Nr. 51
23. July 2018

universität innsbruck
Institut für Mathematik

Preprint-Series: Department of Mathematics - Applied Mathematics

NETT: Solving Inverse Problems with Deep Neural Networks

H. Li, J. Schwab, S. Antholzer, M. Haltmeier

APPLIEDMATHEMATICS

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

NETT: Solving Inverse Problems with Deep Neural Networks

Housen Li[1], Johannes Schwab[2], Stephan Antholzer[2], and Markus Haltmeier[2,⋆]

[1]Institute for Mathematical Stochastics, University of Göttingen, Goldschmidtstrasse 7, 37077 Göttingen, Germany

[2]Department of Mathematics, University of Innsbruck Technikerstrasse 13, 6020 Innsbruck, Austria

[⋆]Correspondence: markus.haltmeier@uibk.ac.at

May 19, 2018

Abstract

Recovering a function or high-dimensional parameter vector from indirect measurements is a central task in various scientific areas. Several methods for solving such inverse problems are well developed and well understood. Recently, novel algorithms using deep learning and neural networks for inverse problems appeared. While still in their infancy, these techniques show astonishing performance for applications like low-dose CT or various sparse data problems. However, theoretical results for deep learning in inverse problems are missing so far. In this paper, we establish such a convergence analysis for the proposed NETT (Network Tikhonov) approach to inverse problems. NETT considers regularized solutions having small value of a regularizer defined by a trained neural network. Opposed to existing deep learning approaches, our regularization scheme enforces data consistency also for the actual unknown to be recovered. This is beneficial in case the unknown to be recovered is not sufficiently similar to available training data. We present a complete convergence analysis for NETT, where we derive well-posedness results and quantitative error estimates, and propose a possible strategy for training the regularizer. Numerical results are presented for a tomographic sparse data problem using the ℓ^q-norm of auto-encoder as trained regularizer, which demonstrate good performance of NETT even for unknowns of different type from the training data.

Keywords: inverse problems, deep learning, convergence analysis, image reconstruction, convolutional neural networks, ℓ^q regularization, total nonlinearity.

AMS subject classifications: 65J20, 65J22, 45F05

1 Introduction

We study the stable solution of inverse problems of the form

$$\text{Estimate } x \in D \text{ from data} \quad y_\delta = \mathbf{F}(x) + \xi_\delta\,. \tag{1.1}$$

Here $\mathbf{F}\colon D \subseteq X \to Y$ is a possibly nonlinear operator between Banach spaces $(X, \|\cdot\|)$ and $(Y, \|\cdot\|)$ with domain D. We thereby allow a possibly infinite-dimensional function space setting, but clearly the approach and results apply to a finite dimensional setting as well. The element $\xi_\delta \in Y$ models the unknown data error (noise) which is assumed to satisfy the estimate $\|\xi_\delta\| \leq \delta$ for some noise level $\delta \geq 0$. We focus on the ill-posed (or ill-conditioned) case where without additional information, the solution of (1.1) is either highly unstable, highly undetermined, or both. Many inverse problems in biomedical imaging, geophysics, engineering sciences, or elsewhere can be written in such a form (see, for example, [12, 29, 34]). For its stable solution one has to employ regularization methods, which are based on approximating (1.1) by neighboring well-posed problems, which enforce stability and uniqueness.

1.1 NETT regularization

Any method for the stable solution of (1.1) uses, either implicitly or explicitly, a-priori information about the unknowns to be recovered. Such information can be that x belongs to a certain set of admissible elements or that x has small value of a regularizer (or regularization functional) $\mathcal{R}\colon X \to [0,\infty]$. In this paper we focus on the latter situation, and assume that the regularizer takes the form

$$\forall x \in X\colon \quad \mathcal{R}(x) = \mathcal{R}(\mathbb{V}, x) := \psi(\boldsymbol{\Phi}(\mathbb{V}, x)) \,. \tag{1.2}$$

Here $\psi\colon \mathbb{X}_L \to [0,\infty]$ is a scalar functional and $\boldsymbol{\Phi}(\mathbb{V}, \cdot\,)\colon X \to \mathbb{X}_L$ a neural network of depth L where $\mathbb{V} \in \mathcal{V}$, for some vector space $\mathcal{V}$, contains free parameters that can be adjusted to available training data (see Section 2.1 for a precise formulation).

With the regularizer (1.2), we approach (1.1) via

$$\mathcal{T}_{\alpha;y_\delta}(x) := \mathcal{D}(\mathbf{F}(x), y_\delta) + \alpha\mathcal{R}(\mathbb{V}, x) \to \min_{x\in D} \,, \tag{1.3}$$

where $\mathcal{D}\colon Y \times Y \to [0,\infty]$ is an appropriate similarity measure in the data space enforcing data consistency. One may take $\mathcal{D}(\mathbf{F}(x), y_\delta) = \|\mathbf{F}(x) - y_\delta\|^2$ but also other distance measures such as the Kullback-Leibler divergence (which, among others, is used in emission tomography) are reasonable choices. Optimization problem (1.3) can be seen as a particular instance of generalized Tikhonov regularization for solving (1.1) with a neural network as regularizer. We therefore name (1.3) network Tikhonov (NETT) approach for inverse problems.

In this paper, we show that under reasonable assumptions, the NETT approach (1.3) is stably solvable. As $\delta \to 0$, the regularized solutions $x_{\alpha,\delta} \in \arg\min_x \mathcal{T}_{\alpha;y_\delta}(x)$ are shown to converge to $\mathcal{R}(\mathbb{V}, \cdot\,)$-minimizing solutions of $\mathbf{F}(x) = y_0$. Here and below $\mathcal{R}(\mathbb{V}, \cdot\,)$-minimizing solutions of $\mathbf{F}(x) = y_0$ are defined as any element

$$x_+ \in \arg\min \{\mathcal{R}(\mathbb{V}, x) \mid x \in D \wedge \mathbf{F}(x) = y_0\} \,. \tag{1.4}$$

Additionally, we derive convergence rates (quantitative error estimates) between $\mathcal{R}(\mathbb{V}, \cdot\,)$-minimizing solutions x_+ and regularized solutions $x_{\alpha,\delta}$. As a consequence, (1.3) provides a stable solution scheme for (1.1) using data consistence and encoding a-priori knowledge via neural networks.

1.2 Possible regularizers

The network regularizer $\mathcal{R}(\mathbb{V}, \cdot\,)$ can either be user-specified, or a trained network, where free parameters are adjusted on appropriate training data. Some examples are as follows.

■ *Non-convex ℓ^q-regularizer:* A simple user-specified instance of the regularizer (1.2) is the convex ℓ^q-regularizer $\mathcal{R}(\mathbb{V}, x) = \sum_{\lambda\in\Lambda} v_\lambda \left|\langle x, \varphi_\lambda\rangle\right|^q$. Here $(\varphi_\lambda)_{\lambda\in\Lambda}$ is a prescribed basis or frame and $(v_\lambda)_{\lambda\in\Lambda}$ are weights. In this case, the neural network is simply given by the analysis operator $\boldsymbol{\Phi}(\mathbb{V}, \cdot\,)\colon X \to \ell^2(\Lambda)\colon x \mapsto \langle x, \varphi_\lambda\rangle$ and NETT regularization reduces to sparse ℓ^q-regularization [11, 16, 17, 26, 31]. This form of the regularizer can also be combined with a training procedure by adjusting the weights $(v_\lambda)_{\lambda\in\Lambda}$ to a class of training data.

In this paper we in particular study a non-convex extension of ℓ^q-regularization, where the regularizer takes the form

$$\mathcal{R}(\mathbb{V}, x) = \sum_{\lambda\in\Lambda} v_\lambda \left|\boldsymbol{\Phi}_\lambda(\mathbb{V}, x)\right|^q , \tag{1.5}$$

with $q \geq 1$ and $\boldsymbol{\Phi}(\mathbb{V}, \cdot\,) = (\boldsymbol{\Phi}_\lambda(\mathbb{V}, \cdot\,))_{\lambda\in\Lambda}$ being a possible nonlinear neural network with multiple layers. In Section 3.4 we present convergence results for this non-convex generalization of ℓ^q-regularization.

■ *CNN regularizer:* The neural network regularizer $\mathcal{R}(\mathbb{V}, \cdot\,)$ in (1.2) may also be defined by a convolutional neural network (CNN) $\boldsymbol{\Phi}(\mathbb{V}, \cdot\,)$, containing free parameters that can be adjusted on appropriate training data. The CNN can be trained in such a way, that the regularizer has small value for elements x in a class of desirable phantoms and large value on a class of un-desirable phantoms. In Section 5, we present a possible regularizer design using an encoder-decoder scheme together with a strategy for training the CNN. We also present numerical results demonstrating that our approach performs well in practice for a sparse tomographic data problem.

1.3 Comparison to previous work

Very recently, several deep learning approaches for inverse problems have been developed (see, for example, [1, 3, 8, 23, 24, 39, 21, 36, 40, 41]). In all these approaches, a reconstruction network $\boldsymbol{\Phi}_{\rm rec}(\mathbb{V}, \cdot\,)\colon Y \to X$ is trained to map measured data to the desired output image.

Most reconstruction networks take the form $\boldsymbol{\Phi}_{\rm rec}(\mathbb{V}, \cdot\,) = \boldsymbol{\Phi}_{\rm CNN}(\mathbb{V}, \cdot\,) \circ \mathbf{B}$, where $\mathbf{B}\colon Y \to X$ maps the data to the reconstruction space (backprojection; no free parameters) and $\boldsymbol{\Phi}_{\rm CNN}(\mathbb{V}, \cdot\,)\colon X \to X$ is a convolutional neural network (CNN) whose free parameters are adjusted to the training data. This basic form allows the use of well established CNNs for image reconstruction [14] and already demonstrates impressing results. Another class of reconstruction networks learn free parameters in iterative schemes. In such approaches, the reconstruction network can be written in the form

$$\boldsymbol{\Phi}_{\rm rec}(\mathbb{V}, y) = \left(\boldsymbol{\Phi}_N(\mathbb{V}_N, \cdot\,) \circ \mathbf{B}_N(y, \cdot\,) \circ \cdots \circ \boldsymbol{\Phi}_1(\mathbb{V}_1, \cdot\,) \circ \mathbf{B}_1(y, \cdot\,)\right)(x_0)\,,$$

where x_0 is the initial guess, $\boldsymbol{\Phi}_k(\mathbb{V}_k, \cdot\,)\colon X \to X$ are CNNs that can be trained, and $\mathbf{B}_k(y, \cdot\,)\colon X \to X$ are iterative updates based on the forward operator and the data. The iterative updates may be defined by a gradient step with respect to the given inverse problem. The free parameters are adjusted to available training data.

Trained iterative schemes repeatedly make use of the forward problem which might yields increased data consistency compared to the first class of methods. Nevertheless, in all existing approaches, no provable non-trivial estimates bounding the data consistency term $\mathcal{D}(\mathbf{F}(x), y)$ are available; data consistency can only be guaranteed for the training data $(\mathbf{F}(z_n), z_n)_{n=1}^N$ for which the parameters in the neural network are optimized. This may results in instability and degraded reconstruction quality if the unknown to be recovered is not similar enough to the class of employed training data. The proposed NETT bounds the data consistency term $\mathcal{D}(\mathbf{F}(x_{\alpha,\delta}), y)$ also for data outside the training set. We expect the combination of the forward problem and a neural network via (1.3) (or, for the noiseless case, (1.4)) to increase reconstruction quality, especially in the case of limited access to a large amount of appropriate training data. Note, further, that the formulation of NETT separates the noise characteristic and the a-priori information of unknowns. This allows us to incorporate the knowledge of data generating mechanism, e.g. Poisson noise or Gaussian noise, by choosing the corresponding log-likelihood as the data consistency term, and also simplifies the training process of $\mathcal{R}(\mathbb{V}, \cdot)$, as it to some extend avoids the impact of noise. Meanwhile, this enhances the interpretability of the resulting approach: we on the one hand require its fidelity to the data, and on the other penalize unfavorable features (e.g. artifacts in tomography).

The results in this paper are a first main step for regularization with neural networks. We propose a new framework in the form of Tikhonov regularization with neural network (the NETT) and present a complete convergence analysis under reasonable assumptions (see Condition 2.2). Many further issues can be addressed in future work. This includes the design of appropriate CNN regularizers, the development of efficient algorithms for minimizing (1.3), and the consideration of other regularization strategies for (1.4). The focus of the present paper is on the theoretical analysis of NETT and demonstrating the feasibility of our approach; detailed comparison with other methods in terms of reconstruction quality, computational performance and applicability to real-world data is beyond our scope here and will be addressed in future work.

1.4 Outline

The rest of this paper is organized as follows. In Section 2, we describe the proposed NETT framework for solving inverse problems. We show its stability and derive convergence in the weak topology (see Theorem 2.3). To obtain the strong convergence of NETT, we introduce a new notion of total nonlinearity of non-convex functionals. For totally nonlinear regularizers, we show norm convergence of NETT (see Section 2.9). Convergence rates (quantitative error estimates) for NETT are derived in Section 3. Among others, we derive a convergence rate result in terms of the absolute Bregman distance (see Proposition 3.3). A framework for learning the regularizer using an auto encoder-decoder strategy is developed in Section 4, and applied to a sparse data problem in photoacoustic tomography in Section 5. The paper concludes with a short summary and outlook presented in Section 6.

2 NETT regularization

In the section we introduce the proposed NETT and analyze its well-posedness (existence, stability and weak convergence). We introduce a new property (total nonlinearity), which is applied to establish convergence of NETT with respect to the norm.

2.1 The NETT framework

Our goal is to solve (1.1) with $\|\xi_\delta\| \leq \delta$ and $\delta > 0$. For that purpose we consider minimizing the NETT functional (1.3), where the regularizer $\mathcal{R}(\mathbb{V}, \cdot)\colon X \to [0, \infty]$ in (1.2) is defined by a neural network of the form

$$\boldsymbol{\Phi}(\mathbb{V}, x) := (\sigma_L \circ \mathbb{V}_L \circ \sigma_{L-1} \circ \mathbb{V}_{L-1} \circ \cdots \circ \sigma_1 \circ \mathbb{V}_1)(x) \,. \tag{2.1}$$

Here L is the depth of the network (the number of layers after the input layer) and $\mathbb{V}_\ell(x) = \mathbb{A}_\ell(x) + b_\ell$ are affine linear operators between Banach spaces $\mathbb{X}_{\ell-1}$ and $\mathbb{X}_{\ell-1/2}$; we take $\mathbb{X}_0 := X$. The operators $\mathbb{A}_\ell \colon \mathbb{X}_{\ell-1} \to \mathbb{X}_{\ell-1/2}$ are the linear parts and $b_\ell \in \mathbb{X}_{\ell-1/2}$ the so-called bias terms. The operators $\sigma_\ell \colon \mathbb{X}_{\ell-1/2} \to \mathbb{X}_\ell$ are possibly nonlinear and the functionals $\psi \colon \mathbb{X}_L \to [0, \infty]$ are possibly non-convex.

As common in machine learning, the affine mappings $\mathbb{V}_\ell$ depend on free parameters that can be adjusted in the training phase, whereas the nonlinearities σ_ℓ are fixed. Therefore $\mathbb{V}_\ell$ and σ_ℓ are treated separately and only the affine part $\mathbb{V} = (\mathbb{V}_\ell)_{\ell=1}^L$ is indicated in the notion of the neural network regularizer $\mathcal{R}(\mathbb{V}, \cdot\,)$. Throughout our theoretical analysis we assume $\mathcal{R}(\mathbb{V}, \cdot\,)$ to be given and all free parameters to be trained before the minimization of (1.3). In Section 4, we present a possible framework for training a neural network regularizer based on an encoder-decoder strategy.

Remark 2.1 (CNNs in Banach space setting)**.** *A typical instance for the neural network in NETT* (1.2)*, is a deep convolutional neural network (CNN). In a possible infinite dimensional setting, such CNNs can be written in the form* (2.1)*, where the involved spaces satisfy* $\mathbb{X}_\ell := \ell^p(\Lambda_\ell, X_\ell)$ *and* $\mathbb{X}_{\ell-1/2} := \ell^p(\Lambda_\ell, X_{\ell-1/2})$ *with* $p \geq 1$*,* X_ℓ *and* $X_{\ell-1/2}$ *being function spaces, and* Λ_ℓ *being an at most countable set that specifies the number of different filters (depth) of the* ℓ*-th layer. The linear operators* $\mathbb{A}_\ell \colon \ell^p(\Lambda_{\ell-1}, X_{\ell-1}) \to \ell^p(\Lambda_\ell, X_{\ell-1/2})$ *are taken as*

$$\forall x \in \ell^p(\Lambda_{\ell-1}, X_{\ell-1}) \; \forall \ell \in \Lambda_\ell \colon \quad \mathbb{A}_\ell(x) = \left(\sum_{\mu \in \Lambda_{\ell-1}} K_{\lambda,\mu}^{(\ell)}(x_\mu) \right)_{\lambda \in \Lambda_\ell} , \tag{2.2}$$

where $K_{\lambda,\mu}^{(\ell)} \colon X_{\ell-1} \to X_{\ell-1/2}$ *are convolution operators.*

We point out, that in the existing machine learning literature, only finite dimensional settings have been considered so far, where $\mathbb{X}_\ell$ *and* $\mathbb{X}_{\ell-1/2}$ *are finite dimensional spaces. In such a finite dimensional setting, we can take* $X_\ell = \mathbb{R}^{N_1^\ell \times N_2^\ell}$*, and* Λ_ℓ *as a set with* N_c^ℓ *elements. One then can identify* $\mathbb{X}_\ell = \ell^p(\Lambda_\ell, X_\ell) \simeq \mathbb{R}^{N_1^\ell \times N_2^\ell \times N_c^\ell}$ *and interpreted its elements as stack of discrete images (the same holds for* $\mathbb{X}_{\ell-1/2}$*). In typical CNNs, either the dimensions* $N_1^\ell \times N_2^\ell$ *of the base space* X_ℓ *are progressively reduced and number of channels* N_c^ℓ *increased, or vice versa. While we are not aware of any infinite dimensional general formulation of CNNs, our proposed formulation* (2.1), (2.2) *is the natural infinite-dimensional Banach space version of CNNs, which reduces to standard CNNs [14] in the finite dimensional setting.*

Basic convex regularizers are sparse ℓ^q-penalties $\mathcal{R}(\mathbb{V}, x) = \sum_{\lambda \in \Lambda} v_\lambda \, |\langle \varphi_\lambda, x \rangle|^q$. In this case one may take (2.1) as a single-layer neural network with $\mathbb{X}_1 = \ell^2(\Lambda_1, \mathbb{R})$, $\sigma = \mathrm{Id}$ and $\boldsymbol{\Phi}(\mathbb{V}, x) = \mathbb{V}(x) = (\langle \varphi_\lambda, x \rangle)_\lambda$ being the analysis operator to some frame $(\varphi_\lambda)_{\lambda \in \Lambda}$. The non-convex functional $\psi(x) = \sum_{\lambda \in \Lambda} v_\lambda \, |x_\lambda|^q$ is a weighted ℓ^q-norm. The frame $(\varphi_\lambda)_{\lambda \in \Lambda}$ may be a prescribed wavelet or curvelet basis [7, 10, 28] or a trained dictionary [2, 19]. In Section 3.4, we analyze a non-convex version of ℓ^q-regularization, where $\langle \phi_\lambda, \cdot\,\rangle$ are replaced by nonlinear functionals.

2.2 Well-posedness and weak convergence

For the convergence analysis of NETT regularization, we use the following assumptions on the regularizer and the data consistency term in (1.3).

Condition 2.2 (Convergence of NETT regularization)**.**

(A1) The regularizer $\mathcal{R}(\mathbb{V},\,\cdot\,)$ is defined by (1.2) and (2.1) and satisfies the following:

- $\mathbb{V}_\ell \colon \mathbb{X}_{\ell-1} \to \mathbb{X}_{\ell-1/2}$ are affine operators of the form $\mathbb{V}_\ell(x) = \mathbb{A}_\ell x + b_\ell$;
- $\mathbb{A}_\ell$ are bounded linear and for some $c_\ell \in [0,\infty)$, we have $\forall x \in X \colon \|x\| \leq c_\ell \|\mathbb{A}_\ell x\|$;
- σ_ℓ are weakly continuous and coercive, that is $\|\sigma_\ell(x)\| \to \infty$ as $\|x\| \to \infty$;
- The functional ψ is lower semi-continuous and coercive.

(A2) The data consistency term satsfies the following:

- For some $\tau \geq 1$ we have $\forall y_0,\, y_1,\, y_2 \in Y \colon \mathcal{D}(y_0, y_1) \leq \tau \mathcal{D}(y_0, y_2) + \tau \mathcal{D}(y_2, y_1)$;
- $\forall y_0,\, y_1 \in Y \colon \mathcal{D}(y_0, y_1) = 0 \iff y_0 = y_1$;
- $\forall (y_k)_{k\in\mathbb{N}} \in Y^{\mathbb{N}} \colon y_k \to y \implies \mathcal{D}(y_k, y) \to 0$;
- The functional $(x, y) \mapsto \mathcal{D}(\mathbf{F}(x), y)$ is sequentially lower semi-continuous.

In CNNs, the spaces $\mathbb{X}_\ell$ and $\mathbb{X}_{\ell-1/2}$ are function spaces (see Remark 2.1) and a standard operation for σ_ℓ is the ReLU (the rectified linear unit), $\operatorname{ReLU}(x) := \max\{x, 0\}$, that is applied component-wise. The plain form of the ReLU is not coercive. However, the slight modification $x \mapsto \max\{x, ax\}$ for some $a \in (0,1)$, named leaky ReLU, is coercive, see [27, 22]. Another coercive standard operation for σ_ℓ in CNNs is max pooling which takes the maximum value $\max\{|x(i)| : i \in I_k\}$ within clusters of transform coefficients.

Theorem 2.3 (Well-posedness of CNN-regularization)**.** *Let Condition 2.2 be satisfied. Then the following assertions hold true:*

(a) Existence: *For all $y \in Y$ and $\alpha > 0$, there exists a minimizer of $\mathcal{T}_{\alpha;y}$;*

(b) Stability: *If $y_k \to y$ and $x_k \in \arg\min \mathcal{T}_{\alpha;y_k}$, then weak accumulation points of $(x_k)_{k\in\mathbb{N}}$ exist and are minimizers of $\mathcal{T}_{\alpha;y}$.*

(c) Convergence: *Let $x \in X$, $y := \mathbf{F}(x)$, $(y_k)_{k\in\mathbb{N}}$ satisfy $\mathcal{D}(y_k, y), \mathcal{D}(y, y_k) \leq \delta_k$ for some sequence $(\delta_k)_{k\in\mathbb{N}} \in (0,\infty)^{\mathbb{N}}$ with $\delta_k \to 0$, suppose $x_k \in \arg\min_x \mathcal{T}_{\alpha(\delta_k)}(x, y_k)$, and let the parameter choice $\alpha \colon (0,\infty) \to (0,\infty)$ satisfy*

$$\lim_{\delta\to 0} \alpha(\delta) = \lim_{\delta\to 0} \frac{\delta}{\alpha(\delta)} = 0\,. \tag{2.3}$$

Then the following holds:

- *Weak accumulation points of $(x_k)_{k\in\mathbb{N}}$ are $\mathcal{R}(\mathbb{V},\,\cdot\,)$-minimizing solutions of $\mathbf{F}(x) = y$;*
- *$(x_k)_{k\in\mathbb{N}}$ has at least one weak accumulation point x_+;*
- *Any weakly convergent subsequence $(x_{k(n)})_{n\in\mathbb{N}}$ satisfies $\mathcal{R}(\mathbb{V}, x_{k(n)}) \to \mathcal{R}(\mathbb{V}, x_+)$;*
- *If the $\mathcal{R}(\mathbb{V},\,\cdot\,)$-minimizing solution of $\mathbf{F}(x) = y$ is unique, then $x_k^\delta \rightharpoonup x_+$.*

Proof. According to [15, 34] it is sufficient to show that the functional $\mathcal{R}(\mathbb{V}, \cdot)$ is weakly sequentially lower semi-continuous and the set $\{x \mid \mathcal{T}_{\alpha;y}(x) \leq t\}$ is sequentially weakly pre-compact for all $t > 0$ and $y \in Y$ and $\alpha > 0$. By the Banach-Alaoglu theorem, the latter condition is satisfied if $\mathcal{R}(\mathbb{V}, \cdot)$ is coercive. The coercivity of $\mathcal{R}(\mathbb{V}, \cdot)$ however is directly implied by Condition 2.2. Also from Condition 2.2 it follows that $\mathcal{R}(\mathbb{V}, \cdot)$ is sequentially lower semi-continuous. □

Remark 2.4. *The boundedness of linear operators $\mathbb{V}_\ell$ seems restrictive, but it is somewhat indispensable. In fact, suppose for the moment that we can drop the boundedness requirement and ensure the first layer is lower semi-continuous. The lower semi-continuity of $\mathcal{R}(\mathbb{V}, \cdot)$ asks for the weights in $\mathbb{V}_\ell$ for $\ell \geq 2$, are all positive, since certain order has to be preserved. This positive weight assumption meanwhile leads to the convexity of $\mathcal{R}(\mathbb{V}, \cdot)$, which, in practice, is often not the case.*

2.3 Absolute Bregman distance and total nonlinearity

For convex regularizers, the notion of Bregman distance is a powerful concept [4, 34]. For non-convex regularizers, the standard definition of the Bregman distance takes negative values. In this paper, we therefore use the notion of absolute Bregman distance. To the best of our knowledge, the absolute Bregman distance has not been used in regularization theory so far.

Definition 2.5 (Absolute Bregman distance)**.** Let $\mathcal{F}\colon D \subseteq X \to \mathbb{R}$ be Gâteaux differentiable at $x \in X$. The *absolute Bregman distance* $\mathcal{B}_\mathcal{F}(\,\cdot\,, x)\colon D \to [0, \infty]$ with respect to $\mathcal{F}$ at x is defined by

$$\forall \tilde{x} \in X\colon \quad \mathcal{B}_\mathcal{F}(\tilde{x}, x) := \left|\mathcal{F}(\tilde{x}) - \mathcal{F}(x) - \mathcal{F}'(x)(\tilde{x} - x)\right|. \tag{2.4}$$

Here $\mathcal{F}'(x)$ denotes the Gâteaux derivative of $\mathcal{F}$ at x.

From Theorem 2.3 we can conclude convergence of $x_{\alpha,\delta}$ to the exact solution in the absolute Bregman distance. Below we show that this implies strong convergence under some additional assumption on the regularization functional. For this purpose we introduce the new total nonlinearity, which has not been studied before.

Definition 2.6 (Total nonlinearity)**.** Let $\mathcal{F}\colon D \subseteq X \to \mathbb{R}$ be Gâteaux differentiable at $x \in D$. We define the *modulus of total nonlinearity* of $\mathcal{F}$ at x as $\nu_\mathcal{F}(x, \cdot) : [0, \infty) \to [0, \infty]$,

$$\forall t > 0\colon \quad \nu_\mathcal{F}(x, t) := \inf\left\{\mathcal{B}_\mathcal{F}(\tilde{x}, x) \mid \tilde{x} \in D \wedge \|\tilde{x} - x\| = t\right\}. \tag{2.5}$$

The function $\mathcal{F}$ is called *totally nonlinear* at x if $\nu_\mathcal{F}(x, t) > 0$ for all $t \in (0, \infty)$.

The notion of total nonlinearity is similar to total convexity [6] for convex functionals. Opposed to total convexity we do not assume convexity of $\mathcal{F}$, and use the absolute Bregman distance instead of the standard Bregman distance. For convex functions, the total nonlinearity reduces to total convexity, as the Bregman distance is always non-negative for convex functionals. For a Gâteaux differentiable function, the total nonlinearity essentially requires that its second derivative at x is bounded away from zero. The functional $\mathcal{F}(x) := \sum_{\lambda\in\Lambda} v_\lambda |x_\lambda|^q$ with $v_\lambda > 0$ is totally nonlinear at every $x = (x_\lambda)_{\lambda\in\Lambda} \in \ell^\infty(\Lambda)$ if $q > 1$.

We have the following result, which generalizes [32, Proposition 2.2] (see also [34, Theorem 3.49]) from the convex to the non-convex case.

Proposition 2.7 (Characterization of total nonlinearity)**.** *For $\mathcal{F}\colon D \subseteq X \to \mathbb{R}$ and any $x \in D$ the following assertions are equivalent:*

(i) The function $\mathcal{F}$ is totally nonlinear at x;

(ii) $\forall\, (x_n)_{n\in\mathbb{N}} \subseteq D^{\mathbb{N}}\colon\ \lim_{n\to\infty} \mathcal{B}_{\mathcal{F}}(x_n, x) = 0 \Rightarrow \lim_{n\to\infty} \|x_n - x\| = 0.$

Proof. The proof of the implication (ii) $\Rightarrow$ (i) is the same as [32, Proposition 2.2]. For the implication (i) $\Rightarrow$ (ii), let (i) hold, let $(x_n)_{n\in\mathbb{N}} \subseteq D^{\mathbb{N}}$ satisfy $\mathcal{B}_{\mathcal{F}}(x_n, x) \to 0$, and suppose $\lim_{n\to\infty} \|x_n - x\| = \delta > 0$ for the moment. For any $\varepsilon > 0$, by the continuity of $\mathcal{B}_{\mathcal{F}}(\,\cdot\,, x)$, there exist $\tilde{x}_n$ with $\|\tilde{x}_n - x\| = \delta$ such that for sufficiently large n

$$\varepsilon \geq \mathcal{B}_{\mathcal{F}}(x_n, x) + \frac{\varepsilon}{2} \geq \mathcal{B}_{\mathcal{F}}(\tilde{x}_n, x) \geq \nu_{\mathcal{F}}(x, \delta).$$

This leads to $\nu_{\mathcal{F}}(x, \delta) = 0$, which contradicts with the total nonlinearity of $\mathcal{F}$ at x. Then, the assertion follows by considering subsequences of $(x_n)_{n\in\mathbb{N}}$. □

Remark 2.8. *We point out that Proposition 2.7 remains true, if we replace the absolute value $|\cdot| : \mathbb{R} \to [0, \infty]$ in (2.4) by the ReLU function, the leaky ReLU function, or any other nonnegative continuous function $\phi\colon \mathbb{R} \to [0, \infty]$ that satisfies $\phi(0) = 0$.*

2.4 Strong convergence of NETT regularization

For totally nonlinear regularizers $\mathcal{R}(\mathbb{V}, \cdot)$ we can prove convergence of NETT with respect to the norm topology.

Theorem 2.9 (Strong convergence of NETT)**.** *Let Condition 2.2 hold and assume additionally that $\mathbf{F}(x) = y$ has a solution, $\mathcal{R}(\mathbb{V}, \cdot)$ is totally nonlinear at $\mathcal{R}(\mathbb{V}, \cdot)$-minimizing solutions, and α satisfies* (2.3)*. Then for every sequence $(y_k)_{k\in\mathbb{N}}$ with $\mathcal{D}(y_k, y), \mathcal{D}(y, y_k) \leq \delta_k$ where $\delta_k \to 0$ and every sequence $x_k \in \arg\min_x \mathcal{T}_{\alpha(\delta_k)}(x, y_k)$, there exist a subsequence $(x_{k(n)})_{n\in\mathbb{N}}$ and an $\mathcal{R}(\mathbb{V}, \cdot)$-minimizing solution x_+ with $\|x_{k(n)} - x_+\| \to 0$. If the $\mathcal{R}(\mathbb{V}, \cdot)$-minimizing solution is unique, then $x_k \to x_+$ with respect to the norm topology.*

Proof. It follows from Theorem 2.3 that there exists a subsequence $(x_{k(n)})_{n\in\mathbb{N}}$ weakly converging to some $\mathcal{R}(\mathbb{V}, \cdot)$-minimizing solution x_+ such that $\mathcal{R}(\mathbb{V}, x_{k(n)}) \to \mathcal{R}(\mathbb{V}, x_+)$. From the weak convergence of $(x_{k(n)})_{n\in\mathbb{N}}$ and the convergence of $(\mathcal{R}(\mathbb{V}, x_{k(n)}))_{n\in\mathbb{N}}$ it follows that $\mathcal{B}_{\mathcal{R}(\mathbb{V},\cdot)}(x_{k(n)}, x_+) \to 0$. Thus it follows from Proposition 2.7, that $\|x_{k(n)} - x_+\| \to 0$. If x_+ is the unique $\mathcal{R}$-minimizing solution, the strong convergence to x_+ again follows from Theorem 2.3 and Proposition 2.7. □

3 Convergence rates

In this section, we derive convergence rates for NETT in terms of general error measures under certain variational inequalities. We discuss instances where the variational inequality is satisfied for the absolute Bregman distance. Additionally, we consider a non-convex generalization of ℓ^q-regularization.

3.1 General convergence rates result

We study convergence rates in terms of a general functional $\mathcal{E} \colon X \times X \to [0, \infty]$ measuring closedness in the space X. For convex $\Psi \colon [0, \infty) \to [0, \infty)$, let $\Psi^* \colon \mathbb{R} \to \mathbb{R}$ denote the Fenchel conjugate of Ψ defined by $\Psi^*(\tau) := \sup\{\tau\, t - \Psi(t) \mid t \geq 0\}$.

Theorem 3.1 (Convergence rates for NETT)**.** *Suppose $\mathcal{E} \colon X \times X \to [0, \infty]$, let $x_+ \in D$ and assume that there exist a concave, continuous and strictly increasing function $\Phi \colon [0, \infty) \to [0, \infty)$ with $\Phi(0) = 0$ and a constant $\beta > 0$ such that for all $\epsilon > 0$ and $x \in D$ with $|\mathcal{R}(\mathbb{V}, x) - \mathcal{R}(\mathbb{V}, x_+)| < \varepsilon$ we have*

$$\beta \mathcal{E}(x, x_+) \leq \mathcal{R}(\mathbb{V}, x) - \mathcal{R}(\mathbb{V}, x_+) + \Phi(\mathcal{D}(\mathbf{F}(x), \mathbf{F}(x_+)))\,. \tag{3.1}$$

Additionally, let Condition 2.2 hold, let $y_\delta \in Y$ and $\delta > 0$ satisfy $\mathcal{D}(y, y_\delta), \mathcal{D}(y_\delta, y) \leq \delta$ and write Φ^{-*} for the Fenchel conjugate of the inverse function Φ^{-1}. Then the following assertions hold true:

(a) For sufficiently small α and δ, we have

$$\forall x_{\alpha,\delta} \in \arg\min \mathcal{T}_{\alpha;y_\delta} \colon \quad \beta \mathcal{E}(x_{\alpha,\delta}, x_+) \leq \frac{\delta}{\alpha} + \Phi(\tau\delta) + \frac{\Phi^{-*}(\tau\alpha)}{\tau\alpha}\,. \tag{3.2}$$

(b) If $\alpha \sim \delta/\Phi(\tau\delta)$, then $\mathcal{E}(x_{\alpha,\delta}, x_+) = \mathcal{O}(\Phi(\tau\delta))$ as $\delta \to 0$.

Proof. (a) By Theorem 2.3 (c), we can assume that $|\mathcal{R}(\mathbb{V}, \cdot)(x_{\alpha,\delta}) - \mathcal{R}(\mathbb{V}, x_+)| \leq \varepsilon$. Note that $\mathcal{D}(\mathbf{F}(x_{\alpha,\delta}), y_\delta) + \alpha\mathcal{R}(\mathbb{V}, \cdot)(x_{\alpha,\delta}) \leq \mathcal{D}(\mathbf{F}(x_+), y_\delta) + \alpha\mathcal{R}(\mathbb{V}, x_+)$. Then from (3.1) we obtain

$$\begin{aligned}
\alpha\beta\mathcal{E}(x_{\alpha,\delta}, x_+) &\leq \delta - \mathcal{D}(\mathbf{F}(x_{\alpha,\delta}), y_\delta) + \alpha\Phi(\mathcal{D}(\mathbf{F}(x_{\alpha,\delta}), \mathbf{F}(x_+))) \\
&\leq \delta - \mathcal{D}(\mathbf{F}(x_{\alpha,\delta}), y_\delta) + \alpha\Phi(\tau\delta + \tau\mathcal{D}(\mathbf{F}(x_{\alpha,\delta}), y_\delta)) \\
&\leq \delta - \mathcal{D}(\mathbf{F}(x_{\alpha,\delta}), y_\delta) + \alpha\Phi(\tau\delta) + \alpha\Phi(\tau\mathcal{D}(\mathbf{F}(x_{\alpha,\delta}), y_\delta)) \\
&\leq \delta + \alpha\Phi(\tau\delta) + \tau^{-1}\Phi^{-*}(\tau\alpha).
\end{aligned}$$

The last estimate is an application of Young's inequality $\alpha\Phi(\tau t) \leq t + \tau^{-1}\Phi^{-*}(\tau\alpha)$.

(b) Elementary computations show $\limsup_{\delta\to 0} \Phi^{-*}\big(\tau\delta/\Phi(\tau\delta)\big)/\delta < \infty$, such that the right hand side of (3.2) is bounded by $\Phi(\tau\delta)$ up to a constant if $\alpha \sim \delta/\Phi(\tau\delta)$. ☐

Remark 3.2. *If $\Phi(t) \leq Ct$ for sufficiently small t, then from (3.2) it follows that $\beta\mathcal{E}(x_{\alpha,\delta}, x) \leq \delta/\alpha + Cs\delta = \mathcal{O}(\delta)$ if $\alpha \leq 1/C$. This says that the regularization parameter α needs not vanish as $\delta \to 0$, which is often referred to as exact penalization (for the convex case, see the discussions in [34, 4]).*

3.2 Rates in the absolute Bregman distance

We next derive conditions under which a variational inequality in form of (3.1) is possible for the absolute Bregman distance as error measure, $\mathcal{E}(x, x_+) := \mathcal{B}_{\mathcal{R}(\mathbb{V},\cdot)}(x, x_+)$.

Proposition 3.3 (Rates in the absolute Bregman distance)**.** *Let X and Y be Hilbert spaces and let $\mathbf{F} \colon X \to Y$ be a bounded linear operator. Assume that $\mathcal{R}(\mathbb{V}, \cdot)$ is Gâteaux differentiable, that $\mathcal{R}'(\mathbb{V}, x_+) \in \operatorname{Ran}(\mathbf{F}^*)$, and that there exist positive constants γ, ε with*

$$\mathcal{R}(\mathbb{V}, x_+) - \mathcal{R}(\mathbb{V}, x) \leq \gamma\|\mathbf{F}(x) - \mathbf{F}(x_+)\| \tag{3.3}$$

for all x satisfying $|\mathcal{R}(\mathbb{V}, x) - \mathcal{R}(\mathbb{V}, x_+)| < \varepsilon$. Then,

$$\mathcal{B}_{\mathcal{R}(\mathbb{V},\,\cdot\,)}(x, x_+) \leq \mathcal{R}(x) - \mathcal{R}(x_+) + C\|\mathbf{F}(x) - \mathbf{F}(x_+)\| \qquad \text{for some constant } C.$$

In particular, for the distance measure $\mathcal{D}(z, y) = \|z - y\|^2$ and under Condition 2.2, Items (a) and (b) of Theorem 3.1 hold true.

Proof. Let ξ satisfy that $\mathcal{R}'(\mathbb{V}, x_+) = \mathbf{F}^*\xi$. Then $|\langle \mathcal{R}'(\mathbb{V}, x_+), x - x_+\rangle| \leq \|\xi\|\|\mathbf{F}(x) - \mathbf{F}(x_+)\|$. Note that $|\mathcal{R}(\mathbb{V}, x) - \mathcal{R}(\mathbb{V}, x_+)| = \mathcal{R}(\mathbb{V}, x) - \mathcal{R}(\mathbb{V}, x_+)$ if $\mathcal{R}(\mathbb{V}, x) \geq \mathcal{R}(\mathbb{V}, x_+)$ and $|\mathcal{R}(\mathbb{V}, x) - \mathcal{R}(\mathbb{V}, x_+)| = \mathcal{R}(\mathbb{V}, x) - \mathcal{R}(\mathbb{V}, x_+) + 2(\mathcal{R}(\mathbb{V}, x_+) - \mathcal{R}(\mathbb{V}, x)) \leq \mathcal{R}(\mathbb{V}, x) - \mathcal{R}(\mathbb{V}, x_+) + 2\gamma\|\mathbf{F}(x) - \mathbf{F}(x_+)\|$ otherwise. This yields

$$\begin{aligned}\mathcal{B}_{\mathcal{R}(\mathbb{V},\,\cdot\,)}(x, x_+) &\leq |\mathcal{R}(\mathbb{V}, x) - \mathcal{R}(\mathbb{V}, x_+)| + |\langle \mathcal{R}'(\mathbb{V}, x_+), x - x_+\rangle| \\ &\leq \mathcal{R}(\mathbb{V}, x) - \mathcal{R}(\mathbb{V}, x_+) + C\|\mathbf{F}(x) - \mathbf{F}(x_+)\|\end{aligned}$$

with the constant $C := \|\xi\| + 2\gamma$, and concludes the proof. □

Remark 3.4. *Proposition 3.3 shows that a variational inequality of the form* (3.1) *with $\beta = 1$ and $\Phi(t) = C\sqrt{t}$ follows from a classical source condition $\mathcal{R}'(x_+) \in \mathrm{Ran}(\mathbf{F}^*)$. By Theorem 3.1, it further implies that $\mathcal{B}_{\mathcal{R}(\mathbb{V},\,\cdot\,)}(x_{\alpha,\delta}, x_+) = \mathcal{O}(\delta)$ if $\|y - y_\delta\| \leq \delta$. Moreover, we point out that the additional assumption* (3.3) *is rather weak, and follows from the classical source condition $\mathcal{R}'(x_+) \in \mathrm{Ran}(\mathbf{F}^*)$ if $\mathcal{R}$ is convex, see [16]. It is clear that a sufficient condition to* (3.3) *is*

$$|\mathcal{R}(\mathbb{V}, x) - \mathcal{R}(\mathbb{V}, x_+) - \langle \mathcal{R}'(\mathbb{V}, x_+), x - x_+\rangle| \leq c|\mathcal{R}(\mathbb{V}, x) - \mathcal{R}(\mathbb{V}, x_+)| \qquad \text{for some } c < 1,$$

which resembles a tangential-cone condition.

3.3 General regularizers

So far we derived well-posedness, convergence and convergence rates for regularizers of the form (1.2). These results can be generalized to Tikhonov regularization

$$\mathcal{T}_{\alpha;y_\delta}(x) := \mathcal{D}(\mathbf{F}(x), y_\delta) + \alpha\mathcal{R}(x) \to \min_x\,, \tag{3.4}$$

where the regularization term is not necessarily defined by a neural network. These results are derived by replacing Condition 2.2 with the following one.

Condition 3.5 (Convergence for general regularizers)**.**

(B1) The functional $\mathcal{R}$ is sequentially lower semi-continuous.

(B2) The set $\{x \mid \mathcal{T}_{\alpha;y}(x) \leq t\}$ is sequentially pre-compact for all t, y and $\alpha > 0$.

(B3) The data consistency term satisfies (A2).

Then we have the following:

Theorem 3.6 (Results for general Tikhonov regularization)**.** *Under Condition 3.5, general Tikhonov regularization* (3.4) *satisfies the following:*

(a) The conclusions from Theorem 2.3 (well-posedness and convergence) hold true.

(b) *If $\mathcal{R}$ is totally nonlinear at $\mathcal{R}$-minimizing solutions, the strong convergence from Theorem 2.9 holds.*

(c) *The convergence rates result from Theorem 3.1 holds.*

(d) *If $\mathbf{F}$ is bounded linear, the assertions of Proposition 3.3 hold for $\mathcal{R}$.*

Proof. All assertions are shown as in the special case $\mathcal{R} = \mathcal{R}(\mathbb{V}, \cdot)$. □

Note that Item (a) in the above theorem is contained in [15]. Items (b)-(d) have not been obtained previously for non-convex regularizers.

3.4 Non-convex ℓ^q-regularization

We now analyze a special instance of NETT regularization (1.2), generalizing ℓ^q-regularization. More precisely, we consider the following non-convex ℓ^q-Tikhonov functional

$$\mathcal{T}_{\alpha,y_\delta}(x) = \|\mathbf{F}(x) - y_\delta\|^2 + \alpha \sum_{\lambda\in\Lambda} v_\lambda \, |\phi_\lambda(x)|^q \qquad \text{with } q > 1. \tag{3.5}$$

Here Λ is a countable set and $\phi_\lambda \colon X \to \mathbb{R}$ are possibly nonlinear functionals. The regularizer $\mathcal{R}(x) := \sum_{\lambda\in\Lambda} v_\lambda \, |\phi_\lambda(x)|^q$ is a particular instance of NETT (1.2), if we take $\sigma_L \circ \mathbb{V}_L \circ \cdots \circ \sigma_1 \circ \mathbb{V}_1 = (\phi_\lambda)_{\lambda\in\Lambda}$, and ψ as a weighted ℓ^q-norm. However, in (3.5) also more general choices for ϕ_λ are allowed (see Condition 3.7).

We assume the following:

Condition 3.7.

(C1) $\mathbf{F} \colon X \to Y$ is a bounded linear operator between Hilbert spaces X and Y.

(C2) $\phi_\lambda \colon X \to \mathbb{R}$ is Gâteaux differentiable for every $\lambda \in \Lambda$.

(C3) There is a $\mathcal{R}(\mathbb{V}, \cdot)$-minimizing solution x_+ with $\mathcal{R}'(\mathbb{V}, x_+) \in \operatorname{Ran}(\mathbf{F}^*)$;

(C4) There exist constants $C, \varepsilon > 0$ such that for all x with $|\mathcal{R}(\mathbb{V}, x) - \mathcal{R}(\mathbb{V}, x_+)| \leq \varepsilon$, it holds that

$$\forall \lambda \in \Lambda\colon \quad \operatorname{sgn}(\phi_\lambda(x_+))(\phi_\lambda(x_+) - \phi_\lambda(x)) \leq C \operatorname{sgn}(\phi_\lambda(x_+))\phi_\lambda'(x_+)(x_+ - x)\,.$$

Here $\operatorname{sgn}(t) = 1$ for $t > 0$, $\operatorname{sgn}(t) = 0$ for $t = 0$, and $\operatorname{sgn}(t) = -1$ otherwise.

Proposition 3.8. *Let Condition 3.7 hold, suppose that $y_\delta \in Y$ is such that $\|y - y_\delta\| \leq \delta$, and let $x_{\alpha,\delta} \in \arg\min \mathcal{T}_{\alpha,y_\delta}$. If choosing $\alpha \sim \delta$, then $\mathcal{B}_{\mathcal{R}}(x_{\alpha,\delta}, x_+) = \mathcal{O}(\delta)$.*

Proof. The convexity of $t \mapsto |t|^q$ implies that

$$\mathcal{R}(x_+) - \mathcal{R}(x) \leq \sum_{\lambda\in\Lambda} v_\lambda |\phi_\lambda(x_+)|^{q-1} \operatorname{sgn}(\phi_\lambda(x_+))(\phi_\lambda(x_+) - \phi_\lambda(x)).$$

By (C4), we obtain $\mathcal{R}(x_+) - \mathcal{R}(x) \leq C\langle \mathcal{R}'(x_+), x_+ - x\rangle$. This together with (C3) implies (3.3). Thus, the assertion follows from Proposition 3.6. □

Remark 3.9. *Consider the case that* $\phi_\lambda(x) := \langle \varphi_\lambda, x\rangle$ *for an orthonormal basis* $(\varphi_\lambda)_{\lambda\in\Lambda}$ *of* X. *It is known that* $\|x - x_+\|^2 = \mathcal{O}\big(\mathcal{B}_{\mathcal{R}(\mathbb{V},\,\cdot\,)}(x, x_+)\big)$, *see e.g. [34]. Then, Proposition 3.8 gives* $\|x_{\alpha,\delta} - x_+\| = \mathcal{O}(\delta^{1/2})$, *which reproduces the result of [34, Theorem 3.54]. This rate can be improved to* $\mathcal{O}(\delta^{1/q})$ *if we further assume sparsity of* x_+ *and restricted injectivity of* $\mathbf{F}$. *It can be shown by Theorem 3.1 because in such a situation* (3.1) *holds with* $\Phi(t) \sim \sqrt{t}$ *and* $\mathcal{E}(x, x_+) = \|x - x_+\|^q$, *see [16] for details.*

4 NETT regularization using auto-encoders

In this section we present a framework for constructing a trained neural network regularizer $\mathcal{R}(\mathbb{V},\,\cdot\,)$ of the form (2.1). The proposed network $\boldsymbol{\Phi}(\mathbb{V},\,\cdot\,)$ has the form of an auto-encoder. Additionally, we develop a strategy for network training and minimizing the NETT functional.

4.1 A trained regularizer

For the regularizer we propose $\mathcal{R}(\mathbb{V},\,\cdot\,) = \sum_{\lambda\in\Lambda_L} \|\boldsymbol{\Phi}_\lambda(\mathbb{V}, x)\|_p^p$ with a network $\boldsymbol{\Phi}(\mathbb{V},\,\cdot\,) = (\boldsymbol{\Phi}_\lambda(\mathbb{V},\,\cdot\,))_{\lambda\in\Lambda}$ of the form (2.1), that itself is part network of the encoder-decoder type,

$$\boldsymbol{\Psi}(\mathbb{W},\,\cdot\,)\circ\boldsymbol{\Phi}(\mathbb{V},\,\cdot\,)\colon X \to X\,. \tag{4.1}$$

Here $\boldsymbol{\Phi}(\mathbb{V},\,\cdot\,)\colon X \to \mathbb{X}_L$ can be interpreted as encoding network and $\boldsymbol{\Psi}\colon \mathbb{X}_L \to X$ as decoding network. Any network with at least one hidden layer can be written in the form (4.1). Training of the network is performed such that $\mathcal{R}(\mathbb{V},\,\cdot\,)$ is small for artifact free images and large for images with artifacts. A possible training strategy is presented below.

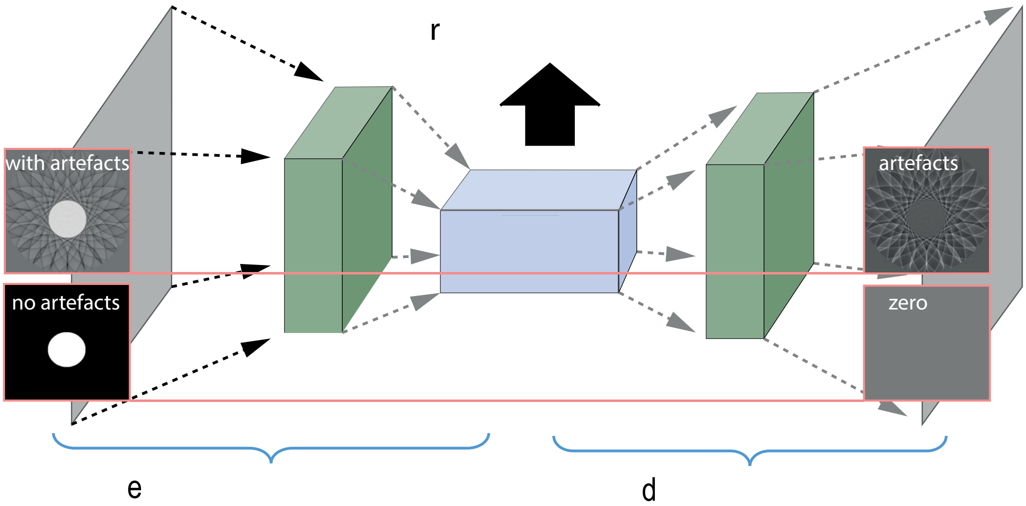

Figure 4.1: **Encoder-decoder scheme and training strategy.** The network consists of the encoder part $\boldsymbol{\Phi}(\mathbb{V},\,\cdot\,)$ and decoder part $\boldsymbol{\Psi}(\mathbb{W},\,\cdot\,)$, and is trained to map any potential solution x to the corresponding artifact part. The norm of $\boldsymbol{\Phi}(\mathbb{V}, x)$ used as trained regularizer.

For suitable network training of the encoder-decoder scheme (4.1), we propose the following strategy (compare Figure 4.1). We choose a set of training phantoms $z_n \in X$ for $n = 1, \dots, 2N$ from which we construct back-projection images $x_n := \mathbf{F}^\mathsf{T}(\mathbf{F}z_n)$ for the first N

training examples, and set $x_n = z_n$ for the last N training images. From this we define the training data $\{(x_n, r_n)\}_{n=1}^{2N}$, where

$$r_n = z_n - x_n = z_n - \mathbf{F}^{\mathsf{T}}(\mathbf{F} z_n) \qquad \text{for } n = 1, \dots, N \tag{4.2}$$

$$r_n = z_n - x_n = 0 \qquad \text{for } n = N+1, \dots, 2N\,. \tag{4.3}$$

The free parameters in (4.1) are adjusted in such a way, that $\boldsymbol{\Psi}(\mathbb{W}, \boldsymbol{\Phi}(\mathbb{V}, x_n)) \simeq r_n$ for any training pair (x_n, r_n). This is achieved by minimizing the error function

$$E_N(\mathbb{V}, \mathbb{W}) := \sum_{n=1}^{N} d(\boldsymbol{\Psi}(\mathbb{W}, \boldsymbol{\Phi}(\mathbb{V}, x_n)), r_n)\,, \tag{4.4}$$

where d is a suitable distance measure (or loss function) that quantifies the error made by the network function on the n-th training sample. Typical choices for d are mean absolute error or mean squared error.

Given an arbitrary unknown $x \in X$, the trained network estimates the artifact part. As a consequence, $\mathcal{R}(\mathbb{V}, x)$ is expected to be large, if x contains severe artifacts and small if it is almost artifact free. If x is similar to elements in the training set, this should produce almost artifact free results with NETT regularization. Even if the true unknown is of different type from the training data, artifacts as well as noise will have large value of the regularizer. Thus our approach is applicable for a wider range of images apart from training ones. This claim is confirmed by our numerical results in Section 5. Note that while we did not explicitly account for the coercivity condition in (A1) during the training phase, the encoder-decoder scheme is expected to automatically yield this property. Investigating this issue in more detail (theoretically and numerically) is an interesting aspect of future research.

4.2 Minimizing the NETT functional

Using the encoder-decoder scheme, regularized solutions are defined as minimizers of the NETT functional

$$\mathcal{T}_{\alpha;y}(x) = \frac{1}{2}\|\mathbf{F}(x) - y_\delta\|^2 + \alpha \sum_{\lambda \in \Lambda_L} \|\boldsymbol{\Phi}_\lambda(\mathbb{V}, x)\|_p^p\,, \tag{4.5}$$

where $\boldsymbol{\Phi}$ is trained as above. The optimization problem (4.5) is non-convex and non-smooth if $p = 1$. Note that the subgradient of the regularization term $\mathcal{R}(\mathbb{V}, x) = \sum_{\lambda \in \Lambda_L} \|\boldsymbol{\Phi}(\mathbb{V}, x)_\lambda\|_p^p$ can be evaluated by standard software for network training with the backpropagation algorithm. We therefore propose to use an incremental gradient method for minimizing the Tikhonov functional (4.5), which alternates between a gradient descent step for $\frac{1}{2}\|\mathbf{F}(x) - y_\delta\|^2$ and a subgradient descent step for the regularizer $\mathcal{R}(\mathbb{V}, x)$.

The resulting minimization procedure is summarized in Algorithm 1.

In practice, we found that Algorithm 1 gives favorable performance, and is stable with respect to tuning parameters. Also other algorithms such as proximal gradient methods [9] or Newton type methods might be used for the minimization of (4.5). A detailed comparison with other algorithms is beyond the scope of this article.

Note that the regularizer may be taken $\mathcal{R}(\mathbb{V}, x) = \|\boldsymbol{\Phi}(x)\|_L$ with an arbitrary norm $\|\cdot\|_L$ on $\mathbb{X}_L$. The concrete training procedure is described below. In the form (4.5), NETT constitutes a non-convex generalization of ℓ^q-regularization.

Algorithm 1 Incremental gradient descent for minimizing NETT

Choose family of step-sizes $(s_i) > 0$
Choose initial iterate x_0
for i = 1 to `maxiter` **do**
$\bar{x}_i \leftarrow x_{i-1} - s_i \mathbf{F}'(x_{i-1})^*(\mathbf{F}(x_{i-1}) - y^\delta)$ {gradient step for $\frac{1}{2}\|\mathbf{F}(x) - y_\delta\|^2$}
$x_i \leftarrow \bar{x}_i - s_i \alpha \nabla_x \mathcal{R}(\mathbb{V}, \cdot)(\bar{x}_i)$ {gradient step for $\mathcal{R}(\mathbb{V}, \cdot)$}
end for

5 Application to sparse data tomography

As a demonstration, we use NETT regularization with the encoder-decoder scheme presented in Section 4 to the sparse data problem in photoacoustic tomography (PAT). PAT is an emerging hybrid imaging method based on the conversion of light in sound, and beneficially combines the high contrast of optical imaging with the good resolution of ultrasound tomography (see, for example, [25, 30, 37, 38]).

5.1 Sparse sampling problem in PAT

The aim of PAT is to recover the initial pressure $p_0 \colon \mathbb{R}^d \to \mathbb{R}$ in the wave equation

$$\begin{cases} \partial_t^2 p(x,t) - \Delta_x p(x,t) = 0\,, & \text{for } (x,t) \in \mathbb{R}^d \times (0,\infty)\,, \\ p\,(x,0) = p_0(x)\,, & \text{for } x \in \mathbb{R}^d\,, \\ \partial_t p\,(x,0) = 0\,, & \text{for } x \in \mathbb{R}^d\,. \end{cases} \tag{5.1}$$

form measurements of p made on an observation surface S outside the support of p_0. Here d is the spatial dimension, Δ_x the spatial Laplacian, and ∂_t the derivative with respect to the time variable t. Both cases $d = 2, 3$ for the spatial dimension are relevant in PAT: The case $d = 3$ corresponds to classical point-wise measurements; the case $d = 2$ to integrating line detectors [5, 30]. In this paper we consider the case of $d = 3$ and assume the initial pressure $p_0 \colon \mathbb{R}^2 \to \mathbb{R}$ vanishes outside the unit disc D_1, the ball of radius 1, and that acoustic data are collected at the boundary sphere $\mathbb{S}^1 = \partial D_1$. In particular, we are interested in the sparse sampling case, where data are only given for a small number of sensor locations on $\mathbb{S}^1$. This is the case that one often faces in practical applications.

In the full sampling case, the discrete PAT forward operator is written as $\mathcal{F} \colon \mathbb{R}^{N_1 \times N_2} \to \mathbb{R}^{M_{\text{full}} \times M_2}$ where M_{full} corresponds to the number of complete spatial sampling points and M_2 to the number of temporal sampling points. Sufficient sampling conditions for PAT in the circular geometry have been derived in [20]. We discretize the exact inversion formula of [13] to obtain an approximation $\mathcal{F}^\sharp \colon \mathbb{R}^{M_{\text{full}} \times M_2} \to \mathbb{R}^{N_1 \times N_2}$ to the inverse of $\mathbf{F}$. In the full data case, application of $\mathcal{F}^\sharp$ to data $\mathcal{F}x \in \mathbb{R}^{M_{\text{full}} \times M_2}$ gives an almost artifact free reconstruction $x \in \mathbb{R}^{N_1 \times N_2}$, see [20]. Note that $\mathcal{F}^\sharp$ is the discretization of the continuous adjoint of $\mathcal{F}$ with respect to a weighted L^2-inner product (see [13]).

In the sparse sampling case, the PAT forward operator is given by

$$\mathbf{F} = \mathbf{S} \circ \mathcal{F} \colon X = \mathbb{R}^{N_1 \times N_2} \to \mathbb{R}^{M_1 \times M_2}\,. \tag{5.2}$$

Here $\mathbf{S} \colon \mathbb{R}^{M_{\text{full}} \times M_2} \to \mathbb{R}^{M_1 \times M_2}$ is the subsampling operator, which restricts the full data in to a small number of spatial sampling points. In the case of spatial under-sampling, the

FBP reconstruction $\mathbf{F}^\sharp := \mathcal{F}^\sharp \circ \mathbf{S}^\mathsf{T}$ yields typical streak-like under-sampling artifacts (see, for example, the examples in Figures 5.1 and 5.2 below).

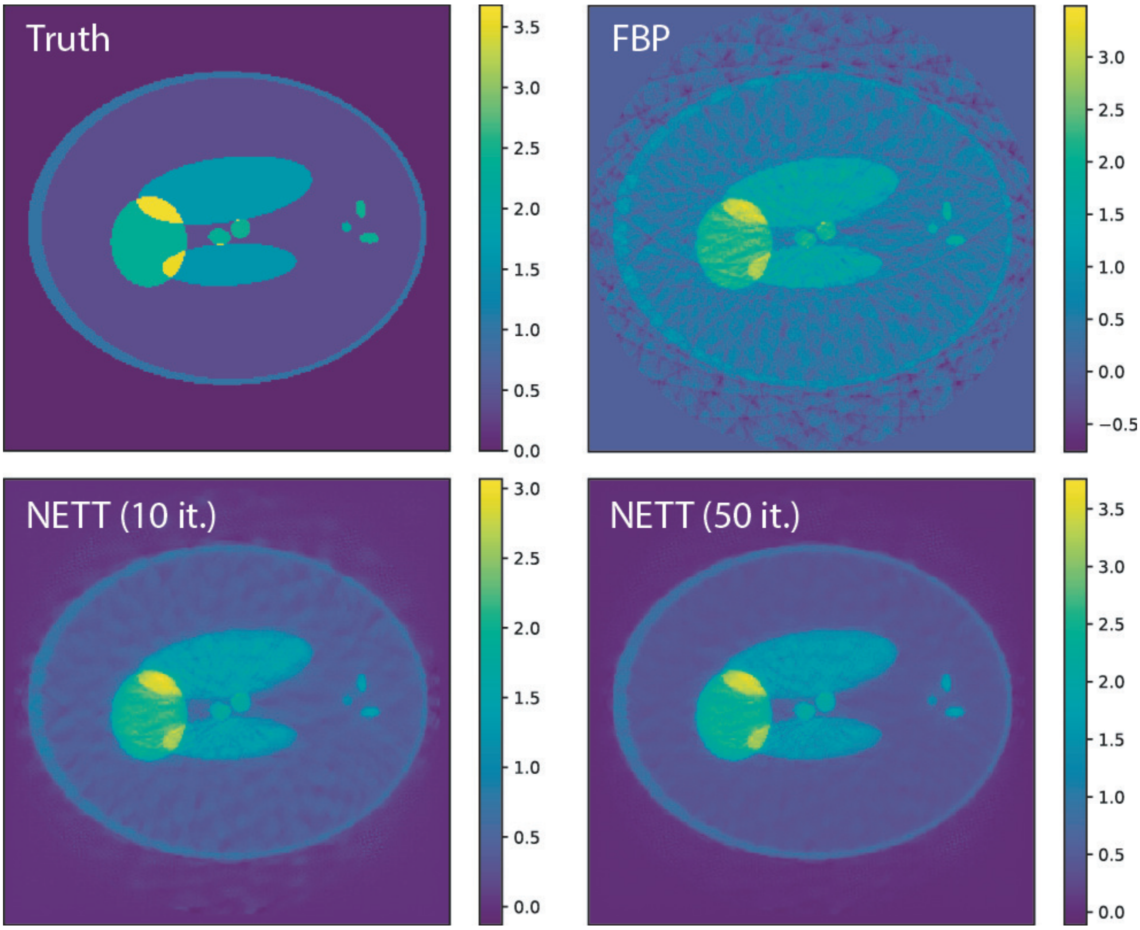

Figure 5.1: Reconstruction results for a phantom of Shepp-Logan type. **Top**: Phantom x (left) and corresponding FBP reconstruction (right); **Bottom**: Iterates x_{10} (left) and x_{50} (right) with the proposed algorithm for minimizing the NETT functional.

5.2 Implementation details

Consider NETT where the regularizer is defined by the encode-decoder framework described in Section 4. The network $\boldsymbol{\Psi}(\mathbb{W}, \cdot) \circ \boldsymbol{\Phi}(\mathbb{V}, \cdot)$ is taken as the Unet, where the $\boldsymbol{\Phi}(\mathbb{V}, x)$ corresponds to the output of the bottom layer with smallest image size and largest depth. The Unet has been proposed in [33] for image segmentation and successfully applied to PAT in [3, 35]. However, we point out, that any network that has the encoder-decoder of the form $\boldsymbol{\Psi}(\mathbb{W}, \cdot) \circ \boldsymbol{\Phi}(\mathbb{V}, \cdot)$ can be used in an analogous manner.

The network was trained on a set of training pairs $\{(x_n, r_n)\}_{n=1}^{2N}$, with $N = 975$, where exactly half of them contained under-sampling artifacts. For generating such training we used (4.2), (4.3) where z_n are taken as randomly generated piecewise constant Shepp-Logan type phantoms. The Shepp-Logan type phantoms have position, angle, shape and intensity of every ellipse chosen uniformly at random under the side constraints that the support of every ellipse lies inside the unit disc and the intensity of the phantom is in the range $[0, 6]$.

In this discrete sparse sampling case, we take the forward operator $\mathbf{F}$ as in (5.2) with $N_1 = N_2 = 256$ and $M_1 = 30$ spatial samples distributed equidistantly on the boundary circle. We used $M_2 = 2000$ times sampled evenly in the interval $[0, 2.5]$. The under-sampling problem in PAT is solved by FBP, and NETT regularization using $\alpha = 1/4$. We minimize (4.5) using Algorithm 1, where we chose a constant step size of $s_i = 0.4$ and take the zero image $x_0 = 0$ for the initial guess.

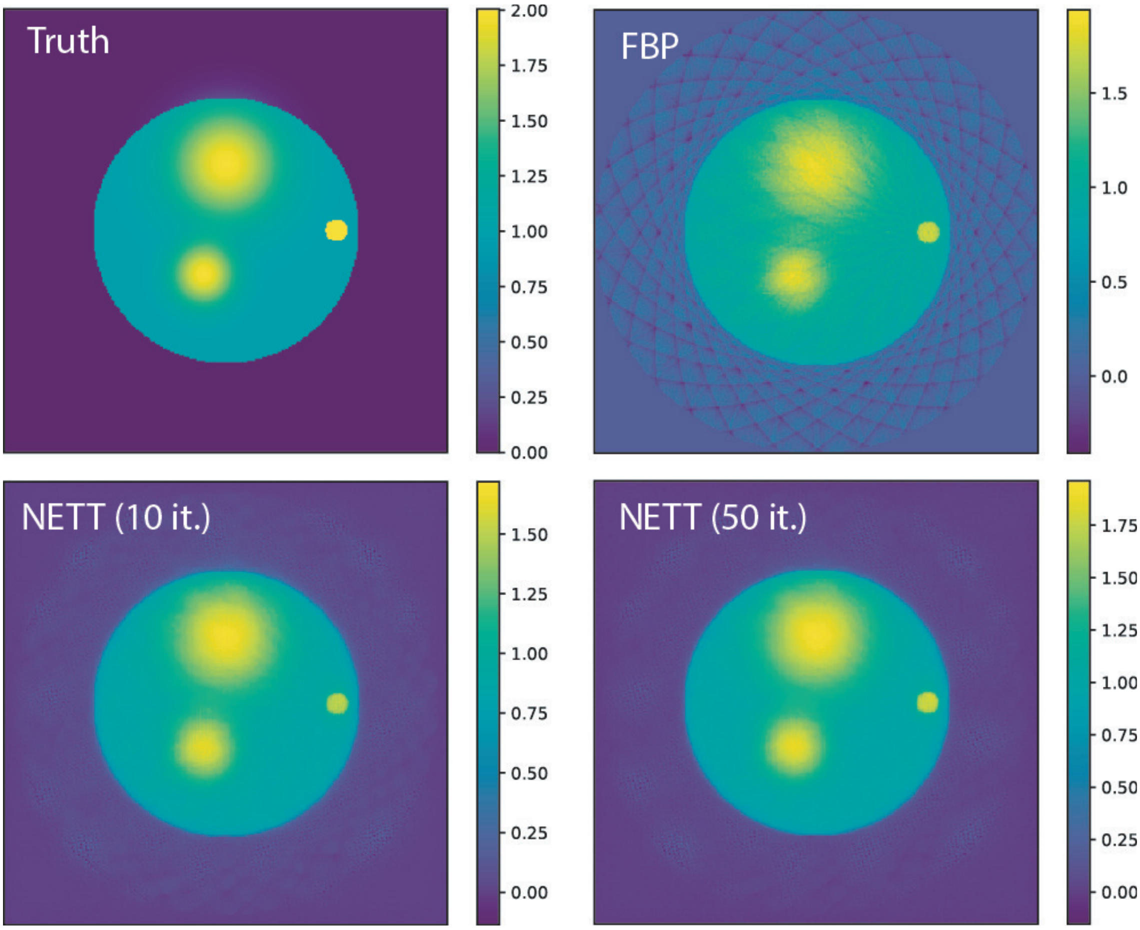

Figure 5.2: Reconstruction results for a phantom of different type from training data. **Top**: Phantom x with smooth blobs (left) and corresponding FBP reconstruction (right); **Bottom**: Iterates x_{10} (left) and x_{50} (right) with the proposed algorithm for minimizing the NETT functional.

5.3 Results and discussion

The top left image in Figure 5.1 shows a Shepp-Logan type phantom $x \colon \mathbb{R}^{256\times256} \to \mathbb{R}$ corresponding to a function on the domain $[-1, 1]^2$. It is of the same type as the training data, but is not contained in the training data. The NETT reconstruction x_{10} and x_{50} with Algorithm 1 after 10 and 50 iterations for the Shepp-Logan type phantom are shown in the bottom row of Figure 5.1. The top right image Figure 5.1 shows the reconstruction $x_{\mathrm{FBP}} = \mathbf{F}^\sharp \mathbf{F} x$ with the FBP algorithm of [13]. The relative L^2-errors $E(z) := \|x - z\|_2 / \|x\|_2$ of the iterates for the Shepp-Logan type phantom are $E(x_{10}) = 0.262$ and $E(x_{50}) = 0.192$, whereas the relative error of the FBP reconstruction is $E(x_{\mathrm{FBP}}) = 0.338$. From Figure 5.1

one recognizes that NETT is able to well remove under-sampling artifacts while preserving high resolution information.

We also consider a phantom image (blobs phantom) that additionally includes smooth parts and is of different type from the phantoms used for training. The blobs phantom as well as the FBP reconstruction $x_{\rm FBP}$ and NETT reconstructions x_{10} and x_{50} are shown in Figure 5.2. For the blobs phantom, the relative reconstruction errors are given by $E(x_{10}) = 0.176$, $E(x_{50}) = 0.102$ and $E(x_{\rm FBP}) = 0.179$. Again, for this phantom different from the training set, NETT removes under-sampling artifacts and at the same time preserves high resolution.

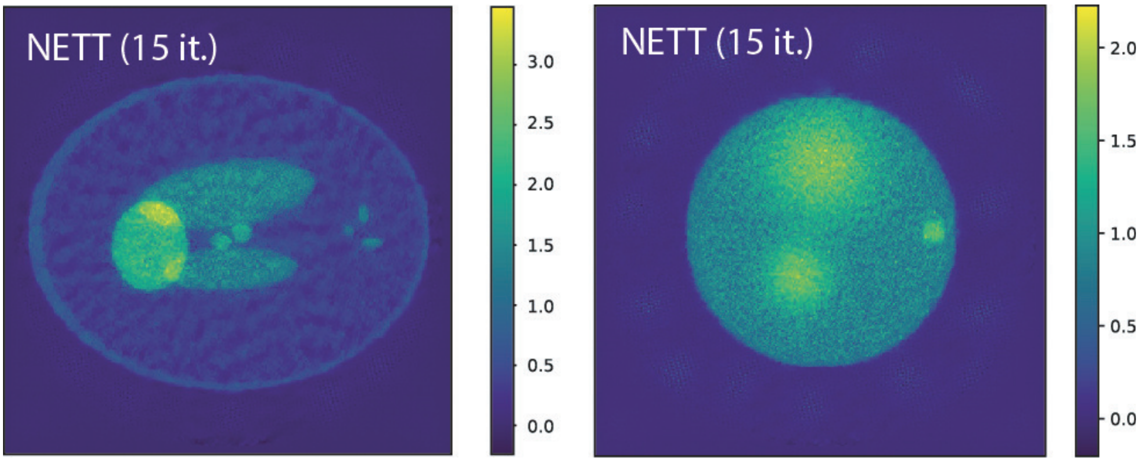

Figure 5.3: Reconstruction results from noisy data using NETT with 15 iterations. **Left**: Shop Logan phantom; **Right**: Blobs phantom.

Finally, Figure 5.3 shows reconstruction results with NETT from noisy data where we added 5 % additive Gaussian noise to the the data. We performed 15 iterations with Algorithm 1. The relative reconstruction errors are $E(x_{15}) = 0.280$ for the Shepp Logan phantom and $E(x_{15}) = 0.210$ for the blobs phantom. Parameters have been taken as in the noiseless data case; no detailed studies for finding optimal values of step site and regularization parameter have been performed. In both cases, the reconstructions for are free from under-sampling artifacts and contain high frequency information, which demonstrates the applicability of NETT for noisy data as well.

The above results demonstrate the proposed NETT regularization using the encoder-decoder framework and with Algorithm 1 for minimization is able to remove under-sampling artifacts. Its gives pretty good results even on images with smooth structures not contained in the training data. This shows that in the NETT framework, learning the regularization functional on one class of training data, can lead to good results even for images beyond that class.

6 Conclusion and outlook

In this paper we developed a new framework for the solution of inverse problems via NETT (1.3). We presented a complete convergence analysis and derived well-posedness and weak convergence (Theorem 2.3), norm-convergence (Theorem 2.9), as well as various convergence rates results (see Section 3). NETT combines deep neural networks with a Tikhonov regularization strategy. The regularizer is defined by a network that might be a user-specified function (generalizing frame based regularization), or might be a CNN trained on a appropriate training data. We have developed a possible strategy for learning a deep CNN (using an encoder-decoder framework, see Section 4). Initial numerical results for a sparse data problem in PAT (see Section 5) demonstrated that NETT with the trained regularizer works well and also yields good results for phantoms different from the class of training data. This may be a result of the fact, that opposed to other deep learning approaches for image reconstruction, the NETT includes a data consistency term as well as the trained network that focuses on identifying artifacts. Detailed comparison with other deep learning methods for inverse problems as well as variational regularization methods (including TV-minimization) is subject of future studies.

Many possible lines of future research arise from the proposed NETT regularization and the corresponding network-minimizing solution concept (1.4). For example, instead of the Tikhonov variant (1.3) one can employ and analyze the residual method (or Ivanov regularization) for approximating (1.4), see [18]. Instead of the simple incremental gradient descent algorithm (cf. Algorithm 1) for minimizing NETT one could investigate different algorithms such as proximal gradient or semi-smooth Newton methods. Studying network designs and training strategies different from the encoder-decoder scheme is a promising aspect of future studies. Finally, application of NETT to other inverse problems is another interesting research direction.

Acknowledgement

SA and MH acknowledge support of the Austrian Science Fund (FWF), project P 30747. The work of HL has been support through SFB 937, subproject A10, and the National Nature Science Foundation of China (61571008).

References

[1] J. Adler and O. Öktem. Solving ill-posed inverse problems using iterative deep neural networks. *Inverse Probl.*, 33(12):124007, 2017.

[2] M. Aharon, M. Elad, and A. Bruckstein. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Trans. Signal Proc.*, 54(11):4311–4322, 2006.

[3] S. Antholzer, M. Haltmeier, and J. Schwab. Deep learning for photoacoustic tomography from sparse data. *arXiv:1704.04587*, 2017.

[4] M. Burger and S. Osher. Convergence rates of convex variational regularization. *Inverse Probl.*, 20(5):1411–1421, 2004.

[5] P. Burgholzer, J. Bauer-Marschallinger, H. Grün, M. Haltmeier, and G. Paltauf. Temporal back-projection algorithms for photoacoustic tomography with integrating line detectors. *Inverse Probl.*, 23(6):S65–S80, 2007.

[6] D. Butnariu and A. N. Iusem. *Totally convex functions for fixed points computation and infinite dimensional optimization*, volume 40 of *Applied Optimization*. Kluwer Academic Publishers, Dordrecht, 2000.

[7] E. J. Candès and D. L. Donoho. New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities. *Comm. Pure Appl. Math.*, 57(2):219–266, 2004.

[8] H. Chen, Y. Zhang, W. Zhang, P. Liao, K. Li, J. Zhou, and G. Wang. Low-dose CT via convolutional neural network. *Biomed. Opt. Express*, 8(2):679–694, 2017.

[9] P. L. Combettes and J.-C. Pesquet. Proximal splitting methods in signal processing. In *Fixed-point algorithms for inverse problems in science and engineering*, pages 185–212. Springer, 2011.

[10] I. Daubechies. Orthonormal bases of compactly supported wavelets. *Comm. Pure Appl. Math.*, 41(7):909–996, 1988.

[11] I. Daubechies, M. Defrise, and C. De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Comm. Pure Appl. Math.*, 57(11):1413–1457, 2004.

[12] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of inverse problems*, volume 375. Kluwer Academic Publishers Group, Dordrecht, 1996.

[13] D. Finch, M. Haltmeier, and Rakesh. Inversion of spherical means and the wave equation in even dimensions. *SIAM J. Appl. Math.*, 68(2):392–412, 2007.

[14] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT Press, 2016. `http://www.deeplearningbook.org`.

[15] M. Grasmair. Generalized Bregman distances and convergence rates for non-convex regularization methods. *Inverse Problems*, 26(11):115014, 16, 2010.

[16] M. Grasmair, M. Haltmeier, and O. Scherzer. Sparse regularization with l^q penalty term. *Inverse Problems*, 24(5):055020, 13, 2008.

[17] M. Grasmair, M. Haltmeier, and O. Scherzer. Necessary and sufficient conditions for linear convergence of ℓ^1-regularization. *Comm. Pure Appl. Math.*, 64(2):161–182, 2011.

[18] M. Grasmair, M. Haltmeier, and O. Scherzer. The residual method for regularizing ill-posed problems. *Appl. Math. Comput.*, 218(6):2693–2710, 2011.

[19] R. Gribonval and K. Schnass. Dictionary identification — sparse matrix-factorization via ℓ_1-minimization. *IEEE Trans. Inf. Theory*, 56(7):3523–3539, 2010.

[20] M. Haltmeier. Sampling conditions for the circular radon transform. *IEEE Trans. Image Process.*, 25(6):2910–2919, 2016.

[21] Y. Han, J. J. Yoo, and J. C. Ye. Deep residual learning for compressed sensing CT reconstruction via persistent homology analysis, 2016. http://arxiv.org/abs/1611.06391.

[22] K. He, X. Zhang, S. Ren, and J. Sun. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proc. ICCV*, pages 1026–1034, 2015.

[23] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser. Deep convolutional neural network for inverse problems in imaging. *IEEE Trans. Image Process.*, 26(9):4509–4522, 2017.

[24] B. Kelly, T. P. Matthews, and M. A. Anastasio. Deep learning-guided image reconstruction from incomplete data. *arXiv:1709.00584*, 2017.

[25] P. Kuchment and L. A. Kunyansky. Mathematics of thermoacoustic and photoacoustic tomography. *Eur. J. Appl. Math.*, 19:191–224, 2008.

[26] D. Lorenz. Convergence rates and source conditions for Tikhonov regularization with sparsity constraints. *J. Inverse Ill-Posed Probl.*, 16(5):463–478, 2008.

[27] A. L. Maas, A. Y. Hannun, and A. Y. Ng. Rectifier nonlinearities improve neural network acoustic models. In *Proc. ICML*, volume 30, 2013.

[28] S. Mallat. *A wavelet tour of signal processing: The sparse way*. Elsevier/Academic Press, Amsterdam, third edition, 2009.

[29] F. Natterer and F. Wübbeling. *Mathematical Methods in Image Reconstruction*, volume 5 of *Monographs on Mathematical Modeling and Computation*. SIAM, Philadelphia, PA, 2001.

[30] G. Paltauf, R. Nuster, M. Haltmeier, and P. Burgholzer. Photoacoustic tomography using a Mach-Zehnder interferometer as an acoustic line detector. *Appl. Opt.*, 46(16):3352–3358, 2007.

[31] R. Ramlau and G. Teschke. A Tikhonov-based projection iteration for nonlinear ill-posed problems with sparsity constraints. *Num. Math.*, 104(2):177–203, 2006.

[32] E. Resmerita. On total convexity, Bregman projections and stability in Banach spaces. *J. Convex Anal.*, 11(1):1–16, 2004.

[33] O. Ronneberge, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 234–241, 2015.

[34] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen. *Variational methods in imaging*, volume 167 of *Applied Mathematical Sciences*. Springer, New York, 2009.

[35] J. Schwab, S. Antholzer, R. Nuster, and M. Haltmeier. DALNet: High-resolution photoacoustic projection imaging using deep learning. *arXiv:1801.06693*, 2018.

[36] G. Wang. A perspective on deep imaging. *IEEE Access*, 4:8914–8924, 2016.

[37] K. Wang and M. A. Anastasio. Photoacoustic and thermoacoustic tomography: image formation principles. In *Handbook of Mathematical Methods in Imaging*, pages 781–815. Springer, 2011.

[38] L. V. Wang. Multiscale photoacoustic microscopy and computed tomography. *Nature Phot.*, 3(9):503–509, 2009.

[39] S. Wang, Z. Su, L. Ying, X. Peng, S. Zhu, F. Liang, D. Feng, and D. Liang. Accelerating magnetic resonance imaging via deep learning. In *IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, pages 514–517, 2016.

[40] T. Würfl, F. C. Ghesu, V. Christlein, and A. Maier. Deep learning computed tomography. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 432–440. Springer, 2016.

[41] H Zhang, L. Li, K. Qiao, L. Wang, B. Yan, L. Li, and G. Hu. Image prediction for limited-angle tomography via deep learning with convolutional neural network. arXiv:1607.08707, 2016.