Constrained Pole-Zero Linear Prediction: an Efficient and Near-Optimal Method for Multi-tone Frequency Estimation

Toon van Waterschoot* and Marc Moonen

EUSIPCO 2008



Katholieke Universiteit Leuven, ESAT-SCD,
Kasteelpark Arenberg 10, B-3001 Leuven, Belgium,
http://homes.esat.kuleuven.be/∼tvanwate
toon.vanwaterschoot@esat.kuleuven.be

*This research work was carried out at the ESAT laboratory of the Katholieke Universiteit Leuven, in the frame of K.U.Leuven Research Council CoE EF/05/006 Optimization in Engineering (OPTEC), the Belgian Programme on Interuniversity Attraction Poles, initiated by the Belgian Federal Science Policy Office IUAP P6/04 (DYSCO, 'Dynamical systems, control and optimization', 2007-2011), and EU/FP7-ICT-2007-1 Project 216785 ("Ultra-wide band real-time interference monitoring and cellular management strategies – UCELLS"), and was supported by the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen). The scientific responsibility is assumed by its authors.

Outline

- Introduction
  - Multi-tone frequency estimation
  - Parametric vs. non-parametric frequency estimation
- Constrained Pole-Zero Linear Prediction (CPZLP)
  - Constrained pole-zero signal model
  - Decoupled parameter estimation
  - Line search optimization method
- Performance
  - Computational complexity
  - Simulation results
- Conclusion and ongoing work

Introduction: Multi-tone frequency estimation

- Problem formulation:
  given an observed data sequence

$$y(t) = \sum_{n=1}^{P} A_n \cos(\omega_n t + \phi_n) + r(t),\ t = 1, \ldots, N$$

  estimate the unknown frequencies $\{\omega_1, \ldots, \omega_P\}$.

- Assumptions:
  1. the number of sinuoids P is known
  2. the amplitudes A_n, frequencies ω_n, and phases $\phi_n,\ n = 1, \ldots, P$ are constant in the observed time interval $t \in [1, N]$

- Note:
  - no explicit noise assumptions
  - amplitudes and phases are not estimated

Introduction: Parametric vs. non-parametric estimation

- Non-parametric frequency estimation
  - = first estimate observed signal spectrum, then estimate unknown frequencies from spectrum estimate
  - \+ conceptually simple (mimics "human way" of frequency estimation)
  - \+ reasonable computational complexity: $O(N \log_2 N)$ multiplications when FFT-based
  - – limited accuracy, certainly for short data records
- Parametric frequency estimation
  - = first estimate set of parameters that model the observed signal spectrum, then map estimated parameters to frequency estimates
  - \+ potentially high accuracy ($\approx$ CRLB), even for finite-length data records
  - \+ does not require full spectrum estimate
  - +/– computational complexity depends on estimation method: $O(N) \longrightarrow O(N^3)$ mult.
  - – requires parametric model of observed signal spectrum
  - – requires map between model parameters and unknown frequencies

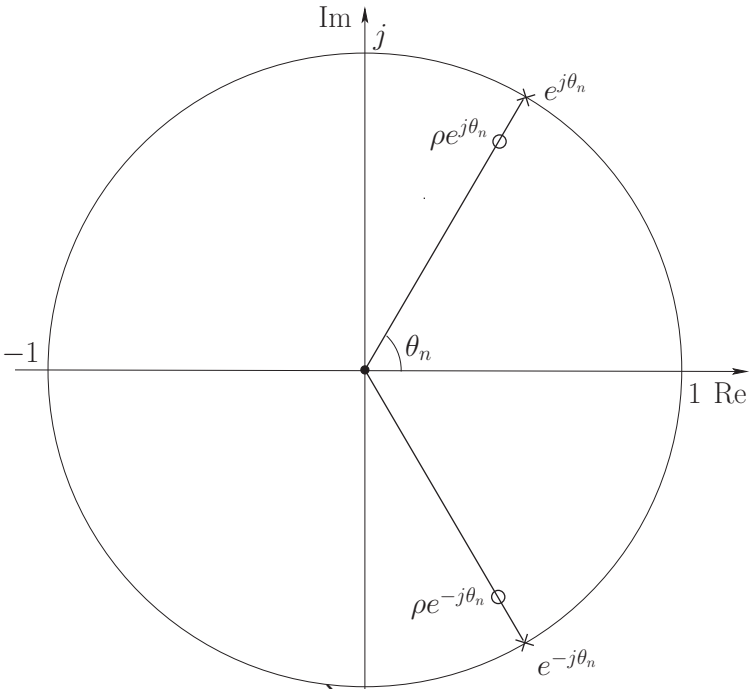
Constrained pole-zero linear prediction: reasonably high accuracy ($\longrightarrow$ CRLB) at $O(N)$ complexity

CPZLP: Constrained pole-zero signal model

- Linear prediction property [Chan *et al.*, 1981]:

$$y(t) = -\sum_{k=1}^{2P} a_k y(t-k) + \sum_{k=1}^{2P} a_k r(t-k) + r(t)$$

with $1 + \sum_{k=1}^{2P} a_k e^{\pm j\omega_n k} = 0,\ n = 1, \ldots, P$



- Constrained pole-zero model [Nehorai, 1985]:
  - cascade of narrowband filters $B \sim (1-\rho)$
  - widely used in adaptive notch filtering (ANF)
  - residual $e(t)$ is not equivalent to noise $r(t)$

$$y(t) = \frac{1 + \sum_{k=1}^{2P} \rho^k a_k z^{-k}}{1 + \sum_{k=1}^{2P} a_k z^{-k}} e(t) = \left(\prod_{n=1}^{P} \frac{1 - 2\rho\cos\theta_n z^{-1} + \rho^2 z^{-2}}{1 - 2\cos\theta_n z^{-1} + z^{-2}} \right) e(t)$$

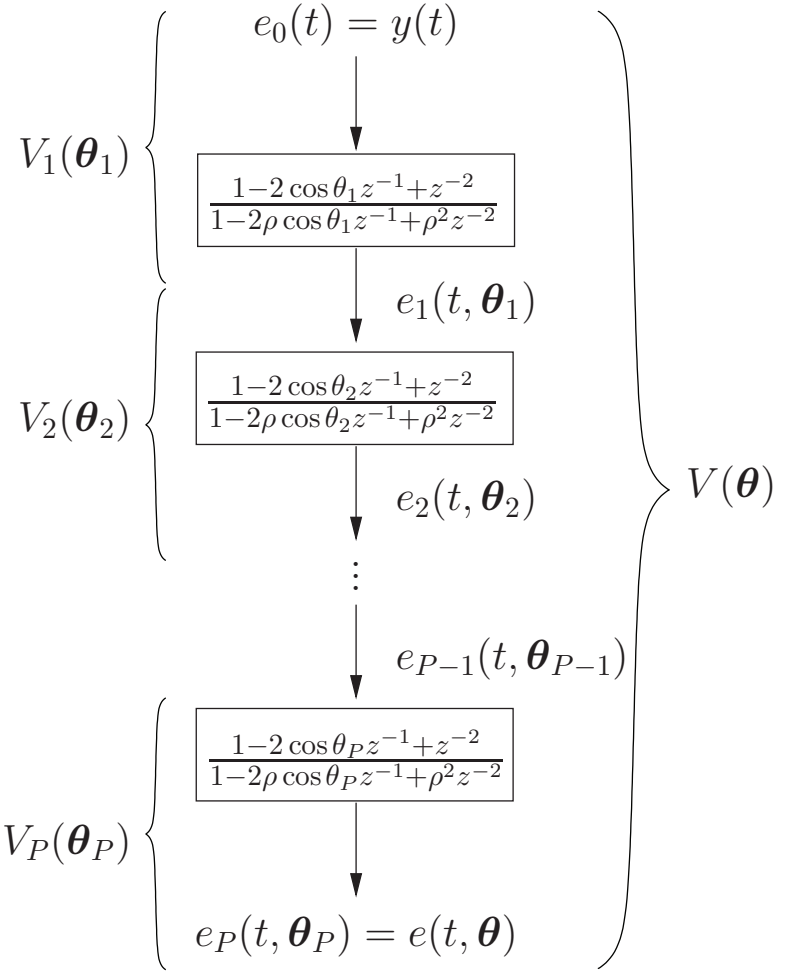
CPZLP: Decoupled parameter estimation

$$\min_{\boldsymbol{\theta}} V(\boldsymbol{\theta}) = \min_{\boldsymbol{\theta}} \frac{1}{N} \sum_{t=1}^{N} e^2(t, \boldsymbol{\theta})$$

divide into $\downarrow$ P subproblems

$$\begin{cases} \min\limits_{\theta_1} V_1(\boldsymbol{\theta}_1) & = & \min\limits_{\theta_1} \dfrac{1}{N} \sum\limits_{t=1}^{N} e_1^2(t, \boldsymbol{\theta}_1) \\ \min\limits_{\theta_2} V_2(\boldsymbol{\theta}_2) & = & \min\limits_{\theta_2} \dfrac{1}{N} \sum\limits_{t=1}^{N} e_2^2(t, \boldsymbol{\theta}_2) \\ \vdots & & \\ \min\limits_{\theta_P} V_P(\boldsymbol{\theta}_P) & = & \min\limits_{\theta_P} \dfrac{1}{N} \sum\limits_{t=1}^{N} e_P^2(t, \boldsymbol{\theta}_P) \end{cases}$$

- $\min V_n(\boldsymbol{\theta}_n)$ w.r.t. θ_n only $\Rightarrow$ scalar subproblems
- solve subproblems consecutively $\Rightarrow$ in subproblem n, estimates of $\theta_1, \ldots, \theta_{n-1}$ are available



CPZLP: Line search optimization method

- Solution to subproblem n: obtained iteratively using a line search optimization approach

$$\theta_n^{(k+1)} = \theta_n^{(k)} + \mu_k p^{(k)}$$

- Step length μ_k: determined using backtracking with Armijo's sufficient decrease condition
- Search direction $p^{(k)}$: obtained with one of the following methods
  - Steepest descent **(SD)**: only gradient information in iteration k
  - Gauss-Newton **(GN)**: Hessian approximation based on gradient information in iteration k
  - Quasi-Newton damped BFGS updating **(BFGS)**: Hessian approximation based on gradient information in iterations 1 to k
- Objective $V_n(\boldsymbol{\theta}_n)$ and gradient $\frac{\partial}{\partial \theta_n} V_n(\boldsymbol{\theta}_n)$: calculated by IIR filtering $e_{n-1}(t, \boldsymbol{\theta}_{n-1})$
  - approximating $\boldsymbol{\theta}_{n-1}$ using solutions of previous subproblems
  - using the entire data record $t = 1, \ldots, N$ of $e_{n-1}(t, \boldsymbol{\theta}_{n-1})$ (key difference with ANF)
  - assuming zero initial conditions ($t \leq 0$) for $e_{n-1}(t, \boldsymbol{\theta}_{n-1})$ and $e_n(t, \boldsymbol{\theta}_n)$

Performance: Computational complexity

- Number of multiplications in iteration k of subproblem n:

calculation of	SD	GN	BFGS
gradient	$10N+6$	$10N+6$	$10N+6$
Hessian	0	N	2
search direction	1	2	2
termination criterion	1	1	1
step length	$(1+\beta_{n,k})(3N+5)$	$(1+\beta_{n,k})(3N+5)$	$(1+\beta_{n,k})(3N+5)$
parameter estimate	1	1	1
TOTAL	$(13+3\beta_{n,k})N+(14+5\beta_{n,k})$	$(14+3\beta_{n,k})N+(15+5\beta_{n,k})$	$(13+3\beta_{n,k})N+(17+5\beta_{n,k})$

- Global number of multiplications:

$$\left.\begin{aligned} M_{\mathsf{SD}} &= \bar{\kappa}_{\mathsf{SD}} P\left[(13+3\bar{\beta}_{\mathsf{SD}})N+(14+5\bar{\beta}_{\mathsf{SD}})\right] \\ M_{\mathsf{GN}} &= \bar{\kappa}_{\mathsf{GN}} P\left[(14+3\bar{\beta}_{\mathsf{GN}})N+(15+5\bar{\beta}_{\mathsf{GN}})\right] \\ M_{\mathsf{BFGS}} &= \bar{\kappa}_{\mathsf{BFGS}} P\left[(13+3\bar{\beta}_{\mathsf{BFGS}})N+(17+5\bar{\beta}_{\mathsf{BFGS}})\right] \end{aligned}\right\} = O(PN)$$

Performance: Simulation results (1)

Monte Carlo simulations:

- varying $N = 64, \ldots, 8192$ (SNR $= 10$ dB)
- varying SNR $= -10, \ldots, 40$ dB ($N = 2048$)
- 100 trials per simulation

Simulation parameters:

- $P = 3$
- $r(t) =$ white/colored noise
- $[A_1, A_2, A_3] = [1, 0.5, 1.5]$
- $[\omega_1, \omega_2, \omega_3] = [0.25, 0.4, 0.7]\pi$
- $[\phi_1, \phi_2, \phi_3] = [0, 0.8, 1.5]\pi$
- fixed $\rho = 0.95$
- $\hat{\theta}_n^{(0)} = \pi/3$

Consistency bounds:

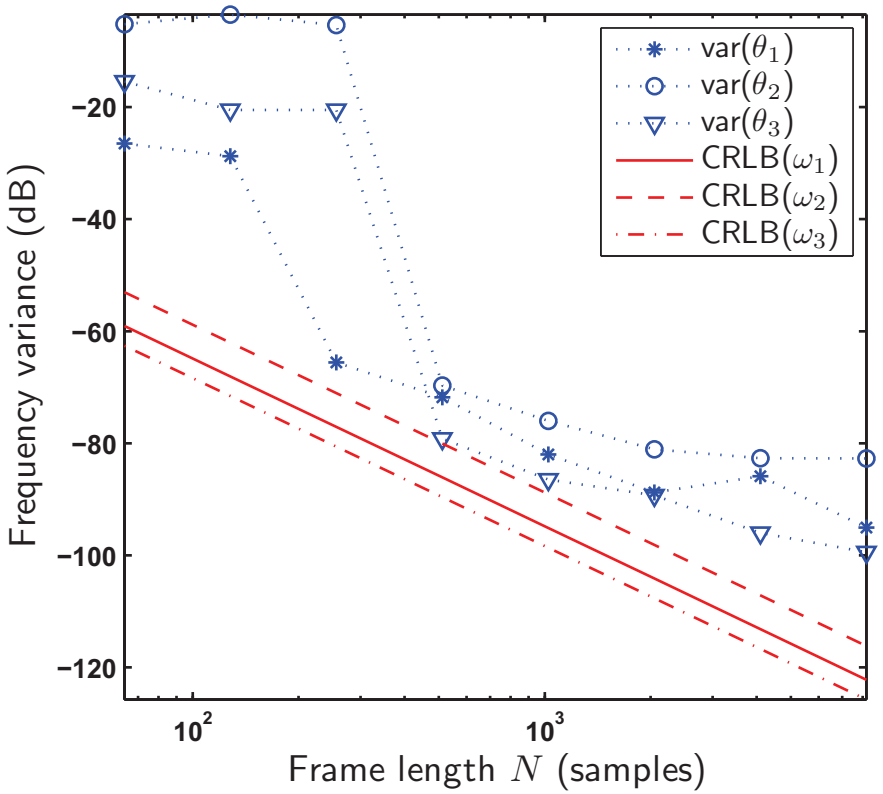
SD	GN	BFGS
$N \geq 256$	$N \geq 512$	$N \geq 512$
SNR≥ 0 dB	SNR≥ 25 dB	SNR≥ 10 dB

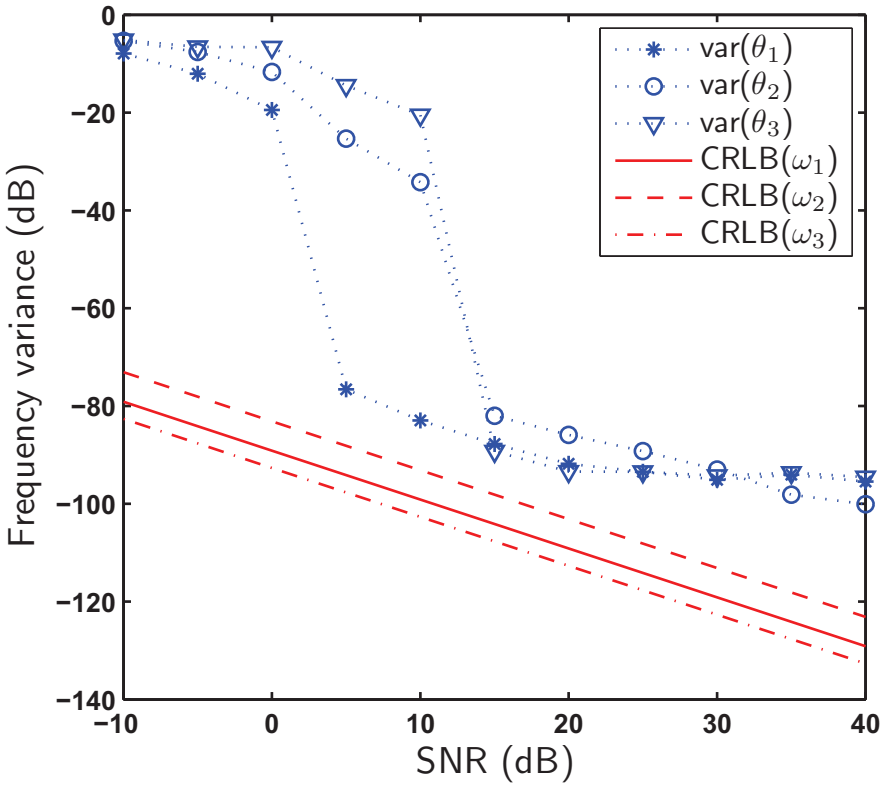
Complexity per sinusoid ($N = 2048$):

SD	GN	BFGS
$\bar{\kappa} = 30 (= k_{\max})$	$\bar{\kappa} = 9.8$	$\bar{\kappa} = 7.8$
$\bar{\beta} = 37.5$	$\bar{\beta} = 0$	$\bar{\beta} = 5.7$
$M = 7.9e6$	$M = 2.8e5$	$M = 4.8e5$

Performance: Simulation results (2)

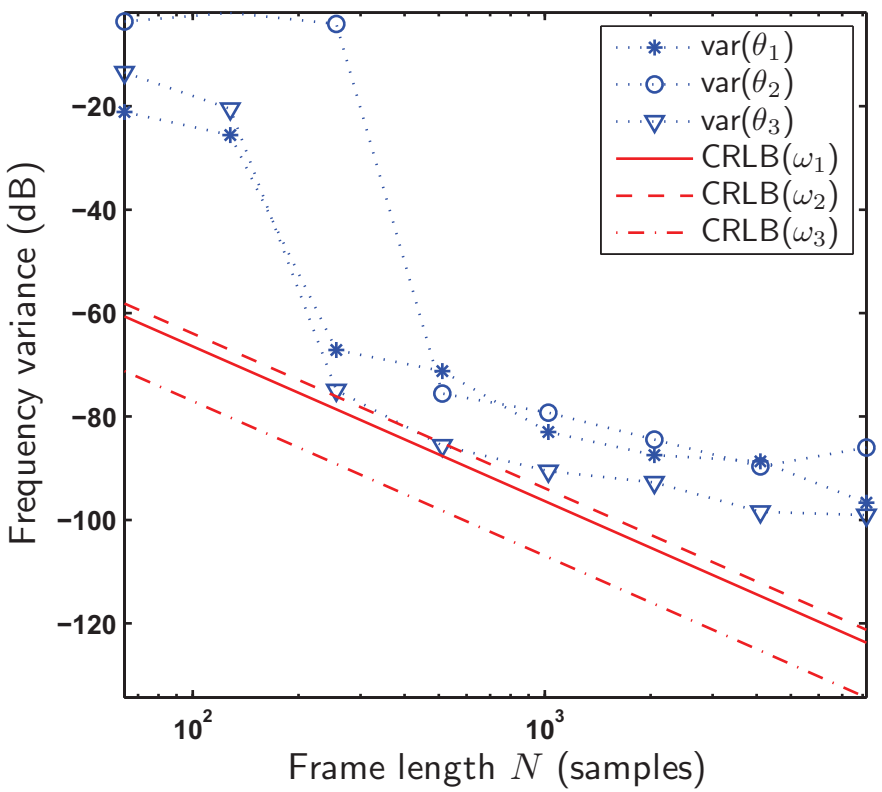
- Monte Carlo simulation results (white noise): frequency variance vs. N and SNR

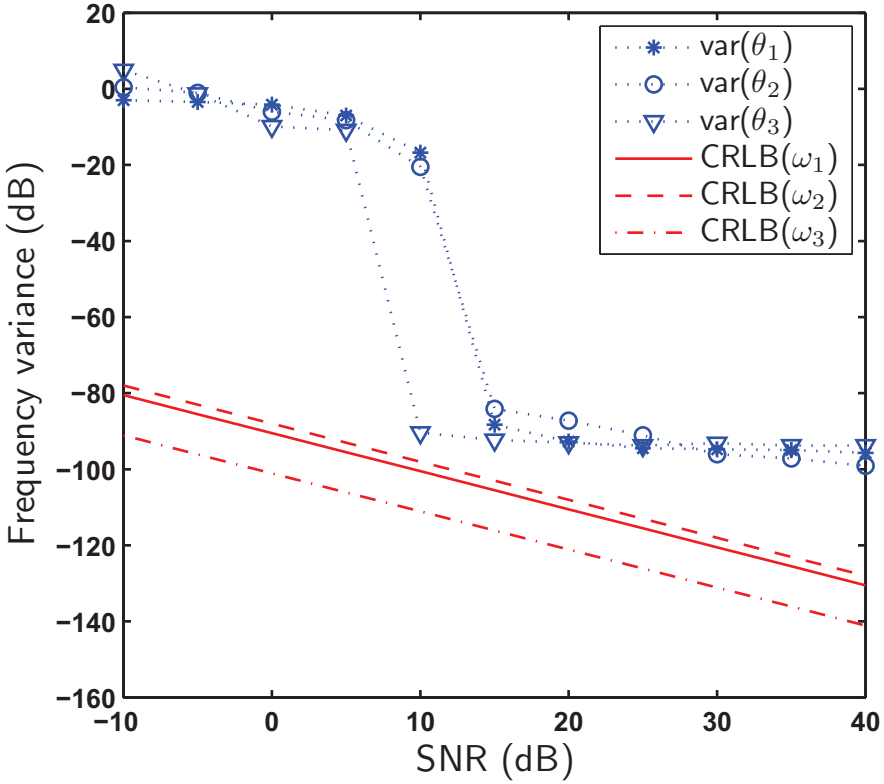




Performance: Simulation results (3)

- Monte Carlo simulation results (colored AR(1) noise): frequency variance vs. N and SNR





Conclusion and ongoing work

- Constrained pole-zero linear prediction (CPZLP):
  - parametric method for multi-tone frequency estimation
  - constrained pole-zero signal model with direct frequency parametrization
  - computational complexity that is linear in data record length and number of sinusoids
  - nearly achieves Cramér-Rao lower bound for reasonable data record length and SNR values
  - no assumptions regarding the noise
- Ongoing work:
  - trust-region i.o. line search optimization method:
    - improve global convergence (by avoiding saddle points and local noise minima)
    - decrease computational complexity (avoid backtracking)
    - easily incorporate bound constraints on parameters
  - joint order and frequency estimation

CPZLP: Objective and gradient calculation

- Objective and gradient:

$$V_n(\boldsymbol{\theta}_n) = \frac{1}{N}\sum_{t=1}^{N} e_n^2(t,\boldsymbol{\theta}_n) \quad \text{and} \quad \frac{\partial}{\partial\theta_n}V_n(\boldsymbol{\theta}_n) = \frac{2}{N}\sum_{t=1}^{N}\left(\frac{\partial}{\partial\theta_n}e_n(t,\boldsymbol{\theta}_n)\right)e_n(t,\boldsymbol{\theta}_n)$$

  can be calculated using $(t = 1, \ldots, N)$

$$e_n(t,\boldsymbol{\theta}_n) = \frac{1 - 2\cos\theta_n z^{-1} + z^{-2}}{1 - 2\rho\cos\theta_n z^{-1} + \rho^2 z^{-2}} e_{n-1}(t,\boldsymbol{\theta}_{n-1})$$

$$\frac{\partial}{\partial\theta_n}e_n(t,\boldsymbol{\theta}_n) = \frac{2(1-\rho)\sin\theta_n z^{-1}(1-\rho z^{-2})}{(1 - 2\rho\cos\theta_n z^{-1} + \rho^2 z^{-2})^2} e_{n-1}(t,\boldsymbol{\theta}_{n-1})$$

- Note:
  - $e_{n-1}(t,\boldsymbol{\theta}_{n-1})$ is available from the final iteration in subproblem $n-1$
  - $e_n(t,\boldsymbol{\theta}_n)$ and $\frac{\partial}{\partial\theta_n}e_n(t,\boldsymbol{\theta}_n)$ can be calculated by IIR filtering $e_{n-1}(t,\boldsymbol{\theta}_{n-1})$, assuming zero initial filter states

Performance: One-shot simulation

Simulation parameters:

- $P = 3$
- $N = 2048$
- $r(t) =$ GWN
- SNR $= 10$ dB
- $[A_1, A_2, A_3] = [1, 0.5, 1.5]$
- $[\omega_1, \omega_2, \omega_3] = [0.25, 0.4, 0.7]\pi$
- $[\phi_1, \phi_2, \phi_3] = [0, 0.8, 1.5]\pi$
- fixed $\rho = 0.95$
- $\hat{\theta}_n^{(0)} = \pi/3$
- method $=$ BFGS

