Mediterr. J. Math. (2022) 19:240
https://doi.org/10.1007/s00009-022-02150-y
1660-5446/22/050001-20
published online September 19, 2022
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2022

Mediterranean Journal of Mathematics

Stochastic Quasi-Interpolation with Bernstein Polynomials

Xingping Sun and Xuan Zhou

Abstract. We introduce the notion "stochastic quasi-interpolation on compact Hausdorff spaces", and establish Gaussian-type L^p-concentration inequalities ($1 \leq p \leq \infty$) for stochastic Bernstein polynomials in terms of the modulus of continuity of a target function $f \in C[0,1]$. For p in the range $1 \leq p < \infty$, these inequalities hold true unconditionally in the sense that no additional assumption on a given target function is required. For the case $p = \infty$, our proof calls for a crucial application of Dvoretzky–Kiefer–Wolfowitz inequality (Dvoretzky et al. in Ann Math Stat 27(3):642–669, 1956; Massart in Ann Probab 18(3):1269–1283, 1990) , and requires a moderate decay condition on the modulus of continuity. Our result for the case $p = \infty$ confirms a similar conjecture raised in Sun and Wu (Proc Am Math Soc 147(2):671–679, 2019). As a corollary, we show that for all $1 \leq p \leq \infty$ the expected L^p-approximation order of stochastic Bernstein polynomials is comparable to that given by the classical Bernstein polynomials.

Mathematics Subject Classification. 60G20, 60H30, 41A17, 41A35.

Keywords. Dvoretzky–Kiefer–Wolfowitz inequality, modulus of continuity, order statistics, stochastic quasi-interpolation.

1. Introduction

Let Ω be a compact Hausdorff space, and let $\mathcal{R}_p(\Omega)$ denote the totality of Radon probability measures on Ω. For an $f \in C(\Omega)$ and $n \in \mathbb{N}$, we prescribe a set of n functions $\psi_{n,1}, \ldots, \psi_{n,n} \in C(\Omega)$, and define *stochastic quasi-interpolants* $Q_n^X(f)$ of f by

$$Q_n^X(f)(x) := \sum_{j=1}^{n} f(X_{n,j})\psi_{n,j}(x), \quad x \in \Omega.$$

Here $\{f(X_{n,j})\psi_{n,j}\}_{j=1}^n$ are $C(\Omega)$-valued random variables obeying respectively laws $\{\nu_{n,j}\}_{j=1}^n \subset \mathcal{R}_p(\Omega)$, the design of which caters to problems on

hand and may require some ingenuity. Ideally, one selects functions $\psi_{n,1}, \ldots,$ $\psi_{n,n} \in C(\Omega)$ following the "partition of unity" principle. That is,

$$\sum_{j=1}^{n} \psi_{n,j}(x) = 1, \quad x \in \Omega, \quad n \in \mathbb{N}.$$

In many theoretical and practical problems, however, an exact partition of unity may be either hard to obtain or cumbersome to implement. One then uses a set of functions that forms an *approximate partition of unity* in the sense that the sequence of functions $\sum_{j=1}^{n} \psi_{n,j}(x)$ approximates 1 within a desirable error bound under an appropriately-selected topology on $C(\Omega)$. We will call this general type of stochastic approximation scheme "stochastic quasi-interpolation".

Let μ be the uniform probability measure on Ω as defined by Niederreiter [27]. For $1 \leq p \leq \infty$, let $L^p(\Omega)$ denote the Banach spaces consisting of all real-valued Borel measurable functions f on Ω for which $\|f\|_p < \infty$. Here $\|f\|_p := \|f\|_{L^p(\Omega)}$ is defined by:

$$\|f\|_{L^p(\Omega)} = \begin{cases} \left(\int_\Omega |f(x)|^p \mathrm{d}\mu(x) \right)^{1/p}, & \text{if } 1 \leq p < \infty, \\ \inf\{C : |f(x)| \leq C \text{ almost surely with respect to } \mu\}, & \text{if } p = \infty. \end{cases}$$

Let $\mathbb{P}(A)$ and $\mathbb{E}(Z)$ denote, respectively, the probability of the event A and the expectation of the random variable Z. A coveted type of estimates in the study of stochastic quasi-interpolation is Gaussian-style L^p-concentration inequalities for some target functions f of interest: For any given $\epsilon >$, there holds

$$\mathbb{P}\{\|Q_n^X(f) - f\|_\infty > \epsilon\} \leq c_1 \exp\left(-c_2 n^\alpha \ \epsilon^2\right).$$

Here $\alpha \geq 1$, and c_1, c_2 are absolute constants, which can be made explicit with efforts. As a corollary of the above inequality, we derive

$$\mathbb{E}\left(\|Q_n(f) - f\|_\infty\right) \leq c_3 \ n^{-\alpha/2}, \tag{1.1}$$

where c_3 is an absolute constant depending on c_1 and c_2. This amounts to saying that the "average approximation order" of the stochastic quasi-interpolation scheme is: $n^{-\alpha/2}$. Of course, one needs to carry out "Bochner integrals" on the underlying Banach space valued random variables en route. Under certain circumstances, it may be impossible to obtain the desired L^∞-concentration inequalities. One then settles on the next best thing: For some p in $[1, \infty)$, and any given $\epsilon > 0$, there holds

$$\mathbb{P}\{\|Q_n^X(f) - f\|_p > \epsilon\} \leq c_1 \exp\left(-c_2 n^\alpha \ \epsilon^2\right).$$

Due to the compactness of Ω, if the above inequality holds true for a $p_1 \in [1, \infty)$, then it does for all p_2 in the range $1 \leq p_2 \leq p_1$. Many theoretical and practical problems may be modeled under this stochastic quasi-interpolation framework. Literature abounds in implicit applications of stochastic quasi-interpolation techniques. Notably, Wagner [37], and Bourgain and Lindenstrauss [7] applied a stochastic quasi-interpolation method to show the existence of certain node sets giving rise to a near optimal number of segments

in Minkowski sums with equal length needed to approximate a euclidean ball under the Hausdorff metric within a prescribed error bound; see also [5,6]. Gao et al [21] studied the case in which $X_{n,1}, \ldots, X_{n,n}$ are n-independent copies of the random variables uniformly distributed in Ω.

To mitigate uncertainties brought on by unreliable sampling sites, Wu et al. [38] proposed and studied a class of stochastic Bernstein polynomials

$$(B_n^X f)(x) := \sum_{k=0}^{n} f(X_{n,k})\ p_{n,k}(x), \quad n \in \mathbb{N}, \tag{1.2}$$

in which $X_{n,0}, X_{n,1}, \cdots, X_{n,n}$ are the *order statistics* (see [9,14]) of $(n+1)$ independent copies of the random variable uniformly distributed in $(0,1)$, and

$$p_{n,k}(x) = \binom{n}{k} x^k (1-x)^{n-k}, \quad 0 \le k \le n, \quad 0 \le x \le 1.$$

These are stochastic cousins of the classical Bernstein polynomial $B_n f$ defined by

$$(B_n f)(x) := \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\ p_{n,k}(x), \quad n \in \mathbb{N}. \tag{1.3}$$

Stochastic Bernstein polynomials have a simple structure and therefore are nimble in a wide range of applications. Furthermore, the inherent randomness makes them suitable for Monte Carlo simulations [34]. Authors of [1,20,33,39] have established algebraic and exponential convergence rates for probabilistic convergence of these stochastic Bernstein polynomials.

Authors of [34] introduced the notion "L^p-probabilistic convergence" $(1 \le p \le \infty)$ of stochastic Bernstein polynomials (Definition 2.1) and established various L^p-probabilistic convergence rates, a highlight of which is as follows. For any given $\varepsilon > 0$, p in the range $1 \le p \le 2$, and any $f \in C[0,1]$, the following inequality:

$$\mathbb{P}\left\{\left\|B_n^X f - f\right\|_p > \varepsilon\right\} \le 2\exp\left[-\frac{\varepsilon^2}{4\ \omega^2\left(f, \frac{1}{\sqrt{n}}\right)}\right], \tag{1.4}$$

holds true under the assumption that

$$\omega\left(f, \frac{1}{\sqrt{n}}\right) < \frac{\varepsilon}{6.2}, \tag{1.5}$$

in which $\omega(f, \cdot)$ denotes the modulus of continuity of f defined by

$$\omega(f,h) := \sup_{\substack{0 \le x, y \le 1 \\ |x-y| \le h}} |f(x) - f(y)|, \quad 0 \le h \le 1.$$

These are Gaussian-type concentration inequalities dressed in modulus of continuity, which in the current article will be called L^p-concentration inequalities (or simply concentration inequalities) for stochastic Bernstein polynomials. However, both the restriction on p $(1 \le p \le 2)$ and the assumption (1.5) are impractical. Concerning the latter, we have received a constructive feedback from practitioners in the field, a gist of which goes as follows.

In modeling many real-world problems, one often finds a priori information on the smoothness of an unknown target function elusive. Thus, assumption (1.5) is sometimes impossible to verify beforehand, which has hindered application and foreshadowed the effectiveness of estimate (1.4).

The proof given in [34] depends on assumption (1.5) because it requires harnessing the approximation power of the classical Bernstein polynomial $B_n f$. In so doing, authors of [34] have inadvertently followed a familiar methodology in a learning theory paradigm advocated by Cucker and Smale [12], and Cucker and Zhou [13], which decomposes an overall error in two parts: the approximation error and the sampling error. Another caveat of the approach is that it implicitly implies that under no circumstance, will any other type of Bernstein polynomials (spawned from their stochastic cousins) outperform the classical ones. But this is not true. Guided by theoretical analysis, we carried out many rounds of numerical simulations, which show that Bernstein polynomials built upon some carefully-designed triangular arrays of knots converge to a Heaviside-type function exponentially fast in the space $L^1[0,1]$. We will study this interesting problem in the near future.

Approximation theorists are interested in knowing the average approximation order of the stochastic Bernstein polynomials (1.2). For this goal, assumption (1.5) is nearly harmless. The main result of Sect. 4 (Theorem 4.2) asserts that inequality (1.4) holds true for all $1 \le p \le \infty$ under a similar assumption to (1.5). This implies the following inequality (see Corollary 4.3):

$$\mathbb{E}\left(\left\|B_n^X f - f\right\|_p\right) \le C\ \omega\left(\frac{1}{\sqrt{n}}\right), \quad 1 \le p \le \infty.$$

where C is an absolute constant, which will be specified in Corollary 4.3, and $\omega(h)$ is an abbreviation for $\omega(f,h)$. (We will use the abbreviation throughout the article.) The above inequality implies that the expected (average) L^p-approximation order ($1 \le p \le \infty$) of stochastic Bernstein polynomials is the same as that achieved by the classical Bernstein polynomial $B_n f$.[1] This implication does not depend on assumption (1.5).

In summary, there are good arguments to be made on both sides (with or without assumption (1.5)) in so far as L^p-concentration inequalities are developed. Henceforth, we will use the phrase "unconditional L^p-concentration inequalities" to refer to estimates obtained without using assumption (1.5). To facilitate broader applications of stochastic Bernstein polynomials, we will push forward as much as we can to prove unconditional L^p-concentration inequalities, using assumption (1.5) only as the last resort. The departure of the assumption (1.5) has increased the level of difficulty. For p in the range $1 \le p \le 2$, Gaussian bounds of mean square beta distributions [3, pp. 125] play a key role in our proof. But the same method does not carry over to the cases $2 < p < \infty$ for which our proof is much involved, calling for a host of techniques including a crucial application of Dvoretzky–Kiefer–Wolfowitz

[1] Let us take a brief moment off the main topic here to reflect on making a list of standard criteria in evaluating the robustness of a stochastic approximation operator. We suggest it be on the list that the expected approximation power of the stochastic operator matches that of a standard deterministic operator.

inequality [17,26]. Finally, we prove an L^∞-concentration inequality using a slightly weaker version of assumption (1.5).

Dvoretzky et al. [17] proved the following remarkable result. Let $X_1, X_2, \ldots, X_n$ be real-valued independent and identically distributed random variables with c.d.f (cumulative distribution function) F. Let F_n denote the empirical distribution function associated with F defined by

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{\{X_i \le x\}}, \qquad x \in \mathbb{R}.$$

Then for any given $\varepsilon > 0$, there holds:

$$\mathbb{P}\Big(\sup_{x\in\mathbb{R}} |F_n(x) - F(x)| > \varepsilon\Big) \le 2\, e^{-2n\varepsilon^2}. \tag{1.6}$$

We remark that Dvoretzky, Kiefer, and Wolfowitz only established the above inequality with an unspecified multiplicative constant C. The inequality with the sharp constant $C = 2$ was proved by Massart [26], which confirms a conjecture due to Birnbaum and McCarty [2], and has since been known as Dvoretzky–Kiefer–Wolfowitz inequality.

The current investigation has drawn inspirations from previous research done in the theory of random approximation of functions; see [10,18,19,23,24,29,30,36], of which we mention particularly the idea of employing Bernstein polynomials for density and distribution estimation on the interval $[0, 1]$:

$$\widetilde{F}_m(x) = \sum_{k=0}^{n} \widehat{F}_m\left(\frac{k}{n}\right)\, p_{n,k}(x),$$

where $m, n \in \mathbb{N}$, and $\widehat{F}_m$ is the empirical distribution of order m. There are close relationships as well as fundamental differences between the two classes of stochastic Bernstein polynomials. Addressing these interesting topics, however, will take us far afield.

The arrangement of the current article is as follows. In Sect. 2 , we prove unconditional L^p-concentration inequalities for stochastic Bernstein polynomials (1.2) for p in the range $1 \le p \le 2$. Among other results, we obtain the unconditional version of inequality (1.4) using theory of mean square beta distribution. In Sect. 3, we prove unconditional L^p-concentration inequalities for p in the range $2 < p < \infty$. In Sect. 4, we prove an L^∞-concentration inequality with an assumption similar to (1.5). As a corollary, we show that for all p in the range $1 \le p \le \infty$, the average L^p approximation order of stochastic Bernstein polynomials is the same as that achieved by the classical Bernstein polynomial $B_n f$ under a mild restriction.

2. Mean Square Beta Distribution and Unconditional L^p-Concentration Inequalities

Let X be the random variable uniformly distributed in $(0, 1)$. Let $X_{n,0}, X_{n,1}, \ldots, X_{n,n}$ be the *order statistics* of $(n + 1)$ independent copies of X. Then

$X_{n,k}$ is distributed according to the beta distribution $\mathrm{Beta}(k+1, n-k+1)$ with parameters $(k+1)$ and $(n-k+1)$ that has density function

$$(n+1)\, p_{n,k}(x), \quad 0 \le x \le 1, \quad 0 \le k \le n.$$

Beta distributions belong to the wider class of sub-Gaussian distributions that enjoy a Gaussian-type probability concentration inequalities; see [4,8,28]. For a fixed $0 \le x \le 1$, we have

$$p_{n,k}(x) \ge 0, \quad \text{and} \quad \sum_{k=0}^{n} p_{n,k}(x) = 1.$$

Thus, the random variable $B_n^X f$ can be studied effectively as a convex linear combination of the $(n+1)$ random variables $f(X_{n,k})$.

For $1 \le p \le \infty$ and $f \in C[0,1]$, we define

$$\|f\|_p := \begin{cases} \left(\int_0^1 |f(x)|^p dx \right)^{1/p}, & \text{if } 1 \le p < \infty, \\ \max_{x \in [0,1]} |f(x)|, & \text{if } p = \infty. \end{cases}$$

Definition 2.1. Let $1 \le p \le \infty$, and $f \in C[0,1]$ be given. If $\forall \varepsilon > 0$,

$$\lim_{n \to \infty} \mathbb{P}\{\|B_n^X f - f\|_p > \varepsilon\} = 0,$$

then we say that $B_n^X f$ converges to f in probability under the L^p-norm. We also refer to such probabilistic convergence as "L^p-probabilistic convergence" of $B_n^X f$ to f.

The bulk of our effort in the current article is devoted to finding exponential decay rates [2] (in terms of the modulus of continuity of an $f \in C[0,1]$) for L^p-probabilistic convergence. We use throughout the article two basic properties of the modulus of continuity. The first one is called "positivity", which asserts that if $f \in C[0,1]$ is not a constant, then $\omega(h) > 0$ for $0 < h \le 1$. The second one is referred to as "subadditivity", as shown in the following inequality:

$$\omega(\lambda\delta) \le (1+\lambda)\, \omega(\delta), \quad 0 \le \delta \le 1, \quad 0 < \lambda < \infty. \tag{2.1}$$

Neither of the two properties is hard to prove. Authors of [34] showed the following result.

Theorem 2.2. *If $X \sim \mathrm{Beta}(\alpha, \beta)$, then*

$$\mathbb{P}\{|X - \mathbb{E}(X)| > r\} \le 2\exp[-2(\alpha+\beta+1)r^2], \quad r > 0. \tag{2.2}$$

Let $\mathcal{B}_n$ denote the function defined by

$$\mathcal{B}_n(x,y) = (n+1) \sum_{k=0}^{n} p_{n,k}(x) p_{n,k}(y), \quad (x,y) \in [0,1]^2.$$

Let (X, Y) be the random vector with the joint density function $\mathcal{B}_n$. The resulted distribution on $[0,1]^2$ is often called "mean square beta distribution";

[2] The type of exponential decay rates we derive here is often referred to as Gaussian tail bounds.

see [3]. With the newly-updated Gaussian tail bound as shown in Theorem 2.2, we reiterate a result of Bobkov and Ledoux [3, Proposition B.12].

Proposition 2.3. *If (X, Y) is the random vector with the joint density function $\mathcal{B}_n$, then for all $r \geq 0$,*

$$\mathbb{P}\{|X - Y| > r\} \leq 2\exp[-(n+3)r^2].$$

We will utilize the random vector (X, Y) as a "majorant" for our stochastic Bernstein polynomials. The best effect of this approach can be seen via its implementation on the uniform distribution in $(0, 1)$. We denote U the c.d.f. of the uniform distribution in $(0, 1)$.

Lemma 2.4. *Let (X, Y) be the random vector that has joint density function $\mathcal{B}_n(x, y)$. Then for any $1 \leq p < \infty$, we have*

$$\mathbb{E}\left(\|B_n^X U - U\|_p^p\right) \leq \mathbb{E}\left(|X - Y|^p\right).$$

Proof. We first use the triangle inequality to write

$$|(B_n^X U)(x) - U(x)| \leq \sum_{j=0}^{n} |X_{n,k} - x| p_{n,k}(x), \quad x \in [0, 1].$$

By Jensen's inequality, we have

$$\begin{aligned}
\mathbb{E}\left(\|B_n^X U - U\|_p^p\right) &\leq \mathbb{E}\left[\int_0^1 \left(\sum_{k=0}^{n} |X_{n,k} - x| p_{n,k}(x)\right)^p dx\right] \\
&\leq \mathbb{E}\left[\int_0^1 \sum_{k=0}^{n} |X_{n,k} - x|^p p_{n,k}(x) dx\right] \\
&= \int_0^1 \int_0^1 |x - y|^p \mathcal{B}_n(x, y) dy dx \\
&= \mathbb{E}\left(|X - Y|^p\right).
\end{aligned}$$

This completes the proof. □

The result of the above lemma amounts to saying that for any $1 \leq p < \infty$, the p-moment of the random variable $\|B_n^X U - U\|_p$ is majorized by that of the random variable $(X - Y)$. The following result shows that the majorization extends to exponential moments.

Lemma 2.5. *For any $1 \leq p < \infty$ and $r \geq 0$, the following inequality holds true:*

$$\mathbb{E}\left[\exp\left(r\|B_n^X U - U\|_p^p\right)\right] \leq \mathbb{E}\left[\exp\left(r|X - Y|^p\right)\right].$$

Proof. Similar to the proof of Lemma 2.4, we first write the following inequality:

$$\|B_n^X U - U\|_p^p \leq \sum_{k=0}^{n} \int_0^1 |X_{n,k} - x|^p p_{n,k}(x) dx.$$

Applying Jensen's inequality twice (first the integral form then the discrete form), we have

$$\begin{aligned}\exp\left(r\|B_n^X U - U\|_p^p\right) &\le \int_0^1 \exp\left(r\sum_{k=0}^n p_{n,k}(x)|X_{n,k}-x|^p\right)dx\\ &\le \int_0^1 \left(\sum_{k=0}^n \exp\left(r|X_{n,k}-x|^p\right)p_{n,k}(x)\right)dx.\end{aligned}$$

It follows that

$$\mathbb{E}\left[\exp\left(r\|B_n^X U - U\|_p^p\right)\right] \le \int_0^1\int_0^1 \exp\left(r|x-y|^p\right)\mathcal{B}_n(x,y)dydx,$$

which is the desired result. □

The majorization of $\|B_n^X U - U\|_p$ by $|X-Y|$ as shown in Lemmas 2.4 and 2.5 can be passed on to the random variable $\|B_n^X f - f\|_p$ for any $f \in C[0,1]$ via manipulation of its modulus of continuity. Our main goal in the current section is to find upper bounds for the quantities

$$\mathbb{P}\{||B_n^X f - f||_p > \varepsilon\}, \quad 1 \le p \le \infty.$$

where $\varepsilon > 0$ and $n \in \mathbb{N}$ are fixed. If f is a constant, then $B_n^X f \equiv f$. In such a case, the upper bounds hold true automatically. We hereby declare once and for all that in the sequel our target functions are not constant, which justifies writing $\omega(\frac{1}{\sqrt{n}})$ in denominators. For aesthetic reasons, we will use the equivalent form $\omega(n^{-1/2})$ for $\omega(\frac{1}{\sqrt{n}})$ in mathematical contexts in which the presence of the latter expression would make unusually-large displayed equations.

Proposition 2.6. *For any $1 \le p < \infty$, and any given $f \in C[0,1]$, the following inequalities hold true:*

$$\mathbb{E}\left(\|B_n^X f - f\|_p^p\right) \le 2^{p-1}\ \omega^p\left(n^{-1/2}\right)\left[1 + n^{p/2}\mathbb{E}\left(|X-Y|^p\right)\right]; \tag{2.3}$$

$$\mathbb{E}\left[\exp\left(\{2\ \omega(n^{-1/2})\}^{-p}\|B_n^X f - f\|_p^p\right)\right] \le \sqrt{e}\ \cdot \mathbb{E}\left[\exp\left(\frac{1}{2}n^{p/2}|X-Y|^p\right)\right]. \tag{2.4}$$

Proof. The proofs of (2.3) and (2.4) use basically the same techniques as those in the proofs of Lemmas 2.4 and 2.5. We will prove (2.4) as it is more involved than the other. For a fixed $x \in [0,1]$, by the subadditivity of modulus of continuity of f and Jensen's inequality, we write

$$\begin{aligned}&|B_n^X f(x) - f(x)|^p\\ &\le \omega^p\left(n^{-1/2}\right)\sum_{k=0}^n \left(1+\sqrt{n}|X_{n,k}-x|\right)^p p_{n,k}(x)\\ &\le 2^{p-1}\ \omega^p\left(n^{-1/2}\right)\left(1 + n^{p/2}\sum_{k=0}^n |X_{n,k}-x|^p p_{n,k}(x)\right).\end{aligned}$$

It follows that

$$\begin{aligned}
&\exp\left(\{2\,\omega(n^{-1/2})\}^{-p}\|B_n^X f - f\|_p^p\right)\\
&\quad\le \int_0^1 \left[\exp\left(\frac{1}{2}+\frac{1}{2}n^{p/2}\sum_{k=0}^{n}|X_{n,k}-x|^p p_{n,k}(x)\right)\right]dx\\
&\quad\le \sqrt{e}\,\int_0^1 \left[\sum_{k=0}^{n}\exp\left(\frac{1}{2}n^{p/2}|X_{n,k}-x|^p\right)p_{n,k}(x)\right]dx,
\end{aligned}$$

which implies that

$$\begin{aligned}
&\mathbb{E}\left[\exp\left(\{2\,\omega(n^{-1/2})\}^{-p}\|B_n^X f - f\|_p^p\right)\right]\\
&\quad\le \sqrt{e}\int_0^1\int_0^1 \exp\left(\frac{1}{2}n^{p/2}|x-y|^p\right)\mathcal{B}_n(x,y)dydx\\
&\quad= \sqrt{e}\;\cdot\mathbb{E}\left[\exp\left(\frac{1}{2}n^{p/2}|X-Y|^p\right)\right].
\end{aligned}$$

This completes the proof. □

In the rest of this section, we will first develop upper bounds for $\mathbb{E}\left(|X-Y|^p\right)$ $(1\le p<\infty)$ (Lemma 2.7 below), and $\mathbb{E}\left[\exp\left(\frac{1}{2}n^{p/2}|X-Y|^p\right)\right]$ $(1\le p\le 2)$ (Lemma 2.9 below). We will then use Chebyshev inequality to obtain unconditional L^p-concentration inequalities for stochastic Bernstein polynomials (Theorems 2.8, 2.11 below). En route we will be using variants of the following identity

$$\mathbb{E}(V)=\int_0^\infty \mathbb{P}(V>r)\mathrm{d}r, \tag{2.5}$$

where V is a nonnegative random variable.

Lemma 2.7. *Let $1\le p<\infty$. Then*

$$\mathbb{E}\left(|X-Y|^p\right)\le \frac{p\,\Gamma\left(\frac{p}{2}\right)}{(n+3)^{p/2}}.$$

Proof. We make use of identity (2.5) and Theorem 2.2 to write

$$\begin{aligned}
&\mathbb{E}\left(|X-Y|^p\right)\\
&\quad= p\int_0^\infty \mathbb{P}\{|X-Y|>r\}r^{p-1}\mathrm{d}r\\
&\quad\le 2p\int_0^\infty e^{-(n+3)r^2}r^{p-1}\mathrm{d}r\\
&\quad= \frac{p\,\Gamma\left(\frac{p}{2}\right)}{(n+3)^{p/2}},
\end{aligned}$$

which is the desired result. □

Theorem 2.8. *For any $\varepsilon>0$, $1\le p<\infty$, and $f\in C[0,1]$, the following inequalities hold true:*

1. $\mathbb{P}\{\|B_n^X f-f\|_p>\varepsilon\}\le \dfrac{2^{p-1}\left[1+p\,\Gamma\left(\frac{p}{2}\right)\right]\omega^p\left(n^{-1/2}\right)}{\varepsilon^p}.$

1. $\mathbb{E}\left(\|B_n^X f - f\|_p\right) \leq 2^{1-\frac{1}{p}}\left[1 + p\,\Gamma\left(\frac{p}{2}\right)\right]^{\frac{1}{p}} \omega\left(n^{-1/2}\right).$

Part 1 of Theorem 2.8 gives an L^p-probabilistic convergence rate ($1 \leq p < \infty$) for stochastic Bernstein polynomials; Part 2 shows that the average L^p-approximation order ($1 \leq p < \infty$) of stochastic Bernstein polynomials is the same as that given by the classical Bernstein polynomials. The result in Part 2 fails to work for the case $p = \infty$. The causation is that the constant, by Sterling's formula, is of the order $\sqrt{p}$ as $p \to \infty$. In Sect. 4, we will obtain an average L^p-approximation order with an absolute constant for stochastic Bernstein polynomials (Corollary 4.3).

Proof. To prove Part 1, we use Chebyshev inequality, inequality (2.3), and Lemma 2.7 to derive

$$\mathbb{P}\{\|B_n^X f - f\|_p > \varepsilon\} \leq \frac{\mathbb{E}\left(\|B_n^X f - f\|_p^p\right)}{\varepsilon^p} \leq \frac{2^{p-1}\left[1 + p\Gamma\left(\frac{p}{2}\right)\right]\omega^p\left(n^{-1/2}\right)}{\varepsilon^p}.$$

To prove Part 2, we use Jensen's inequality, inequality (2.3), and Lemma 2.7 to derive

$$\mathbb{E}\left(\|B_n^X f - f\|_p\right) \leq \left[\mathbb{E}\left(\|B_n^X f - f\|_p^p\right)\right]^{1/p} \leq 2^{1-\frac{1}{p}}\left[1 + p\Gamma\left(\frac{p}{2}\right)\right]^{1/p}\omega\left(n^{-1/2}\right).$$

This completes the proof. □

Lemma 2.9. *Suppose that Z is a random variable satisfying the following inequality:*

$$\mathbb{P}\{|Z| > r\} \leq 2\exp\left(-2nr^2\right), \quad r > 0.$$

Then we have

$$\mathbb{E}\left[\exp\left(n|Z|^2\right)\right] \leq 2.$$

Proof. We use Proposition (2.5) to derive

$$\begin{aligned}
&\mathbb{E}\left[\exp\left(n|Z|^2\right)\right] \\
&= \int_0^\infty \mathbb{P}\{\exp\left(n|Z|^2\right) > r\}\mathrm{d}r \\
&= \int_0^\infty \mathbb{P}\{|Z| > \tau\}\mathrm{d}\exp\left(n\tau^2\right) \\
&\leq 2\int_0^\infty \exp(-2n\tau^2)\mathrm{d}\exp\left(n\tau^2\right) \\
&\leq 2\int_1^\infty \frac{du}{u^2} = 2.
\end{aligned}$$

This completes the proof. □

Proposition 2.3 and Lemma 2.9 yield the following result.

Corollary 2.10. *The following inequality holds true:*

$$\mathbb{E}\left[\exp\left(\frac{1}{2}n|X - Y|^2\right)\right] \leq 2.$$

We now state and prove the main result of the section.

Theorem 2.11. *For any given $\varepsilon > 0$ and p in the range $1 \leq p \leq 2$, and $f \in C[0,1]$, the following inequality holds true.*

$$\mathbb{P}\left\{\left\|B_n^X f - f\right\|_p > \varepsilon\right\} \leq 2\sqrt{e}\ \exp\left[-\frac{\varepsilon^2}{4\ \omega^2\left(n^{-1/2}\right)}\right].$$

Proof. For any given $\varepsilon > 0$ and p in the range $1 \leq p \leq 2$, and $f \in C[0,1]$, we have

$$\mathbb{P}\left\{\left\|B_n^X f - f\right\|_p > \varepsilon\right\} \leq \mathbb{P}\left\{\left\|B_n^X f - f\right\|_2 > \varepsilon\right\}. \tag{2.6}$$

Thus, it suffices to prove the theorem for the case $p = 2$. By (2.4) and Corollary 2.10, we have

$$\mathbb{E}\left[\exp\left\{\left[2\ \omega\left(n^{-1/2}\right)\right]^{-2}\|B_n^X f - f\|_2^2\right\}\right] \leq \sqrt{e}\ \mathbb{E}\left[\exp\left(\frac{1}{2}n|X-Y|^2\right)\right] \leq 2\sqrt{e}.$$

Applying Chebyshev inequality to the right hand side of (2.6), we obtain

$$\begin{aligned}
&\mathbb{P}\left\{\|B_n^X f - f\|_2 > \varepsilon\right\} \\
&\quad\leq \frac{\mathbb{E}\left[\exp\left\{\left[2\ \omega\left(n^{-1/2}\right)\right]^{-2}\|B_n^X f - f\|_2^2\right\}\right]}{\exp\left\{\varepsilon^2\ \left[2\ \omega\left(n^{-1/2}\right)\right]^{-2}\right\}} \\
&\quad\leq 2\sqrt{e}\ \exp\left[-\frac{\varepsilon^2}{4\ \omega^2\left(n^{-1/2}\right)}\right],
\end{aligned}$$

which is the desired result. □

Unlike (1.4), we do not need assumption (1.5) in the proof of Theorem 2.11. The price we paid for this is the multiplicative constant $\sqrt{e}$.

3. Gaussian-Type L^p-Concentration Inequalities $(1 \leq p < \infty)$

In an essential way, the proof of Theorem 2.11 depends on the boundedness of the sequence

$$\mathbb{E}\left(\exp(n^{p/2}|X-Y|^p)\right), \quad n \in \mathbb{N}, \quad 1 \leq p \leq 2.$$

This is no longer true for $p > 2$. To prove unconditional L^p-concentration inequalities for $2 < p < \infty$, we start with the random process $V_n(x)$ defined by

$$V_n(x) := \sum_{k=0}^{n} |X_{n,k} - x| p_{n,k}(x), \quad x \in [0,1].$$

Lemma 3.1. *The following inequality holds true almost surely:*

$$\|B_n^X f - f\|_p^2 \leq 2\ \omega^2\left(n^{-1/2}\right)\left(1 + n\|V_n\|_p^2\right), \quad 1 \leq p < \infty, \quad f \in C[0,1].$$

Proof. By subadditivity of the modulus of continuity, we have

$$\begin{aligned}
&|(B_n^X f)(x) - f(x)| \\
&\leq \sum_{k=0}^{n} \omega(|X_{n,k} - x|)\ p_{n,k}(x) \\
&\leq \omega\left(n^{-1/2}\right)\left(1 + \sqrt{n}\ V_n(x)\right).
\end{aligned}$$

We then apply Minkowski's and Jensen's inequalities to get the desired the result. □

We use a triangle inequality to write

$$V_n(x) \leq \sum_{k=0}^{n} \left|X_{n,k} - \frac{k}{n}\right| p_{n,k}(x) + \sum_{k=0}^{n} \left|x - \frac{k}{n}\right| p_{n,k}(x). \tag{3.1}$$

Let $Y_n(x)$ denote the random process as is expressed by the first sum on the right hand side of (3.1), and W_n the random variable $\max_{0\leq k\leq n} |X_{n,k} - \frac{k}{n}|$. For a fixed $x \in [0,1]$, let $T(x)$ be the random variable taking values 0 and 1, with respective probabilities x and $1-x$. Let $\{T_k(x)\}_{k=0}^{n}$ be $(n+1)$ independent copies of $T(x)$. Let $Z_n = Z_n(x)$ be the random variable defined by

$$Z_n(x) = \frac{1}{n}\sum_{k=0}^{n} T_k(x). \tag{3.2}$$

Then $Z_n(x)$ takes values $\frac{k}{n}$ with respective probabilities $p_{n,k}(x)$, $k = 0, 1, \ldots, n$. These lead to the following result.

Lemma 3.2. *The following inequality holds true almost surely:*

$$\|V_n\|_p^2 \leq 2\ \left(W_n^2 + \|\mathbb{E}\left(|Z_n - \mathbb{E}(Z_n)|\right)\|_p^2\right), \quad 1 \leq p < \infty.$$

Proof. For all $x \in [0,1]$, we have

$$0 \leq Y_n(x) \leq W_n \sum_{k=0}^{n} p_{n,k}(x) = W_n,$$

which implies that $\|Y_n\|_\infty \leq W_n$ almost surely. By Minkowski's and Jensen's inequalities, we have

$$\begin{aligned}
\|V_n\|_p^2 \leq& 2\left\|\sum_{k=0}^{n}\left|X_{n,k} - \frac{k}{n}\right| p_{n,k}\right\|_p^2 + 2\|\mathbb{E}\left(|Z_n - \mathbb{E}(Z_n)|\right)\|_p^2. \\
\leq& 2\left(W_n^2 + \|\mathbb{E}\left(|Z_n - \mathbb{E}(Z_n)|\right)\|_p^2\right),
\end{aligned}$$

which is the desired result.

We will deal with the two terms on the right hand side of the above inequality separately, and start with the first. The key strategy is to tie the random variable W_n to the empirical process of the uniform distribution in $(0,1)$.

Lemma 3.3. *Let U_{n+1} be the empirical distribution function of order $(n+1)$ associated with the uniform distribution in $(0,1)$ defined by*

$$U_{n+1}(x) := \frac{1}{n+1}\sum_{k=0}^{n} \mathbb{1}_{\{X_{n,k}<x\}}. \tag{3.3}$$

Then the following inequality holds true almost surely,

$$W_n \leq \sup_{x\in[0,1]} |U_{n+1}(x) - x|\,. \tag{3.4}$$

Proof. Let $X^*_{n,0} < X^*_{n,1} < \cdots < X^*_{n,n}$ be a set of observed values of the random variables $X_{n,0}, X_{n,1}, \ldots, X_{n,n}$. Let U^*_{n+1} be the corresponding (observed) empirical distribution function for the uniform distribution U on $[0,1]$. Let $I_0 := [0, X^*_{n,0})$, $I_k := [X^*_{n,k-1}, X^*_{n,k})$, $1 \leq k < n$, and $I_n := [X^*_{n,n}, 1]$. Then we have

$$\begin{aligned}
&\sup_{x\in I_0}\left(x - U^*_{n+1}(x)\right) = X^*_{n,0},\\
&\sup_{x\in I_k}\left(x - U^*_{n+1}(x)\right) = X^*_{n,k} - \frac{k}{n+1}, \quad 1 \leq k \leq n,\\
&\sup_{x\in I_n}\left(x - U^*_{n+1}(x)\right) = 0.
\end{aligned}$$

The equations above show that

$$\sup_{x\in[0,1]}\left(x - U^*_{n+1}(x)\right) = \max_{0\leq k\leq n}\left(X^*_{n,k} - \frac{k}{n+1}\right).$$

Since $X^*_{n,k} - \frac{k}{n} < X^*_{n,k} - \frac{k}{n+1}$, we have $= \max_{0\leq k\leq n}\left(X^*_{n,k} - \frac{k}{n}\right) \leq \sup_{x\in[0,1]}$ $\left(x - U^*_{n+1}(x)\right)$. This inequality holds true for every set of such observed values of the order statistics. Hence we conclude that

$$\max_{0\leq k\leq n}\left(X_{n,k} - \frac{k}{n}\right) \leq \sup_{x\in[0,1]}\left(x - U_{n+1}(x)\right) \quad \text{almost surely.} \tag{3.5}$$

Similarly we show that

$$\max_{0\leq k\leq n}\left(\frac{k}{n} - X_{n,k}\right) \leq \sup_{x\in[0,1]}\left(U_{n+1}(x) - x\right) \quad \text{almost surely.} \tag{3.6}$$

Combining (3.5) and (3.6), we get the desired result. □

Lemma 3.4. *Let $\varepsilon > 0$ and $n \in \mathbb{N}$ be given. Then we have the following inequality:*

$$\mathbb{P}\{W_n > \varepsilon\} \leq 2\, e^{-2(n+1)\varepsilon^2}.$$

Proof. By Lemma 3.3 and Dvoretzky–Kiefer–Wolfowitz inequality (see (1.6)), we have

$$\mathbb{P}\{W_n > \varepsilon\} \leq \mathbb{P}\{\sup_{x\in[0,1]} |U_{n+1}(x) - x| > \varepsilon\} \leq 2\, e^{-2(n+1)\varepsilon^2},$$

which is the desired result. □

We now turn our attention to estimating the term $\|\mathbb{E}\left(|Z_n - \mathbb{E}(Z_n)|\right)\|_p^2$ in Lemma 3.2, en route we need a well-known concentration inequality known as Bernstein inequality; see [12].

Theorem 3.5. *Let $c > 0$. Let $X_1, \ldots, X_n$ be independent variables. If $\mathbb{P}(|X_j| \leq c) = 1$, then for any $\varepsilon > 0$,*

$$\mathbb{P}(|S_n - \mu| > \varepsilon) \leq 2\exp\left\{\frac{-n\varepsilon^2}{2\sigma^2 + 2c\varepsilon/3}\right\},$$

in which $S_n := \frac{1}{n}\sum_{j=1}^n X_j$, $\mathbb{E}(S_n) = \mu$, and $\sigma^2 := \frac{1}{n}\sum_{j=1}^n \mathrm{Var}(X_j)$.

In the next lemma, we apply Bernstein inequality to control a certain segment of sums of $p_{n,k}$.

Lemma 3.6. *For a fixed $x \in [0,1]$, the following inequality holds true:*

$$\sum_{|x-\frac{k}{n}|\geq\varepsilon} p_{n,k}(x) \leq 2\exp\left[\frac{-n\varepsilon^2}{2x(1-x) + 3\varepsilon/2}\right].$$

Proof. Consider the random variable Z_n as defined in (3.2). We have $\mathbb{E}(Z_n) = x$, and $\mathrm{Var}(Z_n) = x(1-x)$. We apply Bernstein inequality to the random variable Z_n to get the desired inequality. □

Denote

$$\Delta_{n,x} := \max\left(\sqrt{\frac{x(1-x)}{n}}, \quad \frac{1}{n}\right).$$

Theorem 3.7. *Let $0 < p < \infty$ be given. The central p-moment of the random variable Z_n satisfies the following inequality:*

$$\mathbb{E}\left(|Z_n - \mathbb{E}(Z_n)|^p\right) \leq C_p\ \Delta_{n,x}^p,$$

where

$$C_p := 2\sum_{j=0}^{\infty}(j+1)^p \exp[-j^2/(2+3j/2)]. \tag{3.7}$$

Proof. If $x = 0$ or $x = 1$, then the desired inequality holds true trivially. In the rest of the proof, we assume that $x(1-x) > 0$. Denote

$$E_{j,x} := \left\{k : j\Delta_{n,x} \leq \left|\frac{k}{n} - x\right| \leq (j+1)\Delta_{n,x}\right\}, \quad j = 0, 1, \ldots, J_{n,x},$$

where $J_{n,x} := \lceil \Delta_{n,x}^{-1} \rceil$, and use Lemma 3.6 to write

$$\begin{aligned}
\mathbb{E}|Z_n - \mathbb{E}(Z_n)|^p &= \sum_{k=0}^{n}\left|\frac{k}{n} - x\right|^p p_{n,k}(x) \\
&\leq \sum_{j=0}^{J_{n,x}}\sum_{k\in E_{j,x}}\left|\frac{k}{n} - x\right|^p p_{n,k}(x) \\
&\leq 2\Delta_{n,x}^p \sum_{j=0}^{J_{n,x}}(j+1)^p \exp\left[\frac{-n\left(j\Delta_{n,x}\right)^2}{2x(1-x) + 3j\Delta_{n,x}/2}\right].
\end{aligned}$$

In what follows, we derive an upper bound (independent of x and n) for the sum above by considering two separate cases. Case (i): $x(1-x) \leq \frac{1}{n}$. We have $\Delta_{n,x} \leq 1/n$. It follows that

$$\exp\left[\frac{-n\left(j\Delta_{n,x}\right)^2}{2x(1-x)+3j\Delta_{n,x}/2}\right] \leq \exp[-j^2/(2+3j/2)].$$

Case (ii): $x(1-x) > \frac{1}{n}$. We have

$$\Delta_{n,x} = \sqrt{\frac{x(1-x)}{n}} \leq x(1-x),$$

which yields

$$\begin{aligned}\exp\left[\frac{-n\left(j\Delta_{n,x}\right)^2}{2x(1-x)+3j\Delta_{n,x}/2}\right] &\leq \exp\left[\frac{-n\left(jx(1-x)/n\right)^2}{2x(1-x)+3jx(1-x)/2}\right] \\ &\leq \exp\left[-j^2/(2+3j/2)\right].\end{aligned}$$

The proof is complete. □

Lemma 3.8. *The following inequality holds true:*

$$\|\mathbb{E}\left(|Z_n - \mathbb{E}(Z_n)|\right)\|_p^2 \leq C_p^{2/p}/(4n),$$

where C_p is as defined in (3.7).

Proof. By Jensen's inequality and Theorem 3.7, we have

$$\begin{aligned}&\|\mathbb{E}\left(|Z_n - \mathbb{E}(Z_n)|\right)\|_p^2 \\ &\leq \left[\int_0^1 \left(\sum_{k=0}^n \left|x - \frac{k}{n}\right| p_{n,k}(x)\right)^p dx\right]^{2/p} \\ &\leq \left[\int_0^1 \left(\sum_{k=0}^n \left|x - \frac{k}{n}\right|^p p_{n,k}(x)\right) dx\right]^{2/p} \\ &\leq \left(C_p \int_0^1 \Delta_{n,x}^{p/2}\, dx\right)^{2/p} \leq C_p^{2/p}/(4n),\end{aligned}$$

which is the desired result. □

Here is the main result of the section.

Theorem 3.9. *Let $\varepsilon > 0$ and $n \in \mathbb{N}$ be given. Then the following inequality holds true:*

$$\mathbb{P}\left\{\left\|B_n^X f - f\right\|_p > \varepsilon\right\} \leq 2D_p\ \exp\left[-\frac{\varepsilon^2}{4\ \omega^2\left(n^{-1/2}\right)}\right], f \in C[0,1], 1 \leq p < \infty,$$

in which $D_p := \exp\left[\left(1 + \frac{C_p^{2/p}}{2}\right)/2\right]$.

Proof. By Lemmas 3.1, 3.2, and 3.8, the following inequalities hold true almost surely:

$$\begin{aligned}\left\|B_n^X f - f\right\|_p^2 \leq & 2\ \omega^2\left(n^{-1/2}\right)\left[1 + 2n\left(W_2^2 + \|\mathbb{E}\left(|Z_n - \mathbb{E}(Z_n)|\right)\|_p^2\right)\right] \\ \leq & 2\ \omega^2\left(n^{-1/2}\right)\left(1 + 2n\ W_2^2 + \frac{C_p^{2/p}}{2}\right).\end{aligned}$$

It follows from Lemmas 3.4 and 2.9 that

$$\begin{aligned}&\mathbb{E}\left(\exp\left[\left(4\ \omega^2\left(n^{-1/2}\right)\right)^{-1}\|B_n^X f - f\|_p^2\right]\right) \\ &\leq \mathbb{E}\left(\exp\left[\frac{1}{2}\left(1 + 2n\ W_2^2 + \frac{C_p^{2/p}}{2}\right)\right]\right) \\ &= \exp\left[\left(1 + (C_p^{2/p})/2\right)/2\right]\mathbb{E}\left[\exp(2n\ W_2^2)\right] = 2D_p.\end{aligned}$$

Applying Chebyshev inequality, we get

$$\begin{aligned}&\mathbb{P}\left\{\left\|B_n^X f - f\right\|_p > \varepsilon\right\} \\ &\leq \mathbb{E}\left(\exp\left[\left(4\ \omega^2\left(n^{-1/2}\right)\right)^{-1}\|B_n^X f - f\|_p^2\right]\right) \\ &\quad\cdot\left(\exp\left[\left(4\ \omega^2\left(n^{-1/2}\right)\right)^{-1}\varepsilon^2\right]\right)^{-1} \\ &\leq 2D_p\ \exp\left[-\frac{\varepsilon^2}{4\ \omega^2\left(n^{-1/2}\right)}\right],\end{aligned}$$

which is the desired result. □

Two remarks about Theorem 3.9 are in order.

1. For $p = 2$, the multiplicative constant in the result of Theorem 3.9 is $\exp\left[\left(1 + \frac{C_2}{2}\right)/2\right]$, which is noticeably larger than $\sqrt{e}$ as given in Theorem 2.11.
2. The multiplicative constant D_p is of the order $\exp(\sqrt{p})$ as $p \to \infty$, which is the typical growth rate of the pth order of exponential moments of a sub-Gaussian random variable. This has ruled out the feasibility of obtaining an L^∞-concentration inequality for stochastic Bernstein polynomials by taking the limit $p \to \infty$.

4. Exponential L_∞-Probabilistic Convergence Rates

The approach undertaken in this section needs to harness the deterministic approximation power of classical Bernstein polynomials, which features many publications; see [11,15,16,22,25,35], and the references therein. Sikkema [31] proved the following remarkable result:

$$\sup_n \sup_f \sup_{0\leq x\leq 1} \frac{|B_n f(x) - f(x)|}{\omega(f, n^{-1/2})} = \frac{4306 + 837\sqrt{6}}{5832} = 1.0898871330\cdots.$$

We denote this constant by c_s. Sikkema [32] also showed that the value on the right hand side of the above equation is only reached for $n = 6$. But we will not use the latter result in the current article.

Lemma 4.1. *Let $\varepsilon > 0$ and $f \in C[0,1]$ be given. Suppose that $\omega(\frac{1}{\sqrt{n}}) < \frac{\varepsilon}{2(c_s+1)}$. Then the following inequality holds true:*

$$\mathbb{P}\{\|B_n^X f - f\|_p > \varepsilon\} \leq \mathbb{P}\left\{\|Y_n\|_p > \frac{\varepsilon}{2\sqrt{n}\ \omega\left(n^{-1/2}\right)}\right\}, \quad 1 \leq p \leq \infty. \quad (4.1)$$

Authors of [34, Lemma 2.10] proved a slightly different version of the above lemma. We will omit the proof, and refer interested readers to the proof of Lemma 2.10 in [34]

Theorem 4.2. *Let $\varepsilon > 0$ and $f \in C[0,1]$ be given. Suppose that*

$$\omega\left(\frac{1}{\sqrt{n}}\right) < \frac{\varepsilon}{2(c_s+1)}.$$

Then the following inequality holds true:

$$\mathbb{P}\{\|B_n^X f - f\|_p > \varepsilon\} \leq 2\exp\left[-\frac{\varepsilon^2}{2\ \omega^2(n^{-1/2})}\right], \quad 1 \leq p \leq \infty.$$

Proof. Let $1 \leq p \leq \infty$ be given. We use Lemmas 4.1 and 3.4 in a successive fashion to derive the following inequalities:

$$\begin{aligned}
&\mathbb{P}\{\|B_n^X f - f\|_p > \varepsilon\} \\
&\quad \leq \mathbb{P}\{\|B_n^X f - f\|_\infty > \varepsilon\} \\
&\quad \leq \mathbb{P}\left\{\|Y_n\|_\infty > \frac{\varepsilon}{2\sqrt{n}\ \omega\left(n^{-1/2}\right)}\right\} \\
&\quad \leq \mathbb{P}\left\{W_n > \frac{\varepsilon}{2\sqrt{n}\ \omega\left(n^{-1/2}\right)}\right\} \\
&\quad \leq 2\exp\left[-\frac{\varepsilon^2}{2\omega^2(n^{-1/2})}\right].
\end{aligned}$$

This completes the proof. □

The result of Theorem 4.2 confirms a similar conjecture raised in [33].

Corollary 4.3. *Let $f \in C[0,1]$ be given. Then the following inequality holds true:*

$$\mathbb{E}\left(\|B_n^X f - f\|_p\right) \leq \left[2(c_s+1) + \sqrt{\frac{\pi}{2}}\ \right]\omega(\frac{1}{\sqrt{n}}), \quad 1 \leq p \leq \infty.$$

Proof. Denote $a_n := 2(c_s+1)\omega(n^{-1/2})$. By (2.5) and Theorem 4.2, we have

$$\begin{aligned}
\mathbb{E}\left(\|B_n^X f - f\|_p\right) &= \left\{\int_0^{a_n} + \int_{a_n}^{\infty}\right\}\mathbb{P}\{\|B_n^X f - f\|_p > r\}\mathrm{d}r \\
&\leq 2(c_s+1)\ \omega\left(\frac{1}{\sqrt{n}}\right)
\end{aligned}$$

$$+ 2 \int_0^\infty \exp\left[-\frac{r^2}{2\,\omega^2\left(n^{-1/2}\right)} \right] \mathrm{d}r$$
$$= \left[2(c_s + 1) + \sqrt{\frac{\pi}{2}} \right] \omega\left(\frac{1}{\sqrt{n}}\right).$$

This completes the proof. □

Acknowledgements

We thank Professor Zongmin Wu, Fudan University, for fruitful discussions on topics related to the article.

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

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

References

[1] Adell, J.A., Cardenas-Morales, D.: Stochastic Bernstein polynomials: uniform convergence in probability with rates. Adv. Comput. Math. **46**, 16 (2020)

[2] Birnbaum, Z.W., McCarty, R.C.: A distribution-free upper confidence bound for $\Pr\{Y < X\}$, based on independent samples of X and Y. Ann. Math. Stat. **29**, 558–562 (1958)

[3] Bobkov, S., Ledoux, M.: One-dimensional empirical measures, order statistics and Kantorovich transport distances. Mem. Am. Math. Soc. **261**, 1259 (2019)

[4] Boucheron, S., Lugosi, G., Massart, P.: Concentration Inequalities: A Nonasymptotic Theory of Independence. Oxford University Press, Oxford (2013)

[5] Bourgain, J., Lindenstrauss, J.: Distribution of points on spheres and approximation by zonotopes. Israel J. Math. **64**, 25–31 (1988)

[6] Bourgain, J., Lindenstrauss, J., Milman, V.: Approximation of zonoids by zonotopes. Acta Math. **162**, 73–141 (1989)

[7] Bourgain, J., Lindenstrauss, J.: Approximating the ball by a Minkowski sum of segments with equal length. Discret. Comput. Geom. **9**, 131–144 (1993)

[8] Buldygin, V.V., Kozachenko, I.V.: Metric Characterization of Random Variables and Random Processes, Translations of Mathematical Monographs, vol. 188. American Mathematical Society, Providence (2000)

[9] Casella, G., Berger, R.: Statistical Reference, 2nd ed, Duxbury Advanced Series (2002)

[10] Cenusa, Gh., Sacuiu, I.: On some stochastic approximations for random functions. Rend. Mat. (Roma) Ser. VI **12**, 143–156 (1979)

[11] Cheney, E.W.: Introduction to Approximation Theory, 2nd edn. Chelsea, New York (1982)

[12] Cucker, F., Smale, S.: On the mathematical foundations of learning. Bull. (New Ser.) Am. Math. Soc. **39**, 1–49 (2001)

[13] Cucker, F., Zhou, D.X.: Learning Theory: An Approximation Theory Viewpoint. Cambridge University Press, Cambridge (2007)

[14] David, H.A., Nagaraja, H.N.: Order Statistics, Wiley Series in Probability and Statistics, 3rd edn. Wiley, Hoboken (2003)

[15] Ditzian, Z., Ivanov, K.G.: Strong converse inequalities. J. Anal. Math. **61**, 61–111 (1993)

[16] Ditzian, Z., Totik, V.: Moduli of Smoothness. Springer, Berlin (2012)

[17] Dvoretzky, A., Kiefer, J., Wolfowitz, J.: Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. Ann. Math. Stat. **27**(3), 642–669 (1956)

[18] Gal, S.G.: Jackson type estimates in the approximation of random functions by random polynomials. Rend. Mat. Appl. Roma **4**, 543–556 (1994)

[19] Gal, S.G.: Approximation Theory in Random Setting, Chapter 12 in Handbook of Analytic Computational Methods in Applied Mathematics, pp. 571–616. Chapman and Hall, Boca Raton (2000)

[20] Gal, S.G., Niculescu, P.C.: Approximation of random functions by stochastic Bernstein polynomials in capacity spaces. Carpathian J. Math. **37**(2), 185–194 (2021)

[21] Gao, W., Sun, X., Wu, Z., Zhou, X.: Multivariate Monte Carlo approximation based on scattered data. SIAM J. Sci. Comput. **42**, 2262–2280 (2020)

[22] Gavrea, I., Ivan, M.: The Bernstein Voronovskaja-type theorem for positive linear operators. J. Approx. Theory **192**, 291–296 (2015)

[23] Ignatov, Z.G., Mills, T.M., Tzankova, I.P.: On the rate of approximation of random functions. Serdica Bulg. Math. Publ. **18**, 240–247 (1992)

[24] Kamolov, A.I.: On exact estimates of approximation of random processes (in Russian). Dokl. Akad. Nauk UzSSR **11**, 4–6 (1986)

[25] Lorentz, G.G.: Bernstein Polynomials, Mathematical Expositions, vol. 8. University of Toronto Press, Toronto (1953)

[26] Massart, P.: The tight constant in the Dvoretzky-Kiefer-Wolfowitz inequality. Ann. Probab. **18**(3), 1269–1283 (1990)

[27] Niederrreiter, H.: On the existence of uniformly distributed sequences in compact spaces. Compos. Math. **25**(1), 93–99 (1972)

[28] Marchal, O., Arbel, J.: On the sub-Gaussianity of the Beta and Dirichlet distributions. Electron. Commun. Probab. **22**(54), 1–14 (2017)

[29] Onicescu, O., Istratescu, V.I.: Approximation theorems for random functions. Rend. Mat. (Roma) Ser. VI **8**, 65–81 (1975)

[30] Onicescu, O., Istratescu, V.I.: Approximation theorems for random functions. II. Rend. Mat. (Roma) Ser. VI **11**(4), 585–589 (1978)

[31] Sikkema, P.C.: Der Wert einiger Konstanten in der Theorie der Approximation mit Bernstein-Polynomen, (German). Numer. Math. **3**, 107–116 (1961)

[32] Sikkema, P.C.: Über den Grad der approximation mit Bernstein-Polynomen (German). Numer. Math. **1**, 221–239 (1959)

[33] Sun, X., Wu, Z.: Chebyshev type inequality for stochastic Bernstein polynomials. Proc. Am. Math. Soc. **147**(2), 671–679 (2019)

[34] Sun, X., Wu, Z., Zhou, X.: On probabilistic convergence rates of stochastic Bernstein polynomials. Math. Comp. **90**(328), 813–830 (2021)

[35] Totik, V.: Approximation by Bernstein polynomials. Am. J. Math. **116**, 995–1018 (1994)

[36] Vitale, R.A.: A Bernstein polynomial approach to density function estimation, statistical inference and related topics 2. In: Proceedings of the Summer Research Institute on Statistical Inference for Stochastic Processes, Bloomington, Indiana, July 31–August 9, pp. 87–99 (1975)

[37] Wagner, G.: On a new method for constructing good point sets on spheres. Discret. Comput. Geom. **9**, 111–129 (1993)

[38] Wu, Z., Sun, X., Ma, L.: Sampling scattered data with Bernstein polynomials: stochastic and deterministic error estimates. Adv. Comput. Math. **38**, 187–205 (2013)

[39] Wu, Z., Zhou, X.: Polynomial convergence order of stochastic Bernstein approximation. Adv. Comput. Math. **46**, 25 (2020)

Xingping Sun
Department of Mathematics
Missouri State University
Springfield
MO 65897
USA
e-mail: `xsun@missouristate.edu`

Xuan Zhou
School of Mathematical Sciences
Fudan University
Shanghai
People's Republic of China
e-mail: `15110180025@fudan.edu.cn`

Received: November 25, 2021.
Revised: April 9, 2022.
Accepted: August 6, 2022.