Regularization of inverse problems by filtered diagonal frame decomposition under general source

Regularization of Inverse Problems by Filtered Diagonal Frame Decomposition under general source

Dang Duc Trong[1,2], Nguyen Dang Minh[3], Luu Xuan Thang[4], Luu Dang Khoa[1,2]

October 8, 2025

[1] Faculty of Maths and Computer Science, University of Science, Ho Chi Minh City, Viet Nam.

[2] Vietnam National University, Ho Chi Minh City, Viet Nam.

[3] Faculty of Fundamental Science, Ho Chi Minh City Open University, Viet Nam.

[4] Faculty of Natural Science, University of Khanh Hoa, Viet Nam.

Abstract

This paper addresses the ill-posed inverse problem $\mathbf{K}x = y$ in real or complex Hilbert settings, where data y is contaminated by noise. We propose regularization methods utilizing Diagonal Frame Decomposition (DFD) as a generalization of SVD-based techniques to achieve stable solutions. Our approach introduces a regularization solution through filter-based methods, and we establish comprehensive theoretical results on convergence rates and optimality under a generalized source condition. These findings are applied to the fractional backward problem, specifically examining DFD system construction, relationships between DFD and SVD singular values, and extending existing source conditions for optimal regularization in polynomially and exponentially ill-posed scenarios.

Key words: Ill-posed problem; frame decomposition; convergence rates; Inverse problem.

MSC 2010: 47A52; 47J06.

1 Introduction

Let X and Y be real or complex Hilbert spaces, and let $\mathbf{K} : X \to Y$ be a bounded linear operator. In this paper, we seek a solution $x \in X$ to the inverse problem defined by the operator equation

$$\mathbf{K}x = y. \tag{1}$$

As is customary, we assume that the exact data y is unavailable, and instead, we are given noisy data y^δ with a known noise level δ. Specifically, the noise satisfies

$$\|y^\delta - y\|_Y \leq \delta. \tag{2}$$

[0]Corresponding author: Nguyen Dang Minh, minh.nd@ou.edu.vn

Due to the inherent instability of inverse problems, even a small perturbation in the data can lead to significant errors in the solution, rendering the numerical computation of solutions to the inverse problem challenging. To address this issue, we employ a regularization method for the system defined by equations (1) and (2). One such regularization approach, based on filtering techniques, has been thoroughly developed in [23] and [10]. In this context, $\mathbf{K}$ is assumed to be a compact operator possessing a singular system (u_k, v_k, σ_k). Consequently, $\mathbf{K}$ admits a singular value decomposition (SVD) of the form

$$\mathbf{K}x = \sum_{k=1}^{\infty} \sigma_k \langle x, u_k \rangle_X v_k,$$

where the symbols σ_k denote the singular values, and u_k and v_k are singular functions satisfying

$$\mathbf{K}u_k = \sigma_k v_k \quad \text{and} \quad \mathbf{K}^* v_k = \sigma_k u_k. \tag{3}$$

It is well known that the SVD is a fundamental tool for solving inverse problems. The minimum-norm least-squares solution $x^\dagger$ to equation (1) is then given by the Picard formula

$$x^\dagger := \mathbf{K}^\ddagger y := \sum_{k=1}^{\infty} \frac{\langle y, v_k \rangle_Y}{\sigma_k} u_k,$$

provided the Picard condition holds:

$$\sum_{k=1}^{\infty} \frac{|\langle y, v_k \rangle_Y|^2}{\sigma_k^2} < \infty.$$

When the exact data y is replaced by the noisy data y^δ with a given noise level δ, the approximate solution takes the form

$$x_\alpha^\delta := \mathbf{R}_\alpha y^\delta := \sum_{k=1}^{\infty} \sigma_k g_\alpha(\sigma_k^2) \langle y^\delta, v_k \rangle_Y u_k, \tag{4}$$

where $\alpha > 0$. We define g_α as the filter function, which has the property $g_\alpha(\lambda) \to 1/\lambda$ as $\alpha \to 0$ (see, e.g., [10, 25, 23, 26, 33]). In this context, $\mathbf{R}_\alpha$ serves as a regularization operator for equation (1). Furthermore, the convergence rate and optimality of this regularization are analyzed under the classical source condition $x^\dagger = \varphi(\mathbf{K}^*\mathbf{K})z$ for some $z \in X$ (see, e.g., [18, 30]). Recently, Hofmann et al. introduced the concept of the variational source condition (VSC) as an alternative to the classical source condition (see [17]). However, these investigations fall outside the scope of this paper, so we do not delve deeper into them.

Computing the SVD of an operator is often nontrivial and, in certain cases, computationally expensive, as noted by Ebner, Göppel, and Donoho et al. in [7], [14], and [5], respectively. Additionally, SVD-based regularization may not be well-suited for a variety of problems, as highlighted by Donoho in [5]. Thus, developing more efficient computational methods becomes essential. One approach is to identify a system that partially satisfies the SVD conditions in (3). A notable development in this direction retains the second condition, i.e., $\mathbf{K}^* v_\lambda = \kappa_\lambda u_\lambda$, where λ belongs to a countable index set Λ. This concept underpins the Wavelet-Vaguelette Decomposition (WVD) in [5], and more generally, the Diagonal Frame Decomposition (DFD) in [7], as well as the Translation-Invariant DFD (TI-DFD) in [14]. By leveraging frame theory, such generalizations enable the construction of an expansion analogous to the Picard formula, suggesting that regularization methods tailored to this framework hold significant potential. This topic has only been developed in the last few years, and related papers are still very rare. Among them, we would like to introduce two papers: [7] and [21]. In particular, the recent paper [21] contains many ideas that can be further developed. Our paper is inspired by [7, 21] and the closely related book [10]. Therefore, we frequently reference these documents in our discussions.

To regularize the solution of (2) in the DFD setting, numerous methods have been proposed. Specifically, learned filter methods are employed in [9], while the ℓ^1 -Tikhonov method is presented in [11]. Furthermore, in [12], the authors developed sparse regularization through operator-adapted frame thresholding. Even with the emergence of these very new methods, the filter function approach remains an active area of research. A generalized filter-based formulation was proposed by Ebner and Haltmeier (see [8]) in which the coefficient $\sigma_k^{-1}\langle y, v_k\rangle_Y$ in the Picard formula is replaced by the filter term $\kappa_\lambda^{-1}\varphi_\alpha(\kappa_\lambda, \langle y^\delta, v_\lambda\rangle_Y)$ which could be nonlinear with respect to $\langle y^\delta, v_\lambda\rangle_Y$. Indeed, in [7], the authors reformulate foundational concepts for DFD-based regularization filtering, akin to the SVD filtering presented in [10]. The quantity κ_λ is not a singular value. Therefore, it is generally not bound by the condition $\kappa_\lambda > 0$, and we can typically assume κ_λ is a complex number. For the scope of this paper, we will limit our discussion to linear filter-based methods where $\varphi_\alpha(\kappa_\lambda, \langle y^\delta, v_\lambda\rangle_Y)$ is a linear function of $\langle y^\delta, v_\lambda\rangle_Y$ (see, e.g., [7, 21, 19, 20, 28, 35]).

To estimate regularization errors in the linear filter-based method, the aforementioned papers (specifically paper [7]) adapt the SVD source condition, assuming the solution belongs to a DFD-type source set with a polynomial (or Hölder) form:

$$\mathbf{M} := \left\{ x \in X : \langle x, u_\lambda\rangle_X = \kappa_\lambda^{2\mu} w_\lambda \ \forall \lambda \in \Lambda \text{ and } \sum_{\lambda\in\Lambda} |w_\lambda|^2 \le \rho^2 \right\}, \tag{5}$$

where $\mu, \rho > 0$ and $(u_\lambda, v_\lambda, \kappa_\lambda)$ constitutes a DFD of the operator $\mathbf{K}$. A similarly modified source condition for complex scenarios appears in [21], where the authors explore the optimality of a-posteriori regularization methods. The framework developed in [7, 21] opens up numerous application possibilities. As this theory is still emerging, several natural questions arise:

(i) How can DFD systems be constructed for a specific operator $\mathbf{K}$?

(ii) What is the relationship between the DFD singular values κ_λ and the SVD singular values σ_k? How do these quasi-singular values influence the regularization of ill-posed problems?

(iii) To what extent can the regularization theory for polynomial DFD source conditions be generalized to include non-polynomial forms, specifically logarithmic source conditions for complex scenarios?

(iv) How does the DFD source condition relate to the classical source condition?

(v) Do a priori and a-posteriori regularization methods achieve optimality?

Question (i) is particularly compelling and has been extensively explored in the field of tomography (see [7, 21, 28]), photoacoustic tomography ([11, 12]), atmospheric tomography ([19, 35]). This problem exhibits polynomial ill-posedness, where WVD systems prove effective. In Section 4 of this paper, we examine the backward fractional problem, considering two scenarios: polynomial ill-posedness and exponential ill-posedness. In the latter case, the WVD system appears inadequate, prompting us to propose a specialized DFD system.

The investigation of Question (ii) remains in its early stages. In [7, 21], it is limited to assessing the ill-posedness of the problem $\mathbf{K}x = y$. Our paper advances this inquiry by exploring the "sparseness" or "thickness" of the DFD singular values κ_λ through the set

$$D_{\lambda,\beta} = [\delta_\lambda^*, \beta^{-1}\delta_\lambda^*], \quad \text{where} \quad \delta_\lambda^* \asymp \sqrt{|\kappa_\lambda|^2\Phi(|\kappa_\lambda|^2)},$$

with Φ referred to as a *source function* (detailed in the subsequent paragraphs). This set facilitates assertions regarding the sequential or uniform optimality of regularization methods. A new DFD quasi-minimal property is also proposed by us to ease the minimality requirement of frames in certain scenarios.

Motivated by question (iii), we must devise a method for constructing the filter function and defining the source function that allows for adapting classical regularization theory to the complex framework. We first consider the ideas to construct a general filter function. From Picard's formula, to stabilize the solution, common ideas are of using the relation $1/\sigma_k \approx \sigma_k g_\alpha(\sigma_k^2)$, where $g_\alpha(\lambda) \to 1/\lambda$ as $\alpha \to 0^+$. Hubmer et al. [21] use a similar idea for the case where $\sigma_k > 0$ is replaced by $\kappa_\lambda \in \mathbb{C}$: they substitute $1/\kappa_\lambda$ by $\kappa_\lambda g_\alpha(\kappa_\lambda^2)$, where $g_\alpha : \mathbb{C} \to \mathbb{C}$ satisfies $g_\alpha(\lambda) \to 1/\lambda$ as $\alpha \to 0^+$. This definition works perfectly well for the case when κ_λ is a positive real number. However, when κ_λ is a complex number, SVD filter functions - such as the Tikhonov filter, Landweber filter - are difficult to apply directly. Thus, an approach to constructing filter functions is required to directly apply classical SVD filters to the complex scenario. This paper aims to accomplish this.

Subsequently, we explore generalizations of the source function definition. In [21], the authors used the source set (5) corresponding with the source function $\varphi(\mu) = \mu^\nu$. As with the filter function, this power source function is well-defined if μ is a non-negative real number. However, if μ is a complex number, then μ^ν is a multi-valued function, making it difficult to define the function well and to extend it to other functions like the logarithm in complex framework. To overcome this difficulty, we revisit the SVD source function in [7, 21]. This function is defined on $(0,\infty)$, and we can write $\mu^\nu = |\mu|^\nu$ for every $\mu > 0$. The two formulas are equivalent when μ is a non-negative real number. However, the latter formula seems more suitable for generalizing to complex μ, since, as discussed, μ^ν is then a multi-valued function with respect to μ. To address this, rather than selecting the general source function $\varphi : \mathbb{C} \to \mathbb{C}$, we will restrict our focus to functions of the form $\varphi(\mu) = \Phi(|\mu|)$, where $\Phi : (0,\infty) \to (0,\infty)$ is a positive real-valued function. Using this idea, we extend the results of [21] (and [7]) to a DFD source set defined by a general source function, rather than a polynomial one. Specifically, for a positive constant E,

$$\mathbf{M}_{\Phi,E} := \left\{ x \in X : \sum_{\lambda\in\Lambda} [\Phi(|\kappa_\lambda|^2)]^{-1} |\langle x, u_\lambda\rangle_X|^2 \leq E^2 \right\}, \tag{6}$$

where the "source" function Φ satisfies conditions detailed in subsequent assumptions. Such conditions naturally arise in ill-posed problems, such as tomography with $\Phi(\mu) = |\mu|^{2\nu}$ or the backward problem with $\Phi(\mu) = (-\ln|\mu|)^{-p}$ (see subsequent sections). This topic merits further attention, and in Subsection 4.4 of our paper, we study the latter index function Φ.

To address Question (iv), we present two examples demonstrating that the classical source condition can suffice to derive the DFD source condition. These examples illustrate the connection between classical and DFD source conditions, though these findings are preliminary and warrant deeper investigation in future work.

Question (v) is thoroughly explored in this paper. Investigations into this topic can be found in [7, 11, 21]. In [11], the authors studied the optimal problem with a source condition in Besov spaces. Similarly, a spectral source condition was the subject of study in both [7] and [21]. Inspired by Hubmer et al.'s work [21], which assessed the optimal convergence rates of a-posteriori regularization under a polynomial source condition, our paper addresses a significant gap: the optimal a-posteriori convergence rates for non-polynomial source conditions have not yet been investigated. To assess the optimality of the regularizations, we extended the findings of [7] and derived a lower bound on the worst-case error for the general source function. Our proof refined the condition that the frame be minimal. We instead demonstrated the results for frames satisfying a more general property, provisionally termed DFD quasi-minimal. Moreover, we enhance the existing analysis by specifically addressing a-posteriori strategies, a gap left by [7, 21] and [14]. We further refine the classification of optimality properties, distinguishing between sequential order optimality (as noted in [7, 21]) and global order optimality, the latter of which has not been previously explored.

Structurally, Section 2 provides a review of foundational results on frames and introduces the concept of optimal regularization. Section 3 is dedicated to the paper's core findings, encompassing lower bounds for the modulus of continuity of $\mathbf{K}^{-1}$ on the set

$\mathbf{M}_{\Phi,E}$, along with convergence rates for both a priori and a-posteriori parameter selections. Section 4 illustrates these theoretical findings through specific examples and their corresponding numerical experiments. Finally, Section 5 contains the long and technical proofs of the main results.

2 Some basic notions and notations

2.1 Notions of frames

For $z \in \mathbb{C}$, we denote its conjugate by $\overline{z}$, its modulus by $|z|$ and its real part by Re z. We also have the equality $|z_1 \pm z_2|^2 = |z_1|^2 \pm 2\mathrm{Re}\ z_1\overline{z}_2 + |z_2|^2$. Letting Λ be an at most countable set of indices, we denote

$$l^2(\Lambda) = \left\{ \mathbf{a} = (a_\lambda)_{\lambda\in\Lambda} : a_\lambda \in \mathbb{C}, \sum_{\lambda\in\Lambda} |a_\lambda|^2 < \infty \right\}$$

with the norm $\|\mathbf{a}\|_2 = \left(\sum_{\lambda\in\Lambda} |a_\lambda|^2\right)^{1/2}$.

Before delving into the specific content of the article, we would like to recall some results about frames in a Hilbert space $\mathcal{H}$. These results can be found in [2], [7], and [21]. For convenience, let us introduce the definition of a frame

Definition 1. *A sequence $\mathbf{w} = \{w_\lambda\}_{\lambda\in\Lambda}$ in a Hilbert subspace $\mathbb{H} \subset \mathcal{H}$ is called a frame over $\mathbb{H}$, if and only if there exist frame bounds $0 < A_w, B_w \in \mathbb{R}$ such that for all $x \in \mathbb{H}$ there holds*

$$A_w \|x\|^2_{\mathcal{H}} \leq \sum_{\lambda\in\Lambda} |\langle x, w_\lambda\rangle_{\mathcal{H}}|^2 \leq B_w \|x\|^2_{\mathcal{H}}. \tag{7}$$

If $w_{\lambda_0} \notin \overline{\mathrm{span}\{w_\lambda\}}_{\lambda\neq\lambda_0}$ for every $\lambda_0 \in \Lambda$ then we say that the frame is minimal.

From now on, we denote $\|x\|_w := \sqrt{\sum_{\lambda\in\Lambda} |\langle x, w_\lambda\rangle_{\mathcal{H}}|^2}$ for every $x \in \mathcal{H}$ and

$$|\mathbf{w}|_{\inf} = \inf\{\|x\|_w : x \in \mathbb{H} \text{ and } \|x\|_{\mathcal{H}} = 1\},$$
$$|\mathbf{w}|_{\sup} = \sup\{\|x\|_w : x \in \mathbb{H} \text{ and } \|x\|_{\mathcal{H}} = 1\}.$$

If $|\mathbf{w}|_{\inf} = |\mathbf{w}|_{\sup}$, we say that the frame is tight and denote $|\mathbf{w}|_{\mathrm{fr}} := |\mathbf{w}|_{\sup} = |\mathbf{w}|_{\inf}$.

From the definition, we have $0 < \sqrt{A_w} \leq |\mathbf{w}|_{\inf} \leq |\mathbf{w}|_{\sup} \leq \sqrt{B_w}$ and

$$|\mathbf{w}|_{\inf}\|x\|_{\mathcal{H}} \leq \|x\|_w \leq |\mathbf{w}|_{\sup}\|x\|_{\mathcal{H}} \tag{8}$$

for every $x \in \mathbb{H}$. For $x' \in \mathcal{H}$, we have

$$\|x'\|^2_w = \sum_{\lambda\in\Lambda} |\langle P_{\mathbb{H}}x', w_\lambda\rangle_{\mathcal{H}}|^2 \leq |\mathbf{w}|_{\sup} \|P_{\mathbb{H}}x'\|^2_{\mathcal{H}} \leq |\mathbf{w}|_{\sup} \|x'\|^2_{\mathcal{H}}. \tag{9}$$

Here $P_{\mathbb{H}}$ is the orthogonal projection on $\mathbb{H}$. For a given frame $\{w_\lambda\}_{\lambda\in\Lambda}$, one can define the frame analysis operator F as below

$$F : \mathbb{H} \to l_2(\Lambda), \quad x \mapsto \{\langle x, w_\lambda\rangle_{\mathcal{H}}\}_{\lambda\in\Lambda},$$

and, the synthesis operator F^*, which is given by

$$F^* : l_2(\Lambda) \to \mathbb{H}, \quad (a_\lambda)_{\lambda\in\Lambda} \mapsto \sum_{\lambda\in\Lambda} a_\lambda w_\lambda.$$

From the inequality (7), there holds

$$\sqrt{A_w} \le \|F\| = \|F^*\| \le \sqrt{B_w}.$$

We can define the operator $S := F^*F$, that is,

$$Sx := \sum_{\lambda\in\Lambda} \langle x, w_\lambda\rangle_{\mathcal{H}} w_\lambda.$$

It is worth noting that, in this case, the operator S is a bounded, linear, and invertible operator. Specifically, $A_w I \le S \le B_w I$ and $B_w^{-1} I \le S^{-1} \le A_w^{-1} I$. Therefore, if we set $\widetilde{w}_\lambda := S^{-1} w_\lambda$, $\widetilde{\mathbf{w}} = \{\widetilde{w}_\lambda\}_{\lambda\in\Lambda}$, then

$$B_w^{-1} \|x\|_{\mathcal{H}}^2 \le \|x\|_{\widetilde{w}}^2 \le A_w^{-1} \|x\|_{\mathcal{H}}^2$$

for every $x \in \mathbb{H}$. Consequently, the set $\{\widetilde{w}_\lambda\}_{\lambda\in\Lambda}$ is also a frame over $\mathbb{H}$. As we know, it is referred to as the dual frame of $\{w_k\}_{k\in\mathbb{N}}$. In that case, the analysis and synthesis operators of this frame are as follows. The analysis operator $\widetilde{F}$ is defined as below

$$\widetilde{F} : \mathbb{H} \to l_2(\Lambda), \quad x \mapsto \{\langle x, \widetilde{w}_\lambda\rangle_{\mathcal{H}}\}_{\lambda\in\Lambda}$$

and, the synthesis operator $\widetilde{F}^*$, which is given by

$$\widetilde{F}^* : l_2(\Lambda) \to \mathbb{H}, \quad \{a_\lambda\}_{\lambda\in\Lambda} \mapsto \sum_{\lambda\in\Lambda} a_\lambda \widetilde{w}_\lambda.$$

From the inequalities (7), (8) there also holds

$$\sqrt{B_w^{-1}} \le |\mathbf{w}|_{\sup}^{-1} \le \|\widetilde{F}\| = \|\widetilde{F}^*\| \le |\mathbf{w}|_{\inf}^{-1} \le \sqrt{A_w^{-1}}.$$

It follows that

$$\left\|\sum_{\lambda\in\Lambda} a_\lambda \widetilde{w}_\lambda\right\|_{\mathcal{H}} = \|\widetilde{F}^*(\{a_\lambda\})\|_{\mathcal{H}} \le |\mathbf{w}|_{\inf}^{-1} \left(\sum_{\lambda\in\Lambda} |a_\lambda|^2\right)^{1/2}. \tag{10}$$

Moreover, it can also be proved that $\widetilde{F}^* F = F^* \widetilde{F} = I$, and thus, for any $x \in \mathbb{H}$, it can always be expressed as $x = \sum_{\lambda\in\Lambda} x_\lambda \widetilde{w}_\lambda$ where $x_\lambda = \langle x, w_\lambda\rangle_{\mathbb{H}} + a_\lambda$ with $\mathbf{a} = (a_\lambda) \in N(\widetilde{F}^*)$. Especially, we have

$$x = \sum_{\lambda\in\Lambda} \langle x, w_\lambda\rangle_{\mathbb{H}} \widetilde{w}_\lambda. \tag{11}$$

Generally, the calculation of $\widetilde{w}_\lambda$ is only easy in some special cases. In fact, if $\{w_\lambda\}$ is tight then $\widetilde{w}_\lambda = \frac{1}{|\mathbf{w}|_{\text{fr}}} w_\lambda$ (see, e.g., [2], chap. 5). In general, we always have $\{0\} \subset N(F^*) = N\left(\widetilde{F}^*\right)$, and therefore, the representation of x in (11) is not unique. However, this representation is considered the most economical according to [21]. From [7], we have known that the frame $\{\widetilde{w}_\lambda\}$ is the biorthonormal sequence of $\{w_\lambda\}$, i.e. $\langle w_\lambda, \widetilde{w}_\nu\rangle_{\mathcal{H}} = \delta_{\lambda\nu}$ for $\lambda, \nu \in \Lambda$, is equivalent to $\{w_\lambda\}$ being minimal. In this case, we have $x_\lambda = \langle x, w_\lambda\rangle_{\mathbb{H}}$ and the expansion (11) is unique. Next, we recall the definition of diagonal frame decomposition (see, e.g., [21]).

Definition 2. *Let $\mathbf{K} : X \to Y$ be a bounded linear operator, and Λ is an at most countable index set. We define $(\mathbf{u}, \mathbf{v}, \boldsymbol{\kappa}) = (u_\lambda, v_\lambda, \kappa_\lambda)_{\lambda\in\Lambda}$ as a diagonal frame decomposition (DFD) for the operator $\mathbf{K}$ if the following conditions hold*

(D1) $\{u_\lambda\}_{\lambda\in\Lambda}$ *is a frame over* $(\ker\mathbf{K})^\perp \subset X$.

(D2) $\{v_\lambda\}_{\lambda\in\Lambda}$ *is a frame over* $\overline{\operatorname{ran}\mathbf{K}} \subset Y$.

(D3) $(\kappa_\lambda)_{\lambda\in\Lambda} \in (\mathbb{C} \setminus \{0\})^\Lambda$ *satisfies the quasi-singular relations*

$$\mathbf{K}^* v_\lambda = \overline{\kappa}_\lambda u_\lambda, \ \text{for all } \lambda \in \Lambda.$$

The κ_λ values are called the DFD singular values.

Remark 1. *We can replace conditions (D1), (D2) with*

(D1)' $\{u_\lambda\}_{\lambda\in\Lambda}$ *is a frame over* X_0,

(D2)' $\{v_\lambda\}_{\lambda\in\Lambda}$ *is a frame over* Y_0.

Here X_0 and Y_0 are (closed) subspaces of Hilbert spaces X and Y, respectively, such that $(\ker\mathbf{K})^\perp \subset X_0$ and $\overline{\operatorname{ran}\mathbf{K}} \subset Y_0$. If the frames are chosen in this manner, κ_λ may be equal to 0. As suggested in [21], utilizing conditions (D1)' and (D2)' often allows for the selection of more manageable frames $\{u_\lambda\}$ and $\{v_\lambda\}$, thereby significantly streamlining the computational procedure. A detailed discussion regarding this concept is available for readers in [21].

For $h : \mathbb{R} \to [0, \infty)$, $x \in X$, we define

$$\langle h(\mathbf{K}^*\mathbf{K})x, x\rangle_u = \sum_{\lambda\in\Lambda} h(|\kappa_\lambda|^2)|\langle x, u_\lambda\rangle_Y|^2.$$

From (**D1**), (**D2**) we can find numbers $A_u, A_v, B_u, B_v > 0$ such that

$$A_u\|w\|_X^2 \le \|w\|_u^2 \le B_u\|w\|_X^2, \forall w \in (\ker\mathbf{K})^\perp, \tag{12}$$

$$A_v\|z\|_Y^2 \le \|z\|_v^2 \le B_v\|z\|_Y^2, \forall z \in \overline{\operatorname{ran}\mathbf{K}}. \tag{13}$$

From now on, we always denote by a^* an extended real number such that $a^* > \sup_\lambda |\kappa_\lambda|^2$ if $\sup_\lambda |\kappa_\lambda|^2 < \infty$ and $a^* = \infty$ if $\sup_\lambda |\kappa_\lambda|^2 = \infty$.

2.2 Notions of the worst case error and optimality

Consider the problem (1) and denote the Moore-Penrose operator

$$\mathbf{K}^{\ddagger}(z) = \operatorname{argmin}\{\|h\|_X : h \in X, z \in \operatorname{ran}\mathbf{K}, \mathbf{K}(h) = z\}.$$

We denote the Moore-Penrose solution of (1) by $x^{\dagger} = \mathbf{K}^{\ddagger}y$. Let an operator $\mathbf{R} : Y \to X$ satisfy $\mathbf{R}y \approx x$. We say that $\mathbf{R}$ is an approximation method of the problem (1). Assume that the solution $x^{\dagger}$ of (1) belongs to a subset $\mathbf{M} \subset X$, we recall the definition of the worst-case error of the method $\mathbf{R}$ on $\mathbf{M}$ as below.

$$\Delta(\mathbf{M}, \delta, \mathbf{R}) := \sup \left\{ \|\mathbf{R}y^{\delta} - x^{\dagger}\|_Y : x^{\dagger} \in \mathbf{M} \wedge y^{\delta} \in Y \wedge \|\mathbf{K}x^{\dagger} - y^{\delta}\|_Y \le \delta \right\}.$$

Worst-case error holds a vital role in optimal regularization theory. Many regularization techniques involve a regularization parameter that balances data fidelity with solution stability. The worst-case error framework helps in making an informed choice of this parameter. You can analyze how the worst-case error behaves as this parameter changes, allowing you to select a value that minimizes the maximum possible error, thereby yielding the most robust and accurate solution given the input uncertainties. Many classical references discuss worst-case error (see, [10, 18, 25, 26, 30, 33]). In this paper, we extensively refer to the worst-case error from [7] and [33]. Drawing on the aforementioned worst-case error concept, the previously mentioned documents provide the following definition for the optimality of an approximation method:

Definition 3. *We say that the method* $\mathbf{R}_{opt} : Y \to X$ *is optimal on* $\mathbf{M}$ *if* $\Delta(\mathbf{M}, \delta, \mathbf{R}_{opt}) = \inf_{\mathbf{R}} \Delta(\mathbf{M}, \delta, \mathbf{R})$ *and* $\mathbf{R}_{opt} : Y \to X$ *is order optimal on* $\mathbf{M}$ *if there is a constant* $c > 0$ *independent of* δ *such that* $\Delta(\mathbf{M}, \delta, \mathbf{R}_{opt}) \le c \inf_{\mathbf{R}} \Delta(\mathbf{M}, \delta, \mathbf{R})$.

In our paper, we choose $\mathbf{M} = \mathbf{M}_{\Phi,E}$ defined in (6) and $\mathbf{R}$ is in a regularization method. To assist our readers, we will now recall the concept of regularization method. Let $\mathbf{R}_{\alpha} : Y \to X$, $\alpha > 0$, be a family of bounded operators and let $\alpha^* : (0, \alpha_0) \times Y \to (0, \infty)$. As in [7, 10], we say that $(\mathbf{R}_{\alpha}, \alpha^*)$ is a regularization method if

$$\lim_{\delta \to 0^+} \sup\{\alpha^*(\delta, y^{\delta}) : \; y^{\delta} \in Y \wedge \|y^{\delta} - y\|_Y \le \delta\} = 0,$$
$$\lim_{\delta \to 0^+} \sup\{\|\mathbf{K}^{\ddagger}y - \mathbf{R}_{\alpha^*(\delta, y^{\delta})}\|_X : \; y^{\delta} \in Y \wedge \|y^{\delta} - y\|_Y \le \delta\} = 0.$$

The quantities α and α^* are called the regularization parameter and the admissible parameter choice respectively. In the framework, our goal of finding $\mathbf{R}_{opt}$ now reduces to determining the parameter α^* that optimizes $\Delta(M_{\Phi,E}, \delta, \mathbf{R}_{\alpha})$. Inspired by the classical optimal regularization theory ([30, 33]), we can classify the order optimality for our problem.

Definition 4. *Let* $\mathbf{M} \subset X$*. We say that the regularization method* $(\mathbf{R}_\alpha, \alpha^*)$ *is*

(a) sequential order optimal on $\mathbf{M}$ *if there is a sequence* $\delta_n \to 0^+$ *such that there exists a constant* $c > 0$ *independent of* n *such that* $\Delta(\mathbf{M}, \delta_n, \mathbf{R}_{\alpha^*(\delta_n, y^{\delta_n})}) \le c \inf_{\mathbf{R}} \Delta(\mathbf{M}, \delta_n, \mathbf{R})$,

(b) uniform order optimal on $\mathbf{M}$ *if there is a* δ_0 *and a constant* c *independent of* δ *such that* $\Delta(\mathbf{M}, \delta, \mathbf{R}_{\alpha^*(\delta, y^\delta)}) \le c \inf_{\mathbf{R}} \Delta(\mathbf{M}, \delta, \mathbf{R})$ *for every* $\delta \in (0, \delta_0)$.

Sequential optimal regularization is studied in the recent papers [7, 14, 21], but uniformly optimal regularizations have not been discussed yet.

3 Main results

In this section, we aim to provide an overview of the main results of the paper. Therefore, only a few brief proofs will be presented immediately after the theorem statements. Theorems without immediate proofs are those with lengthy and technically involved proofs. These proofs will be deferred to the final section.

3.1 Pointwise convergence

Let $(\mathbf{u}, \mathbf{v}, \boldsymbol{\kappa})$ be a DFD for $\mathbf{K}$ and y be as in (1), we have

$$\langle y, v_\lambda \rangle_Y = \langle \mathbf{K} x^\dagger, v_\lambda \rangle_Y = \langle x^\dagger, \mathbf{K}^* v_\lambda \rangle_X = \langle x^\dagger, \overline{\kappa}_\lambda u_\lambda \rangle_X = \kappa_\lambda \langle x^\dagger, u_\lambda \rangle_X \text{ for } \lambda \in \Lambda. \tag{14}$$

Hence, from the expansion (11), the Moore-Penrose solution of (1) has the expansion

$$x^\dagger := \mathbf{K}^\ddagger y = \sum_{\lambda \in \Lambda} \frac{1}{\kappa_\lambda} \langle y, v_\lambda \rangle_Y \widetilde{u}_\lambda. \tag{15}$$

The expansion implies that $y \in \operatorname{dom} \mathbf{K}^\ddagger$ if and only if $\sum_\lambda \left| \frac{\langle y, v_\lambda \rangle_Y}{\kappa_\lambda} \right|^2 < \infty$.

Remark 2. *(i) If* $\{u_\lambda\}$ *is tight then we obtain*

$$x^\dagger := \mathbf{K}^\ddagger y = \sum_{\lambda \in \Lambda} \frac{1}{\kappa_\lambda |\mathbf{u}|_{\mathrm{fr}}} \langle y, v_\lambda \rangle_Y u_\lambda.$$

(ii) If $\{u_\lambda\}$ *is a frame over the whole* X *then the latter equation can be replaced by*

$$x^\dagger := \mathbf{K}^\ddagger y = \sum_{\lambda \in \Lambda, \kappa_\lambda \neq 0} \frac{1}{\kappa_\lambda} \langle y, v_\lambda \rangle_Y \widetilde{u}_\lambda.$$

The stability of solution (15) depends on the infimum of $\{|\kappa_\lambda|\}$. In fact we have

Theorem 3.1. *Let* $(u_\lambda, v_\lambda, \kappa_\lambda)_{\lambda \in \Lambda}$ *be a DFD as in Definition 2. We have the equivalence of the following two conditions:*

(i) $\inf_{\lambda \in \Lambda} |\kappa_\lambda| > 0$,

(ii) the operator $\mathbf{K}^\ddagger : \operatorname{ran} \mathbf{K} \to X$ *is bounded and*

$$\inf_{\lambda \in \Lambda} \frac{\|v_\lambda\|_Y}{\|u_\lambda\|_X} > 0. \tag{16}$$

Remark 3. *Our result constitutes an advancement upon those established by [7]. The condition that $\|v_\lambda\|_Y$ is bounded below implies the condition 16. Specifically, since (u_λ) is a frame, we have $\|u_\lambda\|_X^4 = |\langle u_\lambda, u_\lambda\rangle|_X^2 \le |\mathbf{u}|_{\sup}^2 \|u_\lambda\|_X^2$. It follows that $\|u_\lambda\|_X \le |\mathbf{u}|_{\sup}$. Consequently, if $\inf_{\lambda\in\Lambda}\|v_\lambda\|_Y > 0$, then*

$$\inf_{\lambda\in\Lambda} \frac{\|v_\lambda\|_Y}{\|u_\lambda\|_X} \ge \frac{\inf_{\lambda\in\Lambda}\|v_\lambda\|_Y}{\sup_{\lambda\in\Lambda}\|u_\lambda\|_X} > 0.$$

Proof. $(i) \Rightarrow (ii)$: If $\inf_{\kappa_\lambda\in\Lambda}|\kappa_\lambda| \ge \kappa_0 > 0$ then we obtain, in view of (10), that

$$\|\mathbf{K}^\ddagger h\|_X = \left\|\sum_{\lambda\in\Lambda} \frac{1}{\kappa_\lambda}\langle h, v_\lambda\rangle_Y \widetilde{u}_\lambda\right\|_X \le \frac{1}{\kappa_0|\mathbf{u}|_{\inf}}\left(\sum_{\lambda\in\Lambda}|\langle h, v_\lambda\rangle_Y|^2\right)^{1/2} \le \frac{|\mathbf{v}|_{\sup}}{\kappa_0|\mathbf{u}|_{\inf}}\|h\|_Y$$

for every $h \in \mathrm{ran}\mathbf{K}$, i.e., the operator $\mathbf{K}^\ddagger : \mathrm{ran}\mathbf{K} \to X$ is bounded. From the condition (D3) of the DFD, we have

$$\|K^*\|.\|v_\lambda\|_Y \ge \|u_\lambda\|_X \inf_{\lambda\in\Lambda}|\kappa_\lambda|.$$

It follows that

$$\inf_{\lambda\in\Lambda}\frac{\|v_\lambda\|_Y}{\|u_\lambda\|_X} \ge \frac{1}{\|K^*\|}\inf_{\lambda\in\Lambda}|\kappa_\lambda| > 0.$$

$(ii) \Rightarrow (i)$: We verify that $\mathrm{ran}\mathbf{K} = \overline{\mathrm{ran}\mathbf{K}}$. Choosing $y_0 \in \overline{\mathrm{ran}\mathbf{K}}$, we can find $y_n \in \mathrm{ran}\mathbf{K}$ such that $\lim_{n\to\infty}\|y_n - y_0\|_Y = 0$. Denote $x_n = \mathbf{K}^\ddagger y_n$. We have

$$\|x_n - x_m\|_X = \|\mathbf{K}^\ddagger(y_n - y_m)\|_X \le \|\mathbf{K}^\ddagger\|.\|y_n - y_m\|_Y.$$

Because (y_n) converges in $\overline{\mathrm{ran}\mathbf{K}}$, the latter inequality implies that (x_n) converges to an element x_0 in $(\ker\mathbf{K})^\perp$. We can deduce that $\mathbf{K}x_0 = y_0$, i.e. $y_0 \in \mathrm{ran}\mathbf{K}$.

Since $\mathrm{ran}\mathbf{K} = \overline{\mathrm{ran}\mathbf{K}}$, the operator $\mathbf{K} : (\ker\mathbf{K})^\perp \to \overline{\mathrm{ran}\mathbf{K}}$ is bijective. Here the restriction of the operator $\mathbf{K}$ to $(\ker\mathbf{K})^\perp$ is still denoted by $\mathbf{K}$. We deduce that $\mathbf{K}^{-1} = \mathbf{K}^\ddagger : \overline{\mathrm{ran}\mathbf{K}} \to (\ker\mathbf{K})^\perp$ is bounded. Hence we have $\overline{\kappa}_\lambda(\mathbf{K}^\ddagger)^* u_\lambda = v_\lambda$ which implies

$$|\kappa_\lambda|.\|(\mathbf{K}^\ddagger)^*\|.\|u_\lambda\|_X \ge \|v_\lambda\|_Y.$$

It follows that

$$\inf_{\lambda\in\Lambda}|\kappa_\lambda| \ge \|(\mathbf{K}^\ddagger)^*\|^{-1}\inf_{\lambda\in\Lambda}\frac{\|v_\lambda\|_Y}{\|u_\lambda\|_X} > 0.$$

This completes the proof of Theorem 3.1. □

If $\inf_{\lambda\in\Lambda}|\kappa_\lambda| = 0$, the solution $x^\dagger$ of equation (15) can be unstable. As mentioned in Introduction, we present here the idea to obtain our filter functions. Suggested by the equation (15), we can rewrite

$$x^\dagger := \mathbf{K}^\ddagger y = \sum_{\lambda\in\Lambda}\frac{\overline{\kappa}_\lambda}{|\kappa_\lambda|^2}\langle y, v_\lambda\rangle_Y \widetilde{u}_\lambda.$$

Hence, we just need to choose a real function $g_\alpha : (0,\infty) \to \mathbb{R}$ such that $g_\alpha(\mu) \to 1/\mu$ as $\alpha \to 0^+$ to be *a filter function*. Building the latter ideas, and the framework of (4), we construct a filtered regularization of the form

$$x_\alpha^\delta := R_\alpha y^\delta = \sum_{\lambda\in\Lambda} \overline{\kappa}_\lambda g_\alpha(|\kappa_\lambda|^2)\langle y^\delta, v_\lambda\rangle_Y \widetilde{u}_\lambda \tag{17}$$

where $\{\widetilde{u}_\lambda\}_{\lambda\in\Lambda}$ is the dual frame of $\{u_\lambda\}_{\lambda\in\Lambda}$. This regularization formula is entirely compatible with the complex framework and allows us to use SVD regularization principles for this new improvement. The functions $g_\alpha : [0, a^*) \to \mathbb{R}$ will be chosen as standard filter functions (see, e.g., [10, 18, 23, 25, 30, 33]) that satisfy

Assumption C

(C1) For all $\alpha > 0$, $\mu \in [0, a^*)$: $\sqrt{\mu} g_\alpha(\mu) < \infty$,

(C2) There exists a constant $C_g > 0$ such that $\sup\{|\mu g_\alpha(\mu)| : \alpha > 0, 0 \le \mu < a^*\} \le C_g$,

(C3) For all $\mu \in (0, a^*)$ there holds $\lim_{\alpha\to 0} \mu g_\alpha(\mu) = 1$.

To illustrate, we can list the three commonly used filtering functions as such:

$$g_\alpha(\mu) = \begin{cases} (\alpha+\mu)^{-1} & \text{Tikhonov filter,} \\ \mu^{-1}\chi_{[\alpha,\infty)}(\mu) & \text{Truncated filter,} \\ \mu^{-1}(1-(1-\tau\mu)^{1/\alpha}) & \text{Landweber filter,} \end{cases}$$

where $0 < \tau a^* < 1$ (see, e.g., [10, 23, 25, 33]) and $\chi_{[\alpha,\infty)}(\mu) = 1$ if $\mu \ge \alpha$, $\chi_{[\alpha,\infty)}(\mu) = 0$ if $\mu < \alpha$. For the sake of conciseness in the expressions, we will introduce common notations (see, e.g., [21, 33]) in the definition below.

Definition 5. *We denote* $r_\alpha(\mu) = 1 - \mu g_\alpha(\mu)$, $\rho(\alpha) = \sup_{\mu\in(0,a^*)} r_\alpha(\mu)$, $\ell(\alpha) = \sup_{\mu\in(0,a^*)} g_\alpha(\mu)$, $L(\alpha) = \sup_{\mu\in(0,a^*)} \sqrt{\mu} g_\alpha(\mu)$ *and*

$$d_\alpha(h) = \left(\sum_{\lambda\in\Lambda} (r_\alpha(|\kappa_\lambda|^2))^2 |\langle h, v_\lambda\rangle_Y|^2\right)^{1/2} \quad \textit{for every } h \in Y.$$

From Assumption C2, we obtain $0 \le \rho(\alpha) \le \max\{C_g, 1\}$. For convenience, we show here some properties of the function d_α. We note that d_α satisfies the triangle inequality $d_\alpha(h + z) \le d_\alpha(h) + d_\alpha(z)$ for every $h, z \in Y$. Moreover,

$$d_\alpha(h) \le \rho(\alpha)|\mathbf{v}|_{\sup}\|h\|_Y \le \rho(\alpha)\sqrt{B_v}\|h\|_Y \le \max\{C_g, 1\}\sqrt{B_v}\|h\|_Y \text{ for } h \in Y. \tag{18}$$

The inequality can be verified briefly as follows

$$d_\alpha(h)^2 = \sum_{\lambda\in\Lambda} r_\alpha(|\kappa_\lambda|^2)^2 |\langle h, v_\lambda\rangle_Y|^2 \le \rho^2(\alpha)\sum_{\lambda\in\Lambda} |\langle h, v_\lambda\rangle_Y|^2 \le \rho^2(\alpha)|\mathbf{v}|^2_{\sup}\|h\|^2_Y.$$

As shown in [7], we have

Theorem 3.2. *Let Assumption C and* (2) *hold. If*

$$\alpha(\delta) \to 0 \text{ and } \delta L(\alpha(\delta)) \to 0 \text{ as } \delta \to 0^+,$$

then $\lim_{\delta\to 0^+} \|x^\delta_{\alpha(\delta)} - x^\dagger\|_X = 0$.

Remark 4. *Using Assumption C, we have* $L(\alpha) = \sqrt{g_\alpha(\mu)}\sqrt{\mu g_\alpha(\mu)} \le \sqrt{C_g}\sqrt{\ell(\alpha)}$. *Hence, if* $\delta\sqrt{\ell(\alpha)} \to 0$ *(condition in [21]) then* $\delta L(\delta) \to 0$ *(condition in [7]) as* $\delta \to 0$. *Thus, we employ a condition analogous to the one found in [7].*

The proof of Theorem 3.2 could be seen in [7, 21]. For the convenience of our readers, however, we will provide the main ideas of the proof here.

Proof. In line with (17), the regularization solution for equation (1) with noiseless data is given by:

$$x_\alpha := R_\alpha y = \sum_{\lambda\in\Lambda} \overline{\kappa}_\lambda g_\alpha\left(|\kappa_\lambda|^2\right)\langle y, v_\lambda\rangle_Y \widetilde{u}_\lambda. \tag{19}$$

The triangle inequality yields

$$\|x^\delta_\alpha - x^\dagger\|_X \le \|x^\delta_\alpha - x_\alpha\|_X + \|x_\alpha - x^\dagger\|_X. \tag{20}$$

For the first term on the right hand side, using (19) and (17) gives

$$\|x^\delta_\alpha - x_\alpha\|_X = \left\|\sum_{\lambda\in\Lambda} \overline{\kappa}_\lambda g_\alpha\left(|\kappa_\lambda|^2\right)\langle y^\delta - y, v_\lambda\rangle_Y \widetilde{u}_\lambda\right\|_X.$$

From (12) and (9), the inequality (10) yields

$$\begin{aligned}\|x^\delta_\alpha - x_\alpha\|_X &\le \frac{1}{\sqrt{A_u}}\left(\sum_{\lambda\in\Lambda}\sup_\lambda\left(|\kappa_\lambda|^2 g^2_\alpha\left(|\kappa_\lambda|^2\right)\right)\left|\langle y^\delta - y, v_\lambda\rangle_Y\right|^2\right)^{\frac{1}{2}}\\ &\le \frac{L(\alpha)}{\sqrt{A_u}}\left(\sum_{\lambda\in\Lambda}\left|\langle y^\delta - y, v_\lambda\rangle_Y\right|^2\right)^{\frac{1}{2}}\\ &\le \sqrt{\frac{B_v}{A_u}}L(\alpha)\|y^\delta - y\|_Y \le \sqrt{\frac{B_v}{A_u}}\delta L(\alpha).\end{aligned} \tag{21}$$

For the last term in (20), using (15), (19), it follows that

$$\begin{aligned}\|x_\alpha - x^\dagger\|_X &= \left\|\sum_{\lambda\in\Lambda}\left(|\kappa_\lambda|^2 g_\alpha\left(|\kappa_\lambda|^2\right) - 1\right)\langle x^\dagger, u_\lambda\rangle_X \widetilde{u}_\lambda\right\|_X\\ &\le \frac{1}{\sqrt{A_u}}\left(\sum_{\lambda\in\Lambda}\left|r_\alpha\left(|\kappa_\lambda|^2\right)\right|^2\left|\langle x^\dagger, u_\lambda\rangle_X\right|^2\right)^{\frac{1}{2}}.\end{aligned} \tag{22}$$

Here the function $r_\alpha(\mu)$ is defined in (5). From (21) and (22), using the Lebesgue dominated convergence theorem, we can obtain our result. □

3.2 Lower bound of worst case error

To assess the optimality of an approximation method, our initial thought is often to directly compute $\Delta_{\inf}(\mathbf{M}, \delta) := \inf_{\mathbf{R}} \Delta(\mathbf{M}, \delta, \mathbf{R})$, where $\mathbf{R} : Y \to X$ represents any approximation method. However, this calculation is infeasible because $\mathbf{R}$ can be any mapping-linear or nonlinear-that we simply cannot control.

Therefore, we need a different approach. For suggestions, we can draw upon concepts within an SVD framework (see, e.g., [10, 23, 30]). If we can identify a function $\Psi(\delta)$ such that $C\Psi(\delta) \le \Delta_{\inf}(M, \delta)$ (for some constant $C > 0$), and subsequently find an approximation method $R^* : Y \to X$ where $\Delta(\mathbf{M}, \delta, R^*) \le C'\Psi(\delta)$ (for another constant $C' > 0$), this establishes the following relationship:

$$C\Psi(\delta) \le \Delta_{\inf}(\mathbf{M}, \delta) \le \Delta(\mathbf{M}, \delta, R^*) \le C'\Psi(\delta).$$

In this scenario, we can prove that R^* is order optimal. Building on this idea, we will evaluate the optimality of the proposed regularization R_α by first finding a lower bound for the worst-case error. This will be crucial for proving the optimality of the DFD-based regularization method over the source set $\mathbf{M}_{\Phi,E}$ in later theorems. It also provides a basis for choosing appropriate regularization parameters.

Similar to [18, 30], we shall consider the computation of the worst-case error of the regularization operator $R_\alpha : Y \to X$ in the source set $\mathbf{M}_{\Phi,E}$ with the function Φ satisfying **Assumption A1.** Function $\Phi : (0, a^*) \to (0, \infty)$ is continuous and satisfies the following conditions

(i) $\lim_{\mu \to 0} \Phi(\mu) = 0$,

(ii) Function Φ is strictly increasing on $(0, a^*)$,

(iii) Function $\Theta : (0, \Phi(a^*)] \to (0, a^*\Phi(a^*)]$, given by $\Theta(\mu) = \mu\Phi^{-1}(\mu)$, is convex.

Here we denote $\Phi(a^*) = \lim_{\mu \to a^{*-}} \Phi(\mu)$. As demonstrated by the optimal approach in the SVD context, the function $\Theta(\mu) = \mu\Phi^{-1}(\mu)$ is crucial in optimal regularization theory for inverse problems. It helps determine the optimal regularization parameter α by linking it to the smoothness of the true solution (represented by Φ) and the problem's SVD structure (represented by μ). This function balances data fidelity ($\asymp \delta/\sqrt{\alpha}$) and solution regularity ($\asymp \sqrt{\Phi(\alpha)}$), often appearing in analyses of optimal convergence rates (see the proof of Theorem 3.4). Furthermore, Θ is used to represent lower bounds of regularizations, providing theoretical limits on the achievable accuracy. This function is extensively discussed in numerous papers on optimal regularization, such as those by [18, 30] and the references therein. Key properties of this function include its monotonicity and the identity $\Theta(\Phi(z)) = z\Phi(z)$, which highlights its fundamental relationship with the function Φ and the spectral parameter μ. Deriving from the results in an SVD context, we can infer that the lower bound we are looking for is of the form $\asymp \sqrt{\Theta^{-1}(C\delta^2)}$.

Even when utilizing tools as found within an SVD framework, it remains challenging to establish a lower bound result for the worst-case error. This difficulty arises because the frames do not possess characteristics analogous to the eigenvectors of the SVD system. While [7] achieved this by assuming the strong condition that the frame $\mathbf{u}$ is minimal (which consequently guarantees a biorthogonal sequence $\widetilde{\mathbf{u}} = (\widetilde{u}_\lambda)_\lambda$ with $\langle u_\lambda, \widetilde{u}_\nu \rangle = \delta_{\lambda\nu}$, for all $\lambda, \nu \in \Lambda$), our paper endeavors to reduce the restrictiveness of this condition. In fact, we define a new kind of frames

Definition 6. *Let $(u_\lambda, v_\lambda, \kappa_\lambda)$ be a DFD frame of $\mathbf{K} : X \to Y$. For $m_*, m^* \in \mathbb{R}$ such that $0 < m_* \le m^*$, we denote $I_{m_*,m^*} = \{u_\lambda : |\kappa_\lambda| < m_* \text{ or } |\kappa_\lambda| > m^*\}$. If there is a constant $Q \ge 1$ such that $I^\perp_{|\kappa_\lambda|,Q|\kappa_\lambda|} \ne \{0\}$ for every $\lambda \in \Lambda$ then we say that the frame $\mathbf{u} = (u_\lambda)_{\lambda\in\Lambda}$ is* DFD quasi minimal *(with respect to Q).*

Remark 5. *(i) If $\mathbf{u}$ is minimal then $\mathbf{u}$ is DFD quasi minimal with respect to $Q = 1$. In fact, for every $\lambda_0 \in \Lambda$, we have $u_{\lambda_0} \notin \overline{\operatorname{span}\{u_\lambda : \lambda \ne \lambda_0\}}$. Choose $Q = 1$, $I_{|\kappa_{\lambda_0}|,|\kappa_{\lambda_0}|} = \{u_\lambda : |\kappa_\lambda| \ne |\kappa_{\lambda_0}|\} \subset \{u_\lambda : \lambda \ne \lambda_0\}$. Hence $I^\perp_{|\kappa_{\lambda_0}|,|\kappa_{\lambda_0}|} \supset \{u_\lambda : \lambda \ne \lambda_0\}^\perp \ne \{0\}$.*

(ii) There are many frames that are DFD quasi minimal but not minimal. Later in the paper, we will demonstrate such an example. Nevertheless, for the reader's convenience, we also offer a straightforward example here. We choose $X = Y = \ell^2(\mathbb{N})$, $e_k = (\delta_{jk})_{j\in\mathbb{N}}$. For $x = (x_j)_{j\in\mathbb{N}}$ we denote $\mathbf{K}(x) = \left(\frac{x_j}{\sqrt{j+1}}\right)_{j\in\mathbb{N}}$. Choose

$$\mathbf{u} = (e_0, e_0, e_1, e_1, e_2, e_2, \ldots),$$
$$\mathbf{v} = (e_0, e_0, e_1, e_1, e_2, e_2, \ldots),$$
$$\boldsymbol{\kappa} = \left(1, 1, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \ldots\right).$$

It is clear that $\mathbf{u}$ is not minimal. Moreover, $I^\perp_{(j+1)^{-1/2},(j+1)^{-1/2}} \supset \{e_j\}$ for all $j \in \mathbb{N}$. Hence, $\mathbf{u}$ is DFD quasi minimal with respect to $Q = 1$.

From the tools previously mentioned, we can derive a lower bound result that is comparable to the one in SVD theory [18, 30]. In fact, we have

Theorem 3.3. *Let $\delta_0 > 0$, $\delta \in (0, \delta_0)$, $\beta \in (0, 1), Q \ge 1$, let $(\mathbf{u}, \mathbf{v}, \boldsymbol{\kappa})$ be a DFD of $\mathbf{K}$ and let the source sets $\mathbf{M}_{\Phi,E}$ define by (6) . Put*

$$D_{\lambda,\beta} = [\delta^*_\lambda, \beta^{-1}\delta^*_\lambda], \quad \text{where} \quad \delta^*_\lambda = |\mathbf{v}_{\inf}|^{-1} Q E \sqrt{|\kappa_\lambda|^2 \Phi(|\kappa_\lambda|^2)}. \tag{23}$$

Assume that

(a) $\inf_{\lambda\in\Lambda} |\kappa_\lambda| = 0$,

(b) $\mathbf{u}$ is DFD quasi minimal with respect to Q.

If $\delta \in \bigcup_{\lambda\in\Lambda} D_{\lambda,\beta}$ then

$$\inf_{\mathbf{R}} \Delta(\mathbf{M}_{\Phi,E}, \delta, \mathbf{R}) \ge |\mathbf{u}|^{-1}_{\sup} E \sqrt{\Theta^{-1}\left(\frac{\beta^2 |\mathbf{v}|^2_{\inf} \delta^2}{Q^2 E^2}\right)}. \tag{24}$$

In addition, if Θ satisfies the condition

$$\inf_{0<\mu<a^*\Phi(a^*)} \frac{\Theta^{-1}(c\mu)}{\Theta^{-1}(\mu)} := \eta(c) > 0 \text{ for every } c \in (0,1), \tag{25}$$

then

$$\inf_{\mathbf{R}} \Delta(\mathbf{M}_{\Phi,E}, \delta, \mathbf{R}) \geq \sqrt{\eta\left(\frac{\beta^2}{Q^2}\right)} |\mathbf{u}|_{\sup}^{-1} E \sqrt{\Theta^{-1}\left(\frac{|\mathbf{v}|_{\inf}^2 \delta^2}{E^2}\right)}. \tag{26}$$

Especially, if $(0, \delta_0] \subset \bigcup_{\lambda\in\Lambda} D_{\lambda,\beta}$ then (24) holds for every $0 < \delta \leq \delta_0$.

The proof of this theorem involves many technical details and is quite lengthy, so we are moving it to Section 5. This theorem serves as a criterion for determining the order optimality of regularizations, so we will have a bit more commentary on it.

Remark 6. *(i) The condition that the system $\{u_\lambda\}$ is DFD quasi minimal is essential in the proof of the theorem. The investigation of the lower bound when $\{u_\lambda\}$ is not DFD quasi minimal is a worthy topic of study.*

(ii) To show that an approximation method $\mathbf{R} : Y \to X$ is order-optimal, we only need to verify that

$$\Delta(\mathbf{M}_{\Phi,E}, \delta, \mathbf{R}) \leq CE\sqrt{\Theta^{-1}\left(\beta^2 |\mathbf{v}|_{\inf}^2 \delta^2 / Q^2 E^2\right)}.$$

(iii) In fact, we can prove that $\inf_{\mathbf{R}} \Delta(\mathbf{M}_{\Phi,E}, \delta, \mathbf{R}) \geq \sqrt{B_u^{-1}} E \sqrt{\Theta^{-1}\left(\beta^2 A_v \delta^2 / Q^2 E^2\right)}$ for all B_u, A_v satisfy (12), (13). Since

$$|\mathbf{u}|_{\sup}^{-1} E \sqrt{\Theta^{-1}\left(\beta^2 |\mathbf{v}|_{\inf}^2 \delta^2 / Q^2 E^2\right)} \geq \sqrt{B_u^{-1}} E \sqrt{\Theta^{-1}\left(\beta^2 A_v \delta^2 / Q^2 E^2\right)},$$

our lower bound is better.

(iv) In [7], to obtain the lower bound of the worst case error, the authors choose $\delta = \delta_\lambda = \sqrt{A_v^{-1}} E \kappa_\lambda^{2\nu+1}$. The case that the mentioned paper examines corresponds to considering the source function $\Phi(\mu) = \mu^{2\nu}$, $Q = 1$. In this case, we have $\mu\Phi(\mu) = \mu^{2\nu+1}$ and $\delta_\lambda = \sqrt{A_v^{-1}} E \sqrt{|\kappa_\lambda|^2 \Phi(|\kappa_\lambda|^2)}$. For $\beta = \sqrt{A_v}/|\mathbf{v}|_{\inf}$, since $\sqrt{A_v} \leq |\mathbf{v}|_{\inf}$, we have $0 < \beta \leq 1$ and $\delta_\lambda^* \leq \delta_\lambda \leq \beta^{-1}\delta_\lambda^*$ which gives $\delta_\lambda \in \bigcup_{\lambda\in\Lambda} D_{\lambda,\beta}$. Hence, the inequality (26) hold for the chosen sequence $(\delta_\lambda)_{\lambda\in\Lambda}$.

(v) As shown in classical optimal regularization theory ([30, 33]), the optimal property is not true if the singular values of the operator $\mathbf{K}$ are too sparse, e.g., $\lim_{n\to\infty} \sigma_{n+1}/\sigma_n = 0$. The distribution of the singular values affects the classification of the optimization types. Similarly, the optimal result depends on the distribution of δ_λ^*. In fact, we have $(0, \delta_0) \subset \bigcup_{\beta>0} \bigcup_{\lambda\in\Lambda} D_{\lambda,\beta}$. If $(0, \delta_0) \not\subset \bigcup_{\lambda\in\Lambda} D_{\lambda,\beta}$ for every $\beta > 0$ then the distribution of δ_λ^* is very sparse. In this case, the lower bound may be valid for only some subsequences of δ_λ.

(vi) Note that, in the case of Hölder-type source condition, i.e., $\Phi(\mu) = \mu^{2\nu}$, $\mu, \nu > 0$, then $\Theta^{-1}(\mu) = \mu^{\frac{2\nu}{2\nu+1}}$. For $c \in (0,1)$, we have

$$\frac{\Theta^{-1}(c\mu)}{\Theta^{-1}(\mu)} = c^{\frac{2\nu}{2\nu+1}} > 0,$$

i.e., the condition (25) *holds in this case. Specifically, we have* $\Theta^{-1}\left(|\mathbf{v}|_{\inf}\delta^2/E^2\right) = |\mathbf{v}|_{\inf}^{\frac{2\nu}{2\nu+1}}\delta^{\frac{4\nu}{2\nu+1}}E^{\frac{-4\nu}{2\nu+1}}$. *So we get that*

$$\inf_{\mathbf{R}}\Delta(\mathbf{M}_{\Phi,E},\delta,\mathbf{R}) \geq \left(\frac{\beta}{Q}\right)^{\frac{2\nu}{2\nu+1}}\frac{|\mathbf{v}|_{\inf}^{\frac{\nu}{2\nu+1}}}{|\mathbf{u}|_{\sup}}\delta^{\frac{2\nu}{2\nu+1}}E^{\frac{1}{2\nu+1}}.$$

A similar lower bound is stated in Theorem 3.11 in [7] with δ *is in the sequence* $(\delta_\lambda)_{\lambda\in\Lambda}$ *as in Remark (ii).*

(vii) In some problems, we have the logarithmic source condition $\Phi(\mu) = (-\ln\mu)^{-p}$, $p > 0$. *Direct computation yields* $\Theta(\mu) = \mu e^{-\mu^{-1/(2p)}}$ *and* $\sqrt{\Theta^{-1}(\mu)} = \Phi(\mu)(1+o(1))$ *(see, e.g., [18]). So*

$$\inf_{\mathbf{R}}\Delta(\mathbf{M}_{\Phi,E},\delta,\mathbf{R}) \geq \sqrt{B_u^{-1}}E\left(\ln\left(\frac{Q^2E^2}{|\mathbf{v}|_{\inf}^2\beta^2\delta^2}\right)\right)^{-p}(1+o(1)).$$

Since

$$\frac{\Theta^{-1}(c\mu)}{\Theta^{-1}(\mu)} \asymp \frac{\Phi^2(c\mu)}{\Phi^2(\mu)} \xrightarrow{\mu\to 0^+} 1$$

we can verify directly the condition (25) *of* Θ *to obtain a similar form of the inequality* (26).

(viii) If Φ *is concave then Lemma 5.2 implies that*

$$\frac{\Theta^{-1}(c\mu)}{\Theta^{-1}(\mu)} \geq \sqrt{c} \text{ for } c\in(0,1),$$

i.e., the condition (25) *holds.*

3.3 Convergence rate and a priori parameter choice

Returning to the main content of this article, to extend the results of Ebner and colleagues [7] from a polynomial source set to a more general source set, we consider the source function Φ as in the definition of the set $\mathbf{M}_{\Phi,E}$ defined in (6). Next, we investigate issues such as the lower bound of the worst-case error, convergence rate in both the choice of a priori and a-posteriori parameters.

Our initial focus is on the convergence rate of regularization when selecting a priori parameters. Achieving optimal order estimates necessitates the following assumptions for the source function Φ and the filter function g_α, which parallel those found within the SVD framework (see, e.g., [26, 33]).

Assumption A2. There are constants $\gamma_1, \gamma_2 > 0$ such that

(i) $\sup_{0\leq\mu<a^*}\left|\sqrt{\mu}g_\alpha(\mu)\right| \leq \frac{\gamma_1}{\sqrt{\alpha}}$,

(ii) $\sup_{0\leq\mu<a^*}|r_\alpha(\mu)|\sqrt{\Phi(\mu)} \leq \gamma_2\sqrt{\Phi(\alpha)}$.

From Assumption A2, we derive convergence rates which give order optimal estimates on the reconstruction error $\left\|x_\alpha^\delta - x^\dagger\right\|_X$.

Theorem 3.4. *Let $A_v \in (0, |\mathbf{v}|^2_{\inf})$. For $(\mathbf{u}, \mathbf{v}, \boldsymbol{\kappa})$ being a DFD of $\mathbf{K}$, with $\widetilde{\mathbf{u}}$ as a dual frame of $\mathbf{u}$ and $x^\dagger \in \mathbf{M}_{\Phi,E}$. In this case, if we choose the regularization parameter as*

$$\alpha(\delta) = \alpha^*\left(\delta, y^\delta\right) := \Phi^{-1} \circ \Theta^{-1}\left(A_v\delta^2/E^2\right), \tag{27}$$

then the following convergence rate result holds:

$$\left\|x^\delta_{\alpha(\delta)} - x^\dagger\right\|_X \leq \sqrt{A_u^{-1}A_v^{-1}}\left(\gamma_1\sqrt{B_v} + \gamma_2\sqrt{A_v}\right)E\sqrt{\Theta^{-1}\left(|\mathbf{v}|^2_{\inf}\delta^2/E^2\right)}, \tag{28}$$

where A_u, B_v are bounds of $\mathbf{u}$ and $\mathbf{v}$, respectively, and γ_1, γ_2 are constants as in Assumption A2. From the inequality (28) *we obtain*

$$\Delta(\mathbf{M}_{\Phi,E}, \delta, R_{\alpha(\delta)}) \leq \sqrt{A_u^{-1}A_v^{-1}}\left(\gamma_1\sqrt{B_v} + \gamma_2\sqrt{A_v}\right)E\sqrt{\Theta^{-1}\left(|\mathbf{v}|^2_{\inf}\delta^2/E^2\right)},$$

where R_α is defined in (17)*. Moreover,*

(a) If $\mathbf{u}$ is DFD quasi minimal and Θ satisfies the condition (25) *then $R_{\alpha(\delta)}$ is sequential order optimal.*

(b) For a fixed $\beta \in (0, 1)$, if $\mathbf{u}$ is DFD quasi minimal, Θ satisfies (25) *and $(0, \delta_0] \subset \bigcup_{\lambda\in\Lambda} D_{\lambda,\beta}$, then $R_{\alpha(\delta)}$ is uniform order optimal. Here $D_{\lambda,\beta}$ is defined in* (23)*.*

Remark 7. *(i) Note that, for $\Phi(\mu) = \mu^{2\nu}$ with $\nu > 0$, our result aligns with Theorem 3.8 in [7] and Theorem 2.5 in [21].*

(ii) In the case of polynomial and logarithmic sources, the concave condition of Φ can be relaxed.

(iii) Calculating the exact number $|\mathbf{v}|_{\inf}$ is not easy. Therefore, choosing A_v as in the theorem will make the calculation of $\alpha(\delta)$ more feasible. However, if A_v is small, the error will contain A_v^{-1} and so will be large. Therefore, A_v should be chosen such that $\beta|\mathbf{v}|_{\inf} \leq A_v \leq |\mathbf{v}|_{\inf}$.

Proof. From the triangle inequality, for x_α defined in (19), we have

$$\|x^\delta_\alpha - x^\dagger\|_X \leq \|x^\delta_\alpha - x_\alpha\|_X + \|x_\alpha - x^\dagger\|_X. \tag{29}$$

For the first term on the right hand side, using (21) and Assumption A2 (i) gives

$$\|x^\delta_\alpha - x_\alpha\|_X \leq \gamma_1\sqrt{B_vA_u^{-1}}\frac{\delta}{\sqrt{\alpha}}. \tag{30}$$

Denote $\omega_\lambda = \sqrt{\Phi(|\kappa_\lambda|^2)^{-1}}\langle x^\dagger, u_\lambda\rangle_X$. For the last term in (29), combining (15) and (19), we obtain

$$\begin{aligned}
\|x_\alpha - x^\dagger\|_X &\leq \frac{1}{\sqrt{A_u}}\left(\sum_{\lambda\in\Lambda}\left|1 - |\kappa_\lambda|^2 g_\alpha\left(|\kappa_\lambda|^2\right)\right|^2\left|\langle x^\dagger, u_\lambda\rangle_X\right|^2\right)^{\frac{1}{2}} \\
&\leq \frac{1}{\sqrt{A_u}}\left(\sum_{\lambda\in\Lambda}\left|1 - |\kappa_\lambda|^2 g_\alpha\left(|\kappa_\lambda|^2\right)\right|^2\left|\sqrt{\Phi\left(|\kappa_\lambda|^2\right)}\omega_\lambda\right|^2\right)^{\frac{1}{2}} \\
&\leq \frac{\gamma_2}{\sqrt{A_u}}\sqrt{\Phi(\alpha)}\left(\sum_{\lambda\in\Lambda}|\omega_\lambda|^2\right)^{\frac{1}{2}} \leq \frac{\gamma_2}{\sqrt{A_u}}\sqrt{\Phi(\alpha)}E.
\end{aligned} \tag{31}$$

The third line is obtained from Assumption A2 (ii) and the condition $x^\dagger \in \mathbf{M}_{\Phi,E}$.

Combining (30) and (31) yields

$$\|x_\alpha^\delta - x^\dagger\|_X \le \gamma_1 \sqrt{B_v A_u^{-1}} \frac{\delta}{\sqrt{\alpha}} + \frac{\gamma_2}{\sqrt{A_u}} \sqrt{\Phi(\alpha)} E.$$

Based on the parameter choice (27), the regularization parameter is selected as $\alpha = \alpha^*\left(\delta, y^\delta\right) = \Phi^{-1} \circ \Theta^{-1}\left(A_v \delta^2/E^2\right)$. This implies that $\Phi(\alpha) = \Theta^{-1}\left(A_v\delta^2/E^2\right)$. From the latter equality, it follows that $\Theta(\Phi(\alpha)) = A_v\delta^2/E^2$, which yields $\delta^2 = A_v^{-1}E^2\Theta(\Phi(\alpha))$. In combination with Assumption A1 (iii), we obtain that

$$\frac{\delta}{\sqrt{\alpha}} = \sqrt{\frac{\delta^2}{\alpha}} = \sqrt{\frac{A_v^{-1}E^2\Theta(\Phi(\alpha))}{\alpha}} = \sqrt{A_v^{-1}E^2\Phi(\alpha)} = E\sqrt{A_v^{-1}\Theta^{-1}\left(A_v\delta^2/E^2\right)}.$$

Hence, we get that

$$\begin{aligned}
\|x_\alpha^\delta - x^\dagger\|_X &\le \gamma_1\sqrt{A_u^{-1}}\sqrt{A_v^{-1}}\sqrt{B_v}E\sqrt{\Theta^{-1}\left(A_v\delta^2/E^2\right)} + \frac{\gamma_2}{\sqrt{A_u}}E\sqrt{\Theta^{-1}\left(A_v\delta^2/E^2\right)} \\
&= \sqrt{A_u^{-1}}\sqrt{A_v^{-1}}\left(\gamma_1\sqrt{B_v} + \gamma_2\sqrt{A_v}\right)E\sqrt{\Theta^{-1}\left(A_v\delta^2/E^2\right)} \\
&\le \sqrt{A_u^{-1}}\sqrt{A_v^{-1}}\left(\gamma_1\sqrt{B_v} + \gamma_2\sqrt{A_v}\right)E\sqrt{\Theta^{-1}\left(|\mathbf{v}|_{\inf}^2\delta^2/E^2\right)}.
\end{aligned}$$

The above estimate completes the proof. □

3.4 A-posteriori parameter choice

In this subsection, we present the results of the discrepancy between the exact solution and the regularized solution. To achieve this, a preliminary idea to accomplish this is to apply the Morozov discrepancy principle, which involves considering the equation $\|\mathbf{K}x_\alpha^\delta - y^\delta\|_Y = \tau\delta$ where $\tau > 1$. However, calculating $\|\mathbf{K}x_\alpha^\delta - y^\delta\|_Y$ is computationally intensive because it requires knowing $\{\widetilde{u}_\lambda\}, \{v_\lambda\}$. Therefore, we will try to find an alternative formula. We have

$$\|\mathbf{K}x_\alpha^\delta - y^\delta\|_Y^2 = \|\mathbf{K}x_\alpha^\delta - P_{\overline{\mathrm{ran}\mathbf{K}}}y^\delta\|_Y^2 + \|P_{\overline{\mathrm{ran}\mathbf{K}}^\perp}y^\delta\|_Y^2.$$

We know that $\|P_{\overline{\mathrm{ran}\mathbf{K}}^\perp}y^\delta\|_Y \le \|y^\delta - y\|_Y \le \delta$ and

$$|\mathbf{v}|_{\sup}^{-1}\sum_{\lambda\in\Lambda}|\langle\mathbf{K}x_\alpha^\delta - P_{\overline{\mathrm{ran}\mathbf{K}}}y^\delta, v_\lambda\rangle_Y|^2 \le \|\mathbf{K}x_\alpha^\delta - P_{\overline{\mathrm{ran}\mathbf{K}}}y^\delta\|_Y^2 \le |\mathbf{v}|_{\inf}^{-1}\sum_{\lambda\in\Lambda}|\langle\mathbf{K}x_\alpha^\delta - P_{\overline{\mathrm{ran}\mathbf{K}}}y^\delta, v_\lambda\rangle_Y|^2.$$

For better insight, we will examine the special case where the frame $\{\mathbf{u}_\lambda\}$ is minimal and satisfies $\langle u_\nu, \widetilde{u}_\lambda\rangle = \delta_{\nu\lambda}$ for all $\lambda, \nu \in \Lambda$. Using Lemma 5.1 (see Section 5) yields

$$\begin{aligned}
\langle\mathbf{K}x_\alpha^\delta - P_{\overline{\mathrm{ran}\mathbf{K}}}y^\delta, v_\lambda\rangle_Y &= \langle\mathbf{K}x_\alpha^\delta - y^\delta, v_\lambda\rangle_Y = \langle\mathbf{K}x_\alpha^\delta, v_\lambda\rangle_Y - \langle y^\delta, v_\lambda\rangle_Y \\
&= (|\kappa_\lambda|^2 g_\alpha(|\kappa_\lambda|^2) - 1)\langle y^\delta, v_\lambda\rangle_Y = -r_\alpha(|\kappa_\lambda|^2)\langle y^\delta, v_\lambda\rangle_Y.
\end{aligned}$$

Hence

$$\sum_{\lambda\in\Lambda}|\langle\mathbf{K}x_\alpha^\delta - P_{\overline{\mathrm{ran}\mathbf{K}}}y^\delta, v_\lambda\rangle_Y|^2 = \sum_{\lambda\in\Lambda}(r_\alpha(|\kappa_\lambda|^2))^2|\langle y^\delta, v_\lambda\rangle_Y|^2 = d_\alpha(y^\delta)^2.$$

From the above suggestions (see also [21]), we will use the expression $d_\alpha(y^\delta)$ (defined in Definition 5) to replace $\|\mathbf{K}x_\alpha^\delta - y^\delta\|_Y$, even if $\mathbf{u}$ is not minimal. Specifically, the regularization parameter α is selected based on the Morozov principle [26], which provides a criterion for choosing α by solving the equation

$$d_\alpha(y^\delta) = \tau\sqrt{B_v}\delta \text{ with } \quad \tau > 1. \tag{32}$$

Let's additionally assume that the function g_α satisfies the following Assumption B1.

Assumption B1. The function $g_\alpha : (0, a^*] \to (0, \infty)$ satisfies

(i) $\lim_{\alpha\to(a^*)^-} r_\alpha(\mu) = \rho(\mu) \geq \rho > 0$ for a $\rho > 0$, and for each $\mu \in [0, a^*)$,

(ii) $g_{\alpha_n}(\mu) \to g_\alpha(\mu)$ for $\alpha_n \to \alpha > 0$ and for every $\mu \in [0, a^*)$.

Theorem 3.5. *Let $\tau > 1, \delta > 0$, $\mathbf{K}x^\dagger = y \in Y$, $y \neq 0$ and Assumptions C and B1 hold. Assume that*

$$0 < \tau\sqrt{B_v}\delta < \rho\sqrt{A_v}\|P_{\overline{rank\mathbf{K}}}y^\delta\|_Y. \tag{33}$$

Then, there exists a constant $\alpha_D(\delta)$ such that the equation (32) holds. In addition, if $\tau > \max\{C_g, 1\}$ and we have the following assumptions:

(a) $r_\alpha(\mu) \neq 0$ for every $\alpha > 0, \mu \in (0, a^*]$,

(b) there are a $C_p > 0$ and an $\alpha_0 > 0$ such that

$$\ell(\alpha) \sup_{\mu\in(a,a^*)} (r_\alpha(\mu))^2\mu\Phi(\mu) \leq C_p \text{ for every } 0 < \alpha \leq \alpha_0$$

and

$$\lim_{\alpha\to 0^+} \ell(\alpha)(r_\alpha(\mu))^2\mu\Phi(\mu) = 0 \text{ for every } \mu \in (0, a^*),$$

then $\|x^\delta_{\alpha_D(\delta)} - x^\dagger\|_X \to 0$ as $\delta \to 0^+$. Here, as defined in Definition 5, $\ell(\alpha) = \sup_{\mu\in(0,a^*)} g_\alpha(\mu)$.

Remark 8. *(i) Since $\lim_{\delta\to 0} \tau\sqrt{B_v}\delta = 0$, $\lim_{\delta\to 0} \rho\sqrt{A_v}\|P_{\overline{rank\mathbf{K}}}y^\delta\|_Y = \|y\|_Y > 0$, the condition (33) holds for every δ small enough.*

(ii) For convenience of calculation, we can choose the parameter α such that $d_\alpha(y^\delta) \geq \tau'\sqrt{B_v}\delta$ for $\tau' > 1$. Putting $\tau = \sqrt{B_v^{-1}}d_\alpha(y^\delta)$, we obtain the equation $d_\alpha(y^\delta) = \tau\sqrt{B_v}\delta$ and $\tau > \tau' > 1$.

(iii) Numerous filters satisfy the assumption (a), particularly those for which the filter function g_α strictly decreases as the variable α increases (e.g., Tikhonov and Landweber filters). In stark contrast, truncation filters, unexpectedly, do not satisfy this condition. Although investigating this phenomenon is compelling, a comprehensive exploration lies outside the purview of the current discussion.

(iv) To obtain optimal results under a polynomial source condition, the authors in [21] made use of assumptions

$$|r_\alpha(\mu)| \leq C\frac{\alpha^{\nu+1/2}}{\mu^{\nu+1/2}},\ \ell(\alpha) \leq c\alpha^{-1},\ \Phi(\mu) = \mu^{2\nu},\ c, C, \nu, \mu > 0. \tag{34}$$

In this case, we have

$$\ell(\alpha)(r_\alpha(\mu))^2\mu\Phi(\mu) \le cC\alpha^{-1}\frac{\alpha^{2\nu+1}}{\mu^{2\nu+1}}\mu\mu^{2\nu} = c\alpha^{2\nu} \text{ for all } \mu \in (0, a^*).$$

Hence, the assumption (b) holds. This shows that our condition is encompassed by the condition in [21]. Therefore, to prove the convergence (without requiring optimality) of the a-posteriori method, our relaxed condition can be used.

Proof. We know that under the conditions (i) - (ii) of Assumption B1, the function $d_\alpha(y)$ is continuous with respect to α and has the following results:

$$\lim_{\alpha\to 0} d_\alpha(y^\delta) = 0 \quad \text{and} \quad \lim_{\alpha\to(a^*)^-} d_\alpha(y^\delta) \ge \rho\left(\sum_\lambda \left|\langle y^\delta, v_\lambda\rangle_Y\right|^2\right)^{\frac{1}{2}}. \tag{35}$$

On the other hand,

$$0 < \tau\sqrt{B_v}\delta < \rho\sqrt{A_v}\|P_{\overline{\mathrm{ran}\mathbf{K}}}y^\delta\|_Y \le \rho\left(\sum_\lambda \left|\langle y^\delta, v_\lambda\rangle_Y\right|^2\right)^{\frac{1}{2}}. \tag{36}$$

Therefore, under Assumption B1 and (35), the equation (32) has a solution $\alpha = \alpha_D(\delta)$.

We consider the second part of the theorem. To this end, in view of Theorem 3.2, we have to verify

$$\lim_{\delta\to 0^+} \alpha_D(\delta) = 0, \ \lim_{\delta\to 0} \delta\sqrt{\ell(\alpha_D(\delta))} = 0.$$

Step 1. *Prove that $\alpha_D(\delta) \to 0$ as $\delta \to 0^+$.*

For a proof by contradiction, suppose there is a sequence $\{\delta_n\}$ such that $\delta_n \to 0$ and $\alpha_D(\delta_n) \to \alpha^* > 0$. We have

$$d_{\alpha_D(\delta)}(y) \le d_{\alpha_D(\delta)}(y^\delta) + d_{\alpha_D(\delta)}(y - y^\delta) \le C\delta.$$

Since $d_{\alpha_D(\delta_n)}(y^{\delta_n}) = \tau\sqrt{B_v}\delta_n$, letting $n \to \infty$, we obtain $d_{\alpha^*}(y) = 0$, i.e.,

$$\sum_{\lambda\in\Lambda} r_{\alpha^*}(|\kappa_\lambda|^2)^2|\langle y, v_\lambda\rangle_Y|^2 = 0.$$

From the assumption (a), one has $r_{\alpha^*}(|\kappa_\lambda|^2) \ne 0$ for all $\lambda \in \Lambda$, and thus $|\langle y, v_\lambda\rangle| = 0$ for every $\lambda \in \Lambda$. Since $y \in \mathrm{ran}\mathbf{K}$, we obtain $y = 0$, a contradiction. Hence, we have $\lim_{\delta\to 0^+} \alpha_D(\delta) = 0$.

Step 2. *Prove that* $\lim_{\delta\to 0^+} \delta\sqrt{\ell(\alpha_D(\delta))} = 0$.

In this step, we will write $\alpha_D(\delta)$ as α_D for brevity. Using Lemma 5.1 and the condition $x^\dagger \in \mathbf{M}_{\Phi,E}$ yields

$$\begin{aligned}
\sqrt{\ell(\alpha_D)}d_{\alpha_D}(y) &= \sum_{\lambda\in\Lambda} \ell(\alpha_D)(r_{\alpha_D}(|\kappa_\lambda|^2))^2|\langle y, v_\lambda\rangle_Y|^2 \\
&= \sum_{\lambda\in\Lambda} \ell(\alpha_D)(r_{\alpha_D}(|\kappa_\lambda|^2))^2|\kappa_\lambda|^2\Phi(|\kappa_\lambda|^2).[\Phi(|\kappa_\lambda|^2)]^{-1}\left|\frac{\langle y, v_\lambda\rangle_Y}{\kappa_\lambda}\right|^2 \\
&\le C_p\sum_{\lambda\in\Lambda}[\Phi(|\kappa_\lambda|^2)]^{-1}\left|\frac{\langle y, v_\lambda\rangle_Y}{\kappa_\lambda}\right|^2 = C_p\sum_{\lambda\in\Lambda}[\Phi(|\kappa_\lambda|^2)]^{-1}\left|\langle x^\dagger, u_\lambda\rangle_X\right|^2 \le C_pE.
\end{aligned}$$

Hence, using (b) and the Lebesgue dominated convergence theorem gives $\lim_{\delta\to 0^+} \ell(\alpha) d_{\alpha_D(\delta)}(y) = 0$. On the other hand, from the triangle inequality and (18), we obtain

$$\begin{aligned} d_{\alpha_D}(y) \geq d_{\alpha_D}(y^\delta) - d_{\alpha_D}(y^\delta - g) &\geq (\tau\sqrt{B_v} - \max\{C_g, 1\}|\mathbf{v}|_{\sup})\delta \\ &\geq (\tau - \max\{C_g, 1\})\sqrt{B_v}\delta. \end{aligned}$$

It follows that

$$0 \leq (\tau - \max\{C_g, 1\})\sqrt{B_v}\delta\sqrt{\ell(\alpha_D(\delta))} \leq \sqrt{\ell(\alpha_D(\delta))} d_{\alpha_D(\delta)}(y) \to 0 \text{ as } \delta \to 0^+$$

Hence $\lim_{\delta\to 0^+} \delta\sqrt{\ell(\alpha_D(\delta))} = 0$. Applying Theorem 3.2 and Remark 4 gives

$$\|x^\delta_{\alpha_D(\delta)} - x^\dagger\|_X \to 0 \text{ as } \delta \to 0^+.$$

□

Next, we will introduce some additional conditions, inspired by [33], page 75. These assumptions are crucial for establishing the theoretical results that follow, particularly regarding the optimality and convergence rates of our regularization method.

Assumption B2. The function $g_\alpha : (0, a^*] \to \mathbb{R}$ satisfies

(i) $g_\alpha(\mu) \geq 0$,

(ii) $0 \leq r_\alpha(\mu) \leq \frac{g_\alpha(\mu)}{\ell(\alpha)}$ with $\ell(\alpha) := \sup_{0\leq\mu\leq a^*} g_\alpha(\mu)$,

(iii) $\frac{\ell_*}{\alpha} \leq \ell(\alpha) \leq \frac{\ell^*}{\alpha}$ with constants $\ell_*, \ell^* > 0$.

Conditions on the filter function g_α are fundamental in the analysis of a-posteriori regularization methods. They are crucial for proving convergence rates and deriving error estimates. Similarly, in the work by Hubmer et al. [21], the authors also utilized a comparable condition (see Remark 9 (ii) below).

With these critical assumptions in place, particularly Assumption B2 regarding the filter function g_α, we can now establish key properties of our regularization approach. The following theorem provides a bound for the error $\|x^\delta_\alpha - x^\dagger\|_X$ and demonstrates the optimality of the a-posteriori choice rule for the regularization parameter $\alpha_D(\delta)$.

Theorem 3.6. *Let $(\mathbf{u}, \mathbf{v}, \boldsymbol{\kappa})$ be a DFD for the operator $\mathbf{K}$. With $\mathbf{M}_{\Phi,E}$ defined as in (6) and Assumption A1 simultaneously satisfying that x^δ_α is the approximate regularization solution as in (17) with g_α satisfying Assumption B2. Moreover, assuming Assumption B1, (36) is satisfied, and $\alpha_D(\delta)$ is chosen by the Morozov principle (32). If the function Φ is concave, then*

$$\|x^\delta_{\alpha_D(\delta)} - x^\dagger\|_X \leq |\mathbf{u}|^{-1}_{\inf}\sqrt{A_u^{-1}B_u}(\tau + 1)E\sqrt{\Theta^{-1}\left(|\mathbf{v}|^2_{\inf}\delta^2/E^2\right)}, \tag{37}$$

where A_u, B_u are bounds of frame $\mathbf{u}$. Hence, if, in addition, the frame $\{u_\lambda\}$ is DFD quasi optimal then *(i)* $R_{\alpha_D(\delta)}$ *is sequential order optimal over the set* $\mathbf{M}_{\Phi,E}$.

(ii) Moreover, if there exists $\beta \in (0,1)$ *and a* $\delta_0 > 0$ *such that* $(0, \delta_0] \subset \bigcup_{\lambda\in\Lambda} D_{\lambda,\beta}$ *then the regularization method* $R_{\alpha_D(\delta)}$ *is uniform order optimal. Here we recall that* $D_{\lambda,\beta}$ *defined in* (23).

Due to the specialized nature and length of this theorem's proof, we have moved it to Section 5 to maintain a coherent flow in the main presentation. Note that our result will be as in Theorem 2.7 in [21] if the source function Φ is replaced by polynomial function.

Remark 9. *(i) While Tikhonov filter is shown to satisfy Assumption B2 (as verified later in this paper), truncated filtering fails to meet this condition. Nevertheless, by adjusting the truncated filter to use* $g_\alpha(\mu) = \frac{1}{\max\{\alpha,\mu\}}$, *we can derive a filter that fulfills Assumption B2.*

(ii) In the framework of polynomial source functions, the conditions provided by [21] offer an alternative to our Assumption B2 (see (34)*). Notably, their assumption is expressed in terms of the source function, contrasting with our filter-centric approach. Further, a compelling direction for future work is to generalize the assumptions from [21] to cover the arbitrary source function case.*

(iii) Just as in Remark 7, we will illustrate that Assumption (b) in Theorem 3.5 relaxes Assumption B2 when the filter function $g_\alpha(\mu)$ *is strictly decreasing with respect to* μ. *This will prove that the condition in Theorem 3.5 is sensible if optimal rate is not demanded. In fact, if Assumption A2 (i) and Assumption B2 (ii), (iii) hold then*

$$\ell(\alpha)(r_\alpha(\mu))^2\mu\Phi(\mu) \le \ell(\alpha)\left(\frac{g_\alpha(\mu)}{\ell(\alpha)}\right)^2\mu\Phi(\mu) \le \gamma_1^2\Phi(\mu)\frac{1}{\alpha\ell(\alpha)} \le \gamma_1^2\Phi(a^*)\ell_*^{-1}.$$

With the limit in Assumption (b) in Theorem 3.5, it is more difficult to verify. However, if g_α *is a strictly decreasing function of the variable* μ *(e.g., Tikhonov and Landweber filters), then we can check this condition. In this case, put* $\mu_\alpha = g_\alpha^{-1}(\sqrt{\alpha})$ *then we have* $\lim_{\alpha\to0^+}\mu_\alpha = 0$. *Moreover*

$$\ell(\alpha)(r_\alpha(\mu))^2\mu\Phi(\mu) \le \begin{cases} \gamma_1^2\Phi(\mu_\alpha)\ell_*^{-1}, & \mu \ge \mu_\alpha, \\ \sqrt{\alpha}a^*\Phi(a^*)\ell_*^{-1}, & 0 < \mu < \mu_\alpha. \end{cases}$$

Hence, $\lim_{\alpha\to0^+}\ell(\alpha)(r_\alpha(\mu))^2\mu\Phi(\mu) = 0$ *for every* $\mu \in (0, a^*)$. *In the case where* $a^* = \infty$, *the evaluations remain feasible if we alter the condition in Theorem 3.5 slightly. Nevertheless, we will not delve into the specifics of this issue.*

4 Illustrative problems

In this Section, to provide a clear overview of the theory's application, we will consolidate all proofs into the final Section.

4.1 Statement of the problems

We give an example to illustrate our results in previous section. For $\gamma \in (0, 1]$, we consider the fractional heat equation

$$\partial_t^\gamma u(x,t) - u_{xx}(x,t) = 0, \ x \in \mathbb{R}, 0 < t < T. \tag{38}$$

Here

$$\partial_t^\gamma f(t) := \begin{cases} \frac{1}{\Gamma(1-\gamma)} \int_0^t (t-\tau)^{-\gamma} \frac{df}{d\tau}(\tau) d\tau, & 0 < \gamma < 1, \\ \frac{df}{dt}, & \gamma = 1, \end{cases}$$

and $\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt$ is the Gamma function. The equation (38) is vital across diverse fields. They are used in image processing (denoising, restoration), finance (option pricing, volatility), medical imaging (tomography), environmental modeling (pollutant source identification), and material science (parameter identification, non-destructive testing) (see, e.g., [6, 16, 22, 27, 32]).

In this part, we consider the backward problems which aim to find initial conditions from future observations. They are crucial for understanding system history but are ill-posed, meaning solutions can be unstable, non-existent, or non-unique. This ill-posedness, especially with fractional time derivatives modeling anomalous diffusion, presents significant inverse problem challenges that necessitate specialized regularization techniques. Specifically, we find the solution at the initial time $u(x, 0) = \theta_0(x)$ knowing that $u(x,t)$ satisfies the equation (38) subject to the final condition

$$u(x,T) = \theta_T(x), \ x \in \mathbb{R}. \tag{39}$$

Similar to the condition (2), we have to consider the problem (38)-(39) with the unknown exact data θ_T replaced by noisy data θ_T^δ satisfying

$$\|\theta_T^\delta - \theta_T\| \leq \delta.$$

Fractional backward problems have been a very active area of research in recent years (consult [4, 24, 34] and the references within it). Our paper's presentation of fractional backward problems will provide illustrations of the polynomial and non-polynomial source conditions previously discussed in the theoretical part.

4.2 The ill-posed nature and Operator form of the problem

We can use the Fourier transform to solve the equation (38). Here, we recall that the Fourier transform of a function $f(x)$, $x \in \mathbb{R}$, is defined as $\mathcal{F}f(\omega) = \int_{\mathbb{R}} f(x) e^{-i\omega . x} dx$. We denote the inner product of f, g in $L^2(\mathbb{R})$ by $\langle f, g \rangle = \int_{\mathbb{R}} f(x)\overline{g(x)} dx$ and the $L^2(\mathbb{R})$-norm by $\|f\| = \sqrt{\langle f, f \rangle}$. We also define the Hilbert scales by

$$H^p(\mathbb{R}) = \left\{ f \in L^2(\mathbb{R}) : \int_{\mathbb{R}} (1 + |\omega|^2)^p |\mathcal{F}f(\omega)|^2 d\omega < \infty \right\}.$$

Taking the Fourier transform of both sides of (38) yields

$$\partial_t^\gamma \mathcal{F}u(\omega,t) + \omega^2 \mathcal{F}u(\omega,t) = 0.$$

Solving the fractional differential equation (see, [15, 27]) gives $\mathcal{F}u(\omega,t) = E_{\gamma,1}(-\omega^2 t^\gamma)\mathcal{F}u(\omega,0)$ where $E_{\gamma,1}(z)$, $z \in \mathbb{C}$ is the Mittag-Leffler function

$$E_{\alpha,\beta}(z) = \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(\alpha n + \beta)}, E_{1,1}(z) = e^z.$$

Using the Fourier form of the solution formula, we get

$$\mathcal{F}\theta_T(\omega) = E_{\gamma,1}\left(-|\omega|^2 T^\gamma\right)\mathcal{F}\theta_0(\omega), \ \omega \in \mathbb{R}. \tag{40}$$

Equation (40) will be employed to discuss the ill-posedness of the problem. In fact, we can rewrite (40) as

$$\mathcal{F}\theta_0(\omega) = E_{\gamma,1}^{-1}\left(-|\omega|^2 T^\gamma\right)\mathcal{F}\theta_T(\omega), \ \omega \in \mathbb{R}. \tag{41}$$

We first consider the case $0 < \gamma < 1$. From [15, 31], for $0 < \gamma < 1$, there are constants $\underline{c}, \widetilde{c}$, $0 < \underline{c} < \widetilde{c}$, such that

$$\frac{\underline{c}}{1+|z|} \le |E_{\gamma,1}(z)| \le \frac{\widetilde{c}}{1+|z|}, \text{ for } z < 0. \tag{42}$$

Consequently, the factor $E_{\gamma,1}^{-1}(-|\omega|^2 T^\gamma)$ has a polynomial behavior as $\omega \to \infty$ of order $|\omega|^2$. This leads to instability as $|\omega|$ increases. Next, for $\gamma = 1$, the factor $E_{\gamma,1}^{-1}(-|\omega|^2 T^\gamma) = e^{|\omega|^2 T}$. The exponential growth of this factor leads to severe instability as $|\omega|$ is large. In conclusion, the problem is polynomially ill-posed for $\gamma \in (0,1)$ and exponentially ill-posed for $\gamma = 1$, with these two types of ill-posedness being fundamentally different in nature. Therefore, suitable DFD regularization methods need to be developed for each.

Throughout the rest of this subsection, we will introduce the operator form of the problem and explicitly show how v_λ is represented in terms of u_λ in this particular case. From (40), we can write $\mathbf{K}\theta_0 = \theta_T$ where $\mathbf{K} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$,

$$\mathbf{K}h = \mathcal{F}^{-1}((E_{\gamma,1}(-|\omega|^2 T^\gamma))\mathcal{F}h) \text{ for every } h \in L^2(\mathbb{R}). \tag{43}$$

By the definition of $\mathbf{K}$, we claim that $\mathbf{K}^* = \mathbf{K}$. In fact, we have

$$\begin{aligned}\langle \mathbf{K}h, v_0\rangle &= \frac{1}{2\pi}\langle \mathcal{F}\mathbf{K}h, \mathcal{F}v_0\rangle = \frac{1}{2\pi}\langle E_{\gamma,1}\left(-|\omega|^2 T^\gamma\right)\mathcal{F}h, \mathcal{F}v_0\rangle \\ &= \frac{1}{2\pi}\langle \mathcal{F}\theta_0, \overline{E_{\gamma,1}\left(-|\omega|^2 T^\gamma\right)}\mathcal{F}v_0\rangle = \frac{1}{2\pi}\langle \mathcal{F}h, E_{\gamma,1}\left(-|\omega|^2 T^\gamma\right)\mathcal{F}v_0\rangle \\ &= \langle h, \mathbf{K}v_0\rangle \text{ for every } h, v_0 \in L^2(\mathbb{R}).\end{aligned}$$

It follows that $\mathbf{K}^* = \mathbf{K}$. Now, we suppose $(u_\lambda, v_\lambda, \kappa_\lambda)_{\lambda\in\Lambda}$ is a DFD system of the operator $\mathbf{K}$ on $L^2(\mathbb{R})$. Since $\mathbf{K}^* = \mathbf{K}$, we obtain

$$\mathcal{F}\mathbf{K}^* v_\lambda = \mathcal{F}\mathbf{K}v_\lambda = [E_{\gamma,1}(-|\omega|^2 T^\gamma)]\mathcal{F}v_\lambda.$$

From the definition, we have $\mathbf{K}^* v_\lambda = \overline{\kappa}_\lambda u_\lambda$. Combining two latter equations yields

$$\mathcal{F}v_\lambda = \overline{\kappa}_\lambda [E_{\gamma,1}(-|\omega|^2 T^\gamma)]^{-1}\mathcal{F}u_\lambda. \tag{44}$$

4.3 The fractional backward problem

We consider the case $\gamma \in (0,1)$. Denote the $L^2(\mathbb{R})$-wavelet orthonormal basis by $\psi^{j,k}(x) = 2^{j/2}\psi(2^j x - k)$, $(j,k) \in \mathbb{Z}^2$, where ψ is a mother wavelet (see, e.g., [2, 3]). Put $\Lambda = \mathbb{Z}^2$, $\lambda = (\lambda_D, \lambda_T) \in \mathbb{Z}^2$. We consider the wavelet orthonormal basis u_λ in $L^2(\mathbb{R})$ in the form

$$u_\lambda(x) = \psi^\lambda(x),\ \forall \lambda = (\lambda_D, \lambda_T) \in \Lambda, x \in \mathbb{R}. \tag{45}$$

From here, we construct the DFD for the operator $\mathbf{K}$ using the following theorem.

Theorem 4.1. *Let $(u_\lambda)_{\lambda\in\Lambda}$ be defined as in (45) such that supp $(\mathcal{F}\psi) \subset \{\omega \in \mathbb{R} : a_u \le |\omega| \le b_u\}$ where a_u, b_u be positive constants. Then*

(a) $(u_\lambda, v_\lambda, \kappa_\lambda)_{\lambda\in\mathbb{Z}^2}$ be a DFD for $\mathbf{K}$ where

$$\kappa_\lambda = \begin{cases} 2^{-2\lambda_D}, & \text{for } \lambda_D \ge 1, \\ 1, & \text{for } \lambda_D < 1, \end{cases}$$

and $v_\lambda = \overline{\kappa}_\lambda \mathcal{F}^{-1}([E_{\gamma,1}(-|\omega|^2 T^\gamma)]^{-1}\mathcal{F}\psi^\lambda)$.

(b) $\theta_0 \in \mathbf{M}_{\Phi,E}$ for every $\theta_0 \in H^p(\mathbb{R})$ $(p \ge 0)$, where $\Phi(\mu) = \mu^{p/2}$ and E large enough.

(c) There exists a $\delta_0 > 0$ such that $(0,\delta_0] \subset \bigcup_{\lambda\in\Lambda} D_{\lambda,\beta}$ for $\beta = 2^{-(2+p)}$. Here $D_{\lambda,\beta}$ is defined in (23).

Remark 10. *(i) For polynomially ill-posed problems, the WVD system can be used well. We can see that in the tomography problems (see [7, 21]) and the fractional backward problem.*

(ii) The result (b) provides a sufficient condition for the function θ_0 to satisfy the DFD source condition. The function only needs to lie in the Hilbert scales $H^p(\mathbb{R})$.

From Theorem 4.1, we obtain the WVD of the operator $\mathbf{K}$. In particular, that is $(u_\lambda, v_\lambda, \kappa_\lambda)_{\lambda\in\Lambda}$. This allows us to regularize the inverse problem for the fractional heat equation with the source function $\Phi(\mu) = \mu^{p/2}$ and then $\mathbf{M}_{\Phi,E}$ becomes

$$\mathbf{M}_{\Phi,E} := \left\{ \theta_0 \in L^2(\mathbb{R}) : \sum_{\lambda\in\Lambda} [\Phi(|\kappa_\lambda|^2)]^{-1} |\langle \theta_0, u_\lambda\rangle|^2 \le E^2 \right\}.$$

To regularize the problem, we use the Tikhonov filter $g_\alpha(\lambda) = \frac{1}{\alpha+\lambda}$. The chosen $\{u_\lambda\}$ is tight, since it is orthonormal. So $\widetilde{u}_\lambda = u_\lambda$ and (17) can be rewritten as

$$u_{0\alpha}^\delta := R_\alpha(\theta_T^\delta) = \sum_{\lambda\in\Lambda} \frac{2^{-2\lambda_D}}{\alpha + 2^{-4\lambda_D}} \langle \theta_T^\delta, v_\lambda\rangle u_\lambda.$$

The source function Φ and the filter function g_α satisfy Assumptions C, A1, A2, B1, B2 (see the proof of Theorem 4.2). Hence, from Theorems 3.3, 3.4, 3.6,

we obtain the result for both a priori and a-posteriori parameters.

Theorem 4.2. *Let $(u_\lambda, v_\lambda, \kappa_\lambda)_{\lambda\in\Lambda}$ be as in Theorem 4.1 and $\theta_0 \in \mathbf{M}_{\Phi,E}$ for $\Phi(\mu) = \mu^{p/2}, p > 0$.*

(a) (a priori regularizarion) For $0 < p \le 4$, if we choose the regularization parameter as

$$\alpha := \alpha(\delta) = (\delta/E)^{\frac{2}{p+2}}$$

then $R_{\alpha(\delta)}$ is uniform order optimal and the following convergence rate result holds

$$\left\| u^\delta_{0\alpha(\delta)} - \theta_0 \right\| \le C\delta^{\frac{p}{p+2}} E^{\frac{2}{p+2}},$$

(b) (a-posteriori regularization) If $0 < p \le 2$, assume that α_D is chosen by the Morozov principle (32). Then $R_{\alpha_D(\delta)}$ is uniform order optimal over the set $\mathbf{M}_{\Phi,E}$, and

$$\| u^\delta_{0\alpha_D(\delta)} - \theta_0 \| \le C\delta^{\frac{p}{p+2}} E^{\frac{2}{p+2}}.$$

Remark 11. *If $p = 4\nu$, we obtain the error stated in [7].*

4.4 The classical backward problem

Put $B_N = \{\omega \in \mathbb{R} : \sqrt{N} \le |\omega| \le \sqrt{N+1}\}$, $N \in \mathbb{N}$. From here, we construct the DFD for the operator $\mathbf{K}$ using the following theorem.

Theorem 4.3. *Let $(u_\lambda)_{\lambda\in\Lambda}$ be defined as in (45) such that*

$$supp\,(\mathcal{F}\psi) \subset \{\omega \in \mathbb{R} : a_u \le |\omega| \le b_u\}$$

where a_u, b_u are positive constants and where $\lambda = (\lambda_D, \lambda_T)$. Put $u_{\lambda,N} = \mathcal{F}^{-1}(\mathbf{1}_{B_N}\mathcal{F}(u_\lambda))$. Then

(a) $(u_{\lambda,N}, v_{\lambda,N}, \kappa_{\lambda,N})_{\lambda\in\mathbb{Z}^2, N\in\mathbb{N}}$ is a DFD for $\mathbf{K}$ where

$$\kappa_{\lambda,N} = e^{-NT},\ \ N \in \mathbb{N}$$

and $v_\lambda = \kappa_{\lambda,N}\mathcal{F}^{-1}(e^{|\omega|^2 T} u_{\lambda,N})$ defined as in (44). Moreover, $u_{\lambda,N}$ is tight and DFD quasi minimal.

(b) $\theta_0 \in \mathbf{M}_{\Phi,E}$ for every $\theta_0 \in H^p(\mathbb{R})$, where $\Phi(\mu) = (-\ln \mu)^{-p}, p > 0$ and E large enough.

(c) There exists a $\delta_0 > 0$ such that $(0, \delta_0] \subset \bigcup_{(\lambda,N)\in\Lambda\times\mathbb{N}} D_{(\lambda,N),\beta}$ for $\beta = e^{-T}$. Here we recall $D_{(\lambda,N),\beta}$ is defined in (23).

Remark 12. *Using the classical wavelet system as in the previous section, we cannot find a suitable κ_λ. Therefore, it is necessary to construct a suitable DFD system. There are many ways to construct the system mentioned. However, we use a system that inherits the classical wavelet system as presented.*

From Theorem 4.3, we obtain the WVD of the operator $\mathbf{K}$. In particular, that is $(u_{\lambda,N}, v_{\lambda,N}, \kappa_{\lambda,N})_{\lambda\in\Lambda,N\in\mathbb{N}}$. This allows us to regularize the inverse problem for the fractional heat equation with the source function $\Phi(\lambda) = (-\ln\lambda)^{-p}$. Using the Tikhonov regularization for $g_\alpha(\mu) = \frac{1}{\alpha+\mu}$. Then the approximate solution can be written in the form (17). The chosen $\{u_\lambda\}$ is a tight frame, so $\widetilde{u}_\lambda = u_\lambda$ and (17) can be rewritten as

$$u^\delta_{0\alpha} := R_\alpha(\theta^T_\delta) = \sum_{N\in\mathbb{N}}\sum_{\lambda\in\Lambda} \frac{\kappa_{\lambda,N}}{\alpha+\kappa^2_{\lambda,N}} \langle\theta^\delta_T, v_{\lambda,N}\rangle u_{\lambda,N}.$$

The Assumptions A1, A2, B1, and B2 are shown to hold for the Tikhonov filter $g_\alpha(\mu) = \frac{1}{\alpha+\mu}$ and the source function $\Phi(\mu) = (-\ln\mu)^p$, as detailed in the final part of our paper. From Theorem 3.3, 3.4 and Theorem 3.6, we deduce the following consequence.

Theorem 4.4. *let $(u_{\lambda,N}, v_{\lambda,N}, \kappa_{\lambda,N})_{\lambda\in\Lambda,N\in\mathbb{N}}$ be as in Theorem 4.3 and $\theta_0 \in \mathbf{M}_{\Phi,E}$ where $\Phi(\mu) = (-\ln\mu)^p$, $p > 0$.*

(a) (a priori regularization) If we choose the regularization parameter $\alpha = \delta/E$ then $R_{\alpha(\delta)}$ is uniform order optimal and the following convergence rate result holds

$$\left\|u^\delta_{0\alpha} - \theta_0\right\| \le CE\left(-\ln(\delta/E)\right)^{-p}.$$

(b) (a-posteriori regularization) Let $0 < a^* < e^{-1}$ and $0 < p \le \ln(a^*)^{-1} - 1$. Assume that α_D is chosen by the Morozov principle (32). Then $R_{\alpha_D(\delta)}$ is uniform order optimal over the set $\mathbf{M}_{\Phi,E}$, and

$$\|u^\delta_{0\alpha_D(\delta)} - \theta_0\| \le CE\left(-\ln(\delta/E)\right)^{-p}.$$

Remark 13. *(i) While the system $\{u_{\lambda,N}\}$ is unlikely to satisfy the minimal property, its DFD quasi minimal nature allows the application of the optimal results from Theorems 3.3, 3.4, and 3.6.*

(ii) Condition $0 < a^* < e^{-1}$ can be mitigated, but since this is just an illustrative example, we will not go into the details.

4.5 Numerical simulation

4.5.1 Scheme of simulation

Step 1. Data Generation

We consider the following initial value problem:

$$\begin{cases} \partial^\gamma_t u(x,t) - u_{xx}(x,t) = 0, & x\in\mathbb{R}, \quad 0<t<T, \\ u(x,0) = \theta_0(x), & x\in\mathbb{R}. \end{cases}$$

From (41), it readily follows that the solution can be expressed as

$$\theta_T(x) = \mathcal{F}^{-1}\left(E_{\gamma,1}\left(-|\omega|^2T^\gamma\right)\cdot\mathcal{F}\theta_0(\omega)\right), \quad x\in\mathbb{R}.$$

In order to facilitate numerical computation, we restrict our attention to functions θ_0 exhibiting rapid decay to zero. Given this formulation, our primary focus lies in the efficient numerical evaluation of the Fourier transform. Let $A, B \in \mathbb{R}$ with $A < B$, and assume that $f(x) = 0$ for $x \notin (A, B)$. Define the spatial domain and parameters by considering the interval $[A, B]$, partitioned into a vector of L evenly spaced points

$$\mathbf{x} = [x_i]_{i=1}^{L},$$

where $\mathbf{x}$ is a vector of length L, with each element $x_i = A + (i-1)\Delta x$ for $i = \overline{1, L}$, and the spacing is given by

$$\Delta x = \frac{B - A}{L - 1}.$$

Due to the length of the paper, in the numerical examples, we restrict our consideration to the case $\gamma \in (0, 1)$, specifically $\gamma = 0.8$ and final time $T = 1$ in the tests. Compute the Fourier transform $\mathcal{F}\theta_0(\omega)$ and frequencies $w_i = \frac{2\pi}{B-A}\left(i - \frac{L}{2}\right)$ for $i = \overline{1, L}$ using the function `ft_by_fft` ([1]). Then, define the auxiliary array

$$\mathbf{z} = [z_i]_{i=1}^{L},$$

where $\mathbf{z}$ is a vector of length L, with each element $z_i = -|w_i|^2 T^{\gamma}$ for $i = \overline{1, L}$. Calculate the Mittag-Leffler values with MATLAB's `ml_matrix` ([13]) as

$$\mathbf{E}_{\text{val}} = \texttt{ml\_matrix}(\mathbf{z}, \gamma, 1),$$

where $\mathbf{E}_{\text{val}}$ is a vector of length L, with each element $E_{\text{val}}(i) = \texttt{ml\_matrix}(z_i, \gamma, 1)$. Applying a threshold ($10^{-10}$) to avoid near-zero values, and obtain their element-wise inverses

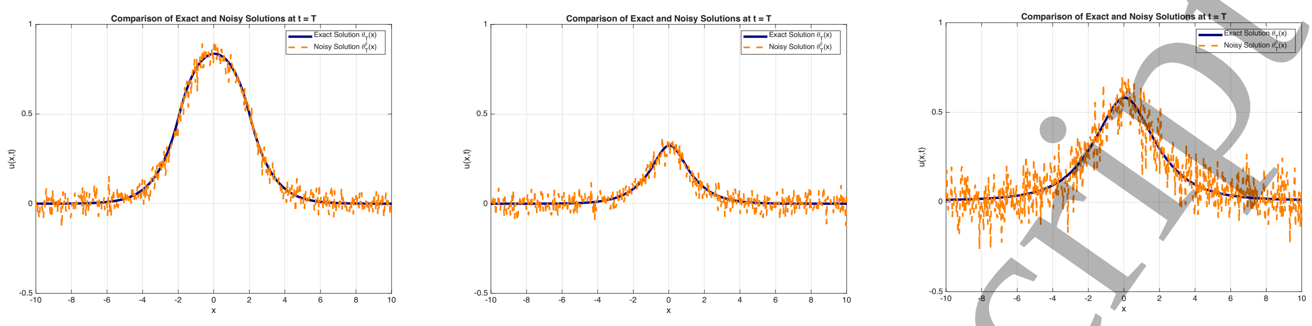
$$\mathbf{E}_{\text{val\_inv}} = \frac{1}{\mathbf{E}_{\text{val}}}.$$

The $\mathbf{E}_{\text{val\_inv}}$ is used to compute v_{λ_J} and v_λ in *Step 2*. The solution at time T is then computed by multiplying the characteristic function with the Mittag-Leffler term in the frequency domain,

$$\theta_T(x_i) = \mathcal{F}^{-1}\big(E_{\gamma,1}(-|\omega_i|^2 T^{\gamma}) \cdot \mathcal{F}\theta_0(\omega_i)\big), \quad i = \overline{1, L}$$

and applying the inverse Fourier transform with `ift_by_fft` ([1]) to obtain $\theta_T(x)$. Finally, adding Gaussian noise yields the data

$$\theta_T^{\delta}(x_i) = \theta_T(x_i) + \delta\texttt{*randn(L)}, \quad i = \overline{1, L}.$$



a. $\theta_0(x) = \chi_{[-2,2]}(x)$ b. $\theta_0(x) = \chi_{[-1,1]}(x)(1-|x|)$ c. $\theta_0(x) = \frac{1}{1+x^2}$

Figure 1: Comparison of the exact data $\theta_T(x)$ and its noisy counterpart $\theta_T^\delta(x)$ at $t = T$ with level noise $\delta = 0.01$, derived from different $\theta_0(x)$.

Step 2. Wavelet transform

Let J denote the initial resolution level and $J_{\max}$ the maximal resolution level employed in this analysis. We adopt the Daubechies wavelet of order 4 (`db4`) with the fixed filter length $M = 8$, chosen for its orthogonality and compact support characteristics, which render it particularly advantageous for Discrete Fourier Decomposition (DFD) within the framework of ill-posed inverse problems. The scaling function $\phi(x)$ and the wavelet function $\psi(x)$ are rigorously defined by the following two-scale relations:

$$\phi(x) = \sqrt{2}\sum_{m=0}^{M-1} h_m\,\phi(2x-m), \quad \psi(x) = \sqrt{2}\sum_{m=0}^{M-1} g_m\,\phi(2x-m),$$

where the sequence $\{h_m\}_{m=0}^{M-1}$ comprises the low-pass filter coefficients associated with $\phi(x)$, and $\{g_m\}_{m=0}^{M-1}$ denotes the corresponding high-pass filter coefficients, derived from $\{h_m\}$ via the relation

$$g_m = (-1)^m h_{M-1-m}, \quad m = \overline{0, M-1}.$$

The wavelet basis functions at scale j and position k are defined by

$$\phi^{j,k}(x) = 2^{\frac{j}{2}}\,\phi(2^j x - k), \quad \psi^{j,k}(x) = 2^{\frac{j}{2}}\,\psi(2^j x - k),$$

for $k = \overline{0, 2^j-1}$, $j = \overline{J, J_{\max}}$ where these functions form an orthonormal basis of the Hilbert space $L^2(\mathbb{R})$. These coefficients are available in standard wavelet toolboxes, such as MATLAB's `wfilters('db4')` or `wavefun` functions [3, 29].

According to Theorem 4.1, the regularized basis functions v_λ, parametrized by

$$\Lambda_1 := \{\lambda_J = (J,k) : k = \overline{0, 2^J-1}\}, \quad \Lambda_2 := \{\lambda = (j,k) : k = \overline{0, 2^j-1}, j = \overline{J, J_{\max}}\}$$

are constructed as

$$v_{\lambda_J} = \kappa(J) \cdot \mathcal{F}^{-1}\left(\left[E_{\gamma,1}\left(-|\omega|^2 T^\gamma\right)\right]^{-1} \cdot \mathcal{F}\phi^{\lambda_J}\right),$$

$$v_{\lambda} = \kappa(j) \cdot \mathcal{F}^{-1}\left(\left[E_{\gamma,1}\left(-|\omega|^2 T^\gamma\right)\right]^{-1} \cdot \mathcal{F}\psi^{\lambda}\right),$$

where κ_j are given in Theorem 4.1. Consequently, for each regularization parameter α, the reconstructed initial condition is expressed as

$$u_0^\delta(x) = \sum_{\lambda_J \in \Lambda_1} \frac{2^{-2J}}{\alpha + 2^{-4J}} \langle \theta_T^\delta, v_{\lambda_J}\rangle\, u_{\lambda_J}(x) + \sum_{\lambda \in \Lambda_2} \frac{2^{-2j}}{\alpha + 2^{-4j}} \langle \theta_T^\delta, v_{\lambda}\rangle\, u_{\lambda}(x),$$

where $u_{\lambda_J}(x) = \phi^{\lambda_J}(x), u_\lambda(x) = \psi^\lambda(x)$. Now, we proceed to the next step of optimal regularization.

Step 3. Optimal regularization

To determine the optimal regularization parameter α_{opt} in accordance with the Morozov discrepancy principle, we select α such that it satisfies the equation

$$d_\alpha(u_0^\delta) = \sum_{\lambda_J \in \Lambda_1} (r_\alpha(|\kappa_{\lambda_J}|^2))^2 |\langle \theta_T^\delta, v_\lambda\rangle|^2 + \sum_{\lambda \in \Lambda_2} (r_\alpha(|\kappa_{\lambda}|^2))^2 |\langle \theta_T^\delta, v_\lambda\rangle|^2 = \tau\delta,$$

where $\tau > 1$. Using the obtained α_{opt}, the reconstructed solution is recalculated as

$$u_{0,\text{opt}}^\delta(x) = \sum_{\lambda_J \in \Lambda_1} \frac{2^{-2J}}{\alpha_{\text{opt}} + 2^{-4J}} \langle \theta_T^\delta, v_{\lambda_J}\rangle\, u_{\lambda_J}(x) + \sum_{\lambda \in \Lambda_2} \frac{2^{-2j}}{\alpha_{\text{opt}} + 2^{-4j}} \langle \theta_T^\delta, v_{\lambda}\rangle\, u_{\lambda}(x).$$

Finally, the L^2-norm of the reconstruction error is computed and reported as

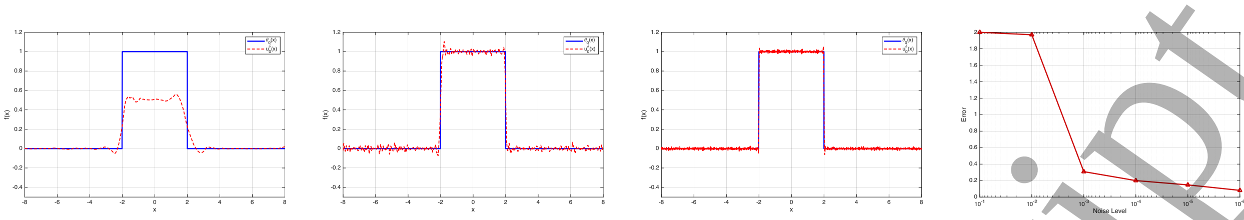
$$\text{Error}_{L^2} = \left(\sum_i \left(\theta_0(x_i) - u_{0,\text{opt}}^\delta(x_i)\right)^2 \Delta x \right)^{1/2},$$

providing a quantitative validation of the regularization method's effectiveness in addressing the fractional backward problem.

4.5.2 Examples

This subsection showcases three distinct numerical experiments, each employing a unique initial condition to test the robustness of our regularization framework. We explore a discontinuous and non-differentiable characteristic function, a continuous yet non-differentiable triangular function, and a smooth function that is both continuous and differentiable.

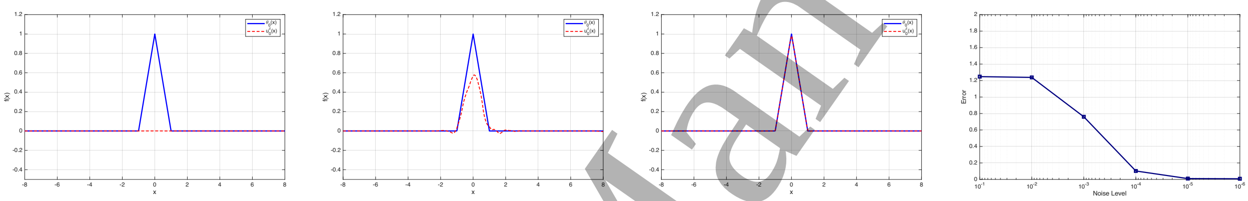
Test 1. We utilize $\theta_0(x) = \chi_{[-2,2]}(x)$, the characteristic function defined over the interval $[-2, 2]$, which belongs to $L^2(\mathbb{R})$. The effectiveness of the regularization is meticulously evaluated through the error profiles presented in Figures 2.



a. $\delta = 10^{-2}$ b. $\delta = 10^{-4}$ c. $\delta = 10^{-6}$ d. L^2 error

Figure 2: Comparison of the regularized solutions and exact solutions for the characteristic function at noise levels $\delta = 10^{-2}, 10^{-4}, 10^{-6}$, along with their respective error plots.

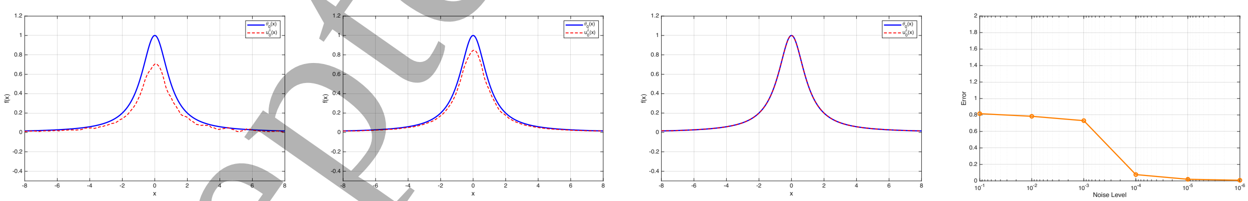
Test 2. We implement a triangular function expressed as $\theta_0(x) = \chi_{[-1,1]}(x)(1 - |x|)$ with the regularization precision quantified by the errors showcased in Figures 3.



a. $\delta = 10^{-2}$ a. $\delta = 10^{-4}$ a. $\delta = 10^{-6}$ d. L^2 error

Figure 3: Comparison of regularized solutions and exact solutions for the triangular function at $\delta = 10^{-2}, 10^{-4}, 10^{-6}$ along with their respective error plots.

Test 3. We employ a smooth function defined by $\theta_0(x) = 1/(1 + x^2)$, and assess the regularization performance through the error metrics illustrated in Figures 4.



a. $\delta = 10^{-2}$ a. $\delta = 10^{-4}$ a. $\delta = 10^{-6}$ d. L^2 error

Figure 4: Comparison of regularized solutions and exact solutions for the smooth function at $\delta = 10^{-2}, 10^{-4}, 10^{-6}$ along with their respective error plots.

Test	10^{-1}		10^{-2}		10^{-3}		10^{-4}		10^{-5}		10^{-6}	
	Abs	Re	Abs	Re	Abs	Re	Abs	Re	Abs	Re	Abs	Re
1	1.9992	0.9996	1.9696	0.9848	0.308684	0.154342	0.2003	0.10015	0.14793	0.073965	0.081661	0.040831
2	1.2476	1.5270	1.2379	1.5160	0.76031	0.9313	0.10276	0.1258	0.010454	0.0128	0.008298	0.01016
3	0.81575	0.6506	0.78415	0.6255	0.73114	0.5836	0.076911	0.06134	0.020003	0.01596	0.0078846	0.00629

Table 1: Comparison of the absolute (Abs) and relative (Re) L^2 errors of the regularized solutions for the three test cases at different noise levels δ.

Comment of the Tests: Table 1 has demonstrated that a result of the theory is plausible: the quality of signal recovery directly depends on the smoothness of the signal. The DFD method proved effective across all three scenarios, but the best results were achieved with smooth functions, with slightly poorer performance for functions with sharp edges or discontinuities. The selection of these three examples served as a relatively comprehensive test, showing that the proposed method not only works under ideal conditions (smooth functions) but is also capable of handling more complex signals, which are common in real-world applications.

On the other hand, when compared to the total variation method (see [34]), a method developed specifically for reconstructing discontinuous original functions, we find that the convergence results are comparable. Conversely, when comparing our results with the Tikhonov method (see [36]), the DFD method results are similar in the case of smooth functions but significantly improve the results at points of discontinuity and non-smoothness. These comparisons allow us to affirm that the DFD method improves upon previously published signal recovery results.

5 Proofs

5.1 Preliminary lemmas

Lemma 5.1. *Let $(\mathbf{u}, \mathbf{v}, \boldsymbol{\kappa})$ be the DFD for the operator $\mathbf{K}$. Then*

$$\langle x^\dagger, u_\lambda \rangle_X = \frac{1}{\kappa_\lambda} \langle y, v_\lambda \rangle_Y.$$

Assume in addition that $\mathbf{u}$ is minimal. Then we have

$$\begin{aligned}
\langle x^\delta_\alpha, u_\lambda \rangle_X &= \overline{\kappa}_\lambda g_\alpha(|\kappa_\lambda|^2) \langle y^\delta, v_\lambda \rangle_Y, \\
\langle \mathbf{K} x^\delta_\alpha, v_\lambda \rangle_Y &= |\kappa_\lambda|^2 g_\alpha(|\kappa_\lambda|^2) \langle y^\delta, v_\lambda \rangle_Y, \\
\langle \mathbf{K} x_\alpha, v_\lambda \rangle_Y &= |\kappa_\lambda|^2 g_\alpha(|\kappa_\lambda|^2) \langle y, v_\lambda \rangle_Y.
\end{aligned}$$

Proof. The first equality is verified in (14). If $\mathbf{u}$ is minimal, since $\langle \widetilde{u}_\nu, u_\lambda \rangle = \delta_{\nu\lambda}$ for every $\nu, \lambda \in \Lambda$, we have

$$\langle x^\delta_\alpha, u_\lambda \rangle_X = \left\langle \sum_{\nu \in \Lambda} \overline{\kappa}_\nu g_\alpha(|\kappa_\nu|^2) \langle y^\delta, v_\nu \rangle_Y \widetilde{u}_\nu, u_\lambda \right\rangle = \overline{\kappa}_\lambda g_\alpha(|\kappa_\lambda|^2) \langle y^\delta, v_\lambda \rangle_Y.$$

From (17), (19),(11) we obtain

$$\langle \mathbf{K}x_\alpha^\delta, v_\lambda\rangle_Y = \langle x_\alpha^\delta, \mathbf{K}^* v_\lambda\rangle_X = \kappa_\lambda \langle x_\alpha^\delta, u_\lambda\rangle_X = |\kappa_\lambda|^2 g_\alpha(|\kappa_\lambda|^2)\langle y^\delta, v_\lambda\rangle_Y,$$
$$\langle \mathbf{K}x_\alpha, v_\lambda\rangle_Y = \langle x_\alpha, \mathbf{K}^* v_\lambda\rangle_X = \kappa_\lambda \langle x_\alpha, u_\lambda\rangle_X = |\kappa_\lambda|^2 g_\alpha(|\kappa_\lambda|^2)\langle y, v_\lambda\rangle_Y.$$

□

Lemma 5.2. *Let $t \in (0,1)$ and let the function Φ satisfy Assumption A1 and $\Theta(\mu) = \mu\Phi^{-1}(\mu)$. If Φ is concave on $(0,a^*)$ then*
(i) $\Theta(t\Phi(\mu)) \le t^2\mu\Phi(\mu)$, for every $\mu \in (0,a^*)$,
(ii) $\Theta^{-1}(t^2 z) \ge t\Theta^{-1}(z)$ for every $z \in (0, a^*\Phi(a^*))$.

Proof. From the definition of the function Θ, we have $\Theta^{-1}(\mu\Phi(\mu)) = \Phi(\mu)$. The concavity of the function Φ combined with the condition $\lim_{\mu\to 0}\Phi(\mu) = 0$ implies that $t\Phi(\mu) \le \Phi(t\mu)$ for $t \in [0,1]$. Equivalently, $\Phi^{-1}(t\Phi(\mu)) \le \lambda t$. We also have, $\Theta(\mu) := \mu\Phi^{-1}(\mu)$, it follows that $\Theta(t\Phi(\mu)) \le t^2\mu\Phi(\mu)$. Hence, $\Theta^{-1}(t^2\mu\Phi(\mu)) \ge t\Phi(\mu) = t\Theta^{-1}(\mu\Phi(\mu))$. Putting $z = \mu\Phi(\mu)$, we obtain the desired inequality.

□

5.2 Proof of Theorem 3.3.

Proof. One of the commonly used methods to find a lower bound for the worst-case error is to compute the modulus of continuity

$$\Omega(\mathbf{M}, \delta) = \sup\{\|x\|_X \mid x \in \mathbf{M} \wedge \|\mathbf{K}x\|_Y \le \delta\}.$$

As is known (see, e.g. [10, 30]), we have

$$\Delta(\mathbf{M}_{\Phi,E}, \delta, \mathbf{R}) \ge \Omega(\mathbf{M}_{\Phi,E}, \delta).$$

The proof is divided into two steps.

Step 1. *Construct an element $x_\nu \in X, \nu \in \Lambda$ such that $x_\nu \in I^{\perp}_{|\kappa_\nu|, Q|\kappa_\nu|}$, $x_\nu \in \mathbf{M}_{\Phi,E}$ and $\|\mathbf{K}x_\nu\|_Y \le \delta$.*

Since $\delta \in \bigcup_{\lambda\in\Lambda} D_{\lambda,\beta}$ we can find an index $\nu \in \Lambda$ such that $\delta \in D_{\nu,\beta}$. For such the index $\nu \in \Lambda$, because that $I^{\perp}_{|\kappa_\nu|, Q|\kappa_\nu|} \ne \emptyset$, we can choose $z_\nu \in I^{\perp}_{|\kappa_\nu|, Q|\kappa_\nu|}$ such that $\|z_\nu\|_X = |\mathbf{u}|^{-1}_{\sup} E$. Setting $x_\nu := \sqrt{\Phi(|\kappa_\nu|^2)} z_\nu$, we obtain $x_\nu \in I^{\perp}_{|\kappa_\nu|, Q|\kappa_\nu|}$. We claim that $x_\nu \in \mathbf{M}_{\Phi,E}$ and that $\|\mathbf{K}x_\nu\|_Y \le \delta$.

We first verify that $x_\nu \in \mathbf{M}_{\Phi,E}$. In fact, since $z_\nu \perp u_\lambda$ for $|\kappa_\lambda| < |\kappa_\nu|$ or $|\kappa_\lambda| > Q|\kappa_\nu|$, we obtain

$$\langle x_\nu, u_\lambda\rangle_X = \begin{cases} \sqrt{\Phi(|\kappa_\nu|^2)}\,\langle z_\nu, u_\lambda\rangle_X, & \text{if } |\kappa_\nu| \le |\kappa_\lambda| \le Q|\kappa_\nu| \\ 0, & \text{else.} \end{cases}$$

Hence, using the inequality $\Phi(|\kappa_\nu|^2) \le \Phi(|\kappa_\lambda|^2)$ for $|\kappa_\lambda| \le |\kappa_\nu|$ gives

$$\sum_{\lambda\in\Lambda}[\Phi(|\kappa_\lambda|^2)]^{-1}|\left\langle x_\nu, u_\lambda\right\rangle_X|^2 = \sum_{|\kappa_\nu|\le|\kappa_\lambda|\le Q|\kappa_\nu|}[\Phi(|\kappa_\lambda|^2)]^{-1}\Phi(|\kappa_\nu|^2)|\left\langle z_\nu, u_\lambda\right\rangle_X|^2$$
$$\le \sum_{\lambda\in\Lambda}|\left\langle z_\nu, u_\lambda\right\rangle_X|^2 \le |\mathbf{u}|^2_{\sup}\|z_\nu\|^2_X \le E^2.$$

Hence, $x_\nu \in \mathbf{M}_{\Phi,E}$.

Next, we verify that $\|\mathbf{K}x_\nu\|_Y \le \delta$. In fact,

$$\begin{aligned}
\|\mathbf{K}x_\nu\|^2_Y &\le \frac{1}{|\mathbf{v}|^2_{\inf}}\sum_{\lambda\in\Lambda}|\langle \mathbf{K}x_\nu, v_\lambda\rangle_Y|^2 = \frac{1}{|\mathbf{v}|^2_{\inf}}\sum_{\lambda\in\Lambda}|\kappa_\lambda|^2\,|\langle x_\nu, u_\lambda\rangle_X|^2 \\
&= \frac{\Phi(|\kappa_\nu|^2)}{|\mathbf{v}|^2_{\inf}}\sum_{\lambda\in\Lambda}|\kappa_\lambda|^2\,|\langle z_\nu, u_\lambda\rangle_X|^2 = \frac{\Phi(|\kappa_\nu|^2)}{|\mathbf{v}|^2_{\inf}}\sum_{|\kappa_\nu|\le|\kappa_\lambda|\le Q|\kappa_\nu|}|\kappa_\lambda|^2\,|\langle z_\nu, u_\lambda\rangle_X|^2 \\
&\le \frac{1}{|\mathbf{v}|^2_{\inf}}Q^2\kappa_\nu^2\Phi\left(\kappa_\nu^2\right)\sum_{\lambda\in\Lambda}|\langle z_\nu, u_\lambda\rangle_X|^2 \le \frac{1}{|\mathbf{v}|^2_{\inf}}Q^2|\kappa_\nu|^2\Phi\left(\kappa_\nu^2\right)|\mathbf{u}|^2_{\sup}\|z_\nu\|^2_X \\
&= \frac{1}{|\mathbf{v}|^2_{\inf}}Q^2|\kappa_\nu|^2\Phi\left(|\kappa_\nu|^2\right)E^2.
\end{aligned}$$

From the definition of $D_{\nu,\beta}$ in (23) and the condition $\delta \in D_{\nu,\beta}$, we obtain

$$\delta^*_\nu \le \delta \le \beta^{-1}\delta^*_\nu \text{ where } \delta^*_\nu = |\mathbf{v}|^{-1}_{\inf}QE\sqrt{|\kappa_\nu|^2\Phi(|\kappa_\lambda|^2)} = |\mathbf{v}|^{-1}_{\inf}QE\sqrt{\Theta(\Phi(|\kappa_\lambda|^2))}. \quad (46)$$

Hence $\|\mathbf{K}x_\nu\|^2_Y \le \delta^{*2}_\nu \le \delta^2$.

Step 2. *Prove the lower bound for* $\Omega\left(\mathbf{M}_{\Phi,E}, \delta\right)$

Using the constructed x_ν, we will find a lower bound for $\Omega\left(\mathbf{M}_{\Phi,E}, \delta_\nu\right)$. Since $x_\nu \in \mathbf{M}_{\Phi,E}$, $\|\mathbf{K}x_\nu\| \le \delta$, the definition of $\Omega\left(\mathbf{M}_{\Phi,E}, \delta\right)$ yields

$$\Omega\left(\mathbf{M}_{\Phi,E}, \delta\right) \ge \|x_\nu\|_X = \sqrt{\Phi(|\kappa_\nu|^2)}\|z_\nu\|_X = \sqrt{\Phi(|\kappa_\nu|^2)}|\mathbf{u}|^{-1}_{\sup}E. \quad (47)$$

We will find a lower bound for $\sqrt{\Phi(|\kappa_\nu|^2)}$. From (46), we have

$$\Theta\left(\Phi\left(|\kappa_\nu|^2\right)\right) = |\mathbf{v}|^2_{\inf}\delta^{*2}_\nu/Q^2E^2 \ge |\mathbf{v}|^2_{\inf}\beta^2\delta^2/Q^2E^2.$$

Since the function Θ is increasing we obtain

$$\Phi\left(|\kappa_\nu|^2\right) \ge \Theta^{-1}\left(|\mathbf{v}|^2_{\inf}\beta^2\delta^2/Q^2E^2\right).$$

Hence

$$\Delta(\mathbf{M}_{\Phi,E}, \delta, \mathbf{R}) \ge \Omega(\mathbf{M}_{\Phi,E}, \delta) \ge E|\mathbf{u}|^{-1}_{\sup}\sqrt{\Theta^{-1}\left(\beta^2|\mathbf{v}|^2_{\inf}\delta^2/Q^2E^2\right)}.$$

Finally, if Θ satisfies the condition (25) then we obtain

$$\Theta^{-1}\left(\beta^2|\mathbf{v}|^2_{\inf}\delta^2/Q^2E^2\right) \ge \eta\left(\frac{\beta^2}{Q^2}\right)\Theta^{-1}\left(|\mathbf{v}|^2_{\inf}\delta^2/E^2\right).$$

Combining the two latter inequalities with (47) yields

$$\Delta(\mathbf{M}_{\Phi,E}, \delta, \mathbf{R}) \ge \Omega(\mathbf{M}_{\Phi,E}, \delta) \ge E|\mathbf{u}|^{-1}_{\sup}\sqrt{\eta\left(\frac{\beta^2}{Q^2}\right)}\sqrt{\Theta^{-1}\left(|\mathbf{v}|^2_{\inf}\delta^2/E^2\right)}.$$

$\square$

5.3 Proof of Theorem 3.6

First, with Assumption B2, using the ideas in [33], page 77, we obtain the following results

Lemma 5.3. *We denote*

$$e_\lambda = \overline{\kappa}_\lambda g_\alpha(|\kappa_\lambda|^2)\langle y^\delta, v_\lambda\rangle_Y - \frac{1}{\kappa_\lambda}\langle y, v_\lambda\rangle_Y.$$

Then $x_\alpha^\delta - x^\dagger = \sum_{\lambda\in\Lambda} e_\lambda \widetilde{u}_\lambda$. *Moreover, we have*

$$\sum_{\lambda\in\Lambda}|e_\lambda|^2 + \ell(\alpha)(d_\alpha^2(y^\delta) - \|y-y^\delta\|_v^2) \le \langle [I - \mathbf{K}^*\mathbf{K}g_\alpha(\mathbf{K}^*\mathbf{K})]\, x^\dagger, x^\dagger\rangle_u. \tag{48}$$

Proof of Lemma 5.3.

Proof. We recall that $r_\alpha(\mu) = 1 - \mu g_\alpha(\mu)$. We claim that

$$\begin{aligned}&|\langle x_\alpha^\delta - x^\dagger, u_\lambda\rangle_X|^2 + \ell(\alpha)|\langle |r_\alpha(|\kappa_\lambda|^2)\langle y^\delta, v_\lambda\rangle_Y|^2 - \ell(\alpha)|\langle y - y^\delta, v_\lambda\rangle_Y|^2\\ &\le (1 - |\kappa_\lambda|^2 g_\alpha(|\kappa_\lambda|^2))|\langle x^\dagger, u_\lambda\rangle_X|^2.\end{aligned} \tag{49}$$

We can write

$$\begin{aligned}|e_\lambda|^2 &= \frac{1}{|\kappa_\lambda|^2}\left||\kappa_\lambda|^2 g_\alpha(|\kappa_\lambda|^2)\langle y^\delta - y, v_\lambda\rangle_Y - r_\alpha(|\kappa_\lambda|^2)\langle y, v_\lambda\rangle_Y\right|^2\\ &= |\kappa_\lambda|^2 g_\alpha^2(|\kappa_\lambda|^2)|\langle y^\delta - y, v_\lambda\rangle_Y|^2 - 2g_\alpha(|\kappa_\lambda|^2)r_\alpha(|\kappa_\lambda|^2)\mathrm{Re}\ \langle y^\delta - y, v_\lambda\rangle_Y\overline{\langle y, v_\lambda\rangle}_Y\\ &+ \frac{r_\alpha^2(|\kappa_\lambda|^2)}{|\kappa_\lambda|^2}|\langle y, v_\lambda\rangle_Y|^2.\end{aligned}$$

On the other hand, we have

$$\begin{aligned}g_\alpha(|\kappa_\lambda|^2)r_\alpha(|\kappa_\lambda|^2)|\langle y^\delta, v_\lambda\rangle_Y|^2 &= g_\alpha(|\kappa_\lambda|^2)r_\alpha(|\kappa_\lambda|^2)|\langle y^\delta - y, v_\lambda\rangle_Y + \langle y, v_\lambda\rangle_Y|^2\\ &= g_\alpha(|\kappa_\lambda|^2)r_\alpha(|\kappa_\lambda|^2)(|\langle y^\delta - y, v_\lambda\rangle_Y|^2 + 2\mathrm{Re}\ \langle y^\delta - y, v_\lambda\rangle_Y\overline{\langle y, v_\lambda\rangle}_Y + |\langle y, v_\lambda\rangle_Y|^2).\end{aligned}$$

Combining two equalities yields

$$|e_\lambda|^2 + g_\alpha(|\kappa_\lambda|^2)r_\alpha(|\kappa_\lambda|^2)|\langle y^\delta, v_\lambda\rangle_Y|^2 = g_\alpha(|\kappa_\lambda|^2)|\langle y^\delta - y, v_\lambda\rangle_Y|^2 + \frac{r_\alpha(|\kappa_\lambda|^2)}{|\kappa_\lambda|^2}|\langle y, v_\lambda\rangle_Y|^2.$$

Since $\ell(\alpha) \ge g_\alpha(\mu) \ge \ell(\alpha)r_\alpha(\mu)$ we obtain

$$\begin{aligned}g_\alpha(|\kappa_\lambda|^2)r_\alpha(|\kappa_\lambda|^2)|\langle y^\delta, v_\lambda\rangle_Y|^2 &\ge \ell(\alpha)|r_\alpha(|\kappa_\lambda|^2)\langle y^\delta, v_\lambda\rangle_Y|^2,\\ g_\alpha(|\kappa_\lambda|^2)|\langle y^\delta - y, v_\lambda\rangle_Y|^2 &\le \ell(\alpha)|\langle y^\delta - y, v_\lambda\rangle_Y|^2.\end{aligned}$$

Hence the inequality (49) holds. Taking the sum of the inequalities (49) with respect to $\lambda \in \Lambda$, we get (48). □

Proof of Theorem 3.6.

Proof. Let $\alpha = \alpha_D$ be the regularization parameter chosen by (32). Using (9) and (2) yields

$$d_\alpha(y^\delta) = \tau\sqrt{B_v}\delta > \tau\sqrt{B_v}\|y - y^\delta\|_Y \geq \|y - y^\delta\|_v.$$

Hence, from Lemma 5.3 we obtain

$$\sum_{\lambda\in\Lambda} |e_\lambda|^2 \leq \langle [I - \mathbf{K}^*\mathbf{K}g_\alpha(\mathbf{K}^*\mathbf{K})]\, x^\dagger, x^\dagger\rangle_u^{\frac{1}{2}} = \|\, [r_\alpha(\mathbf{K}^*\mathbf{K})]^{\frac{1}{2}}\, x^\dagger\|_u. \tag{50}$$

Using the triangle inequality gives

$$d_{\alpha_D(\delta)}(y) \leq d_{\alpha_D(\delta)}(y^\delta) + d_{\alpha_D(\delta)}(y - y^\delta) \leq d_{\alpha_D(\delta)}(y^\delta) + \sqrt{B_v}\|y - y^\delta\|_Y^2.$$

Since $\alpha = \alpha_D$ is a solution of the equation (32), we deduce that

$$\left(\sum_{\lambda\in\Lambda} r^2_{\alpha_D}(|\kappa_\lambda|^2)|\langle y, v_\lambda\rangle|^2\right)^{1/2} \leq \sqrt{B_v}\,(\tau+1)\,\delta. \tag{51}$$

We denote

$$\omega = (\omega_\lambda)_{\lambda\in\Lambda} \text{ with } \omega_\lambda = [\Phi(|\kappa_\lambda|^2)]^{-1/2}\langle x^\dagger, u_\lambda\rangle_X.$$

Using the definition of $\mathbf{M}_{\Phi,E}$, we obtain $\|\omega\|_2 \leq E$. Computing directly yields

$$\begin{aligned}\Theta\left(\frac{\left\|[r_\alpha(\mathbf{K}^*\mathbf{K})]^{\frac{1}{2}}\, x^\dagger\right\|_u^2}{\|\omega\|_2^2}\right) &= \Theta\left(\frac{\sum_{\lambda\in\Lambda}\left|\langle [r_\alpha(|\kappa_\lambda|^2)]^{\frac{1}{2}}\, x^\dagger, u_\lambda\rangle_X\right|^2}{\|\omega\|_2^2}\right)\\ &= \Theta\left(\frac{\sum_{\lambda\in\Lambda} r_\alpha(|\kappa_\lambda|^2)\,\Phi(|\kappa_\lambda|^2)\,|\omega_\lambda|^2}{\|\omega\|_2^2}\right).\end{aligned}$$

Applying Lemma 5.2 for $t = r_\alpha(\mu) := 1 - \mu g_\alpha(\mu)$ gives

$$\Theta(r_\alpha(\mu)\Phi(\mu)) \leq \mu r_\alpha^2(\mu)\Phi(\mu). \tag{52}$$

Combining the convexity of Θ, the Jensen inequality, and the inequality (52), we obtain

$$\begin{aligned}&\Theta\left(\frac{\sum_{\lambda\in\Lambda} r_\alpha(|\kappa_\lambda|^2)\,\Phi(|\kappa_\lambda|^2)\,|\omega_\lambda|^2}{\|\omega\|_2^2}\right) \leq \frac{\sum_{\lambda\in\Lambda}\Theta\left(r_\alpha(|\kappa_\lambda|^2)\,\Phi(|\kappa_\lambda|^2)\right)|\omega_\lambda|^2}{\|\omega\|_2^2}\\ &\leq \frac{\sum_{\lambda\in\Lambda}|\kappa_\lambda|^2 r_\alpha^2(|\kappa_\lambda|^2)\,\Phi(|\kappa_\lambda|^2)\,|\omega_\lambda|^2}{\|\omega\|_2^2} \leq \frac{\sum_{\lambda\in\Lambda}|\kappa_\lambda|^2 r_\alpha^2(|\kappa_\lambda|^2)\,\Phi(|\kappa_\lambda|^2)\,|\omega_\lambda|^2}{\|\omega\|_2^2}.\end{aligned}$$

Using the latter results, the inequality (51) and the bound of frame $\mathbf{u}$, we can infer that

$$\Theta\left(\frac{\left\|[r_\alpha(\mathbf{K}^*\mathbf{K})]^{\frac{1}{2}}\, x^\dagger\right\|_u^2}{\|\omega\|_2^2}\right) \leq \frac{\sum_{\lambda\in\Lambda}\left|\langle \kappa_\lambda r_\alpha(|\kappa_\lambda|^2)\, x^\dagger, u_\lambda\rangle_X\right|^2}{\|\omega\|_2^2} \leq \frac{B_v(\tau+1)^2\delta^2}{\|\omega\|_2^2}. \tag{53}$$

From the inequality $B_v A_v^{-1} \geq 1$ and the definition of the source-set $\mathbf{M}_{\Phi,E}$, we deduce that $\sqrt{B_v A_v^{-1}}(\tau+1)E \geq \|\omega\|_2$. Using the monotonicity of Φ^{-1}, the relation $\Phi^{-1}(\lambda) = \frac{1}{\lambda}\Theta(\lambda)$ and the estimate (53), we obtain

$$\Phi^{-1}\left(\frac{\|r_\alpha(\mathbf{K}^*\mathbf{K})]^{\frac{1}{2}} x^\dagger\|_u^2}{A_v^{-1}B_v(\tau+1)^2E^2}\right) \leq \Phi^{-1}\left(\frac{\left\|[r_\alpha(\mathbf{K}^*\mathbf{K})]^{\frac{1}{2}} x^\dagger\right\|_u^2}{\|\omega\|_2^2}\right)$$

$$= \frac{\|\omega\|_2^2}{\left\|[r_\alpha(\mathbf{K}^*\mathbf{K})]^{1/2} x^\dagger\right\|_u^2}\Theta\left(\frac{\left\|[r_\alpha(\mathbf{K}^*\mathbf{K})]^{\frac{1}{2}} x^\dagger\right\|_u^2}{\|\omega\|_2^2}\right)$$

$$= \frac{\|\omega\|_2^2}{\left\|[r_\alpha(\mathbf{K}^*\mathbf{K})]^{1/2} x^\dagger\right\|_u^2}\frac{B_v(\tau+1)^2\delta^2}{\|\omega\|_2^2} = \frac{B_v(\tau+1)^2\delta^2}{\left\|[r_\alpha(\mathbf{K}^*\mathbf{K})]^{1/2} x^\dagger\right\|_u^2}.$$

Equivalently,

$$\Theta\left(\frac{\left\|[r_\alpha(\mathbf{K}^*\mathbf{K})]^{\frac{1}{2}} x^\dagger\right\|_u^2}{A_v^{-1}B_v(\tau+1)^2E^2}\right) \leq \frac{A_v\delta^2}{E^2}.$$

From here, it follows

$$\sum_{\lambda\in\Lambda}|e_\lambda|^2 \leq \left\|[r_\alpha(\mathbf{K}^*\mathbf{K})]^{\frac{1}{2}} x^\dagger\right\|_u^2 \leq A_v^{-1}B_v(\tau+1)^2E^2\Theta^{-1}\left(\frac{A_v\delta^2}{E^2}\right).$$

This estimate and (50) give us the result (37). From Lemma 5.3, we obtain

$$\begin{aligned}
\|x_\alpha^\delta - x^\dagger\|_X &= \left\|\sum_{\lambda\in\Lambda} e_\lambda \widetilde{u}_\lambda\right\|_X \leq |\mathbf{u}|_{\inf}^{-1}\left(\sum_{\lambda\in\Lambda}|e_\lambda|^2\right)^{1/2} \\
&\leq |\mathbf{u}|_{\inf}^{-1}\sqrt{A_u^{-1}B_u}(\tau+1)E\sqrt{\Theta^{-1}\left(A_v\delta^2/E^2\right)} \\
&\leq |\mathbf{u}|_{\inf}^{-1}\sqrt{A_u^{-1}B_u}(\tau+1)E\sqrt{\Theta^{-1}\left(|\mathbf{v}|_{\inf}^2\delta^2/E^2\right)}.
\end{aligned}$$

□

5.4 Proof of Theorem 4.1

Proof. (a) (**D1**) holds because $\{u_\lambda\}_{\lambda\in\Lambda}$ is an orthonormal basis for $L^2(\mathbb{R})$. (**D3**) is also established by the relationship between $\{v_\lambda\}$ and $\{u_\lambda\}$ shown in (44). Therefore, we only need to check (**D2**). For each $\lambda = (\lambda_D, \lambda_T) \in \Lambda$, we can verify directly that

$$supp\,(\mathcal{F}u_\lambda) \subset \left\{\omega \in \mathbb{R} : 2^{\lambda_D}a_u \leq |\omega| \leq 2^{\lambda_D}b_u\right\}.$$

From here, we deduce that $2^{2\lambda_D}a_u^2T^\gamma \leq |\omega|^2 T^\gamma \leq 2^{2\lambda_D}b_u^2T^\gamma$ for every $\omega \in supp\,(\mathcal{F}u_\lambda)$. Using the monotonicity property of the function $E_{\gamma,1}(z)$, we obtain

$$E_{\gamma,1}\left(-2^{2\lambda_D}b_u^2T^\gamma\right) \leq E_{\gamma,1}\left(-|\omega|^2T^\gamma\right) \leq E_{\gamma,1}\left(-2^{2\lambda_D}a_u^2T^\gamma\right).$$

This follows that

$$\kappa_\lambda E_{\gamma,1}^{-1}\left(-2^{2\lambda_D}a_u^2T^\gamma\right) \leq \left|\kappa_\lambda E_{\gamma,1}^{-1}\left(-|\omega|^2 T^\gamma\right)\right| \leq \kappa_\lambda E_{\gamma,1}^{-1}\left(-2^{2\lambda_D}b_u^2T^\gamma\right).$$

Moreover, from the inequality (42), and $\kappa_\lambda = 2^{-2\lambda_D}$, we can deduce

$$(1+a_u^2T^\gamma)/\widetilde{c} \leq \left|\kappa_\lambda E_{\gamma,1}^{-1}\left(-|\omega|^2 T^\gamma\right)\right| \leq (1+b_u^2T^\gamma)/\underline{c}. \tag{54}$$

We recall that $\{u_\lambda\}$ is a frame, this means that for all $\theta \in L^2(\mathbb{R})$,

$$A_u\|\theta\|^2 \leq \sum_{\lambda\in\Lambda}|\langle\theta, u_\lambda\rangle|^2 \leq B_u\|\theta\|^2.$$

Equivalently,

$$A_u\|\mathcal{F}\theta\|^2 \leq \sum_{\lambda\in\Lambda}|\langle\mathcal{F}\theta, \mathcal{F}u_\lambda\rangle|^2 \leq B_u\|\mathcal{F}\theta\|^2. \tag{55}$$

Taking any $\theta_T \in \mathrm{ran}\mathbf{K}$, we show that there are A_v, B_v satisfying

$$A_v\|\theta_T\|^2 \leq \sum_{\lambda\in\Lambda}|\langle\theta_T, v_\lambda\rangle|^2 \leq B_v\|\theta_T\|^2.$$

This is equivalent to proving

$$A_v\|\mathcal{F}\theta_T\|^2 \leq \sum_{\lambda\in\Lambda}|\langle\mathcal{F}\theta_T, \mathcal{F}v_\lambda\rangle|^2 \leq B_v\|\mathcal{F}\theta_T\|^2. \tag{56}$$

In fact, for every $\lambda \in \Lambda$, from (44), we have

$$|\langle\mathcal{F}\theta_T, \mathcal{F}v_\lambda\rangle|^2 = \left|\kappa_\lambda E_{\gamma,1}^{-1}\left(-|\omega|^2 T^\gamma\right)\right|^2 |\langle\mathcal{F}\theta_T, \mathcal{F}u_\lambda\rangle|^2.$$

Using the inequality (54), it follows that

$$\begin{aligned}\left((1+a_u^2T^\gamma)/\widetilde{c}\right)^2|\langle\mathcal{F}\theta_T, \mathcal{F}u_\lambda\rangle|^2 &\leq \left|\kappa_\lambda E_{\gamma,1}^{-1}\left(-|\omega|^2 T^\gamma\right)\right|^2|\langle\mathcal{F}\theta_T, \mathcal{F}u_\lambda\rangle|^2\\ &\leq \left((1+b_u^2T^\gamma)/\underline{c}\right)^2|\langle\mathcal{F}\theta_T, \mathcal{F}u_\lambda\rangle|^2,\end{aligned}$$

for every $\lambda \in \Lambda$. Hence,

$$\begin{aligned}\left((1+a_u^2T^\gamma)/\widetilde{c}\right)^2\sum_{\lambda\in\Lambda}|\langle\mathcal{F}\theta_T, \mathcal{F}u_\lambda\rangle|^2 &\leq \sum_{\lambda\in\Lambda}|\langle\mathcal{F}\theta_T, \mathcal{F}v_\lambda\rangle|^2\\ &\leq \left(1+b_u^2T^\gamma/\underline{c}\right)^2\sum_{\lambda\in\Lambda}|\langle\mathcal{F}\theta_T, \mathcal{F}u_\lambda\rangle|^2.\end{aligned}$$

Combining with (55), we get

$$\left((1+a_u^2T^\gamma)/\widetilde{c}\right)^2 A_u\|\mathcal{F}\theta_T\|^2 \leq \sum_{\lambda\in\Lambda}|\langle\mathcal{F}\theta_T, \mathcal{F}v_\lambda\rangle|^2 \leq \left((1+b_u^2T^\gamma)/\underline{c}\right)^2 B_u\|\mathcal{F}\theta_T\|^2.$$

Finally, (56) is proved for $A_v = \left((1+a_u^2T^\gamma)/\widetilde{c}\right)^2 A_u$ and $B_v = \left((1+b_u^2T^\gamma)/\underline{c}\right)^2 B_u$.

(b) We find the source condition for the solution θ_0. Naturally, we can assume that $\theta_0 \in H^p(\mathbb{R})$ for $p \geq 0$. Putting

$$G(\lambda_D) = \{\omega \in \mathbb{R} : 2^{\lambda_D} a_u \leq |\omega| \leq 2^{\lambda_D} b_u\}, \tag{57}$$

we note that $supp(u_\lambda) \subset G(\lambda_D)$. For $\lambda = (\lambda_D, \lambda_T)$, we can write

$$\begin{aligned}\langle \theta_0, u_\lambda \rangle &= \frac{1}{2\pi}\langle \mathcal{F}\theta_0, \mathcal{F}u_\lambda \rangle = \frac{1}{2\pi}\int_{\mathbb{R}} \mathcal{F}\theta_0(\xi)\overline{\mathcal{F}u_\lambda(\xi)}d\xi \\ &= \frac{1}{2\pi}\int_{\mathbb{R}} 1_{G(\lambda_D)}(\omega)\mathcal{F}\theta_0(\xi)\overline{\mathcal{F}u_\lambda(\xi)}d\xi.\end{aligned}$$

For $\lambda_D \geq 0$, using the Bessel inequality yields

$$\sum_{\lambda_T \in \mathbb{Z}} |\langle \theta_0, u_\lambda \rangle|^2 \leq \frac{1}{2\pi}\|1_{G(\lambda_D)}\mathcal{F}(\theta_0)\|^2 \leq 2^{-2p\lambda_D}\omega_{\lambda_D}^2 = \kappa_\lambda^p \omega_{\lambda_D}^2$$

where

$$w_{\lambda_D} = a_u^{-p}\|1_{G(\lambda_D)}(1+\omega^2)^p\mathcal{F}(\theta_0)\|.$$

Hence

$$\sum_{\lambda_D \in \mathbb{Z}_+}\sum_{\lambda_T \in \mathbb{Z}} |\kappa_\lambda|^{-p}|\langle \theta_0, u_\lambda \rangle|^2 \leq \sum_{\lambda_D \in \mathbb{Z}_+} w_{\lambda_D}^2 \leq C\|\theta_0\|_{H^p(\mathbb{R})}^2.$$

Here $\mathbb{Z}_+ = \{z \in \mathbb{Z} : z \geq 0\}$. For $\lambda_D < 0$, we have $\kappa_\lambda = 1$ and

$$\langle \theta_0, u_\lambda \rangle = \kappa_\lambda^{p/2}\langle \theta_0, u_\lambda \rangle.$$

Direct computations yields

$$\sum_{\lambda_D \in \mathbb{Z}_-}\sum_{\lambda_T \in \mathbb{Z}} \kappa_\lambda^{-p}|\langle \theta_0, u_\lambda \rangle|^2 \leq \sum_{\lambda \in \mathbb{Z}^2} |\langle \theta_0, u_\lambda \rangle|^2 \leq C\|\theta_0\|_{H^p(\mathbb{R})}^2.$$

Here $\mathbb{Z}_- = \{z \in \mathbb{Z} : z < 0\}$. So the function $\theta_0 \in \mathbf{M}_{\Phi,E}$ where $\Phi(\mu) = \mu^{p/2}$ and E is large enough.

(c) To obtain the order optimal result, we verify the conditions in Theorem 3.3. We have $\delta_\lambda^* = |\mathbf{v}|_{\inf}^{-1}E\sqrt{|\kappa_\lambda|^2\Phi(|\kappa_\lambda|^2)} = |\mathbf{v}|_{\inf}^{-1}E\sqrt{\kappa_\lambda^{2+p}} = |\mathbf{v}|_{\inf}^{-1}E2^{-\lambda_D(2+p)}$. Letting $0 < \delta < |\mathbf{v}|_{\inf}^{-1}E2^{-(2+p)}$, we can choose a λ_δ such that

$$\delta_{\lambda_\delta}^* = |\mathbf{v}|_{\inf}^{-1}E2^{-\lambda_\delta(2+p)} \leq \delta \leq |\mathbf{v}|_{\inf}^{-1}E2^{-(\lambda_\delta-1)(2+p)} = 2^{(2+p)}\delta_{\lambda_\delta}^*.$$

So we have $\delta \in \bigcup_{\lambda \in \Lambda} D_{\lambda, 2^{-(2+p)}}$.

□

5.5 Proof of Theorem 4.2

The fact that the function g_α satisfies Assumptions A, B, C is a known result. However, for the convenience of the reader, we will check these assumptions.

(a) We verify that $g_\alpha(\mu) = \frac{1}{\alpha+\mu}$ and $\Phi(\mu) = \mu^{p/2}$ satisfy Assumptions C, A1, A2. Direct verifying yields that Assumption C holds for g_α. The index function Φ satisfies Assumption A1. We verify Assumption A2. We have

$$\sqrt{\mu} g_\alpha(\mu) = \frac{\sqrt{\mu}}{\alpha+\mu} \le \frac{2\sqrt{\mu}}{\sqrt{\alpha\mu}} = \frac{2}{\sqrt{\alpha}}.$$

We verify Assumption A2 (ii). We have

$$|r_\alpha(\mu)|\sqrt{\Phi(\mu)} = \frac{\alpha\mu^{p/4}}{\alpha+\mu}.$$

Put $H(\mu) = \frac{\alpha\mu^r}{\alpha+\mu}$, $r \in (0,1)$. We have $H'(r) = \alpha\frac{r\mu^{r-1}(\alpha+\mu)-\mu^r}{(\alpha+\mu)^2}$. The function attains its maximum when $r(\alpha+\mu) - \mu = 0$ which gives $\mu = \frac{\alpha}{1-r}$. Choose $r = p/4$, we obtain Assumption A2 (ii) Hence $H(\mu) \le C\alpha^r$. For $p = 4$ we have $r = 1$, $H(\alpha) \le \alpha$ which give Assumption A2 (ii).

(b) We first consider Assumption B1. In fact we have $r_\alpha(\mu) = \frac{\alpha}{\alpha+\mu} \to 1$ as $\alpha \to \infty$. Hence Assumption B1 (i) holds. The function $g_\alpha(.)$ is continuous with respect to α. Hence Assumption B1 (ii) holds.

We verify Assumption B2. As known, $r_\alpha(\mu) = \frac{\alpha}{\alpha+\mu}$. We also have $\ell(\alpha) = \sup_{\mu\ge 0} g_\alpha(\mu) = \frac{1}{\alpha}$ which satisfies Assumption B2 (iii) with $\ell_* = \ell^* = 1$. Finally, we verify that the function Φ is concave. In fact, we have $\Phi''(\mu) = (p/2)(p/2-1)\mu^{p/2-2} < 0$ since $0 < p \le 2$.

5.6 Proof of Theorem 4.3

(a) The proof is divided into three steps.

Step 1. *Prove that $\{u_{\lambda,N}\}$ is a tight frame over $L^2(\mathbb{R})$.*

For $\theta \in L^2(\mathbb{R})$, we have

$$\langle \theta, u_{\lambda,N}\rangle = \frac{1}{2\pi}\langle \mathcal{F}(\theta), \mathcal{F}u_{\lambda,N}\rangle = \frac{1}{2\pi}\langle 1_{B_N}\mathcal{F}(\theta), \mathcal{F}u_\lambda\rangle. \tag{58}$$

Hence, using the Parseval equality gives

$$\begin{aligned}\sum_{N=0}^{\infty}\sum_{\lambda\in\Lambda}|\langle\theta, u_{\lambda,N}\rangle|^2 &= \frac{1}{2\pi}\sum_{N=0}^{\infty}\sum_{\lambda\in\Lambda}|\langle 1_{B_N}\mathcal{F}(\theta), \mathcal{F}u_\lambda\rangle|^2 \\ &= \frac{1}{2\pi}\sum_{N=0}^{\infty}\|1_{B_N}\mathcal{F}(\theta)\|^2 = \frac{1}{2\pi}\|\mathcal{F}(\theta)\|^2 = \|\theta\|^2.\end{aligned}$$

This follows that $\{u_{\lambda,N}\}$ is a tight frame.

Step 2. *Check to ensure that $\{u_\lambda\}$ is DFD quasi minimal.*

For every $(\nu, N_0) \in \mathbb{Z}^2 \times \mathbb{N}$, we choose $Q = 1$ and

$$I_{|\kappa_{\nu,N_0}|,Q|\kappa_{\nu,N_0}|} = I_{\kappa_{\nu,N_0},Q\kappa_{\nu,N_0}} = I_{e^{N_0T},e^{N_0T}}$$
$$= \{u_{\lambda,N} : \lambda \in \mathbb{Z}^2, N \in \mathbb{N}, \kappa_{\lambda,N} < e^{-N_0T} \text{ or } \kappa_{\lambda,N} > e^{-N_0T}\} = \{u_{\nu,N} : N \neq N_0\}.$$

We have $\mathcal{F}u_{\lambda,N} = 1_{B_N}\psi^\lambda \in I_{\kappa_{\nu,N_0},Q\kappa_{\nu,N_0}}$ for $N \neq N_0$. Since $B_N \cap B_{N_0} = \emptyset$ for $N \neq N_0$, we obtain

$$\langle u_{\lambda,N}, u_{\nu,N_0}\rangle = \frac{1}{2\pi}\langle 1_{B_N}\psi^\lambda, 1_{B_{N_0}}\psi^\nu\rangle = 0 \text{ for } N \neq N_0.$$

Hence $u_{\nu,N_0} \in I^\perp_{\kappa_{\nu,N_0},Q\kappa_{\nu,N_0}}$ which follows $I^\perp_{\kappa_{\nu,N_0},Q\kappa_{\nu,N_0}} \neq \emptyset$. We conclude that $\{u_{\lambda,N}\}$ is DFD quasi minimal.

Step 3. *Prove that $\{v_{\lambda,N}\}$ is a frame*

We recall $\mathcal{F}v_{\lambda,N} = \kappa_\lambda \exp(|\omega|^2T)\mathcal{F}(u_{\lambda,N}) = \kappa_\lambda 1_{B_N} \exp(|\omega|^2T)\mathcal{F}(u_\lambda)$. It follows that

$$\sum_{N=0}^{\infty}\sum_{\lambda\in\Lambda} |\langle\theta, v_{\lambda,N}\rangle|^2 = \frac{1}{2\pi}\sum_{N=0}^{\infty}\sum_{\lambda\in\Lambda} |\langle\kappa_{\lambda,N}e^{|\omega|^2T}1_{B_N}\mathcal{F}(\theta), \mathcal{F}u_\lambda\rangle|^2$$
$$= \frac{1}{2\pi}\sum_{N=0}^{\infty} \|\kappa_{\lambda,N}e^{|\omega|^2T}1_{B_N}\mathcal{F}(\theta)\|^2.$$

For $\omega \in B_N$, we have $\exp(NT) \le \exp(|\omega|^2T) \le \exp((N+1)T)$. Hence, since $\kappa_{\lambda,N} = e^{-NT}$ we obtain

$$1 \le \kappa_{\lambda,N}\exp(|\omega|^2T) \le e^T \text{ for } \omega \in B_N. \tag{59}$$

Using (59) gives

$$\frac{1}{2\pi}\sum_{N=0}^{\infty}\|1_{B_N}\mathcal{F}(\theta)\|^2 \le \frac{1}{2\pi}\sum_{N=0}^{\infty}\|\kappa_{\lambda,N}e^{|\omega|^2T}1_{B_N}\mathcal{F}(\theta)\|^2 \le \frac{e^T}{2\pi}\sum_{N=0}^{\infty}\|1_{B_N}\mathcal{F}(\theta)\|^2.$$

But $\frac{1}{2\pi}\sum_{N=0}^{\infty}\|1_{B_N}\mathcal{F}(\theta)\|^2 = \frac{1}{2\pi}\|\mathcal{F}(\theta)\|^2 = \|\theta\|^2$. Hence $\{v_{\lambda,N}\}$ is a frame.

(b) We find the source condition for the solution θ_0. As in the previous theorem, we can assume that $\theta_0 \in H^p(\mathbb{R})$ for $p \ge 0$. For $(\lambda, N) = (\lambda_D, \lambda_T, N)$, we can write

$$\langle\theta_0, u_{\lambda,N}\rangle = \frac{1}{2\pi}\langle\mathcal{F}\theta_0, \mathcal{F}u_\lambda\rangle = \frac{1}{2\pi}\int_{\mathbb{R}} 1_{G(\lambda_D)}1_{B_N}\mathcal{F}\theta_0(\xi)\overline{\mathcal{F}u_\lambda(\xi)}d\xi.$$

Hence, for $\lambda_D > 0$, using the Bessel inequality yields

$$\sum_{\lambda_T\in\mathbb{Z}} |\langle\theta_0, u_{\lambda,N}\rangle|^2 \le \|1_{G_D}1_{B_N}\mathcal{F}\theta_0\|^2 \le (1 + 2^{2\lambda_D}a_u)^{-p}\omega^2_{\lambda_D,N}$$

where

$$\omega_{\lambda_D,N} = \|1_{G_D}1_{B_N}(1+\omega^2)^p\mathcal{F}\theta_0\|.$$

It follows that

$$\sum_{\lambda_T\in\mathbb{Z}} N^p|\langle\theta_0, u_{\lambda,N}\rangle|^2 \le N^p\|1_{G_D}1_{B_N}\mathcal{F}\theta_0\|^2 \le N^p(1 + 2^{2\lambda_D}a_u)^{-p}\omega^2_{\lambda_D,N}.$$

On the other hand, we have $\langle\theta_0, u_{\lambda,N}\rangle \neq 0$ then $B_N \cap G_{\lambda_D} \neq \emptyset$, which gives $\sqrt{N} \leq 2^{\lambda_D} b_u$. So we have

$$\sum_{\lambda_T\in\mathbb{Z}} N^p |\langle\theta_0, u_{\lambda,N}\rangle|^2 \leq 2^{2p\lambda_D} b_u^{2p}(1+2^{2\lambda_D}a_u)^{-p}\omega^2_{\lambda_D,N} \leq b_u^{2p}a_u^{-p}\omega^2_{\lambda_D,N}.$$

Noting that $N = -\frac{1}{2T}\ln\kappa^2_{\lambda,N}$, we have $N^p = \frac{1}{(2T)^p}[\Phi(|\kappa_\lambda|^2)]^{-1}$ and

$$\sum_{\lambda_T\in\mathbb{Z}} [\Phi(|\kappa_\lambda|^2)]^{-1}|\langle\theta_0, u_{\lambda,N}\rangle|^2 \leq (2T)^p b_u^{2p}a_u^{-p}\omega^2_{\lambda_D,N}.$$

It implies that

$$\sum_{N=0}^{\infty}\sum_{\lambda_D\in\mathbb{Z}}\sum_{\lambda_T\in\mathbb{Z}} [\Phi(|\kappa_\lambda|^2)]^{-1}|\langle\theta_0, u_{\lambda,N}\rangle|^2 \leq (2T)^p b_u^{2p}a_u^{-p}\sum_{N=0}^{\infty}\sum_{\lambda_D\in\mathbb{Z}}\omega^2_{\lambda_D,N} \leq (2T)^p b_u^{2p}a_u^{-p}\|\theta_0\|^2_{H^p}.$$

So we obtain $\theta_0 \in \mathbf{M}_{\Phi,E}$ where $E = (2T)^{p/2}b_u^p a_u^{-p/2}\|\theta_0\|^2_{H^p}$, $\Phi(\mu) = (-\ln\mu)^{-p}$ for $\mu \in (0,1)$.

(c) We have $|\kappa_{\lambda,N}| = \kappa_{\lambda,N} = e^{-NT}$, $\Phi(\mu) = (-\ln\mu)^{-p}$. It follows that $\Phi(\kappa^2_{\lambda,N}) = (2NT)^{-p}$. Hence, from (23), we obtain $\delta^*_{\lambda,N} = |\mathbf{v}|^{-1}_{\inf}E\sqrt{\kappa^2_{\lambda,N}\Phi(\kappa^2_{\lambda,N})} = |\mathbf{v}|^{-1}_{\inf}\sqrt{(2NT)^{-p}e^{-2NT}}$. Hence, for every $\delta \in (0,\delta_0)$ where $\delta_0 = |\mathbf{v}|^{-1}_{\inf}\sqrt{(2T)^{-p}e^{-2T}}$, we can find $N_0 \in \mathbb{N}$ such that $\delta^*_{\lambda,N_0} \leq \delta \leq \delta^*_{\lambda,N_0-1}$. We note that

$$\delta^*_{\lambda,N_0-1} = |\mathbf{v}|^{-1}_{\inf}\sqrt{(2(N_0-1)T)^{-p}e^{-2(N_0-1)T}} \leq e^T|\mathbf{v}|^{-1}_{\inf}\sqrt{(2N_0T)^{-p}e^{-2N_0T}} = e^T\delta^*_{\lambda,N_0}.$$

Hence $\delta \in [\delta^*_{\lambda,N_0}, \beta^{-1}\delta^*_{\lambda,N_0}] \subset \bigcup_{(\lambda,N)\in\Lambda\times\mathbb{N}} D_{(\lambda,N),\beta}$ with $\beta = e^{-T}$.

5.7 Proof of Theorem 4.4

We verify that $g_\alpha(\mu) = \frac{1}{\alpha+\mu}$ and $\Phi(\mu) = (-\ln\mu)^{-p}$ satisfy Assumptions C, A1, A2, B1, B2. As shown in the proof of Theorem 4.2, Assumptions C, B1, B2 hold for g_α. The index function Φ satisfies Assumptions A1 (i), (ii). Using Theorem 9.1 in [30] gives that the function Θ is convex on $(0,\infty)$ for $p > 0$, i.e. Assumptions A1 (iii) holds.

Assumption A2 (i) is verified in the proof of Theorem 4.2. We verify Assumption A2 (ii). We have

$$|r_\alpha(\mu)|\sqrt{\Phi(\mu)} = \frac{\alpha\,(-\ln\mu)^{-p/2}}{\alpha+\mu}.$$

Putting $\tau = \alpha/\mu$, we obtain

$$\frac{\alpha\,(-\ln\mu)^{-p/2}}{\alpha+\mu} = \frac{\tau(\ln(\tau\alpha^{-1}))^{-p/2}}{\tau+1}.$$

For $\alpha < \tau \leq \sqrt{\alpha}$, $\eta \in (0,1)$, using the inequality $z^\eta(\ln z)^{-p/2} \to 0$ as $z \to 0$, we have

$$\frac{\tau(\ln(\tau\alpha^{-1}))^{-p/2}}{\tau+1} \leq \frac{\tau^{1-\eta}\alpha^\eta(\tau\alpha^{-1})^\eta(\ln(\tau\alpha^{-1}))^{-p/2}}{\tau+1} \leq C\alpha^\eta \leq C'(-\ln\alpha)^{-p/2}.$$

For $\tau \geq \sqrt{\alpha}$ we have

$$\frac{\tau(\ln(\tau\alpha^{-1}))^{-p/2}}{\tau+1} \leq (\ln(\tau\alpha^{-1}))^{-p/2} \leq (\ln(\alpha^{-1/2}))^{-p/2} \leq C\ln(\alpha^{-1})^{-p/2}.$$

Finally, to apply Theorem 3.6 we show that $\Phi(\mu)$ is concave. We have $\Phi'(\mu) = \frac{p}{\mu^2}(-\ln\mu)^{-p-1}$ and $\Phi''(\mu) = \frac{p}{\mu^2}(-\ln\mu)^{-p-2}(\ln\mu + p + 1) \leq \frac{p}{\mu^2}(-\ln\mu)^{-p-2}(\ln a^* + p + 1) < 0$. Hence $\Phi(\mu)$ is concave.

Since Assumptions C, A1, A2, B1 and B2 hold, we obtain the order optimal property of our a priori and a-posteriori regularization.

6 Conclusion

The paper has investigated DFD regularizations in both a priori and a-posteriori cases. For the case where the $\{u_\lambda\}$ system is DFD quasi minimal, we have proved the sequential order optimality property and the global optimality for DFD regularizations. Some issues that need to be investigated in the future are

-Methods of constructing DFDs for ill-posed problems

-Investigation of the relationship between the classical source condition and the DFD source condition.

-Investigation of optimality in the case where $\{u_\lambda\}$ is not DFD quasi minimal.

-Find the condition of the DFD singular value so that the regularization method is uniformly optimal.

Acknowledgments. We thank the reviewers for their profound feedback, which allowed us to elevate this paper substantially from its original version. This research received support from Vietnam National University (VNU-HCM) under grant number T2024-18-01. The first author also completed a portion of this paper during a sponsored visiting research period at Vietnam Institute for Advanced Studies in Mathematics (VIASM) from July 1 to August 30, 2025.

References

[1] D. H. Bailey, and P. N. Swarztrauber.*A fast method for the numerical evaluation of continuous Fourier and Laplace transforms.* SIAM Journal on Scientific Computing, vol. 15, no. 5, pp. 1105-1110, 1994.

[2] O. Christensen. An Introduction to Frames and Riesz Bases. Birkhäuser Cham, 2009.

[3] I. Daubechies. Ten Lectures on Wavelets. CBMS-NSF Regional Conference Series in Applied Mathematics. Society for Industrial and Applied Mathematics, Philadelphia, PA, 1992.

[4] Z. Deng, X. Yang. *A Discretized Tikhonov regularization method for a fractional backward heat conduction problem.* Abstract and Applied Analysis, 2014, 964373.

[5] D. Donoho. *Nonlinear Solution of Linear Inverse Problem by Wavelet-Vaguelette Decomposition.* Applied and Computational Harmonic Analysis, **02**, 1995, pp. 101-125.

[6] K. Diethelm. The Analysis of Fractional Differential Equations. Springer, Berlin, 2010.

[7] A. Ebner, J. Frikel, D. Lorenz, J. Schwab, M. Haltmeier. *Regularization of Inverse Problems by Filtered Diagonal Frame Decomposition.* Applied and Computational Harmonic Analysis, Volume 62, January 2023, Pages 66-83.

[8] A. Ebner and M. Haltmeier. *Convergence of Non-linear Diagonal Frame Filtering for Regularizing Inverse Problems.* Inverse Problems, 40(5):055009, 2024.

[9] A. Ebner, M. Schwab, and M. Haltmeier. *Error Estimates for Weakly Convex Frame-Based Regularization Including Learned Filters.* SIAM Journal on Imaging Sciences, 18(2):822–850, 2025.

[10] H. W. Engl, M. Hanke, A. Neubauer. Regularization of Inverse Problems,1996.

[11] J. Frikel and M. Haltmeier, *Efficient regularization with wavelet sparsity constraints in photoacoustic tomography.* Inverse Problems, 34(2):024006, 2018.

[12] J. Frikel and M. Haltmeier. *Sparse Regularization of Inverse Problems by Operator-Adapted Frame Thresholding.* In W. Dörfler, M.Hochbruck, D. Hundertmark, W. Reichel, A. Rieder, R. Schnaubelt, and B. Schörkhuber, editors, Mathematics of Wave Phenomena, pages 163–178, Cham, 2020. Springer International Publishing.

[13] R. Garrappa.*Numerical evaluation of two and three parameter Mittag-Leffler functions.* SIAM Journal on Numerical Analysis, vol. 53, no. 3, pp. 1350–1369, 2015.

[14] S. Göppel, J. Frikel, M. Haltmeier. *Translation invariant diagonal frame decomposition of inverse problems and their regularization.* Inverse Problems 39, 065011 (21pp), 2023.

[15] R. Gorenflo, A. A. Kilbas, F. Mainardi, S. V. Rogosin. Mittag-Leffler Functions, Related Topics and Applications. Springer Berlin Heidelberg, 2020.

[16] R. Hilfer. Applications of Fractional Calculus in Physics. World Scientific, Singapore, 2000.

[17] B. Hofmann, B. Kaltenbacher, C. Poeschl, and O. Scherzer. *A convergence rates result for Tikhonov regularization in Banach spaces with non-smooth operators.* Inverse Problems, 23 (2007), p. 987.

[18] T. Hohage. *Regularization of exponentially ill-posed problems.* Numerical Functional Analysis and Optimization, 21:3-4, 2000, 439-464.

[19] S. Hubmer and R. Ramlau. *A frame decomposition of the atmospheric tomography operator.* Inverse Problems, 36(9):094001, 2020.

[20] S. Hubmer and R. Ramlau. *Frame Decompositions of Bounded Linear Operators in Hilbert Spaces with Applications in Tomography.* Inverse Problems, 37(5):055001, 2021.

[21] S. Hubmer and R. Ramlau, L. Weissinger. *On regularization via frame decompositions with applications in tomography.* Inverse Problems, **38**, 2022, pp. 01-28.

[22] A.A. Kilbas, H.M., Srivastava, J.J.Trujillo. Theory and Applications of Fractional Differential Equations. Elsevier, New York, 2006.

[23] A. Kirsch. An Introduction to the Mathematical Theory of Inverse Problems. Springer New York, NY, 2011.

[24] J.J. Liu, M. Yamamoto. *A backward problem for the time-fractional diffusion equation.* Appl. Anal. 11, 1769-1788, 2010.

[25] A. K. Louis. Inverse und schlecht gestellte Probleme. Teubner Studienbücher Mathematik. Vieweg+Teubner Verlag, 1989.

[26] V. A. Morozov. *On the solution of functional equations by the method of regularization.* Soviet Math. Dokl. 7 (1966), 414-417.

[27] I. Podlubny. Fractional Differential Equations. Academic Press, 1999.

[28] M. Quellmalz, L. Weissinger, S. Hubmer, and P. D. Erchinger. *A Frame Decomposition of the Funk-Radon Transform.* In: Scale Space and Variational Methods in Computer Vision, pages 42–54, Cham, 2023. Springer International Publishing.

[29] G. Strang, and T. Nguyen. Wavelets and Filter Banks. Wellesley, Wellesley-Cambridge Press, 1996.

[30] U. Tautenhahn, U. Hamarik, B. Hofmann, Y. Shao. *Conditional Stability Estimates for Ill-posed PDE Problems by Using Interpolation.* Numerical Functional Analysis and Optimization **34:12**, 2013, pp. 1370-1417.

[31] D. D. Trong, E. Nane, N.D. Minh, N. H. Tuan. *Continuity of Solutions of a Class of Fractional Equations*. Potential, Volume 49, (2018), 423–478.

[32] V.V. Uchaikin. Fractional Derivatives for Physicists and Engineers: Background and Theory. Springer, Berlin, 2013.

[33] G. M. Vainikko, A. Y. Veretennikov. Iteration Procedures in Ill-Posed Problems. Nauka, Moscow, 1986.

[34] L. Wang,J. Liu. *Total variation regularization for a backward time-fractional diffusion problem*. Inverse Problems 29, 115013, 2013.

[35] L. Weissinger, S. Hubmer, B. Stadler, and Ronny Ramlau. *Singular value and frame decomposition-based reconstruction for atmospheric tomography*. In: Inverse Problems on Large Scales: Mathematical Modelling and Computational Methods, Berlin, Boston: De Gruyter, 2025. https://doi.org/10.1515/9783111357270

[36] Jun-Gang Wang, Ting Wei, Yu-Bin Zhou. *Tikhonov regularization method for a backward problem for the time-fractional diffusion equation*, Applied Mathematical Modelling 37, pages 8518-8532, 2013.