Selecting the Corner in the L-Curve Approach to Tikhonov Regularization

Peter R. Johnston and Ramesh M. Gulrajani*

Abstract—The performance of two methods for selecting the corner in the L-curve approach to Tikhonov regularization is evaluated via computer simulation. These methods are selecting the corner as the point of maximum curvature in the L-curve, and selecting it as the point where the product of abcissa and ordinate is a minimum. It is shown that both these methods resulted in significantly better regularization parameters than that obtained with an often-used empirical Composite REsidual and Smoothing Operator approach, particularly in conditions where correlated geometry noise exceeds Gaussian measurement noise. It is also shown that the regularization parameter that results with the minimum-product method is identical to that selected with another empirical zero-crossing approach proposed earlier.

Index Terms—Electrocardiography, electroencephalography, inverse problems, L-curve, regularization.

I. INTRODUCTION

Tikhonov regularization is often employed in the ill-posed inverse problems of electrocardiography and electroencephalography in order to stabilize the solution in the face of measurement noise or errors in geometry. In the zero-order version of Tikhonov regularization, the functional to be minimized is given by

$$M(\phi) = \|\mathbf{A}\phi - \phi_m\|^2 + t\|\phi\|^2 \tag{1}$$

where ϕ denotes the $n \times 1$ solution matrix, ϕ_m the $m \times 1$ column matrix of data potentials, and $\mathbf{A}$ the transfer matrix relating the two. The symbol $\|\ \ \|$ denotes the Euclidean norm and the regularization parameter t serves to determine the relative weight accorded to the residual error and the solution norm. The corresponding regularized solution is given by

$$\phi(t) = \left(\mathbf{A}^T\mathbf{A} + t\mathbf{I}\right)^{-1}\mathbf{A}^T\phi_m \tag{2}$$

where $\mathbf{I}$ is the identity matrix and the superscript T denotes the transpose. Too small a value for t results in continued instability of the solution, whereas too large a value results in an overregularized solution that, while stable, has an unnecessarily large residual error. If the true solution ϕ_s is known *a priori* (as in simulation studies), an optimal choice t_{opt} for t can be determined where the relative error (*RE*), defined as $RE(t) = \|\phi - \phi_s\|/\|\phi_s\|$, is a minimum. In a clinical situation, however, an alternative choice for t is needed that ideally should approach t_{opt}.

Two approaches for selecting t that have gained wide acceptance are Composite REsidual and Smoothing Operator (CRESO) [1] and the L-curve [2], [3]. The L-curve approach involves a plot, using a log-log scale, of the norm of the solution $\|\phi\|$ on the ordinate against the norm of the residual $\|\mathbf{A}\phi - \phi_m\|$ on the abscissa, with t as a parameter

along the resulting curve. As long as the uncorrelated measurement noise in ϕ_m dominates the more highly correlated geometry noise in $\mathbf{A}$, this plot is in the form of an L, and the t value at the corner of the L intuitively suggests an appropriate solution wherein both the norm and the residual simultaneously attain low values. Since the corner of the L is often rounded, Hansen and O'Leary [4] suggested that the corner be found as the point of maximum curvature. More recently, Lian *et al.* [5] have proposed that the t value, and hence the corner, be selected as the point on the L-curve where the product

$$P(t) = \|\phi\|\|\mathbf{A}\phi - \phi_m\| \tag{3}$$

is a minimum. This minimum-product criterion is a special case (corresponding to $\lambda = 1$) of a more general criterion that minimizes the product $(\|\phi\|^2)^\lambda(\|\mathbf{A}\phi - \phi_m\|^2)$ where $\lambda > 0$ [6]. The present short communication describes simulation tests that we have done to evaluate the Hansen *et al.* and Lian *et al.* choices for t. Results with both choices are compared to those obtained with the optimal and CRESO choices for t. We also show that the minimum-product criterion yields a t value that is identical to one that results with a "zero-crossing" approach recently proposed by us [7].

II. METHODS

A. Hansen and O'Leary's Point of Maximum Curvature

Many of the expressions needed below are most conveniently determined following a singular value decomposition $\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^T$ of the transfer matrix $\mathbf{A}$. For example, the regularized solution (2) can be written as [7]

$$\phi(t) = \sum_{i=1}^{n} \left(\frac{\sigma_{ii}\alpha_i}{\sigma_{ii}^2 + t}\right) v_i \tag{4}$$

where

σ_{ii} are the elements of the diagonal matrix Σ;
α_i is the scalar product $u_i^T\phi_m$;
u_i and v_i are the columns of $\mathbf{U}$ and $\mathbf{V}$, respectively.

Accordingly, since the matrices $\mathbf{U}$ and $\mathbf{V}$ are orthonormal, we have

$$\|\phi(t)\|^2 = \sum_{i=1}^{n} \frac{\sigma_{ii}^2\alpha_i^2}{(t + \sigma_{ii}^2)^2}. \tag{5}$$

Similarly, the residual can be written as [7]

$$\|\mathbf{A}\phi - \phi_m\|^2 = \sum_{i=1}^{n} \frac{t^2\alpha_i^2}{(t + \sigma_{ii}^2)^2} + \|\mathbf{r}_\perp\|^2 \tag{6}$$

where $\mathbf{r}_\perp$ is the residual vector of a conventional unregularized least-squares-error solution given by $\phi_{lse} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\phi_m$, i.e., $\mathbf{r}_\perp = \mathbf{A}\phi_{lse} - \phi_m$.

Hansen *et al.*'s choice is the point on the L-curve

$$(\zeta(t), \eta(t)) \equiv (\log\|\mathbf{A}\phi - \phi_m\|, \log\|\phi\|) \tag{7}$$

that has maximum curvature. Here, the curvature, κ, is defined as

$$\kappa(t) = \frac{\zeta'\eta'' - \zeta''\eta'}{\{(\zeta')^2 + (\eta')^2\}^{3/2}} \tag{8}$$

where differentiation is with respect to t. Implementation of the above formula is reasonably straightforward. Recall that

$$\eta(t) = \log \|\phi(t)\| = \tfrac{1}{2} \log \|\phi(t)\|^2 \qquad (9)$$

so

$$\eta'(t) = \frac{1}{2} \frac{d\|\phi(t)\|^2}{dt} \bigg/ \|\phi(t)\|^2 \qquad (10)$$

and

$$\eta''(t) = \frac{\dfrac{d^2\|\phi(t)\|^2}{dt^2}\|\phi(t)\|^2 - \left(\dfrac{d\|\phi(t)\|^2}{dt}\right)^2}{2\|\phi(t)\|^4} \qquad (11)$$

with similar expressions for $\zeta'(t)$ and $\zeta''(t)$ in terms of $\|\mathbf{A}\phi - \phi_m\|^2$ and its derivatives with respect to t. The first and second derivatives of $\|\phi\|^2$ and $\|\mathbf{A}\phi - \phi_m\|^2$ can be found by differentiating (5) and (6), respectively. Substituting the expressions for η', η'', ζ' and ζ'' in (8) finally allows evaluation of the curvature $\kappa(t)$. The maximum curvature is found by simply evaluating $\kappa(t)$ over the entire L-curve, and yields the t value denoted t_L in the sequel.

B. The Lian et al. Minimum-Product Criterion for the Corner

We wish to find the value of t that minimizes the product $P(t) = \|\phi\|\|\mathbf{A}\phi - \phi_m\|$. Since this product is always positive, we can look for the value of t that minimizes the squared product $\|\phi\|^2\|\mathbf{A}\phi - \phi_m\|^2$. To find the required value of t, we differentiate the squared product with respect to t, and set the derivative to zero. Employing (5) and (6), we have

$$\frac{d}{dt}[\|\phi\|^2\|\mathbf{A}\phi - \phi_m\|^2]$$

$$= \frac{d}{dt}\left[\left(\sum_{i=1}^n \frac{\sigma_{ii}^2\alpha_i^2}{(t+\sigma_{ii}^2)^2}\right)\left(\sum_{i=1}^n \frac{t^2\alpha_i^2}{(t+\sigma_{ii}^2)^2} + \|\mathbf{r}_\perp\|^2\right)\right]$$

$$= \left[-2\sum_{i=1}^n \frac{\sigma_{ii}^2\alpha_i^2}{(t+\sigma_{ii}^2)^3}\right]\left[\sum_{i=1}^n \frac{t^2\alpha_i^2}{(t+\sigma_{ii}^2)^2} + \|\mathbf{r}_\perp\|^2\right]$$

$$+ \left[2t\sum_{i=1}^n \frac{\sigma_{ii}^2\alpha_i^2}{(t+\sigma_{ii}^2)^2}\right]\left[\sum_{i=1}^n \frac{\sigma_{ii}^2\alpha_i^2}{(t+\sigma_{ii}^2)^3}\right]$$

$$= 2\left[\sum_{i=1}^n \frac{\sigma_{ii}^2\alpha_i^2}{(t+\sigma_{ii}^2)^3}\right]\left[t\sum_{i=1}^n \frac{\alpha_i^2(\sigma_{ii}^2 - t)}{(t+\sigma_{ii}^2)^2} - \|\mathbf{r}_\perp\|^2\right].$$

Since each term in the first summation is positive, it follows that the above derivative is zero when

$$t\sum_{i=1}^n \frac{\alpha_i^2(\sigma_{ii}^2 - t)}{(t+\sigma_{ii}^2)^2} - \|\mathbf{r}_\perp\|^2 = 0 \qquad (12)$$

Equation (12) can be written as $B(t) = 0$, where

$$B(t) = t\sum_{i=1}^n \frac{\alpha_i^2(\sigma_{ii}^2 - t)}{(t+\sigma_{ii}^2)^2} - \|\mathbf{r}_\perp\|^2 = t\|\phi(t)\|^2 - \|\mathbf{A}\phi - \phi_m\|^2. \qquad (13)$$

The zero-crossing approach described by us [7] also involved solving $B(t) = 0$, and taking the smallest t value for which a solution exists as the zero-crossing regularization parameter. Thus, the zero-crossing approach and the minimum-product corner criterion are equivalent, and the determined t value is denoted t_P in the sequel. The parameter t_P can be found either from the smallest zero of $B(t)$ or, equivalently, by plotting $P(t)$ and determining its minimum.

C. The CRESO Regularization Parameter

The CRESO regularization parameter t_{CRE} is determined as the *smallest* value of $t > 0$ that results in a local maximum of the function

$$C(t) = \|\phi(t)\|^2 + 2t\frac{d}{dt}\|\phi(t)\|^2. \qquad (14)$$

Using (5), this CRESO function can be rewritten as

$$C(t) = \sum_{i=1}^n \left[\frac{\sigma_{ii}\alpha_i}{t+\sigma_{ii}^2}\right]^2 \left[1 - \frac{4t}{t+\sigma_{ii}^2}\right]. \qquad (15)$$

This function can be plotted for different values of t and its first local maximum determined.

D. Simulation Protocol

The simulation protocol that we used to compare the performances of the maximum-curvature, minimum-product and CRESO t values is described in our earlier paper [7] and hence is only briefly summarized here. It consists of three radial dipoles (two pointing inward and one outward) placed inside a realistic-geometry epicardium comprising 610 nodes and 1216 triangles. Infinite-medium potential equations were used to approximate the epicardial potential distribution at the 610 node points. This epicardial source distribution was placed inside a homogeneous torso model that also comprised 610 nodes and 1216 triangles and the boundary element method employing a linear variation of potential over each triangle was used to compute a 610×610 $\mathbf{A}$ matrix, and hence, the 610 torso node potentials. The three dipoles resulted in an epicardial potential distribution with one maximum and two minima, but a smoothed torso potential distribution with a single maximum and a single minimum (see [7, Fig. 4]). Next, 168 of these torso potentials were used to inversely compute the potentials at 114 uniformly distributed epicardial nodes. The required 168×114 $\mathbf{A}$ matrix for the inverse solution was recomputed from a coarser 114-node epicardial mesh, with, however, the original 610-node torso, by first computing a 610×114 $\mathbf{A}$ matrix and then extracting rows corresponding to the 168 torso positions. Using the smaller 168×114 reduced $\mathbf{A}$ matrix in the inverse computations introduced an intrinsic amount of geometry noise on account of the simplified epicardial geometry. To this, we added some more geometry noise by offsetting the heart 1 cm in each of two diametrically opposite directions, one inward toward the torso center and the other outward toward the anterior torso surface. An incorrect reduced $\mathbf{A}$ matrix was computed each time and used for the inverse computations. The geometry noise introduced by the offsetting is an even more highly correlated form of noise than the intrinsic geometry noise. A final set of simulations added 0.5% Gaussian measurement noise to the 168 starting torso potentials.

III. RESULTS

As in our earlier study, the initial simulations with geometry noise alone verified that the L-curve approach broke down under these conditions. This is shown more explicitly in Fig. 1 (left column) where we show (from top to bottom) the L-curve, a plot of its curvature, the minimum product function $P(t)$ together with the relative error curve $RE(t)$, and the CRESO function $C(t)$, for simulations with intrinsic geometry noise plus the heart offset 1 cm inward. The L-curve has an indistinct corner. Its curvature plot exhibits three maxima, of which the first is the largest. If the t value (6×10^{-12}) corresponding to this first maximum is selected, the solution will be extremely underregularized, as may be verified by looking at the $RE(t)$ curve which reveals t_{opt} as 8×10^{-4}. Indeed, the t value corresponding to the third (smallest) peak in the curvature plot lies closest to t_{opt} and is the one that should be
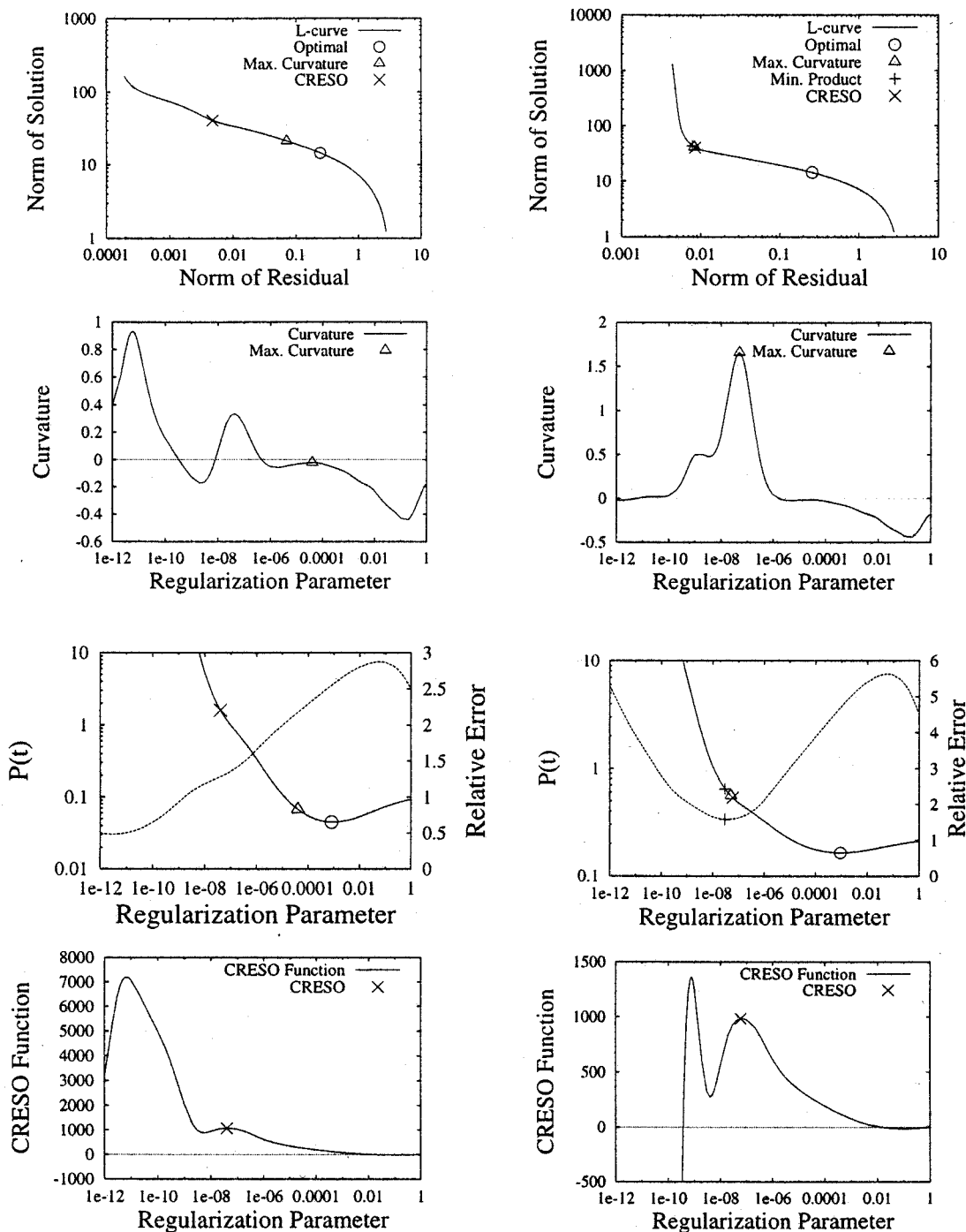
Fig. 1. Inverse solutions with realistic heart-torso geometries. The left-hand column shows the curves for intrinsic geometry noise plus geometry noise obtained by offsetting the heart 1 cm inward toward the center of the torso (see text). The right-hand column shows curves when, in addition to the intrinsic geometry noise and geometry noise due to a 1 cm heart offset inward, 0.5% Gaussian measurement noise is also included. The first row plots the L-curves, the second the curvature of the L-curves as a function of t, the third $P(t)$ and $RE(t)$ (dotted and solid trace, respectively), and the fourth the CRESO function $C(t)$. Values of t_{opt} ($\circ$), t_L ($\triangle$), t_P ($+$) and t_{CRE} ($\times$) are marked as appropriate.

selected, and is, accordingly, the one marked on the curvature and RE plots. Strictly speaking though, the absence of a corner in the L-curve alerts us to the fact that the curvature plot is meaningless, and accordingly that the maximum-curvature criterion for the corner breaks down. What is more interesting is the absence of a minimum in the product function $P(t)$. Thus, this function also serves as an indicator of the absence of a corner in the L-curve, and the minimum-product criterion also breaks down. Finally, the CRESO function reveals that t_{CRE} has to be selected corresponding to the second (rather than the first)

local maximum in order to be closest to t_{opt}. Thus, all three techniques for selecting t break down, with the minimum-product criterion identifying this breakdown by the absence of a minimum in $P(t)$ and the maximum-curvature criterion by the absence of a sharp corner in the L-curve. Very similar situations prevailed for the case of intrinsic geometry noise and that of intrinsic geometry noise plus the heart offset 1 cm outward (see Table I).

This complete breakdown is avoided when Gaussian measurement noise is added, even as little as 0.5% noise. The corner in the L-curve

TABLE I
SUMMARY OF INVERSION PARAMETERS FOR REALISTIC HEART-TORSO GEOMETRIES WITH THREE RADIAL DIPOLES WITHIN THE HEART

Simulation	Solution	t	RE
Intrinsic geometry noise	Optimal	4×10^{-5}	0.521
	Max Curvature*	4×10^{-5}	0.521
	CRESO*	1×10^{-7}	0.997
Intrinsic geometry noise +	Optimal	8×10^{-4}	0.652
Heart offset 1 cm inward	Max Curvature*	4×10^{-5}	0.827
	CRESO*	4×10^{-8}	2.206
Intrinsic geometry noise +	Optimal	8×10^{-6}	0.553
Heart offset 1 cm outward	Max Curvature*	1×10^{-4}	0.560
	CRESO*	2×10^{-8}	0.622
Intrinsic geometry noise +	Optimal	9×10^{-4}	0.652
Heart offset 1 cm inward +	Max Curvature	5×10^{-8}	2.241
0.5% Gaussian noise	Min Product	3×10^{-8}	2.428
	CRESO*	7×10^{-8}	2.140
Intrinsic geometry noise +	Optimal	1×10^{-5}	0.559
Heart offset 1 cm outward +	Max Curvature	2×10^{-7}	0.695
0.5% Gaussian noise	Min Product	2×10^{-7}	0.695
	CRESO	1×10^{-7}	0.749

Asterisks indicate where Maximum Curvature and CRESO techniques broke down. No Minimum Product solutions were obtained under conditions of geometry noise alone.

becomes prominent, the maximum in the curvature plot does identify the corner, and a minimum appears in the product function plot. These features are easily verified in Fig. 1 (right column) which shows the results when, in addition to the intrinsic geometry noise and heart offset inward, 0.5% noise was added to the body surface potentials. Note, however, that while both t_L and t_P fall in the corner region of the L-curve, t_{opt} falls to the right of the corner. Also, once again it is the second local maximum in the CRESO function that is closest to t_{opt}, and one could argue that the CRESO approach again breaks down. With intrinsic geometry noise, heart offset outward and 0.5% noise, however, all three techniques worked as stipulated with t values at the corner of the L-curve, but t_{opt} still remained to the right of the corner (Table I).

IV. DISCUSSION

The demonstration that the Lian *et al.* minimum-product corner criterion results in the same t values as our earlier more empirical zero-crossing approach offers a rational explanation as to why the latter approach for regularization parameter selection was successful, something that we were not able to show earlier, and that we redress here. Quite simply, the zero-crossing approach is equivalent to the L-curve approach employing a minimum-product corner criterion.

Our simulations show that t_L and t_P are generally superior to t_{CRE}, resulting more often than not in a lower *RE*. One advantage of t_P is that when correlated geometry noise dominates measurement noise, no

minimum appears in $P(t)$, thus, explicitly indicating that the method breaks down. The same can also be said for t_L as the corner in the L-curve disappears under these conditions, and when, accordingly, no maximum-curvature criterion is applicable. No such warning is provided by CRESO. Indeed, the first local maximum of the CRESO function under these geometry noise conditions may not yield the most appropriate solution. In the real clinical or experimental solution, however, once measurement noise is present, this drawback of t_{CRE} disappears.

Thus, when Gaussian measurement noise starts to dominate geometry noise, all three methods work. The parameters t_L, t_P and t_{CRE} all lie in the vicinity of the corner in the L-curve with t_{opt} to the right of the corner. Additional simulations described by us earlier [7] showed that with increased Gaussian noise, the corner moves outward, i.e., to the right, and eventually catches up with t_{opt}. At this point, t_L, t_P, t_{CRE}, and t_{opt} are all close to the corner. As the noise increases further and the corner continues to move outward, all four continue to move outward with the corner. Thus, it is mainly in the low-noise setting that significant differences are seen between all four, and it is in this low-noise setting that inverse solutions computed with t_L, t_P, and t_{CRE} exhibit larger relative errors than the optimum solution.

These simulations indicate that an appropriate strategy for inverse solutions obtained by Tikhonov regularization would be to first compute the L-curve and ensure the presence of a corner. The corner can then be identified either by the point of maximum curvature, or in easier fashion by plotting $P(t)$ and finding its minimum. Both these corner criteria will yield solutions that will either be close to the optimal solution or, if t_{opt} does not happen to be at the corner, solutions that will be underregularized.

REFERENCES

[1] P. Colli-Franzone, L. Guerri, B. Taccardi, and C. Viganotti, "Finite element approximation of regularised solutions of the inverse potential problem of electrocardiography and applications to experimental data," *Calcolo*, vol. XXII, no. 1, pp. 91–186, 1985.

[2] P. C. Hansen, "Truncated singular value decomposition solutions to discrete ill-posed problems with ill-determined numerical rank," *SIAM J. Sci. Stat. Comput.*, vol. 11, pp. 503–518, 1990.

[3] ——, "Analysis of discrete ill-posed problems by means of the L-curve," *SIAM Rev.*, vol. 34, pp. 561–580, 1992.

[4] P. C. Hansen and D. P. O'Leary, "The use of the L-curve in the regularization of discrete ill-posed problems," *SIAM J. Sci. Comput.*, vol. 14, pp. 1487–1503, 1993.

[5] J. Lian, D. Yao, and B. He, "A new method for implementation of regularization in cortical potential imaging," in *Proc. 20th Annu. Int. Conf. IEEE/EMBS*, 1998, pp. 2155–2158.

[6] T. Regińska, "A regularization parameter in discrete ill-posed problems," *SIAM J. Sci. Comput.*, vol. 17, pp. 740–749, 1996.

[7] P. R. Johnston and R. M. Gulrajani, "A new method for regularization parameter determination in the inverse problem of electrocardiography," *IEEE Trans. Biomed. Eng.*, vol. 44, pp. 19–39, Jan. 1997.