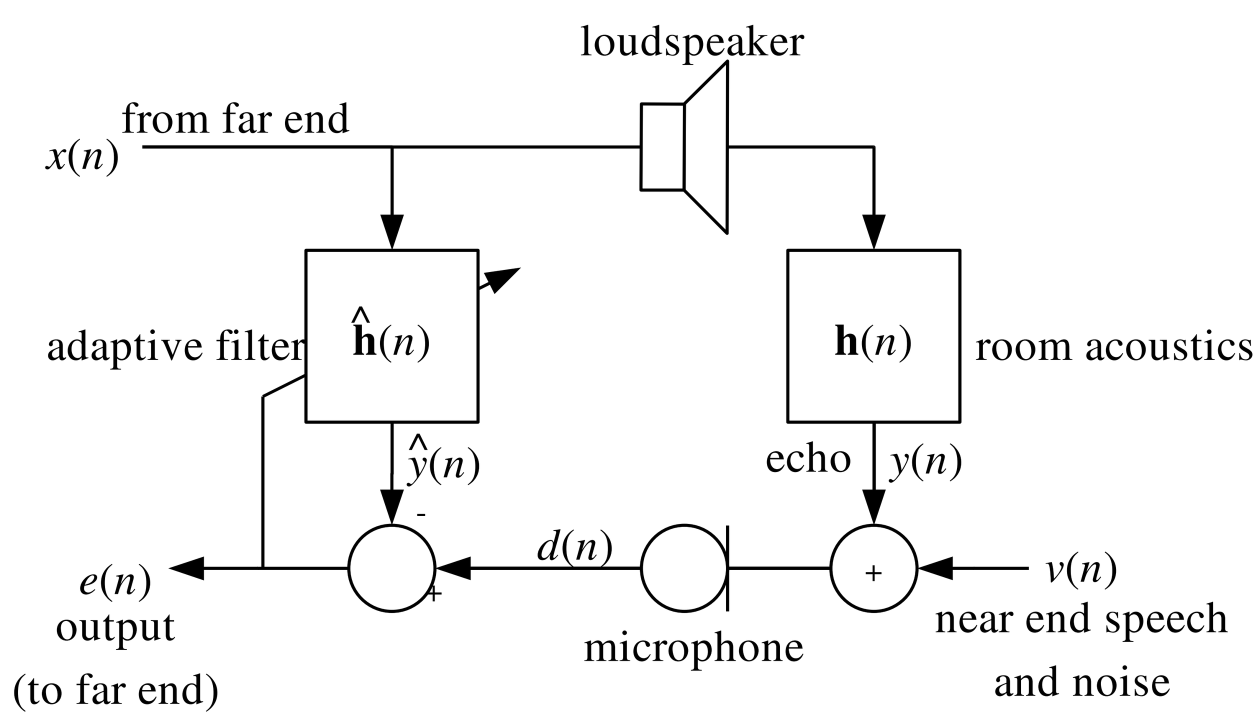
A New Robust Frequency-Domain Echo Canceller With Closed-Loop Learning Rate Adaptation

Context: Acoustic echo cancellation in the presence of double-talk and noise
Problem: How to select the optimal learning rate
Solution: Gradient-adaptive learning rate adaptation



Optimal Learning Rate:

$$\mu_{opt}(n)=\frac{E\left[r^2(n)\right]}{E\left[e^2(n)\right]}$$

- Depends on the Echo Return Loss Enhancement (ERLE)
- ERLE not directly measurable
- Optimal rate never zero when far end is active

Multidelay Block Frequency Domain (MDF) Adaptive Filter

- K blocks of N samples (FFT on $2N$ samples)

MDF algorithm summary

$$\begin{aligned}
\underline{\mathbf{e}}(\ell) &= \underline{\mathbf{d}}(\ell) - \underline{\mathbf{y}}(\ell)\\
\underline{\hat{\mathbf{y}}}(\ell) &= \mathbf{G}_1\underline{\mathbf{X}}(\ell)\,\underline{\hat{\mathbf{h}}}(\ell)\\
\underline{\hat{\mathbf{h}}}(\ell+1) &= \underline{\hat{\mathbf{h}}}(\ell) + \mathbf{G}_2\underline{\boldsymbol{\mu}}(\ell)\,\nabla\underline{\hat{\mathbf{h}}}(\ell)\\
\nabla\underline{\hat{\mathbf{h}}}(\ell) &= \mathbf{\Phi}_{\mathbf{xx}}^{-1}(\ell)\,\underline{\mathbf{X}}^H(\ell)\,\underline{\mathbf{e}}(\ell)
\end{aligned}$$

Proposed Solution

- Frequency-dependent learning rate
- Online optimisation of the learning rate
- Based on the behaviour of the stochastic gradient

$$\mu_k(\ell) = \min\left(\eta(\ell)\frac{\left|\underline{\hat{\mathbf{y}}}_k(\ell)\right|^2}{\left|\underline{\mathbf{e}}_k(\ell)\right|^2}, \mu_0\right)$$

Hard to estimate — Easy to estimate

Main Parameter

$$\eta(\ell+1) = \eta(\ell)\exp\left[\rho\frac{\left|\underline{\hat{\mathbf{y}}}_k(\ell)\right|^2 \Re\left\{\underline{\boldsymbol{\psi}}^H(\ell)\,\nabla\underline{\hat{\mathbf{h}}}(\ell)\right\}}{\left|\underline{\mathbf{e}}_k(\ell)\right|^2\sum_k\left|\Re\left\{\psi_k^*(\ell)\,\nabla\hat{h}_k(\ell)\right\}\right|}\right]$$

$$\underline{\boldsymbol{\psi}}(\ell+1) = \alpha\underline{\boldsymbol{\psi}}(\ell) + \nabla\underline{\hat{\mathbf{h}}}(\ell)$$

smoothed gradient

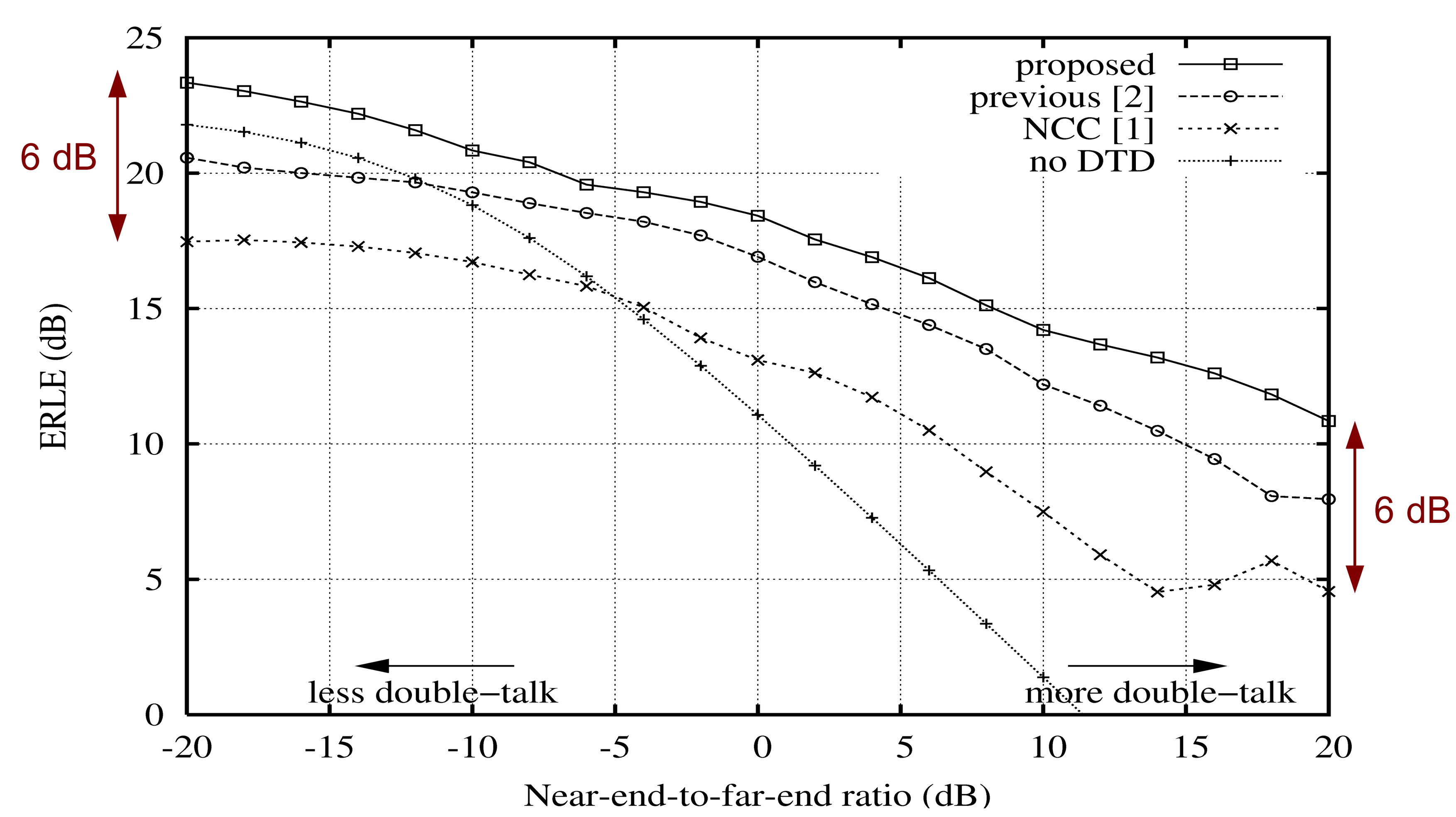
Analysis

- ERLE $\approx 1/\eta$
- When gradient direction is stable, η increases
- When gradient alternates, η decreases
- Double-talk: $e(n)$ increases abruptly, η constant
- Echo path change: gradient stabilises, η increases
- Reacts quickly to double-talk because η only tracks ERLE

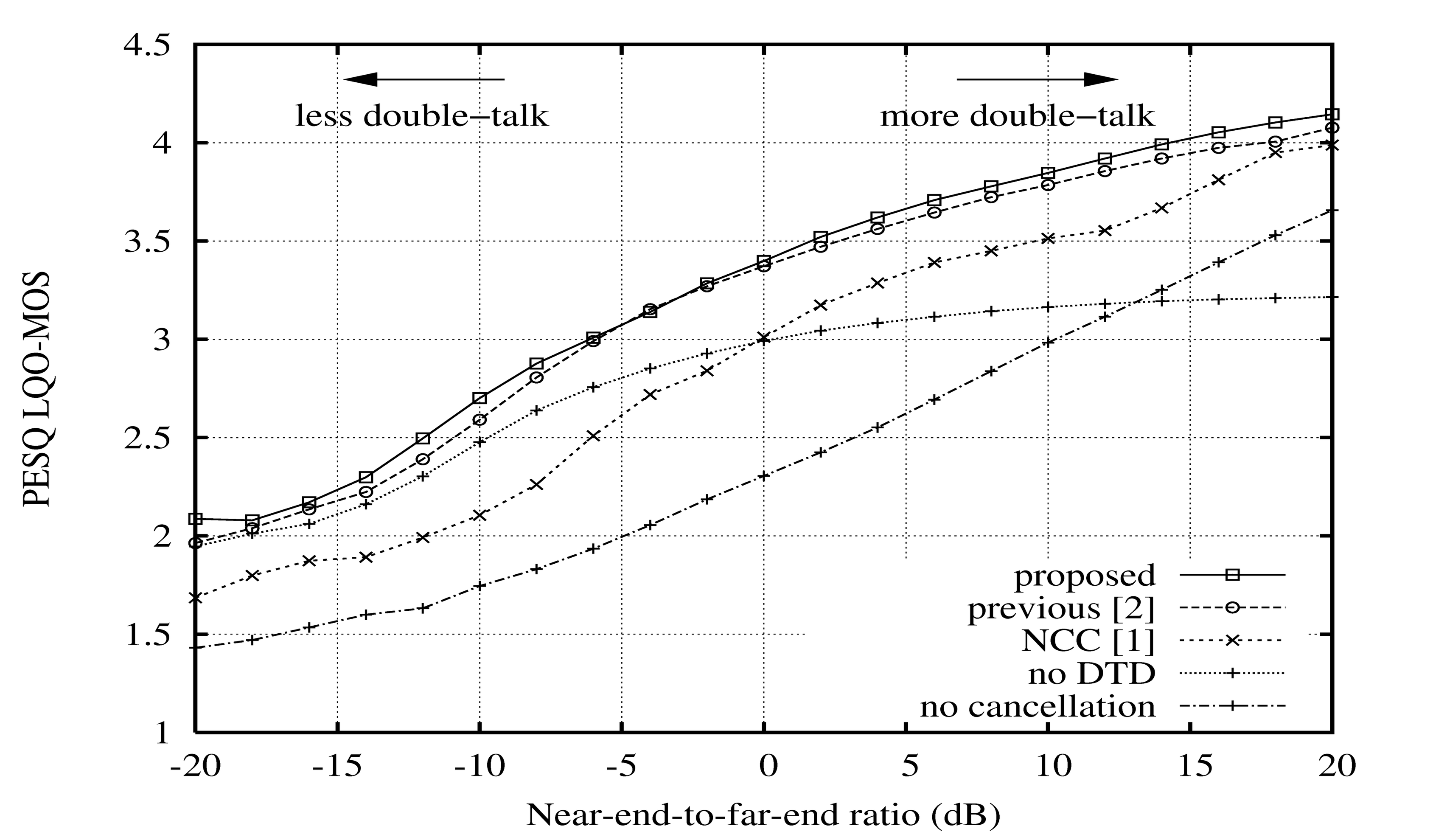
Results

- Tested on a 32-second speech sample
- Background noise and double-talk (-20 dB to +20 dB)
- Echo path change at 16 seconds
- Outperforms explicit double-talk detection by ~6 dB
  - Able to adapt at all time
- Outperforms direct ERLE estimation by ~2 dB
  - Less subject to estimation errors

Steady-state ERLE



PESQ LQ0-MOS



Jean-Marc Valin, Networking Technologies
(jean-marc.valin@csiro.au)
Iain B. Collings, Wireless Technologies
(iain.collings@csiro.au)