Set-Membership Filtering: A Viable Tool for Non-Linear Adaptive Signal Processing

Shirish Nagaraj, Sridhar Gollamudi, Samir Kapoor and Yih-Fang Huang

Laboratory for Image and Signal Analysis
Department of Electrical Engineering
University of Notre Dame, Notre Dame, IN 46556
Email: huang.2@nd.edu

Abstract— This paper formulates, and provides a solution to, a filtering problem based on certain deterministic constraint imposed on the error sequence. The formulation is called *Set-Membership Filtering (SMF)* and is a generalization of Set-Membership Identification (SMI) for system identification. A Set-Membership Decision Feedback Equalizer (SM-DFE) is proposed as an example of its application to non-linear systems. This paper also establishes connections of SM-DFE with the Minimum Mean Square Error DFE (MMSE-DFE). Recursive solutions to the general SMF problem are also presented.

I. INTRODUCTION

Set-membership signal processing and control is a growing field and has seen tremendous increase in research activity in the last decade, see, *e.g.,* [1, 2]. In control systems, for example, design of robust controllers based on identification of a set of feasible plants is now a well established method. In the area of system identification, such methods are termed *Set-Membership Identification* (SMI) procedures. The common thread that binds together all these techniques is the fact that they utilize certain set-theoretic assumption in the estimation procedure because of which, one obtains a set of possible solutions or estimates which are consistent with that assumption. This is in contrast with conventional estimation paradigms which yield point estimates. For instance, one of the most common and often realistic assumption made in SMI is that the corrupting noise at the output is bounded in magnitude for all time.

The restriction on the model dependence of SMI can be removed by posing the problem in a general filtering framework. The mapping between the inputs and the desired outputs is done by designing a filter which achieves a pre-set performance specification. This sets the stage for the so-called *Set-Membership Filtering* (SMF) problem formulation, developed in [3], wherein a specification on the maximum allowable filter output error is imposed in the form of a magnitude bound. Moreover, in many situations, linear-in-input system model descriptions are not accurate and hence are unacceptable. Examples where such non-linearities exist include many communication systems, where non-linear behavior occurs due to signal companding and amplifier saturation effects, among others [4]. Equalizers and estimators based on Volterra series expansions [5] have been developed in literature [4]. Another example of a non-linear system is the celebrated Decision Feedback Equalizer (DFE), where the hard-limited outputs of the equalizer are used in the feedback path, making the overall system non-linear [6]. This paper demonstrates that SMF is a viable tool for estimation of certain non-linear systems, which are linear-in-parameters, such as a Volterra Filter (VF) and under certain assumptions, a DFE. Adaptive solution strategies to the SMF problem have been presented [7]. It is shown here that the algorithms presented are equipped with many features lacking in traditional Recursive Least Squares (RLS) and Least Mean Squares (LMS), for example, a selective updating capability and enhanced tracking properties.

II. SMF FOR NON-LINEAR SYSTEMS

Assume that the input-desired output pairs come from a certain data space, $\mathcal{H} \subset \mathbf{C}^n \times \mathbf{C}$, where $\mathbf{C}^n$ denotes the n-dimensional complex Euclidean space. The problem at hand is the following: *Given $\mathcal{H}$, design a filter, $\mathcal{F}_\theta : \mathcal{H} \to \mathbf{C}$, which maps the pairs $(\mathbf{x}, d)$ to the output error,*

$$e \triangleq \mathcal{F}_\theta(\mathbf{x}, d) = d - \theta^T \mathbf{x} \tag{1}$$

where $\theta \in \mathbf{C}^n$. The objective of SMF is to choose θ such that

$$|e| \leq \gamma \qquad \forall (\mathbf{x}, d) \in \mathcal{H} \tag{2}$$

for a designer-specified positive real number γ. In other words, find the parameter vector, θ, which meets the specification for all input-desired output pairs in $\mathcal{H}$. Conventional performance criteria minimize the squared error, either in a deterministic (least-square) or stochastic (mean-square) settings. For further details on this formulation, refer [3, 8].

The class of filters modeled by the above includes FIR and IIR filters that are commonly used in signal processing and communication systems. It also includes the autoregressive (AR) model in which $\mathbf{x}$ is the regressor vector of the output sequence, as well as several nonlinear systems. Specifically, $\mathbf{x}$ may contain nonlinear functions of past filter outputs, as is the case in a DFE, or might contain linear, quadratic, and other higher order terms of the Volterra series expansion [5] as in a Volterra filter. In this case, the elements of the parameter vector, θ, are the Volterra series kernels [9].

It can be seen that SMF is directly applicable to many non-linear systems. The intuitive justification of the bounded error specification for Volterra filters is that the non-linear (higher order) terms neglected in the VF should not be too large. The only assumption made in SMF is that the filter is of the form (1), *i.e.*, linear-in-parameters. As an example, the model for a quadratic VF, with memory order N, is of the form:

$$\begin{aligned} y(k) &= h_0 + \sum_{i_1=0}^{N} h_1(i_1)u(k-i_1) + \\ & \sum_{i_1,i_2=0}^{N} h_2(i_1,i_2)u(k-i_1)u(k-i_2) + \epsilon(k) \end{aligned}$$

where $y(\cdot)$ is the output of the filter, $u(\cdot)$ is the input, and $\epsilon(\cdot)$ is the error. The coefficients, $h_1(\cdot)$ and $h_2(\cdot,\cdot)$ are the Volterra kernels and are the parameters to be identified. It is seen that this can be cast in the form of a linear-in-parameter filter as in (1).

The bounded error specification in turn leads to a set of parameters satisfying the criteria if it is not too stringent. Therefore, the objective of SMF is to estimate the so-called *feasibility set*, given by:

$$\Theta \triangleq \bigcap_{(\mathbf{x},d)\in\mathcal{H}} \{\theta \in \mathbf{C}^n : |d - \theta^T\mathbf{x}|^2 \leq \gamma^2\} \tag{3}$$

Choice of γ critically determines $\mathbf{\Theta}(n,\gamma)$, which may turn out to be an empty set if γ is inappropriately chosen [3]. Certain sufficient and necessary conditions for the existence of a non-empty feasible set for the case of linear channel equalization have been derived in [8].

III. A SET-MEMBERSHIP DECISION FEEDBACK EQUALIZER

The existence of the MMSE-DFE is guaranteed in all situations. This is not the case with SM-DFE, wherein one can get an empty feasibility set if the performance requirement is too stringent. Therefore, conditions for the non-emptiness of the same are discussed first.

An equivalent complex baseband discrete-time channel model [6] is used. Let $\{a_n\}_{n=-\infty}^{\infty}$ be the transmitted symbol sequence and $\{\nu_n\}_{n=-\infty}^{\infty}$ be the additive noise at the channel output. The output of the channel, $\{u_n\}_{n=-\infty}^{\infty}$, is given by

$$u_n = \sum_{i=-D}^{D} c_i a_{n-i} + \nu_n \tag{4}$$

where $\{c_i\}_{i=-D}^{D}$ is the (finite length) channel impulse response.

Consider a decision feedback equalizer whose output is given by

$$z_n = \mathbf{f}^T\mathbf{x}_n - \mathbf{b}^T\hat{\mathbf{m}}_n \tag{5}$$

where $\mathbf{f} = [f_{-N_f}, \ldots, f_{N_f}]^T$ is the vector of coefficients of the forward filter and $\mathbf{b} = [b_{-N_b}, \ldots, b_{N_b}]^T$ is the (strictly causal) feedback filter tap weight vector. Further, $\mathbf{x}_n$ is the input vector to the equalizer, *i.e.*, the regressor vector comprising of the channel outputs $\{u_k,\ k = n+N_f, \ldots, n-N_f\}$ and $\hat{\mathbf{m}}_n$ is the regressor vector of the previously decoded bits, $\{\hat{a}_k,\ k = n-1, \ldots, n-N_b\}$. These are given by

$$\mathbf{x}_n = C^T\mathbf{s}_n + \mathbf{w}_n \tag{6}$$

where

$$\begin{aligned} \mathbf{s}_n &= [a_{n+K}, \ldots, a_{n-K}]^T; \qquad K = D + N_f, \\ \mathbf{w}_n &= [\nu_{n+N_f}, \ldots, \nu_{n-N_f}]^T, \end{aligned}$$

and C is the channel convolution matrix [8]. Also,

$$\hat{\mathbf{m}}_n = [\hat{a}_{n-1}, \ldots, \hat{a}_{n-N_b}]^T \tag{7}$$

Let the transmitted symbols come from a constellation $\mathcal{C}$. In order to guarantee the output error to be bounded by some $\gamma > 0$, we need to define the design space for which the specification will hold. In the case of equalization, it is intuitively appealing to fix the upper bound on the output error at $d_{min}/2$, where d_{min} is the minimum Euclidean distance of two distinct symbols from the constellation $\mathcal{C}$. By such a design, the SM-DFE will have no decoding errors whenever the data pairs come from the design space $\mathcal{H}$. The design space is completely specified by the Cartesian product space of the regressor vector of transmitted symbols $\mathbf{x}_n$ and the additive noise $\mathbf{w}_n$ for all n.

Define the design space to be all the possible pairs of $\mathbf{s}_n$ and $\mathbf{w}_n$ where $\mathbf{s}_n$ is taken from $\mathcal{C}^{2K+1}$ and $\mathbf{w}_n \in \mathcal{V}^{2N_f+1}$ and

$$\mathcal{V} \triangleq \{v \in \mathbf{C} : |v|^2 \leq \gamma_\nu^2\} \tag{8}$$

for some noise bound $\gamma_\nu > 0$. In other words, designing $\mathcal{H}$ involves fixing the symbol constellation and the value of γ_ν. It can be seen from Equation (6) that

$\mathbf{x}$ depends only on $\mathbf{s}$ and $\mathbf{w}$, whereas $\hat{\mathbf{m}}$ not only depends on the design space, but is a non-linear function of the parameter vectors, $\mathbf{f}$ and $\mathbf{b}$ and the design space (which consists of the input and noise vectors). However, to make the analysis tractable, we impose the so-called *"DFE assumption"* [6]; namely that the previous decoded bits are all correct, which implies that $\hat{\mathbf{m}} = P^T\ \mathbf{s}$ where

$$P^T = \left[\begin{array}{c|c|c} \mathbf{0}_{N_b \times (K+1)} & \mathbf{I}_{N_b \times N_b} & \mathbf{0}_{N_b \times (K-N_b)} \end{array}\right]$$

Define the set of all forward and feedback equalizer weight vectors of length $(2N_f+1)$ and N_b respectively, that satisfy the specification as the *feasible SM-DFE set* $\mathbf{\Theta}(n,\gamma)$,

$$\mathbf{\Theta}(n,\gamma) \triangleq \bigcap_{(\mathbf{s},\mathbf{w})\in\mathcal{H}} \{(\mathbf{f},\mathbf{b}) : |a - (\mathbf{f}^T\mathbf{x} - \mathbf{b}^T\hat{\mathbf{m}})|^2 \leq \gamma^2\} \tag{9}$$

where $n \triangleq N_f + 1 + N_b$ and $\hat{\mathbf{m}} = P^T\ \mathbf{s}$.

Recall that the desired output of the equalizer can be expressed as $a = \mathbf{e}_0^T\mathbf{s}$, where $\mathbf{e}_0 = [0,\ldots,1,\ldots,0]^T$ is the unit vector of dimension $(2K+1)$ with a '1' in the $(K+1)^{st}$ position. For simplicity, assume that the input symbols and noise come from a zero mean independent and identically distributed (*i.i.d.*) sequence (hence white), and that the noise and input symbols are uncorrelated.

Theorem 1 (Existence of SM-DFE) *Consider a linear channel model (4) and a DFE equalizer as described above. The following are true:*

1. $\mathbf{\Theta}(n,\gamma)$ *is non-empty if the channel satisfies*

$$\gamma_{\overline{a}} \sum_{i=-K}^{K} |\epsilon_i^{(1)}| + \gamma_\nu \sum_{i=-N_f}^{N_f} |f_i^{(1)}| \ \leq\ \gamma \tag{10}$$

where $\epsilon^{(1)} = (C\mathbf{f}^{(1)} - P\mathbf{b}^{(1)} - \mathbf{e}_0)$, $\gamma_{\overline{a}}$ *is the maximum amplitude in the constellation* $\mathcal{C}$, $\mathbf{f}^{(1)} = \left(C^H(C - P\ H) + \frac{\gamma_\nu^2}{\gamma_{\overline{a}}^2}I\right)^{-1} C^H\mathbf{e}_0$, *and* $\mathbf{b}^{(1)} = H\ \mathbf{f}^{(1)}$, *where* H *consists of the* $(K+2)$ *to* $(K+N_b+1)$ *rows of* C. *If (10) is satisfied, then* $(\mathbf{f}^{(1)},\mathbf{b}^{(1)}) \in \mathbf{\Theta}(n,\gamma)$.

2. *If* $\mathbf{\Theta}(n,\gamma)$ *is non-empty, then*

$$\|\epsilon^{(2)}\|^2\sigma_a^2 + \|\mathbf{f}^{(2)}\|^2\gamma_\nu^2 \leq \gamma^2 \tag{11}$$

where σ_a *is the variance of the symbols in the constellation* $\mathcal{C}$, $\epsilon^{(2)} = (C\mathbf{f}^{(2)} - P\mathbf{b}^{(2)} - \mathbf{e}_0)$, $\mathbf{f}^{(2)} = \left(C^H(C - P\ H) + \frac{\gamma_\nu^2}{\sigma_a^2}I\right)^{-1} C^H\mathbf{e}_0$, *and* $\mathbf{b}^{(2)} = H\ \mathbf{f}^{(2)}$, *where* H *is as before.*

If the sufficient condition (10) is met, then $(\mathbf{f}^{(1)},\mathbf{b}^{(1)})$ can be used as a member of the feasibility set and hence, corresponds to a valid SM-DFE filter. The structure of this filter is very similar to that of the MMSE-DFE [6, 10] in that the feedback filter cancels part of the combined *post-cursor ISI* effects [6] of the channel and the forward filter. In fact, by setting $N_b = K$, it can be shown that *all* the post-cursor ISI terms are cancelled by the filter $\mathbf{b}^{(1)}$. Hence, we assume $N_b = K$.

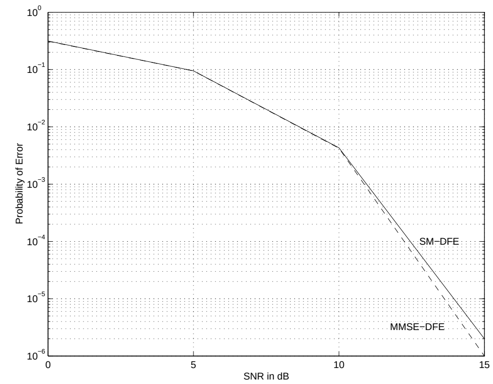

Figure 1: Probability of Error v/s SNR for SM-DFE (solid line) and MMSE-DFE (dashed line)

It turns out that the MMSE-DFE is a special case of SM-DFE in certain situations. This provides an interesting insight into SM-DFE.

Theorem 2 *Consider the following design space,* $\mathcal{H}_* \triangleq \mathcal{C}_*^{(2K+1)} \times \mathcal{V}_*^{(2N_f+1)}$, *where* $\mathcal{C}_*$ *is any symbol constellation and*

$$\mathcal{V}_* \triangleq \{v \in \mathbf{C} : |v|^2 \leq \gamma_{\nu*}^2; \quad \gamma_{\nu*}^2 = \frac{\gamma_{\overline{a}}^2\sigma_\nu^2}{\sigma_a^2}\}$$

where σ_ν^2 *is the variance of the additive noise. Also,* $\gamma_{\overline{a}}$ *is the maximum amplitude and* σ_a^2 *is the variance of the symbols in* $\mathcal{C}_*$.

If the sufficient condition (10) holds for $\mathcal{H}_*$, *then* $(\mathbf{f}^{(1)},\mathbf{b}^{(1)})$, *from (10), is the MMSE-DFE and hence, the MMSE-DFE belongs to the feasibility set,* $\mathbf{\Theta}(n,\gamma)$ *(for* $\mathcal{H}_*$*).*

To illustrate the method, consider a microwave radio channel, obtained from actual field measurements [11]. The Probability of error curves for different values of SNR are plotted for both the SM and MMSE-DFEs in Figure (1). The curve for SM-DFE corresponds to the filter $(\mathbf{f}^{(1)},\mathbf{b}^{(1)})$, as given in (10). It was observed that the performance of the MMSE-DFE was very similar, and almost identical, to that of the MMSE-DFE in the range of 0-20 dB of SNR.

In a lot of practical situations, however, the design space $(\mathcal{H})$ is not known a priori or might be time-varying. In both these cases, it is not possible to design the SM filter off-line. Such a scenario requires the use of recursive schemes to implement the SM filter.

IV. ADAPTIVE SOLUTION OPTIONS

A class of recursive, or on-line, estimation schemes, termed SMART (Set-Membership Adaptive Recursive Techniques) to estimate a parameter in the *feasibility set* have been developed. A typical example of SMART are the Optimal Bounding Ellipsoid (OBE) algorithms [1, 2, 7]. The main properties of such algorithms include a discerning update capability, an "intelligent" and non-linear weighting strategy for the incoming data, and improved estimation error for both stationary and non-stationary environments compared to standards algorithms like RLS. The selective update capability is especially useful: since the data is generally not being processed for more than 80% of the time, it allows the designer to use higher order Volterra filter models without incurring a significantly increased computational burden. It should be emphasized that the convergence properties of these algorithms is comparable to RLS in a stationary input scenario. Many of these algorithms have a proven convergence under certain persistence of excitation conditions on the input sequence [12]. These features make recursive SMF solutions a promising tool for estimation of non-linear systems. Thus, when the design space is unknown, or is time-varying, these algorithms can be used to adaptively converge to a point in the feasibility set. A novel OBE algorithm has been developed recently [7] which shares all the attractive features offered by the family of OBE algorithms, while possessing many other properties like excellent and robust tracking for fast-time varying systems. Moreover, it has been shown that the central point estimates obtained via this algorithm can be derived as a solution of a certain constrained least-squares problem [7].

Another possible adaptive solution is an algorithm called SM-NLMS [3], which is a variable gain normalized LMS algorithm. This algorithm, though slower in convergence, is much faster than conventional LMS while being computationally more efficient than LMS.

V. CONCLUSIONS

It was shown in this paper that the concept of Set Membership Filtering can be used for estimating parameters in several non-linear systems that are linear-in-parameters. As an example, a set membership DFE was introduced. Sufficient and necessary conditions were derived for the existence of this filter. Simulations demonstrate the excellent performance of this DFE. Connections with the well-known MMSE-DFE were also explored. For adaptive implementations, it was shown that the family of algorithms, termed SMART, could be used fruitfully to estimate the parameters of non-linear systems.

VI. ACKNOWLEDGMENTS

This work was supported, in part, by the National Science Foundation under Grant MIP-9705173, in part by the Center for Applied Mathematics, University of Notre Dame, and in part, by the Tellabs Research Center, Mishawaka, IN.

References

[1] E. Fogel and Y. F. Huang, "On the value of Information in System Identification - Bounded Noise Case," *Automatica*, vol. 18, No. 2, pp. 229-238, March 1982.

[2] J.R. Deller, Jr., M. Nayeri, and S.F. Odeh, "Least-squares identification with error bounds for real-time signal processing and control," *Proc. of the IEEE*, vol. 81, No. 6, pp. 815-849, June, 1993.

[3] S. Gollamudi, S. Nagaraj and Y. F. Huang, "SMART: A Toolbox for Set-Membership Filtering," *Proc. European Conference on Circuit Theory and Design (ECCTD)*, Budapest, Hungary, 1997.

[4] J. Lee and V. J. Mathews, "A fast recursive least-squares adaptive second-order Volterra filter and its performance analysis," *IEEE Transactions on Signal Processing*, vol. 41, no. 3, pp. 1087-1102, Mar. 1993.

[5] M. Schetzen, *The Volterra and Wiener theories of non-linear systems*, New York: Wiley, 1980.

[6] J.G. Proakis, *Digital Communications, 2nd ed.* New York: McGraw-Hill, 1989.

[7] S. Nagaraj, S. Gollamudi, S. Kapoor and Y.F. Huang, "Bounded error estimation: Set-theoretic and least-squares formulations," *Proc. 1997 Conf. Information Sciences and Systems*, The Johns Hopkins University, Baltimore, MD, 1997.

[8] S. Gollamudi, S. Kapoor, S. Nagaraj and Y.F. Huang, "Set-membership adaptive equalization and an updator-shared implementation for multiple channel communications systems," Submitted for publication in the *IEEE Trans. Signal Proc.*, Oct., 1996.

[9] I. W. Sandberg, "On Volterra expansions for time-varying non-linear systems," *IEEE Transactions on Circuits and Systems*, vol. CAS-30, no. 2, pp. 61-67, Feb. 1983.

[10] N. Al-Dhahir and J. M. Cioffi, "MMSE Decision-Feedback Equalizers: Finite Length Results," *IEEE Transactions on Information Theory*, vo. 41, no. 4, July 1995.

[11] C.R. Johnson, Jr. et.al., "On Fractionally-Spaced Equalizer Design for Digital Microwave Radio Channels," *Proc. 29th Asilomar Conference*, Pacific Grove, CA, Nov 1995.

[12] S. Dasgupta and Y.F. Huang, "Asymptotically convergent modified recursive least squares with data dependent updating and forgetting factor for systems with bounded noise," *IEEE Trans. Inform. Theory*, vol. IT-33, pp. 383-392, 1987.