Utilizing Random Matrix Theory and Subspace Methods for Solving Ill-posed Inverse Problems

Amartansh Dubey
ECE Department, HKUST
adubey@connect.ust.hk

Samruddhi Deshmukh
ECE Department, HKUST
ssdeshmukh@connect.ust.hk

Nilesh Kumar Jha
ECE Department, HKUST
nkjha@connect.ust.hk

Abstract

This project is the continuation of our mini-project where we presented subspace methods to tune the regularization parameter for solving an ill-posed inverse problem. We have briefly defined the problem again in this report for the sake of completion: An ill-posed inverse problem is the one where system is under-determined (knowns $\ll$ unknowns) and unstable (where small perturbations in measurements can lead to large changes in the solution). A common approach is to impose priors on the variables using regularization to obtain a stable solution. This requires finding optimum regularization parameter/s (λ). A widely used approach to tune λ is *cross-validation* where λ is varied in an interval and its optimum value is selected using a validation dataset. It is a supervised way of tuning λ which might fail if the test data deviates from the distribution followed by training and validation data. This often happens in the field of electromagnetic imaging where a scatterer (object of interest) can take any possible shape, size, location and physical properties (like refractive index) and it is difficult to create a comprehensive training/validation dataset. In the mini-project, we proposed a simple but effective PCA and Sparse-PCA based unsupervised approach to tune λ that does not need any validation dataset. In this report, we propose three more techniques to tune λ based on Random Matrix Theory (RMT). The idea is to reformulate the regularization problem in a way that 1) makes the searchable domain of λ very narrow or 2) utilizes asymptotic results from RMT to remove the need to tune λ at the first place. To demonstrate the performance of these techniques, we have used data from Radio imaging (using electromagnetic simulations of wave scattering). We also compare our results to those obtained using the L-curve method which is another commonly used unsupervised approach for tuning λ and also with our previously proposed approaches in the mini-project.

1 Introduction

Ill-posed problems occur in a wide spectrum of research, commercial and medical fields [Chen (2018); Tearney et al. (1995); Lavrent_ev et al. (1986)]. The major challenge in solving an ill-posed problem is to achieve a unique and stable solution. Such under-determined systems are usually solved by imposing priors (like sparsity) which reduce the effective number of unknowns compared to known observations [Friedman et al. (2001)]. These approaches can be categorized into two classes, 1) maximum a posteriori probability (MAP) estimation and 2) Bayesian estimation. In this work, we focus on MAP estimation. The widely used regularization methods including LASSO, Ridge, Total Variation and Elastic Net are all MAP estimates.

33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Vancouver, Canada.

A general inverse problem can defined in form of an optimization problem as

$$\begin{aligned} \min_{x} \quad & f(y - Ax) \\ \text{s.t.} \quad & f_i(x) < 0,\ i = 1, 2, 3... \end{aligned} \tag{1}$$

where, $y \in \mathbf{C}^{m\times 1}$ is a measurement vector (m known observations) and $x \in \mathbf{C}^{p\times 1}$ is the unknown variable (p unknowns). $A \in \mathbf{C}^{m\times p}$ is a model matrix which transforms the variable vector into the measurement space. Functions f and f_i can be convex or non-convex (in this report we assume them to be convex). Matrix A depends on the domain of the inverse problem and encapsulates the physical behavior of the system under test. When f is simply an l_2 norm and there are no constraints in (1), we can estimate x as follows: When $m = p$ and A is full rank, x can be estimated as $x = A^{-1}y$. For an over-determined system ($m >> p$), the least square estimate (LSE) can give unique and stable solution as $x = (A^H A)^{-1} A^H y$. The challenging part is when the problem is under-determined ($m << p$). In this case the problem is solved by enforcing priors using regularization as

$$\underset{x}{\text{minimize}} \quad ||(y - Ax)||_2^2 + \lambda f(x) \tag{2}$$

where f helps impose priors on x. The selection of f provides a wide range of regularization techniques, for example, $f(x) = ||x||_2^2$ (Ridge), $f(x) = ||x||_1$ (LASSO) and $f(x) = \sum_{\forall i} |x_{i+1} - x_i|$ (Total Variation).

1.1 Brief Introduction to Inverse Scattering Problem

The detailed physics of the imaging problem is given in the mini-project report [Group-7 et al. (2021)]. Here we have given a brief summary: The goal is to reconstruct the shape of any object kept in a indoor region using wireless signals. The wireless signals are transmitted and received using WiFi sensors operating at 2.4 GHz frequency. This is very similar to security scanners in airports which use X-rays or biomedical imaging like ultrasound.

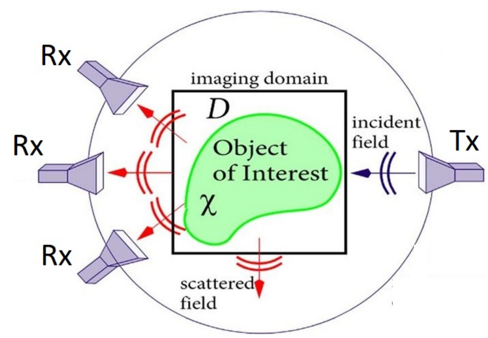


Figure 1: Illustration of General Radio Imaging Setup. Tx and Rx denotes Transmitting and Receiving wireless sensors respectively.

The inverse problem is to find the image of the scattering object given the received signals and can be stated as minimizing the least square error $||y - Ax||^2$, where the measurement vector $y \in \mathbf{R}^{m\times 1}$ is the WiFi signal power at receiver Rx, the model matrix $A \in \mathbf{R}^{m\times p}$ is derived in [Depatla et al. (2015)] and the unknown image variable $x \in \mathbf{R}^{p\times 1}$ which needs to be estimated is the contrast in refractive index ν of the object with respect to air. This forms a linear inverse problem where $y \approx Ax$ is the given system of equations with x being unknown. In this report, we used our forward research models to accurately generate y and A and then solved the corresponding inverse problem to find x when the problem is severely under-determined ($m << p$). Given y and A, x can be estimated using regularized optimization in (2) which requires careful tuning of regularization parameter λ.

1.2 Common Methods to Tune Regularization Parameter (λ)

The value of λ trade-offs between extent of prior being enforced on x and actual data fitting using LSE term. Since, λ is a hyper-parameter, there is no unique solution to (2) and accuracy will depend on the choice of λ. The most common way to tune λ^* is by cross validation, which simply uses a validation dataset to find appropriate value of λ which is then used for the test data. This is effective when enough data is available for training, validation and testing purposes. However, if the data is insufficient or there is a high probability that test data can be out of the validation data distribution, cross-validation will not give good results. This is often true for inverse imaging problems where the object of interest can take any possible shape, size, location, physical properties. For such cases, it is better to solve (2) as a standalone optimization problem by only relying on the given measurement vector y and model matrix A (and not relying on creating training, validation datasets using collection of labeled pairs of measurements and variables). For this, we need *unsupervised* learning approach for tuning λ instead of *supervised* cross-validation method.

One such unsupervised approach is the L-curve method which does not need additional data [(Hansen and O'Leary, 1993; Calvetti et al., 2000)]. Instead, it uses trade-off curve between LSE residual term ($||(y - Ax)||_2^2$) and regularized term $f(x)$ for different regularization parameters λ and then selects the value where the trade-off curve has a sharp knee point (more details of L-curve can be found here [(Hansen and O'Leary, 1993)]). Therefore, *L-curve* method can find λ in a unsupervised manner whereas *cross validation* requires additional labeled data. This helps in generalizing the solution for out-of-training datasets and hence L-curve method is preferred in ill-posed inverse problems in physics.

It has been shown that for severely under-determined large dimensional problems with large measurement noise, L-curve may not guarantee a good selection of λ as shown in [(Xu et al., 2016)]. Furthermore, L-curve method may also give sub-optimal value of λ if variable vector x has intricate correlations (temporal or spatial) and the model matrix is an approximation to actual underlying non-linear model (as we shall show this later in results).

In this report, we propose three techniques to tune λ based on Random Matrix Theory (RMT). The idea is to reformulate the regularization problem in a way that 1) searchable domain of λ becomes very narrow and/or 2) utilize asymptotic results from RMT to remove need of tuning λ at the first place.

2 Problem Formulation

In this work, we focus on a special class of MAP estimator known as Tikhonov regularization. The general form of Tikhonov reg. can be given as

$$\underset{x}{\text{minimize}} \quad ||(y - Ax)||_2^2 + \lambda||\Theta x||_2^2 \tag{3}$$

where Θ is the Tikhonov matrix which operates on the variable vector. If we select $\Theta = I_{[p\times p]}$, it reduces (3) to the well-known Ridge regression (l_2 regularization). Several forms of Θ operators are proposed to enforce desired priors such as sparsity and smoothness (e.g. finite difference or weighted Fourier operator) [Friedman et al. (2001)]. Therefore, Tikhonov is not one regularization method, but a family of regularization methods with different possible priors. Next we try to relate the stability of Tikhonov Reg. to the scaled Sample Covariance Matrix (SCM) ($\Sigma = A^T A$, assuming original matrix A is zero mean) as it will help connect it with the RMT results. Following four equivalent analytical forms can be solution to (3),

[general form]
$$x^*(\lambda) = (A^T A + \lambda \Theta^T \Theta)^{-1} A^T y \tag{4}$$

[In terms of SCM and Shrinkage Target]
$$x^*(\lambda) = (\Sigma + \lambda T)^{-1} A^T y \tag{5}$$

[In terms of Shrinked SCM]
$$x^*(\lambda) = (\Sigma_T(\lambda))^{-1} A^T y \tag{6}$$

[In terms of eigenvalues of SCM]
$$x^*(\lambda) = \sum_{i=1}^{P} \left(\frac{e_i}{e_i + \lambda}\right) \frac{u_i^T y}{\sqrt{e_i}} v_i \tag{7}$$

All the three expressions in (4)-(7) are equivalent. Σ in (5) is the scaled SCM, $T = \Theta^T\Theta$ in (5) is the target matrix, $\Sigma_T = \Sigma + \lambda T$ in (6) is the linear combination of SCM (Σ) and the target matrix T. Finally, e_i in (7) are sample eigenvalues of SCM and u_i and v_i are the left and right singular vectors respectively of model matrix A. Solution in (7) is for when target T in (6) is an identity matrix.

The key interpretation of the aforementioned analytical forms is the manipulation of covariance matrix (or its eigenvalues). For an ill-posed (under-determined) system, the covariance matrix of A will be badly conditioned with either some eigenvalues either being equal to zero or close to zero. Therefore, in the least square solution, inverse of covariance matrix, i.e. $(A^T A)^{-1}$ will either does not exist or will be poorly scaled. This will lead to completely misleading results. The Tikhonov reg. solves this problem which can be explained using following effects:

1. Taking linear combination of the high variance $A^T A$ with a highly structured target λT achieves better estimate of covariance matrix or helps in cleaning the noise in SCM.

2. In terms of eigenvalues, the operation $A^T A + \lambda T$ implies that we are pulling up the small eigenvalues which were close to zero to a constant value decided by λ as shown in denominator of (7). This means bulk of eigenvalues near zero will become larger and hence inverse of this new covariance matrix will be stable.

3 Proposed Method: Cleaning SCM using tools from RMT

To summarize the previous section, a key interpretation of Tikhonov regularization is that it is cleaning or improving a poorly-conditioned SCM by taking its linear combination with a highly structured matrix. This is where we can utilize tools from RMT.

Random Matrix Theory (RMT) is an evolving field of mathematics which deals with estimating behavior of eigenvalues of large random matrices and can help clean SCM by cleaning noise in sample eigenvalues. The elegance of RMT lies in the fact that the eigen-spectrum of random matrices can converge to highly deterministic distributions and limits (similar to Central limit theorem in probability theory). In this section we explore two results from RMT - 1) Shrinkage Transformations 2) MP law

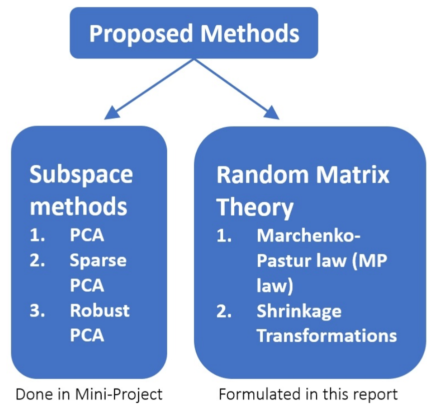


Figure 2

3.0.1 Using Shrinkage Transformations to improve Tikhonov Regularization

Shrinkage transformations shrink extreme eigenvalues of SCM towards their mean value by taking a convex combination of SCM with a highly structured target matrix (such as the identity matrix). Linear Shrinkage differs from Tikhonov regularization in two ways - 1) As opposed to Tikhonov which uses linear combination of SCM with another matrix, Linear Shrinkage uses a convex combination, and 2) The choice of matrices also differs. Since Linear Shrinkage uses a convex combination of SCM with a target matrix, it adds bias to the SCM by pulling small eigenvalues up and large eigenvalue down towards the center of the eigenvalue spectrum. This has shown to significantly reduce the sampling noise in SCM and it also makes SCM well-conditioned [Ledoit and Wolf (2004, 2017)]. The general expression for linear shrinkage (LS) estimator can be written as convex combination of SCM (Σ) with the target matrix F, as

$$\Sigma_{LS} = \alpha F + (1 - \alpha)\Sigma \tag{8}$$

where $\alpha \in [0, 1]$ is the shrinkage intensity. The **choice of shrinkage target** F is based on certain prior knowledge about the true covariance matrix and therefore varies based on the domain of interest. Several shrinkage targets have been proposed over the years. The simplest target is the identity matrix scaled by a constant ($F = \mu I$, where μ is the average of all sample variances). Other proposed target includes a sample variance zero covariance target where F contains sample variances as diagonal elements and average of all covariances as off-diagonal elements [Ledoit and Wolf (2004)]. Other similar targets are proposed, for example using sample variance along diagonal and average covariance along diagonal [Ledoit and Wolf (2004)]. We used Sample variance-Zero Covariance target from now-onwards in this work as it has been shown to be highly effective in handling poorly conditioned SCM [Ledoit and Wolf (2004)]. More advance shrinkage operators are also proposed like non-linear shrinkage [Ledoit and Wolf (2017)] which uses asymptotic results from RMT (non-linear shrinkage is out of scope of this work).

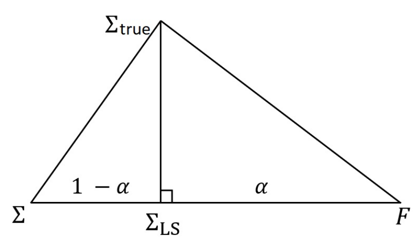


Figure 3: Linear Shrinkage as a projection in Hilbert space

An interesting interpretation of linear shrinkage is in the form of a projection theorem in Hilbert space [Ledoit and Wolf (2004)] as shown in Fig. 3. Assuming F and Σ both lie in an m dimensional space of symmetric random matrices, Σ_{LS} can be seen as the projection of Σ on the line joining F and Σ. Whether Σ_{LS} ends up being closer to F or Σ depends on which one of them is closer to the true covariance matrix Σ_{true}.

Unlike the Tikhonov approach, here shrinkage target matrix is directly designed instead of taking covariance of Tikhonov matrix Θ (and most importantly *convex combination* is used instead of *linear combination*). Shrinking towards these highly structured targets using convex combination can impose prior on the SCM, forcing its extreme covariance cofficients (or sample eigenvalues) to be pulled towards center. This helps in adding bias to the SCM and hence pulls up small eigenvalues towards center of the eigenvalue spectrum and pulling down the extremely large eigenvalues. This has shown to significantly reduce the sampling noise in SCM and also it makes SCM well-conditioned [Ledoit and Wolf (2004, 2017)].

Using (8) we can reformulate original analytical solution to Tikhonov formulation in (5) as,

$$x^*(\lambda) = ((1-\lambda)\Sigma + \lambda F)^{-1} A^T y, \text{ such that } \lambda \in [0,1] \tag{9}$$

The structure of F can be designed as per the type of data and prior information about the model (we have selected Sample variance-Zero Covariance target as F as explained above). Most important difference in (5) and our formulation in (9) is the range of regularization parameter λ. In original Tikhonov reg., $\lambda \in [0, \lambda_{max}]$ and λ_{max} can be arbitrarily large depending on the data. In fact, for most forms of Tikhonov reg, there is no analytical way to estimate λ_{max} and in that case, we just have to set λ_{max} to an arbitrarily large value to achieve maximum imposition of prior like sparsity on the variables. Therefore, to tune λ in original Tikhonov reg. (5), we have to sample many values of λ from interval $[0, \lambda_{max}]$ and use L-curve or cross validation to select any one value. Furthermore, the solution can be highly sensitive to values of λ and hence we have to sample a sufficiently large number of values. On the other hand, when using convex combination instead of linear combination as proposed in (9), the range of λ is limited to just [0,1]. This proves to be extremely advantageous as upper bound on λ is well defined and equal to 1. We show later in results, the L-curve of proposed method (9) gives more accurate information than the L-curve of the original Tikhonov reg. in (5) and it can be further improved with the right choice of F.

3.0.2 Marchenko-Pastur law (MP law)

RMT provides tools to predict the distribution of eigenvalues of large random matrices under asymptotic assumptions ($m, n \to \infty, c = m/n < 1$) [Marčenko and Pastur (1967)]. The elegance of these results lies in the fact that completely random behavior of large matrices can converge to highly deterministic behavior, similar to well known results in probability theory like the central limit theorem. Techniques based on MP law clean the eigenvalues of SCM unlike Linear Shrinkage which directly operates on the SCM. The end result of both RMT and Linear Shrinkage estimators are cleaned (or shrunk) sample eigenvalues and hence better covariance estimates and well-conditioned covariance matrix.

MP law can be stated as: Let $X \in \mathbb{R}^{m \times n}$ be a random matrix with $m, n \to \infty$, and $m/n \to c \in (0, 1]$. Let entries of X are jointly independent real random variables (with zero mean and variance 1). Further assume that the kth moment of entries is bounded, independent of m (As taught in class, first 8 moments should be finite). Under these assumptions, the empirical eigenvalue distribution $f(\tilde{\lambda})$ of XX^T/n converges to MP law given as,

$$\begin{aligned} f(\tilde{e}) &= \frac{\sqrt{(e_+ - \tilde{e})(\tilde{e} - e_-)}}{2\pi c\tilde{e}} \; \mathbf{1}_{[\mathbf{e_+}, \mathbf{e_-}]} \\ e_+ &= (1+\sqrt{c})^2 \\ e_- &= (1-\sqrt{c})^2 \end{aligned} \tag{10}$$

where, $\mathbf{1}_{[\mathbf{e_+}, \mathbf{e_-}]} = \mathbf{1}$ for $\tilde{e} \in [e_+, e_-]$, else equal to zero. The MP law bounds (e_- and e_+) represents the lower and upper bounds of eigenvalues that cannot be differentiated from sampling noise.

As explained earlier, the main problem in achieving the least square estimate of an ill-posed inverse problem is the presence of a non-invertible covariance matrix of model matrix A in the analytical solution. We can use our knowledge of the MP law bounds to clean the covariance matrix and make it comparatively well-conditioned. This can be done as follows:

1. Calculate the sample covariance matrix of A as $\Sigma = \frac{A^T A}{n}$ where A is an appropriately scaled and centered matrix.
2. Obtain the sample correlation matrix C of Σ as $C = D^{-1/2} \Sigma D^{-1/2}$ where D represents the diagonal matrix of variances, i.e. its diagonal elements are the diagonal elements of Σ.

3. Find sample eigenvalues $e_i, i = 1, 2, ..., p$ of C.
4. Calculate the MP law upper bound e_+.
5. Replace all sample eigenvalues of C below e_+ with their mean value e_{mean}

$$e_i^{clip} = \begin{cases} e_i & \text{if } e_i > e_+ \\ e_{mean} & \text{otherwise} \end{cases} \tag{11}$$

   such that $tr(C_{clip}) = tr(C)$
6. Reconstruct the clipped covariance matrix $\Sigma_{MP} = D^{1/2} C_{clip} D^{1/2}$

We can then use this clipped covariance matrix Σ_{MP} in the least square analytical solution as,

$$x^* = (\Sigma_{MP})^{-1} A^T y \tag{12}$$

Through this technique, extremely small eigenvalues which are close to zero get pulled up. This in turn makes the resultant covariance matrix Σ_{MP} well-conditioned. In other words, we are assuming that smaller eigenvalues (below MP law upper bound) are just representing sampling noise and can be removed. This technique has been used in finance and is also known as Eigenvalue Clipping [Bun et al. (2017)].

IMPORTANT NOTE: The solution obtained using Eigenvalue Clipping in (12) does not need any tuning of regularization parameter λ. It only depends on the MP law upper bound which in turn only depends on the dimensionality constant c of the problem. Therefore, our proposed solution using Linear Shrinkage in (9) makes tuning of λ easier by shrinking its domain in tight bounds and our second proposed solution in (12) removes need of regularization parameter λ altogether.

4 Numerical Analysis and Results

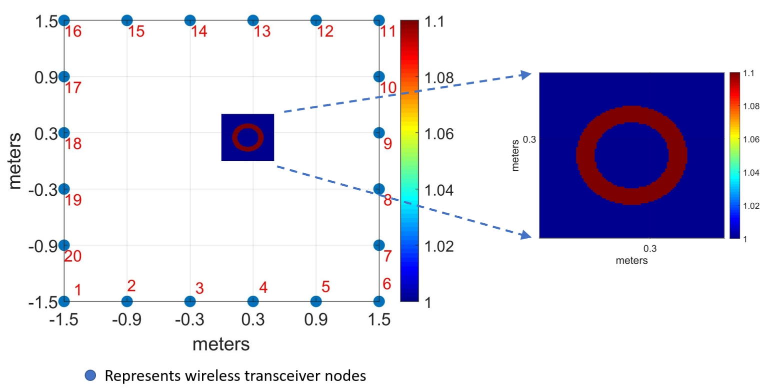


Figure 4: Illustration of Imaging Domain of interest (DOI). The DOI is 3 x 3 m^2 in area and 20 wireless sensors are placed around this DOI (see blue circles placed at the boundary of DOI, numbered from 1 to 20). On right, is the magnified region which consist of the ring shaped scatterer with refractive index = 1.1.

We use similar problem as used in mini project [Group-7 et al. (2021)] to evaluate our new proposed methods. In this section we solve an ill-posed inverse scattering imaging problem (explained in section 1.1) to demonstrate how our proposed methods can provide a stable solution with minimal or no tuning required for λ. Fig. 4 (similar to Fig. 1) shows a geometric setup where a concentric ring shaped scattering object is placed at the center of a DOI. DOI is of dimension 3 x 3 m^2. 20 wireless transceiver sensors are placed at the boundary of the DOI. When one node transmit 2.4 GHz radio waves towards the DOI, the object inside the DOI scatters the waves and these scattered waves are received by other sensors in the receiving mode. There are total $m = 20 \times (20-1)/2 = 190$ unique link (from transmitting to receiving nodes) and hence number of measurements are $m = 190$.Therefore measurement vector y is a 190×1 vector (exact physical process behind measuring y is not relevant to this work but interested readers can see [Group-7 et al. (2021); Depatla et al. (2015)]). The aim is to image part of the DOI where scatterer is located as shown in the magnified area of 0.5 x 0.5 m^2 on right in Fig. 4. For imaging, we discretized the 0.5 x 0.5 m^2 DOI into 50×50 small pixels (total

$p = 2500$ pixels). Each of these pixels has an associated refractive index value which needs to be estimated.

The inverse problem here is to estimate refractive index of all 2500 grids using just 190 measurements. Therefore, knowns are $m = 190$ and unknowns are $p = 2500$ which makes it a severely under-determined problem. The model matrix A is a 190×2500 matrix (derivation of matrix is not relevant here, but interested readers can see [Group-7 et al. (2021); Depatla et al. (2015)]).

NOTE: As shown in previous sections, we formulated our proposed techniques for general form of Tikhonov reg. in (3). In this section, we selected the Tikhonov target matrix T to be identity which gives special limiting case of Tikhonov reg. known as norm-2 regularization. However it is very important to note that our proposed formulations are valid for any choice of T as no assumptions were imposed while deriving it. We selected norm-2 regularization as it gives a lot of noisy coefficients in the reconstruction of DOI profile and testing the proposed methods against highly noisy conditions gives better idea about the performance.

We need some metric to compare the reconstructed image to actual ground truth image of object. We used state of the art image quality assessment metric called the structural similarity index measure (SSIM) [MATLAB R2020a (2014)]. It is shown to be more robust than traditional metric such as mean square error (MSE). SSIM value ranges from 0 to 1. If SSIM is close to 1, it means reconstructed image is exactly similar to ground truth and if SSIM = 0, than no similarity is found. SSIM can be estimated directly using Matlab package [MATLAB R2020a (2014)].

We compared various existing and proposed methods which are summarized in the Fig. 5. All methods uses $m = 190$ measurements to find $p = 2500$ unknowns. Note that any solution from any of the methods will be a vector $x^* \in \mathbf{R}^{p\times 1}$. This estimated vector can be reshaped into $\sqrt{p} \times \sqrt{p}$ 2D image of DOI (where p = 2500, hence image will be 50x50 pixels). For traditional Tikhonov (norm-2 regularization), easiest way to select λ_{max} is to make it large enough such that most coefficients tends to zero. For our inverse scattering problem, we select λ_{max} such that more than 90% of coefficients of $x^*(\lambda_{max})$ are below 10^{-5}. (NOTE: pixels having zero values are shown as pixel value of one in Fig. 4 because actual refractive index image is written as $= 1 + x^*$).

Methods	Analytical solution	Tuning for λ
Traditional Tikhonov Regularization	$\boldsymbol{x}^*_{\boldsymbol{Tikh}}(\boldsymbol{\lambda}) = (A^T A + \lambda\, T)^{-1} A^T y$	Need to tune λ in range $[0, \lambda_{max}]$
Proposed modified Tikhonov using shrinkage transformation	$\boldsymbol{x}^*_{\boldsymbol{ST}}(\boldsymbol{\lambda}) = \left((1-\lambda)A^T A + \lambda\, F\right)^{-1} A^T y$	Need to tune λ in range $[0, 1]$
Proposed method using MP law	$\boldsymbol{x}^*_{\boldsymbol{MP}} = (\Sigma_{MP})^{-1} A^T y$	No tuning required
Proposed PCA based result from mini-project	(Refer to mini-project report for details)	No tuning required
Proposed SPCA based result from mini-project	(Refer to mini-project report for details)	No tuning required

Figure 5: List of various methods (traditional and proposed) which are compared in this section

Fig. 6 to Fig. 11 show results for all the methods listed in Fig. 5. Fig. 6 shows results for traditional Tikhonov ($x^*_{Tikh}(\lambda)$). The reconstructions shown are for 40 values of λ uniformly sampled from $[0, \lambda_{max}]$. The first subplot in Fig. 6 is for $\lambda_1 = 0$ whereas last subplot (bottom-right) is for $\lambda_{40} = \lambda_{max}$. On the other hand, Fig. 7 shows results for proposed shrinkage transformation based approach ($x^*_{ST}(\lambda)$). The reconstructions shown are for 40 values of λ uniformly sampled from $[0, 1]$.

It can be clearly seen from Fig. 6 and Fig. 7 that our proposed method using shrinkage transformation gives high quality reconstructions for majority of the values of λ where traditional Tikhonov only gives good reconstructions for few values of λ (only for λ_{26} to λ_{34}). This shows remarkable improvement due to proposed method in terms of making the inverse covariance matrix stable and well-conditioned (as explained in section 2).

Fig. 8 shows the L-curves for traditional Tikhonov solutions (solutions in Fig. 6). Fig. 8(a) is in linear scale whereas Fig. 8(b) is in log scale. It can be seen that L-curve gives three knee points, at $\lambda_{17}, \lambda_{18}$ and λ_{30}. The reconstructions corresponding to these three values of λ can be seen in Fig. 6.

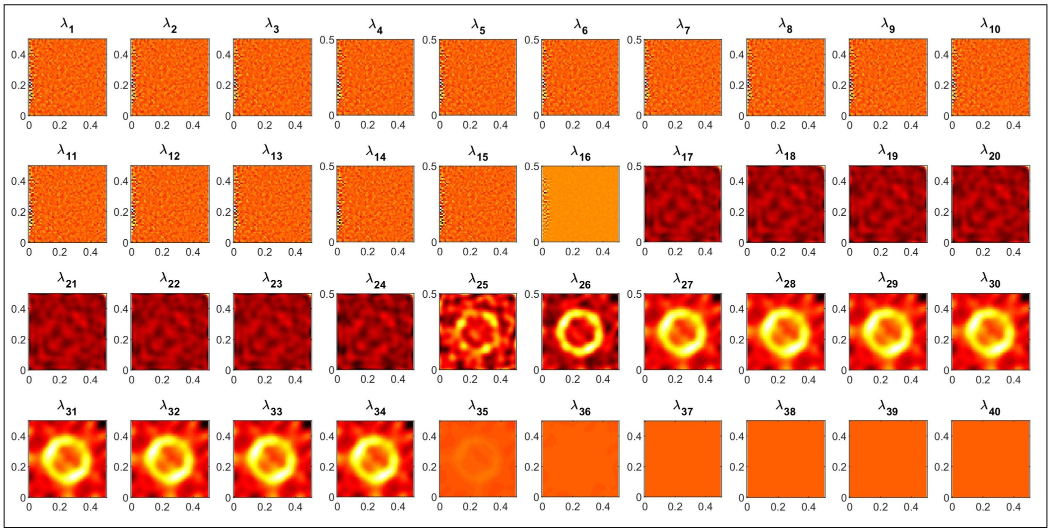


Figure 6: Reconstructions obtained by solving Traditional Tikhonov (T is selected to be identity in (5)) for 40 values of λ uniformly sampled from $[0, \lambda_{max}]$.

s

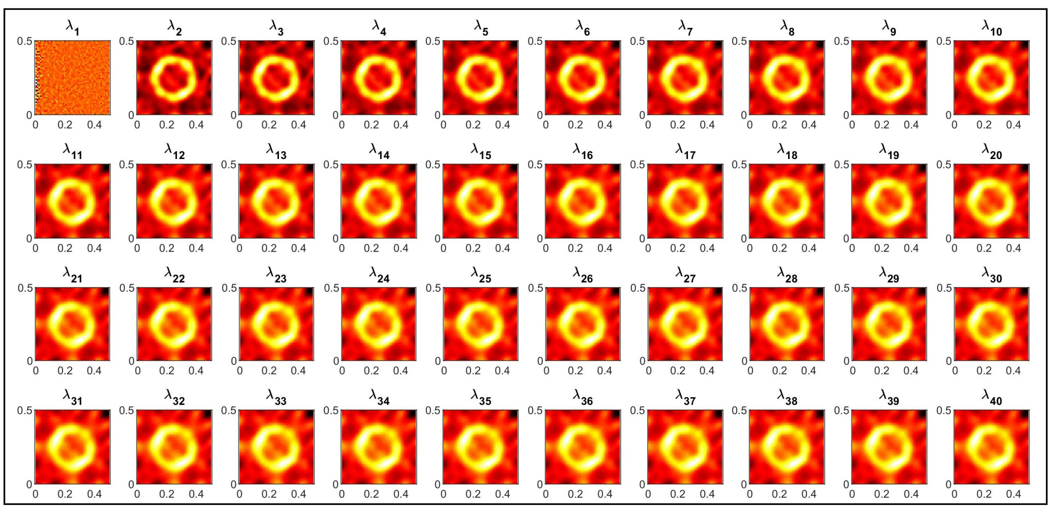
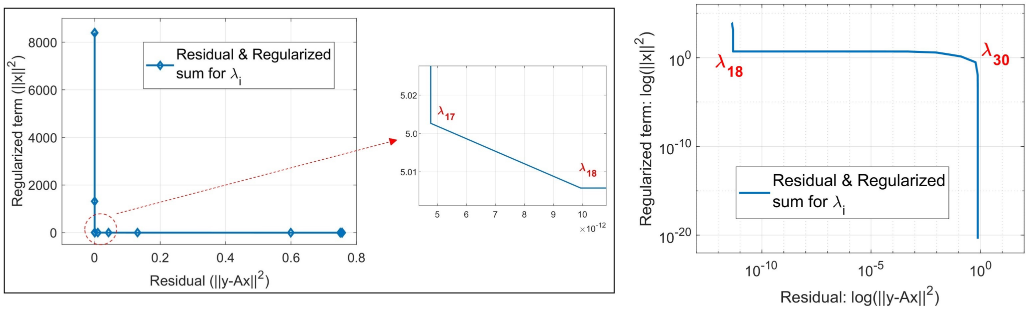


Figure 7: Reconstructions obtained by solving proposed method based on Shrinkage transformation (F is selected to be Sample-Variance Zero Covariance in (9)) for 40 values of λ uniformly sampled from $[0, 1]$.

It is clear that reconstructions for $\lambda_{17}, \lambda_{18}$ are very poor quality whereas for λ_{30} reconstruction is better. However, it is important to note that knee point around λ_{17} is much more steep and hence, without knowledge of ground truth in practical scenarios, one will have to select λ_{17} as optimal value which will lead to wrong result. To overcome this problem of false knee point, more advance versions of L-curve can be explored in future work [(Xu et al., 2016)]. Assuming we avoid this false knee problem, for traditional Tikhonov, the selected final solution can be reconstruction for another knee at λ_{30}. Reconstruction for λ_{30} which can be seen in Fig. 6. The SSIM value for this traditional Tikhonov solution $SSIM(x^*_{Tikh}(\lambda_{30})) = 0.68$.

Fig. 9 shows the L-curve for proposed shrinkage transformation based solution (x^*_{ST}) (solutions shown in Fig. 7). A very clear knee point can be seen at λ_2 and also there are no other false knee (unlike in Fig. 8). This means that proposed method provide better bias-variance trade-off than traditional Tikhonov. The reconstruction corresponding to λ_2 can be seen in Fig. 7. This is much better than solutions given by the traditional Tikhonov. The $SSIM(x^*_{ST}(\lambda_2)) = 0.82$ which is far better than what we achieved in traditional Tikhonov solution ($SSIM(x^*_{Tikh}(\lambda_{30})) = 0.68$).



(a) (b)

Figure 8: (a) L-curve for traditional Tikhonov results (solutions shown in Fig.6) for 40 values of λ in $[0, \lambda_{max}]$. The points around knee point is shown as magnified image on right. The knee point can be seen at λ_{17} and λ_{18} (b) Log scale plot of same L-curve

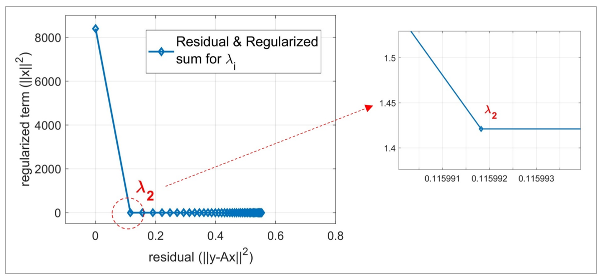
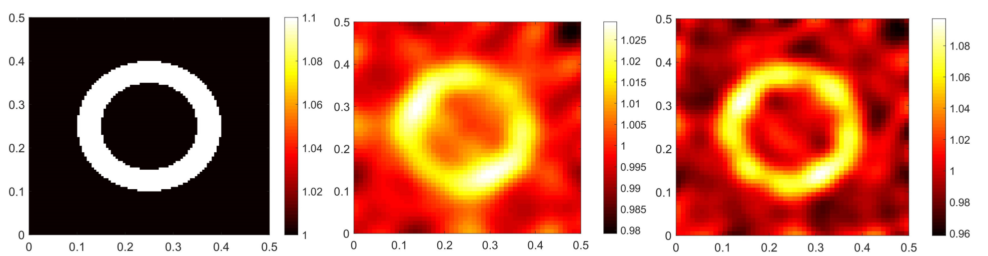


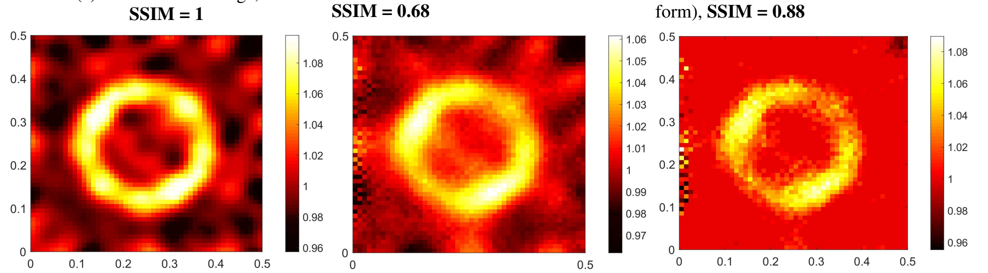
Figure 9: L-curve for proposed shrinkage transformation results (solutions shown in Fig.7) for 40 values of λ in [0,1]. The points around knee point is shown as magnified image on right. A clear knee point can be seen at λ_2.



(a) Ground Turth Image, **SSIM = 1**

(b) Traditional Tikhonov regularization, **SSIM = 0.68**

(c) Proposed method (Shrinkage Transform), **SSIM = 0.88**



(d) Proposed method (MP law), **SSIM = 0.87**

(e) PCA (from the mini-project), **SSIM = 0.74**

(f) SPCA (from the mini-project), **SSIM = 0.77**

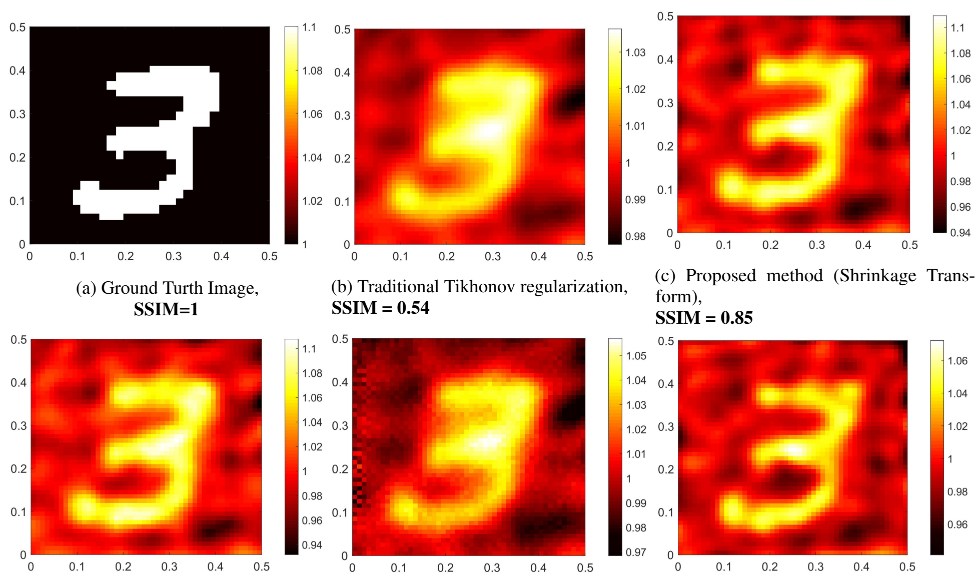
Figure 10: Solutions obtained using various methods listed in Fig. 5. SSIM values are provided for each reconstructions. The ground truth reference is shown in (a).

Overall, shrinkage transformation based solutions makes tuning of λ easier by bounding it in [0, 1] and also it gives more stable L-curve with well-defined unique knee point. Hence provide better bias-variance tradeoff. The SSIM value is also better for proposed method and hence it clearly outperforms traditional tikhonov both in terms of quality and ease of tuning regularization parameter.

Fig. 10 show results of all five methods listed in Fig. 5. Fig. 10(a) is the ground truth image for reference. Fig. 10(b) and (c) shows reconstruction using traditional Tikhonov ($x^*_{Tikh}(\lambda_{30})$ and proposed shrinkage transformation ($x^*_{ST}(\lambda_2)$ respectively, which we also saw previously in Fig. 7 and 6. Fig. 10(d) shows reconstruction using proposed MP law based method (x^*_{MP} (as described in section 3.0.2). Lastly, Fig. 10(e) and (f) shows results from mini-project using PCA and SPCA.

It can be clearly seen in Fig. 10 that in terms of SSIM, both proposed methods (in Fig. 10(c) and (d)) outperforms other techniques. The SSIM values for proposed shrinkage transformation and proposed MP law based estimation are close. But it is important to note that MP law based solution don't even need tuning of any regularization parameter and hence provide huge advantage over other methods. Overall, order of performance in terms of SSIM is **Shrinkage transformation solution** $>$ **MP law based solution** $>$ **SPCA** $>$ **PCA** $>$ **Traditional Tikhonov**. In terms of ease of tuning regularization parameter, MP law based solution, PCA and SPCA solutions do not require tuning at all. The shrinkage transformation requires tuning but in small range of values of λ.

Fig. 11 shows similar results but for different object shape. Here we used more intricate shape of the object by taking one arbitrary sample of digit 3 from the MNIST data-set and treat it like a scattering object kept inside the DOI. Fig.11 shows reconstruction of this profile using same methods as shown in Fig.10. Overall order of performance remains same as in Fig.10. Our proposed methods based on shrinkage transformation and MP law outperforms all other methods.



(a) Ground Turth Image, **SSIM=1**

(b) Traditional Tikhonov regularization, **SSIM = 0.54**

(c) Proposed method (Shrinkage Transform), **SSIM = 0.85**

(d) Proposed method (MP law), **SSIM = 0.83**

(e) PCA (from mini-project), **SSIM = 0.61**

(f) SPCA (from mini-project), **SSIM = 0.75**

Figure 11: (a) Solutions obtained using various methods listed in Fig. 5. SSIM values are provided for each reconstructions. The ground truth reference is shown in (a).

5 Conclusion

In this report, we proposed new methods based on RMT to modify and improve traditional Tikhonov regularization. We propose two methods, one based on shrinkage transformation and other based on MP law. In terms of estimation accuracy, both of the proposed methods outperforms the traditional methods. Furthermore, the shrinkage transformation based solution significantly eases the tuning process for λ by bounding λ in [0,1]. Also,the shrinkage transformation based solutions provide more stable L-curve with well-defined knee point which shows that it is effective in making SCM well-conditioned and in removing the sampling noise from SCM. On the other hand, MP law based proposed method removes the need for tuning any regularization parameter and hence provide much faster and efficient solution. Both the proposed techniques can be used for any class of Tikhonov regularization (for any appropriate priors like sparsity, smoothness, etc). The proposed methods provide unsupervised way of solving ill-posed problem without need of any validation data-set or additional labeled datasets.

6 Future Work

Future work can include:

1. Extend results for other regularization techniques outside Tikhonov family.
2. Combine results from multiple regularization techniques (like sparsity from LASSO and piece-wise continuity from TV regularization).
3. Explore more connections between L-curves and proposed methods and explore better alternatives to L-curve.

References

Xudong Chen. *Computational methods for electromagnetic inverse scattering*. Wiley Online Library, 2018.

GJ Tearney, ME Brezinski, JF Southern, BE Bouma, MR Hee, and JG Fujimoto. Determination of the refractive index of highly scattering human tissue by optical coherence tomography. *Optics Letters*, 20(21):2258–2260, 1995.

Mikhail Mikha_lovich Lavrent_ev, Vladimir Gavrilovich Romanov, and Serge_ Petrovich Shishatski_. *Ill-posed problems of mathematical physics and analysis*, volume 64. American Mathematical Soc., 1986.

Jerome Friedman, Trevor Hastie, Robert Tibshirani, et al. *The elements of statistical learning*, volume 1. Springer series in statistics New York, 2001.

Group-7, Amartansh Dubey, Samruddhi Deshmukh, and Nilesh Jha. *Tuning Regularization Parameter for Solving ill-posed Inverse Problems using Subspace Methods*. 2021. URL `https://github.com/yao-lab/yao-lab.github.io/blob/master/2021.csic5011/project1/Dubey-Deshmukh-Jha/CSIC%205011%20Project%201.pdf`.

Saandeep Depatla, Lucas Buckland, and Yasamin Mostofi. X-ray vision with only WiFi power measurements using rytov wave models. *IEEE Transactions on Vehicular Technology*, 64(4): 1376–1387, 2015.

Per Christian Hansen and Dianne Prost O'Leary. The use of the l-curve in the regularization of discrete ill-posed problems. *SIAM Journal on Scientific Computing*, 14(6):1487–1503, 1993.

D Calvetti, S Morigi, L Reichel, and F Sgallari. Tikhonov regularization and the l-curve for large discrete ill-posed problems. *Journal of Computational and Applied Mathematics*, 123(1-2):423–446, 2000.

Yanbin Xu, Yang Pei, and Feng Dong. An extended l-curve method for choosing a regularization parameter in electrical resistance tomography. *Measurement Science and Technology*, 27(11): 114002, 2016.

Olivier Ledoit and Michael Wolf. Honey, i shrunk the sample covariance matrix. *The Journal of Portfolio Management*, 30(4):110–119, 2004.

Olivier Ledoit and Michael Wolf. Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets goldilocks. *The Review of Financial Studies*, 30(12):4349–4388, 2017.

Vladimir A Marčenko and Leonid Andreevich Pastur. Distribution of eigenvalues for some sets of random matrices. *Mathematics of the USSR-Sbornik*, 1(4):457, 1967.

J Bun, J Bouchaud, and M Potters. Cleaning correlation matrices. *Risk Magazine*, 2017.

MATLAB R2020a. Image quality metrics , image processing toolbox, 2014. URL `https://www.mathworks.com/help/images/image-quality-metrics.html`.

Yuan Yao. *CSIC 5011: Topological and Geometric Data Reduction and Visualization Spring 2021*. 2021. URL `https://yao-lab.github.io/2021.csic5011/`.

CONTRIBUTION		
AMARTANSH DUBEY (Student ID: 20413942)	SAMRUDDHI DESKHMUKH (Student ID: 20783488)	NILESH KUMAR JHA (Student ID: 20617354)
Brainstormed and co-developed the idea of using Subspace methods and RMT for regularization parameter tuning	Brainstormed and co-developed the idea of using Subspace methods and RMT for regularization parameter tuning	Brainstormed and co-developed the idea of using Subspace methods and RMT for regularization parameter tuning
Simulated suitable data for the report using Method of Moments code for electromagnetic scattering	Developed framework to solve the inverse scattering problem for a wide range of regularization parameters	Developed framework to implement Sparse PCA and apply it on the inverse scattering data
Wrote Abstract and Section 2 (Proposed method) for the project report	Wrote Introduction and Numerical section of the project report	Wrote Introduction and Numerical section of the project report
Co-developed subspace method codes for analyzing the data, specifically, performing eigendecomposition on the multiple estimates of inverse solution to provide preliminary verification of idea	Converted Matlab codes related to physics based electromagnetic forward and inverse simulations to Python for utilizing wider range of libraries and all in one platform for analysis	Converted Matlab codes related to physics based electromagnetic forward and inverse simulations to Python for utilizing wider range of libraries and all in one platform for analysis
Literature survey on challenges in tuning regularization parameter, specifically in inverse scattering problems, subspace methods and implement MP law based results.	Literature survey on challenges in solving high dimensional ill-posed problem and faster subspace decomposition. Implemented Shrinkage transformation code.	Literature survey on various prior enforcing optimization techniques. Implemented Shrinkage transformation code.

Figure 12

Madeleine Udell, Corinne Horn, Reza Zadeh, and Stephen Boyd. Generalized low rank models. *arXiv preprint arXiv:1410.0342*, 2014.