

# Objective Selection of Hyperparameter for EIT

**B M Graham, A Adler**

School of Information Technology and Engineering (SITE), University of  
Ottawa, Canada

E-mail: [graham.bm@sympatico.ca](mailto:graham.bm@sympatico.ca), [adler@site.uottawa.ca](mailto:adler@site.uottawa.ca)

**Abstract.** An algorithm for objectively calculating the hyperparameter for the class of linearized one step EIT image reconstruction algorithms is proposed and compared to existing strategies. EIT is an ill-conditioned problem in which regularization is used to calculate a stable and accurate solution by incorporating some form of prior knowledge into the solution. A hyperparameter is used to control the balance between conformance to data and conformance to the prior. A remaining challenge is to develop and validate methods of objectively selecting the hyperparameter. In this paper evaluate an compare and evaluate five different strategies for hyperparameter selection. We propose a calibration based method of objective hyperparameter selection, called *BestRes*, that leads to repeatable and stable image reconstructions that are indistinguishable from heuristic selections. Results indicate: 1) heuristic selections of hyperparameter are inconsistent among experts, 2) Generalized Cross-Validation approaches produce under-regularized solutions, 3) L-Curve approaches are unreliable for EIT, and 4) *BestRes* produces good solutions comparable to expert selections. Additionally, we show that it is possible to reliably detect an *inverse crime* based on analysis of these parameters.

*Keywords:* regularization, EIT, hyperparameter, L-Curve, GCV

## 1. Introduction

EIT attempts to calculate a stable and accurate image of the conductivity or conductivity change within a medium from electrical measurements made on the medium boundary. The image reconstruction problem is essentially underdetermined and characterized by a system matrix with large condition number. Image reconstructions obtained through naïve methods such as least squares are unstable and dominated by noise. The problem has been overcome through the use of various regularization methods which produce useful solutions by imposing additional conditions (priors), such as image smoothness, on the problem (Vauhkonen et al. 1997). The balance between solution conformance to the measured data and conformance to the prior is controlled by a scalar hyperparameter often labelled  $\lambda$ .

A difficulty with experimental and clinical EIT reconstruction algorithms is the tendency of algorithms to rely on subjective methods to select a hyperparameter. The absence of objective hyperparameter selection methods results in several issues which hinder experimental and clinical use of the technique: 1) users of EIT for clinical applications are uncomfortable using “fiddle” adjustments to modify images, 2) comparisons of EIT reconstruction algorithms can be subjective due to the necessity of manual tuning of hyperparameter values, 3) experimental work is not repeatable if disparate researchers cannot objectively recreate the hyperparameter values used

in the work of others, and 4) meta-algorithms, such as detection of electrode errors (Asfaw & Adler 2005), require a method to fix these values.

In order to address this issue, we investigate existing hyperparameter selection methods and propose a new calibration based method called *BestRes* (Best Resolution). By ‘‘calibration’’ we mean that a procedure is defined to select a value for a given EIT system and measurement configuration, rather than for each image or data set. We define a configuration as the combination of current injection pattern, Finite Element Mesh (FEM), assumed prior conductivity ( $\sigma_0$ ) and regularization prior. Hyperparameter selection methods are then compared for several one-step linearized EIT reconstruction algorithms.

## 2. Methods

We consider EIT difference imaging, which is widely understood to improve reconstructed image stability in the presence of problems such as unknown contact impedance, inaccurate electrode positions, nonlinearity, and the use of 2D approximations for 3D electrical fields (Barber & Brown 1988b) (Lionheart 2004). Initially, we address the class of normalized one-step linearized reconstruction algorithms that calculate the proportional change in a finite element conductivity distribution,  $\mathbf{x} = (\boldsymbol{\sigma}_2 - \boldsymbol{\sigma}_1)/\boldsymbol{\sigma}_1$ , due to a proportional change in difference signal,  $\mathbf{z} = (\mathbf{v}_2 - \mathbf{v}_1)/\mathbf{v}_1$ , over a time interval  $(t_1, t_2)$ . Since we do not know  $\boldsymbol{\sigma}_1$ ,  $\mathbf{x}$  is interpreted as the proportional change in conductivity with respect to the unknown initial conductivity  $\mathbf{x} = \Delta\boldsymbol{\sigma}/\boldsymbol{\sigma}_0$ . For small changes around a background conductivity the relationship between  $\mathbf{x}$  and  $\mathbf{z}$  may be linearized as

$$\mathbf{z} = \mathbf{H}\mathbf{x} + \mathbf{n} \quad (1)$$

where  $\mathbf{H}$  is the Jacobian or sensitivity matrix and  $\mathbf{n}$  is the measurement system noise, assumed to be uncorrelated additive white Gaussian (AWGN). Each element  $i, j$ , of  $\mathbf{H}$  is calculated as  $H_{ij} = \left. \frac{\partial z_i}{\partial x_j} \right|_{\sigma_0}$  and relates a small change in the  $i^{th}$  proportional difference measurement to a small change in the proportional conductivity of  $j^{th}$  element.  $\mathbf{H}$  is a function of the finite element mesh (FEM), the current injection pattern, and the background conductivity,  $\boldsymbol{\sigma}_0$ . We use a homogenous background with  $\sigma_0 = 1$  for each of the elements. Normalizing the signal requires that we also normalize the sensitivity matrix, by dividing its columns by  $\mathbf{v}_{\text{ref}}$  which is a vector of reference voltages obtained by solving the forward problem over a homogenous domain.

### 2.1. Regularization

In order to overcome the ill conditioning of  $\mathbf{H}$  we solve (1) using the following regularized inverse

$$\hat{\mathbf{x}} = (\mathbf{H}^T \mathbf{W} \mathbf{H} + \lambda \mathbf{R})^{-1} \mathbf{H}^T \mathbf{W} \mathbf{z} = \mathbf{B} \mathbf{z} \quad (2)$$

where  $\hat{\mathbf{x}}$  is an estimate of the true proportional change in conductivity distribution,  $\mathbf{R}$  is a regularization matrix,  $\lambda$  is a scalar hyperparameter that controls the amount of regularization, and  $\mathbf{W}$  models the system noise. Since noise is uncorrelated in the system,  $\mathbf{W}$  is a diagonal matrix;  $\mathbf{W}_{i,i} = 1/\sigma_i^2$  where  $\sigma_i^2$  is the noise variance for measurement  $i$ .  $\mathbf{W}$  can also be modified to account for variable gain settings on each tomograph channel. However, for this work we assume that all measurements have

equal noise variance, thus  $\mathbf{W}$  becomes a multiple of the identity matrix. With  $\mathbf{R} = \mathbf{I}$  (labelled  $\mathbf{R}_{\text{Tik}}$ ) equation (2) is the 0<sup>th</sup> order Tikhonov algorithm. With  $\mathbf{R} = \text{diag}(\mathbf{H})$  (labelled  $\mathbf{R}_{\text{diag}(\mathbf{H})}$ ) equation (2) is the regularization matrix used in the NOSER algorithm of (Cheney et al. 1991). (Adler & Guardo 1996) modelled  $\mathbf{R}$  as a spatially invariant Gaussian high pass filter (labelled  $\mathbf{R}_{\text{HPF}}$ ) with a cut-off frequency selected so the spatial period is a given fraction of the medium diameter.  $\mathbf{R}_{\text{HPF}}$  reconstructions appear reasonable for cut-off frequencies corresponding to 5%, 10% and 20% diameter. A 16 electrode EIT system, using adjacent measurements not at current injection sites, yields 208 measurements of which 104 are independent. We consider  $\lambda_{\text{HPF}}$  for 10% because it appears better justified in terms of available measurements, in that 10% is roughly  $1/\sqrt{104}$ .

All three of these priors are smoothing filters, however the Gaussian HPF has the advantage of being mesh size independent in that it is a function of the mesh inter-element correlations. The Tikhonov and NOSER are *ad hoc* priors that do not consider correlations between solution mesh elements. While several other one-step regularized inverse algorithms exist for EIT (Barber & Brown 1988a), (Kotre 1994), (Cohen-Bacrie et al. 1997), in this paper we consider equation (2) with the Tikhonov, NOSER, and Gaussian HPF regularization matrices as a representative sample with which to compare hyperparameter selection strategies.

The rest of this paper is organized as follows. First, some existing hyperparameter strategies are discussed. A calibration strategy to objectively calculate an optimal hyperparameter for a given configuration is then described. Next, the effectiveness of each strategy for linearized one-step EIT is compared with respect to heuristic selection. Finally, observations of this work are discussed and we conclude with a recommendation of the *BestRes* hyperparameter selection method.

### 3. Hyperparameter Selection Methods

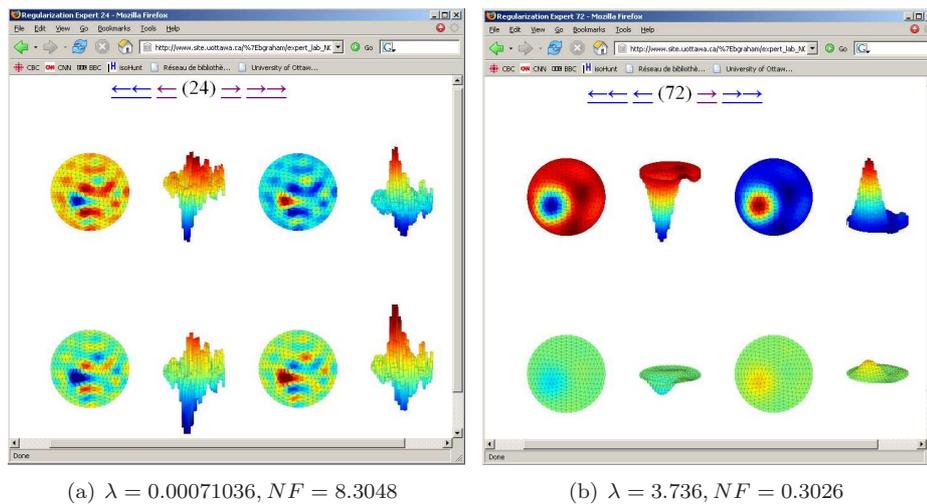
The goal of hyperparameter selection is to produce a “good” reconstruction. Intuitively hyperparameter selection should produce solutions that preserve as much of the measured data as possible by applying the least amount of *a priori* information required to obtain a useful reconstruction.

#### 3.1. Heuristic Selection

The most common method of hyperparameter selection is Heuristic Selection in which researchers examine sets of reconstructions generated over a range of hyperparameter values and select the best image. This method is subjective and not repeatable. To our knowledge no research has specifically evaluated the performance of objective hyperparameter selection for one step solutions.

In this work, heuristic selection was performed by 5 graduate students who were asked to participate in an experiment evaluating human performance in choosing regularization parameters. A web site was set up in which 5 independent data sets were used to generate sequences of reconstructions. Each sequence showed reconstructions as a function of the 77 different values of the hyperparameter. Each web page, such as the example of figure 1(a), showed the same conductivity change solution using 8 visual styles (subimages).

Each pair of images is shown as a 2D false colour representation of the conductivity change image and an associated 3D version where the  $z$  dimension represents



**Figure 1.** Two web pages from the heuristic selection experiment. Images generated from tank data using  $R_{H_{PF}}$  prior with different hyperparameter values

conductivity change. The left half of the page shows the reconstruction as a decrease in conductivity while the right side shows the inverse of the image (we reverse the reference and data frames) so that the reconstruction appears to be a conductivity increase. The top row uses relative colour and z-axis scaling, thus each subimage in the 77 page sequence uses the full range of colours. For the 3D representation the conductivity is scaled to fill the entire z-axis. The bottom row uses an absolute colour and z-axis scaling thus each subimage of the 77 page sequence uses the same colour and vertical axis extent. Consequently highly smoothed images (large hyperparameter values), such as figure 1(b) have little color variation and reduced vertical extent compared to images reconstructed with lower hyperparameter values such as figure 1(a). Students were instructed to choose the best image based on the following definition: the image which shows the best resolution for the contrasting region(s) without excessive contamination by noise.

The same set of students was asked to repeat the experiment four months later. Students were instructed not to look at their earlier results as the aim of the second experiment was to evaluate repeatability. Full details of the experiment are available in (Graham 2005)

### 3.2. L-Curve

Perhaps the most well known method of hyperparameter selection after heuristic selection is the L-Curve method (Hansen 1992). This method plots the semi-norm of the regularized solution,  $\log_{10} |\mathbf{R}\hat{\mathbf{x}}|$ , versus the norm of the corresponding residual vector,  $\log_{10} |\mathbf{H}\hat{\mathbf{x}} - \mathbf{z}|$ , parametrically over  $\lambda$ . The resulting plot, such as figure 2(a), will often have an “L” shape where the optimal value for  $\lambda$  is located at the point of maximum curvature. Hansen describes a method for calculating this “corner” of the L-curve in (Hansen 1997). We call this value  $\lambda_{LC}$ . There are cases where the L-Curve may fail. Figure 2(b) is an L-Curve that does not have a “corner”.

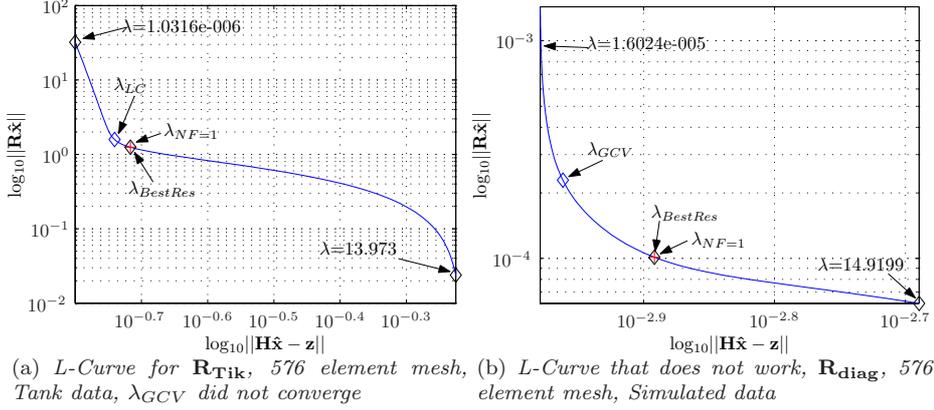


Figure 2. Example L-Curves.

### 3.3. Generalized Cross-Validation

Generalized cross-validation (GCV) is based on the principle that if any arbitrary element of the data,  $\mathbf{z}$ , is left out, then the corresponding regularized solution should predict the missing element (Hansen 1997). Its advantage is that no prior knowledge about the error norm is required. This leads to choosing a regularization parameter which minimizes the GCV function

$$\text{GCV}(\lambda) = \frac{\|\mathbf{H}\hat{\mathbf{x}} - \mathbf{z}\|^2}{\text{trace}(\mathbf{I} - \mathbf{H}\mathbf{B})^2} \quad (3)$$

where  $\mathbf{B}$ ,  $\mathbf{z}$  and  $\hat{\mathbf{x}}$  are as in equations 1 and 2. This method has often been used for the Tikhonov prior. In this work we use it for all three priors.

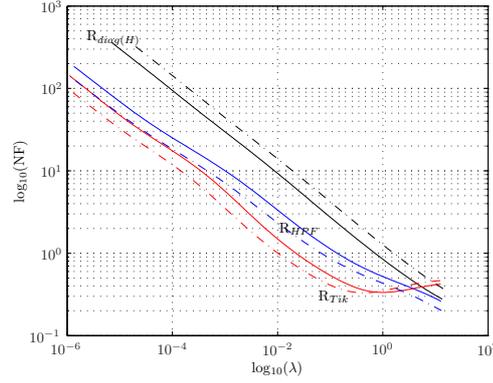
### 3.4. Fixed Noise Figure (NF)

The Fixed NF Method is based on a *Noise Figure* calculation introduced by (Adler & Guardo 1996) where NF is defined as the ratio of signal to noise ratio (SNR) in the measurements to SNR in the image:

$$\text{NF} = \frac{\text{SNR}_{in}}{\text{SNR}_{out}} = \left( \frac{\text{mean}[\mathbf{z}_c]}{\sqrt{\text{var}[\mathbf{n}]}} \right) \bigg/ \left( \frac{\text{mean}[\mathbf{B}\mathbf{z}_c]}{\sqrt{\text{var}[\mathbf{B}\mathbf{n}]}} \right) \quad (4)$$

The signal used in this definition is  $\mathbf{z}_c = \mathbf{H}\mathbf{x}_c$ , where  $\mathbf{x}_c$  is a small contrast in the centre of the medium. The user selects a NF value and the corresponding  $\lambda$  is found using a bisection search technique. The Fixed NF Method substitutes the manual selection of  $\lambda$  with the manual selection of a NF, which the algorithm then maps to a hyperparameter value; the value for  $\text{NF} = 1$  is labelled  $\lambda_{\text{NF}=1}$ . As shown in figure 3, for a given configuration,  $\log(\text{NF})$  is nearly linearly inversely proportionally to  $\log(\lambda)$  throughout the extent where  $\lambda$  yields good solutions.

Experience has shown that noise figures in the range 0.5 to 2 consistently lead to good reconstructions regardless of configuration or data source (simulated or measured) while the associated  $\lambda$  value can range over several orders of magnitude dependent on configuration. The 18 configurations used in this work,  $\lambda$  ranged over 3 orders of magnitude for  $\text{NF} = 1$ . The advantage of the Fixed NF method is that the suitable NF range is not configuration dependent while  $\lambda$  is.



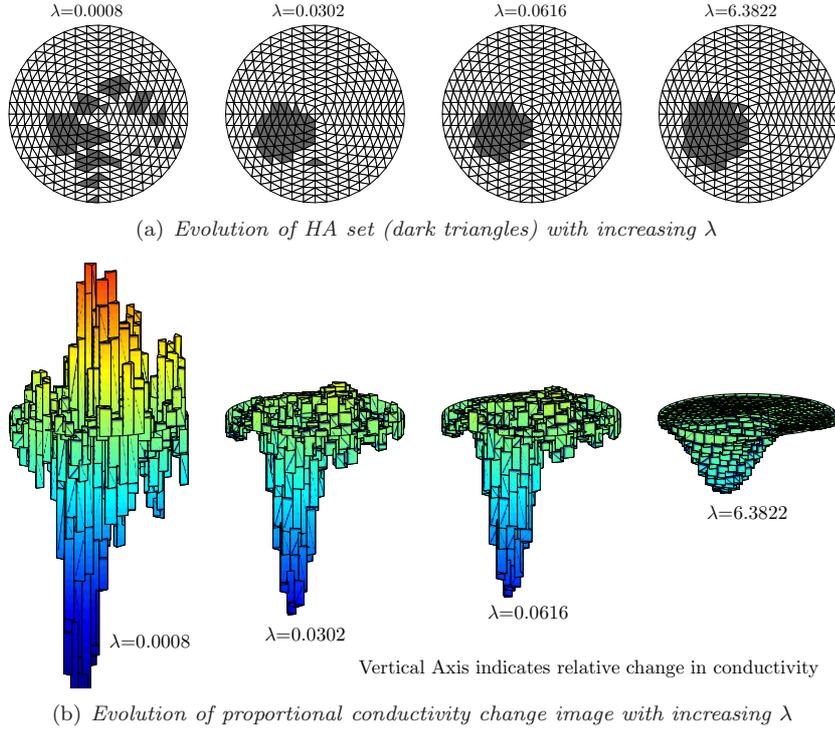
**Figure 3.** Plot showing linear relationship between  $\log(\lambda)$  and  $\log(NF)$ . Solid lines are simulated data on 256 element mesh, dashed lines are tank data on 576 element mesh. Priors are:  $\mathbf{R}_{diag(H)}$  (black),  $\mathbf{R}_{HPF}$  (blue),  $\mathbf{R}_{Tik}$  (red)

### 3.5. Calibration Based Hyperparameter Selection Method

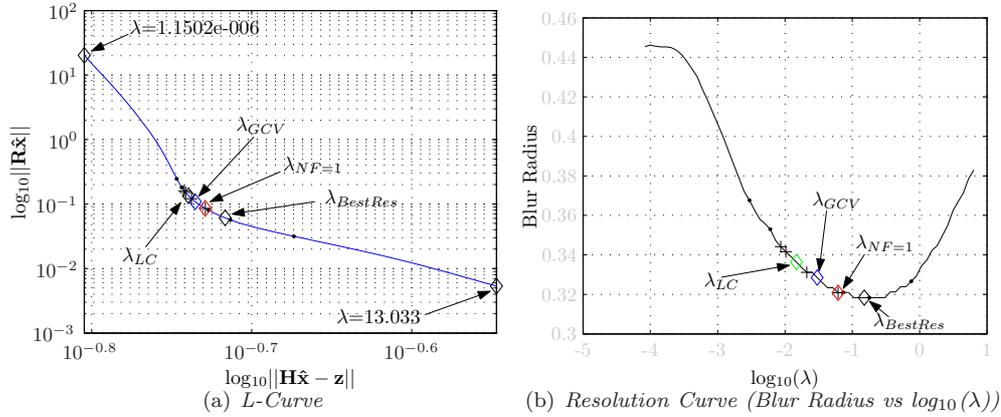
**3.5.1. aside - Figure of Merit** An objective selection method requires a quantitative figure of merit to measure the quality of the reconstructed image. (Wheeler et al. 2002) reviewed several figures of merit for EIT that have been proposed in the literature. The primary figure of merit used in this work is *Resolution* which we calculate in terms of Blur Radius (BR). We define Blur Radius as  $BR = r_z/r_0 = \sqrt{A_z/A_0}$  where  $r_0$  and  $A_0$  are the radius and area respectively of the entire 2D medium and  $r_z$  and  $A_z$  are the radius and area of the reconstructed contrast containing half the magnitude of the reconstructed image (Adler & Guardo 1996). BR calculates the area fraction of the elements that contain 50% of the total image amplitude. We call this the half amplitude (HA) set. Figure 4(a) shows the evolution of the HA set in response to increasing  $\lambda$  for an impulse contrast. Figure 4(b) shows the corresponding impedance change images, here represented with 3D visualization. With insufficient  $\lambda$  the image is dominated by noise and the HA set is composed of spatially disjoint elements. As  $\lambda$  is increased, noise is filtered through the smoothing action of the prior, image energy starts to concentrate, and the HA set starts to cluster. The point at which the HA set is comprised of adjacent elements is termed the “onset of stability” (OS). Excessive regularization blurs the image and expands the now contiguous HA set. A Resolution Curve (plot of Resolution vs  $\lambda$ ) such as figure 5(b) shows a rapid increase in resolution (decreasing BR) to a minimum value that occurs after the onset of stability, followed by a slower increase in BR as filtering starts to blur the image. For an impulse contrast the minimum point of the Resolution curve indicates the best resolution. This value can be considered optimal with respect to both Resolution and stability (slope of curve is low indicating small change in signal for a small change in  $\lambda$ ) for the given data set.

**3.5.2. BestRes Method** The Resolution curve of figure 5(b) suggests the following hyperparameter selection strategy, which we refer to as the “*BestRes*” Method.

- (i) Image an impulse contrast
  - (a) The preferred method is to use imaging equipment to collect a frame of reference data from a homogenous medium. Then collect a data frame



**Figure 4.** Sample Reconstructions of impulse phantom, tank data with  $\mathbf{R}_{\text{HPF}}$  on 576 element mesh.



**Figure 5.** Points indicate original heuristic selections, crosses indicate second set of heuristic selections.

by imaging an impulse contrast using a physical phantom located halfway between the centre and boundary of medium ( $r/2$ ).

- (b) If equipment is not available the method can use simulated data. Again simulate a reference frame using a homogenous medium. Simulate a data

frame by changing the conductivity of a single mesh element located at  $r/2$ . Using real data has the potential to produce a hyperparameter for the given configuration that is tailored to the equipment. In practice the reconstructions obtained using this seemingly more accurate method are not qualitatively improved over those that are generated using simulated data.

- (ii) Reconstruct a series of images as a function of the hyperparameter and plot the associated Resolution curve as in Figure 5(b).
- (iii) Determine  $\lambda_{BestRes}$  as the point for maximum resolution - minimum BR. This value of  $\lambda$  is then used for all subsequent reconstructions using simulated or real data.

If using simulated data then representative noise should be included. We suggest producing several Resolution curves (we used 50) each with a different instance of the representative noise level. Each curve will produce a value of  $\lambda_{BestRes}$ . The mean of this set of  $\lambda_{BestRes}$  is the output of the *BestRes* method.

## 4. Results

### 4.1. Data Sources

Three sources of test data were used to compare the hyperparameter selection methods:

- (i) simulated data, generated using a 2D finite element mesh with 1968 elements using the point electrode model. Data for the reference frame was generated using a homogenous background conductivity with  $\sigma_0 = 1$ . The data frame was generated by reducing the conductivity of a single FEM element (0.05% of medium area) located halfway along the radius of the tank ( $r/2$ ) by 15%.
- (ii) simulated data obtained by adding Gaussian noise to set #1. Noise variance was 0.05% of maximum signal value, and
- (iii) tank data using a Goe-MF II type tomography system (Viasys Healthcare, Höchberg, FRG). The reference frame of the tank data was generated using a homogenous saline solution in a 30cm diameter tank. The data frame was generated using a 2cm diameter non-conductive impulse phantom located at  $r/2$ .

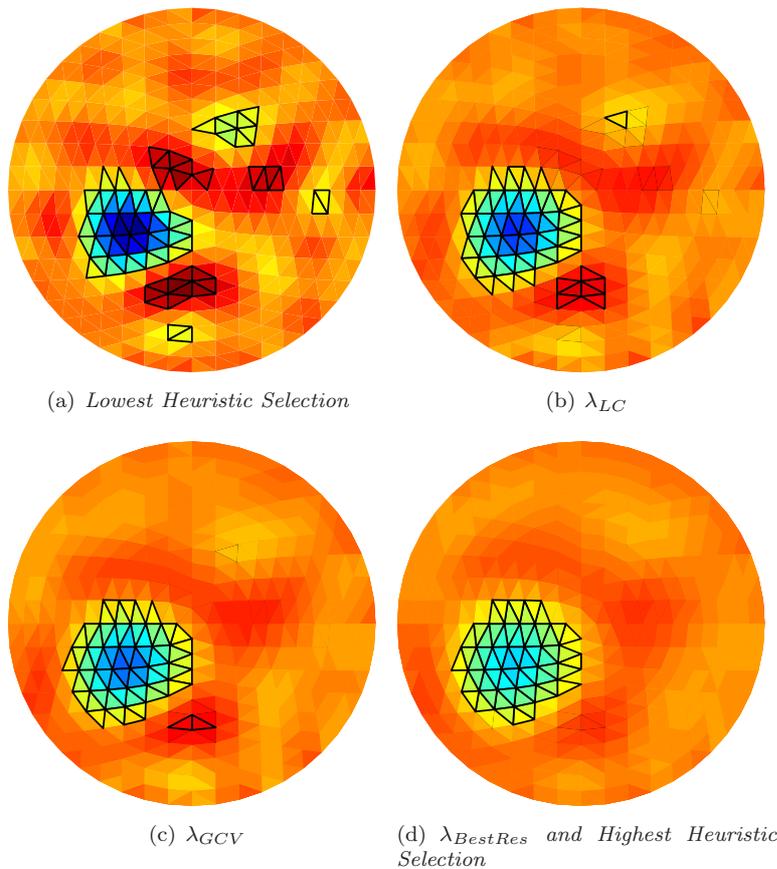
Both simulated and tank data used 16 electrodes equispaced in a plane driven with the adjacent current drive protocol. Measurements involving one or more of the driving electrodes were discarded giving 208 measurements per data frame.

The three data sets were used to reconstruct images using 18 configurations (6 meshes, 3 regularization matrices). The 6 meshes have 64, 256, 492, 576, 1024, and 1968 elements. Reconstructions of simulated data using the 1968 element mesh constituted an inverse crime (Wirgin 2004), which we discuss later. The hyperparameter selection methods are compared using the L-Curve and Resolution Curves of figures 5. Hyperparameter values selected from each of the methods are shown on both curves. Associated reconstructions are shown in figure 6.

### 4.2. Heuristic Results

The points on figures 5 indicate the first set of hyperparameters selected (indirectly) by heuristic selection. The associated reconstructions are found in figure 6. Heuristic

selections varied and were not confined to the minimum region of the Resolution Curve or knee of the L-Curve: no clear preference was shown among images reconstructed using  $\lambda$  from the minimum region of the Resolution Curve. The crosses on figures 5 indicate the hyperparameter values selected by the same experts when the experiment was repeated 4 months later. Results indicate that Heuristic selections of hyperparameter are inconsistent among experts and unrepeatable. Heuristic selections are subject to many biases including the colour scheme used in images, whether impedance changes are shown from a 2D or 3D perspective, the *a priori* expectation of the expert concerning noise levels, desired image properties, and other unknown individual idiosyncrasies. The heuristic results suggests that there is no single preferred value of  $\lambda$ , rather there is a preferred region of  $\lambda$  over which reconstructions are not subjectively distinguishable.



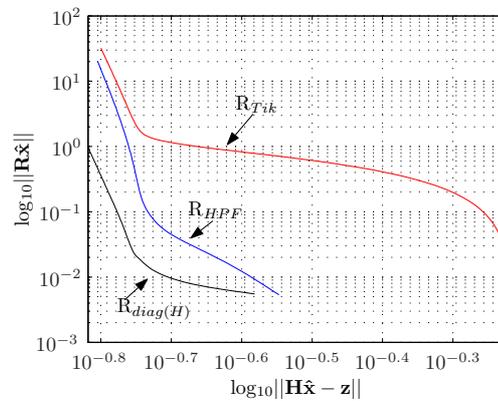
**Figure 6.** Example Reconstructions, Black Bordered triangles are elements of the Half Amplitude Set

#### 4.3. L-Curve Results

Although most L-Curves from this data were able to indicate an optimal tradeoff region, not all curves had a pronounced enough corner to allow unambiguous selection

of  $\lambda$ . In the six Tikhonov configurations the L-Curve always indicated a clear point of maximum curvature. However there were some configurations, such as figure 2(b), where the L-Curve did not exhibit a corner from which a hyperparameter could be calculated. In general the L-Curve indicated a lower value for  $\lambda$  than the Fixed NF and *BestRes* methods. As a result L-Curve derived images were comparatively noisier. In several instances  $\lambda_{LC}$  occurred much earlier than the onset of stability. We make the observation that L-Curves for the NOSER and Gaussian HPF priors are shallower than classic L-Curves discussed in the general inverse problems field such as (Hansen 1992). This is illustrated in figure 7 which compares the relatively sharp corner of the Tikhonov L-Curve to the shallower curves for the Gaussian HPF and NOSER priors.

The L-Curve method also requires the generation of an L-curve for each set of data. We feel it is preferable to be able to calculate a single hyperparameter value suitable for continuous use with a specific configuration which can be done with the *BestRes* and Fixed NF methods.



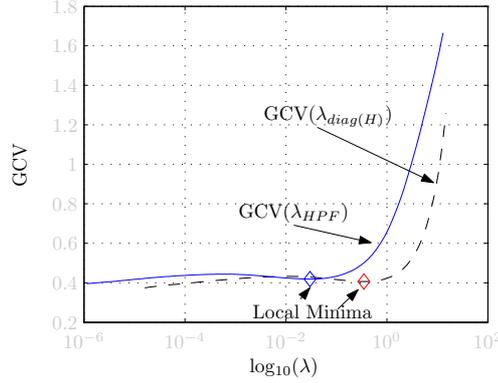
**Figure 7.** L-Curves for different priors on same mesh (576 element, Tank Data)

#### 4.4. GCV

It has been noted (Tenorio 2001) that the GCV method often calculates very small values of  $\lambda$  leading to solutions that are severely under-regularized. As illustrated in figure 8, the GCV function can also be very shallow making it difficult to isolate a clear minimum. In some cases the GCV curve was monotonically increasing, and thus did not have a minimum. For example, the GCV curve for a reconstruction using tank data on the 576 element mesh with the Tikhonov prior failed to exhibit a minimum. Overall the GCV criterion was unreliable in calculating hyperparameters for linearized one-step EIT reconstructions.

#### 4.5. BestRes Results

For both tank and simulated data using all 18 configurations the Resolution Curve exhibited a distinct minimum point at which  $\lambda_{\text{BestRes}}$  could be calculated and subsequently used to obtain a “good” reconstruction. It appears that Resolution is a useful figure of merit for EIT reconstructions.



**Figure 8.** GCV Curves for Tank Data reconstructed on the 576 element mesh

#### 4.6. Fixed NF Results

With both the tank and simulated data  $\lambda_{NF=1}$  was consistently located in the minimal region of the Resolution Curve. Moreover  $\lambda_{NF=1}$  always fell within the boundaries of the hyperparameters selected by the experts (i.e. it was as consistent as the experts). Fixed NF with  $NF = 1$  always calculated a hyperparameter that resulted in a good reconstruction.

## 5. Discussion

This paper has investigated the performance of various hyperparameter selection methods including the *BestRes* method herein introduced. In the course of these studies it became clear that several other aspects of EIT image reconstruction are related to hyperparameter selection. In this section we discuss the effects of noise level, radial position of contrasts, and normalization on hyperparameter selection. We also touch on applicability to non-linear reconstructions and inverse crimes.

### 5.1. Noise and Radial Position

We further explored the behaviour of the Fixed NF and *BestRes* methods by performing two additional experiments. Simulated data was generated with increasing amounts of AWGN noise for various radial positions of the generating impulse contrast. Figure 9(a) shows  $\lambda_{BestRes}$  as a function of increasing noise. Hyperparameter values for all methods except Fixed NF increased as noise level increased resulting in greater noise suppression through increased smoothing. The maximum noise levels used in this work are much larger than found in practice but were used to understand trends. Figure 9(b) shows the effect of  $\lambda$  as a function of radial position of the generating contrast. The noise in this plot is fixed at 0.50%. The best resolution is achieved for contrasts located at 25% of radial distance from the centre. Although  $\lambda$  continues to increase as the radial position increases past 90%, the corresponding resolution starts to decrease due to blurring caused by proximity to the edge.

Since Fixed NF is not a function of the data, hyperparameters selected with Fixed NF do not change with noise. Consideration of data noise will in general require more smoothing (therefore larger  $\lambda$  values). However at realistic noise levels  $\lambda_{NF=1}$  falls

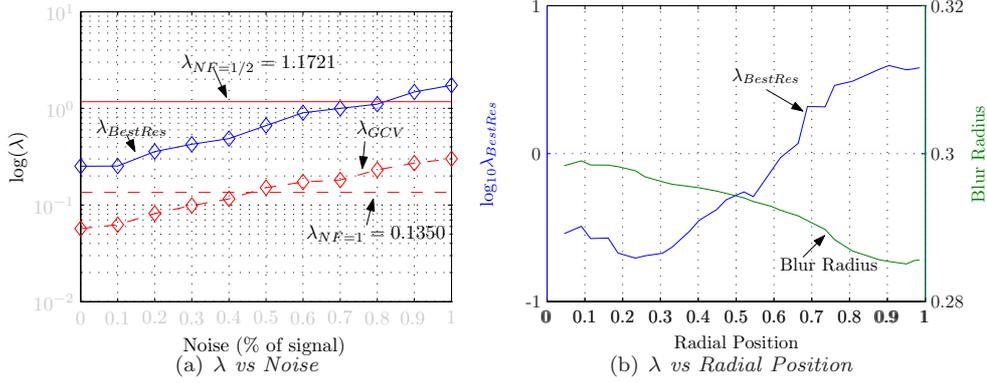


Figure 9. Hyperparameter trends, 256 element mesh, Gaussian HPF Prior

within the minimum region of the Resolution Curve, is indistinguishable from heuristic selections, and consistently results in good reconstructions.

*5.1.1. Normalization* Although primarily concerned with proportional (normalized) difference imaging we also investigated the performance of the Fixed NF method using simple (non-normalized) difference algorithms. The simple difference problem is solved using equation 2 with  $\mathbf{x}$  defined as  $\mathbf{x} = \sigma_2 - \sigma_1$  and  $\mathbf{z}$  defined as  $\mathbf{z} = \mathbf{v}_2 - \mathbf{v}_1$ . ( $\mathbf{H}$  is also modified in that its columns are not divided by  $\mathbf{v}_{\text{ref}}$  as in section 2). Similar to the proportional difference algorithm the Fixed NF method was able to consistently calculate a hyperparameter located in the minimum region of the associated Resolution Curve.

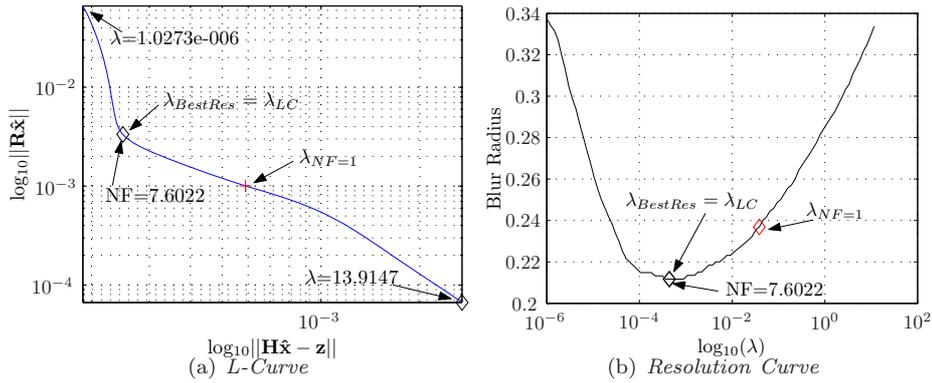
*Comment on Normalization* Normalization has some advantages. 1) By normalizing the data, one does not need to know what value of current was injected in obtaining the boundary data. Thus when the forward problem is solved one is free to use whatever value of injected current makes sense numerically for the algorithm (usually  $\pm 1$ ). 2) The authors of (Metherall et al. 1996) justify normalization with the argument that a normalized sensitivity matrix is less sensitive to the boundary shape of the object and the position of the electrodes, which is a substantial problem in clinical applications. 3) Some EIT systems change the gain settings for each electrode depending on the amplitude of signal expected. Normalization can compensate for these cases as well as instances where the calibrated gain settings are not exact.

Normalization of the sensitivity matrix does not change the spectrum of singular values of  $\mathbf{H}^T \mathbf{H}$ . Although the condition number of  $\mathbf{H}^T \mathbf{H}$  may improve by an order of magnitude or more the relative change is not significant. For the meshes used in this work the condition number of  $\mathbf{H}^T \mathbf{H}$  improved by a factor of about 10. In other work involving non-circular meshes we have seen changes from as low as 1.5 to as high as 1000. In some cases (non-homogenous background conductivity) the condition number increased, so it is not possible to state that, as a rule, normalization always improves the condition number of  $\mathbf{H}^T \mathbf{H}$ .

*5.1.2. Non-Linear Reconstructions* We investigated the use of Fixed NF method for use with an iterative static reconstruction algorithm. In this experiment we used Fixed

NF to calculate a single hyperparameter that was used for each iteration. Running the algorithm to convergence indicated that the  $\lambda_{NF=1}$  was located in the minimum region of the Resolution Curve. It may also possible to use the Fixed NF method to calculate a new hyperparameter for each step of the iterative algorithm.

*5.1.3. Inverse Crime* The act of employing the same model to generate as well as to invert simulated data is known as an *inverse crime* (Wirgin 2004). Consequently, using our simulated data, created using the 1968 element mesh, to reconstruct images over the 1968 element mesh constitute an inverse crime. Such reconstructions had noticeably better resolution than was achieved with other meshes and, as shown in figures 10(a) and 10(b), exhibited  $\lambda_{BestRes}$  and  $\lambda_{LC}$  corresponding to an uncharacteristically high NF. This suggests a method to detect inverse crimes: Using the suspect FEM, and associated data, construct a Resolution Curve or L-Curve with the simulated data and calculate the NF corresponding to  $\lambda_{BestRes}$  or  $\lambda_{LC}$ . If the  $NF \gg 3$  (for example, figure 10 had  $NF > 7$ ) it is likely that the reconstruction algorithm is committing an inverse crime. The method was validated by the observation that reconstructions over the 1968 element mesh using tank data did not exhibit the large NF bias. One explanation for these results is that the geometry matching between forward and inverse solutions is giving a regularizing effect.



**Figure 10.** Curves for Inverse Crime Detection. Solutions for reconstructions of simulated data over the generating mesh.

## 6. Conclusion

This paper proposes a new method of objective hyperparameter selection for use in one-step image reconstructions and compares it to some existing methods including heuristic selection. We present the following observations:

- (i) Heuristic selections of hyperparameter are inconsistent among experts and unrepeatable. This suggests that there is no single preferred value of  $\lambda$ ; rather there is a preferred region of  $\lambda$  over which reconstructions are not subjectively distinguishable. Moreover, it was not possible for observers to differentiate

- reconstructions based on heuristic hyperparameter selections from those produced from the objective methods.
- (ii) The GCV method is unreliable for the class of algorithms used in this work.
  - (iii) The L-Curve is, in general, shallow for EIT applications and is not reliable for all configurations (doesn't indicate a hyperparameter). When the method does work it provides a lower hyperparameter value than the Fixed NF and *BestRes* methods.
  - (iv) With  $NF = 1$ , the Fixed NF Method calculates a hyperparameter that falls in the minimum region of the Resolution Curve. At low noise levels  $\lambda_{NF=1}$  is very close to  $\lambda_{BestRes}$ . As AWGN is added to the simulated data  $\lambda_{BestRes}$  increases while  $\lambda_{NF=1}$  remains constant. However at the noise level found in our EIT equipment a NF of 1 produces good reconstructions that are close to the optimal reconstructions achieved with *BestRes* method.
  - (v) The Fixed NF Method provides a configuration independent method to select  $\lambda$  that is repeatable and is more consistent than expert selection. One could use the Fixed NF method with  $NF = 1$  to calculate a minimum hyperparameter value for any configuration. This method is repeatable and in applications with realistic noise levels will produce consistent stable reconstructions that are as good as heuristic selection.
  - (vi) Hyperparameters taken from the minimum region of the Resolution Curve (*BestRes* method) always produce good solutions that are comparable to, but more consistent than expert selections. Moreover  $\lambda_{BestRes}$  is optimal in terms of our figure of merit.

For the class of regularized reconstruction algorithms used in this work both the Fixed NF and *BestRes* methods provide objective methods to select a good value for  $\lambda$ . The values were indistinguishable from those selected by human experts. Both methods were developed using simulated data but shown to be applicable (validated) using Tank data.

We recommend using the *BestRes* method to calculate the hyperparameter for a specific configuration. The *BestRes* method works well using simulated data, additionally it can be tailored to compensate for realistic noise if impulse data can be obtained from the target equipment. In either case this hyperparameter is then associated with the configuration for all subsequent work using the configuration.

Although these methods do not solve the problem of obtaining an *optimal* hyperparameter value, they do provide a rationale and method for objectively and automatically selecting  $\lambda$ . Using the *BestRes* method with imaging equipment provides a sound engineering method for manufactures or researchers to obtain a configuration dependent hyperparameter that is optimal in terms of resolution. This allows end users to perform impedance imaging without the necessity of having to manually tweak parameters.

## References

- Adler A & Guardo R 1996 *IEEE Trans. Med. Imaging* **15**, 170–179.  
 Asfaw Y & Adler A 2005 *Physiological Measurement* **26**, S175–S183.  
 Barber C & Brown B 1988a *IEEE Trans. Biomedical Engineering* **32**, 100–105.  
 Barber D & Brown B 1988b *Clin. Phys. Physiol. Meas* **9**, 101–104.  
 Cheney M, Isaacson D, Newell J, Simske S & Goble J 1991 *Int. J. Imaging Syst. Technol* **2**, 66–75.

- Cohen-Bacrie C, Goussard Y & Guardo R 1997 *IEEE Trans. Medical Imaging* **16**, 562–571.
- Graham B 2005 Human performance in hyperparameter selection Technical Report TR-2005-03  
School of Information Technology and Engineering (SITE), University of Ottawa, Canada.
- Hansen P 1992 *SIAM Review* **34**, 561–581.
- Hansen P 1997 *Rank-Deficient and Discrete Ill-Posed Problems* 1st edn SIAM Philadelphia.
- Kotre C 1994 *Physiological Measurement* **15**, A125–A136.
- Lionheart W R 2004 *Physiological Measurement* **25**(1).
- Metherall P, Barber D, Smallwood R & Brown B 1996 *Nature* **380**, 509–512.
- Tenorio L 2001 *SIAM Review* **43**, 347–366.
- Vauhkonen M, Kaipio J, Somersalo E & Karjalainen P 1997 *Inverse Problems* **13**, 523–530.
- Wheeler J, Wang W & Tang M 2002 *Physiological Measurement* **23**(1), 169–176.
- Wirgin A 2004 *ArXiv Mathematical Physics e-prints* .