ESTIMATION OF THE L-CURVE VIA LANCZOS BIDIAGONALIZATION

D. CALVETTI *, G. H. GOLUB †, AND L. REICHEL ‡

Abstract. The L-curve criterion is often applied to determine a suitable value of the regularization parameter when solving ill-conditioned linear systems of equations with a right-hand side contaminated by errors of unknown norm. However, the computation of the L-curve is quite costly for large problems; the determination of a point on the L-curve requires that both the norm of the regularized approximate solution and the norm of the corresponding residual vector be available. Therefore, usually only a few points on the L-curve are computed, and these values rather than the L-curve, are used to determine a value of the regularization parameter. We propose a new approach to determine a value of the regularization parameter based on computing an L-ribbon that contains the L-curve in its interior. An L-ribbon can be computed fairly inexpensively by partial Lanczos bidiagonalization of the matrix of the given linear system of equations. A suitable value of the regularization parameter is then determined from the L-ribbon, and we show that an associated approximate solution of the linear system can be computed with little additional work.

Key words. Ill-posed problem, regularization, L-curve criterion, Gauss quadrature.

1. Introduction.

Let

$$A\mathbf{x} = \mathbf{b}, \quad A \in \mathrm{R}^{m\times n}, \quad \mathbf{x} \in \mathrm{R}^n, \quad \mathbf{b} \in \mathrm{R}^m, \tag{1}$$

be a linear system of equations with a very ill-conditioned matrix and a right-hand side vector that is contaminated by errors of unknown size. It is well known that in order to be able to compute a useful approximate solution of (1), the system has to be replaced by a less ill-conditioned nearby system. This replacement is known as regularization. One of the most popular regularization methods is Tikhonov regularization, in which the linear system (1) is replaced by the minimization problem

$$\min_{\mathbf{x}\in\mathrm{R}^n} \{\|A\mathbf{x} - \mathbf{b}\|^2 + \mu\|\mathbf{x}\|^2\}, \tag{2}$$

where $\mu \geq 0$ is the regularization parameter. Throughout this paper $\|\cdot\|$ denotes the Euclidean norm. The solution $\mathbf{x}_\mu$ of (2) also solves the linear system

$$(A^TA + \mu I)\mathbf{x} = A^T\mathbf{b}, \tag{3}$$

which shows that

$$\mathbf{x}_\mu = (A^TA + \mu I)^{-1}A^T\mathbf{b}. \tag{4}$$

A proper choice of the value of the regularization parameter μ is important. If the norm of the error in the right-hand side vector $\mathbf{b}$ is known, then the value of μ can be chosen so that the associated solution (4) of (2) satisfies the Morozov discrepancy principle; see, e.g., [15]. However, in many important applications the norm of the

* Department of Mathematics, Case Western Reserve University, Cleveland, OH 44106. E-mail: dxc57@po.cwru.edu. Research supported in part by NSF grant DMS-9404692.

† Department of Computer Science, Stanford University, Stanford, CA 94305. E-mail na.golub@na-net.ornl.gov. Research supported in part by NSF grant CCR-9505393.

‡ Department of Mathematics and Computer Science, Kent State University, Kent, OH 44242. E-mail: reichel@mcs.kent.edu. Research supported in part by NSF grant DMS-9404706.

error in the given right-hand side is not explicitly known, and other approaches to determining a value of the regularization parameter μ have to be employed.

Lawson and Hanson [14] observed in the '60s that an efficient way to display how the value of the regularization parameter affects the solution $\mathbf{x}_\mu$ of (2) and the residual error $\mathbf{b} - A\mathbf{x}_\mu$ is to plot the curve $(\|\mathbf{x}_\mu\|, \|\mathbf{b} - A\mathbf{x}_\mu\|)$ for $\mu \in (0, +\infty)$. This curve is known as the L-curve, because it is shaped roughly like an "L". Lawson and Hanson [14], and more recently Hansen and O'Leary [11, 13], proposed to choose the value of μ that corresponds to the point $(\|\mathbf{x}_\mu\|, \|\mathbf{b} - A\mathbf{x}_\mu\|)$ at the "vertex" of the "L". We denote this value by μ_L. A heuristic motivation for this value of the regularization parameter is that when $\mu > 0$ is "tiny", the associated solution $\mathbf{x}_\mu$ is likely to be very contaminated by propagated errors due to errors in the right-hand side. On the other hand, when μ is large, the solution $\mathbf{x}_\mu$ of (2) is a poor approximate solution of (1). The choice $\mu = \mu_L$ seeks to balance these sources of errors.

For many problems the L-curve indeed is roughly L-shaped, however, the "vertex" of the curve is marked more or less depending on the spectral properties of the matrix $A^T A$. It has been observed that the use of a log-log plot can be helpful in enhancing the sharpness of the "vertex". An analysis of the L-curve and its properties can be found in [11, 13]. Recent work by Engl and Grever [3], Hanke [10] and Vogel [16] point out limitations of the L-curve criterion. Nevertheless, computational experience shows the L-curve criterion to give suitable values of the regularization parameter for many problems.

A major difficulty when using the L-curve criterion for determining a value of the regularization parameter for large ill-posed problems is that it is expensive to compute points on the L-curve; the determination of each point requires the solution of a minimization problem (2). For large problems, one therefore typically only computes a few points on or close to the L-curve, and uses these points to determine a suitable value of the regularization parameter. Numerical issues related to locating the "vertex" of the L-curve in this manner are discussed in [13].

This paper proposes a new method for approximately determining the location of the "vertex" of the L-curve. We show how rectangular regions that contain points on the L-curve can be determined fairly inexpensively for several values of the regularization parameter without having to solve the corresponding regularized minimization problems (2). The union of these rectangles determine a ribbon-like region which contains the L-curve in its interior. We refer to this ribbon as the L-ribbon. When the L-ribbon is narrow, it is possible to infer the approximate location of the "vertex" of the L-curve from the shape of the L-ribbon.

The paper is organized as follows. In Section 2 we express the coordinates of points on the L-curve in terms of certain matrix functions of the regularization parameter μ. We can determine upper and lower bounds for the abscissa and ordinate for each point on the L-curve by computing upper and lower bounds for these matrix functions. Section 3 describes how Gauss and Gauss-Radau quadrature rules can be employed to compute such bounds, and Section 4 presents an algorithm for constructing the L-ribbon. We also show how to use already computed quantities to inexpensively determine an approximate solution of (3) that corresponds to a point in the L-ribbon. A few illustrative numerical examples are presented in Section 5, and Section 6 contains concluding remarks.

2. The L-curve and matrix functions. Let $\mu > 0$ be a given value of the regularization parameter for Tikhonov regularization, and let $P(\mu) = (\|\mathbf{x}_\mu\|, \|A\mathbf{x}_\mu - \mathbf{b}\|)$ be the coordinates of the corresponding point on the L-curve. It follows from (4) that

$$\|\mathbf{x}_\mu\|^2 = \mathbf{x}_\mu^T \mathbf{x}_\mu = \mathbf{b}^T A(A^T A + \mu I)^{-2} A^T \mathbf{b} \tag{5}$$

and

$$\begin{aligned} \|A\mathbf{x}_\mu - \mathbf{b}\|^2 &= (A(A^T A + \mu I)^{-1} A^T \mathbf{b} - \mathbf{b})^T (A(A^T A + \mu I)^{-1} A^T \mathbf{b} - \mathbf{b}) \\ &= \mathbf{b}^T \mathbf{b} + \mathbf{b}^T A(A^T A + \mu I)^{-1} A^T A(A^T A + \mu I)^{-1} A^T \mathbf{b} \\ &\quad -2\mathbf{b}^T A(A^T A + \mu I)^{-1} A^T \mathbf{b}. \end{aligned} \tag{6}$$

Substituting $K := A^T A$ and $\mathbf{c} := A^T \mathbf{b}$ into (6) yields

$$\|A\mathbf{x}_\mu - \mathbf{b}\|^2 = \mathbf{b}^T \mathbf{b} + \mathbf{c}^T K(K + \mu I)^{-2} \mathbf{c} - 2\mathbf{c}^T (K + \mu I)^{-1} \mathbf{c}.$$

Introduce the functions

$$\phi_1(t) := (t + \mu)^{-2}, \tag{7}$$

$$\phi_2(t) := t(t + \mu)^{-2} - 2(t + \mu)^{-1}, \tag{8}$$

and define the quantities

$$s_i := \mathbf{c}^T \phi_i(K)\mathbf{c}, \qquad i = 1, 2. \tag{9}$$

Then the coordinates of $P(\mu)$ can be expressed in terms of the quantities s_i as follows

$$\|\mathbf{x}_\mu\| = s_1^{1/2}, \tag{10}$$

$$\|A\mathbf{x}_\mu - \mathbf{b}\| = (\mathbf{b}^T \mathbf{b} + s_2)^{1/2}. \tag{11}$$

The s_i can be written as Stieltjes integrals. Introduce the spectral factorization

$$K = U \Lambda U^T, \tag{12}$$

where

$$\Lambda := \mathrm{diag}[\lambda_1, \lambda_2, \ldots, \lambda_n], \qquad 0 \leq \lambda_1 \leq \lambda_2 \ldots \leq \lambda_n,$$

and $U \in \mathrm{R}^{n \times n}$, $U^T U = I_n$. Here I_n denotes the identity matrix of order n. Define

$$\mathbf{h} = [h_1, h_2, \ldots, h_n]^T := U^T \mathbf{c}.$$

Then

$$s_i = \mathbf{h}^T \phi_i(\Lambda)\mathbf{h} = \sum_{k=1}^{n} \phi_i(\lambda_k) h_k^2 = \int_{-\infty}^{\infty} \phi_i(t) d\omega(t), \qquad i = 1, 2. \tag{13}$$

The discrete measure $\omega(t)$ is a non-decreasing step function with jump discontinuities at the eigenvalues λ_k of K. The jump at λ_k is h_k^2. We show in the next section how Gauss and Gauss-Radau quadrature rules can be applied to cheaply compute upper and lower bounds for the s_i.

3. Gauss quadrature. The quantities s_i defined by (9) depend on the parameter μ. Their computation for many different values of μ is not feasible when the matrix A is very large. However, the computation of upper and lower bounds for the s_i for several values of μ can be carried out efficiently by using Gauss quadrature rules. We remark that the computation of upper and lower bounds of certain matrix functions by application of Gauss quadrature rules is well known; see the papers by Golub, Meurant and Strakos [6, 7] and references therein. Recently, Golub and von Matt [9] exploited the connection between the Stieltjes integral in (13) and Gauss quadrature to derive methods different from ours for determining the regularization parameter.

We briefly review some known facts about Gauss quadrature rules and their connection with the Lanczos process. Define the inner product induced by the measure $\omega(\lambda)$ defined in (13),

$$< f, g >:= \int_{-\infty}^{\infty} f(t)g(t)d\omega(t) = \sum_{k=1}^{n} f(\lambda_k)g(\lambda_k)h_k^2 = \mathbf{h}^T f(\Lambda)g(\Lambda)\mathbf{h}, \tag{14}$$

and let $\{q_k\}_{k=0}^{n-1}$ be orthonormal polynomials with respect to this inner product, i.e.,

$$< q_k, q_j >= \begin{cases} 0, & k \neq j, \\ 1, & k = j. \end{cases} \tag{15}$$

The q_k satisfy a three-term recurrence relation of the form

$$tq_{k-1}(t) = \beta_k q_k(t) + \alpha_k q_{k-1}(t) + \beta_{k-1} q_{k-2}(t), \quad k = 1, 2, \ldots , \tag{16}$$

where $q_{-1}(t) := 0$ and $q_0(t) :=< 1, 1 >^{-1/2}$. It is well-known that ℓ steps of the Lanczos process applied to the matrix $K = A^T A$ with initial vector $\mathbf{c} := A^T\mathbf{b}$ yields the symmetric positive definite or semidefinite tridiagonal matrix

$$T_\ell := \begin{bmatrix} \alpha_1 & \beta_1 & & & \\ \beta_1 & \alpha_2 & \beta_2 & & \\ & \ddots & \ddots & \ddots & \\ & & \beta_{\ell-2} & \alpha_{\ell-1} & \beta_{\ell-1} \\ & & & \beta_{\ell-1} & \alpha_\ell \end{bmatrix} \tag{17}$$

whose entries are the first $2\ell - 1$ coefficients in the recurrence relation (16); see, e.g., [6, 7]. We assume here that ℓ is sufficiently small to secure that $\beta_j > 0$ for $1 \leq j < \ell$. Symmetric tridiagonal matrices with positive subdiagonal elements are often referred to as Jacobi matrices.

Our algorithm for computing the L-ribbon only requires the Cholesky factor B_ℓ of T_ℓ. We compute B_ℓ by the Lanczos bidiagonalization algorithm.

Algorithm 1 (Lanczos Bidiagonalization Algorithm).
Input: $\mathbf{c} := A^T\mathbf{b} \in \mathrm{R}^n\backslash\{0\}$, $A \in \mathrm{R}^{m\times n}$, $0 < \ell < n$;
Output: $\{\mathbf{p}_j\}_{j=0}^{\ell-1}$, $\{\mathbf{q}_j\}_{j=0}^{\ell}$, $\{\delta_j\}_{j=1}^{\ell}$, $\{\gamma_j\}_{j=1}^{\ell}$;
$\mathbf{q}_0 := \mathbf{c}/\|\mathbf{c}\|$; $\mathbf{r}_0 := A\mathbf{q}_0$;
$\gamma_1 = \|\mathbf{r}_0\|$; $\mathbf{p}_0 := \mathbf{r}_0/\gamma_1$;
for $j = 1, 2, \ldots, \ell$ **do**
  $\mathbf{s}_j := A^T\mathbf{p}_{j-1} - \gamma_j\mathbf{q}_{j-1}$;
  $\delta_j := \|\mathbf{s}_j\|$; $\mathbf{q}_j := \mathbf{s}_j/\delta_j$;
  if $j = \ell$ **then** *exit*;
  $\mathbf{r}_j := A\mathbf{q}_j - \delta_j\mathbf{p}_{j-1}$;
  $\gamma_{j+1} := \|\mathbf{r}_j\|$; $\mathbf{p}_j := \mathbf{r}_j/\gamma_{j+1}$

end *j;* □

We assume that the parameter ℓ in Algorithm 1 is chosen small enough so that $\gamma_j > 0$ and $\delta_j > 0$ for $1 \leq j \leq \ell$. Then the algorithm determines the $n \times \ell$ matrices $P_\ell = [\mathbf{p}_0, \mathbf{p}_1, \ldots, \mathbf{p}_{\ell-1}]$ and $Q_\ell = [\mathbf{q}_0, \mathbf{q}_1, \ldots, \mathbf{q}_{\ell-1}]$ with orthonormal columns, as well as the upper bidiagonal matrix

$$B_\ell := \begin{bmatrix} \gamma_1 & \delta_1 & & \\ & \ddots & \ddots & \\ & & \gamma_{\ell-1} & \delta_{\ell-1} \\ & & & \gamma_\ell \end{bmatrix}, \tag{18}$$

such that

$$T_\ell = B_\ell^T B_\ell.$$

Let $\mathbf{e}_j = [0, \ldots, 0, 1, 0, \ldots, 0]^T$ be the jth axis vector. Then $\mathbf{e}_j^T \mathbf{q}_k = q_k(\lambda_j) h_j$, and the orthonormality of the vectors $\mathbf{q}_k$ follows from the orthonormality of the polynomials q_k with respect to the inner product (14). The matrices B_ℓ, P_ℓ and Q_ℓ determined by Algorithm 1 satisfy

$$\begin{aligned} AQ_\ell &= P_\ell B_\ell, \\ A^T P_\ell &= Q_\ell B_\ell^T + \delta_\ell \mathbf{q}_\ell \mathbf{e}_\ell^T, \end{aligned}$$

and it follows that

$$KQ_\ell = Q_\ell T_\ell + \beta_\ell \mathbf{q}_\ell \mathbf{e}_\ell^T, \tag{19}$$

where $K = A^T A$ and $\beta_\ell = \gamma_\ell \delta_\ell$. We refer to (19) as a Lanczos decomposition of K.

It is convenient to discuss Gauss quadrature in terms of the matrix T_ℓ and modifications thereof. Introduce the spectral factorization

$$T_\ell = V_\ell \Theta_\ell V_\ell^T, \tag{20}$$

where $V_\ell \in \mathrm{R}^{\ell \times \ell}$, $V_\ell^T V_\ell = I_\ell$ and $\Theta_\ell = \mathrm{diag}[\theta_1, \theta_2, \ldots, \theta_\ell]$ with $\theta_1 < \theta_2 < \ldots < \theta_\ell$. It is well known that the eigenvalues of T_ℓ are the zeros of the orthonormal polynomial $q_\ell(t)$, as well as the nodes of the ℓ-point Gauss quadrature rule

$$\mathcal{G}_\ell(f) := \sum_{i=1}^{\ell} f(\theta_i) \omega_i \tag{21}$$

with respect to the measure $\omega(\lambda)$ introduced in (13). The weights of the quadrature rule (21) are given by

$$\omega_i := \|\mathbf{c}\|^2 (\mathbf{e}_1^T V_\ell \mathbf{e}_i)^2,$$

where $\mathbf{c} := A^T \mathbf{b}$. Thus,

$$\begin{aligned} \mathcal{G}_\ell(f) &= \|\mathbf{c}\|^2 \sum_{i=1}^{\ell} f(\theta_i) (\mathbf{e}_1^T V_\ell \mathbf{e}_i)^2 = \|\mathbf{c}\|^2 \mathbf{e}_1^T V_\ell f(\Theta_\ell) V_\ell^T \mathbf{e}_1 \\ &= \|\mathbf{c}\|^2 \mathbf{e}_1^T f(T_\ell) \mathbf{e}_1. \end{aligned} \tag{22}$$

We remark that (22) yields an efficient way to evaluate the Gauss quadrature rule when the tridiagonal matrix T_ℓ is available.

Analogously, the ℓ-point Gauss-Radau quadrature rule with one assigned node at $a \leq \lambda_1$ can be evaluated as

$$\mathcal{R}_\ell(f) = \|\mathbf{c}\|^2 \mathbf{e}_1^T f(T_\ell^{(a)}) \mathbf{e}_1, \tag{23}$$

where

$$T_\ell^{(a)} := \begin{bmatrix} \alpha_1 & \beta_1 & & & \\ \beta_1 & \alpha_2 & \beta_2 & & \\ & \ddots & \ddots & \ddots & \\ & & \beta_{\ell-2} & \alpha_{\ell-1} & \beta_{\ell-1} \\ & & & \beta_{\ell-1} & \alpha_\ell^{(a)} \end{bmatrix} \tag{24}$$

and

$$\alpha_\ell^{(a)} := a + \beta_{\ell-1}^2 \mathbf{e}_{\ell-1}^T (T_{\ell-1} - aI)^{-1} \mathbf{e}_{\ell-1}; \tag{25}$$

see Golub [5]. In particular, the tridiagonal matrix $T_\ell^{(0)}$ associated with the ℓ-point Gauss-Radau rule with a prescribed node at the origin can be written as

$$T_\ell^{(0)} = (B_\ell^{(0)})^T B_\ell^{(0)},$$

where $B_\ell^{(0)}$ is obtained from B_ℓ by setting $\gamma_\ell = 0$; see Golub and von Matt [8].

Introduce the quadrature error

$$\mathcal{E}_{\mathcal{Q}_\ell}(f) := \int_{-\infty}^{\infty} f(t) d\omega(t) - \mathcal{Q}_\ell(f),$$

where $\mathcal{Q}_\ell = \mathcal{G}_\ell$ or $\mathcal{Q}_\ell = \mathcal{R}_\ell$. In the former case, there exists $\theta_{\mathcal{G}_\ell}$, $\lambda_1 \leq \theta_{\mathcal{G}_\ell} \leq \lambda_n$, such that

$$\mathcal{E}_{\mathcal{G}_\ell}(f) = \frac{f^{(2\ell)}(\theta_{\mathcal{G}_\ell})}{(2\ell)!} \int_{-\infty}^{\infty} \prod_{i=1}^{\ell} (t - \theta_i)^2 d\omega(t), \tag{26}$$

where $f^{(j)}$ denotes the jth derivative of the function f. On the other hand, if $\mathcal{Q}_\ell = \mathcal{R}_\ell$ then there exists $\theta_{\mathcal{R}_\ell}$, $\lambda_1 \leq \theta_{\mathcal{R}_\ell} \leq \lambda_n$, such that

$$\mathcal{E}_{\mathcal{R}_\ell}(f) = \frac{f^{(2\ell-1)}(\theta_{\mathcal{R}_\ell})}{(2\ell-1)!} \int_{-\infty}^{\infty} (t-a) \prod_{i=2}^{\ell} (t - \hat{\theta}_i)^2 d\omega(t), \tag{27}$$

where $a = \hat{\theta}_1 < \hat{\theta}_2 < \ldots < \hat{\theta}_\ell$ are the nodes of the Gauss-Radau quadrature rule and the eigenvalues of the matrix (24).

PROPOSITION 3.1. *Let ϕ_1 and ϕ_2 be defined by (7) and (8). Then, for $t \geq 0$ and $\ell \geq 1$,*

$$\begin{aligned} \phi_1^{(2\ell-1)}(t) < 0, \quad \phi_1^{(2\ell)}(t) > 0, \\ \phi_2^{(2\ell-1)}(t) > 0, \quad \phi_2^{(2\ell)}(t) < 0. \end{aligned} \tag{28}$$

Therefore

$$\begin{array}{ll} \mathcal{E}_{\mathcal{G}_\ell}(\phi_1) > 0, & \mathcal{E}_{\mathcal{R}_\ell}(\phi_1) < 0, \\ \mathcal{E}_{\mathcal{G}_\ell}(\phi_2) < 0, & \mathcal{E}_{\mathcal{R}_\ell}(\phi_2) > 0. \end{array} \tag{29}$$

Proof. The inequalities (28) for ϕ_2 follow from $\phi_2^{(j)}(t) = (-1)^{j+1} j!(t+\mu)^{-j-2}(t + (j+2)\mu)$. Substitution of these inequalities into (26) and (27) shows (29) for ϕ_2. The inequalities involving ϕ_1 are shown similarly. □

It follows from Proposition 3.1 that upper and lower bounds for the quantities s_1 and s_2 can be determined by replacing the Stieltjes integrals in (13) by Gauss and Gauss-Radau quadrature rules.

4. Estimation of the regularization parameter by the L-ribbon. We show how to compute upper and lower bounds for the abscissa and ordinate of one point on the L-curve corresponding to a fixed value of the regularization parameter. Then, we discuss how bounds for the coordinates of other points on the L-curve can be computed without carrying out additional matrix-vector products with the matrices A or A^T. This allows us to inexpensively determine a ribbon-like region, the L-ribbon, which contains the L-curve in its interior. The ribbon is determined by first computing the bidiagonal matrix B_ℓ by Algorithm 1, and its width decreases as ℓ increases; when ℓ in Algorithm 1 is sufficiently large, the L-ribbon has zero width, and coincides with the L-curve. We also show how to cheaply compute an approximate solution $\mathbf{x}_{\mu,\ell}$ of (3) associated with a desired value of the regularization parameter.

Assume, for the moment, that the value of the regularization parameter μ is given, and let $P_\mu = (\|\mathbf{x}_\mu\|, \|\mathbf{b} - A\mathbf{x}_\mu\|)$ be the associated point on the L-curve. After application of Algorithm 1, we can compute upper and lower bounds for $\|\mathbf{x}_\mu\|$ by determining upper and lower bounds for the quantity s_1 associated with this value of the regularization parameter; see (10). The latter are computed by approximating the integral in (9) by ℓ-point Gauss-Radau and Gauss quadrature rules. We have

$$\mathcal{G}_\ell(\phi_1) \leq s_1 \leq \mathcal{R}_\ell(\phi_1).$$

It follows from (22) and (23) that

$$\mathcal{G}_\ell(\phi_1) = \|\mathbf{c}\|^2 \mathbf{e}_1^T (B_\ell^T B_\ell + \mu I)^{-2} \mathbf{e}_1, \tag{30}$$

$$\mathcal{R}_\ell(\phi_1) = \|\mathbf{c}\|^2 \mathbf{e}_1^T ((B_\ell^{(0)})^T B_\ell^{(0)} + \mu I)^{-2} \mathbf{e}_1. \tag{31}$$

These quadrature rules can be evaluated by determining vectors $\mathbf{y}$, such that

$$(B^T B + \mu I)\mathbf{y} = \mathbf{e}_1, \qquad B \in \{B_\ell, B_\ell^{(0)}\}. \tag{32}$$

These vectors can be computed efficiently by solving the least-squares problems

$$\min_{\mathbf{y} \in \mathrm{R}^\ell} \left\| \begin{bmatrix} B \\ \mu^{1/2} I \end{bmatrix} \mathbf{y} - \mu^{-1/2} \mathbf{e}_{\ell+1} \right\|, \qquad B \in \{B_\ell, B_\ell^{(0)}\}, \tag{33}$$

with the aid of Givens rotations; see [2, 4]. For future reference, we denote the solution of (32) corresponding to $B = B_\ell$ by $\mathbf{y}_{\mu,\ell}$, and the solution corresponding to $B = B_\ell^{(0)}$ by $\mathbf{y}_{\mu,\ell}^{(0)}$. Note that the matrices B_ℓ and $B_\ell^{(0)}$ are independent of the regularization

parameter μ. Therefore, given these matrices, the Gauss and Gauss-Radau rules (30) and (31) can be evaluated in only $\mathcal{O}(\ell)$ arithmetic floating point operations for each value of μ. The scaling factor $\|\mathbf{c}\|$ in the quadrature rule is computed in Algorithm 1.

Similarly, we determine bounds for $\|\mathbf{b}-A\mathbf{x}_\mu\|$ by computing bounds for the quantity s_2 defined in (9). It follows from Proposition 3.1 that

$$\mathcal{R}_\ell(\phi_2) \leq s_2 \leq \mathcal{G}_\ell(\phi_2),$$

where

$$\begin{aligned}
\mathcal{G}_\ell(\phi_2) &= \|\mathbf{c}\|^2(\mathbf{e}_1^T B_\ell^T B_\ell (B_\ell^T B_\ell + \mu I)^{-2}\mathbf{e}_1 - 2\mathbf{e}_1^T (B_\ell^T B_\ell + \mu I)^{-1}\mathbf{e}_1),\\
\mathcal{R}_\ell(\phi_2) &= \|\mathbf{c}\|^2(\mathbf{e}_1^T (B_\ell^{(0)})^T B_\ell^{(0)} ((B_\ell^{(0)})^T B_\ell^{(0)} + \mu I)^{-2}\mathbf{e}_1 - 2\mathbf{e}_1^T ((B_\ell^{(0)})^T B_\ell^{(0)} + \mu I)^{-1}\mathbf{e}_1).
\end{aligned}$$

The solutions $\mathbf{y}_{\mu,\ell}$ and $\mathbf{y}_{\mu,\ell}^{(0)}$ of (32) can be used to evaluate these quadrature rules. We obtain

$$\begin{aligned}
\mathcal{G}_\ell(\phi_2) &= \|\mathbf{c}\|^2(\|B_\ell \mathbf{y}_{\mu,\ell}\|^2 - 2\mathbf{e}_1^T \mathbf{y}_{\mu,\ell}),\\
\mathcal{R}_\ell(\phi_2) &= \|\mathbf{c}\|^2(\|B_\ell^{(0)} \mathbf{y}_{\mu,\ell}^{(0)}\|^2 - 2\mathbf{e}_1^T \mathbf{y}_{\mu,\ell}^{(0)}),
\end{aligned}$$

and, given the matrices B_ℓ and $B_\ell^{(0)}$, these formulas can be used to evaluate $\mathcal{G}(\phi_2)$ and $\mathcal{R}(\phi_2)$ in only $\mathcal{O}(\ell)$ arithmetic floating point operations for each value of μ.

In view of (10) and (11), we obtain the following bounds for the point $P_\mu = (\|\mathbf{x}_\mu\|, \|\mathbf{b}-A\mathbf{x}_\mu\|)$ on the L-curve:

$$\begin{aligned}
(\mathcal{G}_\ell(\phi_1))^{1/2} \leq &\quad \|\mathbf{x}_\mu\| &\leq (\mathcal{R}_\ell(\phi_1))^{1/2},\\
(\mathbf{b}^T\mathbf{b} + \mathcal{R}_\ell(\phi_2))^{1/2} \leq &\quad \|\mathbf{b}-A\mathbf{x}_\mu\| &\leq (\mathbf{b}^T\mathbf{b} + \mathcal{G}_\ell(\phi_2))^{1/2}.
\end{aligned}$$

Introduce the quantities

$$\begin{aligned}
x^-(\mu) &:= (\mathcal{G}_\ell(\phi_1))^{1/2}\\
x^+(\mu) &:= (\mathcal{R}_\ell(\phi_1))^{1/2}\\
y^-(\mu) &:= (\mathbf{b}^T\mathbf{b} + \mathcal{R}_\ell(\phi_2))^{1/2}\\
y^+(\mu) &:= (\mathbf{b}^T\mathbf{b} + \mathcal{G}_\ell(\phi_2))^{1/2}
\end{aligned}$$

for $\mu > 0$. We define the L-ribbon as the union of the rectangular regions

$$\bigcup_{\mu>0} \{\{x(\mu), y(\mu)\} : x^-(\mu) \leq x(\mu) \leq x^+(\mu), y^-(\mu) \leq y(\mu) \leq y^+(\mu)\}.$$

The following algorithm determines rectangles associated with the parameter values μ_j, $1 \leq j \leq p$.

Algorithm 2 (L-Ribbon Algorithm).
Input: $\mathbf{b} \in \mathrm{R}^m$, $A \in \mathrm{R}^{m\times n}$, ℓ, $\{\mu_j\}_{j=1}^p$;
Output: $\{x_j^+\}_{j=1}^p$, $\{x_j^-\}_{j=1}^p$, $\{y_j^+\}_{j=1}^p$, $\{y_j^-\}_{j=1}^p$;

i) Apply Algorithm 1 to compute the entries of the bidiagonal matrix B_ℓ. Define the matrix $B_\ell^{(0)}$.

ii) for $j = 1, 2, \ldots, p$ *do*
$\mu := \mu_j$;
% The functions ϕ_1 *and* ϕ_2 *below depend on* μ *%*
Evaluate $\mathcal{G}_\ell(\phi_1)$, $\mathcal{G}_\ell(\phi_2)$, $\mathcal{R}_\ell(\phi_1)$, $\mathcal{R}_\ell(\phi_2)$;
$x_j^- := x^-(\mu_j)$; $x_j^+ := x^+(\mu_j)$;
$y_j^- := y^-(\mu_j)$; $y_j^+ := y^+(\mu_j)$;
endfor

□

The following proposition shows how to inexpensively evaluate an approximate solution of (3) that corresponds to a point in the L-ribbon.

PROPOSITION 4.1. *Let* $\mu > 0$ *be a desired value of the regularization parameter, and let* $\mathbf{x}_\mu$ *be the associated solution of (3). Let* $\mathbf{y}_{\mu,\ell}$ *satisfy (32) with* $B = B_\ell$, *and consider the approximation*

$$\mathbf{x}_{\mu,\ell} := \|\mathbf{c}\| Q_\ell \mathbf{y}_{\mu,\ell} \tag{34}$$

of $\mathbf{x}_\mu$, *where the* $n \times \ell$ *matrix* $Q_\ell = [\mathbf{q}_0, \mathbf{q}_1, \ldots, \mathbf{q}_{\ell-1}]$ *is determined by Algorithm 1. Then* $\|\mathbf{x}_{\mu,\ell}\| = (\mathcal{G}_\ell(\phi_1))^{1/2}$ *and* $\|\mathbf{b} - A\mathbf{x}_{\mu,\ell}\| = (\mathbf{b}^T\mathbf{b} + \mathcal{G}_\ell(\phi_2))^{1/2}$.

Proof. The residual error $\mathbf{r}_{\mu,\ell} := A^T\mathbf{b} - (A^TA + \mu I)\mathbf{x}_{\mu,\ell}$ associated with the approximate solution (34) of (3) is orthogonal to the columns of the matrix Q_ℓ. Thus, $\mathbf{x}_{\mu,\ell}$ is a Galerkin solution of (3). Substitution of (34) into $\|\mathbf{x}_{\mu,\ell}\|$ and $\|\mathbf{b} - A\mathbf{x}_{\mu,\ell}\|$, and using that $\mathbf{y}_{\mu,\ell}$ satisfies (32) with $B = B_\ell$, shows that these norms can be evaluated as ℓ-point Gauss quadrature rules for the integrands ϕ_1 and ϕ_2, respectively. □

The proposition shows that the approximate solution (34) corresponds to a vertex in the rectangular region of the L-ribbon associated with the value μ of the regularization parameter.

5. Computed examples. The computations of this section were carried out with Matlab on an HP9000/777 workstation. We seek to solve the Fredholm integral equation of the first kind

$$\int_0^\pi \exp(s\cos(t))x(t)dt = 2\frac{\sinh(s)}{s}, \qquad 0 \le s \le \frac{\pi}{2},$$

considered by Baart [1]. The integral equation is discretized by a Galerkin method with orthonormal box functions. We used the Matlab code by Hansen [12] for computing the nonsymmetric matrix $A \in \mathrm{R}^{200\times 200}$ and right-hand side vector $\mathbf{b}^{\text{exact}} \in \mathrm{R}^{200}$. A "noise vector" $\mathbf{w} \in \mathrm{R}^{200}$ is generated by first determining a vector with normally distributed random numbers with zero mean and variance one as components, and then scaling this vector to be of desired length. The right-hand side in the linear system (1) to be solved is defined as $\mathbf{b} := \mathbf{b}^{\text{exact}} + \mathbf{w}$.

Example 5.1. The noise vector $\mathbf{w}$ is scaled to have norm $1 \cdot 10^{-2}$. We apply Algorithm 2 with values of the regularization parameter $\mu_j := 5 \cdot 10^{-7+(j-1)/4}$, $1 \le j \le 20$. The number of Lanczos steps ℓ is chosen to be 6. Rectangles of the L-ribbon associated with the values μ_j generated by Algorithm 2 are displayed in Figure 1. The vertices $\{x_j^-, y_j^+\}$ of these rectangles are marked by "x". These vertices correspond to the solutions $\mathbf{x}_{\mu_j,6}$. When the rectangular region is "tiny" only the vertex "x" is visible.

For comparison, Figure 1 also displays points on the L-curve associated with the values μ_j of the regularization parameter. These points are marked by small circles

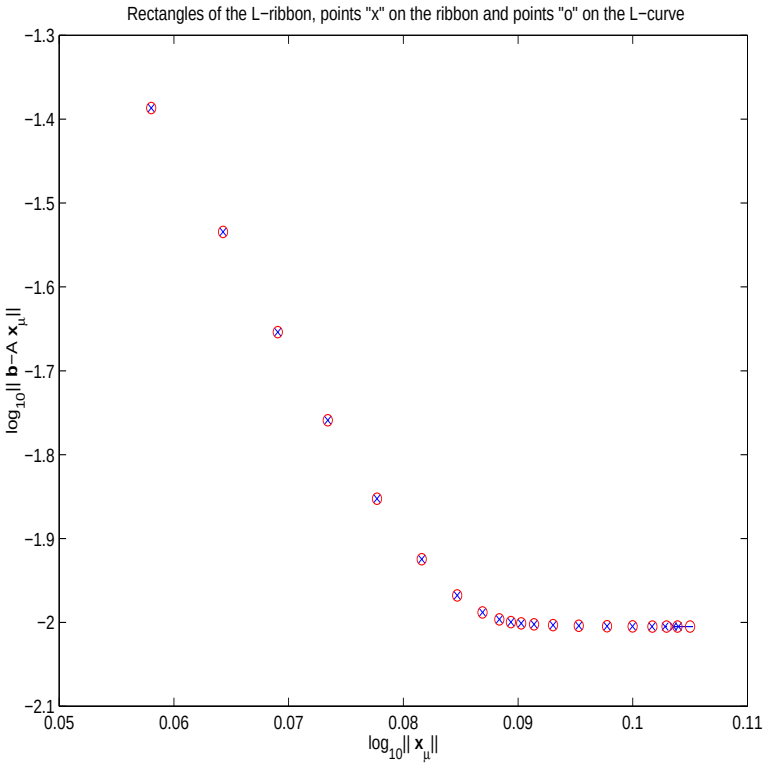

FIG. 1. *Example 5.1:* $\ell = 6$ *in Alg. 2 and* $\|\mathbf{w}\| = 1 \cdot 10^{-2}$

"o". The points "o" on the L-curve on the right-hand side of Figure 1 are associated with the smallest μ_j and the points on the left-hand side of the figure with the largest μ_j. For the larger values μ_j the vertices "x" coincide with the points on the L-curve associated with the same value of the regularization parameter to plotting accuracy.

Figure 1 shows where the "vertex" of the L-curve is located, and can be used to determine a suitable value of the regularization parameter, e.g., $\mu := \mu_{12} = 5 \cdot 10^{-17/4}$. The rectangles of the L-ribbon close to the vertex are "tiny", and therefore the use of only $\ell = 6$ steps of the Lanczos process in Algorithm 2 yields an approximate solution $\mathbf{x}_{\mu_{12},6}$ of (3) of sufficient accuracy. As pointed out in Section 4, this approximate solution is very inexpensive to compute.

We envision the L-ribbon to be used in an interactive manner; a user would plot rectangles of the ribbon associated with certain values of the regularization parameter and a certain number of Lanczos steps ℓ. The plotted output would guide in the selection of new values of the regularization parameters, and would show if the number of Lanczos steps ℓ in Algorithm 2 needs to be increased.

Figure 1 shows the rectangles of the L-ribbon to increase in size for small values of μ, and a more detailed image of this behavior is provided by Figure 2. The latter figure is generated by computing rectangles of the L-curve for $\mu_j := 5 \cdot 10^{-7+(j-1)/40}$, $1 \leq j \leq 40$, using the same quadrature rules as for Figure 1.

Figures 1 and 2 show that the Gauss and Gauss-Radau quadrature rules yield highest accuracy for the larger values μ_j. This depends on that the integrands (7) and (8) have poles at $t = -\mu_j$, and these poles are closer to the spectrum of $K = A^T A$, the closer μ_j is to the origin. The presence of a pole close to the spectrum of K reduces

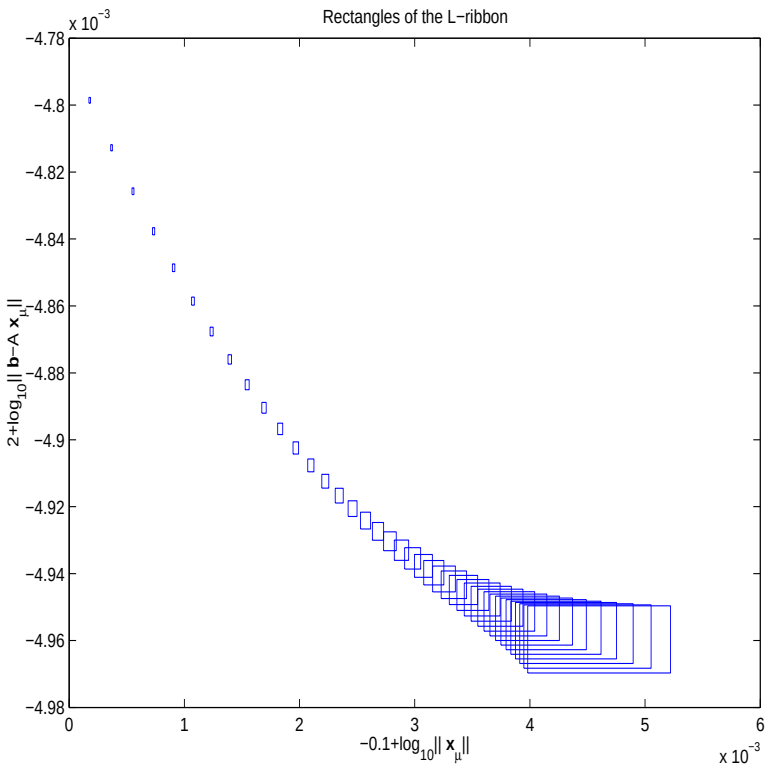

FIG. 2. *Example 5.1:* $\ell = 6$ *in Alg. 2 and* $\|\mathbf{w}\| = 1 \cdot 10^{-2}$

the accuracy of the Gauss and Gauss-Radau quadrature rules.

Figure 3 illustrates the accuracy in the computed solutions $\mathbf{x}_{\mu_j,6}$ for the four smallest values μ_j in Figure 1. The figure shows that the points on the L-curve associated with the values μ_j, marked by small circles "o" are further away from the corresponding vertices "x" the smaller the value μ_j. Thus, if an approximate solution of (1) associated with a small value of the regularization parameter were desired, then an increase of the number of Lanczos steps ℓ may be desirable.

6. Conclusion. We have demonstrated that the L-ribbon is a versatile tool for interactive solution of large-scale linear discrete ill-posed problems. The ribbon often is inexpensive to determine, and yields insight into both how to select a suitable value of the regularization parameter, and how many Lanczos steps to carry out for the computation of an approximate solution of (1).

REFERENCES

[1] M. L. Baart, *The use of auto-correlation for pseudo-rank determination in noisy ill-conditioned least-squares problems*, IMA J. Numer. Anal. 2 (1982), pp. 241–247.
[2] L. Eldén, *Algorithms for the regularization of ill-conditioned least squares problems*, BIT 17 (1977), pp. 134–145.
[3] H. W. Engl and W. Grever, *Using the L-curve for determining optimal regularization parameters*, Numer. Math. 69 (1994), pp. 25–31.
[4] W. Gander, *Least squares with a quadratic constraint*, Numer. Math. 36 (1981), pp. 291–307.
[5] G. H. Golub, *Some modified matrix eigenvalue problems*, SIAM Review 15 (1973), pp. 318–334.
[6] G. H. Golub and G. Meurant, *Matrices, moments and quadrature*, in Numerical Analysis 1993, eds. D. F. Griffiths and G. A. Watson, Longman, Essex, England, 1994, pp. 105–156.

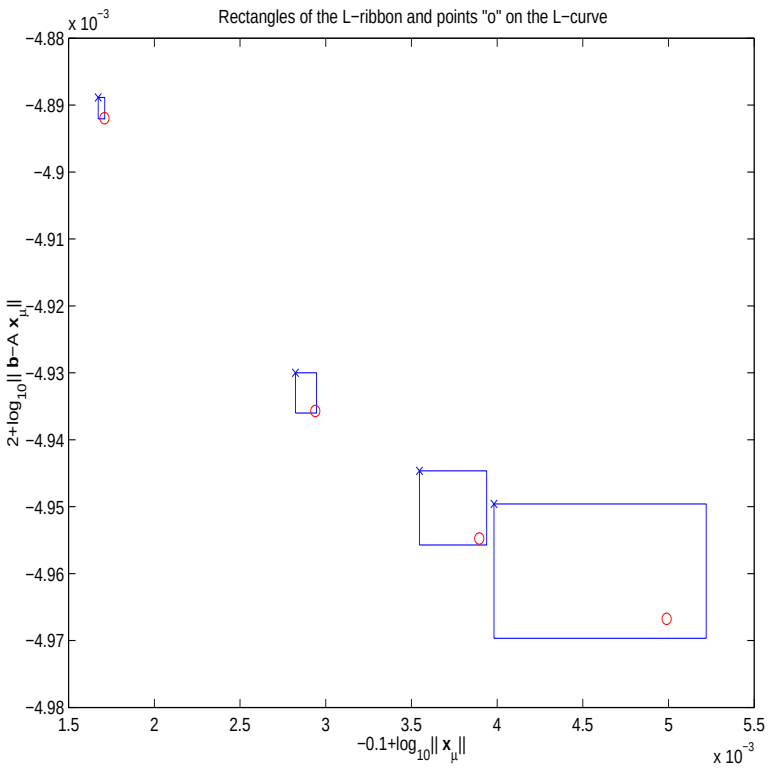

FIG. 3. *Example 5.1:* $\ell = 6$ *in Alg. 2 and* $\|\mathbf{w}\| = 1 \cdot 10^{-2}$

[7] G. H. Golub and Z. Strakos, *Estimates in quadratic formulas*, Numer. Algor. 8 (1994), pp. 241–268.

[8] G. H. Golub and U. von Matt, *Generalized cross-validation for large scale problems*, J. Comput. Graph. Stat. 6 (1997), pp. 1–34.

[9] G. H. Golub and U. von Matt, *Tikhonov regularization for large scale problems*, in Workshop on Scientific Computing, eds. G. H. Golub, S. H. Lui, F. Luk and R. Plemmons, Springer, New York, 1997.

[10] M. Hanke, *Limitations of the L-curve method in ill-posed problems*, BIT 36 (1996), pp. 287–301.

[11] P. C. Hansen, *Analysis of discrete ill-posed problems by means of the L-curve*, SIAM Review 34 (1992), pp. 561–680.

[12] P. C. Hansen, *Regularization tools: A Matlab package for analysis and solution of discrete ill-posed problems*, Numer. Algor. 6 (1994), pp. 1–35. Software is available in Netlib at `http://www.netlib.org`.

[13] 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. 14 (1993), pp. 1487–1503.

[14] C. L. Lawson and R. J. Hanson, Solving Least Squares Problems, SIAM, Philadelphia, PA, 1995.

[15] V. A. Morozov, *On the solution of functional equations by the method of regularization*, Soviet Math. Dokl. 7 (1966), pp. 414–417.

[16] C. R. Vogel, *Non-convergence of the L-curve regularization parameter selection method*, Report, Department of Mathematics, Montana State University, Bozeman, MT, 1996.