

A REGULARITY THEORY FOR STATIC SCHRÖDINGER EQUATIONS ON \mathbb{R}^d IN SPECTRAL BARRON SPACES*

ZIANG CHEN[†], JIANFENG LU[‡], YULONG LU[§], AND SHENGXUAN ZHOU[¶]

Abstract. Spectral Barron spaces have received considerable interest recently, as it is the natural function space for approximation theory of two-layer neural networks with a dimension-free convergence rate. In this paper, we study the regularity of solutions to the whole-space static Schrödinger equation in spectral Barron spaces. We prove that if the source of the equation lies in the spectral Barron space $\mathcal{B}^s(\mathbb{R}^d)$ and the potential function admitting a nonnegative lower bound decomposes as a positive constant plus a function in $\mathcal{B}^s(\mathbb{R}^d)$, then the solution lies in the spectral Barron space $\mathcal{B}^{s+2}(\mathbb{R}^d)$.

Key words. Schrödinger equation, Barron space, neural networks, regularity theory

MSC codes. 68T07, 65N99, 35Q40

DOI. 10.1137/22M1478719

1. Introduction. Numerical methods using neural networks for solving high-dimensional PDEs have achieved great success recently; see e.g., [4, 8, 10, 5]. A key advantage of these neural network-based algorithms is that neural networks can approximate functions in certain classes efficiently, meaning that the complexity grows at most polynomially in the dimension. By contrast, conventional methods suffer from the curse of dimensionality (CoD). For example, the complexity for approximating a d -dimensional function using piecewise constant function with error tolerance ϵ is $\mathcal{O}(\epsilon^{-d})$, which scales exponentially in d .

The efficiency of neural networks for approximating high-dimensional functions can even be observed for simple network structures, e.g., two-layer neural networks,

$$(1.1) \quad g_n(x) = \sum_{j=1}^n a_j \sigma(w_j^\top x + b_j) + c,$$

where σ is the activation function and n is the number of neurons. In the seminal work of Barron [2], he shows that for a function g satisfying

$$(1.2) \quad C_g := \int_{\mathbb{R}^d} |\hat{g}(\xi)| \cdot |\xi| d\xi < \infty,$$

* Received by the editors February 17, 2022; accepted for publication (in revised form) November 14, 2022; published electronically February 23, 2023.

<https://doi.org/10.1137/22M1478719>

Funding: The work of the first and second authors was supported in part by National Science Foundation grant DMS-2012286. The work of the third author was supported by National Science Foundation grant DMS-2107934.

[†] Department of Mathematics, Duke University, Durham, NC 27708 USA (ziang@math.duke.edu).

[‡] Departments of Mathematics, Physics, and Chemistry, Duke University, Durham, NC 27708 USA (jianfeng@math.duke.edu).

[§] Department of Mathematics and Statistics, University of Massachusetts, Amherst, MA 01003 USA (yulonglu@umass.edu).

[¶] Beijing International Center for Mathematical Research, Peking University, Beijing 100871, People's Republic of China (zhoushx19@pku.edu.cn).

it can be approximated by a two-layer neural network g_n of the form (1.1) in the L^2 -norm, where complexity n depends on the dimension d at most polynomially if C_g is viewed as a constant.

Along the direction of [2], several specific types of function classes on which the neural network approximation in various norms does not suffer from CoD have been defined and illustrated in the literature, including particularly the spectral Barron space [14, 18, 19], where the spectral Barron norms generalizing (1.2) are defined as some weighted L^1 -norm of \hat{u} or its discrete version, and the Barron space [6, 1], which generalizes (1.1) into an integral representation with respect to an underlying probability measure on the parameter space and defines the Barron norm as a path norm of the representation. With such function spaces specializing in high-dimensional problems, a natural question for the study of PDEs is this:

If the coefficients of a PDE lie in (spectral) Barron space, can the solution to the PDE also be guaranteed to be in (spectral) Barron space?

In this paper, we give a positive answer to the question above in the context of solving the d -dimensional Schrödinger equation in the whole space:

$$(1.3) \quad -\Delta u + Vu = f \quad \text{in } \mathbb{R}^d.$$

Here $V : \mathbb{R}^d \rightarrow \mathbb{R}$ is the potential energy and $f : \mathbb{R}^d \rightarrow \mathbb{R}$ is the source term. Morally speaking, our main result (Theorem 2.3) shows that if f is a spectral Barron function and V , with a nonnegative lower bound on \mathbb{R}^d , is the sum of a positive constant and a spectral Barron function, then the unique solution u^* to (1.3) is also in the spectral Barron space, with the order of spectral Barron regularity increased by 2. We remark that a function in the spectral Barron spaces we consider has Fourier transform in $L^1(\mathbb{R}^d)$, which implies that the function is bounded and decays to 0 at infinity (by the Riemann–Lebesgue lemma). Therefore, Assumption 2.2 on V implies that V is positive at infinity, which guarantees the uniqueness of the bounded solution and the spectral Barron solution; see Proposition 3.8 for a precise statement. An important consequence of Theorem 2.3 is that there exists a two-layer neural network that approximates u^* without CoD; see Theorem 2.4.

1.1. Related works. Regularity results of PDEs in Barron spaces have been studied in some recent works. In [14], a solution theory for the Poisson equation and the Schrödinger equation on the bounded domain $\Omega = [0, 1]^d$ with a homogeneous Neumann boundary condition is established in a type of spectral Barron space defined on Ω via cosine expansions. The same regularity result is later extended to the regularity estimate of the ground state of the Schrödinger operator in [13]. The work [7] proved regularity results for the screened Poisson equation $-\Delta u + \lambda^2 u = f$ and some time-dependent equations in the Barron space based on integral representation.

Besides the regularity estimates, another direction is to investigate the complexity of approximating PDE solutions using Barron functions or neural networks. Obtained in [3] (representational) are Barron complexity estimates for a general class of whole-space elliptic PDEs. A (deep) neural network complexity estimate for elliptic PDEs with a homogeneous Dirichlet boundary condition is established in [15].

1.2. Notations. We use $|x|$ for the Euclidean norm of a vector $x \in \mathbb{R}^d$ and use $B_r(x) = \{y \in \mathbb{R}^d : |y - x| < r\}$ for the open ℓ_2 ball in \mathbb{R}^d centered at x with radius r . For $i = 1, 2, \dots, d$, let $e_i \in \mathbb{R}^d$ be the vector with the i th entry being 1 and other entries being 0. For $g : \mathbb{R}^d \rightarrow \mathbb{R}$, we denote by \hat{g} its Fourier transform, given by

$$(1.4) \quad \hat{g}(\xi) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} g(x) e^{-ix^\top \xi} dx.$$

This is defined for $g \in L^1(\mathbb{R}^d)$ and can be extended to tempered distributions. Note that we included a multiplicative constant $(2\pi)^{-d}$ in the definition (1.4) for the purpose of getting a neater inverse Fourier transform:

$$(1.5) \quad g(x) = \int_{\mathbb{R}^d} \hat{g}(\xi) e^{ix^\top \xi} d\xi.$$

2. Main results. Recall that regularity estimates for elliptic PDEs in Sobolev spaces are classical: Suppose that $V_{\min} \leq V(x) \leq V_{\max} \forall x \in \mathbb{R}^d$ for some $V_{\min}, V_{\max} \in (0, +\infty)$ and that $f \in H^s(\mathbb{R}^d)$ with $s \geq -1$. Thanks to the Lax–Milgram theorem, the Schrödinger equation (1.3) admits a unique solution $u^* \in H^1(\mathbb{R}^d)$. One can also obtain higher regularity for u^* using a standard bootstrap argument. More specifically, if $u^* \in H^{s'}(\mathbb{R}^d)$ and V is sufficiently smooth, e.g., $V \in W^{s,\infty}(\mathbb{R}^d)$, then it holds that

$$(I - \Delta)u^* = f - u^* - Vu^* \in H^{\min\{s', s\}}(\mathbb{R}^d),$$

which implies that

$$u^* = (I - \Delta)^{-1}(f - u^* - Vu^*) \in H^{\min\{s', s\}+2}(\mathbb{R}^d).$$

Therefore, one can conclude $u^* \in H^{s+2}(\mathbb{R}^d)$ with

$$\|u^*\|_{H^{s+2}(\mathbb{R}^d)} \leq C \|f\|_{H^s(\mathbb{R}^d)},$$

where C is a constant depending on V , d , and s . Thus, u^* has higher regularity than the source term f . Similar regularity results have also been studied for elliptic PDEs on bounded domains; see e.g., [11, Theorem 5.27].

Our focus is to establish regularity results in the spectral Barron spaces for the Schrödinger equation (1.3). Let us first define the spectral Barron spaces as follows.

DEFINITION 2.1. *Given $s \in \mathbb{R}$, for a function $g : \mathbb{R}^d \rightarrow \mathbb{R}$, its spectral Barron norm is defined via*

$$\|g\|_{\mathcal{B}^s(\mathbb{R}^d)} := \int_{\mathbb{R}^d} |\hat{g}(\xi)| \cdot (1 + |\xi|^2)^{\frac{s}{2}} d\xi.$$

The spectral Barron space is the collection of functions with the finite spectral Barron norm:

$$\mathcal{B}^s(\mathbb{R}^d) = \left\{ g : \|g\|_{\mathcal{B}^s(\mathbb{R}^d)} < \infty \right\}.$$

The spectral Barron space with index $s = 1$ was first defined in the seminal work of Barron [2] and has been further developed with general index s in recent literature; see e.g., [14, 18, 19]. Spectral Barron spaces are of particular interest for high-dimensional problems since a spectral Barron function can be efficiently approximated by a two-layer neural network without CoD (see, e.g., [2]) for approximation in the L^2 -norm and (see, e.g., [18]) for approximation in the H^k -norm. A related but different notion of Barron space building upon integral representation was also proposed and studied in [6, 1]; see also [16, 17] for a characterization of such space via Radon transform.

Notice that by definition, the spectral Barron space $\mathcal{B}^s(\mathbb{R}^d)$ is a Banach space, and the completeness follows from the fact that the spectral Barron norm is a weighted L^1 -norm. A key difference between the spectral Barron space and the Sobolev space is that $\|g\|_{\mathcal{B}^s(\mathbb{R}^d)}$ is the L^1 -norm of $\hat{g}(\xi) \cdot (1 + |\xi|^2)^{s/2}$, while $\|g\|_{H^s(\mathbb{R}^d)}$ is the L^2 -norm

of $\hat{g}(\xi) \cdot (1 + |\xi|^2)^{s/2}$. In particular, $\mathcal{B}^s(\mathbb{R}^d)$ is not a Hilbert space. The lack of a Hilbert structure complicates the analysis of the existence and uniqueness of solutions in the spectral Barron space.

To state our regularity theory for PDEs in spectral Barron spaces, we make the following assumption on the potential V .

Assumption 2.2. Assume that the potential function V satisfies the following:

- (i) $V(x) \geq 0 \forall x \in \mathbb{R}^d$;
- (ii) $V = \alpha + W$, where $\alpha > 0$ is a constant and $W \in \mathcal{B}^s(\mathbb{R}^d)$.

Our main theorem can then be stated as follows.

THEOREM 2.3. *Suppose that Assumption 2.2 holds with $s \geq 0$. Then for any $f \in \mathcal{B}^s(\mathbb{R}^d)$, there exists a unique solution u^* in $\mathcal{B}^{s+2}(\mathbb{R}^d)$, and in addition it satisfies*

$$(2.1) \quad \|u^*\|_{\mathcal{B}^{s+2}(\mathbb{R}^d)} \leq C \|f\|_{\mathcal{B}^s(\mathbb{R}^d)},$$

where C is a constant depending on V , d , and s .

As a direct corollary of Theorem 2.3, the solution to the Schrödinger equation (1.3) can be approximated efficiently by a two-layer neural network on any bounded domain.

THEOREM 2.4. *Under the same assumptions of Theorem 2.3, let C be the constant in Theorem 2.3. Then for any $f \in \mathcal{B}^s(\mathbb{R}^d)$, any bounded domain $\Omega \subset \mathbb{R}^d$, and any $n \in \mathbb{N}_+$, there is a cosine-activated two-layer neural network with n hidden neurons,*

$$u_n(x) = \frac{1}{n} \sum_{j=1}^n a_j \cos(w_j^\top x + b_j),$$

satisfying

$$(2.2) \quad \|u_n - u^*\|_{H^1(\Omega)} \leq \frac{C \sqrt{m(\Omega)} \|f\|_{\mathcal{B}^s(\mathbb{R}^d)}}{n^{1/2}},$$

where u^* is the unique solution to (1.3) in $\mathcal{B}^{s+2}(\mathbb{R}^d)$ and $m(\Omega)$ is the Lebesgue measure of Ω .

Remark 2.5. The approximation in Theorem 2.4 is stated in the H^1 -norm. However, stronger results could be expected. Since $u^* \in \mathcal{B}^{s+2}(\mathbb{R}^d)$, it holds that $u^* \in H^{s+2}(\Omega)$, and approximation in the H^{s+2} -norm could hold. We refer the reader to [18] for details.

3. Proofs. This section is devoted to the proof of Theorem 2.3. Due to the lack of a Hilbert structure in the spectral Barron space $\mathcal{B}^s(\mathbb{R}^d)$, the standard Lax–Milgram theorem used to prove the well-posedness of elliptic equations in Sobolev spaces cannot be applied to the Barron spaces. Instead, we follow [14] and rewrite (1.3) as an integral equation of the second kind,

$$(3.1) \quad u + \mathcal{T}_{\alpha, W}(u) = (\alpha - \Delta)^{-1} f,$$

where

$$(3.2) \quad \mathcal{T}_{\alpha, W}(u) = (\alpha - \Delta)^{-1}(Wu).$$

Our approach is to apply the Fredholm alternative to the integral equation (3.1); thus, the existence of solution follows from its uniqueness. To this end, the essential

step is to prove the compactness of $\mathcal{T}_{\alpha,W}$. The compactness was established in [14] for PDEs on bounded domains, but it becomes more challenging for unbounded domains, e.g., the whole space in our setting. In fact, when $W = 1$, it is well known that $\mathcal{T}_{\alpha,1} = (\alpha - \Delta)^{-1}$ with $\alpha > 0$ is not compact on $L^2(\mathbb{R}^d)$ since it has a continuous spectrum. On the contrary, we shall show that $\mathcal{T}_{\alpha,W}$ is indeed compact on $\mathcal{B}^s(\mathbb{R}^d)$ provided that $\hat{W} \in L^1(\mathbb{R}^d)$, which is implied by Assumption 2.2 (ii) with $s \geq 0$.

3.1. Preliminary lemmas. We first present some preliminary lemmas for properties of the spectral Barron spaces and the operator $\mathcal{T}_{\alpha,W}$.

LEMMA 3.1. *The following embeddings hold:*

- (i) $\mathcal{B}^0(\mathbb{R}^d) \hookrightarrow L^\infty(\mathbb{R}^d)$ with $\|g\|_{L^\infty(\mathbb{R}^d)} \leq \|g\|_{\mathcal{B}^0(\mathbb{R}^d)}$;
- (ii) $\mathcal{B}^{s'}(\mathbb{R}^d) \hookrightarrow \mathcal{B}^s(\mathbb{R}^d)$ with $\|g\|_{\mathcal{B}^s(\mathbb{R}^d)} \leq \|g\|_{\mathcal{B}^{s'}(\mathbb{R}^d)}$ if $s \leq s'$.

Proof. They follow directly from Definition 2.1 and (1.5). □

LEMMA 3.2. *If $\alpha > 0$, then for any $g \in \mathcal{B}^s(\mathbb{R}^d)$, it holds that*

$$\|(\alpha - \Delta)^{-1}g\|_{\mathcal{B}^s(\mathbb{R}^d)} \leq \frac{1}{\alpha} \|g\|_{\mathcal{B}^s(\mathbb{R}^d)}$$

and that

$$(\alpha - \Delta)^{-1}g_{\mathcal{B}^{s+2}(\mathbb{R}^d)} \leq \frac{1}{\min\{\alpha, 1\}} \|g\|_{\mathcal{B}^s(\mathbb{R}^d)}.$$

Proof. Denote $h = (\alpha - \Delta)^{-1}g$. Then $\hat{h}(\xi) = \frac{1}{\alpha + |\xi|^2} \hat{g}(\xi)$. One can hence compute that

$$\|h\|_{\mathcal{B}^s(\mathbb{R}^d)} = \int_{\mathbb{R}^d} |\hat{h}(\xi)| \cdot (1 + |\xi|^2)^{\frac{s}{2}} d\xi \leq \frac{1}{\alpha} \int_{\mathbb{R}^d} |\hat{g}(\xi)| \cdot (1 + |\xi|^2)^{\frac{s}{2}} d\xi = \frac{1}{\alpha} \|g\|_{\mathcal{B}^s(\mathbb{R}^d)}$$

and that

$$\begin{aligned} \|h\|_{\mathcal{B}^{s+2}(\mathbb{R}^d)} &= \int_{\mathbb{R}^d} |\hat{h}(\xi)| \cdot (1 + |\xi|^2)^{\frac{s+2}{2}} d\xi = \int_{\mathbb{R}^d} |\hat{g}(\xi)| \cdot \frac{1 + |\xi|^2}{\alpha + |\xi|^2} \cdot (1 + |\xi|^2)^{\frac{s}{2}} d\xi \\ &\leq \frac{1}{\min\{\alpha, 1\}} \int_{\mathbb{R}^d} |\hat{g}(\xi)| \cdot (1 + |\xi|^2)^{\frac{s}{2}} d\xi = \frac{1}{\min\{\alpha, 1\}} \|g\|_{\mathcal{B}^s(\mathbb{R}^d)}. \end{aligned} \quad \square$$

Remark 3.3. Since $g \in \mathcal{B}^s(\mathbb{R}^d)$ is a real-valued function, $(\alpha - \Delta)^{-1}g$ with $\alpha > 0$ must also be real-valued. This is because $-\Delta u + \alpha u = 0$ only has a trivial solution in the space of tempered distributions, which can be seen directly by taking Fourier transform $\hat{u} = \frac{0}{\alpha + |\xi|^2} = 0$.

LEMMA 3.4. *Suppose that $W \in \mathcal{B}^s(\mathbb{R}^d)$ with $s \geq 0$. Then for any $u \in \mathcal{B}^s(\mathbb{R}^d)$, it holds that*

$$\|Wu\|_{\mathcal{B}^s(\mathbb{R}^d)} \leq 2^{\frac{s}{2}} \|W\|_{\mathcal{B}^s(\mathbb{R}^d)} \|u\|_{\mathcal{B}^s(\mathbb{R}^d)}.$$

Proof. It follows from

$$\widehat{Wu}(\eta) = \hat{W} * \hat{u}(\eta) = \int_{\mathbb{R}^d} \hat{W}(\xi) \hat{u}(\eta - \xi) d\xi$$

that

$$\begin{aligned}
 \|Wu\|_{\mathcal{B}^s} &\leq \int_{\mathbb{R}^d \times \mathbb{R}^d} |\widehat{W}(\xi)| \cdot |\widehat{u}(\eta - \xi)| \cdot (1 + |\eta|^2)^{\frac{s}{2}} d\xi d\eta \\
 &\leq \int_{\mathbb{R}^d \times \mathbb{R}^d} |\widehat{W}(\xi)| \cdot |\widehat{u}(\eta - \xi)| \cdot (1 + 2|\xi|^2 + 2|\eta - \xi|^2)^{\frac{s}{2}} d\xi d\eta \\
 &\leq 2^{\frac{s}{2}} \int_{\mathbb{R}^d \times \mathbb{R}^d} |\widehat{W}(\xi)| \cdot |\widehat{u}(\eta - \xi)| \cdot (1 + |\xi|^2)^{\frac{s}{2}} \cdot (1 + |\eta - \xi|^2)^{\frac{s}{2}} d\xi d\eta \\
 &= 2^{\frac{s}{2}} \int_{\mathbb{R}^d} |\widehat{W}(\xi)| \cdot (1 + |\xi|^2)^{\frac{s}{2}} d\xi \int_{\mathbb{R}^d} |\widehat{u}(\xi)| \cdot (1 + |\xi|^2)^{\frac{s}{2}} d\xi \\
 &= 2^{\frac{s}{2}} \|W\|_{\mathcal{B}^s} \|u\|_{\mathcal{B}^s}. \quad \square
 \end{aligned}$$

3.2. Compactness of $\mathcal{T}_{\alpha,W}$. Lemmas 3.2 and 3.4 imply that the operator $\mathcal{T}_{\alpha,W}$ defined in (3.2) is bounded in $\mathcal{B}^s(\mathbb{R}^d)$ if $\alpha > 0$ and $W \in \mathcal{B}^s(\mathbb{R}^d)$ with $s \geq 0$. We now show that this operator is compact with a more careful analysis.

PROPOSITION 3.5. *Suppose that $\alpha > 0$ and $W \in \mathcal{B}^s(\mathbb{R}^d)$ with $s \geq 0$. Then the operator $\mathcal{T}_{\alpha,W} : \mathcal{B}^s(\mathbb{R}^d) \rightarrow \mathcal{B}^s(\mathbb{R}^d)$ defined in (3.2) is compact.*

To prove that $\mathcal{T}_{\alpha,W} : \mathcal{B}^s(\mathbb{R}^d) \rightarrow \mathcal{B}^s(\mathbb{R}^d)$ is compact, it suffices to show that the image of the closed unit ball in $\mathcal{B}^s(\mathbb{R}^d)$,

$$(3.3) \quad \left\{ \mathcal{T}_{\alpha,W}(u) : \|u\|_{\mathcal{B}^s(\mathbb{R}^d)} \leq 1 \right\},$$

is relatively compact in $\mathcal{B}^s(\mathbb{R}^d)$. Notice that $\mathcal{B}^s(\mathbb{R}^d)$ is complete, which implies that a subset of $\mathcal{B}^s(\mathbb{R}^d)$ is relatively compact if and only if it is totally bounded. Therefore, it suffices to prove the total boundedness of

$$(3.4) \quad \mathcal{F} := \left\{ \widehat{\mathcal{T}_{\alpha,W}(u)}(\xi) \cdot (1 + |\xi|^2)^{\frac{s}{2}} : \|u\|_{\mathcal{B}^s(\mathbb{R}^d)} \leq 1 \right\} \subset L^1(\mathbb{R}^d),$$

where we translate the \mathcal{B}^s -norm into the usual L^1 -norm. The following Kolmogorov–Riesz theorem will be useful for establishing the total boundedness.

THEOREM 3.6 (Kolmogorov–Riesz theorem [12, Theorem 5]). *For $p \in [1, \infty)$, a subset $\mathcal{F} \subset L^p(\mathbb{R}^d)$ is totally bounded if and only if the following three conditions hold:*

- (i) \mathcal{F} is bounded;
- (ii) for any $\epsilon \rightarrow 0$, there exists $R \rightarrow 0$ such that

$$\int_{|x| \rightarrow R} |f(x)|^p dx < \epsilon^p \quad \forall f \in \mathcal{F};$$

- (iii) for any $\epsilon \rightarrow 0$, there exists $\delta \rightarrow 0$ such that

$$\int_{\mathbb{R}^d} |f(x+y) - f(x)|^p dx < \epsilon^p \quad \forall f \in \mathcal{F}, |y| < \delta.$$

We then prove Proposition 3.5 using Theorem 3.6.

Proof of Proposition 3.5. As discussed above, the compactness of $\mathcal{T}_{\alpha,W}$ follows from the total boundedness of \mathcal{F} defined in (3.4). Therefore, it suffices to verify the three conditions in Theorem 3.6 for \mathcal{F} . We verify them one by one.

- *Verification of condition (i) in Theorem 3.6:* For any $\|u\|_{\mathcal{B}^s(\mathbb{R}^d)} \leq 1$, since

$$\widehat{\mathcal{T}_{\alpha,W}(u)}(\xi) = \frac{1}{\alpha + |\xi|^2} \widehat{Wu}(\xi),$$

we have that

$$\begin{aligned} \left\| \widehat{\mathcal{T}_{\alpha,W}(u)}(\xi) \cdot (1 + |\xi|^2)^{\frac{s}{2}} \right\|_{L^1(\mathbb{R}^d)} &\leq \|(\alpha - \Delta)^{-1}Wu\|_{\mathcal{B}^s(\mathbb{R}^d)} \\ &\leq \frac{1}{\alpha} \|Wu\|_{\mathcal{B}^s(\mathbb{R}^d)} \\ &\leq \frac{2^{\frac{s}{2}}}{\alpha} \|W\|_{\mathcal{B}^s(\mathbb{R}^d)} \|u\|_{\mathcal{B}^s(\mathbb{R}^d)} \\ &\leq \frac{2^{\frac{s}{2}}}{\alpha} \|W\|_{\mathcal{B}^s(\mathbb{R}^d)}, \end{aligned}$$

where we used Lemmas 3.2 and 3.4. Therefore, \mathcal{F} is bounded in $L^1(\mathbb{R}^d)$.

- *Verification of condition (ii) in Theorem 3.6:* For any $\epsilon \rightarrow 0$, there exists $R \rightarrow 0$ such that $\frac{1}{\alpha + |\xi|^2} < \epsilon$ for any $|\xi| \rightarrow R$. Then for any $\|u\|_{\mathcal{B}^s(\mathbb{R}^d)} \leq 1$, it holds that

$$\begin{aligned} \int_{|\xi| \rightarrow R} |\widehat{\mathcal{T}_{\alpha,W}(u)}(\xi)| \cdot (1 + |\xi|^2)^{\frac{s}{2}} d\xi &\leq \int_{|\xi| \rightarrow R} \frac{1}{\alpha + |\xi|^2} \cdot |\widehat{Wu}(\xi)| \cdot (1 + |\xi|^2)^{\frac{s}{2}} d\xi \\ &\leq \epsilon \int_{|\xi| \rightarrow R} |\widehat{Wu}(\xi)| \cdot (1 + |\xi|^2)^{\frac{s}{2}} d\xi \\ &\leq \epsilon \|Wu\|_{\mathcal{B}^s(\mathbb{R}^d)} \\ &\leq \epsilon 2^{\frac{s}{2}} \|W\|_{\mathcal{B}^s(\mathbb{R}^d)} \|u\|_{\mathcal{B}^s(\mathbb{R}^d)} \\ &\leq \epsilon 2^{\frac{s}{2}} \|W\|_{\mathcal{B}^s(\mathbb{R}^d)}, \end{aligned}$$

where we also used Lemma 3.4.

- *Verification of condition (iii) in Theorem 3.6:* Since condition (ii) in Theorem 3.6 holds, for any $\epsilon \rightarrow 0$, there exists $R \rightarrow 0$ such that

$$(3.5) \quad \int_{|\xi| \rightarrow R} |\widehat{\mathcal{T}_{\alpha,W}(u)}(\xi)| \cdot (1 + |\xi|^2)^{\frac{s}{2}} d\xi \leq \epsilon \quad \forall \|u\|_{\mathcal{B}^s(\mathbb{R}^d)} \leq 1.$$

Set

$$(3.6) \quad L_1 := \max_{|\xi| \leq 2R} \frac{(1 + |\xi|^2)^{\frac{s}{2}}}{\alpha + |\xi|^2} \quad \text{and} \quad L_2 := \int_{|\xi| \leq 2R} d\xi.$$

It follows from $W \in \mathcal{B}^s(\mathbb{R}^d)$ with $s \geq 0$ that $\hat{W} \in L^1(\mathbb{R}^d)$. According to [9, Proposition 8.17], there exists $\varphi \in C_c^\infty(\mathbb{R}^d)$ satisfying

$$(3.7) \quad \left\| \hat{W} - \varphi \right\|_{L^1(\mathbb{R}^d)} \leq \frac{\epsilon}{L_1}.$$

Note that $\xi \mapsto \frac{(1 + |\xi|^2)^{\frac{s}{2}}}{\alpha + |\xi|^2}$ is continuous and is hence uniformly continuous on any compact subsets of \mathbb{R}^d . One also has that $\varphi \in C_c^\infty(\mathbb{R}^d)$ is uniformly continuous on \mathbb{R}^d . Thus, there exists some $\delta < R$, such that

$$(3.8) \quad \left| \frac{(1 + |\xi|^2)^{\frac{s}{2}}}{\alpha + |\xi|^2} - \frac{(1 + |\xi'|^2)^{s/2}}{\alpha + |\xi'|^2} \right| \leq \epsilon \quad \forall |\xi| \leq 3R, |\xi'| \leq 3R, |\xi - \xi'| < \delta$$

and that

$$(3.9) \quad |\varphi(\xi) - \varphi(\xi')| < \frac{\epsilon}{L_1 L_2} \quad \forall \xi, \xi' \in \mathbb{R}^d, \quad |\xi - \xi'| < \delta.$$

Consider any $|y| < \delta < R$ and any $\|u\|_{\mathcal{B}^s(\mathbb{R}^d)} \leq 1$. We have that

$$(3.10) \quad \begin{aligned} & \int_{\mathbb{R}^d} \left| \widehat{\mathcal{T}_{\alpha, W}(u)}(\xi + y) \cdot (1 + |\xi + y|^2)^{\frac{\alpha}{2}} - \widehat{\mathcal{T}_{\alpha, W}(u)}(\xi) \cdot (1 + |\xi|^2)^{\frac{\alpha}{2}} \right| d\xi \\ & \leq \int_{|\xi| \rightarrow 2R} |\widehat{\mathcal{T}_{\alpha, W}(u)}(\xi + y)| \cdot (1 + |\xi + y|^2)^{\frac{\alpha}{2}} d\xi \\ & \quad + \int_{|\xi| \rightarrow 2R} |\widehat{\mathcal{T}_{\alpha, W}(u)}(\xi)| \cdot (1 + |\xi|^2)^{\frac{\alpha}{2}} d\xi \\ & \quad + \int_{|\xi| \leq 2R} \left| \widehat{\mathcal{T}_{\alpha, W}(u)}(\xi + y) \cdot (1 + |\xi + y|^2)^{\frac{\alpha}{2}} - \widehat{\mathcal{T}_{\alpha, W}(u)}(\xi) \cdot (1 + |\xi|^2)^{\frac{\alpha}{2}} \right| d\xi \\ & \leq 2 \int_{|\xi| \rightarrow R} |\widehat{\mathcal{T}_{\alpha, W}(u)}(\xi)| \cdot (1 + |\xi|^2)^{\frac{\alpha}{2}} d\xi \\ & \quad + \int_{|\xi| \leq 2R} \left| \widehat{\mathcal{T}_{\alpha, W}(u)}(\xi + y) \cdot (1 + |\xi + y|^2)^{\frac{\alpha}{2}} - \widehat{\mathcal{T}_{\alpha, W}(u)}(\xi) \cdot (1 + |\xi|^2)^{\frac{\alpha}{2}} \right| d\xi \\ & \leq 2\epsilon + \int_{|\xi| \leq 2R} \left| \widehat{\mathcal{T}_{\alpha, W}(u)}(\xi + y) \cdot (1 + |\xi + y|^2)^{\frac{\alpha}{2}} - \widehat{\mathcal{T}_{\alpha, W}(u)}(\xi) \cdot (1 + |\xi|^2)^{\frac{\alpha}{2}} \right| d\xi, \end{aligned}$$

where the last inequality follows from (3.5). Then we estimate the second term in the last line above. We have that

$$(3.11) \quad \begin{aligned} & \left| \widehat{\mathcal{T}_{\alpha, W}(u)}(\xi + y) \cdot (1 + |\xi + y|^2)^{\frac{\alpha}{2}} - \widehat{\mathcal{T}_{\alpha, W}(u)}(\xi) \cdot (1 + |\xi|^2)^{\frac{\alpha}{2}} \right| \\ & = \left| \widehat{Wu}(\xi + y) \cdot \frac{(1 + |\xi + y|^2)^{\frac{\alpha}{2}}}{\alpha + |\xi + y|^2} - \widehat{Wu}(\xi) \cdot \frac{(1 + |\xi|^2)^{\frac{\alpha}{2}}}{\alpha + |\xi|^2} \right| \\ & \leq |\widehat{Wu}(\xi + y)| \cdot \left| \frac{(1 + |\xi + y|^2)^{\frac{\alpha}{2}}}{\alpha + |\xi + y|^2} - \frac{(1 + |\xi|^2)^{\frac{\alpha}{2}}}{\alpha + |\xi|^2} \right| + \frac{(1 + |\xi|^2)^{\frac{\alpha}{2}}}{\alpha + |\xi|^2} \cdot |\widehat{Wu}(\xi + y) - \widehat{Wu}(\xi)|. \end{aligned}$$

According to (3.8) and Lemma 3.4, it holds that

$$(3.12) \quad \begin{aligned} & \int_{|\xi| \leq 2R} |\widehat{Wu}(\xi + y)| \cdot \left| \frac{(1 + |\xi + y|^2)^{\frac{\alpha}{2}}}{\alpha + |\xi + y|^2} - \frac{(1 + |\xi|^2)^{\frac{\alpha}{2}}}{\alpha + |\xi|^2} \right| d\xi \\ & \leq \epsilon \int_{|\xi| \leq 2R} |\widehat{Wu}(\xi + y)| d\xi \leq \epsilon \|Wu\|_{\mathcal{B}^0(\mathbb{R}^d)} \leq \epsilon \|W\|_{\mathcal{B}^0(\mathbb{R}^d)} \|u\|_{\mathcal{B}^0(\mathbb{R}^d)} \\ & \leq \epsilon \|W\|_{\mathcal{B}^s(\mathbb{R}^d)} \|u\|_{\mathcal{B}^s(\mathbb{R}^d)} \leq \epsilon \|W\|_{\mathcal{B}^s(\mathbb{R}^d)}. \end{aligned}$$

By (3.6), (3.7), and (3.9), it holds that

$$\begin{aligned}
 & \int_{|\xi| \leq 2R} \frac{(1 + |\xi|^2)^{\frac{s}{2}}}{\alpha + |\xi|^2} \cdot |\widehat{W}u(\xi + y) - \widehat{W}u(\xi)| d\xi \\
 &= \int_{|\xi| \leq 2R} \frac{(1 + |\xi|^2)^{\frac{s}{2}}}{\alpha + |\xi|^2} \cdot \left| \int_{\mathbb{R}^d} \hat{u}(\eta) \hat{W}(\xi + y - \eta) d\eta - \int_{\mathbb{R}^d} \hat{u}(\eta) \hat{W}(\xi - \eta) d\eta \right| d\xi \\
 &\leq L_1 \int_{|\xi| \leq 2R} \int_{\mathbb{R}^d} |\hat{u}(\eta)| \cdot |\hat{W}(\xi + y - \eta) - \hat{W}(\xi - \eta)| d\eta d\xi \\
 (3.13) \quad &\leq L_1 \int_{|\xi| \leq 2R} \int_{\mathbb{R}^d} |\hat{u}(\eta)| \cdot |\varphi(\xi + y - \eta) - \varphi(\xi - \eta)| d\eta d\xi \\
 &\quad + L_1 \int_{|\xi| \leq 2R} \int_{\mathbb{R}^d} |\hat{u}(\eta)| \cdot |\hat{W}(\xi + y - \eta) - \varphi(\xi + y - \eta)| d\eta d\xi \\
 &\quad + L_1 \int_{|\xi| \leq 2R} \int_{\mathbb{R}^d} |\hat{u}(\eta)| \cdot |\hat{W}(\xi - \eta) - \varphi(\xi - \eta)| d\eta d\xi \\
 &\leq \frac{\epsilon}{L_2} \int_{|\xi| \leq 2R} d\xi \int_{\mathbb{R}^d} |\hat{u}(\eta)| d\eta + 2L_1 \left\| \hat{W} - \varphi \right\|_{L^1(\mathbb{R}^d)} \int_{\mathbb{R}^d} |\hat{u}(\eta)| d\eta \\
 &\leq 3\epsilon \int_{\mathbb{R}^d} |\hat{u}(\eta)| d\eta \leq 3\epsilon \|u\|_{\mathcal{B}^s(\mathbb{R}^d)} \leq 3\epsilon.
 \end{aligned}$$

Combining (3.11), (3.12), and (3.13), we obtain that

$$\begin{aligned}
 & \int_{|\xi| \leq 2R} \left| \widehat{\mathcal{T}_{\alpha, W} u}(\xi + y) \cdot (1 + |\xi + y|^2)^{\frac{s}{2}} - \widehat{\mathcal{T}_{\alpha, W} u}(\xi) \cdot (1 + |\xi|^2)^{\frac{s}{2}} \right| d\xi \\
 &\leq (3 + \|W\|_{\mathcal{B}^s(\mathbb{R}^d)}) \cdot \epsilon,
 \end{aligned}$$

which, combined with (3.10), yields that

$$\begin{aligned}
 & \int_{\mathbb{R}^d} \left| \widehat{\mathcal{T}_{\alpha, W} u}(\xi + y) \cdot (1 + |\xi + y|^2)^{\frac{s}{2}} - \widehat{\mathcal{T}_{\alpha, W} u}(\xi) \cdot (1 + |\xi|^2)^{\frac{s}{2}} \right| d\xi \\
 &\leq (5 + \|W\|_{\mathcal{B}^s(\mathbb{R}^d)}) \cdot \epsilon
 \end{aligned}$$

for any $|y| < \delta$ and $\|u\|_{\mathcal{B}^s(\mathbb{R}^d)} \leq 1$. This completes the proof. \square

3.3. Proof of the main results. We finish the proof of Theorem 2.3 in this subsection. We first need to establish the existence of the solution to (3.1) that is equivalent to the original PDE (1.3) in $\mathcal{B}^s(\mathbb{R}^d)$.

PROPOSITION 3.7. *Suppose that Assumption 2.2 holds with $s \geq 0$. Then the operator*

$$(I + \mathcal{T}_{\alpha, W})^{-1} : \mathcal{B}^s(\mathbb{R}^d) \rightarrow \mathcal{B}^s(\mathbb{R}^d)$$

is bounded.

Since $\mathcal{T}_{\alpha, W} : \mathcal{B}^s(\mathbb{R}^d) \rightarrow \mathcal{B}^s(\mathbb{R}^d)$ has been proved as compact in Proposition 3.5, $I + \mathcal{T}_{\alpha, W}$ is a Fredholm operator. Therefore, to show that $I + \mathcal{T}_{\alpha, W}$ has a bounded inverse, it suffices to show that $I + \mathcal{T}_{\alpha, W}$ is injective, which is established in the following proposition.

PROPOSITION 3.8. *Suppose that Assumption 2.2 holds with $s \geq 0$. Then the operator*

$$I + \mathcal{T}_{\alpha, W} : \mathcal{B}^s(\mathbb{R}^d) \rightarrow \mathcal{B}^s(\mathbb{R}^d)$$

is injective.

Proof. Suppose that there exists some $u \in \mathcal{B}^s(\mathbb{R}^d)$ such that

$$u + \mathcal{T}_{\alpha, W}(u) = 0,$$

which is equivalent to

$$-\Delta u + Vu = 0,$$

where $V = \alpha + W$. Since $s \geq 0$, we have that $u \in L^\infty(\mathbb{R}^d)$ by Lemma 3.1. Furthermore, u and V are both continuous, as the Fourier transform of a function in $\mathcal{B}^s(\mathbb{R}^d)$ is in $L^1(\mathbb{R}^d)$.

Suppose that u is not identically zero, which means that $u(x_0) \neq 0$ holds for some $x_0 \in \mathbb{R}^d$. It follows from $\hat{W} \in L^1(\mathbb{R}^d)$ and the Riemann–Lebesgue lemma that $\lim_{|x| \rightarrow \infty} W(x) = 0$, which implies that there exists some $R \geq |x_0|$ such that

$$\inf_{|x| \geq R} V(x) \geq \frac{\alpha}{2}.$$

Note that Assumption 2.2 (i) states that $V(x) \geq 0 \forall x \in \mathbb{R}^d$. According to the weak maximal principle, we have for $r \rightarrow 0$ that

$$(3.14) \quad \sup_{|x| \leq r} |u(x)| = \sup_{|x|=r} |u(x)|.$$

By (3.14), there is a sequence $\{x_k\}_{k=1}^\infty \subset \mathbb{R}^d$ with

$$(3.15) \quad |x_k| = R + k \quad \text{and} \quad |u(x_k)| \geq |u(x_0)|.$$

Let us set

$$\psi_k(r) = \int_{\partial B_1(0)} u^2(rx + x_k) dS = r^{-(d-1)} \int_{\partial B_r(x_k)} u^2 dS \geq 0.$$

Then it holds that

$$\begin{aligned} \psi_k'(r) &= \int_{\partial B_1(0)} \frac{\partial}{\partial r} u^2(rx + x_k) dS = \int_{\partial B_1(0)} \langle \nabla u^2(rx + x_k), x \rangle dS \\ &= r^{-(d-1)} \int_{\partial B_r(0)} \left\langle \nabla u^2(x + x_k), \frac{x}{r} \right\rangle dS = r^{-(d-1)} \int_{B_r(x_k)} \Delta(u^2) dx \\ &= 2r^{-(d-1)} \int_{B_r(x_k)} (u \Delta u + |\nabla u|^2) dx \geq 2r^{-(d-1)} \int_{B_r(x_k)} u \Delta u dx \\ &= 2r^{-(d-1)} \int_{B_r(x_k)} Vu^2 dx. \end{aligned}$$

Note that $V(x) \geq \frac{\alpha}{2}$ holds on $B_k(x_k)$. We have for any $r \leq k$ that

$$\begin{aligned} \psi_k'(r) &\geq 2r^{-(d-1)} \int_{B_r(x_k)} Vu^2 dx \geq \alpha r^{-(d-1)} \int_{B_r(x_k)} u^2 dx \\ &= \alpha r^{-(d-1)} \int_0^r \int_{\partial B_t(x_k)} u^2 dS dt = \alpha \int_0^r \left(\frac{t}{r}\right)^{d-1} \psi_k(t) dt \geq 0, \end{aligned}$$

which implies that ψ_k is monotonically increasing on $[0, k]$. For any $n \in \{1, 2, \dots, k-1\}$ and any $r \in [n, n+1]$, we have that

$$\begin{aligned} \psi_k'(r) &\geq \alpha \int_0^r \left(\frac{t}{r}\right)^{d-1} \psi_k(t) dt \geq \alpha \int_n^r \left(\frac{t}{r}\right)^{d-1} \psi_k(t) dt \\ &\geq \alpha(r-n) \cdot \left(\frac{n}{n+1}\right)^{d-1} \psi_k(n) \end{aligned}$$

and hence that

$$\begin{aligned} \psi_k(n+1) &= \psi_k(n) + \int_n^{n+1} \psi_k'(r) dr \\ &\geq \left(1 + \alpha \int_n^{n+1} (r-n) dr \cdot \left(\frac{n}{n+1}\right)^{d-1}\right) \psi_k(n) \\ &\geq \left(1 + \frac{\alpha}{2^d}\right) \cdot \psi_k(n). \end{aligned}$$

Thus, it holds that

$$\begin{aligned} (3.16) \quad \psi_k(k) &\geq \left(1 + \frac{\alpha}{2^d}\right)^{k-1} \cdot \psi_k(1) \geq \left(1 + \frac{\alpha}{2^d}\right)^{k-1} \cdot \psi_k(0) \\ &= \left(1 + \frac{\alpha}{2^d}\right)^{k-1} \cdot u(x_k)^2 \geq \left(1 + \frac{\alpha}{2^d}\right)^{k-1} \cdot u(x_0)^2 \quad \forall k \in \mathbb{N}_+, \end{aligned}$$

where we used the monotone property of ψ_k on $[0, k]$ and (3.15).

Note that $u \in L^\infty(\mathbb{R}^d)$. So $\{\psi(k)\}_{k=1}^\infty$ must be bounded, which contradicts (3.16), as $u(x_0) \neq 0$. We therefore can conclude that $u = 0$, which proves the injectivity of $I + \mathcal{T}_{\alpha, W}$. \square

Remark 3.9. We remark that the standard proof of the uniqueness of H^1 -solutions to elliptic PDEs does not apply to the Barron solutions. In fact, the uniqueness in $H^1(\mathbb{R}^d)$ of solutions of the equation $-\Delta u + Vu = 0$ follows from a standard energy estimate. Noticing that $-\Delta u + Vu = 0 \in H^{-1}(\mathbb{R}^d)$, which is the dual space of $H^1(\mathbb{R}^d)$, one has that $0 = \int_{\mathbb{R}^d} (-\Delta u + Vu)u = \int_{\mathbb{R}^d} |\nabla u|^2 + Vu^2$, which implies $u = 0$. However, such an energy estimate in general does not apply to Barron functions in $\mathcal{B}^s(\mathbb{R}^d)$ since in general $\mathcal{B}^s(\mathbb{R}^d) \not\subseteq H^1(\mathbb{R}^d)$. To give a concrete example, let us consider the function u , whose Fourier transform is defined by

$$\hat{u}(\xi_1, \dots, \xi_d) = \begin{cases} |\xi_1|^{-\frac{1}{2}} & \text{if } \xi \in ([-1, 1] \setminus \{0\}) \times [-1, 1]^{d-1}, \\ 0 & \text{otherwise.} \end{cases}$$

Then $\int_{\mathbb{R}^d} |\hat{u}(\xi)| \cdot (1 + |\xi|^2)^{s/2} d\xi < \infty$, while $\int_{\mathbb{R}^d} |\hat{u}(\xi)|^2 \cdot (1 + |\xi|^2) d\xi = \infty$, i.e., $u \in \mathcal{B}^s(\mathbb{R}^d) \setminus H^1(\mathbb{R}^d)$.

Proposition 3.7 is then a direct corollary.

Proof of Proposition 3.7. The result follows directly from Propositions 3.5 and 3.8 and the Fredholm alternative. \square

We can finally prove Theorem 2.3.

Proof of Theorem 2.3. Let u^* be the unique solution to (1.3) or (3.1). Notice that by Proposition 3.7 and Lemma 3.2,

$$u^* = (I + \mathcal{T}_{\alpha, W})^{-1} ((\alpha - \Delta)^{-1} f) \in \mathcal{B}^s(\mathbb{R}^d),$$

with

$$(3.17) \quad \begin{aligned} \|u^*\|_{\mathcal{B}^s(\mathbb{R}^d)} &= \|(I + \mathcal{T}_{\alpha, W})^{-1} ((\alpha - \Delta)^{-1} f)\|_{\mathcal{B}^s(\mathbb{R}^d)} \\ &\leq \frac{1}{\alpha} \|(I + \mathcal{T}_{\alpha, W})^{-1}\|_{\mathcal{B}^s(\mathbb{R}^d) \rightarrow \mathcal{B}^s(\mathbb{R}^d)} \|f\|_{\mathcal{B}^s(\mathbb{R}^d)}. \end{aligned}$$

It follows from Lemmas 3.2 and 3.4 and (1.3) that

$$(3.18) \quad \begin{aligned} \|u^*\|_{\mathcal{B}^{s+2}(\mathbb{R}^d)} &\leq \frac{1}{\min\{\alpha, 1\}} \|(\alpha - \Delta)u^*\|_{\mathcal{B}^s(\mathbb{R}^d)} \\ &= \frac{1}{\min\{\alpha, 1\}} \|Wu^* - f\|_{\mathcal{B}^s(\mathbb{R}^d)} \\ &\leq \frac{1}{\min\{\alpha, 1\}} (\|Wu^*\|_{\mathcal{B}^s(\mathbb{R}^d)} + \|f\|_{\mathcal{B}^s(\mathbb{R}^d)}) \\ &\leq \frac{1}{\min\{\alpha, 1\}} \left(2^{\frac{s}{2}} \|W\|_{\mathcal{B}^s(\mathbb{R}^d)} \|u^*\|_{\mathcal{B}^s(\mathbb{R}^d)} + \|f\|_{\mathcal{B}^s(\mathbb{R}^d)} \right). \end{aligned}$$

Combining (3.17) and (3.18), we obtain that

$$\|u^*\|_{\mathcal{B}^{s+2}(\mathbb{R}^d)} \leq \frac{1}{\min\{\alpha, 1\}} \left(\frac{2^{\frac{s}{2}} \|W\|_{\mathcal{B}^s(\mathbb{R}^d)}}{\alpha} \|(I + \mathcal{T}_{\alpha, W})^{-1}\|_{\mathcal{B}^s(\mathbb{R}^d) \rightarrow \mathcal{B}^s(\mathbb{R}^d)} + 1 \right) \|f\|_{\mathcal{B}^s}.$$

Hence, (2.1) holds with

$$C = \frac{1}{\min\{\alpha, 1\}} \left(\frac{2^{\frac{s}{2}} \|W\|_{\mathcal{B}^s(\mathbb{R}^d)}}{\alpha} \|(I + \mathcal{T}_{\alpha, W})^{-1}\|_{\mathcal{B}^s(\mathbb{R}^d) \rightarrow \mathcal{B}^s(\mathbb{R}^d)} + 1 \right),$$

which completes the proof. \square

Theorem 2.4 then follows directly from Theorem 2.3 and some techniques for establishing approximation without CoD in previous literature.

Proof of Theorem 2.4. This proof uses techniques from [2, 6] and is similar to [3, Theorem 2.5]. Note that $s \geq 0$. According to Theorem 2.3 and Lemma 3.1, it holds that

$$\|u^*\|_{\mathcal{B}^0(\mathbb{R}^d)} \leq \|u^*\|_{\mathcal{B}^2(\mathbb{R}^d)} \leq \|u^*\|_{\mathcal{B}^{s+2}(\mathbb{R}^d)} \leq C \|f\|_{\mathcal{B}^s(\mathbb{R}^d)}.$$

Denote $\hat{u}^*(\xi) = |\hat{u}^*(\xi)|e^{i\theta(\xi)}$, and let μ be a probability distribution on \mathbb{R}^d with density being $|\hat{u}^*(\xi)|/|\hat{u}^*|_{L^1(\mathbb{R}^d)} = |\hat{u}^*(\xi)|/\|u^*\|_{\mathcal{B}^0(\mathbb{R}^d)}$. Then the real-valued function u^* can be represented as

$$\begin{aligned} u^*(x) &= \int_{\mathbb{R}^d} \hat{u}^*(\xi) e^{i\xi^T x} d\xi = \int_{\mathbb{R}^d} |\hat{u}^*(\xi)| e^{i(\xi^T x + \theta(\xi))} d\xi \\ &= \int_{\mathbb{R}^d} |\hat{u}^*(\xi)| \cos(\xi^T x + \theta(\xi)) d\xi = \|u^*\|_{\mathcal{B}^0(\mathbb{R}^d)} \mathbb{E}_{\xi \sim \mu} [\cos(\xi^T x + \theta(\xi))]. \end{aligned}$$

Note that $\hat{u}^* \in \mathcal{B}^2(\mathbb{R}^d)$, which implies that μ has finite first-order and second-order moments. Therefore,

$$\frac{\partial}{\partial x_k} u^*(x) = -\|u^*\|_{\mathcal{B}^0(\mathbb{R}^d)} \mathbb{E}_{\xi \sim \mu} [\langle \xi, e_k \rangle \sin(\xi^T x + \theta(\xi))].$$

Let $\xi_1, \xi_2, \dots, \xi_n$ be independently and identically distributed samples from μ , and let

$$u_n(x) = \frac{1}{n} \sum_{j=1}^n a_j \cos(w_j^\top x + b_j),$$

where $a_j = \|u^*\|_{\mathcal{B}^0(\mathbb{R}^d)}$, $w_j = \xi_j$, and $b_j = \theta(\xi_j)$. Then it holds that

$$\begin{aligned} & \mathbb{E}_{\mu^{\otimes n}} \|u^* - u_n\|_{H^1(\Omega)}^2 \\ &= \mathbb{E}_{\mu^{\otimes n}} \left[\int_{\Omega} |u^*(x) - u_n(x)|^2 dx + \sum_{k=1}^d \int_{\Omega} \left| \frac{\partial}{\partial x_k} u^*(x) - \frac{\partial}{\partial x_k} u_n(x) \right|^2 dx \right] \\ &= \|u^*\|_{\mathcal{B}^0(\mathbb{R}^d)}^2 \int_{\Omega} \text{Var}_{\mu^{\otimes n}} \left[\frac{1}{n} \sum_{j=1}^n \cos(\xi_j^\top x + \theta(\xi_j)) \right] dx \\ &\quad + \|u^*\|_{\mathcal{B}^0(\mathbb{R}^d)}^2 \sum_{k=1}^d \int_{\Omega} \text{Var}_{\mu^{\otimes n}} \left[\frac{1}{n} \sum_{j=1}^n \langle \xi_j, e_k \rangle \sin(\xi_j^\top x + \theta(\xi_j)) \right] dx \\ &= \frac{\|u^*\|_{\mathcal{B}^0(\mathbb{R}^d)}^2}{n} \int_{\Omega} \text{Var}_{\xi \sim \mu} [\cos(\xi^\top x + \theta(\xi))] dx \\ &\quad + \frac{\|u^*\|_{\mathcal{B}^0(\mathbb{R}^d)}^2}{n} \sum_{k=1}^d \int_{\Omega} \text{Var}_{\xi \sim \mu} [\langle \xi, e_k \rangle \sin(\xi^\top x + \theta(\xi))] dx \\ &\leq \frac{\|u^*\|_{\mathcal{B}^0(\mathbb{R}^d)}^2}{n} \int_{\Omega} \mathbb{E}_{\xi \sim \mu} \left[1 + \sum_{k=1}^d \langle \xi, e_k \rangle^d \right] dx \\ &= \frac{\|u^*\|_{\mathcal{B}^0(\mathbb{R}^d)}}{n} \int_{\Omega} \int_{\mathbb{R}^d} |\hat{u}^*(\xi)| \cdot (1 + |\xi|^2) d\xi dx \\ &\leq \frac{m(\Omega) \|u^*\|_{\mathcal{B}^0(\mathbb{R}^d)} \|u^*\|_{\mathcal{B}^2(\mathbb{R}^d)}}{n} \\ &\leq \frac{m(\Omega) \cdot C^2 \|f\|_{\mathcal{B}^s(\mathbb{R}^d)}^2}{n}, \end{aligned}$$

which implies (2.2). \square

Acknowledgment. Z. Chen and S. Zhou thank Zexing Li for helpful discussions and communications.

REFERENCES

- [1] F. BACH, *Breaking the curse of dimensionality with convex neural networks*, J. Mach. Learn. Res., 18 (2017), pp. 629–681.
- [2] A. R. BARRON, *Universal approximation bounds for superpositions of a sigmoidal function*, IEEE Trans. Inform. Theory, 39 (1993), pp. 930–945.
- [3] Z. CHEN, J. LU, AND Y. LU, *On the representation of solutions to elliptic PDEs in Barron spaces*, in Advances in Neural Information Processing Systems 34, M. Ranzato, A. Beygelzimer, Y. Dauphin, P. Liang, and J. W. Vaughan, eds., Curran Associates, Red Hook, NY, 2021, pp. 6454–6465.
- [4] E. WEINAN, J. HAN AND A. JENTZEN, *Deep learning-based numerical methods for high-dimensional parabolic partial differential equations and backward stochastic differential equations*, Commun. Math. Stat., 5 (2017), pp. 349–380.
- [5] E. WEINAN, J. HAN AND A. JENTZEN, *Algorithms for solving high dimensional PDEs: From nonlinear Monte Carlo to machine learning*, Nonlinearity, 35 (2021), p. 278.

- [6] E. WEINAN, C. MA AND L. WU, *The Barron space and the flow-induced function spaces for neural network models*, Constr. Approx., (2021), pp. 1–38.
- [7] E. WEINAN, AND S. WOJTOWYTSCH, *Some observations on high-dimensional partial differential equations with Barron data*, in Proceedings of the 2nd Mathematical and Scientific Machine Learning Conference, J. Bruna, J. Hesthaven, and L. Zdeborova, eds., Proceedings of Machine Learning Research, PMLR 145, 2022, pp. 253–269.
- [8] E. WEINAN, AND B. YU, *The deep Ritz method: A deep learning-based numerical algorithm for solving variational problems*, Commun. Math. Stat., 6 (2018), pp. 1–12.
- [9] G. B. FOLLAND, *Real Analysis: Modern Techniques and Their Applications*, 2nd ed., John Wiley, New York, 1999.
- [10] J. HAN, A. JENTZEN, AND W. E, *Solving high-dimensional partial differential equations using deep learning*, Proc. Natl. Acad. Sci. USA, 115 (2018), pp. 8505–8510.
- [11] Q. HAN AND F.-H. LIN, *Elliptic Partial Differential Equations*, 2nd ed., Courant Lecture Notes, Courant Institute of Mathematical Sciences, Robotics Lab, New York University, New York, 2011.
- [12] H. HANCHE-OLSEN AND H. HOLDEN, *The Kolmogorov-Riesz compactness theorem*, Expo. Math., 28 (2010), pp. 385–394.
- [13] J. LU AND Y. LU, *A priori generalization error analysis of two-layer neural networks for solving high dimensional Schrödinger eigenvalue problems*, Comm. Amer. Math. Soc., 2 (2022), pp. 1–21.
- [14] Y. LU, J. LU, AND M. WANG, *A priori generalization analysis of the deep Ritz method for solving high dimensional elliptic partial differential equations*, in Proceedings of 34th Conference on Learning Theory, M. Belkin and S. Kpotufe, eds., Proceedings of Machine Learning Research, PMLR 134, 2021, pp. 3196–3241.
- [15] T. MARWAH, Z. LIPTON, AND A. RISTESKI, *Parametric complexity bounds for approximating PDEs with neural networks*, in Advances in Neural Information Processing Systems 34, M. Ranzato, A. Beygelzimer, Y. Dauphin, P. Liang, and J. W. Vaughan, eds., Curran Associates, Red Hook, NY, 2021, pp. 15044–15055.
- [16] R. PARHI AND R. D. NOWAK, *Banach space representer theorems for neural networks and ridge splines*, J. Mach. Learn. Res., 22 (2021), pp. 1–40.
- [17] R. PARHI AND R. D. NOWAK, *Near-Minimax Optimal Estimation with Shallow ReLU Neural Networks*, preprint, arXiv:2109.08844 [stat.ML], 2022.
- [18] J. W. SIEGEL AND J. XU, *High-order approximation rates for shallow neural networks with cosine and ReLU^k activation functions*, Appl. Comput. Harmon. Anal., 58 (2022), pp. 1–26.
- [19] J. W. SIEGEL AND J. XU, *Sharp bounds on the approximation rates, metric entropy, and n-widths of shallow neural networks*, Found. Comput. Math., (2022), pp. 1–57.