

# Implicit Control of Noise Canceller for Speech Enhancement

Julien Bourgeois, Jürgen Freudenberger

Guillaume Lathoud

DaimlerChrysler Research and Technology  
P.O. Box 23 60 – 89013 Ulm – Germany

julien.bourgeois@daimlerchrysler.com  
juergen.freudenberger@daimlerchrysler.com

IDIAP Research Institute  
CP 592 – 1920 Martigny – Switzerland

lathoud@idiap.ch

## Abstract

Widrow’s interference canceller adapted by the normalized LMS (NLMS) is a standard approach for separating signals from multiple speakers, for example from the driver (target) and the codriver (interferer) in a car. In practice, the adaptation must be carried out only when the interferer is dominant, i.e. only when some estimate of the signal-to-interference ratio (SIR) is below a certain threshold. In this paper, we present the implicitly controlled LMS (ILMS), a modification of the NLMS. ILMS adaptation is performed continuously using a variable step-size, whose design implicitly detects dominance of the interferer over target activity. Specific measures are taken to guarantee the stability during adaptation. Theoretical analysis of the ILMS transient convergence and stability conditions prove significant improvement with respect to the original NLMS. Experimental results on real in-car data assess the predicted behavior.

## 1. Introduction

When several persons speak at the same time in front of a speech-based human-machine interface, the separation of their individual speech becomes necessary since only one signal should be passed as input to a speech recognition component. The adaptive interference canceller (AIC) is a widely used approach to this problem [4] and can be applied with an array of directional microphones. One microphone, oriented to the target, provides a target reference signal  $x_0(t)$  but still contains interferer components. The other  $M'$  microphones provide interferer reference signals  $x_{B,m}(t)$ ,  $m = 1, \dots, M'$ , contaminated by target components. Note that  $x_0$  and  $x_{B,m}$  may result from some spatial preprocessing, like with the generalized side-lobe canceller (GSC) [3]. Introducing the input data vectors

$$\begin{aligned} \mathbf{x}_{B,m}(t) &= (x_{B,m}(t), \dots, x_{B,m}(t-L+1))^T, \\ \mathbf{x}_B(t) &= (\mathbf{x}_{B,1}^T(t), \dots, \mathbf{x}_{B,M'}^T(t))^T, \end{aligned}$$

the AIC output  $y(t)$  is defined as  $y(t) = x_0(t) + \mathbf{a}^T(t)\mathbf{x}_B(t)$ , where the vector  $\mathbf{a}(t)$  contains the  $M'L$  coefficients of the noise canceller. The optimal time-dependent noise canceller  $\mathbf{a}_{opt}(t)$  minimizes the interferer power at the output  $y(t)$ .  $\mathbf{a}_{opt}(t)$  is given by the Wiener solution computed on the interferer contribution [4]. Now the question is how to adapt the noise canceller  $\mathbf{a}(t)$  to track  $\mathbf{a}_{opt}(t)$ .

Our algorithmic starting point for the adaptation of  $\mathbf{a}(t)$  is the well-known normalized least-mean-square (NLMS) algorithm [4]:

$$\mathbf{a}(t+1) = \mathbf{a}(t) - \mu_{NLMS} \frac{y(t)\mathbf{x}_B(t)}{\|\mathbf{x}_B(t)\|^2}, \quad (1)$$

where  $\|\mathbf{x}\|^2 = \mathbf{x}^T\mathbf{x}$  and  $\mu_{NLMS}$  denotes the step-size. The convergence and stability are guaranteed in the mean if the simple following sufficient condition is fulfilled

$$0 < \mu_{NLMS} < 2. \quad (2)$$

NLMS minimizes the output power  $\mathbf{E}\{y^2(t)\}$  at time  $t$  but also the correlation  $\mathbf{E}\{y(t)\mathbf{x}_B(t)\}$ . Since  $\mathbf{x}_B(t)$  contains target components, minimizing the correlation  $\mathbf{E}\{y(t)\mathbf{x}_B(t)\}$  reduces the target components in  $y(t)$ , if the target is not silent. Then the adaptation should be, depending on the input signal-to-interference ratio (SIR), slowed down with a smaller step-size or stopped with  $\mu_{NLMS} = 0$ . This is usually done with an all-or-nothing control that allows adaptation only when an estimate of the input SIR is below a certain threshold (e.g. [1]).

In section 2, we modify the NLMS with a variable step-size that performs most adaptation when the interferer is dominant. This is the implicitly controlled NLMS (ILMS) algorithm introduced in [1] and successfully tested in a block-processing fashion. The main contributions of this paper are in sections 3 and 5. On the theoretical side, we show in section 3 that ILMS converges faster and diverges slower than NLMS, and we give an interpretation of the ILMS stability. In section 4, we set a customary constraint on  $\|\mathbf{a}(t)\|$  to further limit the target cancellation. Experimental results on real data are given in section 5, and confirms the predicted behavior. Compared to [1], the sample-by-sample update improves ILMS significantly.

## 2. Implicit Adaptation Control

Observe the standard NLMS equation (1): the norm of the adaptation term

$$\frac{1}{\|\mathbf{x}_B(t)\|^2} \|y(t)\mathbf{x}_B(t)\| \quad (3)$$

decreases with the power of the interferer reference signals  $\|\mathbf{x}_B(t)\|^2$ . When the interferer has much power, the adaptation becomes slower. This effect assures the stability of the algorithm, but leads to slower adaptation in favorable conditions, i.e. when the input signal to interference ratio (SIR) is low. Define  $\mathbf{y}(t) = (y(t), \dots, y(t-L+1))^T$  and assume that  $\|\mathbf{y}(t)\|^2/L$  is a fairly good estimate of the target power. It makes sense to use a large step-size when the target has less power. This can be obtained with an adaptation term with norm

$$\frac{1}{M'\|\mathbf{y}(t)\|^2} \|y(t)\mathbf{x}_B(t)\|.$$

The factor  $M'$  in the denominator is introduced for consistency with (3). This yields also small adaptation terms when the target is loud, which reduces the risk of target cancellation. In other

words, an implicitly controlled adaptation is obtained replacing the NLMS algorithm with:

$$\mathbf{a}(t+1) = \mathbf{a}(t) - \mu_0 \frac{y(t)\mathbf{x}_B(t)}{M'\|\mathbf{y}(t)\|^2}. \quad (4)$$

Equations (4) and (1) are equivalent if we replace  $\mu_{NLMS}$  with a variable step-size  $\mu(t)$ ,

$$\mu(t) = \mu_0 \frac{\|\mathbf{x}_B(t)\|^2}{M'\|\mathbf{y}(t)\|^2}. \quad (5)$$

The condition  $0 < \mu_0 < 2$  does obviously not guarantee that (4) is stable (in the mean) and specific measures must be taken in the implementation to assure its stability. Roughly speaking, (4) can become unstable if  $\|\mathbf{y}(t)\|^2$  is small, since the gradient term  $y(t)\mathbf{x}_B(t)$  is then multiplied by a large factor. A common approach to ensure stability consists in increasing the denominator of the update term in (4) by a fixed regularization term  $\delta$  [2]

$$\mathbf{a}(t+1) = \mathbf{a}(t) - \mu_0 \frac{y(t)\mathbf{x}_B(t)}{\|\mathbf{y}(t)\|^2 + \delta}.$$

However, the fixed regularization scheme generally decreases the convergence speed. We consider rather explicitly the domain of stability (2) expressed with (5):

$$0 < \mu_0 \frac{\|\mathbf{x}_B(t)\|^2}{M'\|\mathbf{y}(t)\|^2} < 2. \quad (6)$$

If the condition (6) is not fulfilled, the adaptation is switched to the standard NLMS with step-size  $\mu_0$ . It must be mentioned that it is safer in practice to consider a stability condition that is more conservative than (6) with  $\mu_{max} < 2$  and

$$0 < \mu_0 \frac{\|\mathbf{x}_B\|^2}{M'\|\mathbf{y}\|^2} < \mu_{max}.$$

To summarize, the ILMS algorithm is stated as

$$\mathbf{a}(t+1) = \mathbf{a}(t) - \begin{cases} \mu_0 \frac{y(t)\mathbf{x}_B(t)}{M'\|\mathbf{y}(t)\|^2} & \text{if } \mu_0 \frac{\|\mathbf{x}_B(t)\|^2}{M'\|\mathbf{y}(t)\|^2} < \mu_{max}, \\ \mu_0 \frac{y(t)\mathbf{x}_B(t)}{\|\mathbf{x}_B(t)\|^2} & \text{otherwise.} \end{cases} \quad (7)$$

In the next section, we examine the transient behavior of the ILMS and give an interpretation of the ILMS stability condition (6).

### 3. ILMS transient behavior and stability

The time  $t$  is fixed and may be dropped for clarity. Let us denote by  $\sigma_1^2$  the variance of the target at the output  $y$  and denote by  $\sigma_2^2$  the variance of the interferer at the interferer reference  $\mathbf{x}_B$ . Assuming uncorrelated target and interference, the variance of  $\mathbf{x}_B$  and  $y$  are decomposed as

$$\mathbf{E}\{\|\mathbf{x}_B\|^2\} = M'L(\varepsilon_{leakage}\sigma_1^2 + \sigma_2^2), \quad (8)$$

$$\mathbf{E}\{\|\mathbf{y}\|^2\} = L\sigma_y^2 = L(\sigma_1^2 + \varepsilon_{mismatch}\sigma_2^2). \quad (9)$$

The factor  $0 < \varepsilon_{leakage} < 1$  represents the amount of target leakage into the interferer reference. The factor  $0 < \varepsilon_{mismatch} < 1$  controls the interferer power at the beamformer output, and is roughly speaking proportional to the mismatch  $\mathbf{m}(t) = \mathbf{a}(t) - \mathbf{a}_{opt}(t)$ . Assuming that  $\|\mathbf{x}_B\|^2$  and  $\|\mathbf{y}\|^2$  are uncorrelated, the expected step-size (5) is then:

$$\mathbf{E}\{\mu(t)\} = \mu_0 \frac{\varepsilon_{leakage}\sigma_1^2 + \sigma_2^2}{\sigma_1^2 + \varepsilon_{mismatch}\sigma_2^2}. \quad (10)$$

#### 3.1. Transient convergence and divergence

The ILMS algorithm is similar to the NLMS algorithm in that sense that they share at each step the same gradient direction,  $y(t)\mathbf{x}_B(t)$ . Both algorithms converge (resp. diverge) during target activity (resp. silence). But their speed of convergence and divergence differ. In this subsection we derive upper bounds on the transient convergence and divergence rates of the ILMS and the NLMS algorithms. The convergence rate is evaluated when the target is silent, the worst-case divergence rate is obtained when the interference is silent.

We approximate the interferer references as white signals so that the correlation matrix  $\mathbf{R}_{\mathbf{x}_B\mathbf{x}_B} = \mathbf{E}\{\mathbf{x}_B(t)\mathbf{x}_B^T(t)\} \propto \mathbf{I}$ , where  $\mathbf{I}$  denotes the identity matrix. It is also assumed we are in the stable case, i.e.  $\mathbf{E}\{\mu(t)\} < 2$ , since otherwise, the ILMS is switched to NLMS and both algorithms behave identically. Assuming that  $y(t)\mathbf{x}_B(t)$  and  $\|\mathbf{y}(t)\|^2$  are decorrelated, it can be shown that the expectation of the ILMS adaptation step (7) at time  $t$  is:

$$\mathbf{E}\{\mathbf{m}(t+1)\} = \frac{\mathbf{E}\{\mathbf{m}(t)\} - \mathbf{E}\{\mu(t)\} \times \frac{\mathbf{E}\{b(t)\mathbf{x}_B(t)\} + \mathbf{R}_{\mathbf{x}_B\mathbf{x}_B}\mathbf{E}\{\mathbf{m}(t)\}}{\mathbf{E}\{\|\mathbf{x}_B(t)\|^2\}}}{\mathbf{E}\{\|\mathbf{x}_B(t)\|^2\}}. \quad (11)$$

with  $b(t) = x_0^{(d)}(t) + \mathbf{a}_{opt}(t)\mathbf{x}_B^{(d)}(t)$ . The superscript  $(d)$  in  $x^{(d)}(t)$  refers to the target signal contribution in the signal  $x(t)$ .

*ILMS convergence.* Convergence of the mismatch  $\mathbf{m}(t)$  to zero occurs in particular when the target is silent, i.e. when  $\sigma_1^2 = 0$  and  $\sigma_2^2 > 0$ . With (11), (10),  $\mathbf{R}_{\mathbf{x}_B\mathbf{x}_B} = \sigma_2^2\mathbf{I}$  and  $b(t) = 0$ , this yields

$$\mathbf{E}\{\mathbf{m}(t+1)\} = \underbrace{\left(1 - \frac{\mu_0}{M'L\varepsilon_{mismatch}}\right)}_{\text{contraction factor } \alpha} \mathbf{E}\{\mathbf{m}(t)\}.$$

Let us denote by  $\tilde{\mu}_0$  the step-size normalized to the filter length and the number of interferer references:  $\tilde{\mu}_0 = \mu_0/(M'L)$ . The contraction factor  $\alpha = 1 - \frac{\tilde{\mu}_0}{\varepsilon_{mismatch}}$  controls the initial convergence of  $\mathbf{E}\{\mathbf{m}(t)\}$  to zero. It should be as close to zero as possible and its absolute value should be always smaller than one. Note that if an estimate of  $\varepsilon_{mismatch}$  would be available, setting  $\tilde{\mu}_0 = \varepsilon_{mismatch}$  would lead to the fastest convergence (under the assumptions that are made, convergence in the mean would be achieved in one step). Assuming that  $\varepsilon_{mismatch} < 1$ , the contraction factor is upperbounded by

$$1 - \frac{\tilde{\mu}_0}{\varepsilon_{mismatch}} < 1 - \tilde{\mu}_0.$$

$\mathbf{E}\{\mathbf{m}(t)\}$  decreases to zero faster than  $(1 - \tilde{\mu}_0)^t$ . Thus, we define (an overestimation of) the ILMS convergence contraction factor  $\alpha_{ILMS}(\tilde{\mu}_0) = 1 - \tilde{\mu}_0$ .

*ILMS divergence.* The worst-case situation happens when the target speaks and the interferer is silent, i.e. when  $\sigma_1^2 > 0$  and  $\sigma_2^2 = 0$ . Then the mean mismatch  $\mathbf{E}\{\mathbf{m}(t)\}$  diverges from zero and converges to another point denoted by  $\mathbf{m}_d(t)$ . Set  $\mathbf{m}'(t) = \mathbf{m}(t) - \mathbf{m}_d(t)$  and assume  $\mathbf{m}_d(t+1) = \mathbf{m}_d(t)$ . It can be shown that the ILMS adaptation in terms of  $\mathbf{m}'(t)$  is then with (11), (10), and  $\mathbf{R}_{\mathbf{x}_B\mathbf{x}_B} = \varepsilon_{leakage}\sigma_1^2\mathbf{I}$ :

$$\mathbf{E}\{\mathbf{m}'(t+1)\} = \left(1 - \mu_0 \frac{\varepsilon_{leakage}}{M'L}\right) \mathbf{E}\{\mathbf{m}'(t)\}.$$

With the normalized step-size  $\tilde{\mu}_0$ , we define the ILMS divergence contraction factor

$$b_{ILMS}(\tilde{\mu}_0) = 1 - \mu_0 \frac{\varepsilon_{leakage}}{M'L} = 1 - \tilde{\mu}_0\varepsilon_{leakage}.$$

*NLMS convergence and divergence.* Similarly, it can be shown that the convergence of the NLMS with constant step-size  $\mu_{NLMS}$  is described in the best-case scenario by  $\mathbf{E}\{\mathbf{m}(t+1)\} = \mathbf{E}\{\mathbf{m}(t)\}(1 - \tilde{\mu}_{NLMS})$ , and in the worst-case scenario by  $\mathbf{E}\{\mathbf{m}'(t+1)\} = \mathbf{E}\{\mathbf{m}'(t)\}(1 - \tilde{\mu}_{NLMS})$ , with  $\tilde{\mu}_{NLMS} = \mu_{NLMS}/M'L$ . We define the NLMS convergence factor  $a_{NLMS}(\tilde{\mu}) = 1 - \tilde{\mu}$  and the divergence factor  $b_{NLMS}(\tilde{\mu}) = 1 - \tilde{\mu}$ .

*Comparison.* To represent jointly the convergence and divergence speeds, we depict in Fig. 1 the divergence contraction factor  $b(\mu)$  as a function of the convergence contraction factor  $a(\mu)$ , obtaining a convergence/divergence pattern. The best convergence behavior would be obtained with  $b(\mu) = 1, \forall a(\mu)$ : the mismatch would not increase, even in the worst-case scenario. Conversely, the worst behavior would be obtained with  $a(\mu) = 1, \forall b(\mu)$ : the mismatch would not reduce, even in best-case scenario. For the NLMS, the convergence and divergence

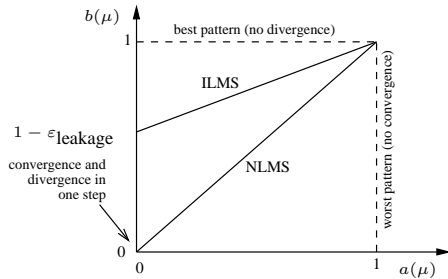


Figure 1: NLMS and ILMS convergence/divergence patterns for  $0 < \epsilon_{\text{leakage}} < 1$ .

contraction terms are equal,  $b_{NLMS}(\mu) = a_{NLMS}(\mu)$ . In other words, the NLMS converges and diverges with the same speed. On the other hand we have

$$b_{ILMS}(\mu) = a_{ILMS}(\mu)\epsilon_{\text{leakage}} + (1 - \epsilon_{\text{leakage}}).$$

If  $\epsilon_{\text{leakage}} < 1$ , then the convergence/divergence pattern of the ILMS is "better" than the NLMS pattern.

### 3.2. About the stability

Let us examine more closely in which circumstances instability can be detected. Assume the stability condition (6) is not fulfilled in the mean, so

$$\mu_0 \frac{\epsilon_{\text{leakage}}\sigma_1^2 + \sigma_2^2}{\sigma_1^2 + \epsilon_{\text{mismatch}}\sigma_2^2} > 2. \quad (12)$$

Rearranging (12) yields

$$\epsilon_{\text{mismatch}} < \frac{\mu_0}{2} - \frac{\sigma_1^2}{\sigma_2^2} \left(1 - \epsilon_{\text{leakage}} \frac{\mu_0}{2}\right) \quad (13)$$

For  $\mu_0$  sufficiently small, the term  $1 - \epsilon_{\text{leakage}}\mu_0/2$  is positive, so that (13) implies

$$\begin{cases} (i) & \epsilon_{\text{mismatch}} < \frac{\mu_0}{2} \\ (ii) & \frac{\sigma_1^2}{\sigma_2^2} < \frac{\frac{\mu_0}{2} - \epsilon_{\text{mismatch}}}{1 - \epsilon_{\text{leakage}} \frac{\mu_0}{2}} < \left(\frac{2}{\mu_0} - \epsilon_{\text{leakage}}\right)^{-1} \end{cases} \quad (14)$$

The equations (14) tell us that the ILMS algorithms is switched to the standard NLMS only if

- (i) the mismatch factor is smaller<sup>1</sup> than  $\mu_0/2$ ,

<sup>1</sup>Note that typical step-sizes have order of magnitude  $\mu_0 = 0.01$  or smaller.

- (ii) and if the input SIR, defined here by  $\frac{\sigma_1^2}{\sigma_2^2}$ , is close to zero.

This important result reveals that the ILMS property of implicit silence detection is not lost by switching to the standard NLMS when instability is detected, since *the stability condition (6) acts as a target silence detector.*

## 4. Quadratic inequality constraint

In most applications where prior information about the target position is available, the SIR at the target reference signal  $x_0(t)$  (resp. interferer reference  $x_B(t)$ ) is positive (resp. negative). Therefore, cancelling the target at the beamformer output requires large noise canceller coefficients whereas cancelling the interferer does not. Thus, we enforce an upper bound on  $\|\mathbf{a}(t)\|$  with the projection:

$$\mathbf{a}(t) \leftarrow \frac{a}{\|\mathbf{a}(t)\|} \mathbf{a}(t) \text{ if } \|\mathbf{a}(t)\| > a. \quad (15)$$

where  $a > 0$  is a constant. The smaller  $a$ , the lower are the target and interference reductions. Assume that the optimal noise canceller  $\|\mathbf{a}_{\text{opt}}(t)\| < a$ . Then, (15) also limits the mismatch, since  $\|\mathbf{m}(t)\| = \|\mathbf{a}(t) - \mathbf{a}_{\text{opt}}(t)\| < 2a$ . The constant  $a$  depends on the position of the sources and on the microphone arrangement. Thus, one must determine experimentally the smallest  $a$  that does not impair the interference reduction.

## 5. Experiments

We present experimental results obtained with real recordings sampled at  $f_s = 16000$  Hz. The following algorithms are compared:

- **NLMS:** The adaptation is performed after (1) with step-size  $\mu_{NLMS}$  when speech activity is detected. The speech activity detection<sup>2</sup> is obtained comparing the short-term input power, averaged on the  $M$  microphones, with a fixed threshold. This threshold is determined as the microphone power during the speech-free first 200 ms, averaged on the  $M$  microphones.
- **ILMS:** The adaptation is performed after (7) with step-size  $\mu_0$  when speech activity is detected. The stability threshold  $\mu_{\text{max}}$  is set to  $\mu_{\text{max}} = \frac{1}{2}$ .
- **VAD-NLMS:** Adaptation of the noise canceller is performed after (1) with step-size  $\mu_{NLMS}$  when speech activity is detected and the target is silent. The voice activity detector (VAD) is obtained with a sector-based approach (see [1] and the references therein).

The filter length is set to  $L = 256$ . The three algorithms are implemented with the quadratic inequality constraint (15). The algorithms performances are quantified by the desired signal reduction  $SR(t)$  and the interference reduction  $IR(t)$ , measured with respect to the target reference  $x_0(t)$ .

### 5.1. Experiment with directional microphones

A two directional microphones arrangement is mounted in a Mercedes S320 vehicle. The microphones are installed on the roof with an inter-microphone spacing of 17 cm and oriented to the driver and the codriver. The microphone that is oriented to the driver provides the target reference signal  $x_0(t)$ , the other one is the interferer reference  $x_B(t)$  ( $M' = 1$ ).

<sup>2</sup>No distinction is made between speech of the target and speech of the interferer.

*Determination of the constants  $\mu_0$  and  $a$ .* For the tuning of  $\mu_0$  and  $a$ , a quality measure  $Q$  is defined as the interference reduction evaluated with respect to the first microphone:

$$Q = \frac{\sum_{\tau=1}^{2f_s} (x_0^{(i)}(\tau))^2}{\sum_{\tau=1}^{2f_s} y^2(\tau)},$$

where  $x_0^{(i)}(t)$  denotes the interferer components in the target reference signal.  $Q(\mu_0)$  is presented on Fig. 2 together with

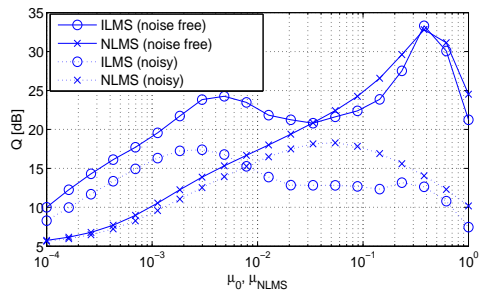


Figure 2: Left:  $Q$  as a function of  $\mu_0$  and  $\mu_{NLMS}$ .

$Q(\mu_{NLMS})$ , without and with road noise.  $Q(\mu_0)$  exhibit two local maxima. The relevant one is the first, at  $\mu_0 = 0.005$ . The second maximum of  $Q(\mu_0)$  for  $\mu_0 > 0.1$  appears because of switching to NLMS. It can be observed that the step-size  $\mu_0$  that maximizes  $Q$  is practically independent of the background noise. In this respect, the robustness of the ILMS adaptation is remarkable. In contrast, the NLMS step-size  $\mu_{NLMS}$  must be adjusted to the noise level:  $\mu_{NLMS} = 0.2$  without background noise,  $\mu_{NLMS} = 0.05$  in noisy conditions. Concerning the upper bound  $a$  of quadratic inequality constraint, we found out that the conservative  $a = 0.5$  does not impair the interference reduction.

*Online performance.* The input signals consist of recordings of two male speakers and exhibit a significant degree of overlap (double-talk). Online performances without road background noise are illustrated on Fig. 3, those with road background noise are not presented but they have similar characteristics. It can be

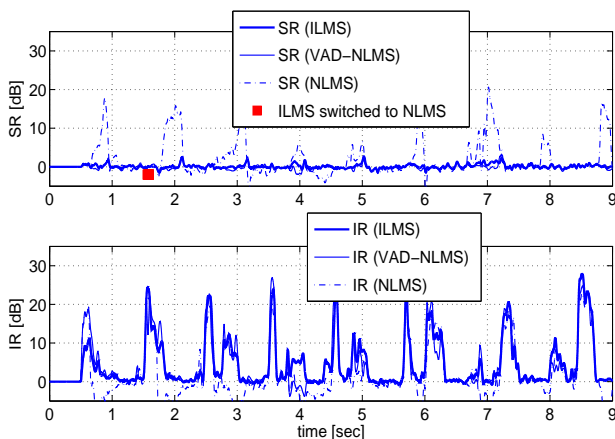


Figure 3: Online performance. The step-size  $\mu_{NLMS}$  is set to  $\mu_{NLMS} = 0.2$ .

observed that **NLMS** leads to significant target reduction. Although the prior information about the source position is used at the physical level with directive microphone and with conservative constraint  $\|a(t)\| < 0.5$ , target cancellation occurs. This

influences negatively the **NLMS** interference reduction, since the degrees of freedom that are allocated for target cancellation are not available for the interference reduction. In contrast, the **ILMS** method produces no noticeable target reduction. The periods of time when the stability condition is not fulfilled are indicated below the  $SR(t)$  curves. As predicted in section 3.2, ILMS switches in noise free conditions to NLMS only if the target is silent (i.e., where no  $SR(t)$  is measured). The fastest convergence is reached with the **VAD-NLMS** method. However, the interference reduction is limited during double-talk since no adaptation occurs if the target is active. At last, our results somewhat differ from those presented in [1]: instead of a bloc-based update, we use a sample-by-sample adaptation, which improves of performance of ILMS significantly.

## 6. Summary and conclusions

Adaptative noise canceller tracks not only changes of the acoustic responses but also the spectral changes of the interferer, exploiting at best its actual colored<sup>3</sup> spectrum. For the NLMS algorithm, large step-sizes are desirable since they allow a rapid tracking of the spectral changes. However, a large step-size also leads to an important target signal reduction, e.g. during double-talk. The ILMS algorithm solves this contradiction: it is shown to converge faster and diverge slower than the NLMS algorithm. It is also shown how the associated stability condition acts as a target silence detector. To further increase the robustness against target cancellation, we integrate a quadratic inequality constraint. The performance of the ILMS algorithm is then examined experimentally and compared to the traditional NLMS algorithm. The theoretical results are confirmed. Moreover, it appears that the step-size  $\mu_0$  of the ILMS does not require to be adjusted to the background noise level. Without extra computational cost, the SIR improvement provided by ILMS overtakes that of the uncontrolled NLMS algorithm. With the considered microphone arrangement, the ILMS performance is similar to that of the VAD-controlled NLMS, while not requiring any external adaptation control. Thus, ILMS is with this setup very robust against target leakage and has a high practical relevance.

*Acknowledgements.* Thanks to the European Community's Human Potential Programme under contract HPRN-CT-2002-00276 HOARSE for funding.

## 7. References

- [1] G. Lathoud, J. Bourgeois and J. Freudenberger. "Sector-Based Detection for Speech Enhancement in Cars", IDIAP Research Report RR-04-67, December 2004.
- [2] E. Hänsler, G. U. Schmidt, "Acoustic Echo and Noise Control: A Practical approach." Wiley, 2004.
- [3] B.D. Van Veen and K.M. Buckley, "Beamforming: a versatile approach to spatial filtering." IEEE Acoustics, Speech, Signal Processing Magazine, vol. 5, pp 4–24, April 1988.
- [4] B. Widrow, S. D. Stearn, "Adaptive Signal Processing." Prentice-Hall Signal Processing Series, 1985.

<sup>3</sup>Colored interferer signals lead to small eigenvalues of the interference correlation matrix  $\mathbf{R}_{x_B \times x_B}$ . In general, these small eigenvalues lead to a slower transient convergence but they have a relatively small contribution to the output power.