



**Dipartimento di Matematica**

A. FAVINI, L. PANDOLFI

**MULTISCALE LAVRENTEV METHOD FOR SYSTEMS  
OF VOLTERRA EQUATIONS OF THE FIRST KIND**

Rapporto interno N. 12, aprile 2007

**Politecnico di Torino**

Corso Duca degli Abruzzi, 24-10129 Torino-Italia

# Multiscale Lavrentev Method for Systems of Volterra Equations of the first kind\*

A. Favini<sup>†</sup>      L. Pandolfi,<sup>‡</sup>

## Abstract

We study the singular perturbation approach proposed by Lavrentev for the regularization of systems of Volterra integral equations of the first kind, in the case that the kernel  $K(t)$  is not invertible for  $t = 0$  and without assuming  $K(t) \sim t^\nu I$ . We single out a class of kernels, which we call “diagonally dominant”. We show that when the kernel belongs to this class then it is possible to regularize the problem using a multiscale singular perturbation method.

## 1 Introduction

We informally describe the problem to be studied in this paper. Precise assumptions will then be introduced in Section 1.2. We consider a linear causal time invariant input-output system described by the Volterra integral equation

$$y(t) = \int_0^t K(t-s)u(s) \, ds, \quad t \in [0, T] \quad (1)$$

where  $y$  and  $u$  are  $n$ -vectors and  $K(t)$  is an  $n \times n$  matrix. Our goal is the reconstruction of  $u$  on the basis of measures taken on  $y$ . It is well known that this problem is ill posed so that regularization algorithms must be used for its solution. Moreover, we want a *causal* reconstruction algorithm, i.e. we want that an estimate  $\mathbf{v}$  of  $u$  be constructed in such a way that its value  $\mathbf{v}(t)$  at time  $t$  depends on the measures taken on  $y$  at previous instants. Lavrentev method and its variants have been widely used to achieve this

---

\*This paper fits the research programs of GNAMPA-INDAM

<sup>†</sup>University of Bologna, Department of Mathematics, Piazza di Porta San Donato 5, Bologna, Italy [favini@dm.unibo.it](mailto:favini@dm.unibo.it)

<sup>‡</sup>Politecnico di Torino, Dipartimento di Matematica, Corso Duca degli Abruzzi 24, 10129 Torino, Italy [luca.pandolfi@polito.it](mailto:luca.pandolfi@polito.it)

goal, see the references given below. The idea behind Lavrentev method is to study the following singular perturbation problem

$$\epsilon \mathbf{v}(t) + \int_0^t K(t-s) \mathbf{v}(s) \, ds = y(t), \quad t \in [0, T].$$

This problem is now well posed for every  $\epsilon$ . Let  $\mathbf{v}(t) = \mathbf{v}_\epsilon(t)$  (dependence on  $\epsilon$  will not be explicitly indicated below) be its solution. Our goal is to give conditions under which

$$\lim_{\epsilon \rightarrow 0^+} \mathbf{v}_\epsilon(t) = \lim_{\epsilon \rightarrow 0^+} \mathbf{v}(t) = u(t).$$

The limit must be taken in an appropriate sense. We shall be interested in pointwise, uniform or  $L^2$  convergence.

Note that in practice the measure taken on  $y$  are always affected by errors. Here we are ignoring this fact, which will be studied in Section 2.1.

The standard assumption on  $K(t)$  in order to achieve (1) is that  $\det K(0) \neq 0$  so that it is possible to assume (after a coordinate transformation)  $K(0) = I$ . If instead  $K(0) = t^n K_0(t)$  and  $K_0(0) = I$  then a suitable number of derivatives can in principle be used in order to reduce the problem to the case  $\det K(0) \neq 0$ , although in practice this need not be a very efficient method, due to the noise in the measures.

In this paper we study the case  $\det K(0) = 0$ , but when the order of the zero is not the same for every entry of  $K(t)$ . For reasons which will appear below, we shall call “diagonally dominant” the class of systems which fit the assumptions of our paper.

The idea is to use a multiscale singular perturbation problem, fitted to the structure of  $K(t)$  as  $t \rightarrow 0^+$ .

**Remark 1** The proofs in this paper make use of smoothness assumptions on the kernel  $K(t)$  and piecewise smoothness of the unknown input  $u(t)$ . However, numerical computation of the derivatives is not required by the algorithm we are going to present. ■

Before we go on to describe the technical assumptions of this paper, we present an example to which the results can be applied.

## 1.1 An example

We consider two interacting electrical lines of infinite length, described by

$$\begin{cases} \eta_{tt} &= \eta_{ss} + \rho \xi + u(t), \\ \xi_{tt} &= \xi_{ss} + \psi(s) v(t). \end{cases} \quad (2)$$

Initial conditions are  $\eta(0, s) = 0$ ,  $\xi(0, s) = 0$ .  
Here  $\eta$  and  $\xi$  are potentials and the output is

$$y(t) = \begin{bmatrix} \eta(t, 0) \\ \xi_t(t, 0) \end{bmatrix}.$$

Note that the model is not so realistic since when the second line acts on the first one, there should be a reaction which is not taken into account by the model. Moreover, for simplicity we assume that  $\rho$  is constant and the coefficient of  $u$  is 1.

It is easy to check that

$$\begin{aligned} \xi(t, s) &= \frac{1}{2} \int_0^t \left[ \int_{s-t+r}^{s+t-r} \psi(\nu) \, d\nu \right] v(r) \, dr \\ &= \frac{1}{2} \int_0^t [\Psi(s+t-r) - \Psi(s-t+r)] v(r) \, dr, \end{aligned}$$

where  $\Psi'(s) = \psi(s)$  and

$$\begin{aligned} \eta(t, s) &= \eta_1(t, s) + \eta_2(t, s), \\ \eta_1(t, s) &= \int_0^t (t-r)u(r) \, dr, \\ \eta_2(t, s) &= \frac{\rho}{4} \int_0^t \left\{ \int_{s-t+r}^{s+t-r} \left[ \int_0^r (\Psi(\nu+r-\mu) - \Psi(\nu-r+\mu)) v(\mu) \, d\mu \right] \, d\nu \right\} \, dr. \end{aligned}$$

Hence the components  $y_1(t)$  and  $y_2(t)$  of the observation are

$$\begin{aligned} y_2(t) &= \int_0^t G(t-\nu)v(\nu) \, d\nu, \quad G(t) = \frac{1}{2}[\psi(t) + \psi(-t)] \\ y_1(t) &= \frac{\rho}{4} \int_0^t (t-r)u(r) \, dr + \int_0^t H(t-r)v(r) \, dr \\ H(t) &= \int_0^t \left[ \int_{\tau-t}^{\tau+t} [\Psi(\nu+\tau) - \Psi(\nu-\tau)] \, d\nu \right] \, d\tau. \end{aligned}$$

The important property to be noted is that  $H(t)$  has a zero, for  $t = 0$  of higher order than that of  $t$ , the kernel which appears in the expression of  $y_1(t)$ . Hence, the output has the following form

$$\begin{bmatrix} \int_0^t [(t-s)^{r_1} H_1(t-s)u(s) + (t-s)^{r_2} H_2(t-s)v(s)] \, ds \\ \int_0^t [(t-s)^{\chi_1} L_1(t-s)u(s) + (t-s)^{\chi_2} L_2(t-s)v(s)] \, ds \end{bmatrix}$$

and

$$\begin{cases} r_1 < r_2, \\ \chi_1 > \chi_2, \\ \chi_2 < r_1 \end{cases}$$

(in fact in our example  $L_1(t) = 0$ .) We shall see that, under the conditions above, the diagonal terms of the  $2 \times 2$  matrix of the kernels have a dominant role.

Kernels like this we shall call “diagonally dominant” and the problem we are going to study is the deconvolution problem for diagonally dominant systems.

## 1.2 Regularity assumptions and diagonally dominant matrices

The assumptions we make are both regularity assumptions and assumptions on the structure of the matrix  $K(t)$ . As to the regularity assumptions on the input, we shall assume that  $u(t)$  is piecewise of class  $C^\infty$ , although we shall see that a finite number of derivatives will suffice for the proofs. We note that with this assumption  $u(t)$  may have jumps so that the assumption is usually satisfied in the applications.

In order to get a Volterra equation, i.e. a convolution with the upper limit of integration equal to  $t$ , we must have  $K(t) = 0$  for  $t < 0$ . We shall say that  $K(t)$  is “smooth” when it is the restriction to  $[0, T)$  of a function  $\tilde{K}(t)$  which is smooth on  $\mathbb{R}$ . The smoothness we require is that  $\tilde{K}(t)$  be of class  $C^\infty$  and that its Taylor series converges to  $\tilde{K}(t)$  near 0.

Let now  $\mathbf{r}_i(t)$  be the rows of  $K(t)$ ,

$$K(t) = \begin{bmatrix} \mathbf{r}_1(t) \\ \vdots \\ \mathbf{r}_n(t) \end{bmatrix}, \quad \mathbf{r}_i(t) = \left[ k_{i,1}(t) \quad k_{i,2}(t) \quad \dots \quad k_{i,n-1}(t) \quad k_{i,n}(t) \right],$$

$$k_{i,j}(t) = \sum_{\mu=0}^{+\infty} k_{i,j;(\mu)} t^\mu$$

(the series expansion converges in a neighborhood of  $t = 0$ ).

We define

$$\nu_{i,j} = \min_{\mu} \{k_{i,j;(\mu)} \neq 0\} \quad \text{or} \quad \nu_{i,j} = +\infty \quad \text{if} \quad k_{i,j;(\mu)} = 0 \quad \text{for every} \quad \mu.$$

and

$$\nu_i = \min\{\nu_{i,j}, 1 \leq j \leq n\}.$$

We shall call the number  $\nu_i$  the  $i$ -th *row degree* of  $K(t)$ .

**Remark 2** Actually this is a variant of the usual definition of the order of the zero at  $\infty$  for a row vector. When  $K(t)$  is a row vector whose entries are exponential polynomials, then we can consider its Laplace transform  $\hat{K}(\lambda)$ . Let its entries be

$$\hat{k}_{i,j}(\lambda) = \sum_{\mu=0}^{+\infty} k_{i,j;(\mu)} \frac{1}{\lambda^\mu}.$$

The number  $\nu_i$  defined as above is the order of  $\infty$  as a zero of  $\hat{\mathbf{r}}(\lambda)$ , see [12]. ■

We introduce the following sets:

$$I_1 = \{j : \nu_{1,j} = \nu_1\}, \quad I_2 = \{j : \nu_{2,j} = \nu_2\}, \dots, \quad I_n = \{j : \nu_{n,j} = \nu_n\}$$

and

**Assumption 3** For  $k$  and  $s$  between 1 and  $n$ , either  $I_k \cap I_s = \emptyset$  or  $I_k = I_s$  and then  $\nu_k = \nu_s$ . ■

I.e. we assume that the sets  $I_j$  are a partition of the set  $\{1, 2, \dots, n\}$ .

We shall apply a permutation matrix  $Q$  to  $K(t)$  on the left, so to order the row indices in increasing order,

$$\nu_1 \leq \nu_2 \leq \dots \leq \nu_n, \quad \nu_1 \geq 0.$$

This does not change the problem to be studied.

Let us consider now the first row  $\mathbf{r}_1(t)$  of  $K(t)$  and let us consider the entries of this row in  $I_1$ , i.e. such that  $\nu_{1,j} = \nu_1$ . There exists a permutation matrix  $P$  which moves these entries in the first positions:

$$\mathbf{r}_1(t)P = \left[ r_{1,1}(t) \quad \dots \quad r_{1,j_1}(t) \quad r_{1,j_1+1}(t) \quad \dots \quad r_{1,n}(t) \right]$$

and

$$\nu_{1,i} = \nu_1, \quad 1 \leq i \leq j_1, \quad \nu_{1,i} > \nu_1 \quad \text{if } i > j_1.$$

Due to the assumption 3, either  $I_1 = I_2$  and  $\nu_1 = \nu_2$  or  $\nu_{2,i} > \nu_2$  if  $i \in I_1$ .

Now we do the same for the second larger exponent etc. We end up with a block form

$$K(t)P = \begin{bmatrix} \tilde{K}_{1,1}(t) & \tilde{K}_{1,2}(t) & \dots & \tilde{K}_{1,s}(t) \\ \tilde{K}_{2,1}(t) & \tilde{K}_{2,2}(t) & \dots & \tilde{K}_{2,s}(t) \\ \vdots & & & \\ \tilde{K}_{i,1}(t) & \tilde{K}_{i,2}(t) & \dots & \tilde{K}_{i,s}(t) \\ \vdots & & & \end{bmatrix}$$

and every entry of  $\tilde{K}_{ii}(t)$  has the same exponent, let us call it  $\nu_i$  from now on, which is larger than the exponents in the entries in the previous and subsequent blocks in the same line.

A second key assumption is:

**Assumption 4** The diagonal blocks  $\tilde{K}_{ii}$  are square and not degenerate. ■

We sum up, the matrix  $K(t)$  can be reduced to the following block form:

$$QK(t)P = \begin{bmatrix} t^{\nu_1} K_{1,1}(t) & t^{\nu_1+1} K_{1,2}(t) & \dots & t^{\nu_1+1} K_{1,s}(t) \\ t^{\nu_2+1} K_{2,1}(t) & t^{\nu_2} K_{2,2}(t) & \dots & t^{\nu_2+1} K_{2,s}(t) \\ \vdots & & & \\ t^{\nu_s+1} K_{s,1}(t) & t^{\nu_s+1} K_{s,2}(t) & \dots & t^{\nu_s} K_{s,s}(t) \end{bmatrix} \quad (3)$$

and (possibly after a last coordinate transformation)

$$K_{i,i}(0) = \frac{1}{\nu!} I.$$

Matrices with the properties stated in Assumptions 3 and 4 will be called *diagonally dominant* for obvious reasons.

We shall see that for diagonally dominant kernels the deconvolution problem can be solved recursively, using a multiscale singular perturbation approach.

### 1.3 Comments on the literature

A singular perturbations approach to inverse problems has been advocated by Lavrentev, see [9]. Chapter 5 in this book is devoted to the application of this idea to a deconvolution problem. As it turns out, Lavrentev method when applied to integral equations of Volterra type gives an on-line reconstruction algorithm. A different on-line reconstruction algorithms has been advocated in [1] and the algorithm proposed here has been rigorously studied in [6] in the case  $K(0) = 1$ . This analysis is then extended to the case  $K(t) \sim t^\nu \tilde{K}(t)$ ,  $\tilde{K}(0) = I$  in [7].

Volterra integral equations may represent input output relations of systems in state space form. On-line algorithms for input identifications when the state space description of the system is known and the full state space is available to measures are in [11]. In fact, this book also considers special instances of output measures. This case however has been solved in general in [2]. The algorithm in [2] has been extended to general classes of Volterra equations in [3, 4], and will be extended more in general in the present paper.

The use of on-line deconvolution problems in regulation and control applications can be found in [5, 8, 13].

## 2 The multiscale singular perturbation and the deconvolution problem

In this section we show how the deconvolution problem can be solved for the case of diagonally dominant kernels. We examine the noiseless case here, while the effect of the noise will be studied in Section 2.1.

We assume that the exchanges of rows and columns of  $K(t)$  have already been performed so that we can take  $Q$  and  $P$  in (3) to be the identity matrices.

In order to be as clear as possible, we study here the case that  $K(t)$  is a  $2 \times 2$  block matrix, with  $\nu_1 = 0$ . The general case is then considered in section 3. So, we now consider the case that

$$K(t) = \begin{bmatrix} K_1(t) & tK_2(t) \\ t^{\nu+1}K_3(t) & t^\nu K_4(t) \end{bmatrix}, \quad K_1(0) = I, \quad K_4(0) = \frac{1}{\nu!}I. \quad (4)$$

Note that  $K_2(0)$  and  $K_3(0)$  might be zero.

We introduce the following notations:  $I$  is the identity operator,  $D$  denotes the derivative while  $J$  is the integration operator

$$(Iu)(t) = u(t), \quad (Du)(t) = u'(t), \quad (Ju)(t) = \int_0^t u(s) ds$$

so that

$$\left\{ \begin{array}{l} (J^{\nu+1}u)(t) = \frac{1}{\nu!} \int_0^t (t-s)^\nu u(s) ds, \quad D^r J^\nu = \begin{cases} J^{\nu-r} & \text{if } r < \nu \\ I & \text{if } r = \nu \\ D^{r-\nu} & \text{if } r > \nu, \end{cases} \\ D^r \int_0^t (t-s)^\nu K(t-s)u(s) ds = \\ \int_0^t \{D^r[(t-s)^\nu K(t-s)]\} u(s) ds, \quad r < \nu + 1, \\ D^{\nu+1} \int_0^t (t-s)^\nu K(t-s)u(s) ds = \nu!K(0)u(t) + \int_0^t \Phi(t-s)u(s) ds \end{array} \right. \quad (5)$$

where, using Leibniz formula, it is easily seen that

$$\begin{aligned} \Phi(t) = \nu!K'(t) + \sum_{j=0}^{\nu-1} \binom{\nu}{j} \frac{\nu!}{(\nu-j)!} \{(\nu-j)t^{\nu-j-1}K^{(\nu-j)}(t) \\ + t^{\nu-j}K^{(\nu-j+1)}(t)\}. \end{aligned} \quad (6)$$

The harmless assumption  $K_4(0) = I/\nu!$  has been made in order to simplify the factor  $\nu!$  in the previous expression.

We now consider the following two-scale singular perturbation problem. The first row essentially represents the Lavrentev singular perturbation approach which is used in the case  $K(0)$  is invertible. We prefer to write  $\{(\epsilon I + J) - J\}v_1$  instead of simply  $v_1$ , for analogy with the second line. The second line corresponds to a singular perturbation problem of higher order. This choice is suggested by the Morse quasi-canonical form used in [2] for the approximate construction of the inverse system of a linear finite dimensional control system.

$$\begin{aligned}
& \{(\epsilon I + J) - J\} v_1 + \int_0^t \{K_1(t-s)v_1(s) + (t-s)K_2(t-s)v_2(s)\} ds \\
& = y_1(t) = \int_0^t \{K_1(t-s)u_1(s) + (t-s)K_2(t-s)u_2(s)\} ds
\end{aligned} \tag{7}$$

and

$$\begin{aligned}
& \{(\epsilon I + J)^{\nu+1} - J^{\nu+1}\} v_2 \\
& + \int_0^t \{(t-s)^{\nu+1}K_3(t-s)v_1(s) + (t-s)^\nu K_4(t-s)v_2(s)\} ds \\
& = y_2(t) = \int_0^t \{(t-s)^{\nu+1}K_3(t-s)u_1(s) + (t-s)^\nu K_2(t-s)u_2(s)\} ds.
\end{aligned} \tag{8}$$

We recall that  $v_1$  and  $v_2$  are vectors, so that differentiation and integration are applied to every component. Now the estimate  $\mathbf{v}(t)$  of  $u(t)$  is

$$\mathbf{v}(t) = \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix}$$

and we study the convergence of  $\mathbf{v}$  to  $u$ . It is convenient to introduce the error

$$e(t) = \begin{bmatrix} e_1(t) \\ e_2(t) \end{bmatrix} = \begin{bmatrix} v_1(t) - u_1(t) \\ v_2(t) - u_2(t) \end{bmatrix}$$

so that

$$\epsilon e_1(t) + \int_0^t K_1(t-s)e_1(s) ds = - \int_0^t (t-s)K_2(t-s)e_2(s) ds - \epsilon u_1(t) \tag{9}$$

and

$$\begin{aligned}
& \{(\epsilon I + J)^{\nu+1} - J^{\nu+1}\} e_2 + \int_0^t (t-s)^\nu K_4(t-s)e_2(s) ds \\
& = - \int_0^t (t-s)^{\nu+1} K_3(t-s)e_1(s) ds - \sum_{r=0}^{\nu} \binom{\nu+1}{r} \epsilon^{\nu+1-r} J^r u_2
\end{aligned} \tag{10}$$

We shall prove the following Theorem:

**Theorem 5** *Let  $[0, T]$  be a fixed interval and let  $\mathbf{v}_\epsilon$  be constructed as above. We have that  $\mathbf{v}_\epsilon \rightarrow u$  in  $L^2(0, T)$ . Furthermore,  $\mathbf{v}_\epsilon(t)$  and its derivatives converge respectively to  $u$  and to the corresponding derivatives of  $u$ , uniformly on every interval  $[a, b] \subseteq (0, T]$  such that  $u$  is smooth on  $(a - \sigma, b)$  for some number  $\sigma > 0$ .*

In the course of the proof, we shall compute derivatives of  $e(t)$ . This is only possible at the points where  $u(t)$  is differentiable. So, let  $t_0$  be such that  $u$  is smooth on  $(0, t_0)$ . We shall work first on this interval and then we shall show how the results can be extended. As we noted already, derivatives are only used as a device in the proofs.

We apply  $D$  and  $D^{\nu+1}$  respectively to (9) and (10) and we find

$$\epsilon e_1' = -e_1 - \int_0^t L_1(t-s)e(s) ds - \epsilon u_1'(t) \quad (11)$$

where

$$L_1(t) = \left[ K_1'(t) \quad D\{tK_2(t)\} \right]$$

We note that

$$\begin{aligned} & D^{\nu+1} \left\{ J^{\nu+1} e - \int_0^t (t-s)^\nu K_4(t-s)e_2(s) ds \right\} \\ &= \int_0^t \Phi(t-s)e_2(s) ds \end{aligned}$$

where  $\Phi(t)$  is given by (6) with  $K(t)$  replaced by  $K_2(t)$ . This shows that, after  $\nu + 1$  derivatives, Equation (10) takes the following form:

$$\begin{aligned} (\epsilon D + I)^{\nu+1} e_2 &= - \int_0^t L_2(t-s)e(s) ds \\ &\quad - \sum_{r=0}^{\nu} \binom{\nu+1}{r} \epsilon^{\nu+1-r} D^{\nu+1-r} u_2. \end{aligned} \quad (12)$$

This equality holds on those intervals over which  $u$  is smooth. Here

$$L_2(t) = \left[ D^{\nu+1} \{t^{\nu+1}K_3(t)\} \quad \Phi(t) \right].$$

We see from (11) that

$$e_1(t) = e^{-t/\epsilon} e_1(0) - \int_0^t e^{-(t-s)/\epsilon} u_1'(s) ds - \int_0^t \frac{1}{\epsilon} e^{-(t-s)/\epsilon} \int_0^s L_1(s-r)e(r) dr ds. \quad (13)$$

Here  $e_1(0) = u_1(0)$ . It is clear that  $e^{-t/\epsilon} e_1(0) = e^{-t/\epsilon} u_1(0)$  tends to zero in  $L^2(0, T)$  and uniformly on  $[\sigma, t_0]$  for every  $\sigma > 0$  while

$$\left\| \int_0^t e^{-(t-s)/\epsilon} u_1'(s) ds \right\| < M\epsilon.$$

The constant  $M$  depends on  $u_1'(t)$ . Even more, the  $L^2$  norms of these two terms are of the order of  $\sqrt{\epsilon}$  (the norm of the integral is of the order of  $\epsilon$ ,

but we consider  $\epsilon \in (0, 1)$  so that  $\sqrt{\epsilon}$  dominates). Hence we can write

$$\begin{aligned} \int_0^t \|e_1(s)\|^2 ds &\leq M_1\epsilon + 2 \int_0^t \left[ \int_0^r \frac{1}{\epsilon} e^{-(r-s)/\epsilon} \int_0^s L_1(s-\mu)e(\mu) d\mu ds \right]^2 dr \\ &\leq M_1\epsilon + 2 \int_0^t \left[ \int_0^r \sqrt{\frac{1}{\epsilon}} e^{-(r-s)/\epsilon} \sqrt{\frac{1}{\epsilon}} e^{-(r-s)/\epsilon} \int_0^s L_1(s-\mu)e(\mu) d\mu ds \right]^2 dr \\ &\leq M_1\epsilon + 2 \int_0^t \left[ \int_0^r \frac{1}{\epsilon} e^{-s/\epsilon} ds \right] dr \cdot \\ &\int_0^r \left[ \frac{1}{\epsilon} e^{-(r-s)/\epsilon} \left( \int_0^s \|L_1(\mu)\|^2 d\mu \int_0^s \|e(\mu)\|^2 d\mu \right) ds \right] dr \end{aligned} \quad (14)$$

$$\begin{aligned} &\leq M_1\epsilon + 2M_2 \int_0^t \int_0^r \left[ \int_\mu^r \frac{1}{\epsilon} e^{-(r-s)/\epsilon} ds \right] \|e(\mu)\|^2 d\mu dr \\ &\leq M_1\epsilon + 2M_3 \int_0^t \left( \int_0^r \|e(\mu)\|^2 d\mu \right) dr. \end{aligned} \quad (15)$$

From now on we shall use  $M$  for a generic constant, so that the constants  $M_i$  above,  $i = 1, 2, 3$  will be simply denoted as  $M$ .

We now try to give an analogous estimate for the second component  $e_2(t)$ .

Before we do this we introduce an observation, whose interest will appear later on.

Let  $\Psi_1(t; \tau, h)$ ,  $t \in [0, T]$ , be a family of functions of the parameters  $\tau$  and  $h$ , such that

$$\|\Psi_1'(t; \tau, h)\| < \chi_1(\tau, h) \quad \text{a.e. } t \in [0, T] \quad (16)$$

and let  $\Psi_1(t; \tau, h)$  appear as a further addendum on the right side of (7), hence of (9). In this case the following terms shows up on the right hand sides of, respectively, (11) and (13):

$$\Psi_1'(t; \tau, h), \quad \int_0^t \frac{1}{\epsilon} e^{-(t-s)/\epsilon} \Psi_1'(s; \tau, h) ds.$$

Consequently, we have an additional term

$$M \cdot \chi_1^2(\tau, h)$$

on the right hand side of (15) which now reads

$$\int_0^t \|e_1(s)\|^2 ds \leq M \left\{ \left[ \epsilon + \chi_1^2(\tau, h) \right] + \int_0^t \left( \int_0^r \|e(\mu)\|^2 d\mu \right) dr \right\} \quad (17)$$

for a suitable choice of the constant  $M$ .

We shall use this inequality in Section 2.1.

Now we estimate the second component  $e_2(t)$ . It is easier if we pass to the Laplace transform. We recall the following formulas:

$$\begin{cases} \mathcal{L}(D^k e) = \lambda^k \hat{e}(\lambda) - \sum_{j=0}^{k-1} \lambda^{k-1-j} e^{(j)}(0) \\ \mathcal{L}\left(\frac{1}{k!} t^k e^{at}\right) = \frac{1}{(\lambda - a)^{k+1}} \\ \int t^k e^{-at} dt = -e^{-at} \left[ \frac{k!}{a^{k+1}} + \sum_{m=0}^{k-1} \frac{k!}{(k-m)!} \frac{t^{k-m}}{a^{m+1}} \right]. \end{cases} \quad (18)$$

Moreover, we note that  $D^j v(0) = 0$  if  $j \leq \nu$  so that

$$e^{(j)}(0) = -u^{(j)}(0) \quad \text{if } j \leq \nu.$$

This is easily seen from (8) and the fact that for  $j \leq \nu$ ,  $D^j$  and the integrals commute.

We take the Laplace transform of both the sides of (12). We note

$$(\epsilon D + I)^{\nu+1} e_2 = e_2 + \sum_{k=1}^{\nu+1} \binom{\nu+1}{k} \epsilon^k D^k e_2,$$

so that its Laplace transform is

$$\begin{aligned} & \hat{e}_2(\lambda) + \sum_{k=1}^{\nu+1} \binom{\nu+1}{k} \epsilon^k \left\{ \lambda^k \hat{e}(\lambda) - \sum_{j=0}^{k-1} \lambda^j e_2^{(k-1-j)}(0) \right\} \\ &= (\epsilon \lambda + 1)^{\nu+1} \hat{e}_2(\lambda) - \sum_{j=0}^{\nu} \lambda^j \left[ \sum_{k=j+1}^{\nu+1} \binom{\nu+1}{k} \epsilon^k e_2^{(k-1-j)}(0) \right]. \end{aligned}$$

The important fact here is that  $j < \nu + 1$ .

Proceeding analogously we get

$$\begin{aligned} & \hat{e}_2(\lambda) \\ &= \frac{1}{(\epsilon \lambda + 1)^{\nu+1}} \sum_{j=0}^{\nu} \lambda^j \left[ \sum_{k=j+1}^{\nu+1} \binom{\nu+1}{k} \epsilon^k e_2^{(k-1-j)}(0) \right] \end{aligned} \quad (19)$$

$$- \frac{1}{(\epsilon \lambda + 1)^{\nu+1}} \sum_{r=0}^{\nu} \binom{\nu+1}{r} \epsilon^{\nu+1-r} \mathcal{L}(D^{\nu+1-r} u_2)(\lambda) \quad (20)$$

$$- \frac{1}{(\epsilon \lambda + 1)^{\nu+1}} \hat{L}_2(\lambda) \hat{e}(\lambda) ds. \quad (21)$$

Our goal now is to prove that the time functions which correspond to the lines (19) and (20) tend to zero in  $L^2(0, T)$  and uniformly on  $[\sigma, T]$  and

that an estimate similar to (15) holds for the term in (21). The first goal is achieved recalling that  $e^j(0) = -u^j(0)$ . We consider the individual terms separately and we keep precise track of the range of variation of the indices.

Let us consider the individual terms in the summations in (19). We ignore inessential constants and we see that these terms have the general form

$$\frac{\epsilon^k \lambda^j}{(\epsilon \lambda + 1)^{\nu+1}}, \quad \begin{cases} 0 \leq j \leq \nu \\ j+1 \leq k \leq \nu+1. \end{cases}$$

We expand  $\lambda^j$  as

$$\left( \lambda + \frac{1}{\epsilon} - \frac{1}{\epsilon} \right)^j$$

and we see that the individual terms now are

$$\frac{1}{\epsilon^{\nu+1-k+j-r}} \frac{1}{(\lambda + 1/\epsilon)^{\nu+1-r}}, \quad \begin{cases} 0 \leq j \leq \nu \\ j+1 \leq k \leq \nu+1 \\ 0 \leq r \leq j. \end{cases}$$

Hence we must give estimates on the functions

$$\frac{1}{\epsilon^{\nu+1-k+j-r}} t^{\nu-r} e^{-t/\epsilon}, \quad \begin{cases} 0 \leq j \leq \nu \\ j+1 \leq k \leq \nu+1 \\ 0 \leq r \leq j. \end{cases}$$

Clearly these functions tend uniformly to zero on compact sets of  $(0, +\infty)$  while the *square* of the  $L^2$  norm on  $(0, +\infty)$  is

$$\epsilon^{2(k-j)-1}.$$

This is of the order of  $\epsilon$  at least, because  $k - j \geq 1$ .

We consider now the row (20). We recall that we are working on the interval  $[0, t_0)$  over which  $u$  is smooth and in order to estimate the norms on this interval we can extend  $u$  smoothly to  $\mathbb{R}$ , so to have compact support, thanks to the causal nature of the problem (7) and (8), i.e. (9) and (10). We must now give an estimate for

$$\frac{1}{\epsilon^r} \int_0^t (t-s)^\nu e^{-(t-s)/\epsilon} f(s) \, ds$$

where  $f$  is a derivative of  $u$  and  $r \leq \nu$ . Young inequalities imply that the  $L^2$ -norm is less than

$$\|f\|_{L^2} \cdot \frac{1}{\epsilon^r} \int_0^{+\infty} t^\nu e^{-t/\epsilon} \, dt \asymp \epsilon^{\nu-r+1}$$

so that the  $L^2$  norm of the convolution is less then

$$M\epsilon$$

where  $M$  is a bound of the derivatives of  $u$ . Even more, using Schwartz inequality we see that the convolution converges to zero uniformly on  $[0, t_0]$ , of the order  $\epsilon$ . As  $\epsilon \rightarrow 0$ , the estimates we found are of the order  $\sqrt{\epsilon}$ .

Now we consider the term (21). In the time domain this is

$$\frac{1}{\epsilon^{\nu+1}} \int_0^t (t-s)^\nu e^{-(t-s)/\epsilon} \int_0^s L_2(s-r)e(r) dr ds.$$

We take the integral of the square of this function. We have:

$$\begin{aligned} & \int_0^t \left[ \frac{1}{\epsilon^{\nu+1}} \int_0^\tau (\tau-s)^\nu e^{-(\tau-s)/\epsilon} \int_0^s L_2(s-r)e(r) dr ds \right]^2 d\tau \\ & \leq M \int_0^t \int_0^s |e(r)|^2 dr ds. \end{aligned}$$

Consequently, we have the following estimates, which hold on an interval  $[0, t_0)$  over which  $u(t)$  is smooth:

$$\int_0^t \|e_1(s)\|^2 ds \leq M\epsilon + 2M \int_0^t \left( \int_0^r \|e(\mu)\|^2 d\mu \right) dr, \quad (22)$$

$$\int_0^t \|e_2(s)\|^2 ds \leq M\epsilon + 2M \int_0^t \left( \int_0^r \|e(\mu)\|^2 d\mu \right) dr. \quad (23)$$

Here, as already said,  $M$  is a generic constant. We sum and we find

$$\int_0^t \|e(s)\|^2 ds \leq M\epsilon + 2M \int_0^t \left( \int_0^r \|e(\mu)\|^2 d\mu \right) dr \quad (24)$$

so that, from Gronwall inequality,

$$\int_0^{t_0} \|e(s)\|^2 ds \leq M\epsilon. \quad (25)$$

This gives convergence in  $L^2(0, t_0)$ , where  $t_0$  is the first point in which  $u$  is not regular. In order to extend the result to a larger interval we need the following result on uniform convergence:

**Theorem 6** *We have  $e_1 \rightarrow 0$  uniformly on  $(\sigma, t_0]$  for every  $\sigma > 0$ .*

**Proof.** This is obvious for the first two terms in (13) while the last integral is less then

$$M \int_0^t \left[ \int_r^t \frac{1}{\epsilon} e^{-(t-s)/\epsilon} ds \right] \|e(r)\| dr.$$

The inner integral is bounded while, as we told already, the  $L^2$  norm of  $e(r)$  converges to zero. ■

We need a similar statements concerning  $e_2(t)$  and its derivatives up to the order  $\nu$ . The expression in (19) corresponds to a sum of terms of the type

$$(\text{const}) \cdot \frac{1}{\epsilon^s t^r} e^{-t/\epsilon},$$

which has the required properties. Line (20) corresponds to a linear combination of terms of the type

$$\frac{1}{\epsilon^r} \int_0^t s^\nu e^{-s/\epsilon} f(t-s) ds$$

where  $f$  is a suitable derivative of  $u_2$ . Boundedness of the derivatives of  $u_2$  shows that these integrals are dominated by

$$\text{const} \cdot \epsilon^{\nu+1-r}, \quad r \leq \nu,$$

as wanted.

Consequently, for  $j \leq \nu$ , we have

$$D^j e(t) = \psi_\epsilon(t) + \int_0^t \Psi(t-s) \int_0^s L(s-r) e(r) dr ds$$

where  $\Psi(t) = D^j [t^\nu e^{-t/\epsilon}]$ . Using the already known fact that the  $L^2$ -norm of  $e(t)$  converges to 0 we deduce:

**Theorem 7** *The component  $e_2(t)$  of  $e(t)$  converges to zero together with its derivatives of order up to  $\nu$ , uniformly on  $[\sigma, t_0]$ , for every  $\sigma > 0$ .*

This proves our result in the first interval  $[0, t_0]$ , up to the first point  $t_0$  where the unknown input  $u$  loses regularity. Let now  $u$  be smooth on  $(t_0, t_1)$ . We can extend our result to this second interval  $[t_0, t_1]$  too. In fact, we observe that  $\mathbf{v}$  is a smooth function so that the jump of  $e$ , and its derivatives up to  $\nu$  at  $t_0$ , is the same as the corresponding jump of  $u$ . This allows to repeat the previous arguments on  $[t_0, t_1]$ , where  $t_1$  is the second point at which  $u$  is not regular. The details are as in [14].

This completes the proof of Theorem 5.

Finally, we need a detour analogous to the one we have seen for the first component  $e_1(t)$ .

Let an additional error  $\Psi_2(t, \tau, h)$  be present on the right hand side of (8). The parameters  $h$  and  $\tau$  here introduced will be of use below. For the moment we can think to them as fixed. We assume

$$|D^{(\nu+1)} \Psi_2(t, \tau, h)| \leq \chi_2(\tau, h).$$

This gives a new term on the right side of (23), which is dominated by

$$M \cdot \chi_2(\tau, h)$$

so that inequality (23) is replaced by

$$\int_0^t \|e_2(s)\|^2 ds \leq M \left\{ \epsilon + \chi_2(\tau, h) + \int_0^t \left( \int_0^r \|e(\mu)\|^2 d\mu \right) dr \right\}.$$

We take this new inequality and the corresponding inequality for  $e_1$  into account and we find the following final estimate for  $e$ :

$$\int_0^{t_0} \|e(s)\|^2 ds \leq M \left[ \epsilon + \chi_1^2(\tau, h) + \chi_2^2(\tau, h) \right]. \quad (26)$$

## 2.1 Noisy measures

Noisy measures are considered now. We still denote  $y(t)$  the measure without errors while the noisy measure is denoted  $\xi(t)$ . The noise is  $\theta(t)$ , so that

$$\xi(t) = y(t) + \theta(t) = \begin{bmatrix} \xi_1(t) \\ \xi_2(t) \end{bmatrix} = \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} + \begin{bmatrix} \theta_1(t) \\ \theta_2(t) \end{bmatrix}.$$

We assume that  $\theta(t)$  is measurable and it is bounded by  $h$  either in  $L^\infty$  or in  $L^2$  norm.

We adapt the idea in [14], inspired by the mollification method in [10]. We introduce a second regularization parameter  $\tau$  ( $\tau = \epsilon$  is a possible choice) and we replace  $y_1(t)$  on the right hand side of Eq. (7) with

$$\frac{1}{\tau} \int_0^t e^{-(t-s)/\tau} \xi(s) ds$$

which we represent as

$$y_1(t) + \left[ \frac{1}{\tau} \int_0^t e^{-(t-s)/\tau} y_1(s) ds - y_1(t) \right] + \frac{1}{\tau} \int_0^t e^{-(t-s)/\tau} \theta_1(s) ds$$

so that Eq. (9) maintains its same form, with a further addendum on the right hand side, which is

$$\Psi_1(t, \tau, h) = \left[ \frac{1}{\tau} \int_0^t e^{-(t-s)/\tau} y_1(s) ds - y_1(t) \right] + \frac{1}{\tau} \int_0^t e^{-(t-s)/\tau} \theta_1(s) ds$$

(note the usual abuse of language:  $\Psi_1$  does depend on  $\theta$  but an estimate for it will only depend on  $h$  so that we use the notation above). It is now sufficient to obtain an estimate for  $\Psi_1$ , of the form (16). Consider the two addenda separately. As to the second one, its derivative is

$$\frac{1}{\tau} \theta_1(t) + \frac{1}{\tau^2} \int_0^t e^{-(t-s)/\tau} \theta(s) ds$$

Both the  $L^\infty(0, T)$  and the  $L^2(0, T)$  norms of this term are less than

$$(\text{const}) \cdot \frac{h}{\tau}.$$

As to the first term, it is clear that

$$\begin{aligned} & \left\| \frac{d}{dt} \left[ \frac{1}{\tau} \int_0^t e^{-(t-s)/\tau} y(s) ds - y(t) \right] \right\| \\ &= \left\| \frac{1}{\tau} \int_0^t e^{-(t-s)/\tau} y'(s) ds - y'(t) \right\| \leq M\chi_0(\tau) \end{aligned}$$

and

$$\lim_{t \rightarrow 0^+} \chi_0(\tau) = 0,$$

because the family of kernels which are zero for  $t < 0$ , and equal to  $(1/\tau)e^{-t/\tau}$ , is an approximate identity. This estimate holds for every  $t$  on each interval over which  $y$  is differentiable; it holds for  $L^2$  norms in general.

Now we proceed analogously for the second component  $y_2$ .

It is convenient to introduce a second penalization parameter  $\sigma > 0$ . It might be  $\sigma = \tau$  but we are not forced to this choice and we shall see that this additional degree of freedom is useful.

We choose the kernel

$$K_\sigma(t) = \frac{\sigma^{\nu+1}}{\nu!} t^\nu e^{-t/\sigma} \quad t > 0$$

(zero for negative times). This family of kernels is an approximate identity so that

$$\begin{aligned} & \left\| D^{\nu+1} \left[ \int_0^t K_\sigma(t-s)y(s) ds - y(t) \right] \right\| \\ &= \left\| \left[ \int_0^t K_\sigma(t-s)y^{\nu+1}(s) ds - y^{(\nu+1)}(t) \right] \right\| \leq M_2\chi_2(\sigma) \end{aligned}$$

with

$$\lim_{t \rightarrow 0^+} \chi_2(\sigma) = 0,$$

The effect of the disturbance  $\theta_2$  is now estimated as follows: we replace  $y_2(t)$  in the right hand side of (8) with  $K_\sigma \star \xi_2$  (we use  $\star$  to denote the convolution) which we represent as

$$(K_\sigma \star \xi_2)(t) = y_2(t) + \Psi_2(t; \sigma, h), \quad \Psi_2(t; \sigma, h) = [(K_\sigma \star y_2)(t) - y_2(t)] + (K_\sigma \star \theta_2)(t).$$

and we find that  $\chi_2(\sigma, h)$  is now of the order

$$\frac{h}{\sigma^{\nu+1}}.$$

This shows that the following consistency result holds:

**Theorem 8** *Let  $\epsilon \rightarrow 0+$ ,  $\tau \rightarrow 0+$  and  $\sigma \rightarrow 0+$  so to respect the conditions*

$$\frac{h}{\tau} \rightarrow 0, \quad \frac{h}{\sigma^{\nu+1}} \rightarrow 0.$$

*We have that  $\mathbf{v}$  converges to  $u$  in  $L^2(0, T)$  for each  $T > 0$ . The convergence is uniform on each interval  $[a, b]$  provided that  $u$  and its derivatives have no jump neither at  $a$  nor at the interior points of  $[a, b]$ .*

### 3 The general case

It is now easily seen that the ideas presented in the previous sections can be adapted to the general case. We must consider first the case that  $\nu_1 > 0$  and then the case that the matrix  $K(t)$  has a larger number of blocks.

Looking back at the computations in Sect. 2 we see that the following property has a crucial role: the order of the zero at 0 of the kernel of the integral on the right hand side of (9) is larger than  $\nu_1$ , the order of the zero of  $K_1(t)$ . We explicitly presented the computations in the case  $\nu_1 = 0$ , but this is not a crucial point. If  $\nu_1 > 0$  then the same inequality (15), i.e. (22), is obtained provided that  $\{(\epsilon I + J) - J\}v_1$ , i.e.  $\epsilon v_1$ , is replaced by  $\{(\epsilon I + J)^{\nu_1+1} - J^{\nu_1+1}\}v_1$ . Computations are precisely as those leading to (23). The second equation is not affected by this modification since  $\nu_2 > \nu_1$ . The case  $\nu_1 = 0$  has been singled out solely for the sake of clarity, since the analysis of this case suggests how to approach the general case.

So, we have indicated how the condition  $\nu_1 = 0$  can be removed. As to the fact that we considered a  $2 \times 2$  block matrix, this is not an issue at all, as long as  $K(t)$  is diagonally dominant, since in the proofs every block row is manipulated independently of the others. Note in fact that the integrals in the right hand sides of (11) and (12) depend on every component of  $e(t)$ . So, we conclude that in the general case of a matrix  $K(t)$  equal to the right hand side of (3), we construct  $\mathbf{v}(t)$  from the equation

$$H_\epsilon \mathbf{v}(t) = -(K \star \mathbf{v})(t) + (K \star u)(t)$$

( $\star$  denotes convolution) where  $H_\epsilon$  is a block operator matrix. The sizes of its blocks are compatible with those of  $K(t)$ ; the off-diagonal blocks are zero while the  $i$ -th block on the diagonal is

$$\left\{ (\epsilon I + J)^{\nu_i+1} - J^{\nu_i+1} \right\}.$$

Precisely the same proofs as in Sect. 2, applied to every block row, lead to the inequality (24) and we are done.

## References

- [1] J.V. Beck, B. Blackwell, C.R. St. Clair, *Inverse heat conduction*, Wiley-Interscience, 1985.
- [2] F. Fagnani, L. Pandolfi, A singular perturbation approach to a recursive deconvolution problem, *SIAM J. Control Optim.* **40** 1384-1405, 2002.
- [3] F. Fagnani, L. Pandolfi, A recursive algorithm for the approximate solution of Volterra integral equations of the first kind of convolution type, *Inverse Problems* **19** 23-47, 2003.
- [4] F. Fagnani, L. Pandolfi, On the solution of a class of Volterra integral equations of the first kind. *J. Inverse Ill-Posed Probl.* **11** 485-503, 2003.
- [5] F. Fagnani, V. Maksimov, L. Pandolfi, A recursive deconvolution approach to disturbance reduction, *IEEE Trans. Automat. Control*, **49** 907-921, 2004.
- [6] P.K. Lamm, Future-sequential regularization methods for ill-posed Volterra equations, *J. Mathematical Analysis Appl.* **195** 465-494, 1995.
- [7] P.K. Lamm, Regularized inversion of finitely smoothing Volterra operators: predictor-corrector regularization methods, *Inverse Problems*, **13** 375-402, 1997.
- [8] P.K. Lamm, Variable smoothing regularization methods for inverse problems, in *Theory and Practice of Control and Systems*, A. Tornambè, G. Conte, A.M. Perdon Ed.s, World Scientific, Singapore, 331-336, 1999.
- [9] M.M. Lavrentev, V.G. Romanov, S.P. Shishatskiĭ, *Ill-posed problems of mathematical physics and analysis*, Translations of Mathematical Monographs n. 64 American Mathematical Society, Providence, RI, 1986.
- [10] D.A. Murio, *The mollification method and the numerical solution of ill-posed problems*, John Wiley & Sons, Inc., New York, 1993.
- [11] Yu. S. Osipov, A.V. Kryazhimskii, *Inverse problems for ordinary differential equations: dynamical solutions*, Gordon and Breach, London, 1995.
- [12] L. Pandolfi, The structure of the singularities of holomorphic matrices, *Proc. Roy. Soc. Edinburgh Sect. A*, **107** 1-14 1987.
- [13] L. Pandolfi, Adaptive recursive deconvolution and adaptive noise cancellation, *Int. J. Control*, **80** 403-415, 2007.

- [14] Perri, F., Pandolfi, L., Input identification to a class of nonlinear input-output causal systems, *Comput. Math. Appl.* **51** 1773–1788, 2006.