

ROBUST BLIND SIMO CHANNEL ESTIMATION USING ADATRON

Dongho Han, Jose C. Principe, Liuqing Yang

University of Florida
Dept. of Electrical & Computer Engineering
Gainesville, FL 32611-6130, USA

Ignacio Santamaría, Javier Vía

University of Cantabria,
Dept. of Communications Engineering
39005 Santander, Spain

ABSTRACT

In this paper we apply the structural risk minimization (SRM) principle to derive a blind single-input multiple-output (SIMO) channel estimation algorithm, which is robust to channel order overestimation. Specifically, the blind estimation is formulated as a support vector regression (SVR) problem in which the channel coefficients are the Lagrange multipliers of the dual problem. In this paper, we show that the SRM principle pushes to zero the small leading and trailing terms of the channel impulse response even when its order is highly overestimated. The main drawback of this approach is the high computational cost of the resulting quadratic programming (QP) problem. To alleviate this, in this paper we propose to use a simple and fast algorithm called the Adatron to solve the QP problem. Simulation results are provided to demonstrate the performance of our channel estimator.

1. INTRODUCTION

Blind estimation of single-input multiple-output (SIMO) channels is a widely studied problem with many signal processing applications. Since the work of [1], it has been well known that second order statistics (SOS) are sufficient for blind identification when the input signal is informative enough and the channels do not share any common roots. Widely used SOS-based methods include the subspace (SS) approach, the least squares (LS) technique and the linear prediction (LP) methods. However, a common drawback of SS and LS techniques is their poor performance when the channel order is overestimated. Recently some robust techniques have been proposed to mitigate this problem (see e.g., [2, 3]). Although these methods offer increased robustness, they still fail when the channel order is highly overestimated.

The structural risk minimization (SRM) principle is a criterion that establishes a trade-off between the complexity of the solution and the closeness to the data. In particular, the support vector machine (SVM) technique, which can be

derived from the SRM principle, typically provides a robust solution. The work in [4] was the first attempt to apply an SVM-based approach to the blind identification of SIMO channels. However, the sparsity provided by the SVM solution was not explicitly exploited in [4]. Later work in this direction was presented in [5], in which a new blind identification algorithm based on support vector regression and specifically tailored for sparse SIMO channels was proposed. The main idea of [5] is that the sparse SIMO channel identification can be reformulated as a set of regression problems in which the channel coefficients play the role of the Lagrange multipliers. By using the ϵ -insensitive Vapnik's loss function in the regression problem, a large number of Lagrange multipliers (and, therefore, a large number of filter coefficients) become zero, thus yielding a sparse filter estimate.

In this paper we extend the previous work [4, 5] in the following directions: first, we derive a robust algorithm for the blind estimation of a non-sparse channel when the channel order has been highly overestimated; secondly, to avoid the high computational cost in solving a QP problem, we use a fast and simple algorithm called the Adatron [6]; and finally, simulation results are provided to verify that the proposed algorithm outperforms existing robust methods even when the channel order is highly overestimated.

2. BLIND SIMO ESTIMATION

SIMO channels emerge either when the signal is oversampled at the receiver or from the use of an array of antennas. Without loss of generality, in this work we focus on the one-input, two-output SIMO system shown in Fig. 1. In blind channel identification, we need to identify the unknown channel responses, $\mathbf{h}_1, \mathbf{h}_2$, from the received signals only. If the order of the channels is M , then the received signal $\mathbf{x}_i(n)$ from the i th channel is

$$x_i(n) = \sum_{k=0}^M h_i(k)s(n-k) + n_i(n), \quad i = 1, 2. \quad (1)$$

This work was partially supported by NSF grant ECS-0300340 and by MEC (Spain) grant TEC2004-06451-C05-02.

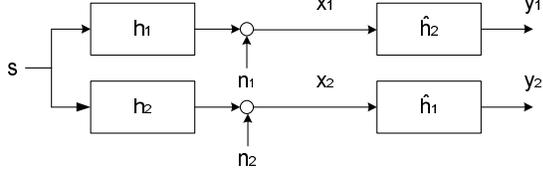


Fig. 1. Single-input two-output channel.

When we cast $x_i(n)$, $h_i(k)$, $s(n)$, $n_i(n)$ into vectors \mathbf{x}_i , \mathbf{h}_i , \mathbf{s} and \mathbf{n}_i respectively, Eq. (1) becomes

$$\mathbf{x}_i = \mathbf{h}_i * \mathbf{s} + \mathbf{n}_i, \quad i = 1, 2 \quad (2)$$

where $*$ denotes convolution. As shown in Fig. 1, using the channel outputs $(\mathbf{x}_1, \mathbf{x}_2)$ and the channel estimates $(\hat{\mathbf{h}}_1, \hat{\mathbf{h}}_2)$, one can obtain the following matrix-vector form:

$$\mathbf{y}_1 = \mathbf{X}_1 \hat{\mathbf{h}}_2 = \mathbf{X}_2 \hat{\mathbf{h}}_1 = \mathbf{y}_2, \quad (3)$$

where \mathbf{X}_i 's are Toeplitz matrices defined as

$$\mathbf{X}_i = \begin{pmatrix} x_i(M) & \cdots & x_i(0) \\ x_i(M+1) & \cdots & x_i(1) \\ \vdots & \ddots & \vdots \\ x_i(M+N-1) & \cdots & x_i(N-1) \end{pmatrix}, \quad (4)$$

or equivalently,

$$\mathbf{X} \hat{\mathbf{h}} = \mathbf{0}, \quad (5)$$

where

$$\mathbf{X} = [\mathbf{X}_2 \quad -\mathbf{X}_1], \quad \hat{\mathbf{h}} = \begin{bmatrix} \hat{\mathbf{h}}_1 \\ \hat{\mathbf{h}}_2 \end{bmatrix}.$$

If we solve (5) by minimizing $\hat{\mathbf{h}}^H \mathbf{X}^H \mathbf{X} \hat{\mathbf{h}}$ with the constraint $\|\hat{\mathbf{h}}\| = 1$, then $\hat{\mathbf{y}}$ is the LS solution which is the eigenvector corresponding to the minimum eigenvalue of $\mathbf{X}^H \mathbf{X}$. Based on (5), we will next develop a SVM based robust blind identification method even when the channel order is highly overestimated.

3. SVM BASED APPROACH

3.1. Support Vector Regression

In [5], the authors proposed a blind iterative procedure for sparse SIMO channels. This procedure offers an increased robustness in comparison to [2] when the channel is sparse. We can also exploit explicitly the sparsity provided by SVM solution to estimate the non-sparse channels when the channel order is highly overestimated. In this situation, the SRM principle pushes to zero the small leading and trailing terms of the impulse response.

From (3), we can formulate the following regression problems:

$$\mathbf{X}_1 \hat{\mathbf{h}}_2 \simeq \mathbf{y}_d, \quad (6)$$

$$\mathbf{X}_2 \hat{\mathbf{h}}_1 \simeq \mathbf{y}_d, \quad (7)$$

where the desired output is constructed as $\mathbf{y}_d = \frac{\mathbf{y}_1 + \mathbf{y}_2}{2}$. This is an intuitive and simple choice, because it drags the actual outputs \mathbf{y}_1 and \mathbf{y}_2 closer to each other in order to achieve the equality in (3). To fully exploit the sparse approximation characteristics provided by SVMs, each of the regression problems is pre-multiplied by its conjugate transposed input matrix \mathbf{X}_i^H to yield [c.f. (6),(7)]:

$$\underbrace{\mathbf{X}_1^H \mathbf{X}_1}_{\mathbf{w}_1} \hat{\mathbf{h}}_2 = \underbrace{\mathbf{X}_1^H \mathbf{y}_d}_{\tilde{\mathbf{y}}_1}, \quad (8)$$

$$\underbrace{\mathbf{X}_2^H \mathbf{X}_2}_{\mathbf{w}_2} \hat{\mathbf{h}}_1 = \underbrace{\mathbf{X}_2^H \mathbf{y}_d}_{\tilde{\mathbf{y}}_2}. \quad (9)$$

The resultant regression problems have input matrices that are simply the conjugate transposed input matrix \mathbf{X}_i^H , and the corresponding desired output vectors become $\mathbf{X}_i^H \mathbf{y}_d$. Moreover, the new regressor \mathbf{w}_i admits an expansion in terms of the filter coefficients, which, in this way, become the Lagrange multipliers of the SVM formulation.

The SVM method minimizes the following cost function

$$J(\mathbf{w}) = C \sum_{n=1}^M (\xi_n + \xi_n^*) + \frac{1}{2} \|\mathbf{w}\|^2 \quad (10)$$

subject to

$$\begin{aligned} \tilde{y}(n) - \mathbf{w}^H \mathbf{x}(n) &\leq \epsilon + \xi_n, & n = 1, \dots, M \\ \mathbf{w}^H \mathbf{x}(n) - \tilde{y}(n) &\leq \epsilon + \xi_n^*, & n = 1, \dots, M \\ \xi_n &\geq 0, & n = 1, \dots, M \\ \xi_n^* &\geq 0, & n = 1, \dots, M \end{aligned}$$

where ξ and ξ^* are positive slack variables introduced by SVM procedure and $\mathbf{x}(n)$ denotes the n -th column of \mathbf{X}_i for $i = 1, 2$.

In (10), the regularization parameter C controls the trade-off between the training error and the complexity of the solution. On the other hand, ϵ is a parameter that determines the precision of the regression and therefore controls the sparseness of the final solution. Then, the solution is a linear combination of input data

$$\mathbf{w} = \sum_{n=1}^M (\alpha_n^* - \alpha_n) \mathbf{x} \quad (11)$$

where α_n^* , α_n are two different Lagrange multipliers.

In (11), only a small number of Lagrange multipliers $(\alpha_n^* - \alpha_n)$ will be nonzero which corresponds to the channel coefficients $h(0), h(1), \dots, h(M-1)$. Accordingly, the overestimated channel coefficients will be zeros by the SRM principle.

3.2. The Adatron algorithm

The computational cost in solving the QP problem in Eq. (10) is the main drawback of applying the SVM technique to practical estimation problems. Several techniques have been proposed to solve this problem, including the use of iterative reweighted least squares (IRWLS) techniques [7, 8] and the Adatron algorithm [6, 9]. The IRWLS requires a matrix inversion at each iteration so the computational burden could be considerably high even for a moderate number of data. On the other hand, the Adatron algorithm is a much simpler least mean square (LMS)-like adaptive algorithm and its convergence rate is exponential with the number of iterations. However, it is a memory intensive method because all the kernel products need to be precomputed and saved. Until [10], the Adatron algorithm has only been used for classification problems.

In this paper, we use the Adatron algorithm to solve the re-formulated regression problems in Eqs. (8) and (9). In dual representation, the optimization problem in Eq. (10) can be written as

$$L = -\frac{1}{2} \sum_{n=1}^M \sum_{m=1}^M (\alpha_n^* - \alpha_n)(\alpha_m^* - \alpha_m) \langle \mathbf{x}_n, \mathbf{x}_m \rangle - \epsilon \sum_{n=1}^M (\alpha_n^* + \alpha_n) + \sum_{n=1}^M y_n (\alpha_n^* - \alpha_n) + \sum_{n=1}^M \lambda (\alpha_n - \alpha_n^*)$$

subject to $\alpha_n, \alpha_n^* \in [0, C]$.

The Adatron algorithm maximizes the above Lagrangian with gradient ascent techniques. Specifically, the Lagrange multipliers are updated according to

$$\delta\alpha_n = \eta \left(-\sum_{m=1}^M (\alpha_m - \alpha_m^*) \langle \mathbf{x}_n, \mathbf{x}_m \rangle + y_n - \epsilon - \lambda \right), \quad (12)$$

$$\delta\alpha_n^* = \eta \left(\sum_{m=1}^M (\alpha_m - \alpha_m^*) \langle \mathbf{x}_n, \mathbf{x}_m \rangle + y_n - \epsilon + \lambda \right), \quad (13)$$

followed by updating α_n and α_n^* with $(\alpha_n + \delta\alpha_n)^\dagger$ and $(\alpha_n^* + \delta\alpha_n^*)^\dagger$, respectively, where $a^\dagger \triangleq \max\{a, 0\}$. In addition, the bias b is updated as $b + \delta\alpha_n - \delta\alpha_n^*$ and the evolution of this bias value can be used to check the convergence of the algorithm. As we can see in Eqs. (12) and (13), the Adatron algorithm is very simple to implement especially in DSP hardware. To run this algorithm, all one needs is just an adder and a multiplier. Furthermore, the computation time of Adatron increases linearly with the number of data while the conventional QP's increases exponentially. This simplicity is its main advantage.

Using this algorithm, we propose a blind channel estimation method as summarized under Algorithm 1, which is robust to channel order overestimation.

Algorithm 1 Summary of the robust blind SIMO channel estimation method

initialize $\epsilon, C, \hat{h}_1, \hat{h}_2$

while Convergence criterion does not meet **do**

 Calculate $y_1, y_2: \mathbf{X}_1 \hat{h}_2 = y_1$ and $\mathbf{X}_2 \hat{h}_1 = y_2$

 Calculate $y_d = (y_1 + y_2)/2$

 Premultiply $\mathbf{X}_1^H, \mathbf{X}_2^H$

 Solve (10) for $i=1,2$ using Adatron algorithm (12),(13)

 Obtain \mathbf{h}_i for $i=1,2$ from (11)

end while

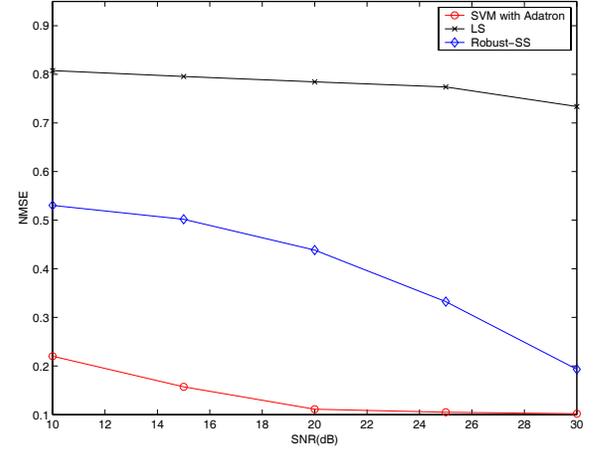


Fig. 2. Performance comparison when the channel order is overestimated by 20 taps (raised-cosine pulse followed by a multipath channel).

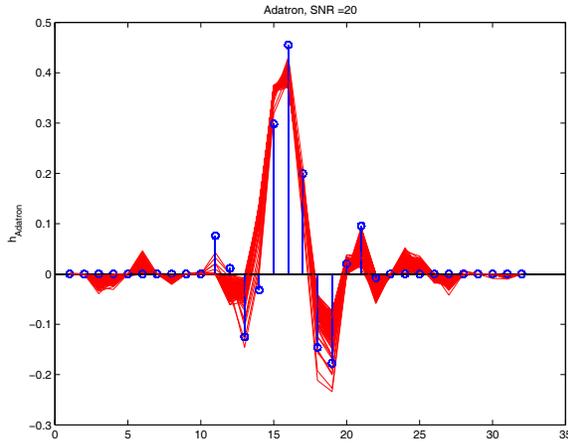
4. SIMULATION RESULTS

Several simulations have been conducted to test the performance of our proposed algorithm. The performance is measured in terms of the normalized mean squared error (NMSE) defined as [3]:

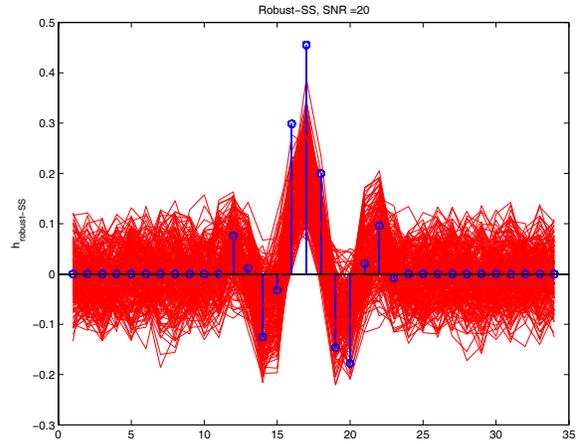
$$NMSE = \frac{1}{\|\mathbf{h}\|^2} \min_{\alpha, k \geq 0} \left\| \alpha \hat{\mathbf{h}} - \begin{bmatrix} \mathbf{0}_{k,1} \\ \mathbf{h} \\ \mathbf{0}_{M'-M-k} \end{bmatrix} \right\|^2$$

where $M' \geq M$ is the estimated channel order.

In this example we consider a raised-cosine pulse with duration $4T$ (T is the symbol period) with a roll-off factor 0.1 and the multipath channel is $h(t) = \delta(t) - 0.7\delta(t - \frac{T}{4})$. The input signal is i.i.d. BPSK signal and the received data is sampled at twice the symbol rate to obtain a SIMO system. Fig. 2 depicts the performance at different SNRs. Note that the performance of our proposed method using the Adatron algorithm is much better than other methods, especially at low SNR. In Fig. 3, 50 trials of our proposed algorithm and the robust method proposed in [2], it is clear that the estimation of our proposed method at zero tap coefficients is much better than the robust-SS method.



(a) SVM based method using the Adatron



(b) Robust-SS method

Fig. 3. 50 trials of the SVM based method using the Adatron and robust-SS method when the channel order is overestimated by 20 taps and SNR=20dB.

5. CONCLUSIONS

In this paper we extend our previous work [5] to robust blind estimation of non-sparse SIMO channels when the channel order is highly overestimated. To reduce the computational cost of solving the QP problem, we use a fast and simple algorithm called the Adatron. We have shown that the proposed algorithm is easy to implement in DSP hardware and the good generalization performance of SVM leads to a remarkable improvement over other robust methods, mainly when the channel order is highly overestimated and for moderate or low SNRs. Finding a true sample by sample (on-line) training algorithm will be an interesting further research.

6. REFERENCES

- [1] G. Xu, H. Liu, L. Tong, and T. Kailath, "Blind identification and equalization based on second-order statistics: A time domain approach," *IEEE Tran. on Information Theory*, vol. 40, pp. 340–349, Mar 1994.
- [2] H. Gazzah, P. A. Regalia, J. Delmas, and K. Abed-Meraim, "A blind multichannel identification algorithm robust to order overestimation," *IEEE Tran. on Signal Processing*, vol. 50, no. 6, pp. 1449–1458, June 2002.
- [3] A. Safavi and K. Abed-Meraim, "Blind channel identification robust to order overestimation: A constant modulus approach," in *Proceedings of ICASSP'03*, Hong Kong, China, Apr 2003, vol. IV, pp. 313–316.
- [4] I. Santamaría, J. Vía, and C. C. Gaudes, "Robust blind identification of SIMO channels: A support vector regression approach," in *Proceedings of ICASSP'04*, 2004, pp. 673–676.
- [5] D. Han, J. C. Principe, L. Yang, I. Santamaría, and J. Vía, "Structural risk minimization for robust blind identification of sparse simo channels," in *Signal Processing Advances in Wireless Communication, IEEE 6th Workshop*, New York, 2–8 June 2005, pp. 806–810.
- [6] I. Santamaría, R. González, C. Pantaleón, and J.C. Principe, "Maximum margin equalizers trained with the adatron algorithm," *Signal Processing*, vol. 83, no. 3, pp. 593–602, 2003.
- [7] F. Pérez-Cruz, A. Navia-Vázquez, P. Alarcón-Diana, and A. Artés-Rodríguez, "An irwls procedure for svr," in *Proceedings of the EUSIPCO'00*, Tampere, Finland, Sept 2000.
- [8] C.C. Gaudes, J. Via, and I. Santamaria, "An irwls procedure for robust beamforming with sidelobe control," in *Proceedings of the Third IEEE Sensor Array and Multichannel Signal Processing Workshop*, Sitges, Spain, July 2004.
- [9] J.K. Anlauf and M. Biehl, "The adatron: an adaptive perceptron algorithm," *Europhys. Lett.* 10, pp. 687–692, 1989.
- [10] Veropoulos K., *Machine Learning Approaches to Medical Decision Making*, Ph.D. thesis, University of Bristol, Bristol, UK, 2001.