}(a_{m_n} \cdot)$, respectively, hence it is not difficult to see that $Q_n' = \varpi(Q_n)$ and $Q_n' = \varpi(\hat{Q}_n)$.

**Lemma E.1.** For any $\eta \in (0, \sup_{x \in S} g_{\rho_0,1}(x|H_0))$, it holds that

(i) $\lim_{n \to \infty} G_{\rho_0,1}\left(\{x \in \hat{S} : g_{\rho_0,1}(x|H_0) \leq \eta\} \triangle \{x \in \hat{S}_{m_n}' : f_{m_n}(x) \leq \eta\} \mid H_0\right) = 0$;

(ii) $\lim_{n \to \infty} G_{\rho_0,1}\left(\{x \in \hat{S} : g_{\rho_0,1}(x|H_0) \leq \eta\} \triangle \{x \in \hat{S}_{m_n}' : \hat{g}_{n}^{(\infty)}(a_{m_n}x) \prod_{j=1}^{d} a_{m_n,j} \leq \eta\} \mid H_0\right) = 0$, $F_0^{(\infty)} - \text{as.}$
PROOF. Under Condition 4.1 and (26)-(27), $\tilde{S}_{\rho_n} \subset \tilde{S}$ and

$$\lim_{n \to \infty} \mathcal{G}_T(G_{\rho_0,1}(\cdot|H_0), \tilde{T}_{F_0}^{-1}) = 0,$$

moreover $g_{\rho_0,1}(\cdot|H_0)$ is continuous and satisfy

$$\lim_{\delta \downarrow 0} G_{\rho_0,1}(\{x \in \tilde{S} : |g_{\rho_0,1}(x) - \eta| < \delta\}) = 0.$$

By these facts, for a small $\epsilon > 0$ there exists $\delta > 0$, such that, for all sufficiently large $n$,

$$G_{\rho_0,1}(\{x \in \tilde{S} : g_{\rho_0,1}(x|H_0) \leq \eta\} \triangle \{x \in \tilde{S}_{\rho_n} : f_{\rho_n}(x) \leq \eta\}|H_0)$$

$$\leq G_{\rho_0,1}(\{x \in \tilde{S}_{\rho_n} : g_{\rho_0,1}(x|H_0) > \eta\} \triangle \{x \in \tilde{S}_{\rho_n} : f_{\rho_n}(x) \leq \eta\}|H_0)$$

$$+ G_{\rho_0,1}(\{x \in \tilde{S}_{\rho_n} : f_{\rho_n}(x|H_0) \leq \eta\} \triangle \{x \in \tilde{S}_{\rho_n} : f_{\rho_n}(x) > \eta\}|H_0)$$

$$+ G_{\rho_0,1}(\tilde{S} \setminus \tilde{S}_{\rho_n}|H_0)$$

$$\leq G_{\rho_0,1}(\{x \in \tilde{S}_{\rho_n} : g_{\rho_0,1}(x|H_0) > \eta + \epsilon\} \triangle \{x \in \tilde{S}_{\rho_n} : f_{\rho_n}(x) \leq \eta\}|H_0)$$

$$+ G_{\rho_0,1}(\{x \in \tilde{S}_{\rho_n} : f_{\rho_n}(x|H_0) \leq \eta - \epsilon\} \triangle \{x \in \tilde{S}_{\rho_n} : f_{\rho_n}(x) > \eta\}|H_0) + \epsilon.$$

Furthermore, denoting $k_1 = \sqrt{1 + \delta / \eta}$ and $k_2 = \sqrt{\eta/(\eta - \delta)}$, by Lemma 2 in [80], as $n \to \infty$

$$G_{\rho_0,1}(\{x \in \tilde{S}_{\rho_n} : g_{\rho_0,1}(x|H_0) > \eta + \epsilon\} \triangle \{x \in \tilde{S}_{\rho_n} : f_{\rho_n}(x) \leq \eta\}|H_0)$$

$$\leq G_{\rho_0,1}(\{x \in \tilde{S}_{\rho_n} : f_{\rho_n}(x)/g_{\rho_0,1}(x|H_0) \leq \eta/(\eta + \epsilon)\})$$

$$\leq (1 - 1/k_1)^{-2} \mathcal{G}_H^2(f_{\rho_n}, g_{\rho_0,1}(\cdot|H_0))$$

$$\to 0$$

and

$$G_{\rho_0,1}(\{x \in \tilde{S}_{\rho_n} : g_{\rho_0,1}(x|H_0) \leq \eta - \epsilon\} \triangle \{x \in \tilde{S}_{\rho_n} : f_{\rho_n}(x) > \eta\}|H_0)$$

$$\leq F_0^{-1} \circ \omega^{-1}_{\rho_n}(\{x \in \tilde{S}_{\rho_n} : g_{\rho_0,1}(x|H_0)/f_{\rho_n}(x) \leq (\eta - \epsilon)/\eta\})$$

$$\leq (1 - 1/k_2)^{-2} \mathcal{G}_H^2(f_{\rho_n}, g_{\rho_0,1}(\cdot|H_0))$$

$$\to 0.$$

Since $\epsilon$ can be chosen arbitrarily small, the result at point (i) now follows. The result at point (ii) can be derived following similar steps. \hfill \Box

**Lemma E.2.** We have that $\lim_{n \to \infty} \alpha_{\rho_n} = \alpha$ and

$$\lim_{n \to \infty} \alpha_{\rho_n}' = \alpha, \quad F^{(\infty)}_0 - \text{as}.$$  

**Proof.** The two results can be shown by contradiction. We explicitly establish the first one, the second one follows by an analogous reasoning. Note that, by (147), as $n \to \infty$

$$G_{\rho_0,1}(Q) = p = F_0^{-1} \circ \omega^{-1}_{\rho_n}(Q_n) = G_{\rho_0,1}(Q_n) + o(1).$$

Assume $\lim_{n \to \infty} \alpha_{\rho_n}' \neq \alpha$, then there exists $\delta > 0$ such that for any $n'$ there is $n > n'$ satisfying $\alpha_{\rho_n}' < \alpha - \delta$ or $\alpha_{\rho_n}' > \alpha + \delta$. Assume the first case holds true: for any arbitrarily small $\epsilon > 0$, by Lemma 1(i) we thus have

$$G_{\rho_0,1}(Q_{\rho_n}'|H_0) \leq G_{\rho_0,1}(\{x \in \tilde{S}_{\rho_n} : f_{\rho_n}(x) \leq \alpha - \delta\}|H_0)$$

$$\leq G_{\rho_0,1}(\{x \in \tilde{S} : g_{\rho_0,1}(x|H_0) \leq \alpha - \delta\}|H_0) + \epsilon/2$$

$$= p - G_{\rho_0,1}(\{x \in \tilde{S} : \alpha - \delta < g_{\rho_0,1}(x|H_0) \leq \alpha\}|H_0) + \epsilon/2.$$

By Condition 4.1 and (26)-(27), for all $y_1, y_2$ satisfying

$$\inf_{x \in S} g_{\rho_0,1}(x|H_0) \leq y_1 < y_2 \leq \sup_{x \in S} g_{\rho_0,1}(x|H_0)$$

$$\alpha_{\rho_n}' \leq \frac{\inf_{x \in S} g_{\rho_0,1}(x|H_0)}{\sup_{x \in S} g_{\rho_0,1}(x|H_0)}$$

$$\leq \alpha_{\rho_n}.$$  

Since $\inf_{x \in S} g_{\rho_0,1}(x|H_0) > 0$, we have $\alpha_{\rho_n}' \leq \alpha_{\rho_n}$.

\hfill \Box
it holds that

\[ G_{\rho_0,1}(\{x \in \hat{S} : y_1 < g_{\rho_0,1}(x|H_0) \leq y_2 \}|H_0) > 0. \]

We are therefore allowed to select a positive \( \epsilon \) via

\[ \epsilon = G_{\rho_0,1}(\{x \in \hat{S} : \alpha - \delta < g_{\rho_0,1}(x|H_0) \leq \alpha \}|H_0) > 0, \]

which would render the right-hand side of (150) equal to \( p - \epsilon / 2 \). Since \( n' \) (and thus \( n \)) can be made arbitrarily large, the inequality \( G_{\rho_0,1}(\hat{Q}_n'|H_0) \leq p - \epsilon / 2 \) would end up contradicting (149). A similar contradiction is obtained in the case where \( \alpha'_n > \alpha + \delta \), hence it must be that \( \lim_{n \to \infty} \alpha'_n = \alpha \).

**Lemma E.3.** We have that

(i) \( \lim_{n \to \infty} G_{\rho_0,1}(\hat{Q}_n'|\{x \in \hat{S}_{m_n} : f_{m_n}(x) \leq \alpha \}|H_0) = 0; \)

(ii) \( \lim_{n \to \infty} G_{\rho_0,1}(\hat{Q}_n'|\{x \in \hat{S}_{m_n} : \hat{g}_n(a_{m_n}x) \prod_{j=1}^{d} a_{m_n,j} \leq \alpha \}|H_0) = 0, \]

**Proof.** To establish the result at point (i), observe that, by (148), for any \( \epsilon > 0 \) there exists \( \delta > 0 \) satisfying

\[ G_{\rho_0,1}(\{x \in \hat{S} : \alpha - \delta < g_{\rho_0,1}(x|H_0) \leq \alpha \}|H_0) < \epsilon, \]

thus by Lemma E.2 as \( n \to \infty \) we have \( \alpha'_n \in (\alpha - \delta, \alpha + \delta) \) and

\[ G_{\rho_0,1}(\hat{Q}_n'|\{x \in \hat{S}_{m_n} : f_{m_n}(x) \leq \alpha \}|H_0) \]
\[ \leq G_{\rho_0,1}(\{x \in \hat{S}_{m_n} : \alpha < f_{m_n}(x) \leq \alpha + \delta \}|H_0) 
\]
\[ + G_{\rho_0,1}(\{x \in \hat{S}_{m_n} : \alpha - \delta < f_{m_n}(x) \leq \alpha \}|H_0) \]
\[ \leq G_{\rho_0,1}(\{x \in \hat{S} : \alpha - \delta < g_{\rho_0,1}(x|H_0) \leq \alpha + \delta \}|H_0) + \epsilon \]
\[ \leq 2\epsilon, \]

where the inequality in third line follows from Lemma E.1(i). Since \( \epsilon \) can be chosen arbitrarily small, the result at point (i) now follows. The result at point (ii) can be established in an analogous fashion.

**Proof of (146).** Observe that

\[ F_{0,m}^m(\hat{Q}_n \triangle \hat{Q}_n) = \hat{F}_{0,m}^m \circ \pi_m^{-1}(\hat{Q}_n' \triangle \hat{Q}_n') \]
\[ \leq \hat{F}_{TV}(G_{\rho_0,1}(\cdot|H_0), F_{0,m}^m \circ \pi_m^{-1}) + G_{\rho_0,1}(\hat{Q}_n' \triangle \hat{Q}_n'|H_0). \]

By Condition 4.1, the first term on the right-hand side converges to zero as \( n \to \infty \). As for the second term, we have that

\[ G_{\rho_0,1}(\hat{Q}_n' \triangle \hat{Q}_n'|H_0) \leq G_{\rho_0,1}(\hat{Q}_n' \triangle \hat{Q}|H_0) + G_{\rho_0,1}(\hat{Q} \triangle \hat{Q}_n'|H_0) \]

On one hand, by Lemma E.1(i) and Lemma E.3(i), as \( n \to \infty \)

\[ G_{\rho_0,1}(\hat{Q}_n' \triangle \hat{Q}|H_0) \leq G_{\rho_0,1}(\hat{Q}_n' \triangle \{x \in \hat{S}_{m_n} : f_{m_n}(x) \leq \alpha \}|H_0) \]
\[ + G_{\rho_0,1}(\{x \in \hat{S}_{m_n} : f_{m_n}(x) \leq \alpha \} \triangle \hat{Q}|H_0) \]
\[ \to 0. \]

On the other hand, by Lemma E.1(ii) and Lemma E.3(ii), almost surely as \( n \to \infty \)

\[ G_{\rho_0,1}(\hat{Q} \triangle \hat{Q}_n'|H_0) \leq G_{\rho_0,1}(\hat{Q} \triangle \{x \in \hat{S}_{m_n} : \hat{g}_n(a_{m_n}x) \prod_{j=1}^{d} a_{m_n,j} \leq \alpha \}|H_0) \]
\[ + G_{\rho_0,1}(\{x \in \hat{S}_{m_n} : \hat{g}_n(a_{m_n}x) \prod_{j=1}^{d} a_{m_n,j} \leq \alpha \} \triangle \hat{Q}_n'|H_0) \]
\[ \to 0. \]

The result now follows.
E.3.2. Technical derivations for Example 5.4. Once more, the limiting Fréchet case is considered, postulating that Condition 4.1 and (25) hold true. Adaptations to the other cases are outlined at the end. To show that

\[
\lim_{n \to \infty} \left( F_{0,j}^{\omega_n} \right)^{\times (1 - p)} = 1, \quad j = 1, \ldots, d, \quad \forall p \in (0, 1), \quad F_0^{(\infty)} - as
\]

we resort to the following two lemmas.

**Lemma E.4.** For \( j = 1, \ldots, d \), let \( q_j := G_{\rho_0,j}^{\omega_n}(1 - p) \) and \( q_n,j := (F_{0,j}^{\omega_n})^{\times (1 - p)} \), then

\[
\lim_{n \to \infty} \frac{q_n,j}{a_{m_n,j}} = q_j.
\]

**Proof.** We use a reasoning by contradiction. Fix \( j \in \{1, \ldots, d\} \), then under Condition 4.1 for all \( \delta > 0 \) there exists \( n_\delta \in \mathbb{N} \) such that

\[
\mathcal{D}_{KS}(F_{0,j}^{\omega_n}, G_{\rho_0,j,1}) < \delta, \quad \forall n \geq n_\delta.
\]

In particular,

\[
F_{0,j}^{\omega_n}(q_n,j) = F_{0,j}^{\omega_n}(a_{m_n,j}(q_n,j/a_{m_n,j})) = 1 - p,
\]

thus

\[
|G_{\rho_0,j,1}(q_n,j/a_{m_n,j}) - (1 - p)| < \delta, \quad \forall n \geq n_\delta.
\]

Now, if (152) did not hold true, for all \( \epsilon > 0 \) and \( n' \in \mathbb{N} \) there would exist \( n \geq n' \) such that \( q_n,j/a_{m_n,j} < (1 - \epsilon)q_j \) or \( q_n,j/a_{m_n,j} > (1 + \epsilon)q_j \). Assume the second case holds true, then for the specific choice of \( \delta = G_{\rho_0,j,1}(1 + (1 + \epsilon)q_j) - G_{\rho_0,j,1}(q_j) \) we would have

\[
G_{\rho_0,j,1}(q_n,j/a_{m_n,j}) > G_{\rho_0,j,1}(1 + (1 + \epsilon)q_j) > \delta + 1 - p,
\]

which would end up contradicting (153), since \( n' \) (and thus \( n \)) can be made arbitrarily large. A similar contradiction obtains in the case where \( q_n,j < (1 - \epsilon)q_j \) via the inequality \( G_{\rho_0,j,1}(q_n,j/a_{m_n,j}) < 1 - p - \delta \), under the specific choice \( \delta = G_{\rho_0,j,1}(q_j) - G_{\rho_0,j,1}(q_j(1 - \epsilon)) \). The proof is now complete.

**Lemma E.5.** For \( j = 1, \ldots, d \), let \( q_j \) be as in Lemma E.4 and define \( \hat{q}_n,j := (G_{\omega_n}^{\infty})^{\times (1 - p)} \), then

\[
\lim_{n \to \infty} \frac{\hat{q}_n,j}{a_{m_n,j}} = q_j, \quad F_0^{(\infty)} - as.
\]

**Proof.** Observe that (25) can be rephrased as

\[
\lim_{n \to \infty} \mathcal{D}_H \left( \hat{g}_n^\omega(a_{m_n,j}) \prod_{j=1}^d a_{m_n,j} \cdot f_{m_n} \right) = 0, \quad F_0^{(\infty)} - as.
\]

On the other hand, Condition 4.1 entails that \( \lim_{n \to \infty} \mathcal{D}_H \left( q_{\rho_0,1}(|H_0), f_{m_n} \right) = 0 \). Thus, by triangular inequality

\[
\lim_{n \to \infty} \mathcal{D}_H \left( g_n^\omega(a_{m_n,j}) \prod_{j=1}^d a_{m_n,j} \cdot q_{\rho_0,1}(|H_0) \right) = 0, \quad F_0^{(\infty)} - as.
\]

Consequently, for any given \( j \in \{1, \ldots, d\} \) and \( \delta > 0 \), eventually almost surely \( \mathcal{D}_{KS}((\hat{G}_{n,j}^{\omega_n})(a_{m_n,j}), G_{\rho_0,j,1}) < \delta \). The result can be now proved by contradiction, reasoning as in the proof of Lemma E.4.

The result in (151) now immediately obtains by combining Lemmas E.4-E.5. The result obtains analogously in the Gumbel and Weibull limiting cases. Indeed, by Condition 4.1, for all \( j = 1, \ldots, d \), the Kolmogorov-Smirnov distance between \( F_{0,j}^{\omega_n}(a_{m_n,j} + b_{m_n,j}) \) and \( \hat{G}_{0,1} \) (Gumbel case) or \( G_{\omega_0,j,1,0} \) (Weibull case) converges to
zero. Consequently, defining $q_{n,j}$ as above and $q_j$ as $G_{0,1}^+(1 - p)$ or $G_{\omega_{0,j},1,0}^+(1 - p)$, arguments similar to those in the proof of Lemma E.4 yield that
\[
\lim_{n \to \infty} \frac{q_{n,j} - b_{m_{n,j}}}{a_{m_{n,j}}} = q_j.
\]
The Kolmogorov-Smirnoff distance between $\hat{G}_{n,j}^o(a_{m_{n,j}} \cdot + b_{m_{n,j}})$ and $G_{0,1}$ (Gumbel case) or $G_{\omega_{0,j},1,0}$ (Weibull case) converges to zero too, either almost surely or in probability, depending whether Hellinger consistency of $\hat{g}_n^o$ holds almost surely or in probability. Accordingly, defining $\hat{q}_{n,j}$ as above, it can be shown that
\[
\frac{\hat{q}_{n,j} - b_{m_{n,j}}}{a_{m_{n,j}}} \to 0
\]
as $n \to \infty$, either almost surely or in probability. Such a result parallels Lemma E.5. From the two displays above and the properties of the norming sequences chosen as in (18) it can be finally deduced that
\[
\frac{q_{n,j}}{\hat{q}_{n,j}} \to 1
\]
as $n \to \infty$, either almost surely or in probability.