



Delft University of Technology

Document Version

Final published version

Licence

Dutch Copyright Act (Article 25fa)

Citation (APA)

Gili, F., Jongbloed, G., & Van Der Vaart, A. (2026). Semiparametric Bernstein–Von Mises Phenomenon Via Isotonized Posterior In Wicksell’s Problem. *Annals of Statistics*, 54(1), 383-407. <https://doi.org/10.1214/25-AOS2571>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

In case the licence states “Dutch Copyright Act (Article 25fa)”, this publication was made available Green Open Access via the TU Delft Institutional Repository pursuant to Dutch Copyright Act (Article 25fa, the Taverne amendment). This provision does not affect copyright ownership.
Unless copyright is transferred by contract or statute, it remains with the copyright holder.

Sharing and reuse

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.

SEMIPARAMETRIC BERNSTEIN–VON MISES PHENOMENON VIA ISOTONIZED POSTERIOR IN WICKSELL’S PROBLEM

BY FRANCESCO GILI^a, GEURT JONGBLOED^b AND AAD VAN DER VAART^c

Delft Institute of Applied Mathematics (DIAM), Delft University of Technology, ^aF.Gili@tudelft.nl, ^bG.Jongbloed@tudelft.nl, ^cA.W.vanderVaart@tudelft.nl

In this paper, we propose a novel Bayesian approach for nonparametric estimation in Wicksell’s problem. This has important applications in astronomy for estimating the distribution of the positions of the stars in a galaxy given projected stellar positions and in materials science to determine the 3D microstructure of a material, using its 2D cross-sections. We deviate from the classical Bayesian nonparametric approach, which would place a Dirichlet Process (DP) prior on the distribution function of the unobservables, by directly placing a DP prior on the distribution function of the observables. Our method offers computational simplicity due to the conjugacy of the posterior and allows for asymptotically efficient estimation by projecting the posterior onto the $\mathbb{L}_2$ subspace of increasing, right-continuous functions. Indeed, the resulting Isotonized Inverse Posterior (IIP) satisfies a Bernstein–von Mises (BvM) phenomenon with minimax asymptotic variance $g_0(x)/2\gamma$, where $\gamma > 1/2$ reflects the degree of Hölder continuity of the true cdf at x . Since the IIP gives automatic uncertainty quantification, it eliminates the need to estimate γ . Our results provide the first semiparametric Bernstein–von Mises theorem for projection-based posteriors with a DP prior in inverse problems.

1. Introduction. In the field of stereology, scientists study three-dimensional properties of materials and objects by interpreting their two-dimensional cross-sections. This approach allows the estimation of certain quantities without making costly 3D reconstructions. Consider now these two stereological problems. The first one: astronomers are interested in the distribution of stars in the universe. Globular clusters—spherical aggregations of stars held together by gravity (cf. [1, 38])—are a topic of particular focus (see Figure 1). But how to determine this distribution when given only a 2D photo of the clusters? The second one: having a material that presents a globular microstructure that cannot be scanned (or only at high cost, e.g., steel) but only sectioned, how can we determine the distribution of particle sizes inside the material when given only a limited number of 2D cross-sections (cf. [12, 30])? Despite their seemingly distant nature, these problems share a common thread: the inherent structure of an inverse problem, originally conceptualized by Wicksell (cf. [17, 22, 43, 44]).

The setting of the opaque material having globular microstructure is formalized by assuming that the embedded particle squared radii are i.i.d. realizations from a random variable $X^* \sim F_0^*$ and our observations consist of cross-sections of the medium (i.e., “planes cutting the medium”). Note that the spheres with bigger radii are more likely to be sampled. Thus the distribution of the squared radii of the spheres actually cut is not F_0^* but a biased version of it; we write for the squared radius of the sphere actually cut: $X \sim F_0$. Since we assume that the samples are gathered uniformly, by the Pythagorean theorem, the observable squared circle radius Z is related to X via: $Z = (1 - U^2)X = YX$, where $U \sim U[0, 1]$ so $Y \sim \text{Be}(1, 1/2)$, with density $y \mapsto \frac{1}{2}(1 - y)^{-1/2}\mathbf{1}_{(0,1)}(y)$. So, the observed squared radii Z

Received March 2025; revised July 2025.

MSC2020 subject classifications. Primary 62F15, 62C10, 62G20, 62G05; secondary 62C20, 62E20.

Key words and phrases. Semiparametric estimation, Bernstein–von Mises, Bayesian nonparametrics, inverse problems, isotonic estimation, efficiency theory.

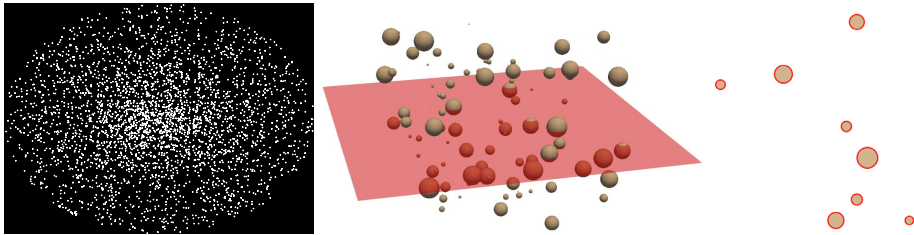


FIG. 1. Two examples of sampling in stereological inverse problems. Left panel: The Messier 4 (also designated NGC 6121) is a globular cluster in the constellation of Scorpius (cf. [1]). Center and right panel: simulation 3D configuration with spheres and the intersecting plane in material microstructures (from [17]).

are a random fraction of the true squared radii depending on where the plane cuts through the sphere. Consequently, since both X and Y are nonnegative, a version of the density g_0 of Z can, for $z \geq 0$, be written as

$$(1.1) \qquad g_0(z) = \frac{d}{dz} \int_{(0,\infty)} \mathbb{P}(Y \leq z/x) dF_0(x) = \int_z^\infty \frac{dF_0(x)}{2\sqrt{x^2 - xz}}.$$

The problem in the astronomical setting can be modeled as follows: let $\mathbf{Z} = (Z_{p_1}, Z_{p_2}, Z_{p_3})$ be a spherically symmetric random vector (the three-dimensional position of a star in a galaxy) of which only (Z_{p_1}, Z_{p_2}) can be observed. Under the assumption that the joint of $\mathbf{Z}$ is a function of the distance to the center of the galaxy, astronomers are interested in estimating the distribution of the squared distance of a star to the center of the globular cluster or galaxy $X = Z_{p_1}^2 + Z_{p_2}^2 + Z_{p_3}^2 \sim F_0$ using $Z = Z_{p_1}^2 + Z_{p_2}^2$. From the assumed spherical symmetry, $\mathbf{Z} \sim O\mathbf{Z}$ for all orthogonal matrices $O \in \text{O}(3)$. Therefore, $\mathbf{U} := \mathbf{Z}/\|\mathbf{Z}\|$ (note $X = \|\mathbf{Z}\|^2$) is uniform on the unit sphere $\mathbb{S}^2$. It follows that the random variable Z_{p_3} (the third coordinate) is distributed as $\sqrt{X} \cdot U_3$ where $U_3 \sim U[-1, 1]$ independent of X . Note also that $Y \equiv 1 - U_3^2 \sim \text{Be}(1, 1/2)$ as above. Because

$$\mathbb{P}(Z \leq z) = \mathbb{P}(Z_{p_1}^2 + Z_{p_2}^2 + Z_{p_3}^2 \leq z + Z_{p_3}^2) = \mathbb{P}(X \leq z + X \cdot U_3^2) = \mathbb{P}(YX \leq z),$$

it follows that both for the application in materials science and the one in astronomy, the density of the observations $Z \sim G_0$ is the one given on the right-hand side of (1.1).

By using Abel-integral equation's theory, we can invert (1.1) and obtain (cf. [38])

$$(1.2) \qquad F_0(x) = 1 - \frac{2}{\pi} \sqrt{x} V_0(x) - \frac{2}{\pi} \int_x^\infty \frac{V_0(s)}{2\sqrt{s}} ds,$$

where

$$(1.3) \qquad V_0(x) := \int_x^\infty \frac{g_0(z)}{\sqrt{z-x}} dz.$$

(The proof of (1.2) is provided in Lemma 1.1 in the Supplementary Material [18]). In Corollary 2.3, we formally show that the functional $x \mapsto -\frac{2}{\pi} \int_x^\infty V_0(s)/(2\sqrt{s}) ds$ can be estimated at the rate $\sqrt{n}$, whereas the functional $V_0(x)$ can be estimated (nonparametrically) only at the slower rate $\sqrt{n}/\sqrt{\log n}$. This can be intuitively seen because, by plugging (1.3) into the second term of (1.2) and using Fubini's theorem to exchange the order of integration:

$$(1.4) \qquad \int_x^\infty \frac{V_0(s)}{2\sqrt{s}} ds = \frac{\pi}{2} - \int_0^\infty \sin^{-1} \sqrt{1 \wedge \frac{x}{z}} g_0(z) dz$$

and $\text{Var}[\sin^{-1} \sqrt{1 \wedge (x/Z)}] < \infty$, for $Z \sim G_0$, in contrast with $\text{Var}[(Z-x)^{-1/2}] = \infty$. It thus follows that the estimation of V_0 and F_0 are essentially equivalent. Furthermore, as shown in

[22], the relation $F_0^*(x) = 1 - \frac{V_0(x)}{V_0(0)}$ holds, which implies that also in the materials science setting, the primary target of estimation is V_0 . Accordingly, our focus in this paper will first be on estimating $V_0(x)$, and we will subsequently show that these results naturally extend to the estimation of $F_0(x)$. The latter will be the main object of interest in the astronomical application discussed in the simulation study.

We aim to develop an easy-to-compute and statistically efficient nonparametric Bayesian method for estimating $V_0(x)$ at a given point $x \geq 0$. We begin by placing a Dirichlet process prior directly on the distribution function of the observables: $G \sim \text{DP}(\alpha)$. Given an observed sample $Z_1, \dots, Z_n$ from cdf G_0 (associated with density g_0), the posterior distribution is (cf. [15], Theorem 4.6 in [16]):

$$(1.5) \quad G \mid Z_1, \dots, Z_n \sim \text{DP}(\alpha + n\mathbb{G}_n), \quad \mathbb{G}_n := \frac{1}{n} \sum_{i=1}^n \delta_{Z_i}.$$

This choice offers computational simplicity due to the conjugacy of the posterior. Leveraging the inverse relations inherent to the problem, we can estimate the functional of interest using (1.2) by

$$(1.6) \quad V_G(x) := \int_x^\infty \frac{dG(z)}{\sqrt{z-x}}.$$

This initial step yields the *Naive Bayes Posterior* (NBP). We show below that this satisfies a BvM phenomenon with asymptotic variance $g_0(x)$. However, the realizations V_G of this plug-in posterior are not necessarily monotone functions. To remedy this, we project them into the $\mathbb{L}_2$ subspace of decreasing and right-continuous functions. It turns out that this not only corrects the nonmonotonicity, but also improves the estimator. We show that the resulting *Isotonized Inverse Posterior* (IIP) $\hat{V}_G(x)$ satisfies a BvM theorem with improved variance that depends on the local smoothness of V_0 . The smoothness at x is measured through a parameter $\gamma > \frac{1}{2}$ that can be loosely understood as the degree of Hölder continuity at x . Using (1.2), we can define the IIP to target the estimation of $F_0(x)$, which we will denote by $\hat{F}_G(x)$. The formal definitions of $\hat{V}_G(x)$ and $\hat{F}_G(x)$, being quite technical, are given in (2.9) and (2.11) below, respectively.

Thus the approach in this paper has three advantages:

- (i) *Computational speed*: The Dirichlet posterior (1.5) is explicit and its projection can be effectively computed by the Pool-Adjacent-Violators Algorithm (PAVA, cf. [3]).
- (ii) *Efficiency and adaptation*: The IIP achieves the asymptotic variance of the efficient estimators introduced in [17] and [21] and it is adaptive to the level of local smoothness of V_0 .
- (iii) *Uncertainty quantification*: The Bayesian framework inherently provides uncertainty quantification without the need to estimate γ or other parameters. We show below that this is asymptotically correct in the frequentist sense.

The Naive Bayes Posterior (NBP) is analogous to the naive estimator introduced in [22] and [43], which is constructed as follows: given a sample $Z_1, \dots, Z_n$ from the density g_0 and for $\mathbb{G}_n$ the empirical distribution function based on this sample,

$$(1.7) \quad V_n(x) := \int_x^\infty \frac{d\mathbb{G}_n(z)}{\sqrt{z-x}}.$$

In [17], the authors showed that the projection of the naive estimator onto the set of monotone functions, the *Isotonized Inverse Estimator* (IIE) $\hat{V}_n$, is asymptotically efficient and the attained asymptotic variance depends on the local smoothness of V_0 in the same way as for the IIP.

Projecting the naive posterior distribution departs from a classical Bayesian framework, which might assign a Dirichlet Process (DP) prior to F and employ a hierarchical model to sample from the density of the observations. A Bernstein–von Mises (BvM) result for this approach is currently unavailable. Moreover, the available algorithms for posterior sampling rely on Gibbs sampling procedures, which are much slower than the method we propose here.

From a practical perspective, the Bernstein–von Mises (BvM) result established in this paper assures practitioners that the Bayesian methodology based on DP prior provides all the efficiency benefits of the Isotonic Inverse Estimator (IIE) of [17], but without requiring the estimation of difficult-to-access nuisance quantities such as the density of the observations $g_0(x)$ or the local Hölder smoothness $\gamma > 1/2$. In applications, where one aims to recover the distribution of squared distances of stars from their projected two-dimensional positions or the distribution of the embedded particles' squared radii, this is especially valuable: the underlying smoothness of these distributions (and thus of V_0) may vary across radii, and the BvM result holds for all local smoothness levels $\gamma > 1/2$. This significantly improves over existing bootstrap-based uncertainty guarantees of [38], which only hold under the stronger assumption that V_0 is differentiable at the estimation point x (which is even stronger than $\gamma = 1$).

Motivation and connections with the literature. There is a vast literature on Wicksell's problem; see, for instance, [11, 13, 17, 20–24, 26, 38, 39, 43]. Wicksell's problem has present-day applications, where it is common to see the use of the so-called Saltykov method [12], and numerical discretization methods [30], often lacking theoretical guarantees. These methods are far from being efficient and in some cases not even consistent. The estimation of V_0 given data from g_0 is of interest not only for stereological procedures, but also from the mathematical point of view, both because of the unusual rates of convergence and because of the nonstandard efficiency theory.

In the nonparametric frequentist approach to Wicksell's problem, the standard rate of convergence for pointwise asymptotic normality in estimating V_0 is $(\log n)^{-1/2}\sqrt{n}$ (cf. [17, 22, 43]). The first time the naive plug-in estimator was studied is in [43]. Later in [22], the proof of the asymptotic normality of this estimator was simplified considerably. However, viewed as a function, the naive plug-in estimator is severely ill-behaved due to infinite discontinuities and nonmonotonicity, and this led to the definition of the IIE in [17, 22]. The nonparametric maximum likelihood estimator is studied in [11, 26]. For Wicksell's problem, the usual definition of the NPMLE is also ill-posed, as the log-likelihood can attain value infinity. However, work has been done, in [11] and [26], to obtain likelihood based estimators. Confidence interval results can be found in [13, 38]. The results of [13] provide confidence bands by developing a methodology for isotonization procedures, which often lead to nuisance parameters in the limiting distributions that tackles this issue. However, they do not cover the setting considered in the present paper. Extending their results to our case appears nontrivial, as their proof techniques rely on at least the differentiability of the estimated object—an assumption we do not make. The bootstrap results in [38] are more directly relevant and will be discussed in detail in the simulation section, including a comprehensive numerical comparison. The main advantage of the methodology proposed in this paper, if compared to the bootstrap methods of [38], is that it relies on significantly less stringent assumptions. In particular, our results do not depend on the assumption that the underlying distribution function F_0 is differentiable at x . Instead, our methodology is provably consistent for all local smoothness levels $\gamma > 1/2$ of V_0 , where γ is defined in condition (2.13) below.

In [21], the authors investigate efficient estimation in the classical Wicksell problem, focusing on the scenario where the distribution function F_0^* of interest belongs to a Hölder class of smoothness with degree $\gamma > \frac{1}{2}$. They propose a kernel density-based estimator that,

under appropriate bandwidth selection, achieves asymptotic efficiency. However, a notable limitation is that the estimator requires prior knowledge or guesstimation of the smoothness parameter γ , making it nonadaptive. Moreover, efficiency is only guaranteed under global smoothness assumptions, that is, the entire distribution function must belong to a global Hölder class. Despite these limitations, their work remains the first foundational contribution to semiparametric efficiency in Wicksell's problem. [17] builds upon earlier work of [21] and [22] by addressing both the adaptivity and locality of efficient estimation in Wicksell's problem. The key innovation consists in showing that the Isotonized Inverse Estimator (IIE) achieves asymptotic efficiency under the same local smoothness assumptions on the target function F_0^* of the present paper. Specifically, instead of assuming that $F_0^* \in \Lambda_\gamma$ globally as in [21], the analysis in [17] only requires that F_0^* satisfies local Hölder-type smoothness at the evaluation point x and at the origin. The estimator adapts automatically to the local regularity of F_0^* , offering a more flexible framework for practitioners. Additionally, [17] provides a local asymptotic minimax lower bound, showing that the IIE attains the asymptotically optimal risk in this refined setting. The estimator thus achieves the same efficiency as the KDE-based approaches, but without requiring tuning of smoothness parameters, and without imposing strong global assumptions. Finally, other results concerning efficient estimation in Wicksell's problem are contained in [19] and [20]. [19] shows that in the case the underlying cdf is constant on a *known* interval of strictly positive length the IIE evaluated at x is not efficient, for x in the interval of constancy. However, it behaves in practice very closely to informed estimators that make use of the fact that the cdf is constant on that interval. Finally, one of the authors of [21] did some further work in [20] where it was shown how to attain efficient smoothing for the plug-in estimator under squared losses.

While Bayesian methods for inverse problems are widely adopted, their theoretical guarantees remain scarce, particularly in infinite-dimensional settings. Although recent years have seen significant advances in this area [6–8, 25, 28, 29, 31–36], the technical frameworks developed there do not readily apply to the specific inverse problem and Dirichlet process prior setting of this paper. We provide a detailed discussion of this related work in Supplementary Material [18]. A similar methodology to the one proposed in this paper—using an unconstrained conjugate prior for the distribution function of the observables and projecting the posterior samples into the space of monotone functions—has been very recently adopted in isotonic regression settings (cf. [9, 41, 42]), isotonic density estimation (cf. [10]) and ODEs settings (cf. [4, 5]). [4, 5] develop Bayesian two-step methods for parameter inference in ODE models by embedding them in a nonparametric regression framework with B-spline priors. The parameter posterior is obtained by minimizing a discrepancy between the estimated regression function and the ODE solution, either via derivative matching or Runge–Kutta approximations. [5] further establishes a Bernstein–von Mises theorem for the Runge–Kutta-based method, showing that the resulting Bayes estimator is asymptotically efficient. However, their framework relies on a very different type of prior modeling and the projection method does not entail isotonization. In [9], the authors study a nonparametric Bayesian regression model for a response variable Y with respect to a predictor variable $X \in [0, 1]$, given by $Y = f(X) + \varepsilon$, where f is a monotone increasing function and ε is a mean-zero random error. They propose a “projection-posterior” estimator f^* , constructed using a finite random series of step functions with normal basis coefficients as a prior for f . Their main finding is that when f^* is centered at the maximum likelihood estimator (MLE) and rescaled appropriately, the Bernstein–von Mises (BvM) theorem does not hold. However, when f^* is instead centered at the true function f_0 , a BvM result is obtained. This phenomenon resonates with the known inconsistency of the bootstrap for the Grenander estimator in isotonic regression (cf. [37]) and represents a key difference with the present work. In our case, the “projection-posterior” IIP $\hat{V}_G$, when centered at the IIE $\hat{V}_n$, is shown to converge in distribution. This may

also not be surprising given the existing results in [38], where bootstrap methods are shown to be consistent for Wicksell’s problem.

One of the key advantages our approach shares with the works in [4, 5, 9, 41, 42] is the computational efficiency, driven by posterior conjugacy, as well as the relative simplicity of the asymptotic analysis. However, unlike in [9], our method ensures that the limiting distribution in the Bernstein–von Mises (BvM) theorem coincides with that of the asymptotically efficient estimator—the IIE of [17].

2. Main results. We start by proving a BvM result for the *Naive Bayes Posterior* (NBP), analogous to the result obtained for the naive estimator (cf. Theorem 2 in [22]). For our first result, we need our version of g_0 to be well defined and bounded, that is,

(2.1)
$$\|g_0\|_\infty := \sup_{z \geq 0} \left| \int_z^\infty \frac{dV_0(s)}{\pi \sqrt{s-z}} \right| < \infty.$$

Assumption (2.1) asserts that projected 2D star positions do not become arbitrarily concentrated. From a mathematical perspective, without this assumption the inverse problem becomes highly unstable: small errors in the data can cause arbitrarily large errors in estimating V_0 . For instance, if g_0 is unbounded at x , this leads to instability in estimating $V_0(x)$, reflected in an infinite asymptotic variance. Thus, boundedness of g_0 ensures stability of the inverse mapping. Mathematically, if F_0 is Hölder continuous with degree $\gamma > 1/2$, then g_0 is bounded (cf. Lemma A.3 in [17]).

In equations (2.2) and (2.15) below, we use the notion of conditional weak convergence in distribution (in probability). For its precise definition, we refer the reader to (A.1)–(A.2) in Appendix A.

THEOREM 2.1 (BvM for NBP). *Let $x \geq 0$ and suppose that (2.1) holds true. Let $G \sim \text{DP}(\alpha)$, with $\int_x^\infty (s-x)^{-1/2} d\alpha(s) < \infty$. Then, in probability under G_0 ,*

(2.2)
$$\frac{\sqrt{n}}{\sqrt{\log n}} (V_G(x) - V_n(x)) \mid Z_1, \dots, Z_n \rightsquigarrow N(0, g_0(x)).$$

PROOF. For a measure P and a function f , we denote by Pf the integral $\int f dP$, this holding also for empirical measures like $\mathbb{G}_n = \frac{1}{n} \sum_{i=1}^n \delta_{Z_i}$.

Since $G \mid Z_1, \dots, Z_n \sim \text{DP}(\alpha + n\mathbb{G}_n)$ by Proposition G.10 in [16], for $Q \sim \text{DP}(\alpha)$, $\mathbb{B}_n \sim \text{DP}(n\mathbb{G}_n)$, $V_n \sim \text{Be}(|\alpha|, n)$ and $f(z) = 1/\sqrt{(z-x)_+}$, $x \geq 0$, conditionally on $Z_1, \dots, Z_n$, we have the following representation in distribution:

(2.3)
$$Gf = V_n Qf + (1 - V_n) \mathbb{B}_n f,$$

where $Q, \mathbb{B}_n, V_n$ are independent. First, note that by assumption Qf is a.s. well defined (cf. Remark 4.4 in [16]). By Proposition G.2 in [16] for $\varepsilon_i \stackrel{\text{i.i.d.}}{\sim} \text{Exp}(1)$ and $W_{ni} := \varepsilon_i/\bar{\varepsilon}_n$, we have $(W_{n1}, \dots, W_{nn}) \sim \text{Dir}(n; 1/n, \dots, 1/n)$. Since $\mathbb{B}_n \sim \text{DP}(n\mathbb{G}_n)$, for δ_{Z_i} Dirac measure at Z_i , we have the following representation in distribution of $\mathbb{B}_n$ (cf. Example 3.7.9 in [40]):

$$\mathbb{B}_n = \frac{1}{n} \sum_{i=1}^n W_{ni} \delta_{Z_i}.$$

Because of this representation (cf. Section 3.7.2. in [40]), $\mathbb{B}_n$ is also known as *Bayesian bootstrap*. Conclude that

$$\frac{\sqrt{n}}{\sqrt{\log n}} (\mathbb{B}_n - \mathbb{G}_n) f = \frac{1}{\sqrt{n \log n}} \sum_{i=1}^n (W_{ni} - 1) f(Z_i).$$

Because $|\bar{\varepsilon}_n - 1| = O_p(1/\sqrt{n})$, we study $(n \log n)^{-1/2} \sum_{i=1}^n (\varepsilon_i - 1) f(Z_i)$. Now note that by Lemma 2.4 in the Supplementary Material [18]

$$(2.4) \quad \frac{1}{\sqrt{n \log n}} \sum_{i=1}^n \frac{(\varepsilon_i - 1)}{\sqrt{Z_i - x}} \mathbf{1}_{\{Z_i > x\}} \mid Z_1, \dots, Z_n \rightsquigarrow N(0, g_0(x)) \quad \text{under } G_0.$$

Because Qf is a.s. well defined, we obtain the claim by Slutsky's lemma upon noticing $\frac{\sqrt{n}}{\sqrt{\log n}} V_n Qf \xrightarrow{G_0} 0$. $\square$

We consider a projection-type posterior by isotonizing draws from the obtained naive posterior (1.6) as follows. First, define, for $G \sim \text{DP}(\alpha + n\mathbb{G}_n)$ and $\mathbb{G}_n, G_0$ as above:

$$(2.5) \quad U_G(x) := \int_0^x V_G(y) dy = 2 \int_0^\infty \sqrt{z} dG(z) - 2 \int_x^\infty \sqrt{z-x} dG(z),$$

$$(2.6) \quad U_n(x) := \int_0^x V_n(y) dy = 2 \int_0^\infty \sqrt{z} d\mathbb{G}_n(z) - 2 \int_x^\infty \sqrt{z-x} d\mathbb{G}_n(z),$$

$$(2.7) \quad U_0(x) := \int_0^x V_0(y) dy = 2 \int_0^\infty \sqrt{z} dG_0(z) - 2 \int_x^\infty \sqrt{z-x} dG_0(z).$$

Furthermore, for a continuous function $k : [0, \infty) \rightarrow [0, \infty)$, the Least Concave Majorant (LCM) and its right-hand side derivative, are given for $x \in [0, \infty)$ by

$$k^*(x) := \inf \{ f(x) : f \geq k \text{ on } [0, \infty), f \text{ concave on } [0, \infty) \},$$

$$(k^*)'_+(x) := \lim_{h \rightarrow 0^+} \frac{k^*(x+h) - k^*(x)}{h}.$$

We denote by $\hat{V}_n$ (IIE) the right-hand side derivative of U_n^* , the LCM of U_n . Similarly, $\hat{V}_G$ is the right-hand side derivative of U_G^* , the LCM of U_G , that is, the isotonization of a draw from the naive Bayes posterior. Thus, for G as in (1.5),

$$(2.8) \quad \hat{V}_n(x) := (U_n^*)'_+(x),$$

$$(2.9) \quad \hat{V}_G(x) := (U_G^*)'_+(x).$$

By Lemma 2 in [22], $\hat{V}_n$ and $\hat{V}_G$ can be viewed as (equivalent to) $\mathbb{L}_2$ -projections of the naive estimators V_n and V_G into the space of nonnegative, decreasing and right-continuous functions on $[0, \infty)$. We can finally define the Naive posterior and the IIP to target the estimation of F_0 as follows:

$$(2.10) \quad F_G(x) := 1 - \frac{2}{\pi} \sqrt{x} V_G(x) - \frac{2}{\pi} \int_x^\infty \frac{V_G(s)}{2\sqrt{s}} ds,$$

$$(2.11) \quad \hat{F}_G(x) := 1 - \frac{2}{\pi} \sqrt{x} \hat{V}_G(x) - \frac{2}{\pi} \int_x^\infty \frac{\hat{V}_G(s)}{2\sqrt{s}} ds,$$

and analogously we define $F_n(x)$, with V_n instead of V_G in (2.10) and $\hat{F}_n(x)$, with $\hat{V}_n$ instead of $\hat{V}_G$ in (2.11).

We now state two additional conditions needed for Theorem 2.2. First, a natural assumption is a moment constraint for Z (which can be translated into a moment condition on X , cf. [17]):

$$(2.12) \quad \int_0^\infty z g_0(z) dz < \infty.$$

This is a mild assumption, which is practically always satisfied, given that materials studied in materials science are of bounded size and globular clusters typically do not extend indefinitely. Additionally, assume $\exists K > 0$ and $\gamma > \frac{1}{2}$ such that

(2.13)
$$\int_0^1 (V_0(x) - V_0(x + u\delta)) \, du \sim \operatorname{sgn}(\delta) |\delta|^\gamma K, \quad \text{as } \delta \rightarrow 0,$$

where $\sim$ means that the ratio of the left-hand side and of the right-hand side goes to 1 as $\delta \rightarrow 0$. The constant γ indicates the level of smoothness at x of V_0 . This is analogous to equation (5) from [17]. Note that it is possible to see what this condition means in terms of the underlying F_0 , using (1.2), we obtain that $\exists K > 0$ and $\gamma > \frac{1}{2}$ such that, as $\delta \rightarrow 0$,

(2.14)
$$\int_0^1 (\sqrt{x} V_0(x) - \sqrt{x + u\delta} V_0(x + u\delta)) \, du \sim \operatorname{sgn}(\delta) |\delta|^\gamma K - \frac{V_0(x)}{4\sqrt{x}} \operatorname{sgn}(\delta) |\delta|,$$

(upon using a Taylor expansion for $\int_0^1 (\sqrt{x} - \sqrt{x + u\delta}) \, du$). Because $\frac{2}{\pi} \int_x^\infty \frac{V_0(s)}{2\sqrt{s}} \, ds$ is Lipschitz continuous, the local smoothness of F_0 at x is $\min(\gamma, 1)$. Therefore, for the most relevant levels of smoothness, that is, $\gamma \in (\frac{1}{2}, 1]$, the smoothness of V_0 carries over to F_0 . Because $F_0^*(x) = 1 - \frac{V_0(x)}{V_0(0)}$, equation (5) from [17] is equivalent to (2.13) (note that in that paper F denotes the current F_0^*). Thus, for a better mathematical understanding of the role of γ , we refer the interested reader to the discussion given in [17], in particular Examples 2.1, 2.2, 2.3 and Remark 2.1.

From a physical point of view, the role of γ is also interpretable. In a globular cluster, different regions (e.g., dense core vs. sparse halo) may correspond to different levels of smoothness of F_0 , and the semiparametric theory reflects this. The local adaptivity of the Bayesian estimator—encoded in the limiting variance of the Bernstein–von Mises—mirrors this phenomenon. If there is a sharp increase in the concentration of stars with size close to x (i.e., γ close to $1/2$), then the 2D observations are much noisier. This makes the estimation more difficult and results in a higher asymptotic variance.

THEOREM 2.2 (BvM for IIP). *Let $x \geq 0$, suppose that (2.1), (2.12) and (2.13) hold true. Moreover, let $G \sim \text{DP}(\alpha)$, where the base measure α has bounded density. Then, in probability under G_0 ,*

(2.15)
$$\frac{\sqrt{n}}{\sqrt{\log n}} (\hat{V}_G(x) - \hat{V}_n(x)) \mid Z_1, \dots, Z_n \rightsquigarrow N\left(0, \frac{g_0(x)}{2\gamma}\right).$$

PROOF. The proof essentially consists of three parts. The first part is to rewrite the event

$$\left\{ \sqrt{\frac{n}{\log n}} (\hat{V}_G(x) - \hat{V}_n(x)) \leq a \right\}$$

given $Z_1, \dots, Z_n$ in terms of the location of the maximum ($\arg \max$) of a stochastic process $t \mapsto \tilde{Z}_n^G(t) - at$. It turns out that the process $\tilde{Z}_n^G(t)$ can be decomposed into three essential parts. The second step consists in studying these processes. One process can be interpreted as the Bayesian bootstrap counterpart of the process studied in the proof of Theorem 2.1 in [17] (cf. equation (14) in [17]). This process converges conditionally in distribution to a Gaussian process $\mathbb{Z}$, which we show can be written as $\mathbb{Z}(t) = tX$, where $X \sim N(0, g_0(x)/(2\gamma))$. A second process (see (2.22)), which is a deterministic process, whose behavior is controlled by the local smoothness of V_0 at x and which converges to $-|t|^{\gamma+1}K$ (with γ and K as in (2.13)). Lastly, a third remainder process which converges uniformly to zero with respect to the topology of uniform convergence on compacta. The final step is to prove that from the

conditional convergence in distribution of $t \mapsto \tilde{Z}_n^G(t) - at$ to $\mathbb{Z}(t) - at$ it follows that the $\arg \max$ of $(\tilde{Z}_n^G(t) - at)$ converges in distribution to the $\arg \max$ of $\mathbb{Z}(t) - at$. Since the $\arg \max$ functional is certainly not continuous on the space of continuous functions equipped with the uniform norm, the simplest form of the continuous mapping theorem cannot be applied. For this reason, we prove in Lemma A.3, specifically for this application, a conditional argmax continuous mapping theorem (extension of Theorem 3.2.2 in [40]). From (2.29) below, it follows that the sequence of locations of the maximum of $(\tilde{Z}_n(t) - at)$ is conditionally uniformly tight. This enables us to use Lemma A.3. Moreover, since the limiting process turns out almost surely to have a unique maximum, the asserted conditional convergence in distribution in (2.15) follows.

For sequences $\delta_n = \sqrt{\log n}/\sqrt{n}$ and $\delta_n^* = \delta_n^{1/\gamma}$ where $\gamma > 1/2$ satisfies condition (2.13), define

$$(2.16) \quad Z_n^{G_0}(t) := \delta_n^{-1}(\delta_n^*)^{-1} \{U_n(x + \delta_n^* t) - U_n(x) - U_0(x + \delta_n^* t) + U_0(x)\}$$

$$(2.17) \quad = 2\delta_n^{-1}(\delta_n^*)^{-1} \int \{\sqrt{(z-x)_+} - \sqrt{(z-x-\delta_n^* t)_+}\} d(\mathbb{G}_n - G_0)(z).$$

By Lemma A.1 in Appendix A, it suffices to show, for $X \sim N(0, \frac{g_0(x)}{2\gamma})$ and every $a \in \mathbb{R}$,

$$(2.18) \quad \Pi_n \left(\sqrt{\frac{n}{\log n}} (\hat{V}_G(x) - V_0(x)) - Z_n^{G_0}(1) \leq a \mid Z_1, \dots, Z_n \right) \xrightarrow{G_0} \mathbb{P}(X \leq a).$$

Define the process T_G for $a > 0$,

$$T_G(a) := \inf\{t \geq 0 : U_G(t) - at \text{ is maximal}\}.$$

By the switch relation (see Lemma 3.2 in [23]),

$$T_G(a) \leq x \iff \hat{V}_G(x) \leq a.$$

Therefore, for sufficiently large n , conditionally on $Z_1, \dots, Z_n$,

$$\delta_n^{-1}(\hat{V}_G(x) - V_0(x)) - Z_n^{G_0}(1) \leq a \iff (\delta_n^*)^{-1}(T_G(V_0(x) + \delta_n(a + Z_n^{G_0}(1))) - x) \leq 0.$$

Because the location of a maximum of a function is equivariant under translations and invariant under multiplication by a positive number and addition of a constant:

$$\begin{aligned} & (\delta_n^*)^{-1}(T_G(V_0(x) + \delta_n(a + Z_n^{G_0}(1))) - x) \\ &= (\delta_n^*)^{-1}(\inf\{x + t \geq 0 : U_G(x + t) - V_0(x)(x + t) \\ & \quad - \delta_n(a + Z_n^{G_0}(1))(x + t) \text{ is max}\} - x) \\ (2.19) \quad &= \inf\{t \geq -(\delta_n^*)^{-1}x : U_G(x + \delta_n^* t) - U_G(x) - U_n(x + \delta_n^* t) + U_n(x) \\ & \quad + U_0(x + \delta_n^* t) - U_0(x) - V_0(x)\delta_n^* t - \delta_n\delta_n^*(at + Z_n^{G_0}(1)t - Z_n^{G_0}(t)) \text{ is max}\} \\ &= \inf\{t \geq -(\delta_n^*)^{-1}x : \tilde{Z}_n^G(t) - at \text{ is max}\}, \end{aligned}$$

where we define the localized process $\tilde{Z}_n^G$ by

$$(2.20) \quad \tilde{Z}_n^G(t) := \delta_n^{-1}(\delta_n^*)^{-1} \{U_G(x + \delta_n^* t) - U_G(x) - U_n(x + \delta_n^* t) + U_n(x)\}$$

$$(2.21) \quad + \delta_n^{-1}(\delta_n^*)^{-1} \{U_0(x + \delta_n^* t) - U_0(x) - V_0(x)\delta_n^* t\} + Z_n^{G_0}(t) - tZ_n^{G_0}(1)$$

for $t \in I_x$, where $I_x := [0, \infty)$ if $x = 0$, and $I_x := \mathbb{R}$ if $x > 0$.

We determine the asymptotic behavior of the process $\tilde{Z}_n^G$ by studying separately the components in (2.20) and (2.21), and next we conclude by applying Slutsky's lemma. First, note that $|Z_n^{G_0}(1)t - Z_n^{G_0}(t)| = o_{G_0}(1)$ uniformly in t in compacta by the proof of Theorem 2.1 in [17] (cf. display after equation (16)). Furthermore, using assumption (2.13) (as in equation (13) in [17]), uniformly in t in compacta

$$(2.22) \quad \delta_n^{-1}(\delta_n^*)^{-1} \{U_0(x + \delta_n^*t) - U_0(x) - V_0(x)\delta_n^*t\} = -|t|^{\gamma+1}(K + o(1)).$$

This determines the behavior of (2.21). Now we study the posterior behavior of the process in (2.20), which can be interpreted as a *Bayesian bootstrap* counterpart of the process (2.16) studied in the proof of Theorem 2.1 in [17] (cf. equation (14) in [17]):

$$\begin{aligned} Z_n^G(t) &:= \delta_n^{-1}(\delta_n^*)^{-1} \{U_G(x + \delta_n^*t) - U_G(x) - U_n(x + \delta_n^*t) + U_n(x)\} \\ &= 2\delta_n^{-1}(\delta_n^*)^{-1} \int \{\sqrt{(z-x)_+} - \sqrt{(z-x-\delta_n^*t)_+}\} d(G - \mathbb{G}_n)(z). \end{aligned}$$

As in proof of Theorem 2.1 in equation (2.3), we have by Proposition G.10 in [16], for independent $Q \sim \text{DP}(\alpha)$, $\mathbb{B}_n \sim \text{DP}(n\mathbb{G}_n)$, and $V_n \sim \text{Be}(|\alpha|, n)$, the following representation in distribution:

$$(2.23) \quad Gf_t^n = V_n Qf_t^n + (1 - V_n)\mathbb{B}_n f_t^n,$$

where $f_t^n(z) := 2(\delta_n^*)^{-1} \{\sqrt{(z-x)_+} - \sqrt{(z-x-\delta_n^*t)_+}\}$, and thus the following representation in distribution of the process $Z_n^G(t)$:

$$Z_n^G(t) = \delta_n^{-1} V_n (Q - \mathbb{G}_n) f_t^n + \delta_n^{-1} (1 - V_n) (\mathbb{B}_n - \mathbb{G}_n) f_t^n.$$

This representation is crucial for the asymptotic analysis of the process $Z_n^G(t)$. The first thing to do is to identify which parts of the process $Z_n^G(t)$ do not give any contribution to limiting distribution of the process, as they converge conditionally to zero in probability. The second part is to identify the limiting distribution of the remaining parts of the process, which will be a Gaussian process. In Appendix A, we give the [proof](#) that uniformly in t in compacta

$$(2.24) \quad Z_n^G(t) = \mathbb{Z}_n^*(t) + o_p(1),$$

where for $\varepsilon_i \stackrel{\text{i.i.d.}}{\sim} \text{Exp}(1)$, the process $\mathbb{Z}_n^*$ is defined as

$$(2.25) \quad \mathbb{Z}_n^*(t) := \frac{1}{\sqrt{n \log n}} \sum_{i=1}^n (\varepsilon_i - 1) f_t^n(Z_i).$$

Then by Slutsky's lemma, if we determine the conditional weak limit of $(\mathbb{Z}_n^*(t) : t \in I_x)$, then we determine the one of $(Z_n^G(t) : t \in I_x)$. We now study the marginals of $\mathbb{Z}_n^*$. Note that by Lemma 2.5 in the Supplementary Material [18] for fixed $t \in I_x$, in probability under G_0 ,

$$(2.26) \quad \mathbb{Z}_n^*(t) \mid Z_1, \dots, Z_n \rightsquigarrow N(0, t^2 g_0(x)/(2\gamma)),$$

and again by Lemma 2.5 in the Supplementary Material [18] that

$$(2.27) \quad \text{Cov}(\mathbb{Z}_n^*(t), \mathbb{Z}_n^*(s) \mid Z_1, \dots, Z_n) = \frac{1}{2\gamma} g_0(x)st \left(1 - O_{G_0}\left(\frac{\log \log n}{\log n}\right)\right) + O_{G_0}\left(\frac{1}{\log n}\right).$$

By the conditional Chebyshev inequality and (2.27), we conclude $\forall \varepsilon > 0$

$$\mathbb{P}(|\mathbb{Z}_n^*(t) - t\mathbb{Z}_n^*(1)| > \varepsilon \mid Z_1, \dots, Z_n) \leq \frac{1}{\varepsilon^2} \text{Var}(\mathbb{Z}_n^*(t) - t\mathbb{Z}_n^*(1) \mid Z_1, \dots, Z_n) \xrightarrow{G_0} 0.$$

By (2.24), conclude that (2.26) holds true with $\mathbb{Z}_n^*$ replaced by Z_n^G . Therefore, the finite-dimensional distributions of Z_n^G converge conditionally in distribution to the finite-dimensional distributions of the process: $\mathbb{Z}(t) = tX$, where $X \sim N(0, g_0(x)/(2\gamma))$.

Lemma A.2 gives the conditional asymptotic equicontinuity of the process $\mathbb{Z}_n^*$. The conditional asymptotic equicontinuity together with the conditional convergence in distribution of the finite-dimensional distributions, implies in probability under G_0 as $n \rightarrow \infty$:

$$(2.28) \quad \mathbb{Z}_n^* \mid Z_1, \dots, Z_n \rightsquigarrow \mathbb{Z} \quad \text{in } \ell^\infty(K)$$

for every compact $K \subset I_x$ and ℓ^∞ being the space of bounded functions on K equipped with the supremum norm. For a review of the conditional weak convergence in ℓ^∞ , see [40] (cf. Example 1.5.1 and Theorem 1.5.4). By (2.24) and (2.26), conclude that (2.28) holds true with $\mathbb{Z}_n^*$ replaced by Z_n^G , that is, the conditional weak limit of Z_n^G relative to the topology of uniform convergence on compacta is the same Gaussian process. Altogether, recalling also (2.22), we can conclude that conditionally on $Z_1, \dots, Z_n$ the weak limit of $(\tilde{Z}_n^G(t) - at : t \in K)$ in ℓ^∞ is given by $(tX - |t|^{\gamma+1}K - at : t \in K)$ for every compact $K \subset I_x$.

In order to apply Lemma A.3, notice that $\mathbb{Z}$ has a.s. continuous sample paths and a.s. unique maxima (see Lemma 2.6 in [27]) and that $\hat{t}_n^G := \inf\{t \geq -(\delta_n^*)^{-1}x : \tilde{Z}_n^G(t) - at \text{ is max}\}$ is an argmax process conditionally on $Z_1, \dots, Z_n$, satisfying the second condition on Lemma A.3. It is thus enough to show that there exists a conditionally stochastically bounded sequence $(\tilde{M}_n)$ that satisfies

$$(2.29) \quad \Pi_n((\delta_n^*)^{-1} |T_G(V_0(x) + \delta_n(a + Z_n^{G_0}(1))) - x| \leq \tilde{M}_n \mid Z_1, \dots, Z_n) \xrightarrow{G_0} 1.$$

The proof of (2.29) is given in Appendix A. Thus by Lemma A.3, we can conclude

$$\begin{aligned} \Pi_n\left(\sqrt{\frac{n}{\log n}}(\hat{V}_G(x) - V_0(x)) - Z_n^{G_0}(1) \leq a \mid Z_1, \dots, Z_n\right) \\ \xrightarrow{G_0} \mathbb{P}\{\arg \max_{t \in I_x} \{tX - |t|^{\gamma+1}K - at\} \leq 0\} = \mathbb{P}\{X \leq a\}, \end{aligned}$$

because $t \mapsto t(X - a) - |t|^{\gamma+1}K$ attains its maximum on $(-\infty, 0]$ iff $X - a \leq 0$. $\square$

We conclude this section by showing the implications of the above results for the uncertainty quantification of the posterior. Under the conditions of Theorem 2.2 (cf. [17]),

$$\frac{\sqrt{n}}{\sqrt{\log n}}(\hat{V}_n(x) - V_0(x)) \rightsquigarrow N\left(0, \frac{g_0(x)}{2\gamma}\right).$$

Thus as consequence of (1.2) and the analogous definition of $\hat{F}_n(x)$, we have that (see proof of Corollary 2.3 in the [18]):

$$(2.30) \quad \frac{\sqrt{n}}{\sqrt{\log n}}(\hat{F}_n(x) - F_0(x)) \rightsquigarrow N\left(0, \frac{2x}{\pi^2\gamma}g_0(x)\right).$$

Together with (2.32) below, this will imply that the credible intervals constructed from the Isotonized Inverse Posterior (IIP) of $\hat{F}_G(x)$ are (asymptotic) frequentist confidence sets. The following Corollary 2.3, whose proof is given in the Supplementary Material [18], summarizes formally the above discussion. Note that because the uncertainty quantification based on the IIP is always preferable to the one based on the naive Bayes posterior, we only propose uncertainty quantification using the latter.

COROLLARY 2.3 (BvM for $\hat{F}_G$ and F_G and UQ for F_0). *Let F_G as in (2.10) and $\hat{F}_G$ as in (2.11) and analogously let $F_n(x)$, with V_n instead of V_G in (2.10) and $\hat{F}_n(x)$, with $\hat{V}_n$ instead of $\hat{V}_G$ in (2.11). Under the conditions of Theorem 2.1, in probability under G_0 as $n \rightarrow \infty$,*

$$(2.31) \quad \frac{\sqrt{n}}{\sqrt{\log n}}(F_G(x) - F_n(x)) \mid Z_1, \dots, Z_n \rightsquigarrow N\left(0, \frac{4x}{\pi^2} g_0(x)\right).$$

Under the same conditions of Theorem 2.2, in probability under G_0 as $n \rightarrow \infty$,

$$(2.32) \quad \frac{\sqrt{n}}{\sqrt{\log n}}(\hat{F}_G(x) - \hat{F}_n(x)) \mid Z_1, \dots, Z_n \rightsquigarrow N\left(0, \frac{2x}{\pi^2 \gamma} g_0(x)\right).$$

Further, if we define the posterior quantiles for isotonized posterior $\hat{F}_G$, for $\alpha \in [0, 1]$, as

$$q_\alpha := \inf\{t \in \mathbb{R} : \Pi_n(\hat{F}_G(x) \leq t \mid Z_1, \dots, Z_n) \geq \alpha\}$$

then under G_0 as $n \rightarrow \infty$,

$$(2.33) \quad \mathbb{P}(F_0(x) \in [q_{\alpha/2}, q_{1-\alpha/2}]) \rightarrow 1 - \alpha.$$

3. Simulation study and discussion about uncertainty quantification. In this section, we present simulations to illustrate the practical behavior of the IIP in comparison to the IIE studied in [17]. Furthermore, we present a simulation study that compares the IIP-uncertainty quantification to the one proposed in [38]. We consider three distinct settings. The first two are *synthetic* settings where the true underlying distribution function F_0 is known. The first two simulation settings serve as valuable benchmarks for assessing the quality of our methodology. In particular, the first one verifies that the variance attained by the IIP coincides with the theoretical one that we derived in Theorem 2.3. The second *synthetic* setting compares the uncertainty quantification of our methodology to the bootstrap-based one derived in [38]. The last simulation setting is a *real-data* application involving the positions of stars in the Globular Cluster Messier 62. In the last case, the ground truth distribution of the stars is unknown. Also in this setting, we compare our methodology with the bootstrap of [38].

In Figure 2, we analyze the setting where the underlying cdf F_0 is exponential with rate 1.2. Samples $Z_1, \dots, Z_n$ used to compute the projection posterior of the IIP are generated hierarchically: $X_1, \dots, X_n \sim F_0$, $Y_1, \dots, Y_n \sim \text{Be}(1, 1/2)$, with $Z_i = X_i \cdot Y_i$ for all i . The plots use the following color scheme: green for estimates of F based on prior draws from $G \sim \text{DP}(M\alpha)$ (parameters M and α specified in the plots); purple for draws from the isotonized posterior; red for the average of these draws (approximating the posterior mean); yellow for the Isotonic Inverse Estimator (IIE) of [17]; and orange for F_0 . The first row of Figure 2 compares the behavior of the IIP and the IIE as the sample size goes from 200 to 2000. With $n = 200$, the IIE and the average of the IIP draws already behave very similarly and with $n = 2000$, the uncertainty significantly reduces and the two estimation procedures essentially coincide. The average IIP appears to be somewhat smoother, which will typically be seen as an advantage. The middle row focuses on $n = 2000$, providing a zoomed-in view and highlighting $x = 1.5$, the point used for subsequent analysis of the normality. The final row examines normality ($n = 2000$): histograms, QQ-plots and statistical tests strongly confirm normality of the gathered samples. Furthermore, the Bayesian-estimated standard deviation aligns closely with the theoretical value attained by the frequentist procedure (cf. [17], note that $\gamma = 1$ because F_0 is Lipschitz). In all the plots, we employ 300 posterior draws.

In Figure 3, we analyze six different settings where at $x = 5$ the underlying cdfs F_0 have varying degrees of smoothness, making them locally Hölder smooth of degree $\gamma \in \{0.55, 0.6, 0.65, 0.7, 0.8, 1.5\}$. In particular, for an appropriate constant $K > 0$, $F_0(y) = \frac{1}{2} - K(5 - y)^\gamma$, for $y \in [0, 5]$ and $F_0(y) = \frac{1}{2} + K(y - 5)^\gamma$, for $y \in [5, 10]$. The plots use the

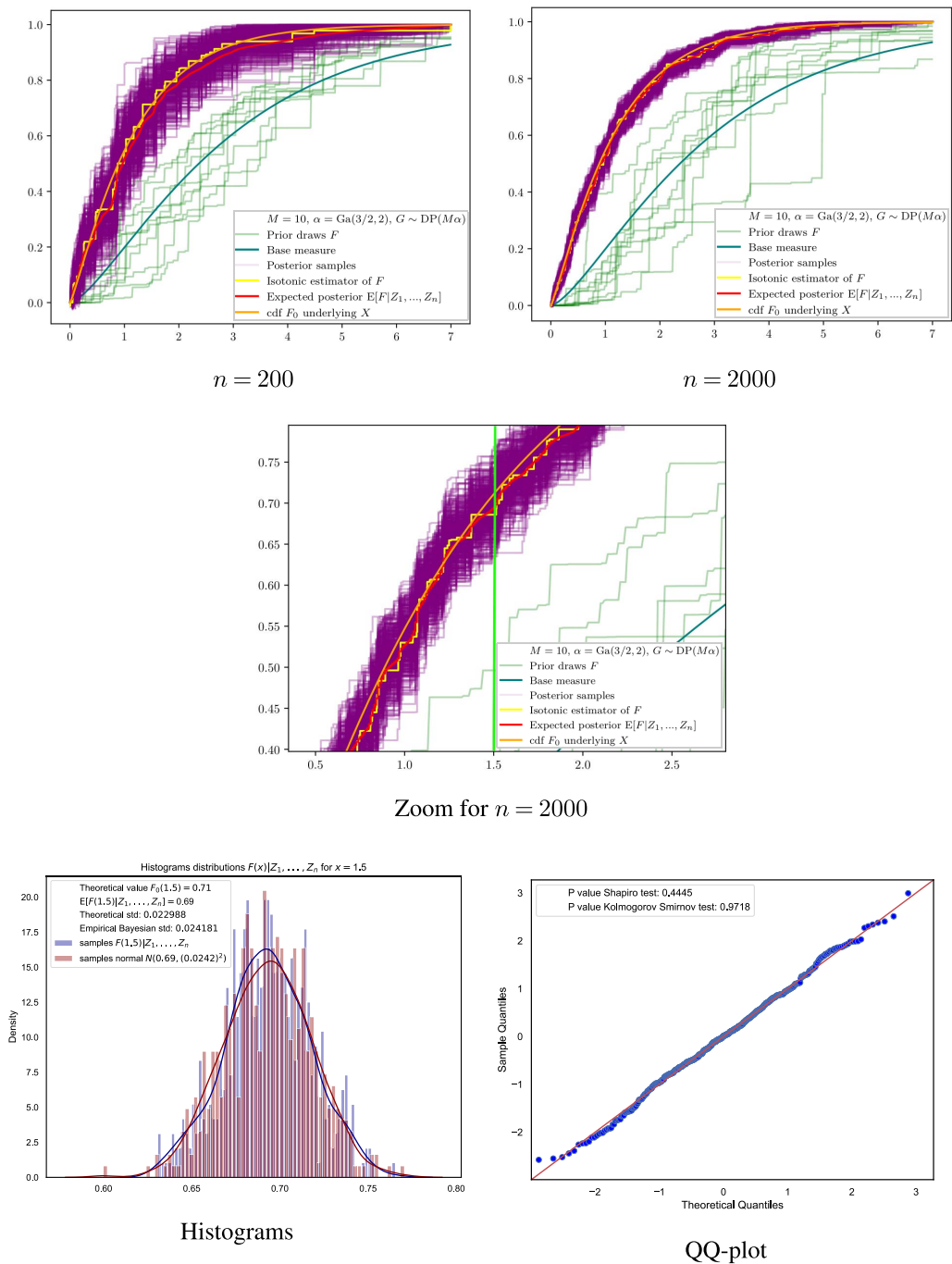


FIG. 2. Synthetic setting 1: Uncertainty for different sample sizes and analysis of normality at $x = 1.5$, based on 300 draws from the posterior.

TABLE 1
Estimated widths of CI for different levels of γ for Bootstrap and Bayesian procedures

	$\gamma = 0.55$	$\gamma = 0.6$	$\gamma = 0.65$	$\gamma = 0.7$	$\gamma = 0.8$	$\gamma = 1.5$
Bootstrap CI width at $x = 5$	0.116	0.110	0.104	0.108	0.100	0.073
Bayesian CI width at $x = 5$	0.114	0.121	0.103	0.101	0.099	0.077

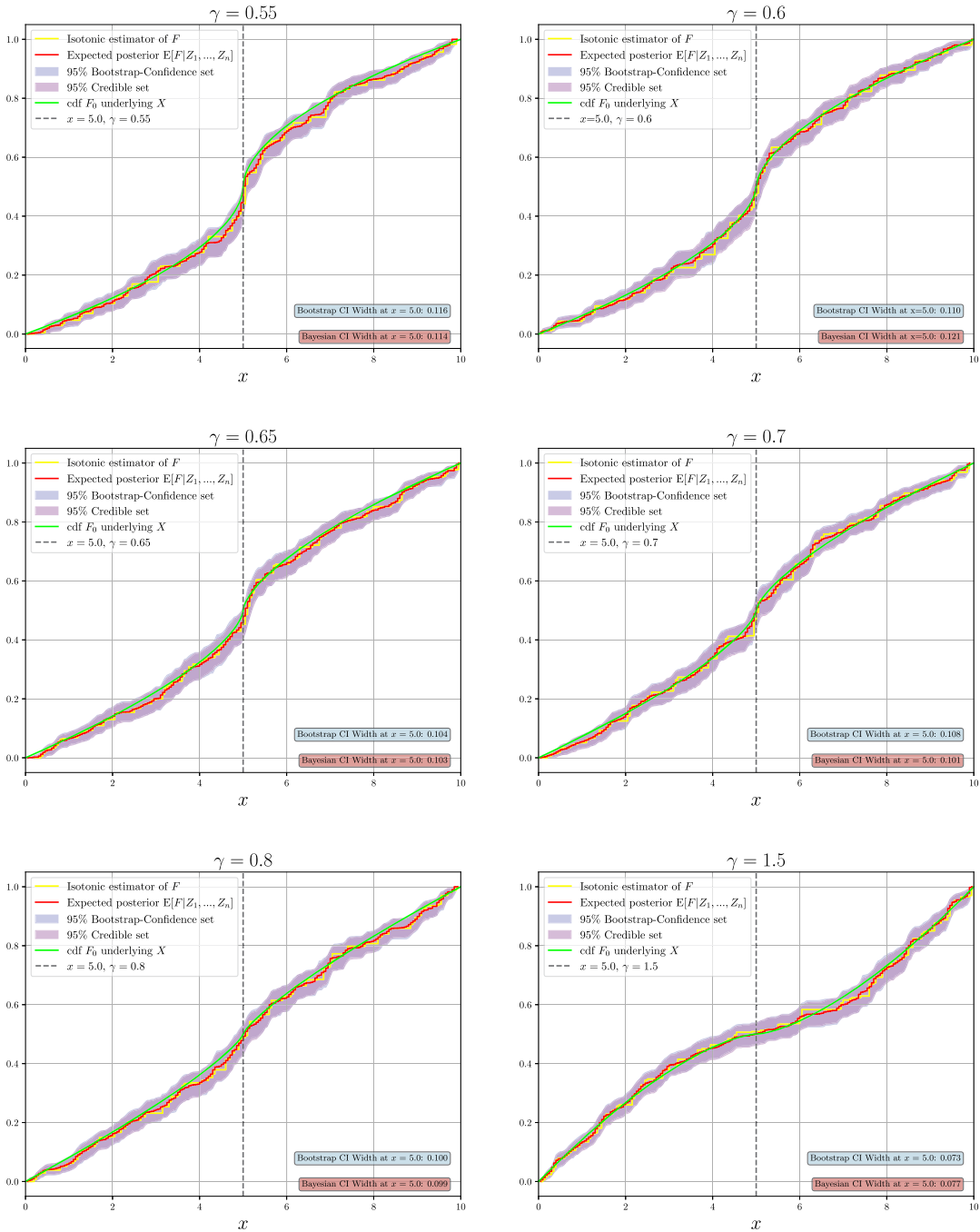


FIG. 3. Synthetic setting 2: Comparison uncertainty quantification IIP and bootstrap method from [38].

following color scheme: pink for the Bayesian-based credible sets; red for the approximated posterior mean; light blue for the bootstrap-based confidence sets of [38]; yellow for the Isotonic Inverse Estimator (IIE) of [17]; and lime green for F_0 . The prior base measure used for these simulation is the same used in the previous simulation study. In all cases $n = 2000$, and we employed 300 posterior-based and bootstrap-based draws. This comparison confirms that the IIP delivers valid uncertainty quantification for varying degrees of smoothness and that the uncertainty quantification delivered by [38] behaves closely (see Table 3), even though the authors formally showed it works only for a differentiable underlying truth F_0 .

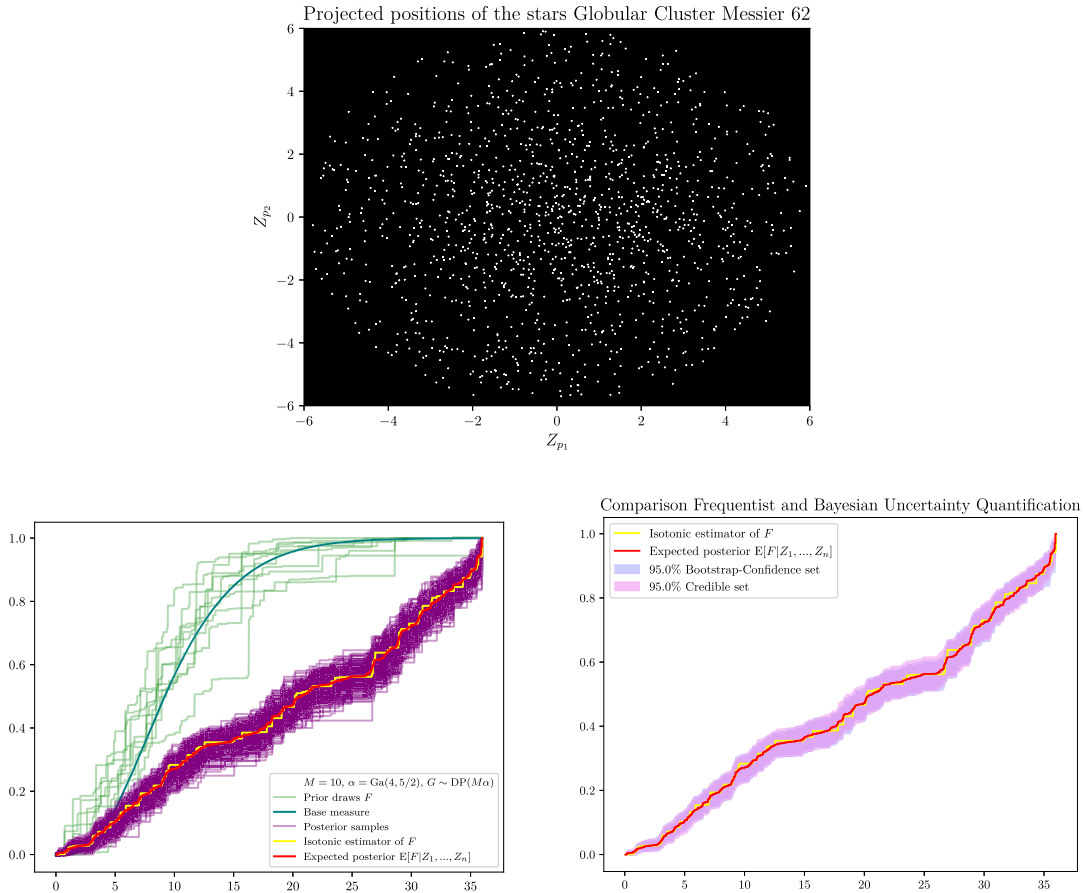


FIG. 4. Real-data application: $n = 1400$ projected star positions. On the left-hand side, the IIE from [17] and the IIP with 300 posterior samples; on the right-hand side, a comparison between the 95% bootstrap confidence set for the IIE from [38] and the 95% credible sets for the IIP.

In Figure 4, we apply the methodology to star position data from globular cluster Messier 62. Let $\mathbf{Z} = (Z_{p_1}, Z_{p_2}, Z_{p_3})$ represent a star’s three-dimensional position, where only (Z_{p_1}, Z_{p_2}) are observed. Here, $X = Z_{p_1}^2 + Z_{p_2}^2 + Z_{p_3}^2 \sim F_0$, and $Z = Z_{p_1}^2 + Z_{p_2}^2$. Beyond serving as a compelling application of our methodology, this example highlights a few key conclusions. First of all, also for real data the IIE closely resembles the IIP. However, because in this case the true γ is unknown, it is not immediately clear how the frequentist uncertainty quantification compares with the Bayesian one. From the work of [38], we know that bootstrap procedures work for Wicksell’s problem, at least when the underlying truth is differentiable at x . We thus compared our methodology with the one proposed in [38] and we observed that they behave very similarly with 1000 or more samples. In Figure 4, the reader can observe an instance with $n = 1400$.

Discussion about uncertainty quantification. The main limitation of the approach in [17] is the lack of uncertainty quantification, which is essential for practical applications. The asymptotic variance of the limiting distribution achieved by the IIE involves two nuisance parameters: the local smoothness γ and the density $g_0(x)$ of the observed data at the evaluation point x . While $g_0(x)$ may be estimable (for most practical applications the rate of convergence would be slower than $\sqrt{n}/\sqrt{\log n}$), estimating γ is particularly challenging, as it requires recovering the local smoothness of the target distribution function F_0 at a specific point x using samples from g_0 —effectively a separate inverse problem. To our knowledge,

no practical methods currently exist for this task. Moreover, because of the method proposed in this paper, estimating γ offers little practical benefit, and thus makes this research direction of limited value. This highlights the key advantage of the Bayesian approach proposed in this paper: it bypasses the need to estimate either γ or $g_0(x)$, as the posterior distribution inherently provides valid uncertainty quantification over a wide range of local smoothness levels, that is, all $\gamma > 1/2$. For $\gamma \leq 1/2$, the inverse problem becomes severely ill-behaved due to the unboundedness of the density of the observations g_0 . This might even lead to a slower rate of convergence both for the naive plug-in estimator and for the IIE. Therefore, for all practically relevant cases with $\gamma > 1/2$, practitioners can apply our method without needing to estimate nuisance parameters or make strong assumptions about the form of F_0 , thanks to its adaptive nature and the automatic uncertainty quantification.

The results displayed in Figure 3 suggest that the bootstrap method of [38] behaves very closely to our method. We can thus make the empirically grounded hypothesis that the bootstrap method proposed in [38] might work under the less stringent assumptions of the current paper. The value of this simulation study has to be taken carefully, as a formal mathematical proof is still needed. However, proving this hypothesis is nontrivial and we therefore leave this as potential future work.

From a practitioner’s perspective, thus the key advantage of our Bayesian methodology lies in its provable robustness across a broad class of smoothness levels: it provides valid uncertainty quantification for all $\gamma > 1/2$. In contrast, the bootstrap approach in [38] is provably consistent only when F_0 is differentiable at x , a substantially stronger condition (stricter than $\gamma = 1$). This distinction is crucial, since determining whether F_0 is differentiable at a given point based solely on samples from g_0 constitutes a challenging inverse problem—both theoretically and in practice. By avoiding this requirement, our method offers a more broadly applicable and practitioner-friendly solution.

4. Conclusion. We have introduced a novel Bayesian approach for nonparametric estimation in Wicksell’s problem, leveraging a Dirichlet process prior on the distribution of observables rather than unobservables. This deviation from the classical Bayesian framework in inverse problems allows for computational efficiency through posterior conjugacy. By projecting the posterior draws onto the $\mathbb{L}_2$ subspace of increasing, right-continuous functions, we obtained the Isotonized Inverse Bayes Posterior (IIP) and we verified that it allows for a Bernstein–von Mises theorem with minimax asymptotic variance, reflecting the Hölder continuity of the true distribution. This proves that our methodology gives asymptotically correct uncertainty quantification under the frequentist point of view and it highlights its potential for Bayesian nonparametrics in inverse problems.

Regarding potential future work, several interesting directions emerge. First, it would be important to verify whether the bootstrap procedure of [38] remains valid for all degrees of smoothness of the underlying truth, not just for F_0 differentiable at x . Second, an open question is whether a Bernstein–von Mises (BvM) result can be obtained for Wicksell’s problem within the classical Bayesian framework—that is, by placing a Dirichlet Process (DP) prior F on F_0 . Formally, obtaining a nonparametric hierarchical Bayesian model for the observations $Z_1, \dots, Z_n$ is as follows:

- (i) A probability distribution F on $\mathbb{R}_+$ is generated from the Dirichlet process prior $\text{DP}(\alpha)$.
- (ii) An i.i.d. sample $X_1, \dots, X_n$ is generated from F .
- (iii) An i.i.d. sample $Y_1, \dots, Y_n$ is generated from a density of a beta distribution $\text{Be}(1, 1/2)$, independent of the other samples $X_1, \dots, X_n$.
- (iv) The observations are $Z_i = Y_i \cdot X_i$, for $i = 1, \dots, n$, with density $g_F(z) := \int_z^\infty \frac{dF(x)}{2\sqrt{x^2 - xz}}$.

In the above model, the posterior is a Mixture of Dirichlet Processes (MDP), with distribution:

$$(4.1) \quad F \mid Z_1, \dots, Z_n \sim \int \text{DP}\left(\alpha + \sum_{i=1}^n \delta_{X_i}\right) dP(X_1, \dots, X_n \mid Z_1, \dots, Z_n),$$

where (4.1) follows directly from Corollary 3.1 in [2]. The authors of the present paper have obtained unpublished preliminary results in this model, showing:

a. denoting by h Hellinger distance, there exists a constant M such that, in G_0 -probability:

$$\Pi_n(F : h(g_F, g_0) \geq Mn^{-1/4}(\log n)^{1/4} \mid Z_1, \dots, Z_n) \xrightarrow{G_0} 0,$$

b. denoting by V_F the posterior of V with the above specified prior, for every $x \geq 0$, there exists a constant M such that, in G_0 -probability:

$$\Pi_n(F : |V_F(x) - V_0(x)| \geq Mn^{-1/4}(\log n)^{3/4} \mid Z_1, \dots, Z_n) \xrightarrow{G_0} 0.$$

The contraction rate $n^{-1/4}$ for the Hellinger distance is equal to the $\mathbb{L}_2$ -minimax rate for estimating a density in a Sobolev space of order $1/2$ (see [14]). Because the densities g_0 are in $H^{1/2}[0, \infty)$ if F_0 possesses a bounded density, this suggests that the rate $n^{-1/4}$ is close to optimal. The pointwise contraction rate for the functional V_0 follows from the contraction rate in Hellinger distance, combined with a proof technique based on a Laplace transform rewriting of the posterior as in equation (2.13) in [8], a local perturbation and a LAN expansion as in [17]. This perturbation, which is believed to correspond to the efficient influence function for the functional of interest, is essential for handling arguments involving contraction rates. However, the local nature of the perturbation in Wicksell's problem comes at the cost of contracting at the Hellinger distance rate h , which is too slow (and not $h^2 = O_p(\sqrt{\log n}/\sqrt{n})$ as in the proof style of [8]). Even though the authors have made significant efforts to overcome this issue, this challenge remains unsolved.

We conclude by remarking that the work of the present paper introduces a fresh perspective on Bayesian inverse problems by employing the concept of the *projection posterior*, first proposed in [9]. The methods proposed here are also significantly computationally faster if compared to the standard hierarchical Bayes model illustrated above, because they do not rely on computationally expensive Gibbs samplers. Finally, a valuable direction for future research is to identify other problems where this approach could be effectively utilized.

APPENDIX A: COMPLEMENTARY RESULTS

Measure-theoretical framework. In this paper, the sample space is $(\mathbb{R}^+, \mathcal{B}(\mathbb{R}^+))$, and $\mathfrak{M}$ (with Borel σ -field $\mathcal{M}$ for the weak topology) is the collection of all Borel probability measures on $(\mathbb{R}^+, \mathcal{B}(\mathbb{R}^+))$. The Dirichlet process prior is a probability measure on $\mathfrak{M}$. A prior on the set of probability measures $\mathfrak{M}$ is a probability measure on a sigma-field $\mathcal{M}$ of subsets of $\mathfrak{M}$. Alternatively, it can be viewed as map from some probability space $(\Omega, \mathcal{U}, \text{Pr})$ into $(\mathfrak{M}, \mathcal{M})$. We can think of the hierarchy of the spaces as follows: $(\Omega, \mathcal{U}, \text{Pr}) \rightarrow (\mathfrak{M}, \mathcal{M}) \rightarrow (\mathbb{R}^+, \mathcal{B}(\mathbb{R}^+))$. We put a Dirichlet process prior over G and write $G \sim \Pi = \text{DP}(\alpha)$ for some base measure α and prior precision $|\alpha|$. This means that for any measurable set $B \in \mathcal{M}$ we have $\Pi(B) = \text{Pr}(G \in B)$. Throughout the paper, we use the notations P and E , which should be understood as the outer probability measure and outer expectation, respectively, whenever issues of measurability in the considered processes arise (see [40] for a rigorous treatment of these topics).

Furthermore, we use the notion of conditional convergence in distribution. This happens if the bounded Lipschitz distance between the conditional law of Borel measurable X_n given

$\mathcal{B}_n$ and a Borel probability measure L tends to zero, where the convergence can be in outer probability. The sequence X_n tends to L conditionally given $\mathcal{B}_n$ in outer probability if

$$(A.1) \quad d_{BL}(\mathcal{L}(X_n | \mathcal{B}_n), L) := \sup_{f \in BL_1} \left| \mathbb{E}(f(X_n) | \mathcal{B}_n) - \int f dL \right| \xrightarrow{P} 0,$$

where BL_1 is the space of bounded Lipschitz functions with Lipschitz constant 1.

In particular if $Z_1, \dots, Z_n \stackrel{\text{i.i.d.}}{\sim} P_Z$, then $X_n | Z_1, \dots, Z_n \rightsquigarrow N(0, \sigma^2)$ means

$$(A.2) \quad d_{BL}(P(X_n \in \cdot | Z_1, \dots, Z_n), N(0, \sigma^2)) \xrightarrow{P_Z} 0.$$

LEMMA A.1. (2.18) implies (2.15).

PROOF. By Theorem 1.13.1(i) and (ii) in [40], the claim (2.15) is obtained as soon as we show (cf. (A.1)), for $X \sim N(0, \frac{g_0(x)}{2\gamma})$:

$$\sup_{f \in BL_1} \left| \mathbb{E} \left(f \left(\frac{\sqrt{n}}{\sqrt{\log n}} (\hat{V}_G(x) - \hat{V}_n(x)) \right) | Z_1, \dots, Z_n \right) - \mathbb{E}(f(X)) \right| \xrightarrow{G_0} 0.$$

By Lemma 2.1 given in the Supplementary Material [18], for any bounded 1-Lipschitz function $f: \mathbb{R} \mapsto \mathbb{R}$ and $Z_n^{G_0}$ as in (2.16):

$$\begin{aligned} & \left| \mathbb{E} \left[f \left(\frac{\sqrt{n}}{\sqrt{\log n}} (\hat{V}_G(x) - \hat{V}_n(x)) \right) - f \left(\frac{\sqrt{n}}{\sqrt{\log n}} (\hat{V}_G(x) - V_0(x)) - Z_n^{G_0}(1) \right) | Z_1, \dots, Z_n \right] \right| \\ & \leq \left| \frac{\sqrt{n}}{\sqrt{\log n}} (\hat{V}_n(x) - V_0(x)) - Z_n^{G_0}(1) \right| \xrightarrow{G_0} 0. \end{aligned}$$

This means that the claim is obtained as soon as we show

$$\sup_{f \in BL_1} \left| \mathbb{E} \left(f \left(\frac{\sqrt{n}}{\sqrt{\log n}} (\hat{V}_G(x) - V_0(x)) - Z_n^{G_0}(1) \right) | Z_1, \dots, Z_n \right) - \mathbb{E}(f(X)) \right| \xrightarrow{G_0} 0.$$

By Theorem 1.13.1(iii) in [40], the last display is implied by (2.18). $\square$

PROOF OF (2.24). First, note that by assumption Qf is a.s. well defined (cf. Remark 4.4 in [16]). Moreover, note that $\mathbb{E}V_n = |\alpha|/(|\alpha| + n)$ and $|f_n^t(Z)| \leq 2\sqrt{|t|/\delta_n^*}$ and $\forall \varepsilon > 0$, by independence, that

$$\begin{aligned} & P(\delta_n^{-1} |V_n(Q - \mathbb{G}_n)f_t^n| > \varepsilon | Z_1, \dots, Z_n) \\ & \leq \delta_n^{-1} \varepsilon^{-1} \mathbb{E}V_n \mathbb{E}[(Q + \mathbb{G}_n)|f_t^n| | Z_1, \dots, Z_n] \\ & \leq 4\delta_n^{-1} \varepsilon^{-1} \frac{|\alpha|}{|\alpha| + n} \frac{\sqrt{|t|}}{\sqrt{\delta_n^*}} = O(\delta_n^{-1} n^{-1} (\delta_n^*)^{-1/2}). \end{aligned}$$

As $\gamma > \frac{1}{2}$, note that $\delta_n^{-1} n^{-1} (\delta_n^*)^{-1/2} = n^{-1} (n/\log n)^{\frac{1}{2} + \frac{1}{4\gamma}} \ll n^{-1} (\sqrt{n}/\sqrt{\log n})^2 \rightarrow 0$. Similarly, $P(\delta_n^{-1} |V_n(\mathbb{B}_n - \mathbb{G}_n)f_t^n| > \varepsilon | Z_1, \dots, Z_n) \xrightarrow{G_0} 0$.

By Proposition G.2 in [16] for $\varepsilon_i \stackrel{\text{i.i.d.}}{\sim} \text{Exp}(1)$ and $W_{ni} := \varepsilon_i/\bar{\varepsilon}_n$, we have $(W_{n1}, \dots, W_{nn}) \sim \text{Dir}(n; 1/n, \dots, 1/n)$. Since $\mathbb{B}_n \sim \text{DP}(n\mathbb{G}_n)$, for δ_{Z_i} Dirac measure at Z_i , we have the following representation in distribution of $\mathbb{B}_n$ (cf. Example 3.7.9 in [40]):

$$\mathbb{B}_n = \frac{1}{n} \sum_{i=1}^n W_{ni} \delta_{Z_i}.$$

Because of this representation (cf. Section 3.7.2. in [40]), $\mathbb{B}_n$ is also known as *Bayesian bootstrap*. Conclude that

$$\frac{\sqrt{n}}{\sqrt{\log n}}(\mathbb{B}_n - \mathbb{G}_n)f_t^n = \frac{1}{\sqrt{n \log n}} \sum_{i=1}^n (W_{ni} - 1)f_t^n(Z_i).$$

Because $|\bar{\varepsilon}_n - 1| = O_p(1/\sqrt{n})$, for almost every sequence $Z_1, \dots, Z_n$ conclude that, uniformly in t in compacta,

$$(A.3) \quad \frac{1}{\sqrt{n \log n}} \sum_{i=1}^n (W_{ni} - 1)f_t^n(Z_i) = \frac{1}{\sqrt{n \log n}} \sum_{i=1}^n (\varepsilon_i - 1)f_t^n(Z_i) + o_p(1),$$

conclude by recalling the definition of $\mathbb{Z}_n^*$ in (2.25). $\square$

LEMMA A.2. *The process $\mathbb{Z}_n^*$ is conditionally asymptotically equicontinuous, that is, $\forall \eta > 0$ as $n \rightarrow \infty$ followed by $\delta > 0$:*

$$\mathbb{P} \left\{ \sup_{|s-t|<\delta} |\mathbb{Z}_n^*(s) - \mathbb{Z}_n^*(t)| > \eta \mid Z_1, \dots, Z_n \right\} \xrightarrow{G_0} 0.$$

PROOF. Take the class of functions for $\delta > 0, M > 0$:

$$(A.4) \quad \mathcal{F}_{n,\delta}^M := \{(\log n)^{-1/2}(\varepsilon - 1)(f_s^n - f_t^n)(z) : s, t \in I_x, |s - t| < \delta, \max(|s|, |t|) < M\}$$

and the class of functions $\mathcal{H}_{n,\delta}^M$, whose functions for $f \in \mathcal{F}_{n,\delta}^M$ are given by $f/(\varepsilon - 1)$, thus without the $\varepsilon - 1$. These classes have square integrable natural envelopes $F_{n,\delta}$ and $H_{n,\delta}$ that satisfy, for all $\delta > 0$ small (cf. Lemma 5 in [17]),

$$\|F_{n,\delta}\|_2^2 = \mathbb{E}(\varepsilon - 1)^2 \|H_{n,\delta}\|_{2,G_0}^2 = \|H_{n,\delta}\|_{2,G_0}^2 \leq g_0(x + \delta_n^* s \wedge t) \delta^2 (\log n)^{-1} (\log(\delta \delta_n^*)^{-2}).$$

For fixed $s, t \in I_x$, we prove conditional asymptotic equicontinuity of $\mathbb{Z}_n^*$, that is, $\forall \eta > 0$ as $n \rightarrow \infty$ followed by $\delta > 0$, the right-hand side of the display below converges to 0 under G_0 :

$$\mathbb{P} \left\{ \sup_{|s-t|<\delta} |\mathbb{Z}_n^*(s) - \mathbb{Z}_n^*(t)| > \eta \mid Z_1, \dots, Z_n \right\} \leq \frac{1}{\eta} \mathbb{E} \left[\sup_{|s-t|<\delta} |\mathbb{Z}_n^*(s) - \mathbb{Z}_n^*(t)| \mid Z_1, \dots, Z_n \right].$$

By the maximal inequality in Theorem 2.14.1 in [40],

$$\mathbb{E} \left[\sup_{|s-t|<\delta} |\mathbb{Z}_n^*(s) - \mathbb{Z}_n^*(t)| \right] \lesssim J(1) \|F_{n,\delta}\|_2 \lesssim \sqrt{\|g_0\|_\infty \delta^2 (\log n)^{-1} (\log(\delta \delta_n^*)^{-2})} \rightarrow 0,$$

where $J(1) := J(1, \mathcal{F}_{n,\delta}^M, \mathbb{L}_2) = \sup_Q \int_0^1 \sqrt{\log N(\eta \|F_{n,\delta}\|_2, \mathcal{F}_{n,\delta}^M, \mathbb{L}_2(Q))} d\eta$ and where the square $\mathbb{L}_2(Q)$ -distance is given for s, t and $s', t' \in I_x$ by: $\frac{1}{\log n} \int ((f_s^n - f_t^n)(z) - (f_{s'}^n - f_{t'}^n)(z))^2 dQ(z)$. (Note that we did not take $\varepsilon - 1$ into account as $\mathbb{E}(\varepsilon - 1)^2 = 1$). Furthermore, we used that $J(1, \mathcal{F}_{n,\delta}^M, \mathbb{L}_2) < \infty$ as a consequence of Lemma 2.1 in [17], where it is shown that the function classes under consideration are VC of uniform bounded index. $\square$

In the following lemma, we make use of the concept of weak convergence of a process in $\ell^\infty(K)$ for every compact $K \subset \mathbb{R}$ and ℓ^∞ being the space of bounded functions on K equipped with the supremum norm. For a review of the conditional weak convergence in ℓ^∞ , see [40] (cf. Example 1.5.1 and Theorem 1.5.4).

LEMMA A.3 (Conditional argmax continuous mapping). *Let $\mathbb{Z}_n, \mathbb{Z}$ be stochastic processes indexed by $H \subseteq \mathbb{R}$ such that, in probability, given a sequence of random variables $Z_1, \dots, Z_n$,*

$$\mathbb{Z}_n \mid Z_1, \dots, Z_n \rightsquigarrow \mathbb{Z} \quad \text{in} \quad \ell^\infty(K)$$

for every compact $K \subset H$. Suppose that almost all sample paths $t \mapsto \mathbb{Z}(t)$ are upper semi-continuous and possess a unique maximum at a (random) point $\hat{t}$, which as a random map in H is tight. If $\forall n \in \mathbb{N}$, $\mathbb{Z}_n(\hat{t}_n) = \sup_t \mathbb{Z}_n(t)$ a.s. and there exist a compact set $\tilde{K} \subset H$ such that, under $\mathbb{P}$,

$$\mathbb{P}(\hat{t}_n \notin \tilde{K} \mid Z_1, \dots, Z_n) \xrightarrow{\mathbb{P}} 0,$$

then $\hat{t}_n \mid Z_1, \dots, Z_n \rightsquigarrow \hat{t}$ in probability in H .

PROOF. By the continuous mapping theorem, for every closed set F and A, B arbitrary sets of H : $\sup_{t \in F \cap A} \mathbb{Z}_n(t) - \sup_{t \in B} \mathbb{Z}_n(t)$ converges conditionally on $Z_1, \dots, Z_n$ in distribution to $\sup_{t \in F \cap A} \mathbb{Z}(t) - \sup_{t \in B} \mathbb{Z}(t)$, where $\mathbb{Z}(t)$ has a.s. continuous sample paths and a.s. unique maxima. Thus, using the assumption on $\hat{t}_n$,

$$\begin{aligned} \limsup_{n \rightarrow \infty} \mathbb{P}(\hat{t}_n \in F \cap A \mid Z_1, \dots, Z_n) &\leq \limsup_{n \rightarrow \infty} \mathbb{P}\left(\sup_{t \in F \cap A} \mathbb{Z}_n(t) > \sup_{t \in B} \mathbb{Z}_n(t) \mid Z_1, \dots, Z_n\right) \\ &= \mathbb{P}\left(\sup_{t \in F \cap A} \mathbb{Z}(t) > \sup_{t \in B} \mathbb{Z}(t)\right) \leq \mathbb{P}(\hat{t} \in F \cup B^c), \end{aligned}$$

where in the last inequality we use that $\mathbb{Z}(\hat{t}) > \sup_{t \notin F^c, t \in A} \mathbb{Z}(t)$, and thus the event

$$\left\{ \sup_{t \in F \cap A} \mathbb{Z}(t) > \sup_{t \in B} \mathbb{Z}(t) \right\} \subseteq \{\hat{t} \in F \cup B^c\}.$$

Because $\mathbb{Z}_n$ converges in distribution conditionally on $Z_1, \dots, Z_n$ to $\mathbb{Z}$ in $\ell^\infty(K)$ for every compact K , take now $A = B = K$. Then for every open set G containing $\hat{t}$, it holds

$$\mathbb{Z}(\hat{t}) > \sup_{t \notin G, t \in K} \mathbb{Z}(t),$$

because of the continuity of $\mathbb{Z}$ and the uniqueness of $\hat{t}$. Thus, the above yields

$$\begin{aligned} \limsup_{n \rightarrow \infty} \mathbb{P}(\hat{t}_n \in F \mid Z_1, \dots, Z_n) &\leq \limsup_{n \rightarrow \infty} \mathbb{P}(\hat{t}_n \in F \cap K \mid Z_1, \dots, Z_n) \\ &\quad + \limsup_{n \rightarrow \infty} \mathbb{P}(\hat{t}_n \in K^c \mid Z_1, \dots, Z_n) = \mathbb{P}(\hat{t} \in F \cup K^c). \end{aligned}$$

Note: $\mathbb{P}(\hat{t} \in F \cup K^c) \leq \mathbb{P}(\hat{t} \in F) + \mathbb{P}(\hat{t} \in K^c)$, where by the stochastic boundedness of $\hat{t}_n \mid Z_1, \dots, Z_n$ and the choice of K^c , the terms $\mathbb{P}(\hat{t} \in K^c)$ can be made arbitrarily small and $\limsup_{n \rightarrow \infty} \mathbb{P}(\hat{t}_n \in K^c \mid Z_1, \dots, Z_n)$ converges to 0. $\square$

LEMMA A.4 (Consistency). *Assume (2.13) holds true. The IIP is consistent, that is, $\forall \varepsilon > 0$ as $n \rightarrow \infty$,*

$$(A.5) \quad \Pi_n(|\hat{V}_G(x) - V_0(x)| > \varepsilon \mid Z_1, \dots, Z_n) \xrightarrow{G_0} 0.$$

Moreover,

$$(A.6) \quad \Pi_n(|T_G(V_0(x) + \delta_n(a + Z_n^{G_0}(1))) - x| > \varepsilon \mid Z_1, \dots, Z_n) \xrightarrow{G_0} 0.$$

PROOF. We begin by showing (A.5). Let $\psi_x(z) := 2\sqrt{z} - 2\sqrt{(z-x)_+}$, using Proposition 4.3 in [16] with ψ_x we obtain

$$\begin{aligned} \mathbb{E}(U_G(x) \mid Z_1, \dots, Z_n) &= \frac{1}{|\alpha| + n} \int \psi_x d\left(\alpha + \sum_{i=1}^n \delta_{Z_i}\right), \\ \text{Var}(U_G(x) \mid Z_1, \dots, Z_n) &\lesssim \frac{1}{1 + |\alpha| + n} \frac{1}{|\alpha| + n} \int \psi_x^2 d\left(\alpha + \sum_{i=1}^n \delta_{Z_i}\right). \end{aligned}$$

Thus, using the conditional Chebyshev inequality,

$$\Pi_n(|U_G(x) - U_0(x)| > \varepsilon \mid Z_1, \dots, Z_n) \rightarrow 0, \quad \text{a.s.}$$

Note now that the functions U_G are nondecreasing, U_0 is bounded continuous nondecreasing, thus $\forall \varepsilon > 0$ by Exercise 3.2 in [23] we conclude

$$\Pi_n(\|U_G - U_0\|_\infty > \varepsilon \mid Z_1, \dots, Z_n) \rightarrow 0, \quad \text{a.s.}$$

Thus, by Marshall's inequality (cf. Exercise 3.1 in [23]), with U_G^* least concave majorant of U_G ,

$$\Pi_n(\|U_G^* - U_0\|_\infty > \varepsilon \mid Z_1, \dots, Z_n) \xrightarrow{G_0} 0.$$

Now fix $\varepsilon > 0$ and $h > 0$ such that $|(U_0(x) - U_0(x-h))/h - V_0(x)| < \varepsilon$. By concavity of U_G^* , we have

$$V_G(x) \leq \frac{U_G^*(x) - U_G^*(x-h)}{h}.$$

Thus, by Lemma 3.1 in [23],

$$\begin{aligned} \Pi_n(V_G(x) \geq V_0(x) + 2\varepsilon \mid Z_1, \dots, Z_n) \\ \leq \Pi_n\left(\frac{U_G^*(x) - U_G^*(x-h)}{h} \geq \frac{U_0(x) - U_0(x-h)}{h} + \varepsilon \mid Z_1, \dots, Z_n\right) \rightarrow 0, \quad \text{a.s.} \end{aligned}$$

Similarly, $V_G(x) \geq (U_G^*(x+h) - U_G^*(x))/h$ gives $\Pi_n(V_G(x) \leq V_0(x) - 2\varepsilon \mid Z_1, \dots, Z_n) \rightarrow 0$, a.s. This being true for every $\varepsilon > 0$ gives (A.5).

From (A.5) follows consistency of T_G ; by the switch relation, $\forall \varepsilon > 0$,

$$\begin{aligned} \Pi_n(|T_G(V_0(x) + \delta_n(a + Z_n^{G_0}(1))) - x| > \varepsilon \mid Z_1, \dots, Z_n) \\ \leq \Pi_n(T_G(V_0(x) + \delta_n(a + Z_n^{G_0}(1))) > x + \varepsilon \mid Z_1, \dots, Z_n) \\ \quad + \Pi_n(T_G(V_0(x) + \delta_n(a + Z_n^{G_0}(1))) < x - \varepsilon \mid Z_1, \dots, Z_n) \\ \leq \Pi_n(\hat{V}_G(x + \varepsilon) > V_0(x) + \delta_n(a + Z_n^{G_0}(1)) \mid Z_1, \dots, Z_n) \\ \quad + \Pi_n(\hat{V}_G(x - \varepsilon) < V_0(x) + \delta_n(a + Z_n^{G_0}(1)) \mid Z_1, \dots, Z_n) \xrightarrow{G_0} 0, \end{aligned} \tag{A.7}$$

where by (A.5) above $\hat{V}_G(x \pm \varepsilon) \mid Z_1, \dots, Z_n \rightarrow V_0(x \pm \varepsilon)$, $Z_n^{G_0}(1) \rightsquigarrow \mathbb{Z}(1)$ and $\delta_n \downarrow 0$, thus $\delta_n(a + Z_n^{G_0}(1)) \xrightarrow{G_0} 0$ and V_0 is strictly decreasing at x , thus we have the strict inequalities $V_0(x + \varepsilon) < V_0(x) < V_0(x - \varepsilon)$. $\square$

PROOF OF (2.29). The claim is implied by the unconditional version of the convergence given in (2.29) and this entails proving that the rate for $|T_G(V_0(x) + \delta_n(a + Z_n^{G_0}(1))) - x|$ is $(\delta_n^*)^{-1}$. Because of Lemma A.4, we can proceed by verifying the rest of the conditions of Theorem 3.2.5 in [40] (or see Theorem 3.4.1 for a more general version).

The argmax process we are interested in is $\tilde{T}_n := \arg \max_t \{\tilde{Z}_n^G(t) - at\}$ where $\tilde{Z}_n^G$ is given in (2.20)–(2.21). Thus, note that we can rewrite it as $\tilde{T}_n = (\delta_n^*)^{-1} \arg \max_t \{\mathbb{M}_n(t)\}$, where

$$(A.8) \quad \mathbb{M}_n(t) := 2 \int \left(\sqrt{(z-x)_+} - \sqrt{(z-x-t)_+} - \frac{1}{2} V_0(x)t \right) d(G - \mathbb{G}_n + G_0)(z)$$

$$(A.9) \quad - 2 \frac{t}{\delta_n^*} \int \left(\sqrt{(z-x)_+} - \sqrt{(z-x-\delta_n^*)_+} \right) d(\mathbb{G}_n - G_0)(z)$$

$$(A.10) \quad + 2 \int \left(\sqrt{(z-x)_+} - \sqrt{(z-x-t)_+} \right) d(\mathbb{G}_n - G_0)(z) - at\delta_n.$$

Moreover, define the deterministic function

$$\mathbb{M}(t) := 2 \int \left(\sqrt{(z-x)_+} - \sqrt{(z-x-t)_+} - \frac{1}{2} V_0(x)t \right) dG_0(z).$$

Note that $\mathbb{M}_n(0) = \mathbb{M}(0) = 0$. In [17] (cf. proof of (18)), it is shown that $\exists \delta_0 > 0$ such that $\forall |t| < \delta_0$

$$(A.11) \quad \mathbb{M}(t) - \mathbb{M}(0) \leq -K|t|^{\gamma+1}.$$

Following Theorem 3.2.5 in [40], define thus the metric $d(s, t) := |s - t|^{\frac{\gamma+1}{2}}$.

Consider the class of functions $h_t : [0, \infty) \rightarrow \mathbb{R}$

$$\mathcal{H} := \{h_t : h_t(z) = 2\sqrt{(z-x)_+} - 2\sqrt{(z-x-t)_+} - V_0(x)t, t \geq 0\}$$

and, for $\eta > 0$, its subclass $\mathcal{H}_\eta := \{h_t \in \mathcal{H} : |t| \leq \eta\}$. In [22], it is shown that the envelope of $\mathcal{H}_\eta$, $H_\eta(z) := 2\sqrt{(z-x)_+} - 2\sqrt{(z-x-\eta)_+} - V_0(x)\eta$, satisfies for all η small enough (cf. (11) in [22]):

$$(A.12) \quad \int H_\eta^2(z) g_0(z) dz \leq -2g_0(x)\eta^2 \log(\eta).$$

Consider now the class of functions $\tilde{\mathcal{H}}_n := \{\tilde{h}_t^n : t \geq 0\}$ where, for any fixed $t \geq 0$,

$$\tilde{h}_t^n(z) := 2 \left(\frac{t}{\delta_n^*} - 1 \right) \sqrt{(z-x)_+} - 2 \frac{t}{\delta_n^*} \sqrt{(z-x-\delta_n^*)_+} + 2\sqrt{(z-x-t)_+},$$

and, for $\eta > 0$, its subclass $\tilde{\mathcal{H}}_{n,\eta} := \{\tilde{h}_t^n \in \tilde{\mathcal{H}}_n : |t| \leq \eta\}$. Using the same computations as in [22], it follows that the envelope of $\tilde{\mathcal{H}}_{n,\eta}$, $\tilde{H}_{n,\eta}(z) := \sup_{|t| \leq \eta} |\tilde{h}_t^n(z)|$, satisfies for all η small enough:

$$(A.13) \quad \int \tilde{H}_{n,\eta}^2(z) g_0(z) dz \leq -2g_0(x)\eta^2 (\log(\eta) + \log(\delta_n^*)).$$

Using the representation (2.23), we have

$$\begin{aligned} \mathbb{M}_n(t) - \mathbb{M}(t) &= (1 - V_n) \int h_t(z) d(\mathbb{B}_n - \mathbb{G}_n)(z) \\ &\quad + V_n \int h_t(z) d(Q - \mathbb{G}_n)(z) - at\delta_n \\ &\quad - \int \tilde{h}_t^n(z) d(\mathbb{G}_n - G_0)(z), \end{aligned}$$

and, by applying the maximal inequality in Theorem 2.14.1 in [40] for sufficiently small $\delta > 0$, we conclude $(d(t, 0) < \delta \iff |t| < \delta^{\frac{2}{1+\gamma}})$

$$\begin{aligned} \mathbb{E} \left[\sup_{|t| < \delta^{\frac{2}{1+\gamma}}} \sqrt{n} |(\mathbb{M}_n - \mathbb{M})(t)| \right] &\lesssim J(1, \mathcal{H}_{\delta^{\frac{2}{1+\gamma}}}^{\frac{2}{1+\gamma}}, \mathbb{L}_2) \sqrt{\int H_{\delta^{\frac{2}{1+\gamma}}}^2 dG_0} + \sqrt{n} \frac{1}{n} \delta^{\frac{2}{1+\gamma}} \\ &\quad + |a| \delta^{\frac{2}{1+\gamma}} \delta_n \sqrt{n} + J(1, \tilde{\mathcal{H}}_{n, \delta^{\frac{2}{1+\gamma}}}^{\frac{2}{1+\gamma}}, \mathbb{L}_2) \sqrt{\int \tilde{H}_{n, \delta^{\frac{2}{1+\gamma}}}^2 dG_0} \\ &\lesssim \delta^{\frac{2}{1+\gamma}} (\sqrt{\log \delta} + \sqrt{\log n}). \end{aligned}$$

The first term after applying the maximal inequality comes from the $(1 - V_n)(\mathbb{B}_n - \mathbb{G}_n)$ -term in the representation above of $\mathbb{M}_n - \mathbb{M}$; the second one comes from the $V_n(Q - \mathbb{G}_n)$ -term, the third one from the $at\delta_n$ -term and the last one from the $\mathbb{G}_n - G_0$ -term (having used also for this one the maximal inequality). The last inequality is just a consequence of (A.12), (A.13) and the facts that $J(1, \mathcal{H}_{\delta^{\frac{2}{1+\gamma}}}^{\frac{2}{1+\gamma}}, \mathbb{L}_2) < \infty$ by the proof of Lemma 2.1 in [17] and by the same reasonings $J(1, \tilde{\mathcal{H}}_{n, \delta^{\frac{2}{1+\gamma}}}^{\frac{2}{1+\gamma}}, \mathbb{L}_2)$ is uniformly bounded.

The above computations and Theorem 3.2.5 in [40] imply that the rate for d is determined as follows: $\delta^{\frac{2}{1+\gamma}} (\sqrt{\log \delta} + \sqrt{\log n}) \leq \sqrt{n} \delta^2$, hence $\delta \gtrsim (\frac{\log n}{n})^{\frac{1+\gamma}{4\gamma}}$. Because $d(s, t) = |s - t|^{\frac{\gamma+1}{2}}$, the rate for the distance $|\cdot|$ is given by

$$\left(\frac{\log n}{n} \right)^{\frac{1+\gamma}{4\gamma} \cdot \frac{2}{1+\gamma}} = \left(\frac{\log n}{n} \right)^{\frac{1}{2\gamma}} \equiv \delta_n^*.$$

Altogether the above proves that the rate for $\arg \max_t \{\mathbb{M}_n(t)\}$ is $(\delta_n^*)^{-1}$. $\square$

Acknowledgments. We thank the anonymous reviewers and the Associate Editor for their thoughtful comments and constructive suggestions, which have helped us improve the clarity and quality of this work.

SUPPLEMENTARY MATERIAL

Supplement to the article: “Semiparametric Bernstein–von Mises phenomenon via isotonized posterior in Wicksell’s problem” (DOI: [10.1214/25-AOS2571SUPP](https://doi.org/10.1214/25-AOS2571SUPP); .pdf). Available at [18].

REFERENCES

- [1] ALONSO, J., MATIO, M. and SEN, B. (2007). Uncloaking globular clusters of the inner galaxy. *Proceedings of the International Astronomical Union* **3** 359–360. <https://doi.org/10.1088/0004-6256/143/3/70>
- [2] ANTONIAK, C. (1974). Mixtures of Dirichlet processes with applications to Bayesian nonparametric problems. *The Annals of Statistics* **2** 1152–1174. <https://doi.org/10.1214/aos/1176342871>
- [3] AYER, M., BRUNK, H. D., EWING, G. M., REID, W. T. and SILVERMAN, E. (1955). An empirical distribution function for sampling with incomplete information. *Ann. Math. Statist.* **26** 641–647. [MR0073895 https://doi.org/10.1214/aoms/1177728423](https://doi.org/10.1214/aoms/1177728423)
- [4] BHAUMIK, P. and GHOSAL, S. (2015). Bayesian two-step estimation in differential equation models. *Electron. J. Stat.* **9** 3124–3154. [MR3453972 https://doi.org/10.1214/15-EJS1099](https://doi.org/10.1214/15-EJS1099)
- [5] BHAUMIK, P. and GHOSAL, S. (2017). Efficient Bayesian estimation and uncertainty quantification in ordinary differential equation models. *Bernoulli* **23** 3537–3570. [MR3654815 https://doi.org/10.3150/16-BEJ856](https://doi.org/10.3150/16-BEJ856)
- [6] CASTILLO, I. and NICKL, R. (2013). Nonparametric Bernstein-von Mises theorems in Gaussian white noise. *Ann. Statist.* **41** 1999–2028. [MR3127856 https://doi.org/10.1214/13-AOS1133](https://doi.org/10.1214/13-AOS1133)

- [7] CASTILLO, I. and NICKL, R. (2014). On the Bernstein-von Mises phenomenon for nonparametric Bayes procedures. *Ann. Statist.* **42** 1941–1969. [MR3262473](#) <https://doi.org/10.1214/14-AOS1246>
- [8] CASTILLO, I. and ROUSSEAU, J. (2015). A Bernstein–von Mises theorem for smooth functionals in semi-parametric models. *Ann. Statist.* **43** 2353–2383. [MR3405597](#) <https://doi.org/10.1214/15-AOS1336>
- [9] CHAKRABORTY, M. and GHOSAL, S. (2021). Coverage of credible intervals in nonparametric monotone regression. *Ann. Statist.* **49** 1011–1028. [MR4255117](#) <https://doi.org/10.1214/20-aos1989>
- [10] CHAKRABORTY, M. and GHOSAL, S. (2022). Rates and coverage for monotone densities using projection-posterior. *Bernoulli* **28** 1093–1119. [MR4388931](#) <https://doi.org/10.3150/21-bej1379>
- [11] CHAN, K. C. G. and QIN, J. (2016). Nonparametric maximum likelihood estimation for the multisample Wicksell corpuscle problem. *Biometrika* **103** 273–286. [MR3509886](#) <https://doi.org/10.1093/biomet/asw011>
- [12] CUZZI, J. and OLSON, D. (2017). Recovering 3D particle size distributions from 2D sections. *Meteoritics & Planetary Science* **52** 532–545. <https://doi.org/10.1111/maps.12812>
- [13] DENG, H., HAN, Q. and ZHANG, C.-H. (2021). Confidence intervals for multiple isotonic regression and other monotone models. *Ann. Statist.* **49** 2021–2052. [MR4319240](#) <https://doi.org/10.1214/20-aos2025>
- [14] EFROMOVICH, S. (2010). A sharp minimax lower bound for the nonparametric estimation of Sobolev densities of order $\frac{1}{2}$. *Statist. Probab. Lett.* **80** 77–81. [MR2563888](#) <https://doi.org/10.1016/j.spl.2009.08.011>
- [15] FERGUSON, T. S. (1974). Prior distributions on spaces of probability measures. *Ann. Statist.* **2** 615–629. <https://doi.org/10.1214/aos/1176342752>
- [16] GHOSAL, S. and VAN DER VAART, A. (2017). *Fundamentals of Nonparametric Bayesian Inference*. Cambridge Series in Statistical and Probabilistic Mathematics **44**. Cambridge Univ. Press, Cambridge. [MR3587782](#) <https://doi.org/10.1017/9781139029834>
- [17] GILI, F., JONGBLOED, G. and VAN DER VAART, A. (2025). Adaptive and efficient isotonic estimation in Wicksell’s problem. *J. Nonparametr. Stat.* **37** 491–531. [MR4948920](#) <https://doi.org/10.1080/10485252.2024.2397680>
- [18] GILI, F., JONGBLOED, G. and VAN DER VAART, A. (2026). Supplement to “Semiparametric Bernstein–von Mises phenomenon via Isotonized Posterior in Wicksell’s problem.” <https://doi.org/10.1214/25-AOS2571SUPP>
- [19] GILI, F., JONGBLOED, G. and VAN DER VAART, A. (2026). Asymptotically efficient estimation under local constraint in Wicksell’s problem. *J. Statist. Plann. Inference* **240** Paper No. 106299, 20. [MR4908067](#) <https://doi.org/10.1016/j.jspi.2025.106299>
- [20] GOLUBEV, G. K. and ENIKEEVA, F. N. (2001). Asymptotically efficient smoothing in the Wicksell problem under squared losses. *Probl. Inf. Transm.* **37** 12–21. <https://doi.org/10.1023/A:1010495609901>
- [21] GOLUBEV, G. K. and LEVIT, B. Y. (1998). Asymptotically efficient estimation in the Wicksell problem. *Ann. Statist.* **26** 2407–2419. [MR1700238](#) <https://doi.org/10.1214/aos/1024691477>
- [22] GROENEBOOM, P. and JONGBLOED, G. (1995). Isotonic estimation and rates of convergence in Wicksell’s problem. *The Annals of Statistics* **23** 1528–1542. <https://doi.org/10.1214/aos/1176324310>
- [23] GROENEBOOM, P. and JONGBLOED, G. (2014). *Nonparametric Estimation Under Shape Constraints: Estimators, Algorithms and Asymptotics*. Cambridge Series in Statistical and Probabilistic Mathematics **38**. Cambridge Univ. Press, New York. [MR3445293](#) <https://doi.org/10.1017/CBO9781139020893>
- [24] HALL, P. and SMITH, R. L. (1988). The kernel method for unfolding sphere size distributions. *J. Comput. Phys.* **74** 409–421. [MR0930210](#) [https://doi.org/10.1016/0021-9991\(88\)90085-X](https://doi.org/10.1016/0021-9991(88)90085-X)
- [25] HJORT, N. L. and PETRONE, S. (2007). Nonparametric quantile inference using Dirichlet processes. In *Advances in Statistical Modeling and Inference. Ser. Biostat.* **3** 463–492. World Scientific, Hackensack, NJ. [MR2416129](#) https://doi.org/10.1142/9789812708298_0023
- [26] JONGBLOED, G. (2001). Sieved maximum likelihood estimation in Wicksell’s problem and related deconvolution problems. *Scand. J. Statist.* **28** 161–183. [MR1844355](#) <https://doi.org/10.1111/1467-9469.00230>
- [27] KIM, J. and POLLARD, D. (1990). Cube root asymptotics. *Ann. Statist.* **18** 191–219. <https://doi.org/10.1214/aos/1176347498>
- [28] KIM, Y. (2006). The Bernstein–von Mises theorem for the proportional hazard model. *Ann. Statist.* **34** 1678–1700. <https://doi.org/10.1214/009053606000000533>
- [29] KIM, Y. and LEE, J. (2004). A Bernstein-von Mises theorem in the nonparametric right-censoring model. *Ann. Statist.* **32** 1492–1512. [MR2089131](#) <https://doi.org/10.1214/009053604000000526>
- [30] LOPEZ-SANCHEZ, M. A. and LLANA-FÚNEZ, S. (2016). An extension of the saltykov method to quantify 3D grain size distributions in mylonites. *Journal of Structural Geology* **93** 149–161. <https://doi.org/10.1016/j.jsg.2016.10.008>
- [31] MONARD, F., NICKL, R. and PATERNAIN, G. P. (2019). Efficient nonparametric Bayesian inference for X-ray transforms. *Ann. Statist.* **47** 1113–1147. [MR3909962](#) <https://doi.org/10.1214/18-AOS1708>

- [32] NICKL, R. (2020). Bernstein–von Mises theorems for statistical inverse problems I: Schrödinger equation. *J. Eur. Math. Soc. (JEMS)* **22** 2697–2750. [MR4118619](#) <https://doi.org/10.4171/JEMS/975>
- [33] NICKL, R. and SÖHL, J. (2019). Bernstein–von Mises theorems for statistical inverse problems II: Compound Poisson processes. *Electron. J. Stat.* **13** 3513–3571. [MR4013745](#) <https://doi.org/10.1214/19-ejs1609>
- [34] RAY, K. (2017). Adaptive Bernstein–von Mises theorems in Gaussian white noise. *Ann. Statist.* **45** 2511–2536. [MR3737900](#) <https://doi.org/10.1214/16-AOS1533>
- [35] RAY, K. and VAN DER VAART, A. (2020). Semiparametric Bayesian causal inference. *Ann. Statist.* **48** 2999–3020. [MR4152632](#) <https://doi.org/10.1214/19-AOS1919>
- [36] RIVOIRARD, V. and ROUSSEAU, J. (2012). Bernstein–von Mises theorem for linear functionals of the density. *Ann. Statist.* **40** 1489–1523. [MR3015033](#) <https://doi.org/10.1214/12-AOS1004>
- [37] SEN, B., BANERJEE, M. and WOODROOFE, M. (2010). Inconsistency of bootstrap: The Grenander estimator. *Ann. Statist.* **38** 1953–1977. [MR2676880](#) <https://doi.org/10.1214/09-AOS777>
- [38] SEN, B. and WOODROOFE, M. (2012). Bootstrap confidence intervals for isotonic estimators in a stereological problem. *Bernoulli* **18** 1249–1266. [MR2995794](#) <https://doi.org/10.3150/12-BEJ378>
- [39] STOYAN, D. and KENDALL, W. S. (1987). *Stochastic Geometry and Its Applications*. Wiley, New York.
- [40] VAN DER VAART, A. W. and WELLNER, J. A. (2023). *Weak Convergence and Empirical Processes—with Applications to Statistics*. *Springer Series in Statistics*. Springer, Cham. [MR4628026](#) <https://doi.org/10.1007/978-3-031-29040-4>
- [41] WANG, K. and GHOSAL, S. (2023). Coverage of credible intervals in Bayesian multivariate isotonic regression. *Ann. Statist.* **51** 1376–1400. [MR4630953](#) <https://doi.org/10.1214/23-aos2298>
- [42] WANG, K. and GHOSAL, S. (2023). Posterior contraction and testing for multivariate isotonic regression. *Electron. J. Stat.* **17** 798–822. [MR4554662](#) <https://doi.org/10.1214/23-ejs2115>
- [43] WATSON, G. S. (1971). Estimating functionals of particle size distributions. *Biometrika* **58** 483–490. [MR0312639](#) <https://doi.org/10.1093/biomet/58.3.483>
- [44] WICKSELL, S. D. (1925). The corpuscle problem. *Biometrika* **17** 84–99. <https://doi.org/10.1093/biomet/17.1-2.84>