

An Iterative Algorithm for Nonlinear Inverse Problems with Joint Sparsity Constraints in Vector Valued Regimes and an Application to Color Image Inpainting

Gerd Teschke* and Ronny Ramlau†

March 20, 2007

Abstract

This paper is concerned with nonlinear inverse problems where data and solution are vector valued and, moreover, where the solution is assumed to have a sparse expansion with respect to a preassigned frame. We especially focus on such problems where the different components of the solution exhibit a common or so-called *joint* sparsity pattern. Joint sparsity means here that the measure (typically constructed as weighted ℓ_1 norms of componentwise ℓ_q norms of the frame coefficients) promotes a coupling of non-vanishing components. Quite recently, an iterative strategy for linear inverse problems with such joint sparsity constraints was presented. Here we develop an iterative concept for nonlinear inverse problems with joint sparsity constraints for which we show convergence and regularization properties. Moreover, we demonstrate the capabilities of the proposed algorithm in the field of color image inpainting.

MSC: 15A29, 47A52, 68U10, 94A08, 94A40

Keywords: nonlinear ill-posed inverse problems, vector valued data and solution, (joint) sparsity measures, frames, iterative approximation, color image inpainting

1 Introduction

We address the problem of computing an approximation to a solution of a nonlinear operator equation $T(x) = y$, where $T : X \rightarrow Y$ is a nonlinear operator between Hilbert spaces X and Y . Since in many relevant cases only noisy data y^δ with $\|y^\delta - y\| \leq \delta$ are available, we are often faced with the problem of ill-posedness (in the sense of a discontinuous dependency of the solution on the data) and therefore with regularization issues. In the last decade, many of the well known regularization methods for linear inverse and ill-posed problems have been generalized to nonlinear problems incorporating at most quadratic regularization terms. Quite recently, during the time when highly

*G.T. is with the Zuse Institute Berlin (ZIB), Takustr. 7, 14195 Berlin, Germany.

†R.R. is with the Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Sciences, Altenbergerstr. 69, 4040 Linz, Austria.

redundant frame expansions and sparsity issues became important, attempts have been made to involve frame theory and sparsity concepts into linear as well as nonlinear inverse problems. The first breakthrough was made by Daubechies et.al. [4]. In this paper, the authors consider linear inverse problems with ‘classical’ sparsity constraints (no mixed or joint measures), and they have proposed a Landweber iteration with a shrinkage operation applied in each iteration step. Encouraged by [4], several efforts have been made in order to provide a theory for more generalized linear settings such as vector valued problems and the involvement of (highly) redundant dictionaries of frames, see e.g. [1, 17], mixed smoothness and sparsity and BV constraints, see e.g. [5, 6, 7, 8], and also for nonlinear problems with sparsity constraints, see [2, 14, 15, 16].

However, when dealing with vector valued problems, most of the recent literature is addressed to the recovery of sparse scalar functions only. But multi-channel signals may not only possess sparse frame expansions for each individual channel, but additionally the different channels may also exhibit joint sparsity pattern. A first result on this new aspect on joint sparsity in combination with solving linear inverse problems was achieved in [11]. In the present paper, we address the joint sparsity issue when dealing with nonlinear inverse problems, i.e. we aim to find a minimizer of a functional of the form

$$J_\alpha(x) = \|y^\delta - T(x)\|_Y^2 + 2\alpha\Psi(x) , \quad (1.1)$$

where Ψ is a suitable sparsity measure (to be specified below) promoting joint sparsity between all the channels of x , i.e. promoting sort of coupling of non-vanishing components of x .

As the main result of this paper we provide an iteration scheme to approach a minimizer of variational functional (1.1) when Ψ is a suitable measure for joint sparsity. Moreover, we show norm convergence of the proposed scheme and regularization properties for a-priori parameter rules. The potential of the proposed iteration is demonstrated in the context of color image inpainting.

The remaining paper is organized as follows: In Section 2, we explain in detail the setup of the problem and describe the iteration procedure. In Section 3, we explicitly compute the necessary conditions for a minimum of (1.1). The result on norm convergence (convergence towards critical points) is shown in Section 4, whereas in Section 5 we prove the regularization properties of the proposed scheme. We finish this paper with Section 6 in which numerical results on color image inpainting are shown.

2 Setup of the Problem

2.1 Representation by Frames

Assume we are given a preassigned system of functions $\{\phi_\lambda\}_{\lambda \in \Lambda} \subset X$ for which we may consider the operator

$$F : X \rightarrow \ell_2 \text{ via } x \mapsto \{\langle x, \phi_\lambda \rangle\}_{\lambda \in \Lambda}$$

with adjoint

$$F^* : \ell_2 \rightarrow X \text{ via } \mathbf{g} \mapsto \sum_{\lambda \in \Lambda} g_\lambda \phi_\lambda.$$

The system $\{\phi_\lambda\}_{\lambda \in \Lambda}$ is called a frame if there exists constants A, B with $0 < A \leq B < \infty$ such that for all $x \in X$,

$$A\|x\|_X^2 \leq \sum_{\lambda \in \Lambda} |\langle x, \phi_\lambda \rangle|^2 \leq B\|x\|_X^2. \quad (2.1)$$

The latter condition implies invertibility of F^*F and ensures thus that each $x \in X$ can be reconstructed from its moments $\langle x, \phi_\lambda \rangle$ or $\langle x, (F^*F)^{-1}\phi_\lambda \rangle$, i.e. we have

$$\sum_{\lambda \in \Lambda} \langle x, (F^*F)^{-1}\phi_\lambda \rangle \phi_\lambda = F^*F(F^*F)^{-1}x = x = (F^*F)^{-1}F^*Fx = \sum_{\lambda \in \Lambda} \langle x, \phi_\lambda \rangle (F^*F)^{-1}\phi_\lambda.$$

Frames are typically ‘overcomplete’, i.e. for a given $x \in X$, one may find many different sequences $\mathbf{g} \in \ell_2$ such that

$$x = \sum_{\lambda \in \Lambda} g_\lambda \phi_\lambda.$$

A few of them have special properties for which they are preferred, e.g. a sequence with minimal ℓ_2 norm. As mentioned in Section 1, we are interested in sequences that are sparser than minimum ℓ_2 norm sequences. For instance, one may know a priori that x is a noisy version of a linear combination of ϕ_λ with a coefficient sequence with small ℓ_p norm (e.g. for some $p < 2$). To this end, we have to choose for Ψ in (1.1) an ℓ_q sparsity constraint that promotes the desired sparsity on \mathbf{g} .

2.2 Inverse Problem and Joint Sparsity

In order to cast the inverse problem by means of some variational functional close to (1.1), we firstly have to setup the vector valued structure. To this end, let \mathcal{X}_i , $i = 1, \dots, n$, and \mathcal{Y}_j , $j = 1, \dots, m$, be Hilbert spaces. For sake of short notation, we introduce Hilbert spaces

$$X = \bigotimes_{i=1}^n \mathcal{X}_i \quad \text{and} \quad Y = \bigotimes_{j=1}^m \mathcal{Y}_j$$

equipped with the usual inner product, i.e. for $x, h \in X$, $\langle x, h \rangle = \sum_{i=1}^n \langle x_i, h_i \rangle$, with $x_i, h_i \in \mathcal{X}_i$, and for $y, v \in Y$, $\langle y, v \rangle = \sum_{j=1}^m \langle y_j, v_j \rangle$, with $y_j, v_j \in \mathcal{Y}_j$. Assume now the data $y \in Y$ are related to $x \in X$ via the nonlinear relationship \tilde{T} ,

$$\tilde{T}(x) = y, \quad \text{i.e.} \quad (y_1, \dots, y_m) = \tilde{T}(x_1, \dots, x_n).$$

Assume furthermore, for each individual \mathcal{X}_i we have a frame $\{\phi_\lambda^i\}_{\lambda \in \Lambda_i}$, then each x_i can be expressed by

$$x_i = \sum_{\lambda \in \Lambda_i} \mathbf{g}^i \phi_\lambda^i.$$

To simplify the notation, we suppose the frame $\{\phi_\lambda^i\}_{\lambda \in \Lambda_i}$ to be the same for each \mathcal{X}_i (note that all the results stated below hold true also without this restriction). Thus, we may naturally define for $\mathbf{g} \in (\ell_2)^n$,

$$T(\mathbf{g}) = T(\mathbf{g}^1, \dots, \mathbf{g}^n) := \tilde{T}(F^*\mathbf{g}^1, \dots, F^*\mathbf{g}^n) = \tilde{T}(x_1, \dots, x_n). \quad (2.2)$$

One might consider the following situations that fit into this setting:

- Nonlinear-Linear mixing: $T(\mathbf{g}) = T\left(\left\{\sum_{l=1}^r A_{l,i} F^* \mathbf{g}^l\right\}_{i=1,\dots,n}\right)$, for linear and bounded mappings $A_{l,i} : \mathcal{H}_l \rightarrow \mathcal{X}_i$.

Simple cases:

- Nonlinear scalar valued: $T(\mathbf{g}) = T(\mathbf{g}^1, \dots, \mathbf{g}^n) = T\left(\sum_{i=1}^n F^* \mathbf{g}^i\right)$.
- Purely linear: T some linear and bounded operator.

For our purpose of treating vector valued data and solutions, we cast the variational problem of the inverse problem now as follows: find a vector of sequences $\mathbf{g} = (\mathbf{g}^1, \dots, \mathbf{g}^n) \in (\ell_2)^n$, where $\mathbf{g}^i = \{(\mathbf{g}^i)_\lambda\}_{\lambda \in \Lambda}$, such that

$$J_\alpha(\mathbf{g}) = \|y^\delta - T(\mathbf{g})\|_Y^2 + 2\alpha\Psi(\mathbf{g}) \quad (2.3)$$

is minimized, where $\Psi(\mathbf{g})$ is a constraint promoting *joint sparsity* of the vector $\mathbf{g} \in (\ell_2)^n$. In order to utilize recent results of [14], it would be preferable to assume Ψ to be some positive, weakly lower semi-continuous and convex penalty. Then, several important lemmas and theorems of [14] can then be conveniently applied in this vector valued framework and only those statements must be proven where the particular structure of Ψ plays a role. Following the arguments in [11] on joint sparsity, a reasonable measure that forces a coupling of non-vanishing components is of the form

$$\Psi(\mathbf{g}) = \sum_{\lambda \in \Lambda} \omega_\lambda \|\mathbf{g}_\lambda\|_q^p, \quad (2.4)$$

with $p \in \{1, q\}$ and $\mathbf{g}_\lambda = ((\mathbf{g}^1)_\lambda, \dots, (\mathbf{g}^n)_\lambda)$ denoting the vector of the n channel components for some fixed frame index λ , $1 \leq q \leq 2$ and $\omega_\lambda \geq c > 0$. The case $p = 1$ coincides with the suggestion in [11]. In order to apply general concepts of convex optimization for nonlinear inverse problems as elaborated in [14], we have to verify at first that (2.4) is indeed a weak lower semi-continuous functional.

Lemma 1 *Let $\{\mathbf{x}_k\}$ be a sequence for which $\mathbf{x}_k \in \ell_2 \cap \ell_1$ and $\mathbf{x}_k \xrightarrow{\ell_2} \mathbf{x}$. Then $\mathbf{x} \in \ell_1$ with $\|\mathbf{x}\|_{\ell_1} \leq \liminf_k \|\mathbf{x}_k\|_{\ell_1}$.*

Proof. Given $\mathbf{x}_k = ((\mathbf{x}_k)_1, (\mathbf{x}_k)_2, \dots) \in \ell_2 \cap \ell_1$. Define $f_k(t) = \sum_{i=1}^\infty (\mathbf{x}_k)_i \chi_{(i-1, i]}(t)$. Then, clearly $\|f_k\|_{L_1} = \|\mathbf{x}_k\|_{\ell_1}$ and $\|f_k\|_{L_2} = \|\mathbf{x}_k\|_{\ell_2}$. Assume now $\mathbf{x}_k \xrightarrow{\ell_2} \mathbf{x}$ and construct also some $f \in L_2$ with respect to \mathbf{x} . Consequently, $\|f_k - f\|_{L_2} = \|\mathbf{x}_k - \mathbf{x}\|_{\ell_2} \rightarrow 0$. Since the coefficients of \mathbf{x}_k converge to the coefficients of \mathbf{x} , one has point-wise convergence of f_k towards f and therefore also $|f_k| \rightarrow |f|$ point-wise with $\|f_k\|_{L_1} = \|\mathbf{x}_k\|_{\ell_1}$ and $\|f\|_{L_1} = \|\mathbf{x}\|_{\ell_1}$. Hence, with Fatou's Lemma, we have

$$\|\mathbf{x}\|_{\ell_1} = \int |f(t)| dt = \int \lim |f_k(t)| dt \leq \liminf \int |f_k(t)| dt = \liminf \|\mathbf{x}_k\|_{\ell_1} .$$

□

Lemma 2 *Let $p = 1$ and let ω denote the weight sequence $\{\omega_\lambda\}_{\lambda \in \Lambda}$ in (2.4), then the functional Ψ as defined in (2.4) is weakly lower semi-continuous with respect to weighted sequence space $\ell_{2,\omega}$.*

Proof. Ψ is obviously convex. Therefore it is enough to show lower semi-continuity. Assume $\mathbf{g}_k^i \xrightarrow{\ell_2, \omega} \mathbf{g}^i$. Define $(\mathbf{x}_k)_\lambda := \omega_\lambda \|(\mathbf{g}_k)_\lambda\|_q$ and $\mathbf{x}_\lambda := \omega_\lambda \|\mathbf{g}_\lambda\|_q$. Then, $\mathbf{x}_k \xrightarrow{\ell_2} \mathbf{x}$. This is true because

$$\begin{aligned} \|\mathbf{x}_k - \mathbf{x}\|_{\ell_2}^2 &= \sum_{\lambda \in \Lambda} (\omega_\lambda \|\mathbf{g}_k\|_q - \omega_\lambda \|\mathbf{g}\|_q)^2 \\ &\leq \sum_{\lambda \in \Lambda} (\omega_\lambda \|\mathbf{g}_k - \mathbf{g}\|_q)^2 = \sum_{\lambda \in \Lambda} \omega_\lambda^2 \left(\sum_{i=1}^n |(\mathbf{g}_k^i)_\lambda - (\mathbf{g}^i)_\lambda|^p \right)^{2/p} \\ &\leq n^{2/p} \sum_{i=1}^n \sum_{\lambda \in \Lambda} \omega_\lambda^2 |(\mathbf{g}_k^i)_\lambda - (\mathbf{g}^i)_\lambda|^2 = n^{2/p} \sum_{i=1}^n \|\mathbf{g}_k^i - \mathbf{g}^i\|_{\ell_2, \omega}^2 \rightarrow 0. \end{aligned}$$

With the help of Lemma 2, it now follows

$$\Psi(\mathbf{g}) = \sum_{\lambda \in \Lambda} |\mathbf{x}_\lambda| = \|\mathbf{x}\|_{\ell_1} \leq \liminf_k \|\mathbf{x}_k\|_{\ell_1} = \liminf_k \Psi(\mathbf{g}_k).$$

□

Lemma 3 *Let $p = q$, then the functional Ψ as defined in (2.4) is weakly lower semi-continuous with respect to ℓ_2 .*

Proof. Since we do not have the triangle inequality, we prove the statement directly. The index set Λ is isomorphic to \mathbb{Z} . Therefore,

$$\Psi(\mathbf{g}) = \sum_{m \in \mathbb{Z}} \tilde{\omega}_m \sum_{i=1}^n |\tilde{\mathbf{g}}_m^i|^q,$$

where $\tilde{\omega}_m$ and $\tilde{\mathbf{g}}_m^i$ are re-ordered and re-indexed versions of ω_λ and \mathbf{g}_λ^i . We define the function $f(x) = \sum_{m \in \mathbb{Z}} f_m(x)$ with

$$f_m(x) = \begin{cases} 0 & \text{for } x \notin (m, m+1] \\ f_{m,i}(x) & \text{for } x \in (m, m+1], i = 1, \dots, n \end{cases},$$

where $f_{m,i}(x)$ is defined by

$$f_{m,i}(x) = \begin{cases} n \tilde{\omega}_m |\tilde{\mathbf{g}}_m^i|^q & \text{for } x \in (m + \frac{i-1}{n}, m + \frac{i}{n}] \\ 0 & \text{elsewhere} \end{cases}. \quad (2.5)$$

Therefore,

$$\int_{\mathbb{R}} f(x) dx = \sum_{m \in \mathbb{Z}} \tilde{\omega}_m \sum_{i=1}^n |\tilde{\mathbf{g}}_m^i|^q = \Psi(\mathbf{g}).$$

Analogously we may define a sequence of functions $f^{(k)}(x) = \sum_{m \in \mathbb{Z}} f_m^{(k)}(x)$ by replacing $\tilde{\mathbf{g}}_m^i$ with $(\tilde{\mathbf{g}}_k)_m^i$ in (2.5). If now $(\mathbf{g}_k^i) \xrightarrow{w-\ell_2} \mathbf{g}^i$ for $i = 1, \dots, n$, we have in particular $(\mathbf{g}_k)_\lambda^i \rightarrow (\mathbf{g})_\lambda^i$ and thus for fixed m, i it follows that $\tilde{\omega}_m |(\tilde{\mathbf{g}}_k)_m^i|^q \rightarrow \tilde{\omega}_m |\tilde{\mathbf{g}}_m^i|^q$ as $k \rightarrow \infty$. Consequently, we observe $f^{(k)}(x) \rightarrow f(x)$ pointwise. With the help of Fatou's lemma,

$$\Psi(\mathbf{g}) = \int_{\mathbb{R}} f(x) dx \leq \liminf \int_{\mathbb{R}} f^{(k)}(x) dx = \liminf \Psi(\mathbf{g}_k)$$

□

Therefore the basic principles suggested in [13, 14] up to the statements that directly involve the structure of Ψ (necessary condition for a minimizer, convergence and regularization results) seem to be also adequate when dealing with this special class of joint sparsity measures.

Summarizing the considerations made above yields the following adjustment of (1.1),

$$J_\alpha(\mathbf{g}) = J_{\alpha,q}(\mathbf{g}) = \|\mathbf{y}^\delta - T(\mathbf{g})\|_Y^2 + 2\alpha \sum_{\lambda \in \Lambda} \omega_\lambda \|\mathbf{g}_\lambda\|_q^p. \quad (2.6)$$

The general idea for solving the nonlinear inverse problem, i.e. to find a way to approach a minimizer of (2.6), is to replace (2.6) by a sequence of functionals from which we hope that they are much easier to treat and that the sequence of minimizers converge to a critical point of (2.6), at least. To be more concrete, for $\mathbf{g} \in (\ell_2)^n$ and some auxiliary $\mathbf{a} \in (\ell_2)^n$, we define

$$J_{\alpha,q}^s(\mathbf{g}, \mathbf{a}) := J_{\alpha,q}(\mathbf{g}) + C\|\mathbf{g} - \mathbf{a}\|_{(\ell_2)^n}^2 - \|T(\mathbf{g}) - T(\mathbf{a})\|_Y^2 \quad (2.7)$$

and create an iteration process by:

1. Pick $\mathbf{g}_0 \in (\ell_2)^n$ and some proper constant $C > 0$
2. Derive a sequence $\{\mathbf{g}_k\}_{k=0,1,\dots}$ by the iteration:

$$\mathbf{g}_{k+1} = \arg \min_{\mathbf{g}_k \in (\ell_2)^n} J_{\alpha,q}^s(\mathbf{g}, \mathbf{g}_k) \quad k = 0, 1, 2, \dots \quad (2.8)$$

In order to ensure norm convergence of the iterates \mathbf{g}_k and regularization properties, we have to restrict ourselves to a class of nonlinear problems for which the following requirements hold true,

$$\begin{aligned} \mathbf{g}_k \xrightarrow{w} \mathbf{g} &\implies T(\mathbf{g}_k) \rightarrow T(\mathbf{g}) \quad , \\ T'(\mathbf{g}_k)^* z &\rightarrow T'(\mathbf{g})^* z \quad , \text{ for all } z \quad , \\ \|T'(\mathbf{g}) - T'(\mathbf{g}')\|_Y &\leq L\|\mathbf{g} - \mathbf{g}'\|_{(\ell_2)^n}. \end{aligned} \quad (2.9)$$

The latter requirements exhibit no significant restriction of the proposed approach and are of common use in the literature. In particular, they characterize continuity and smoothness properties of T and T' , respectively.

3 Minimization of the Surrogate Functional

Since a very similar functional (up to the structure of Ψ) was under consideration in [14], we just review the basic properties and refer the interested reader to [14] for proofs and rigorous arguments. Then, for the computation of the necessary condition for a minimum of $J_{\alpha,q}^s$ (as defined in (2.7)), the specific structure of Ψ comes into play. At this point, the coupling of the frame coefficients via the ℓ_q -norm becomes an important aspect.

Let us start by reviewing some basic properties of our functional. Assume, we are given some $\alpha > 0$ and some initial $\mathbf{g}_0 \in (\ell_2)^n$, then we may define a ball

$$K_r := \{\mathbf{g} \in (\ell_2)^n : \Psi(\mathbf{g}) \leq r\},$$

for which the radius r is determined by

$$r := J_{\alpha,q}(\mathbf{g}_0)/(2\alpha), \quad (3.1)$$

ensuring $\mathbf{g}_0 \in K_r$. As the most important ingredient in the definition of (2.7), we define the constant C ,

$$C := 2B^2 \max \left\{ \left(\sup_{\mathbf{g} \in K_r} \|T'(\mathbf{g})\| \right)^2, L \sqrt{\|y^\delta - T(\mathbf{g}_0)\|^2 + 2\alpha\Psi(\mathbf{g}_0)} \right\}, \quad (3.2)$$

where L is the Lipschitz constant of the Fréchet derivative of T . We suppose that \mathbf{g}_0 was chosen such that $r < \infty$ and $C < \infty$.

Lemma 4 *Let r and C be chosen by (3.1), (3.2). Then, for all $\mathbf{g} \in K_r$,*

$$C\|\mathbf{g} - \mathbf{g}_0\|_{(\ell_2)^n}^2 - \|T(\mathbf{g}) - T(\mathbf{g}_0)\|_Y^2 \geq 0 \quad (3.3)$$

and thus, $J_{\alpha,q}(\mathbf{g}) \leq J_{\alpha,q}^s(\mathbf{g}, \mathbf{g}_0)$.

The proof essentially relies on the Taylor expansion of T ,

$$T(\mathbf{g} + \mathbf{h}) = T(\mathbf{g}) + \int_0^1 T'(\mathbf{g} + \tau\mathbf{h})\mathbf{h} d\tau$$

and can be retraced, e.g., from [14]. Moreover, one has

Proposition 5 *Let \mathbf{g}_0, α be given and r, C be defined by (3.1), (3.2). Then the functionals $J_{\alpha,q}^s(\mathbf{g}, \mathbf{g}_k)$ are bounded from below for all $k \in \mathbb{N}$ and have thus minimizers. For the minimizer \mathbf{g}_{k+1} of $J_{\alpha,q}^s(\mathbf{g}, \mathbf{g}_k)$ holds $\mathbf{g}_{k+1} \in K_r$.*

As an immediate consequence we have

Corollary 6 *The sequences of functional values $\{J_{\alpha,q}(\mathbf{g}_k)\}_{k \in \mathbb{N}}$ and $\{J_{\alpha,q}^s(\mathbf{g}_{k+1}, \mathbf{g}_k)\}_{k \in \mathbb{N}}$ are non-increasing.*

In the next lemma, we characterize a minimizer of the functional $J_{\alpha,q}^s(\mathbf{g}, \mathbf{a})$ by means of the necessary condition for which it can be shown that it can be cast as a fixed point equation with contractive fixed point map. The associated fixed point iteration can be understood as an inner iteration of iteration (2.8) in order to derive the iterates \mathbf{g}_k .

Lemma 7 *The necessary condition for a minimum of $J_{\alpha,q}^s(\mathbf{g}, \mathbf{a})$ is given by*

$$0 \in -T'(\mathbf{g})^*(y^\delta - T(\mathbf{a})) + C\mathbf{g} - C\mathbf{a} + \alpha\partial\Psi(\mathbf{g}). \quad (3.4)$$

The latter condition is written in the notion of subgradients. The structure of $\partial\Psi(\mathbf{g})$ depends now on the explicit choice of p . For $p = 1$, the functional Ψ is one-homogeneous. Therefore, by duality arguments, Ψ can be associated with a closed convex set \mathcal{C} , which goes as follows: Let Ψ^* denote the Fenchel transform (or so-called dual functional) of Ψ . Then there is an associated convex set \mathcal{C} such that $\Psi^* = \chi_{\mathcal{C}}$, and, moreover, $\Psi = (\chi_{\mathcal{C}})^*$. Hence, in the situation $p = 1$, we have the following characterization of the necessary condition.

Lemma 8 *Let $p = 1$, then the necessary condition (3.4) can be rewritten as*

$$\mathbf{g} = \frac{\alpha}{C} (I - P_{\mathcal{C}}) \left(\frac{C}{\alpha} (T'(\mathbf{g})^*(y^\delta - T(\mathbf{a}))/C + \mathbf{a}) \right). \quad (3.5)$$

where $P_{\mathcal{C}}$ is the orthogonal projection onto a convex set \mathcal{C} .

Since the techniques to show this statement are of importance when deriving the explicit structure of the projection, we give the proof in full detail.

Proof. Let $M(\mathbf{g}, \mathbf{a}) := T'(\mathbf{g})^*(y^\delta - T(\mathbf{a}))/C + \mathbf{a}$. Then we may rewrite (3.4) as follows,

$$\frac{M(\mathbf{g}, \mathbf{a}) - \mathbf{g}}{\frac{\alpha}{C}} \in \partial\Psi(\mathbf{g}),$$

and thus, by means of the dual functional,

$$\frac{C}{\alpha} \mathbf{g} \in \frac{C}{\alpha} \partial\Psi^* \left(\frac{M(\mathbf{g}, \mathbf{a}) - \mathbf{g}}{\frac{\alpha}{C}} \right).$$

The latter formula can then be expanded as follows,

$$\begin{aligned} \frac{M(\mathbf{g}, \mathbf{a})}{\frac{\alpha}{C}} &\in \frac{M(\mathbf{g}, \mathbf{a}) - \mathbf{g}}{\frac{\alpha}{C}} + \frac{C}{\alpha} \partial\Psi^* \left(\frac{M(\mathbf{g}, \mathbf{a}) - \mathbf{g}}{\frac{\alpha}{C}} \right) \\ &= \left(I + \frac{C}{\alpha} \partial\Psi^* \right) \left(\frac{M(\mathbf{g}, \mathbf{a}) - \mathbf{g}}{\frac{\alpha}{C}} \right), \end{aligned}$$

which yields

$$\left(I + \frac{C}{\alpha} \partial\Psi^* \right)^{-1} \left(\frac{M(\mathbf{g}, \mathbf{a})}{\frac{\alpha}{C}} \right) = \frac{M(\mathbf{g}, \mathbf{a}) - \mathbf{g}}{\frac{\alpha}{C}}.$$

The operator $\left(I + \frac{C}{\alpha} \partial\Psi^* \right)^{-1}$ is nothing than the orthogonal projection on an associated convex set \mathcal{C} , and hence the assertion follows,

$$\mathbf{g} = \frac{\alpha}{C} (I - P_{\mathcal{C}}) \left(\frac{M(\mathbf{g}, \mathbf{a})}{\frac{\alpha}{C}} \right).$$

□

Lemma 8 provides us with a fixed point problem for which the following result can be easily deduced.

Lemma 9 *Let the fixed point map*

$$\Phi(\mathbf{g}, \mathbf{a}) = \frac{\alpha}{C}(I - P_{\mathcal{C}}) \left(\frac{M(\mathbf{g}, \mathbf{a})}{\frac{\alpha}{C}} \right)$$

be given. For some generic anchor \mathbf{a} , the operator $\Phi(\cdot, \mathbf{a})$ is a contraction as long as $L/C\sqrt{J_{\alpha,q}(\mathbf{a})} < 1$.

Proof. As it can be retrieved, e.g. from [7] or [14], the operator $I - P_{\mathcal{C}}$ is non-expansive as long as \mathcal{C} is a closed and convex set. Then, with the help of the Lipschitz-continuity of T' , we have

$$\begin{aligned} \|\Phi(\mathbf{g}, \mathbf{a}) - \Phi(\tilde{\mathbf{g}}, \mathbf{a})\|_{(\ell_2)^n}^2 &= \frac{\alpha^2}{C^2} \left\| (I - P_{\mathcal{C}}) \left(\frac{M(\mathbf{g}, \mathbf{a})}{\frac{\alpha}{C}} \right) - (I - P_{\mathcal{C}}) \left(\frac{M(\tilde{\mathbf{g}}, \mathbf{a})}{\frac{\alpha}{C}} \right) \right\|_{(\ell_2)^n}^2 \\ &\leq \|M(\mathbf{g}, \mathbf{a}) - M(\tilde{\mathbf{g}}, \mathbf{a})\|_{(\ell_2)^n}^2 \\ &\leq \frac{1}{C^2} \|T'(\mathbf{g}) - T'(\tilde{\mathbf{g}})\|^2 \|y^\delta - T(\mathbf{a})\|_Y^2 \\ &\leq \frac{J_\alpha(\mathbf{a})L^2}{C^2} \|\mathbf{g} - \tilde{\mathbf{g}}\|_{(\ell_2)^n}^2 \end{aligned}$$

and the assertion follows. \square

Setting now $\mathbf{a} = \mathbf{g}_0$, it is clear that $\Phi(\cdot, \mathbf{g}_0)$ is a contraction with $\sqrt{J_\alpha(\mathbf{g}_0)}L/C \leq 1/2 < 1$. With the help of Corollary 6, i.e. since

$$\sqrt{J_\alpha(\mathbf{g}_k)}L/C \leq \sqrt{J_\alpha(\mathbf{g}_{k-1})}L/C \leq \dots \leq \sqrt{J_\alpha(\mathbf{g}_0)}L/C ,$$

we deduce that for each \mathbf{g}_k the fixed point map $\Phi(\cdot, \mathbf{g}_k)$ is a contraction.

Let us now explicitly evaluate the application of $I - P_{\mathcal{C}}$ by incorporating the special structure of Ψ for $p = 1$. To this end, we index the fixed point iterates with l . Note, the meaning of the index l is different than k , the k indexing stands for iteration (2.8) whereas the l indexing stands for the fixed point iteration which is used to derive the iterates of (2.8),

$$\mathbf{g}_{l+1} = \frac{\alpha}{C}(I - P_{\mathcal{C}}) \left(\frac{M(\mathbf{g}_l, \mathbf{a})}{\frac{\alpha}{C}} \right) . \quad (3.6)$$

Proposition 10 *Let $p = 1$ and $1 \leq q \leq \infty$. The coefficients of the fixed point iterates in (3.6) are given by*

$$(\mathbf{g}_{l+1})_\lambda = ((\mathbf{g}_{l+1}^1)_\lambda, \dots, (\mathbf{g}_{l+1}^n)_\lambda) = (I - P_{B_{q'}(C^{-1}\alpha\omega_\lambda)})(M_\lambda(\mathbf{g}_l, \mathbf{a})) ,$$

where $P_{B_{q'}(C^{-1}\alpha\omega_\lambda)}$ denotes the orthogonal projection onto the ball

$$B_{q'}(C^{-1}\alpha\omega_\lambda) = \{\mathbf{v} \in \mathbb{C}^n : \|\mathbf{v}\|_{q'} \leq C^{-1}\alpha\omega_\lambda\}$$

with $1 = 1/q + 1/q'$.

Proof. Similar to the proof of Lemma 8, formula (3.6) can be equivalently cast as,

$$-M(\mathbf{g}_l, \mathbf{a}) + \mathbf{g}_{l+1} + \frac{\alpha}{C} \partial \Psi(\mathbf{g}_{l+1}) \ni 0 .$$

which is the necessary condition for the functional

$$\Omega(\mathbf{u}) = \|\mathbf{u} - \mathbf{v}\|_{(\ell_2)^n}^2 + 2\gamma \sum_{\lambda \in \Lambda} \omega_\lambda \|\mathbf{u}_\lambda\|_q , \quad (3.7)$$

with the substitutes $\mathbf{u} = \mathbf{g}_{l+1}$, $\mathbf{v} = M(\mathbf{g}_l, \mathbf{a})$ and $\gamma = \alpha/C$. Taking into account the vector valued structure of the problem, (3.7) reads as

$$\begin{aligned} \Omega(\mathbf{u}) &= \sum_{i=1}^n \sum_{\lambda \in \Lambda} |(\mathbf{u}^i)_\lambda - (\mathbf{v}^i)_\lambda|_{\ell_2}^2 + 2\gamma \sum_{\lambda \in \Lambda} \omega_\lambda \left(\sum_{i=1}^n |(\mathbf{u}^i)_\lambda|^q \right)^{1/q} \\ &= \sum_{\lambda \in \Lambda} \left\{ \sum_{i=1}^n |(\mathbf{u}^i)_\lambda - (\mathbf{v}^i)_\lambda|_{\ell_2}^2 + 2\gamma \omega_\lambda \left(\sum_{i=1}^n |(\mathbf{u}^i)_\lambda|^q \right)^{1/q} \right\} \end{aligned} \quad (3.8)$$

which can thus be minimized with respect to each individual component, i.e. for each λ we have to minimize

$$\|\mathbf{u}_\lambda - \mathbf{v}_\lambda\|^2 + 2\gamma \omega_\lambda \|\mathbf{u}_\lambda\|_q .$$

By standard arguments, the minimizer is given by

$$(\mathbf{u}^*)_\lambda = \gamma \omega_\lambda (Id - P_{\mathcal{C}})((\gamma \omega_\lambda)^{-1} \mathbf{v}_\lambda) ,$$

where the associated convex set \mathcal{C} is given by the unit ball $B_{q'}(1)$ in the dual norm of $\|\cdot\|_q$ and the proof is complete. \square

In general, the evaluation of $P_{B_{q'}(C^{-1}\alpha\omega_\lambda)}$ is rather difficult and only for individual cases explicitly given, e.g. $q' \in \{1, 2, \infty\}$, see Section 6. Let us now consider the case $p = q$ which can be directly evaluated without operating with duality arguments.

Lemma 11 *Let $p = q$ and $M_\lambda(\mathbf{g}, \mathbf{a}) = (M_\lambda^1(\mathbf{g}, \mathbf{a}), \dots, M_\lambda^n(\mathbf{g}, \mathbf{a}))'$ as before, then the necessary condition (3.4) can be expressed componentwise,*

$$\mathbf{g}_\lambda^i + \frac{\alpha}{C} q \omega_\lambda \text{sgn}(\mathbf{g}_\lambda^i) |\mathbf{g}_\lambda^i|^{q-1} = M_\lambda^i(\mathbf{g}, \mathbf{a}) . \quad (3.9)$$

Proof. A direct computation of the derivative. \square

In order to solve (3.9), we may write as before a fixed point iteration,

$$(\mathbf{g}_{l+1}^i)_\lambda + \frac{\alpha}{C} q \omega_\lambda \text{sgn}((\mathbf{g}_{l+1}^i)_\lambda) |(\mathbf{g}_{l+1}^i)_\lambda|^{q-1} = M_\lambda^i(\mathbf{g}_l, \mathbf{a}) .$$

As it can be retraced in [4], the map $F_{\alpha\omega_\lambda/C, q}(x) = x + \frac{\alpha}{C} q \omega_\lambda \text{sgn}(x) |x|^{q-1}$ is for $\omega_\lambda \geq 0$ and any $q > 1$ a one-to-one map from \mathbb{R} to itself, we thus find that

$$(\mathbf{g}_{l+1}^i)_\lambda = S_{\alpha\omega_\lambda/C, q}(M_\lambda^i(\mathbf{g}_l, \mathbf{a})) ,$$

where $S_{\alpha\omega_\lambda/C,q}$ is defined by $S_{\alpha\omega_\lambda/C,q} = (F_{\alpha\omega_\lambda/C,q})^{-1}$ for $q > 1$. For $q = 1$, $S_{\alpha\omega_\lambda/C,1}$ is nothing than the well-known soft-shrinkage operator with threshold $\alpha\omega_\lambda/C$. Introducing the operator

$$\mathbf{S}_{\alpha\omega/C,q}(\mathbf{g}) = \{(S_{\alpha\omega_\lambda/C,q}(\mathbf{g}_\lambda^1), \dots, S_{\alpha\omega_\lambda/C,q}(\mathbf{g}_\lambda^n))\}_{\lambda \in \Lambda}$$

and taking into account that the operators $S_{\alpha\omega_\lambda/C,q}$ are non-expansive (as it was shown in [4]), Lemma 9 holds true also for $p = q$.

Lemma 12 *For each k , the fixed point maps*

$$\Phi(\mathbf{g}, \mathbf{g}_k) = \mathbf{S}_{\alpha\omega/C,q}(M(\mathbf{g}, \mathbf{g}_k))$$

are contractions with damping factor $L/C\sqrt{J_{\alpha,q}(\mathbf{g}_k)} < 1$.

Therefore, iteration (2.8) is derived for $p = 1$ and $p = q$ in the same way. The only difference is the evaluation of the inner fixed point iteration. We finally have:

Proposition 13 *The necessary equation (3.5) for a minimum of the functional $J_{\alpha,q}^s(\mathbf{g}, \mathbf{g}_k)$ has a unique fixed point, and the fixed point iteration converges towards the minimizer.*

4 Norm Convergence of the Scheme

Within this section we show that the sequence of iterates $\{\mathbf{g}_k\}$ converges strongly towards a critical point of J_α . The first two steps can be retraced from [14]:

1. Since the iterates \mathbf{g}_k are uniformly bounded in the $(\ell_2)^n$ topology, one may extract a weakly convergent subsequence \mathbf{g}_{k_l} with weak limit \mathbf{g}_* , see Lemma 13.
2. The weak limit \mathbf{g}_* satisfies the necessary condition for a minimizer of J_α , see Lemma 16.

It remains to show that \mathbf{g}_{k_l} converges in the $(\ell_2)^n$ norm and that the convergence carries over to the whole sequence \mathbf{g}_k .

Theorem 14 *Let $\{\mathbf{g}_{k_l}\} \subset \{\mathbf{g}_k\}$ with $\mathbf{g}_{k_l} \xrightarrow{w} \mathbf{g}_*$. Assume, moreover, that uniformly for all \mathbf{g}_k ,*

$$\Psi(\mathbf{g}_k) = \sum_{\lambda} \omega_\lambda \|(\mathbf{g}_k)_\lambda\|_q^p \leq \tilde{c} < \infty \quad (4.1)$$

with $\omega_\lambda \geq c > 0$ for all $\lambda \in \Lambda$. Then, for $1 \leq q \leq 2$, the subsequence $\{\mathbf{g}_{k_l}\}$ converges in the $(\ell_2)^n$ -norm.

The existence of some finite constant \tilde{c} in (4.1) is due to Corollary 6 automatically satisfied.

Proof. For the case $p = q$ we may proceed as in the proof of Theorem 18 in [14]. We just have to replace $|\cdot|^p$ with $\|\cdot\|_q^p$ and have to make the obvious modifications. The strategy of this proof but with some technical changes applies to $p = 1$ in the following way: Let \mathbf{g}_l be the shorthand notation for \mathbf{g}_{k_l} . Weak convergence for $\mathbf{g}_l = (\mathbf{g}_l^1, \dots, \mathbf{g}_l^n)'$ means that each channel \mathbf{g}_l^i of the vector converges weakly. Therefore, for $i = 1, \dots, n$,

each component $(\mathbf{g}_l^i)_\lambda$ converges strongly towards $(\mathbf{g}_\star^i)_\lambda$ and consequently, denoting with $(\mathbf{g}_l)_\lambda$ the vector $((\mathbf{g}_l^1)_\lambda, \dots, (\mathbf{g}_l^n)_\lambda)'$, it follows that

$$(\mathbf{g}_l)_\lambda \xrightarrow{\|\cdot\|_q} (\mathbf{g}_\star)_\lambda$$

as $l \rightarrow \infty$. The goal is now to show that

$$\|\mathbf{g}_l\|_{(\ell_2)^n}^2 \longrightarrow \|\mathbf{g}_\star\|_{(\ell_2)^n}^2 \quad \text{as } l \rightarrow \infty.$$

To this end, consider the discrepancy

$$D := \left| \sum_{\lambda} \|(\mathbf{g}_l)_\lambda\|_2^2 - \sum_{\lambda} \|(\mathbf{g}_\star)_\lambda\|_2^2 \right|.$$

Divide now the index set Λ into $\tilde{\Lambda}$ for which $|\tilde{\Lambda}| = N$ holds and the complement $\Lambda \setminus \tilde{\Lambda}$. Then, by the triangle inequality,

$$D \leq \left| \sum_{\lambda \in \tilde{\Lambda}} (\|(\mathbf{g}_l)_\lambda\|_2^2 - \|(\mathbf{g}_\star)_\lambda\|_2^2) \right| + \sum_{\lambda \in \Lambda \setminus \tilde{\Lambda}} \|(\mathbf{g}_l)_\lambda\|_2^2 + \sum_{\lambda \in \Lambda \setminus \tilde{\Lambda}} \|(\mathbf{g}_\star)_\lambda\|_2^2. \quad (4.2)$$

Since $1 \leq q \leq 2$, the $(\ell_2)^n$ -norm of \mathbf{g}_l satisfying $\Psi(\mathbf{g}_l) < \infty$ can be bounded by Hölder's inequality (with $1/q' + 1/q = 1$, implying $q' \geq q$),

$$\sum_{\lambda} \sum_{i=1}^n |(\mathbf{g}_l)_\lambda^i|^2 \leq \sum_{\lambda} \omega_{\lambda} \|(\mathbf{g}_l)_\lambda\|_q c^{-1} \|(\mathbf{g}_l)_\lambda\|_{q'}.$$

To find a uniform bound (with respect to l and λ) for the ℓ_p -norm (which exists because of $\ell_q \subseteq \ell_{q'}$), consider

$$\|(\mathbf{g}_l)_\lambda\|_{q'}^{q'} \leq \max_{1 \leq i \leq n} |(\mathbf{g}_l)_\lambda^i|^{q'-q} \|(\mathbf{g}_l)_\lambda\|_q^q.$$

Since $\omega_{\lambda} \geq c > 0$, it follows that

$$\sum_{\lambda \in \Lambda} \|(\mathbf{g}_l)_\lambda\|_q \leq c^{-1} \Psi(\mathbf{g}_l) \leq c^{-1} \tilde{c}$$

and therefore $\|(\mathbf{g}_l)_\lambda\|_q^q \leq (c^{-1} \tilde{c})^q$ and $\max_i |(\mathbf{g}_l)_\lambda^i|^{q'-q} \leq (c^{-1} \tilde{c})^{q'-q}$, implying $\|(\mathbf{g}_l)_\lambda\|_{q'} \leq c^{-1} \tilde{c}$. Consequently, defining $C_{\Psi} := c^{-2} \tilde{c}$, it holds $\|\mathbf{g}_l\|_{(\ell_2)^n}^2 \leq C_{\Psi} \Psi(\mathbf{g}_l)$. For some fixed $\varepsilon > 0$, choose now N such that

$$\sum_{\lambda \in \Lambda \setminus \tilde{\Lambda}} \omega_{\lambda} \|(\mathbf{g}_\star)_\lambda\|_q \leq \frac{\varepsilon}{5} C_{\Psi}^{-1}. \quad (4.3)$$

It follows immediately that

$$\sum_{\lambda \in \Lambda \setminus \tilde{\Lambda}} \|(\mathbf{g}_\star)_\lambda\|_2^2 \leq C_{\Psi} \sum_{\lambda \in \Lambda \setminus \tilde{\Lambda}} \omega_{\lambda} \|(\mathbf{g}_\star)_\lambda\|_q \leq \frac{\varepsilon}{5}. \quad (4.4)$$

Choosing now the iteration index l large enough such that

$$\sum_{\lambda \in \Lambda} \omega_{\lambda} \|(\mathbf{g}_l)_\lambda\|_q = \sum_{\lambda \in \Lambda} \omega_{\lambda} \|(\mathbf{g}_\star)_\lambda\|_q + \tilde{\varepsilon} \quad (4.5)$$

and for $\lambda \in \tilde{\Lambda}$

$$\omega_\lambda \|(\mathbf{g}_l)_\lambda\|_q = \omega_\lambda \|(\mathbf{g}_\star)_\lambda\|_q + \tilde{\varepsilon}/N \quad , \quad \|(\mathbf{g}_l)_\lambda\|_2^2 = \|(\mathbf{g}_\star)_\lambda\|_2^2 + \tilde{\varepsilon}/N \quad , \quad (4.6)$$

with $|\tilde{\varepsilon}| \leq C_\Psi^{-1}\varepsilon/5$ and $|\tilde{\varepsilon}| \leq \varepsilon/5$. (4.5) is assured by Lemma 17, [14], and (4.6) can be fulfilled as N is fixed and \mathbf{g}_l converges weakly to \mathbf{g}_\star . It follows

$$\begin{aligned} \sum_{\lambda \in \Lambda \setminus \tilde{\Lambda}} \|(\mathbf{g}_l)_\lambda\|_2^2 &\leq C_\Psi \sum_{\lambda \in \Lambda \setminus \tilde{\Lambda}} \omega_\lambda \|(\mathbf{g}_l)_\lambda\|_q = C_\Psi \left\{ \sum_{\lambda \in \Lambda} \omega_\lambda \|(\mathbf{g}_l)_\lambda\|_q - \sum_{\lambda \in \tilde{\Lambda}} \omega_\lambda \|(\mathbf{g}_l)_\lambda\|_q \right\} \\ &\stackrel{(4.5)(4.6)}{\leq} C_\Psi \left\{ \sum_{\lambda \in \Lambda} \omega_\lambda \|(\mathbf{g}_\star)_\lambda\|_q + |\tilde{\varepsilon}| - \sum_{\lambda \in \tilde{\Lambda}} \omega_\lambda \|(\mathbf{g}_\star)_\lambda\|_q + N \frac{|\tilde{\varepsilon}|}{N} \right\} \\ &= C_\Psi \left\{ \sum_{\lambda \in \Lambda \setminus \tilde{\Lambda}} \omega_\lambda \|(\mathbf{g}_\star)_\lambda\|_q + 2|\tilde{\varepsilon}| \right\} \stackrel{(4.3)}{\leq} \frac{3}{5}\varepsilon \quad . \end{aligned} \quad (4.7)$$

Moreover, by (4.6),

$$\left| \sum_{\lambda \in \tilde{\Lambda}} \{ \|(\mathbf{g}_l)_\lambda\|_2^2 - \|(\mathbf{g}_\star)_\lambda\|_2^2 \} \right| \leq \sum_{\lambda \in \tilde{\Lambda}} |\tilde{\varepsilon}|/N \leq \frac{\varepsilon}{5} \quad (4.8)$$

we obtain by combining estimates (4.4), (4.7) and (4.8) into (4.2), $D \leq \varepsilon$ and consequently,

$$\lim_{l \rightarrow \infty} \|\mathbf{g}_l\|_{(\ell_2)^n} = \|\mathbf{g}_\star\|_{(\ell_2)^n} \quad .$$

□

The latter theorem ensures norm convergence for each weakly convergent subsequence. However, the limits of different convergent subsequences $\{\mathbf{g}_{k_l}\} \subset \{\mathbf{g}_k\}$ may differ. Choose now some arbitrary norm convergent subsequence $\{\mathbf{g}_{k_l}\}$ and denote with \mathbf{g}'_{k_l} the predecessor of \mathbf{g}_{k_l} in $\{\mathbf{g}_k\}$. Then, $J_{\alpha,q}^s(\mathbf{g}_{k_l}, \mathbf{g}'_{k_l}) \rightarrow J_{\alpha,q}(\mathbf{g}_\star)$. Moreover, since $J_{\alpha,q}^s(\mathbf{g}_{k+1}, \mathbf{g}_k) \leq J_{\alpha,q}^s(\mathbf{g}_k, \mathbf{g}_{k-1})$ for all k , it turns out that the value of the Tikhonov functional for every limit \mathbf{g}_\star of a convergent subsequence remains the same. In the next theorem we give a simple criterion that ensures strong convergence of the whole sequence $\{\mathbf{g}_k\}$ towards a critical point of $J_{\alpha,q}$.

Theorem 15 *Assume that there exists at least one isolated limit \mathbf{g}_\star of a subsequence $\{\mathbf{g}_{k_l}\} \subset \{\mathbf{g}_k\}$. Then $\mathbf{g}_k \rightarrow \mathbf{g}_\star$ as $k \rightarrow \infty$. The accumulation point \mathbf{g}_\star fulfills the necessary condition for a minimizer of $J_{\alpha,q}$.*

Proof. The first assertion can be directly taken from [13]. The second assertion is just repeating Lemma 16, [14]. □

5 Regularization Result for A–Priori Parameter Rules

We achieve regularization results at two levels. At the first level, which holds true for $p = \{1, q\}$, we verify that for some a–priori parameter choice rule for α the sequence of

minimizers (for $\delta \rightarrow 0$) contains a weakly convergent subsequence for which the limit is a solution for $T(\mathbf{g}) = y$ with minimal Ψ value. At the second level, we show norm convergence. This result is so far restricted to $p = q$.

The first regularization result is very general and applies to Tikhonov functionals of the form

$$J_{\alpha,q}(x) = \|y^\delta - T(x)\|^2 + 2\alpha\Psi(x - x_0) ,$$

with a-priori guess x_0 , a strongly continuous operator T between Banach spaces X and Y , and a penalty functional $\Psi : X \rightarrow \mathbb{R}$ that is weakly lower semi-continuous with respect to weak convergence in X . Moreover, we assume that every sequence $x_n \in X$ with bounded value of $\Psi(x_n - x_0)$ has a weakly convergent subsequence. This setting fits in particular with the mentioned functional (2.6) for $p = \{1, q\}$ and $x_0 = 0$.

Theorem 16 *Let $y^\delta \in Y$ be given with $\|y^\delta - y\| \leq \delta_k$ and $\delta_k \rightarrow 0$ for $k \rightarrow \infty$. If the regularization parameter is chosen such that $\alpha_k \rightarrow 0$ and $\delta_k^2/\alpha_k \rightarrow 0$ as $k \rightarrow \infty$, then every sequence of minimizers $x_{\alpha_k}^{\delta_k}$ has a weakly convergent subsequence in X . Every weak limit is a solution of the equation $T(x) = y$ with minimal distance to x_0 with respect to Ψ . If the solution of the equation is unique, then the whole sequence converges weakly.*

Proof. This proof generalizes results given in [9]. Let $x_{\alpha_k}^{\delta_k}$ be a sequence of minimizers according to the above given parameter rule, and x^\dagger a solution of $T(x) = y$ with minimal value of Ψ . The minimum property yields

$$\begin{aligned} \|y^{\delta_k} - T(x_{\alpha_k}^{\delta_k})\|^2 &\leq \|y^{\delta_k} - T(x_{\alpha_k}^{\delta_k})\|^2 + \alpha_k \Psi(x_{\alpha_k}^{\delta_k} - x_0) \leq \|y^{\delta_k} - T(x^\dagger)\|^2 + \alpha_k \Psi(x^\dagger - x_0) \\ &\leq \delta_k^2 + \alpha_k \Psi(x^\dagger - x_0) , \end{aligned}$$

and thus,

$$\lim_{k \rightarrow \infty} T(x_{\alpha_k}^{\delta_k}) = y . \quad (5.1)$$

Moreover, we have

$$\Psi(x_{\alpha_k}^{\delta_k} - x_0) \leq \frac{\delta_k^2}{\alpha_k} + \Psi(x^\dagger - x_0) , \quad (5.2)$$

which shows that $\Psi(x_{\alpha_k}^{\delta_k} - x_0)$ is bounded. Therefore we have at least a weakly convergent subsequence of $x_{\alpha_k}^{\delta_k}$ with weak limit x_* . For convenience, we will denote this subsequence also by $x_{\alpha_k}^{\delta_k}$. Strongly continuous operators transform weak convergent sequences into strongly convergent ones, and thus it follows

$$y = \lim_{k \rightarrow \infty} T(x_{\alpha_k}^{\delta_k}) = T(x_*) , \quad (5.3)$$

and x_* is also a solution. By the weakly lower semi-continuity of Ψ we conclude

$$\Psi(x_* - x_0) = \liminf \Psi(x_{\alpha_k}^{\delta_k} - x_0) \leq \limsup \Psi(x_{\alpha_k}^{\delta_k} - x_0) \quad (5.4)$$

$$\stackrel{(5.2)}{\leq} \lim_{k \rightarrow \infty} \left(\frac{\delta_k^2}{\alpha_k} + \Psi(x^\dagger - x_0) \right) = \Psi(x^\dagger - x_0) , \quad (5.5)$$

and x_* is also a solution with minimal distance to x_0 with respect to Ψ . □

Lemma 17 *Let $p = 1$ in the penalty constraint of (2.6). Then the $\|\mathbf{g}_{\alpha_k}^{\delta_k}\|_{(\ell_{2,\omega})^n}^2$ are uniformly bounded and thus there exists a subsequence of $\{\mathbf{g}_{\alpha_k}^{\delta_k}\}$ that converges weakly with respect to $(\ell_{2,\omega})^n$.*

Proof. Inequality (5.2) reads as $\Psi(\mathbf{g}_{\alpha_k}^{\delta_k}) \leq c_k := \delta_k^2/\alpha_k + \Psi(\mathbf{g}^\dagger) < \infty$. Similar as in estimate (4.4), we have by Hölders inequality,

$$\|\mathbf{g}_{\alpha_k}^{\delta_k}\|_{(\ell_{2,\omega})^n}^2 = \sum_{i=1}^n \sum_{\lambda} \omega_{\lambda} |(\mathbf{g}_{\alpha_k}^{\delta_k})_{\lambda}^i|^2 \leq \sum_{\lambda} \omega_{\lambda} \|(\mathbf{g}_{\alpha_k}^{\delta_k})_{\lambda}\|_q \|(\mathbf{g}_{\alpha_k}^{\delta_k})_{\lambda}\|_{q'}.$$

Since $\omega_{\lambda} \geq c > 0$, it follows that

$$\sum_{\lambda} \|(\mathbf{g}_{\alpha_k}^{\delta_k})_{\lambda}\|_q \leq c^{-1} c_k \quad \text{and thus, } |(\mathbf{g}_{\alpha_k}^{\delta_k})_{\lambda}^i| \leq c^{-1} c_k$$

and consequently, due to $q' \geq q$, it follows $\|(\mathbf{g}_{\alpha_k}^{\delta_k})_{\lambda}\|_{q'} \leq c^{-1} c_k$. And therefore, we obtain

$$\limsup_{k \rightarrow \infty} \|\mathbf{g}_{\alpha_k}^{\delta_k}\|_{(\ell_{2,\omega})^n}^2 \leq \limsup_{k \rightarrow \infty} c^{-1} c_k^2 = c^{-1} \Psi^2(\mathbf{g}^\dagger) < \infty,$$

i.e. $\|\mathbf{g}_{\alpha_k}^{\delta_k}\|_{(\ell_{2,\omega})^n}^2$ are uniformly bounded and therefore there exists a subsequence of $\{\mathbf{g}_{\alpha_k}^{\delta_k}\}$ that converges weakly with respect to $(\ell_{2,\omega})^n$. \square

Lemma 18 *Let $p = q$ in the penalty constraint of (2.6). Then the $\|\mathbf{g}_{\alpha_k}^{\delta_k}\|_{(\ell_2)^n}$ are uniformly bounded and thus there exists a subsequence of $\{\mathbf{g}_{\alpha_k}^{\delta_k}\}$ that converges weakly with respect to $(\ell_2)^n$.*

Proof. As in the proof of Lemma 17, we may define the constant c_k . Then, since now

$$|(\mathbf{g}_{\alpha_k}^{\delta_k})_{\lambda}^i| \leq (c^{-1} c_k)^{1/q},$$

we have the following estimate,

$$\begin{aligned} \|\mathbf{g}_{\alpha_k}^{\delta_k}\|_{(\ell_2)^n}^2 &\leq c^{-1} \sum_{\lambda} \omega_{\lambda} \sum_{i=1}^n |(\mathbf{g}_{\alpha_k}^{\delta_k})_{\lambda}^i|^2 \\ &\leq c^{-1} \sum_{\lambda} \omega_{\lambda} \max_i |(\mathbf{g}_{\alpha_k}^{\delta_k})_{\lambda}^i|^{2-q} \sum_{i=1}^n |(\mathbf{g}_{\alpha_k}^{\delta_k})_{\lambda}^i|^q \leq c^{-1} (c^{-1} c_k)^{\frac{2-q}{q}} c_k = (c^{-1} c_k)^{2/q}. \end{aligned}$$

Hence,

$$\limsup_{k \rightarrow \infty} \|\mathbf{g}_{\alpha_k}^{\delta_k}\|_{(\ell_2)^n}^2 \leq \limsup_{k \rightarrow \infty} (c^{-1} c_k)^{2/q} = c^{-2/q} \Psi^{2/q}(\mathbf{g}^\dagger) < \infty,$$

and the proof is complete. \square

It remains to show strong convergence of the weakly convergent subsequences. This result, however, is so far limited to the case $p = q$.

Theorem 19 *Let $p = q$ and $\mathbf{g}_{\alpha_k}^{\delta_k}$ be a weakly convergent sequence of minimizers of the Tikhonov functional with weak limit \mathbf{g}_* , obtained with the a-priori parameter rule as in Theorem 16. Then the sequence converges also with respect to Ψ , i.e. also in $(\ell_2)^n$.*

Proof. For ease of notation we will assume the a priori guess to be $\mathbf{g}_0 = 0$. Otherwise, $\mathbf{g}_{\alpha_k}^{\delta_k}$ and \mathbf{g}_* have to be replaced in the whole proof by $\mathbf{g}_{\alpha_k}^{\delta_k} - \mathbf{g}_0$ and $\mathbf{g}_* - \mathbf{g}_0$. In order to estimate $\Psi(\mathbf{g}_{\alpha_k}^{\delta_k} - \mathbf{g}_*)$ we need to rewrite the absolute value of a real number. Defining

$$\varphi(x, h) = \begin{cases} -\operatorname{sgn}(x) \cdot h & \text{if } 0 \neq \operatorname{sgn}(x) = \operatorname{sgn}(h) \text{ and } |x| > |h| \\ (\operatorname{sgn}(x) \cdot h - 2|x|) & \text{if } 0 \neq \operatorname{sgn}(x) = \operatorname{sgn}(h) \text{ and } |x| \leq |h| \\ |h| & \text{if } 0 \neq \operatorname{sgn}(x) = -\operatorname{sgn}(h) \\ |h| & \text{if } x = 0 . \end{cases} \quad (5.6)$$

We obtain

$$|x - h| = |x| + \varphi(x, h) . \quad (5.7)$$

Clearly, we have $|x| \geq 0$ and $|x| + \varphi(x, h) \geq 0$. For $a, a + b > 0$ and $1 \leq q \leq 2$ holds $(a + b)^q \leq \kappa_q(a^q + \operatorname{sgn}(b)|b|^q)$ (see e.g. Lemma 2 in [12]) and thus we get

$$(|x| + \varphi(x, h))^q \leq \kappa_q(|x|^q + \operatorname{sgn}(\varphi)|\varphi(x, h)|^q) . \quad (5.8)$$

Setting $x = (\mathbf{g}_{\alpha_k}^{\delta_k})_\lambda^i$, $h = (\mathbf{g}_*)_\lambda^i$ yields

$$\begin{aligned} \Psi(\mathbf{g}_{\alpha_k}^{\delta_k} - (\mathbf{g}_*)) &= \sum_{\lambda \in \Lambda} \omega_\lambda \sum_{i=1}^n |(\mathbf{g}_{\alpha_k}^{\delta_k})_\lambda^i - (\mathbf{g}_*)_\lambda^i|^q \\ &\stackrel{(5.7)}{=} \sum_{\lambda \in \Lambda} \omega_\lambda \sum_{i=1}^n (|(\mathbf{g}_{\alpha_k}^{\delta_k})_\lambda^i| + \varphi((\mathbf{g}_{\alpha_k}^{\delta_k})_\lambda^i, (\mathbf{g}_*)_\lambda^i))^q \\ &\stackrel{(5.8)}{\leq} \kappa_q \left(\sum_{\lambda \in \Lambda} \omega_\lambda \sum_{i=1}^n |(\mathbf{g}_{\alpha_k}^{\delta_k})_\lambda^i|^q + \sum_{\lambda \in \Lambda} \omega_\lambda \sum_{i=1}^n \operatorname{sgn}(\varphi) |\varphi((\mathbf{g}_{\alpha_k}^{\delta_k})_\lambda^i, (\mathbf{g}_*)_\lambda^i)|^q \right) \\ &= \kappa_q \left(\Psi(\mathbf{g}_{\alpha_k}^{\delta_k}) + \sum_{\lambda \in \Lambda} \omega_\lambda \sum_{i=1}^n \operatorname{sgn}(\varphi) |\varphi((\mathbf{g}_{\alpha_k}^{\delta_k})_\lambda^i, (\mathbf{g}_*)_\lambda^i)|^q \right) \\ &\leq \kappa_q \left(\Psi(\mathbf{g}_*) + \sum_{\lambda \in \Lambda} \omega_\lambda \sum_{i=1}^n \operatorname{sgn}(\varphi) |\varphi((\mathbf{g}_{\alpha_k}^{\delta_k})_\lambda^i, (\mathbf{g}_*)_\lambda^i)|^q \right) . \end{aligned}$$

By the definition of $\varphi(x, h)$ in (5.6) it follows,

$$|\varphi(x, h)| = \begin{cases} |-\operatorname{sgn}(x) \cdot h| \leq |h| & \text{if } 0 \neq \operatorname{sgn}(x) = \operatorname{sgn}(h) \text{ and } |x| > |h| \\ |\operatorname{sgn}(x) \cdot h - 2|x|| \leq 3|h| & \text{if } 0 \neq \operatorname{sgn}(x) = \operatorname{sgn}(h) \text{ and } |x| \leq |h| \\ |h| & \text{if } 0 \neq \operatorname{sgn}(x) = -\operatorname{sgn}(h) \\ |h| & \text{if } x = 0 . \end{cases} \quad (5.9)$$

i.e.

$$|\varphi((\mathbf{g}_{\alpha_k}^{\delta_k})_\lambda^i, (\mathbf{g}_*)_\lambda^i)|^q \leq 3^q |(\mathbf{g}_*)_\lambda^i|^q$$

and thus

$$\sum_{\lambda \in \Lambda} \omega_\lambda \sum_{i=1}^n |\varphi((\mathbf{g}_{\alpha_k}^{\delta_k})_\lambda^i, (\mathbf{g}_*)_\lambda^i)| \leq 3^q \sum_{\lambda \in \Lambda} \omega_\lambda \sum_{i=1}^n |(\mathbf{g}_*)_\lambda^i|^q = 3^q \Psi(\mathbf{g}_*) ,$$

i.e. $3^q \sum_{\lambda \in \Lambda} \omega_\lambda \sum_{i=1}^n |(\mathbf{g}_*)_\lambda^i|^q$ dominates $\sum_{\lambda \in \Lambda} \omega_\lambda \sum_{i=1}^n |\varphi((\mathbf{g}_{\alpha_k}^{\delta_k})_\lambda^i, (\mathbf{g}_*)_\lambda^i)|^q$, and we can interchange limit and sum,

$$\lim_{k \rightarrow \infty} \sum_{\lambda \in \Lambda} \omega_\lambda \sum_{i=1}^n \operatorname{sgn}(\varphi) |\varphi((\mathbf{g}_{\alpha_k}^{\delta_k})_\lambda^i, (\mathbf{g}_*)_\lambda^i)|^q = \sum_{\lambda \in \Lambda} \omega_\lambda \sum_{i=1}^n \lim_{k \rightarrow \infty} \operatorname{sgn}(\varphi) |\varphi((\mathbf{g}_{\alpha_k}^{\delta_k})_\lambda^i, (\mathbf{g}_*)_\lambda^i)|^q. \quad (5.10)$$

As $(\mathbf{g}_{\alpha_k}^{\delta_k})^i \xrightarrow{w-\ell_2} (\mathbf{g}_*)^i$, we have in particular, $(\mathbf{g}_{\alpha_k}^{\delta_k})_\lambda^i \rightarrow (\mathbf{g}_*)_\lambda^i$ for $k \rightarrow \infty$. Now assume $(\mathbf{g}_*)_\lambda^i \neq 0$, then there exists k_0 s.t. $(\mathbf{g}_{\alpha_k}^{\delta_k})_\lambda^i \neq 0$ and $\operatorname{sgn}((\mathbf{g}_{\alpha_k}^{\delta_k})_\lambda^i) = \operatorname{sgn}((\mathbf{g}_*)_\lambda^i)$ for all $k \geq k_0$. According to the definition (5.6) of φ , we have for $k \geq k_0$

$$\varphi((\mathbf{g}_{\alpha_k}^{\delta_k})_\lambda^i, (\mathbf{g}_*)_\lambda^i) = \begin{cases} -\operatorname{sgn}((\mathbf{g}_{\alpha_k}^{\delta_k})_\lambda^i) \cdot (\mathbf{g}_*)_\lambda^i = -|(\mathbf{g}_*)_\lambda^i| \\ \operatorname{sgn}((\mathbf{g}_{\alpha_k}^{\delta_k})_\lambda^i) \cdot (\mathbf{g}_*)_\lambda^i - 2|(\mathbf{g}_{\alpha_k}^{\delta_k})_\lambda^i| = |(\mathbf{g}_*)_\lambda^i| - 2|(\mathbf{g}_{\alpha_k}^{\delta_k})_\lambda^i| \end{cases},$$

which gives

$$\lim_{k \rightarrow \infty} \varphi((\mathbf{g}_{\alpha_k}^{\delta_k})_\lambda^i, (\mathbf{g}_*)_\lambda^i) = -|(\mathbf{g}_*)_\lambda^i|,$$

and

$$\lim_{k \rightarrow \infty} \operatorname{sgn}(\varphi) |\varphi((\mathbf{g}_{\alpha_k}^{\delta_k})_\lambda^i, (\mathbf{g}_*)_\lambda^i)|^q = -|(\mathbf{g}_*)_\lambda^i|^q.$$

Note that we do not have to consider the case $(\mathbf{g}_*)_\lambda^i = 0$, as $\varphi((\mathbf{g}_{\alpha_k}^{\delta_k})_\lambda^i, (\mathbf{g}_*)_\lambda^i) = 0$ holds then anyway. Consequently,

$$\begin{aligned} 0 &\leq \lim_{k \rightarrow \infty} \Psi(\mathbf{g}_{\alpha_k}^{\delta_k} - \mathbf{g}_*) \leq \kappa_q \left(\Psi(\mathbf{g}_*) + \lim_{k \rightarrow \infty} \sum_{\lambda \in \Lambda} \omega_\lambda \sum_{i=1}^n \operatorname{sgn}(\varphi) |\varphi((\mathbf{g}_{\alpha_k}^{\delta_k})_\lambda^i, (\mathbf{g}_*)_\lambda^i)|^q \right) \\ &= \kappa_q \left(\Psi(\mathbf{g}_*) + \sum_{\lambda \in \Lambda} \omega_\lambda \sum_{i=1}^n \lim_{k \rightarrow \infty} \operatorname{sgn}(\varphi) |\varphi((\mathbf{g}_{\alpha_k}^{\delta_k})_\lambda^i, (\mathbf{g}_*)_\lambda^i)|^q \right) \\ &= \kappa_q \left(\Psi(\mathbf{g}_*) - \sum_{\lambda \in \Lambda} \omega_\lambda \sum_{i=1}^M |(\mathbf{g}_*)_\lambda^i|^q \right) = 0, \end{aligned}$$

which proves $\mathbf{g}_{\alpha_k}^{\delta_k} \rightarrow (\mathbf{g}_*)$ with respect to Ψ and since

$$\|\mathbf{g}\|_{(\ell_2)^n}^q \leq c^{-1} \Psi(\mathbf{g})$$

also with respect to $(\ell_2)^n$. □

6 Application to Color Image Inpainting

In this section, we apply the algorithm developed in this paper to the restoration of color images from limited color information and gray levels where the colors are missing, see Figure 1. The reconstruction capabilities of the model are shown for piecewise and randomly given color data, i.e. the restoration is obtained from few sparse complete



Figure 1: Top: original color image (left); gray value image \bar{v} on the whole domain Ω (middle); color fragment image \bar{x} on $D \subset \Omega$ (right); bottom: original color image (left); gray value image \bar{v} on the whole domain Ω (middle); randomly sampled points of the color image \bar{x} on $D \subset \Omega$ (right)

samples of the function and from significant incomplete information. The latter is assumed as the result of a nonlinear distortion and with values in a lower dimensional manifold.

In what follows, we describe how this image restoration problem fits into our setting and how the resulting algorithm performs the color reconstruction. The goal is to reconstruct a vector valued function x that belongs to $(L_2(\Omega))^3$, i.e. $x : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$, on the basis of an observed couple of functions (\bar{x}, \bar{v}) . The observed function \bar{x} is assumed to represent the correct information on $D \subset \Omega$, and \bar{v} is the result of a nonlinear distortion $\tilde{T} : \mathbb{R}^3 \rightarrow \mathbb{R}$ on Ω . Given now a frame $\{\phi_\lambda\}_{\lambda \in \Lambda}$ for $L_2(\Omega)$ (which is in this particular case a Daubechies-3 wavelet frame), we associate to $x = (x^1, x^2, x^3)$, corresponding to the three color channels (*red, green, blue*), a triple of vectors of frame coefficients, $\mathbf{g} = (\mathbf{g}^1, \mathbf{g}^2, \mathbf{g}^3)$, such that

$$x = F^* \mathbf{g} = (F^* \mathbf{g}^1, F^* \mathbf{g}^2, F^* \mathbf{g}^3).$$

Introducing $T(\mathbf{g}) := \tilde{T}(F^* \mathbf{g})$, a suitable cost functional for the color inpainting problem can be cast as follows

$$J_{\alpha, q}(\mathbf{g}) = \|(T(\mathbf{g}) - \bar{v})\|_Y^2 + \mu \|(F^* \mathbf{g} - \bar{x})\chi_D\|_{(L_2(\Omega))^3}^2 + 2\alpha \sum_{\lambda \in \Lambda} \omega_\lambda \|\mathbf{g}_\lambda\|_q^p, \quad (6.1)$$

which is quite close to a cost functional suggested in [10]. The first data misfit term penalizes the distance to the gray value information on Ω . The middle term forces the solution to match the given color information on $D \subset \Omega$. The last penalty is a joint sparsity measure. The choice of this regularization measure is encouraged by the fact that most of currently inpainting approaches deal with BV-like penalties. Since in 2 dimensions, ℓ_1 sparse wavelet expansions are known to give ‘near’ BV solutions, see, e.g.,

[3, 6], this kind of sparsity constraint seems to be well suited for our purpose. Forcing, moreover, for common sparsity pattern (e.g. edges in color images), a coupling of the different color channels is advantageous and can be achieved when setting $p = 1$. Up to the middle term, which is a simple quadratic misfit term, problem (6.1) fits with (2.6). Therefore, we hope that the proposed iterative approach (2.8) yields a reasonable approximation for a minimizer of (6.1).

In order to elaborate the iterative algorithm (2.8) that reconstructs the color image on the basis of (\bar{x}, \bar{v}) , we have to characterize T in full detail and to derive the necessary condition of (2.7). The nonlinear distortion operator T can be described as follows:

$$T(\mathbf{g}) = T(F^* \mathbf{g}) = f(\eta_1 F^* \mathbf{g}^1 + \eta_2 F^* \mathbf{g}^2 + \eta_3 F^* \mathbf{g}^3) = f(\eta \cdot F^* \mathbf{g}) ,$$

where all $\eta_i > 0$ and $\sum_i \eta_i = 1$, and $f : \mathbb{R} \rightarrow \mathbb{R}$ is a non-negative function which is assumed smooth, nonlinear, and normally non-convex and non-concave. Typically, f and the vector η are unknown and have therefore to be estimated in advance by fitting a distribution of data from real color fragments. In this case, there is an area $D \subset \Omega$ of the domain Ω where some fragments with colors are placed and complete information is available, and the full domain (including the inpainting region) where the gray value information is known, modeled as the image of T . In our example, we assume that f and η are given (in order to provide the gray value image). In particular, we set

$$f(t) = (1 - \cos(\pi t))/2 .$$

To derive the necessary condition of (2.7), we set $\mathcal{X}_1 = \mathcal{X}_2 = \mathcal{X}_3 = L_2(\Omega)$ (see Section 2.2), and assume F to be the frame map such that $F^* \mathbf{g}^i \in L_2(\Omega)$. For the distortion map and its Fréchet derivative at \mathbf{g} it holds

$$T : (\ell_2)^3 \rightarrow L_2(\Omega) ,$$

and

$$T'(\mathbf{g}) = \tilde{T}'(F^* \mathbf{g}) F^* = f'(\eta \cdot F^* \mathbf{g}) \eta \cdot F^* : (\ell_2)^3 \rightarrow L_2(\Omega) .$$

Therefore,

$$T'(\mathbf{g})^* = \eta F(f'(\eta \cdot F^* \mathbf{g})(\cdot)) : L_2(\Omega) \rightarrow (\ell_2)^3 ,$$

and hence, the λ -componentwise necessary condition for a minimizer of the surrogate functional (2.7) reads with $\mathbf{g}_\lambda = ((\mathbf{g}^1)_\lambda, (\mathbf{g}^2)_\lambda, (\mathbf{g}^3)_\lambda)$ and anchor \mathbf{a} ,

$$\mathbf{g}_\lambda = (I - P_{B_{q'}(C^{-1}\alpha\omega)}) ([(T'(\mathbf{g})^*(\bar{v} - T(\mathbf{a})) - \mu F(F^* \mathbf{g} - \bar{x})\chi_D)/C + \mathbf{a}]_\lambda) , \quad (6.2)$$

where the projection $P_{B_{q'}(C^{-1}\alpha\omega)}$ (acting on the 3 channels) on the dual of the ℓ_q -ball is for $q = 2$ explicitly given by

$$P_{B_2(C^{-1}\alpha\omega)}(\mathbf{g}_\lambda) = \begin{cases} \mathbf{g}_\lambda & \text{if } \|\mathbf{g}_\lambda\|_2 = \sqrt{\sum_{i=1}^2 |(\mathbf{g}^i)_\lambda|^2} \leq C^{-1}\alpha\omega_\lambda \\ \frac{C^{-1}\alpha\omega_\lambda}{\|\mathbf{g}_\lambda\|_2} \mathbf{g}_\lambda & \text{otherwise} \end{cases} .$$

Denoting with $S_{C^{-1}\alpha\omega}^2$ the λ -componentwise projection $I - P_{B_2(C^{-1}\alpha\omega)}$, we may define for facility of inspection the shorthand notation $\mathbf{S}_{C^{-1}\alpha\omega}^2$ by

$$\mathbf{S}_{C^{-1}\alpha\omega}^2(\mathbf{g}) = \{\mathbf{S}_{C^{-1}\alpha\omega}^2(\mathbf{g}_\lambda)\}_{\lambda \in \Lambda} .$$

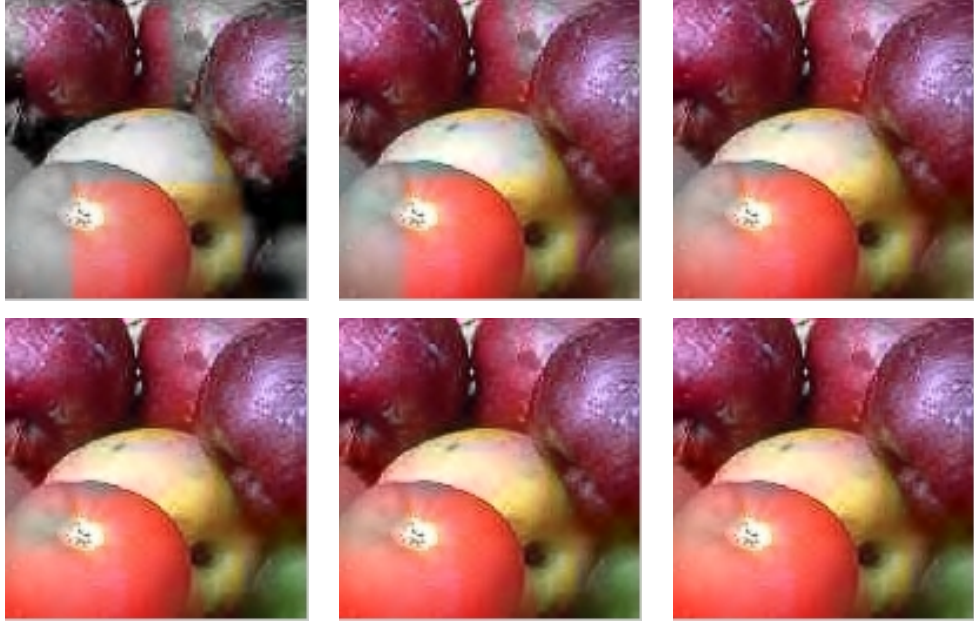


Figure 2: Image inpainting restorations $F^* \mathbf{g}_k = (F^* \mathbf{g}_k^1, F^* \mathbf{g}_k^2, F^* \mathbf{g}_k^3)$ for iteration indices $k = 1, 10, 30, 70, 120, 200$ (from top left to bottom right).

Then, we may write instead of (6.2) shortly

$$\mathbf{g} = \mathbf{S}_{C^{-1}\alpha\omega}^2 ((T'(\mathbf{g})^*(\bar{v} - T(\mathbf{a})) - \mu F(F^* \mathbf{g} - \bar{x})\chi_D)/C + \mathbf{a}) .$$

As $\mathbf{S}_{C^{-1}\alpha\omega}^2$ is non-expansive, it can be easily deduced with the help of Lemma 9 that the associated fixed point map,

$$\Phi(\mathbf{g}, \mathbf{a}) = \mathbf{S}_{C^{-1}\alpha\omega}^2 ((T'(\mathbf{g})^*(\bar{v} - T(\mathbf{a})) - \mu F(F^* \mathbf{g} - \bar{x})\chi_D)/C + \mathbf{a}) ,$$

is contractive as long as

$$\frac{L\sqrt{J_\alpha(\mathbf{g}_0)} + \mu B}{C} < 1 ,$$

where L is the Lipschitz constant of T' and B is the upper frame bound, see (2.1). Since $L\sqrt{J_\alpha(\mathbf{g}_0)}/C \leq 1/2$ it follows that the choice of μ is restricted to $\mu B/C < 1/2$; which can be easily achieved by choosing $\mu > 0$ adequately.

We summarize now our findings and fix the iterative algorithm (2.8) for the purpose of color image inpainting:

1. Pick $\mathbf{g}_0 \in (\ell_2)^3$ and some proper constants $C > 0$ and $\mu > 0$
2. Derive a sequence $\{\mathbf{g}_k\}_{k=0,1,\dots}$ by the iteration:

$$\mathbf{g}_{k+1} = \mathbf{S}_{C^{-1}\alpha\omega}^2 ((T'(\mathbf{g}_{k+1})^*(\bar{v} - T(\mathbf{g}_k)) - \mu F(F^* \mathbf{g}_{k+1} - \bar{x})\chi_D)/C + \mathbf{g}_k) \quad (6.3)$$

where each iterate can be evaluated via a simple fixed point iteration.

The color restoration/inpainting results that are obtained with the proposed algorithm (6.3) are visualized in Figures 2 and 3.



Figure 3: Image inpainting restorations $F^* \mathbf{g}_k = (F^* \mathbf{g}_k^1, F^* \mathbf{g}_k^2, F^* \mathbf{g}_k^3)$ for iteration indices $k = 1, 11, 31, 61, 101, 161$ (from top left to bottom right).

Acknowledgement

Both authors thank M. Fornasier for directing our attention to joint sparsity measures and the color image inpainting problem and for fruitful discussions on the topic. G.T. gratefully acknowledge partial support by Deutsche Forschungsgemeinschaft Grants TE 354/1-2, TE 354/3-1, TE 354/4-1, TE 354/5-1.

References

- [1] S. Anthoine. Different wavelet-based approaches for the separation of noisy and blurred mixtures of components. application to astrophysical data. *PhD Thesis*, 2005.
- [2] K. Bredies, D.A. Lorenz, and P. Maass. A generalized conditional gradient method and its connection to an iterative shrinkage method. *to appear in: Computational Optimization and Application*, 2005.
- [3] A. Cohen, R. DeVore, P. Petrushev, and H. Xu. Nonlinear Approximation and the Space $BV(\mathbb{R}^2)$. *American Journal of Mathematics*, (121):587–628, 1999.
- [4] I. Daubechies, M. DeFrise, and C. DeMol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Comm. Pure Appl. Math*, 57:1413–1541, 2004.
- [5] I. Daubechies and G. Teschke. Wavelet-based image decomposition by variational functionals. *Proc. SPIE Vol. 5266, p. 94-105, Wavelet Applications in Industrial Processing; Frederic Truchetet; Ed.*, Feb. 2004.

- [6] I. Daubechies and G. Teschke. Variational image restoration by means of wavelets: simultaneous decomposition, deblurring and denoising. *Applied and Computational Harmonic Analysis*, 19(1):1–16, 2005.
- [7] I. Daubechies, G. Teschke, and L. Vese. Iteratively solving linear inverse problems with general convex constraints. *Inverse Problems and Imaging*, 1(1):29–46, 2007.
- [8] M. Defrise and C. DeMol. Inverse imaging with mixed penalties. *Conference proceedings*, 2004.
- [9] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems*. Kluwer, Dordrecht, 1996.
- [10] M. Fornasier and R. March. Restoration of color images by vector valued by functions and variational calculus. *Preprint*, 2006.
- [11] M. Fornasier and H. Rauhut. Recovery algorithms for vector valued data with joint sparsity constraint. *Preprint*, 2006.
- [12] R. Ramlau. Regularization Properties of Tikhonov Regularization with Sparsity Constraints. *preprint, under review*, 2007.
- [13] R. Ramlau and G. Teschke. Tikhonov Replacement Functionals for Iteratively Solving Nonlinear Operator Equations. *Inverse Problems*, 21:1571–1592, 2005.
- [14] R. Ramlau and G. Teschke. A Projection Iteration for Nonlinear Operator Equations with Sparsity Constraints. *Numerische Mathematik*, 104:177–203, 2006.
- [15] G. Teschke. Multi-Frames in Thresholding Iterations for Nonlinear Operator Equations with Mixed Sparsity Constraints. *DFG-SPP-1114 preprint 131*, 2005.
- [16] G. Teschke. Operator Equations, Mixed Constraints, Frame Based Iterative Concepts and Application. *Habilitation thesis, Bremen*, 2005.
- [17] G. Teschke. Multi-Frame Representations in Linear Inverse Problems with Mixed Multi-Constraints. *Applied and Computational Harmonic Analysis*, 22:43 – 60, 2007.