Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2014, Article ID 454375, 10 pages
http://dx.doi.org/10.1155/2014/454375

Research Article

Analysis of Approximation by Linear Operators on Variable $L_\rho^{p(\cdot)}$ Spaces and Applications in Learning Theory

Bing-Zheng Li[1] and Ding-Xuan Zhou[2]

[1] *Department of Mathematics, Zhejiang University, Hangzhou 310027, China*
[2] *Department of Mathematics, City University of Hong Kong, Kowloon, Hong Kong*

Correspondence should be addressed to Ding-Xuan Zhou; mazhou@cityu.edu.hk

Received 7 May 2014; Accepted 30 June 2014; Published 16 July 2014

Academic Editor: Uno Hämarik

This paper is concerned with approximation on variable $L_\rho^{p(\cdot)}$ spaces associated with a general exponent function p and a general bounded Borel measure ρ on an open subset Ω of $\mathbb{R}^d$. We mainly consider approximation by Bernstein type linear operators. Under an assumption of log-Hölder continuity of the exponent function p, we verify a conjecture raised previously about the uniform boundedness of Bernstein-Durrmeyer and Bernstein-Kantorovich operators on the $L_\rho^{p(\cdot)}$ space. Quantitative estimates for the approximation are provided for high orders of approximation by linear combinations of such positive linear operators. Motivating connections to classification and quantile regression problems in learning theory are also described.

1. Introduction

Approximation by Bernstein type positive linear operators has a long history and is an important topic in approximation theory. It started with Bernstein operators [1] for proving the Weierstrass theorem about the denseness of the set of polynomials in the space $C[0,1]$ of continuous functions on the interval $[0,1]$. These classical operators are defined as $B_n(f,x) = \sum_{k=0}^n f(k/n)p_{n,k}(x)$ for $x \in [0,1]$ and $f \in C[0,1]$ with the Bernstein basis given by $p_{n,k}(x) = \binom{n}{k} x^k(1-x)^{n-k}$. The Bernstein operators have been extended in various forms for the purpose of approximating discontinuous functions, by replacing the point evaluation functionals by some integrals. The classical examples for approximation in $L^p[0,1]$ (with $1 \le p < \infty$), the Banach space of all integrable functions f on $[0,1]$ with the norm $\|f\|_{L^p} = (\int_0^1 |f(x)|^p dx)^{1/p}$, are Bernstein-Kantorovich operators [2]

$$K_n(f,x) = \sum_{k=0}^{n} (n+1) \int_{k/(n+1)}^{(k+1)/(n+1)} f(t)\, dt\, p_{n,k}(x), \quad x \in [0,1] \tag{1}$$

and Bernstein-Durrmeyer operators [3]

$$D_n(f,x) = \sum_{k=0}^{n} (n+1) \int_0^1 p_{n,k}(t) f(t)\, dt\, p_{n,k}(x), \quad x \in [0,1]. \tag{2}$$

Quantitative estimates for approximation by Bernstein type positive linear operators in $C[0,1]$ or $L^p[0,1]$ have been presented in a large literature (e.g., [4, 5]). See the book [6] and references therein for details and extensions to infinite intervals and linear combinations of positive operators for achieving high orders of approximation.

In this paper we provide a general framework for approximation by linear operators on variable $L_\rho^{p(\cdot)}(\Omega)$ spaces on an open subset Ω of $\mathbb{R}^d$. Here $p : \Omega \to [1,\infty)$ is a measurable function called the *exponent function* and ρ is a positive bounded Borel measure on Ω. The variable space $L_\rho^{p(\cdot)}(\Omega)$ is a generalization of the weighted L^p spaces with a constant exponent $p \in [1,\infty]$. It consists of all the measurable functions f on Ω such that $\int_\Omega (|f(x)|/\lambda)^{p(x)} d\rho < \infty$ for some $\lambda > 0$. The norm is defined by

$$\|f\|_{L_\rho^{p(\cdot)}} = \inf\left\{\lambda > 0 : \int_\Omega \left(\frac{|f(x)|}{\lambda}\right)^{p(x)} d\rho \le 1\right\}. \quad (3)$$

The space $L_\rho^{p(\cdot)}(\Omega)$ is a Banach space [7]. The idea of variable $L^{p(\cdot)}$ spaces was introduced by Orlicz [8]. Motivated by connections to variational integrals with nonstandard growth related to modeling of electrorheological fluids [9], these function spaces have been developed in analysis and research topics include boundedness of maximal operators, continuity of translates, and denseness of smooth functions. We will not go into details which can be found in [7, 10] and references therein. Instead, we only mention the following core condition on the log-Hölder continuity of the exponent function which leads to the boundedness of Hardy-Littlewood maximal operators and the rich theory of the variable $L_\rho^{p(\cdot)}(\Omega)$ spaces.

Definition 1. We say that the exponent function $p : \Omega \to [1,\infty)$ is log-Hölder continuous if there exist positive constants $A_p > 0$ such that

$$|p(x) - p(y)| \le \frac{A_p}{-\log|x-y|}, \quad x, y \in \Omega, |x-y| < \frac{1}{2}. \quad (4)$$

We say that p is log-Hölder continuous at infinity (when Ω is unbounded) if there holds

$$|p(x) - p(y)| \le \frac{A_p}{\log(e+|x|)}, \quad x, y \in \Omega, |y| \ge |x|. \quad (5)$$

Denote

$$p_- = \inf_{x\in\Omega} p(x), \qquad p_+ = \sup_{x\in\Omega} p(x). \quad (6)$$

The issue of approximation by Bernstein type positive linear operators on variable $L_\rho^{p(\cdot)}(\Omega)$ spaces was raised by the second author in [11]. It turned out that the variety of the exponent function p creates technical difficulty in the study of approximation. In particular, the uniform boundedness of the Bernstein-Kantorovich operators (1) and Bernstein-Durrmeyer operators (2) is already a difficult problem. The key analysis in [11] is to show that the Bernstein-Kantorovich operators and Bernstein-Durrmeyer operators are uniformly bounded when the exponent function p is Lipschitz α for some $\alpha \in (0,1]$. It was conjectured there that the uniform boundedness still holds when p is log-Hölder continuous. The first main result of this paper is to confirm this conjecture in Theorem 6 below.

Our second main result is to abandon the positivity and present quantitative estimates for the high order approximation by linear operators including linear combinations of Bernstein type positive linear operators, extending the results in [11] for the first order approximation by positive operators.

2. Motivations from Learning Theory

Our main motivation for considering the approximation of functions by linear operators on variable $L_\rho^{p(\cdot)}(\Omega)$ spaces is from learning theory. Besides the example of extending the Bernstein-Durrmeyer operators (2) to those associated with a general probability measure ρ on Ω in [12, 13] for the multivariate case, we mention two learning theory settings here. Since error analysis for concrete learning algorithms in terms of the introduced noise conditions involves sample error estimates which are out of the scope of this paper, we leave the detailed error bounds to our further study.

2.1. Noise Conditions for Classification and Approximation. The first learning theory setting related to approximation on variable $L_\rho^{p(\cdot)}(\Omega)$ spaces is noise conditions for binary classification. Here Ω is an input space consisting of possible events while the output space is denoted as $Y = \{1, -1\}$. A Borel probability measure P on the product space $\Omega \times Y$ can be decomposed into its marginal distribution P_Ω on Ω and conditional distributions $P(\cdot \mid x)$ for $x \in \Omega$. A binary classifier $f : \Omega \to Y$ makes predictions $f(x) \in Y$ for future events $x \in \Omega$. The best classifier f_c, called Bayes rule, is given by $f_c(x) = 1$ if $P(1 \mid x) > 1/2$ and -1 otherwise. The probability measure P fits the binary classification problem well if the conditional probabilities $P(1 \mid x)$ and $P(-1 \mid x)$ are well separated from the boundary $1/2$ for most events x. Their separations are equivalent to the separation of the value $f_P(x)$ of the regression function $f_P(x) = \int_Y y\,dP(y \mid x) = P(1 \mid x) - P(-1 \mid x)$ from 0 and can be measured in various quantitative ways. The Tsybakov noise condition [14] with noise exponent $q \in (0,\infty]$ asserts that for some constant $c_q > 0$, there holds

$$P_\Omega\left(\left\{x \in \Omega : 0 < |f_P(x)| \le c_q t\right\}\right) \le t^q, \quad \forall t > 0. \quad (7)$$

When $q = \infty$, Tsybakov noise condition (7) means $|f_P(x)| \ge c_q$ almost surely, and $f_P(x)$ is well separated from 0. The case $q < \infty$ means the measure of the set of events x with $f_P(x)$ not well separated from 0 decays polynomially fast as the threshold $c_q t$ tends to 0. More details about the Tsybakov noise condition, the so-called Tsybakov function, and its applications to the study of classification problems can also be found in [15]. Here we introduce a noise condition by allowing some noise situations measured by an exponent function p.

Example 2. We say that the probability measure P satisfies the noise condition associated with an exponent function $p : \Omega \to [0,\infty)$ if for some $\lambda > 0$, there holds

$$\int_\Omega \left(\frac{|f_P(x)|}{\lambda}\right)^{p(x)} dP_\Omega < \infty. \quad (8)$$

Remark 3. The above condition can be applied to the regression setting for dealing with unbounded regression functions. When p takes values on $[1,\infty)$, the above condition is equivalent to the requirement $f_P \in L_\rho^{p(\cdot)}(\Omega)$ with $\rho = P_\Omega$.

When $d = 1$ and $\Omega = \mathbb{R}$, we apply the classical identity $\int_{\mathbb{R}} g(x)dP_\Omega = \int_0^\infty P_\Omega(\{x \in \Omega : g(x) \geq t\})dt$ to the nonnegative function $g(x) = (|f_P(x)|/\lambda)^{p(x)}$ and find that the above condition is equivalent to

$$\int_0^\infty P_\Omega\left(\left\{x \in \Omega : |f_P(x)| \geq \lambda t^{1/p(x)}\right\}\right) dt < \infty. \quad (9)$$

This illustrates some similarity between the noise condition (8) and Tsybakov noise condition (7).

The following is an example to show some differences.

Example 4. Let $d \in \mathbb{N}$ and $\Omega = \{x \in \mathbb{R}^d : |x| \leq 1/2\}$. If P_Ω is the normalized Lebesgue measure on Ω and $f_P(x) = e^{-1/|x|}$, then the measure P satisfies the noise condition associated with the exponent function $p(x) = |x|$ but does not satisfy the Tsybakov noise condition (7) with any $q \in (0, \infty]$. In fact we have $\int_\Omega |f_P(x)|^{p(x)} dP_\Omega = 1/e < \infty$ while for any $q \in (0, \infty)$ and $c_q > 0$, we have $P_\Omega(\{x \in \Omega : 0 < |f_P(x)| \leq c_q t\}) = (2/-\log(c_q t))^d > t^q$ for $t \in (0, \min\{t^*, 1/c_q e^2\})$ with t^* being the positive solution to the equation $t^{q/d} \log 1/c_q t = 2$.

2.2. Noise Conditions for Quantile Regression and Approximation. The second learning theory setting related to approximation on variable $L_\rho^{p(\cdot)}(\Omega)$ spaces is noise conditions for quantile regression. Here the output space is $Y = \mathbb{R}$. Similar to the least squares regression [16] for learning means of conditional distributions $P(\cdot \mid x)$ but providing richer information [17] about response variables such as stretching or compressing tails, the learning problem for quantile regression aims at estimating quantiles of conditional distributions. With a quantile parameter $0 < \tau < 1$, the value of a quantile regression function $f_{P,\tau}$ at $x \in \Omega$ is defined by its value $f_{P,\tau}(x)$ as a τ-quantile of $P(\cdot \mid x)$, that is, a value $t^* \in Y$ satisfying

$$\begin{gathered} P\left(\left\{y \in Y : y \leq t^*\right\} \mid x\right) \geq \tau, \\ P\left(\left\{y \in Y : y \geq t^*\right\} \mid x\right) \geq 1 - \tau. \end{gathered} \quad (10)$$

Quantile regression has been studied by kernel-based regularization schemes in a learning theory literature (e.g., [18, 19]). For optimal error analysis of these learning algorithms, asymptotic behaviors of the conditional distributions near the τ-quantiles are needed. In particular, one is interested in how slow the following function decays as t decreases:

$$\begin{aligned} F_{P,\tau}(x) = \min\{&P\left(\left\{y \in Y : t^* - t < y < t^*\right\} \mid x\right), \\ &P\left(\left\{y \in Y : t^* + t > y > t^*\right\} \mid x\right)\}. \end{aligned} \quad (11)$$

A noise condition was introduced in [18] by requiring lower bounds $F_{P,\tau}(x) \geq b_x t^{q-1}$ for every $t \in [0, a_x]$ and some $q \in (1, \infty)$, $p \in (0, \infty)$ and constants b_x, $a_x > 0$ satisfying $(b_x a_x^{q-1})^{-1} \in L_{P_\Omega}^p$. This condition was extended to a logarithmic bound in [19] by replacing t^{q-1} by $(\log(1/t))^{-q}$ and a_x^{q-1} by $(\log(1/a_x))^{-q}$. Here we introduce the following noise condition which is more general than the one in [18] by allowing the indices q, p to depend on the events $x \in \Omega$.

Example 5. We say that the probability measure P satisfies the quantile noise condition associated with exponent functions $q : \Omega \to (1, \infty)$ and $p : \Omega \to (0, \infty)$ if for every $x \in X$, there exist a τ-quantile $t^* \in \mathbb{R}$ and constants $a_x \in (0, 2]$, $b_x > 0$ such that for each $u \in [0, a_x]$

$$F_{P,\tau}(x) \geq b_x t^{q(x)-1} \quad (12)$$

and that for some $\lambda > 0$, there holds $\int_\Omega (1/\lambda b_x a_x^{q(x)-1})^{p(x)} dP_\Omega < \infty$.

While the lower bounds (12) imply polynomial decays of the conditional distributions near the τ-quantiles with a power index depending on the event, the finiteness of the integral is equivalent to the requirement that the function $1/b_x a_x^{q(x)-1}$ lies in the variable $L_{P_\Omega}^{p(\cdot)}(\Omega)$ space (when p takes values in $[1, \infty)$).

3. Main Results for Approximation on $L_\rho^{p(\cdot)}$

Our first theorem is about the uniform boundedness of a sequence of linear operators on the variable $L_\rho^{p(\cdot)}$ spaces. These operators take the form

$$L_n(f, x) = \int_\Omega K_n(x, t) f(t) d\rho(t), \quad x \in \Omega, f \in L_\rho^{p(\cdot)} \quad (13)$$

in terms of their kernels $\{K_n(x,t)\}_{n=1}^\infty$ defined on $\Omega \times \Omega$. We assume that the kernels satisfy the following three conditions with some positive constants $C_0 \geq 1$, b, $\overline{C}_b$, and C_r (depending on $r \in \mathbb{N}$)

$$\begin{gathered} \sup_{t \in \Omega} \int_\Omega |K_n(x,t)| d\rho(x) \leq C_0, \\ \sup_{x \in \Omega} \int_\Omega |K_n(x,t)| d\rho(t) \leq C_0, \end{gathered} \quad (14)$$

$$\sup_{x,t \in \Omega} |K_n(x,t)| \leq \overline{C}_b n^b, \quad \forall n \in \mathbb{N}, \quad (15)$$

$$\int_\Omega \int_\Omega |K_n(x,t)| |t - x|^{2r} d\rho(t) d\rho(x) \leq C_r n^{-r}, \quad \forall n \in \mathbb{N}, r \in \mathbb{N}. \quad (16)$$

Then the uniform boundedness follows, which will be proved in Section 5.

Theorem 6. *Let $\Omega \subseteq \mathbb{R}^d$ be an open set, and an exponent function $p : \Omega \to (1, \infty)$ satisfy $1 < p_- < p_+ < \infty$ and the log-Hölder continuity condition* (4). *If the kernels $\{K_n(x,t)\}_{n=1}^\infty$ satisfy conditions* (14), (15), *and* (16), *then the operators $\{L_n\}_{n=1}^\infty$ on $L_\rho^{p(\cdot)}$ defined by* (13) *are uniformly bounded as*

$$\|L_n\| \leq M_{p,b}, \quad \forall n \in \mathbb{N} \quad (17)$$

by a positive constant $M_{p,b}$ (depending on p and the constants in (14), (15), *and* (16), *given explicitly in the proof).*

Our second theorem gives orders of approximation when the approximated function has some smoothness stated in terms of a $\mathscr{K}$-functional. Define a Hölder space with index $r \in \mathbb{N}$ on Ω by

$$W_p^{r,\infty} = \left\{ g \in L_\rho^{p(\cdot)} : \|g\|_{p,r,\infty} < \infty \right\}, \tag{18}$$

where $\|g\|_{p,r,\infty}$ is the norm given by $\|g\|_{p,r,\infty} = \|g\|_{L_\rho^{p(\cdot)}} + \sum_{|\alpha|_1 \le r} \|D^\alpha g\|_\infty$ with $D^\alpha g = \partial^{|\alpha|_1} / \partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}$ and $|\alpha|_1 = \sum_{j=1}^d \alpha_j$ for $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{Z}_+^d$. The $\mathscr{K}$-functional $\mathscr{K}_r(f,t)_{p(\cdot)}$ is defined by

$$\mathscr{K}_r(f,t)_{p(\cdot)} = \inf_{g \in W_p^{r,\infty}} \left\{ \|f-g\|_{L_\rho^{p(\cdot)}} + t\|g\|_{p,r,\infty} \right\}, \quad t > 0. \tag{19}$$

Denote $C_0^\infty(\Omega)$ as the space of all compactly supported C^∞ functions on Ω. From [7], we know that when $p^+ < \infty$, $C_0^\infty(\Omega)$ is dense in $L_\rho^{p(\cdot)}(\Omega)$. Hence for any $f \in L_\rho^{p(\cdot)}(\Omega)$, there holds $\mathscr{K}_r(f,t)_{p(\cdot)} \to 0$ as $t \to 0$.

The following theorem, to be proved in Section 5 and extending the results for $r = 1$ in [11], gives orders of approximation by linear operators on $L_\rho^{p(\cdot)}(\Omega)$ when the $\mathscr{K}$-functional has explicit decay rates.

Theorem 7. *Under the assumption of Theorem 6, if Ω is convex, $r \in \mathbb{N}$ and the kernels satisfy $rp_- \ge 2$, and*

$$\int_\Omega K_n(x,t)(t-x)^\alpha d\rho(t) = \delta_{\alpha,0}, \quad \forall \alpha \in \mathbb{Z}^d \text{ with } |\alpha|_1 < r \tag{20}$$

for almost every $x \in \Omega$, then there holds for any $f \in L_\rho^{p(\cdot)}$,

$$\|L_n(f) - f\|_{L_\rho^{p(\cdot)}} \le A_{p,b,d} \mathscr{K}_r\left(f, n^{-r_-/p_+}\right)_{p(\cdot)}, \tag{21}$$

where r_- is the integer part of $rp_-/2$ and the constant $A_{p,b,d}$ is independent of $f \in L_\rho^{p(\cdot)}$ (given explicitly in the proof).

The vanishing moment assumption (20) corresponds to Strang-Fix type conditions in the literature of shift-invariant spaces, for example [20, 21]. It has appeared in the literature of Bernstein type operators when linear combinations are considered, as described by (34) in the next section.

4. Approximation by Bernstein Type Operators

In this section we apply our main results to Bernstein type positive linear operators and give high orders of approximation by linear combinations of these operators on variable $L_\rho^{p(\cdot)}(\Omega)$ spaces. We demonstrate the analysis for the general Bernstein-Durrmeyer operators in detail and describe briefly results for the general Bernstein-Kantorovich operators as an example of other families of operators.

The Bernstein-Durrmeyer operators on an open simplex

$$\Omega = S = \left\{ x \in \mathbb{R}_+^d : |x|_1 < 1 \right\} \tag{22}$$

associated with a general positive Borel measure ρ on S are defined as

$$D_n(f,x) = \sum_{|\alpha|_1 \le n} \frac{\int_S f(t) p_{n,\alpha}(t) d\rho(t)}{\int_S p_{n,\alpha}(t) d\rho(t)} p_{n,\alpha}(x), \quad f \in L_\rho^1, x \in S, \tag{23}$$

where for $x = (x_1, \ldots, x_d) \in S \subset \mathbb{R}^d$, $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{Z}_+^d$, we denote

$$p_{n,\alpha}(x) = \frac{n!}{\prod_{j=1}^d \alpha_j! (n - |\alpha|_1)!} \prod_{j=1}^d x_j^{\alpha_j} (1 - |x|_1)^{n - |\alpha|_1}. \tag{24}$$

The classical Bernstein-Durrmeyer operators (2) on $\Omega = (0,1)$ ($d = 1$) with $d\rho(x) = dx$ have been well studied (e.g., [22]) and extended to a multivariate form with respect to Jacobi weights $d\rho(x) = \prod_{j=1}^d x_j^{q_j} dx$ (e.g., [23]). Bernstein-Durrmeyer operators on $\Omega = (0,1)$ with respect to an arbitrary Borel probability measure were introduced in [12] and applied to error analysis of learning algorithms for support vector machine classifications. The multidimensional version of such linear operators (23) was introduced in [13]. In [24], the first author showed for a constant exponent function $p(x) \equiv p \in [1,\infty)$ that $\lim_{n\to\infty} \|f - D_n(f)\|_{L_\rho^p} = 0$ for any $f \in L_\rho^p$. The case $p \equiv \infty$ was studied in [25, 26]. Here we consider the case with a general exponent function satisfying $1 < p_- < p_+ < \infty$ and the log-Hölder continuity condition (4).

By applying Theorem 6, we can prove the uniform boundedness of the Bernstein-Durrmeyer operators (23).

Proposition 8. *Let $\Omega = S$, ρ be a Borel probability measure on S, and an exponent function $p : S \to (1,\infty)$ satisfy $1 < p_- < p_+ < \infty$ and the log-Hölder continuity condition (4). If there exist positive constants b and C_b^* such that $\int_S p_{n,\alpha}(t) d\rho(t) \ge C_b^* n^{-b}$ for $n \in \mathbb{N}$ and $|\alpha|_1 \le n$, then for the Bernstein-Durrmeyer operators defined on $L_\rho^{p(\cdot)}(\Omega)$ by (23), there exists a positive constant $M_{p,b}$ depending only on p, b, C_b^* such that*

$$\|D_n\| \le M_{p,b}, \quad \forall n \in \mathbb{N}. \tag{25}$$

Proof. Define a sequence of kernels $\{\Psi_n\}$ on $S \times S$ by

$$\Psi_n(x,t) = \sum_{|\alpha|_1 \le n} \frac{p_{n,\alpha}(t) p_{n,\alpha}(x)}{\int_S p_{n,\alpha}(t) d\rho(t)}. \tag{26}$$

Then the Bernstein-Durrmeyer operators (23) can be written as

$$D_n(f,x) = \int_S \Psi_n(x,t) f(t) d\rho(t). \tag{27}$$

So we only need to check the three conditions (14), (15), and (16) of Theorem 6.

Since $\Psi_n(x,t) \geq 0$, we know that

$$\begin{aligned}\int_S |\Psi_n(x,t)|\, d\rho(x) &= \int_S \Psi_n(x,t)\, d\rho(x) \\ &= \sum_{|\alpha|_1\leq n} \frac{p_{n,\alpha}(t)\int_S p_{n,\alpha}(x)\, d\rho(x)}{\int_S p_{n,\alpha}(t)\, d\rho(t)} \\ &= \sum_{|\alpha|_1\leq n} p_{n,\alpha}(t) \equiv 1.\end{aligned} \tag{28}$$

The same is true for $\int_S |\Psi_n(x,t)| d\rho(t) \equiv 1$. So we know that condition (14) holds with the constant $C_0 = 1$.

Applying the lower bound $\int_S p_{n,\alpha}(t) d\rho(t) \geq C_b^* n^{-b}$ and the inequality $p_{n,\alpha}(t) \leq 1$, we see that for any $n \in \mathbb{N}$, and $x, t \in S$

$$|\Psi_n(x,t)| \leq \sum_{|\alpha|_1\leq n} \frac{p_{n,\alpha}(x)}{C_b^* n^{-b}} = \frac{1}{C_b^*} n^b. \tag{29}$$

Hence condition (15) holds with $\overline{C}_b = 1/C_b^*$.

As for the last condition, we separate $t - x$ into $t - (\alpha/n) + (\alpha/n) - x$ and find that for any $r \in \mathbb{N}$ and $n \in \mathbb{N}$

$$\begin{aligned}&\int_S\int_S |\Psi_n(x,t)||t-x|^{2r} d\rho(t)\, d\rho(x) \\ &\leq 2^{2r}\Bigg\{\int_S\int_S \sum_{|\alpha|_1\leq n} \frac{p_{n,\alpha}(t)\, p_{n,\alpha}(x)}{\int_S p_{n,\alpha}(t)\, d\rho(t)} \times \left|t - \frac{\alpha}{n}\right|^{2r} d\rho(t)\, d\rho(x) \\ &\quad + \int_S\int_S \sum_{|\alpha|_1\leq n} \frac{p_{n,\alpha}(t)\, p_{n,\alpha}(x)}{\int_S p_{n,\alpha}(t)\, d\rho(t)} \times \left|\frac{\alpha}{n} - x\right|^{2r} d\rho(t)\, d\rho(x)\Bigg\} \\ &= 2^{2r}\Bigg\{\int_S \sum_{|\alpha|_1\leq n} p_{n,\alpha}(t)\left|t-\frac{\alpha}{n}\right|^{2r} d\rho(t) + \int_S \sum_{|\alpha|_1\leq n} p_{n,\alpha}(x)\left|\frac{\alpha}{n} - x\right|^{2r} d\rho(x)\Bigg\} \\ &= 2^{2r+1}\int_S B_n\left((\cdot - x)^{2r}, x\right) d\rho(x),\end{aligned} \tag{30}$$

where B_n is the multidimensional Bernstein operators on the closure $\overline{S}$ of S defined by

$$B_n(f,x) = \sum_{|\alpha|_1\leq n} f\left(\frac{\alpha}{n}\right) p_{n,\alpha}(x), \quad x \in S, f \in C(\overline{S}). \tag{31}$$

It is well known [6] for the multidimensional Bernstein operators that there exists a constant $C_{d,r}$ depending only on d and r such that

$$B_n\left((\cdot - x)^{2r}, x\right) \leq C_{d,r} n^{-r}, \quad \forall x \in \overline{S}. \tag{32}$$

It follows that

$$\int_S\int_S |\Psi_n(x,t)|\, |t-x|^{2r} d\rho(t)\, d\rho(x) \leq 2^{2r+1} C_{d,r} n^{-r} \tag{33}$$

and condition (16) holds true with $C_r = 2^{2r+1} C_{d,r}$.

With all the three conditions verified, the desired uniform bound (25) for the Bernstein-Durrmeyer operators follows from Theorem 6. This proves the proposition. □

The Bernstein-Durrmeyer operators (23) are positive, which prevent from achieving high order approximation due to a saturation phenomenon. Linear combinations of such operators can be used to get high orders of approximation. The idea and literature review of this method can be found in [6] while further developments will not be mentioned here. The linear combinations are defined as

$$L_{n,r}(f,x) = \sum_{i=0}^{m_{d,r}} C_i(n)\, D_{n_i}(f,x), \tag{34}$$

where $m_{d,r} = (d + r - 1)!/d!(r - 1)!$ is the dimension of the space of polynomials of degree at most $r - 1$, and with two positive constants $\widetilde{B}_1, \widetilde{B}_2$ independent of n, we have

$$\begin{gathered} n = n_0 < n_1 < \cdots < n_{m_{d,r}} \leq \widetilde{B}_1 n, \qquad \sum_{i=0}^{m_{d,r}} |C_i(n)| \leq \widetilde{B}_2, \\ \sum_{i=0}^{m_{d,r}} C_i(n)\, D_{n_i}\left((\cdot - x)^{\alpha}, x\right) = \delta_{\alpha,0}, \quad \forall 0 \leq |\alpha|_1 \leq r - 1. \end{gathered} \tag{35}$$

For the classical Bernstein-Durrmeyer operators with respect to the Lebesgue measure (or even the Jacobi weights), the existence of the above linear combinations can be seen and found in the literature. The existence of such linear combinations with respect to the arbitrary measure ρ is a nontrivial problem and deserves intensive study. This technical question is out of the scope of this paper and will be discussed in our further work. Here we concentrate on the variable $L_\rho^{p(\cdot)}(\Omega)$ spaces and state the following result for the high orders of approximation under the condition (35) which is an immediate consequence of Theorem 7.

Proposition 9. *Under the assumption of Proposition 8, if* $2 \leq r \in N$ *and the operators* $\{L_{n,r}\}_{n\in\mathbb{N}}$ *defined by* (34) *satisfy* (35), *then for any* $f \in L_\rho^{p(\cdot)}$, *we have*

$$\|L_{n,r}(f) - f\|_{L_\rho^{p(\cdot)}} \leq A_{p,b,d}\mathscr{K}_r\left(f, n^{-r_-/p_+}\right)_{p(\cdot)}, \tag{36}$$

where r_- *is the integer part of* $rp_-/2$ *and the constant* $A_{p,b,d}$ *is independent of* $f \in L_\rho^{p(\cdot)}$.

Let us now briefly describe approximation results for the Bernstein-Kantorovich operators on S defined [27] as

$$BK_n(f,x) = \sum_{|\alpha|_1\leq n} \frac{\int_{S_{n,\alpha}} f(t)\, d\rho(t)}{\rho(S_{n,\alpha})} p_{n,\alpha}(x), \quad f \in L_\rho^1, x \in S, \tag{37}$$

where $\{S_{n,\alpha}\}_\alpha$ are subdomains of S defined by

$$S_{n,\alpha} = \left\{ x \in S : x \in \prod_{i=1}^{d} \left[\frac{\alpha_i}{n+1}, \frac{\alpha_i}{n+1} \right), |x|_1 \le \frac{|\alpha|_1 + 1}{n+1} \right\}, \quad |\alpha|_1 \le n. \tag{38}$$

In the same way as for the Bernstein-Durrmeyer operators, we have the following results for the Bernstein-Kantorovich operators.

Proposition 10. *Under the assumption of Proposition 8 for p, if there exist positive constants b and C_b^* such that $\rho(S_{n,\alpha}) \ge C_b^* n^{-b}$ for $n \in \mathbb{N}$ and $|\alpha|_1 \le n$, then for the Bernstein-Kantorovich operators defined on $L_\rho^{p(\cdot)}(\Omega)$ by (37), there exists a positive constant $M_{p,b}$ depending only on p, b, C_b^* such that*

$$\|BK_n\| \le M_{p,b}, \quad \forall n \in \mathbb{N}. \tag{39}$$

If $2 \le r \in N$ and, with D_{n_i} replaced by BK_{n_i}, the operators $\{L_{n,r}\}_{n\in\mathbb{N}}$ defined by (34) *satisfy* (35), *then*

$$\|L_{n,r}(f) - f\|_{L_\rho^{p(\cdot)}} \le A_{p,b,d} \mathscr{K}_r\left(f, n^{-r_-/p_+}\right)_{p(\cdot)}, \quad \forall f \in L_\rho^{p(\cdot)}. \tag{40}$$

5. Proof of Main Results

In this section we give detailed proof of our main results. Let us first prove Theorem 6.

Proof of Theorem 6. Let $f \in L^{p(\cdot)}$ have norm 1, which implies $\int_\Omega |f(t)|^{p(t)} d\rho(t) \le 1$. Choose $\gamma = b/(p_- - 1) > 0$.

For $x \in \Omega$, we define two subsets $\Omega_{n,x}$ and $\Omega'_{n,x}$ of Ω as

$$\Omega_{n,x} = \left\{ t \in \Omega : |f(t)| \le n^\gamma, |t - x| \le n^{-1/4} \right\},$$
$$\Omega'_{n,x} = \left\{ t \in \Omega : |f(t)| \le n^\gamma, |t - x| > n^{-1/4} \right\}. \tag{41}$$

Set

$$L_{n,1}(f,x) = \int_{\Omega\setminus(\Omega_{n,x}\cup\Omega'_{n,x})} K_n(x,t) f(t) \, d\rho(t),$$
$$L_{n,2}(f,x) = \int_{\Omega_{n,x}} K_n(x,t) f(t) \, d\rho(t), \tag{42}$$
$$L_{n,3}(f,x) = \int_{\Omega'_{n,x}} K_n(x,t) f(t) \, d\rho(t).$$

Then the value $L_n(f,x)$ can be decomposed into three parts as

$$L_n(f,x) = L_{n,1}(f,x) + L_{n,2}(f,x) + L_{n,3}(f,x), \quad x \in \Omega, \tag{43}$$

and we have

$$\int_\Omega |L_n(f,x)|^{p(x)} d\rho(x) \le 3^{p_+} \sum_{j=1}^{3} \int_\Omega |L_{n,j}(f,x)|^{p(x)} d\rho(x). \tag{44}$$

In the following we estimate the three terms in (44) separately.

Step 1. Estimating the First Term of (44). By the definition of p_-, we have $p(t) \ge p_- > 1$. For $t \in \Omega \setminus (\Omega_{n,x} \cup \Omega'_{n,x})$, we have $|f(t)| > n^\gamma$ and thereby

$$\begin{aligned} |f(t)| &\le |f(t)| \left(|f(t)| n^{-\gamma} \right)^{p(t)-1} = |f(t)|^{p(t)} n^{\gamma(1-p(t))} \\ &\le |f(t)|^{p(t)} n^{\gamma(1-p_-)} = n^{-b} |f(t)|^{p(t)}. \end{aligned} \tag{45}$$

It follows from condition (15) that

$$\begin{aligned} |L_{n,1}(f,x)| &\le n^{-b} \int_{\Omega\setminus(\Omega_{n,x}\cup\Omega'_{n,x})} |K_n(x,t)| |f(t)|^{p(t)} d\rho(t) \\ &\le n^{-b} \overline{C}_b n^b \int_{\Omega\setminus(\Omega_{n,x}\cup\Omega'_{n,x})} |f(t)|^{p(t)} d\rho(t). \end{aligned} \tag{46}$$

So by the assumption $\int_\Omega |f(t)|^{p(t)} d\rho(t) \le 1$,

$$|L_{n,1}(f,x)| \le \overline{C}_b \int_\Omega |f(t)|^{p(t)} d\rho(t) \le \overline{C}_b. \tag{47}$$

Consequently,

$$\int_\Omega |L_{n,1}(f,x)|^{p(x)} d\rho(x) \le \rho(\Omega) \left(\overline{C}_b^{\,p_-} + \overline{C}_b^{\,p_+} \right). \tag{48}$$

Step 2. Estimating the Second Term of (44). By the condition $\int_\Omega |K_n(x,t)| d\rho(t) \le C_0$ in (14) with $C_0 \ge 1$, we know by the Hölder inequality

$$\begin{aligned} &\int_\Omega |L_{n,2}(f,x)|^{p(x)} d\rho(x) \\ &\le \int_\Omega \left\{ \int_{\Omega_{n,x}} |K_n(x,t)| |f(t)|^{p(x)} d\rho(t) \times \left(\int_{\Omega_{n,x}} |K_n(x,t)| d\rho(t) \right)^{p(x)-1} \right\} d\rho(x) \\ &\le C_0^{p_+-1} \int_\Omega \int_{\Omega_{n,x}} |K_n(x,t)| |f(t)|^{p(t)} |f(t)|^{p(x)-p(t)} \times d\rho(t) \, d\rho(x) \\ &= C_0^{p_+-1} \int_\Omega \int_{\Omega_{n,x}} J_n(f,x,t) |f(t)|^{p(t)} d\rho(t) \, d\rho(x), \end{aligned} \tag{49}$$

where

$$J_n(f,x,t) := |K_n(x,t)| |f(t)|^{p(x)-p(t)}, \quad t \in \Omega_{n,x}, x \in \Omega. \tag{50}$$

For $t \in \Omega_{n,x}$, we have a bound $|f(t)| \le n^\gamma$ and the restriction $|x - t| \le n^{-1/4}$. From the log-Hölder continuity of the exponent function p, there exists a constant A_p only dependent on p such that

$$p(x) - p(t) \le \frac{A_p}{-\log n^{-1/4}} = \frac{4A_p}{\log n}. \tag{51}$$

When $|f(t)| \geq 1$, we find

$$\left|f(t)\right|^{p(x)-p(t)} \leq \left(n^{\gamma}\right)^{4A_p/\log n} \leq n^{4\gamma A_p/\log n} \leq \widehat{M}_p, \tag{52}$$

where the constant number $\widehat{M}_p$ defined by

$$\widehat{M}_p = \sup_{n\in\mathbb{N}} n^{4\gamma A_p/\log n}, \tag{53}$$

is finite because

$$\log n^{4\gamma A_p/\log n} = \frac{4\gamma A_p \log n}{\log n} \longrightarrow 4\gamma A_p, \quad \text{as } n \longrightarrow \infty. \tag{54}$$

When $|f(t)| < 1$, we simply use $|f(t)|^{p(x)-p(t)} \leq |f(t)|^{-p(t)}$. Applying these bounds, we can estimate the core part of Step 2 as

$$\begin{aligned}
&\int_{\Omega}\int_{\Omega_{n,x}} J_n(f,x,t)\left|f(t)\right|^{p(t)} d\rho(t)\, d\rho(x) \\
&\leq \int_{\Omega}\int_{\Omega_{n,x}} \left|K_n(x,t)\right|\left|f(t)\right|^{p(t)}\left(\widehat{M}_p + \left|f(t)\right|^{-p(t)}\right) \times d\rho(t)\, d\rho(x) \\
&= \widehat{M}_p \int_{\Omega}\int_{\Omega_{n,x}} \left|K_n(x,t)\right|\left|f(t)\right|^{p(t)} d\rho(t)\, d\rho(x) \\
&\quad + \int_{\Omega}\int_{\Omega_{n,x}} \left|K_n(x,t)\right| d\rho(t)\, d\rho(x) \\
&\leq \widehat{M}_p C_0 \int_{\Omega} \left|f(t)\right|^{p(t)} d\rho(t) \\
&\quad + C_0\rho(\Omega) \leq C_0\widehat{M}_p + C_0\rho(\Omega).
\end{aligned} \tag{55}$$

Here we have used the assumption $\int_{\Omega}|f(t)|^{p(t)}d\rho(t) \leq 1$ and (14). So the second term of (44) can be estimated as

$$\int_{\Omega}\left|L_{n,2}(f,x)\right|^{p(x)} d\rho(x) \leq C_0^{p_+}\left(\widehat{M}_p + \rho(\Omega)\right). \tag{56}$$

Step 3. Estimating the Third Term of (44). For $t \in \Omega'_{n,x}$, we have $|f(t)| \leq n^{\gamma}$ and $|x-t| > n^{-1/4}$ yielding $n^{1/4}|x-t| > 1$. It follows that

$$\left|L_{n,3}(f,x)\right| \leq \int_{\Omega'_{n,x}} \left|K_n(x,t)\right| n^{\gamma}\left(n^{1/4}|x-t|\right)^{2r/p(x)} d\rho(t), \tag{57}$$

where r is the integer part of $2\gamma p_+ + 1$. Applying the Hölder inequality and (14), we see that

$$\begin{aligned}
&\left|L_{n,3}(f,x)\right|^{p(x)} \\
&\quad \leq \int_{\Omega'_{n,x}} n^{\gamma p(x)+(r/2)}\left|K_n(x,t)\right| |x-t|^{2r} d\rho(t)\, C_0^{p(x)-1}.
\end{aligned} \tag{58}$$

So by the bound $p(x) \leq p_+$ and condition (16), the third term of (44) can be bounded as

$$\begin{aligned}
\int_{\Omega}\left|L_{n,3}(f,x)\right|^{p(x)} d\rho(x) &\leq C_0^{p_+-1} n^{\gamma p_+ + (r/2)} C_r n^{-r} \\
&= C_0^{p_+-1} C_r n^{\gamma p_+ - (r/2)} \leq C_0^{p_+-1} C_r.
\end{aligned} \tag{59}$$

Finally, we put the estimates (48), (56), and (59) into (44) to conclude

$$\begin{aligned}
&\int_{\Omega}\left|L_n(f,x)\right|^{p(x)} d\rho(x) \\
&\quad \leq 3^{p_+}\left(\rho(\Omega)\left(\overline{C}_b^{p_-} + \overline{C}_b^{p_+}\right)\right. \\
&\qquad \left. + C_0^{p_+}\left(\widehat{M}_p + \rho(\Omega)\right) + C_0^{p_+-1}C_r\right).
\end{aligned} \tag{60}$$

Take

$$\begin{aligned}
\lambda = M_{p,b} := 1 + 3^{p_+}&\left(\rho(\Omega)\left(\overline{C}_b^{p_-} + \overline{C}_b^{p_+}\right)\right. \\
&\left. + C_0^{p_+}\left(\widehat{M}_p + \rho(\Omega)\right) + C_0^{p_+-1}C_r\right);
\end{aligned} \tag{61}$$

we find

$$\begin{aligned}
&\int_{\Omega}\left(\frac{\left|L_n(f,x)\right|}{\lambda}\right)^{p(x)} d\rho(x) \\
&\quad \leq \left(\frac{1}{\lambda}\right)^{p_-}\int_{\Omega}\left|L_n(f,x)\right|^{p(x)} d\rho(x) \leq 1.
\end{aligned} \tag{62}$$

This implies $\|L_n(f)\|_{L_\rho^{p(\cdot)}} \leq M_{p,b}$. The bound $M_{p,b}$ is independent of f. So we have $\|L_n\| \leq M_{p,b}$. The proof Theorem 6 is complete. □

We are now in a position to prove Theorem 7.

Proof of Theorem 7. We follow the standard procedure in approximation theory and consider the error $L_n(g,x) - g(x)$ for $g \in W_p^{r,\infty}$. Apply the Taylor expansion

$$\begin{aligned}
g(t) = g(x) &+ \sum_{1\leq|\alpha|_1\leq r-1} \frac{D^{\alpha}g(x)}{\alpha!}(t-x)^{\alpha} \\
&+ R_{g,r}(x,t), \quad x,t \in \Omega,
\end{aligned} \tag{63}$$

where the remainder term $R_{g,r}(x,t)$ is given by

$$\begin{aligned}
&R_{g,r}(x,t) \\
&\quad = \int_0^1 (1-u)^{r-1} \sum_{|\alpha|_1=r} \frac{D^{\alpha}g(x+u(t-x))}{\alpha!}(t-x)^{\alpha} du.
\end{aligned} \tag{64}$$

We see from the vanishing moment condition (20) that

$$
\begin{aligned}
L_n(g,x) - g(x) &= \int_\Omega K_n(x,t)\left\{ g(x) + \sum_{1\le|\alpha|_1\le r-1} \frac{D^\alpha g(x)}{\alpha!}(t-x)^\alpha \right. \\
&\qquad \left. + R_{g,r}(x,t) \right\} d\rho(t) - g(x) \\
&= \int_\Omega K_n(x,t)\left\{ \int_0^1 (1-u)^{r-1} \sum_{|\alpha|_1=r} \frac{D^\alpha g(x+u(t-x))}{\alpha!} \right. \\
&\qquad \left. \times (t-x)^\alpha du \right\} d\rho(t).
\end{aligned} \tag{65}
$$

Since Ω is convex, $x+u(t-x)\in\Omega$ for any $u\in[0,1]$, $x,t\in\Omega$. So $|D^\alpha g(x+u(t-x))| \le \|g\|_{p,r,\infty}$ and we have

$$
\begin{aligned}
|L_n(g,x) - g(x)| &\le \int_0^1 (1-u)^{r-1} \\
&\quad \times \sum_{|\alpha|_1=r} \frac{\|g\|_{p,r,\infty}}{\alpha!}\left\{ \int_\Omega |K_n(x,t)|\,|t-x|^{|\alpha|} d\rho(t)\right\} du \\
&\le d^r \|g\|_{p,r,\infty} \int_\Omega |K_n(x,t)|\,|t-x|^r d\rho(t).
\end{aligned} \tag{66}
$$

By (14) and the Hölder inequality,

$$
\begin{aligned}
&\left(\int_\Omega |K_n(x,t)|\,|t-x|^r d\rho(t)\right)^{p(x)} \\
&\quad \le \int_\Omega |K_n(x,t)|\,|t-x|^{rp(x)} d\rho(t)\, C_0^{p(x)-1}.
\end{aligned} \tag{67}
$$

If we denote the largest integer r_- satisfying $2r_- \le rp_-$ as r_-, and the smallest integer r_+ satisfying $2r_+ \ge rp_+$, we find

$$
|t-x|^{rp(x)} \le \begin{cases} |t-x|^{rp_+} \le |t-x|^{2r_+}, & \text{if } |t-x|\ge 1, \\ |t-x|^{rp_-} \le |t-x|^{2r_-}, & \text{if } |t-x| < 1. \end{cases} \tag{68}
$$

We combine this with (16) and see that for $\lambda = d^r\|g\|_{p,r,\infty}$,

$$
\begin{aligned}
&\int_\Omega \left(\frac{|L_n(g,x) - g(x)|}{\lambda}\right)^{p(x)} d\rho(x) \\
&\le C_0^{p_+-1} \int_\Omega \int_\Omega |K_n(x,t)| \left\{ |t-x|^{2r_+} + |t-x|^{2r_-}\right\} d\rho(t)\, d\rho(x) \\
&\le C_0^{p_+-1}\left(C_{r_+} n^{-r_+} + C_{r_-} n^{-r_-}\right) \le C_0^{p_+-1}\left(C_{r_+} + C_{r_-}\right) n^{-r_-}.
\end{aligned} \tag{69}
$$

When $n \ge C_0^{(p_+-1)/r_-}(C_{r_+}+C_{r_-})^{1/r_-}$, we take

$$
\tilde{\lambda} = d^r \|g\|_{p,r,\infty}\left(C_0^{p_+-1}\left(C_{r_+} + C_{r_-}\right) n^{-r_-}\right)^{1/p_+} \tag{70}
$$

and find

$$
\int_\Omega \left(\frac{|L_n(g,x) - g(x)|}{\tilde{\lambda}}\right)^{p(x)} d\rho(x) \le 1, \tag{71}
$$

which implies

$$
\begin{aligned}
&\|L_n(g) - g\|_{L_\rho^{p(\cdot)}} \\
&\quad \le \tilde{\lambda} \le d^r C_0^{1-(1/p_+)}\left(C_{r_+} + C_{r_-}\right)^{1/p_+} n^{-r_-/p_+} \|g\|_{p,r,\infty}.
\end{aligned} \tag{72}
$$

Thus by Theorem 6 and taking infimum over $g\in W_p^{r,\infty}$, we have

$$
\begin{aligned}
&\|L_n(f) - f\|_{L_\rho^{p(\cdot)}} \\
&\quad \le \inf_{g\in W_p^{r,\infty}} \left\{ \|L_n(f-g)\|_{L_\rho^{p(\cdot)}} + \|L_n(g) - g\|_{L_\rho^{p(\cdot)}} \right. \\
&\qquad\qquad \left. + \|g-f\|_{L_\rho^{p(\cdot)}} \right\} \\
&\quad \le \inf_{g\in W_p^{r,\infty}} \left\{ \left(M_{p,b}+1\right) \|f-g\|_{L_\rho^{p(\cdot)}} \right. \\
&\qquad\qquad \left. + d^r C_0^{1-(1/p_+)}\left(C_{r_+} + C_{r_-}\right)^{1/p_+} n^{-r_-/p_+} \|g\|_{p,r,\infty}\right\} \\
&\quad \le A_{p,b,d} \mathscr{K}_r\left(f, n^{-r_-/p_+}\right)_{p(\cdot)},
\end{aligned} \tag{73}
$$

where the constant $A_{p,b,d}$ is given by

$$
A_{p,b,d} = M_{p,b} + 1 + d^r C_0^{1-(1/p_+)}\left(C_{r_+} + C_{r_-}\right)^{1/p_+}. \tag{74}
$$

When $n < C_0^{(p_+-1)/r_-}(C_{r_+}+C_{r_-})^{1/r_-}$, then from the inequality $\|f-g\|_{L_\rho^{p(\cdot)}} + t\|g\|_{p,r,\infty} \ge t\|f\|_{L_\rho^{p(\cdot)}}$ valid for $t\le 1$ we observe

$$
\mathscr{K}_r\left(f, n^{-r_-/p_+}\right)_{p(\cdot)} \ge n^{-r_-/p_+} \|f\|_{L_\rho^{p(\cdot)}}, \tag{75}
$$

and applying Theorem 6 directly yields

$$
\begin{aligned}
\|L_n(f) - f\|_{L_\rho^{p(\cdot)}} &\le \left(M_{p,b}+1\right) \|f\|_{L_\rho^{p(\cdot)}} \\
&\le \left(M_{p,b}+1\right) n^{r_-/p_+} \mathscr{K}_r\left(f, n^{-r_-/p_+}\right)_{p(\cdot)}
\end{aligned} \tag{76}
$$

and thereby (21) by setting

$$
A_{p,b,d} = \left(M_{p,b}+1\right) C_0^{1-(1/p_+)}\left(C_{r_+} + C_{r_-}\right)^{1/p_+}. \tag{77}
$$

The proof of Theorem 7 is complete. □

6. Further Topics and Discussion

Approximation by linear operators is an important topic in approximation theory. It mainly consists of two families of approximation schemes: Bernstein type positive linear operators and quasi-interpolation type linear operators in multivariate approximation.

In this paper, we mainly consider Bernstein type positive linear operators. We verified a conjecture in [11] about the uniform boundedness of Bernstein-Durrmeyer and Bernstein-Kantorovich operators with respect to an arbitrary Borel measure on $(0, 1)$ on the variable $L_\rho^{p(\cdot)}$ space under the assumption of log-Hölder continuity of the exponent function p. We also provide quantitative estimates for high orders of approximation on the variable $L_\rho^{p(\cdot)}$ by linear combinations of Bernstein type positive linear operators.

The study of quasi-interpolation type linear operators started with the classical work of Schoenberg on cardinal interpolation by B-splines. It has been developed significantly due to important applications in the areas of finite element methods, cardinal interpolation for multivariate approximation, and wavelet analysis. A large class of linear operators for approximating functions on $\mathbb{R}^d$ take the form

$$T(f, x) = \int_{\mathbb{R}^d} \Phi(x, t) f(t) \, dt, \quad x \in \mathbb{R}^d, \tag{78}$$

where $\Phi : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ is a window function satisfying $\int_{\mathbb{R}^d} \Phi(x, t) dt = 1$ and some conditions for decays of $|\Phi(x, t)|$ as $|x - t|$ increases. Quantitative estimates for the approximation of functions in $C(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ or $L^p(\mathbb{R}^d)$ with $1 \le p < \infty$ can be found in a large literature of multivariate approximation (see, e.g., [20, 21, 28]). Establishing analysis for approximation by quasi-interpolation type linear operators on the variable $L_\rho^{p(\cdot)}$ spaces would be an interesting topic. An immediate barrier we meet with such analysis is the boundedness assumption of the measure ρ ($\rho(\Omega) < \infty$). This assumption is not satisfied for most quasi-interpolation type linear operators or the classical Weierstrass (or Gaussian convolution) operators $\mathscr{W}_n(f) = (\sqrt{n/2\pi})^d \int_{\mathbb{R}^d} \exp\{-n\|t - x\|_2^2/2\} f(t) dt$, for which ρ is often the Lebesgue measure on $\mathbb{R}^d$. It is desirable to overcome the technical difficulty and establish error analysis for linear operators with respect to unbounded measures.

We described motivations of our study in learning theory. It would be interesting to implement detailed error analysis for some related learning algorithms in classification and quantile regression by means of our results on orders of approximation by linear operators.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors would like to thank the anonymous referees for their constructive suggestions and comments. The work described in this paper is supported partially by the Research Grants Council of Hong Kong (Project no. CityU 105011). The corresponding author is Ding-Xuan Zhou.

References

[1] S. N. Bernstein, "Démonstration du téoréme de Weirerstrass, fondée sur le calcul des probabilités," *Communications of the Kharkov Mathematical Society*, vol. 13, pp. 1–2, 1913.

[2] L. V. Kantorovich, "Sur certaines developments suivant les polynômes de la forme de S. Bernstein I-II," *Comptes Rendus de l'Académie des Sciences de L'URSS A*, vol. 563–568, pp. 595–600, 1930.

[3] J. L. Durrmeyer, *Une formule d'inversion de la transformée Laplace: applications á la théorie des moments [Thése de 3e cycle: Sciences]*, Faculté des Sciences, l'Université Paris, Paris, France, 1967.

[4] H. Berens and G. G. Lorentz, "Inverse theorems for Bernstein polynomials," *Indiana University Mathematics Journal*, vol. 21, pp. 693–708, 1972.

[5] H. Berens and R. A. DeVore, "Quantitative Korovkin theorems for positive linear operators on *LP*-spaces," *Transactions of the American Mathematical Society*, vol. 245, pp. 349–361, 1978.

[6] Z. Ditzian and V. Totik, *Moduli of Smoothness*, vol. 9 of *Springer Series in Computational Mathematics*, Springer, New York, NY, USA, 1987.

[7] L. Diening, P. Harjulehto, P. Hästö, and M. Ruzicka, *Lebesgue and Sobolev Spaces with Variable Exponents*, Springer, Berlin, Germany, 2011.

[8] W. Orlicz, "Über konjugierte Exponentenfolgen," *Studia Mathematica*, vol. 3, pp. 200–211, 1931.

[9] E. Acerbi and G. Mingione, "Regularity results for a class of functionals with non-standard growth," *Archive for Rational Mechanics and Analysis*, vol. 156, no. 2, pp. 121–140, 2001.

[10] O. Kovácik and J. Rákosnk, "On spaces $L^{p(x)}$ and $W^{1,p(x)}$," *Czechoslovak Mathematical Journal*, vol. 41, no. 116, pp. 592–618, 1991.

[11] D. X. Zhou, "Approximation by positive linear operators on variables $L^{p(x)}$ spaces," *Journal of Applied Functional Analysis*, vol. 9, no. 3-4, pp. 379–391, 2014.

[12] D. X. Zhou and K. Jetter, "Approximation with polynomial kernels and SVM classifiers," *Advances in Computational Mathematics*, vol. 25, no. 1–3, pp. 323–344, 2006.

[13] E. E. Berdysheva and K. Jetter, "Multivariate Bernstein-Durrmeyer operators with arbitrary weight functions," *Journal of Approximation Theory*, vol. 162, no. 3, pp. 576–598, 2010.

[14] A. B. Tsybakov, "Optimal aggregation of classifiers in statistical learning," *The Annals of Statistics*, vol. 32, no. 1, pp. 135–166, 2004.

[15] S. Smale and D. X. Zhou, "Learning theory estimates via integral operators and their approximations," *Constructive Approximation*, vol. 26, no. 2, pp. 153–172, 2007.

[16] S. Smale and D. X. Zhou, "Shannon sampling and function reconstruction from point values," *The American Mathematical Society: Bulletin*, vol. 41, no. 3, pp. 279–305, 2004.

[17] T. Hu, J. Fan, Q. Wu, and D. X. Zhou, "Regularization schemes for minimum error entropy principle," *Analysis and Applications*, 2014.

[18] I. Steinwart and A. Christmann, "Estimating conditional quantiles with the help of the pinball loss," *Bernoulli*, vol. 17, no. 1, pp. 211–225, 2011.

[19] D. H. Xiang, "A new comparison theorem on conditional quantiles," *Applied Mathematics Letters*, vol. 25, no. 1, pp. 58–62, 2012.

[20] J. Lei, R. Jia, and E. W. Cheney, "Approximation from shift-invariant spaces by integral operators," *SIAM Journal on Mathematical Analysis*, vol. 28, no. 2, pp. 481–498, 1997.

[21] K. Jetter and D. X. Zhou, "Order of linear approximation from shift-invariant spaces," *Constructive Approximation*, vol. 11, no. 4, pp. 423–438, 1995.

[22] M. Derriennic, "On multivariate approximation by Bernstein-type polynomials," *Journal of Approximation Theory*, vol. 45, no. 2, pp. 155–166, 1985.

[23] H. Berens and Y. Xu, "On Bernstein-Durrmeyer polynomials with Jacobi weights," in *Approximation Theory and Functional Analysis*, C. K. Chui, Ed., pp. 25–46, Academic Press, Boston, Mass, USA, 1991.

[24] B.-Z. Li, "Approximation by multivariate Bernstein-Durrmeyer operators and learning rates of least-squares regularized regression with multivariate polynomial kernels," *Journal of Approximation Theory*, vol. 173, pp. 33–55, 2013.

[25] E. E. Berdysheva, "Uniform convergence of Bernstein-Durrmeyer operators with respect to arbitrary measure," *Journal of Mathematical Analysis and Applications*, vol. 394, no. 1, pp. 324–336, 2012.

[26] E. E. Berdysheva, "Bernstein–Durrmeyer operators with respect to arbitrary measure, II: pointwise convergence," *Journal of Mathematical Analysis and Applications*, vol. 418, no. 2, pp. 734–752, 2014.

[27] D. X. Zhou, "Converse theorems for multidimensional Kantorovich operators," *Analysis Mathematica*, vol. 19, no. 1, pp. 85–100, 1993.

[28] C. de Boor, R. A. DeVore, and A. Ron, "Approximation from shift-invariant subspaces of $L_2(\mathbb{R}^d)$," *Transactions of the American Mathematical Society*, vol. 341, no. 2, pp. 787–806, 1994.