INTERNATIONAL JOURNAL OF ADAPTIVE CONTROL AND SIGNAL PROCESSING, VOL. 9, 3–17 (1995)

DECONVOLUTION WITH BOUNDED UNCERTAINTY

PATRICK L. COMBETTES*

Department of Electrical Engineering, City College and Graduate School, City University of New York, New York, NY 10031, U.S.A.

AND

H. JOEL TRUSSELL

Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, NC 27695, U.S.A.

SUMMARY

In deconvolution problems there are two primary sources of uncertainty in the data formation mechanism, namely measurement noise and errors in the model of the system. In this paper we develop an abstract set theoretic deconvolution framework for problems in which the only information available about these sources of uncertainty consists of bounds. Iterative methods based on projections are used to generate solutions consistent with these bounds, the output data signal and *a priori* knowledge about the input signal. An example of application of this general framework to discrete signal recovery is demonstrated.

KEY WORDS Signal deconvolution Bounded–error Set theoretic estimation Convex sets Projections

1. INTRODUCTION

The generic signal deconvolution problem is to estimate the input of a known linear shift-invariant system from its output. This problem arises in a wide range of applications, including channel equalization, seismology, medical imaging and astronomy. Traditionally, deconvolution has been approached as an optimization problem, either within the framework of regularization techniques for deterministic models[1] or within that of point estimation theory (primarily maximum likelihood and Bayesian methods) for probabilistic models.[2,3] The practical value of the solutions thus obtained is often questionable, be it because of the relevance of the optimality criterion or because of the adequacy of the underlying statistical hypotheses.[4]

An alternative route to deconvolution is provided by set theoretic estimation. In this framework the requirement for a single estimate is relaxed and the end product is a set of equally valid solutions defined as those objects consistent with all the information available about the problem, i.e. consistent with the *a priori* knowledge as well as with the data.[5] The set of signals consistent with a particular piece of information is called a property set; the intersection of all property sets is the set of feasible signals, i.e. the solution set. The set theoretic approach was first employed in signal deconvolution in Reference 6. In that study it was assumed that a complete probabilistic description of the noise was available: it was taken

* Author to whom correspondence should be addressed.

CCC 0890–6327/95/010003–15

to be zero-mean, white and Gaussian with known power. From this knowledge various property sets were constructed. In Reference 7 some of these sets were re-examined in the context of fuzzy set theory. The use of a wide range of probabilistic properties of the noise in general set theoretic estimation problems was addressed in Reference 8. In Reference 9 the analysis of Reference 6 was extended to incorporate instances in which the convolution kernel is random with known probabilistic attributes. More specific applications of set theoretic deconvolution can be found in Reference 10–12.

In this paper we investigate the problem of set theoretic deconvolution when the only information available about the uncertainty consists of error bounds. Two sources of uncertainty are considered: the convolution kernel has unknown variations about a known component and additive measurement noise is present. In contrast with the set theoretic deconvolution studies mentioned above, no statistical description of these two sources of uncertainty is assumed and only bounds are known. In systems theory a significant amount of work has been devoted to set theoretic estimation with unknown but bounded uncertainty. This approach originated in the field of state estimation [13] and applications and extensions have been reported in various areas such as speech processing, [14] spectral estimation, [15] control, [16] system identification [17,18] and filtering. [19] In all these studies the feasibility set is approximated (from inside or outside) by a geometrically simple set (e.g. an ellipsoid or a hyperparallelepiped) which is updated with each data sample. Methods have also been developed to characterize the feasibility set exactly for particular models. [20–22] In signal deconvolution the bounding ellipsoid approach has been utilized in Reference 23, where the centre of an ellipsoid bounding the intersection of the property sets was proposed as a solution. It is important to note that such a methodology is not guaranteed to yield a feasible solution, especially if the size of the signal to be recovered is large. Our set theoretic approach to deconvolution departs from the above setting. A single feasible solution will be obtained via the theory of convex projections in Hilbert spaces. This framework, which has been employed in various signal recovery applications, [4] allows one to compute exact feasible signals for a wide class of property sets and thereby offers great flexibility in the incorporation of *a priori* knowledge.

The paper is organized as follows. The general bounded–error deconvolution problem is posed in Section 2 and its set theoretic formulation is derived in Section 3. The theoretical and practical aspects of bounded–error deconvolution and its relation to existing work are discussed in Section 4. The approach is illustrated in Section 5 with an application to discrete signal recovery. Section 6 concludes the paper.

2. BOUNDED ERROR MODEL

2.1. *Hilbert space setting*

For the sake of generality we pose the deconvolution problem in an abstract Hilbert function space. Particular cases of the analysis relevant to applied work will be discussed in Section 2.3. Before proceeding with the analysis, we shall introduce our notation and recall a few basic facts about $\mathscr{L}^p$ spaces. Background on measure theory and integration can be found in References 24 and 25.

Let $(\Omega, \mathscr{F}, \mu)$ be a measure space and let $\mathscr{L}^0(\Omega, \mathscr{F}, \mu)$ be the space of (classes of equivalence of μ-a.e. equal) real-valued, $\mathscr{F}$-measurable functions defined on the set Ω. The $\mathscr{L}^p(\Omega, \mathscr{F}, \mu)$

spaces ($\mathscr{L}^p$ for short) are defined as

$$\mathscr{L}^p = \begin{cases} \{a \in \mathscr{L}^0(\Omega, \mathscr{F}, \mu) \mid \int_\Omega |a|^p \, d\mu < \infty\} & \text{if } 1 \leqslant p < \infty \\ \{a \in \mathscr{L}^0(\Omega, \mathscr{F}, \mu) \mid |a| < \infty \text{ a.e.}\} & \text{if } p = \infty \end{cases} \tag{1}$$

$\mathscr{L}^p$ is a Banach space with norm

$$(\forall a \in \mathscr{L}^p) \quad \| a \|_p = \begin{cases} (\int_\Omega |a|^p \, d\mu)^{1/p} & \text{if } 1 \leqslant p < \infty \\ \inf\{\delta \in \mathbb{R} \mid |a| \leqslant \delta \text{ a.e.}\} & \text{if } p = \infty \end{cases} \tag{2}$$

($\| a \|_\infty$ is called the essential supremum of a). In particular, $\mathscr{L}^2$ is a Hilbert space whose scalar product is given by $\langle a \mid b \rangle = \int_\Omega ab \, d\mu$. Its norm will simply be denoted by $\| \cdot \|$ instead of $\| \cdot \|_2$.

Throughout this paper h and $\bar{T}$ are two functions in $\mathscr{L}^2$. Mathematically the deconvolution problem can be stated as that of determining h given $\bar{T}$ and a function x related to h via the linear transformation

$$(\forall \omega \in \Omega) \quad x(\omega) = \int_\Omega \bar{T}(\omega - \eta) h(\eta) \mu(d\eta) \tag{3}$$

which will simply be written as

$$x = \bar{T} * h \tag{4}$$

The physical interpretation of this equation is that $\bar{T}$ is the kernel of a linear shift-invariant system whose input is the signal h and whose output is the signal x. The variable ω represents a physical dimension such as space or time. In practice the ideal model (4) is never exact, since the data generation mechanism is inherently uncertain. A more realistic model should also comprise an error term, i.e.

$$x = \bar{T} * h + u \tag{5}$$

The signal u will be referred to as the error signal. A shortcoming of such a global modelling of the uncertainty surrounding the true system is that little information on u is likely to be available. In order to structure the error, we shall assume that two major sources of uncertainty are present: inaccuracies in the kernel of the linear system and additive measurement noise. The exact kernel T of the system is taken to be of the form

$$(\forall (\omega, \eta) \in \Omega^2) \quad T(\omega, \eta) = \bar{T}(\omega - \eta) + \tilde{T}(\omega, \eta) \tag{6}$$

In other words, $\tilde{T}$ is an unknown shift-variant component which represents the unmodelled dynamics of the system. This leads to the expression for the error signal

$$(\forall \omega \in \Omega) \quad u(\omega) = \int_\Omega \tilde{T}(\omega, \eta) h(\eta) \mu(d\eta) + v(\omega) \tag{7}$$

where v is the measurement noise. We can rewrite (7) more compactly as

$$(\forall \omega \in \Omega) \quad u(\omega) = \langle \tilde{T}_\omega \mid h \rangle + v(\omega) \tag{8}$$

where $\tilde{T}_\omega : \eta \mapsto \tilde{T}(\omega, \eta)$.

2.2. *Bounds on the error signal*

The basic assumption vis-à-vis the uncertainty is that the terms $\tilde{T}$ and v are totally unknown except for bounds. The amplitude of the noise signal is assumed to be bounded by a known function β, i.e.

$$(\forall \omega \in \Omega) \quad |v(\omega)| \leqslant \beta(\omega) < \infty \tag{9}$$

As to the component of u involving $\tilde{T}$, several bounds are possible depending on the nature of the *a priori* knowledge. Let us define

$$p' = \begin{cases} \infty & \text{if } p = 1 \\ p/(p-1) & \text{if } 1 < p < \infty \\ 1 & \text{if } p = \infty \end{cases} \tag{10}$$

Now let us fix an arbitrary ω in Ω. Assume that, for some p in $[1, \infty]$, h lies in $\mathscr{L}^2 \cap \mathscr{L}^p$ and that there exists a number $\alpha_p(\omega)$ such that

$$\| \tilde{T}_\omega \|_{p'} \leqslant \alpha_p(\omega) < \infty \tag{11}$$

Then, by invoking Hölder's inequality,[24,25] we get

$$|\langle \tilde{T}_\omega | h \rangle| \leqslant \| \tilde{T}_\omega h \|_1 \leqslant \| \tilde{T}_\omega \|_{p'} . \| h \|_p \leqslant \alpha_p(\omega) . \| h \|_p \tag{12}$$

It then follows from (8) and (9) that

$$|u(\omega)| \leqslant \alpha_p(\omega) . \| h \|_p + \beta(\omega) = \gamma(\omega) \tag{13}$$

Therefore the error signal is pointwise bounded: $|u| \leqslant \gamma$. At this point a few remarks are in order apropos of the construction of the bounding signal γ. First, it is noticed that to allow greater flexibility in the modelling of the disturbances $\tilde{T}$ and v, the bound γ is actually a function of ω. In many applications, though, the bound will be shift-invariant and γ will reduce to a constant. Now let us consider the type of information needed to compute $\gamma(\omega)$. Besides the bound β on the noise, one needs information on the disturbance $\tilde{T}$ and the input signal h. Thus for $p = 1$ one needs the $\mathscr{L}^1$-norm of h and a bound on the essential supremum of $\tilde{T}_\omega$; for $p = 2$ one needs the energy of h and a bound on the energy of $\tilde{T}_\omega$; finally, for $p = \infty$ one needs the essential supremum of h and a bound on the $\mathscr{L}^1$-norm of $\tilde{T}_\omega$. Of course, if $\alpha_p(\omega)$ and $\| h \|_p$ are available for more than one value of p, one will choose that which yields the tightest bounding. On the other hand, in some problems the only information on $\tilde{T}_\omega$ and h may be $\alpha_p(\omega)$ and $\| h \|_q$, with $p \neq q$. If $\mu(\Omega) < \infty$, a bound can still be obtained in some cases by noting that[24,25]

$$(\forall (r,s) \in [1,\infty]^2)(\forall a \in \mathscr{L}^s) \quad r \leqslant s \quad \Rightarrow \quad \| a \|_r \leqslant \| a \|_s \mu(\Omega)^{(1/r)-(1/s)} \tag{14}$$

Hence, if $p < q$, one can use $\| h \|_q \mu(\Omega)^{(1/p)-(1/q)}$ in lieu of $\| h \|_p$ in (12).

2.3. *Specific models*

A quite general modelling has been carried out based on the powerful theory of $\mathscr{L}^p$ spaces. We shall now consider two specific measure spaces $(\Omega, \mathscr{F}, \mu)$ which are often employed in applied work.

First let Ω be a subset of $\mathbb{R}^k$ (the set of real k-tuples), $\mathscr{F}$ the Borel subsets of Ω, and μ the k-dimensional Lebesgue measure. Then $\mathscr{L}^2$ is the space L^2 of square-integrable functions of k real variables $t = (t_1, \dots, t_k)$ on a domain Ω,[24,25] and h models a k-dimensional finite energy analogue signal. Moreover, (5) and (7) lead to the model

$$(\forall t \in \Omega) \quad x(t) = \int_\Omega \bar{T}(t-s) h(s)\, ds + \int_\Omega \tilde{T}(t,s) h(s)\, ds + v(t) \tag{15}$$

Now let $\Omega = \mathbb{Z}^k$ (the set of k-tuples of integers), $\mathscr{F}$ be the class of all subsets of Ω, and μ the counting measure. Then $\mathscr{L}^2$ is the space l^2 of square-summable k-dimensional sequences[24,25] indexed on $n = (n_1, \dots, n_k)$, and h models a k-dimensional finite energy discrete signal. Moreover, (5) and (7) lead to the model

$$(\forall n \in \mathbb{Z}^k) \quad x(n) = \sum_{m \in \mathbb{Z}^k} \bar{T}(n-m)h(m) + \sum_{m \in \mathbb{Z}^k} \tilde{T}(n,m)h(m) + v(n) \tag{16}$$

Since in practice the deconvolution is eventually performed on a digital computer, one must consider discrete signals of finite extent and replace $\mathbb{Z}^k$ by a finite subset Ω. The signals x, h and v can then be represented by finite k-dimensional tensors. One can then stack the columns of the tensors to obtain vectors and put (16) in the vector–matrix form

$$x = \bar{T}h + \tilde{T}h + v \tag{17}$$

3. SET THEORETIC FORMULATION

The information available about the deconvolution problem is associated with a collection of property sets $(S_\iota)_{\iota \in I}$ in $\mathscr{L}^2$. $(\mathscr{L}^2, (S_\iota)_{\iota \in I})$ is called a set theoretic formulation and $S = \bigcap_{\iota \in I} S_\iota$ the feasibility set.[5] To construct the sets $(S_\iota)_{\iota \in I}$, one can exploit information pertaining to the error signal u as well as *a priori* knowledge about the input signal h.

3.1. Sets based on the error signal

For every signal a in $\mathscr{L}^2$ we can define a residual signal on Ω as

$$y_a = x - \bar{T} * a \tag{18}$$

It follows at once from (5) that

$$y_h = u \tag{19}$$

Therefore an estimate a of h can be constrained to yield a residual y_a consistent with every known property of u. Let $\Delta \subset \Omega$ denote the domain over which the data are observed. Then, upon imposing the boundedness property (13) of the error signal for every data point, the estimates are forced to lie in the set

$$S_b = \bigcap_{\omega \in \Delta} S_\omega, \quad \text{where } S_\omega = \{a \in \mathscr{L}^2 \mid |x(\omega) - (\bar{T} * a)(\omega)| \leqslant \gamma(\omega)\} \tag{20}$$

According to (13), the construction of these sets also requires some knowledge of the input signal, namely $\| h \|_p$.

Proposition

S_ω (and therefore S_b) is closed and convex.

Proof. Let $H\colon a \mapsto (\bar{T} * a)(\omega)$ and $K\colon a \mapsto |x(\omega) - H(a)|$ be two operators on $\mathscr{L}^2$. From the Cauchy–Schwarz inequality we get $|H(a)| = |(\bar{T} * a)(\omega)| \leqslant \| \bar{T} \| \cdot \| a \|$, whence H is bounded, since $\| H \| = \sup\{ |H(a)| \mid \| a \| = 1\} \leqslant \| \bar{T} \| < \infty$. Since H is linear and bounded, it is continuous and so is K. Therefore $S_\omega = K^{-1}]-\infty, \gamma(\omega)]$ is closed as the continuous inverse image of a closed interval. To show that S_ω is convex, it is enough to show that K is convex.

This follows from the linearity of H: let $0 < \rho < 1$ and let a and b be two points in $\mathscr{L}^2$; then $K(\rho a + (1-\rho)b) = |\rho(x(\omega) - H(a)) + (1-\rho)(x(\omega) - H(b))| \leqslant \rho K(a) + (1-\rho)K(b)$. □

3.2. *Sets based on the input signal*

There is a wide range of prior knowledge on the input signal h that can easily be associated with property sets in $\mathscr{L}^2$. First of all, h may be partially known in the spatial and/or Fourier domains. The information which falls in this category includes region of support, Fourier phase, Fourier magnitude, band limitedness as well as lower and upper bounds on amplitude, energy and entropy (some of these sets are described in Reference 26). Subjective attributes such as sparsity, smoothness, impulsiveness or maximum deviation from a prototype signal can also be included.[5,7,12]

3.3. *Synthesis of a set theoretic solution*

The solution of the set theoretic deconvolution problem is the set of all signals consistent with all available information, i.e. the feasibility set. In this subsection we describe algorithms to compute feasible signals. Given the set theoretic formulation $(\mathscr{L}^2, (S_\iota)_{\iota \in I})$, this problem is written as

$$\text{find } a \in S = \bigcap_{\iota \in I} S_\iota \tag{21}$$

In the literature various methods have been proposed to solve (21) which depend on the properties of the sets $(S_\iota)_{\iota \in I}$, the dimension of $\mathscr{L}^2$ and the cardinality of I (see Reference 5 for a detailed account). Since (21) can usually not be solved in one step, these procedures are iterative and proceed by activating the sets individually. For our Hilbertian deconvolution problem, methods based on projections are well suited. Some of these projection algorithms have already been employed in signal recovery problems.[5,26]

Let us recall that a projection of a point a onto a subset S_ι of $\mathscr{L}^2$ is a point $P_\iota(a)$ in S_ι such that

$$\| a - P_\iota(a) \| = \inf\{ \| a - b \| \mid b \in S_\iota \} \tag{22}$$

If S_ι is closed and convex, $P_\iota(a)$ exists and is unique. Moreover, in $\mathscr{L}^2$ a sequence $(a_n)_{n \geqslant 0}$ is said to converge to a strongly if $(\| a_n - a \|)_{n \geqslant 0}$ converges to zero and weakly if, for every b in $\mathscr{L}^2$, $(\langle a_n - a \mid b \rangle)_{n \geqslant 0}$ converges to zero. Strong convergence implies weak convergence, and if $\mathscr{L}^2$ has finite dimension, the converse is true.

Let us first assume that all the S_ι are closed and convex and that their intersection S is non-empty. Let $(\lambda_n)_{n \geqslant 0}$ be a sequence of numbers in $[\varepsilon_1, \varepsilon_2] \subset]0, 2[$ and let a_0 be an arbitrary signal in $\mathscr{L}^2$. Consider the sequence $(a_n)_{n \geqslant 0}$ defined by the recursion

$$a_{n+1} = a_n + \lambda_n (P_{\iota_n}(a_n) - a_n) \tag{23}$$

In other words, the update a_{n+1} lies on the line segment between the current iterate a_n and its reflection $2P_{\iota_n}(a_n) - a_n$ with respect to S_ι. Then, if one lets ι_n be the index the most remote set from a_n, $(a_n)_{n \geqslant 0}$ converges weakly to a signal in S.[27] If the cardinal of I is finite, say $I = \{1, \ldots, m\}$, this conclusion also holds for other projection schemes, e.g. when the sets are activated in a cyclic manner, i.e. $\iota_n = n$ (modulo m) + 1.[27] In practice it may happen that $S = \emptyset$, e.g. when the bounds on the perturbations are underspecified. In such instances the convergence properties of serial algorithms are not satisfactory.[5] An alternative scheme is to activate the sets in parallel at each iteration rather than in series. Such a method is described by the recursion

$$a_{n+1} = a_n + \lambda_n \left(\sum_{\iota=1}^{m} w_\iota P_\iota(a_n) - a_n \right) \tag{24}$$

where the relaxation sequence $(\lambda_n)_{n \geqslant 0}$ lies in $[\varepsilon_1, \varepsilon_2] \subset]0, 2[$ and the weights satisfy

$$\sum_{\iota=1}^{m} w_\iota = 1, \qquad (\forall \iota \in \{1, \ldots, m\}) \quad w_\iota > 0 \tag{25}$$

In Reference 28 it was shown that this parallel algorithm converges weakly to a minimizer of the functional Φ: $a \mapsto \sum_{\iota=1}^{m} w_\iota d(a, S_\iota)^2$. Hence, if $S \neq \emptyset$, it converges weakly to a feasible signal, whereas if $S = \emptyset$, it converges weakly to a weighted least squares solution of the inconsistent system of constraints. Another advantage of (24) is that it is easily implementable on concurrent processors. However, it usually does not converge as fast as (23) in the consistent case. Let us mention that strong convergence of these algorithms can also be achieved when $\mathscr{L}^2$ has infinite dimension under additional conditions on the sets (see References 28 and 29 for details). Finally, if at least one of the sets is not convex, convergence to a feasible point for most of the above algorithms will hold locally, i.e. when a_0 belongs to a prescribed region of the space.[5,11]

4. DISCUSSION

4.1. Deterministic versus probabilistic bounds

To derive a bound on the error signal of (8), we have assumed that deterministic bounds were available for the norm of the unknown component of the kernel, the input signal and the amplitude of the noise. In some problems, though, these quantities cannot be assumed to be bounded and, strictly speaking, the above set theoretic framework cannot be applied. We shall see that bounds can still be derived in such cases provided that some probabilistic information is available to describe the uncertainty.

Suppose that the data process (5) has been observed over a finite domain of N points, say $\Delta = \{\omega_1, , \ldots, \omega_N\}$, and assume that the samples $(u(\omega_i))_{1 \leqslant i \leqslant N}$ of the error signal are independent, zero-mean, unbounded random variables with known distributions. Then for every i in $\{1, \ldots, N\}$ a number $\gamma(\omega_i)$ can be found such that the interval $[0, \gamma(\omega_i)]$ will contain $|u(\omega_i)|$ with probability $1 - \varepsilon$, where ε is a small positive number. Therefore the hyperrectangle $\times_{i=1}^{N} [0, \gamma(\omega_i)]$ will contain $(|u(\omega_i)|)_{1 \leqslant i \leqslant N}$ with probability $(1 - \varepsilon)^N$. It then follows from (19) that an estimate a of h can be constrained, to within a confidence coefficient $(1 - \varepsilon)^N$, to yield residual samples $(|y_a(\omega_i)|)_{1 \leqslant i \leqslant N}$ which lie in $\times_{i=1}^{N} [0, \gamma(\omega_i)]$. Hence S_b in (20) now becomes the confidence region

$$S_b = \bigcap_{i=1}^{N} S_i, \quad \text{where } S_i = \{a \in \mathscr{L}^2 \mid |x(\omega_i) - \bar{T} * a)(\omega_i)| \leqslant \gamma(\omega_i)\} \tag{26}$$

If the distribution of $u(\omega_i)$ is not known but one of its absolute moments, say $\mathsf{E}|u(\omega_i)|^p$, is available, a probabilistic bound can still be derived via Markov's inequality, which states that the probability of $|u(\omega_i)|$ exceeding $\gamma(\omega_i)$ is at most $\varepsilon = \mathsf{E}|u(\omega_i)|^p / \gamma(\omega_i)^p$.

In order to produce accurate set theoretic estimates, one should seek to obtain the smallest possible S_b in terms of an appropriate notion of set size (set functions relevant to the evaluation of set theoretic estimates are discussed in Reference 5). If, as in Section 2, $u(\omega_i)$ were bounded,

h would belong to S_b with probability one, regardless of N. Since S_b decreases monotonically as N increases, one should therefore use as many samples as possible. This conclusion no longer holds with an unbounded error and a trade-off arises as to the choice of N. The objective is to achieve the smallest S_b for a given confidence coefficient $c = (1 - \varepsilon)^N$ (typically $0{\cdot}95 \leqslant c \leqslant 0{\cdot}99$). To maintain c constant as N increases, one must decrease ε and therefore increase the bounds $(\gamma(\omega_i))_{1 \leqslant i \leqslant N}$. This makes S_b the intersection of a larger number of bigger sets. Likewise, as N decreases, S_b is the intersection of fewer smaller sets. In general the optimal value of N will depend on the distributions of the $u(\omega_i)$ and the set function selected to evaluate the size of S_b.

4.2. Relation to set membership identification

The set theoretic approach in system identification with bounded but unknown uncertainty has been the focus of a great deal of research and is often referred to as 'set membership identification' (see References 18 and 30 for recent developments). The system identification problem is closely related to that of deconvolution. For instance, in the context of the linear model (4), the goal of the former is to estimate $\overline{T}$ whereas the goal of the latter is to estimate h. In this section we compare the proposed set theoretic deconvolution framework with set membership identification.

First of all, it is important to remark that the two frameworks share the same philosophy. Both produce estimates based on a criterion of consistency with the data and prior knowledge of bounds on the uncertain components of the model. In fact, in the identification problem the property sets arising from the boundedness constraint take the same functional form as (20). For example, for a discrete system with M parameters and N observations of the output we would get

$$S_b = \bigcap_{n=1}^{N} \{\overline{T} \in \mathbb{R}^M \mid |x(n) - (\overline{T} * h)(n)| \leqslant \gamma(n)\} \tag{27}$$

However, the two frameworks differ slightly in that set membership identification is more geared towards determining the entire feasibility set whereas the basic objective in set theoretic deconvolution is to produce a single feasible solution. In that respect the former gives a more complete picture of the set theoretic estimation problem.

In set membership identification the estimation of the feasibility set is done sequentially and the estimate is updated with the property set associated with each data sample. Recursive schemes have been devised to compute the update efficiently and make the method amenable to on-line processing. In the presence of a linear model the boundedness constraints yield the polytope (27), which can be estimated exactly.[20-22] If the complexity of a direct evaluation of the polytope is unsuitable, it can be approximated by a monotone sequence of simpler supersets such as ellipsoids or hyperparallelepipeds. A shortcoming of this approach is that even if the data record is long, the resulting approximation of the actual feasibility set may be quite loose. Hence the feasibility of a point estimate is not guaranteed. On the other hand, in the proposed projection-based set theoretic deconvolution method the property sets are activated repeatedly (in series or in parallel) to construct a sequence converging to their intersection and feasibility of a solution is therefore guaranteed. However, no particular knowledge of the actual feasibility set is required. This feature is valuable as it makes the incorporation of additional *a priori* information about the problem easy (it amounts to adding more property sets to the set theoretic formulation). On the other hand, this approach gives little information about the solution set and therefore little insight into the quality of a feasible solution, which constitutes a separate issue.[5] It should be noted that the method requires that all the sets be available simultaneously, which implies that the whole data record is needed before processing. This is a minor problem, since deconvolution is usually done off-line. Moreover, in many applications (e.g. image restoration) the whole data record is obtained at once.

In system identification the size of the data record (the output signal) is usually large compared with that of the unknown (the system kernel). However, in signal deconvolution the data record has typically the same extent as the unknown (the input signal) (in digital image-processing applications a number of unknowns $N = 1024 \times 1024$ is common). Therefore, although a loose approximation of the exact feasibility set may be acceptable for the former problem, it will not be for the latter and a solution method which guarantees a feasible signal is required.

Finally, let us remark that in set membership identification a bound on the global error signal u is supposed to be known *a priori*. In this paper a bound has been derived by breaking u up into a structural component which depends on the unknown and a measurement noise component which models external, input-independent disturbances.

4.3. *Applications*

There are countless practical inverse problems in which the underlying linear system is partially known and the data are corrupted by additive noise. Measurement noise is a common instance in any data-collecting process and it need not be further discussed. Inaccurately modelled linear systems arise in many situations. For instance, in geophysics the characteristics of the system, in this case the earth, are partially known. In communications, linear channels are often subject to uncontrollable factors which make their exact modelling impossible. In image restoration the point spread function of the imaging system is often subject to unknown perturbations, e.g. the random fluctuations of the index of refraction of the light propagation medium in atmospheric imaging or the random scattering of the quanta in X-ray imaging.

The application of the proposed framework to such problems is contingent on the availability of a bounding function for the error signal. If the sole source of uncertainty is the noise (i.e. $\tilde{T} = 0$ in (6)), only the bound β of (9) is needed. If the model of the system is uncertain, then the $\mathscr{L}^p$-norm of the input signal h and a bound on the $\mathscr{L}^{p'}$-norm of the disturbance $\tilde{T}_\omega$ are also needed for each observed data sample $x(\omega)$. Of course, such information is application-dependent. Fortunately, in a wide range of problems it will be available. As an example, let us consider image deblurring. In this case h represents the original image, whose norm may be available through the *a priori* knowledge or may be inferred from the degraded image x. If the blur is due to an out-of-focus lens, the focus of an object is a function of the distance to the lens, which varies widely in many common photographic scenes. With information about the scene it is possible to bound the variation of the blur. If the blur is caused by uniform motion, bounds on its extent can be made from knowledge about the possible speed of objects or the relative motion of the camera. For example, a television system set to monitor a highway can expect cars to be travelling within a few km h^{-1} of the speed limit. With this information, bounds on any $\mathscr{L}^p$-norm of $\tilde{T}_\omega$ can be computed.

It is important to note that the amount of information required to construct the function γ bounding the error in (13) is usually less than is required by classical methods. For instance, in the deconvolution approaches for uncertain models based on Wiener filtering[31] or Bayesian analysis,[32] a statistical description of the input signal h, the disturbance $\tilde{T}$ and the noise v is needed. Statistical knowledge of $\tilde{T}$ and v is also required in the set theoretic approach of

Reference 9. In addition, in order to be computationally efficient, these methods must be implemented in the frequency domain via the fast Fourier transform, which imposes that the uncertainty be stationary. This assumption can be relaxed in the bounded error framework, thus providing a more general and realistic modelling of the uncertainty, since in most practical instances the unknown fluctuations of the system are shift-variant.

5. A DISCRETE SIGNAL RECOVERY APPLICATION

The purpose of this section is to provide a detailed example of application of the above analysis to a concrete deconvolution problem, namely the restoration of a blurred and noise-corrupted one-dimensional discrete signal. The length of the signals is set to $N = 512$ and the Hilbertian solution space is $\mathbb{R}^N$ endowed with the Euclidean distance. From (16) the nth sample of the degraded signal x for a causal blur is given by

$$x(n) = \sum_{m=0}^{n} \bar{T}(n-m)h(m) + \sum_{m=0}^{n} \tilde{T}(n,m)h(m) + v(n) \tag{28}$$

The original signal h is depicted in Figure 1. It is assumed that h is non-negative with maximum value $\| h \|_\infty$. The blur is causal, of length $L = 16$, and its known component is rectangular and shift-invariant.

$$(\forall n \in \{0, \ldots, N-1\}) \quad \bar{T}(n) = \begin{cases} 1/L & \text{if } 0 \leqslant n < L-1 \\ 0 & \text{otherwise} \end{cases} \tag{29}$$

Physically $\bar{T}$ can be regarded as a one-dimensional linear motion blur. The unknown component $\tilde{T}$ of the blur is shift-variant with the same region of support as $\bar{T}$ and, for every n, $\| \tilde{T}_n \|_1 = \alpha_\infty < \infty$. The noise samples $(v(n))_{0 \leqslant n \leqslant N-1}$ are uniformly distributed in the interval $[-\beta, \beta]$, where $\beta = 0\cdot1$. From this information the bound on the error signal samples

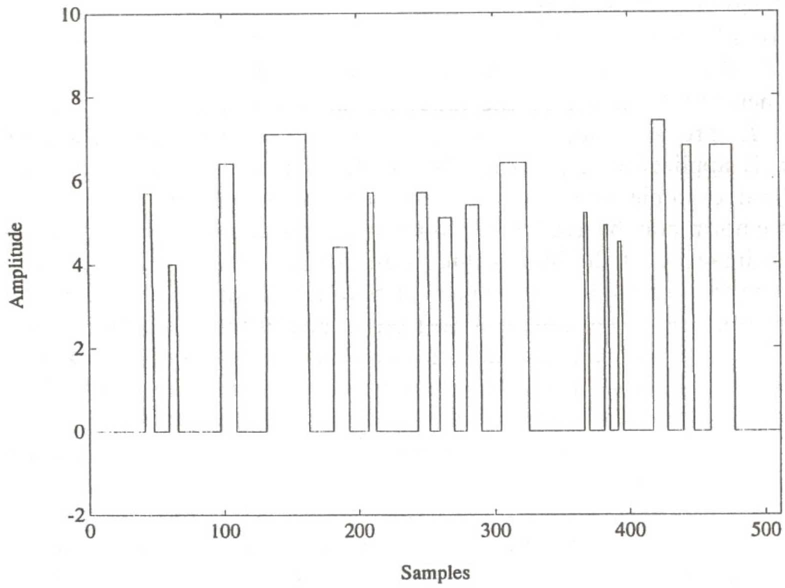

Figure 1. Original signal

given by (13) is $\gamma(n) = \gamma = \alpha_\infty \,.\, \| h \|_\infty + \beta$. The closed and convex property sets of (20) are

$$S_n = \{a \in \mathbb{R}^N \mid \, | x(n) - (\bar{T} * a)(n) | \leqslant \gamma\} \quad \text{for } 0 \leqslant n \leqslant N-1 \tag{30}$$

Now let $\bar{T}_n$ be the reflected and n-shifted version of $\bar{T}$ so that $(\bar{T} * a)(n) = \langle \bar{T}_n \mid a \rangle$. We can then write $S_n = \{a \in \mathbb{R}^N \mid x(n) - \gamma \leqslant \langle \bar{T}_n \mid a \rangle \leqslant x(n) + \gamma\}$, which is simply a closed hyperslab delimited by two hyperplanes. Hence the projection onto S_n of a vector not in S_n is given by the projection onto the nearest hyperplane. Since the projection of a point a' onto the hyperplane $\{a \in \mathbb{R}^k \mid \langle a \mid b \rangle = \delta\}$ is $a' + [(\delta - \langle a' \mid b \rangle)/\| b \|^2]\, b$, the projection operator onto S_n is given by

$$P_n(a') = \begin{cases} a' + [(x(n) + \gamma - \langle \bar{T}_n' \mid a' \rangle)/\| \bar{T}_n' \|^2]\, \bar{T}_n' & \text{if } \langle \bar{T}_n' \mid a' \rangle > x(n) + \gamma \\ a' + [(x(n) - \gamma - \langle \bar{T}_n' \mid a' \rangle)/\| \bar{T}_n' \|^2]\, \bar{T}_n' & \text{if } \langle \bar{T}_n' \mid a' \rangle < x(n) - \gamma \\ a' & \text{otherwise} \end{cases} \tag{31}$$

The last closed and convex set in the set theoretic formulation is based on the above information on the amplitude of the input signal, namely

$$S_N = \{a \in \mathbb{R}^N \mid (\forall n \in \{0, \ldots, N-1\}) \quad 0 \leqslant a_n \leqslant \| h \|_\infty\} \tag{32}$$

The projection operator onto this set is easily seen to be

$$(\forall n \in \{0, \ldots, N-1\}) \quad [P_N(a')]_n = \begin{cases} 0 & \text{if } a_n' < 0 \\ \| h \|_\infty & \text{if } a_n' > \| h \|_\infty \\ a_n' & \text{otherwise} \end{cases} \tag{33}$$

A signal in the feasibility set $S = \bigcap_{\iota=0}^{N} S_\iota$ is obtained via algorithm (23), activated under cyclic control with $\lambda_n = 1$ and the degraded signal as an initial estimate. Convergence of $(a_{n(N+1)})_{n \geqslant 0}$ was typically achieved within 10–30 iterations. First we simulate an instance when the blur is known exactly, i.e. $\alpha_\infty = 0$. The degraded signal is shown in Figure 2 and the restoration in

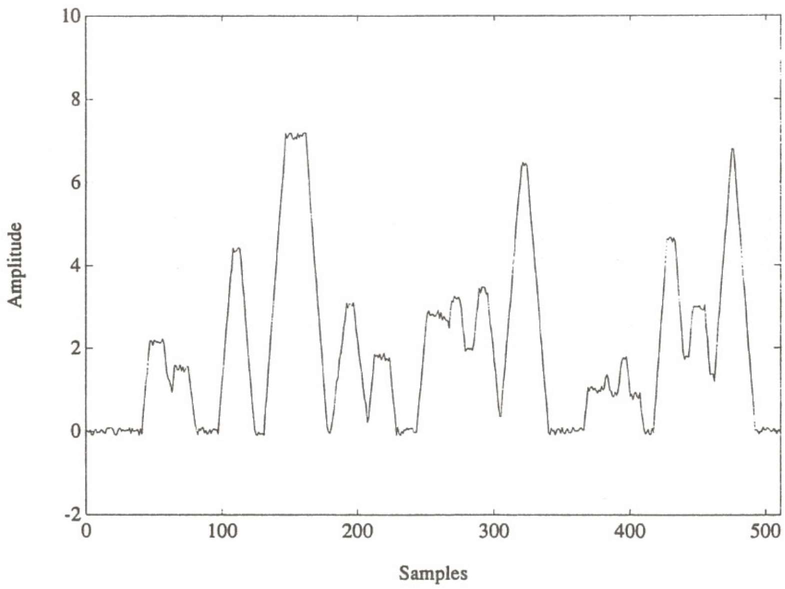

Figure 2. Degraded signal — known blur

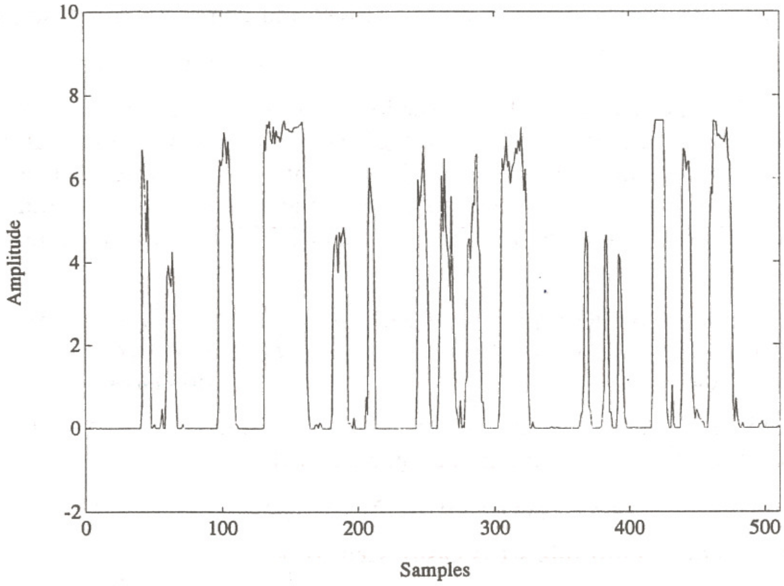

Figure 3. Restored signal — known blur

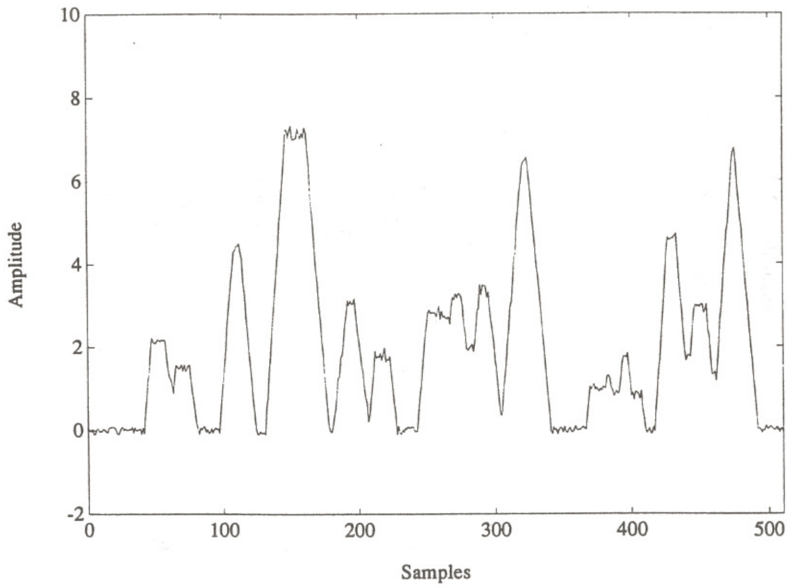

Figure 4. Degraded signal — perturbed blur

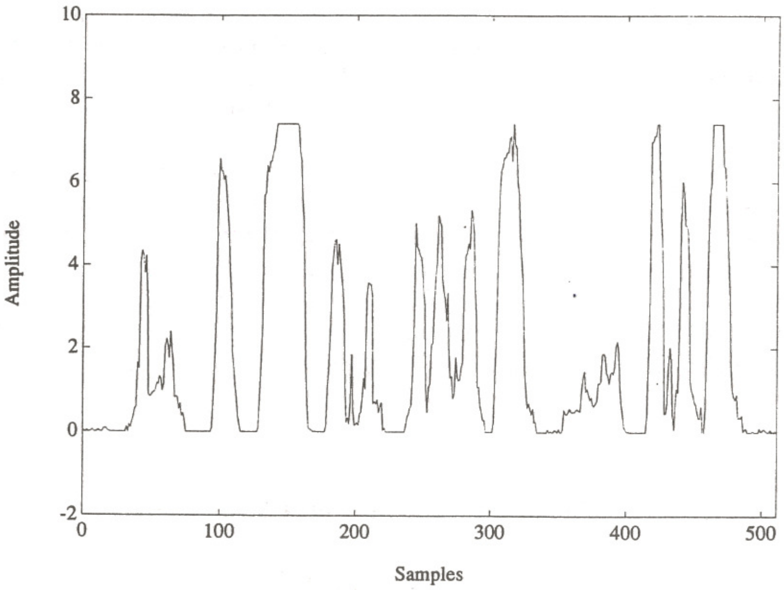

Figure 5. Restored signal — perturbed blur

Figure 3. It is seen that a large amount of information has been recovered. Then we introduce shift-varying perturbations in the blurring kernel, with $\alpha_\infty = 0{\cdot}04$, to obtain the degraded signal of Figure 4, whose restoration is shown in Figure 5. Clearly the added uncertainty has increased the bound γ and therefore the feasibility set, which results in a poorer restoration.

Besides the knowledge of the component $\overline{T}$ of the kernel, the information used in these set theoretic restorations is limited to upper and lower bounds on the input signal and the noise and an upper bound on the l^1-norm of the disturbances of the kernel. It is noted that no statistical assumption has been made and that the only conventional deconvolution method that could be implemented with such little information would be inverse filtering, which is known to give unacceptable results.[2]

6. CONCLUSIONS

In this paper we have considered the general signal deconvolution problem in the presence of unknown system kernel disturbances and measurement noise. From the prior knowledge of bounds on these two sources of uncertainty, amplitude bounds on the error in the data formation mechanism were derived. By constraining the estimation residual to be consistent with these error bounds, sets were constructed in the solution space. Upon adding these sets to those based on other *a priori* knowledge, one obtains a collection of sets whose intersection is the set of feasible signals for the deconvolution problem. Projection methods were described as a computationally efficient scheme to generate feasible solutions.

Such a framework for deconvolution is of particular interest whenever a probabilistic description of uncertainty is either not appropriate or not available and bounds on the unknown components of the data model are known. While the set theoretic formulation is conceptually easy to formulate, the most important part of obtaining a meaningful solution is the determination of accurate bounds for the error signal. If the bounds are too conservative, the intersection of the property sets will be large and the quality of a solution will be poor. If the bounds are too small, the constraints on the residual samples will be inconsistent and the feasibility set may be empty. Hence the task is to derive from the *a priori* knowledge the tightest possible bounds that guarantee that the true solution is within the feasibility set.

The method offers great flexibility with regard to the incorporation of *a priori* knowledge and, unlike existing methods, can handle shift-variant perturbations on the kernel of the system. The ability of the method to use only the knowledge that is available without requiring questionable assumptions or estimations is another advantage. This paper has also discussed the disadvantages of the method which may make it unsuitable for certain types of problems. While no method is optimal for all problems, the set theoretic approach definitely constitutes a reliable and efficient tool that will prove indispensable in a wide range of applications.

ACKNOWLEDGEMENTS

The work of Patrick L. Combettes was supported by a PSC-CUNY grant.

REFERENCES

1. Roubine, E., *Distributions — Signal*, 2nd edn, Eyrolles, Paris, 1990.
2. Andrews, H. C., and B. R. Hunt, *Digital Image Restoration*, Prentice-Hall, Englewood Cliffs, NJ, 1977.
3. Mendel, J. M., *Optimal Seismic Deconvolution: An Estimation-Based Approach*, Academic, New York, 1983.
4. Combettes, P. L., 'Convex set theoretic image recovery: history, current status, and new directions', *J. Visual Commun. Image Represent.*, **3**, 307–315 December (1992).
5. Combettes, P. L., 'The foundations of set theoretic estimation', *Proc. IEEE*, **81**, 182–208 (1993).
6. Trussell, H. J., and M. R. Civanlar, 'The feasible solution in signal restoration', *IEEE Trans. Acoust. Speech, Signal Process.*, **ASSP-32**, 201–212 (1984).
7. Civanlar, M. R. and H. J. Trussell, 'Digital signal restoration using fuzzy sets', *IEEE Trans. Acoust., Speech, Signal Process.*, **ASSP-34**, 919–936 (1986).
8. Combettes, P. L., and H. J. Trussell, 'The use of noise properties in set theoretic estimation', *IEEE Trans. Signal Process.*, **SP-39**, 1630–1641 (1991).
9. Combettes, P. L. and H. J. Trussell, 'Methods for digital restoration of signals degraded by a stochastic impulse response', *IEEE Trans. Acoust., Speech, Signal Process.*, **ASSP-37**, 393–401 (1989).
10. Chin, R. T., C. L. Yeh and W. S. Olson, 'Restoration of multichannel microwave radiometric images', *IEEE Trans. Pattern Anal. Machine Intell.*, **PAMI-7**, 475–484 (1985).
11. Combettes, P. L., and H. J. Trussell, 'Method of successive projections for finding a common point of sets in metric space', *J. Optim. Theory App.*, **67**, 487–507 (1990).
12. Sezan, M. I., and H. J. Trussell, 'Prototype image constraints for set-theoretic image restoration', *IEEE Trans. Signal Process.*, **SP-39**, 2275–2285 (1991).
13. Schweppe, F. C., 'Recursive state estimation: unknown but bounded errors and system inputs', *IEEE Trans. Automatic Control*, **AC-13**, 22–28 (1968).
14. Deller Jr., J. R., and T. C. Luk, 'Linear prediction analysis of speech based on set-membership theory', *Comput. Speech Lang.*, **3**, 301–327 (1989).
15. Huang, Y. F., 'A Recursive estimation algorithm using selective updating for spectral analysis and adaptive signal processing', *IEEE Trans. Acoust., Speech, Signal Process.*, **ASSP-34**, 1331–1334 (1986).
16. Kosut, R. L., 'Adaptive control via parameter set estimation', *Int. J. Adapt. Control Signal Process.*, **2**, 371–399 (1988).
17. Norton, J. P., 'Identification and application of bounded-parameter models', *Automatica*, **23**, 497–507 (1987).
18. Walter, E. (guest ed.), 'Parameter identifications with error bounds', *Math. Comput. Simul.*, **32**, (5–6) (1990).
19. Tempo, R., 'Robust estimation and filtering in the presence of bounded noise', *IEEE Trans. Automatic Control*, **AC-33**, 864–867 (1988).
20. Broman, V., and M. J. Shensa, 'A compact algorithm for the intersection and approximation of *N*-dimensional polytopes', *Math. Comput. Simul.*, **32**, 469–480 (1990).
21. Mo, S. H., and J. P. Norton, 'Fast and robust algorithm to compute exact polytope parameter bounds', *Math. Comput. Simul.*, **32**, 481–493 (1990).
22. Piet-Lahanier, H., and E. Walter, 'Exact recursive characterization of feasible parameter sets in the linear case', *Math. Comput. Simul.*, **32**, 495–504 (1990).

23. Katsaggelos, A. K., J. Biemond, R. W. Schafer and R. M. Mersereau, 'A regularized iterative image restoration algorithm', *IEEE Trans. Signal Process.*, **SP-39**, 914–929 (1991).
24. Ash, R. B., *Real Analysis and Probability*, Academic, New York, 1972.
25. Schwartz, L., *Analyse III — Calcul Intégral*, Hermann, Paris, 1993.
26. Youla, D. C. and H. Webb, 'Image restoration by the method of convex projections: Part 1 — Theory', *IEEE Trans. Med. Imag.*, **MI-1**, 81–94 (1982).
27. Gubin, L. G., B. T. Polyak and E. V. Raik, 'The method of projections for finding the common point of convex sets', *USSR Comput. Math. Math. Phys.*, **7**, (6), 1–24 (1967),
28. Combettes, P. L. and Puh, H., 'Iterations of parallel convex projections in Hilbert spaces', *Numer. Funct. Anal. Optim.*, **15**, 225–243 (1994).
29. Ottavy, N., 'Strong convergence of projection-like methods in Hilbert spaces', *J. Optim. Theory Appl.*, **56**, 433–461 (1988).
30. Deller Jr., J. R., M. Nayeri and S. F. Odeh, 'Least-square identification with error bounds for real-time signal processing and control', *Proc. IEEE*, **81**, 815–849 (1993).
31. Slepian, D., 'Linear least-squares filtering of distorted images', *J. Opt. Soc. Am.*, **57**, 918–922 (1967).
32. Guan, L., and R. K. Ward, 'Restoration of randomly blurred images via the maximum *a posteriori* criterion', *IEEE Trans. on Image Process.*, **IP-1**, 256–262 (1992).