A Fast Convergence Two-step Procedure for AR Modeling of Power Spectral Densities

uOttawa
L'Université canadienne
Canada's university

Frédéric Mustiere, Martin Bouchard, Miodrag Bolić
{mustiere,bouchard,mbolic}@site.uottawa.ca

Problem context

- Properly matching the envelope of a discrete lines power spectral density (PSD), with an all-pole auto-regressive (AR) model.
- Such discrete lines PSDs occur for voiced speech, for periodic or quasi-periodic harmonic spectra, etc.
- Simple Yule-Walker (YW) solution does not perform well for such discrete lines PSDs:
  - bias of modeled envelope (or formants) towards the harmonic frequencies
  - large error between AR model response and original harmonic spectral points.
- Other cost functions and approaches have been proposed to better match the harmonic envelope, based on the minimization of other spectral distances (e.g. Itakura-Saito IS, COSH, RMS-log)
- For such cost functions, only simple one-step gradient converging methods have been reported in the literature so far, with the exception of a Quasi-Newton minimization of the IS and COSH distances.

Our Contributions

- To show that the minimization of all popular cost functions can be cast in the exact same framework.
- To propose a fast converging two-step minimization algorithm common to all cost functions, considering as variables the set of normalized AR coefficients and the residual variance.
- To propose a convenient way to compute the gradient for each cost function.

Notation

$P(n)$ two-sided discrete Power Spectral Density with tonal components, sampling the envelope at a subset of N peaks indexed by n, at frequencies ω_n

$\tilde{\mathbf{a}}$ length M+1 vector of unnormalized AR coefficients $\tilde{a}_m$, producing the induced envelope :

$$\hat{P}(n) = \left|\sum_{m=0}^{M} \tilde{a}_m e^{-j\omega_n m}\right|^{-2}$$

$\mathbf{a}$ length M vector of normalized AR coefficients a_m, with a residual variance σ^2, producing the induced envelope:

$$\hat{P}(n) = \sigma^2 \left|1 - \sum_{m=1}^{M} a_m e^{-j\omega_n m}\right|^{-2}$$

$T(n)$ any real symmetric sequence with the dimensions of $P(n)$, with the following operations defined on $T(n)$:

$$r_T(m) = \frac{1}{N}\sum_{n=1}^{N} T(n)\cos(\omega_n m) \quad 0 \le m \le M$$

$$\mathbf{r}_{T,M} = \left[r_T(1) \quad r_T(2) \quad \dots \quad r_T(M)\right]^T$$

$$\mathbf{R}_{T,M} = Toeplitz\left(r_T(0:M-1)\right)$$

Some cost functions

$$E_{YW} = \frac{1}{N}\sum_n \frac{P(n)}{\hat{P}(n)}$$ **Yule Walker**

$$E_{IS} = \frac{1}{N}\sum_n \left[\frac{P(n)}{\hat{P}(n)} - \log\frac{P(n)}{\hat{P}(n)} - 1\right]$$ **Ikatura-Saito**

$$E_{\log RMS} = \frac{1}{2N}\sum_n \left[\log\frac{P(n)}{\hat{P}(n)}\right]^2$$ **RMS-log ratio**

$$E_{COSH} = \frac{1}{N}\sum_n \left[\frac{P(n)}{\hat{P}(n)} + \frac{\hat{P}(n)}{P(n)}\right]$$ **COSH distance**

$$E_{MSE} = \frac{1}{2N}\sum_n \left[P(n) - \hat{P}(n)\right]^2$$ **MS error**

Optimization framework

- Gradient-based descent corresponding to the minimization of the cost functions above always leads to the same gradient form:
  - one for the one-step approach (minimization of the unnormalized AR coefficients)
  - one for the proposed two-step update approach (minimization of the normalized AR coefficients + residual variance).
- For the proposed two-step update, faster convergence is expected because:
  - residual variance is exactly updated conditioned on the latest updated value of the normalized AR coefficients
  - gradient descent is done in a smaller dimension (M rather than M+1).
- Descent scheme for the one-step approach (k is the iteration index and E is one of the cost functions above):

  Until convergence, compute:

$$\tilde{\mathbf{a}}(k+1) = \tilde{\mathbf{a}}(k) - \alpha\nabla E(\tilde{\mathbf{a}}(k))$$

- Descent scheme for the proposed two-step approach: Until convergence, compute:

$$\sigma(k) = \arg\min\left(E(\sigma \mid \mathbf{a}(k))\right) = f_E(\sigma \mid \mathbf{a}(k))$$

$$\mathbf{a}(k+1) = \mathbf{a}(k) - \alpha\nabla E(\mathbf{a}(k), \sigma(k))$$

  where the first step is an exact update.

Cost functions, gradients, and residual variances

- Gradient computations are conveniently carried out by introducing auxiliary sequences called Q for the Itakura-Saito, Log-RMS, COSH, and MSE cost functions:

$$Q_{IS} = P - \hat{P} \qquad Q_{\log RMS} = \hat{P}\log\frac{P}{\hat{P}}$$

$$Q_{COSH} = P - \frac{\hat{P}^2}{P} \qquad Q_{MSE} = (P - \hat{P})\hat{P}^2$$

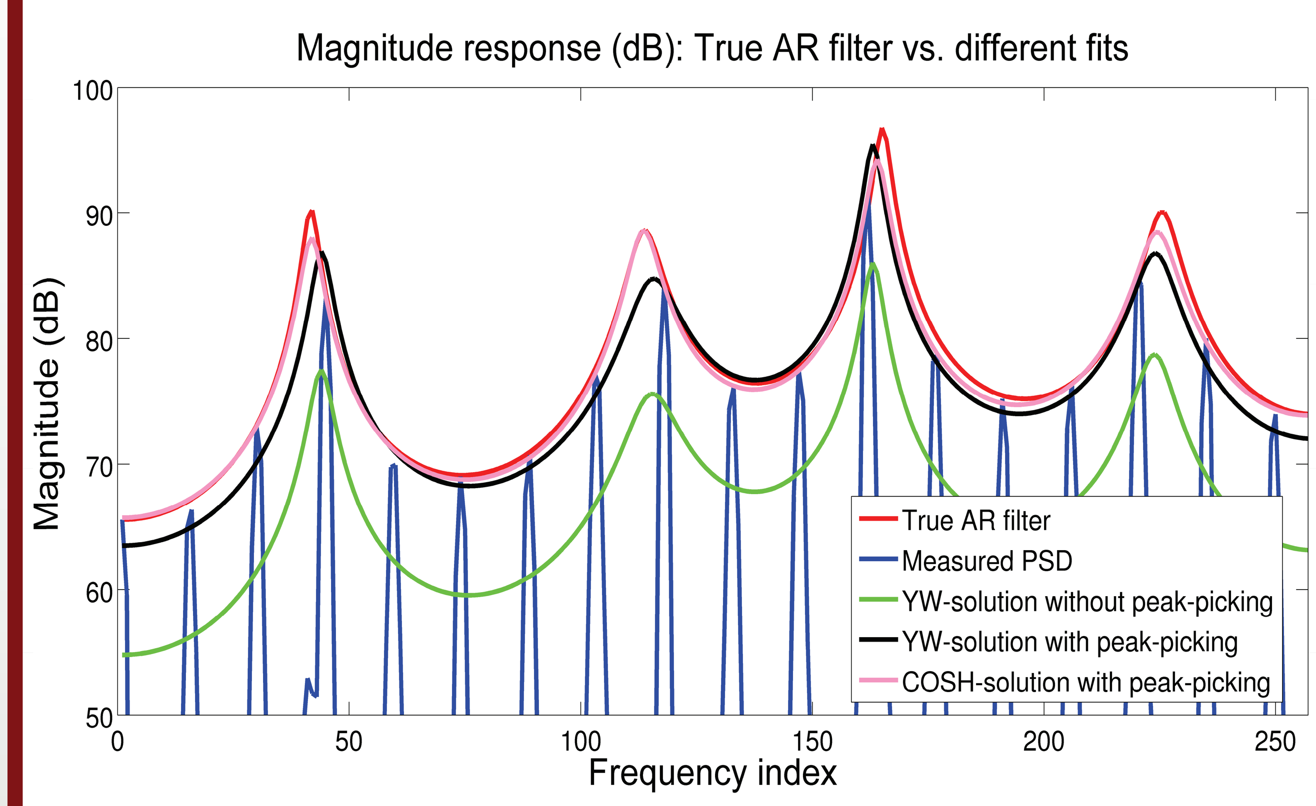
- Gradient computations then simply become:

$$\nabla E(\tilde{\mathbf{a}}(k)) = \mathbf{R}_{Q,M+1}\tilde{\mathbf{a}}$$ (one step approach)

$$\nabla E(\mathbf{a}(k), \sigma(k)) = \sigma^{-2}(k)\left(\mathbf{R}_{Q,M}\mathbf{a} - \mathbf{r}_{Q,M}\right)$$ (two steps approach).

- For the proposed two-steps approach, the functions $\sigma(k) = f_E(\sigma \mid \mathbf{a}(k))$ for the exact computation of the residual variance also have a simple form for the different cost-functions, and they can be found in the paper.

Illustration of spectral matching

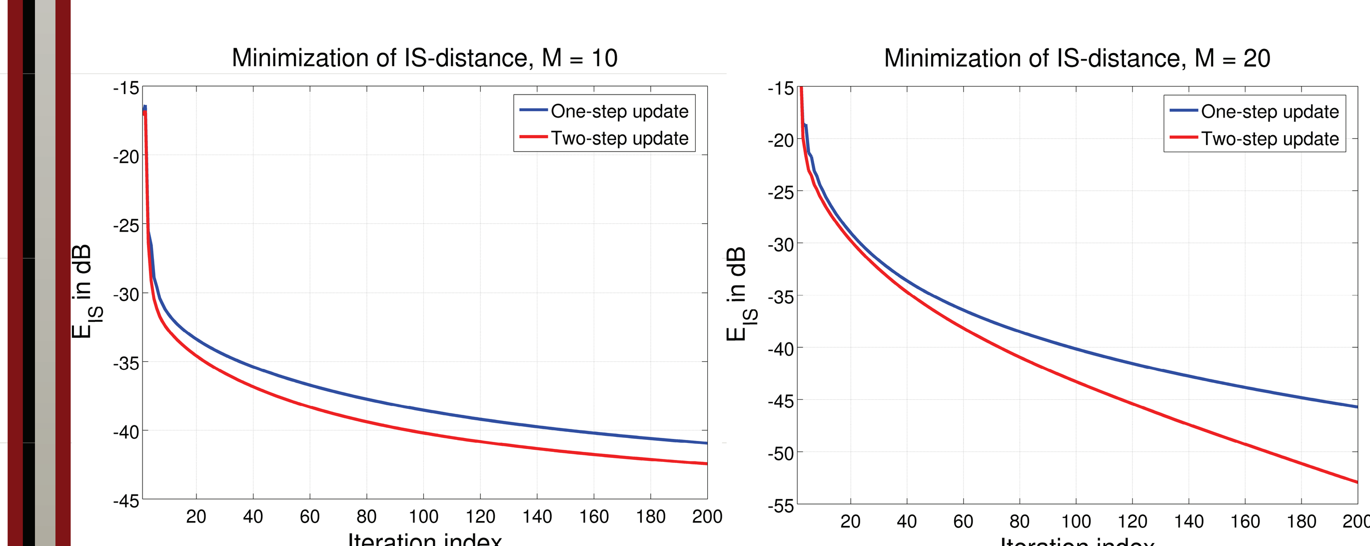


Some convergence results

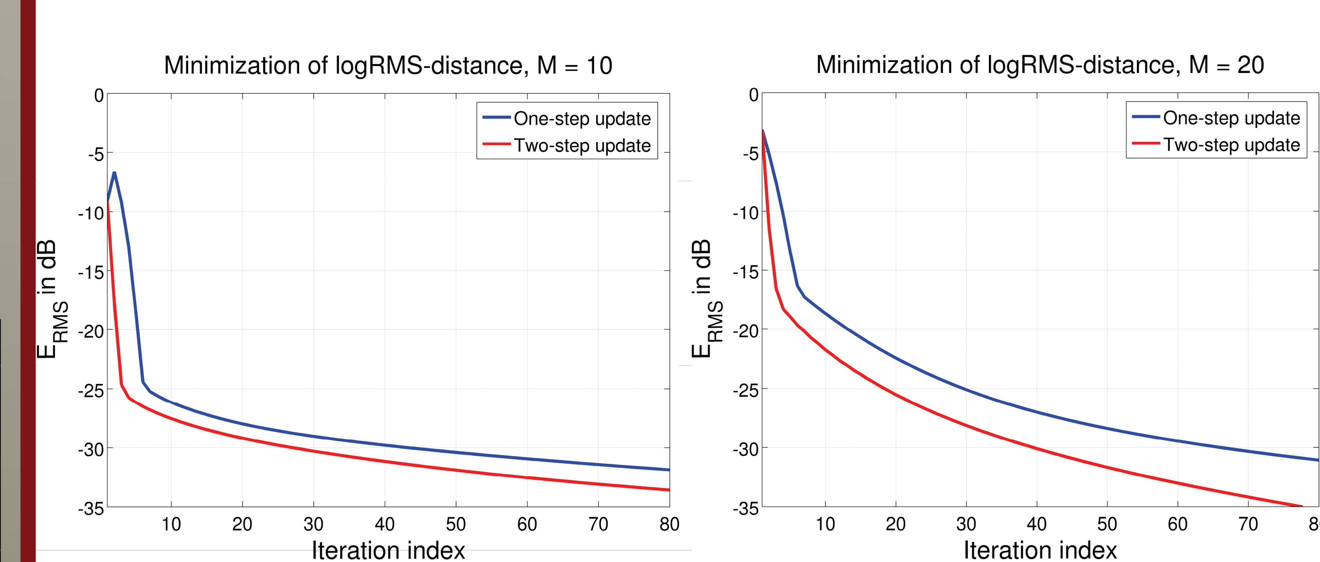
Experimental setup:

- synthetic AR filter excited by periodic impulsions; spectrum of size 512 DFT points (figure above), with peak-picking to select useful harmonic frequency bins (36 peaks)
- comparison of one-step and two-steps gradient descent approaches, with step-sizes for each case adjusted to avoid divergence while achieving fast convergence
- YW solution used as initial condition

Example 1: IS-distance with M = 10 and M = 20



Example 2: RMS-log ratio, M = 10 and M = 20



- The two-step update outperforms the one-step update in terms of convergence properties.
- As the AR order grows, the difference in convergence rate is more pronounced.
- Eventually (after a few thousand iterations) the two methods stabilize to the same value.

Main references

J. Makhoul, "Linear Prediction: A Tutorial Review," *Proceedings of the IEEE*, vol. 63, no. 4, pp. 561–580, 1975.

A. El-Jaroudi and J. Makhoul, "Discrete All-pole Modeling," *IEEE Transactions on Signal Processing*, vol. 39, no. 2, pp. 441–412, 1991.

B. Wei and J. Gibson, "A new Discrete Spectral Modeling method and an application to CELP coding," *IEEE Signal Processing Letters*, vol. 10, pp. 101–103, 2003.