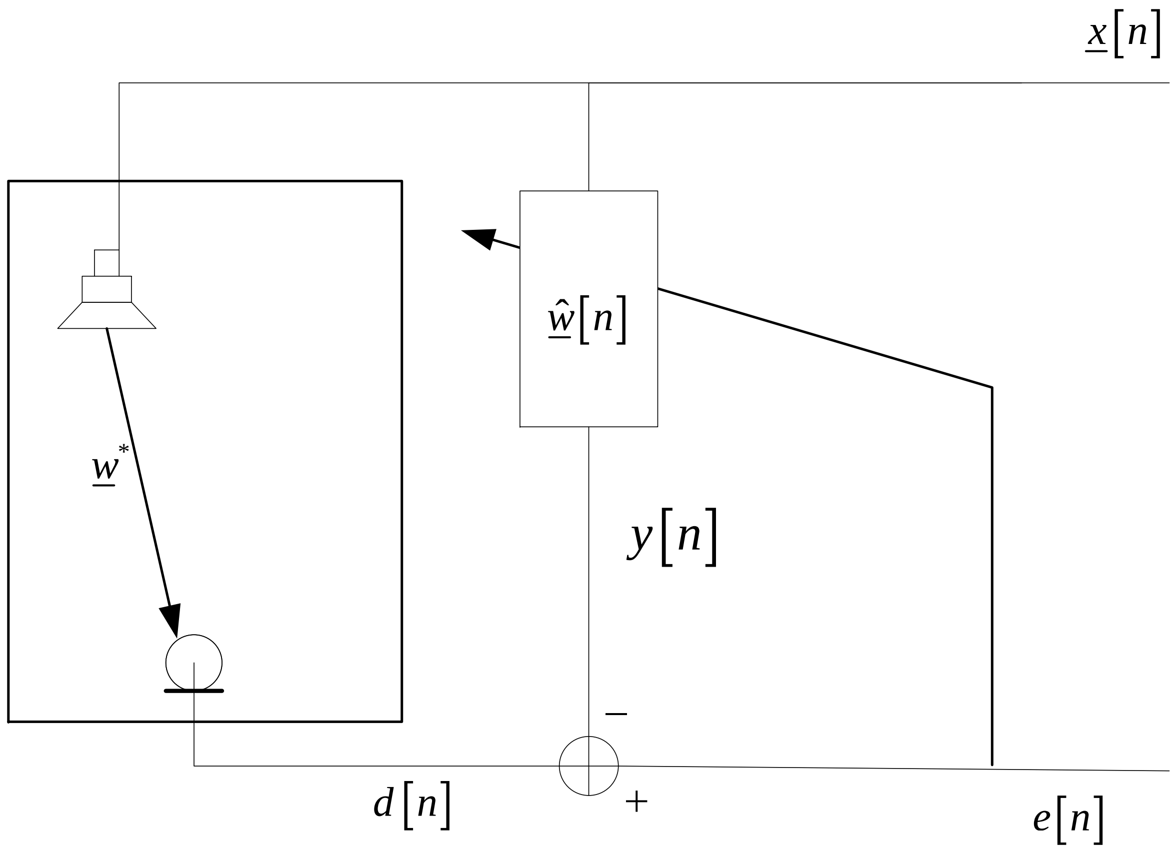
Stereophonic Acoustic Echo Cancellation

Overview

- SAEC introduction
  - Introduction to stereophonic communication
  - Problem formulation
- Current solutions
  - IS
  - POWER
- Proposed modification
  - Requirement
  - Unified convergence limit algorithm
  - Proposed amendment

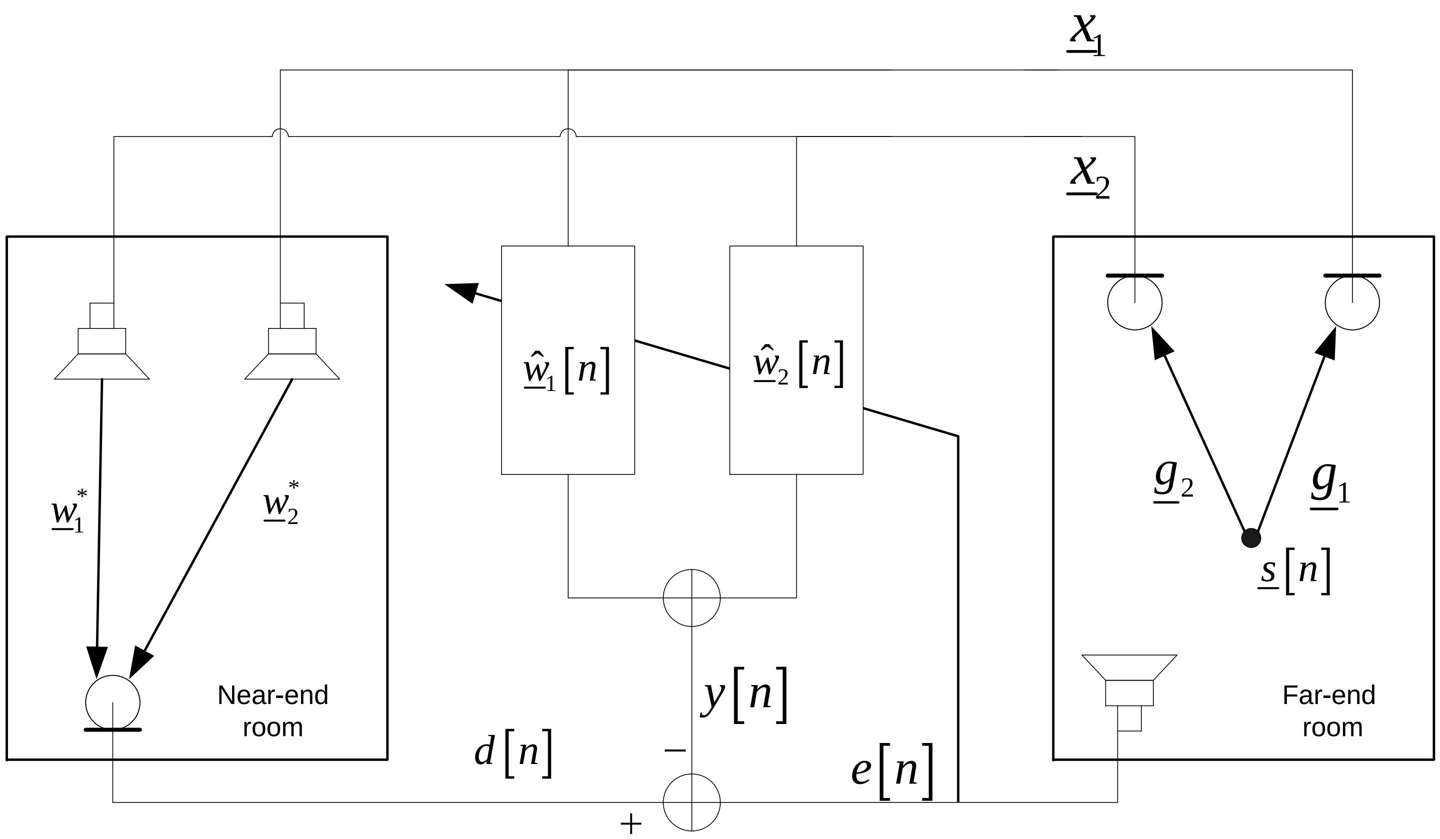
Classic AEC Scheme



Single Channel EC

- Has been researched for some 30 years
- Has many robust efficient solutions
- Lacks spatial identification
- This reason motivates the addition of another channel

Classic SAEC Scheme



SAEC Problems – Convergence Rate

- Autocorrelation matrix of the input is not full rank
- Slow convergence rate even for white input

$$\underline{x}[n] = \left[\underline{x}_1^T[n] \quad \underline{x}_2^T[n]\right]^T$$

$$\underline{w}^* = \left[\underline{w}_1^{*T} \quad \underline{w}_2^{*T}\right]^T$$

$$\hat{\underline{w}}[n] = \left[\hat{\underline{w}}_1^T[n] \quad \hat{\underline{w}}_2^T[n]\right]^T$$

$$E\left\{e^2[n]\right\} = E\left\{\left(\underline{w}^* - \hat{\underline{w}}[n]\right)^T \underline{x}[n]\underline{x}^T[n]\left(\underline{w}^* - \hat{\underline{w}}[n]\right)\right\} = \left(\underline{w}^* - \hat{\underline{w}}[n]\right)^T R\left(\underline{w}^* - \hat{\underline{w}}[n]\right)$$

$$R = E\left\{\left[\underline{x}_1^T[n] \quad \underline{x}_2^T[n]\right]^T \left[\underline{x}_1^T[n] \quad \underline{x}_2^T[n]\right]\right\} = \begin{bmatrix} R_{x_1x_1} & R_{x_1x_2} \\ R_{x_2x_1} & R_{x_2x_2} \end{bmatrix}$$

$$\underline{g}_1^T \underline{x}_2[n] = \underline{g}_2^T \underline{x}_1[n]$$

SAEC Problems - Nonuniqueness (1)

- Defining the impulse response matrix of the Far-end room as:

$$e[n] = \left(\underline{w}_1^* - \underline{\hat{w}}_1[n]\right)^T \underline{x}_1[n] + \left(\underline{w}_2^* - \underline{\hat{w}}_2[n]\right)^T \underline{x}_2[n]$$

$$\underline{x}_i[n] = G_i \underline{s}[n]$$

$$G_i = \begin{bmatrix} g_{i,0} & g_{i,1} & \cdots & g_{i,M-1} & 0 & \cdots & 0 & 0 \\ 0 & g_{i,0} & \cdots & g_{i,M-2} & g_{i,M-1} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \cdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & g_{i,0} & g_{i,1} & \cdots & g_{i,M-1} \end{bmatrix}$$

SAEC Problems - Nonuniqueness (2)

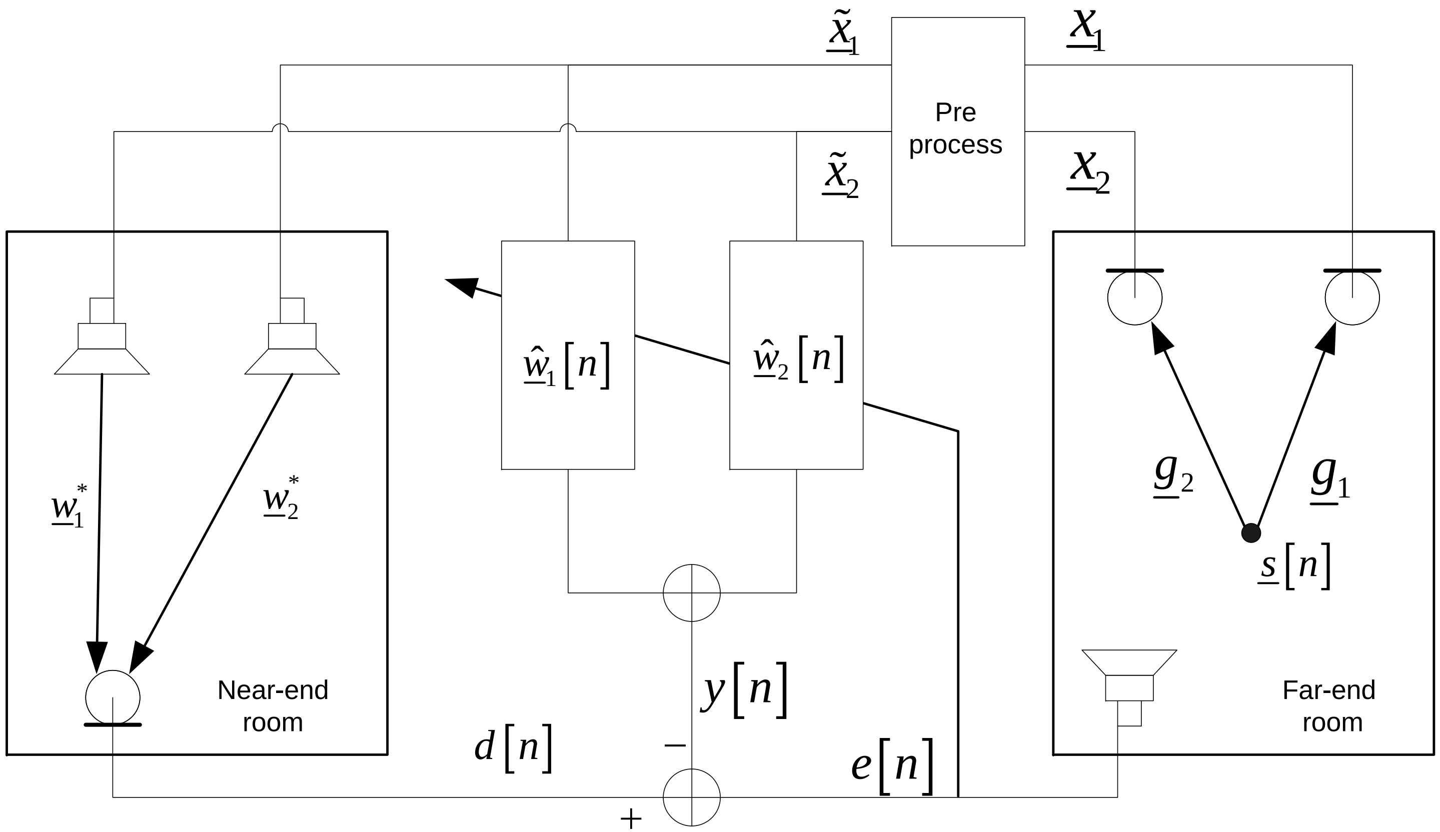
- Algorithms converge to local minima point dependent on the **far-end** impulse response

$$e[n]=\left[\left(\underline{w}_1^*-\hat{\underline{w}}_1[n]\right)^T G_1+\left(\underline{w}_2^*-\hat{\underline{w}}_2[n]\right)^T G_2\right]\underline{s}[n]=\underline{p}^T[n]\underline{s}[n]$$

$$\nabla=\frac{dE\left\{e^2[n]\right\}}{d\underline{p}[n]}=\underline{p}[n]R_{ss}=0\Rightarrow\underline{p}[n]=0$$

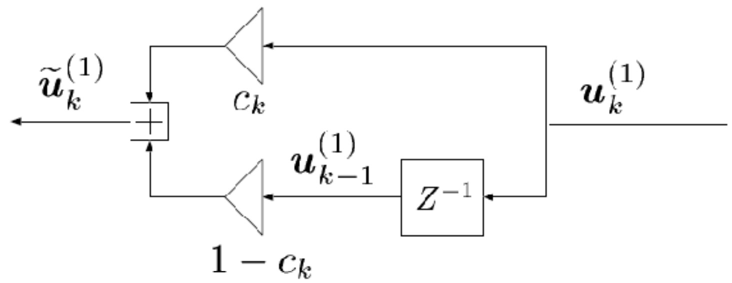
$$\left[\left(\underline{w}_1^*-\hat{\underline{w}}_1[n]\right)^T G_1+\left(\underline{w}_2^*-\hat{\underline{w}}_2[n]\right)^T G_2\right]=0\Leftrightarrow\underline{w}_1^*\neq\hat{\underline{w}}_1[n],\underline{w}_2^*\neq\hat{\underline{w}}_2[n]$$

Current Solutions



Input Slide

- The IS unit is connected to one input signals
- c_k is a periodical signal – equals 1 for half the period, 0 for the other
- The processed input signal is delayed for half the period

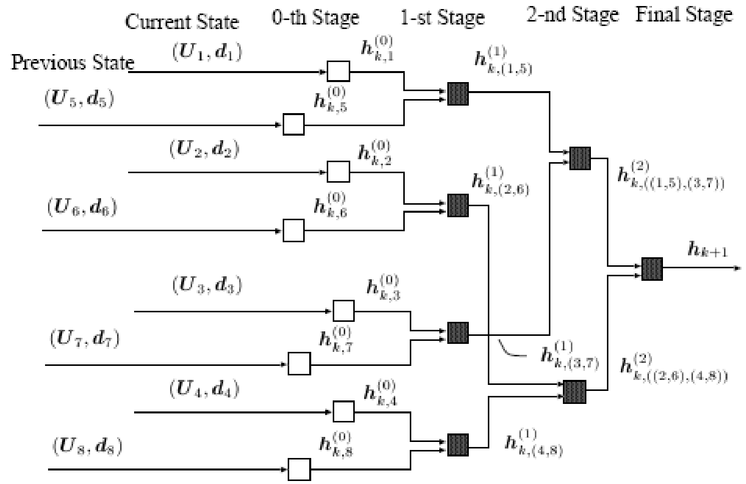


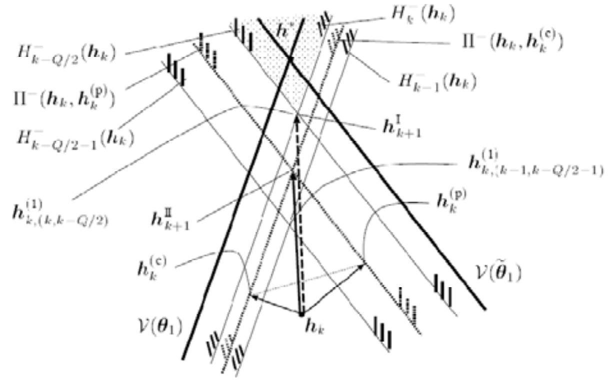
Input Slide Problems

- Quick transitions between states causes audible clicks – can be solved with a smoother change function
- Short c_k period will cause aliasing and degrade speech quality, longer period will slow algorithm convergence rate

POWER – a Swifter Solution

- The POWER algorithm uses data from the two states simultaneously to project the current solution using NLMS based APA to the next proposed solution
- Then the POWER algorithm finds an intersection point between the two solutions to obtain the next solution



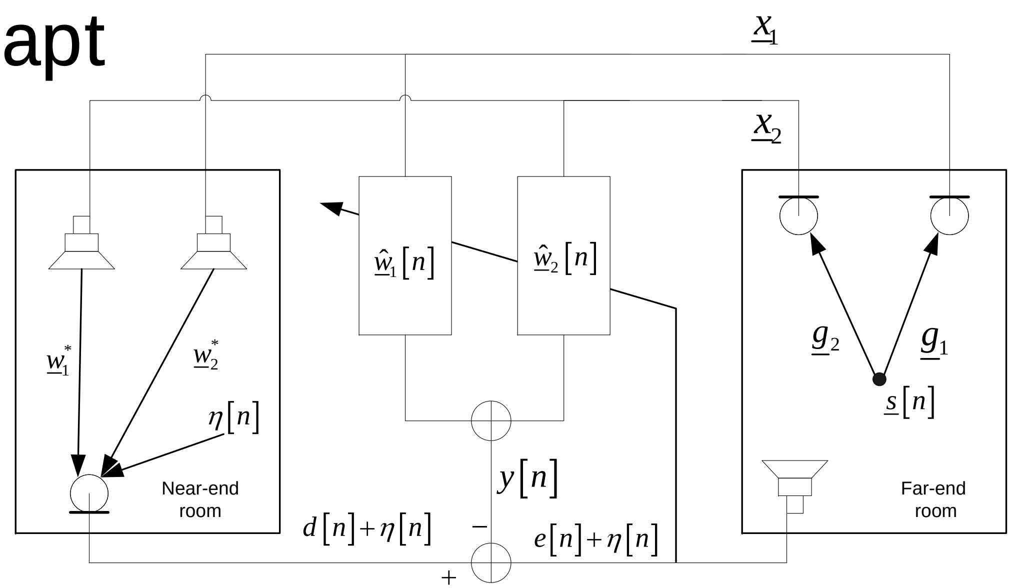


Further Convergence Enhancement

- Use of IS and POWER algorithm provides good SAEC solution
- IS solves the Non-uniqueness problem
- POWER provides convergence rate acceleration

Further Convergence Enhancement (1)

- Both method relay on DTD device to freeze adaptation during DT situation
- In the case of constant near-end room interference the current solutions will be unable to adapt



Further Convergence Enhancement (2)

- The SAEC adaptive filters work under a noisier environment when compared to the single channel case
- Added noise is due to the common error signal and is the result of:
  - additional adaptive filter – additional tap mismatch noise
  - Additional IR tail

Further Convergence Enhancement (3)

- IR tail – the part of the impulse response which the adaptive filter cannot adapt to, due to its shorter length:
  - Acoustic IR are very long (thousands of taps)
  - The adaptive filters length is limited by calculation power
  - The length mismatch induces additional system noise
- Additional IR results in additional IR tail noise

Further Convergence Enhancement (4)

- In real life Stereo environment, the Acoustic IR energies wont always be symmetrical
- In the case of a dominant channel, the weak channel will converge in the presence of the dominant channel IR tail noise
- This results in a poorer convergence limit when compared to the single channel case

Possible Solution

- A possible solution for the above problem can be the use of a noise robust adaptive algorithm instead of NLMS for the projection process of the POWER algorithm
- Benefits of using a noise robust algorithm:
  - Will lower the IR tail effect for the non-dominant channel – thus improving the overall misadjustment of the system
  - Will enhance the system misalignment for the high noise environments
  - Can be used in a DT situation to continue convergence during the interference period

Candidate for DT Robust Alg.

- NLMS: $\underline{\hat{w}}[n+1] = \underline{\hat{w}}[n] + \frac{\mu \underline{x}[n] e[n]}{\left\| \underline{x}[n] \right\|^2}$

  poor candidate – unbounded gradient

- SA: $\underline{\hat{w}}[n+1] = \underline{\hat{w}}[n] + \mu \underline{x}[n] sign(e[n])$

  better choice – bounded gradient

- SA convergence limit is affected by disturbance energy

$$\sigma^2_{\varepsilon_{SA-clean}}[n] \geq \frac{\pi}{8} \mu^2 \sigma_x^4 N^2$$

$$\sigma^2_{\varepsilon_{SA-noisy}}[n] \geq \sqrt{\frac{\pi}{8}} \mu N \sigma_x^2 \sigma_\eta$$

Deriving Opt. Step Size for SA In a Noisy Environment

Tap update equation:

$$\underline{\hat{w}}[n+1] = \underline{\hat{w}}[n] + \mu \underline{x}[n] sign(e[n])$$

Assume that the error is composed of the residual echo and the near-end interference

$$e[n] = \varepsilon[n] + \eta[n]$$

Model the signals as uncorrelated Gaussian RVs:

$$\eta[n] \sim N\left(0, \sigma_\eta^2[n]\right)$$

$$\varepsilon[n] \sim N\left(0, \sigma_\varepsilon^2[n]\right)$$

$$x[n] \sim N\left(0, \sigma_x^2\right)$$

$$\varepsilon[n] \amalg \eta[n] \amalg x[n]$$

Opt Step Size (1)

Define the misadjustment expression

$$\underline{m}[n]=\underline{w}^{*}-\underline{\hat{w}}[n] \quad \underline{m}[n]\amalg \underline{x}[n]$$

Express the misadjustment update eq. using the above:

$$\underline{m}[n+1]=\underline{m}[n]-\mu\underline{x}[n]sign\left(\varepsilon[n]+\eta[n]\right)$$

Using inner product express the misadjustment l2 norm:

$$\left\|\underline{m}[n+1]\right\|^{2}=\left\|\underline{m}[n]\right\|^{2}+\mu^{2}\left\|\underline{x}[n]\right\|^{2}-2\mu\left\langle\underline{m}[n],\underline{x}[n]\right\rangle sign\left(\varepsilon[n]+\eta[n]\right)$$

Opt Step Size (2)

Remember: $\langle m[n], x[n] \rangle = \varepsilon[n]$

$$\left\|\underline{m}[n+1]\right\|^2 = \left\|\underline{m}[n]\right\|^2 + \mu^2 \left\|\underline{x}[n]\right\|^2 - 2\mu\varepsilon[n]sign\left(\varepsilon[n]+\eta[n]\right)$$

Take statistical expectation:

$$E\left\{\left\|\underline{m}[n+1]\right\|^2\right\} = E\left\{\left\|\underline{m}[n]\right\|^2\right\} + \mu^2 E\left\{\left\|\underline{x}[n]\right\|^2\right\} - 2\mu E\left\{\varepsilon[n]sign\left(\varepsilon[n]+\eta[n]\right)\right\}$$

The following facts can be used:

$$\sigma_\varepsilon^2[n] = E\left\{\varepsilon^2[n]\right\} = E\left\{\underline{m}^T[n]\underline{x}[n]\underline{x}^T[n]\underline{m}[n]\right\} = \sigma_x^2 E\left\{\left\|\underline{m}[n]\right\|^2\right\}$$

$$E\left\{\left\|\underline{x}[n]\right\|^2\right\} = E\left\{\underline{x}^T[n]\underline{x}[n]\right\} = E\left\{\sum_{i=0}^{N-1} x^2[n-i]\right\} = \sum_{i=0}^{N-1} E\left\{x^2[n-i]\right\} = N\sigma_x^2$$

Opt Step Size (3)

Thus multiplying by σ_x^2 leaves us with only one expression to resolve:

$$\sigma_\varepsilon^2[n+1] = \sigma_\varepsilon^2[n] + \mu^2 N \sigma_x^4 - 2\mu\sigma_x^2 E\left\{\varepsilon[n]\,sign\left(\varepsilon[n]+\eta[n]\right)\right\}$$

This expression is simplified by using basic math:

$$E\left\{\varepsilon[n]\,sign\left(\varepsilon[n]+\eta[n]\right)\right\} \overset{\Delta}{=}$$

$$\overset{\Delta}{=} \frac{1}{2\pi\sigma_\varepsilon[n]\sigma_\eta[n]} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \varepsilon[n]\,sign\left(\varepsilon[n]+\eta[n]\right) \exp\left(-\frac{1}{2}\left(\frac{\varepsilon[n]}{\sigma_\varepsilon^2[n]} + \frac{\eta[n]}{\sigma_\eta^2[n]}\right)\right) d\varepsilon[n]\, d\eta[n] =$$

$$= \cdots = \sqrt{\frac{2}{\pi}} \frac{\sigma_\varepsilon^2[n]}{\sqrt{\sigma_\varepsilon^2[n] + \sigma_\eta^2[n]}} = \sqrt{\frac{2}{\pi}} \frac{\sigma_\varepsilon^2[n]}{\sigma_e[n]}$$

Opt Step Size (4)

We receive the final expression:

$$\sigma_\varepsilon^2[n+1] = \sigma_\varepsilon^2[n] + \mu^2 N \sigma_x^4 - \mu\sqrt{\frac{8}{\pi}}\frac{\sigma_x^2\sigma_\varepsilon^2[n]}{\sigma_e[n]}$$

Noticing that the equation is quadratic w.r.t. mue, it can be derived to produce the optimum step size:

$$\frac{d}{d\mu}\left(\sigma_\varepsilon^2[n+1]\right) = \frac{d}{d\mu}\left(\sigma_\varepsilon^2[n] + \mu^2 N \sigma_x^4 - \sqrt{\frac{8}{\pi}}\frac{\mu\sigma_x^2\sigma_\varepsilon^2[n]}{\sigma_e[n]}\right) = 2\mu N \sigma_x^4 - \sqrt{\frac{8}{\pi}}\frac{\sigma_x^2\sigma_\varepsilon^2[n]}{\sigma_e[n]} = 0 \Rightarrow$$

$$\Rightarrow \mu^{OPT} = \sqrt{\frac{2}{\pi}}\frac{\sigma_\varepsilon^2[n]}{N\sigma_x^2\sigma_e[n]}$$

Suboptimal Step Size

- The optimal step size expression $\sqrt{\frac{2}{\pi}}\frac{\sigma_{\varepsilon}^{2}[n]}{N\sigma_{x}^{2}\sigma_{e}[n]}$ includes the residual echo energy level $\sigma_{\varepsilon}^{2}[n]$ which is hard to estimate under persistent DT situation

- Since the steady state ERLE for SA is constant, we can replace the following expression with a constant: $\sqrt{\frac{2}{\pi}}\frac{\sigma_{\varepsilon}^{2}[n]}{N\sigma_{x}^{2}}$

- To produce a suboptimal step size: $\frac{\mu}{\sigma_{e}[n]}$

Derived Algorithm

This finally produces the proposed algorithm:

$$\underline{\hat{w}}[n+1] = \underline{\hat{w}}[n] + \mu[n] \cdot \underline{x}[n] \cdot sign(e[n])$$

$$\mu[n] = \frac{\mu}{\hat{\sigma}_e[n]}$$

$$\hat{\sigma}_e^2[n] = \lambda \hat{\sigma}_e^2[n-1] + (1-\lambda)|e[n]|^2$$

Unified Convergence Limit (1)

Continuing from the residual echo energy levels update equation:

$$\sigma_\varepsilon^2[n+1] = \sigma_\varepsilon^2[n] + \mu^2 N \sigma_x^4 - \sqrt{\frac{8}{\pi}} \frac{\mu \sigma_x^2 \sigma_\varepsilon^2[n]}{\sigma_e[n]}$$

Replacing the step size with the suboptimal step size:

$$\sigma_\varepsilon^2[n+1] = \sigma_\varepsilon^2[n] + \frac{\mu^2 N \sigma_x^4}{\sigma_e^2[n]} - \sqrt{\frac{8}{\pi}} \frac{\mu \sigma_x^2 \sigma_\varepsilon^2[n]}{\sigma_e^2[n]}$$

A geometric series can be derived:

$$\sigma_\varepsilon^2[n+1] = \sigma_\varepsilon^2[n]\left(1 + \frac{\mu^2 N \sigma_x^4}{\sigma_\varepsilon^2[n]\sigma_e^2[n]} - \sqrt{\frac{8}{\pi}} \frac{\mu \sigma_x^2}{\sigma_e^2[n]}\right)$$

Unified Convergence Limit (2)

For the series convergence a step size smaller then 1 is required:

$$\frac{\mu^2 N\sigma_x^4}{\sigma_\varepsilon^2[n]\sigma_e^2[n]} < \sqrt{\frac{8}{\pi}}\frac{\mu\sigma_x^2}{\sigma_e^2[n]}$$

Hence a lower bound on the residual echo energy level is:

$$\sigma_\varepsilon^2[n] > \sqrt{\frac{\pi}{8}}\mu N\sigma_x^2$$

Unified Convergence Limit (3)

- The resulting convergence limit is **independent** of the interference energy levels
- This implies that the algorithm will converge toward the same boundary for a DT situation or in a noisy environment
- This result distinguish the proposed algorithm from the SA or NLMS which have different convergence limits for the clean and noisy cases

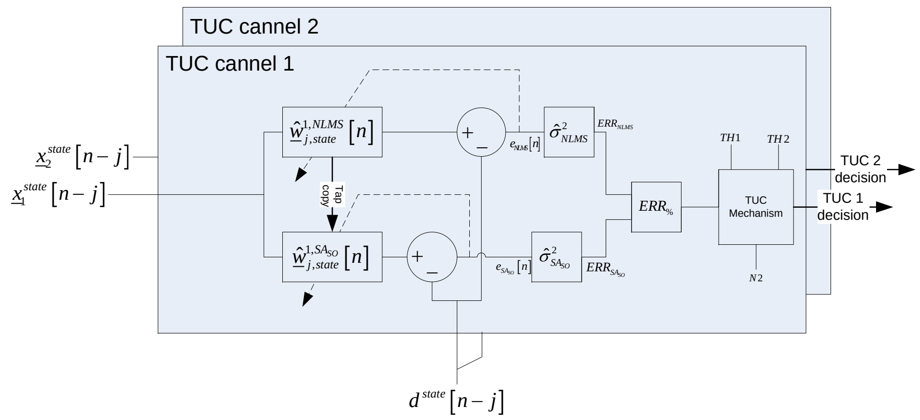
Unified Convergence Limit Algorithm Disadvantages

- The main disadvantage of the algorithm is its initial convergence and re-convergence (lack of) speed caused by the division with $\hat{\sigma}_e[n]$
- To solve this the proposed algorithm can be integrated into the speedy POWER algorithm

POWER and Unified Convergence Limit Integration

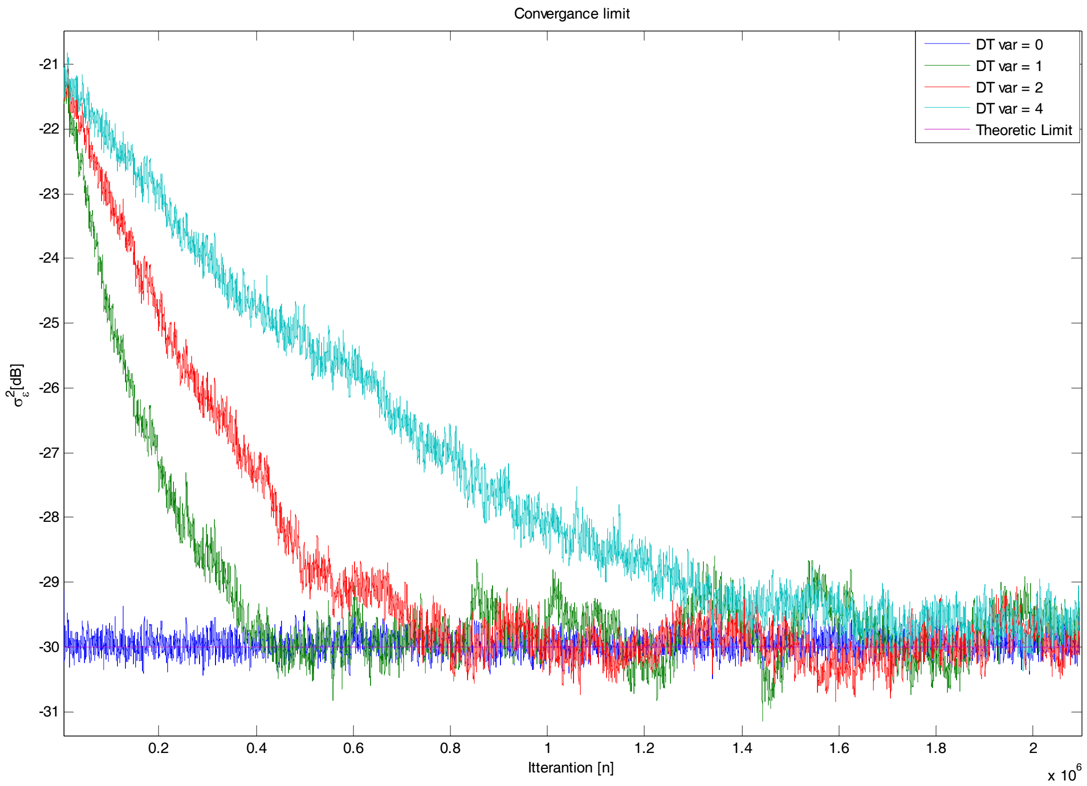
- Since the NLMS algorithm converges fast but will introduce a higher then expected steady state error, the proposed algorithm should be applied as the POWER projection method once the NLMS algorithm has achieved initial convergence or when DT situation is encountered
- Once a re-convergence state was detected the NLMS algorithm should also be reapplied

Projection Method Distinction

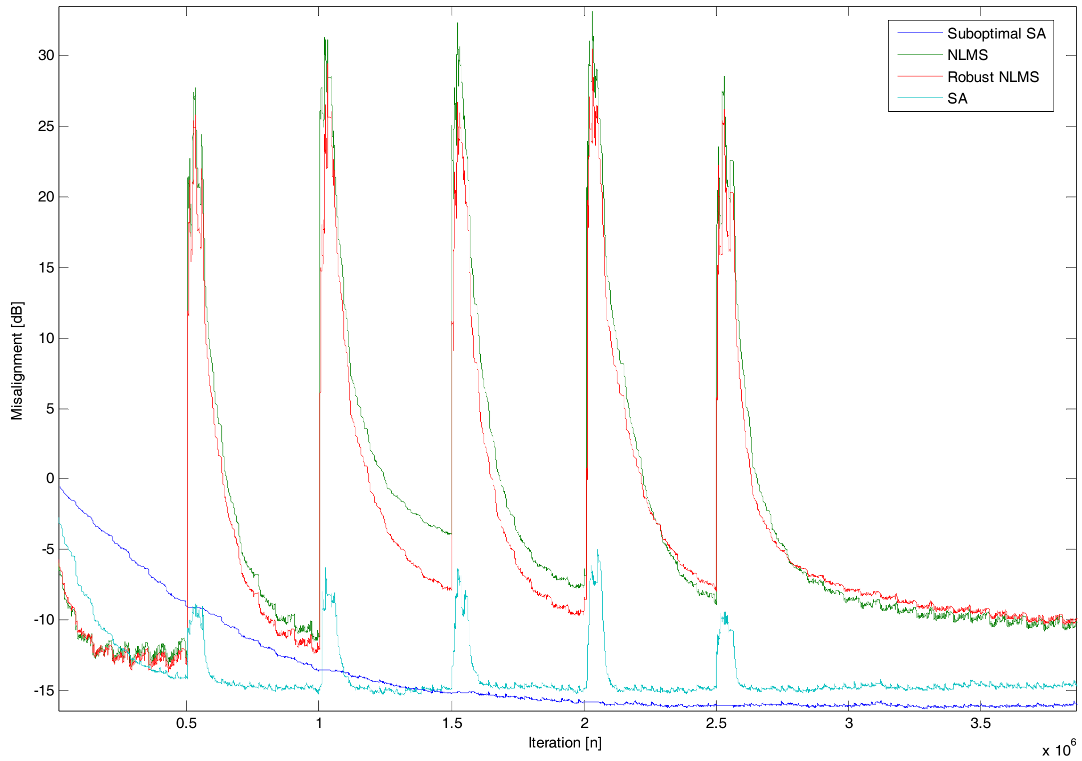


- To distinguish between different states Tap Update Control (TUC) should be added to each POWER Projection block
- The TUC consists of two ghost filters each updated with either NLMS or the unified convergence limit algorithm
- The lower error variance ghost filter will determine the projection state

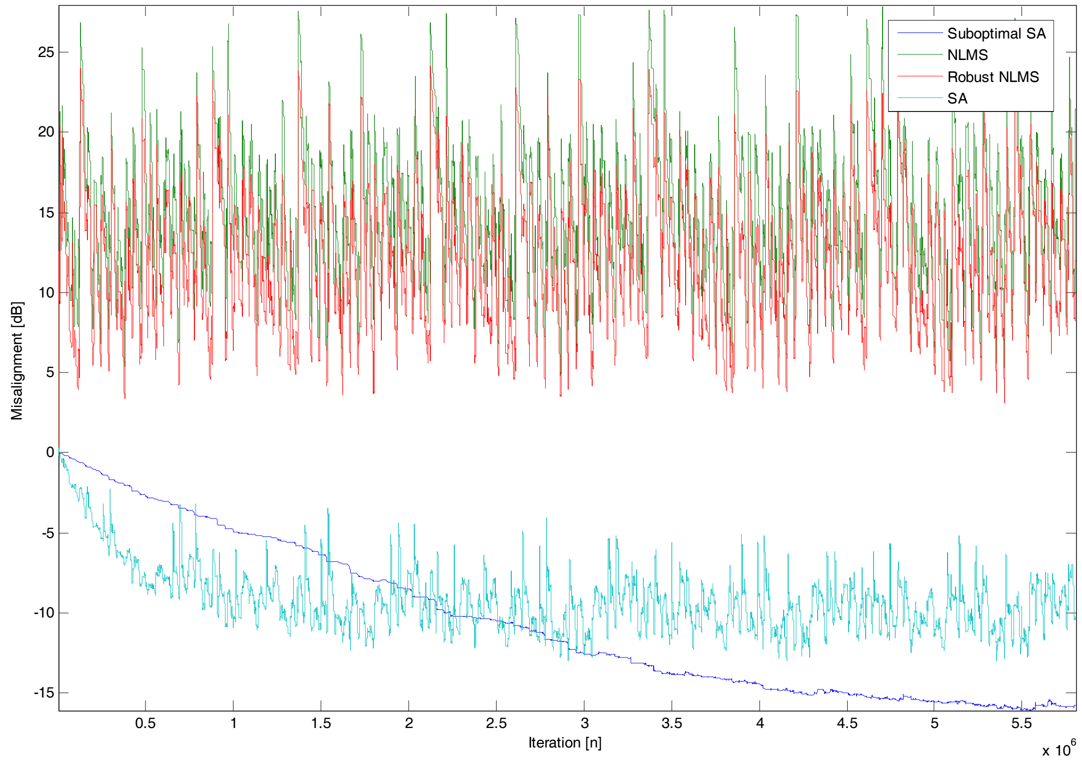
Unified Convergence Limit



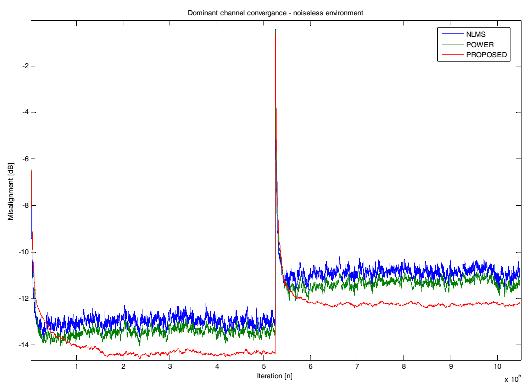
U.C.L. Algorithm under DT

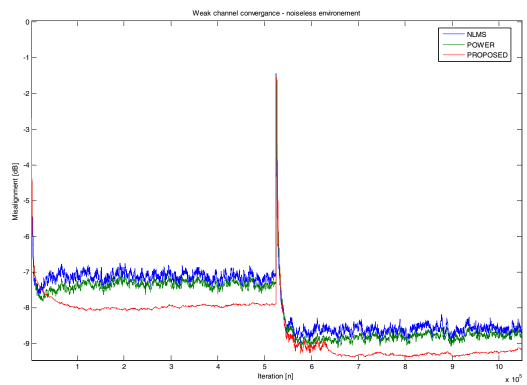


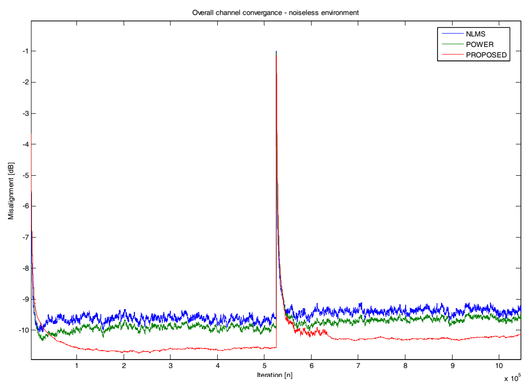
Continuous DT

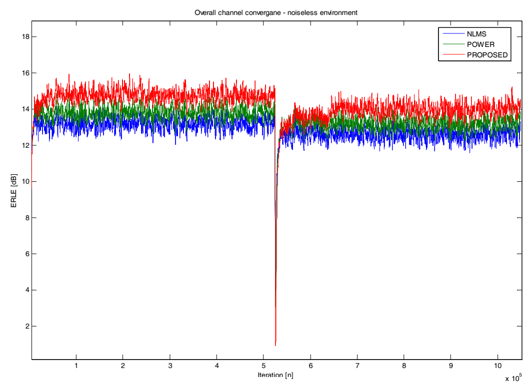


Clean reconvergence

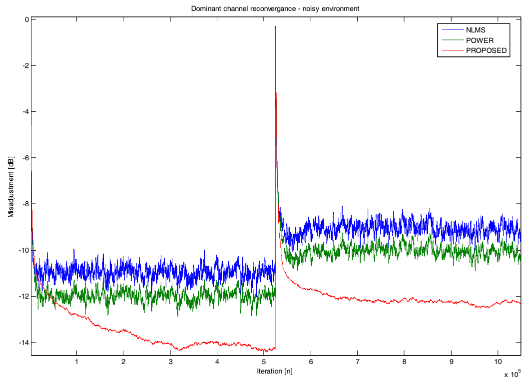


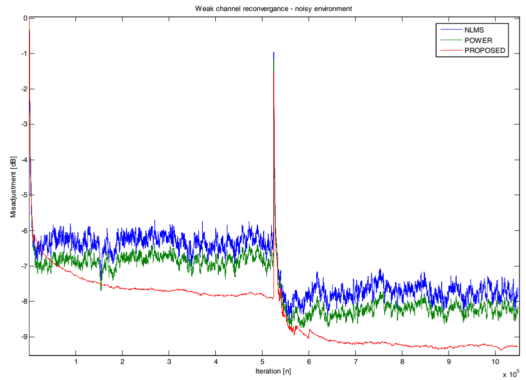


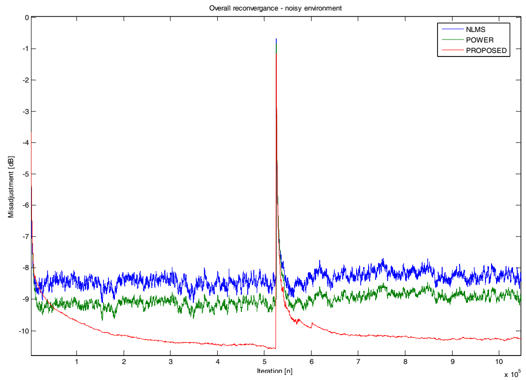


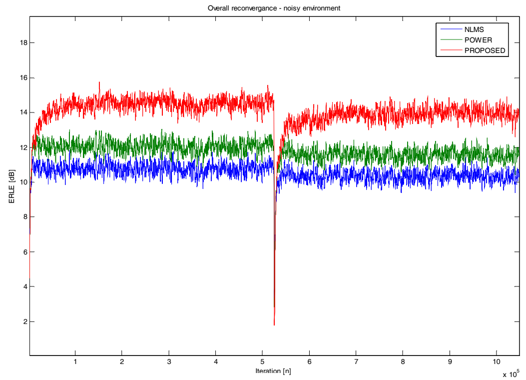


SNR 10 [dB] reconvergence

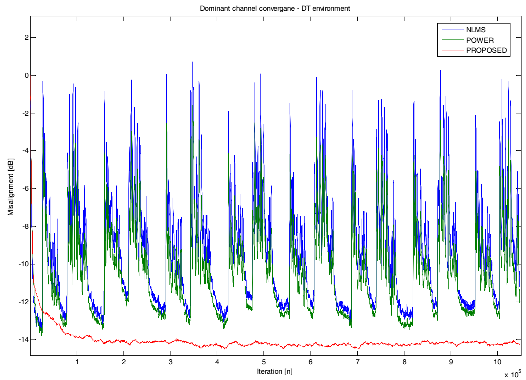


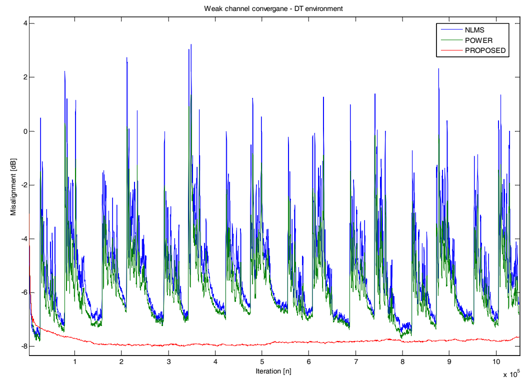


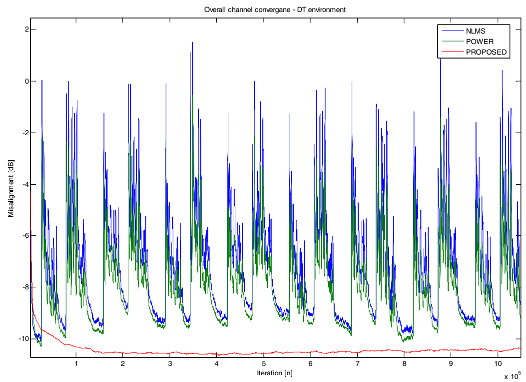


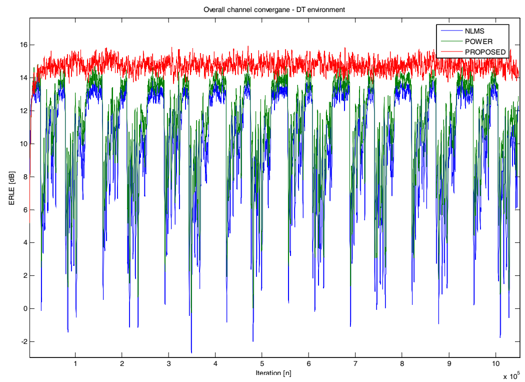


DT SNR 5 [dB] convergence



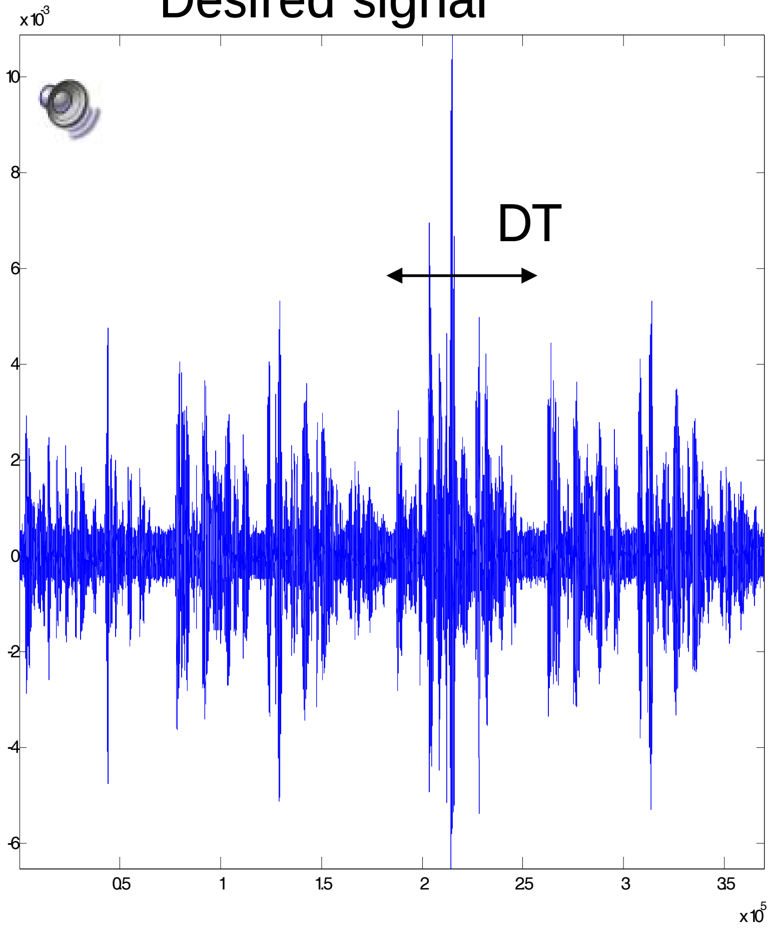




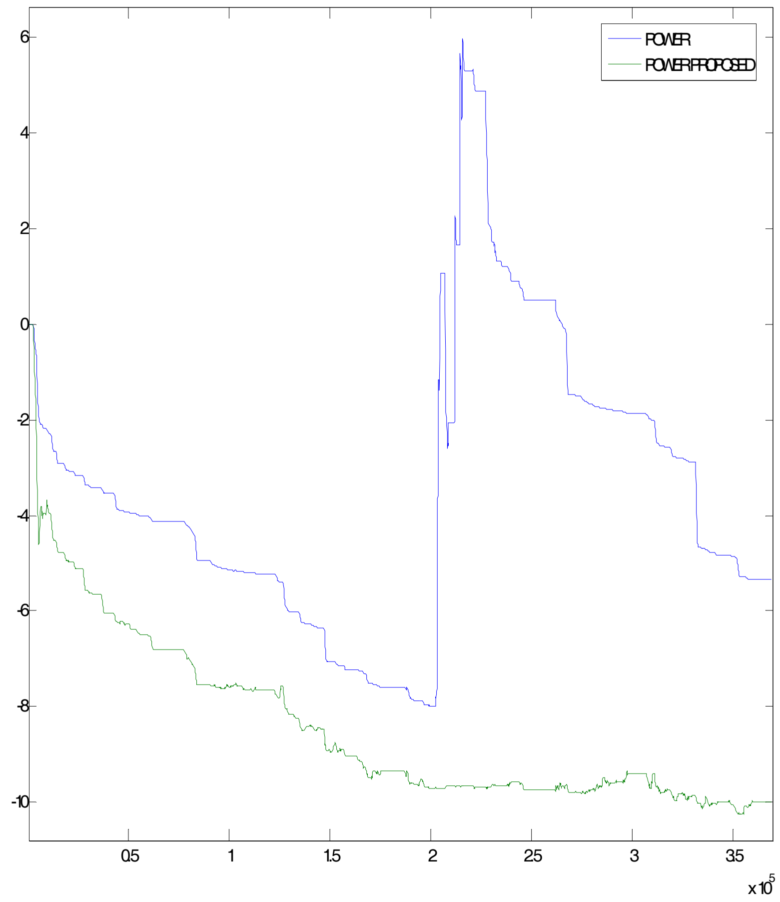


Audio example SNR 10 [dB]

Desired signal

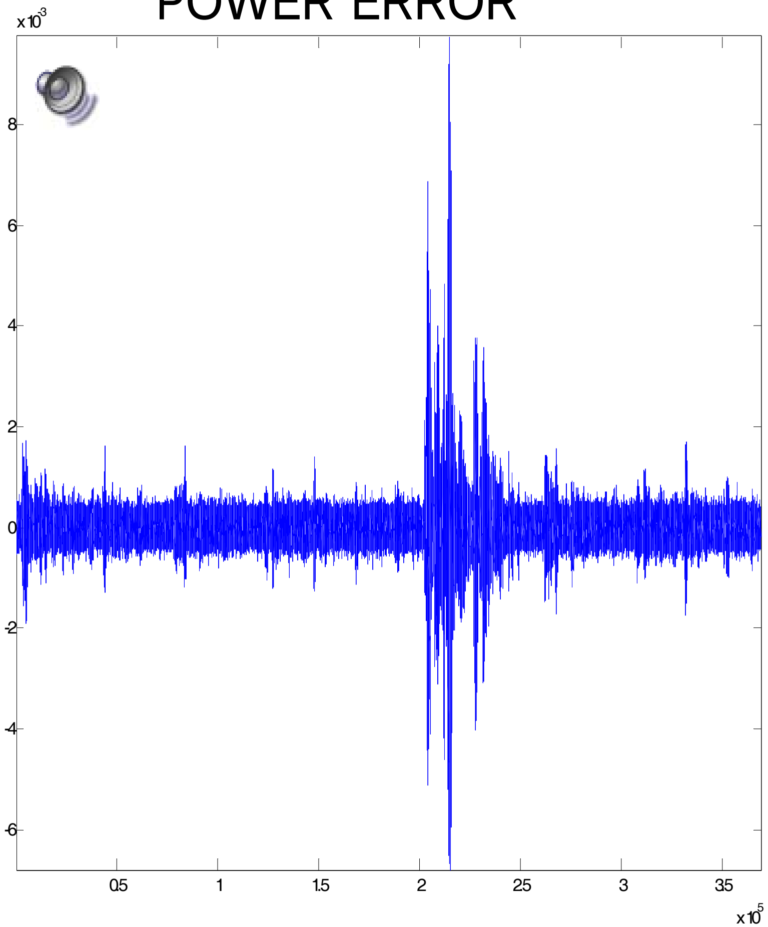


Misalignment

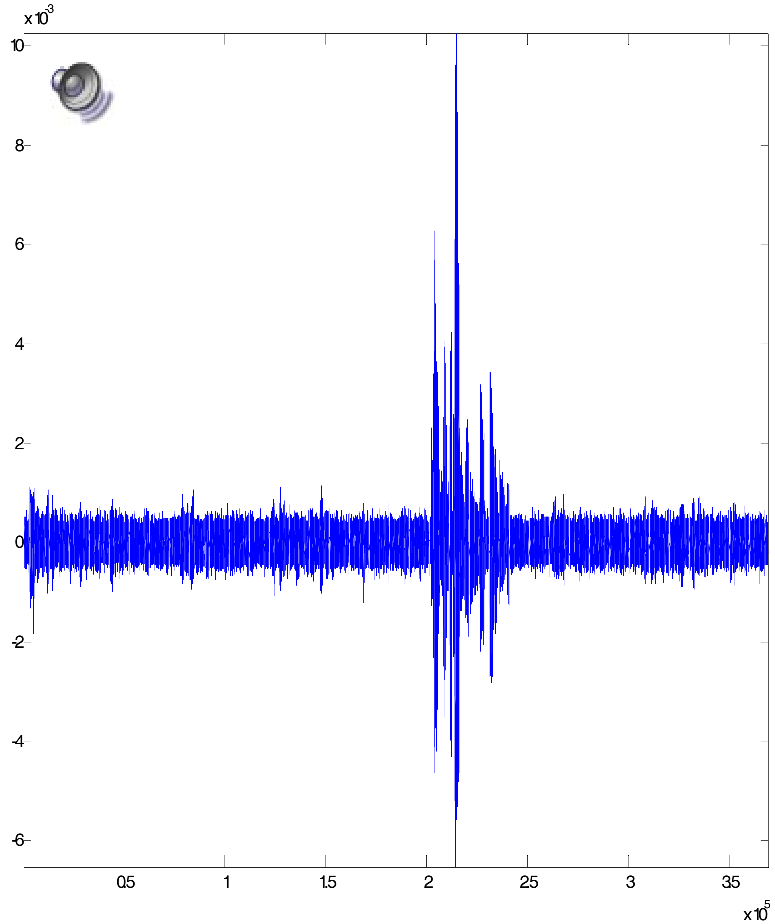


Audio example – Error signals

POWER ERROR



PROPOSED ERROR



Conclusions

- Simpler methods of determining reconvergence scenario
- Use of robust NLMS to avoid divergence during DT misdetection
- Methods of accelerating initial convergence
- Single channel applications – AEC in a noisy environment
- Other voice enhancement algorithms (GSC)

Questions?