Optimal control as a regularization method for ill-posed problems

Stefan Kindermann, Carmeliza Navasca

Department of Mathematics
University of California
Los Angeles, CA 90095-1555
USA

{kinder,navasca}@math.ucla.edu

Abstract. We describe two regularization techniques based on optimal control for solving two types of ill-posed problems. We include convergence proofs of the regularization method and error estimates. We illustrate our method through problems in signal processing and parameter identification using an efficient Riccati solver. Our numerical results are compared to the same examples solved using Tikhonov regularization.

Keywords: Riccati equation, ill-posed problems, regularization

1. Introduction

In this work we consider control-theoretic methods for the construction of regularization operators for ill-posed problems. As a main tool we make use of the well-known Riccati equation and efficient numerical methods for solving it.

We consider two types of ill-posed problem. At first we construct regularization operators for abstract ill-posed problems in Hilbert spaces. We reformulate these problems in the spirit of the dynamic regularization approach [8]. However, we take a different technique to regularize them by casting it as an infinite-horizon optimal control problem.

For the second type of ill-posed problem, we consider a specific parameter estimation problem for the heat equation. This particularly problem includes a naturally defined evolution eqution, namely, the heat equation. Our approach takes advantage of this system and we propose a dynamic regularization which propagates in time.

The common theme between these two methods is to formulate the problem as a quadratic minimization problem with a linear dynamics constraint. The solutions for both problems require solving the Riccati equation. In addition, we develop numerical regularization algorithms based on an efficient Riccati solver [10].

We consider the following ill-posed problems:

Problem 1: Approximate the solution u of

$$Fu = y \tag{1}$$

where F is a bounded linear operator between Hilbert spaces U to X, representing an ill-posed problem.

Problem 2: Reconstruct a solution to the heat equation from overdetermined boundary data.

A simple example of problem 1 is an integral equation of the first kind

$$Fu = \int_0^1 k(s-t)u(t)dt. \tag{2}$$

where F is a blurring operator with a convolution kernel. We want to reconstruct u from possibly noisy data y. Because we assume ill-posedness of the problem the inverse of F does not exist or is not continuous. This is the typical case for equation (2) and for many inverse problems.

A standard treatment for ill-posedness is to look at a generalized solution defined by a solution to the least square problem

$$u = \arg\min_{u \in U} \|Fu - y\|^2. \tag{3}$$

With the least square formulation above, a minimizer in (3) solves the normal equations

$$F^*Fu = F^*y.$$

This problem does not have a solution for all y nor is a solution unique if it exists at all. But for a dense set of data y a solution exists, and we can define a unique one by a so called minimal-norm solution:

$$u^\dagger = \arg\min_{u \text{ sol.to(3)}} \|u\|.$$

The mapping $F^\dagger : y \to u^\dagger$ defines a generalized inverse, the Moore-Penrose inverse $F^\dagger$ [5]. However, for ill-posed problems, such as equation (2), the Moore-Penrose inverse, is not continuous. To obtain a feasible numerical algorithm it has to be approximated by a sequence of stable operators, the regularization operators. Regularization is the approximation of an ill-posed problem by a family of well-posed problems. In 1963, Tikhonov introduced a stable method for numerically computing solutions to inverse problems. He proposed to minimize the Tikhonov functional

$$\|Fu - y\|^2 + \alpha\|u\|^2 \tag{4}$$

for some $\alpha > 0$. Thus, the solution u of Problem 1 is approximated by a family of solutions,

$$u_\alpha = (F^*F + \alpha I)^{-1}F^*y.$$

Another choice of functionals is to minimize

$$\|Fu - y\|^2 + \alpha\|\nabla u\|_{L^2}^2$$

for some $\alpha > 0$.

It can be shown that under reasonable conditions u_α converges to $u^\dagger$ in the limit $\alpha \to 0$. This technique can be generalized by choosing other norms or seminorms as penalty functional. One important choice in image processing is the bounded variation seminorm, which leads to the Rudin-Osher-Fatemi method [12].

Our approach is based on dynamic regularization; we approximate the solution $u^\dagger$ by a dynamic process which converges to $u^\dagger$ in the limit as time goes to infinity. The dynamic process is constructed by minimizing the following functional which has some resemblance to (4).

$$\int_0^\infty \|Fu(t) - y\|^2 + \|u'(t)\|^2 \, dt.$$

In Section 2, we formulate this problem as a linear quadratic control problem where a regularization parameter is introduced through the dynamics of the problem. By control theory, a solution is constructed from an algebraic Riccati equation. In Section 3 we investigate the regularization properties of this procedure and prove convergence and convergence rates of this method.

In Problem 2, we consider a Cauchy problem for the heat equation:

$$w_t = \Delta w \quad \text{on } \Omega \times [0,\infty),$$

subject to the initial and boundary conditions

$$\begin{aligned} w(x,0) &= w_0(x) \quad x \in \Omega \\ w(x,t) &= 0 \quad \text{on } \partial\Omega \setminus \Gamma_2 \times [0,\infty) \\ \frac{\partial w(x,t)}{\partial n} &= v(t) \quad \text{on } \Gamma_2 \times [0,\infty) \end{aligned}$$

We want to reconstruct the heat flux $v(t)$ on Γ_2 from the measurements of the heat flux on another part of the boundary Γ_1, where $x \in \Omega$, $\Gamma_1, \Gamma_2 \subset \partial\Omega$ and $\Gamma_1 \cap \Gamma_2 = \emptyset$. This particular problem can be written as an inverse problem where F is the operator

$$\begin{aligned} F : H^{-1/2}(\Gamma_2) \times [0,\infty) &\rightarrow H^{-1/2}(\Gamma_1) \\ v &\rightarrow \frac{\partial w(x,t)}{\partial n} \end{aligned} \tag{5}$$

and thus, can be solved similarly as for problem 1. However, such a formulation neglects the inherent time-structure of the heat equation. Treating the unknown solution v as control variable and denoting $y(t) := Fv$ as the given data allows to apply the optimal control paradigm. In Section 4 we use again control-theoretic methods to find a regularization for this problem. Finally we outline the numerical aspects of our algorithm in Section 5 and give some numerical results for the regularization method for Problem 1 and Problem 2 in Section 6.

2. Regularization Problem I

Let X and Y be Hilbert spaces. Consider the inverse problem 1 of recovering u from the data y where $F : X \rightarrow Y$ is a linear bounded operator.

In [8] this problem was treated by dynamic regularization, i.e. a regularized solution was constructed by introducing an artificial time-variable $u(t)$ and minimizing the functional

$$\int_0^T \|Fu(t) - y\|^2 + \|u'(t)\|^2 \, dt.$$

Although the original problem is stationary, the new time-dependent problem is convenient as it allows a formulation of a quadratic optimization problem with a linear dynamics. Regularization is introduced by penalizing the time derivative of u, contrary to the usual Tikhonov regularization, where a spacial norm (such as $L^2(\Omega)$) or derivatives on the spacial variable is penalized. It was also shown in [8] that the minimizer $u(T)$ converges to the minimal norm solution $u^\dagger$ for noise-free data as $T \rightarrow \infty$.

In our new approach we consider a similar functional but with $T = \infty$; i.e. we consider an infinite-time horizon problem

$$J(u) := \frac{1}{2}\int_0^\infty \|Fu(t) - y\|^2 + \|u'(t)\|^2 dt. \tag{6}$$

As in the finite-time case this functional leads to a linear quadratic control problem. By introducing the new variables $\epsilon(t) := Fu(t) - y$ and $v(t) := u'(t)$, the minimization over (6) is now equivalent to the following problem

$$\min_{v} \frac{1}{2} \int_0^\infty \|\epsilon\|^2 + \|v\|^2 \, dt \tag{7}$$

subject to the dynamics

$$\begin{aligned} \dot{\epsilon}(t) &= Fv \\ \epsilon(0) &= Fu_0 - y =: \epsilon_0. \end{aligned} \tag{8}$$

We know that a minimizer at infinity, $u(\infty) = F^\dagger y$, satisfying the problem above is equivalent to computing the Moore-Penrose inverse $F^\dagger$, which is unbounded for ill-posed problem. Then the problem (7-8) provides a method for computing a discontinuous operator which cannot be done in a stable way. Therefore, this problem is not stable.

In order to deal with the instability we have to replace (7-8) by an approximate stable solution. Keeping T finite is one possibility, but we propose another regularization by replacing the dynamics (8) by

$$\epsilon'_\alpha(t) = A_\alpha \epsilon_\alpha + F v_\alpha$$

where A_α is bounded linear operator depending on a regularization parameter α such that $\lim_{\alpha \to 0} A_\alpha = 0$.

Then the new regularized problem is the following:

$$\min_{v_\alpha} \frac{1}{2} \int_0^\infty \|\epsilon_\alpha\|^2 + \|v_\alpha\|^2 \, dt \tag{9}$$

subject to the dynamics

$$\begin{aligned} \dot{\epsilon}_\alpha(t) &= A_\alpha \epsilon_\alpha + F v_\alpha \\ \epsilon_\alpha(0) &= Fu_0 - y = \epsilon_0. \end{aligned} \tag{10}$$

The solution of the regularized optimal control problem on an infinite-time horizon is closely related to stationary solution of the following operator algebraic Riccati equation [15, 4],

$$A_\alpha^* P + P A_\alpha - P F F^* P + I = 0.$$

where P is the unknown. We give some basic definitions and theorems which ensure existence (and uniqueness) of the operator P (see [4],[15]).

Definition 2.1. *The pair (A_α, F) is said to be exponentially stabilizable if there exists a linear, bounded K such that $A_\alpha + FK$ generates an exponentially stable semigroup; i.e. if in the (10), the control is $v_\alpha = K\epsilon_\alpha(t)$, then*

$$\dot{\epsilon}_\alpha(t) = (A_\alpha + FK)\epsilon_\alpha$$

and

$$\epsilon_\alpha(t) = S(t)\epsilon_0 \longrightarrow 0 \text{ as } t \longrightarrow \infty$$

where $S(t)$ is semigroup for $t \geq 0$.

Definition 2.2. *The pair (A_α, I) is said to be exponentially detectable if there exists a linear, bounded L such that $A_\alpha + L$ generates an exponential stable semigroup.*

Theorem 2.3. *If the pair (A_α, F) is exponentially stabilizable, then the Riccati equation has at least one linear, bounded, and nonnegative solution P.*

Theorem 2.4. *If the pair (A_α, F) is exponentially stabilizable and (A_α, I) is exponentially detectable, then the Riccati equation has a unique linear, bounded, and nonnegative P and $A_\alpha + FK$ generates an exponential stable semigroup where $K = -F^*P$.*

It follows that $A_\alpha + FK$ is a bounded linear operator and hence, the exponentially stable semigroup is $e^{(A_\alpha+FK)t}$; i.e.

$$\epsilon_\alpha(t) = e^{(A_\alpha+FK)t}\epsilon_\alpha(0) \to 0$$

as $t \to \infty$.

Moreover, since $u'_\alpha(t) = v_\alpha(t)$, then the integrated optimal control is the following:

$$\begin{aligned} u^*_\alpha &= u_0 + \int_0^\infty v_\alpha(s)\ ds \\ &= u_0 + \int_0^\infty K\epsilon_\alpha(s)\ ds \\ &= u_0 + K\int_0^\infty e^{(A_\alpha+FK)s}\epsilon_\alpha(0)\ ds. \end{aligned}$$

For simplicity we choose $u_0 = 0$. Then it follows

$$u^*_\alpha = R_\alpha y$$

with the regularization operator

$$R_\alpha = F^*P(A_\alpha - FF^*P)^{-1}. \tag{11}$$

In the next section we will show that this is indeed a regularization, and that $u^*_\alpha \to u^\dagger$ as $\alpha \to 0$ for the choice $A_\alpha = -\alpha I$.

Furthermore, the ill-posedness of problem (8) can be further discussed using the following control-theoretic arguments.

Theorem 2.5. *The linear system (8) is not exponentially stabilizable; i.e. there does not exists a linear, bounded K such that FK generates an exponentially stable semigroup.*

Proof: Suppose there exists a linear, bounded K such that FK generates an exponentially stable semigroup. Then, the pair $(0, F)$ is exponentially stabilizable. By Theorem 2.3, there exists linear, bounded, and nonnegative P satisfying

$$PFF^*P = I,$$

the Riccati equation corresponding to the linear-quadratic regulator problem (7-8). Observe that $P = (FF^*)^{-\frac{1}{2}}$. However, for ill-posed problems 0 is in $\sigma(FF^*)$, the spectrum of FF^*, thus the eigenvalues of $(FF^*)^{-1}$ are unbounded. Thus, P is an unbounded operator. Therefore, we get a contradiction. ∎

3. Regularization Properties

In this section we express the operator $R_\alpha : y \to u_\alpha$ through the spectral filter function g_α. In the spirit of [5], we prove convergence as well as calculate the rates of convergence.

3.1. **Spectral Filter Function.** Let

$$A = -\alpha I$$

in the operator Riccati equation where $\alpha > 0$ is the regularization parameter and I is the identity operator. Then P is the positive solution of

$$-2\alpha P - PFF^*P + I = 0. \tag{12}$$

Denote by F_λ the spectral family associated with FF^*. From (12) it follows that P has a representation via the spectral measure

$$P = \int_0^{\|F\|^2} p(\lambda) dF_\lambda.$$

with an appropriate function $p(\lambda)$. Recall that the regularization operator is defined as $R_\alpha = F^*P(A_\alpha - FF^*P)^{-1}$, which yields

$$R_\alpha = F^* g_\alpha(FF^*)$$

where

$$g_\alpha(FF^*) = \int_0^{\|F\|^2} \frac{p(\lambda)}{\alpha + \lambda p(\lambda)} dF_\lambda. \tag{13}$$

From the operator Riccati equation (12), we have the condition

$$0 = \int_0^{\|F\|^2} 1 - 2\alpha p(\lambda) - \lambda p(\lambda)^2 dF_\lambda.$$

By simple algebra this can be solved to

$$p_\pm(\lambda) = \frac{1}{\lambda}\left(-\alpha \mp \sqrt{\alpha^2 + \lambda}\right).$$

In control theory only the positive P is meaningful and computed, hence we set

$$p(\lambda) = \frac{1}{\lambda}\left(-\alpha + \sqrt{\alpha^2 + \lambda}\right).$$

The spectral filter function (13) becomes

$$g_\alpha(FF^*) = \int_0^{\|F\|^2} \frac{-\alpha + \sqrt{\alpha^2 + \lambda}}{\lambda\sqrt{\alpha^2 + \lambda}} dF_\lambda.$$

The following identity, which can easily be proven by the Weierstrass aproximation theorem, is well-known in spectral theory:

$$F^* g_\alpha(FF^*) = g_\alpha(F^*F)F^*.$$

Finally this leads to the representation of the regularization operator

$$R_\alpha = g_\alpha(F^*F)F. \tag{14}$$

3.2. **Estimates and Convergence.** We denote the regularized solution from the exact data $y \in R(F)$ ($R(F)$ is the range of F) and the noisy data y_δ as u^*_α and $u^*_{\alpha,\delta}$, respectively.

Recall that

$$g_\alpha(\lambda) = \frac{-\alpha + \sqrt{\alpha^2+\lambda}}{\lambda\sqrt{\alpha^2+\lambda}}.$$

Observe that

$$|\lambda g_\alpha(\lambda)| = \left|\frac{-\alpha + \sqrt{\alpha^2+\lambda}}{\sqrt{\alpha^2+\lambda}.}\right| \le 1. \tag{15}$$

In addition, for $\lambda > 0$ fixed we have

$$\lim_{\alpha\to 0} g_\alpha(\lambda) = \frac{1}{\lambda}. \tag{16}$$

Since g_α is a continuous function on $[0, \|F\|^2]$ satisfying the properties (15-16), Theorem 4.1 in [5], is applicable, which gives the convergence result for noise-free data:

$$\|u^*_\alpha - u^\dagger\| \to 0 \text{ as } \alpha \to 0, \qquad \text{for } y \in R(F). \tag{17}$$

Moreover, Theorem 4.2 in [5] with

$$G_\alpha := \sup_{\lambda\in[0,\|F\|^2]} |g_\alpha(\lambda)| = \frac{1}{2\alpha^2} \tag{18}$$

calculated using L'Hopital's rule leads to

$$\|u^*_\alpha - u^*_{\alpha,\delta}\| \le \delta\sqrt{G_\alpha} \le \frac{\delta}{\sqrt{2}\alpha} \tag{19}$$

whenever $\|y - y_\delta\| \le \delta$ with $y \in R(F)$. The estimate (19) shows the stability of the method when noise is present. Notice that the equation (17) only gives the convergence of the u^*_α to $u^\dagger$. In the following theorems, we find the rates of convergence. Note that convergence rates for regularization operators can only be find if an additional abstract smoothness condition on the exact data, the so called source condition (20) is satisfied [5]. In Theorem 3.1 we state the precise convergence rates results under such a condition:

Theorem 3.1. *Let $u^*_\alpha := R_\alpha y$, with R_α as in* (11). *If the data are exact, i.e. $y = Fu^\dagger$, and the exact solution satisfies a source condition*

$$u^\dagger \in \mathcal{R}((F^*F)^\mu) \tag{20}$$

for $\mu > 0$, then we have the error estimate

$$\|u^*_\alpha - u^\dagger\| \le C\alpha^{2\mu^*} \qquad \textit{for} \qquad \mu^* = \min(\mu, \tfrac{1}{2}).$$

Proof: First, by the spectral representation (14) we have

$$u^\dagger - u^*_\alpha = u^\dagger - g_\alpha(F^*F)F^*y,$$

which implies

$$u^\dagger - g_\alpha(F^*F)F^*y = (I - g_\alpha(F^*F)F^*F)u^\dagger.$$

Define

$$r_\alpha(\lambda) := 1 - \lambda g_\alpha(\lambda).$$

By (20),

$$u^\dagger - u_\alpha^* = r_\alpha(F^*F)u^\dagger = r_\alpha(F^*F)(F^*F)^\mu w$$

for some $w \in X$. It follows that

$$\|u^\dagger - u_\alpha^*\| \leq C\lambda^\mu |r_\alpha(\lambda)|,$$

with $C = \|w\|$. According to [5], Theorem 4.3, it suffices to find a bound $\omega_\mu(\alpha)$, such that

$$\lambda^\mu |r_\alpha(\lambda)| \leq \omega_\mu(\alpha) \qquad \forall \lambda \in [0, \|F\|^2] \tag{21}$$

to establish the rates

$$\|u_\alpha^* - u^\dagger\| \leq \omega_\mu(\alpha).$$

We find that maximum on the left hand side in (21) over $\lambda \in \mathbb{R}^+$ is obtained at $\lambda = \frac{2\alpha^2\mu}{1-2\mu}$ for $\mu \leq \frac{1}{2}$, resulting in

$$\omega_\mu(\alpha) = \begin{cases} \mathcal{O}(\alpha^{2\mu}) & \mu \leq \frac{1}{2} \\ \mathcal{O}(\alpha) & \mu \geq \frac{1}{2} \end{cases}$$

Hence,

$$\|u_\alpha^* - u^\dagger\| \leq C\alpha^{2\mu^*}$$

where $\mu^* = \min(\mu, \frac{1}{2})$ for some $C > 0$. ∎

The previous theorem establishes convergence rates for exact data in terms of the regularization parameter α. For the case of noisy data it is a general result [5] that no convergence holds for $\alpha \to 0$. Instead the regularization parameter has to be coupled to the noise level $\delta = \|y - y_\delta\|$. A function $\alpha(\delta)$ which relates the regularization parameter α to the noise level is named a parameter choice rule. Having an appropriate parameter choice rule the correct notion of convergence (and convergence rates) is to look for estimates $\|u^*_{\alpha(\delta),\delta} - u^\dagger\|$ in terms of δ. Similar as for the noise free case again a source condition is necessary for this. Following the proofs in [5] we can show that a certain parameter choice rule gives convergence rates for noisy data:

Theorem 3.2. *Given the assumptions in Theorem 3.1. If the data are noisy with noise level δ, i.e. $\|y - y_\delta\| \leq \delta$ with $y \in \mathcal{R}(F)$, and the source condition* (20) *is satisfied, then the parameter choice rule $\alpha \sim \delta^{\frac{1}{2\mu^*+1}}$ yields the the total error*

$$\|u^*_{\alpha,\delta} - u^\dagger\| \leq \tilde{C}\delta^{\frac{2\mu^*}{2\mu^*+1}}$$

for $\mu^* = \min(\mu, \frac{1}{2})$.

Proof: From the Theorem 3.1, we obtain

$$\omega_\mu(\alpha) = \mathcal{O}(\alpha^{2\mu}) \tag{22}$$

and from equation (18) we have

$$G_\alpha = \mathcal{O}\left(\frac{1}{\alpha^2}\right). \tag{23}$$

By the Corollary 4.4 in [5] the estimates (22-23) suggest an optimal order parameter choice rule

$$\alpha \sim \delta^{\frac{1}{2\mu^*+1}}$$

for $\mu^* = \min(\mu, \frac{1}{2})$. It follows from Theorem (3.1) that

$$\|u_\alpha^* - u^\dagger\| \leq C\delta^{\frac{2\mu^*}{2\mu^*+1}}. \tag{24}$$

Hence, by the bound on the propagated data error (19) and the estimate (24), we obtain the estimate for the total error

$$\|u_{\alpha,\delta}^* - u^\dagger\| \leq C\delta^{\frac{2\mu^*}{2\mu^*+1}} + \frac{\delta}{\sqrt{2\alpha}} \leq \tilde{C}\delta^{\frac{2\mu^*}{2\mu^*+1}}$$

for some $\tilde{C} > 0$. ∎

The result in Theorem 3.2 gives convergence rates, which are of optimal order for $\mu^* \leq \frac{1}{2}$. It also shows that this optimal order result breaks down for $\mu > \frac{1}{2}$. This phenomenon is called saturation; it is quite common in other regularization methods, such as, Tikhonov regularization [5]. The point of break down is termed the qualification of the method. Thus our case the qualification is $\frac{1}{2}$.

In the next section we focus on problem 2, the heat equation problem with overdetermined boundary conditions:

4. Regularization Problem II

4.1. **Cauchy Problem for the Heat Equation.** Let Ω be a smooth domain, and $\Gamma_1, \Gamma_2 \subset \partial\Omega$ where $\Gamma_1 \cap \Gamma_2 = \emptyset$. Consider the homogeneous heat equation problem

$$w_t = \Delta w \quad \text{on } \Omega \times [0, \infty), \tag{25}$$

subject to the initial and boundary conditions

$$\begin{aligned} w(\cdot, 0) &= w_0(x) && (26)\\ w(\cdot, t) &= 0 \quad \text{on } \partial\Omega \setminus \Gamma_2 \times [0, \infty) && (27)\\ \tfrac{\partial}{\partial n} w(\cdot, t) &= v(t) \quad \text{on } \Gamma_2 \times [0, \infty) && (28)\end{aligned}$$

Suppose Γ_2 is part of the boundary which is not accessible, but we can measure the heat flux on another part Γ_1; i.e., the data $y(t)$ for our problem are

$$y(t) = \frac{\partial}{\partial n} w(\cdot, t) \qquad \text{on} \qquad \Gamma_1 \times [0, \infty) \tag{29}$$

The inverse problem is to find $v(t)$ and $w(t)$ such that (25)-(29) are satisfied.

This problem can be formulated within the operator theoretic framework, consider an operator F which maps

$$v \to \frac{\partial}{\partial n} w(\cdot, t)$$

from $H^{-1/2}(\Gamma_2) \times [0, \infty)$ to $H^{-1/2}(\Gamma_1) \times [0, \infty)$ and w satisfies (25-28). This leads to the ill-posed operator equation

$$Fv = y.$$

Now we could use the previous techniques to tackle this equation by introducing an artificial time variable and a dynamics. However there is no need for this, as the problem already has a naturally defined dynamics by the heat equation. Note that for Problem 1 we used a linear quadratic control problem. This involved a pair of state and control variable $(\epsilon(t), v(t))$, a linear evolution equation and a quadratic functional for them. We use a similar idea here by considering $w(t)$ as state variable and $z = (v(t), y(t))$ as control variable. The dynamics of the system is then governed by the equation in weak form

$$(w_t, \phi) = -(\nabla w, \nabla \phi) + \int_{\Gamma_1} y(t)\phi ds + \int_{\Gamma_2} v(t)\phi ds \quad \forall \phi \in Z, \tag{30}$$

where $Z = \{\phi \in H^1(\Omega) \,|\, \phi = 0 \text{ on } \Omega \setminus (\Gamma_1 \cup \Gamma_2)\}$. This equation has the similar structure as (8) where w corresponds to e_α and the pair $z = (y(t), v(t))$ corresponds to v_α. What remains to be found is a functional to minimize. The difference to the ideas in Section 2 is that the dynamics is already stable, so there is no need to include a regularization. Instead a penalty term should be included into the functional. The function $w(t)$ is completely determined by the evolution equation (30) and the initial conditions (26). The additional condition that we impose on a solution to the problem is (27). Then, it follows that w is a minimizer of the functional

$$J(u) := \frac{1}{2} \int_0^\infty \int_{\Gamma_1} w^2 dx \; dt.$$

As mentioned above we add a penalty term involving the control variable $z = (v, y)$ as a regularization, then we arrive at the new functional

$$\begin{aligned} J_\alpha(u) &:= \frac{1}{2} \int_0^\infty \left[\int_{\Gamma_1} w^2 dx + \int_{\Gamma_2} \langle z, Rz \rangle dx \right] \; dt \\ &= \frac{1}{2} \int_0^\infty \langle u, u \rangle_{L^2(\Gamma 1)} + \langle z, Rz \rangle_{L^2(\Gamma 2)} \; dt. \end{aligned}$$

where

$$R = \begin{bmatrix} I & 0 \\ 0 & \alpha I \end{bmatrix}.$$

Given the functional J_α in (31) and the heat equation (30) we have the ingredients for formulating the problem as an optimal control problem in the infinite dimensional space where w is the state variable and $z = (v, y)$ is the control variable. Summing up, we now consider the linear quadratic regulator problem of minimizing a cost functional

$$J(u) = \int_0^\infty \langle u, u \rangle_{L^2(\Gamma 1)} + \alpha \langle z, Rz \rangle_{L^2(\Gamma 2)} \; dt \tag{31}$$

subject to

$$\begin{aligned} w_t &= \Delta w \\ \frac{\partial}{\partial n} w(., t) &= y(t) \text{ on } \Gamma_1 \times [0, \infty) \\ \frac{\partial}{\partial n} w(., t) &= v(t) \text{ on } \Gamma_2 \times [0, \infty) \qquad (32) \\ w &= 0 \quad \text{ on } \partial\Omega \setminus (\Gamma_2 \cup \Gamma_1) \times [0, \infty). \end{aligned}$$

This has the structure we want and as in Section (2), we can define the feedback operator K

$$z = (v(t), y(t)) = Kw(t).$$

Note that K is independent of the time-variable t, as only to be calculated once by solving the corresponding Riccati equation. Hence if $w(t)$ is known we can eliminate the control from (30) and solve the resulting equation. This gives an evolution equation which can be solved progressing in time.

After a spacial discretization the optimization problem can be written as

$$J_{n,\alpha} = \int_0^\infty (w, Cw) + (z, \tilde{R}z)dt$$

subject to the constraint

$$w_t = Aw + B_1 y(t) + B_2 v(t). \tag{33}$$

Here, C denotes a matrix corresponding to the discrete trace operator $T : w \to w|_{\Gamma_1}$, B_1, B_2 correspond to the trace operator, which maps w to the normal derivatives on Γ_1 and Γ_2, $\tilde{R}$ is the trace operator which maps $z = (y, v) \to z|_{\Gamma_2}$. Finally, A is the discretization of the Laplacian. We denote by the matrix $B = [B_1\ B_2]$. Then the feedback operator is obtained by solving the Riccati equation,

$$A'P + PA + K'\tilde{R}^{-1}K + C = 0 \tag{34}$$

for P where feedback matrix

$$K = -B'P.$$

This yields

$$(y, v) = Kw(t) = (K_1 w(t), K_2 w(t)) \tag{35}$$

However, since $y(t)$ is given we only need the values for $v(t)$, leading to the evolution equation

$$w_t = Aw + B_1 y(t) + B_2 K_2 w. \tag{36}$$

The unknown control $v(t)$ is now removed from the system and the equation can be solved for w in a usual way. This equation can be discretized in time by an appropriate explicit or implicit scheme. The regularized solution for the heat flux on Γ_2 can be found by applying K to the computed solution $w(t)$ at each time step via (35).

In comparison to an operator-theoretic and least-squares approach the heat flux v can be found by progressing in time. That means that $v(t)$ can be computed independent of the values of $w(t')$ for $t' > t$. This is in contrast to an usual iterative regularization scheme using (6) (e.g. Landweber iteration or Conjugate Gradient Method), where in each iteration step an evolution equation has to be solved.

In the next sections we look at the numerical approximations and results of our method.

5. Numerical Method For The Riccati Equation

Both ill-posed problems considered in the previous sections require numerical solution of the algebraic Riccati equation. The computational techniques of finite element and quadrature rule were used to approximate the functionals in optimization problems (9-10) and (31-32). The partial differential equations (10) and (32) were discretized using finite difference schemes. These numerical approximations reduce the optimization problems to linear-quadratic regulator problems in finite dimensional space.

The linear-quadratic regulator problem is

$$V(x) = \min_u \frac{1}{2} \int_0^\infty x'Qx + u'Rudt$$

subject to the dynamics

$$\begin{aligned} \dot{x} &= Ax + Fu \\ x(0) &= x_0 \end{aligned}$$

where the state variable $x \in \mathbb{R}^n$ and the control $u \in \mathbb{R}^m$. We assume Q is symmetric and nonnegative matrix and R is symmetric and positive definite matrix. It is well known that the solution we seek has the following form: the minimum cost is a quadratic function

$$V(x) = x'Px \tag{37}$$

and the optimal control feedback is a linear function

$$u = Kx \tag{38}$$

where K is the control law.

Through the dynamic programming technique [1], the solutions (37) and (38) of the regulator problem are realized through the algebraic Riccati equation,

$$A'P + PA - K'RK + I = 0 \tag{39}$$

and

$$K = -R^{-1}F'P \tag{40}$$

Then the Riccati equation (39) can be recasted as a Lyapunov equation

$$(A' + K'F')P + P(A + FK) = D'D \tag{41}$$

where $D' = [PF\ Q^{1/2}]$. If the spectrum of $A + FK$ lies in the left-half complex plane, $\mathcal{C}_-$ and $D'D$ is positive definite, then a positive definite solution P exists for the Lyapunov equation (41).

From the equations (41) and (40), the matrices P and K can be solved simultaneously. The well known iterative technique is the Newton-Kleinman method [7]. Other variants of the Newton-Kleinman can be found in [2, 13, 10, 11]. We use an efficient algorithm [10] to approximate the matrix P solution of the Riccati equation of high order. Large Riccati equation often arises from approximations of K and P from partial differential equations. The inverse Problem 1 typically has a large data size n resulting in a high order Riccati equation in search for the unknown matrix P whose dimension is $n \times n$. I n practice, the number of unknown states of u is set to the number of data n. Current methods based on eigenvalue estimation fail to give good approximation in high dimension because high frequency modes are not included. The Riccati solver in [10, 11] is based on the Newton-Kleinman iteration [7] with the Cholesky-ADI scheme Lyapunov solver [14, 9].

The matrix solution P is approximated iteratively using the following scheme:

$$A_i'P_i + P_iA_i = D_i'D_i,$$

where

$$A_i = (A + FK_i), \tag{42}$$

$$K_i = -R^{-1}F'P_{i-1}, \tag{43}$$

and

$$D_i = [P_{i-1}F\ Q^{1/2}], \text{ for } i = 0, 1, \ldots \tag{44}$$

Note that the subscript i denotes the iterates. The initial guess K_0 is selected such that $\sigma(A - FK_0) \subset \mathcal{C}_-$.

It has been shown in [7] that the iterative scheme above converges quadratically where the sequence of iterates $\{P_i\}_{i=0}^{\infty}$ is monotonically decreasing, i.e.

$$P \leq P_{i+1} \leq P_i \leq \dots, \quad \text{for } i = 0, 1, \dots$$

with

$$\lim_{i \to \infty} P_i = P.$$

Another scheme can be found in [11] where it bypasses the calculation of P and directly approximates K. For most control problems, only the feedback gain K is necessary. By directly calculating the gain matrix K, the method describes in [11] gives considerable amount of savings in computation and memory.

The complexity of the Lyapunov solver for dense A and sparse A are $\mathcal{O}(J_L n_d n^2)$ and $\mathcal{O}(J_L n_d n)$, respectively, where J_L is the number of Lyapunov iterations, n_d is the number of columns in D, and n is the size of A. The Riccati solver requires two matrix multiplications per iteration in (42-44). Generally, the number of Riccati iterations $J_R << n$. Thus, the Cholesky-ADI scheme [14, 9] has a complexity that is $\mathcal{O}(n^2)$ for full A and $\mathcal{O}(n)$ for sparse A when $J_L, J_R, n_d << n$.

6. Numerical Results

6.1. **Deblurring Problem.** We first consider the deblurring problem. This problems falls into the category of ill-posed equations of problem 1. We take an integral operator with the mildly smoothing kernel

$$k(x,y) = \begin{cases} 1 - \frac{(x-y)^2}{0.1}, & |x-y|^2 \leq 0.1 \\ 0, & \text{elsewhere} \end{cases}$$

as the blurring operator. The forward operator F is defined as the integral operator associated with this kernel. To obtain the discretized F, we use the Galerkin approximation with piecewise constant basis; i.e. each entries of the matrix is

$$F_{i,j} = (\phi_i, F\phi_j) \sim \frac{1}{n^2} k(x_i, x_j),$$

where

$$\phi_i(x) = \begin{cases} 1, & \frac{i-1}{n} \leq x \leq \frac{i}{n} \\ 0, & \text{elsewhere} \end{cases}$$

Then we derive the associated Riccati equation with this discretized problem. Calculating P, the solution to the Riccati equation, with the solver described in the previous section, allows us to approximate the regularization operator,

$$R_\alpha = -F^* P (A_\alpha - FF^* P)^{-1}. \tag{45}$$

Our method is quite robust in sense that the computed R_α is independent of the data y. This is of course an advantage to the other iterative methods which are data specific. If only one solution has to be calculated, then the iterative methods are in general more efficient. We compare our results with Tikhonov regularization because it is also a multiple data solver. The equation (45) reminds us closely of Tikhonov regularization, which is defined as

$$R_T = (F^* F + \alpha I)^{-1} F^*. \tag{46}$$

Observe that if $P = I$ and $A_\alpha = -\alpha I$, then the regularization operator (45) is exactly the Tikhonov regularization. Since we use an efficient Riccati solver with complexity

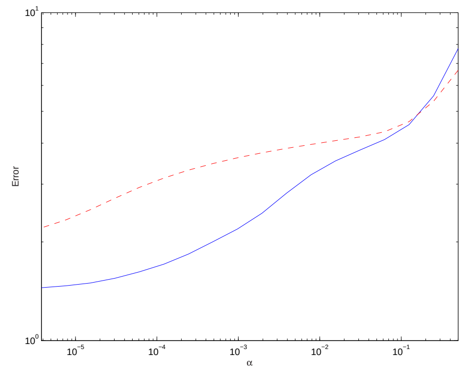

FIGURE 1. Error vs. α

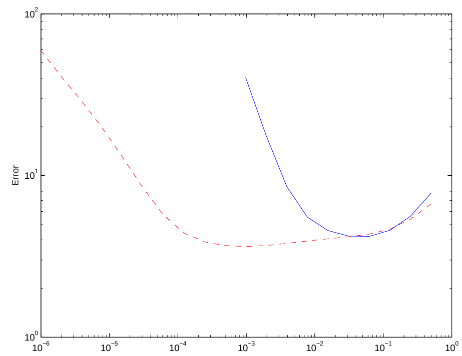

FIGURE 2. Error vs. α (noisy data)

$\mathcal{O}(n^2)$ flops our method is dominated mainly by th$\mathcal{O}(n^3)$ operations in the matrix-inversion. Thus, our method has the same complexity as of Tikhonov's method where matrix-inversion is performed.

We consider the exact solution

$$u^\dagger = 2\chi_{[0.3,0.8]}$$

as a test example where χ denotes the characteristic function.

From Theorem 3.1 the error $\|u_\alpha - u^\dagger\|$ behaves as $\alpha^{2\mu}$. This error is better than the error of Tikhonov regularization where the order is α^μ ([5]). Hence we may choose a large regularization parameter in (46) which implies that the inversion in (45) is more stable than the inversion in (46).

Figures 1 and 2 illustrate these results. It shows the error versus the regularization parameter α for our approach (blue –) and for Tikhonov regularization (red - -) in the cases of the noise-free data and the noisy data on a log-log scale. In both pictures we see that the error decreases faster for our method. In figure 2, the optimal parameter α in the noisy case is larger than the parameter for Tikhonov regularization, which is in agreement with Theorem 3.1, because an optimal parmeter-choice rule for (46) is $\alpha_{TR} \sim \delta^{\frac{2}{2\mu+1}}$ whereas in Theorem 3.1 the parameter-choice rule is $\alpha \sim \delta^{\frac{1}{2\mu+1}}$ for $\mu \in [0, \frac{1}{2}]$. For Tikhonov regularization, a large parameter-choice α means inverting a more stable matrix $F^*F + \alpha I$. In summary, we can say that the operator P acts as a preconditioner for problem 1 with the effect that a more stable operator has to be inverted than the matrix inverted for Tikhonov regularization.

6.2. **Heat Equation.** We now turn to the Cauchy problem for the heat equation. We consider the problem on the unit square $\Omega = [0,1]^2$, where Γ_1 is the right boundary $\Gamma_1 = \{0\} \times [0,1]$, and Γ_2 is the left boundary $\Gamma_2 = \{1\} \times [0,1]$. The heat equation is discretized on a uniform $n \times n$-grid by finite differences. This gives a problem similar to (33) where all the matrices A, B, C are sparse.

To obtain the feedback operator K so that we can solve equation (36), we first solve the Riccati equation. This equation is discretized in time by an implicit scheme:

$$w_{m+1} = w_m + \Delta t(Aw_{m+1} + B_1 y(t_m) + B_2 K_2 w_{m+1}), \qquad t_m = \frac{m}{\Delta t}.$$

The solution to the inverse problem is then obtained by

$$v(t_m) = Kw_m.$$

The first example we want to identify an oscillating function in time and space:

$$v_1(t,x) = \sin(20t)\sin(2\pi x)$$

The second test example is a moving "hot spot:"

$$v_2(t,x) = \exp(-100(x - 0.1\sin(40t) - 0.5)^2)$$

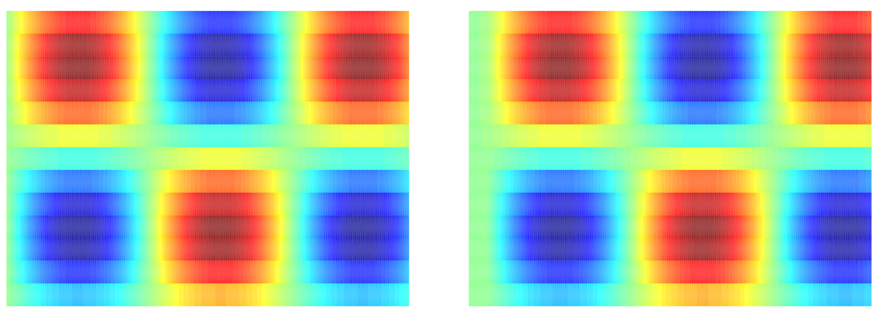

FIGURE 3. exact and computed solutions for v_1

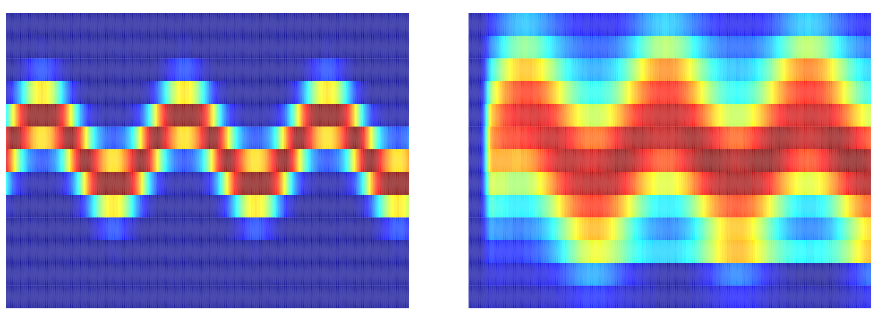

FIGURE 4. exact and computed solutions for v_2

Figure 3 shows the exact and the computed solution for the first example v_1 on the time interval $[0, 0.5]$. Figure 4 shows the result for the second example v_2.

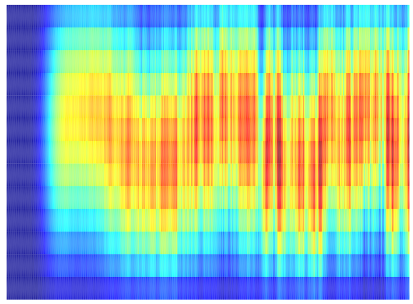

FIGURE 5. solution for example 2 with noise

Finally figure 5 shows the result in the case with noisy data. The data are perturbed with 5% random noise. The result shows the oscillating pattern of the hot spot has noisy components in the time direction. The oscillations are expected from the regularization procedure because this method only penalizes the L^2-norm of the control v; the values of v at different times are hardly correlated. An improvement of this would be achieved by using a norm which penalizes the time-derivate of v.

Acknowledgements

The research is supported in part by ONR MURI N00014-02-1-0720. C.N. is partially supported by NSF grant DMS-9983726. S.K. is supported by NSF grant DMI-0327077.

References

[1] B. D. O. Anderson and J. B. Moore, "Optimal Control, Linear Quadratic Methods," Prentice Hall, Englewood Cliffs, NJ, 1990.

[2] H. Banks and K. Ito, "A Numerical Algorithm for Optimal Feedback Gains in High Dimensional Linear Quadratic Regulator Problems," SIAM J. Control and Optim., 29 (3), (1991), 499-515.

[3] H. Banks and K. Kunisch, "The linear regulator problem for parabolic systems," SIAM J. Control and Optim., 22, (1984), 684-696.

[4] R.R. Curtain and H.J. Zwart "An Introduction to Infinite-Dimensional Linear Systems Theory," Springer-Verlag, New York (1995).

[5] H.W. Engl, M. Hanke, A. Neubauer, "Regularization of Inverse Problems," Kluwer Academic Publishers, Dordrecht (1996).

[6] J. S. Gibson, "The Riccati integral equations for optimal control problems on Hilbert spaces," SIAM J. Control and Optim., 17, (1979), 537-565.

[7] D. Kleinman, "On an Iterative Technique for Riccati Equation Computations," IEEE Transactions on Automat. Control, 13, (1968), 114-115.

[8] S. Kindermann, A. Leitao, "On regularization methods based on dynamic programming techniques," SFB-Report 2003-13, University Linz, (2003).

[9] J-R. Li and J. White, "Low Rank Solution of Lyapunov Equations," SIAM J. Matrix Anal. Appl., 24 (1), (2002), 260-280, 2002.

[10] K. Morris and C. Navasca, "Solution of Algebraic Riccati Equations Arising in Control of Partial Differential Equations," In J. P. Zolessio and J. Cagnol, eds., Control and Boundary Analysis, Marcel Dekker, New York, (2004).

[11] K. Morris and C. Navasca, "Iterative Solution of Algebraic Riccati Equations using a Modified Newton-Kleinman Method," Proceedings of 16th Mathematical Theory of Networks and Systems, Belgium, (2004).

[12] L. I. Rudin, S. Osher, and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Phys. D, 60, (1992), 259-268.

[13] I. G. Rosen and C. Wang, "A Multi-level technique for the approximate solution of operator Lyapunov and Riccat equations," SIAM J. Numer. Anal., 32 (1994), 514-541.

[14] T. Penzl, "A Cyclic Low-Rank Smith Method for Large Sparse Lyapunov Equations," SIAM J. Sci. Comput., 21 (4), (2000), 1401-1418.

[15] J. Zabczyk, "Mathematical Control Theory: An Introduction," Birkhauser, Boston (1992).