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

Toon van Waterschoot

SISTA Seminar 21/08/08

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



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
  - Objective and gradient calculation
- 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

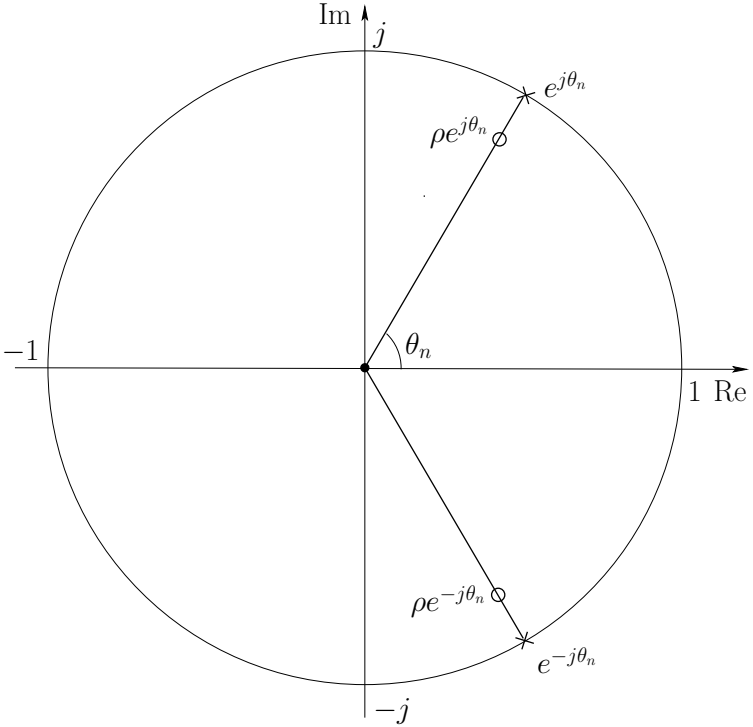
- Non-parametric frequency estimation
  - = first estimate observed signal spectrum, then estimate unknown frequencies from spectrum estimate
  - \+ conceptually simple (mimics "human way" of frequency estimation)
  - \+ reasonable computational complexity: $O(N \log_2 N)$ multiplications when FFT-based
  - – 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, even for short data records
  - \+ does not require full spectrum estimate
  - +/– computational complexity depends on estimation method: $O(N) \rightarrow O(N^3)$ multiplications
  - – requires parametric model of observed signal spectrum
  - – requires map between model parameters and unknown frequencies

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
  - 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

- Global least-squares (LS) parameter estimation problem:

$$\min_{\boldsymbol{\theta}} V(\boldsymbol{\theta}) = \min_{\boldsymbol{\theta}} \frac{1}{N} \sum_{t=1}^{N} e^2(t, \boldsymbol{\theta}) \text{ with } \left\{ \begin{array}{rcl} \boldsymbol{\theta} & = & \begin{bmatrix} \theta_1 & \dots & \theta_N \end{bmatrix}^T \\ e(t, \boldsymbol{\theta}) & = & \left(\prod_{n=1}^{P} \frac{1 - 2\cos\theta_n z^{-1} + z^{-2}}{1 - 2\rho\cos\theta_n z^{-1} + \rho^2 z^{-2}} \right) y(t) \end{array} \right.$$

  can be divided into N subproblems as follows,

- Decoupled LS parameter estimation problem: $(n = 1, \dots, P)$

$$\min_{\theta_n} V_n(\boldsymbol{\theta}_n) = \min_{\theta_n} \frac{1}{N} \sum_{t=1}^{N} e_n^2(t, \boldsymbol{\theta}_n) \text{ with } \left\{ \begin{array}{rcl} \boldsymbol{\theta}_n & = & \begin{bmatrix} \theta_1 & \dots & \theta_n \end{bmatrix}^T \\ e_n(t, \boldsymbol{\theta}_n) & = & \left(\prod_{l=1}^{n} \frac{1 - 2\cos\theta_l z^{-1} + z^{-2}}{1 - 2\rho\cos\theta_l z^{-1} + \rho^2 z^{-2}} \right) y(t) \end{array} \right.$$

- Note:
  - in subproblem n, the minimization is performed w.r.t. θ_n only
  - if the subproblems are solved consecutively, then estimates of $\theta_1, \dots, \theta_{n-1}$ are available when solving subproblem n

CPZLP: Line search optimization method

- The solution to subproblem n is obtained iteratively using a line search optimization approach:

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

- The step length μ_k is determined using backtracking with Armijo's sufficient decrease condition.
- The search direction $p^{(k)}$ can be obtained with one of the following methods:
  - Steepest descent (SD):

$$p^{(k)} = -\frac{\partial}{\partial \theta_n} V_n\left(\hat{\boldsymbol{\theta}}_n^{(k)}\right) \text{ with } \hat{\boldsymbol{\theta}}_n^{(k)} = \begin{bmatrix} \hat{\theta}_1^{(\kappa_1)} & \dots & \hat{\theta}_{n-1}^{(\kappa_{n-1})} & \hat{\theta}_n^{(k)} \end{bmatrix}^T$$

  - Gauss-Newton (GN):

$$p^{(k)} = -\frac{\left(\frac{\partial}{\partial \theta_n}\mathbf{e}_n\left(\hat{\boldsymbol{\theta}}_n^{(k)}\right)\right)^T \mathbf{e}_n\left(\hat{\boldsymbol{\theta}}_n^{(k)}\right)}{\left(\frac{\partial}{\partial \theta_n}\mathbf{e}_n\left(\hat{\boldsymbol{\theta}}_n^{(k)}\right)\right)^T \left(\frac{\partial}{\partial \theta_n}\mathbf{e}_n\left(\hat{\boldsymbol{\theta}}_n^{(k)}\right)\right)} \text{ with } \mathbf{e}_n(\boldsymbol{\theta}_n) = \begin{bmatrix} e_n(1, \boldsymbol{\theta}_n) & \dots & e_n(N, \boldsymbol{\theta}_n) \end{bmatrix}^T$$

– Quasi-Newton with damped BFGS updating (BFGS):

$$p^{(k)} = -B_k^{-1} \frac{\partial}{\partial \theta_n} V_n\big(\hat{\boldsymbol{\theta}}_n^{(k)}\big)$$

with

$$\begin{aligned} s_k &= \hat{\theta}_n^{(k+1)} - \hat{\theta}_n^{(k)} \\ v_k &= \frac{\partial}{\partial \theta_n} V_n\big(\hat{\boldsymbol{\theta}}_n^{(k+1)}\big) - \frac{\partial}{\partial \theta_n} V_n\big(\hat{\boldsymbol{\theta}}_n^{(k)}\big) \\ B_{k+1} &= \max\Big(\frac{v_k}{s_k}, \gamma B_k\Big) \end{aligned}$$

and $\gamma \in (0, 1)$ (usually, $\gamma = 0.2$ [Powell, 1977]).

- The iterative algorithm is terminated at iteration κ_n, either when $\kappa_n = k_{\max}$, or

$$\left| p^{(\kappa n)} \frac{\partial}{\partial \theta_n} V_n\big(\hat{\boldsymbol{\theta}}_n^{(\kappa n)}\big) \right| \leq \tau$$

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: 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	$(13 + 3\beta_{n,k})N + (14 + 5\beta_{n,k})$	$(14 + 3\beta_{n,k})N + (15 + 5\beta_{n,k})$	$(13 + 3\beta_{n,k})N + (17 + 5\beta_{n,k})$

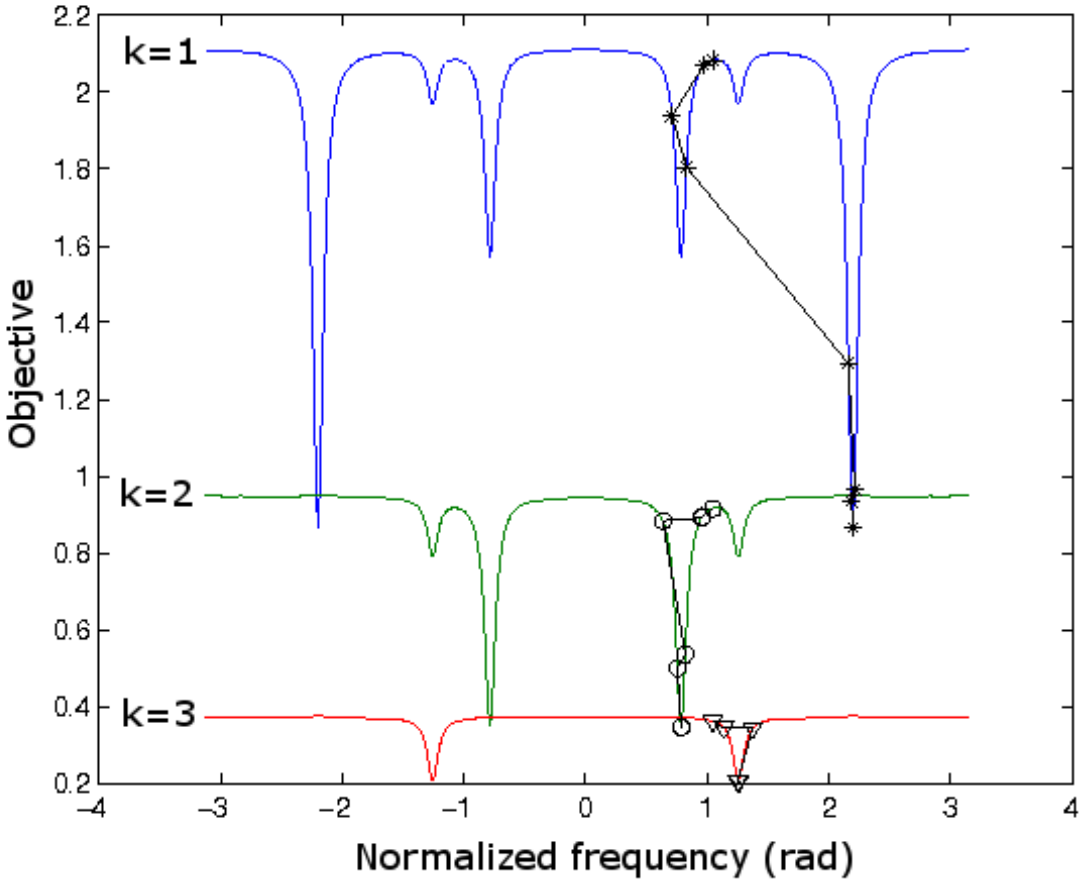
- Global number of multiplications:

$$M_{\text{SD}} = \bar{\kappa} P\big[(13 + 3\bar{\beta})N + (14 + 5\bar{\beta})\big] = O(PN)$$
$$M_{\text{GN}} = \bar{\kappa} P\big[(14 + 3\bar{\beta})N + (15 + 5\bar{\beta})\big] = O(PN)$$
$$M_{\text{BFGS}} = \bar{\kappa} P\big[(13 + 3\bar{\beta})N + (17 + 5\bar{\beta})\big] = O(PN)$$

Performance: Simulation results (1)

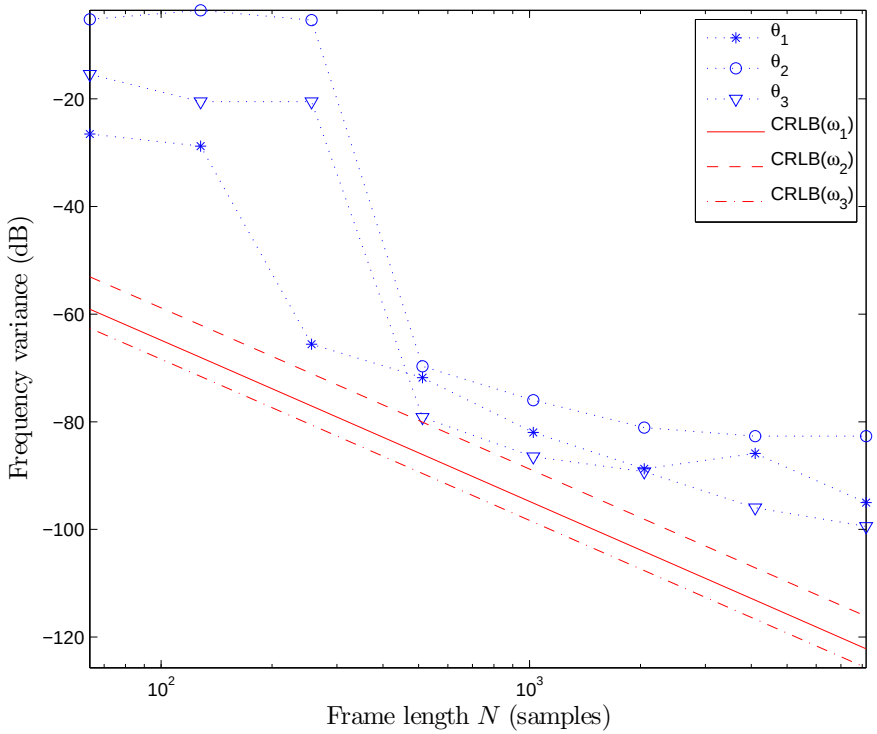
Simulation parameters:

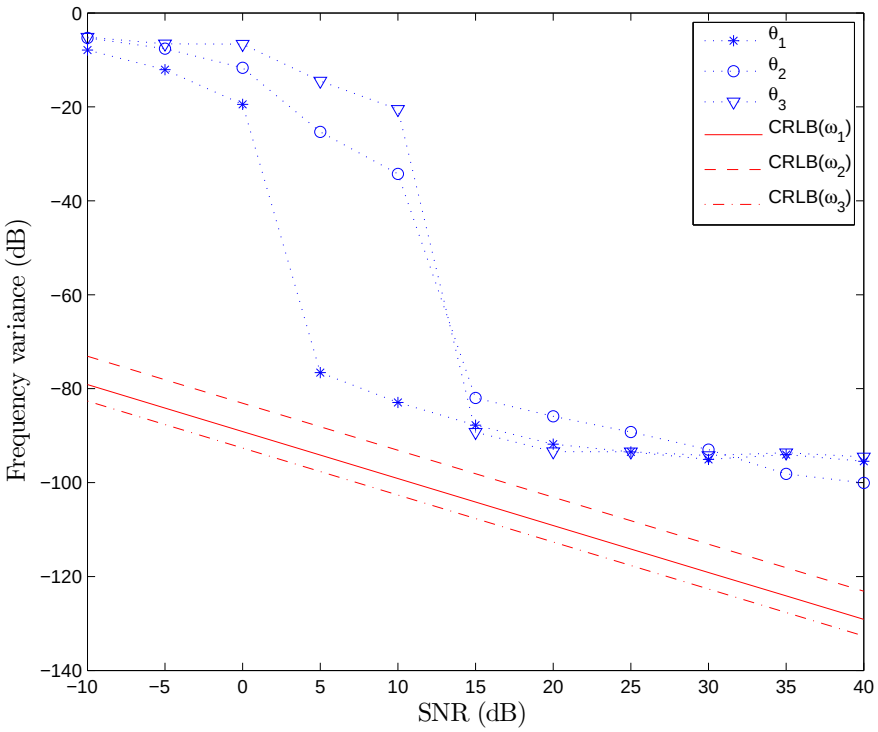
- $P = 3$
- $N = 2048$
- $r(t) =$ GWN
- SNR $= 10$ dB
- $[A_1, A_2, A_3] = [1, 0.5, 1.5]$
- $[\omega_1, \omega_2, \omega_3] = [0.25, 0.4, 0.7]\pi$
- $[\phi_1, \phi_2, \phi_3] = [0, 0.8, 1.5]\pi$
- fixed $\rho = 0.95$
- $\hat{\theta}_n^{(0)} = \pi/3$
- method $=$ BFGS



CPZLP: Simulation results (2)

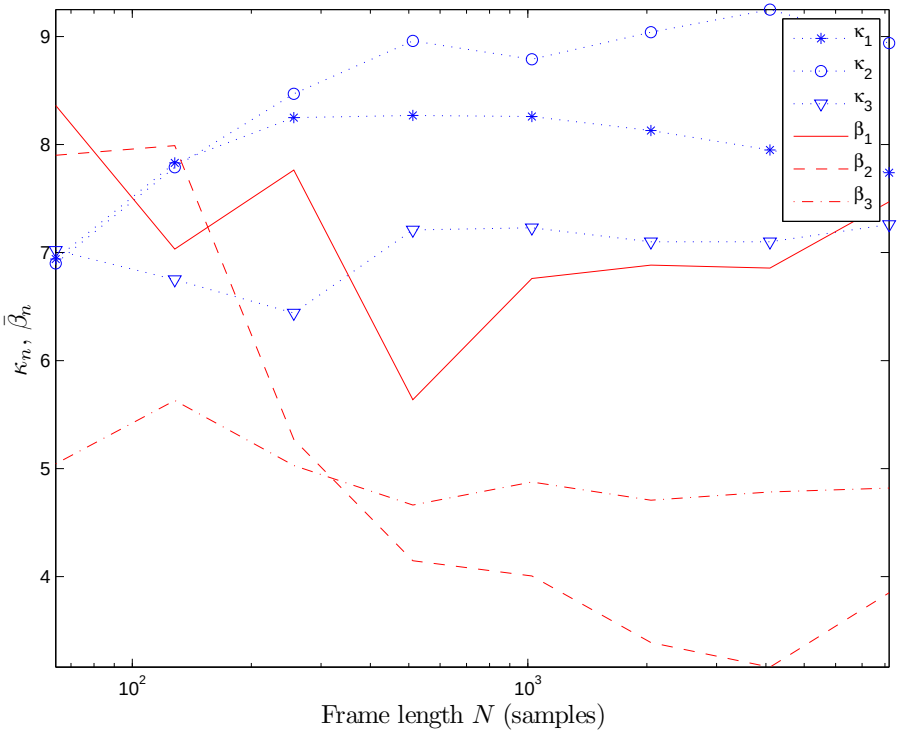
- Monte Carlo simulation results (100 trials per simulation): frequency variance vs. L and SNR

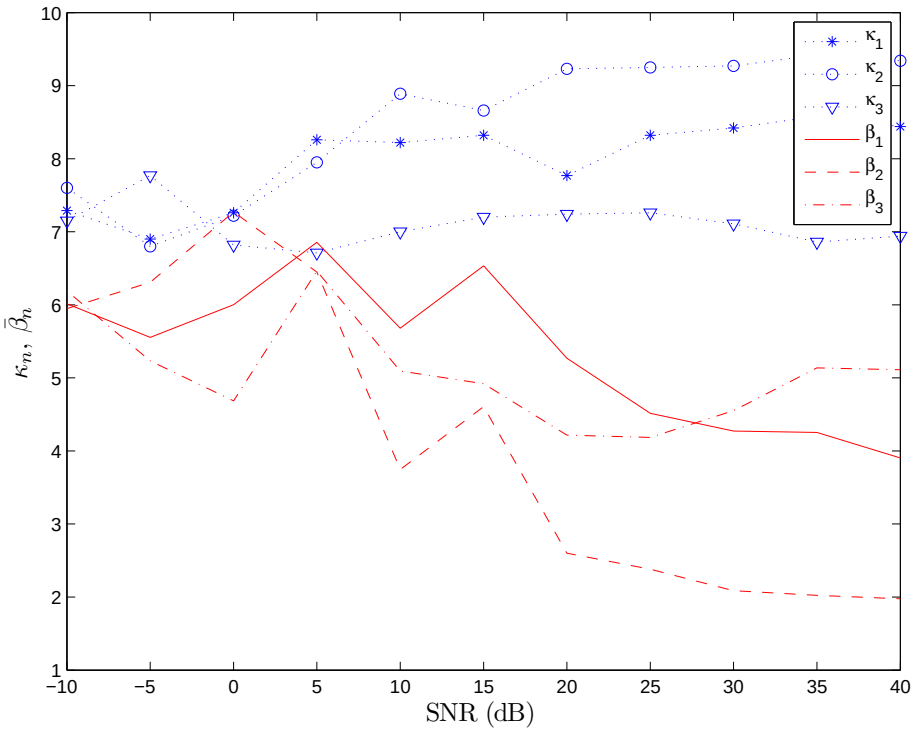




CPZLP: Simulation results (3)

- Monte Carlo simulation results (100 trials per simulation): average number of iterations per subproblem and average number of backtracking steps per iteration vs. L 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
  - power and bandwidth estimation