Nr. 98
30. Aug 2023

universität innsbruck
Institut für Mathematik

Preprint-Series: Department of Mathematics - Applied Mathematics

Convergence of non-linear diagonal frame filtering for regularizing inverse problems

A. Ebner, M. Haltmeier

AppliedMathematics

Technikerstraße 13 - 6020 Innsbruck - Austria
Tel.: +43 512 507 53803 Fax: +43 512 507 53898
https://applied-math.uibk.ac.at

Convergence of non-linear diagonal frame filtering for regularizing inverse problems

Andrea Ebner

Department of Mathematics, University of Innsbruck
Technikerstrasse 13, 6020 Innsbruck, Austria
E-mail: andrea.ebner@uibk.ac.at

Markus Haltmeier

Department of Mathematics, University of Innsbruck
Technikerstrasse 13, 6020 Innsbruck, Austria
E-mail: markus.haltmeier@uibk.ac.at

August 30, 2023

Abstract

Inverse problems are core issues in several scientific areas, including signal processing and medical imaging. As inverse problems typically suffer from instability with respect to data perturbations, a variety of regularization techniques have been proposed. In particular, the use of filtered diagonal frame decompositions has proven to be effective and computationally efficient. However, the existing convergence analysis applies only to linear filters and a few non-linear filters such as soft thresholding. In this paper, we analyze the filtered diagonal frame decomposition with general non-linear filters. In particular, our results generalize SVD-based spectral filtering from linear to non-linear filters as a special case. We present three strategies to demonstrate convergence. The first two strategies relate non-linear diagonal frame filtering to variational regularization and plug-and-play regularization, respectively. The third strategy allows us to relax the assumptions involved and still obtain a full convergence analysis.

Keywords: Inverse problems, non-linear filtering, diagonal frame decomposition, regularization, convergence analysis, spectral filtering.

1 Introduction

Let $\mathbf{A}\colon \mathbb{X} \to \mathbb{Y}$ be a bounded linear operator between two real Hilbert spaces $\mathbb{X}$ and $\mathbb{Y}$. We consider the inverse problem of recovering $x^+ \in \mathbb{X}$ from noisy data

$$y^\delta = \mathbf{A}x^+ + z\,, \tag{1.1}$$

where z is the data perturbation with $\|z\| \leq \delta$ for some noise level $\delta > 0$. Inverting the operator $\mathbf{A}$ is often ill-posed in the sense that the the Moore-Penrose inverse $\mathbf{A}^+$ is discontinuous. Thus, small errors in the data are significantly amplified by use of exact solution methods. To address this problem, regularization methods have been developed with the aim of finding approximate but stable solution strategies [4, 12, 25].

1.1 Diagonal frame filtering

Diagonal frame decompositions in combination with regularizing filters are a flexible and efficient regularization concept for (1.1). Suppose $\mathbf{A}$ has a diagonal frame decomposition (DFD) giving the representations

$$\mathbf{A} = \sum_{\lambda\in\Lambda} \kappa_\lambda \langle \cdot , u_\lambda \rangle \bar{v}_\lambda$$
$$\mathbf{A}^+ = \sum_{\lambda\in\Lambda} \kappa_\lambda^{-1} \langle \cdot , v_\lambda \rangle \bar{u}_\lambda \,.$$

Here $(u_\lambda)_{\lambda\in\Lambda}$ and $(v_\lambda)_{\lambda\in\Lambda}$ are frames of $\ker(\mathbf{A})^\perp$ and $\overline{\operatorname{ran}(\mathbf{A})}$, respectively, with corresponding dual frames $(\bar{u}_\lambda)_{\lambda\in\Lambda}$ and $(\bar{v}_\lambda)_{\lambda\in\Lambda}$. In the special case that $(u_\lambda)_\lambda$ and $(v_\lambda)_\lambda$ are orthonormal, $\mathbf{A} = \sum_{\lambda\in\Lambda} \kappa_\lambda \langle \cdot, u_\lambda \rangle \bar{v}_\lambda$ is a singular value decomposition (SVD) for $\mathbf{A}$. More general frame decompositions have first been studied by Candés and Donoho [6, 8] in the context of statistical estimation and recently in [13, 14, 9, 16, 24, 20, 21] in the context of regularization theory. Specifically, in [9, Theorem 2.10] it has been shown that if $(\kappa_\lambda)_{\lambda\in\Lambda}$ accumulate at zero then $\mathbf{A}^+$ is unbounded. In such a situation regularization methods have to applied for the solution of (1.1).

Linear filtered DFD methods are based on a regularizing filter family $(f_\alpha)_{\alpha>0}$ and defined by $\mathbf{B}_\alpha(y^\delta) := \sum_{\lambda\in\Lambda} f_\alpha(\kappa_\lambda)\langle y^\delta, v_\lambda\rangle \bar{u}_\lambda$; see [9]. Each factor $f_\alpha(\kappa_\lambda)\langle y^\delta, v_\lambda\rangle$ is a damped version of the exact coefficient inverse $\kappa_\lambda^{-1}\langle y^\delta, v_\lambda\rangle$. The filtering process $\varphi_\alpha(\kappa_\lambda, \langle y^\delta, v_\lambda\rangle) = (f_\alpha(\kappa_\lambda)\kappa_\lambda) \cdot \langle y^\delta, v_\lambda\rangle$ is linear in $\langle y^\delta, v_\lambda\rangle$ and represented by the damping factors $f_\alpha(\kappa_\lambda)\kappa_\lambda$. Further note that $\mathbf{B}_\alpha$ reduces to the well-known spectral filtering technique which in the regularization context is also referred to as filter-based regularization. The convergence analysis in this special case is well known and can be found for example in [12, 18]. The more general case of filtered DFDs has been first analyzed in [9] and later in [21].

1.2 Non-linear extension

A major drawback of linear regularizing filters is that the damping factor depends only on the quasi-singular values and is independent of the data. In practice, certain filters that depend non-linearly on $\langle y^\delta, v_\lambda\rangle$ tend to perform better in filtering out noise than linear methods; see [1, 15, 22]. The aim of this paper is to analyze general non-linear frame-based diagonal filtering

$$\mathbf{B}_\alpha(y^\delta) = \sum_{\lambda\in\Lambda} \kappa_\lambda^{-1} \varphi_\alpha(\kappa_\lambda, \langle y^\delta, v_\lambda\rangle)\bar{u}_\lambda \,, \tag{1.2}$$

where $(\varphi_\alpha)_{\alpha>0}$ is a non-linear filter rigorously introduced in Definition 3.1. The reconstruction mappings $\mathbf{B}_\alpha \colon \mathbb{Y} \to \mathbb{X}$ come with a clear interpretation: In order to avoid noise amplification due to multiplication with κ_λ^{-1}, the filter decreases each coefficient according to the filter function $\varphi_\alpha(\kappa_\lambda, \langle y^\delta, v_\lambda\rangle)$ prior to the inversion. By taking $\varphi_\alpha(\kappa_\lambda, \langle y^\delta, v_\lambda\rangle) = (f_\alpha(\kappa_\lambda)\kappa_\lambda) \cdot \langle y^\delta, v_\lambda\rangle$ we recover linear filtering analyzed in [9]. An example of a non-linear filter is the soft thresholding filter defined by $\varphi_\alpha(\kappa, c) = \operatorname{sign}(c)(|c| - \alpha/\kappa)_+$. In [14] it has been shown that this filter yields a convergent regularization method. Affine filters applied with the SVD have recently been studied in [2]. Our paper extends these special cases to general classes of non-linear filters.

Specifically, we derive the following main results. In these theorems we use the additional Assumptions A, B, and C on the regularization filters, to be introduced in Section 4. In particular, we will show that Assumption A allows the application of variational regularization [26], Assumption B the application of Plug-and-Play (PnP) regularization [10], and with Assumption C we exploit the specific diagonal structure.

Theorem 1.1 (Stability)**.** *Let $(\varphi_\alpha)_{\alpha>0}$ be a non-linear regularizing filter, $\alpha > 0$ and $y^\delta, y^k \in \mathbb{Y}$ be with $y^k \to y^\delta$. If Assumption A or C holds, then $\mathbf{B}_\alpha(y^k) \rightharpoonup \mathbf{B}_\alpha(y^\delta)$. If instead Assumption B holds, then $\mathbf{B}_\alpha(y^k) \to \mathbf{B}_\alpha(y^\delta)$.*

Theorem 1.2 (Convergence)**.** *Let $(\varphi_\alpha)_{\alpha>0}$ be a non-linear regularizing filter, $(\delta_k)_k, (\alpha_k)_k$ null sequences and $y \in \operatorname{ran}(\mathbf{A})$. Suppose $\left(\int_0^{\langle y, v_\lambda\rangle} \max(\varphi_{\tilde{\alpha},\lambda}^{-1}(y))dy\right)_\lambda \in \ell^1(\Lambda)$ for some $\tilde{\alpha} > 0$. Let $(y^k)_k \in \mathbb{Y}^\mathbb{N}$ with $\|y^k - y\| \leq \delta_k$ and suppose that one of the following holds:*

- *Assumption A and $\delta_k^2/\alpha_k \to 0$.*
- *Assumption B and $(1 - \ell_{\alpha_k})/\ell_{\alpha_k} \gtrsim \delta_k$.*
- *Assumption C and $\alpha_k \gtrsim \delta_k$ and $1 - \ell_{\alpha_k} \gtrsim \delta_k$.*

Then $\mathbf{B}_{\alpha_k}(y^k) \rightharpoonup \mathbf{A}^+ y$ as $k \to \infty$.

Proofs of Theorems 1.1 and 1.2 will be given in Section 4.

Note that even in the case where we reduce our analysis to variational regularization, the non-linear filtered regularization technique is related but different from variational regularization with separable constraints [17, 5]. This is discussed in detail for special case of soft thresholding in [14], where the diagonal frame filtering has been opposed to frame-analysis and frame-synthesis regularization.

1.3 Outline

In Section 2 we present preliminaries in the form of a notion, auxiliary results, and some technical lemmas. In Section 3 we rigorously introduce non-linear filters and non-linear filtered diagonal frame decompositions. The main results of this paper are presented in Section 4, where we provide three approaches to proving convergence: by linking to existing theories of variational regularization and PnP regularization, and by a direct proof. Finally, in the concluding section, we summarize our findings and offer some future research directions.

2 Preliminaries

Let $\mathbb{X}$, $\mathbb{Y}$ be Hilbert spaces. If $\mathbf{B}\colon \operatorname{dom}(\mathbf{B}) \subseteq \mathbb{X} \to \mathbb{Y}$, then $\operatorname{dom}(\mathbf{B})$ denotes the domain and $\operatorname{ran}(\mathbf{B}) = \mathbf{B}(\operatorname{dom}(\mathbf{B}))$ the range of $\mathbf{B}$. If $\mathbf{B}$ is linear bounded with $\operatorname{dom}(\mathbf{B}) = \mathbb{X}$ we write $\mathbf{B}\colon \operatorname{dom}(\mathbf{B}^+) \subseteq \mathbb{Y} \to \mathbb{X}$ with $\operatorname{dom}(\mathbf{B}^+) \coloneqq \operatorname{ran}(\mathbf{B}) \oplus \operatorname{ran}(\mathbf{B})^\perp$ for the Moore-Penrose inverse of $\mathbf{B}$. Recall that $\Phi\colon \mathbb{X} \to \mathbb{X}$ is nonexpansive if $\forall x, y \in \mathbb{X}\colon \|\Phi(x) - \Phi(y)\| \leq \|x - y\|$.

2.1 Functionals and proximity operators

Functionals on $\mathbb{X}$ will be written as $\mathcal{R}\colon \mathbb{X} \to \mathbb{R} \cup \{\infty\}$ and we usually use s to denote a functional when $\mathbb{X} = \mathbb{R}$. We define the domain of $\mathcal{R}$ by $\operatorname{dom}(\mathcal{R}) \coloneqq \{x \in \mathbb{X} \mid \mathcal{R}(x) < \infty\}$ and for $q \in \mathbb{R}$ we define the lower level set of $\mathcal{R}$ with bound q by $\mathcal{L}(\mathcal{R}, q) \coloneqq \{x \in \mathbb{X} \mid \mathcal{R}(x) \leq q\}$. The functional $\mathcal{R}$ is called proper if $\operatorname{dom}(\mathcal{R}) \neq \emptyset$ and convex if $\mathcal{R}(tx + (1 - t)y) \leq$

$t\mathcal{R}(x) + (1-t)\mathcal{R}(y)$ for all $x, y \in \mathbb{X}$ and $t \in [0,1]$. It is called lower semi-continuous if $\mathcal{L}(\mathcal{R}, q)$ is closed for all $q \in \mathbb{R}$. For convex functions sequentially, weak as well as weak sequential lower semi-continuity are equivalent to strong lower semi-continuity. We call $\mathcal{R}$ norm-coercive if $\mathcal{R}(x^k) \to \infty$ for all $(x^k)_{k\in\mathbb{N}} \in \operatorname{dom}(\mathcal{R})^{\mathbb{N}}$ with $\|x^k\| \to \infty$.

We define $\Gamma_0(\mathbb{X})$ as the set of all $\mathcal{R}\colon \mathbb{X} \to \mathbb{R} \cup \{\infty\}$ that are proper, convex and lower semi-continuous. The subdifferential of $\mathcal{R} \in \Gamma_0(\mathbb{X})$ is a set-valued operator $\partial\mathcal{R}\colon \mathbb{X} \to 2^{\mathbb{X}}$ defined by

$$\partial\mathcal{R}(x) \coloneqq \{u \in \mathbb{X} \mid \forall y \in \mathbb{X}\colon \langle y - x, u\rangle + \mathcal{R}(x) \leq \mathcal{R}(y)\}\,.$$

The elements of $\partial\mathcal{R}(x)$ are called subgradients of $\mathcal{R}$ at x.

Let Λ be a countable index set and $s_\lambda\colon \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ be non-negative for $\lambda \in \Lambda$. The functional $\mathcal{R} = \bigoplus_{\lambda\in\Lambda} s_\lambda\colon \ell^2(\Lambda) \to \mathbb{R} \cup \{\infty\}$ is defined by $\mathcal{R}((x_\lambda)_\lambda) = \sum_{\lambda\in\Lambda} s_\lambda(x_\lambda)$.

For $\mathcal{R} \in \Gamma_0(\mathbb{X})$, the proximity operator $\operatorname{prox}_{\mathcal{R}}\colon \mathbb{X} \to \mathbb{X}$ is well defined by $\operatorname{prox}_{\mathcal{R}}(x) \coloneqq \operatorname{argmin}_{y\in\mathbb{X}} \|x - y\|^2/2 + \mathcal{R}(y)$.

Lemma 2.1 (Properties of proximity operators)**.** *Let $\mathcal{R} \in \Gamma_0(\mathbb{X})$ and $\phi\colon \mathbb{R} \to \mathbb{R}$.*

(a) $\operatorname{prox}_{\mathcal{R}} = (\operatorname{id} + \partial\mathcal{R})^{-1}$.

(b) $\operatorname{prox}_{\mathcal{R}}$ *is nonexpansive.*

(c) Let $\mathcal{S}\colon \mathbb{X} \to \mathbb{R}$ be convex and Fréchet differentiable. Then, for all $\gamma \neq 0$,

$$\operatorname{argmin}(\mathcal{S} + \mathcal{R}) = \operatorname{Fix}(\operatorname{prox}_{\gamma\mathcal{R}} \circ (\operatorname{id} - \gamma\nabla\mathcal{S}))\,.$$

(d) If $\mathbb{X} = \mathbb{R}$ then $\operatorname{dom}(\mathcal{R})$ is a closed interval and $\mathcal{R}$ is continuous on $\operatorname{dom}(\mathcal{R})$.

(e) $\exists q \in \Gamma_0(\mathbb{R})\colon \phi = \operatorname{prox}_q \Leftrightarrow \phi$ is nonexpansive and increasing.

(f) Let Λ be at most countable and $(r_\lambda)_{\lambda\in\Lambda} \in \Gamma_0(\mathbb{R})^\Lambda$ with $\forall\lambda\colon r_\lambda \geq r_\lambda(0) = 0$. Then $\mathcal{R} \coloneqq \bigoplus_{\lambda\in\Lambda} r_\lambda$ is contained in $\Gamma_0(\ell^2(\Lambda))$, the proximity operator $\operatorname{prox}_{\mathcal{R}}$ is weakly sequentially continuous, and $\forall (y_\lambda)_\lambda \in \ell^2(\Lambda)\colon \operatorname{prox}_{\mathcal{R}}((y_\lambda)_\lambda) = (\operatorname{prox}_{r_\lambda}(y_\lambda))_\lambda$.

Proofs of all claims in Lemma 2.1 can be found in [3].

2.2 Diagonal frame decomposition

Let $\mathbf{A}\colon \mathbb{X} \to \mathbb{Y}$ be a bounded linear operator and Λ an at most countable index set. The filter techniques analysed in this paper use a diagonal frame decomposition (DFD) $(\mathbf{u}, \mathbf{v}, \boldsymbol{\kappa}) = (u_\lambda, v_\lambda, \kappa_\lambda)_{\lambda\in\Lambda}$ as introduced [9].

Definition 2.2 (Diagonal drame decomposition, DFD)**.** The triple $(\mathbf{u}, \mathbf{v}, \boldsymbol{\kappa}) = (u_\lambda, v_\lambda, \kappa_\lambda)_{\lambda\in\Lambda}$ is called diagonal frame decomposition (DFD) for $\mathbf{A}$ if the following holds:

(D1) $(u_\lambda)_{\lambda\in\Lambda}$ is a frame for $(\ker \mathbf{A})^\perp \subseteq \mathbb{X}$.

(D2) $(v_\lambda)_{\lambda\in\Lambda}$ is a frame for $\overline{\operatorname{ran}\mathbf{A}} \subseteq \mathbb{Y}$.

(D3) $(\kappa_\lambda)_{\lambda\in\Lambda} \in (0,\infty)^\Lambda$ satisfies the quasi-singular relations $\forall\lambda \in \Lambda\colon \mathbf{A}^* v_\lambda = \kappa_\lambda u_\lambda$.

We call $(\kappa_\lambda)_{\lambda\in\Lambda}$ the family of quasi-singular values and $(u_\lambda)_{\lambda\in\Lambda}$, $(v_\lambda)_{\lambda\in\Lambda}$ the corresponding quasi-singular systems.

The notion of a DFD reduces to the singular value decomposition (SVD) when $(u_\lambda)_{\lambda\in\Lambda}$ and $(v_\lambda)_{\lambda\in\Lambda}$ are orthonormal bases. In particular, DFDs exist in quite general settings. They can also exist where there is no SVD, for example when the spectrum of $\mathbf{A}$ is continuous. However, the main advantage of DFDs over SVDs is that the quasi-singular systems can provide better approximation properties than the singular systems from the SVD. An example of this is when $(u_\lambda)_{\lambda\in\Lambda}$ can be taken as the wavelet basis, as in the case of the Radon transform [8, 9, 21].

Using the DFD, the Moore-Penrose inverse of $\mathbf{A}$ can be written as

$$\forall y \in \operatorname{dom}(\mathbf{A}^+)\colon \quad \mathbf{A}^+(y) = \sum_{\lambda\in\Lambda} \frac{1}{\kappa_\lambda} \langle y, v_\lambda\rangle \, \bar{u}_\lambda = \mathbf{T}_{\bar{\mathbf{u}}} \circ \mathbf{M}_{\boldsymbol{\kappa}}^+ \circ \mathbf{T}_{\mathbf{v}}^*(y) \,, \tag{2.1}$$

where $\bar{\mathbf{u}}$ is the dual frame of $\mathbf{u}$, $\mathbf{T}_{\bar{\mathbf{u}}}$ and $\mathbf{T}_{\mathbf{v}}^*$ are the synthesis and analysis operator of the frame $\bar{\mathbf{u}}$ and $\mathbf{v}$, respectively, and $\mathbf{M}_{\boldsymbol{\kappa}}$ is the component-wise multiplication operator $\mathbf{M}_{\boldsymbol{\kappa}}((x_\lambda)_{\lambda\in\Lambda}) = (\kappa_\lambda x_\lambda)_{\lambda\in\Lambda}$. Since the frame operators are continuous and invertible, diagonalizing $\mathbf{A}$ with a DFD basically reduces the inverse problem (1.1) to an inverse problem with a diagonal forward operator from $\ell^2(\Lambda)$ to $\ell^2(\Lambda)$.

Due to the ill-posedness of inverting $\mathbf{A}$, the values of $(\kappa_\lambda)_{\lambda\in\Lambda}$ accumulate at zero [9, Thm. 7], which means that $(1/\kappa_\lambda)_\lambda$ is unbounded. As a result, small errors in the data can be significantly amplified using (2.1). To reduce error amplification, we use regularizing filters aiming to damp noisy coefficients. Unlike the widely studied linear regularizing filters [6, 9, 19, 21], we investigate non-linear filters that can depend on both the operator and the data in a non-linear matter.

2.3 Technical lemmas

Lemma 2.3. *Let $s\colon \mathbb{R} \to \mathbb{R}^+$ be convex and lower semi-continuous with $s(0) = 0$, and let $\kappa, \alpha > 0$. Suppose there exist $b, c > 0$ such that*

$$\forall x \in \mathbb{R}\colon \quad |x| \leq \min(\operatorname{prox}_{\alpha s}^{-1}(c\kappa)) \Rightarrow |\operatorname{prox}_{\alpha s}(x)| \leq \frac{\kappa^2}{\kappa^2 + \alpha b}|x| \,. \tag{2.2}$$

Then for all $y \in \mathbb{R}$ the following holds:

- *If $|y| \leq c\kappa$, then $s(y) \geq (b/2) \cdot |y/\kappa|^2$.*
- *If $|y| > c\kappa$, then $s(y) \geq bc\,|y/\kappa| - bc^2/2$.*

Proof. Fix $b, c > 0$ such that (2.2) is satisfied. We start by considering the case where $y \geq 0$ with $y \leq c\kappa$. Since s is nonnegative and $s(0) = 0$, we have $\operatorname{prox}_{\alpha s}(0) = 0$ and $\operatorname{prox}_{\alpha s}$ is positive for positive arguments and negative for negative arguments. From (2.2) for $x > 0$ it follows $\operatorname{prox}_{\alpha s}^{-1}(y) \geq y(\kappa^2 + \alpha b)/\kappa^2$. Note that $\operatorname{prox}_{\alpha s}^{-1}(y)$ can be set-valued in which case the last inequality is understood to be satisfied for all elements. By Lemma 2.1(a) we get $(\operatorname{id} + \alpha\partial s)(y) \geq y + y\alpha b/\kappa^2$ which is equivalent to $\partial s(y) \geq yb/\kappa^2 = \partial(y^2 b/(2\kappa^2))$. With $s(0) = 0$ we get $s(y) \geq (b/2) \cdot |y/\kappa|^2$ for all $y \in [0, c\kappa]$. Now consider the case $y > c\kappa$. Since s is convex, the sub-gradients increase. At point κc the sub-gradients of s are grater than bc/κ, hence for all $y \geq c\kappa$ we have $s(y) \geq y\, bc/\kappa + a$ and by taking $y = \kappa c$ we get $a \geq -bc^2/2$. Similar arguments for $y < 0$ complete the proof. ☐

Lemma 2.4. *Let $(\varphi_\alpha)_{\alpha>0}$ be a family of increasing and nonexpansive functions $\varphi_\alpha\colon \mathbb{R} \to \mathbb{R}$ with $\varphi_\alpha(0) = 0$. The following statements are equivalent:*

(1) $\forall x \in \mathbb{R}\colon \lim_{\alpha\to 0} \varphi_\alpha(x) = x$.

(2) $\forall x \in \mathbb{R}$: $\lim_{\alpha\to 0} \varphi_\alpha^{-1}(x) = \{x\}$.

Here the limit in (2) is defined by $\lim_{\alpha\to 0} \varphi_\alpha^{-1}(x) = \{x\} :\Leftrightarrow \lim_{\alpha\to 0} \sup_{y\in\varphi_\alpha^{-1}(x)} |y-x| = 0$, *and we use the convention* $\sup_{y\in\emptyset} |y-x| := \infty$.

Proof. Since φ_α is increasing and $\varphi_\alpha(0) = 0$, the pre-image is either empty or a closed interval of the form $\varphi_\alpha^{-1}(x) = [p_\alpha, q_\alpha]$ depending on α where we have three cases depending on x. If $x \geq 0$ then $0 \leq p_\alpha \leq q_\alpha < \infty$, if $x \leq 0$ then $-\infty < p_\alpha \leq q_\alpha \leq 0$ and if $x = 0$ then $\infty < p_\alpha \leq 0 \leq q_\alpha < \infty$. Furthermore, since φ_α is nonexpansive we have $|p_\alpha|, |q_\alpha| \geq |x|$.

Suppose (1) holds. By assumption we have $0 \in \varphi_\alpha^{-1}(0)$. Let $x \neq 0$ and suppose there is a null sequence α_k such that $\forall k \in \mathbb{N}$: $\varphi_{\alpha_k}^{-1}(x) = \emptyset$. Since φ_α is nonexpansive, we have $|\varphi_{\alpha_k}(2x)| < |x|$ and therefore $|\varphi_{\alpha_k}(2x) - 2x| > |x|$ for all $k \in \mathbb{N}$, which is a contradiction. Thus, for all $x \in \mathbb{R}$ there exists $\tilde{\alpha} > 0$ such that $\forall \alpha \in (0, \tilde{\alpha})$: $\varphi_\alpha^{-1}(x) \neq \emptyset$. Now fix any $x > 0$. Then for all $\alpha < \tilde{\alpha}$

$$\sup_{y\in\varphi_\alpha^{-1}(x)} |y-x| = \sup_{y\in[p_\alpha,q_\alpha]} |y-x| = |q_\alpha - x|$$

Suppose that $q_\alpha \nrightarrow x$, then $\limsup_{\alpha\to 0} s_\alpha =: q > x$ (and $q < \infty$, otherwise there would exist a null sequence α_k such that $\lim_{k\to\infty} \varphi_{\alpha_k}(y) = y \leq x$ for all $y \in \mathbb{R}$, which is a contradiction). Then there exists a null sequence α_k such that $|\varphi_{\alpha_k}(q_{\alpha_k}) - \varphi_{\alpha_k}(q)| \leq |q_{\alpha_k} - q| \to 0$ as $k \to \infty$. But $|\varphi_{\alpha_k}(q_{\alpha_k}) - \varphi_{\alpha_k}(q)| = |x - \varphi_{\alpha_k}(q)| \to |x - q| > 0$, as $k \to \infty$, which is a contradiction.

To show the converse implication, suppose that (2) holds and let $x \in \mathbb{R}$ be fixed. With $z := \varphi_\alpha(x)$, we have $x \in \varphi_\alpha^{-1}(z)$ and $|\varphi_\alpha(x) - x| = |z - x| \leq \sup_{x\in\varphi_\alpha^{-1}(z)} |z - x| \to 0$. ◻

Lemma 2.5. *Let* $s \in \Gamma_0(\mathbb{R})$ *and* $\alpha > 0, \gamma, \kappa > 0$ *such that* $\gamma\kappa^2 < 1$. *Suppose there exist* $t \in [0,1)$ *such that*

$$\forall x \in \mathbb{R}: \quad |x| \leq \alpha/\kappa \Rightarrow |\operatorname{prox}_s(x)| \leq \frac{\gamma\kappa^2 t}{1 - t(1-\gamma\kappa^2)} |x| .$$

Then for all $x \in \mathbb{R}$ *the following hold*

(a) $|x| \leq \gamma\alpha \Rightarrow |\operatorname{prox}_{\gamma s(\kappa(\cdot))}(x)| \leq t|x|$.

(b) $|x| > \gamma\alpha \Rightarrow |\operatorname{prox}_{\gamma s(\kappa(\cdot))}(x) - x| > (1-t)\gamma\alpha$.

Proof. Let $y \in \mathbb{R}$ be arbitrary. We have

$$\begin{aligned}\operatorname{prox}_r^{-1}(y) &= y + \frac{1}{\gamma}\partial\gamma s(y) = y + \frac{1}{\gamma\kappa}\partial(\gamma s(\kappa(\cdot)))(y/\kappa) \\ &= y + \frac{1}{\gamma\kappa}(\operatorname{prox}_{\gamma s(\kappa(\cdot))}^{-1} - \operatorname{id})(y/\kappa) = y - y/(\gamma\kappa^2) + \frac{1}{\gamma\kappa}\operatorname{prox}_{\gamma s(\kappa(\cdot))}^{-1}(y/\kappa).\end{aligned}$$

Now let $|x| \leq \gamma\alpha$ and define $a := \kappa \operatorname{prox}_{\gamma r(\kappa(\cdot))}(x)$. Then,

$$b := \kappa\big(1 - 1/(\gamma\kappa^2)\big)\operatorname{prox}_{\gamma s(\kappa(\cdot))}(x) + x/(\gamma\kappa) \in \operatorname{prox}_s^{-1}(a)$$

and $|b| = (\gamma\kappa)^{-1}|x - (1-\kappa^2\gamma)\operatorname{prox}_{\gamma s(\kappa(\cdot))}(x)| \leq |x|/(\gamma\kappa) \leq \alpha/\kappa$. Thus, we have

$$\begin{aligned}\left|\operatorname{prox}_{\gamma s(\kappa(\cdot))}(x)\right| = \left|\frac{a}{\kappa}\right| = \frac{1}{\kappa}|\operatorname{prox}_s(b)| &\leq \frac{\gamma\kappa t}{1 - t(1-\gamma\kappa^2)}|b| \\ &\leq \frac{t}{1-t(1-\gamma\kappa^2)}\big(|x| - (1-\gamma\kappa^2)|\operatorname{prox}_{\gamma s(\kappa(\cdot))}(x)|\big),\end{aligned}$$

from which it follows that $t^{-1}|\operatorname{prox}_{\gamma s(\kappa(\cdot))}(x)| \leq |x|$. ◻

3 Non-linear diagonal frame filtering

Throughout this paper $\mathbb{X}$, $\mathbb{Y}$ denote Hilbert spaces and $\mathbf{A}\colon \mathbb{X} \to \mathbb{Y}$ is a bounded linear operator with an unbounded Moore-Penrose inverse $\mathbf{A}^+$. We assume that $\mathbf{A}$ has a diagonal frame decomposition (DFD) $(\mathbf{u}, \mathbf{v}, \boldsymbol{\kappa}) = (u_\lambda, v_\lambda, \kappa_\lambda)_{\lambda\in\Lambda}$ where $\sup_{\lambda\in\Lambda} \kappa_\lambda < \infty$.

3.1 Non-linear filtered DFD

The following two definitions are central for this work.

Definition 3.1 (Non-linear regularizing filter)**.** We call a family $(\varphi_\alpha)_{\alpha>0}$ of functions $\varphi_\alpha\colon \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ a non-linear regularizing filter if for all $\alpha, \kappa > 0$, the following holds

- (F1) $\varphi_\alpha(\kappa, \cdot)$ is monotonically increasing.
- (F2) $\varphi_\alpha(\kappa, \cdot)$ is nonexpansive.
- (F3) $\varphi_\alpha(\kappa, 0) = 0$.
- (F4) $\forall c \in \mathbb{R}\colon \lim_{\alpha\to 0} \varphi_\alpha(\kappa, c) = c$.

The properties required in the definition of non-linear regularizing filters are quite natural. Recall that the field of imaging often exploits the fact that natural images or signals, other than noise, are sparse in certain frames, which means that the majority of the coefficients are zero. Assuming the original coefficients to be small and the noisy coefficient to be zero, it makes sense to leave them at zero after filtering. Furthermore, without specific knowledge of the noise structure, it seems reasonable to preserve the order after coefficient filtering which is the monotonicity. In addition, the distance between the original coefficients should not exceed the distance between the noisy coefficients, which is nonexpansive. The last property is a technical one to show the convergence of the non-linear filtered DFD defined next to an inverse of $\mathbf{A}$.

Definition 3.2 (Non-linear filtered DFD)**.** Let $(\varphi_\alpha)_{\alpha>0}$ be a non-linear regularizing filter. The non-linear filtered DFD $(\mathbf{B}_\alpha)_{\alpha>0}$ with $\mathbf{B}_\alpha\colon\ \mathrm{dom}(\mathbf{B}_\alpha) \to \mathbb{X}$ is defined by

$$\mathrm{dom}(\mathbf{B}_\alpha) \coloneqq \{y \in \mathbb{Y} \mid \Phi_{\alpha,\boldsymbol{\kappa}} \circ \mathbf{T}_{\mathbf{v}}^*(y) \in \mathrm{dom}(\mathbf{M}_{\boldsymbol{\kappa}}^+)\}\,, \tag{3.1}$$

$$\mathbf{B}_\alpha(y) \coloneqq \sum_{\lambda\in\Lambda} \frac{1}{\kappa_\lambda} \varphi_\alpha(\kappa_\lambda, \langle y, v_\lambda\rangle)\bar{u}_\lambda = \mathbf{T}_{\bar{\mathbf{u}}} \circ \mathbf{M}_{\boldsymbol{\kappa}}^+ \circ \Phi_{\alpha,\boldsymbol{\kappa}} \circ \mathbf{T}_{\mathbf{v}}^*(y)\,, \tag{3.2}$$

where $\Phi_{\alpha,\boldsymbol{\kappa}}\colon \ell^2(\Lambda) \to \ell^2(\Lambda)\colon (c_\lambda)_{\lambda\in\Lambda} \mapsto (\varphi_\alpha(\kappa_\lambda, c_\lambda))_{\lambda\in\Lambda}$.

According to (F2), $\Phi_{\alpha,\boldsymbol{\kappa}}$ is well-defined and nonexpansive.

Remark 3.3 (Linear case)**.** If $\varphi_\alpha(\kappa, c) = \kappa f_\alpha(\kappa) c$ for a linear regularizing filter $(f_\alpha)_{\alpha>0}$ (as defined in [9, Definition 3.1]), then φ_α is linear in the second component and (3.2) reduces to the linear filtered DFD. If $f_\alpha \geq 0$, $\sup\{|\kappa f_\alpha(\kappa)| \mid \alpha, \kappa > 0\} \leq 1$ and $\lim_{\alpha\to 0} f_\alpha(\kappa) = 1/\kappa$, then $(\varphi_\alpha)_{\alpha>0}$ satisfies (F1)-(F4). Linear filtered DFDs has been analyzed and shown to be a regularization method in [9]. Here we extend the analysis to filters that are non-linear in the second component.

The goal of this paper is to show stability and convergence of $(\mathbf{B}_\alpha)_{\alpha>0}$. Since $\mathbf{T}_{\bar{\mathbf{u}}}$ and $\mathbf{T}_{\mathbf{v}}^*$ are continuous, according to (2.1), (3.2) it is sufficient to analyze stability and convergence of $(\mathbf{M}_{\boldsymbol{\kappa}}^+ \circ \Phi_{\alpha,\boldsymbol{\kappa}})_{\alpha>0}$.

3.2 Filters as proximity operators

In the next lemma we demonstrate that regularizing filters are proximity operators of proper, convex, and lower semi-continuous functionals.

Lemma 3.4 (Filters are proximity operators)**.** *Let $(\varphi_\alpha)_{\alpha>0}$ satisfy (F1)-(F3) and set $\varphi_{\alpha,\lambda} \coloneqq \varphi_\alpha(\kappa_\lambda,\cdot)$. Then the following hold*

(a) There exist $s_{\alpha,\lambda} \in \Gamma_0(\mathbb{R})$ with $s_{\alpha,\lambda} \geq s_{\alpha,\lambda}(0) = 0$ such that $\varphi_{\alpha,\lambda} = \operatorname{prox}_{s_{\alpha,\lambda}}$.

(b) $\mathcal{R}_\alpha \coloneqq \bigoplus_{\lambda\in\Lambda} s_{\alpha,\lambda}(\kappa_\lambda(\cdot))$ is non-negative and contained in $\Gamma_0(\ell^2(\Lambda))$ for all $\alpha > 0$.

(c) $\forall (z_\lambda)_{\lambda\in\Lambda} \in \ell^2(\Lambda)\colon \operatorname{prox}_{\mathcal{R}_\alpha}((z_\lambda)_{\lambda\in\Lambda}) = (\operatorname{prox}_{s_{\alpha,\lambda}(\kappa_\lambda(\cdot))}(z_\lambda))_{\lambda\in\Lambda}$.

(d) For all $z \in \operatorname{dom}(\mathbf{M}_{\boldsymbol{\kappa}}^+ \circ \Phi_{\alpha,\boldsymbol{\kappa}})$ and any fixed $\gamma > 0$ we have

$$\mathbf{M}_{\boldsymbol{\kappa}}^+ \circ \Phi_{\alpha,\boldsymbol{\kappa}}(z) = \operatorname*{argmin}_{x\in\ell^2(\Lambda)} \|\mathbf{M}_{\boldsymbol{\kappa}} x - z\|^2/2 + \mathcal{R}_\alpha(x)\,, \tag{3.3}$$

$$\mathbf{M}_{\boldsymbol{\kappa}}^+ \circ \Phi_{\alpha,\boldsymbol{\kappa}}(z) = \operatorname{Fix}\bigl(\operatorname{prox}_{\gamma\mathcal{R}_\alpha} \circ (\operatorname{id} - \gamma\nabla\|\mathbf{M}_{\boldsymbol{\kappa}}(\cdot) - z\|^2/2)\bigr)\,. \tag{3.4}$$

Proof. Since $\varphi_{\alpha,\lambda}\colon \mathbb{R}\to\mathbb{R}$ is increasing and nonexpansive, according to a Lemma 2.1(e), $\varphi_{\alpha,\lambda} = \operatorname{prox}_{s_{\alpha,\lambda}}$ with $s_{\alpha,\lambda} \in \Gamma_0(\mathbb{R})$. By (F3), $\operatorname{prox}_{s_{\alpha,\lambda}}(0) = 0$ which means $0 \in \operatorname{argmin} s_{\alpha,\lambda}$ and thus $s_{\alpha,\lambda}(0) \geq s_{\alpha,\lambda}$. In particular, we can choose $s_{\alpha,\lambda}$ with $s_{\alpha,\lambda}(0) = 0$ which yields (a). Clearly $\mathcal{R}_\alpha \coloneqq \bigoplus_{\lambda\in\Lambda} s_{\alpha,\lambda}(\kappa_\lambda(\cdot))$ is non-negative. By Lemma 2.1(f), $\mathcal{R}_\alpha \in \Gamma_0(\ell^2(\Lambda))$ and $(\operatorname{prox}_{s_{\alpha,\lambda}(\kappa_\lambda(\cdot))}(z_\lambda))_{\lambda\in\Lambda} = \operatorname{prox}_{\mathcal{R}_\alpha}((z_\lambda)_{\lambda\in\Lambda})$ which shows (b), (c). One easily verifies $\kappa_\lambda^{-1}\varphi_\alpha(\kappa_\lambda, z_\lambda) = \operatorname{argmin}_x |\kappa_\lambda x - z_\lambda|^2/2 + s_{\alpha,\lambda}(\kappa_\lambda x)$. Thus, for $\mathbf{M}_{\boldsymbol{\kappa}}^+ \circ \Phi_{\alpha,\boldsymbol{\kappa}}(z) \in \ell^2(\Lambda)$,

$$\begin{aligned}
\mathbf{M}_{\boldsymbol{\kappa}}^+ \circ \Phi_{\alpha,\boldsymbol{\kappa}}(z) &= \bigl(\kappa_\lambda^{-1}\varphi_\alpha(\kappa_\lambda, z_\lambda)\bigr)_{\lambda\in\Lambda} \\
&= \bigl(\operatorname*{argmin}_{x_\lambda\in\mathbb{R}}(|\kappa_\lambda x_\lambda - z_\lambda|^2/2 + s_{\alpha,\lambda}(\kappa_\lambda x_\lambda))\bigr)_{\lambda\in\Lambda} \\
&= \operatorname*{argmin}_{x\in\ell^2(\Lambda)}\Bigl(\sum_{\lambda\in\Lambda} |\kappa_\lambda x_\lambda - z_\lambda|^2/2 + s_{\alpha,\lambda}(\kappa_\lambda x_\lambda)\Bigr) \\
&= \operatorname*{argmin}_{x\in\ell^2(\Lambda)} \|\mathbf{M}_{\boldsymbol{\kappa}} x - z\|^2/2 + \mathcal{R}_\alpha(x)\,,
\end{aligned}$$

which is (3.3). Since $\sup_{\lambda\in\Lambda}\kappa_\lambda < \infty$, the operator $\mathbf{M}_{\boldsymbol{\kappa}}$ is bounded and thus $\|\mathbf{M}_{\boldsymbol{\kappa}}(\cdot) - z\|^2/2$ is convex and Fréchet differentiable. Together with Lemma 2.1(c) we get (3.4). □

Let $(\varphi_\alpha)_{\alpha>0}$ be a non-linear regularizing filter. Unless otherwise stated, we denote by $s_{\alpha,\lambda}$ and $\mathcal{R}_\alpha$ the functionals induced by $(\varphi_\alpha)_{\alpha>0}$ according to Lemma 3.4. According to Lemma 2.1(c), the domain of $s_{\alpha,\lambda}$ is a closed interval and contains zero. In particular, $s_{\alpha,\lambda}$ is continuous on the interior of $\operatorname{dom}(s_{\alpha,\lambda})$ and

$$\operatorname{ran}(\varphi_{\alpha,\lambda}) = \operatorname{ran}(\operatorname{prox}_{s_{\alpha,\lambda}}) \subseteq \operatorname{dom}(s_{\alpha,\lambda})\,,$$

where $\operatorname{ran}(\varphi_{\alpha,\lambda})$ is also an interval containing zero. For all $z \in \operatorname{dom}(\mathbf{M}_{\boldsymbol{\kappa}}^+ \circ \Phi_{\alpha,\boldsymbol{\kappa}})$, Equation (3.3) implies that $\mathbf{M}_{\boldsymbol{\kappa}}^+ \circ \Phi_{\alpha,\boldsymbol{\kappa}}(z) \in \operatorname{dom}(\mathcal{R}_\alpha)$.

Lemma 3.5. *Let $(\varphi_\alpha)_{\alpha>0}$ be a non-linear regularizing filter and $\alpha > 0$. Then for all $x = (x_\lambda)_\lambda \in \ell^2(\Lambda)$ with $\kappa_\lambda x_\lambda \in \operatorname{dom}(s_{\alpha,\lambda})$ for all $\lambda \in \Lambda$, the following holds:*

$$\mathcal{R}_\alpha(x) = \left\| \left(\int_0^{\kappa_\lambda x_\lambda} \max(\varphi_{\alpha,\lambda}^{-1}(y))dy \right)_{\lambda\in\Lambda} \right\|_1 - \frac{1}{2}\|\mathbf{M}_{\boldsymbol{\kappa}} x\|_2^2\,. \tag{3.5}$$

In particular, $x \in \operatorname{dom}(\mathcal{R}_\alpha)$ if and only if $\bigl(\int_0^{\kappa_\lambda x_\lambda} \max(\varphi_{\alpha,\lambda}^{-1}(y))dy\bigr)_{\lambda\in\Lambda} \in \ell^1(\Lambda)$, where we set $\max(\emptyset) \coloneqq \infty$.

Proof. Recall $\mathcal{R}_\alpha = \bigoplus_{\lambda\in\Lambda} s_{\alpha,\lambda}(\kappa_\lambda(\cdot))$, where $s_{\alpha,\lambda}|_{\mathrm{dom}(s_{\alpha,\lambda})^\circ}$ is convex, continuous, and $s_{\alpha,\lambda} \geq s_{\alpha,\lambda}(0) = 0$ for all $\lambda \in \Lambda$. In particular, $s_{\alpha,\lambda}$ is locally Lipschitz continuous and, by Rademacher's theorem, differentiable almost everywhere. From $\mathrm{prox}_{s_{\alpha,\lambda}} = \varphi_{\alpha,\lambda}$ and Lemma 2.1(a) we get $\partial s_{\alpha,\lambda} = \varphi_{\alpha,\lambda}^{-1} - \mathrm{id}$. For $x_\lambda \in \mathrm{dom}(\partial s_{\alpha,\lambda}) \supseteq \mathrm{dom}(s_{\alpha,\lambda})^\circ$, define $q'_{\alpha,\lambda}(x_\lambda) \coloneqq \max(\partial s_{\alpha,\lambda}(x_\lambda)) = \max(\varphi_{\alpha,\lambda}^{-1}(x_\lambda) - x_\lambda)$. By the fundamental theorem of calculus we have

$$s_{\alpha,\lambda}(\kappa_\lambda x_\lambda) = \int_0^{\kappa_\lambda x_\lambda} q'_{\alpha,\lambda}(y)\, dy = \int_0^{\kappa_\lambda x_\lambda} \max(\varphi_{\alpha,\lambda}^{-1}(y))dy - \frac{1}{2}(\kappa_\lambda x_\lambda)^2 \geq 0\,.$$

Summation over $\lambda \in \Lambda$ gives (3.5). □

4 Convergence Analysis

In this section we investigate stability and convergence of $(\mathbf{M}_{\boldsymbol{\kappa}}^+ \circ \Phi_{\alpha,\boldsymbol{\kappa}})_{\alpha>0}$, which, according to (3.2), also applies to $(\mathbf{B}_\alpha)_{\alpha>0}$. To do so, we consider three cases of additional assumptions on the filter $(\varphi_\alpha)_{\alpha>0}$. The first two cases relate the filtered DFD to known regularization concepts and the third case directly analyzed the regularizing filters.

4.1 Case A: Reduction to variational regularization

Variational regularization uses minimizers of the generalized Tikhonov functional $\mathcal{T}_{\alpha,y}(x) = \|\mathbf{A}x - y\|^2/2 + \alpha\mathcal{R}(x)$, where $\mathcal{R}$ is a regularizing functional and $\alpha > 0$ the regularization parameter. It is well-investigated by numerous works, such as [11, 12, 23, 26]. In this paper we use the following convergence result.

Lemma 4.1 (Variational regularization)**.** *Let $\mathcal{R} \in \Gamma_0(\mathbb{X})$ be norm-coercive. Suppose $\alpha_k, \alpha > 0$ and $y, y^k \in \mathbb{Y}$ with $y^k \to y$.*

- *Existence: $\mathcal{T}_{\alpha,y}$ has at least one minimizer.*
- *Stability: For $x^k \in \operatorname{argmin} \mathcal{T}_{\alpha,y^k}$, there exists a subsequence $(x^{k(\ell)})_{\ell\in\mathbb{N}}$ and some $x_\alpha \in \operatorname{argmin} \mathcal{T}_{\alpha,y}$ with $x^{k(\ell)} \rightharpoonup x_\alpha$ and $\mathcal{R}(x^{k(\ell)}) \to \mathcal{R}(x_\alpha)$ as $\ell \to \infty$. If the minimizer of $\mathcal{T}_{\alpha,y}$ is unique, then $x^k \rightharpoonup x_\alpha$ as $k \to \infty$.*
- *Convergence: Assume $y \in \mathrm{ran}(\mathcal{R})$ and $\|y^k - y\| \leq \delta_k$ and $\delta_k, \alpha_k, \delta_k^2/\alpha_k \to 0$. For $x^k \in \operatorname{argmin} \mathcal{T}_{\alpha_k,y^k}$ there exists a subsequence $(x^{k(\ell)})_{\ell\in\mathbb{N}}$ and a solution x^+ of $\mathbf{A}x = y$ with $x^{k(\ell)} \rightharpoonup x^+$ and $\mathcal{R}(x^{k(\ell)}) \to \mathcal{R}(x^+)$. If the solution is unique, then $x^k \rightharpoonup x^+$.*

Proof. See [26, Theorems 3.22, 3.23, 3.26]. □

We will show that variational regularization already covers a large class of non-linear filter-based methods. An example of this relation was provided in [14], where it was shown that the soft thresholding filter yields a regularization method.

Assumption A (Reduction to variational regularization)**.** *Let $(\varphi_\alpha)_{\alpha>0}$ be a non-linear regularizing filter which satisfies the following conditions:*

(A1) $\forall \alpha > 0\ \forall \lambda \in \Lambda\ \forall y \in \mathbb{R}$: $\varphi_{\alpha,\lambda}^{-1}(y) = y + \alpha I_\lambda(y)$, where $I_\lambda(y)$ is a closed interval.

(A2) For some $\alpha > 0$ there exist $b, c > 0$ such that for all $\lambda \in \Lambda$, we have

$$\forall x \in \mathbb{R}: \quad |x| \leq \min(\varphi_{\alpha,\lambda}^{-1}(c\kappa_\lambda)) \Rightarrow |\varphi_{\alpha,\lambda}(x)| \leq \frac{\kappa_\lambda^2}{\kappa_\lambda^2 + \alpha b}|x|\,.$$

Assumption A implies that $\varphi_{\alpha,\lambda}$ is surjective. If this were not the case, there would exist $y \in \mathbb{R}$ with $I_\lambda(y) = \emptyset$ and then $\lim_{\alpha\to 0} \varphi_{\alpha,\lambda}^{-1}(y) = \emptyset$, which contradicts Lemma 2.4.

Remark 4.2 (Generation from φ_1)**.** Opposed to the general case, filters satisfying Assumption A are uniquely determined by a single filter function $\varphi_1 \colon \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$. Let $\varphi_1(\kappa,\cdot)$ be monotonically increasing and nonexpansive with $\varphi_1 \geq \varphi_1(\kappa, 0) = 0$ for all $\kappa > 0$, and assume there exist $b, c > 0$ such that $|\varphi_1(\kappa, x)| \leq |x|\, \kappa^2/(\kappa^2 + b)$ for all $\kappa > 0$ and $x \in \mathbb{R}$ with $|x| \leq \min(\varphi_1(\kappa,\cdot)^{-1}(c\kappa))$. Then, φ_α is uniquely defined by

$$\varphi_\alpha(\kappa, x) = ((1-\alpha)\operatorname{id} + \alpha\varphi_1(\kappa,\cdot)^{-1})^{-1}(x)\,.$$

For all $\alpha \in (0,1)$ the function φ_α is well defined, and, by Lemma 2.4, the family $(\varphi_\alpha)_{\alpha>0}$ is a non-linear regularizing filter. Define for all $\lambda \in \Lambda$ the closed interval $I_\lambda(y) = \varphi_{1,\lambda}^{-1}(y) - y$, then we easily see that $\varphi_{\alpha,\lambda}^{-1}(y) = y + \alpha I_\lambda(y)$.

Example 4.3. For fixed $b, d > 0$ consider the function

$$\varphi_1(\kappa, x) = \begin{cases} \frac{\kappa^2}{\kappa^2+b}\, x & |x| \leq d/\kappa \\ x - \operatorname{sign}(x)\frac{d\,b}{\kappa(\kappa^2+b)} & |x| > d/\kappa\,. \end{cases}$$

Then, for all $\kappa > 0$, $\varphi_1(\kappa,\cdot)$ is monotonically increasing, nonexpansive, and $\varphi_1(\kappa, 0) = 0$. Set $c = d/(\max_\lambda \kappa_\lambda^2 + b)$. Then, $\varphi_1(\kappa_\lambda,\cdot)^{-1}(c\kappa_\lambda) = ((b + \kappa_\lambda^2)/\kappa_\lambda^2)\, c\kappa_\lambda \leq d/\kappa_\lambda$ and by definition, $|\varphi_1(\kappa_\lambda, x)| \leq \kappa_\lambda^2/(\kappa_\lambda^2 + b)|x|$ for all $|x| \leq d/\kappa_\lambda$. By Remark 4.2

$$\varphi_\alpha(\kappa, x) = \begin{cases} \frac{\kappa^2}{\kappa^2+\alpha b}\, x & |x| \leq d\frac{\kappa^2+\alpha b}{\kappa(\kappa^2+b)} \\ x - \operatorname{sign}(x)\frac{db\alpha}{\kappa(\kappa^2+b)} & |x| > d\frac{\kappa^2+\alpha b}{\kappa(\kappa^2+b)}. \end{cases}$$

defines a non-linear regularizing filter $(\varphi_\alpha)_{\alpha>0}$ satisfying Assumption A.

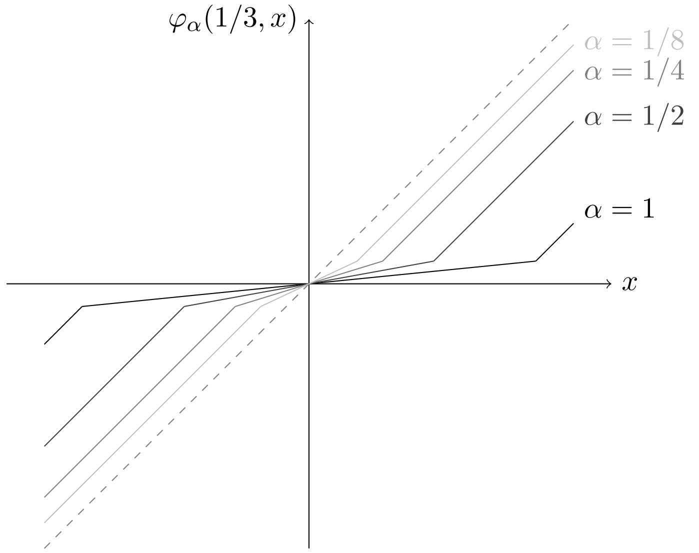

Figure 1: Illustration of the filter $(\varphi_\alpha)_{\alpha>0}$ from Example 4.3 that is generated by a single filter function φ_1 and satisfies Assumption A.

The following Lemma reduces the non-linear filtered DFD to variational regularization.

Lemma 4.4 (Reduction to variational regularization)**.** *Let $(\varphi_\alpha)_{\alpha>0}$ be a non-linear regularizing that satisfies Assumption A. Then* $\operatorname{dom}(\mathbf{M}_{\boldsymbol{\kappa}}^+ \circ \Phi_{\alpha,\boldsymbol{\kappa}}) = \ell^2(\Lambda)$ *and there exist* $s_\lambda \in \Gamma_0(\mathbb{R})$ *with* $s_\lambda(0) = 0$ *and* $\varphi_{\alpha,\lambda} = \operatorname{prox}_{\alpha s_\lambda}$*. Define* $\mathcal{R} = \bigoplus_{\lambda\in\Lambda} s_\lambda(\kappa_\lambda(\cdot))$*, then* $\mathcal{R} \in \Gamma_0(\ell^2(\Lambda))$*,* $\mathcal{R}$ *is norm-coercive and*

$$\forall z \in \ell^2(\Lambda): \quad \mathbf{M}_{\boldsymbol{\kappa}}^+ \circ \Phi_{\alpha,\boldsymbol{\kappa}}(z) = \operatorname*{argmin}_{x\in\ell^2(\Lambda)} \frac{1}{2}\|\mathbf{M}_{\boldsymbol{\kappa}} x - z\|^2 + \alpha\mathcal{R}(x)\,. \tag{4.1}$$

Proof. The identity $\operatorname{dom}(\mathbf{M}_{\boldsymbol{\kappa}}^+ \circ \Phi_{\alpha,\boldsymbol{\kappa}}) = \ell^2(\Lambda)$ will be shown shown under more general assumptions in Lemma 4.13 below. By Lemma 3.4, there exist $f_{\alpha,\lambda} \in \Gamma_0(\mathbb{R})$ with $f_{\alpha,\lambda} \geq f_{\alpha,\lambda}(0) = 0$ such that $\varphi_{\alpha,\lambda} = \operatorname{prox}_{f_{\alpha,\lambda}} = (\operatorname{id} + \partial f_{\alpha,\lambda})^{-1}$, so

$$y + \partial f_{\alpha,\lambda}(y) = \varphi_{\alpha,\lambda}^{-1}(y) = y + \alpha I_\lambda(y).$$

In particular, $\partial(f_{\alpha,\lambda}/\alpha)(y) = I_\lambda(y)$ is independent of α. With $s_\lambda \coloneqq f_{\alpha,\lambda}/\alpha$ we have $\varphi_{\alpha,\lambda} = (\operatorname{id} + \partial(\alpha s_\lambda))^{-1} = \operatorname{prox}_{\alpha s_\lambda}$. Note that s_λ is also proper, convex, and lower semi-continuous and $s_\lambda(0) = 0$. Moreover, $\mathcal{R} = \bigoplus_{\lambda\in\Lambda} s_\lambda(\kappa_\lambda(\cdot))$ is positive, $\mathcal{R} \in \Gamma_0(\ell^2(\Lambda))$ and satisfies (4.1). To show the norm-coercivity of $\mathcal{R}$ consider Lemma 2.3. For $y = (y_\lambda)_\lambda \in \ell^2(\Lambda)$,

$$\begin{aligned}
\mathcal{R}(y) &= \sum_{\lambda\in\Lambda} s_\lambda(\kappa_\lambda y_\lambda) = \sum_{|\kappa_\lambda y_\lambda| \leq c\kappa_\lambda} s_\lambda(\kappa_\lambda y_\lambda) + \sum_{|\kappa_\lambda y_\lambda| > c\kappa_\lambda} s_\lambda(\kappa_\lambda y_\lambda) \\
&\geq \sum_{|y_\lambda|\leq c} b|y_\lambda|^2/2 + \sum_{|y_\lambda|>c} \left(bc|y_\lambda| - bc^2/2\right) \geq \frac{b}{2} \sum_{|y_\lambda|\leq c} |y_\lambda|^2 + (bc^2/2)\,|\{\lambda \mid |y_\lambda| > c\}|\,.
\end{aligned}$$

Now let $(y^k)_k \in \operatorname{dom}(\mathcal{R})^{\mathbb{N}}$ be a sequence with $\|y^k\|_2 \to \infty$ as $k \to \infty$ and define $N_k \coloneqq |\{\lambda \mid |y_\lambda^k| > c\}|$. We now show that $(y^k)_k$ can be covered by subsequences τ_i with $\mathcal{R}(y^{\tau_i(k)})_k \to \infty$ for $i = 1, 2, 3$. For all subsequences $(\tau_1(k))_k$ with $\left\|y^{\tau_1(k)}\right\|_\infty \to \infty$ we have for k large enough that

$$\mathcal{R}(y^{\tau_1(k)}) \geq bc \sum_{|y_\lambda^{\tau_1(k)}|>c} \left(|y_\lambda^{\tau_1(k)}| - \frac{c}{2}\right) \geq bc\left(\|y^{\tau_1(k)}\|_\infty - \frac{c}{2}\right) \to \infty.$$

For subsequences $(\tau_2(k))_k$ with $N_{\tau_2(k)} \to \infty$ we have $\mathcal{R}(y^{\tau_2(k)}) \geq bc^2 N_{\tau_2(k)}/2 \to \infty$. Finally, the rest are subsequences $(\tau_3(k))_k$, where there exists a constant $M > 0$ such that $\|y^{\tau_3(k)}\|_\infty \leq M$ and $N_{\tau_3(k)}$ is bounded, then

$$\mathcal{R}(y^{\tau_3(k)}) \geq \frac{b}{2}\left(\|y^{\tau_3(k)}\|_2^2 - \sum_{|y_\lambda^{\tau(k)}|>c} |y_\lambda^{\tau(k)}|^2\right) \geq \frac{b}{2}\|y^{\tau_3(k)}\|_2^2 - \frac{b}{2} M N_{\tau_3(k)} \to \infty.$$

Hence $\mathcal{R}(y^k) \to \infty$, which shows that $\mathcal{R}$ is coercive. □

Due to Assumption A, the function $s_{\alpha,\lambda}$ is continuous on $\mathbb{R}$ and we have a simpler formulation for the domain of $\mathcal{R}$. In fact, in the setting of Lemma 4.4,

$$x \in \operatorname{dom}(\mathcal{R}) \Leftrightarrow \left(\int_0^{\kappa_\lambda x_\lambda} \max(I_\lambda(y))\, dy\right)_{\lambda\in\Lambda} \in \ell^1(\Lambda)\,.$$

Moreover, from Lemmas 4.4 and 4.1 we obtain the following.

Proposition 4.5 (non-linear filtered DFD is a regularization, Case A)**.** *Let* $(\varphi_\alpha)_{\alpha>0}$ *be a non-linear regularizing filter such that Assumption A is satisfied. Suppose* $\alpha > 0$*, and* $z, z^k \in \ell^2(\Lambda)$ *for* $k \in \mathbb{N}$ *with* $(z^k)_{k\in\mathbb{N}} \to z$ *and* $c_\alpha \coloneqq \mathbf{M}_{\boldsymbol{\kappa}}^+ \circ \Phi_{\alpha,\boldsymbol{\kappa}}(z)$.

- *Existence:* $\operatorname{dom}(\mathbf{M}_{\boldsymbol{\kappa}}^+ \circ \Phi_{\alpha,\boldsymbol{\kappa}}) = \ell^2(\Lambda)$.
- *Stability: With* $c^k \coloneqq \mathbf{M}_{\boldsymbol{\kappa}}^+ \circ \Phi_{\alpha,\boldsymbol{\kappa}}(z^k)$ *we have* $c^k \rightharpoonup c_\alpha$ *and* $\mathcal{R}(c^k) \to \mathcal{R}(c_\alpha)$ *as* $k \to \infty$.
- *Convergence: Let* $z = \mathbf{M}_{\boldsymbol{\kappa}} c^+$ *with* $c^+ \in \operatorname{dom}(\mathcal{R})$ *and* $\|z^k - z\| \le \delta_k$ *where* $\delta_k \to 0$*. Consider* $\alpha_k \to 0$ *such that* $\delta_k^2/\alpha_k \to 0$ *and define* $c^k = \mathbf{M}_{\boldsymbol{\kappa}}^+ \circ \Phi_{\alpha_k,\boldsymbol{\kappa}}(z^k)$*. Then* $c^k \rightharpoonup c^+$ *and* $\mathcal{R}(c^k) \to \mathcal{R}(c^+)$ *as* $k \to \infty$.

Proof. Follows from Lemmas 4.1 and 4.4 and the injectivity of $\mathbf{M}_{\boldsymbol{\kappa}}$. □

4.2 Case B: Reduction to PnP regularization

The concept in PnP regularization is to find a fixed point of the operator

$$\mathbf{T}_{\alpha,y} \coloneqq \mathbf{D}_\alpha \circ \left(\operatorname{id} - \gamma \nabla \|\mathbf{A}(\cdot) - y\|^2/2\right) \colon \mathbb{X} \to \mathbb{X}\,, \tag{4.2}$$

where $(\mathbf{D}_\alpha)_{\alpha>0}$ is a suitable family of regularization operators (or denoisers). The PnP framework has been applied successfully various fields including image restoration [7, 29] and inverse imaging problems [28, 27].

Definition 4.6 (Family of denoisers)**.** We call $(\mathbf{D}_\alpha)_{\alpha>0}$ admissible family of denoisers if the following hold:

(D1) $\forall \alpha > 0\colon \operatorname{Lip}(\mathbf{D}_\alpha) < 1$.

(D2) $\forall x \in \mathbb{E} \coloneqq \bigcup_{\alpha>0} \operatorname{ran}(\mathbf{D}_\alpha)\colon \mathbf{D}_\alpha(x) \to x$ as $\alpha \to 0$.

(D3) $\forall B \subseteq \mathbb{E}$ bounded $: \forall z \in \mathbb{X}\colon \sup_{x\in B} \langle \mathbf{D}_\alpha(x) - x, z\rangle \to 0$.

(D4) $\forall x \in \mathbb{E}\colon \exists M_x < \infty\colon \forall \alpha > 0\colon \|x - \mathbf{D}_\alpha(x)\|/(1 - \operatorname{Lip}(\mathbf{D}_\alpha)) \le M_x$.

In [10], the following results for PnP as a regularization method have been derived.

Lemma 4.7 (PnP regularization)**.** *Suppose* $(\mathbf{D}_\alpha)_{\alpha>0}$ *is an admissible family of denoisers, let* $\gamma \in (0, 2/\|\mathbf{A}\|^2)$ *and suppose* $y, y^k \in \mathbb{Y}$ *with* $y^k \to y$ *as* $k \to \infty$*. Then we have:*

- *Existence:* $\mathbf{T}_{\alpha,y}$ *has exactly one fixed point.*
- *Stability:* $\operatorname{Fix}(\mathbf{T}_{\alpha,y^k}) \to \operatorname{Fix}(\mathbf{T}_{\alpha,y})$ *as* $k \to \infty$.
- *Convergence: Assume* $y = \mathbf{A}x$ *for* $x \in \mathbb{E}$ *and suppose* $\|y^k - y\| \le \delta_k$ *where* $\delta_k \to 0$*. Consider* $\alpha_k \to 0$ *such that* $(1 - \operatorname{Lip}(\mathbf{D}_{\alpha_k}))/\operatorname{Lip}(\mathbf{D}_{\alpha_k}) \gtrsim \delta_k$ *and take* $x^k = \operatorname{Fix}(\mathbf{T}_{\alpha_k,y^k})$*. There exists a subsequence* $(x^{k(\ell)})_{l\in\mathbb{N}}$ *and a solution* x^+ *of* $\mathbf{A}x = y$ *such that* $x^{k(\ell)} \rightharpoonup x^+$*. If the solution is unique, then* $x^k \rightharpoonup x^+$.

According to Lemma 3.4, we will apply Proposition 4.7 with $\mathbf{D}_\alpha = \operatorname{prox}_{\gamma\mathcal{R}_\alpha}$ and $\mathbf{A} = \mathbf{M}_{\boldsymbol{\kappa}}$ to establish a regularization result for non-linear filtered DFD. For that purpose we assume the non-linear regularizing filter to satisfy the following assumption.

Assumption B (Reduction to PnP regularization)**.** *Let* $(\varphi_\alpha)_{\alpha>0}$ *be a non-linear regularizing filter and* $\gamma \in (0, 1/\max_\lambda \kappa_\lambda^2)$ *such that for all* $\alpha > 0$ *there exist* $\ell_\alpha \in (0,1)$ *with*

(B1) $\forall \lambda \in \Lambda\colon \operatorname{Lip}(\varphi_{\alpha,\lambda}) \le \gamma\kappa_\lambda^2 \ell_\alpha/(1 - \ell_\alpha(1 - \gamma\kappa_\lambda^2))$.

(B2) $\forall \lambda \in \Lambda\colon \exists C \in [1,\infty)\colon 1/(1-\ell_\alpha) \geq C \;\Rightarrow\; |\varphi_{\alpha,\lambda}(x)| \geq \frac{\gamma\kappa_\lambda^2(1-C(1-\ell_\alpha))}{1+(1-C(1-\ell_\alpha))(1-\gamma\kappa_\lambda^2)}|x|$.

Next we give an example of a family $(\varphi_\alpha)_{\alpha>0}$ that satisfies Assumption B and does not satisfy Assumption A.

Example 4.8. Set $\ell_\alpha = 1/(1+\alpha)$ and consider the function

$$\varphi_\alpha(\kappa, x) = \begin{cases} \frac{\gamma\kappa^2\ell_\alpha}{1-\ell_\alpha(1-\gamma\kappa^2)}\, x & |x| \leq \alpha/3 \\ \frac{\gamma\kappa^2\ell_\alpha}{1-\ell_\alpha(1-\gamma\kappa^2)}\, \operatorname{sign}(x)\alpha/3 & \alpha/3 \leq |x| \leq 2\alpha/3 \\ \frac{\gamma\kappa^2\ell_\alpha}{1-\ell_\alpha(1-\gamma\kappa^2)}\, (x - \operatorname{sign}(x)\alpha/3) & |x| \leq 2\alpha/3\,. \end{cases}$$

Then $(\varphi_\alpha)_{\alpha>0}$ is a non-linear regularizing filter and satisfies Assumption B. Since $(|\varphi_\alpha(\kappa, x)|)_{\alpha>0}$ is not monotone for some $x \in \mathbb{R}$, $(\varphi_\alpha)_{\alpha>0}$ does not satisfy Assumption A.

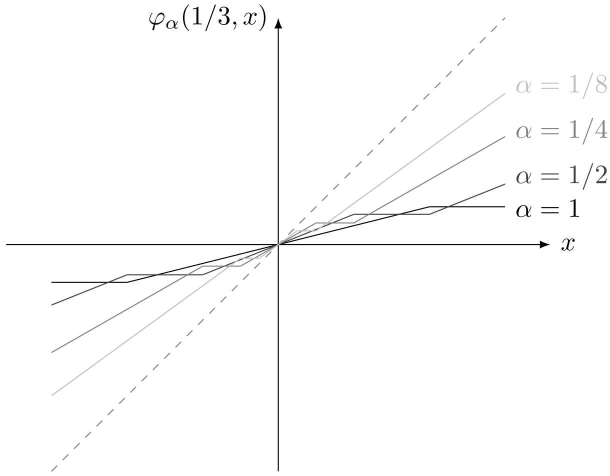

Figure 2: Illustration of the filter $(\varphi_\alpha)_{\alpha>0}$ from Example 4.8 that satisfies Assumption B and does not satisfy Assumption A.

The following Lemma is the key to the convergence in this section.

Lemma 4.9 (Reduction to PnP regularization)**.** *Let* $(\varphi_\alpha)_{\alpha>0}$ *be non-linear regularizing filter satisfying Assumption B. Then,* $(\operatorname{prox}_{\gamma\mathcal{R}_\alpha})_{\alpha>0}$ *is an admissible family of denoisers in the sense of Definition 4.6.*

Proof. Let $y \in \mathbb{R}$. For $s \in \Gamma_0(\mathbb{R})$ and $\gamma\kappa^2 < 1$ we have

$$\begin{aligned} \operatorname{prox}_s^{-1}(y) &= y + \frac{1}{\gamma}\partial\gamma s(y) = y + \frac{1}{\gamma\kappa}\partial(\gamma s(\kappa(\cdot)))\left(\frac{y}{\kappa}\right) \\ &= y + \frac{1}{\gamma\kappa}(\operatorname{prox}^{-1}_{\gamma s(\kappa\cdot)} - \mathrm{id})\left(\frac{y}{\kappa}\right) = y - \left(\frac{y}{\gamma\kappa^2}\right) + \frac{1}{\gamma\kappa}\operatorname{prox}^{-1}_{\gamma s(\kappa(\cdot))}\left(\frac{y}{\kappa}\right) \end{aligned}$$

and therefore $\operatorname{prox}^{-1}_{\gamma s(\kappa(\cdot))}(y) - y = \gamma\kappa(\operatorname{prox}_s^{-1}(\kappa y) - \kappa y)$. Now let $x \in \mathbb{R}$ be arbitrary. If $|\operatorname{prox}_s(x)| \geq (\gamma\kappa^2 t)/(1+t(1-\gamma\kappa^2)))|x|$ we derive $|z - \kappa y| \leq ((1-t)/(t\gamma\kappa))|x|$ for all $z \in \operatorname{prox}_s^{-1}(\kappa y)$. Similar to the proof of Lemma 2.5 it follows that

$$\left|x - \operatorname{prox}_{\gamma s(\kappa(\cdot))}(x)\right| \leq (1-t)|x|\,. \tag{4.3}$$

Also following the proof of Lemma 2.5 one can show that for all $t \in [0,1)$,

$$\forall x, y \in \mathbb{R}\colon\ |\operatorname{prox}_s(x) - \operatorname{prox}_s(y)| \leq \frac{\gamma\kappa^2 t}{1 - t(1-\gamma\kappa^2)}|x-y|$$
$$\Rightarrow \forall x, y \in \mathbb{R}\colon\ \left|\operatorname{prox}_{\gamma s(\kappa(\cdot))}(x) - \operatorname{prox}_{\gamma s(\kappa(\cdot))}(y)\right| \leq t|x-y|\,.$$

Let $c = (c_\lambda)_\lambda, d = (d_\lambda)_\lambda \in \ell^2(\Lambda)$ be arbitrary and $\alpha > 0$ be fixed. Then,

$$\begin{aligned}\|\operatorname{prox}_{\gamma\mathcal{R}_\alpha}(c) - \operatorname{prox}_{\gamma\mathcal{R}_\alpha}(d)\|^2 &= \sum_{\lambda\in\Lambda}\left|\operatorname{prox}_{\gamma s_{\alpha,\lambda}(\kappa_\lambda(\cdot))}(c_\lambda) - \operatorname{prox}_{\gamma s_{\alpha,\lambda}(\kappa_\lambda(\cdot))}(d_\lambda)\right|^2 \\ &\leq \sum_{\lambda\in\Lambda}\ell_\alpha^2\,|c_\lambda - d_\lambda|^2 = \ell_\alpha^2\|c-d\|^2.\end{aligned}$$

Since $\ell_\alpha^2 < 1$, the operator $\operatorname{prox}_{\gamma\mathcal{R}_\alpha}$ is a contraction for all $\alpha > 0$, which is (D1). To show (D2), let $x = (x_\lambda)_\lambda \in \ell^2(\Lambda)$. By assumption, for all $y \in \mathbb{R}$ and for all $\lambda \in \Lambda$, we have $\lim_{\alpha\to 0}\varphi_{\alpha,\lambda}(y) = y$ and by Lemma 2.4, we also have $\lim_{\alpha\to 0}\varphi_{\alpha,\lambda}^{-1}(y) = \{y\}$. It follows that

$$\operatorname{prox}_{\gamma s_{\alpha,\lambda}(\kappa_\lambda(\cdot))}^{-1}(y) = \kappa_\lambda\varphi_{\alpha,\lambda}^{-1}(\kappa_\lambda y) - \kappa_\lambda^2 y + y \to \{\kappa_\lambda^2 y - \kappa_\lambda^2 y + y\} = \{y\} \quad \text{ as } \alpha \to 0\,.$$

Since $\operatorname{prox}_{\gamma s_{\alpha,\lambda}(\kappa_\lambda(\cdot))}$ is increasing and nonexpansive with $\operatorname{prox}_{\gamma s_{\alpha,\lambda}(\kappa_\lambda(\cdot))}(0) = 0$, according to Lemma 2.4, it holds $\lim_{\alpha\to 0}\operatorname{prox}_{\gamma s_{\alpha,\lambda}(\kappa_\lambda(\cdot))}(y) = y$ for all $y \in \mathbb{R}$. Now let $\epsilon > 0$. Then there exists a finite subset $\Omega \subseteq \Lambda$ such that $\sum_{\lambda\in\Lambda\setminus\Omega}|x_\lambda|^2 < \epsilon/2$ and there exist $\tilde{\alpha} > 0$ small enough such that for all $\alpha < \tilde{\alpha}$, we have

$$\sum_{\lambda\in\Omega}\left|\operatorname{prox}_{\gamma s_{\alpha,\lambda}(\kappa_\lambda(\cdot))}(x_\lambda) - x_\lambda\right|^2 < \frac{\epsilon}{2}.$$

From Assumption 3.1 it follows that $|\operatorname{prox}_{\gamma s_{\alpha,\lambda}(\kappa_\lambda(\cdot))}(y) - y| \leq |y|$, hence we have

$$\begin{aligned}\|\operatorname{prox}_{\gamma\mathcal{R}_\alpha}(x) - x\|^2 &= \sum_{\lambda\in\Omega}\left|\operatorname{prox}_{\gamma s_{\alpha,\lambda}(\kappa_\lambda(\cdot))}(x_\lambda) - x_\lambda\right|^2 + \sum_{\lambda\in\Lambda\setminus\Omega}\left|\operatorname{prox}_{\gamma s_{\alpha,\lambda}(\kappa_\lambda(\cdot))}(x_\lambda) - x_\lambda\right|^2 \\ &\leq \sum_{\lambda\in\Omega}\left|\operatorname{prox}_{\gamma s_{\alpha,\lambda}(\kappa_\lambda(\cdot))}(x_\lambda) - x_\lambda\right|^2 + \sum_{\lambda\in\Lambda\setminus\Omega}|x_\lambda|^2 < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon\,.\end{aligned}$$

Hence $\lim_{\alpha\to 0}\operatorname{prox}_{\gamma\mathcal{R}_\alpha}(x) = x$ for all $x \in \ell^2(\Lambda)$. Since $\operatorname{prox}_{\gamma\mathcal{R}_\alpha} \to \mathrm{id}$, we have $1 = \lim_{\alpha\to 0}\operatorname{Lip}(\operatorname{prox}_{\mathcal{R}_\alpha}) \leq \lim_{\alpha\to 0}\ell_\alpha \leq 1$ which is (D2). Let $z \in \ell^2(\Lambda)$ and B be a bounded set. By (4.3) for α small enough we have

$$\begin{aligned}|\langle\operatorname{prox}_{\gamma\mathcal{R}_\alpha}(x) - x, z\rangle| &\leq \sum_{\lambda\in\Lambda}|\operatorname{prox}_{\gamma s_{\alpha,\lambda}(\kappa_\lambda(\cdot))}(x_\lambda) - x_\lambda|\,|z_\lambda| \\ &\leq \sum_{\lambda\in\Lambda}C(1-\ell_\alpha)|x_\lambda|\,|z_\lambda| \leq C(1-\ell_\alpha)\|x\|\|z\| \to 0\end{aligned}$$

since $\ell_\alpha \to 1$ as $\alpha \to 0$. This shows (D3). By (4.3) for all $x \in \ell^2(\Lambda)$ and α small enough we have

$$\begin{aligned}\frac{\left\|\operatorname{prox}_{\gamma\mathcal{R}_\alpha}(x) - x\right\|^2}{(1 - \operatorname{Lip}(\operatorname{prox}_{\gamma\mathcal{R}_\alpha}))^2} &= \frac{\sum_{\lambda\in\Lambda}|\operatorname{prox}_{\gamma s_{\alpha,\lambda}(\kappa_\lambda(\cdot))}(x_\lambda) - x_\lambda|^2}{(1-\ell_\alpha)^2} \\ &\leq \frac{\sum_{\lambda\in\Lambda}C^2(1-\ell_\alpha)^2|x_\lambda|^2}{(1-\ell_\alpha)^2} = C^2\|x\|^2\end{aligned}$$

which shows (D4). □

Proposition 4.10 (non-linear filtered DFD is a regularization, Case B)**.** *Let $(\varphi_\alpha)_{\alpha>0}$ be a non-linear regularizing filter such that Assumption B is satisfied. Suppose $\alpha, \alpha^k > 0$, and $z, z^k \in \ell^2(\Lambda)$ for $k \in \mathbb{N}$ with $(z^k)_{k\in\mathbb{N}} \to z$.*

- *Existence:* $\operatorname{dom}(\mathbf{M}_{\boldsymbol{\kappa}}^+ \circ \Phi_{\alpha,\boldsymbol{\kappa}}) = \ell^2(\Lambda)$.
- *Stability:* $\mathbf{M}_{\boldsymbol{\kappa}}^+ \circ \Phi_{\alpha,\boldsymbol{\kappa}}(z^k) \to \mathbf{M}_{\boldsymbol{\kappa}}^+ \circ \Phi_{\alpha,\boldsymbol{\kappa}}(z)$ *as* $k \to \infty$.
- *Convergence: Assume* $z = \mathbf{M}_{\boldsymbol{\kappa}} c^+$ *for* $c^+ \in \bigcup_{\alpha>0} \operatorname{dom}(\mathcal{R}_\alpha)$ *and suppose* $\|z^k - z\| \le \delta_k$. *Let* $\delta_k, \alpha_k \to 0$ *and* $(1-\ell_{\alpha_k})/\ell_{\alpha_k} \gtrsim \delta_k$. *Then* $\mathbf{M}_{\boldsymbol{\kappa}}^+ \circ \Phi_{\alpha_k,\boldsymbol{\kappa}}(z^k) \rightharpoonup c^+$ *as* $k \to \infty$.

Proof. Follows from Lemmas 4.9 and 4.7 and the injectivity of $\mathbf{M}_{\boldsymbol{\kappa}}$. □

4.3 Case C: Direct analysis

Assumption (B1) requires $\varphi_{\alpha,\lambda}$ to be a contraction for all $\alpha, \lambda > 0$. In this section we weaken this condition and exploit the diagonality of $\mathbf{M}_{\boldsymbol{\kappa}} \circ \Phi_{\alpha,\boldsymbol{\kappa}}$.

Assumption C. *Suppose $\gamma \in (0, 1/\max \kappa_\lambda^2)$. Let $(\varphi_\alpha)_{\alpha>0}$ be a non-linear regularizing filter which satisfies the following conditions:*

(C1) $\forall \lambda \in \Lambda \colon \forall x \in \mathbb{R} \colon (|\varphi_{\alpha,\lambda}(x)|)_{\alpha>0}$ *is monotonically increasing for* $\alpha \downarrow 0$.

(C2) $\exists d > 0\ \exists \ell_\alpha \in [0,1)\ \forall \alpha > 0\ \forall x \in \mathbb{R} \colon |x| \le d\alpha/\kappa_\lambda \Rightarrow |\varphi_{\alpha,\lambda}(x)| \le \frac{\gamma\kappa_\lambda^2 \ell_\alpha}{1-\ell_\alpha(1-\gamma\kappa_\lambda^2)}|x|$.

Next we show that Assumption A is weaker than Assumption C, hence the results in this section generalize the results of variational regularization.

Lemma 4.11. *If the non-linear regularizing filter $(\varphi_\alpha)_{\alpha>0}$ satisfies Assumption A with constants $b, c > 0$, then it satisfies Assumption C with $\ell_\alpha = 1/(1+\alpha b\gamma)$ and $d = bc$.*

Proof. Let $(\varphi_\alpha)_{\alpha>0}$ satisfy Assumption A with constants $b, c > 0$ and let $\alpha > 0$, $\lambda \in \Lambda$ and $y \in \mathbb{R}$ be arbitrary. By Lemma 4.4 and Lemma 2.1 we have $\varphi_{\alpha,\lambda}^{-1}(y) = y + \alpha\partial s_\lambda(y)$, where $s_\lambda \in \Gamma_0(\mathbb{R})$. Since $\partial s_\lambda(y)$ is a closed interval, the maximum value of $|\partial s_\lambda(y)|$ exists and $\max|\varphi_{\alpha,\lambda}^{-1}(y)| < \max|\varphi_{\tilde\alpha,\lambda}^{-1}(y)|$ for all $\alpha < \tilde\alpha$. Therefore, $|\varphi_{\alpha,\lambda}(y)| < |\varphi_{\tilde\alpha,\lambda}(y)|$ which shows (C1). Let $\lambda \in \Lambda$ be arbitrary. Choose $\tilde\alpha > 0$ such that $|\varphi_{\tilde\alpha,\lambda}(x)| \le |x|\kappa_\lambda^2/(\kappa_\lambda^2 + \tilde\alpha b)$ for all $x \in \mathbb{R}$ with $|x| \le \min \varphi_{\tilde\alpha,\lambda}^{-1}(c\kappa_\lambda) = c\kappa_\lambda + \tilde\alpha \min \partial s_\lambda(c\kappa_\lambda)$. Following the proof of Lemma 2.3, one shows that if this holds for one $\tilde\alpha > 0$, it holds for all $\alpha > 0$. Now set $\ell_\alpha := 1/(1+\alpha b\gamma)$ and $d := bc$. Then

$$|\varphi_{\alpha,\lambda}(x)| \le \frac{\kappa_\lambda^2}{\kappa_\lambda^2 + \tilde\alpha b}|x| = \frac{\kappa_\lambda^2 \gamma \ell_\alpha}{1-\ell_\alpha(1-\kappa_\lambda^2\gamma)}|x|$$

for all $x \in \mathbb{R}$ with $|x| \le c\kappa_\lambda + \alpha \min \partial s_\lambda(c\kappa_\lambda)$, whereas $\min \partial s_\lambda(c\kappa_\lambda) \ge d/\kappa_\lambda$. Hence the inequality holds for $|x| \le \alpha d/\kappa_\lambda$, showing (C2). □

Example 4.12. Fix the constants $b, d > 0$ and consider the function

$$\varphi_\alpha(\kappa, x) = \begin{cases} \frac{\kappa^2}{\kappa^2+\alpha b}\, x & |x| \le d\alpha/\kappa \\ x - \operatorname{sign}(x)\frac{db\alpha^2}{\kappa(\kappa^2+\alpha b)} & |x| > d\alpha/\kappa\,. \end{cases}$$

$(\varphi_\alpha)_{\alpha>0}$ is a non-linear regularizing filter and satisfies Assumption C. Note that φ_1 is the same function as in Example 4.3 but for $\alpha \neq 1$ the filter functions differ. By Remark 4.2 the construction of the non-linear regularizing filter satisfying Assumption A is uniquely given by φ_1. Thus, $(\varphi_\alpha)_{\alpha>0}$ does not satisfy Assumption A.

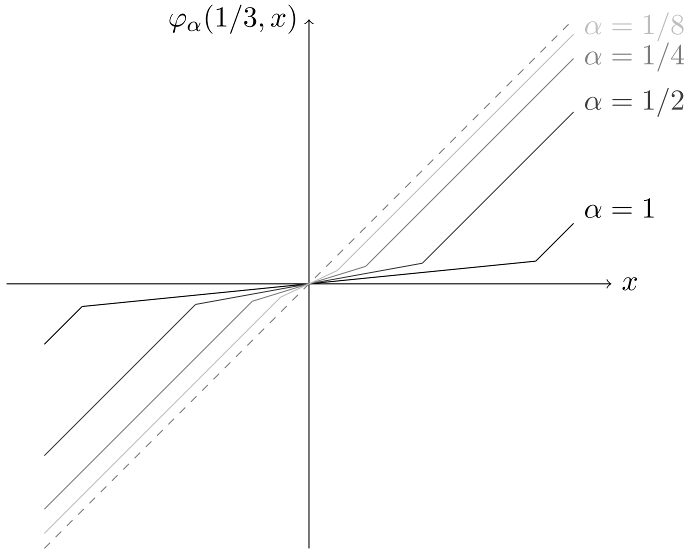

Figure 3: Illustration of the filter of Example 4.12 that satisfies Assumption C but not Assumption A.

Proposition 4.13 (Existence, Case C). *Let $(\varphi_\alpha)_{\alpha>0}$ be non-linear regularizing filter such that (C2) is satisfied. Then,* $\operatorname{dom}(\mathbf{M}_{\boldsymbol{\kappa}}^+ \circ \Phi_{\alpha,\boldsymbol{\kappa}}) = \ell^2(\Lambda)$.

Proof. Let $\alpha > 0$, $x = (x_\lambda)_{\lambda\in\Lambda} \in \ell^2(\Lambda)$ and $d > 0$, ℓ_α such that $\forall \alpha > 0$: $|\varphi_\alpha(z)| \leq |z|\gamma\kappa_\lambda^2\ell_\alpha/(1-\ell_\alpha(1-\gamma\kappa_\lambda^2))$ for all $z \in \mathbb{R}$ with $|z| \leq d\alpha/\kappa_\lambda$. Since $\sup_{\lambda\in\Lambda}\kappa_\lambda < \infty$ there are only finitely many λ such that $|x_\lambda| > d\alpha/\kappa_\lambda$. Thus,

$$\left\|\mathbf{M}_{\boldsymbol{\kappa}}^+ \circ \Phi_{\alpha,\boldsymbol{\kappa}}(x)\right\|^2 = \sum_{\lambda\in\Lambda} \frac{|\varphi_{\alpha,\lambda}(x_\lambda)|^2}{\kappa_\lambda^2} = \sum_{|x_\lambda|\leq d\alpha/\kappa_\lambda} \frac{|\varphi_{\alpha,\lambda}(x_\lambda)|^2}{\kappa_\lambda^2} + \sum_{|x_\lambda|> d\alpha/\kappa_\lambda} \frac{|\varphi_{\alpha,\lambda}(x_\lambda)|^2}{\kappa_\lambda^2}$$
$$\leq \sum_{|x_\lambda|\leq d\alpha/\kappa_\lambda} \frac{\gamma^2\kappa_\lambda^2\ell_\alpha^2}{(1-\ell_\alpha(1-\gamma\kappa_\lambda^2))^2}|x_\lambda|^2 + C < \infty\,,$$

because $\gamma^2\kappa_\lambda^2\ell_\alpha^2/(1-\ell_\alpha(1-\gamma\kappa_\lambda^2))^2 \leq \gamma(\ell_\alpha/(1-\ell_\alpha))^2$. □

Under Assumption C, we can show that the mappings $\mathbf{M}_{\boldsymbol{\kappa}} \circ \Phi_{\alpha,\boldsymbol{\kappa}}$ are strong-weak continuous and converges to $\mathbf{M}_{\boldsymbol{\kappa}}$ as $\alpha \to 0$.

Proposition 4.14 (Stability, Case C). *Let $(\varphi_\alpha)_{\alpha>0}$ be non-linear regularizing filter such that Assumption C is satisfied. Let $\alpha > 0$ be fixed and $z, z^k \in \ell^2(\Lambda)$ such that $\|z-z^k\|_2 \to 0$. Then $\mathbf{M}_{\boldsymbol{\kappa}}^+ \circ \Phi_{\alpha,\boldsymbol{\kappa}}(z^k) \rightharpoonup \mathbf{M}_{\boldsymbol{\kappa}}^+ \circ \Phi_{\alpha,\boldsymbol{\kappa}}(z)$ as $k \to \infty$.*

Proof. Fix the constant d in Assumption C and set $\Lambda_k \coloneqq \{\lambda \in \Lambda \mid |z_\lambda^k| > d\alpha/\kappa_\lambda\}$ for $n \in \mathbb{N}$. Then, for all $\lambda \in \Lambda_k$,

$$\kappa_\lambda \geq \frac{d\alpha}{|z_\lambda^k|} \geq \frac{d\alpha}{|z_\lambda^k| + \delta_k} \geq \frac{d\alpha}{\|z\|_\infty + \max_k \delta_k} =: a > 0\,.$$

Since a is independent of n, it follows $\inf_{\lambda\in\Lambda_k}\kappa_\lambda \geq a$ for all $n \in \mathbb{N}$. Therefore

$$\|x^k\|^2 = \sum_{\kappa_\lambda\geq a}\left|\frac{1}{\kappa_\lambda}\varphi_{\alpha,\lambda}(z_\lambda^k)\right|^2 + \sum_{\kappa_\lambda< a}\left|\frac{1}{\kappa_\lambda}\varphi_{\alpha,\lambda}(z_\lambda^k)\right|^2$$

$$\leq \frac{L^2}{a^2} \sum_{\kappa_\lambda \geq a} |z_\lambda^k|^2 + \sum_{\kappa_\lambda < a} \Big| \frac{\gamma \kappa_\lambda \ell_\alpha}{1 - \ell_\alpha (1 - \gamma \kappa_\lambda^2)} \Big|^2 |z_\lambda^k|^2$$
$$\leq \Big(\frac{L^2}{a^2} + \Big| \frac{\ell_\alpha \gamma a}{1 - \ell_\alpha} \Big|^2 \Big) \cdot (\|z\| + \max_k \delta_k)^2 .$$

Hence $(x^k)_k$ is bounded and has a weakly convergent subsequence $(x^{n(\ell)})_\ell$. By Lemma 3.4, there exist $s_{\alpha,\lambda} \in \Gamma_0(\mathbb{R})$ with $s_{\alpha,\lambda}(0) = 0$, $\varphi_{\alpha,\lambda} = \operatorname{prox}_{s_{\alpha,\lambda}}$, $\mathcal{R}_\alpha := \bigoplus_{\lambda \in \Lambda} s_{\alpha,\lambda}(\kappa_\lambda(\cdot)) \in \Gamma_0(\ell^2(\Lambda))$ positive and $(\operatorname{prox}_{\gamma s_{\alpha,\lambda}(\kappa_\lambda(\cdot))}(x_\lambda))_{\lambda \in \Lambda} = \operatorname{prox}_{\gamma \mathcal{R}_\alpha}((x_\lambda)_{\lambda \in \Lambda})$. Thus,

$$\begin{aligned} x^k := \mathbf{M}_{\boldsymbol{\kappa}}^+ \circ \Phi_{\alpha,\boldsymbol{\kappa}}(z^k) &= \operatorname{Fix}\big(\operatorname{prox}_{\gamma \mathcal{R}_\alpha} \circ (\operatorname{id} - \gamma \nabla \|\mathbf{M}_{\boldsymbol{\kappa}}(\cdot) - z^k\|^2/2)\big) \\ &= \big(\operatorname{Fix}\big(\operatorname{prox}_{\gamma s_{\alpha,\lambda}(\kappa_\lambda(\cdot))}((1 - \gamma \kappa_\lambda^2)(\cdot) + \gamma \kappa_\lambda z_\lambda^k)\big)\big)_\lambda . \end{aligned}$$

Assume $x^{n(\ell)} \rightharpoonup x^+$ and $y \in \ell^2(\Lambda)$. With Lemma 2.1(f),

$$\begin{aligned} &|\langle \operatorname{prox}_{\gamma \mathcal{R}_\alpha} \circ \big(\operatorname{id} - \gamma \mathbf{M}_{\boldsymbol{\kappa}}^2 + \gamma \mathbf{M}_{\boldsymbol{\kappa}} z\big)(x^+) - x^+, y \rangle| \\ &\leq |\langle \operatorname{prox}_{\gamma \mathcal{R}_\alpha} \circ (\operatorname{id} - \gamma \mathbf{M}_{\boldsymbol{\kappa}}^2 + \gamma \mathbf{M}_{\boldsymbol{\kappa}} z)(x^+) - \operatorname{prox}_{\gamma \mathcal{R}_\alpha} \circ (\operatorname{id} - \gamma \mathbf{M}_{\boldsymbol{\kappa}}^2 + \gamma \mathbf{M}_{\boldsymbol{\kappa}} z^{n(\ell)})(x^+), y \rangle| \\ &+ |\langle \operatorname{prox}_{\gamma \mathcal{R}_\alpha} \circ (\operatorname{id} - \gamma \mathbf{M}_{\boldsymbol{\kappa}}^2 + \gamma \mathbf{M}_{\boldsymbol{\kappa}} z^{n(\ell)})(x^+) - \operatorname{prox}_{\gamma \mathcal{R}_\alpha} \circ (\operatorname{id} - \gamma \mathbf{M}_{\boldsymbol{\kappa}}^2 + \gamma \mathbf{M}_{\boldsymbol{\kappa}} z^{n(\ell)})(x^{n(\ell)}), y \rangle| \\ &+ |\langle x^{n(\ell)} - x^+, y \rangle| \to 0. \end{aligned}$$

Thus, $x^+ = \mathbf{M}_{\boldsymbol{\kappa}}^+ \circ \Phi_{\alpha,\boldsymbol{\kappa}}(z)$. □

Proposition 4.15 (Convergence, Case C)**.** *Let $(\varphi_\alpha)_{\alpha > 0}$ be non-linear regularizing filter such that Assumption C is satisfied. Let $z \in \operatorname{ran}(\mathbf{M}_{\boldsymbol{\kappa}})$ such that $\big(\int_0^{z_\lambda} \max\{\varphi_{\tilde{\alpha},\lambda}^{-1}(y)\} dy\big)_\lambda \in \ell^1$ for some $\tilde{\alpha} > 0$. Let $z^k \in \ell^2(\Lambda)$ such that $\|z - z^k\|_2 \leq \delta_k$ and let $\delta_k, \alpha_k \to 0$ such that $\alpha_k \gtrsim \delta_k$ and $1 - \ell_{\alpha_k} \gtrsim \delta_k$. Then $\mathbf{M}_{\boldsymbol{\kappa}}^+ \circ \Phi_{\alpha_k,\boldsymbol{\kappa}}(z^k) \rightharpoonup \mathbf{M}_{\boldsymbol{\kappa}}^+ z$.*

Proof. Fix constant d of Assumption C and define $\Lambda_k := \{\lambda \in \Lambda \mid |z_\lambda^k| \geq d\alpha_k/(\kappa_\lambda)\}$ for $n \in \mathbb{N}$. Then, $\|z\|_\infty + \max_k \delta_k \geq |z_\lambda| + \delta_k \geq |z_\lambda^k| \geq d\alpha_k/\kappa_\lambda \geq C\delta_k/\kappa_\lambda$ for all $\lambda \in \Lambda_k$ and some constant $C > 0$. Thus $(\delta_k / \inf_{\lambda \in \Lambda_k} \kappa_\lambda)_{k \in \mathbb{N}}$ is bounded, as otherwise one could create a contradiction to $z \in \ell^2(\Lambda)$. Therefore, for some constant $D > 0$,

$$\begin{aligned} \|x^k\|^2 &= \sum_{\lambda \in \Lambda_k} \Big| \frac{1}{\kappa_\lambda} \varphi_{\alpha,\lambda}(z_\lambda^k) \Big|^2 + \sum_{\lambda \notin \Lambda_k} \Big| \frac{1}{\kappa_\lambda} \varphi_{\alpha,\lambda}(z_\lambda^k) \Big|^2 \\ &\leq \sum_{\lambda \in \Lambda_k} L^2 \Big(\frac{|z_\lambda^k|}{\kappa_\lambda} \Big)^2 + \sum_{\lambda \notin \Lambda_k} \Big(\frac{\ell_{\alpha_k} \gamma \kappa_\lambda}{1 - \ell_{\alpha_k}(1 - \gamma \kappa_\lambda^2)} |z_\lambda^k| \Big)^2 \\ &\leq L^2 \sum_{\lambda \in \Lambda_k} \Big(\frac{|z_\lambda|}{\kappa_\lambda} + \frac{|z_\lambda^k - z_\lambda|}{\kappa_\lambda} \Big)^2 + \sum_{\lambda \notin \Lambda_k} \Big(\frac{|z_\lambda|}{\kappa_\lambda} + \frac{\ell_{\alpha_k}}{1 - \ell_{\alpha_k}} \gamma \sup_{\lambda \in \Lambda} \kappa_\lambda |z_\lambda^k - z_\lambda| \Big)^2 \\ &\leq L^2 \Big(\big\| \mathbf{M}_{\boldsymbol{\kappa}}^+ z \big\| + \frac{\delta_k}{\inf_{\lambda \in \Lambda_k} \kappa_\lambda} \Big)^2 + \Big(\big\| \mathbf{M}_{\boldsymbol{\kappa}}^+ z \big\| + \gamma \sup_{\lambda \in \Lambda} \kappa_\lambda \frac{\delta_k}{1 - \ell_{\alpha_k}} \Big)^2 \\ &\leq D . \end{aligned}$$

Thus $(x^k)_k$ is bounded and has a weakly convergent subsequence.

Now $\lambda \in \Lambda$. By Lemma 2.4, we have $\varphi_{\alpha,\lambda}^{-1}(y) = y + \partial s_{\alpha,\lambda}(y) \to \{y\}$ and therefore $\partial s_{\alpha,\lambda}(y) \to \{0\}$ as $\alpha \to 0$ for all $y \in \mathbb{R}$. Let $y \in \mathbb{R}$ be arbitrary and for every $\alpha > 0$ choose $z_\alpha \in \partial s_{\alpha,\lambda}(y)$. The definition of the sub-differential implies that $s_{\alpha,\lambda}(y) \leq z_\alpha y$. Since $s_{\alpha,\lambda}$

is positive and $z_\alpha \to 0$ as $\alpha \to 0$, we have that $s_{\alpha,\lambda}(y) \to 0$. Consequently, for a fixed $\lambda \in \Lambda$, $s_{\alpha,\lambda}$ converge point-wise to the zero function. By Lemma 3.5, we have

$$\mathcal{R}_\alpha(\mathbf{M}_{\boldsymbol{\kappa}}^+ z) = \sum_{\lambda\in\Lambda} s_{\alpha,\lambda}(z_\lambda) = \sum_{\lambda\in\Lambda} \int_0^{z_\lambda} \max\{\varphi_{\alpha,\lambda}^{-1}(y)\}\, dy - z_\lambda^2/2\,.$$

Then the theorem of monotone convergence implies $\mathcal{R}_\alpha(\mathbf{M}_{\boldsymbol{\kappa}}^+ z) \to 0$ as $\alpha \to 0$. In particular, $\mathbf{M}_{\boldsymbol{\kappa}}^+ z \in \operatorname{dom}(\mathcal{R}_\alpha)$ for all $\alpha \in (0, \tilde{\alpha})$.

Let $(x^{n(\ell)})_\ell$ be a weakly convergent subsequence with weak limit x^+. By Lemma 3.4, we have $x^{n(\ell)} \in \operatorname{argmin}_x \left\|\mathbf{M}_{\boldsymbol{\kappa}} x - z^{n(\ell)}\right\|^2/2 + \mathcal{R}_{\alpha_{n(\ell)}}(x)$ and, as $\ell \to \infty$,

$$\begin{aligned}\frac{1}{2}\left\|\mathbf{M}_{\boldsymbol{\kappa}} x^{n(\ell)} - z^{n(\ell)}\right\|^2 + \mathcal{R}_{\alpha_{n(\ell)}}(x^{n(\ell)}) &\leq \frac{1}{2}\left\|z - z^{n(\ell)}\right\|^2 + \mathcal{R}_{\alpha_{n(\ell)}}(\mathbf{M}_{\boldsymbol{\kappa}}^+ z) \\ &\leq \delta_{n(\ell)}^2/2 + \mathcal{R}_{\alpha_{n(\ell)}}(\mathbf{M}_{\boldsymbol{\kappa}}^+ z) \to 0\,.\end{aligned}$$

Thus $\|\mathbf{M}_{\boldsymbol{\kappa}} x^{n(\ell)} - z^{n(\ell)}\| \to 0$ and since $z^{n(\ell)} \to z$ and $\mathbf{M}_{\boldsymbol{\kappa}}$ is linear and bounded, we conclude $\mathbf{M}_{\boldsymbol{\kappa}} x^{n(\ell)} \rightharpoonup \mathbf{M}_{\boldsymbol{\kappa}} x^+$. Therefore, $z = \mathbf{M}_{\boldsymbol{\kappa}} x^+$ and $x^+ = \mathbf{M}_{\boldsymbol{\kappa}}^+ z$. Because this holds for every weakly convergent subsequence, we conclude $x^k \rightharpoonup \mathbf{M}_{\boldsymbol{\kappa}}^+ z$. □

Decomposing an operator with a diagonal frame decomposition can significantly simplify the inverse problem. Another application of this technique is in the analysis of a variant of the Plug-and-Play (PnP) regularization method [10]. In this approach, after diagonalizing the operator $\mathbf{A}$, component-wise denoising is performed using functions $d_{\alpha,\lambda}\colon \mathbb{R} \to \mathbb{R}$ that satisfy the following conditions:

(a) $\forall \alpha, \lambda\colon d_{\alpha,\lambda}$ is monotonically increasing, nonexpansive, and $d_{\alpha,\lambda}(0) = 0$.

(b) $\forall x, \lambda\colon (|d_{\alpha,\lambda}(x)|)_{\alpha>0}$ is increasing as $\alpha \to 0$.

(c) $\forall \alpha, \lambda \exists \ell_\alpha \in [0, 1)\colon\ |d_{\alpha,\lambda}(x)| \leq \ell_\alpha |x|$ for $|x| \leq \gamma\alpha$ and $\alpha \lesssim 1 - \ell_\alpha$.

Then $d_{\alpha,\lambda} = \operatorname{prox}_{s_{\alpha,\lambda}}$ with $s_{\alpha,\lambda} \in \Gamma_0(\mathbb{R})$ and $\mathbf{D}_\alpha = \operatorname{prox}_{\mathcal{R}_\alpha}$ with $\mathcal{R}_\alpha = \bigoplus_{\lambda\in\Lambda} s_{\alpha,\lambda} \in \Gamma_0(\ell^2(\Lambda))$. Now, we can prove convergence analogously to the previous argument. Let $y \in \operatorname{ran}(\mathbf{A})$, $(y^k)_k \in \mathbb{Y}^{\mathbb{N}}$ with $\|y^k - y\| \leq \delta_k$. If $\delta_k, \alpha_k \to 0$ and $\alpha_k \gtrsim \delta_k^{1/2}$ and $c^k = \operatorname{Fix}(\mathbf{D}_{\alpha_k} \circ (\operatorname{id} - \gamma\nabla\|\mathbf{M}_{\boldsymbol{\kappa}}(\cdot) - \mathbf{T}_{\mathbf{v}}^*(y^k)\|^2/2)$, then $\mathbf{T}_{\bar{\mathbf{u}}} c^k$ converges weakly to the minimal norm solution of $y = \mathbf{A}x$.

4.4 Proof of main theorems

Using the results above for $\mathbf{M}_{\boldsymbol{\kappa}}^+ \circ \Phi_{\alpha_k,\boldsymbol{\kappa}}$ under the three different assumptions we can derive the main Theorems 1.1 and 1.2 for the non-linear filter-based reconstruction method $(\mathbf{B}_\alpha)_{\alpha>0}$ stated in the introduction.

The stability Theorem 1.1, in fact, follows from the representation $\mathbf{B}_\alpha = \mathbf{T}_{\bar{\mathbf{u}}} \circ \mathbf{M}_{\boldsymbol{\kappa}}^+ \circ \Phi_{\alpha,\boldsymbol{\kappa}} \circ \mathbf{T}_{\mathbf{v}}^*$ and the stability results in Propositions 4.5, 4.10 and 4.15 combined with the continuity of $\mathbf{T}_{\bar{\mathbf{u}}}$ and $\mathbf{T}_{\mathbf{v}^*}$.

To show Theorem 1.2 write $x^+ = \mathbf{A}^+ y$ and $c^k = \mathbf{M}_{\boldsymbol{\kappa}}^+ \circ \Phi_{\alpha_k,\boldsymbol{\kappa}} \circ \mathbf{T}_{\mathbf{v}}^*(y^k)$. By the definition of the DFD we have

$$\|\mathbf{T}_{\mathbf{v}}^* y_k - \mathbf{M}_{\boldsymbol{\kappa}} \mathbf{T}_{\mathbf{u}}^* x^+\| = \left\|\mathbf{T}_{\mathbf{v}}^* y_k - \mathbf{T}_{\mathbf{v}}^* \mathbf{A} x^+\right\| \leq \|\mathbf{T}_{\mathbf{v}}\|\delta_k\,.$$

Application of the convergence results of Propositions 4.5, 4.10 and 4.15 shows $c^k \rightharpoonup \mathbf{M}_{\boldsymbol{\kappa}}^+ \mathbf{M}_{\boldsymbol{\kappa}} \mathbf{T}_{\mathbf{u}}^* x^+ = \mathbf{T}_{\mathbf{u}}^* x^+$. Thus for any $z \in \mathbb{X}$, we have

$$\langle z, \mathbf{T}_{\bar{\mathbf{u}}} c^k - x^+ \rangle = \langle s, \mathbf{T}_{\bar{\mathbf{u}}} c^k - \mathbf{T}_{\bar{\mathbf{u}}} \mathbf{T}_{\mathbf{u}}^* x^+ \rangle = \langle \mathbf{T}_{\bar{\mathbf{u}}}^* z, c^k - \mathbf{T}_{\mathbf{u}}^* x^+ \rangle \to 0\,.$$

Thus $x^k = \mathbf{T}_{\bar{\mathbf{u}}} c^k \rightharpoonup x^+$, which concludes the proof of Theorem 1.2.

5 Conclusion

In this paper, we have analyzed non-linear diagonal frame filtering for the regularization of inverse problems. Compared to the previously analyzed linear filters, non-linear filters can better exploit the specific structure of the target signal and the noise. We have presented three different approaches to prove convergence in the context of regularization theory. Future research might focus on convergence in terms of norm topology and the derivation of convergence rates.

References

[1] Anestis Antoniadis and Jianqing Fan. Regularization of wavelet approximations. *Journal of the American Statistical Association*, 96(455):939–955, 2001.

[2] Clemens Arndt, Alexander Denker, Sören Dittmer, Nick Heilenkötter, Meira Iske, Tobias Kluth, Peter Maass, and Judith Nickel. Invertible residual networks in the context of regularization theory for linear inverse problems. *arXiv:2306.01335*, 2023.

[3] Heinz Bauschke and Patrick Combettes. *Convex analysis and monotone operator theory in Hilbert spaces*. CMS Books in Mathematics. Springer, Cham, second edition, 2017.

[4] Martin Benning and Martin Burger. Modern regularization methods for inverse problems. *Acta Numerica*, 27(1):1–111, 2018.

[5] Kristian Bredies and Dirk Lorenz. Regularization with non-convex separable constraints. *Inverse Problems*, 25(8):085011, 2009.

[6] Emmanuel Candès and David Donoho. Recovering edges in ill-posed inverse problems: Optimality of curvelet frames. *Annals of Statistics*, 30(3):784–842, 2002.

[7] Stanley Chan, Xiran Wang, and Omar Elgendy. Plug-and-Play ADMM for Image Restoration: Fixed-Point Convergence and Applications. *IEEE Transactions on Computational Imaging*, 3(1):84–98, 2016.

[8] David Donoho. Nonlinear solution of linear inverse problems by wavelet–vaguelette decomposition. *Applied and Computational Harmonic Analysis*, 2(2):101–126, 1995.

[9] Andrea Ebner, Jürgen Frikel, Dirk Lorenz, Johannes Schwab, and Markus Haltmeier. Regularization of inverse problems by filtered diagonal frame decomposition. *Applied and Computational Harmonic Analysis*, 62:66–83, 2023.

[10] Andrea Ebner and Markus Haltmeier. Plug-and-play image reconstruction is a convergent regularization method. *arXiv:2212.06881*, 2022.

[11] Ivar Ekeland and Roger Temam. *Convex analysis and variational problems.* SIAM, 1999.

[12] Heinz Engl, Martin Hanke, and Andreas Neubauer. *Regularization of Inverse Problems.* Springer Netherlands, 1996.

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

[14] Jürgen Frikel and Markus Haltmeier. Sparse regularization of inverse problems by operator-adapted frame thresholding. In *Mathematics of Wave Phenomena*, pages 163–178, Cham, 2020. Springer International Publishing.

[15] Ye Gao. Wavelet shrinkage denoising using the non-negative garrote. *Journal of Computational and Graphical Statistics*, 7(4):469–488, 1998.

[16] Simon Göppel, Jürgen Frikel, and Markus Haltmeier. Translation invariant diagonal frame decomposition of inverse problems and their regularization. *Inverse Problems*, 39(6):065011, 2023.

[17] Markus Grasmair. Linear convergence rates for Tikhonov regularization with positively homogeneous functionals. *Inverse Problems*, 27(7):075014, 2011.

[18] Charles Groetsch. The theory of Tikhonov regularization for Fredholm equations. *104p, Boston Pitman Publication*, 1984.

[19] Charles Groetsch and Curtis Vogel. Asymptotic theory of filtering for linear operator equations with discrete noisy data. *Mathematics of Computation*, 49(180):499–506, 1987.

[20] Simon Hubmer and Ronny Ramlau. Frame decompositions of bounded linear operators in Hilbert spaces with applications in tomography. *Inverse Problems*, 37(5):055001, 2021.

[21] Simon Hubmer, Ronny Ramlau, and Lukas Weissinger. On regularization via frame decompositions with applications in tomography. *Inverse Problems*, 38(5):055003, 2022.

[22] Stéphane Mallat. *A wavelet tour of signal processing.* Elsevier, 1999.

[23] Gisela Mazzieri, Ruben Spies, and Karina Temperini. Existence, uniqueness and stability of minimizers of generalized tikhonov–phillips functionals. *Journal of Mathematical Analysis and Applications*, 396(1):396–411, 2012.

[24] Michael Quellmalz, Lukas Weissinger, Simon Hubmer, and Paul Erchinger. A frame decomposition of the Funk-Radon transform. In *International Conference on Scale Space and Variational Methods in Computer Vision*, pages 42–54. Springer, 2023.

[25] Andreas Rieder. *Keine Probleme mit inversen Problemen: eine Einführung in ihre stabile Lösung.* Springer-Verlag, 2013.

[26] Otmar Scherzer, Markus Grasmair, Harald Grossauer, Markus Haltmeier, and Frank Lenzen. *Variational methods in imaging*, volume 167. Springer, 2009.

[27] Yu Sun, Brendt Wohlberg, and Ulugbek S. Kamilov. An Online Plug-and-Play Algorithm for Regularized Image Reconstruction. *IEEE Transactions on Computational Imaging*, 5(3):395–408, 2019.

[28] Kaixuan Wei, Angelica Aviles-Rivero, Jingwei Liang, Ying Fu, Carola-Bibiane Schönlieb, and Hua Huang. Tuning-free plug-and-play proximal algorithm for inverse imaging problems. In *International Conference on Machine Learning*, pages 10158–10169. PMLR, 2020.

[29] Kai Zhang, Yawei Li, Wangmeng Zuo, Lei Zhang, Luc Van Gool, and Radu Timofte. Plug-and-play image restoration with deep denoiser prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(10):6360–6376, 2021.