Rates of convergence in de Finetti's representation theorem, and Hausdorff moment problem

EMANUELE DOLERA[1] and STEFANO FAVARO[2,1]

[1]*Department of Mathematics, University of Pavia, Pavia, 27100 Italy. E-mail: emanuele.dolera@unipv.it*
[2]*Department of Economics and Statistics, University of Torino and Collegio Carlo Alberto, Torino, 10134 Italy. E-mail: stefano.favaro@unito.it*

Given a sequence $\{X_n\}_{n \geq 1}$ of exchangeable Bernoulli random variables, the celebrated de Finetti representation theorem states that $\frac{1}{n} \sum_{i=1}^{n} X_i \xrightarrow{a.s.} Y$ for a suitable random variable $Y : \Omega \to [0, 1]$ satisfying $\mathsf{P}[X_1 = x_1, \ldots, X_n = x_n | Y] = Y^{\sum_{i=1}^{n} x_i} (1 - Y)^{n - \sum_{i=1}^{n} x_i}$. In this paper, we study the rate of convergence in law of $\frac{1}{n} \sum_{i=1}^{n} X_i$ to Y under the Kolmogorov distance. After showing that a rate of the type of $1/n^\alpha$ can be obtained for any index $\alpha \in (0, 1]$, we find a sufficient condition on the distribution of Y for the achievement of the optimal rate of convergence, that is $1/n$. Besides extending and strengthening recent results under the weaker Wasserstein distance, our main result weakens the regularity hypotheses on Y in the context of the Hausdorff moment problem.

Keywords: de Finetti's law of large numbers; de Finetti's representation theorem; Edgeworth expansions; exchangeability; Hausdorff moment problem; Kolmogorov distance; Wasserstein distance

1. Introduction

This paper contributes to the study of the rate of convergence of the law of large numbers for exchangeable random variables (r.v.'s) in the sense of de Finetti [5]. For ease in exposition, we confine ourselves to an infinite sequence $\{X_n\}_{n \geq 1}$ of Bernoulli variables defined on the probability space $(\Omega, \mathcal{F}, \mathsf{P})$. The sequence $\{X_n\}_{n \geq 1}$ satisfies the exchangeability condition if

$$\mathsf{P}[X_1 = x_1, \ldots, X_n = x_n] = \mathsf{P}[X_1 = x_{\sigma_n(1)}, \ldots, X_n = x_{\sigma_n(n)}]$$

for all $n \in \mathbb{N}$ $(x_1, \ldots, x_n) \in \{0, 1\}^n$ and permutation σ_n of the set $\{1, \ldots, n\}$. De Finetti [5] proved a strong law of large numbers for the exchangeable X_i's, that is, $\frac{1}{n} \sum_{i=1}^{n} X_i \xrightarrow{a.s.} Y$ for a suitable r.v. $Y : \Omega \to [0, 1]$ satisfying $\mathsf{P}[X_1 = x_1, \ldots, X_n = x_n | Y] = Y^{s_n} (1 - Y)^{n - s_n}$, where $s_n := \sum_{i=1}^{n} x_i$. This identity yields the so-called de Finetti representation theorem [4], which reads as

$$\mathsf{P}[X_1 = x_1, \ldots, X_n = x_n] = \int_0^1 \theta^{s_n} (1 - \theta)^{n - s_n} \mu(\mathrm{d}\theta) \tag{1.1}$$

---

[1]Also affiliated to IMATI-CNR "Enrico Magenes" (Milan, Italy).

for all $n \in \mathbb{N}$ and $(x_1, \ldots, x_n) \in \{0, 1\}^n$, where μ, the so-called de Finetti measure (or prior measure), stands for the probability distribution (p.d.) of Y. See the monograph by Aldous [1] and references therein for a comprehensive treatment of exchangeability and de Finetti's theorem. The above law of large number entails that the p.d. of $\frac{1}{n} \sum_{i=1}^{n} X_i$, say μ_n, converges weakly to μ as $n \to +\infty$ ($\mu_n \Rightarrow \mu$ in symbols), meaning that $\lim_{n \to \infty} \int_0^1 \psi(\theta) \mu_n(d\theta) = \int_0^1 \psi(\theta) \mu(d\theta)$ is valid for all $\psi : [0, 1] \to \mathbb{R}$ bounded and continuous.

The study of the rate of convergence of the empirical measure μ_n to μ requires the choice of a suitable distance that induces the weak convergence. A reasonable choice of such a distance yields an explicit evaluation of the discrepancy between μ_n and μ, as a function of the sample size n, and also a practical interpretation of the approximation from various points of view. In this paper, we focus on the Kolmogorov distance which, for any pair (ν_1, ν_2) of probability measures (p.m.'s) on $([0, 1], \mathcal{B}([0, 1]))$, is defined as follows

$$d_K(\nu_1; \nu_2) := \sup_{x \in [0,1]} \left| \nu_1([0, x]) - \nu_2([0, x]) \right| = \sup_{x \in [0,1]} \left| F_1(x) - F_2(x) \right|.$$

Recall that d_K metrizes the weak convergence on the space $\mathcal{P}(0, 1)$ of all p.m.'s on $([0, 1], \mathcal{B}([0, 1]))$ when the limiting p.m. has a continuous distribution function (d.f.). The next theorem is the main result of the present paper. It provides the first quantitative version, with respect to d_K, of de Finetti's law of large numbers for μ_n. We denote by $L^\infty(0, 1)$ the space of essentially bounded function on $(0, 1)$, with respect to the Lebesgue measure.

Theorem 1.1. *If μ has density function f with distributional derivative f' such that*

$$[\theta(1 - \theta)]^\gamma \left| f'(\theta) \right| \in L^\infty(0, 1), \tag{1.2}$$

for some $\gamma \in (0, 1)$, then

$$d_K(\mu_n; \mu) \le \frac{C(\mu)}{n} \tag{1.3}$$

is valid for all $n \in \mathbb{N}$, with $C(\mu)$ that depends on μ only through $\mathrm{ess.\,sup}_{\theta \in [0,1]} f(\theta)$ and $\mathrm{ess.\,sup}_{\theta \in [0,1]} [\theta(1 - \theta)]^\gamma |f'(\theta)|$.

Despite the long history of the celebrated de Finetti representation theorem, the study of the rate of convergence of μ_n to μ has been initiated very recently in the work of Mijoule, Peccati and Swan [19]. In particular, they proved a quantitative version of de Finetti's law of large numbers with respect to the Kantorovich distance (or Wasserstein distance of order 1) $d_W(\nu_1; \nu_2) := \int_0^1 |F_1(x) - F_2(x)| \, dx$. In particular, they showed that for any $n \in \mathbb{N}$

$$\frac{C_1(\mu)}{n} \le d_W(\mu_n; \mu) \le \sqrt{\frac{C_1(\mu)}{n}}$$

hold for any $\mu \in \mathcal{P}(0, 1)$, where $C_1(\mu) := \int_0^1 \theta(1 - \theta) \mu(d\theta)$. Thus, $1/n$ is precisely the best possible rate of convergence to zero also for $d_K(\mu_n; \mu)$, since $d_K(\mu_n; \mu) \ge d_W(\mu_n; \mu)$, and it follows that $d_W(\mu_n; \mu)$ goes to zero at least as fast as $1/\sqrt{n}$. Mijoule Peccati and Swan [19] also

provided sufficient conditions for the achievement of the best rate of convergence. In particular, if the p.d. μ is absolutely continuous with a density f satisfying $\int_0^1 \theta(1-\theta)|f'(\theta)| \, d\theta < +\infty$, where f' stands for the (distributional) derivative of f, then

$$d_W(\mu_n; \mu) \leq \frac{C_2(\mu)}{n}$$

holds for any $n \in \mathbb{N}$ with an explicit constant $C_2(\mu)$. Finally, it is proved that, for every $\delta \in [\frac{1}{2}, 1]$, there exists a suitable $\mu \in \mathcal{P}(0, 1)$ for which $d_W(\mu_n; \mu) \sim 1/n^\delta$ as $n \to +\infty$. Recall that the beta distribution with parameters $(a, b) \in (0, +\infty)^2$ is the element of $\mathcal{P}(0, 1)$ with density function

$$(0, 1) \ni \theta \mapsto \beta(\theta; a, b) := \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \theta^{a-1} (1-\theta)^{b-1}. \tag{1.4}$$

The work of Mijoule Peccati and Swan [19] generalizes bound of the form $d_W(\mu_n; \mu) \leq C/n$, which was originally obtained in Goldstein and Reinert [15] under the assumption that μ is a beta distribution. See also Döbler [8] and references therein for some related results.

Our study of $d_K(\mu_n; \mu)$ is more challenging than the study of $d_W(\mu_n; \mu)$. While results in Mijoule Peccati and Swan [19] rely on classical Berry-Esseen bounds for the Gaussian approximation in the central limit theorem, the proof of Theorem 1.1 relies on novel refined version of these Berry-Esseen bounds, usually known as Edgeworth expansions. See Chapter 5 and Chapter 6 in Petrov [22], Chapter 3 in Ibragimov and Linnik [17] and references therein. To highlight the difference between the study of $d_K(\mu_n; \mu)$ and $d_W(\mu_n; \mu)$, we state a simple result on beta prior measures, whose proof can be obtain by direct computation. This shows that any rate $n^{-\alpha}$, with $\alpha \in (0, 1]$, is actually achieved by $d_K(\mu_n; \mu)$.

Proposition 1.1. *If μ is the beta distribution with parameter $(\alpha, 1)$ or $(1, \alpha)$, with $\alpha > 0$, then there exists a constant C_α for which, for any $n \in \mathbb{N}$,*

$$d_K(\mu_n; \mu) \leq C_\alpha \left(\frac{1}{n}\right)^{\alpha \wedge 1} \tag{1.5}$$

is fulfilled, where $\wedge$ denotes the minimum value.

It is important to stress that, in view of results in Goldstein and Reinert [15] and Mijoule Peccati and Swan [19], for the beta prior we have $d_W(\mu_n; \mu) \sim 1/n$. This shows that the asymptotic behaviour of $d_K(\mu_n; \mu)$ is different from the asymptotic behaviour of $d_W(\mu_n; \mu)$. In addition, we notice that it seems not convenient at all to resort to the inequality

$$d_K(\nu_1; \nu_2) \leq C(\nu_2)\sqrt{d_W(\nu_1; \nu_2)}, \tag{1.6}$$

which is valid whenever ν_2 has a bounded density function. See, for example, Gibbs and Su [13]. Indeed, for the beta distribution with parameters $(\alpha, 1)$ or $(1, \alpha)$, with $\alpha \geq 1$, (1.6) would lead to an upper bound like $C/\sqrt{n}$, which is worse than the upper bound C/n given by (1.5).

Besides extending and strengthening results obtained in Goldstein and Reinert [15] and Mijoule Peccati and Swan [19], Theorem 1.1 is also connected to the celebrated Hausdorff moment

problem. The Hausdorff moment problem is known to be closely related to de Finetti's theorem. See, for example, Ressel [24], Diaconis and Freedman [6], Diaconis and Freedman [7] and references therein. Within the context of the Hausdorff moment problem, the main result of Mnatsakanov [20] shows the rate of convergence of $d_K(\mu_n; \mu)$ under a certain assumption on the prior measure μ. In the equivalent reformulation of our problem as the finding of the rate of approximation in the Hausdorff moment problem, Theorem 2 in Mnatsakanov [20] provides the existence of another constant $C_*(\mu)$ for which (1.3) holds for any $n \in \mathbb{N}$, with another constant $C_*(\mu)$ in the place of our $C(\mu)$. In spite of a very direct proof, the main difference is that $C_*(\mu)$ depends on $\mathrm{ess.}\sup_{\theta \in [0,1]} |f'(\theta)|$, which is tantamount to requiring that the density f of μ belongs to $W^{1,\infty}(0, 1)$, the Sobolev space of essentially bounded functions on $(0, 1)$ with an essentially bounded distributional derivative.

The comparison of Theorem 2 in Mnatsakanov [20] with Proposition 1.1 shows that the assumption $f \in W^{1,\infty}(0, 1)$ is indeed too strong and far from capturing the whole class of prior distributions for which (1.3) is met. Theorem 1.1 fills this gap by providing the general sufficient condition (1.2) for the achievement of the best rate $1/n$. This leads to a remarkable improvement of Theorem 2 in Mnatsakanov [20]. For prior distributions μ with a support strictly contained in $(0, 1)$, condition (1.2) boils down to the assumption that $f \in W^{1,\infty}(0, 1)$, but, without this restriction, it is evident that Theorem 1.1 improves Theorem 2 in Mnatsakanov [20] by allowing $f'(\theta)$ to diverge moderately at $\theta = 0$ and $\theta = 1$. As a final remark, we show that Mnatsakanov's result cannot be re-adapted via a smoothing argument to obtain our sharper bound, thus justifying the effort of providing a longer and more complex proof. In fact, given any prior μ with density $f \notin W^{1,\infty}(0, 1)$, one could try to smooth it by introducing a family of new priors, say $\{\mu_\varepsilon\}_{\varepsilon > 0}$, each with density $f_\varepsilon \in W^{1,\infty}(0, 1)$, so that $\mu_\varepsilon \Rightarrow \mu$ as $\varepsilon \downarrow 0$. Letting $\{X_n^{(\varepsilon)}\}_{n \geq 1}$ be a new sequence of exchangeable Bernoulli r.v.'s having μ_ε as de Finetti's measure, one could argue by resorting to

$$d_K(\mu_n; \mu) \leq d_K(\mu_n; \mu_{n,\varepsilon}) + d_K(\mu_{n,\varepsilon}; \mu_\varepsilon) + d_K(\mu_\varepsilon; \mu),$$

where $\mu_{n,\varepsilon}(\cdot) := \mathsf{P}[\frac{1}{n} \sum_{i=1}^{n} X_i^{(\varepsilon)} \in \cdot]$. Since $d_K(\mu_{n,\varepsilon}; \mu_\varepsilon) \leq C_*(\mu_\varepsilon)/n$ would follow from Theorem 2 in Mnatsakanov [20], the achievement of (1.3) through this line of reasoning would entail $\mathrm{ess.}\sup_{\theta \in [0,1]} |f_\varepsilon'(\theta)| \leq C < +\infty$ for all $\varepsilon > 0$. Moreover, in order to have $d_K(\mu_\varepsilon; \mu) \leq C/n$, the parameter ε should depend on n, yielding $\varepsilon = \varepsilon(n) \downarrow 0$ as $n \to +\infty$. But now, the Ascoli-Arzelà theorem implies that $f \in W^{1,\infty}(0, 1)$, leading to a contradiction.

As a corollary, we state a result for μ being a beta distribution. It is worth pointing out that this result agrees with Proposition 1.1, thus capturing exactly the elements of this class of priors for which the bound (1.3) is valid. Indeed, since the beta density function belongs to $W^{1,\infty}(0, 1)$ if and only if $a, b \geq 2$, we improve the assumption by the following corollary.

Corollary 1.1. *If μ is the beta distribution with parameters (a, b), then* (1.3) *is fulfilled if and only if $a, b \geq 1$.*

To conclude, we state a proposition that deals with a larger class of priors than the class considered in Theorem 1.1. We show that a rate of convergence for $d_K(\mu_n; \mu)$, although not sharp, can be obtained also for a non absolutely continuous prior μ, provided that its d.f. F is Hölder con-

tinuous. This happens, for instance, if F coincides with the Cantor function. The determination of the sharp rate in this setting remains an interesting open problem.

Proposition 1.2. *If $\mu \in \mathcal{P}(0, 1)$ has a γ-Hölder continuous d.f. for some $\gamma \in (0, 1]$, then there exists a suitable constant $L_\gamma(\mu)$ for which, for any $n \in \mathbb{N}$*

$$d_K(\mu_n; \mu) \leq \frac{L_\gamma(\mu)}{n^{\gamma/2}} \tag{1.7}$$

The paper is structured as follows. Section 2 contains the proof of Theorem 1.1, which requires a few preliminary lemmas. Proofs of Proposition 1.1 and Proposition 1.2 are deferred to the appendix.

2. Proof of Theorem 1.1

We start with some preliminary lemmas. First, a decomposition lemma for probability density functions which will be used to justify the introduction of the additional hypothesis $f(0) = f(1) = 0$ in the first part of the proof of Theorem 1.1. As to notation, any relation involving the symbols $A_\pm$ and $f_\pm$ must be intended as a short-hand for the two analogous relations which hold with A_+, f_+ and with A_-, f_-, respectively, in place of $A_\pm$, $f_\pm$.

Lemma 2.1. *Given a probability density function f on $[0, 1]$ which is expressed by a polynomial, then there exist three non-negative constants A_∞, A_+, A_- and three continuous probability density functions f_∞, f_+, f_- on $[0, 1]$ such that:*

(i) $A_\infty \leq \|f\|_\infty := \sup_{\theta \in [0,1]} |f(\theta)|$ *and* $A_\pm \leq 1 + \|f\|_\infty$;

(ii) $f_\infty \in W^{1,\infty}(0, 1)$ *with* $A_\infty \|f_\infty\|_\infty \leq \|f\|_\infty$ *and* $A_\infty \|f'_\infty\|_\infty \leq 2\|f\|_\infty$;

(iii) $f_+, f_- \in W^{1,\infty}(0, 1)$, $f_+(0) = f_+(1) = f_-(0) = f_-(1) = 0$, $A_\pm \|f_\pm\|_\infty \leq 2\|f\|_\infty$ *and, for any $\gamma \in (0, 1)$,*

$$A_\pm \sup_{\theta \in [0,1]} [\theta(1-\theta)]^\gamma |f'_\pm(\theta)| \leq \frac{2}{4^\gamma} \|f\|_\infty + \sup_{\theta \in [0,1]} [\theta(1-\theta)]^\gamma |f'(\theta)|; \tag{2.1}$$

(iv) $f(\theta) = A_\infty f_\infty(\theta) + A_+ f_+(\theta) - A_- f_-(\theta)$ *for all* $\theta \in [0, 1]$.

Proof. If $f(0) = f(1) = 0$, the thesis is trivial. Otherwise, we put $A_\infty = \frac{f(1)+f(0)}{2} \leq \|f\|_\infty$ and $f_\infty(\theta) = \frac{[f(1)-f(0)]\theta + f(0)}{A_\infty}$, so that (ii) holds trivially. Then, recalling that, for any $a, b \in \mathbb{R}$, $a = b + (a - b)_+ - (b - a)_+$, where $x_+ := \max\{0, x\}$, we set $g_+(\theta) := (f(\theta) - A_\infty f_\infty(\theta))_+$ and $g_-(\theta) := (A_\infty f_\infty(\theta) - f(\theta))_+$. Thus, we put $A_\pm = \int_0^1 g_\pm(\theta) \, d\theta$ and $f_\pm(\theta) = g_\pm(\theta)/A_\pm$ with the proviso that, if $A_+ = 0$ ($A_- = 0$, respectively), the definition of f_+ (f_-, respectively) is arbitrary and can be chosen equal to $6x(1 - x)$. By definition, point iv) is met along with $f_+(0) = f_+(1) = f_-(0) = f_-(1) = 0$. Moreover, we have $A_\pm \leq 1 + A_\infty$ and point (i) follows. To prove that $f_+, f_- \in W^{1,\infty}(0, 1)$, it is enough to notice that $f(\theta) = A_\infty f_\infty(\theta)$ for finitely many θ's, by virtue of the fundamental theorem of algebra, and, in the complement of this set,

both f_+ and f_- are again two polynomials. To check that $A_\pm \|f_\pm\|_\infty \leq 2\|f\|_\infty$, it is enough to observe that $\|g_\pm\|_\infty \leq \|f\|_\infty + A_\infty \|f_\infty\|_\infty$. Finally, we just note that, except on the finite set $\{\theta \in [0,1] | f(\theta) = A_\infty f_\infty(\theta)\}$, we have $A_\pm |f'_\pm(\theta)| \leq |f'(\theta)| + A_\infty |f'_\infty(\theta)|$, so that we can deduce the validity of (2.1) from the previous bounds. $\qquad\square$

Another preliminary result deals with further regularity properties of the densities that satisfy (1.2). We observe that, since $[x(1-x)]^{-\gamma} \in L^1(0,1)$ if $\gamma \in (0,1)$, the validity of (1.2) entails $f \in W^{1,1}(0,1)$ and, hence, the existence of a continuous version of the same density on the whole set $[0,1]$. See, for example, Theorem 8.2 in Brezis [2].

Lemma 2.2. *Let f be a probability density function satisfying* (1.2) *for some $\gamma \in (0,1)$, and let* $F(x) = \int_0^x f(y)\, dy$. *Then, there exists a positive constant $R(\gamma)$ such that*

$$\sup_{\substack{x \in (0,1), \\ 0 < w < x(1-x)}} x(1-x) \left| \frac{F(x+w) - 2F(x) + F(x-w)}{w^2} \right| \leq R(\gamma)|f|_{1,\gamma} \tag{2.2}$$

is fulfilled with $|f|_{1,\gamma} := \text{ess.}\sup_{\theta \in [0,1]}[\theta(1-\theta)]^\gamma |f'(\theta)|$. Moreover, if the additional condition $f(0) = f(1) = 0$ (referred to the continuous representative of f) is in force, then:

(i) $f(\theta) \leq M(f)\theta^{1-\gamma}$, $f(\theta) \leq M(f)(1-\theta)^{1-\gamma}$ *hold for all $\theta \in [0,1]$, with $M(f) := \frac{2^\gamma}{1-\gamma}|f|_{1,\gamma} + 2^{1-\gamma}\|f\|_\infty$;*

(ii) $F(x) \leq M(f)\frac{x^{2-\gamma}}{2-\gamma}$, $1 - F(x) \leq M(f)\frac{(1-x)^{2-\gamma}}{2-\gamma}$ *hold for all $x \in [0,1]$.*

Proof. Since $f \in W^{1,1}(0,1)$ by virtue of (1.2), the Taylor formula with integral remainder can be applied to obtain

$$F(x+w) - 2F(x) + F(x-w) = \int_x^{x+w} (x+w-t)f'(t)\, dt + \int_x^{x-w} (x-w-t)f'(t)\, dt$$

for all $x \in (0,1)$ and w satisfying $0 < w < x(1-x)$. Whence,

$$\left| F(x+w) - 2F(x) + F(x-w) \right|$$
$$\leq \int_x^{x+w} (w+t-x)\left| f'(t) \right| dt + \int_{x-w}^x (w+x-t)\left| f'(t) \right| dt. \tag{2.3}$$

At this stage, we show explicitly how to bound the former integral when $x \in (0, 1/2]$, the other cases being analogous. Since $0 < x < x + w < 3/4$, then we get $|f'(t)| \leq 4^\gamma |f|_{1,\gamma} t^{-\gamma}$ for all $t \in [x, x+w]$, leading to

$$\int_x^{x+w} (w+t-x)\left| f'(t) \right| dt$$
$$\leq 4^\gamma |f|_{1,\gamma} \left[w\frac{(x+w)^{1-\gamma} - x^{1-\gamma}}{1-\gamma} + \frac{(x+w)^{2-\gamma} - x^{2-\gamma}}{2-\gamma} - x\frac{(x+w)^{1-\gamma} - x^{1-\gamma}}{1-\gamma} \right].$$

Then, we put $\eta := w/x$ and we observe that $\eta \in (0, 1/2)$, so that the expression inside the brackets can be written as

$$wx^{1-\gamma} \frac{(1+\eta)^{1-\gamma} - 1}{1-\gamma}$$

$$+ x^{2-\gamma} \left[\frac{(1+\eta)^{2-\gamma} - 1 - (2-\gamma)\eta}{2-\gamma} - \frac{(1+\eta)^{1-\gamma} - 1 - (1-\gamma)\eta}{1-\gamma} \right]. \qquad (2.4)$$

We conclude this argument by noticing that, for any η satisfying $|\eta| \leq 1/2$ and any $\alpha > 0$ there exists a constant $H(\alpha)$ such that $|(1+\eta)^\alpha - 1 - \alpha\eta| \leq H(\alpha)\eta^2$. This remark implies that the expression in (2.4) is bounded by the following quantity

$$\left[1 + \frac{H(2-\gamma)}{2-\gamma} + \frac{3H(1-\gamma)}{2(1-\gamma)} \right] w^2 x^{-\gamma},$$

yielding

$$\sup_{\substack{x \in (0, 1/2], \\ 0 < w < x(1-x)}} \frac{x(1-x)}{w^2} \int_x^{x+w} (w+t-x) |f'(t)| \, dt$$

$$\leq 2^{3\gamma-2} \left[1 + \frac{H(2-\gamma)}{2-\gamma} + \frac{3H(1-\gamma)}{2(1-\gamma)} \right] |f|_{1,\gamma}.$$

As recalled, the treatment of the latter integral on the right-hand side of (2.3) for $x \in (0, 1/2]$ is analogous. Lastly, when $x \in [1/2, 1)$, it is enough to change the variable $t = 1 - s$, obtaining $\int_x^{x+w} (w+t-x)|f'(t)| \, dt = \int_{1-x-w}^{1-x} (w+1-x-s)|f'(1-s)| \, ds$ and $\int_{x-w}^x (w+x-t)|f'(t)| \, dt = \int_{1-x}^{1-x+w} (w+s-1+x)|f'(1-s)| \, ds$, where the integrals in the new variable s are exactly the integrals studied above. To prove (i), we just write $f(\theta) = \int_0^\theta f'(y) \, dy$ and, confining to the case that $\theta \in [0, 1/2]$, we exploit (1.2) in the form $|f'(\theta)| \leq 2^\gamma |f|_{1,\gamma} \theta^{-\gamma}$. This proves the first bound, after noticing that $f(\theta) \leq 2^{1-\gamma} \|f\|_\infty \theta^{1-\gamma}$ is valid for any $\theta \in [1/2, 1]$. For the latter bound, we start from $-f(\theta) = \int_\theta^1 f'(y) \, dy$ and we argue in an analogous way. To prove ii), it is enough to integrate the bounds obtained in point i). □

The last preliminary result provides with a refinement of the well-known estimates of Berry–Esseen type for the characteristic function of a normalized sum of i.i.d., centered r.v.'s. In fact, the following statement can be viewed as a generalization of Lemma 4 in Chapter VI of Petrov [22] and Theorem 3.2.1(2) in Chapter 3 of Ibragimov and Linnik [17] in the case that the summands possess the $3 + \delta$ absolute moment for some $\delta \in (0, 1)$ but, in general, not the fourth moment.

Lemma 2.3. *Let $\{V_n\}_{n \geq 1}$ be a sequence of i.i.d. r.v's defined on $(\Omega, \mathcal{F}, \mathsf{P})$ such that $\beta_{3+\delta} := \mathsf{E}[|V_1|^{3+\delta}] < +\infty$ holds for some $\delta \in (0, 1)$, along with $\mathsf{E}[V_1] = 0$ and $\mathsf{E}[V_1^2] =: \sigma^2 > 0$. Upon putting $\alpha_3 := \mathsf{E}[V_1^3]$ and $\psi_n(\xi) := \mathsf{E}[\exp\{i\xi(\sum_{k=1}^n V_k)/\sqrt{n\sigma^2}\}]$, there holds*

$$\left| \psi_n(\xi) - e^{-\xi^2/2} \left\{ 1 + \frac{\alpha_3}{6\sqrt{n}\sigma^3} (i\xi)^3 \right\} \right| \leq Q(\delta) \frac{\beta_{3+\delta}}{n^{(1+\delta)/2}\sigma^{3+\delta}} |\xi|^{3+\delta} (1 + |\xi|^4) e^{-\xi^2/4} \qquad (2.5)$$

for any ξ satisfying $|\xi| \leq \frac{1}{4}\sqrt{n}(\frac{\sigma^{3+\delta}}{\beta_{3+\delta}})^{1/(1+\delta)}$, *where $Q(\delta)$ is a numerical constant independent of ξ and the p.d. of V_1.*

Proof. The proof is based on the arguments used to prove Lemma A.2 in Dolera, Gabetta and Regazzini [9] and Lemma 3.1 in Dolera and Regazzini [10]. First, we put $\psi(\xi) := \mathsf{E}[e^{i\xi V_1}]$ and we observe that $\psi(\xi) = 1 - \frac{\sigma^2}{2}\xi^2 + \frac{\alpha_3}{6}(i\xi)^3 + \rho_\delta(\xi)$, where

$$\left|\rho_\delta(\xi)\right| \leq \frac{2^{1-\delta}\beta_{3+\delta}}{(1+\delta)(2+\delta)(3+\delta)}|\xi|^{3+\delta}.$$

See, for example, Theorem 1 in Section 8.4 of Chow and Teicher [3]. Whence,

$$\psi\left(\frac{\xi}{\sqrt{n\sigma^2}}\right) = 1 - \frac{1}{2n}\xi^2 + \frac{\alpha_3}{6\sigma^3 n^{3/2}}(i\xi)^3 + \rho_{n,\delta}(\xi)$$

with

$$\left|\rho_{n,\delta}(\xi)\right| \leq \frac{2^{1-\delta}\beta_{3+\delta}}{(1+\delta)(2+\delta)(3+\delta)\sigma^{3+\delta}n^{(3+\delta)/2}}|\xi|^{3+\delta}.$$

Now, we notice that Lyapunov's inequality entails $\sigma^{3+\delta} \leq \beta_{3+\delta}$, while Hölder's inequality shows that

$$\frac{|\alpha_3|}{\sigma^3}\left(\frac{\sigma^{3+\delta}}{\beta_{3+\delta}}\right)^{3/(1+\delta)} \leq \frac{\beta_3}{\sigma^3}\left(\frac{\sigma^{3+\delta}}{\beta_{3+\delta}}\right)^{3/(1+\delta)} \leq \frac{\beta_3}{\sigma^3}\left(\frac{\sigma^{3+\delta}}{\beta_{3+\delta}}\right)^{1/(1+\delta)} \leq 1$$

where $\beta_3 := \mathsf{E}[|V_1|^3]$. Therefore, for any ξ satisfying $|\xi| \leq \frac{1}{4}\sqrt{n}(\frac{\sigma^{3+\delta}}{\beta_{3+\delta}})^{1/(1+\delta)}$, we have

$$\left|-\frac{1}{2n}\xi^2 + \frac{\alpha_3}{6\sigma^3 n^{3/2}}(i\xi)^3 + \rho_{n,\delta}(\xi)\right| \leq \frac{5}{128}.$$

Thanks to this bound, we are allowed to consider the principal logarithm $\mathrm{Log}(1+z) := -\sum_{k=1}^{\infty}\frac{(-z)^k}{k}$, $|z| < 1$ and, since $\psi_n(\xi) = [\psi(\frac{\xi}{\sqrt{n\sigma^2}})]^n$, we have:

$$\psi_n(\xi) = \exp\left\{n\,\mathrm{Log}\left[\psi\left(\frac{\xi}{\sqrt{n\sigma^2}}\right)\right]\right\}$$

$$= \exp\left\{n\,\mathrm{Log}\left[1 - \frac{1}{2n}\xi^2 + \frac{\alpha_3}{6\sigma^3 n^{3/2}}(i\xi)^3 + \rho_{n,\delta}(\xi)\right]\right\}$$

$$= e^{-\xi^2/2}\exp\left\{\frac{\alpha_3}{6\sqrt{n}\sigma^3}(i\xi)^3\right\}e^{\tau_{n,\delta}(\xi)}$$

where $\tau_{n,\delta}(\xi)$ is defined to be

$$n\rho_{n,\delta}(\xi) + n\left[-\frac{1}{2n}\xi^2 + \frac{\alpha_3}{6\sigma^3 n^{3/2}}(i\xi)^3 + \rho_{n,\delta}(\xi)\right]^2 \Upsilon\left(-\frac{1}{2n}\xi^2 + \frac{\alpha_3}{6\sigma^3 n^{3/2}}(i\xi)^3 + \rho_{n,\delta}(\xi)\right)$$

with $\Upsilon(z) := -\sum_{k=2}^{\infty} \frac{(-z)^{k-2}}{k}$ for $|z| < 1$. At this stage, we put $u_3(\xi) := \frac{\alpha_3}{6\sqrt{n}\sigma^3}(i\xi)^3$ and $\Theta(z) :=$ $e^z - 1 - z$, and we exploit the elementary inequality $|e^z - 1| \le |z|e^{|z|}$ to obtain

$$
\left| \psi_n(\xi) - e^{-\xi^2/2} \left\{ 1 + \frac{\alpha_3}{6\sqrt{n}\sigma^3}(i\xi)^3 \right\} \right|
$$
$$
\le e^{-\xi^2/2} e^{|u_3(\xi)|} \left| e^{\tau_{n,\delta}(\xi)} - 1 \right| + e^{-\xi^2/2} \left| \Theta(u_3) \right|
$$
$$
\le e^{-\xi^2/2} e^{|u_3(\xi)| + |\tau_{n,\delta}(\xi)|} \left| \tau_{n,\delta}(\xi) \right| + e^{-\xi^2/2} \Theta(|u_3|).
$$

To conclude, it is enough to notice that, for any ξ satisfying $|\xi| \le \frac{1}{4}\sqrt{n}(\frac{\sigma^{3+\delta}}{\beta_{3+\delta}})^{1/(1+\delta)}$, we have $|u_3(\xi)| \le \frac{1}{384}$, $|\tau_{n,\delta}(\xi)| \le \frac{1}{4}\xi^2$,

$$
\left| \tau_{n,\delta}(\xi) \right| \le Q_1(\delta) \frac{\beta_{3+\delta}}{n^{(1+\delta)/2}\sigma^{3+\delta}} |\xi|^{3+\delta} \left(1 + |\xi|^4 \right)
$$

and

$$
\Theta(|u_3|) \le Q_2(\delta) \frac{\beta_{3+\delta}}{n^{(1+\delta)/2}\sigma^{3+\delta}} |\xi|^{3(1+\delta)}
$$

for suitable constants $Q_1(\delta)$ and $Q_2(\delta)$ independent of ξ and the p.d. of V_1. $\square$

The way is now paved for the study of Theorem 1.1. The first part of the proof, which requires the major effort, is devoted to proving (1.3) when the density f of μ, in addition to (1.2), satisfies $f(0) = f(1) = 0$. We recall again that (1.2) entails $f \in W^{1,1}(0,1)$ and, hence, the existence of a continuous version of this density on the whole set $[0,1]$, by virtue of Theorem 8.2 in Brezis [2]. Obviously, the additional assumption $f(0) = f(1) = 0$ is referred to this version. After these preliminaries, we get into the real proof by defining $I(n,\gamma) := [\overline{x}_{n,\gamma}, 1 - \overline{x}_{n,\gamma}]$, with $\overline{x}_{n,\gamma} := (1/n)^{\frac{1}{2-\gamma}}$, which is a proper interval provided that $n \ge 4$. Then, after denoting by F_n (F, respectively) the d.f. associated to μ_n (μ, respectively), we split the original quantity as follows:

$$
d_K(\mu_n; \mu) \le \sup_{x \in [0,\overline{x}_{n,\gamma}]} \left| F_n(x) - F(x) \right| + \sup_{x \in I(n,\gamma)} \left| F_n(x) - F(x) \right|
$$
$$
+ \sup_{x \in [1-\overline{x}_{n,\gamma},1]} \left| F_n(x) - F(x) \right|
$$
$$
\le F_n(\overline{x}_{n,\gamma}) + F(\overline{x}_{n,\gamma}) + \sup_{x \in I(n,\gamma)} \left| F_n(x) - F(x) \right|
$$
$$
+ \left[1 - F_n(1 - \overline{x}_{n,\gamma}) \right] + \left[1 - F(1 - \overline{x}_{n,\gamma}) \right]. \tag{2.6}
$$

To bound $F(\overline{x}_{n,\gamma})$ and $[1 - F(1 - \overline{x}_{n,\gamma})]$, we use point ii) of Lemma 2.2, which gives:

$$
F(\overline{x}_{n,\gamma}) + \left[1 - F(1 - \overline{x}_{n,\gamma}) \right] \le \frac{2M(f)}{2-\gamma} \cdot \frac{1}{n} \tag{2.7}
$$

for all $n \geq 4$. To bound $\mathsf{F}_n(\overline{x}_{n,\gamma})$ and $1 - \mathsf{F}_n(1 - \overline{x}_{n,\gamma})$, we invoke equation (A.1) in the proof of Proposition 1.1 which, in combination with point ii) of Lemma 2.2, yields

$$
\mathsf{F}_n(\overline{x}_{n,\gamma}) \leq M(f) \int_0^1 \beta\big(y; \lfloor n\overline{x}_{n,\gamma} \rfloor + 1, n - \lfloor n\overline{x}_{n,\gamma} \rfloor\big) y^{2-\gamma} \, dy
$$

$$
= M(f) \frac{\Gamma(n+1)}{\Gamma(\lfloor n\overline{x}_{n,\gamma} \rfloor + 1)\Gamma(n - \lfloor n\overline{x}_{n,\gamma} \rfloor)} \frac{\Gamma(\lfloor n\overline{x}_{n,\gamma} \rfloor + 3 - \gamma)\Gamma(n - \lfloor n\overline{x}_{n,\gamma} \rfloor)}{\Gamma(n+3-\gamma)}
$$

where β is the same as in (1.4) and $\lfloor \cdot \rfloor$ denotes the integral part. The last expression can be majorized by means of *Wendel's inequalities* (see (5) in Qi and Luo [23]) as

$$
M(f) \frac{\lfloor n\overline{x}_{n,\gamma} \rfloor + 2 - \gamma}{n + 2 - \gamma} \left(\frac{\lfloor n\overline{x}_{n,\gamma} \rfloor + 1}{n + 1} \right)^{1-\gamma} \left(\frac{n + 2 - \gamma}{n + 1} \right)^{\gamma}
$$

which, for all $n \geq 4$, is less than $\frac{12}{5} M(f)/n$. To study $1 - \mathsf{F}_n(1 - \overline{x}_{n,\gamma})$, we argue as in the proof of Lemma 2.2 by considering the exchangeable sequence $\{\overline{X}_n\}_{n \geq 1}$, where $\overline{X}_n := 1 - X_n$ for all $n \in \mathbb{N}$. Since the de Finetti measure of this new sequence is the element of $\mathcal{P}(0, 1)$ associated to the d.f. $1 - \mathsf{F}(1 - x)$, we resort again to (A.1) to obtain

$$
1 - \mathsf{F}_n(1 - \overline{x}_{n,\gamma}) \leq \mathsf{P}\left[\frac{1}{n} \sum_{i=1}^n \overline{X}_i \leq \overline{x}_{n,\gamma} \right]
$$

$$
= \int_0^1 \beta\big(y; \lfloor n\overline{x}_{n,\gamma} \rfloor + 1, n - \lfloor n\overline{x}_{n,\gamma} \rfloor\big)\big[1 - \mathsf{F}(1 - y)\big] dy.
$$

Thus, by using the latter bound stated in point ii) of Lemma 2.2 and arguing exactly as above, we conclude that $1 - \mathsf{F}_n(1 - \overline{x}_{n,\gamma}) \leq \frac{12}{5} M(f)/n$.

Now, we study $\sup_{x \in I(n,\gamma)} |\mathsf{F}_n(x) - \mathsf{F}(x)|$. First, we get $\mathsf{P}[\sum_{i=1}^n X_i = k | Y = \theta] = \binom{n}{k} \theta^k (1 - \theta)^{n-k}$ for any $k \in \{0, \ldots, n\}$, thanks to de Finetti's representation. Since

$$
\mathsf{P}\left[\sum_{i=1}^n X_i \leq nx | Y = \theta \right] = \mathsf{P}\left[\frac{\sum_{i=1}^n (X_i - \theta)}{\sqrt{n\theta(1 - \theta)}} \leq \frac{n(x - \theta)}{\sqrt{n\theta(1 - \theta)}} | Y = \theta \right],
$$

we put

$$
u := u(x, \theta, n) := \frac{n(x - \theta)}{\sqrt{n\theta(1 - \theta)}}
$$

and

$$
\mathsf{B}_n(y; \theta) := \mathsf{P}\left[\sum_{i=1}^n (X_i - \theta) \leq y\sqrt{n\theta(1 - \theta)} | Y = \theta \right]
$$

for $y \in \mathbb{R}$. To study $\mathsf{B}_n(\cdot; \theta)$, we make the key remark that it coincides with the d.f. of a normalized sum of i.i.d., centered r.v.'s, so that we can employ well-known results pertinent to the

central limit theorem, as stated in Chapter 8 of Gnedenko and Kolmogorov, [14], Chapters 5–6 of Petrov [22], Chapter 3 of Ibragimov and Linnik [17], and in Osipov [21]. In particular, mimicking the main theorem in Osipov [21], we introduce the functions $\mathsf{G}_n(y;\theta) := \Phi(y) + \mathsf{H}_n(y;\theta)$ and

$$
\mathsf{H}_n(y;\theta) := \frac{1}{\sqrt{2\pi n\theta(1-\theta)}} e^{-\frac{1}{2}y^2} \left\{ \frac{1}{6}(1-2\theta)(1-y^2) \right.
$$
$$
\left. + S\big(n\theta + y\sqrt{n\theta(1-\theta)}\big)\left[1 + \frac{1-2\theta}{6\sqrt{n\theta(1-\theta)}}(y^3-3y)\right]\right\}, \qquad (2.8)
$$

where $\Phi(y) := \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\,dx$ and $S(x) := \lfloor x \rfloor - x + \frac{1}{2}$. Now, Theorem 2 in Chapter 8 of Gnedenko and Kolmogorov [14] (see also Theorem 2b in Chapter II of Esseen [11]) entails $\sup_{y\in\mathbb{R}} |\mathsf{B}_n(y;\theta) - \mathsf{G}_n(y;\theta)| \le \epsilon_n(\theta)$ and provides the existence of three numerical constants $\lambda_1, \lambda_2 > 0$ and $n_0 \in \mathbb{N}$ (independent of n and θ) such that

$$
\epsilon_n(\theta) = \lambda_1 \int_{-n}^{n} \left| \frac{\hat{\mathsf{B}}_n(\xi;\theta) - \hat{\mathsf{G}}_n(\xi;\theta)}{\xi} \right| d\xi + \frac{\lambda_2}{n} \sup_{y\notin\mathcal{Y}(n,\theta)} \left| \frac{\partial}{\partial y}\mathsf{G}_n(y;\theta) \right|
$$
$$
\le \lambda_1 \int_{-n}^{n} \left| \frac{\hat{\mathsf{B}}_n(\xi;\theta) - \hat{\mathsf{G}}_n(\xi;\theta)}{\xi} \right| d\xi + \frac{\lambda_2 G}{n\theta(1-\theta)} \qquad (n \ge n_0) \qquad (2.9)
$$

where: $\hat{\mathsf{B}}_n(\xi;\theta)$ and $\hat{\mathsf{G}}_n(\xi;\theta)$ are Fourier–Stieltjes transforms of $\mathsf{B}_n(\cdot;\theta)$ and $\mathsf{G}_n(\cdot;\theta)$, that is, $\int_{-\infty}^{\infty} e^{i\xi y}\,d_y\mathsf{B}_n(y;\theta)$ and $\int_{-\infty}^{\infty} e^{i\xi y}\,d_y\mathsf{G}_n(y;\theta)$, respectively; $\mathcal{Y}(n,\theta) := \{\frac{k-n\theta}{\sqrt{n\theta(1-\theta)}} | k \in \mathbb{Z}\}$ is the set of the discontinuities of both $\mathsf{B}_n(y;\theta)$ and $\mathsf{G}_n(y;\theta)$; $G > 0$ is another constant (independent of n and θ). Thus, since $\mathsf{F}_n(x) = \int_0^1 \mathsf{B}_n(u(x,\theta,n);\theta)f(\theta)\,d\theta$, we write

$$
\sup_{x\in I(n,\gamma)} \left| \mathsf{F}_n(x) - \mathsf{F}(x) \right|
$$
$$
\le \sup_{x\in I(n,\gamma)} \left| \int_0^1 \Phi\big(u(x,\theta,n)\big)f(\theta)\,d\theta - \mathsf{F}(x) \right|
$$
$$
+ \sup_{x\in I(n,\gamma)} \int_0^1 \left| \mathsf{H}_n\big(u(x,\theta,n);\theta\big) \right| f(\theta)\,d\theta + \int_0^1 \epsilon_n(\theta)f(\theta)\,d\theta \qquad (2.10)
$$

and we try to bound each term on the right-hand side. Apropos of the first term on the right-hand side of (2.10), we introduce a Gaussian r.v. $Z_n : \Omega \to \mathbb{R}$ with zero mean and variance $1/n$, independent of Y, so that

$$
\int_0^1 \Phi\big(u(x,\theta,n)\big)f(\theta)\,d\theta = \mathsf{E}\big[\mathsf{P}\big[Y + Z_n\sqrt{Y(1-Y)} \le x | Y\big]\big]
$$
$$
= \mathsf{P}\big[Y + Z_n\sqrt{Y(1-Y)} \le x\big].
$$

This d.f. (in the x variable) plays an important role also in Mijoule Peccati and Swan [19], where its closeness to F is proved with respect to the Kantorovich distance (see Proposition 4.1

therein). In any case, the proof in Mijoule Peccati and Swan [19] is strongly based on a dual representation of d_W, which does not have any analog for d_K. Therefore, we tackle the problem by a direct computation which, after exchanging the order of conditioning in the above identity and using some elementary algebra, leads to

$$P\left[Y + Z_n\sqrt{Y(1-Y)} \leq x\right] = \int_0^{+\infty} \sqrt{\frac{n}{2\pi}} \exp\left\{-\frac{n}{2}z^2\right\} \cdot \left[F\big(\theta_1(x,z)\big) + F\big(\theta_2(x,z)\big)\right] dz$$

where

$$\theta_1(x,z) := \frac{2x + z^2 - z\sqrt{z^2 + 4x(1-x)}}{2(z^2 + 1)}$$

and

$$\theta_2(x,z) := \frac{2x + z^2 + z\sqrt{z^2 + 4x(1-x)}}{2(z^2 + 1)}$$

It is routine to check that $\theta_1(x,z), \theta_2(x,z) \in [0,1]$ whenever $x \in [0,1]$ and $z > 0$. Whence,

$$\left|\int_0^1 \Phi\big(u(x,\theta,n)\big) f(\theta)\, d\theta - F(x)\right|$$

$$\leq \int_0^{+\infty} \sqrt{\frac{n}{2\pi}} \exp\left\{-\frac{n}{2}z^2\right\} \left|F\big(\theta_1(x,z)\big) + F\big(\theta_2(x,z)\big) - 2F(x)\right| dz$$

so that, introducing $\delta_n := \sqrt{\frac{2\log(n+1)}{n}}$, we have

$$\int_{\delta_n}^{+\infty} \sqrt{\frac{n}{2\pi}} \exp\left\{-\frac{n}{2}z^2\right\} \left|F\big(\theta_1(x,z)\big) + F\big(\theta_2(x,z)\big) - 2F(x)\right| dz$$

$$\leq 2\int_{\delta_n}^{+\infty} \sqrt{\frac{n}{2\pi}} \exp\left\{-\frac{n}{2}z^2\right\} dz \leq \frac{1}{(n+1)\sqrt{\pi\log(n+1)}}. \tag{2.11}$$

It remains to study the integral on $[0, \delta_n]$, by noticing that, after this splitting, we can consider the variable z^2 much smaller than x and $1-x$, whenever $x \in I(n,\gamma)$. More precisely, given $\gamma \in (0,1)$ it is possible to find an integer $N(\gamma) \geq 4$ for which $\delta_n \leq \overline{x}_{n,\gamma}(1 - \overline{x}_{n,\gamma})$ for all $n \geq N(\gamma)$. Therefore, we have that both $\overline{\theta}_1(x,z) := x - z\sqrt{x(1-x)}$ and $\overline{\theta}_2(x,z) := x + z\sqrt{x(1-x)}$ belong to $[0,1]$ whenever $z \in [0, \delta_n]$, $x \in I(n,\gamma)$ and $n \geq N(\gamma)$. We now have

$$\left|F(\theta_1) + F(\theta_2) - 2F(x)\right|$$

$$\leq \left|F(\theta_1) - F(\overline{\theta}_1)\right| + \left|F(\overline{\theta}_1) + F(\overline{\theta}_2) - 2F(x)\right| + \left|F(\overline{\theta}_2) - F(\theta_2)\right|$$

$$\leq \|f\|_\infty \big[|\theta_1 - \overline{\theta}_1| + |\theta_2 - \overline{\theta}_2|\big] + z^2 \left|\frac{F(\overline{\theta}_1) - 2F(x) + F(\overline{\theta}_2)}{z^2}\right|$$

where both $|\theta_1 - \bar{\theta}_1|$ and $|\theta_2 - \bar{\theta}_2|$ are bounded from above by $\frac{3}{2}z^2$. To check this bound, it is enough to consider the quantities

$$\frac{|2x + z^2 \pm z\sqrt{z^2 + 4x(1-x)} - 2x(1+z^2) \mp z(1+z^2)\sqrt{4x(1-x)}|}{2(1+z^2)}$$

which are less than $\frac{|1-2x|}{2}z^2 + \frac{1}{2}z^3 + \frac{1}{2}z[\sqrt{z^2 + 4x(1-x)} - \sqrt{4x(1-x)}]$. The desired result now follows by observing that $|1 - 2x| \le 1$, $z \in (0, 1)$ whenever $n \ge N(\gamma)$, and $\sqrt{z^2 + 4x(1-x)} - \sqrt{4x(1-x)} \le z$. To conclude this argument, we note that $|\frac{F(\bar{\theta}_1) - 2F(x) + F(\bar{\theta}_2)}{z^2}|$ is bounded by virtue of (2.2) with $w = z\sqrt{x(1-x)}$, since $z\sqrt{x(1-x)} < x(1-x)$ holds under the restrictions $z \in [0, \delta_n]$, $x \in I(n, \gamma)$ and $n \ge N(\gamma)$. Whence,

$$\int_0^{\delta_n} \sqrt{\frac{n}{2\pi}} \exp\left\{-\frac{n}{2}z^2\right\} \left|F(\theta_1(x, z)) + F(\theta_2(x, z)) - 2F(x)\right| dz$$

$$\le (3\|f\|_\infty + R(\gamma)|f|_{1,\gamma}) \int_0^{\delta_n} z^2 \sqrt{\frac{n}{2\pi}} \exp\left\{-\frac{n}{2}z^2\right\} dz$$

$$= \frac{1}{2}(3\|f\|_\infty + R(\gamma)|f|_{1,\gamma}) \int_{\mathbb{R}} z^2 \sqrt{\frac{n}{2\pi}} \exp\left\{-\frac{n}{2}z^2\right\} dz$$

$$= \frac{3\|f\|_\infty + R(\gamma)|f|_{1,\gamma}}{2n}. \tag{2.12}$$

To bound the expression in (2.10) that contains H_n, we can exploit the inequalities $|1 - y^2| \le e^{y^2/3}$ and $|S(x)| \le 1/2$, valid for any $x, y \in \mathbb{R}$, to get

$$\left|H_n(y; \theta)\right| \le \frac{2}{3\sqrt{2\pi n\theta(1-\theta)}}e^{-y^2/6} + \frac{\lambda_3}{n\theta(1-\theta)}$$

for all $y \in \mathbb{R}$ and $\theta \in (0, 1)$, where $\lambda_3 := \frac{1}{12\sqrt{2\pi}} \sup_{y \in \mathbb{R}} e^{-y^2/2}|y^3 - 3y|$. Then, after writing

$$\int_0^1 \left|H_n(u(x, \theta, n); \theta)\right| f(\theta) d\theta \le \frac{2\|f\|_\infty}{3\sqrt{2\pi n}} \int_0^1 \frac{e^{-u(x,\theta;n)^2/6}}{\sqrt{\theta(1-\theta)}} d\theta + \frac{\lambda_3}{n} \int_0^1 \frac{f(\theta)}{\theta(1-\theta)} d\theta, \tag{2.13}$$

we have only to show that the former integral on the right-hand side is $O(1/n)$ since, for the latter integral, it is enough to notice that

$$\int_0^1 \frac{f(\theta)}{\theta(1-\theta)} d\theta \le \frac{2^{\gamma+1}}{1-\gamma}|f|_{1,\gamma} \int_0^{1/2} \theta^{-\gamma} d\theta + \frac{2^{\gamma+1}}{1-\gamma}|f|_{1,\gamma} \int_{1/2}^1 (1-\theta)^{-\gamma} d\theta$$

$$\le \frac{2^{\gamma+2}}{(1-\gamma)^2}|f|_{1,\gamma} \tag{2.14}$$

by virtue of the same arguments used to prove point i) of Lemma 2.2. Now, the above-mentioned claim about the former integral follows after checking the boundedness of the expressions

$$\Im_n(a, b; x) := \sqrt{n} \int_a^b \exp\left\{-\frac{n(x-\theta)^2}{6\theta(1-\theta)}\right\} \frac{d\theta}{\sqrt{\theta(1-\theta)}}$$

by letting n and x vary in $\mathbb{N}$ and $I(n, \gamma)$, respectively, where $[a, b]$ coincides with either $[0, 1/2]$ or $[1/2, 1]$. Therefore, taking the former case as reference, we have

$$\Im_n\left(0, \frac{1}{2}; x\right) \leq \sqrt{\frac{n}{2}} \int_0^\infty \exp\left\{-\frac{n(x-\theta)^2}{6\theta}\right\} \frac{d\theta}{\sqrt{\theta}}$$

$$= \sqrt{3} e^{nx/3} \int_0^\infty \exp\left\{-\frac{(\frac{1}{6}nx)^2}{y} - y\right\} \frac{dy}{\sqrt{y}}$$

$$= \sqrt{2}(nx)^{1/2} e^{nx/3} K_{1/2}(nx/3)$$

where $K_{1/2}$ stands for the modified Bessel function of the second kind. See formula 3.471.12 in Gradshtein and Ryzik [16]. Since $x \in I(n, \gamma)$ implies that $nx \geq n^{\frac{1-\gamma}{2-\gamma}} \geq 1$, we notice that the expression $\sqrt{2}(nx)^{1/2} e^{nx/3} K_{1/2}(nx/3)$ is bounded, in view of the asymptotic expansion $K_{1/2}(z) \sim \sqrt{\frac{\pi}{2z}} e^{-z}$, which is valid as $z \to +\infty$. Since an analogous bound holds also for $\Im_n(\frac{1}{2}, 1; x)$, we can combine (2.13)–(2.14) with the analytical study of $\Im_n(a, b; x)$ to obtain that

$$\sup_{x \in I(n,\gamma)} \int_0^1 \left|H_n\left(u(x, \theta, n); \theta\right)\right| f(\theta) \, d\theta \leq \left[\|f\|_\infty + |f|_{1,\gamma}\right] \frac{\lambda_4}{n} \quad (n \geq 4) \qquad (2.15)$$

is valid with a numerical constant λ_4, independent of f and n.

We conclude the first part of the proof with the analysis of the last term on the right-hand side of (2.10). Taking account of (2.9), we immediately realize that the latter summand yields $\frac{\lambda_2 G}{n} \int_0^1 f(\theta)[\theta(1-\theta)]^{-1} d\theta$, which is of order $O(1/n)$ by virtue of (2.14). The study of the former summand in (2.9) is more laborious, and it will be conducted by mimicking the argument used in Ibragimov and Linnik [17] to prove formula (3.3.10). As first step, we borrow from Section 3.3 of Ibragimov and Linnik [17] the explicit expression of the Fourier-Stieltjes transform $d_n(t)$ (see page 101 therein) and we combine it with the formulae displayed in Section VI.1 of Petrov [22], to obtain

$$\hat{D}_n(\xi; \theta) := \int_\mathbb{R} e^{i\xi y} \, d_y D_n(y; \theta)$$

$$= \frac{-\xi}{\sqrt{n\theta(1-\theta)}} \sum_{r \in \mathbb{Z}\setminus\{0\}} \frac{e^{2\pi i r \theta n}}{2\pi r} \exp\left\{-\frac{1}{2}[\xi + 2\pi r\sqrt{n\theta(1-\theta)}]^2\right\}$$

$$\times \left[1 + \frac{1-2\theta}{6\sqrt{n\theta(1-\theta)}}\left(i\xi + 2\pi i r\sqrt{n\theta(1-\theta)}\right)^3\right] \qquad (2.16)$$

where

$$\mathsf{D}_n(y;\theta) := \frac{1}{\sqrt{2\pi n\theta(1-\theta)}} e^{-\frac{1}{2}y^2} S\big(n\theta + y\sqrt{n\theta(1-\theta)}\big)\left[1 + \frac{1-2\theta}{6\sqrt{n\theta(1-\theta)}}(y^3 - 3y)\right].$$

Then, we split the integral in (2.9) into five terms, by dividing the domain $[-n, n]$ into suitable subdomains whose definitions depend on $T_1(n,\theta) := \pi\sqrt{n\theta(1-\theta)}$ and

$$T_2(n,\theta) := \sqrt{\frac{n\theta(1-\theta)}{1-3\theta+3\theta^2}}.$$

We observe that, since $1 - 3\theta + 3\theta^2 \geq 1/4$ for any $\theta \in [0, 1]$, the relation $T_2(n,\theta) \leq T_1(n,\theta)$ is always in force, whereas $T_1(n,\theta) \leq n$ holds whenever $n > \pi^2/4$, which we now assume. Therefore, the desired bound for the integral in (2.9) follows from

$$\int_{-n}^{n} \left| \frac{\hat{\mathsf{B}}_n(\xi;\theta) - \hat{\mathsf{G}}_n(\xi;\theta)}{\xi} \right| d\xi$$

$$\leq \int_{-T_2(n,\theta)}^{T_2(n,\theta)} \left| \frac{\hat{\mathsf{B}}_n(\xi;\theta) - \hat{\mathsf{V}}_n(\xi;\theta)}{\xi} \right| d\xi + \int_{-T_1(n,\theta)}^{T_1(n,\theta)} \left| \frac{\hat{\mathsf{D}}_n(\xi;\theta)}{\xi} \right| d\xi$$

$$+ \int_{\{T_2(n,\theta) \leq |\xi| \leq n\}} \left| \frac{\hat{\mathsf{V}}_n(\xi;\theta)}{\xi} \right| d\xi + \int_{\{T_2(n,\theta) \leq |\xi| \leq T_1(n,\theta)\}} \left| \frac{\hat{\mathsf{B}}_n(\xi;\theta)}{\xi} \right| d\xi$$

$$+ \int_{\{T_1(n,\theta) \leq |\xi| \leq n\}} \left| \frac{\hat{\mathsf{B}}_n(\xi;\theta) - \hat{\mathsf{D}}_n(\xi;\theta)}{\xi} \right| d\xi \tag{2.17}$$

where

$$\hat{\mathsf{V}}_n(\xi;\theta) := \int_{\mathbb{R}} e^{i\xi y}\, d_y\left(\Phi(y) + \frac{1}{\sqrt{2\pi n}} e^{-\frac{1}{2}y^2} \frac{1-2\theta}{6\sqrt{\theta(1-\theta)}}(1-y^2)\right)$$

$$= e^{-\frac{1}{2}\xi^2}\left[1 + \frac{1-2\theta}{6\sqrt{n\theta(1-\theta)}}(i\xi)^3\right].$$

For the derivation of $\hat{\mathsf{V}}(\xi;\theta)$, see Section VI.1 of Petrov [22]. Moreover, with reference to that very same section, we note that the term $\frac{1-2\theta}{\sqrt{\theta(1-\theta)}}$ coincides with the ratio between the third cumulant and the third power of the standard deviation of a centered Bernoulli variable with parameter θ, while $y^2 - 1$ coincides with the Chebyshev–Hermite polynomial of degree 2. In addition, $T_2(n,\theta)$ coincides with the product between $\sqrt{n}$ and the square root of the ratio between the fourth power of the standard deviation and the fourth moment of the same centered Bernoulli variable.

In view of these remarks, we provide a bound for the first integral on the right-hand side of (2.17) by an application of Lemma 4 in Chapter VI of Petrov [22] with $s = 4$, namely

$$\int_{-T_2(n,\theta)}^{T_2(n,\theta)} \left| \frac{\hat{\mathsf{B}}_n(\xi;\theta) - \hat{\mathsf{V}}_n(\xi;\theta)}{\xi} \right| d\xi \leq \lambda_5 \frac{1-3\theta+3\theta^2}{n\theta(1-\theta)} \int_{-T_2(n,\theta)}^{T_2(n,\theta)} \big(|\xi|^3 + |\xi|^9\big) e^{-\frac{1}{12}\xi^2}\, d\xi$$

where λ_5 is a numerical constant specified in the proof of the quoted lemma. Whence,

$$\int_{-T_2(n,\theta)}^{T_2(n,\theta)} \left| \frac{\hat{B}_n(\xi;\theta) - \hat{V}_n(\xi;\theta)}{\xi} \right| d\xi \leq \frac{\lambda_6}{n\theta(1-\theta)} \tag{2.18}$$

where λ_6 is another numerical constant (independent of n and θ).

Then, we study the second integral on the right-hand side of (2.17) by resorting to the explicit expression of $\hat{D}_n(\xi;\theta)$, to obtain

$$\int_{-T_1(n,\theta)}^{T_1(n,\theta)} \left| \frac{\hat{D}_n(\xi;\theta)}{\xi} \right| d\xi \leq \frac{1}{\sqrt{n\theta(1-\theta)}} \sum_{r \in \mathbb{Z} \setminus \{0\}} \frac{1}{2\pi|r|} \left[1 + 6\pi^3 |r|^3 n\theta(1-\theta) \right]$$

$$\times \int_{-T_1(n,\theta)}^{T_1(n,\theta)} \exp\left\{ -\frac{1}{2} [\xi + 2\pi r \sqrt{n\theta(1-\theta)}]^2 \right\} d\xi. \tag{2.19}$$

At this stage, we exploit that $r^2 - |r| \geq \frac{1}{2}r^2$ if $|r| \geq 2$ to write

$$\left[\xi + 2\pi r \sqrt{n\theta(1-\theta)} \right]^2 \geq \xi^2 + \left(2\pi r \sqrt{n\theta(1-\theta)} \right)^2 - 4T_1(n,\theta)\pi |r| \sqrt{n\theta(1-\theta)}$$

$$= \xi^2 + 4\pi^2 (r^2 - |r|) n\theta(1-\theta) \geq \xi^2 + 2\pi^2 r^2 n\theta(1-\theta).$$

After removing the two terms corresponding to $r = \pm 1$, the series on the right-hand side of (2.19) can be bounded by

$$\left(\frac{2}{\pi n\theta(1-\theta)} \right)^{1/2} \sum_{r=2}^{+\infty} \frac{1 + 6\pi^3 r^3 n\theta(1-\theta)}{r} e^{-\pi^2 n\theta(1-\theta)r^2} \tag{2.20}$$

and then, taking cognizance that there is a suitable constant $K(\beta)$ such that $\sum_{r=2}^{+\infty} r^\beta e^{-\lambda r^2} \leq K(\beta)\lambda^{-(\beta+1)/2}$ holds for all $\lambda > 0$ if $\beta \geq 0$, we have that the expression in (2.20) can be bounded by $\lambda_7/[n\theta(1-\theta)]$, where λ_7 is a constant (independent of n and θ). To handle also the terms of the series corresponding to $r = \pm 1$, we take account that $(x \pm 2)^2 \geq \frac{1}{2}x^2 + \frac{1}{3}$ holds for all $x \in [-1, 1]$, to write $[\xi \pm 2\pi \sqrt{n\theta(1-\theta)}]^2 \geq \frac{1}{2}\xi^2 + \frac{\pi^2}{3}n\theta(1-\theta)$ for all $\xi \in [-T_1(n,\theta), T_1(n,\theta)]$. Lastly, the sum of the two terms in (2.19) corresponding to $r = \pm 1$ can be bounded by

$$2 \left(\frac{1}{\pi n\theta(1-\theta)} \right)^{1/2} [1 + 6\pi^3 n\theta(1-\theta)] e^{-\frac{\pi^2}{6}n\theta(1-\theta)}.$$

Therefore, we can conclude that

$$\int_{-T_1(n,\theta)}^{T_1(n,\theta)} \left| \frac{\hat{D}_n(\xi;\theta)}{\xi} \right| d\xi \leq \frac{\lambda_8}{n\theta(1-\theta)} \tag{2.21}$$

holds with a suitable numerical constant λ_8 (independent of n and θ).

As for the third integral on the right-hand side of (2.17), we just use the explicit expression of $\hat{V}_n(\xi; \theta)$ to write

$$\int_{\{T_2(n,\theta)\leq|\xi|\leq n\}}\left|\frac{\hat{V}_n(\xi;\theta)}{\xi}\right|d\xi$$

$$\leq 2\int_{\sqrt{n\theta(1-\theta)}}^{+\infty}\frac{e^{-\frac{1}{2}\xi^2}}{\xi}\,d\xi + \frac{1}{3\sqrt{n\theta(1-\theta)}}\int_{\sqrt{n\theta(1-\theta)}}^{+\infty}\xi^2 e^{-\frac{1}{2}\xi^2}\,d\xi.$$

Using that $x^p e^{-x} \leq (p/e)^p$, which is valid whenever $x, p > 0$, we show that

$$\int_{\{T_2(n,\theta)\leq|\xi|\leq n\}}\left|\frac{\hat{V}_n(\xi;\theta)}{\xi}\right|d\xi \leq \frac{\lambda_9}{n\theta(1-\theta)} \tag{2.22}$$

holds with a suitable numerical constant λ_9 (independent of n and θ).

We now consider the fourth integral on the right-hand side of (2.17). By definition, we have

$$\int_{\{T_2(n,\theta)\leq|\xi|\leq T_1(n,\theta)\}}\left|\frac{\hat{B}_n(\xi;\theta)}{\xi}\right|d\xi$$

$$= \int_{\{T_2(n,\theta)\leq|\xi|\leq T_1(n,\theta)\}}|\xi|^{-1}\left|\mathsf{E}\left[\exp\left\{\frac{i\xi}{\sqrt{n\theta(1-\theta)}}\left(\sum_{j=1}^{n}X_j - n\theta\right)\right\}\Big|Y=\theta\right]\right|d\xi$$

and, after changing the variable by the rule $u = \xi/\sqrt{n\theta(1-\theta)}$ and noticing that $(1 - 3\theta + 3\theta^2)^{-1/2} \geq 1$ is valid for any $\theta \in [0,1]$, we provide the following upper bound

$$\int_{\{1\leq|u|\leq\pi\}}\left|\mathsf{E}\left[\exp\left\{iu\sum_{j=1}^{n}X_j\right\}\Big|Y=\theta\right]\right|du.$$

Now, we just utilize the explicit form of the characteristic function of the binomial distribution with parameters n and θ to write

$$\left|\mathsf{E}\left[\exp\left\{iu\sum_{j=1}^{n}X_j\right\}\Big|Y=\theta\right]\right| = \left|1 - \theta + \theta e^{iu}\right|^n = \left[1 - 2\theta(1-\theta)(1-\cos u)\right]^{n/2}.$$

Whence,

$$\int_{\{T_2(n,\theta)\leq|\xi|\leq T_1(n,\theta)\}}\left|\frac{\hat{B}_n(\xi;\theta)}{\xi}\right|d\xi \leq 2(\pi-1)\left[1 - 2\theta(1-\theta)(1-\cos 1)\right]^{n/2}$$

$$\leq \frac{2(\pi-1)}{e(1-\cos 1)}\frac{1}{n\theta(1-\theta)}. \tag{2.23}$$

To study of the last integral on the right-hand side of (2.17), we introduce the characteristic function $\phi(\cdot;\theta)$ of the r.v. $(X_1 - \theta)$ given $Y = \theta$, that is $\phi(\xi;\theta) = [(1-\theta) + \theta e^{i\xi}]e^{-i\xi\theta}$, so

that we have $\hat{\mathsf{B}}_n(\xi;\theta) = [\phi(\xi/\sqrt{n\theta(1-\theta)};\theta)]^n$. After changing the variable in that integral according to $u = \xi/\sqrt{n\theta(1-\theta)}$ and recalling that $\phi(-\xi;\theta) = \overline{\phi(\xi;\theta)}$, we get

$$
\int_{\{T_1(n,\theta)\leq|\xi|\leq n\}} \left| \frac{\hat{\mathsf{B}}_n(\xi;\theta) - \hat{\mathsf{D}}_n(\xi;\theta)}{\xi} \right| d\xi
$$

$$
= 2 \int_\pi^{\sqrt{n/[\theta(1-\theta)]}} \left| \frac{[\phi(u;\theta)]^n - \hat{\mathsf{D}}_n(u\sqrt{n\theta(1-\theta)};\theta)}{u} \right| du.
$$

At this stage, we introduce the quantity

$$
\bar{r}(n;\theta) := \left\lfloor \frac{1}{2} \left(\frac{1}{\pi} \sqrt{\frac{n}{\theta(1-\theta)}} - 1 \right) \right\rfloor
$$

and we notice that $(2\bar{r}(n;\theta)+1)\pi \leq \sqrt{\frac{n}{\theta(1-\theta)}} < (2\bar{r}(n;\theta)+3)\pi$. In this notation, we have

$$
\int_\pi^{\sqrt{n/[\theta(1-\theta)]}} \left| \frac{[\phi(u;\theta)]^n - \hat{\mathsf{D}}_n(u\sqrt{n\theta(1-\theta)};\theta)}{u} \right| du \leq \sum_{k=1}^{\bar{r}(n;\theta)+1} \mathfrak{J}_k(n;\theta) \qquad (2.24)
$$

where

$$
\mathfrak{J}_k(n;\theta) := \int_{(2k-1)\pi}^{(2k+1)\pi} \left| \frac{[\phi(u;\theta)]^n - \hat{\mathsf{D}}_n(u\sqrt{n\theta(1-\theta)};\theta)}{u} \right| du.
$$

To bound the integrals $\mathfrak{J}_k$'s, we first isolate from the series (2.16) defining $\hat{\mathsf{D}}_n(u\sqrt{n\theta(1-\theta)};\theta)$ the term corresponding to $r = -k$, which reads

$$
\frac{ue^{-2\pi ikn\theta}}{2k\pi} \exp\left\{ -\frac{n\theta(1-\theta)}{2}(u-2k\pi)^2 \right\} \cdot \left[1 + \frac{1-2\theta}{6} n\theta(1-\theta)(iu - 2\pi ki)^3 \right]
$$

$$
=: \hat{\delta}_{n,k}(u;\theta),
$$

so that we obtain

$$
\mathfrak{J}_k(n;\theta) \leq \int_{(2k-1)\pi}^{(2k+1)\pi} \left| \frac{\hat{\mathsf{D}}_n(u\sqrt{n\theta(1-\theta)};\theta) - \hat{\delta}_{n,k}(u;\theta)}{u} \right| du
$$

$$
+ \int_{(2k-1)\pi}^{(2k+1)\pi} \left| \frac{[\phi(u;\theta)]^n - \hat{\delta}_{n,k}(u;\theta)}{u} \right| du
$$

$$
=: \mathfrak{J}_k^{(1)}(n;\theta) + \mathfrak{J}_k^{(2)}(n;\theta).
$$

To analyze $\mathfrak{J}_k^{(1)}(n;\theta)$, we write

$$\mathfrak{J}_k^{(1)}(n;\theta) \le \left(\sum_{r=1}^{\infty} + \sum_{r=-\infty}^{-(k+2)} + \sum_{r\in\{-(k+1)\}} + \sum_{r\in\{-k+1\}} + \sum_{r=-k+2}^{-1} \right) \frac{1}{2\pi|r|}$$
$$\times \int_{(2k-1)\pi}^{(2k+1)\pi} \exp\left\{ -\frac{n\theta(1-\theta)}{2}(u+2\pi r)^2 \right\}$$
$$\times \left[1 + \frac{1}{6}n\theta(1-\theta)|u+2\pi r|^3 \right] du \tag{2.25}$$

with the proviso that both the fourth and the fifth sum are void when $k=1$, and that the fifth sum is void when $k=2$.

To deal with the series in (2.25) limited to $r \in \mathbb{N}$, we observe that $(u+2\pi r)^2 \ge u^2 + 4\pi^2 r^2$ if $u \in [(2k-1)\pi, (2k+1)\pi]$ and we take account that $|z_1 + z_2|^3 \le 4(|z_1|^3 + |z_2|^3)$, so that we deduce, for the series at issue, the upper bound

$$\sum_{r=1}^{\infty} \frac{1}{2\pi r} e^{-2\pi^2 n\theta(1-\theta)r^2}$$
$$\times \int_{(2k-1)\pi}^{(2k+1)\pi} \exp\left\{ -\frac{n\theta(1-\theta)}{2}u^2 \right\} \left[1 + \frac{2}{3}n\theta(1-\theta)(u^3 + 8\pi^3 r^3) \right] du.$$

At this stage, recalling (2.24), we conclude that the sum over the index k of the last expression is majorized by

$$\sum_{r=1}^{\infty} \frac{1}{2\pi r} e^{-2\pi^2 n\theta(1-\theta)r^2} \int_0^{+\infty} \exp\left\{ -\frac{n\theta(1-\theta)}{2}u^2 \right\} \left[1 + \frac{2}{3}n\theta(1-\theta)(u^3 + 8\pi^3 r^3) \right] du$$
$$= \sum_{r=1}^{\infty} \frac{1}{2\pi r} e^{-2\pi^2 n\theta(1-\theta)r^2} \left[\frac{1}{2}\sqrt{\frac{2\pi}{n\theta(1-\theta)}} + \frac{4}{3n\theta(1-\theta)} + \frac{8\sqrt{2\pi}}{3}\pi^3 r^3 \sqrt{n\theta(1-\theta)} \right].$$

Then, we use $x^p e^{-x} \le (p/e)^p$, valid for any $x, p > 0$, with $p = \frac{1}{2} + \varepsilon(\gamma), \varepsilon(\gamma), \frac{3}{2} + \varepsilon(\gamma)$, respectively, and $\varepsilon(\gamma) := \frac{1-\gamma}{2}$, to produce the global bound $S_1(\gamma)[n\theta(1-\theta)]^{-(1+\varepsilon(\gamma))}$ for the last series, where

$$S_1(\gamma) := \frac{1}{2\pi}\left[\frac{\sqrt{2\pi}}{2}\left(\frac{\frac{1}{2}+\varepsilon(\gamma)}{2\pi^2 e} \right)^{\frac{1}{2}+\varepsilon(\gamma)} \zeta\left(2[1+\varepsilon(\gamma)] \right) + \frac{4}{3}\left(\frac{\varepsilon(\gamma)}{2\pi^2 e} \right)^{\varepsilon(\gamma)} \zeta\left(1 + 2\varepsilon(\gamma) \right) \right.$$
$$\left. + \frac{8\pi^3\sqrt{2\pi}}{3}\left(\frac{\frac{3}{2}+\varepsilon(\gamma)}{2\pi^2 e} \right)^{\frac{3}{2}+\varepsilon(\gamma)} \zeta\left(1 + 2\varepsilon(\gamma) \right) \right]$$

$\zeta(\cdot)$ denoting the Riemann zeta function. Now, we come back to (2.25) and we consider the remaining four sums. The change of variable $u = s + 2k\pi$ in the integral and the inequality

$|z_1 + z_2|^3 \leq 4(|z_1|^3 + |z_2|^3)$ lead us to rewrite the expression inside the sums in (2.25) as

$$\frac{1}{2\pi |r|}\left[1 + 6\pi^3 n\theta(1-\theta)|r+k|^3\right]\int_{-\pi}^{\pi} \exp\left\{-\frac{n\theta(1-\theta)}{2}\left[s + 2\pi(k+r)\right]^2\right\} ds. \qquad (2.26)$$

Therefore, for the second series in (2.25), relative to the set $r \leq -(k+2)$, we have

$$\sum_{r=k+2}^{\infty} \frac{1}{2\pi r}\left[1 + 6\pi^3 n\theta(1-\theta)|k-r|^3\right]\int_{-\pi}^{\pi} \exp\left\{-\frac{n\theta(1-\theta)}{2}\left[s + 2\pi(k-r)\right]^2\right\} ds$$

$$\leq \frac{1}{2\pi k}\sum_{h=2}^{\infty}\left[1 + 6\pi^3 n\theta(1-\theta)h^3\right]\int_{-\pi}^{\pi} \exp\left\{-\frac{n\theta(1-\theta)}{2}(s-2\pi h)^2\right\} ds.$$

After noticing that $(s - 2\pi h)^2 \geq s^2 + 2\pi^2 h^2$ for $s \in [-\pi, \pi]$ and $h \geq 2$, we get the new upper bound

$$\frac{1}{2\pi k}\int_{-\pi}^{\pi} \exp\left\{-\frac{n\theta(1-\theta)}{2}s^2\right\} ds \times \sum_{h=2}^{\infty}\left[1 + 6\pi^3 n\theta(1-\theta)h^3\right]e^{-\pi^2 n\theta(1-\theta)h^2}$$

which is less or equal than $\frac{1}{k} S_2(\gamma)[n\theta(1-\theta)]^{-(1+\varepsilon(\gamma))}$ where, by another application of $x^p e^{-x} \leq (p/e)^p$ for $p = \frac{1}{2} + \varepsilon(\gamma)$ and $2 + \varepsilon(\gamma)$, respectively,

$$S_2(\gamma) := \frac{1}{2\pi}\left[\sqrt{2\pi}\left(\frac{\frac{1}{2} + \varepsilon(\gamma)}{\pi^2 e}\right)^{\frac{1}{2}+\varepsilon(\gamma)} \zeta\left(1 + 2\varepsilon(\gamma)\right) \right.$$

$$\left. + 12\pi^4\left(\frac{2 + \varepsilon(\gamma)}{\pi^2 e}\right)^{2+\varepsilon(\gamma)} \zeta\left(2[1 + \varepsilon(\gamma)]\right)\right].$$

For $r = -(k+1)$ the expression in (2.26) is majorized by

$$\frac{1 + 6\pi^3 n\theta(1-\theta)}{k+1} \exp\left\{-\frac{\pi^2 n\theta(1-\theta)}{2}\right\}$$

which is, in turn, less or equal than $\frac{1}{k+1} S_3(\gamma)[n\theta(1-\theta)]^{-(1+\varepsilon(\gamma))}$ with

$$S_3(\gamma) := \left(\frac{2[1 + \varepsilon(\gamma)]}{\pi^2 e}\right)^{1+\varepsilon(\gamma)} + 6\pi^3\left(\frac{2[2 + \varepsilon(\gamma)]}{\pi^2 e}\right)^{2+\varepsilon(\gamma)}.$$

Analogously, for any $k \geq 2$, the expression in (2.26) with $r = -k + 1$ is majorized by

$$\frac{1 + 6\pi^3 n\theta(1-\theta)}{k-1} \exp\left\{-\frac{\pi^2 n\theta(1-\theta)}{2}\right\}$$

which is, in turn, less or equal than $\frac{1}{k-1} S_3(\gamma)[n\theta(1-\theta)]^{-(1+\varepsilon(\gamma))}$. Finally, for $k \geq 3$, it remains to provide an upper bound for the sum of the expression (2.26) as r varies from $-k+2$ to -1.

Changing the variable in the sum, according to $h = k + r$, we obtain the equivalent expression

$$\sum_{h=2}^{k-1} \frac{1}{2\pi(k-h)} \left[1 + 6\pi^3 n\theta(1-\theta)h^3\right] \int_{-\pi}^{\pi} \exp\left\{-\frac{n\theta(1-\theta)}{2}(s+2\pi h)^2\right\} ds$$

which is majorized by virtue of the inequality $(s+2\pi h)^2 \geq s^2 + 2\pi^2 h^2$, valid for any $s \in [-\pi, \pi]$ and $h \geq 2$. Then, we arrive at

$$\int_{-\pi}^{\pi} \exp\left\{-\frac{n\theta(1-\theta)}{2}s^2\right\} ds \times \sum_{h=2}^{k-1} \frac{e^{-\pi^2 n\theta(1-\theta)h^2}}{2\pi(k-h)} \left[1 + 6\pi^3 n\theta(1-\theta)h^3\right]$$

and we now realize that, in view of (2.24), we can exchange the order of summation according to $\sum_{k=3}^{\bar{r}(n;\theta)+1} \sum_{h=2}^{k-1} = \sum_{h=2}^{\bar{r}(n;\theta)} \sum_{k=h+1}^{\bar{r}(n;\theta)+1}$. At this stage, for the inner sum, we have

$$\sum_{k=h+1}^{\bar{r}(n;\theta)+1} \frac{1}{k-h} \leq \sum_{r=1}^{\bar{r}(n;\theta)} \frac{1}{r} \leq \frac{1}{2} \log\left(\frac{n}{\theta(1-\theta)}\right).$$

Using this upper bound, we pass to the outer sum, which is majorized by

$$\int_{-\pi}^{\pi} \exp\left\{-\frac{n\theta(1-\theta)}{2}s^2\right\} ds$$

$$\times \frac{1}{4\pi} \log\left(\frac{n}{\theta(1-\theta)}\right) \sum_{h=2}^{\infty} e^{-\pi^2 n\theta(1-\theta)h^2} \left[1 + 6\pi^3 n\theta(1-\theta)h^3\right].$$

The series in the above expression has been already treated above, yielding the further upper bound

$$\frac{S_2(\gamma)}{2} \log\left(\frac{n}{\theta(1-\theta)}\right) \left[n\theta(1-\theta)\right]^{-(1+\varepsilon(\gamma))}.$$

Therefore, gathering all the bounds that follow formula (2.25), we get

$$\sum_{k=1}^{\bar{r}(n;\theta)+1} \mathfrak{J}_k^{(1)}(n;\theta) \leq \frac{S_1(\gamma)}{[n\theta(1-\theta)]^{(1+\varepsilon(\gamma))}} + \frac{S_2(\gamma) + S_3(\gamma)}{[n\theta(1-\theta)]^{(1+\varepsilon(\gamma))}} \log\left(\frac{n}{\theta(1-\theta)}\right). \qquad (2.27)$$

Now, we pass to analyze the integrals $\mathfrak{J}_k^{(2)}$'s. We start again from the change of variable $u = s + 2k\pi$ and we exploit the fact that $\phi(s + 2k\pi; \theta) = \phi(s; \theta)e^{-2\pi ik\theta}$, to obtain

$$\int_{(2k-1)\pi}^{(2k+1)\pi} \left| \frac{[\phi(u;\theta)]^n - \hat{\delta}_{n,k}(u;\theta)}{u} \right| du$$

$$= \int_{-\pi}^{\pi} \frac{\left| [\phi(s;\theta)]^n - (1 + \frac{s}{2k\pi})\exp\{-\frac{n\theta(1-\theta)}{2}s^2\}[1 + \frac{1-2\theta}{6}n\theta(1-\theta)(is)^3] \right|}{s + 2k\pi} ds.$$

The last integral is majorized by

$$\frac{1}{\pi(2k-1)} \int_{-\pi}^{\pi} \left| [\phi(s;\theta)]^n - \exp\left\{-\frac{n\theta(1-\theta)}{2}s^2\right\} \left[1 + \frac{1-2\theta}{6}n\theta(1-\theta)(is)^3\right] \right| ds$$

$$+ \frac{1}{\pi^2 k(2k-1)} \int_0^{\infty} s \exp\left\{-\frac{n\theta(1-\theta)}{2}s^2\right\} \left[1 + \frac{\pi}{6}n\theta(1-\theta)s^2\right] ds. \tag{2.28}$$

For the first summand in (2.28), we change again the variable according to $s = \xi/\sqrt{n\theta(1-\theta)}$ to obtain the equality with

$$\frac{1}{\pi(2k-1)\sqrt{n\theta(1-\theta)}} \int_{-T_1(n,\theta)}^{T_1(n,\theta)} \left|\hat{\mathsf{B}}_n(\xi;\theta) - \hat{\mathsf{V}}_n(\xi;\theta)\right| d\xi$$

which can be bounded, as before, the splitting the above integral as

$$\int_{-T_2(n,\theta)/4}^{T_2(n,\theta)/4} \left|\hat{\mathsf{B}}_n(\xi;\theta) - \hat{\mathsf{V}}_n(\xi;\theta)\right| d\xi$$

$$+ 2\int_{T_2(n,\theta)/4}^{T_1(n,\theta)} \left|\hat{\mathsf{V}}_n(\xi;\theta)\right| d\xi$$

$$+ 2\int_{T_2(n,\theta)/4}^{T_1(n,\theta)} \left|\hat{\mathsf{B}}_n(\xi;\theta)\right| d\xi.$$

In fact, it is now crucial to observe that the expression

$$T_3(n,\theta,\delta) := \frac{1}{4}\sqrt{n}\left(\frac{[\theta(1-\theta)]^{(3+\delta)/2}}{\theta(1-\theta)[(1-\theta)^{2+\delta} + \theta^{(2+\delta)}]}\right)^{1/(1+\delta)},$$

corresponding to the limitation for $|\xi|$ given in Lemma 2.3 when the V_n's are i.i.d., centered Bernoulli variables, is not less than $T_2(n,\theta)/4$, by virtue of the Hölder inequality. In fact, it is enough to observe that, for a centered r.v. V, we have $\sqrt{\sigma^4/\beta_4} \leq (\sigma^{3+\delta}/\beta_{3+\delta})^{1/(1+\delta)}$ where $\sigma^2 := \mathsf{E}[V^2]$ and $\beta_s := \mathsf{E}[|V|^s]$. Therefore, we can apply Lemma 2.3 with $\delta = 2\varepsilon(\gamma) = 1 - \gamma$, to get

$$\int_{-T_2(n,\theta)/4}^{T_2(n,\theta)/4} \left|\hat{\mathsf{B}}_n(\xi;\theta) - \hat{\mathsf{V}}_n(\xi;\theta)\right| d\xi \leq \frac{\lambda_{10}}{[n\theta(1-\theta)]^{1/2+\varepsilon(\gamma)}}.$$

Since an analogous bound is in force also for $\int_{T_2(n,\theta)/4}^{T_1(n,\theta)} |\hat{\mathsf{V}}_n(\xi;\theta)| d\xi$ and for $\int_{T_2(n,\theta)/4}^{T_1(n,\theta)} |\hat{\mathsf{B}}_n(\xi;\theta)| d\xi$, in view of the argument already used to prove (2.22)–(2.23), we conclude that

$$\sum_{k=1}^{\bar{r}(n;\theta)+1} \frac{1}{\pi(2k-1)} \int_{-\pi}^{\pi} \left| [\phi(s;\theta)]^n - \exp\left\{-\frac{n\theta(1-\theta)}{2}s^2\right\} \cdot \left[1 + \frac{1-2\theta}{6}n\theta(1-\theta)(is)^3\right] \right| ds$$

$$\leq \frac{\lambda_{11}}{[n\theta(1-\theta)]^{1+\varepsilon(\gamma)}} \log\left(\frac{n}{\theta(1-\theta)}\right).$$

As to the latter summand in (2.28), it is enough to notice that it equals

$$\frac{1}{\pi^2 k(2k-1)}\left(2+\frac{\pi}{12}\right)\frac{1}{n\theta(1-\theta)}$$

yielding in the end that

$$\sum_{k=1}^{\bar{r}(n;\theta)+1}\mathfrak{I}_k^{(1)}(n;\theta)\leq\frac{\lambda_{12}}{n\theta(1-\theta)}+\frac{\lambda_{13}}{[n\theta(1-\theta)]^{1+\varepsilon(\gamma)}}\log\left(\frac{n}{\theta(1-\theta)}\right).\qquad(2.29)$$

At this stage, we notice that, for any $\eta>0$, $[\theta(1-\theta)]^\eta\log(\frac{1}{\theta(1-\theta)})$ is bounded by a constant which depends only on η, and we can choose $\eta=\eta(\gamma)=(1-\gamma)/4$. Then, we collect (2.17)–(2.18)–(2.21)–(2.22)–(2.23)–(2.24)–(2.27)–(2.29) to draw the important conclusion that

$$\int_0^1\epsilon_n(\theta)f(\theta)\,\mathrm{d}\theta\leq\lambda_{14}\frac{\|f\|_\infty+|f|_{1,\gamma}}{n}\qquad(2.30)$$

holds with a suitable constant λ_{14} which is independent of n and f, thanks to point (i) in Lemma 2.2 and

$$\int_0^1\frac{1}{[\theta(1-\theta)]^{1+\varepsilon(\gamma)+\eta(\gamma)}}f(\theta)\,\mathrm{d}\theta\leq\frac{2^{3+\gamma+\varepsilon(\gamma)+\eta(\gamma)}}{1-\gamma}M(f).$$

Therefore, the achievement of the bound (2.30) concludes the first part of the proof, culminating in the validity of (1.3) with a suitable constant $C(\mu)$ proportional to $1+\|f\|_\infty+|f|_{1,\gamma}$, under the the additional hypothesis $f(0)=f(1)=0$, thanks to the combination of (2.6)–(2.7), the two bound $9M(f)/n$ for both $\mathsf{F}_n(\bar{x}_{n,\gamma})$ and $1-\mathsf{F}_n(1-\bar{x}_{n,\gamma})$, and (2.10)–(2.11)–(2.12)–(2.15)–(2.30). For completeness, we note that we have proved (1.3) only for $n\geq N_*:=\max\{4,n_0,N(\gamma),\lfloor\pi^2/4\rfloor+1\}$, but now it is immediate to extend the validity of (1.3) to all the set of positive integer: we just add the term N_*/n to the right-hand side of (1.3) and we rename the new constant as $C(\mu)$.

After proving the theorem under the additional hypothesis $f(0)=f(1)=0$, we show how to get rid of this extra-condition. First, we assume that f is given by a polynomial, with generic values of $f(0)$ and $f(1)$, and we apply Lemma 2.1. Since $\mathsf{F}_n(x)=\mathsf{P}[S_n\leq nx]=\int_0^1\mathsf{P}[S_n\leq nx|Y=\theta]\mu(\mathrm{d}\theta)$, we obtain $\mathsf{F}_n(x)=A_\infty\mathsf{F}_{\infty,n}(x)+A_+\mathsf{F}_{+,n}(x)-A_-\mathsf{F}_{-,n}(x)$ for all $x\in[0,1]$, where $\mathsf{F}_{\star,n}(x):=\int_0^1\mathsf{P}[S_n\leq nx|Y=\theta]f_\star(\theta)\,\mathrm{d}\theta$, for $\star=\infty,+$ and $-$, respectively. Whence,

$$\mathrm{d}_K(\mu_n;\mu)\leq A_\infty\sup_{x\in[0,1]}\left|\mathsf{F}_{\infty,n}(x)-\mathsf{F}_\infty(x)\right|+A_+\sup_{x\in[0,1]}\left|\mathsf{F}_{+,n}(x)-\mathsf{F}_+(x)\right|$$
$$+A_-\sup_{x\in[0,1]}\left|\mathsf{F}_{-,n}(x)-\mathsf{F}_-(x)\right|\qquad(2.31)$$

where $\mathsf{F}_\star(x):=\int_0^x f_\star(\theta)\,\mathrm{d}\theta$ for all $x\in[0,1]$ and $\star=\infty,+$ and $-$, respectively. At this stage, from Theorem 2 in Mnatsakanov [20], we can find a constant $C(f_\infty)$, proportional to $1+\|f_\infty\|_\infty+\|f_\infty'\|_\infty$, such that $\sup_{x\in[0,1]}|\mathsf{F}_{\infty,n}(x)-\mathsf{F}_\infty(x)|\leq C(f_\infty)/n$ is in force for all

$n \in \mathbb{N}$. For the last two terms on the right-hand side of (2.31), since $f_{\pm}(0) = f_{\pm}(1) = 0$, the problem is traced back to the first part of the proof. Hence, the inequality (1.3) holds for all $n \in \mathbb{N}$ with a constant $C(\mu)$ proportional to $1 + \|f\|_{\infty} + |f|_{1,\gamma}$, thanks to the bounds provided in points (i)–(ii)–(iii) of Lemma 2.1.

The final act consists in removing the regularity of f by some approximation arguments. First, we start from a probability density f belonging to $C^{1}([0, 1])$ and we consider an approximating family of probability densities $f^{(\delta)}$ expressed by a polynomial which converges to f uniformly with the first derivative, as $\delta \to 0$. See, e.g., Lorentz [18] for classical results about this kind of approximation. Since $\|f^{(\delta)}\|_{\infty} \to \|f\|_{\infty}$ and $|f^{(\delta)}|_{1,\gamma} \to |f|_{1,\gamma}$ are obvious, we pass to analyze the behavior of $d_{K}(\mu_{n}; \mu)$ under the approximation. After fixing n, for any $x \notin \{0, \frac{1}{n}, \ldots, 1\}$, we have

$$\left|F_{n}(x) - F(x)\right| = \lim_{\delta \to 0}\left|F_{n}^{(\delta)}(x) - F^{(\delta)}(x)\right|$$

$$\leq \lim_{\delta \to 0} C \frac{1 + \|f^{(\delta)}\|_{\infty} + |f^{(\delta)}|_{1,\gamma}}{n} = C \frac{1 + \|f\|_{\infty} + |f|_{1,\gamma}}{n},$$

where the inequality follows from the previous argument. This relation entails the validity of (1.3) for all $n \in \mathbb{N}$ and $f \in C^{1}([0, 1])$, with a constant $C(\mu)$ proportional to $1 + \|f\|_{\infty} + |f|_{1,\gamma}$. Finally, the removal of the $C^{1}([0, 1])$-regularity follows by standard arguments based on the convolution of a regularizing kernel.

Appendix

A.1. Proof of Proposition 1.1

We start by dealing with the case of a beta distribution with parameters $(\alpha, 1)$. First, we note that the associated density belongs to $W^{1,\infty}(0, 1)$ if $\alpha \in \{1\} \cup [2, +\infty)$, so that (1.5) follows as a direct application of Theorem 2 in Mnatsakanov [20]. Therefore, we treat the case $\alpha \in (0, 1) \cup (1, 2)$ by starting from the direct computation of the d.f. F_{n} associated to μ_{n}, namely

$$F_{n}(x) = \sum_{k=0}^{\lfloor nx \rfloor} \binom{n}{k} \int_{0}^{1} \theta^{k}(1-\theta)^{n-k} \mu(d\theta) = \int_{0}^{1} \beta\big(y; \lfloor nx \rfloor + 1, n - \lfloor nx \rfloor\big) F(y) \, dy \quad \text{(A.1)}$$

$$= \frac{\Gamma(n+1)}{\Gamma(n+1+\alpha)} \frac{\Gamma(\lfloor nx \rfloor + 1 + \alpha)}{\Gamma(\lfloor nx \rfloor + 1)} \quad \text{(A.2)}$$

for all $x \in (0, 1)$, where β is the same as in (1.4) and $\lfloor \cdot \rfloor$ denotes the integral part. For the validity of the second identity in (A.1), see formulae (13)–(14) in Mnatsakanov [20], or Problems 44–45 at the end of Chapter VI of Feller [12]. Now, for $\alpha \in (0, 1)$, we invoke *Wendell's inequalities* (see formula (5) in Qi and Luo [23]) to obtain

$$L_{\alpha}\big(\lfloor nx \rfloor, n\big) \leq F_{n}(x) \leq U_{\alpha}\big(\lfloor nx \rfloor, n\big)$$

for all $n \in \mathbb{N}$ and $x \in (0, 1)$, where $L_\alpha(\lfloor nx \rfloor, n) := (\frac{\lfloor nx \rfloor + 1}{\lfloor nx \rfloor + 1 + \alpha})^{1-\alpha}(\frac{\lfloor nx \rfloor + 1}{n+1})^\alpha$ and $U_\alpha(\lfloor nx \rfloor, n) := (\frac{n+1+\alpha}{n+1})^{1-\alpha}(\frac{\lfloor nx \rfloor + 1}{n+1})^\alpha$. Then, we observe that we can write

$$
d_K(\mu_n; \mu) = \max_{k \in \{1,\dots,n\}} \sup_{x \in [\frac{k-1}{n}, \frac{k}{n})} \left| F_n(x) - F(x) \right|
$$

$$
= \max_{k \in \{1,\dots,n\}} \sup_{x \in [\frac{k-1}{n}, \frac{k}{n})} \left| F_n\left(\frac{k-1}{n}\right) - x^\alpha \right|
$$

and that, for any $k \in \{1, \dots, n\}$,

$$
\sup_{x \in [\frac{k-1}{n}, \frac{k}{n})} \left| F_n\left(\frac{k-1}{n}\right) - x^\alpha \right|
$$

$$
\leq \left[U_\alpha(k-1, n) - L_\alpha(k-1, n) \right] + \left| U_\alpha(k-1, n) - \left(\frac{k}{n+1}\right)^\alpha \right|
$$

$$
+ \left[\left(\frac{k}{n}\right)^\alpha - \left(\frac{k}{n+1}\right)^\alpha \right] + \left[\left(\frac{k}{n}\right)^\alpha - \left(\frac{k-1}{n}\right)^\alpha \right]. \tag{A.3}
$$

For the first two summands on the above right-hand side, we can write

$$
\left[U_\alpha(k-1, n) - L_\alpha(k-1, n) \right] + \left| U_\alpha(k-1, n) - \left(\frac{k}{n+1}\right)^\alpha \right|
$$

$$
\leq 2\left[\left(\frac{n+1+\alpha}{n+1}\right)^{1-\alpha} - 1 \right] + \left(\frac{k}{n+1}\right)^\alpha \left[1 - \left(\frac{k}{k+\alpha}\right)^{1-\alpha} \right]. \tag{A.4}
$$

At this stage, we observe that $(\frac{n+1+\alpha}{n+1})^{1-\alpha} - 1 \leq \frac{\alpha(1-\alpha)}{n+1}$ holds for all $n \in \mathbb{N}$, while for the latter summand on the right-hand side of (A.4) we get

$$
\left(\frac{k}{n+1}\right)^\alpha \left[1 - \left(\frac{k}{k+\alpha}\right)^{1-\alpha} \right] \leq \alpha(1-\alpha)(1+\alpha)^\alpha \left(\frac{1}{n+1}\right)^\alpha.
$$

Moreover, for the third summand on the right-hand side of (A.3) we have $(\frac{k}{n})^\alpha - (\frac{k}{n+1})^\alpha \leq \frac{\alpha}{n}$, while for the last summand on the right-hand side of the same relation we obtain $(\frac{k}{n})^\alpha - (\frac{k-1}{n})^\alpha \leq (\frac{1}{n})^\alpha$. Putting these bounds together via (A.3)–(A.4), we get (1.5) for $\alpha \in (0, 1)$. When $\alpha \in (1, 2)$, we start again from (A.2), which can be equivalently rewritten as

$$
\frac{\Gamma(n+1)}{\Gamma(n+1+\delta)} \frac{\Gamma(\lfloor nx \rfloor + 1 + \delta)}{\Gamma(\lfloor nx \rfloor + 1)} \frac{\lfloor nx \rfloor + 1 + \delta}{n+1+\delta}
$$

with $\delta := \alpha - 1$. Whence,

$$
\frac{\lfloor nx \rfloor + \alpha}{n + \alpha} L_\delta(\lfloor nx \rfloor, n) \leq F_n(x) \leq \frac{\lfloor nx \rfloor + \alpha}{n + \alpha} U_\delta(\lfloor nx \rfloor, n)
$$

for all $n \in \mathbb{N}$ and $x \in (0, 1)$. Then, we can establish a bound similar to (A.3), namely

$$
\sup_{x \in [\frac{k-1}{n}, \frac{k}{n})} \left| \mathsf{F}_n\left(\frac{k-1}{n}\right) - x^\alpha \right|
$$

$$
\leq \frac{k+\delta}{n+\alpha} \left[U_\delta(k-1, n) - L_\delta(k-1, n) \right] + \frac{k+\delta}{n+\alpha} \left| U_\delta(k-1, n) - \left(\frac{k}{n+1}\right)^\delta \right|
$$

$$
+ \frac{k+\delta}{n+\alpha} \left[\left(\frac{k}{n}\right)^\delta - \left(\frac{k}{n+1}\right)^\delta \right] + \left(\frac{k}{n}\right)^\delta \left| \frac{k+\delta}{n+\alpha} - \frac{k}{n} \right|
$$

$$
+ \left[\left(\frac{k}{n}\right)^\alpha - \left(\frac{k-1}{n}\right)^\alpha \right]. \tag{A.5}
$$

We analyze the first two summands on the above right-hand side by resorting to (A.4), to obtain

$$
\frac{k+\delta}{n+\alpha} \left[U_\delta(k-1, n) - L_\delta(k-1, n) \right] + \frac{k+\delta}{n+\alpha} \left| U_\delta(k-1, n) - \left(\frac{k}{n+1}\right)^\delta \right|
$$

$$
\leq 2 \left[\left(\frac{n+1+\delta}{n+1}\right)^{1-\delta} - 1 \right] + 2 \left(\frac{k}{n+1}\right)^\alpha \left[1 - \left(\frac{k}{k+\delta}\right)^{1-\delta} \right].
$$

Now, the former summand on the above right-hand side has been already bounded by $\frac{2\alpha(1-\alpha)}{n+1}$, so that we can focus the attention on the latter. Arguing as above, we get

$$
2 \left(\frac{k}{n+1}\right)^\alpha \left[1 - \left(\frac{k}{k+\delta}\right)^{1-\delta} \right] \leq 2\delta(1-\delta)(1+\delta)^\delta \frac{1}{n+1}.
$$

Lastly, the very same arguments used to handle the case $\alpha \in (0, 1)$ lead to conclude that also the last three terms on the right-hand side of (A.5) are bounded from above by a term of the type C_α/n for some constant C_α independent of k. The proof of (1.5) is therefore complete in the case that the prior is a beta distribution with parameters $(\alpha, 1)$.

The case of a beta distribution with parameters $(1, \alpha)$ is easily reformulated in terms a beta distribution with parameters $(\alpha, 1)$, in view of the following symmetry argument. First, we note that the d.f. $\mathsf{F}(x)$ of a beta with parameters $(1, \alpha)$ coincides with $1 - \mathsf{F}^*(1 - x)$, where F^* is the d.f. of a beta with parameters $(\alpha, 1)$. An analogous argument is true for the d.f. $\mathsf{F}_n(x)$, in the sense that it coincides, for all $x \in [0, 1] \setminus \{0, \frac{1}{n}, \frac{2}{n}, \ldots, 1\}$, with the d.f. $\mathsf{F}_n^*(x)$ of the r.v. $\frac{1}{n} \sum_{i=1}^n \overline{X}_i$, where $\overline{X}_i := 1 - X_i$ for all $i \in \mathbb{N}$. The conclusion is reached by observing that the de Finetti measure of the (exchangeable) sequence $\{\overline{X}_i\}_{i \geq 1}$ is exactly the beta with parameters $(\alpha, 1)$, whenever the de Finetti measure of the sequence $\{X_i\}_{i \geq 1}$ is the beta with parameters $(1, \alpha)$, and that

$$
d_K(\mu_n; \mu) = \sup_{x \in [0,1] \setminus \{0, \frac{1}{n}, \frac{2}{n}, \ldots, 1\}} \left| \mathsf{F}_n(x) - \mathsf{F}(x) \right| = \sup_{x \in [0,1] \setminus \{0, \frac{1}{n}, \frac{2}{n}, \ldots, 1\}} \left| \mathsf{F}_n^*(x) - \mathsf{F}^*(x) \right|
$$

$$
= \sup_{x \in [0,1]} \left| \mathsf{F}_n^*(x) - \mathsf{F}^*(x) \right|.
$$

A.2. Proof of Proposition 1.2

By resorting once again to formula (A.1) in the proof of Proposition 1.1, we get

$$
d_K(\mu_n; \mu) \leq \sup_{x \in [0,1)} \int_0^1 \beta\big(\theta; \lfloor nx \rfloor + 1, n - \lfloor nx \rfloor\big) \big| F(x) - F(\theta) \big| \, d\theta
$$

$$
\leq H_\gamma(F) \sup_{x \in [0,1)} \int_0^1 \beta\big(\theta; \lfloor nx \rfloor + 1, n - \lfloor nx \rfloor\big) |x - \theta|^\gamma \, d\theta
$$

where $H_\gamma(F)$ denotes the Hölder constant of F. Now, we exploit that $|x - \theta|^\gamma \leq |x - \eta_x|^\gamma + |\eta_x - \theta|^\gamma$, where $\eta_x := \frac{\lfloor nx \rfloor + 1}{n+1} = \int_0^1 \theta \beta(\theta; \lfloor nx \rfloor + 1, n - \lfloor nx \rfloor) \, d\theta$, to get

$$
d_K(\mu_n; \mu) \leq H_\gamma(F) \left[\sup_{x \in [0,1)} |x - \eta_x|^\gamma + \sup_{x \in [0,1)} \int_0^1 |\eta_x - \theta|^\gamma \beta\big(\theta; \lfloor nx \rfloor + 1, n - \lfloor nx \rfloor\big) \, d\theta \right].
$$

Since $\sup_{x \in [0,1)} |x - \eta_x| \leq \frac{2}{n+1}$ follows from direct computation, we can focus on the second summand on the above right-hand side, which can be bounded by means of the Jensen inequality as follows:

$$
\int_0^1 |\eta_x - \theta|^\gamma \beta\big(\theta; \lfloor nx \rfloor + 1, n - \lfloor nx \rfloor\big) \, d\theta \leq \left(\int_0^1 |\eta_x - \theta|^2 \beta\big(\theta; \lfloor nx \rfloor + 1, n - \lfloor nx \rfloor\big) \, d\theta \right)^{\gamma/2}.
$$

The proof is completed by observing that the integral $\int_0^1 |\eta_x - \theta|^2 \beta(\theta; \lfloor nx \rfloor + 1, n - \lfloor nx \rfloor) \, d\theta$ represents the variance of the beta distribution with parameters $(\lfloor nx \rfloor + 1, n - \lfloor nx \rfloor)$, which, being equal to $\frac{(\lfloor nx \rfloor + 1)(n - \lfloor nx \rfloor)}{(n+1)^2(n+2)}$, is less than $\frac{1}{n+2}$ for any $x \in [0, 1]$.

Acknowledgements

References

[1] Aldous, D.J. (1985). Exchangeability and related topics. In *École D'été de Probabilités de Saint-Flour, XIII*—1983. *Lecture Notes in Math.* **1117** 1–198. Berlin: Springer. MR0883646 https://doi.org/10.1007/BFb0099421

[2] Brezis, H. (2011). *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. *Universitext*. New York: Springer. MR2759829

[3] Chow, Y.S. and Teicher, H. (1997). *Probability Theory*: *Independence*, *Interchangeability*, *Martingales*, 3rd ed. *Springer Texts in Statistics*. New York: Springer. MR1476912 https://doi.org/10.1007/978-1-4612-1950-7

[4] de Finetti, B. (1930). Funzione caratteristica di un fenomeno aleatorio. *Atti R. Accad. Naz. Lincei* **4** 86–133.

[5] de Finetti, B. (1933). La legge dei grandi numeri nel caso dei numeri aleatori equivalenti. *Atti R. Accad. Naz. Lincei* **18** 203–207.

[6] Diaconis, P. and Freedman, D. (2004). The Markov moment problem and de Finetti's theorem. I. *Math. Z.* **247** 183–199. MR2054525 https://doi.org/10.1007/s00209-003-0636-6

[7] Diaconis, P. and Freedman, D. (2004). The Markov moment problem and de Finetti's theorem. II. *Math. Z.* **247** 201–212. MR2054526 https://doi.org/10.1007/s00209-003-0636-6

[8] Döbler, C. (2015). Stein's method of exchangeable pairs for the beta distribution and generalizations. *Electron. J. Probab.* **20** no. 109, 34. MR3418541 https://doi.org/10.1214/EJP.v20-3933

[9] Dolera, E., Gabetta, E. and Regazzini, E. (2009). Reaching the best possible rate of convergence to equilibrium for solutions of Kac's equation via central limit theorem. *Ann. Appl. Probab.* **19** 186–209. MR2498676 https://doi.org/10.1214/08-AAP538

[10] Dolera, E. and Regazzini, E. (2010). The role of the central limit theorem in discovering sharp rates of convergence to equilibrium for the solution of the Kac equation. *Ann. Appl. Probab.* **20** 430–461. MR2650038 https://doi.org/10.1214/09-AAP623

[11] Esseen, C.-G. (1945). Fourier analysis of distribution functions. A mathematical study of the Laplace–Gaussian law. *Acta Math.* **77** 1–125. MR0014626 https://doi.org/10.1007/BF02392223

[12] Feller, W. (1968). *An Introduction to Probability Theory and Its Applications*. *Vol. I. Third Edition*. New York: Wiley. MR0228020

[13] Gibbs, A.L. and Su, F.E. (2002). On choosing and bounding probability metrics. *Int. Stat. Rev.* **70** 419–435.

[14] Gnedenko, B.V. and Kolmogorov, A.N. (1968). *Limit Distributions for Sums of Independent Random Variables*. *Translated from the Russian*, *Annotated*, *and Revised by K. L. Chung*. *With Appendices by J. L. Doob and P. L. Hsu. Revised Edition*. Reading: Addison-Wesley. MR0233400

[15] Goldstein, L. and Reinert, G. (2013). Stein's method for the beta distribution and the Pólya–Eggenberger urn. *J. Appl. Probab.* **50** 1187–1205. MR3161381 https://doi.org/10.1239/jap/1389370107

[16] Gradshteyn, I.S. and Ryzhik, I.M. (1965). *Table of Integrals*, *Series*, *and Products*. *Fourth Edition Prepared by Ju. V. Geronimus and M. Ju. Ceĭtlin*. *Translated from the Russian by Scripta Technica*, *Inc*. *Translation Edited by Alan Jeffrey*. New York: Academic Press. MR0197789

[17] Ibragimov, I.A. and Linnik, Yu.V. (1971). *Independent and Stationary Sequences of Random Variables*. Groningen: Wolters-Noordhoff. MR0322926

[18] Lorentz, G.G. (1986). *Bernstein Polynomials*, 2nd ed. New York: Chelsea. MR0864976

[19] Mijoule, G., Peccati, G. and Swan, Y. (2016). On the rate of convergence in de Finetti's representation theorem. *ALEA Lat. Am. J. Probab. Math. Stat.* **13** 1165–1187. MR3582913

[20] Mnatsakanov, R.M. (2008). Hausdorff moment problem: Reconstruction of distributions. *Statist. Probab. Lett.* **78** 1612–1618. MR2453814 https://doi.org/10.1016/j.spl.2008.01.011

[21] Osipov, L.V. (1969). Asymptotic expansions of the distribution function of a sum of independent lattice random variables. *Teor. Veroyatn. Primen.* **14** 468–475. MR0259992

[22] Petrov, V.V. (1975). *Sums of Independent Random Variables*. New York: Springer. MR0388499

[23] Qi, F. and Luo, Q.-M. (2013). Bounds for the ratio of two gamma functions: From Wendel's asymptotic relation to Elezović–Giordano–Pečarić's theorem. *J. Inequal. Appl.* **2013** 542. MR3339476 https://doi.org/10.1186/1029-242x-2013-542

[24] Ressel, P. (1985). De Finetti-type theorems: An analytical approach. *Ann. Probab.* **13** 898–922.
     MR0799427