Reconstruction of Current Densities from Magnetic Images by Bayesian Inference

Colin B. Clement,\textsuperscript{1} James P. Sethna,\textsuperscript{1} and Katja C. Nowack\textsuperscript{1}
\textsuperscript{1}Laboratory of Atomic and Solid-State Physics, Cornell University, Ithaca, NY 14853, United States
(Dated: July 7, 2021)

Electronic transport is at the heart of many phenomena in condensed matter physics and material science. Magnetic imaging is a non-invasive tool for detecting electric current in materials and devices. A two-dimensional current density can be reconstructed from an image of a single component of the magnetic field produced by the current. In this work, we approach the reconstruction problem in the framework of Bayesian inference, i.e. we solve for the most likely current density given an image obtained by a magnetic probe. To enforce a sensible current density priors are used to associate a cost with unphysical features such as pixel-to-pixel oscillations or current outside the device boundary. Beyond previous work, our approach does not require analytically tractable priors and therefore creates flexibility to use priors that have not been explored in the context of current reconstruction. Here, we implement several such priors that have desirable properties. A challenging aspect of imposing a prior is choosing the optimal strength. We describe an empirical way to determine the appropriate strength of the prior. We test our approach on numerically generated examples. Our code is released in an open-source python package called \texttt{pysquid}.

I. INTRODUCTION

Two-dimensional (2D) materials host a variety of electronic transport phenomena, many of which are associated with a non-trivial spatial structure of the current density in the material. A non-invasive way to image a 2D current density is to image the stray magnetic field produced by the current and infer the current density. To date, numerous magnetic imaging techniques have been used to image current densities including scanning SQUID\textsuperscript{1,2} scanning Hall probe\textsuperscript{3}, magneto-optic\textsuperscript{4}, nitrogen-vacancy (NV) centers in diamond\textsuperscript{5,6} and levitated Bose-Einstein condensates\textsuperscript{7–9}.

Most magnetic imaging techniques probe a single magnetic field component in a plane at a fixed height above the sample. (A notable exception are NV centers in diamond, which can realize vector magnetic field imaging\textsuperscript{[12]}) The relation between the current density and the measured magnetic image is defined through two convolutions: the Biot-Savart law relates the magnetic field to the current density and a convolution of the magnetic field with the point spread function (PSF) of the magnetic sensor relates the magnetic field to the output of the sensor. To obtain the current density from a magnetic image, the two convolutions have to be inverted. If the current density only varies in two dimensions, this inversion is in principle possible because current conservation relates the two in-plane components of the current. In practice, the inversion is a non-trivial task because the problem is ill-posed: experimental images contain noise and the finite scan height and PSF lead to a loss of spatial information. As a consequence noise dominates the reconstructed image at high spatial frequencies. There are many solutions that predict the data including noise perfectly, but most of these solutions are not physical. Therefore, a criterion for what constitutes a physically sensible solution is required. This criterion is imposed through different so-called priors or equivalently a regularization.

A detailed overview of existing methods for current reconstruction is given by Meltzer et al\textsuperscript{[13]}. The most intuitive method is to invert the convolutions directly in Fourier space\textsuperscript{[12]}, filtering high spatial frequencies that otherwise cause instability. However, the shape and cutoff frequencies of the applied filters limit the resolution of the reconstructed image in an uncontrolled way. Iterative conjugate gradient methods have also been employed\textsuperscript{[13]}, which are more stable to noise. However, the regularization is not well controlled. Feldmann\textsuperscript{[10]} and Meltzer et al\textsuperscript{[13]} have reported reconstruction procedures using Tikhonov regularization penalizing the Laplacian of the current dipole field (defined below), combined with a cross-validation-based choice of the regularization strength. Tikhonov regularization is an attractive method because it is analytically tractable in Fourier space, i.e. it corresponds to a filter in Fourier space that can be expressed in Fourier space (see below). This allows for computationally efficient solutions and theoretically motivated methods of choosing the regularization strength.

In the wider image reconstruction literature, a variety of priors have been developed that are not analytically tractable. To date, these have not been applied to current reconstruction. For example, a prior based on total variation of the signal penalizes oscillations in a solution, but not sharpness like Tikhonov regularization does. An additional complication when reconstructing current densities is that typically some current leaves and enters the imaged field of view. At the points along the image boundary where this happens, the current density is not conserved. This violates the assumption of conserved current, without which the problem is underconstrained. Meltzer et al\textsuperscript{[13]} have implemented mirror boundary conditions for accommodating currents which enter or leave the image. This method has again the advantage of being analytically tractable, but it is not faithful to the sample geometry unless the sample is mirror-symmetric.

Here we describe a procedure to reconstruct current
density from magnetic images that enables the use of a wider class of priors than previous work, and can accommodate currents crossing the image boundaries. We formulate the reconstruction problem in a probabilistic framework suitable for Bayesian inference that utilizes a generative model of the data. This approach offers significant flexibility to make use of prior information about the current density, including the sample geometry. Previous methods have penalized the Laplacian of the current dipole field (equivalent to Tikhonov regularization), from which the current density is computed, or the components of the currents themselves. Here, we discuss a set of requirements such as rotational invariance which a prior ideally should obey. From this we show that a prior based on the Frobenius of the Hessian is better motivated than the commonly used Laplacian prior. The approach presented here could be extended to other quantitative imaging problems in physics, where a generative model connecting the image and the underlying physical quantities is known.

The priors we explicitly discuss include a Gaussian prior (equivalent to Tikhonov regularization) penalizing the Laplacian and the Frobenius of the Hessian, a total prior (equivalent to Tikhonov regularization), penalizing strong fluctuations in the current density, but does not necessarily blur sharp edges, and a finite support prior, which allows the user to specify areas in the field of view where no current flows. In addition, we show that we can accommodate currents crossing the image boundary through modeling current densities flowing outside the field of view based on the sample geometry. This reconstruction problem ultimately leads to a convex optimization problem which we solve using the Alternating Difference Method of Multipliers (ADMM).

An important aspect of implementing a prior is to choose the strength with which it is imposed. If the prior is too weak, any reconstruction method tends to overfit the data such that even the noise is reproduced. If the prior is too strong, the resulting reconstruction is typically too smooth or has other undesirable and unphysical features. For some regularizations, one can theoretically predict the optimal strength of a prior for a given model. Typically, these are based on considering a single functional of a relatively simple metric such as the standard deviation of the residuals. A similar approach is not available for most priors discussed here. Instead, we choose the strength of the prior through inspecting both the standard deviation of the residuals as well as their spatial structure in real and Fourier space. The approach is more empirical and as implemented here requires visual inspection. However, an advantage is that a sense of the accuracy of the model can be gained through inspection of the spatial structure of the residuals.

The paper is organized as follows. In Sec. II, we define the forward problem and describe how we use Bayesian inference. In Sec. III, we propose requirements a prior should fulfill, derive a new prior which satisfies them, and compare it to a previously studied prior. We explore Gaussian priors and introduce a total variation and a finite support prior. Finally (Sec. IVD), we discuss how we choose the strength of a prior. In Sec. IV, we describe how we account for currents flowing outside the field of view. We benchmark our method using numerical results throughout. Details of the inversion algorithm, the implementation of finite support priors, and the external current models are presented in the appendix. The code is organized in a python module called pysquid, which is publicly available in a github repository.

II. BAYESIAN INFERENCE FORMULATION OF THE CURRENT RECONSTRUCTION PROBLEM

A. Forward problem

First, we describe the forward problem: calculating the magnetic image resulting from a known 2D current density. We assume that the magnetic sensor probes the component of the magnetic field perpendicular to and at a fixed height above the plane in which the current flows. We assume a sensitive area of the sensor that is small compared to the scan height so that the PSF can be ignored. However, the methods we present here can be generalized to include a PSF and this capability is included in the code. Furthermore, it is straightforward to apply our reconstruction procedure to other magnetic field components, and to allow for a finite thickness of the current carrying sheet, as long as the current density has only in-plane component and does not change along the vertical direction.

We choose the coordinate system such that the 2D sheet lies in the $x - y$ plane at $z = 0$. Then, the magnetic field produced by a 2D sheet current $j$ at a position $r = (x, y, z)$ above the sample is given by the Biot-Savart law. Assuming no current sources and drains are present in the field of view, the $x-$ and $y-$components of the current density $j_{x,y}$ obey current conservation: $\partial_x j_x + \partial_y j_y = 0$. We explicitly enforce current conservation by introducing a single scalar field $g(s)$ which only depends on two dimensions. Here $s = (x', y', 0)$ because the current density is constrained to the $x - y$ plane at $z = 0$. From this scalar field, we calculate the current density as $j(s) = \nabla \times g(s) \hat{z}$ with $\hat{z}$ the unit vector in the $z$-direction.

The Biot-Savart law as a function of $g$ can be written as

$$B(r) = \frac{1}{4\pi} \int_V ds \ g(s) \frac{3\hat{z} \cdot \hat{n} - \hat{z}}{|r-s|^3},$$  \hspace{1cm} (1)$$

where $\hat{n} = (r-s)/|r-s|$. The kernel convoluted with $g$ in Eq. 1 is the magnetic field of a point dipole oriented along the $z$-direction. $g$ can therefore be viewed as a decomposition of a 2D current density into circulating currents, which is why we refer to $g$ as the current dipole field.
The magnetic sensor probes the $z$-component, $B_z(r)$, of the magnetic field in Eq. (1).

**B. Inverse Problem and Bayesian Inference**

In the following, we consider a discrete rectangular grid in the $x-y$ plane with pixels centered at coordinates $\{r_i\}$. We define a magnetic image vector $\phi$ with values $\phi_i = B_z(r_i)$ and a current dipole field vector $g$ with values $g_i = g(r_i)$ sampled on the same lateral coordinates. The linearity of the Biot-Savart law allows us to write the relationship between $\phi$ and $g$ as $\phi = Mg$ for some suitable linear operator $M \in \mathbb{R}^{N \times N}$ where $N$ is the number of pixels in the image $\phi$.

Each value $g_i$ of the discrete current dipole field corresponds to the amplitude and orientation of a current that circulates along the boundary of the pixel centered at $r_i$. From Eq. (1) we can directly calculate the elements of $M$ as a function of height (see details in the appendix). While the matrix $M$ is impractical to store for any reasonable image size, the product $Mg$ can be efficiently computed using a Fast Fourier Transform (FFT). A PSF can be incorporated in $M$. This capability is included in the code, but a detailed discussion of this aspect will be presented elsewhere.

We assume that the experimental noise is independent and identically distributed for each pixel, so that our model for a measured magnetic image is

$$\phi = Mg + \eta$$

(2)

where $\eta_i$ is the noise for each pixel. We assume Gaussian noise with variance $\sigma^2$, i.e. the probability distribution of the noise is a normal distribution with zero mean and standard deviation $\sigma$: $p(\eta_i) \sim N(0, \sigma^2)$. The noise causes the data to fluctuate around the model with characteristic distance $\sigma$. We can therefore define the likelihood $p(\phi|g)$ of measuring $\phi$ given $g$:

$$p(\phi|g) = \frac{1}{(2\pi \sigma^2)^{N/2}} \exp\left(-\frac{1}{2\sigma^2} \|Mg - \phi\|^2\right),$$

(3)

where $N$ is the number of pixels in the image $\phi$, and $\|\cdot\|^2$ is the Euclidean $L_2$ norm. This likelihood $p(\phi|g)$ is our model of the data and will allow us to infer the current dipole field.

Our goal is to learn $g$ after having measured $\phi$. We therefore need $p(g|\phi)$ (called the posterior probability). Bayes’ Theorem tells us how to reverse the conditional probability in Eq. (3):

$$p(g|\phi) = \frac{p(\phi|g)p(g)}{p(\phi)}.$$  \hspace{1cm} (4)

Here $p(g)$ is the prior probability, encoding a criterion for preferable and physically sensible solutions. $p(\phi)$ is called the evidence and is useful for quantitatively justifying the selection of one model over another\cite{23}. The maximum likelihood solution, $g^*$, to the reconstruction problem is then the most likely $g$ given $\phi$:

$$g^* = \max_g p(g|\phi) = \max_g p(\phi|g)p(g).$$  \hspace{1cm} (5)

At this stage of inference the evidence $p(\phi)$ can be left out as it is independent of $g$. Full treatment of Bayesian inference including optimal model selection is described in detail by Mackay\cite{23}, but is beyond the scope of this work.

Next, we assume that the prior probability (‘prior’ for short) can be written as $p(g) \propto \exp(-\lambda^2 \ell(g))$ for some non-negative cost function $\ell$ and real number $\lambda$ which controls the strength of the prior. Combined with Eqs. (5) and (3) we find the maximum likelihood solution for a given $\lambda$ and $\ell$ as

$$g_\lambda = \min_g \frac{1}{2} \|Mg - \phi\|^2 + (\lambda\sigma)^2 \ell(g),$$  \hspace{1cm} (6)

Inference of currents is now cast as minimizing the negative log-posterior, or minimizing the distance between our model $Mg$ and the data $\phi$, constrained by a cost function $\ell(g)$.

It is instructive to demonstrate the necessity of a non-trivial prior $p(g)$. If we consider all solutions as equally preferable, i.e. $p(g) \propto 1$, the solution of Eq. (5) is given by

$$g = (M^TM)^{-1}M^T\phi.$$  \hspace{1cm} (7)

Here $(M^TM)^{-1}M^T$ is the pseudoinverse, i.e. the ‘closest’ inverse to the singular $M$, which is calculated only from the eigenvectors of $M$ with non-zero eigenvalue. $M$ has at least one zero eigenvalue, as adding any constant to $g$ does not change $\phi$. The pseudoinverse ignores this symmetry, but since the Biot-Savart law is long-range, $M$ has in addition exponentially small eigenvalues corresponding to high spatial frequencies. The noise has support in all frequencies. As a consequence, the solution to Eq. (7) is highly unstable as the pseudoinverse amplifies any amount of noise. More specifically, the pseudoinverse can yield solutions which fit the data and noise perfectly. There is a huge space of solutions $g$ that overfit the data like this, and most of them oscillate rapidly throughout the image. The role of a non-trivial prior $p(g)$ is to restrict this space by using physical arguments to specify which solutions are more likely.

**III. CONSTRUCTING PHYSICALLY MOTIVATED PRIORS**

**A. Gaussian Priors**

A common choice for a prior is a Gaussian with a cost function that depends on a linear transform of $g$:

$$p(g) \propto e^{-\lambda^2 \ell(g)} = \exp(-\lambda^2 \|Mg\|^2),$$  \hspace{1cm} (8)
where $\Gamma$ is a linear operator. Gaussian priors are the conjugate prior to a Gaussian likelihood. This allows to write the explicit solution to Eq. (9)

$$g_\lambda = (M^T M + (\sigma \lambda)^2 \Gamma^T \Gamma)^{-1} M^T \phi.$$ (9)

In this form, we can see that the role of $\Gamma$ is to overwrite the exponentially small eigenvalues of $M^T M$, regularizing the instability of the pseudoinverse in eqn. (7). The variance of the noise $\sigma$ sets the scale for the regularization strength $\lambda$. We will discuss in detail how to choose $\lambda$ in section III D. Tikhonov regularization as discussed in refs. [13][19] for current reconstruction as well as the optimal Wiener filter[23] are equivalent to choosing a Gaussian prior with corresponding choices of $\Gamma$.

The simplest choice for $\Gamma$ is the identity $\Gamma = \mathbb{I}$. In this case, the prior favors a small-magnitude solution, which is not often physically motivated. If $\Gamma$ corresponds to derivatives, the prior prefers smooth solutions. The Laplacian $\Gamma = D_x^2 + D_y^2$, where $D_x^2/y$ are the second derivatives operator in the $x/y$-directions, is a common choice for image reconstruction problems and has been discussed in the context of current reconstruction.[13][19] The Laplacian is translation invariant, prefers small accumulated curvature, and its solution given in Eq. 9 can be efficiently computed in Fourier space.[13][19] In the following we refer to the corresponding cost function as the Gaussian Laplacian (GL) cost function $\ell_{GL}(g)$.

We can interpret $\ell_{GL}(g)$ by using $j = \nabla \times g \hat{z} = \partial_y g \hat{x} - \partial_x g \hat{y}$ and writing $\ell_{GL}(g)$ in the continuum limit:

$$\ell_{GL}(g) = -\lambda^2 \int d^2 r \left( \partial_x^2 g + \partial_y^2 g \right)^2$$

$$= -\lambda^2 \int d^2 r \ | \nabla \times j |^2.$$ (10)

The second line assumes that current only varies in the $x - y$ plane. We see that this cost function prefers solutions with small accumulated circulation of current. However, it is not clear why we should penalize only the circulation of current.

This leads us to more broadly explore choices for a Gaussian prior in the context of current reconstruction. For a prior that is chosen to suppress unphysical fluctuations in the reconstructed current density, it is desirable that the cost function is a functional of $g$ with physically-motivated symmetries and properties:

1. Invariance under current inversions $g \rightarrow -g$,
2. Invariance under rotations and reflections,
3. Equally penalizing all variations in currents, i.e. first derivatives of $j$ and thus second derivatives of $g$.

The cost function $\ell_{GL}$ satisfies almost all of these; it is the integral of $(\nabla^2 g)^2$; due to the quadratic it is invariant under $g \rightarrow -g$, and it is rotation and reflection invariant (see also below). However, following Eq. (10) we see that not all possible variations of the current are penalized.

We now derive a different cost function that satisfies all above constraints. The combination of the symmetry under current inversion and only allowing second derivative of $g$ constrains the functional to containing products of two elements of the second derivative matrix or so-called Hessian, $H$. This means that the prior must depend on the elements of the Hessian matrix $H_{\alpha,\beta} = \partial_\alpha \partial_\beta g$, where $\alpha \in \{x, y\}$ and $\beta \in \{x, y\}$. One can show that the only way to construct a rotation invariant functional from the Hessian is through contraction of the indices[22] (following Einstein summation conventions). There are only two ways to do this[22] for products of two elements of $H$: (1) $H_{\alpha\alpha} H_{\beta\beta} = (\text{Tr} H)^2$ and (2) $H_{\alpha\beta} H_{\alpha\beta} = \text{Tr} H^T H$, where $\text{Tr}$ denotes the trace of a matrix. The former is the Laplacian; the latter is the square of the Frobenius norm of the Hessian. This leads us to a new Gaussian Frobenius (GF) prior and corresponding cost function $\ell_{GF}(g)$, which satisfies our first two criterion by construction. Following a

Figure 1. Ground truth current density $|j|$ with uniform profile (a) and parabolic profile (e). The image size is 100×100 pixels. The magnetic images $\phi$ was calculated at a height of 4 pixel widths. Noise of $\sigma = 0.05$ was added. The reconstructed current density is shown for the uniform profile (b-d) and the parabolic profile (f-h) using a Gaussian prior penalizing the Laplacian (b,f), the Gaussian prior penalizing the Frobenius Hessian (e,g), and the total variation of the Frobenius Hessian (d,h). The regularization strength for each was chosen by the Bayesian discrepancy principle desribed in the main text. The data $\phi$ was re-scaled to have unit peak-to-peak range, and $\lambda = 1.4$ was used for the TVF prior reconstruction, and $\lambda = 2$ was used for the GL and GF priors.
the current density inside the uniform annulus. Therefore, a prior which penalizes oscillations without penalizing sharp changes in the solution is desirable. In general image reconstruction, this is achieved by a so-called total variation (TV) prior, which depends on the sum of the absolute values of the derivatives. Since we want to penalize derivatives of \( j \), we need to penalize second derivatives of \( g \). This leads us to a first possible choice for a total variation prior (ignoring momentarily the manifest violation of rotation invariance):

\[
\ell_{\text{TV}}(g) = -\lambda^2 \int \! d^2r \left| \partial_x^2 g \right| + \left| \partial_y^2 g \right|
\]

We can gain some intuition about this original total variation cost function \( \ell_{\text{TV}} \) by considering a one dimensional profile. Fig. 2(a) shows three hypothetical variations in current that are monotonic along one spatial direction. The TV prior considers all three equally probable regardless of their smoothness. The TV prior will suppress oscillations in the solution and, unlike Gaussian priors, remain agnostic to the sharpness of the transition. Fig. 2(b) shows both the GL and TV cost functions as functions of the second derivatives of \( g \). The GL cost function is more permissive of small variations of \( j \). In contrast, the absolute value of the TV causes any amount of variations \( j \) to be penalized. Therefore, \( \ell_{\text{TV}} \) prefers solutions of \( g \) with regions of constant \( j \) and allows sharp edges.

Figure 2. (a) Three different current \( j \) profiles which are equally probable under a total variation prior. (b) Illustration to compare Gaussian and total variation cost function. The Gaussian cost functions tolerates small variations in current, while the total variation cost function penalizes any non-zero amount.

In Sec. III we argued that a cost function should have several properties including rotational invariance. However, the TV cost function in Eq. 12 is not rotation invariant. We identified that ideally a cost function should be a functional of the Frobenius norm of the Hessian of \( g \): \( \text{Tr} H^T H \). This leads us to a TV Frobenius (TVF) cost function:

\[
\ell_{\text{TVF}}(g) = -\lambda^2 \int \! d^2r \sqrt{H_{\alpha\beta} H_{\alpha\beta}}
\]

\[
= -\lambda^2 \int \! d^2r \left( |\partial_x^2 g|^2 + |\partial_y^2 g|^2 + 2|\partial_x \partial_y g|^2 \right),
\]
where the square root of a sum of squares gives us a rotation invariant absolute value similar to the original TV cost function.

Fig. 3(d) and (h) show the result of the TVF prior on the reconstructions of the uniform and parabolic current annuli. In both reconstructions, the background is uniformly zero as desired, since \( \ell_{\text{TVF}} \) penalizes any variation and zero current (a constant \( g \) field) is effective at explaining the data. Likewise, in Fig. 3(d), the current density is more uniform in the interior of the annulus. In the case of a parabolic current density profile, the TVF prior (g) and Gaussian prior (h) perform comparably.

C. Finite Support Prior

In many experiments lithographically defined devices are imaged for which the geometry is known in detail. If the field of view of the magnetic image contains regions of the device in which no current can flow, it is desirable to use this knowledge to improve the current reconstruction. Here, we describe and implement a prior that enforces zero current density in pre-defined regions of the image. Regions with zero current density correspond to regions of constant current dipole field \( g \).

In order to impose regions with constant \( g \), we define an image mask based on the device that identifies regions with zero current. The image mask \( m \) is assumed to be of the same shape as \( \phi \). It takes the value 0 in regions where current is unrestricted and therefore \( g \) can vary. It takes the value 1 where the current is zero and therefore \( g \) is constant.

For each contiguous region in the mask with value 1, there is only one free parameter for the value of \( g \) in that region. However, regions that are not connected can have different values. The total number of free values in \( g \) is then reduced to the number of contiguous regions of 0’s in \( m \) plus the number of 1’s in \( m \). We define \( \tilde{g} \) as a vector containing all free parameters in \( g \). There is a linear operator \( F \) such that \( g = F\tilde{g} \), where \( F \in \mathbb{R}^{N \times P} \) for \( N \) pixels in the image plane and \( P \) free current dipole field parameters. \( F \) is given by

\[
F_{jk} = \begin{cases} 
1 & \text{if free parameter } \tilde{g}_k \text{ sets } g_j \\
0 & \text{else.} 
\end{cases}
\]

By replacing \( g \rightarrow F\tilde{g} \) in Eq. 5 we can impose regions with zero current as identified in the mask. We call this the ‘finite support’ (FS) prior. It reduces the number of degrees of freedom and highly constrains the solution space. Since it can be implemented with a linear operator, it is straightforward to include it with both Gaussian priors and TVF prior discussed above.

Fig. 3 shows the result of adding the FS prior to the reconstruction of the numerical examples studied in Fig. 1. The values of \( m \) outside and inside of the annulus are set to 1 and the interior of the annulus to 0. One step of binary erosion is added in order to model experimental uncertainty in aligning the data to the lithography.

D. Choosing the strength of the prior

There are several methods for choosing the strength of the prior, including Bayesian evidence maximization,\textsuperscript{29} the so-called \( l \)-curve method,\textsuperscript{30} cross-validation,\textsuperscript{29} and the discrepancy principle.\textsuperscript{25} Unfortunately, none of these methods works well for every prior. In fact, most methods require analytically tractable (Gaussian) priors. Since the TVF and FS priors are not analytically tractable, a more general method for setting the prior
strength is needed. Here, we describe a modification of the discrepancy principle.

Figure 4. (Top) Example of standard deviation of the reconstruction error std(∥M^g_\lambda - \phi∥^2) as a function of the strength of the prior \(\lambda\) for the annulus with uniform profile and the GF prior from Fig. 1(c). Fourier amplitudes of the residuals (middle panels) and real-space residuals (bottom panels) are shown for the values of \(\lambda\) as indicated by the symbols. The largest value of \(\lambda\) for which the real space residuals have no spatial structure is denoted by a blue star. The black cross marks the value of \(\lambda\) that satisfies the Bayesian discrepancy principle as discussed in the main text. In order from left to right, we demonstrate the spatial properties of the reconstruction error for over-fitting of noise (\(\lambda \approx 0\)), the result of the Bayesian discrepancy principal, under-fitting of the standard discrepancy principle, and finally an exaggeration of under-fitting.

Without a prior we can find a \(g\) for which the residuals \(r = M^g_\lambda - \phi\) are arbitrarily small despite noise in the image. However, the corresponding \(g\) will have non-physical properties such as rapid oscillations to reproduce the image including the experimental noise. The discrepancy principle is based on the observation that in an ideal situation the residuals should be given by the experimental noise. In its simplest form, the discrepancy principle therefore prescribes that the strength of the prior should be increased until the residuals have the same spectrum as the experimental noise. For Gaussian noise with variance \(\sigma^2\), this implies adjusting \(\lambda\) such that std(∥\(r\)∥^2) = \(\sigma\). In Fig. 4 (top) we show the standard deviation of the residuals using the GF prior as a function of \(\lambda\) for the annulus with uniform current density (see Fig. 1(a)). When inspecting images of the real-space residuals, we see that the residuals have spatial structure for the \(\lambda\) that fulfills the discrepancy principle (black cross). The residuals should be independently distributed noise. This indicates that this simplest choice of the prior strength is too strong and produces a too smooth solution, which (partially) fails to reproduce sharp features in the image. Fig. 4 (black hexagon) demonstrates the effect of an even stronger prior by showing the residuals for a larger \(\lambda\).

The discrepancy principle as described above can lead to over-smoothed solutions. The cause is that the finite height in the Biot-Savart Kernel leads to some loss of spatial information. Therefore the true number of degrees of freedom that determine the image is actually lower than the number of pixels in the image. This can be taken into account by modifying the discrepancy principle. Let \(\gamma < 1\) be such that \(\gamma N\) is the effective number of degrees of freedom, where \(N\) is the number of pixels. Then the strength of the prior should fulfill std(∥\(M^g_\lambda - \phi\)∥^2) = \(\gamma\sigma\).

For Gaussian priors the \(\gamma\) can in principle be estimated, but not for general priors.

We therefore adopt a more empirical method to determine the optimum strength of the prior. We find the regularization that satisfies the discrepancy principle, then reduce it until the residuals have minimal spatial structure in both real and Fourier space (as indicated by the blue star in fig. 4). Both are important: spatial structure in real space is indicative of a too strong prior, but for a too weak prior the residuals show no noticeable spatial structure - only their distribution becomes more narrow. The FFT of the residuals shows the distribution of residuals across spatial frequencies. We see an increasingly pronounced spatial structure for weak strength of the prior. As the strength of the prior is reduced, noise corresponding to increasing spatial frequencies is fitted by the reconstruction. Therefore in Fourier space we can see intensity in the residuals missing up to \(k\)-value that increases as the prior strength gets weaker. We refer to this modified discrepancy principle as Bayesian discrepancy principle, because it is guided by empirically identifying the most likely solution \(g_\lambda\).

In the example shown in Fig. 4 we identify the value of \(\lambda\) denoted by the blue star as the optimum prior strength. Some structure remains in the Fourier space residuals. In fact, there is no value of \(\lambda\) at which there is truly no structure in the real-space and the Fourier transformed residuals. This is because our model including the prior is imperfect. Finally, if we choose \(\lambda\) very small (indicated by the black triangle of fig. 4), we observe that the reconstruction fits the noise.

IV. MODELING CURRENTS OUTSIDE THE FIELD OF VIEW

A key assumption we have made so far is current conservation. We compute \(j\) from the current dipole field \(g\) as \(j = \nabla \times g\). Therefore, our model describes currents that circulate within the field of view. However, when imaging a device, current enters and leaves the field of view in some parts of the image. One way to accommodate data in which current crosses the field of view is to assume mirror boundary conditions. Using mirror boundary conditions is analytically tractable, however in most cases it is not faithful to the sample geometry. In
We subtract the resulting field the leads connected by a thin strip within the field of view. In order to remove currents crossing the field of view, we calculate the field from the current density in pink. In order to remove currents crossing the field of view, we assume uniform resistivity in the square, but our device has a uniform resistivity, we build a loop of the device are long and widen which is typical to the leads have a uniform resistivity, we build a loop which enters and leaves the field of view, canceling as much as possible the currents incident on the edges. As an example, we consider a device with a van der Pauw geometry (Fig. 5). The active region of interest in the device is the central square and we assume that the field of view is slightly larger than the square as shown. The leads of the device are long and widen which is typical to make it easy to attach wires to the device. We assume an externally applied current to flow from the top left to the top right lead. From this we calculate a current dipole field \( \mathbf{g} \) for the full device and an image \( \phi \). In our example, we assume uniform resistivity in the square, but our method works for any current density in the device within the field of view. We build a model (shown in pink) that includes the leads and a segment within the field of view connecting them. We compute a corresponding current dipole field \( \mathbf{g}_{\text{ext}} \) and magnetic image \( \phi_{\text{ext}} \) and subtract the latter from \( \phi \). The linearity of all equations guarantees that we will be only trying to recover the difference \( \mathbf{g} - \mathbf{g}_{\text{ext}} \), which should have conserved current in the field of view.

The external model requires extra variations in the optimal \( \mathbf{g} \) which will be penalized by the cost function. Therefore, the cost function needs to be modified to be a function of \( \mathbf{g} + \mathbf{g}_{\text{ext}} \). We modify Eq. 5 as follows:

\[
\mathbf{g}_{\lambda} = \mathbf{g}_{\text{ext}} + \min_{\mathbf{g}} \left[ \frac{1}{2} \| M \mathbf{g} - (\phi - \phi_{\text{ext}}) \|^2 + (\lambda \sigma)^2 \| \mathbf{g} + \mathbf{g}_{\text{ext}} \| \right].
\]  

(15)

where \( \ell \) is any cost function. Appendix Sec. V C explains how we accommodate these modifications for the TVF priors.

Using Eq. (15), we can apply this method to include current leads for the device shown in Fig. 5. Fig. 6(a) shows the current density in the field of view calculated for the entire device shown in Fig. 6 with voltage applied to the top two contact pads. For this simulation, we used a tool that solves simple resistor networks included in the pysquid package. From the current density, we calculate \( \phi \) assuming a height above the plane of 4 pixel and add noise of magnitude \( \sigma = 0.05 \). Fig. 6(b) shows a reconstruction using the GP prior (see Eq. 11), with mirror boundary conditions applied. The reconstruction shows some significant ringing in the current (see also Fig. 7). Figure 6(c) shows the reconstruction using the TVF prior of eqn. 13 with mirror boundary conditions. We find significant current density outside the boundaries of the device, which is an artifact of the mirror boundary conditions since the device lacks mirror symmetry. Finally, Fig. 6(d) shows the reconstruction using the external model of Fig. 15 with the TVF prior. This method significantly reduces the current density outside the boundaries of the device.

V. CONCLUSIONS

The reconstruction of a current density from magnetic images is an increasingly important problem as more local magnetic probes are developed and applied to a variety of materials and devices. Optimal and flexible reconstruction methods are desirable to achieve optimal spatial resolution in the reconstructed image for a given magnetic probe. Even for a point-like magnetic probe, the height dependence of the Biot-Savart law introduces blurring. Therefore, current reconstruction requires strong regularization for stability. We followed the methods of the literature [9-12], defining the current dipole field \( \mathbf{g} \) such that \( \mathbf{j} = \nabla \times \mathbf{g} \), and defined the Biot-Savart kernel \( M \) such that \( \phi = Mg \). Framing the reconstruction problem in terms of Bayesian inference, we defined the negative log-posterior in Eq. 6 the maximization of which provides a solution.

We discussed the importance of prior information, typically called regularization, and derived a new prior, the Frobenius of the Hessian, which improved the standard Gaussian prior. Real experimental data often contains sharp edges at which the current drops to zero (e.g. the device boundaries) and areas of zero current (corresponding to areas where the device is absent). To improve the reconstruction of regions of constant current, we investigated a total variation prior and contrasted it with...
Figure 6. (a) Calculated current density for the sample shown in Fig. 5 with uniform resistivity. Current enters and leaves at the top of the image. From this the magnetic field is calculated at a height of 4 pixel widths above the sample. Noise of $\sigma = 0.05$ was added. The reconstructed current density is shown using (b) the GF prior with mirror-symmetric boundary conditions, (c) the TVF prior with mirror boundary conditions and (c) the TVF prior using an external model. The bottom row shows the residuals between the reconstructed current density and the ground truth in (a). The data $\phi$ was re-scaled to have unit peak-to-peak range, and $\lambda = 0.9$ was used for both TVF prior reconstructions and $\lambda = 0.8$ was used for the GF prior.

Gaussian priors which permit unnecessary oscillations. To leverage information about the device geometry we developed a finite support prior which can enforce where currents are guaranteed to be zero. Finally, we described a procedure to use a model for the currents outside of the field of view in order to accommodate the violation of current conservation in images of realistic samples.

With the total variation and finite support priors we moved beyond analytically tractable priors. We discussed how to choose the strength of the regularization using a Bayesian discrepancy principle. This method relies on well-defined criteria, but does require manual inspection. We argued that this is not a disadvantage compared to analytical tractable priors, for which an optimal prior strength can be explicitly estimated. Requiring manual inspection forces the user to engage with the data and the residuals. This can help to identify the quality of the model including the prior that is used. The flexibility to easily use different priors will then allow the user to evaluate how robust features in the reconstructed current density are for different priors and other parameters of the reconstruction.

In future work, we will develop methods to infer the PSF of a given imaging device. In principle, the PSF can be obtained from imaging a known source of magnetic field, e.g., a magnetic dipole or a superconducting vortex. However, this is another ill-posed deconvolution problem in itself. The open-source pysquid code is already able to use a PSF into account. The package can be readily adapted to use additional priors. In particular, new opportunities arise from recent developments of machine learning such as deep priors and random projectors.

ACKNOWLEDGMENTS

The authors thank Veit Elser, Matthew Bierbaum, Charlie van Loan and G. M. Ferguson for insightful discussions. This work was primarily supported by the NSF (DMR-1719490). Support by the President’s Council of Cornell Women (PCCW) Affinito/Stewart Program grants is acknowledged as well.

VI. APPENDIX

A. Numerically Implementing the Biot-Savart Kernel on a Discrete Grid

In this section we derive the elements of the circulant matrix $M$ introduced in Eq. 2, which represents the convolution of the current dipole field with the Biot-Savart kernel. An image is defined through pixels arranged on a discrete, rectangular grid. We represent the sheet current density by rectangular pixels centered at $z = 0$ below the position at which we detect the magnetic field. Rectangles of constant $g$ correspond to rectangular loop of current present only at the edges. We assume that the $z$-component of the magnetic field is detected. For a pixel centered at $s_0 = (x_0, y_0, 0)$ with a value of $g = 1$ the $z$-component of the magnetic field at $r$ generated by
this pixel is given by
\[
B_z^1(r, s_0) = \frac{1}{4\pi} \int_{x_0 - \frac{a}{2}}^{x_0 + \frac{a}{2}} \int_{y_0 - \frac{b}{2}}^{y_0 + \frac{b}{2}} ds \frac{3z^2 - (r - s)^2}{|r - s|^5},
\]
where \(a\) and \(b\) are the widths of the rectangle. Following Ref. we find the magnetic field due to a rectangle of constant \(g\):
\[
B_z^r(r, s_0) = \frac{1}{4\pi} \left[ I(x_0 - x + \frac{a}{2}, y_0 - y + \frac{a}{2}, z) - I(x_0 - x - \frac{a}{2}, y_0 - y - \frac{a}{2}, z) - I(x_0 - x + \frac{a}{2}, y_0 - y - \frac{a}{2}, z) + I(x_0 - x - \frac{a}{2}, y_0 - y + \frac{a}{2}, z) \right],
\]
where we defined
\[
I(x, y, z) = \frac{xy(2z^2 + x^2 + y^2)}{(z^2 + x^2)(z^2 + y^2)}.
\]
Next we write the magnetic field at position \(\{x_k, y_i\}\) at height \(z\) above a current distribution as a sum over the contributions from each individual pixel:
\[
B_z(x_k, y_i, z) = \sum_i \sum_j B_z^i(x_k, y_i, z', y_j) g(x', y'),
\]
where \(g_{ij} = g(x_i, y_j)\) for some set of pixel centers \(\{x_i, y_j\}\). Since Eq. only depends on relative distances, we observe that Eq. 19 is a discrete convolution. This means that we can write our model more as \(b = Mg\), where \(b\) is the unraveled magnetic field image, \(M\) is a circulant matrix representing the discrete convolution, and \(g\) is our unraveled current dipole density image.

Because \(M\) is circulant it is diagonalized by plane waves, and matrix-vector products like \(Mg\) can be computed efficiently using Fourier transforms. A model for the PSF can in principle be included in the definition of \(M\), as two discrete convolutions are a sequence of multiplications in Fourier space.

**B. Defining Linear Operators to Numerically Implement Priors**

Here we discuss how we numerically implement the Gaussian Laplacian prior of \(\text{eqn.10}\), the Gaussian Frobenius prior of \(\text{eqn.11}\), and the TV Frobenius prior of \(\text{eqn.13}\). All priors contain partial derivatives \(\partial_x^2, \partial_y^2\), and \(\partial_x \partial_y\). To obtain discrete representations of these operators, we use finite-difference derivatives, encoded as a sparse matrix, where the interior of the image is computed using centered finite differences[21], and the edges use forward or backward finite differences, moving away from the edges. This way we can estimate the derivatives using only information that we have. As an example, we write the image \(g\) as a two-index matrix \(g_{x,y}\), and derive \(D_x^2\) explicitly. For a pixel not at the edge, and assuming that the distance between adjacent pixels is \(\Delta x\),
\[
(D_x^2)_{x',y',x,y} = \frac{\delta_{x',x-1} - 2\delta_{x',x} + \delta_{x',x+1}}{(2\Delta x)^2},
\]
whereas for example at the left edge,
\[
(D_x^2)_{0,y',x,y} = \frac{\delta_{0,x-2} - 2\delta_{0,x-1} + \delta_{0,x}}{(2\Delta x)^2}.
\]
Here \(\delta_{x',x}\) denotes the Kronecker delta. We can similarly write discrete linear operators \(D_x^2\) and \(D_y^2\) and the cross-derivative \(D_{xy}\). Products \((D_x^2 g)_i = \sum_j (D_x^2 g)_{ij} g_j\) are unraveled using appropriate \(i = (x, y)\) and \(j = (x, y)\). All implementations are included in the pysquid source code.

With these linear operators defined, we can write the discrete representations of the Gaussian Laplace priors given in Eq. 10 as
\[
\ell_{GL}(g) = -\lambda^2 \| \Gamma g \|^2, \tag{22}
\]
where \(\Gamma = D_x^2 + D_y^2\), and \(\| \cdot \|^2 = \sum_i \cdot_i^2\).

Next, we find the operator \(\Gamma\) representing the Frobenius of the Hessian, \(\text{Tr}H^T H\) where \(H_{\alpha \beta} = \partial_\alpha \partial_\beta g\). While the Laplacian Gaussian integrand in Eq. 10 is a square of a sum, the Frobenius prior is a sum of squares. In order to write this in terms of some operator \(\Gamma\), we need to stack the operators
\[
\Gamma = \begin{pmatrix} D_x^2 & D_x^2 \\ D_y^2 & \sqrt{2} D_{xy}^2 \end{pmatrix}
\]
Written this way, we can write the discrete generalization of the Gaussian Frobenius prior in Eq. 11 as
\[
\ell_{GF}(g) = -\lambda^2 \| \Gamma g \|^2 = -\lambda^2 (\Gamma g)^T \Gamma g = -\lambda^2 \left[ \sum_i (D_x^2 g)_i^2 + \sum_i (D_y^2 g)_i^2 + \sum_i (D_{xy}^2 g)_i^2 \right]. \tag{24}
\]
Finally, the total variation prior can be expressed by modifying Eq. 24 by taking a square root of the sum to obtain
\[
\log p(g) \propto -\lambda^2 \left[ \sum_i (D_x^2 g)_i^2 + \sum_i (D_y^2 g)_i^2 + \sqrt{\sum_i (D_{xy}^2 g)_i^2} \right]. \tag{25}
\]

**C. ADMM for Total Variation Deconvolution**

Alternating Difference Method of Multipliers (ADMM) is a convex optimization algorithm which
solves
\[
\min_{x,z} F(x) + G(z)
\]
subject to \(Ax + Bz = c\),
\[
F(x) + G(z) \equiv G(z)
\]
for some scalar functions \(F\) and \(G\), appropriately-sized matrices \(A\) and \(B\), and vector \(c\). The only requirement for ADMM to provably solve Eq. 26 is that \(F\) and \(G\) be convex in their arguments.

We can cast our problem defined in Eqn. 6 into the standard form of Eqn. 26 by identifying \(x = g\), and by setting \(F = M\), \(A = D_x^2\) and \(G = \phi\) to be the Hessian:
\[
\begin{pmatrix}
D_x^2 & D_y^2 \\
D_y^2 & D_{xy}^2
\end{pmatrix}
\]

Then the function \(G\) is the regularization term set by \(-\log p(z)\). Our isotropic total variation prior penalizes second derivatives of the \(g\)-field. Identifying \(A\) with the \(x\)- and \(y\)-derivative matrices \(D_x^2\) and \(D_y^2\) and \(D_{xy}^2\):
\[
A = \begin{pmatrix}
D_x^2 \\
D_y^2 \\
D_{xy}^2
\end{pmatrix}
\]

setting \(B = -I\), and \(c = 0\), the constraint \(Ag + Bz = c\) is equivalent to
\[
\begin{pmatrix}
D_x^2 & D_y^2 \\
D_y^2 & D_{xy}^2
\end{pmatrix} g = z = \begin{pmatrix} z_x \\ z_y \\ z_{xy} \end{pmatrix},
\]
where \(z\) is twice as long as \(g\), containing both the horizontal and vertical second derivatives of \(g\). The final piece is then the total variation of the Frobenius norm of the Hessian:
\[
G(z) = \lambda^2 \sum_i \sqrt{z_{x,i}^2 + z_{y,i}^2 + z_{xy,i}^2}.
\]

We can modify ADMM to include finite support priors by replacing \(g \rightarrow F_\theta g\), and optimizing \(\theta\) instead of \(g\):
\[
\min_{g,\theta} F(F_\theta g) + (z) 
\]
subject to \(AF_\theta g = z\),
\[
(31)
\]
where \(F\) is still as in Eq. 27, \(G\) is as in Eq. 30, and \(A\) is as in Eq. 28. Equivalently, we can modify the kernel matrix \(M \rightarrow MF\) and the second-derivative matrix \(A \rightarrow AF\). The latter is how it is implemented in the code.

To use an external model as in Eq. 15, we modify ADMM by setting \(c = -Ag_{\text{ext}}\) for the external model current dipole field in the field of view \(g_{\text{ext}}\).

---

1. Beena Kalisky, Eric M. Spanton, Hilary Noad, John R Kirtley, Katja C. Nowack, Christopher Bell, Hiroki K. Sato, Masayuki Hosoda, Yanwu Xie, Yasuyuki Hikita, et al., “Locally enhanced conductivity due to the tetragonal domain structure in LaAlO3/SrTiO3 heterointerfaces,” Nature Materials 12, 1091–1095 (2013).
2. Katja C. Nowack, Eric M. Spanton, Matthias Baenninger, Markus König, John R. Kirtley, Beena Kalisky, Christopher Ames, Philipp Leubner, Christoph Brüne, Hartmut Buhmann, et al., “Imaging currents in hgte quantum wells in the quantum spin hall regime,” Nature Materials 12, 787–791 (2013).
3. Denis Vasyukov, Yonathan Analhory, Lior Embon, Dorri Halbertal, Jo Cuppens, Lior Neeman, Amit Finkler, Yehonathan Segev, Yuri Myasoedov, Michael L. Rappaport, Martin E Huber, and Eli Zeldov, “A scanning superconductive quantum interference device with single electron spin sensitivity,” Nature Nanotechnology 8, 639–644 (2013).
4. Rafael B. Dinner, Kathryn A. Moler, M. R. Beasley, and D. Matthew Feldmann, “Enhanced current flow through meandering grain boundaries in YBa2Cu3O7−δ films,” Applied Physics Letters 90, 212507 (2007).
5. A. E. Pashitski, A. Gurevich, A. A. Polyanskii, D. C. Larbalestier, A. Goyal, E. D. Specht, D. M. Kroeger, J. A. DeLuca, and J. E. Tkaczyk, “Reconstruction of current flow and imaging of current-limiting defects in polycrystalline superconducting films,” Science 275, 367–369 (1997).
6. K. Chang, A. Eichler, J. Rhensius, L. Lorenzelli, and C. L. Degen, “Nanoscale Imaging of Current Density with a Single-Spin Magnetometer,” Nano Letters 17, 2367–2373 (2017).
7. Jean-Philippe Tetienne, Nikolai Dontschuk, David A. Broadway, Alastair Stacey, David A. Simpson, and Lloyd C. L. Hollenberg, “Quantum imaging of current flow in graphene,” Science Advances 3, e1602429 (2017).
8. Mark J. H. Ku, Tony X. Zhou, Qing Li, Young J Shin, Jing K. Shi, Claire Burch, Huiling Zhang, Francesco Cascola, Takashi Taniguchi, Kenji Watanabe, Philip Kim, Amir Yacoby, and Ronald L. Walsworth, “Imaging viscous flow of the Dirac fluid in graphene using a quantum spin magnetometer,” Nature 583, 537–541 (2020).
9. L. Thiel, D. Rohner, M. Ganzhorn, P. Appel, E. Neu, B. Müller, R. Kleiner, D. Koelle, and P. Maletinsky, “Quantitative nanoscale vortex imaging using a cryogenic quantum magnetometer,” Nature Nanotechnology 11, 677–681 (2016).
10. Fan Yang, Alicia J. Kollár, Stephen F. Taylor, Richard W. Turner, and Benjamin L. Lev, “Scanning Quantum Cryogenic Atom Microscope,” Physical Review Applied 7, 034026 (2017).
11. Fan Yang, Stephen F. Taylor, Stephen D. Edkins, Johanna C. Palmstrom, Ian R. Fisher, and Benjamin L. Lev, “Nematic transitions in iron pnictide superconductors imaged with a quantum gas,” Nature Physics 16, 514–519 (2020).
12. D.A. Broadway, S.E. Lillie, S.C. Scholten, D. Rohner, N. Dontschuk, P. Maletinsky, J.-P. Tetienne, and L.C.L. Hollenberg, “Improved current density and magnetiza-
tion reconstruction through vector magnetic field measurements,” Phys. Rev. Applied 14, 024076 (2020).

13 Alexander Y Meltzer, Eitan Levin, and Eli Zeldov, “Direct reconstruction of two-dimensional currents in thin films from magnetic-field measurements,” Physical Review Applied 8, 064030 (2017).

14 Bradley J Roth, Nestor G Sepulveda, and John P Wikswo Jr, “Using a magnetometer to image a two-dimensional current distribution,” Journal of Applied Physics 65, 361–372 (1989).

15 Rinke J Wijngaarden, K Heeck, HJW Spoelder, R Surdeanu, and R Griessen, “Fast determination of 2d current patterns in flat conductors from measurement of their magnetic field,” Physica C: Superconductivity 295, 177–185 (1998).

16 D. M. Feldmann, “Resolution of two-dimensional currents in superconductors from a two-dimensional magnetic field measurement by the method of regularization,” Physical Review B 69, 144515 (2004) [0309678]

17 Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, Jonathan Eckstein, et al., “Distributed optimization and statistical learning via the alternating direction method of multipliers,” Foundations and Trends® in Machine learning 3, 1–122 (2011).

18 https://github.com/colinclement/pysquid

19 Rinke J Wijngaarden, HJW Spoelder, R Surdeanu, and R Griessen, “Determination of two-dimensional current patterns in flat superconductors from magneto-optical measurements: An efficient inversion scheme,” Physical Review B 54, 6742 (1996).

20 David JC MacKay, “Bayesian interpolation,” Neural computation 4, 415–447 (1992).

21 William H Press, Brian P Flannery, Saul A Teukolsky, William T Vetterling, et al., Numerical recipes, Vol. 3 (Cambridge University Press, Cambridge, 1989).

22 Predrag Cvitanović, Group theory: birdtracks, Lie’s, and exceptional groups (Princeton University Press, 2008).

23 David Gottlieb and Chi-Wang Shu, “On the gibbs phenomenon and its resolution,” SIAM review 39, 644–668 (1997).

24 Curtis R Vogel and Mary E Oman, “Fast, robust total variation-based reconstruction of noisy, blurred images,” IEEE transactions on image processing 7, 813–824 (1998).

25 Stanley Osher, Martin Burger, Donald Goldfarb, Jinjun Xu, and Wotao Yin, “An iterative regularization method for total variation-based image restoration,” Multiscale Modeling & Simulation 4, 460–489 (2005).

26 Per Christian Hansen, “Analysis of discrete ill-posed problems by means of the l-curve,” SIAM review 34, 561–580 (1992).

27 Gene H Golub, Michael Heath, and Grace Wahba, “Generalized cross validation as a method for choosing a good ridge parameter,” Technometrics 21, 215–223 (1979).

28 Nikolas P. Galatsanos and Aggelos K. Katsaggelos, “Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation,” IEEE Transactions on image processing 1, 322–336 (1992).

29 Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky, “Deep image prior,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2018) pp. 9446–9454.

30 Sidharth Gupta, Konik Kothari, Maarten V. de Hoop, and Ivan Dokmanić, “Random mesh projectors for inverse problems,” arXiv:1805.11718 (2018).