Adaptive Sinusoidal Modeling

Yannis Stylianou

University of Crete, Computer Science Dept., Greece,

yannis@csd.uoc.gr

SPCC 2015

Sinusoidal Models

- Sinusoidal Model:

$$x(t) = \sum_{k=-K}^{K} \gamma_k e^{j2\pi f_k t}$$

- Special case, Harmonic Model:

$$x(t) = \sum_{k=-K}^{K} \gamma_k e^{j2\pi k f_0 t}$$

- Estimation of parameters (Linear approach):

$$\gamma = \mathcal{F}_{f_k} \mathbf{s}$$

- Model Evaluation, Mean-Squared Error (MSE):

$$\epsilon = \int_w |s(t) - x(t)|^2 \, dt$$

Sinusoidal Models

- Sinusoidal Model:

$$x(t) = \sum_{k=-K}^{K} \gamma_k e^{j2\pi f_k t}$$

- Special case, Harmonic Model:

$$x(t) = \sum_{k=-K}^{K} \gamma_k e^{j2\pi k f_0 t}$$

- Estimation of parameters (Linear approach):

$$\gamma = \mathcal{F}_{f_k} \mathbf{s}$$

- Model Evaluation, Mean-Squared Error (MSE):

$$\epsilon = \int_{w} |s(t) - x(t)|^2 \, dt$$

Sinusoidal Models

- Sinusoidal Model:

$$x(t) = \sum_{k=-K}^{K} \gamma_k e^{j2\pi f_k t}$$

- Special case, Harmonic Model:

$$x(t) = \sum_{k=-K}^{K} \gamma_k e^{j2\pi k f_0 t}$$

- Estimation of parameters (Linear approach):

$$\gamma = \mathcal{F}_{f_k} \mathbf{s}$$

- Model Evaluation, Mean-Squared Error (MSE):

$$\epsilon = \int_w |s(t) - x(t)|^2 \, dt$$

Sinusoidal Models

- Sinusoidal Model:

$$x(t) = \sum_{k=-K}^{K} \gamma_k e^{j2\pi f_k t}$$

- Special case, Harmonic Model:

$$x(t) = \sum_{k=-K}^{K} \gamma_k e^{j2\pi k f_0 t}$$

- Estimation of parameters (Linear approach):

$$\gamma = \mathcal{F}_{f_k} \mathbf{s}$$

- Model Evaluation, Mean-Squared Error (MSE):

$$\epsilon = \int_{w} |s(t) - x(t)|^2 \, dt$$

iQHM, aQHM, eaQHM, aHM

Towards Adaptive Sinusoidal Models

Estimating Sinusoidal parameters

- Sinusoidal representation for a speech/signal frame:

$$x(t) = \left(\sum_{k=-K}^{K} a_k e^{j2\pi f_k t} \right) w(t)$$

- Methods:
  - FFT-based methods (i.e., QIFFT [Abe et al., 2004–05, [2] [3]])
  - Subspace methods
  - Least Squares (LS) method
- Frequency mismatch:

$$\hat{f}_k = f_k + \eta_k$$

Estimating Sinusoidal parameters

- Sinusoidal representation for a speech/signal frame:

$$x(t) = \left(\sum_{k=-K}^{K} a_k e^{j2\pi f_k t} \right) w(t)$$

- Methods:
  - FFT-based methods (i.e., QIFFT [Abe et al., 2004-05, [2] [3]])
  - Subspace methods
  - Least Squares (LS) method
- Frequency mismatch:

$$\hat{f}_k = f_k + \eta_k$$

Estimating Sinusoidal parameters

- Sinusoidal representation for a speech/signal frame:

$$x(t) = \left(\sum_{k=-K}^{K} a_k e^{j2\pi f_k t} \right) w(t)$$

- Methods:
  - FFT-based methods (i.e., QIFFT [Abe et al., 2004-05, [2] [3]])
  - Subspace methods
  - Least Squares (LS) method
- Frequency mismatch:

$$\hat{f}_k = f_k + \eta_k$$

Estimating Sinusoidal parameters

- Sinusoidal representation for a speech/signal frame:

$$x(t) = \left(\sum_{k=-K}^{K} a_k e^{j2\pi f_k t} \right) w(t)$$

- Methods:
  - FFT-based methods (i.e., QIFFT [Abe et al., 2004-05, [2] [3]])
  - Subspace methods
  - Least Squares (LS) method
- Frequency mismatch:

$$\hat{f}_k = f_k + \eta_k$$

Estimating Sinusoidal parameters

- Sinusoidal representation for a speech/signal frame:

$$x(t) = \left(\sum_{k=-K}^{K} a_k e^{j2\pi f_k t} \right) w(t)$$

- Methods:
  - FFT-based methods (i.e., QIFFT [Abe et al., 2004-05, [2] [3]])
  - Subspace methods
  - Least Squares (LS) method
- Frequency mismatch:

$$\hat{f}_k = f_k + \eta_k$$

Estimating Sinusoidal parameters

- Sinusoidal representation for a speech/signal frame:

$$x(t) = \left(\sum_{k=-K}^{K} a_k e^{j2\pi f_k t} \right) w(t)$$

- Methods:
  - FFT-based methods (i.e., QIFFT [Abe et al., 2004-05, [2] [3]])
  - Subspace methods
  - Least Squares (LS) method
- Frequency mismatch:

$$\hat{f}_k = f_k + \eta_k$$

Quasi-Harmonic Model, QHM [6]

- Frequency mismatch:

$$x(t) = \left(\sum_{k=-K}^{K} a_k e^{j2\pi \hat{f}_k t} \right) w(t)$$

- QHM (de Prony 1795, Laroche [4] (1989), Stylianou 1993, Pantazis [5] (2008, 2011)):

$$x(t) = \left(\sum_{k=-K}^{K} (a_k + t b_k) e^{j2\pi \hat{f}_k t} \right) w(t)$$

QUASI-HARMONIC MODEL, QHM [6]

- Frequency mismatch:

$$x(t) = \left(\sum_{k=-K}^{K} a_k e^{j2\pi \hat{f}_k t} \right) w(t)$$

- QHM (de Prony 1795, Laroche [4] (1989), Stylianou 1993, Pantazis [5] (2008, 2011)):

$$x(t) = \left(\sum_{k=-K}^{K} (a_k + t b_k) e^{j2\pi \hat{f}_k t} \right) w(t)$$

A few details of QHM

- QHM in the frequency domain:

$$X_k(f) = a_k W(f - \hat{f}_k) + j\frac{b_k}{2\pi} W'(f - \hat{f}_k)$$

- **Decomposition of** b_k: $b_k = \rho_{1,k} a_k + \rho_{2,k} j a_k$
- Then

$$X_k(f) = a_k \left[W(f - \hat{f}_k) - \frac{\rho_{2,k}}{2\pi} W'(f - \hat{f}_k) + j\frac{\rho_{1,k}}{2\pi} W'(f - \hat{f}_k) \right]$$

- and taking into account:

$$W(f - \hat{f}_k - \frac{\rho_{2,k}}{2\pi}) = W(f - \hat{f}_k) - \frac{\rho_{2,k}}{2\pi} W'(f - \hat{f}_k) + O(\rho_{2,k}^2 W''(f - \hat{f}_k))$$

- **Approximation of** the *kth* component of QHM

$$X_k(f) \approx a_k \left[W(f - \hat{f}_k - \frac{\rho_{2,k}}{2\pi}) + j\frac{\rho_{1,k}}{2\pi} W'(f - \hat{f}_k) \right]$$

A few details of QHM

- QHM in the frequency domain:

$$X_k(f) = a_k W(f - \hat{f}_k) + j\frac{b_k}{2\pi} W'(f - \hat{f}_k)$$

- **Decomposition of** b_k: $b_k = \rho_{1,k} a_k + \rho_{2,k} j a_k$
- Then

$$X_k(f) = a_k \left[W(f - \hat{f}_k) - \frac{\rho_{2,k}}{2\pi} W'(f - \hat{f}_k) + j\frac{\rho_{1,k}}{2\pi} W'(f - \hat{f}_k) \right]$$

- and taking into account:

$$W(f - \hat{f}_k - \frac{\rho_{2,k}}{2\pi}) = W(f - \hat{f}_k) - \frac{\rho_{2,k}}{2\pi} W'(f - \hat{f}_k) + O(\rho_{2,k}^2 W''(f - \hat{f}_k))$$

- **Approximation of** the *kth* component of QHM

$$X_k(f) \approx a_k \left[W(f - \hat{f}_k - \frac{\rho_{2,k}}{2\pi}) + j\frac{\rho_{1,k}}{2\pi} W'(f - \hat{f}_k) \right]$$

A few details of QHM

- QHM in the frequency domain:

$$X_k(f) = a_k W(f - \hat{f}_k) + j\frac{b_k}{2\pi} W'(f - \hat{f}_k)$$

- **Decomposition of** b_k: $b_k = \rho_{1,k} a_k + \rho_{2,k} j a_k$
- Then

$$X_k(f) = a_k \left[W(f - \hat{f}_k) - \frac{\rho_{2,k}}{2\pi} W'(f - \hat{f}_k) + j\frac{\rho_{1,k}}{2\pi} W'(f - \hat{f}_k) \right]$$

- and taking into account:

$$W(f - \hat{f}_k - \frac{\rho_{2,k}}{2\pi}) = W(f - \hat{f}_k) - \frac{\rho_{2,k}}{2\pi} W'(f - \hat{f}_k) + O(\rho_{2,k}^2 W''(f - \hat{f}_k))$$

- **Approximation of** the *kth* component of QHM

$$X_k(f) \approx a_k \left[W(f - \hat{f}_k - \frac{\rho_{2,k}}{2\pi}) + j\frac{\rho_{1,k}}{2\pi} W'(f - \hat{f}_k) \right]$$

A few details of QHM

- QHM in the frequency domain:

$$X_k(f) = a_k W(f - \hat{f}_k) + j\frac{b_k}{2\pi} W'(f - \hat{f}_k)$$

- **Decomposition of** b_k: $b_k = \rho_{1,k} a_k + \rho_{2,k} j a_k$
- Then

$$X_k(f) = a_k \left[W(f - \hat{f}_k) - \frac{\rho_{2,k}}{2\pi} W'(f - \hat{f}_k) + j\frac{\rho_{1,k}}{2\pi} W'(f - \hat{f}_k) \right]$$

- and taking into account:

$$W(f - \hat{f}_k - \frac{\rho_{2,k}}{2\pi}) = W(f - \hat{f}_k) - \frac{\rho_{2,k}}{2\pi} W'(f - \hat{f}_k) + O(\rho_{2,k}^2 W''(f - \hat{f}_k))$$

- **Approximation of** the *kth* component of QHM

$$X_k(f) \approx a_k \left[W(f - \hat{f}_k - \frac{\rho_{2,k}}{2\pi}) + j\frac{\rho_{1,k}}{2\pi} W'(f - \hat{f}_k) \right]$$

A few details of QHM

- QHM in the frequency domain:

$$X_k(f) = a_k W(f - \hat{f}_k) + j\frac{b_k}{2\pi} W'(f - \hat{f}_k)$$

- **Decomposition of** b_k: $b_k = \rho_{1,k} a_k + \rho_{2,k} j a_k$
- Then

$$X_k(f) = a_k \left[W(f - \hat{f}_k) - \frac{\rho_{2,k}}{2\pi} W'(f - \hat{f}_k) + j\frac{\rho_{1,k}}{2\pi} W'(f - \hat{f}_k) \right]$$

- and taking into account:

$$W(f - \hat{f}_k - \frac{\rho_{2,k}}{2\pi}) = W(f - \hat{f}_k) - \frac{\rho_{2,k}}{2\pi} W'(f - \hat{f}_k) + O(\rho_{2,k}^2 W''(f - \hat{f}_k))$$

- **Approximation of** the *kth* component of QHM

$$X_k(f) \approx a_k \left[W(f - \hat{f}_k - \frac{\rho_{2,k}}{2\pi}) + j\frac{\rho_{1,k}}{2\pi} W'(f - \hat{f}_k) \right]$$

Approximations in QHM

- we found previously:

$$X_k(f) \approx a_k \left[W(f - \hat{f}_k - \frac{\rho_{2,k}}{2\pi}) + j\frac{\rho_{1,k}}{2\pi} W'(f - \hat{f}_k) \right]$$

- Back to the time-domain:

$$x_k(t) \approx a_k \left[e^{j(2\pi\hat{f}_k + \rho_{2,k})t} + \rho_{1,k} t e^{j2\pi\hat{f}_k t} \right] w(t)$$

- Initially, we assumed:

$$x_k(t) = a_k \left[e^{j(2\pi(\hat{f}_k + \eta_k))t} \right] w(t)$$

- in other words, **it is suggested**:

$$\hat{\eta}_k = \rho_{2,k}/2\pi$$

Approximations in QHM

- we found previously:

$$X_k(f) \approx a_k \left[W(f - \hat{f}_k - \frac{\rho_{2,k}}{2\pi}) + j\frac{\rho_{1,k}}{2\pi} W'(f - \hat{f}_k) \right]$$

- Back to the time-domain:

$$x_k(t) \approx a_k \left[e^{j(2\pi\hat{f}_k + \rho_{2,k})t} + \rho_{1,k} t e^{j2\pi\hat{f}_k t} \right] w(t)$$

- Initially, we assumed:

$$x_k(t) = a_k \left[e^{j(2\pi(\hat{f}_k + \eta_k))t} \right] w(t)$$

- in other words, **it is suggested**:

$$\hat{\eta}_k = \rho_{2,k}/2\pi$$

Approximations in QHM

- we found previously:

$$X_k(f) \approx a_k \left[W(f - \hat{f}_k - \frac{\rho_{2,k}}{2\pi}) + j\frac{\rho_{1,k}}{2\pi} W'(f - \hat{f}_k) \right]$$

- Back to the time-domain:

$$x_k(t) \approx a_k \left[e^{j(2\pi\hat{f}_k + \rho_{2,k})t} + \rho_{1,k} t e^{j2\pi\hat{f}_k t} \right] w(t)$$

- Initially, we assumed:

$$x_k(t) = a_k \left[e^{j(2\pi(\hat{f}_k + \eta_k))t} \right] w(t)$$

- in other words, **it is suggested**:

$$\hat{\eta}_k = \rho_{2,k}/2\pi$$

Approximations in QHM

- we found previously:

$$X_k(f) \approx a_k \left[W(f - \hat{f}_k - \frac{\rho_{2,k}}{2\pi}) + j\frac{\rho_{1,k}}{2\pi} W'(f - \hat{f}_k) \right]$$

- Back to the time-domain:

$$x_k(t) \approx a_k \left[e^{j(2\pi \hat{f}_k + \rho_{2,k})t} + \rho_{1,k} t e^{j2\pi \hat{f}_k t} \right] w(t)$$

- Initially, we assumed:

$$x_k(t) = a_k \left[e^{j(2\pi(\hat{f}_k + \eta_k))t} \right] w(t)$$

- in other words, **it is suggested**:

$$\hat{\eta}_k = \rho_{2,k}/2\pi$$

CONSTRAINTS

$$W(f - \hat{f}_k - \frac{\rho_{2,k}}{2\pi}) = W(f - \hat{f}_k) - \frac{\rho_{2,k}}{2\pi} W'(f - \hat{f}_k) + O(\rho_{2,k}^2 W''(f - \hat{f}_k))$$

Influence of the Window

- We would like small value for $W''(f)$ at f_k
- $W''(f)$ is influenced by the length, T, of the analysis window, i.e., for rectangular window: $W''(f) \propto T^3$
- So ... we would like a **short analysis window**
- But ... **long analysis window** provides robustness
- Length of the window $\rightleftarrows$ bandwidth

Influence of the Window

- We would like small value for $W''(f)$ at f_k
- $W''(f)$ is influenced by the length, T, of the analysis window, i.e., for rectangular window: $W''(f) \propto T^3$
- So ... we would like a **short analysis window**
- But ... **long analysis window** provides robustness
- Length of the window $\rightleftarrows$ bandwidth

Influence of the Window

- We would like small value for $W''(f)$ at f_k
- $W''(f)$ is influenced by the length, T, of the analysis window, i.e., for rectangular window: $W''(f) \propto T^3$
- So ... we would like a **short analysis window**
- But ... **long analysis window** provides robustness
- Length of the window $\rightleftarrows$ bandwidth

Influence of the Window

- We would like small value for $W''(f)$ at f_k
- $W''(f)$ is influenced by the length, T, of the analysis window, i.e., for rectangular window: $W''(f) \propto T^3$
- So ... we would like a **short analysis window**
- But ... **long analysis window** provides robustness
- Length of the window $\rightleftarrows$ bandwidth

Influence of the Window

- We would like small value for $W''(f)$ at f_k
- $W''(f)$ is influenced by the length, T, of the analysis window, i.e., for rectangular window: $W''(f) \propto T^3$
- So ... we would like a **short analysis window**
- But ... **long analysis window** provides robustness
- Length of the window $\rightleftarrows$ bandwidth

Influence of frequency mismatch

- Assume

$$y(t) = \alpha e^{j(2\pi \hat{f} t + \eta t)}$$

- modeled by QHM as

$$x(t) = (a + tb)e^{j(2\pi \hat{f} t)} \quad -T \leq t \leq T$$

- LS solution provides:

$$a = \alpha \frac{sin(\eta T)}{\eta T}$$

$$b = \alpha \ 3j(\frac{sin(\eta T)}{\eta^2 T^3} - \frac{cos(\eta T)}{\eta T^2})$$

- Frequency mismatch estimate:

$$\hat{\eta} = 3 \left(\frac{1}{\eta T^2} - \frac{cot(\eta T)}{T} \right)$$

Influence of frequency mismatch

- Assume

$$y(t) = \alpha e^{j(2\pi \hat{f} t + \eta t)}$$

- modeled by QHM as

$$x(t) = (a + tb)e^{j(2\pi \hat{f} t)} \quad -T \leq t \leq T$$

- LS solution provides:

$$a = \alpha \frac{sin(\eta T)}{\eta T}$$

$$b = \alpha \ 3j(\frac{sin(\eta T)}{\eta^2 T^3} - \frac{cos(\eta T)}{\eta T^2})$$

- Frequency mismatch estimate:

$$\hat{\eta} = 3\left(\frac{1}{\eta T^2} - \frac{cot(\eta T)}{T}\right)$$

INFLUENCE OF FREQUENCY MISMATCH

- Assume

$$y(t) = \alpha e^{j(2\pi \hat{f} t + \eta t)}$$

- modeled by QHM as

$$x(t) = (a + tb)e^{j(2\pi \hat{f} t)} \quad -T \leq t \leq T$$

- LS solution provides:

$$a = \alpha \frac{sin(\eta T)}{\eta T}$$

$$b = \alpha \; 3j(\frac{sin(\eta T)}{\eta^2 T^3} - \frac{cos(\eta T)}{\eta T^2})$$

- Frequency mismatch estimate:

$$\hat{\eta} = 3\left(\frac{1}{\eta T^2} - \frac{cot(\eta T)}{T}\right)$$

Influence of frequency mismatch

- Assume

$$y(t) = \alpha e^{j(2\pi \hat{f} t + \eta t)}$$

- modeled by QHM as

$$x(t) = (a + tb)e^{j(2\pi \hat{f} t)} \quad -T \leq t \leq T$$

- LS solution provides:

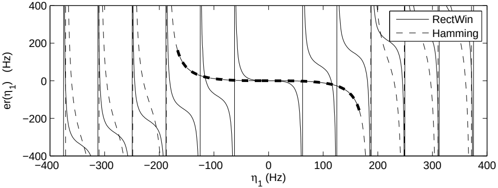
$$a = \alpha \frac{sin(\eta T)}{\eta T}$$

$$b = \alpha \; 3j\left(\frac{sin(\eta T)}{\eta^2 T^3} - \frac{cos(\eta T)}{\eta T^2}\right)$$

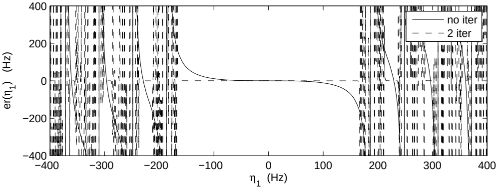
- Frequency mismatch estimate:

$$\hat{\eta} = 3\left(\frac{1}{\eta T^2} - \frac{cot(\eta T)}{T}\right)$$

Combining Influences



(a)



(b)

- Iteratively, the bias can be removed when $|\eta| < B/3$, where B is the bandwidth of the squared analysis window.

Robustness against additive noise

- Signal contaminated by noise:

$$y(t) = \sum_{k=1}^{4} a_k e^{j2\pi f_k} + v(t)$$

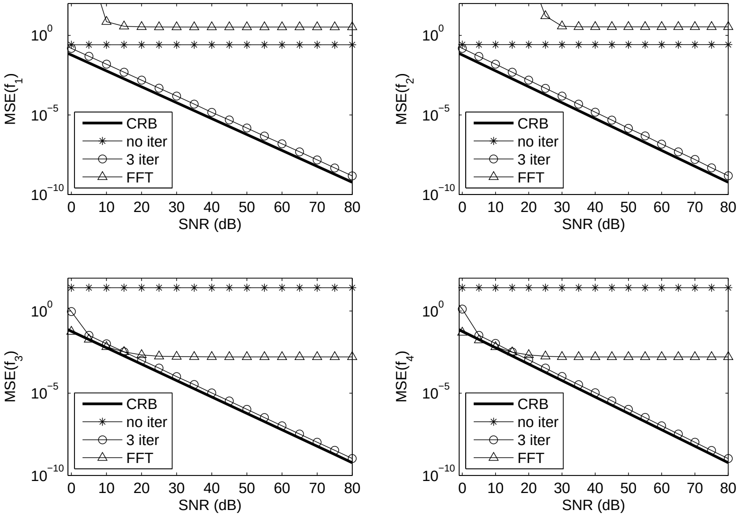
- Mean Squared Error (MSE):

$$MSE\{\hat{f}_k\} = \frac{1}{M}\sum_{i=1}^{M} |\hat{f}_k(i) - f_k|^2$$

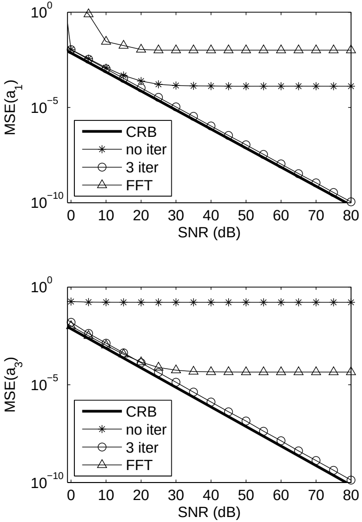
$$MSE\{\hat{a}_k\} = \frac{1}{M}\sum_{i=1}^{M} |\hat{a}_k(i) - a_k|^2$$

- Comparison with Cramer-Rao Bounds (CRB) and QIFFT (Abe et al. 2004)
- 10000 Monte Carlo simulations

MSE of frequencies as a function of SNR.



MSE of amplitudes as a function of SNR.



Notes on QHM

- The approximation made in QHM is valid provided that the frequency mismatch lies in a specific interval.
- This interval is a function of the bandwidth of the analysis window.
- Robustness of QHM against noise was tested and verified.
- It can be shown that iterative QHM is equivalent to (an approximate) Gauss-Newton method.
- There are also relations between QHM and Reassigned Spectrum

Notes on QHM

- The approximation made in QHM is valid provided that the frequency mismatch lies in a specific interval.
- This interval is a function of the bandwidth of the analysis window.
- Robustness of QHM against noise was tested and verified.
- It can be shown that iterative QHM is equivalent to (an approximate) Gauss-Newton method.
- There are also relations between QHM and Reassigned Spectrum

Notes on QHM

- The approximation made in QHM is valid provided that the frequency mismatch lies in a specific interval.
- This interval is a function of the bandwidth of the analysis window.
- Robustness of QHM against noise was tested and verified.
- It can be shown that iterative QHM is equivalent to (an approximate) Gauss-Newton method.
- There are also relations between QHM and Reassigned Spectrum

Notes on QHM

- The approximation made in QHM is valid provided that the frequency mismatch lies in a specific interval.
- This interval is a function of the bandwidth of the analysis window.
- Robustness of QHM against noise was tested and verified.
- It can be shown that iterative QHM is equivalent to (an approximate) Gauss-Newton method.
- There are also relations between QHM and Reassigned Spectrum

Notes on QHM

- The approximation made in QHM is valid provided that the frequency mismatch lies in a specific interval.
- This interval is a function of the bandwidth of the analysis window.
- Robustness of QHM against noise was tested and verified.
- It can be shown that iterative QHM is equivalent to (an approximate) Gauss-Newton method.
- There are also relations between QHM and Reassigned Spectrum

QHM and Gauss-Newton method

- Assuming we have N samples of $x(t) = ce^{j\omega t}w(t)$
- Approximate Gauss-Newton solution for frequency

$$\omega^{(1)} = \omega^{(0)} - \mathcal{R}\left\{ \frac{\sum_{t=-N}^{N} tw^2(t)x(t)e^{-j\omega^{(0)}t}}{c^{(0)}\sum_{t=-N}^{N} t^2 w^2(t)} \right\}$$

- QHM

$$x(t) = (c + tb)e^{j\omega t}w(t)$$

with

$$\begin{aligned} \omega^{(1)} &= \omega^{(0)} + \rho_2 \\ &= \omega^{(0)} - \mathcal{R}\left\{ \frac{jb^{(0)}}{c^{(0)}} \right\} \end{aligned}$$

where

$$b^{(0)} = \frac{\sum_{t=-N}^{N} tw^2(t)x(t)e^{-j\omega^{(0)}t}}{\sum_{t=-N}^{N} t^2 w^2(t)}$$

QHM and Gauss-Newton method

- Assuming we have N samples of $x(t) = ce^{j\omega t}w(t)$
- Approximate Gauss-Newton solution for frequency

$$\omega^{(1)} = \omega^{(0)} - \mathcal{R}\left\{ \frac{\sum_{t=-N}^{N} tw^2(t)x(t)e^{-j\omega^{(0)}t}}{c^{(0)}\sum_{t=-N}^{N} t^2 w^2(t)} \right\}$$

- QHM

$$x(t) = (c + tb)e^{j\omega t}w(t)$$

with

$$\begin{aligned} \omega^{(1)} &= \omega^{(0)} + \rho_2 \\ &= \omega^{(0)} - \mathcal{R}\left\{ \frac{jb^{(0)}}{c^{(0)}} \right\} \end{aligned}$$

where

$$b^{(0)} = \frac{\sum_{t=-N}^{N} tw^2(t)x(t)e^{-j\omega^{(0)}t}}{\sum_{t=-N}^{N} t^2 w^2(t)}$$

QHM and Gauss-Newton method

- Assuming we have N samples of $x(t) = ce^{j\omega t}w(t)$
- Approximate Gauss-Newton solution for frequency

$$\omega^{(1)} = \omega^{(0)} - \mathcal{R}\left\{ \frac{\sum_{t=-N}^{N} tw^2(t)x(t)e^{-j\omega^{(0)}t}}{c^{(0)}\sum_{t=-N}^{N} t^2 w^2(t)} \right\}$$

- QHM

$$x(t) = (c + tb)e^{j\omega t}w(t)$$

with

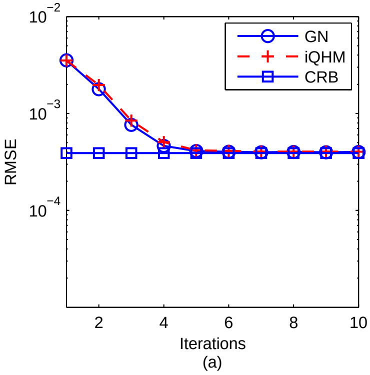
$$\begin{aligned} \omega^{(1)} &= \omega^{(0)} + \rho_2 \\ &= \omega^{(0)} - \mathcal{R}\left\{ \frac{jb^{(0)}}{c^{(0)}} \right\} \end{aligned}$$

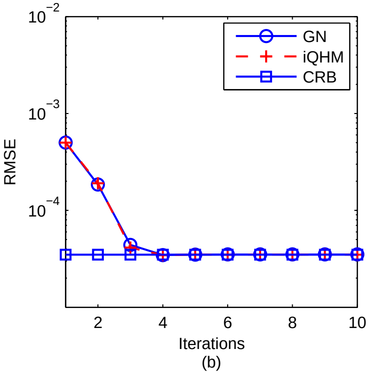
where

$$b^{(0)} = \frac{\sum_{t=-N}^{N} tw^2(t)x(t)e^{-j\omega^{(0)}t}}{\sum_{t=-N}^{N} t^2 w^2(t)}$$

Evaluation: RMSE for frequency

▷ $SNR = 0$dB, $N = 100$ (left), $N = 500$ (right).
CRB: Cramer-Rao lower bound for frequency estimation

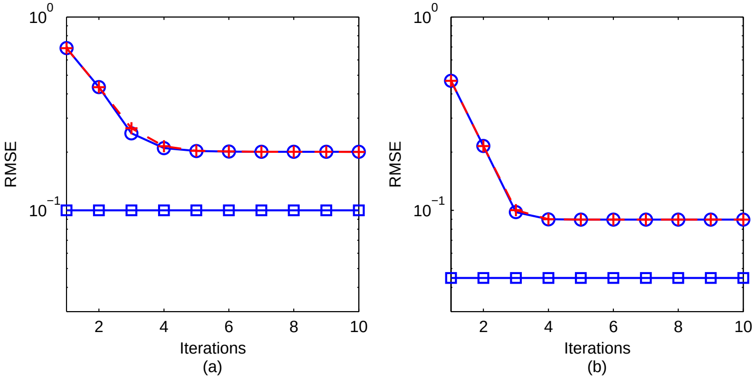




Evaluation: RMSE for amplitude

- $SNR = 0$dB, $N = 100$ (left), $N = 500$ (right).

CRB: Cramer-Rao lower bound for amplitude estimation



Reassigned Spectrum

- Time relocation:

$$\hat{\tau} = -\frac{\partial \phi(\tau, \omega)}{\partial \omega} = \tau - \Re\left(\frac{X_{Tw}(\tau, \omega)}{X(\tau, \omega)}\right)$$

where

$$X_{Tw}(\tau, \omega) = -\sum_{t=-N}^{N} t w(t) x(t+\tau) e^{-j\omega(t+\tau)}$$

- Frequency relocation:

$$\hat{\omega} = \omega + \frac{\partial \phi(\tau, \omega)}{\partial \tau} = \omega + \Im\left(\frac{X_{\mathcal{D}w}(\tau, \omega)}{X(\tau, \omega)}\right)$$

where

$$X_{\mathcal{D}w}(\tau, \omega) = -\sum_{t=-N}^{N} \frac{dw(t)}{dt} x(t+\tau) e^{-j\omega(t+\tau)}$$

Reassigned Spectrum and QHM

- QHM: $x(t) = (a + tb)e^{j\omega t}w(t)$ with $b = \rho_1 a + \rho_2 j a_1$
- Then, it can be shown:

$$\hat{\tau} = \tau + \frac{W_2}{W_0}\rho_1$$

and (in case of using Gaussian windows):

$$\hat{\omega} = \omega + \frac{W_2}{W_0}\rho_2$$

where

$$W_k = \sum_{t=-N}^{N} t^k w(t)$$

Reassigned Spectrum and QHM

- QHM: $x(t) = (a + tb)e^{j\omega t}w(t)$ with $b = \rho_1 a + \rho_2 j a_1$
- Then, it can be shown:

$$\hat{\tau} = \tau + \frac{W_2}{W_0}\rho_1$$

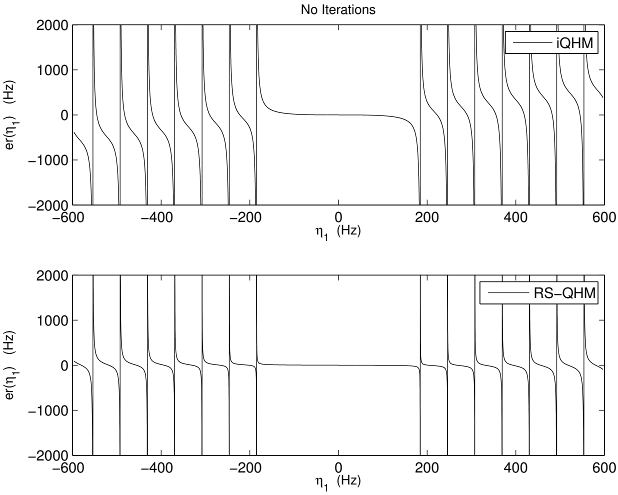
and (in case of using Gaussian windows):

$$\hat{\omega} = \omega + \frac{W_2}{W_0}\rho_2$$

where

$$W_k = \sum_{t=-N}^{N} t^k w(t)$$

Frequency mismatch estimation error



From QHM to Adaptive QHM, aQHM [7]

- QHM (**stationarity assumption**):

$$x(t) = \left(\sum_{k=-K}^{K} (a_k + tb_k) e^{2\pi j f_k t} \right) w(t)$$

- Adaptive QHM (aQHM):

$$x(t) = \left(\sum_{k=-K}^{K} (a_k + tb_k) e^{j\tilde{\phi}_k(t)} \right) w(t)$$

where

$$\tilde{\phi}_k(t) = 2\pi \int_0^t f_k(s) ds + \varphi, \;\; t \in [0, T]$$

is the estimated instantaneous phase.

From QHM to Adaptive QHM, aQHM [7]

- QHM (**stationarity assumption**):

$$x(t) = \left(\sum_{k=-K}^{K} (a_k + t b_k) e^{2\pi j f_k t} \right) w(t)$$

- Adaptive QHM (aQHM):

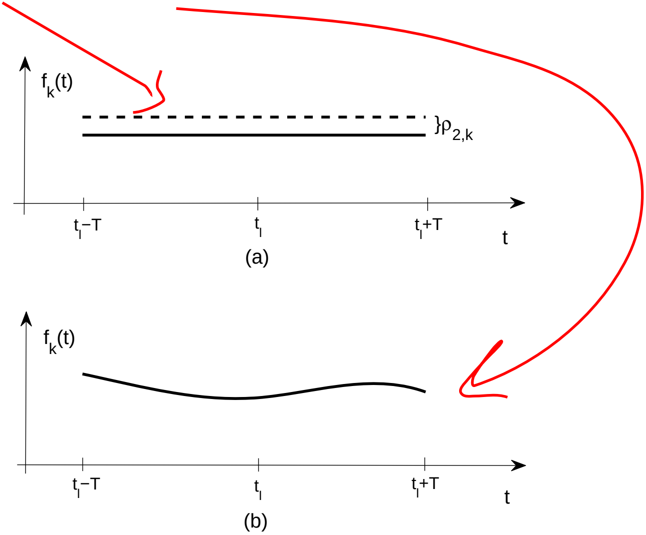
$$x(t) = \left(\sum_{k=-K}^{K} (a_k + t b_k) e^{j \tilde{\phi}_k(t)} \right) w(t)$$

where

$$\tilde{\phi}_k(t) = 2\pi \int_0^t f_k(s) ds + \varphi, \;\; t \in [0, T]$$

is the estimated instantaneous phase.

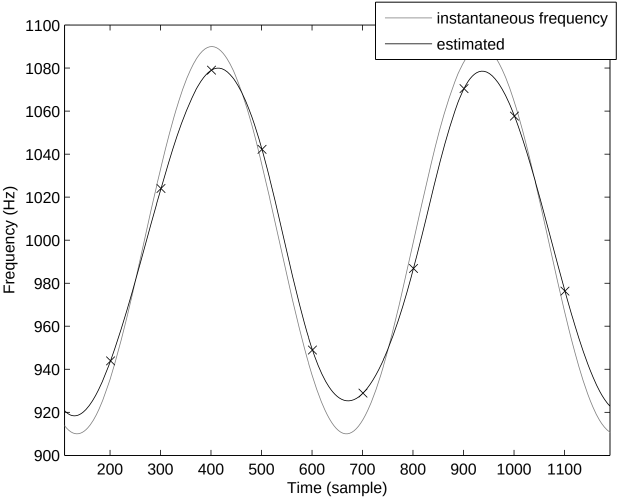
From QHM to aQHM; Graphically



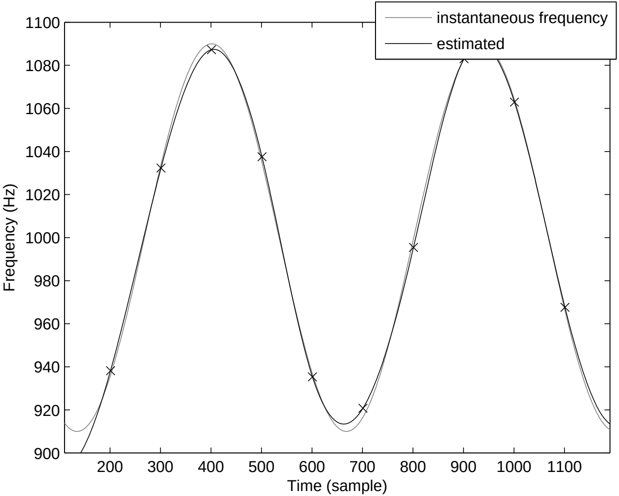
Frame Rate in aQHM

- One sample: no interpolation between estimations
- Higher rates (i.e., 5ms, 10ms): Interpolation between estimates is required:
  - Amplitudes are linearly interpolated
  - Frequencies are interpolated with splines
  - Phases are interpolated by integration of instantaneous frequency

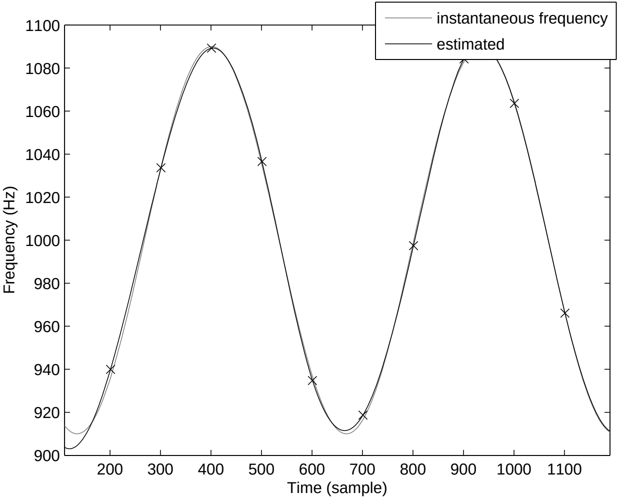
Example of estimation in aQHM: no iteration (QHM)



Example of estimation in aQHM: one iteration



Example of estimation in aQHM: two iterations



Validation

- Let us consider as example:

$$x(t) = (11 - 340t + 4000t^2)e^{j2\pi(280t+19500t^2)}$$

- and frequency mismatch: $|f_k - \zeta_k| = 35Hz$, using Hamming window
- Consider simplified approaches: Sinusoidal-based analysis (SM)
- Mean Absolute Error (MAE) [frame rate: one sample]

	AM	FM (Hz)
QHM, 10ms	0.46	2.15
aQHM (1)	0.008	0.006
SM, 10ms	0.44	3.08

VALIDATION

- Let us consider as example:

$$x(t) = (11 - 340t + 4000t^2)e^{j2\pi(280t+19500t^2)}$$

- and frequency mismatch: $|f_k - \zeta_k| = 35Hz$, using Hamming window
- Consider simplified approaches: Sinusoidal-based analysis (SM)
- Mean Absolute Error (MAE) [frame rate: one sample]

	AM	FM (Hz)
QHM, 10ms	0.46	2.15
aQHM (1)	0.008	0.006
SM, 10ms	0.44	3.08

VALIDATION

- Let us consider as example:

$$x(t) = (11 - 340t + 4000t^2)e^{j2\pi(280t+19500t^2)}$$

- and frequency mismatch: $|f_k - \zeta_k| = 35Hz$, using Hamming window
- Consider simplified approaches: Sinusoidal-based analysis (SM)
- Mean Absolute Error (MAE) [frame rate: one sample]

	AM	FM (Hz)
QHM, 10ms	0.46	2.15
aQHM (1)	0.008	0.006
SM, 10ms	0.44	3.08

VALIDATION

- Let us consider as example:

$$x(t) = (11 - 340t + 4000t^2)e^{j2\pi(280t+19500t^2)}$$

- and frequency mismatch: $|f_k - \zeta_k| = 35Hz$, using Hamming window
- Consider simplified approaches: Sinusoidal-based analysis (SM)
- Mean Absolute Error (MAE) [frame rate: one sample]

	AM	FM (Hz)
QHM, 10ms	0.46	2.15
aQHM (1)	0.008	0.006
SM, 10ms	0.44	3.08

QHM REMINDER

- QHM:

$$x(t) = \left(\sum_{k=-K}^{K} (a_k + tb_k) e^{2\pi j f_k t} \right) w(t)$$

- Instantaneous parameters:

$$\begin{aligned} A_k(t) &= \sqrt{(a_k^R + tb_k^R)^2 + (a_k^I + tb_k^I)^2} \\ \Phi_k(t) &= 2\pi f_k t + atan \frac{a_k^I + tb_k^I}{a_k^R + tb_k^R} \\ F_k(t) &= f_k + \frac{1}{2\pi} \frac{a_k^R b_k^I - a_k^I b_k^R}{M_k^2(t)} = f_k + \frac{1}{2\pi} \rho_{2,k} \end{aligned}$$

QHM reminder

- QHM:

$$x(t) = \left(\sum_{k=-K}^{K} (a_k + t b_k) e^{2\pi j f_k t} \right) w(t)$$

- Instantaneous parameters:

$$\begin{aligned}
A_k(t) &= \sqrt{(a_k^R + t b_k^R)^2 + (a_k^I + t b_k^I)^2} \\
\Phi_k(t) &= 2\pi f_k t + atan \frac{a_k^I + t b_k^I}{a_k^R + t b_k^R} \\
F_k(t) &= f_k + \frac{1}{2\pi} \frac{a_k^R b_k^I - a_k^I b_k^R}{M_k^2(t)} = f_k + \frac{1}{2\pi} \rho_{2,k}
\end{aligned}$$

QHM and AM-FM decomposition

- AM-FM signal

$$y(t) = \sum_{k=1}^{K(t)} a_k(t) cos(\phi_k(t)),$$

- Taylor series expansion of the instantaneous phase of kth component:

$$\phi_k(t) = 2\pi\zeta_k t + \sum_{i=0}^{\infty} \phi_{k,i} \frac{t^i}{i!}$$

- Instantaneous frequency of the kth component at $t = 0$:

$$\nu_k(0) = \zeta_k + \frac{\phi_{k,1}}{2\pi}$$

- ... and previously we had:

$$F_k(0) = f_k + \frac{\rho_{2,k}}{2\pi}$$

QHM and AM-FM decomposition

- AM-FM signal

$$y(t) = \sum_{k=1}^{K(t)} a_k(t) cos(\phi_k(t)),$$

- Taylor series expansion of the instantaneous phase of kth component:

$$\phi_k(t) = 2\pi\zeta_k t + \sum_{i=0}^{\infty} \phi_{k,i} \frac{t^i}{i!}$$

- Instantaneous frequency of the kth component at $t = 0$:

$$\nu_k(0) = \zeta_k + \frac{\phi_{k,1}}{2\pi}$$

- ... and previously we had:

$$F_k(0) = f_k + \frac{\rho_{2,k}}{2\pi}$$

QHM and AM-FM decomposition

- AM-FM signal

$$y(t) = \sum_{k=1}^{K(t)} a_k(t) cos(\phi_k(t)),$$

- Taylor series expansion of the instantaneous phase of kth component:

$$\phi_k(t) = 2\pi\zeta_k t + \sum_{i=0}^{\infty} \phi_{k,i} \frac{t^i}{i!}$$

- Instantaneous frequency of the kth component at $t = 0$:

$$\nu_k(0) = \zeta_k + \frac{\phi_{k,1}}{2\pi}$$

- ... and previously we had:

$$F_k(0) = f_k + \frac{\rho_{2,k}}{2\pi}$$

QHM AND AM-FM DECOMPOSITION

- AM-FM signal

$$y(t) = \sum_{k=1}^{K(t)} a_k(t) cos(\phi_k(t)),$$

- Taylor series expansion of the instantaneous phase of kth component:

$$\phi_k(t) = 2\pi\zeta_k t + \sum_{i=0}^{\infty} \phi_{k,i} \frac{t^i}{i!}$$

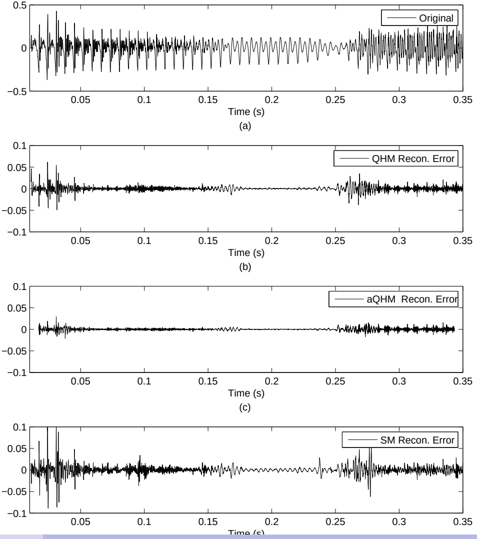
- Instantaneous frequency of the kth component at $t = 0$:

$$\nu_k(0) = \zeta_k + \frac{\phi_{k,1}}{2\pi}$$

- ... and previously we had:

$$F_k(0) = f_k + \frac{\rho_{2,k}}{2\pi}$$

Reconstruction errors with QHM, aQHM, SM



aQHM: AM-FM decomposition of Voiced Speech

- Signal Reconstruction

$$\hat{x}(t) = \sum_{k=1}^{K} \hat{A}_k(t) \cos\left(\hat{\Phi}_k(t)\right)$$

- Test on 5 minutes of 20 female and male voiced speech (TIMIT)
- Average Signal-to-Error Reconstruction Ratio (in dB)

	Male	Female
QHM	23.9	29.1
aQHM (3)	29.1	34.1
SM	17.2	21.1

aQHM: AM-FM decomposition of Voiced Speech

- Signal Reconstruction

$$\hat{x}(t) = \sum_{k=1}^{K} \hat{A}_k(t) \cos(\hat{\Phi}_k(t))$$

- Test on 5 minutes of 20 female and male voiced speech (TIMIT)
- Average Signal-to-Error Reconstruction Ratio (in dB)

	Male	Female
QHM	23.9	29.1
aQHM (3)	29.1	34.1
SM	17.2	21.1

aQHM: AM-FM decomposition of Voiced Speech

- Signal Reconstruction

$$\hat{x}(t) = \sum_{k=1}^{K} \hat{A}_k(t) \cos\left(\hat{\Phi}_k(t)\right)$$

- Test on 5 minutes of 20 female and male voiced speech (TIMIT)
- Average Signal-to-Error Reconstruction Ratio (in dB)

	Male	Female
QHM	23.9	29.1
aQHM (3)	29.1	34.1
SM	17.2	21.1

Extended aQHM

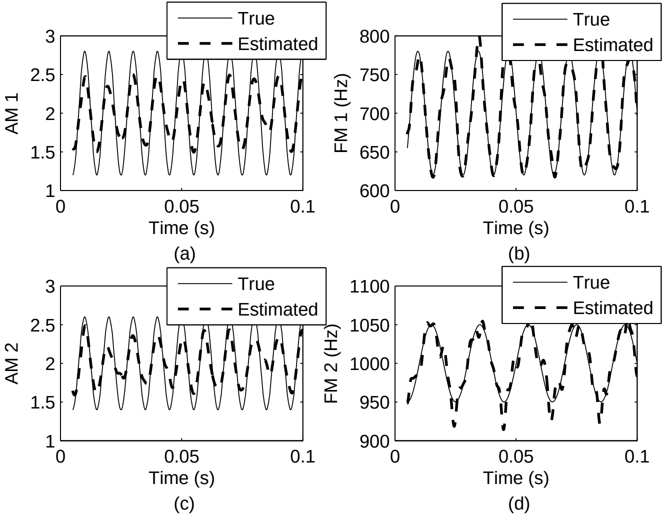
- Recall aQHM:

$$x(t) = \sum_{k=-K}^{K} (a_k + tb_k) e^{j\tilde{\phi}_k(t)}$$

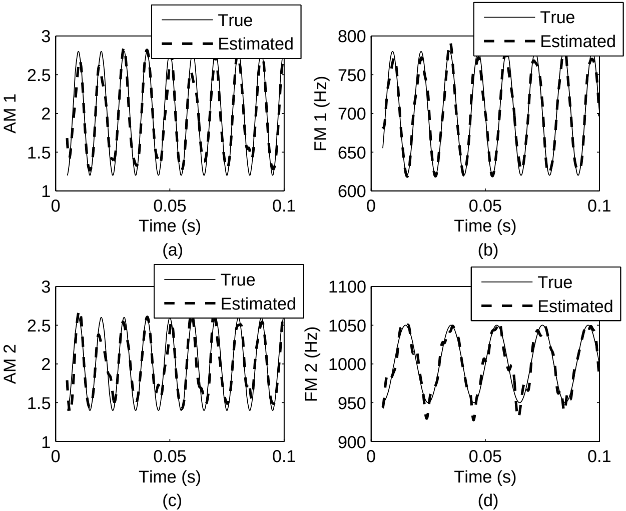
- Extended aQHM:

$$x(t) = \sum_{k=-K}^{K} (a_k + tb_k) \alpha(t) e^{j\tilde{\phi}_k(t)}$$

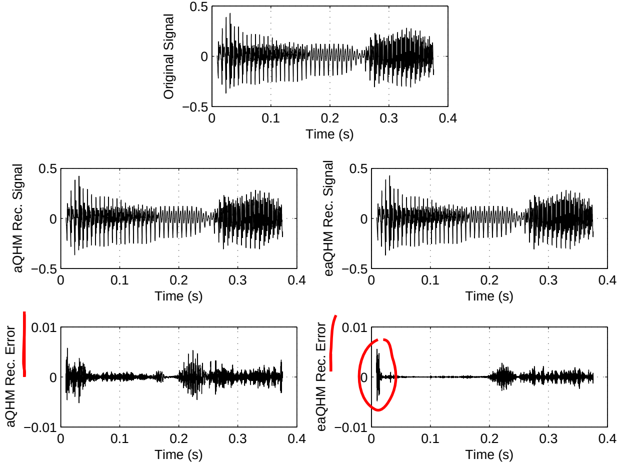
AM-FM modeling: aQHM



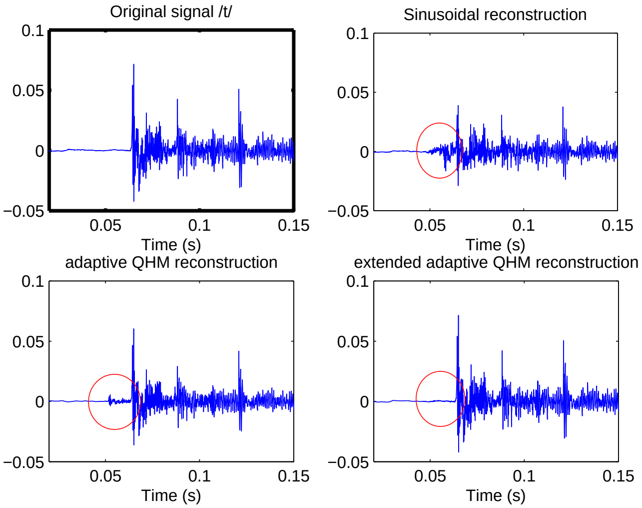
AM-FM modeling: extended aQHM



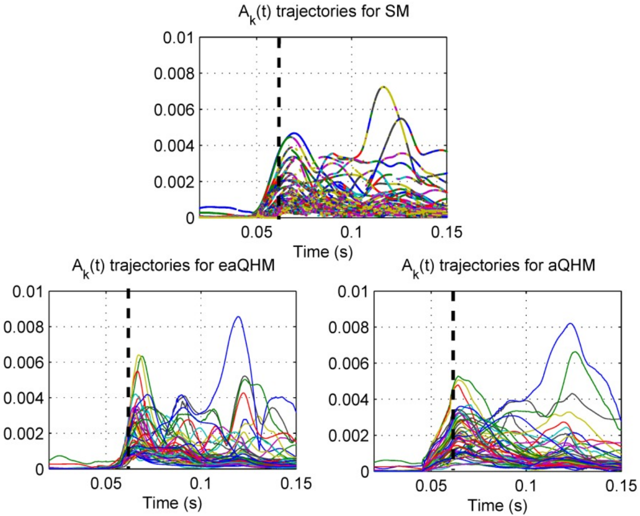
Comparing Adaptive Models



Stop modeling



Stop modeling



Large scale evaluation

Global Signal to Reconstruction Error Ratio (dB)						
Model	/*p*/	/*t*/	/*k*/	/*b*/	/*d*/	/*g*/
SM	12.7	12.8	12.4	16.6	14.9	15.3
aQHM	19.9	20.6	21.7	28.3	26.9	27.5
eaQHM	25.4	25.7	27.2	32.9	32.2	32.9

Speech Modifications

- High quality pitch scale modification
- Within glottal cycle formant tracking and modifications towards voice conversion
- Free pre-echo effect in time scaling
- Emotion detection and control

Speech Modifications

- High quality pitch scale modification
- Within glottal cycle formant tracking and modifications towards voice conversion
- Free pre-echo effect in time scaling
- Emotion detection and control

Speech Modifications

- High quality pitch scale modification
- Within glottal cycle formant tracking and modifications towards voice conversion
- Free pre-echo effect in time scaling
- Emotion detection and control

Speech Modifications

- High quality pitch scale modification
- Within glottal cycle formant tracking and modifications towards voice conversion
- Free pre-echo effect in time scaling
- Emotion detection and control

Speech Synthesis

- Unit selection based: fast synthesis, smooth trajectories, high quality prosodic modifications, voice conversion
- Statistical modeling: *naturally* smooth trajectories, articulators tracking, adaptation and controllability, modulations
- Sinusoidal models work nicely with DNNs

Speech Synthesis

- Unit selection based: fast synthesis, smooth trajectories, high quality prosodic modifications, voice conversion
- Statistical modeling: *naturally* smooth trajectories, articulators tracking, adaptation and controllability, modulations
- Sinusoidal models work nicely with DNNs

Speech Synthesis

- Unit selection based: fast synthesis, smooth trajectories, high quality prosodic modifications, voice conversion
- Statistical modeling: *naturally* smooth trajectories, articulators tracking, adaptation and controllability, modulations
- Sinusoidal models work nicely with DNNs

Speech Recognition

- Robust to noise filter bank estimation
- Data driven dynamic information
- High-resolution modulations
- Fits well to the DNN framework.

Speech Recognition

- Robust to noise filter bank estimation
- Data driven dynamic information
- High-resolution modulations
- Fits well to the DNN framework.

Speech Recognition

- Robust to noise filter bank estimation
- Data driven dynamic information
- High-resolution modulations
- Fits well to the DNN framework.

Speech Recognition

- Robust to noise filter bank estimation
- Data driven dynamic information
- High-resolution modulations
- Fits well to the DNN framework.

Speech Pathology

- Jitter and shimmer estimators
- Vocal fatigue
- Robust estimation in spasmodic dysphonia
- Vocal tremor.

Speech Pathology

- Jitter and shimmer estimators
- Vocal fatigue
- Robust estimation in spasmodic dysphonia
- Vocal tremor.

Speech Pathology

- Jitter and shimmer estimators
- Vocal fatigue
- Robust estimation in spasmodic dysphonia
- Vocal tremor.

Speech Pathology

- Jitter and shimmer estimators
- Vocal fatigue
- Robust estimation in spasmodic dysphonia
- Vocal tremor.

Key papers to read

- R. McAulay and T. Quatieri: *Speech analysis/Synthesis based on a sinusoidal representation* IEEE TASSP, Vol.34, No.4, Aug 1986, pp 744-754.
- Y. Pantazis, O. Rosec, and Y. Stylianou: *Adaptive AM-FM Signal Decomposition with Application to Speech Analysis* IEEE TASLP, Vol.19, No.2, Feb 2011, pp 290-300.
- Y. Pantazis, O. Rosec, and Y. Stylianou: *Iterative Estimation of Sinusoidal Signal Parameters*. IEEE SPL, Vol.17, No.5, May 2010, pp. 461-464
- P. Babu and P. Stoica: *Comments on Iterative Estimation of Sinusoidal Signal Parameters.* Comments on IEEE SPL, pp.1022-1023, Vol.17, No.12, Dec 2010
- Y. Pantazis, O. Rosec, and Y. Stylianou: *Reply to "Comments on 'Iterative Estimation of Sinusoidal Signal Parameters' "* IEEE SPL, pp. 1024-1026, Vol.17, No.12, Dec 2010
- G. Degottex and Y. Stylianou: *Analysis and Synthesis of Speech using an Adaptive Full-band Harmonic Model* IEEE TASLP, Vol.21, Issue 10, pp 2085-2095, 2013

T. F. Quatieri, *Discrete-Time Speech Signal Processing*.
Engewood Cliffs, NJ: Prentice Hall, 2002.

M. Abe and J. S. III, "CQIFFT: Correcting Bias in a Sinusoidal Parameter Estimator based on Quadratic Interpolation of FFT Magnitude Peaks," Tech. Rep. STAN-M-117, Stanford University, California, Oct 2004.

M. Abe and J. S. III, "AM/FM Estimation for Time-varying Sinusoidal Modeling," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, (Philadelphia), pp. III 201–204, 2005.

J. Laroche, "A new analysis/synthesis system of musical signals using Prony's method. Application to heavily damped percussive sounds.," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, (Glasgow, UK), pp. 2053–2056, May 1989.

Y. Pantazis, O. Rosec, and Y. Stylianou, "On the Properties of a Time-Varying Quasi-Harmonic Model of Speech," (Brisbane), Sep 2008.

Y. Pantazis, O. Rosec, and Y. Stylianou, "Iterative Estimation of Sinusoidal Signal Parameters," vol. 17, no. 5, pp. 461–464, 2010.

Y. Pantazis, O. Rosec, and Y. Stylianou, "Adaptive AMFM signal decomposition with application to speech analysis," *IEEE Trans. on Audio, Speech, and Language Processing*, vol. 19, pp. 290–300, 2011.

Acknowledgments

- My colleagues: Yannis Pantazis, Olivier Rosec (Voxygen), Vassilis Tsiaras, Gilles Degottex
- My ex-PhD Student: Giorgos Kafentzis
- Tom Quatieri and Prentice Hall for gave me the permission to use figures from Tom's book[1]

THANK YOU

for your attention on this part as well!