Sparse $\ell^q$-regularization of inverse problems with deep learning

Markus Haltmeier
Department of Mathematics, University of Innsbruck
Technikerstrasse 13, 6020 Innsbruck, Austria
markus.haltmeier@uibk.ac.at

Linh Nguyen
Department of Mathematics, University of Idaho
Moscow, ID 83844,
lnnguyen@uidaho.edu

Daniel Obmann
Department of Mathematics, University of Innsbruck
Technikerstrasse 13, 6020 Innsbruck, Austria
daniel.obmann@uibk.ac.at

Johannes Schwab
Department of Mathematics, University of Innsbruck
Technikerstrasse 13, 6020 Innsbruck, Austria
Johannes.Schwab@uibk.ac.at

Abstract
We propose a sparse reconstruction framework for solving inverse problems. Opposed to existing sparse reconstruction techniques that are based on linear sparsifying transforms, we train an encoder-decoder network $D \circ E$ with $E$ acting as a nonlinear sparsifying transform. We minimize a Tikhonov functional which used a learned regularization term formed by the $\ell^q$-norm of the encoder coefficients and a penalty for the distance to the data manifold. For this augmented sparse $\ell^q$-approach, we present a full convergence analysis, derive convergence rates and describe a training strategy. As a main ingredient for the analysis we establish the coercivity of the augmented regularization term.
1 Introduction

Various applications in medical imaging, remote sensing and elsewhere require solving inverse problems of the form

\[ y = Ax + z, \quad (1.1) \]

where \( A : X \to Y \) is a linear operator between Hilbert spaces \( X, Y \), and \( z \) is the data distortion. Inverse problems are well analyzed and several established approaches for its solution exist, including filter-based methods or variational regularization techniques \(^{[1, 2]}\). In the very recent years, neural networks (NN) and deep learning appeared as new paradigms for solving inverse problems, and demonstrate impressive performance. Several approaches have been developed, including two-step networks \(^{[3, 4, 5]}\), variational networks \(^{[6]}\), iterative networks \(^{[7, 8]}\) and regularizing networks \(^{[9]}\).

Standard deep learning approaches may lack data consistency for unknowns very different from the training images. To address this issue, in \(^{[10]}\) a deep learning approach has been introduced where minimizers

\[ x_\alpha \in \arg \min_x \| A(x) - y \|_Y^2 + \alpha \phi(E(x)) \quad (1.2) \]

are investigated. Here \( E : X \to \Xi \) is a trained NN, \( \Xi \) a Hilbert space, \( \phi : \Xi \to [0, \infty) \) a functional, and \( \alpha > 0 \) the regularization parameter. The resulting reconstruction approach has been named NETT (for network Tikhonov regularization), as it is a generalized form of Tikhonov regularization using a NN as trained regularizer.

In \(^{[10]}\) it is shown that under suitable assumption, NETT yields a convergent regularization method. Moreover, in that paper a training strategy has been proposed, where \( E \) is trained to favor artifact-free reconstructions selected from a set of training images from a certain data manifold \( \mathcal{M} \); see \(^{[11]}\) for a simplified training strategy.

Coercive variant of NETT

One of the main assumptions in the analysis of \(^{[10]}\) is the coercivity of the regularizer \( \phi \circ E \). For the general form used in \(^{[3, 11]}\), this requires special care in the design and training of the network. In order to overcome this limitation, in this paper we propose a modified form of the regularizer for which we are able to rigorously proof its coercivity. More precisely, we consider

\[ x_\alpha \in \arg \min_x \| A(x) - y \|_Y^2 + \alpha \left( \phi(E(x)) + \frac{\beta}{2} \| x - (D \circ E)(x) \|_2^2 \right). \quad (1.3) \]

Here, \( D \circ E : X \to X \) is an encoder-decoder network trained so such that for any \( x \in \mathcal{M} \) we have \( x \simeq DE(x) \) and that \( \phi(E(x)) \) is small. The term \( \phi(E(x)) \)
implements learned prior knowledge. The additional term \( \| x - (D \circ E)(x) \|_2^2 \) forces \( x \) to be close to data manifold \( \mathcal{M} \) and, as we shall prove, also guarantees coercivity of the regularization functional.

In particular, in this paper we investigate the case where \( \Xi = \ell^2(\Lambda) \) for some index set \( \Lambda \) and \( \phi \) is a weighted \( \ell^q \)-norm used as sparsity prior. To construct an appropriate network, we train a (modified) tight frame U-net \([12]\) of the form \( D \circ E \) using the \( \ell^q \)-norm of \( E(x) \) during training, and take the encoder part as analysis network.

Outline

This paper is organized as follows. In Section 2, we present a convergence analysis for the augmented \( \ell^q \)-NETT (see (2.1)). In particular, as main auxiliary result, we establish the coercivity of the regularization term. In Section 3, we derive convergence rates which provide quantitative estimates for the reconstruction accuracy. In Section 4, we present a suggested network structure using a modified tight frame U-net and a corresponding training strategy. The paper concludes with a short summary and outlook given in Section 5.

2 Well-posedness and convergence

2.1 Augmented \( \ell^q \)-NETT

To solve the inverse problem (1.1) we propose and analyze the augmented \( \ell^q \)-NETT, which considers minimizers of

\[
\mathcal{T}_{\alpha,Y}(x) := \| A x - y \|_Y^2 + \alpha \left( \sum_{\lambda \in \Lambda} w_{\lambda} |(E(x))_{\lambda}|^q + \frac{\beta}{2} \| x - (D \circ E)(x) \|_2^2 \right). \tag{2.1}
\]

Here \( \alpha \) is the regularization parameter, \( E : X \to \ell^2(\Lambda) \) is called encoder network, \( D : \ell^2(\Lambda) \to X \) is called decoder network, \( \Lambda \) a countable index set, \( w_{\lambda} \) are positive weights, \( \beta > 0 \) is a tuning parameter, and \( q \in [1,2] \) describes the used norm. The case \( q < 2 \) yields a sparsity promoting regularization term \( \sum_{\lambda \in \Lambda} w_{\lambda}|(E(x))_{\lambda}|^q \), frequently studied when \( E \) is a basis of frame \([13, 14, 15, 16, 17, 18]\). In the present paper, we allow \( D \) and \( E \) to be non-linear mappings.

For our convergence analysis, we use the following assumptions, that we assume to be satisfied throughout this section.

**Condition 2.1** (Augmented \( \ell^q \)-NETT).

\((A1)\) \( A : X \to Y \) is bounded linear;

\((A2)\) \( E : X \to \ell^2(\Lambda) \) is weakly sequentially continuous;
(A3) $D: \ell^2(\Lambda) \to \mathbb{X}$ is weakly sequentially continuous;

$\text{(A4)} \ w_{\text{min}} \triangleq \inf\{w_\lambda \mid \lambda \in \Lambda\} > 0.$

The first term in the considered regularizer

$$R_{q,w}(x) := \sum_{\lambda \in \Lambda} w_\lambda |(E(x))_\lambda|^q + \frac{\beta}{2} \|x - (D \circ E)(x)\|^2$$

(2.2)

was proposed in [10] to impose a sparsity condition on the signal $x$. In this paper, we add the extra term $\frac{\beta}{2} \|x - (D \circ E)(x)\|^2$ forcing the minimizers of $T_{\alpha,y}$ being close to the solution manifold $M$. This term also allows to prove the coercivity of $R_{q,w}$ (see the argument in the proof of Theorem 2.2), which is essential to our analysis.

2.2 Well-posedness

**Theorem 2.2 (Existence).** For all $y \in \mathbb{Y}$ and all $\alpha > 0$, the augmented $\ell^q$-NETT functional (2.1) has at least one minimizer.

**Proof.** Let us first prove that $R_{q,w}$ is coercive. Indeed, let us assume that there exists a sequence $(x_k)_k$ such that $x_k \to \infty$ and $(R_{q,w}(x_k))_k$ is bounded. Then, $(E(x_k))_k$ is bounded in $\ell^q(\Lambda)$. Since $2/q \geq 1$, we obtain

$$\|E(x_k)\|_q^2 = \sum_{\Lambda} |(E(x_k))_\lambda|^q \leq \left( \sum_{\Lambda} |(E(x_k))_\lambda|^q \right)^{2/q} = \|E(x_k)\|_q^2.$$ 

Therefore, $(E(x_k))_k$ is also bounded in $\ell^2(\Lambda)$, too. Now, since $D$ is weakly sequentially continuous, this implies that also $((D \circ E)(x_k))_k$ is a bounded sequence. From the estimate

$$\|x_k\|^2 \leq 2 \left( \|x_k - (D \circ E)(x_k)\|^2 + \|D \circ E)(x_k)\|^2 \right) \leq \frac{4}{\beta} R_{q,w}(x_k) + 2 \|E \circ D)(x_k)\|^2$$

it follows that $(x_k)_k$ is a bounded sequence. This is a contradiction and finishes the proof that $R_{q,w}$ is coercive.

Because the network $D \circ E$ is weakly sequentially continuous, the functional $\|\cdot - (D \circ E)(\cdot)\|^2$ is weakly lower semi-continuous. Therefore, $R_{q,w}$ is weakly lower semi-continuous, too. Since $T_{\alpha,y}$ is bounded from below by 0, it has an infimum $c \geq 0$. Let $(x_k)_k$ be a sequence such that $T_{\alpha,y}(x_k) \to c$. Since $R_{q,w}$ is coercive, the sequence $(x_k)_k$ is bounded, and hence, has an accumulation point in the weak topology, denoted by $x_*$. Because $T_{\alpha,y}$ is sequentially lower semi-continuous, it follows that $T_{\alpha,y}(x_*) \leq \liminf T_{\alpha,y}(x_k) = c$. Therefore, $x_*$ is a minimizer of $T_{\alpha,y}$. \[\square\]
Theorem 2.3 (Stability). Let \( y \in \mathbb{Y}, \alpha > 0, (y_k)_k \in \mathbb{Y}^N \) with \( y_k \to y \), and \( x_k \in \text{arg min} \mathcal{T}_{\alpha, y_k} \). Then weak accumulation points of \((x_k)_k \) exist and are minimizers of \( \mathcal{T}_{\alpha, y} \). For any weak accumulation point \( x \), and subsequence \((x_{k(\ell)})_\ell \to x \), of \((x_k)_k \), it holds that \( \lim_{\ell \to \infty} \mathcal{R}_{q, w}(x_{k(\ell)}) = \mathcal{R}_{q, w}(x) \).

Proof. The proof follows the lines of [2, Theorem 3.23]. We note that the convexity of the regularizer assumed in [2] is not needed in that proof. For the sake of completeness, we sketch here a proof for the non-convex regularizer \( \mathcal{R}_{q, w} \).

Fix \( x \in X \). Then, for all \( k \in \mathbb{N} \), we have \( \mathcal{T}_{\alpha, y_k}(x_k) \leq \mathcal{T}_{\alpha, y_k}(x) \), which implies
\[
\alpha \mathcal{R}_{q, w}(x_k) \leq \mathcal{T}_{\alpha, y_k}(x) \leq 2(\|Ax\|^2 + \|y_k\|^2) + \alpha \mathcal{R}_{q, w}(x).
\]
Consequently, \((\mathcal{R}_{q, w}(x_k))_k \) is bounded and therefore has a weakly convergent subsequence \( x_{k(\ell)} \to x \). Let us prove that each such accumulation point satisfies \( x_\ell \in \text{arg min} \mathcal{T}_{\alpha, y} \). Indeed, given any \( x \in X \), we have \( \mathcal{T}_{\alpha, y_k}(x_k) \leq \mathcal{T}_{\alpha, y_k}(x) \) which implies \( \lim inf_k \mathcal{T}_{\alpha, y_k}(x_k) \leq \lim_k \mathcal{T}_{\alpha, y_k}(x) \) and therefore \( \mathcal{T}_{\alpha, y}(x_\ell) \leq \lim \mathcal{T}_{\alpha, y}(x) \).

Since this holds for all \( x \in X \), we obtain \( x_\ell \in \text{arg min} \mathcal{T}_{\alpha, y} \). It now remains to prove \( \lim \mathcal{R}_{q, w}(x_k) = \mathcal{R}_{q, w}(x_\ell) = \mathcal{R}_{q, w}(x) \). For that purpose, write \( \alpha \mathcal{R}_{q, w}(x_{k(\ell)}) = \mathcal{T}_{\alpha, y_{k(\ell)}}(x) - \|Ax_{k(\ell)} - y_{k(\ell)}\|^2 \). Then \( \lim sup \alpha \mathcal{R}_{q, w}(x_{k(\ell)}) \leq \lim sup \alpha \mathcal{T}_{\alpha, y_{k(\ell)}}(x) - \lim inf \|Ax_{k(\ell)} - y_{k(\ell)}\|^2 \), which implies
\[
\lim sup_\ell \alpha \mathcal{R}_{q, w}(x_{k(\ell)}) \leq \mathcal{T}_{\alpha, y}(x) - \|Ax - y\|^2 = \alpha \mathcal{R}_{q, w}(x).
\]
Together with the weak sequential lower-continuity of the regularizer \( \mathcal{R}_{q, w} \), this yields \( \lim_\ell \mathcal{R}_{q, w}(x_{k(\ell)}) = \mathcal{R}_{q, w}(x) \) and concludes the proof.

2.3 Convergence

We call \( x_* \) an \( \mathcal{R}_{q, w} \)-minimizing solution of the equation \( Ax = y \) if
\[
x_* \in \text{arg min} \{ \mathcal{R}_{q, w}(x) \mid x \in X \wedge Ax = y \}.
\]
As in the convex case [2], one shows that an \( \mathcal{R}_{q, w} \)-minimizing solution exists whenever \( Ax = y \) is solvable.

Theorem 2.4 (Weak Convergence). Let \( x \in X \), set \( y := A(x) \), let \( (y_k)_k \in \mathbb{Y}^N \) satisfy \( \|y_k - y\| \leq \delta_k \) for some sequence \( (\delta_k)_k \in (0, \infty)^N \) with \( \delta_k \to 0 \), suppose \( x_k \in \text{arg min} \mathcal{T}_{\alpha(\delta_k), y_k}(x) \), and let the parameter choice \( \alpha: (0, \infty) \rightarrow (0, \infty) \) satisfy
\[
\lim_{\delta \to 0} \alpha(\delta) = \lim_{\delta \to 0} \frac{\delta^2}{\alpha(\delta)} = 0.
\]

Then the following hold:
(a) \( (x_k)_k \in \mathbb{N} \) has at least one weak accumulation point \( x_* \);
(b) Every weak accumulation point of \((x_k)_{k \in \mathbb{N}}\) is an \(R_{q,w}\)-minimizing solution of \(Ax = y\);

(c) Every weakly convergent subsequence \((x_{k(n)})_{n \in \mathbb{N}}\) satisfies \(R_{q,w}(x_{k(n)}) \to R_{q,w}(x)\);

(d) If the \(R_{q,w}\)-minimizing solution of \(Ax = y\) is unique, then \(x_k \rightharpoonup x^*\).

Proof. This follows along the lines of [2, Theorem 3.26].

Next we derive the strong convergence. For that purpose, let us recall the notions of absolute Bregman distance and total nonlinearity, defined in [10].

Definition 2.5 (Absolute Bregman distance). Let \(F : \mathbb{D} \subset X \to \mathbb{R}\) be Gâteaux differentiable at \(x \in \mathbb{D}\). The absolute Bregman distance \(\Delta_F(\cdot, x) : \mathbb{D} \to [0, \infty]\) at \(x\) with respect to \(F\) is defined by

\[
\Delta_F(\tilde{x}, x) = |F(\tilde{x}) - F(x) - F'(x)(\tilde{x} - x)| \quad \text{for } \tilde{x} \in X.
\]

Here \(F'(x)\) denotes the Gâteaux derivative of \(F\) at \(x\).

Definition 2.6 (Total nonlinearity). Let \(F : \mathbb{D} \subset X \to \mathbb{R}\) be Gâteaux differentiable at \(x \in \mathbb{D}\). We define the modulus of total nonlinearity of \(F\) at \(x\) as the function \(\nu_F(x, \cdot) : [0, \infty) \to [0, \infty)\) given by

\[
\nu_F(x, t) = \inf \{\Delta_F(\tilde{x}, x) : \tilde{x} \in D \wedge \|\tilde{x} - x\| = t\}.
\]

We call \(F\) totally nonlinear at \(x\), if \(\nu_F(x, t) > 0\) for all \(t \in (0, \infty)\).

The following convergence result in the norm topology holds.

Theorem 2.7 (Strong Convergence). Assume that \(Ax = y\) has a solution, let \(R_{q,w}\) be totally nonlinear at \(x\), and let \((x_k)_{k \in \mathbb{N}}, (y_k)_{k \in \mathbb{N}}, (\alpha_k)_{k \in \mathbb{N}}, (\delta_k)_{k \in \mathbb{N}}\) be as in Theorem [2.2]. Then there is a subsequence \((x_{k(l)})_{l \in \mathbb{N}}\) of \((x_k)_{k \in \mathbb{N}}\) and an \(R_{q,w}\)-minimizing solution \(x^*\) of \(Ax = y\) such that \(\lim_{l \to \infty} \|x_{k(l)} - x^*\| = 0\). Moreover, if the \(R_{q,w}\)-minimizing solution of \(Ax = y\) is unique, then \(x \rightharpoonup x^*\), in the norm topology.

Proof. Follows from [10, Theorem 2.8].

2.4 Example: Sparse analysis regularization with a dictionary

A simple application of the above results is the case where \(E : X \to \ell^2(\Lambda)\) is a bounded linear operator with closed range. We can write \(E(x)(\lambda) = \langle e_\lambda, x \rangle\) for so-called atoms \(e_\lambda \in X^*\) and interpret \(E\) as (analysis) dictionary. Moreover, we take the decoder network \(D = E^*\) as the pseudoinverse of \(E\).
We have $D \circ E = P_{\ker(E)}$ and the regularizer takes the form
\[ R_{q,w}(x) := \sum_{\lambda \in \Lambda} w_\lambda |\langle e_\lambda, x \rangle|^q + \frac{\beta}{2} \|P_{\ker(E)}(x)\|^2 \] (2.4)

Clearly the conditions (A2), (A3) are satisfied, which implies that existence, stability and weak convergence for sparse analysis dictionary regularization with (2.4) hold. Following [2, Theorem 3.49] one also derives the strong convergence.

Note that if $E$ is a frame of $X$, then $\ker(E) = \{0\}$ in which case (2.4) yields the standard sparse regularizer $R_{q,w}(x) = \sum_{\lambda \in \Lambda} w_\lambda |\langle e_\lambda, x \rangle|^q$. However, for a general trained dictionary we will typically have $\ker(E) \neq \{0\}$. This is even the case for overcomplete dictionaries, because the dictionary is only trained on elements in a small subset of $X$ which are supposed to satisfy a sparse analysis prior. In this case, the additional term $\frac{\beta}{2} \|P_{\ker(E)}(x)\|^2$ in (2.4) ensures coercivity of the regularizer, which is essential for the convergence of Tikhonov regularization.

3 Convergence rates

Let us now prove a convergence rate in the absolute Bregman distance. For that purpose, we consider general Tikhonov regularization
\[ T_{\alpha,\gamma}(x) := \|A(x) - y\|^2 + \alpha R(x) \rightarrow \min_x . \] (3.1)

Here $\mathcal{R}: \mathbb{X} \rightarrow [0, \infty]$ is a general, possibly non-convex, regularizer, and $A: \mathbb{X} \rightarrow \mathbb{Y}$ the linear forward operator.

The convergence rates will be derived under the following assumptions:

(B1) $A$ is a bounded linear with finite-dimensional range.

(B2) $\mathcal{R}$ is coercive and weakly sequentially lower semi-continuous;

(B3) $\mathcal{R}$ is Lipschitz,

(B4) $\mathcal{R}$ is Gâteaux differentiable.

Remark 3.1. Note that the regularizer $\mathcal{R} = R_{q,w}$ of the augmented $\ell^q$-NETT (1.3) satisfies [B2][B4] as long $q > 1$ and the activation functions of the encoder-decoder network $D \circ E$ are differentiable (such as the sigmoid function or smooth versions of ReLU). Condition [B3] can be relaxed to a local Lipschitz property.

The main restriction in the above list of assumptions is that $A$ has finite-dimensional range. However, this assumption holds true in practical applications such as sparse data tomography. Unlike [10], we do not assume that $\mathcal{R}'(x) \in \text{Range}(A^\ast)$, which is quite difficult to validate in practice. Modified provable conditions will be studied in future work.
We start our analysis with the following result.

**Proposition 3.2.** Let \([B1][B4]\) be satisfied and assume that \(x_*\) is an \(\mathcal{R}\)-minimizing solution of \(Ax = y\). Then there exists a constant \(C > 0\) such that

\[
\forall x \in \mathbb{X}: \quad \Delta_\mathcal{R}(x, x_*) \leq \mathcal{R}(x) - \mathcal{R}(x_*) + C\|A(x) - A(x_*)\|.
\]

**Proof.** Let us first prove that for some constant \(\gamma \in (0, \infty)\) it holds

\[
\forall x \in \mathbb{X}: \quad \mathcal{R}(x_*) - \mathcal{R}(x) \leq \gamma \|A(x_*) - A(x)\|. \quad (3.2)
\]

Indeed, let \(P\) be the orthogonal projection onto \(\ker(A)\) and define \(x_0 = (x_* - Px_*) + Px\). Then, \(A(x_0) = A(x_*)\) and \(x - x_0 \in \ker(A)\). Since the restricted operator \(A|_{\ker(A)} : \ker(A) \to \mathbb{Y}\) is injective and has finite-dimensional range, it is bounded from below by a constant \(\gamma_0\). Therefore,

\[
\|A(x_*) - A(x)\| = \|A(x_0) - A(x)\| = \|A(x_0 - x)\| \geq \gamma_0\|x_0 - x\|. \quad (3.3)
\]

On the other hand, since \(x_*\) is the \(\mathcal{R}\)-minimizing solution of \(Ax = y\) and \(\mathcal{R}\) is Lipschitz, we have \(\mathcal{R}(x_*) - \mathcal{R}(x) \leq \mathcal{R}(x_0) - \mathcal{R}(x) \leq L\|x_0 - x\|.\) Together with (3.3) we obtain (3.2).

Next we prove that there is a constant \(\gamma_1\) such that

\[
\langle \mathcal{R}'(x_*), x_* - x \rangle \leq \gamma_1 \|A(x_*) - A(x)\|. \quad (3.4)
\]

Since \(x_*\) is an \(\mathcal{R}\)-minimizing solution of \(Ax = y\) and \(\mathcal{R}\) is Gâteaux differentiable, we obtain \(\langle \mathcal{R}'(x_*), x_* - x \rangle \leq 0\) for \(x_* - x \in \ker(A)\). On the other hand, if \(x_* - x \in \ker(A)\), we have \(|\langle \mathcal{R}'(x_*), x_* - x \rangle| \leq \|\mathcal{R}'(x_*)\||x_* - x|\) and \(\|A(x_*) - A(x)\| \geq \gamma_0\|x_* - x\|\). This finishes the proof of (3.4).

The proof now follows that in [10, Proposition 3.3]. Indeed, We note that

- \(\mathcal{R}(x_*) \leq \mathcal{R}(x) \Rightarrow |\mathcal{R}(x_*) - \mathcal{R}(x)| = \mathcal{R}(x) - \mathcal{R}(x_*)\)
- \(\mathcal{R}(x_*) \geq \mathcal{R}(x) \Rightarrow |\mathcal{R}(x_*) - \mathcal{R}(x)| = \mathcal{R}(x) - \mathcal{R}(x_*) + 2(\mathcal{R}(x_*) - \mathcal{R}(x))\).

Therefore, using (3.2) and (3.4), we obtain

\[
\Delta_\mathcal{R}(x, \bar{x}) \leq |\mathcal{R}(x_*) - \mathcal{R}(x)| + |\langle \mathcal{R}'(x_*), x - x_* \rangle| \leq \mathcal{R}(x) - \mathcal{R}(x_*) + (2\gamma + \gamma')\|Ax - Ax_*\|,
\]

which concludes our proof with \(C := 2\gamma_0 + \gamma_1\). \(\square\)

The following results is our main convergence rates result. It is similar to Proposition [10, Theorem 3.1], but uses different assumptions.

**Theorem 3.3** (Convergence rates results). Let \([B1][B4]\) be satisfied and suppose \(\alpha \sim \delta\). Then \(\Delta_\mathcal{R}(x_{\alpha, \delta}, x_*) = \mathcal{O}(\delta)\) as \(\delta \to 0\).
Proof. From Proposition 3.2 we obtain

\[
\alpha \Delta_\mathcal{R}(x_{\alpha, \delta}, x_\star) \leq \alpha \mathcal{R}(x_{\alpha, \delta}) - \alpha \mathcal{R}(x_\star) + C \alpha \|A(x_{\alpha, \delta}) - A(x_\star)\|
\]

\[
= \mathcal{T}_{\alpha, \delta}(x_{\alpha, \delta}) - \|A(x_{\alpha, \delta}) - y_\delta\|^2 - \left(\mathcal{T}_{\alpha, \delta}(x_\star) - \|A(x_\star) - y_\delta\|^2\right)
\]

\[
+ C \alpha \|A(x_{\alpha, \delta}) - A(x_\star)\|
\]

\[
\leq \delta^2 + C \alpha \delta - \|A(x_{\alpha, \delta}) - y_\delta\|^2 + C \alpha \|A(x_{\alpha, \delta})\|.
\]

Cauchy’s inequality gives \(\alpha \Delta_\mathcal{R}(x_{\alpha, \delta}, x_\star) \leq \delta^2 + C \alpha \delta + \frac{C^2 \delta^2}{4}\). For \(\alpha \sim \delta\), we easily conclude \(\Delta_\mathcal{R}(x_{\alpha, \delta}, x_\star) = \mathcal{O}(\delta)\). \(\square\)

4 Network design and training

For the encoder-decoder type network required for the augmented \(\ell^1\) regularizer (2.2) we propose a modified tight frame U-net \(\mathcal{N} = D \circ E\) together with a sparse training strategy.

The tight frame-U-net has been introduced in [12] and is less smoothing than the classical U-net [19] in image reconstruction. The tight frame U-net of [12] uses a residual (or by-pass) connection, that is not well suited for our purpose. We therefore work with a modified tight frame U-net that has been used in [20] for sparse synthesis regularization with neural networks.

Figure 4.1: Tight frame U-net architecture. It starts by a standard multiple-input-multiple-output (MIMO) convolution layer. Then each channel is filtered and subsampled by factor 2 using four different wavelet filters (HH, HL, LH, LL). The LL part (low frequency), is recursively used as input for the next layer. After the downsampling to the coarsest resolution, we upsample by applying the transposed wavelet filters. Next we concatenate the layers and use a MIMO convolution layer to obtain the final output.
4.1 Modified tight-frame U-net

For simplicity we assume that $X_0 = \mathbb{R}^{n_0 \times n_0 \times c_0}$ is already a finite dimensional space and contains 2D images of size $n_0 \times n_0$ and $c_0$ channels.

The architecture of the modified tight frame U-net $\mathcal{N} = \mathcal{N}_0 = D \circ E$ is shown in Figure 4.1. It uses a hierarchical multi-scale representation defined by

$$\mathcal{N}_\ell = G_\ell \circ \left( \begin{bmatrix} H_h \circ H_h^T \\ H_d \circ H_d^T \\ H_v \circ H_v^T \\ L \circ \mathcal{N}_{\ell+1} \circ L^T \end{bmatrix} \right) \circ \mathcal{F}_\ell$$

for $\ell \in \{0, \ldots, L - 1\}$, (4.1)

with $\mathcal{N}_L = id$. Here

- $L \leq \log(n_0)$ is the number of used scales;
- $\mathcal{F}_\ell : \mathbb{R}^{n_\ell \times n_\ell \times c_\ell} \to \mathbb{R}^{n_\ell \times n_\ell \times d_\ell}$ and $G_\ell : \mathbb{R}^{n_\ell \times n_\ell \times 4d_\ell} \to \mathbb{R}^{n_\ell \times n_\ell \times c_\ell}$ are convolutional layers followed by a nonlinearity;
- $H_h, H_v, H_d$ are horizontal, vertical and diagonal high-pass filters and $L$ is a low-pass filter such that the tight frame property is satisfies,

$$H_h H_h^T + H_v H_v^T + H_d H_d^T + LL^T = id.$$ (4.2)

Following [12] we define the filters by applying the tensor products $HH$, $HL$, $LH$ and $LL$ of the Haar wavelet low-pass $L = 2^{-1/2} [1, 1]^T$ and high-pass $H = 2^{-1/2} [1, -1]^T$ filters separately in each channel.

We write the tight frame U-net defined by (4.1) in the form $\mathcal{N} = D \circ E$ where $E$ is the encoder and $D$ the decoder part. The encoder part

$$E(x) = (E^{(\ell)}(x))_{\ell=0,\ldots,L}$$

maps the image $x$ to the high frequency parts $H_h, H_v, H_d$ at the $\ell$th scale, denoted by $E^{(\ell)}(x)$ for $\ell = 0, \ldots, L - 1$, and to the low frequency part $L$ at the coarsest scale, denoted by $E^{(L)}(x)$. The decoder $D$ then synthesizes the image $x$ from $E(x)$ recursively via (4.2).

4.2 Sparse network training

To enforce sparsity in the encoded domain we will use a combination of mean-squared-error and an $\ell^1$-penalty of the filter coefficients as loss-function for training. The idea is to thereby enhance the sparsity in the high-pass filtered images.

Given a set of training images $\mathcal{M} = \{x_1, \ldots, x_n\}$ we aim for an encoder-decoder network reproducing $x_i \in \mathcal{M}$. For that purpose, we take the encoder-decoder
pair \((E, D)\) as the minimizer of the loss function (the empirical risk)

\[
R_N(E_\theta, D_\eta) = \frac{\beta}{2N} \sum_{i=1}^{N} \| (D_\eta \circ E_\theta)(x_i) - x_i \|^2_2 + \frac{1}{N} \sum_{i=1}^{N} \sum_{\ell=0}^{L} w_\ell \| E_\theta^{(\ell)}(x_i) \|^2_q + \gamma_1 \| \theta \|^2_2 + \gamma_2 \| \eta \|^2_2,
\]

Here \(\theta\) and \(\eta\) are the adjustably parameters in the tight frame U-net architecture (specifically, in the convolution layers \(F_\ell\) and \(G_\ell\)). The first term of the loss-function is supposed to train the network to reproduce the training images \(x_i \in M\). Following the sparse regularization strategy, the second term forces the network to learn convolutions such that high-pass filtered coefficients are sparse. The additional term \(\gamma_1 \| \theta \|^2_2 + \gamma_2 \| \eta \|^2_2\) ensures the coercivity of the loss-function and balances the size of the weights \(\theta\) and \(\eta\).

Results for the sparse network training described above can be found in [20]. Actual application of the trained network and the augmented \(\ell^q\)-NETT to limited data problems in CT and elsewhere is subject of current work.

5 Conclusion and outlook

In this paper, we proposed and analyzed \(\ell^q\) regularization (called augmented \(\ell^q\)-NETT) using the encoder of a encoder-decoder network \(D \circ E\) as sparsifying transform. In order to obtain the coercivity of the regularizer, we augmented \(\sum_{\lambda \in A} w_\lambda \| (E(x))_\lambda \|^q\) with an additional penalty \(\frac{\beta}{2} \| x - (D \circ E)(x) \|^2\), which can be seen as a measure for the distance of \(x\) from the ideal data manifold. We were able to prove well-posedness and convergence of the augmented \(\ell^q\)-NETT and derived convergence rates in the absolute Bregman distance. We proposed the modified tight frame U-net for the network architecture together with an appropriate sparse training strategy.

Application to sparse data tomography is subject of current work. Theoretical comparison with frame and dictionary based sparse regularization methods will be studied. Moreover, we work on the derivation of additional provable convergence rates results of the augmented \(\ell^q\)-NETT.

Acknowledgments

D.O. and M.H. acknowledge support of the Austrian Science Fund (FWF), project P 30747-N32. The research of L.N. has been supported by the National Science Foundation (NSF) Grants DMS 1212125 and DMS 1616904.


References

[1] H. W. Engl, M. Hanke, and A. Neubauer, Regularization of inverse problems, ser. Mathematics and its Applications. Dordrecht: Kluwer Academic Publishers Group, 1996, vol. 375.

[2] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen, Variational methods in imaging, ser. Applied Mathematical Sciences. New York: Springer, 2009, vol. 167.

[3] D. Lee, J. Yoo, and J. C. Ye, “Deep residual learning for compressed sensing MRI,” in IEEE 14th International Symposium on Biomedical Imaging, 2017, pp. 15–18.

[4] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, “Deep convolutional neural network for inverse problems in imaging,” IEEE Trans. Image Process., vol. 26, no. 9, pp. 4509–4522, 2017.

[5] S. Antholzer, M. Haltmeier, and J. Schwab, “Deep learning for photoacoustic tomography from sparse data,” Inverse Probl. Sci. and Eng., vol. in press, pp. 1–19, 2018.

[6] E. Kobler, T. Klatzer, K. Hammernik, and T. Pock, “Variational networks: connecting variational methods and deep learning,” in German Conference on Pattern Recognition. Springer, 2017, pp. 281–293.

[7] J. R. Chang, C.-L. Li, B. Poczos, and B. V. Kumar, “One network to solve them all–solving linear inverse problems using deep projection models,” in IEEE International Conference on Computer Vision (ICCV), 2017, pp. 5889–5898.

[8] J. Adler and O. Öktem, “Solving ill-posed inverse problems using iterative deep neural networks,” Inverse Probl., vol. 33, p. 124007, 2017.

[9] J. Schwab, S. Antholzer, and M. Haltmeier, “Deep null space learning for inverse problems: convergence analysis and rates,” Inverse Probl., vol. 35, no. 2, p. 025008, 2019.

[10] H. Li, J. Schwab, S. Antholzer, and M. Haltmeier, “NETT: Solving inverse problems with deep neural networks,” 2018, arXiv:1803.00092.

[11] S. Antholzer, J. Schwab, J. Bauer-Marschallinger, P. Burgholzer, and M. Haltmeier, “NETT regularization for compressed sensing photoacoustic tomography,” in Photons Plus Ultrasound: Imaging and Sensing 2019, vol. 10878, 2019, p. 108783B.

[12] Y. Han and J. C. Ye, “Framing U-Net via deep convolutional framelets: Application to sparse-view CT,” IEEE Trans. Med. Imag., vol. 37, pp. 1418–1429, 2018.
[13] I. Daubechies, M. Defrise, and C. De Mol, “An iterative thresholding algorithm for linear inverse problems with a sparsity constraint,” *Comm. Pure Appl. Math.*, vol. 57, pp. 1413–1457, 2004.

[14] R. Ramlau and G. Teschke, “A tikhonov-based projection iteration for non-linear ill-posed problems with sparsity constraints,” *Numerische Mathematik*, vol. 104, no. 2, pp. 177–203, 2006.

[15] M. Grasmair, M. Haltmeier, and O. Scherzer, “Sparse regularization with $l^1$ penalty term,” *Inverse Probl.*, vol. 24, no. 5, pp. 055020, 13, 2008.

[16] ——, “Necessary and sufficient conditions for linear convergence of $l^1$-regularization,” *Comm. Pure Appl. Math.*, vol. 64, no. 2, pp. 161–182, 2011.

[17] S. Vaiter, G. Peyré, C. Dossal, and J. Fadili, “Robust sparse analysis regularization,” *IEEE Transactions on information theory*, vol. 59, no. 4, pp. 2001–2016, 2012.

[18] M. Burger, J. Flemming, and B. Hofmann, “Convergence rates in $l^1$-regularization if the sparsity assumption fails,” *Inverse Problems*, vol. 29, no. 2, p. 025013, 2013.

[19] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.

[20] D. Obmann, J. Schwab, and M. Haltmeier, “Sparse synthesis regularization with deep neural networks,” *arXiv:1902.00390*, 2019.