Automatic and efficient tomographic reconstruction algorithms

Lagerwerf, M.J.

Chapter 2

An interpolation approach for determining regularization parameters

2.1 Introduction

Tomography is a generic 3D imaging technique for reconstructing the interior of an object from a series of its projections. Projections can be acquired using a broad variety of modalities, such as Computed Tomography (CT) [Nat01], Magnetic Resonance Imaging (MRI) [Fes10] and Electron Microscopy (EM) [Mid+01]. The resulting image reconstruction problems all have a similar mathematical problem structure: given a set of tomographic measurements and a description of the physics process, determine a reconstruction of the measured object. If many projections are available over a full angular range around the object, and if the projections have low noise, the reconstruction problem can be solved in a straightforward way by closed-form inversion techniques. For an overview see [Nat01; KS01]. In practice, however, the number of tomographic measurements is typically limited and the measurements can contain substantial noise.

This chapter is based on:

An Efficient Interpolation Approach for Exploring the Parameter Space of Regularized Tomography Algorithms. *MJ Lagerwerf, WJ Palenstijn, F Bleichrodt, KJ Batenburg.* Fundamenta Informaticae (Volume: 172), number 2, pp. 143–167, 2020.

In such limited data cases the information from the measurements and the geometry of the acquisition setup is not sufficient to solve the reconstruction problem accurately and some form of prior knowledge about the object must be incorporated in the solution process. One way of incorporating such prior knowledge in the reconstruction method is the use of regularization, where the balance between the prior knowledge and solving the original problem is usually determined by a regularization parameter [ROF92; BKP10; Goc16].

The choice for this regularization parameter depends on many properties of the problem, such as the measured data and its noise level, or the reconstruction method and its implementation. Moreover, the desired choice of regularization parameter may also be application-specific, for instance with the aim of reconstructing particular image features as sharply as possible at acceptable noise levels, creating a reconstruction that is well-suited for subsequent segmentation, etc.

Consider for example the Total Variation (TV) reconstruction algorithm implemented with a Primal-Dual Hybrid Gradient (PDHG) algorithm initially proposed by Chambolle, Pock, Bischof and Cremers, as described in [SJP12]. Given all the information about the reconstruction algorithm and its implementation, the effect of a particular choice of the regularization parameter on the reconstructed image still varies significantly for different instances of the reconstruction problems. If it is not possible to specifically define and model additional information about the object, it is common that the algorithm user computes a series of reconstructions for different choices of the regularization parameter and chooses the preferred setting by visual inspection. Although there have been strategies proposed to handle such parameter space exploration in a structural manner, such as described in [Sed+14; Pre+11], these strategies do not specifically address the key drawbacks we encounter here. The key drawbacks of such a trial-and-error method are twofold: (i) computing many reconstructions for different regularization parameter values is computationally intensive, as even computing a single reconstruction can already be computationally demanding; (ii) with only a small number of reconstruction evaluations, it is difficult to choose a regularization parameter in a consistent manner as the actual desired value may lie somewhere in between the sampling points.

For variational methods, of which TV is a well known example, there has been extensive theoretical work on how to determine the “optimal” regularization parameter, where each approach has a different concept of optimality. For example, for the Tikhonov method, explicit analytic expressions are found based on the singular value decomposition and discrepancy principle, see chapter 3 and 7 of [Sch+09], and [Vai82]. Moreover, substantial analytical results have been obtained on how the properties of a reconstruction change depending on the value of the regularization parameters [Bur+13; Bur+16; Bri+18]. Although these results are powerful, they also make strong assumptions on the reconstruction algorithm, such as continuity of the solution with respect to the regularization parameter, full convergence of the iterative algorithm, or the availability of certain prior knowledge about the problem such as the noise level, which are not always valid in practice. Other more general strategies, such as the discrepancy principle and the L-curve criterion [Vai82; BM12; Han92; HO93; Han99], have been developed. These methods also rely on a specific definition of the "optimal" reconstruction and require many evaluations of the reconstruction algorithm. A key limitation of all mentioned approaches is that they do not take the application-specific needs into account. The criterion of optimality is based on a mathematical problem formulation without involving the particular requirements of the user.

In this chapter we propose an algorithmic approach for computationally efficient exploration of the regularization parameter space. Once a relatively small number of reconstructions have been computed for a sparse sampling of the regularization parameters, an approximation of the reconstructed image for other parameter values can be computed with very high efficiency (linear time in the number of pixels). In the case of manual selection of the regularization parameter, our approach makes it possible to present the user with a real-time interface where parameters can be adjusted on-the-fly and immediate visual feedback is obtained on the effect of parameter changes on the reconstruction. In the case of automated selection, the output of our approximation method can be used as input for any image-based quality metric that one wants to optimize for.

Accurate approximation of the output of general regularization reconstruction methods is a difficult problem. However, we found that if the output of the reconstruction algorithm is available for just a small number of regularization parameter values, a pixel-wise interpolation scheme is highly suitable for such approximations. The choice of the sampling scheme is of particular importance to the effectiveness of our approach. Through computational experiments we found that although the major changes of the reconstructed image with respect to the regularization occur in a relatively narrow region of the space of regularization parameter values, a logarithmic sampling and corresponding interpolation scheme results in relatively smooth behavior of the pixel values with respect to the regularization parameter choice.

Our experimental results on simulated data for the parallel beam tomography problem demonstrate that for three common variational reconstruction methods, our approach results in accurate approximations of the reconstructed image and that it can be used in combination with existing approaches for choosing optimal regularization parameters. We also provide results for an experimental X-ray CT dataset. The approach is presented in such a manner that it can easily be extended to different modalities and reconstruction methods.

This chapter is structured as follows. In **Section 2.2** we introduce the problem, related notation and mathematical descriptions of the methods used in this chapter. In **Section 2.3** we discuss our proposed method and how it can be used in existing methods. Details about the implementation and experiments are discussed in **Section 2.4**. The results are shown in **Section 2.5** and in **Section 2.6** we summarize and conclude the chapter.

2.2 Notation and mathematical preliminaries

2.2.1 Problem Definition

We focus here on the two dimensional (2D) parallel beam tomography problem, which we define below, and three different types of variational methods; see **Section 2.2.2**. Our approach can be used for other tomography geometries (both 2D and 3D) and other reconstruction methods in a straightforward manner.

The 2D parallel beam tomography problem entails reconstructing a two-dimensional unknown object from its parallel beam projection data. We will consider the discrete version of this problem:

$$W\mathbf{x} = \mathbf{y}, \tag{2.1}$$

with $\mathbf{x} \in \mathbb{R}^{N^2}$ the unknown object, defined on a $N \times N$ pixel grid; $\mathbf{y} \in \mathbb{R}^{N_\theta N_d}$ the parallel beam projection data, N_θ the number of projection angles, N_d the number of detector pixels, and $W : \mathbb{R}^{N^2} \to \mathbb{R}^{N_\theta N_d}$ a discrete version of the Radon transform. A more in depth description of this problem can be found in [Nat01; KS01].

We define a reconstruction method $F : \mathbb{R}^{N_\theta N_d} \times \mathbb{R}^{N_\lambda} \to \mathbb{R}^{N^2}$ with N_λ real-valued regularization parameters for the problem (2.1) as a type of black-box operator:

$$F(\mathbf{y}, \lambda) = \mathbf{x}_F^\lambda, \tag{2.2}$$

with $\lambda \in \mathbb{R}^{N_\lambda}$ a vector containing all the regularization parameters λ_i with $i = 0, \ldots, N_\lambda - 1$.

This definition fits general variational methods that incorporate regularization as discussed in Section 2.2.2, but also fits the well-known Filtered Backprojection (FBP) algorithm [Nat01], where bandwidth of a low-pass filter can be considered as the regularization parameter, and the Simultaneous Iterative Reconstruction Technique (SIRT) [VV90] with the number of iterations as a regularization parameter. As we consider F as a black-box operator, we will only make use of the result of applying F to the projection data for different values of λ, but will not make use of specific properties of the operator F.

The key contribution of this chapter is to propose a computationally efficient approach for approximating $F(\mathbf{y}, \lambda)$ for many values of λ. Having this ability, it provides a way to choose the "optimal" value of the regularization parameter with respect to any user-defined quality criterion: '*Determine* $\lambda^\star$ *such that* $\mathbf{x}_F^{\lambda^\star}$ *is the optimal solution to* (2.1).' We do not specify here what an optimal solution is, because this varies per problem, application or even user of the reconstruction method.

2.2.2 Variational methods

In this section we discuss the choices for reconstruction methods F that we consider in this chapter. The methods we consider are all variational methods and instead of solving (2.1) directly these methods solve a related minimization problem. The following problem formulation is specifically for one regularization parameter:

$$\mathbf{x}_\lambda^\star = \underset{\mathbf{x}\in\mathbb{R}^{N^2}}{\operatorname{argmin}} \{\mathscr{D}(W\mathbf{x}, \mathbf{y}) + \lambda R(\mathbf{x})\}, \tag{2.3}$$

where $\mathscr{D}$ is the *data fidelity*, R is the *regularization* term or the *regularizer* and $\lambda \in \mathbb{R}$ the *regularization parameter*.

The data fidelity term encodes the information of the original problem (2.1). It determines the distance between the input data and the solution $\mathbf{x}$. In this chapter we will only consider the least squares norm as data fidelity:

$$\mathscr{D}(W\mathbf{x}, \mathbf{y}) = \tfrac{1}{2} \parallel W\mathbf{x} - \mathbf{y} \parallel_2^2 . \tag{2.4}$$

The prior knowledge for our inverse problem is encoded in the regularization term. The idea is to define a functional that is small when the image $\mathbf{x}$ has a certain preferred property, such as smoothness or sparsity with respect to a certain set of basis functions.

We consider three types of regularizers in this chapter:

Sobolev and Total Variation regularization These regularizers penalize the gradient of the reconstruction $\mathbf{x}$ in the L^2—sense and L^1—sense, respectively. In mathematical terms:

$$R_{\mathrm{S}}(\mathbf{x}) = \|\nabla\mathbf{x}\|_2^2, \qquad R_{\mathrm{TV}}(\mathbf{x}) = \|\nabla\mathbf{x}\|_1 . \tag{2.5}$$

Note that the resulting minimization problem (2.3) now has a single scalar value $\lambda \in \mathbb{R}$ as the regularization parameter. For a more in depth discussion on these regularizers see [Goc16; ROF92].

Total Generalized Variation regularization The idea for Total Generalized Variation is that one can split the reconstruction into parts with a different order of regularity. In this chapter we will consider the version which splits it into two parts: TGV_λ^2. The priority between these parts is balanced by a minimization problem with a second regularization parameter:

$$R_{\mathrm{TGV}}(\mathbf{x},\lambda)=\operatorname*{argmin}_{\mathbf{v}\in\mathbb{R}^{2\times N^2}}\{\|\nabla\mathbf{x}-\mathbf{v}\|_1+\lambda\,\|\mathscr{E}\mathbf{v}\|_1\}\,, \tag{2.6}$$

here $\mathscr{E}:\mathbb{R}^{2\times N^2}\to\mathbb{R}^{4\times N^2}$ is the distributional symmetrized derivative.

The minimization problem (2.3) in this case has two regularization parameters and two objects to minimize for:

$$\mathbf{x}_\lambda^\star,\mathbf{v}_\lambda^\star=\operatorname*{argmin}_{\mathbf{x}\in\mathbb{R}^{N^2},\mathbf{v}\in\mathbb{R}^{2\times N^2}}\left\{\tfrac{1}{2}\,\|\,W\mathbf{x}-\mathbf{y}\,\|_2^2+\lambda_1\left(\|\nabla\mathbf{x}-\mathbf{v}\|_1+\lambda_2\,\|\mathscr{E}\mathbf{v}\|_1\right)\right\}, \tag{2.7}$$

with $\lambda=(\lambda_1,\lambda_2)$. For a more in-depth discussion on this regularizer see [BKP10].

At this point we have only described the mathematical functions to be minimized for a particular variational method. We use a Primal-Dual Hybrid Gradient (PDHG) method presented in [CP11]. Note that such an algorithm will typically be terminated before the solution has fully converged, and will therefore not reach the exact solution $\mathbf{x}_\lambda^\star$. Taking this all into consideration we define our black-box reconstruction method as follows:

$$F(\mathbf{y},\lambda)=\mathbf{x}^{\lambda}_{\mathrm{PDHG}_{\mathrm{VM}}}, \tag{2.8}$$

with the variational method VM $\in$ {S, TV, TGV}.

Details about the implementation and parameter choices for the PDHG algorithm are discussed in **Section 2.4**.

2.2.3 Parameter optimization methods

Two well known methods for choosing the regularization parameter for single parameter variational methods are the discrepancy principle and the L-curve criterion, which we will briefly introduce here.

Discrepancy principle For the discrepancy principle one chooses the maximal regularization parameter λ such that the projection data $\mathbf{y}$ satisfies:

$$\|\,WF(\mathbf{y},\lambda)-\mathbf{y}\,\|_2^2\le\epsilon, \tag{2.9}$$

where the specific norm is the same as the one used in the data fidelity and ϵ is a parameter that corresponds to the expected noise level in the projection data. For more information see [Vai82; BM12].

L-curve criterion The L-curve criterion chooses the regularization parameter $\lambda^\star$ such that the log-log curve of the data fidelity and the regularizer has a maximum curvature. We define $\rho(\lambda)$ and $\eta(\lambda)$ as the logarithm of the data fidelity and the regularizer, respectively. The log-log curve can then be described as follows:

$$(\rho(\lambda), \eta(\lambda)) = (\log(\mathscr{D}(W\mathbf{x}_\lambda, \mathbf{y})), \log(R(\mathbf{x}_\lambda))), \tag{2.10}$$

Using these expressions we define the $\lambda^\star$ for the L-curve criterion as the λ for which the curvature κ attains its maximum:

$$\lambda^\star = \mathrm{argmax}_\lambda \{\kappa(\lambda)\} = \mathrm{argmax}_\lambda \left\{ \frac{\eta''\rho' - \rho''\eta'}{((\rho')^2 + (\eta')^2)^{3/2}} \right\}. \tag{2.11}$$

The idea is that for any other λ the relative increase of the data fidelity is higher than the relative decrease in the regularization term and vice versa. Therefore, this $\lambda^\star$ provides an "optimal" balance between the two terms. A more in depth discussion with theoretical analysis is given in [Han92; HO93; Han99].

We point out that both the discrepancy principle and the L-curve criterion implicitly assume that the reconstruction method only has one regularization parameter. Moreover, as will also become clear from our simulation results, the regularization parameter values obtained using both methods can be substantially different. No application-specific properties of the reconstructed image are taken into account, and these methods may not yield optimal results if the algorithm user wants to optimize for a particular type of image quality metric.

2.3 Method description

In this section we propose our method for efficiently computing a large number of reconstructions of a given set of measurements $\mathbf{y}$ while varying the regularization parameter. The idea is to evaluate the reconstruction method on a coarse regularization parameter grid and interpolate between these reconstructions in such a way that the resulting interpolation scheme is computationally efficient and accurately approximates the actual reconstructed image for a broad range of parameter values.

2.3.1 Interpolation scheme

Approximating the output of general tomographic reconstruction methods by interpolation is in general a challenging problem. However, if we fix the measurements $\mathbf{y}$ and consider the behavior of one pixel in the reconstructed image for varying λ, we observe that for many relevant reconstruction algorithms, including the variational methods described in Section 2.2.2, the pixel value changes smoothly with the value of λ and the curve seems to be suitable for approximation using a relatively simple model. Therefore we propose to perform pixel-wise interpolation using cubic spline interpolation. The idea is that if the approximation per pixel is sufficiently accurate, combining these single-pixel approximations into an approximation of the reconstructed image will be an accurate approximation of the true reconstructed image for a given value of λ.

First of all let us consider the case with one regularization parameter, $N_\lambda = 1$, and define a coarse grid of N_{ip} interpolation points:

$$\Lambda = \{\lambda_0, \lambda_1, ..., \lambda_{N_{\mathrm{ip}}-1}\}, \qquad \text{with,} \qquad 0 < \lambda_0 < \lambda_1 <, ..., < \lambda_{N_{\mathrm{ip}}-1}, \tag{2.12}$$

for this regularization parameter. We evaluate the reconstruction method $F(\mathbf{y}, \lambda_i) = \mathbf{x}_F^{\lambda_i}$. An important observation is that the range of possible choices for λ can be very large (e.g. between 0 and 10^5), while the actual range where the interesting changes of pixel values take place is usually much narrower. Therefore a linear spacing between the interpolation points $\{\lambda_0, \lambda_1, ..., \lambda_{N_{\mathrm{ip}}-1}\}$ does not cover the transitions of pixel values well. Instead, we found that choosing a scheme where the values $\log \lambda_i$ are equidistantly sampled results in more accurate capturing of the transitions[1]. This requires, however, that also the interpolation between the sampling points respects this logarithmic scale. Specifically: given a set of regularization parameters with corresponding reconstructions from the reconstruction method F, $\{(\log(\lambda_i), \mathbf{x}_F^{\lambda_i})\}_{i=0}^{N_{\mathrm{ip}}-1}$, we can compute the cubic spline interpolation $S_p : \mathbb{R} \to \mathbb{R}$ for a pixel p,

$$S_p(\lambda) = S_p^i(\lambda), \qquad \text{with } \lambda_{i-1} \leq \lambda \leq \lambda_i, i = 1, ..., N_{\mathrm{ip}-1}, \tag{2.13}$$

such that the following statements are satisfied:

$$S_p^i(\lambda) = a_i + b_i \log(\lambda) + c_i \log(\lambda)^2 + d_i \log(\lambda)^3, \qquad \text{with } d_i \neq 0, \tag{2.14}$$

$$S_p(\lambda) = (\mathbf{x}_F^{\lambda})_p, \qquad \text{with } \lambda \in \Lambda. \tag{2.15}$$

These conditions are not sufficient to uniquely compute $S_p(\lambda)$, so we also need boundary conditions. Let us consider the regularization parameter interval Λ broad enough, such that $\mathbf{x}_F^{\lambda_0}$ will be under-regularized and $\mathbf{x}_F^{\lambda_{N_{\mathrm{ip}}-1}}$ will be over-regularized. This means that taking a lower or higher λ, respectively, will not result in significant changes to the reconstruction. Therefore, if we assume the regularization parameter interval broad enough, *clamped* boundary conditions will

[1]Further discussion on how to determine the grid Λ is given in **Section 2.4.1**.

be satisfied on the left and right boundaries, *i.e.*,

$$\frac{dS_p(\lambda)}{d\lambda}(\lambda_0) = 0, \qquad \frac{dS_p(\lambda)}{d\lambda}(\lambda_{N_{\text{ip}}-1}) = 0. \tag{2.16}$$

If all the pixel-wise spline interpolations $S_p(\lambda)$ are computed, we can consider the cubic spline interpolation function $S_F : \mathbb{R} \to \mathbb{R}^{N^2}$ for the full object and reconstruction method F:

$$S_F(\lambda) = \begin{bmatrix} S_0(\lambda) & \cdots & S_{N-1}(\lambda) \\ \vdots & \ddots & \vdots \\ S_{(N-1)N}(\lambda) & \cdots & S_{N^2-1}(\lambda) \end{bmatrix} \qquad \text{with } \lambda \in [\lambda_0, \lambda_{N_{\text{ip}}-1}]. \tag{2.17}$$

To summarize, our proposed method for one regularization parameter λ is step-by-step described in **Algorithm 1**.

Algorithm 1 Pixel-wise spline interpolation

1: Determine λ_0 and $\lambda_{N_{\text{ip}}-1}$, s.t. $\mathbf{x}_F^{\lambda_0}$ and $\mathbf{x}_F^{\lambda_{N_{\text{ip}}-1}}$, are under- and over-regularized, respectively.
2: Define a coarse grid Λ on $[\lambda_0, \lambda_{N_{\text{ip}}-1}]$, s.t. $\log(\lambda_i)$ are equidistantly spaced.
3: **for** $i = \{0, 1, ..., N_{\text{ip}} - 1\}$ **do**
4:     Compute $F(\lambda_i, \mathbf{y}) = \mathbf{x}_F^{\lambda_i}$.
5: **for** $p = \{0, 1, .., N^2 - 1\}$ **do**
6:     Compute the spline interpolation $S_p(\lambda)$ for pixel p such that (2.14), (2.15) and (2.16) are satisfied.
7: Combine the $S_p(\lambda)$, as described in (2.17), to get the spline interpolation function $S_F(\lambda)$.

The proposed method can easily be extended to two regularization parameters. In this case the pixel-wise interpolation becomes a two-dimensional interpolation problem. This means that the coarse grid of regularization parameters Λ is also a two-dimensional grid for which $N_{\text{ip},1}N_{\text{ip},2}$ evaluations of the reconstruction method are needed, increasing the computational effort significantly.

2.3.2 Optimizing the regularization parameter

Once the reconstruction algorithm F has been evaluated on the coarse grid of interpolation points, we can sample the approximations of the reconstructions from the interpolation function $S_F(\lambda)$ and use these to determine the optimal regularization parameter $\lambda^\star$ according to the specific requirements of the algorithm, user, and the application. Here we discuss four strategies for using our interpolation technique to optimize the regularization parameter.

Parameter space exploration When there is no reference image available we simply determine the "visually optimal" $\lambda^\star$ through inspection or exploration of the approximations given by the interpolation function $S_F(\lambda)$. A good general framework for strategies such as this is described in [Sed+14], here one can replace the sampling of the original function with the interpolation function $S_F(\lambda)$. Additionally, one can embed the interpolation scheme in a visual tool where the user can adjust the regularization parameter on-the-fly and receive immediate visual feedback on the approximation of the resulting reconstruction.

Quantitative measure optimization In this strategy the approximations $S_F(\lambda)$ are compared to a ground truth or a high quality reconstruction with respect to a certain quantitative measure. We can define the "optimal" regularization parameter $\lambda^\star$ as follows:

$$\lambda^\star = \underset{\lambda}{\operatorname{argmin}} \{\mathrm{QM}(S_F(\lambda), \mathbf{x}_{\mathrm{ref}})\}, \tag{2.18}$$

with $\mathrm{QM} : \mathbb{R}^{N^2} \to \mathbb{R}$, a quantitative measure on the reconstruction space. A simple example of such a function is the root Mean Squared Error (rMSE). We will discuss quantitative measures further in **Section 2.4.4**.

Discrepancy principle This strategy assumes that an estimate ϵ of the noise level on the projection data is available. We define our $\lambda^\star$ as the largest λ for which $\|WS_F(\lambda) - \mathbf{y}\|_2^2 \leq \epsilon$ is true.

L-curve criterion We can compute the L-curve by plugging in the spline interpolation function $S_F(\lambda)$ in (2.10):

$$(\rho(\lambda), \eta(\lambda)) = (\log(\mathscr{D}(S_F(\lambda), \mathbf{y}), \log(R(S_F))). \tag{2.19}$$

To compute the curvature $\kappa(\lambda)$ we use numerical approximations for the gradient based on spline interpolation.

2.4 Experiments

2.4.1 Implementation

Code All methods are implemented using Python 3.6.5, Numpy 1.14.3 [WCV11], SciPy 1.1.0 [JOP+01], ODL [AKÖ17]. The tomography operators are implemented on the GPU using the ASTRA-toolbox [Van+16], where for performance reasons the forward and backprojection are not exactly each other's adjoint. The code used for this chapter is available on GitHub [Lagc].

Reconstruction method setup To normalize the range of the regularization parameters we scale the data fidelity term and the regularization term in (2.3) to the same range. Without loss of generality we set $\lambda = \hat{\lambda}\frac{\|W\|}{\|\nabla\|}$. For Sobolev regularization this gives:

$$\mathbf{x}^\star_\lambda = \underset{\mathbf{x}\in\mathbb{R}^{N^2}}{\operatorname{argmin}} \left\{ \frac{1}{2} \left\| W\mathbf{x} - \mathbf{y}_2^2 \right\| + \hat{\lambda}\frac{\|W\|}{\|\nabla\|} \|\nabla \mathbf{x}\|_2^2 \right\}, \tag{2.20}$$

$$= \underset{\mathbf{x}\in\mathbb{R}^{N^2}}{\operatorname{argmin}} \left\{ \frac{1}{2} \frac{\left\| W\mathbf{x} - \mathbf{y}_2^2 \right\|}{\|W\|} + \hat{\lambda}\frac{\|\nabla \mathbf{x}\|_2^2}{\|\nabla\|} \right\}, \tag{2.21}$$

with $\|\cdot\|$ the operator norm. Similar scaling can be done for TV and TGV regularization. Further reference to the regularization parameter will be to this normalized parameter $\hat{\lambda}$.

As stated before we use the PDHG algorithm to compute the reconstructions. For this algorithm we need to set the step-size parameters τ and σ, the relaxation parameter θ and the number of iterations. To ensure convergence we must have, $\tau\sigma \|L_{\mathrm{VM}}\|^2 < 1$, therefore we take the step-size parameters as follows:

$$\tau = \frac{0.1}{\|L_{\mathrm{VM}}\|}, \qquad \sigma = \frac{0.99}{\tau \|L_{\mathrm{VM}}\|^2}, \tag{2.22}$$

with L_{VM} the operator related to the PDHG implementation of the variational method VM, more specifically:

$$L_{\mathrm{S}} = L_{\mathrm{TV}} = \begin{bmatrix} W \\ \nabla \end{bmatrix}, \qquad L_{\mathrm{TGV}} = \begin{bmatrix} W & 0 \\ \nabla & -I \\ 0 & \mathcal{E} \end{bmatrix}. \tag{2.23}$$

Lastly, we set the relaxation parameter, $\theta = 1$, and the number of iterations, $n_{\mathrm{iter}} = 500$, if not mentioned otherwise.

Parameter grid choice To assess the accuracy of our approximations, we will compute reconstructions for all the regularization parameters on a fine grid Λ^f. For our interpolation method we use a coarse grid $\Lambda \subset \Lambda^f$.

For Λ^f we take N_s logarithmically sampled points on the interval $[\lambda_0, \lambda_{N_{\text{ip}}-1}]$ containing the endpoints, such that we have

$$\log(\lambda_i) = \frac{\log(\lambda_{N_{\text{ip}}-1}) - \log(\lambda_0)}{N_s - 1} i + \log(\lambda_0). \tag{2.24}$$

for $\lambda_i \in \Lambda^f$ and $\lambda_{N_s-1} = \lambda_{N_{\text{ip}}-1}$.

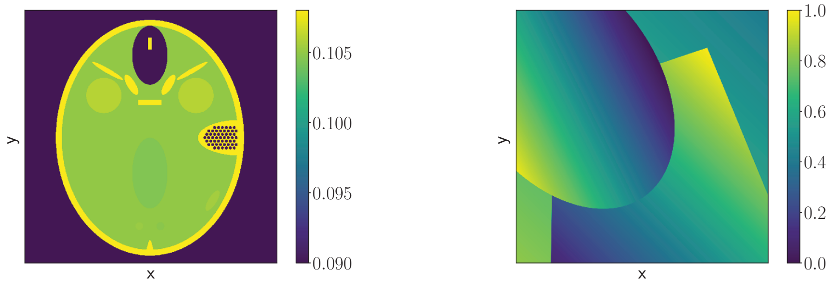
By fixing the number of points N_{ip} we use for the interpolation, we get Λ:

$$\Lambda = \left\{\lambda_j \in \Lambda^f \mid j = k\lceil \tfrac{N_s}{N_{\text{ip}}-1} \rceil, k = 0, .., N_{\text{ip}} - 2\right\} \cup \left\{\lambda_{N_s-1}\right\}, \tag{2.25}$$

which means that the first $N_{\text{ip}} - 1$ points are chosen such that the exponents are equidistant, and the last point coincides with the endpoint.

2.4.2 Computer simulated data

We consider two computer simulated phantoms, shown in **Figure 2.1**, to test the performance of our method. These phantoms are defined independent of a pixel grid. The reconstructions are defined on a 1024×1024 uniform pixel grid. The projection data is defined as 2048 detector elements per projection angle. To avoid the so-called inverse crime the projection data is generated using 2048×2048 phantoms. The resulting projection data with 4096 detector elements per projection angle is rebinned to 2048 detector elements by taking the average of two neighboring elements.



(a) FORBILD head phantom (b) Gradient objects phantom

Figure 2.1: Computer simulated phantoms. The FORBILD head phantom is presented in [LB]. Note that the range of the figure is not the actual range of the phantom; this is to visualize the low contrast objects. The gradient objects phantom is a standard phantom in the ODL package [AKÖ17].

We will consider zero mean additive Gaussian noise with the variance equal to a percentage of the maximum value of the projection data, *i.e.*,

$$\mathbf{y} = \mathbf{y}_{\mathrm{GT}} + \delta, \qquad \delta \sim \mathcal{N}(0, V_\delta), \tag{2.26}$$

with $V_\delta = n_l \cdot \max_i \{\mathbf{y}_i\}$ the variance of the noise and $n_l \geq 0$ the noise level.

2.4.3 Experimental data

The experimental data is acquired from a low-dose scan of a pomegranate. The original scan is a 3D circular cone-beam CT scan of which we took the central detector row for all projection angles, to get a 2D circular fan-beam reconstruction problem. A detector row is 145.34 mm long and contains 1536 detector elements per projection angle and the dataset contains 500 equiangular spaced projection angles. The scans were done using the custom-built and highly flexible FleX-ray CT scanner, developed by XRE NV and located at CWI. Additionally, we use a high-dose scan of the same pomegranate with 2000 equiangular spaced projection angles, from which we compute a *gold standard reconstruction*[2], $\mathbf{x}_{GS}$, that can be used as a reference reconstruction $\mathbf{x}_{\mathrm{ref}}$. Further details about the original scans can be found here [CLB18].

2.4.4 Quantitative measures

To test the accuracy of the approximations of our method compared to the original reconstructions or the ground truth we use two quantitative measures: relative Mean Squared Error (rMSE) and the structural similarity index (SSIM).

The rMSE is defined as follows

$$\mathrm{rMSE}(\mathbf{x}, \mathbf{x}_{\mathrm{ref}}) = \frac{\|\mathbf{x} - \mathbf{x}_{\mathrm{ref}}\|_2^2}{\|\mathbf{x}_{\mathrm{ref}}\|_2^2}, \tag{2.27}$$

which measures the distance between the object $\mathbf{x}$ and the reference object $\mathbf{x}_{\mathrm{ref}}$ in the L^2-sense.

The SSIM measures the luminance, contrast and structure between the samples $\mathbf{x}$ and $\mathbf{x}_{\mathrm{ref}}$. We use the implementation from ODL [AKÖ17]. We set the constants as suggested in [Wan+04], except for L, which we set equal to the dynamic range of the pixel values.[3] The mean and variance are computed with a Gaussian filter with

[2] The reconstruction method used to compute the gold standard reconstruction is a SIRT reconstruction with 300 iterations and a non-negativity constraint.

[3] The dynamic range is the difference between the maximum pixel value and the minimum pixel value of the reference reconstruction $\mathbf{x}_{\mathrm{ref}}$.

width of 11 pixels. We chose these settings because in this case SSIM reflected our own observations of the relative quality of the reconstructions. The SSIM ranges between -1 and 1, where 1 indicates that the image and reference image are identical.

2.5 Results and discussion

The results are structured as follows: we investigate the quality of the approximations in **Section 2.5.1**, in **Section 2.5.2** we investigate the use of our approach for selecting an "optimal" regularization parameter in various scenarios for simulated data and in **Section 2.5.3** for experimental data.

2.5.1 Method validation

As stated in **Section 2.4.2** we consider two computer simulated phantoms (**Figure 2.1**). For these phantoms we will consider two reconstruction problems: a *sparse view reconstruction problem*, with 64 equidistant projection angles and no noise ($n_l = 0$); and a *noisy reconstruction problem*, with 740 projection angles and a noise level $n_l = 0.1$.

Single regularization parameter methods

In this section we will validate the proposed method for the Sobolev and TV regularization reconstruction methods. For the experiments with these methods we took $N_s = 301$ sample points on the interval $[10^{-3}, 1]$, unless mentioned otherwise.

We will mainly look at a sparse view problem with **Figure 2.1a** as phantom reconstructed with TV regularization, however, the shown results are similar for the other cases described. In **Figure 2.2** we show the rMSE and SSIM of the approximations $S_{\text{TV}}(\lambda)$ with respect to the reconstructions $\mathbf{x}_{\text{TV}}^{\lambda}$ for a varying number of interpolation points N_{ip}. We observe that the rMSE and SSIM lie close to 0 and 1, respectively, which indicates that the approximations are close to the reconstructions. Moreover, we see that worst approximations lie in the middle between two interpolation points and that taking more interpolation points results in better approximations. In **Figure 2.3** the approximations for several pixels p are shown. We observe that for $N_{\text{ip}} = 6$ the approximations are of relatively low accuracy. Moreover, we see that the assumption on the boundary conditions (2.16) are not always valid on the right hand side of the interval, which also influences the accuracy of the approximation.

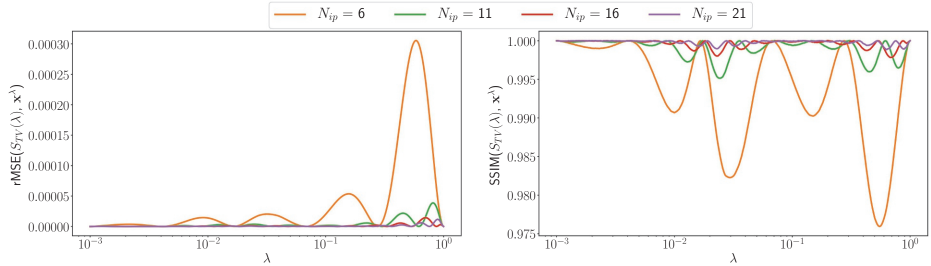


Figure 2.2: The rMSE (Left) and SSIM (Right) of the interpolated approximations with respect to the reconstructions as a function of the regularization parameter λ for varying number of interpolation points N_{ip}. Here we consider the sparse view reconstruction problem as defined in **Section 2.5.1** for the phantom shown in **Figure 2.1a** and the reconstructions are computed with the TV method.

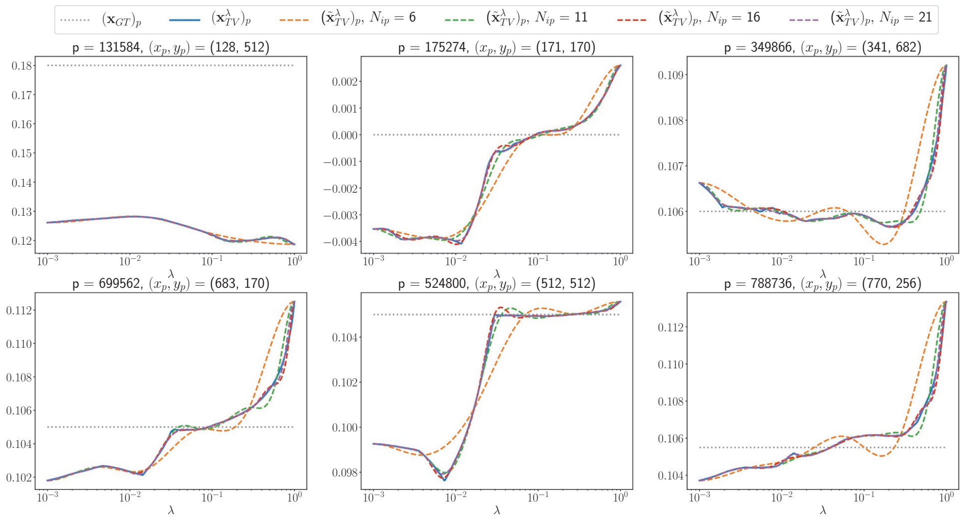


Figure 2.3: Pixel values of the phantom, the reconstruction and the interpolated approximations with varying number of interpolation points N_{ip} as a function of the regularization parameter for several pixels in the image. Here p indicates the pixel position in the vector and (x_p, y_p) indicates the pixel position in the image. Here we consider the sparse view reconstruction problem as defined in **Section 2.5.1** for the phantom shown in **Figure 2.1a** and the reconstructions are computed with the TV method.

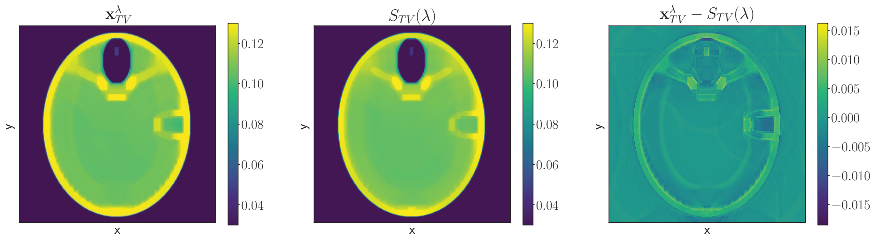


Figure 2.4: (Left) The worst case pixel-wise interpolated approximation with $N_{\text{ip}} = 6$ interpolation points of the TV reconstruction with regularization parameter $\lambda = 10^{-0.26}$ (Middle) and their difference (Right). Here we consider the sparse view reconstruction problem as defined in **Section 2.5.1** for the phantom shown in **Figure 2.1a** and the reconstructions are computed with the TV method.

In **Figure 2.4** we show the worst approximation for the full reconstruction with respect to the rMSE, the corresponding reconstruction and their difference. We observe that the approximation is worst around the high contrast parts of the object. However, the general behavior and properties of the reconstruction are accurately represented.

In **Figure 2.5** the average and the worst approximation of the rMSE and SSIM with respect to the reconstructions are shown for all the cases we stated at the beginning of this section. To avoid a positive bias the interpolation points are not taken into consideration for the statistics. Again we observe that the more interpolation points are used, the better the approximations are, and that the worst approximations are still close to the original.

Two regularization parameter method

For two regularization parameters computing reference reconstructions at high resolution to validate our methods is prohibitively expensive. Therefore, we only consider the noisy reconstruction problem as defined in **Section 2.5.1** for the phantom shown in **Figure 2.1b** and take a smaller reconstruction problem in terms of pixels and angles: 256×256, with 360 equidistant projection angles. For the regularization parameter grid we take $\lambda_1 \in [10^{-4}, 10^2]$ and $\lambda_2 \in [10^{-2}, 10^2]$ and $N_{s,1} = 181$ and $N_{s,2} = 121$.

The quantitative measures of the approximations with respect to the reconstructions are shown in the left and middle column of **Figure 2.6** and in the right column the quantitative measures of the reconstruction with respect to the ground truth. We observe lower accuracy of the approximation between the grid points of

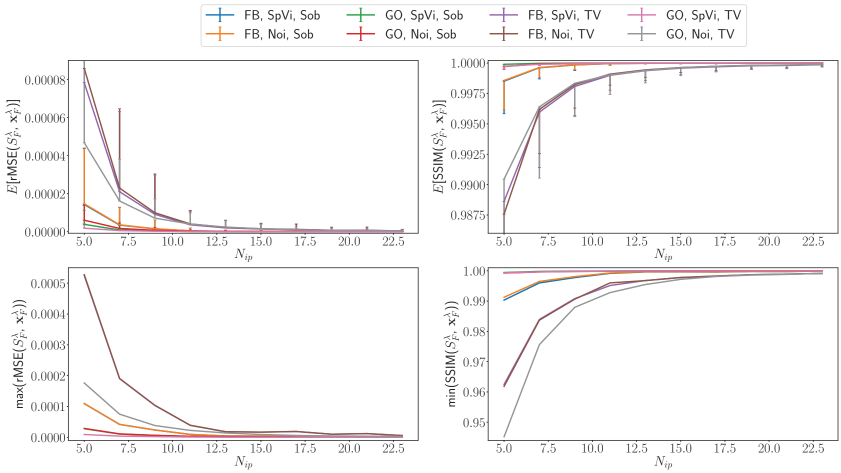


Figure 2.5: The average and standard deviation (Top row) and worst cases (Bottom row) for the rMSE (Left column) and SSIM (Right column) of the pixel-wise interpolated approximations with respect to the reconstructions for varying number of interpolation points N_{ip}, reconstruction problems and methods. To avoid cluttering of the figure the standard deviation bars are only plotted in one direction.

λ_1, as we also saw for one regularization parameter. However, the approximations do not vary in accuracy in λ_2-direction. In **Figure 2.7** we show the worst approximation, and the difference with respect to the SSIM, $\lambda_1 = 10^{-0.267}$, $\lambda_2 = 10^{-0.3}$. In the top row we show $N_{\text{ip},1} = 5$, $N_{\text{ip},2} = 5$ and in the bottom row $N_{\text{ip},1} = 10$, $N_{\text{ip},2} = 5$. We see that taking more points in the λ_1-direction results in more accurate reconstructions.

In **Table 2.1** the average and standard deviation of the rMSE and SSIM with respect to the reconstructions are shown. Again we observe that a finer grid in the λ_1 direction results in more accurate approximations. Moreover, we observe that taking a finer grid for λ_2 has a limited influence on the accuracy of the approximations.

2.5.2 Parameter optimization with simulated data

Single parameter regularization methods

In this section we use the approximations $S_F(\lambda)$ to determine the “optimal” regularization parameter $\lambda^\star$, as determined by a number of objective measures or criteria. We compare our methods to evaluating the method using only the reconstructions available at the interpolation points Λ and at the fully sampled grid Λ^f.

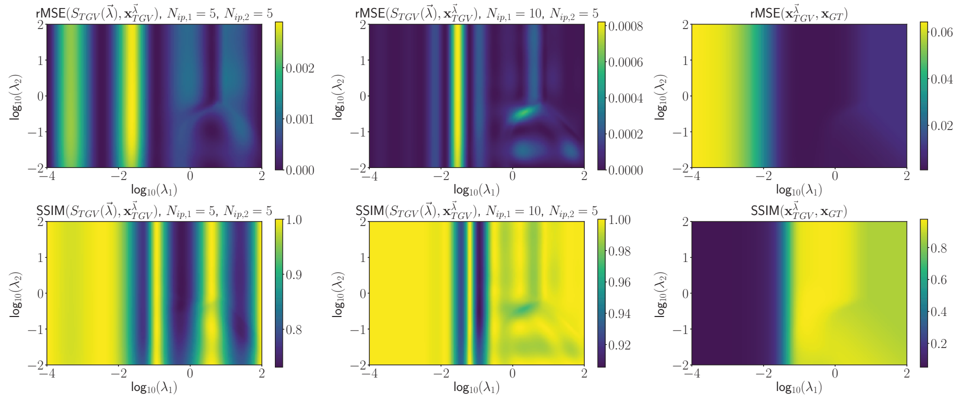


Figure 2.6: (Right and middle column) Quantitative measures of the interpolated approximations with respect to the reconstructions for varying λ_1 and λ_2. The approximations are done with $N_{\text{ip},1} = 5$, $N_{\text{ip},2} = 5$ and $N_{\text{ip},1} = 10$, $N_{\text{ip},2} = 5$ interpolation points in respectively the left and middle column. The right column shows the quantitative measures of the reconstructions with respect to the ground truth.

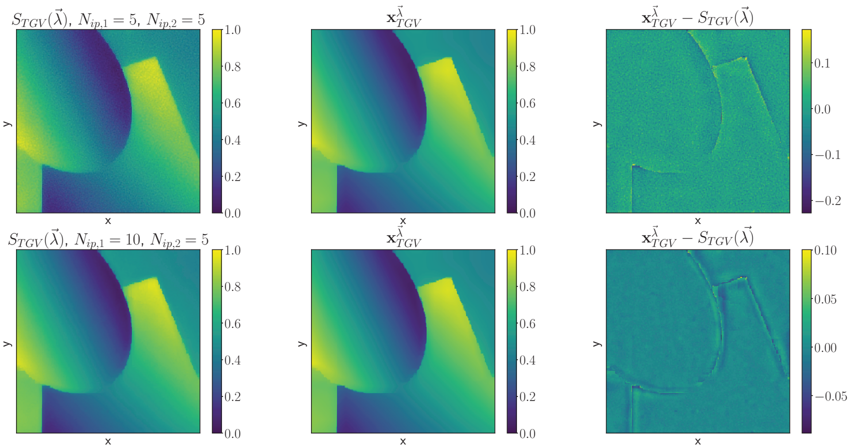


Figure 2.7: (Left column) Worst case pixel-wise interpolated approximation with $N_{\text{ip},1} = 5$, $N_{\text{ip},1} = 5$ (Top) and $N_{\text{ip},1} = 10$, $N_{\text{ip},1} = 5$ (Bottom) interpolation points of the TGV reconstruction with regularization parameters $\lambda_1 = 10^{-0.267}$, $\lambda_2 = 10^{-0.3}$ (Middle column) and their difference (Right column).

	rMSE		
	$N_{\text{ip},1} = 5$	$N_{\text{ip},1} = 10$	$N_{\text{ip},1} = 15$
$N_{\text{ip},2} = 5$	$(8.99 \pm 0.95) \cdot 10^{-4}$	$(1.05 \pm 0.04) \cdot 10^{-4}$	$(4.53 \pm 0.08) \cdot 10^{-5}$
$N_{\text{ip},2} = 10$	$(8.87 \pm 0.96) \cdot 10^{-4}$	$(8.91 \pm 0.4) \cdot 10^{-5}$	$(2.58 \pm 0.03) \cdot 10^{-5}$
$N_{\text{ip},2} = 15$	$(8.88 \pm 0.96) \cdot 10^{-4}$	$(8.85 \pm 0.4) \cdot 10^{-5}$	$(2.46 \pm 0.03) \cdot 10^{-5}$
	SSIM		
$N_{\text{ip},2} = 5$	$.907 \pm 0.12$	$.988 \pm 0.06$	$.996 \pm 0.01$
$N_{\text{ip},2} = 10$	$.908 \pm 0.12$	$.989 \pm 0.06$	$.997 \pm 0.01$
$N_{\text{ip},2} = 15$	$.908 \pm 0.12$	$.989 \pm 0.06$	$.997 \pm 0.01$

Table 2.1: The average and standard deviation rMSE and SSIM of the interpolated approximations with respect to the TGV reconstructions for a varying number of interpolation points $N_{\text{ip, 1}}, N_{\text{ip, 2}}$.

Parameter space exploration **Figure 2.8** shows the process of exploring the parameter space for the noisy reconstruction problem as defined in **Section 2.5.1** for the phantom shown in **Figure 2.1a** with the TV regularization reconstruction method, using 6 reconstructions ($N_{\text{ip}} = 6$). Here we determine the “visually optimal” parameter $\lambda^\star = 10^{-1.58}$ based on our visual inspection of the approximations. The top row are the 6 available reconstructions (red border), from which we see that the visually optimal parameter should lie in the interval $[10^{-1.78}, 10^{-1.17}]$. Knowing this, we take a number of interpolated approximations in this interval and compare them (the first 5 images on the bottom row, orange border). We observe that the approximation with $\lambda = 10^{-1.58}$ has a good trade-off between sharpness and artifacts in the background. The actual reconstruction is shown in the bottom right (green border). Note that without the interpolations the only information available would be the top row, meaning that one would have to choose between $\lambda = 10^{-1.78}$ and $\lambda = 10^{-1.17}$ or do additional computations.

Quantitative measure optimization In the case that there is a ground truth or high quality reconstruction available we determine the “optimal” regularization parameter with respect to a quantitative measure, by optimizing the QM curves (recall (2.18)). In **Figure 2.9** we show the resulting curves for the noisy reconstruction problem as defined in **Section 2.5.1** for the phantom shown in **Figure 2.1a** reconstructed with TV regularization. For comparison the curve computed with the actual reconstructions and direct spline interpolations through the points $\text{QM}(\mathbf{x}_{TV}^{\Lambda}, \mathbf{x}_{GT})$ are shown. We see that both interpolation methods give good approximations, although the pixel-wise interpolations are less accurate at the boundaries. This is most likely due to the assumption of clamped boundary conditions not being accurate enough.

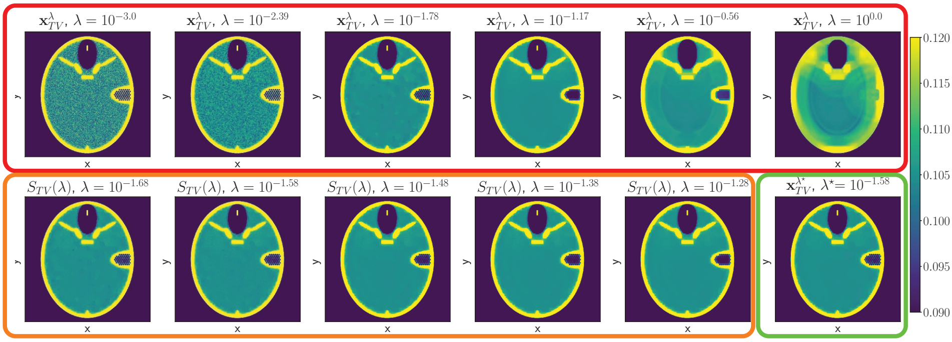


Figure 2.8: Visualization of the parameter space exploration. The top row with the red border are the reconstructions on the coarse grid Λ that are used for the interpolations. The approximations in the window of interest are shown in the bottom row with the orange border and the reconstruction with the "visually optimal" regularization parameter $\lambda^\star = 10^{-1.58}$ is shown in the bottom right with the green border. Here we consider the noisy reconstruction problem as defined in **Section 2.5.1** for the phantom shown in **Figure 2.1a** and the reconstructions are computed with the TV method.

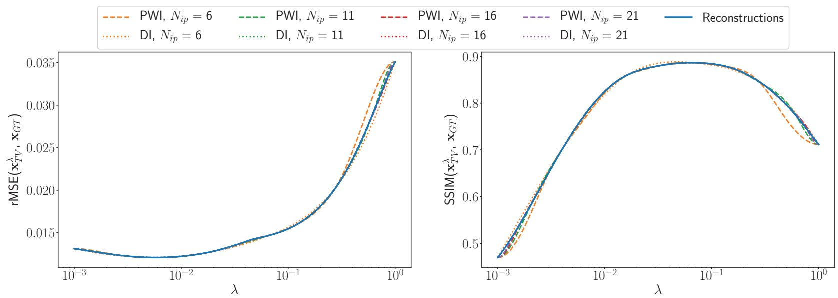


Figure 2.9: The rMSE (Left) and SSIM (Right) curves for the reconstructions and the pixel-wise interpolated (PWI) approximations with respect to the ground truth and the QM curves directly interpolated (DI) from the QM values on the coarse grid Λ for varying number of interpolation points N_{ip}. Here we consider the noisy reconstruction problem as defined in **Section 2.5.1** for the phantom shown in **Figure 2.1a** and the reconstructions are computed with the TV method.

In **Table 2.2** we show the estimated regularization parameters for the 8 different cases we considered also in the previous section. We observe that estimations of both the pixel-wise interpolation and the direct interpolation methods are close to the "optimal" parameters. For the cases where the estimated parameter is less accurate, we inspected the curves and observed that the curve was at a plateau around the optimal value, making it more sensitive to errors. Lastly, we observe that the "optimal" regularization parameter varies depending on which quantitative measure is used.

	Sobolev regularization				TV regularization			
Case: Noisy	FORBILD		Gradient objects		FORBILD		Gradient objects	
Method	rMSE	SSIM†	rMSE	SSIM†	rMSE	SSIM	rMSE	SSIM
$\log_{10}(\lambda^{\star})$	-2.02	-0.06	-0.9	0.05	-2.24	-1.19	-0.63	-0.36
$\log_{10}(\lambda^{\star}_{\mathrm{PWI}})$, $N_{\mathrm{ip}} = 6$	-2.03	-0.39	-0.88	-0.02	-2.25	-1.25	-0.6	-0.59
$\log_{10}(\lambda^{\star}_{\mathrm{PWI}})$, $N_{\mathrm{ip}} = 11$	-2.02	-0.15	-0.9	0.04	-2.23	-1.21	-0.63	-0.15
$\log_{10}(\lambda^{\star}_{\mathrm{PWI}})$, $N_{\mathrm{ip}} = 16$	-2.02	-0.05	-0.9	0.05	-2.24	-1.19	-0.62	-0.35
$\log_{10}(\lambda^{\star}_{\mathrm{PWI}})$, $N_{\mathrm{ip}} = 21$	-2.02	-0.06	-0.9	0.05	-2.24	-1.19	-0.63	-0.35
$\log_{10}(\lambda^{\star}_{\mathrm{DI}})$, $N_{\mathrm{ip}} = 6$	-2.02	-0.05	-0.89	0.07	-2.27	-1.33	-0.84	-0.84
$\log_{10}(\lambda^{\star}_{\mathrm{DI}})$, $N_{\mathrm{ip}} = 11$	-2.02	-0.06	-0.9	0.05	-2.24	-1.16	-0.61	-0.39
$\log_{10}(\lambda^{\star}_{\mathrm{DI}})$, $N_{\mathrm{ip}} = 16$	-2.02	-0.06	-0.9	0.05	-2.24	-1.19	-0.63	-0.35
$\log_{10}(\lambda^{\star}_{\mathrm{DI}})$, $N_{\mathrm{ip}} = 21$	-2.02	-0.06	-0.9	0.05	-2.24	-1.19	-0.63	-0.35

Table 2.2: Estimated and "optimal" regularization parameters based on quantitative measure optimization for two reconstruction methods, two phantoms and the rMSE and SSIM. We only show the noisy reconstruction problem, the sparse view problems have similar results. The estimations of the regularization parameters are done based on the pixel-wise interpolations (PWI) and the direct interpolation (DI) of the QM curves. For the cases denoted with a † the range of the regularization parameter is changed to $[10^{-2}, 10^{1}]$ to ensure that the "optimal" parameter is within the considered range.

Discrepancy principle & L-curve criterion To ensure capturing the desired behavior we take a larger interval for the regularization parameter, $[10^{-4}, 10^{2}]$, and scale the sampling points accordingly, $N_s = 601$. In **Figure 2.10** we show show results for reconstructions, pixel-wise interpolations and direct interpolations. The top row shows the values of the data-fidelity as a function of the regularization parameter λ. The results shown are for the noisy reconstruction problem as defined in **Section 2.5.1** for the phantom shown in **Figure 2.1b**. Here we observe that the significant changes of the functions are in a relative small window of λ. This results in bad approximations if there is no reconstruction available in this window. Moreover, we see that the noise level ϵ intersects the data-fidelity curve at a plateau, which might result in inaccurate estimates for the regularization parameter (see the second to last column of **Table 2.3**).

In the bottom row of **Figure 2.10** we show the L-curve and its curvature. Here we observe that inaccuracies in the initial approximations result in inaccuracies in the L-curve, and even more in its curvature. In the extreme case we observe a change of sign of a peak for the pixel-wise interpolations with $N_{\text{ip}} = 6$. However, even with these inaccuracies, the approximations follow the general behavior of the reference curve and through visual inspection of the approximations and curves one can stil determine the optimal parameter.

The “optimal” parameters for discrepancy principle and the L-curve criterion for the 8 cases considered are shown in **Table 2.3**. We observe more volatility in the “optimal” parameters compared to the results in **Table 2.2**, which coincides with the observations in **Figure 2.10**. Additionally, for two cases the direct interpolations resulted in negative values for the TV term resulting in an undefined L-curve. Lastly, we again see that the optimal value of the regularization parameter varies depending on the used method.

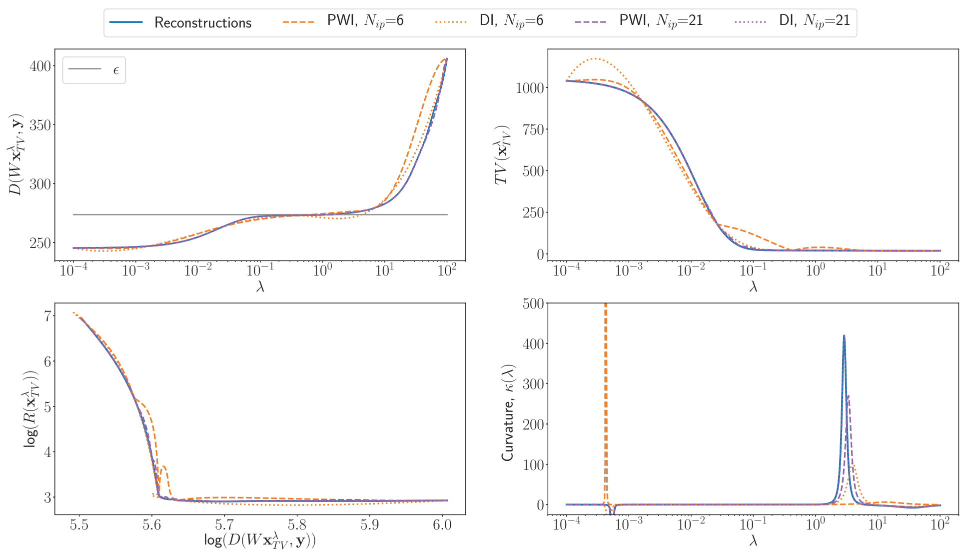


Figure 2.10: (Top) The data-fidelity (Left) and TV functional (Right) as a function of the regularization parameter λ. (Bottom) The L-curve (Left) and its curvature (Right). These curves are computed with the reconstructions, the pixel-wise interpolated (PWI) approximations and through direct interpolation (DI). Here we consider the noisy reconstruction problem as defined in **Section 2.5.1** for the phantom shown in **Figure 2.1b** and the reconstructions are computed with the TV method.

Case: Noisy	Sobolev regularization				TV regularization			
	FORBILD		Gradient objects		FORBILD		Gradient objects	
Method	DP	LC	DP	LC	DP	LC	DP	LC
$\log_{10}(\lambda^{\star})$	-2.33	0.13	-0.58	0.83	-2.95	-0.64	0.09	0.46
$\log_{10}(\lambda^{\star}_{\mathrm{PWI}})$, $N_{\mathrm{ip}} = 6$	-2.34	0.07	-0.58	0.82	-3.01	-0.56	-0.18	$1.12^{\dagger}$
$\log_{10}(\lambda^{\star}_{\mathrm{PWI}})$, $N_{\mathrm{ip}} = 11$	-2.33	0.13	-0.58	$.82^{\dagger}$	-2.96	-0.61	0.09	0.72
$\log_{10}(\lambda^{\star}_{\mathrm{PWI}})$, $N_{\mathrm{ip}} = 16$	-2.33	0.13	-0.58	0.83	-2.95	-0.63	0.09	0.58
$\log_{10}(\lambda^{\star}_{\mathrm{PWI}})$, $N_{\mathrm{ip}} = 21$	-2.33	0.13	-0.58	0.84	-2.95	-0.63	0.09	0.51
$\log_{10}(\lambda^{\star}_{\mathrm{DI}})$, $N_{\mathrm{ip}} = 6$	-2.39	-	-0.62	-	-2.89	-0.55	0.63	0.6
$\log_{10}(\lambda^{\star}_{\mathrm{DI}})$, $N_{\mathrm{ip}} = 11$	-2.33	0.14	-0.58	0.83	-2.97	-0.62	0.05	0.45
$\log_{10}(\lambda^{\star}_{\mathrm{DI}})$, $N_{\mathrm{ip}} = 16$	-2.33	0.13	-0.58	0.83	-2.95	-0.64	0.09	0.46
$\log_{10}(\lambda^{\star}_{\mathrm{DI}})$, $N_{\mathrm{ip}} = 21$	-2.33	0.13	-0.58	0.83	-2.95	-0.63	0.09	0.46

Table 2.3: Estimated and "optimal" regularization parameters based on discrepancy principle (DP) and L-curve criterion (LC) for two reconstruction methods and two phantoms. We only show the noisy reconstruction problem, because these methods are not feasible for the sparse view problem. The estimations of the regularization parameters are done based on the pixel-wise interpolations (PWI) and the direct interpolation (DI) of the QM curves. For the cases denoted with a † the maximum curvature is at another value of λ, however, through closer inspection of the curve (in a similar manner as for **Figure 2.10**) this parameter is chosen.

Two regularization parameter method

Taking into consideration the observations from **Section 2.5.1** we use the same settings for the reconstruction problem and we take $N_{\mathrm{ip},1} = 10$, $N_{\mathrm{ip},2} = 5$ interpolation points for our interpolation scheme.

Quantitative measure optimization In **Figure 2.11** we show the quantitative measures of the approximations (left column) and the reconstructions (right column) with respect to the ground truth. The "optimal" regularization parameters and their respective quantitative measures determined from these figures are given in the caption. We observe that the found "optima" all lie in a plateau region which has relatively small differences in the quantitative measures. This indicates that the proposed method arrives at similar results as the original method.

Parameter space exploration Upon visual inspection of the reconstructions we concluded that the "visually optimal" set of parameters lies in $\log_{10}(\lambda_1) \in [-0.5, 0.2]$ and $\log_{10}(\lambda_2) \in [0.07, 1.1]$ (red border in **Figure 2.12**). Inspection of the approximations (orange border) in this interval resulted the optimal parameter set $(\lambda_1, \lambda_2) = (10^{-0.33}, 10^{0.07})$. We observe minimal differences between the approximation and the actual reconstruction (green border) for this "visually optimal" set of parameters.

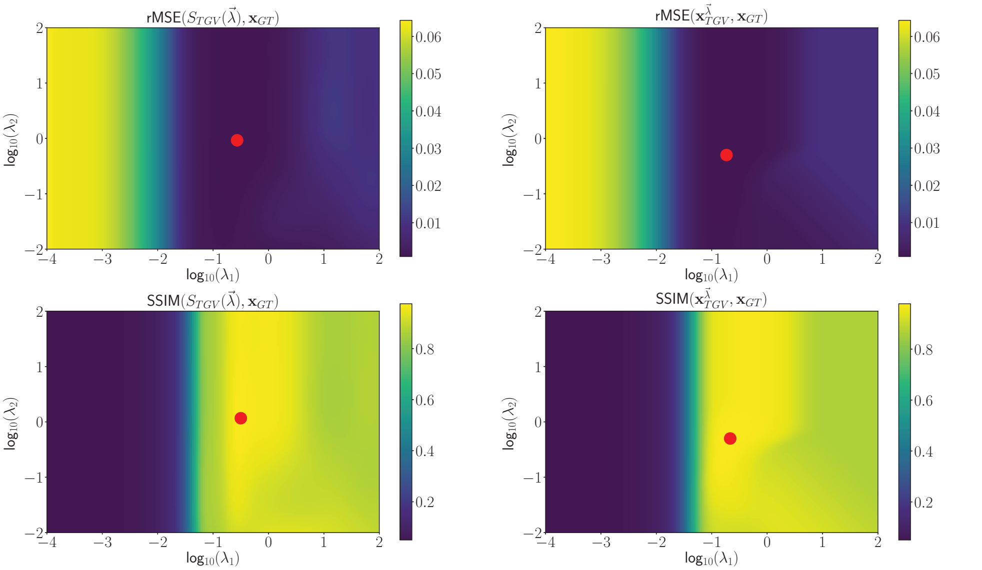


Figure 2.11: Quantitative measure figures for the TGV reconstruction method with respect to the ground truth. The red dot indicates the extremum of the figure.
(Top left) rMSE of the approximations with $\lambda^\star = (10^{-0.57}, 10^{-0.033})$ and min(rMSE) $= 9.1 \cdot 10^{-4}$.
(Top right) rMSE of the reconstructions with $\lambda^\star = (10^{-0.73}, 10^{-0.30})$ and min(rMSE) $= 8.8 \cdot 10^{-4}$.
(Bottom left) SSIM of the approximations with $\lambda^\star = (10^{-0.5}, 10^{0.067})$ and max(SSIM) $= 0.9755$.
(Bottom right) SSIM of the reconstructions with $\lambda^\star = (10^{-0.67}, 10^{-0.30})$ and max(SSIM) $= 0.9779$.

2.5.3 Parameter optimization with experimental data

In this section we show results for our method on experimental fan beam CT data. For this case we took the TV regularization reconstruction method implemented with the PDHG algorithm, using $N_{\text{iter}} = 2500$ iterations, $\tau = \frac{0.1}{\|L_{\text{TV}}\|}$, regularization parameter range $[10^{-4}, 1]$, $N_s = 401$ sample points and $N_{\text{ip}} = 11$ interpolation points.

The parameter space exploration for the experimental data and the TV reconstruction method is shown in **Figure 2.13**. Again only a selection of the approximations and initial reconstructions is shown for a clearer visualization. In **Figure 2.14** we show in the top row the TV reconstruction with the "visually

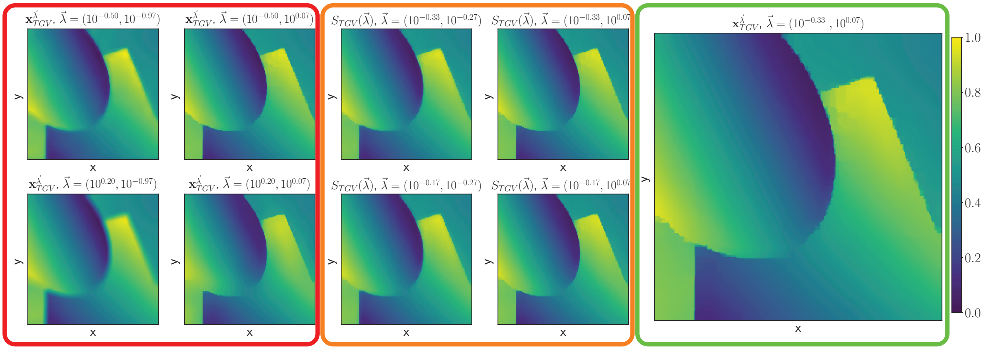


Figure 2.12: Partial visualization of the parameter space exploration for the TGV reconstruction method. The first and second column (red border), show several close to “visually optimal” reconstructions on the coarse grid Λ. The third and fourth column (orange border) partially show the further exploration through the approximations with in the top right the “visually optimal” approximation and the image on the right (green border) shows the reconstruction with this “visually optimal” regularization parameter.

optimal” regularization parameter, a FBP reconstruction and the gold standard reconstruction. We can conclude from the FBP reconstruction[4] that the noise in the data is quite severe and that it would be surprising if any reconstruction method can retrieve the small features at the boundary of the pomegranate and inside the seeds (observed in the gold standard reconstruction). Taking this into consideration the choice of regularization parameter results in an adequate TV reconstruction. We observe in the difference between the gold standard reconstruction and the TV reconstruction (Bottom right in **Figure 2.14**) the loss of the smaller details and a general loss of contrast, which is a known property of the TV method. Lastly, we consider the quantitative measures shown in **Table 2.4**. These results confirm the earlier conclusions; the pixel-wise approximation is very good and the chosen regularization parameter results in an adequate TV reconstruction.

2.6 Conclusion

In this chapter we have proposed an algorithmic approach for computationally efficient exploration of the regularization parameter space, based on a pixel-wise

[4]The FBP reconstruction is done with a Ram-Lak filter, with no windowing to reduce the noise in the input data [Nat01].

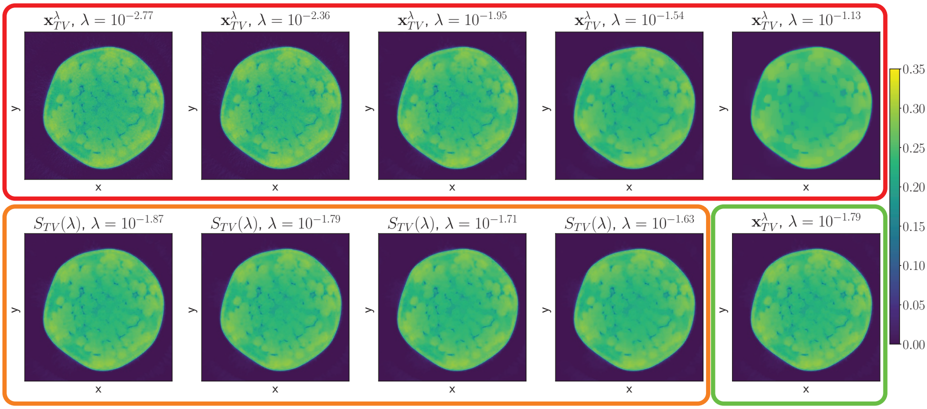


Figure 2.13: Partial visualization of the parameter space exploration. The top row (red border) shows several reconstructions on the coarse grid Λ. The images on the bottom row (orange border) show the further exploration of the parameter space through the use of approximations and on the bottom right (green border) the reconstruction with the "visually optimal" regularization parameter is shown. The difference between the approximation of the reconstruction is shown in **Figure 2.14**.

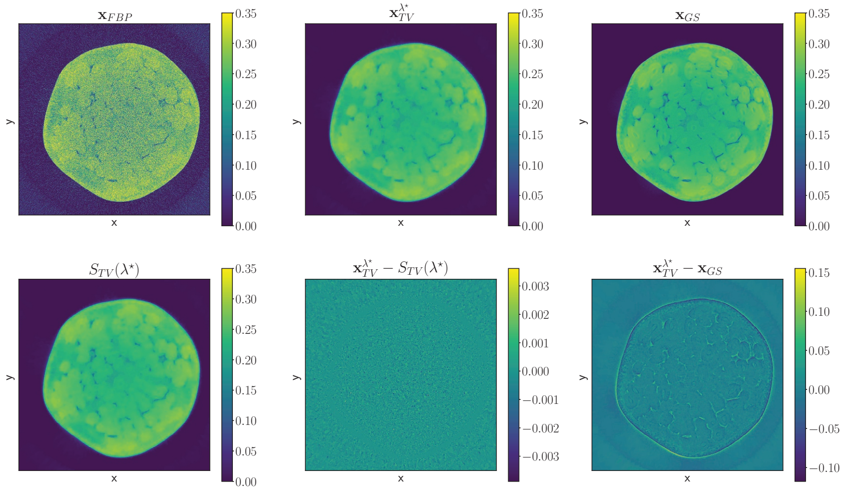


Figure 2.14: (Top left) FBP method reconstruction, (Top middle) TV method reconstruction, (Top right) Gold standard reconstructions. (Bottom left) Pixel-wise approximation to the TV reconstruction, (Bottom middle) Difference between TV approximation and reconstruction, (Bottom right) Difference between the TV and gold standard reconstruction.

Metric	$\mathbf{x} = S_{TV}(\lambda^\star)$, $\mathbf{x}_{\text{ref}} = \mathbf{x}_{TV}^{\lambda^\star}$	$\mathbf{x} = \mathbf{x}_{TV}^{\lambda^\star}$, $\mathbf{x}_{\text{ref}} = \mathbf{x}_{GS}$	$\mathbf{x} = \mathbf{x}_{\text{FBP}}$, $\mathbf{x}_{\text{ref}} = \mathbf{x}_{GS}$
rMSE($\mathbf{x}, \mathbf{x}_{\text{ref}}$)	$1.6328 \cdot 10^{-6}$	$2.2750 \cdot 10^{-3}$	$8.5259 \cdot 10^{-2}$
SSIM($\mathbf{x}, \mathbf{x}_{\text{ref}}$)	0.9988	0.6986	0.0186

Table 2.4: Quantitative measures for the experimental data results. Here $\mathbf{x}_{\text{FBP}}$ indicates the FBP reconstruction [Nat01] with a Ram-Lak filter.

interpolation scheme. Given a relatively small number of reconstructions on a sparsely sampled parameter grid, our method can be used to quickly compute an approximation to a reconstruction for any regularization parameter within the sampled range.

We have shown for three common variational reconstruction methods, Sobolev, TV and TGV regularization, that our method produces accurate approximations for simulated and experimental data. Moreover, we have shown that the approximations can be used in existing parameter optimization methods.

To conclude, our method enables developing computationally efficient tools that provide real-time visualization of the regularization parameter space, and automated parameter selection based on existing optimization criteria.