**ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ**
**ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ**

# Ψηφιακή Επεξεργασία Φωνής

**Διάλεξη:** Προσαρμόσιμα Ημιτονοειδή Μοντέλα

Παρουσίαση: Γιώργος Καφεντζής

Στυλιανού Ιωάννης
Τμήμα Επιστήμης Υπολογιστών

# Adaptive Sinusoidal Models

## A tutorial

George Kafentzis
Ph.D. student
University of Crete
University of Rennes I

November 2012

# Outline

# Introduction

- Sinusoidal modeling of speech

# Introduction

- Sinusoidal modeling of speech
- Dates back to the early 80ies

# Introduction

- Sinusoidal modeling of speech
- Dates back to the early 80ies
  - (Almeida et al, Quatieri et al, Xerra et al)

# Introduction

- Sinusoidal modeling of speech
- Dates back to the early 80ies
  - (Almeida et al, Quatieri et al, Xerra et al)
- Modeling of speech as a sum of sinusoids

# Introduction

- Sinusoidal modeling of speech
- Dates back to the early 80ies
  - (Almeida et al, Quatieri et al, Xerra et al)
- Modeling of speech as a sum of sinusoids
- Parameters: amplitude, frequency, and phase

# Introduction

- Sinusoidal modeling of speech
- Dates back to the early 80ies
  - (Almeida et al, Quatieri et al, Xerra et al)
- Modeling of speech as a sum of sinusoids
- Parameters: amplitude, frequency, and phase
- Essential assumption: **local stationarity!**

# Introduction

- Sinusoidal modeling of speech
- Dates back to the early 80ies
  - (Almeida et al, Quatieri et al, Xerra et al)
- Modeling of speech as a sum of sinusoids
- Parameters: amplitude, frequency, and phase
- Essential assumption: **local stationarity!**
  - Speech is considered stationary in short time intervals

# Introduction

- Sinusoidal modeling of speech
- Dates back to the early 80ies
  - (Almeida et al, Quatieri et al, Xerra et al)
- Modeling of speech as a sum of sinusoids
- Parameters: amplitude, frequency, and phase
- Essential assumption: **local stationarity!**
  - Speech is considered stationary in short time intervals
  - (it is not, but it is a convenient assumption :-) )

# Sinusoidal modeling - Quatieri, McAulay, 1986

- Each frame is modeled as a sum of sinusoids:

$$s(t) = \sum_{k=-K}^{K} a_k e^{j(2\pi f_k t)}$$

# Sinusoidal modeling - Quatieri, McAulay, 1986

- Each frame is modeled as a sum of sinusoids:

$$s(t) = \sum_{k=-K}^{K} a_k e^{j(2\pi f_k t)}$$

- $a_k$: complex amplitudes of the $k^{th}$ sinusoid

# Sinusoidal modeling - Quatieri, McAulay, 1986

- Each frame is modeled as a sum of sinusoids:

$$s(t) = \sum_{k=-K}^{K} a_k e^{j(2\pi f_k t)}$$

- $a_k$: complex amplitudes of the $k^{th}$ sinusoid
- $f_k$: frequency of the $k^{th}$ sinusoid

# Sinusoidal modeling - Quatieri, McAulay, 1986

- Each frame is modeled as a sum of sinusoids:

$$s(t) = \sum_{k=-K}^{K} a_k e^{j(2\pi f_k t)}$$

- $a_k$: complex amplitudes of the $k^{th}$ sinusoid
- $f_k$: frequency of the $k^{th}$ sinusoid
    - Estimation of the sinusoidal parameters: FFT + peak picking

# Sinusoidal modeling - Quatieri, McAulay, 1986

- Each frame is modeled as a sum of sinusoids:

$$s(t) = \sum_{k=-K}^{K} a_k e^{j(2\pi f_k t)}$$

- $a_k$: complex amplitudes of the $k^{th}$ sinusoid
- $f_k$: frequency of the $k^{th}$ sinusoid
    - Estimation of the sinusoidal parameters: FFT + peak picking
    - Various improvements (e.g. quadratic interpolation)

# Sinusoidal modeling - Quatieri, McAulay, 1986

- Each frame is modeled as a sum of sinusoids:

$$s(t) = \sum_{k=-K}^{K} a_k e^{j(2\pi f_k t)}$$

- $a_k$: complex amplitudes of the $k^{th}$ sinusoid
- $f_k$: frequency of the $k^{th}$ sinusoid
  - Estimation of the sinusoidal parameters: FFT + peak picking
  - Various improvements (e.g. quadratic interpolation)
- Highlight: **no distinction between voiced and unvoiced frames**

# Sinusoidal model - Quatieri, McAulay, 1986

Signal reconstruction:

- Overlap Add method: $\hat{s}(t) = \sum_{k=-K}^{K} \hat{a}_k e^{j2\pi \hat{f}_k t}$

# Sinusoidal model - Quatieri, McAulay, 1986

Signal reconstruction:

- Overlap Add method: $\hat{s}(t) = \sum_{k=-K}^{K} \hat{a}_k e^{j2\pi \hat{f}_k t}$
- Frame by frame parameter interpolation (PI):

# Sinusoidal model - Quatieri, McAulay, 1986

Signal reconstruction:

- Overlap Add method: $\hat{s}(t) = \sum_{k=-K}^{K} \hat{a}_k e^{j2\pi \hat{f}_k t}$
- Frame by frame parameter interpolation (PI):
  - Linear amplitude interpolation

# Sinusoidal model - Quatieri, McAulay, 1986

Signal reconstruction:

- Overlap Add method: $\hat{s}(t) = \sum_{k=-K}^{K} \hat{a}_k e^{j2\pi \hat{f}_k t}$
- Frame by frame parameter interpolation (PI):
  - Linear amplitude interpolation
  - Cubic phase interpolation

# Sinusoidal model - Quatieri, McAulay, 1986

- Pros:

# Sinusoidal model - Quatieri, McAulay, 1986

- Pros:
  - Fast

# Sinusoidal model - Quatieri, McAulay, 1986

- Pros:
  - Fast
  - Good signal reconstruction

# Sinusoidal model - Quatieri, McAulay, 1986

- Pros:
  - Fast
  - Good signal reconstruction
- Cons:

# Sinusoidal model - Quatieri, McAulay, 1986

- Pros:
  - Fast
  - Good signal reconstruction
- Cons:
  - Local stationarity assumption holds

# Sinusoidal model - Quatieri, McAulay, 1986

- Pros:
  - Fast
  - Good signal reconstruction
- Cons:
  - Local stationarity assumption holds
  - Not good modifications

# Sinusoidal model - Quatieri, McAulay, 1986

- Pros:
  - Fast
  - Good signal reconstruction
- Cons:
  - Local stationarity assumption holds
  - Not good modifications
  - Requires large windows

# Harmonic + Noise model - Stylianou, 1993-1996

- Separate signal into periodic (deterministic) and aperiodic (stochastic) components:

# Harmonic + Noise model - Stylianou, 1993-1996

- Separate signal into periodic (deterministic) and aperiodic (stochastic) components:
- $s(t) = s_d(t) + s_s(t)$

# Harmonic + Noise model - Stylianou, 1993-1996

- Separate signal into periodic (deterministic) and aperiodic (stochastic) components:
- $s(t) = s_d(t) + s_s(t)$
- Highlight: **In voiced frames, $\mathbf{f_k = kf_0}$ : deterministic $\longrightarrow$ harmonic**

$$s_d(t) = \sum_{k=-K}^{K} a_k e^{j2\pi k \hat{f}_0 t} w(t)$$

# Harmonic + Noise model - Stylianou, 1993-1996

- Separate signal into periodic (deterministic) and aperiodic (stochastic) components:
- $s(t) = s_d(t) + s_s(t)$
- Highlight: **In voiced frames, $\mathbf{f_k = kf_0}$ : deterministic $\longrightarrow$ harmonic**

$$s_d(t) = \sum_{k=-K}^{K} a_k e^{j2\pi k\hat{f}_0 t} w(t)$$

- Least Squares method for finding amplitudes and phases ($f_0$ is considered as known)

# Harmonic + Noise model - Stylianou, 1993-1996

- Separate signal into periodic (deterministic) and aperiodic (stochastic) components:
- $s(t) = s_d(t) + s_s(t)$
- Highlight: **In voiced frames,** $\mathbf{f_k = kf_0}$ **: deterministic** $\longrightarrow$ **harmonic**

$$s_d(t) = \sum_{k=-K}^{K} a_k e^{j2\pi k \hat{f}_0 t} w(t)$$

- Least Squares method for finding amplitudes and phases ($f_0$ is considered as known)
- Window length in HNM < Window length in SM

# Harmonic + Noise model - Stylianou, 1993-1996

- Separate signal into periodic (deterministic) and aperiodic (stochastic) components:
- $s(t) = s_d(t) + s_s(t)$
- Highlight: **In voiced frames, $\mathbf{f_k = kf_0}$ : deterministic $\longrightarrow$ harmonic**

$$s_d(t) = \sum_{k=-K}^{K} a_k e^{j2\pi k \hat{f}_0 t} w(t)$$

- Least Squares method for finding amplitudes and phases ($f_0$ is considered as known)
- Window length in HNM $<$ Window length in SM
- Reconstruction:

# Harmonic + Noise model - Stylianou, 1993-1996

- Separate signal into periodic (deterministic) and aperiodic (stochastic) components:
- $s(t) = s_d(t) + s_s(t)$
- Highlight: **In voiced frames, $\mathbf{f_k = kf_0}$ : deterministic $\longrightarrow$ harmonic**

$$s_d(t) = \sum_{k=-K}^{K} a_k e^{j2\pi k \hat{f}_0 t} w(t)$$

- Least Squares method for finding amplitudes and phases ($f_0$ is considered as known)
- Window length in HNM $<$ Window length in SM
- Reconstruction:
    - OLA or PI for harmonic part

# Harmonic + Noise model - Stylianou, 1993-1996

- Separate signal into periodic (deterministic) and aperiodic (stochastic) components:
- $s(t) = s_d(t) + s_s(t)$
- Highlight: **In voiced frames, $\mathbf{f_k = kf_0}$ : deterministic $\longrightarrow$ harmonic**

$$s_d(t) = \sum_{k=-K}^{K} a_k e^{j2\pi k \hat{f}_0 t} w(t)$$

- Least Squares method for finding amplitudes and phases ($f_0$ is considered as known)
- Window length in HNM < Window length in SM
- Reconstruction:
  - OLA or PI for harmonic part
  - OLA for stochastic part

# Harmonic + Noise model - Stylianou, 1993-1996

- Pros:

# Harmonic + Noise model - Stylianou, 1993-1996

- Pros:
  - Pitch synchronous analysis

# Harmonic + Noise model - Stylianou, 1993-1996

- Pros:
  - Pitch synchronous analysis
  - Smaller window lengths

# Harmonic + Noise model - Stylianou, 1993-1996

- Pros:
  - Pitch synchronous analysis
  - Smaller window lengths
  - Convenient for modifications

# Harmonic + Noise model - Stylianou, 1993-1996

- Pros:
  - Pitch synchronous analysis
  - Smaller window lengths
  - Convenient for modifications
- Cons:

# Harmonic + Noise model - Stylianou, 1993-1996

- Pros:
  - Pitch synchronous analysis
  - Smaller window lengths
  - Convenient for modifications
- Cons:
  - Local stationarity assumption holds

# Harmonic + Noise model - Stylianou, 1993-1996

- Pros:
  - Pitch synchronous analysis
  - Smaller window lengths
  - Convenient for modifications
- Cons:
  - Local stationarity assumption holds
  - Depends on good $f_0$ estimation

# Harmonic + Noise model - Stylianou, 1993-1996

- Pros:
  - Pitch synchronous analysis
  - Smaller window lengths
  - Convenient for modifications
- Cons:
  - Local stationarity assumption holds
  - Depends on good $f_0$ estimation
  - **Speech is not purely harmonic** ($\mathbf{f_k \approx kf_0}$)

# Harmonic + Noise model - Stylianou, 1993-1996

- Pros:
  - Pitch synchronous analysis
  - Smaller window lengths
  - Convenient for modifications
- Cons:
  - Local stationarity assumption holds
  - Depends on good $f_0$ estimation
  - **Speech is not purely harmonic ($\mathbf{f_k \approx kf_0}$)**
    - ...this last observation is the motivation for the following model...

# Outline

# Quasi-Harmonic model - Pantazis, Stylianou, Rosec, 2007-2010

- $s_d(t) = \sum_{k=-K}^{K} (a_k + tb_k) e^{j2\pi \hat{f}_k t} w(t)$

# Quasi-Harmonic model - Pantazis, Stylianou, Rosec, 2007-2010

- $s_d(t) = \sum_{k=-K}^{K} (a_k + tb_k) e^{j2\pi \hat{f}_k t} w(t)$
- $a_k, b_k$ are complex numbers

# Quasi-Harmonic model - Pantazis, Stylianou, Rosec, 2007-2010

- $s_d(t) = \sum_{k=-K}^{K} (a_k + tb_k) e^{j2\pi \hat{f}_k t} w(t)$
- $a_k, b_k$ are complex numbers
- usually $f_k = kf_0$, where $f_0$ is considered as known

# Quasi-Harmonic model - Pantazis, Stylianou, Rosec, 2007-2010

- $s_d(t) = \sum_{k=-K}^{K} (a_k + tb_k) e^{j2\pi \hat{f}_k t} w(t)$
- $a_k, b_k$ are complex numbers
- usually $f_k = kf_0$, where $f_0$ is considered as known
- $w(t)$ is the analysis window

# Quasi-Harmonic model - Pantazis, Stylianou, Rosec, 2007-2010

- $s_d(t) = \sum_{k=-K}^{K} (a_k + tb_k)e^{j2\pi \hat{f}_k t} w(t)$
- $a_k, b_k$ are complex numbers
- usually $f_k = kf_0$, where $f_0$ is considered as known
- $w(t)$ is the analysis window
- Again: Least Squares method for finding amplitudes

# Quasi-Harmonic model - Pantazis, Stylianou, Rosec, 2007-2010

- $s_d(t) = \sum_{k=-K}^{K} (a_k + tb_k) e^{j2\pi \hat{f}_k t} w(t)$
- $a_k, b_k$ are complex numbers
- usually $f_k = kf_0$, where $f_0$ is considered as known
- $w(t)$ is the analysis window
- Again: Least Squares method for finding amplitudes
- Window length $\approx$ 3 pitch periods

# Quasi-Harmonic model - Pantazis, Stylianou, Rosec, 2007-2010

- $s_d(t) = \sum_{k=-K}^{K} (a_k + tb_k) e^{j2\pi \hat{f}_k t} w(t)$
- $a_k, b_k$ are complex numbers
- usually $f_k = kf_0$, where $f_0$ is considered as known
- $w(t)$ is the analysis window
- Again: Least Squares method for finding amplitudes
- Window length $\approx$ 3 pitch periods
- Reconstruction:

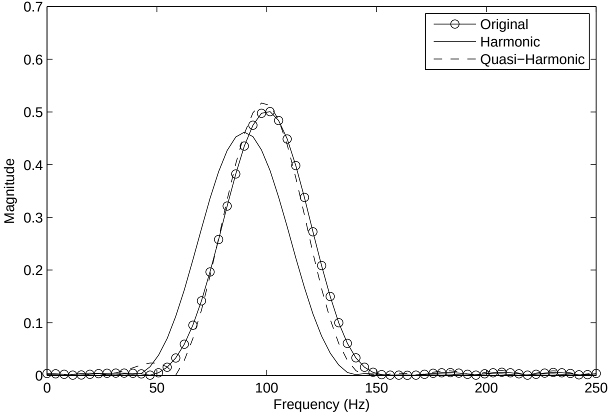
$$\hat{s}_d(t) = \sum_{k=-K}^{K} (\hat{a}_k + t\hat{b}_k) e^{j2\pi \hat{f}_k t} \tag{1}$$

# Quasi-Harmonic model

- HM versus QHM in frequency estimation - pure tone @ 100 Hz

# Quasi-Harmonic model

- HM versus QHM in frequency estimation - pure tone @ 100 Hz
- given frequency for both models: 90 Hz



# Quasi-Harmonic model

Time domain properties:

- Inst. amplitude:
  $M_k(t) = |a_k + tb_k| = \sqrt{(a_k^R + tb_k^R)^2 + (a_k^I + tb_k^I)^2}$

# Quasi-Harmonic model

Time domain properties:

- Inst. amplitude:
  $M_k(t) = |a_k + tb_k| = \sqrt{(a_k^R + tb_k^R)^2 + (a_k^I + tb_k^I)^2}$
- Inst. phase: $\Phi_k(t) = 2\pi \hat{f}_k t + tan^{-1} \dfrac{a_k^I + tb_k^I}{a_k^R + tb_k^R}$

# Quasi-Harmonic model

Time domain properties:

- Inst. amplitude:
$$M_k(t) = |a_k + tb_k| = \sqrt{(a_k^R + tb_k^R)^2 + (a_k^I + tb_k^I)^2}$$
- Inst. phase: $\Phi_k(t) = 2\pi\hat{f}_k t + tan^{-1}\dfrac{a_k^I + tb_k^I}{a_k^R + tb_k^R}$
- Inst. frequency: $F_k(t) = \dfrac{1}{2\pi}\Phi'(t) = \hat{f}_k + \dfrac{1}{2\pi}\dfrac{a_k^R b_k^I - a_k^I b_k^R}{M_k^2(t)}$

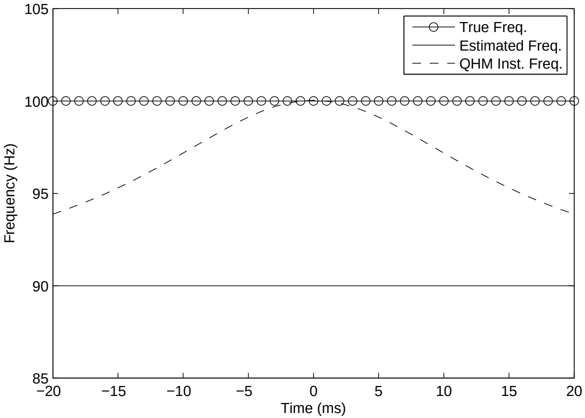
# Quasi-Harmonic model

Time domain properties:

- Inst. amplitude:
  $M_k(t) = |a_k + tb_k| = \sqrt{(a_k^R + tb_k^R)^2 + (a_k^I + tb_k^I)^2}$
- Inst. phase: $\Phi_k(t) = 2\pi\hat{f}_k t + tan^{-1}\frac{a_k^I + tb_k^I}{a_k^R + tb_k^R}$
- Inst. frequency: $F_k(t) = \frac{1}{2\pi}\Phi'(t) = \hat{f}_k + \frac{1}{2\pi}\frac{a_k^R b_k^I - a_k^I b_k^R}{M_k^2(t)}$
- where $x^R, x^I$ denote the real and imaginary part of $x$

# Quasi-Harmonic model

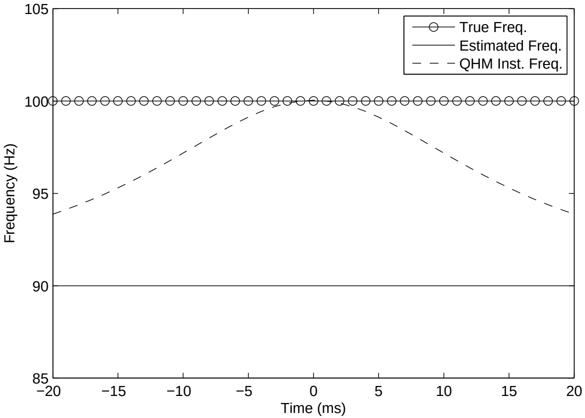
HM vs QHM on frequency tracks - pure tone @ 100 Hz:




- Highlight: **frequency correction mechanism**

# Quasi-Harmonic model

HM vs QHM on frequency tracks - pure tone @ 100 Hz:




- Highlight: **frequency correction mechanism**
  - Let's discuss a bit on that...

# Frequency mismatch correction

Frequency domain view:

- Fourier Transform of the model:

# Frequency mismatch correction

Frequency domain view:

- Fourier Transform of the model:
- Reminder:

$$x_k(t) = a_k e^{j2\pi f_k t} w(t) + t b_k e^{j2\pi f_k t} w(t)$$

# Frequency mismatch correction

Frequency domain view:

- Fourier Transform of the model:
- Reminder:

$$x_k(t) = a_k e^{j2\pi f_k t} w(t) + t b_k e^{j2\pi f_k t} w(t)$$

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

# Frequency mismatch correction

Frequency domain view:

- Fourier Transform of the model:
- Reminder:

$$x_k(t) = a_k e^{j2\pi f_k t} w(t) + t b_k e^{j2\pi f_k t} w(t)$$

  - $$X_k(f) = a_k W(f - \hat{f}_k) + j\frac{b_k}{2\pi} W'(f - \hat{f}_k) \quad (2)$$
  - where $W(f) = FT\{w(t)\}$ and $W'(f) = dW(f)/df$.

# Frequency mismatch correction

Frequency domain view:

- Fourier Transform of the model:
- Reminder:

$$x_k(t) = a_k e^{j2\pi f_k t} w(t) + t b_k e^{j2\pi f_k t} w(t)$$

  - $$X_k(f) = a_k W(f - \hat{f}_k) + j\frac{b_k}{2\pi} W'(f - \hat{f}_k) \quad (2)$$
  - where $W(f) = FT\{w(t)\}$ and $W'(f) = dW(f)/df$.

- Projecting $b_k$ to $a_k$:

$$b_k = \rho_{1,k} a_k + \rho_{2,k} j a_k \quad (3)$$

# Frequency mismatch correction

Frequency domain view:

- Fourier Transform of the model:
- Reminder:

$$x_k(t) = a_k e^{j2\pi f_k t} w(t) + t b_k e^{j2\pi f_k t} w(t)$$

  - $$X_k(f) = a_k W(f - \hat{f}_k) + j\frac{b_k}{2\pi} W'(f - \hat{f}_k) \quad (2)$$
  - where $W(f) = FT\{w(t)\}$ and $W'(f) = dW(f)/df$.

- Projecting $b_k$ to $a_k$:

$$b_k = \rho_{1,k} a_k + \rho_{2,k} j a_k \quad (3)$$

- Then,

$$X_k(f) = a_k \left[ W(f - \hat{f}_k) - \frac{\rho_{2,k}}{2\pi} W'(f - \hat{f}_k) + j\frac{\rho_{1,k}}{2\pi} W'(f - \hat{f}_k) \right] \quad (4)$$

# Frequency mismatch correction

- Taylor series expansion of $W(f - \hat{f}_k - \frac{\rho_{2,k}}{2\pi})$:

# Frequency mismatch correction

- Taylor series expansion of $W(f - \hat{f}_k - \frac{\rho_{2,k}}{2\pi})$:
    - $$W(f - \hat{f}_k - \frac{\rho_{2,k}}{2\pi}) = W(f - \hat{f}_k) - \frac{\rho_{2,k}}{2\pi} W'(f - \hat{f}_k) + O(\rho_{2,k}^2 W''(f - \hat{f}_k)) \quad (5)$$

# Frequency mismatch correction

- Taylor series expansion of $W(f - \hat{f}_k - \frac{\rho_{2,k}}{2\pi})$:
    - $$W(f - \hat{f}_k - \frac{\rho_{2,k}}{2\pi}) = W(f - \hat{f}_k) - \frac{\rho_{2,k}}{2\pi} W'(f - \hat{f}_k) + O(\rho_{2,k}^2 W''(f - \hat{f}_k)) \quad (5)$$
    - If the value of term $W''(f)$ at $f_k$ is small, then for small values of $\rho_{2,k}$, it is:

# Frequency mismatch correction

- Taylor series expansion of $W(f - \hat{f}_k - \frac{\rho_{2,k}}{2\pi})$:
  - $$W(f - \hat{f}_k - \frac{\rho_{2,k}}{2\pi}) = W(f - \hat{f}_k) - \frac{\rho_{2,k}}{2\pi} W'(f - \hat{f}_k) + O(\rho_{2,k}^2 W''(f - \hat{f}_k)) \quad (5)$$
  - If the value of term $W''(f)$ at $f_k$ is small, then for small values of $\rho_{2,k}$, it is:
  - $$W(f - \hat{f}_k - \frac{\rho_{2,k}}{2\pi}) \approx W(f - \hat{f}_k) - \frac{\rho_{2,k}}{2\pi} W'(f - \hat{f}_k) \quad (6)$$

# Frequency mismatch correction

- Taylor series expansion of $W(f-\hat{f}_k-\frac{\rho_{2,k}}{2\pi})$:
  - $$W(f-\hat{f}_k-\frac{\rho_{2,k}}{2\pi}) = W(f-\hat{f}_k)-\frac{\rho_{2,k}}{2\pi}W'(f-\hat{f}_k)+ O(\rho_{2,k}^2 W''(f-\hat{f}_k)) \quad (5)$$
  - If the value of term $W''(f)$ at $f_k$ is small, then for small values of $\rho_{2,k}$, it is:
  - $$W(f-\hat{f}_k-\frac{\rho_{2,k}}{2\pi}) \approx W(f-\hat{f}_k)-\frac{\rho_{2,k}}{2\pi}W'(f-\hat{f}_k) \quad (6)$$
- So,

$$X_k(f) \approx a_k\left[W(f-\hat{f}_k-\frac{\rho_{2,k}}{2\pi})+j\frac{\rho_{1,k}}{2\pi}W'(f-\hat{f}_k)\right] \quad (7)$$

# Frequency mismatch correction

- which goes back in time domain as...

# Frequency mismatch correction

- which goes back in time domain as...
- $$x_k(t) \approx a_k \left[ e^{j(2\pi \hat{f}_k + \rho_{2,k})t} + \rho_{1,k} t e^{j2\pi \hat{f}_k t} \right] w(t) \qquad (8)$$

# Frequency mismatch correction

- which goes back in time domain as...
- $$x_k(t) \approx a_k \left[ e^{j(2\pi \hat{f}_k + \rho_{2,k})t} + \rho_{1,k} t e^{j2\pi \hat{f}_k t} \right] w(t) \tag{8}$$
- Then, $\rho_{2,k}/2\pi$ can be an estimator of the frequency error $\eta_k$:

$$\hat{\eta}_k = \rho_{2,k}/2\pi = \frac{1}{2\pi} \frac{a_k^R b_k^I - a_k^I b_k^R}{|a_k|^2} \tag{9}$$

# Frequency mismatch correction

- which goes back in time domain as...
- $$x_k(t) \approx a_k \left[ e^{j(2\pi \hat{f}_k + \rho_{2,k})t} + \rho_{1,k} t e^{j2\pi \hat{f}_k t} \right] w(t) \tag{8}$$
- Then, $\rho_{2,k}/2\pi$ can be an estimator of the frequency error $\eta_k$:

$$\hat{\eta}_k = \rho_{2,k}/2\pi = \frac{1}{2\pi} \frac{a_k^R b_k^I - a_k^I b_k^R}{|a_k|^2} \tag{9}$$

- In other words, QHM suggests a frequency correction to the input frequencies $\hat{f}_k$ (or a frequency estimator). This suggestion is however conditional on the magnitude of $\rho_{2,k}$ and the value of term $W''(f)$ at $f_k$

# Frequency mismatch correction

- which goes back in time domain as...
- $$x_k(t) \approx a_k \left[ e^{j(2\pi \hat{f}_k + \rho_{2,k})t} + \rho_{1,k} t e^{j2\pi \hat{f}_k t} \right] w(t) \qquad (8)$$
- Then, $\rho_{2,k}/2\pi$ can be an estimator of the frequency error $\eta_k$:

$$\hat{\eta}_k = \rho_{2,k}/2\pi = \frac{1}{2\pi} \frac{a_k^R b_k^I - a_k^I b_k^R}{|a_k|^2} \qquad (9)$$

- In other words, QHM suggests a frequency correction to the input frequencies $\hat{f}_k$ (or a frequency estimator). This suggestion is however conditional on the magnitude of $\rho_{2,k}$ and the value of term $W''(f)$ at $f_k$
- Also, the correction term depends on the window mainlobe width

# Outline

# iterative QHM - Pantazis, Stylianou, Rosec, 2007-2010

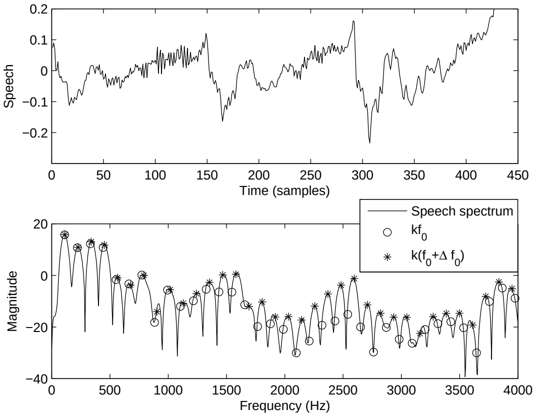
- This frequency updating mechanism provides frequencies which can be used in the model iteratively and result in better parameter estimation $(a_k, b_k)$

# iterative QHM - Pantazis, Stylianou, Rosec, 2007-2010

- This frequency updating mechanism provides frequencies which can be used in the model iteratively and result in better parameter estimation $(a_k, b_k)$
- This iterative parameter estimation is referred to as the *iterative QHM*

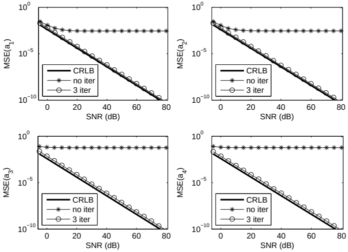
# iterative QHM

HM versus iQHM in frequency estimation - speech signal:

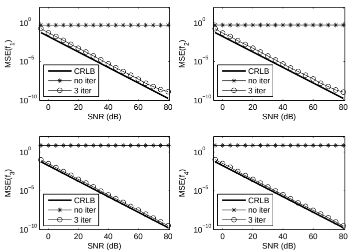


# iterative QHM

Noise robustness:




(a) MSE for amplitudes




(b) MSE for frequencies

Figure: Noise Robustness

# iterative QHM

- Pros:

# iterative QHM

- Pros:
  - Linear amplitude evolution

# iterative QHM

- Pros:
  - Linear amplitude evolution
  - Frequency mismatch correction: $\eta_k = f_k - \hat{f}_k$

# iterative QHM

- Pros:
  - Linear amplitude evolution
  - Frequency mismatch correction: $\eta_k = f_k - \hat{f}_k$
- Cons:

# iterative QHM

- Pros:
  - Linear amplitude evolution
  - Frequency mismatch correction: $\eta_k = f_k - \hat{f}_k$
- Cons:
  - Needs larger analysis window

# iterative QHM

- Pros:
  - Linear amplitude evolution
  - Frequency mismatch correction: $\eta_k = f_k - \hat{f}_k$
- Cons:
  - Needs larger analysis window
  - And what about...

# iterative QHM

- Pros:
  - Linear amplitude evolution
  - Frequency mismatch correction: $\eta_k = f_k - \hat{f}_k$
- Cons:
  - Needs larger analysis window
  - And what about...
    - **local stationarity??**

# iterative QHM

- Pros:
  - Linear amplitude evolution
  - Frequency mismatch correction: $\eta_k = f_k - \hat{f}_k$
- Cons:
  - Needs larger analysis window
  - And what about...
    - **local stationarity??**
    - Speech is non stationary even in very short time intervals

# iterative QHM

- Pros:
  - Linear amplitude evolution
  - Frequency mismatch correction: $\eta_k = f_k - \hat{f}_k$
- Cons:
  - Needs larger analysis window
  - And what about...
    - **local stationarity??**
    - Speech is non stationary even in very short time intervals
    - iQHM still holds the local stationary assumption

# Outline

# Adaptive Quasi-Harmonic Model

- Adaptive models tackle the problem of local non stationarity[5] - How?

# Adaptive Quasi-Harmonic Model

- Adaptive models tackle the problem of local non stationarity[5] - How?
  - By projecting the signal onto non-stationary basis functions $e^{j\phi_k(t)}$!

# Adaptive Quasi-Harmonic Model

- Adaptive models tackle the problem of local non stationarity[5] - How?
  - By projecting the signal onto non-stationary basis functions $e^{j\phi_k(t)}$!
- Model:

$$s_d(t) = \sum_{k=-K}^{K} (a_k + tb_k)\left(e^{j\phi_k(t)}\right) w(t)$$

# Adaptive Quasi-Harmonic Model

- Adaptive models tackle the problem of local non stationarity[5] - How?
  - By projecting the signal onto non-stationary basis functions $e^{j\phi_k(t)}$!
- Model:

$$s_d(t) = \sum_{k=-K}^{K} (a_k + tb_k)\left(e^{j\phi_k(t)}\right)w(t)$$

- complex amplitudes $a_k, b_k$: estimated via Least Squares

# Adaptive Quasi-Harmonic Model

- Adaptive models tackle the problem of local non stationarity[5] - How?
  - By projecting the signal onto non-stationary basis functions $e^{j\phi_k(t)}$!
- Model:
$$s_d(t) = \sum_{k=-K}^{K} (a_k + t b_k)\left(e^{j\phi_k(t)}\right) w(t)$$
- complex amplitudes $a_k, b_k$: estimated via Least Squares
  -
$$\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = (E^H W^H W E)^{-1} E^H W^H W s \qquad (10)$$

# Adaptive Quasi-Harmonic Model

- Adaptive models tackle the problem of local non stationarity[5] - How?
  - By projecting the signal onto non-stationary basis functions $e^{j\phi_k(t)}$!
- Model:

$$s_d(t) = \sum_{k=-K}^{K} (a_k + tb_k)\left(e^{j\phi_k(t)}\right) w(t)$$

- complex amplitudes $a_k, b_k$: estimated via Least Squares
  -

$$\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = (E^H W^H W E)^{-1} E^H W^H W s \tag{10}$$

  - where $E = [E_0|E_1]$:

# Adaptive Quasi-Harmonic Model

- Adaptive models tackle the problem of local non stationarity[5] - How?
    - By projecting the signal onto non-stationary basis functions $e^{j\phi_k(t)}$!
- Model:

$$s_d(t) = \sum_{k=-K}^{K} (a_k + tb_k)\left(e^{j\phi_k(t)}\right) w(t)$$

- complex amplitudes $a_k, b_k$: estimated via Least Squares
    - $$\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = (E^H W^H W E)^{-1} E^H W^H W s \qquad (10)$$
    - where $E = [E_0|E_1]$:
    - $(E_0)_{n,k} =) e^{j\phi_k(t_n)}$

# Adaptive Quasi-Harmonic Model

- Adaptive models tackle the problem of local non stationarity[5] - How?
  - By projecting the signal onto non-stationary basis functions $e^{j\phi_k(t)}$!
- Model:

$$s_d(t) = \sum_{k=-K}^{K} (a_k + tb_k)\left(e^{j\phi_k(t)}\right) w(t)$$

- complex amplitudes $a_k, b_k$: estimated via Least Squares
  -
$$\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = (E^H W^H W E)^{-1} E^H W^H W s \qquad (10)$$
  - where $E = [E_0 | E_1]$:
  - $(E_0)_{n,k} =) e^{j\phi_k(t_n)}$
  - $(E_1)_{n,k} = t_n e^{j\phi_k(t_n)} = t_n (E_0)_{n,k}$

# Adaptive Quasi-Harmonic Model

- Definition of phase:

$$\breve{\phi}_k(t) = \hat{\phi}_k(t_{l-1}) + \int_{t_{l-1}}^{t} 2\pi \hat{f}_k(u)du$$

# Adaptive Quasi-Harmonic Model

- Definition of phase:

$$\breve{\phi}_k(t) = \hat{\phi}_k(t_{l-1}) + \int_{t_{l-1}}^{t} 2\pi \hat{f}_k(u) du$$

- where $\hat{f}_k(t)$ is the estimated instantaneous frequency

# Adaptive Quasi-Harmonic Model

- Definition of phase:

$$\breve{\phi}_k(t) = \hat{\phi}_k(t_{l-1}) + \int_{t_{l-1}}^{t} 2\pi \hat{f}_k(u)du$$

- where $\hat{f}_k(t)$ is the estimated instantaneous frequency
- Estimation error due to non stationarity is reduced [5]

# Adaptive Quasi-Harmonic Model

- Definition of phase:

$$\breve{\phi}_k(t) = \hat{\phi}_k(t_{l-1}) + \int_{t_{l-1}}^{t} 2\pi \hat{f}_k(u)du$$

- where $\hat{f}_k(t)$ is the estimated instantaneous frequency
- Estimation error due to non stationarity is reduced [5]
- Signal is reconstructed using its instantaneous components $\hat{A}_k(t), \hat{f}_k(t), \hat{\phi}_k(t)$:

# Adaptive Quasi-Harmonic Model

- Definition of phase:

$$\breve{\phi}_k(t) = \hat{\phi}_k(t_{l-1}) + \int_{t_{l-1}}^{t} 2\pi \hat{f}_k(u)du$$

- where $\hat{f}_k(t)$ is the estimated instantaneous frequency
- Estimation error due to non stationarity is reduced [5]
- Signal is reconstructed using its instantaneous components $\hat{A}_k(t), \hat{f}_k(t), \hat{\phi}_k(t)$:
  - $\hat{x}(t) = \sum_{k=-K}^{K} \hat{A}_k(t) e^{j\hat{\phi}_k(t)}$

# Adaptive Quasi-Harmonic Model

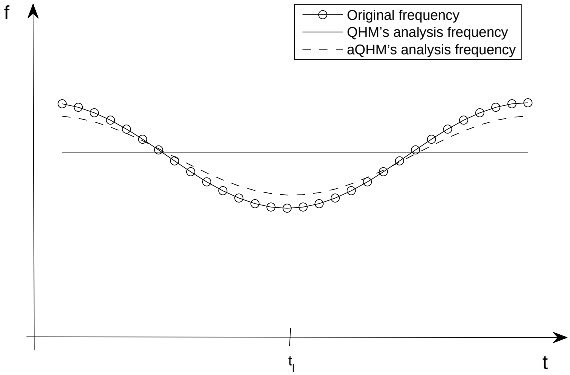
- Definition of phase:

$$\breve{\phi}_k(t) = \hat{\phi}_k(t_{l-1}) + \int_{t_{l-1}}^{t} 2\pi \hat{f}_k(u) du$$

- where $\hat{f}_k(t)$ is the estimated instantaneous frequency
- Estimation error due to non stationarity is reduced [5]
- Signal is reconstructed using its instantaneous components $\hat{A}_k(t), \hat{f}_k(t), \hat{\phi}_k(t)$:
  - $\hat{x}(t) = \sum_{k=-K}^{K} \hat{A}_k(t) e^{j\hat{\phi}_k(t)}$
- Adaptation algorithm can be found in [5]

# Adaptive Quasi-Harmonic model

QHM vs aQHM:



# Adaptation algorithm for aQHM

- Let $\hat{A}_k(t_l), \hat{f}_k(t_l), \hat{\phi}_k(t_l)$ denote the inst. amplitude, frequency, and phase at time instant $t_l$ of the $k^{th}$ component, with $l = 1, \cdots, L$, where $L$ is the number of frames:

# Adaptation algorithm for aQHM

- Let $\hat{A}_k(t_l), \hat{f}_k(t_l), \hat{\phi}_k(t_l)$ denote the inst. amplitude, frequency, and phase at time instant $t_l$ of the $k^{th}$ component, with $l = 1, \cdots, L$, where $L$ is the number of frames:
- Initialization: (QHM)
  1) $\hat{f}_k^0(t_l) = \hat{f}_k^0(t_{l-1}) + \rho_{2,k}^0/2\pi$
  2) $\hat{A}_k^0(t_l) = |a_k^0|, \;\; \hat{\phi}_k^0(t_l) = \angle a_k^0$
  3) $\hat{f}_k^0(t_{l+1}) = \hat{f}_k^0(t_l)$

# Adaptation algorithm for aQHM

- Let $\hat{A}_k(t_l), \hat{f}_k(t_l), \hat{\phi}_k(t_l)$ denote the inst. amplitude, frequency, and phase at time instant $t_l$ of the $k^{th}$ component, with $l = 1, \cdots, L$, where $L$ is the number of frames:
- Initialization: (QHM)
  1) $\hat{f}_k^0(t_l) = \hat{f}_k^0(t_{l-1}) + \rho_{2,k}^0/2\pi$
  2) $\hat{A}_k^0(t_l) = |a_k^0|, \ \hat{\phi}_k^0(t_l) = \angle a_k^0$
  3) $\hat{f}_k^0(t_{l+1}) = \hat{f}_k^0(t_l)$
- FOR adaptation $i = 1, 2, \cdots$
  FOR frame $l = 1, 2, \cdots, L$

END

# Adaptation algorithm for aQHM

- Let $\hat{A}_k(t_l), \hat{f}_k(t_l), \hat{\phi}_k(t_l)$ denote the inst. amplitude, frequency, and phase at time instant $t_l$ of the $k^{th}$ component, with $l = 1, \cdots, L$, where $L$ is the number of frames:
- Initialization: (QHM)
  1) $\hat{f}_k^0(t_l) = \hat{f}_k^0(t_{l-1}) + \rho_{2,k}^0 / 2\pi$
  2) $\hat{A}_k^0(t_l) = |a_k^0|, \;\; \hat{\phi}_k^0(t_l) = \angle a_k^0$
  3) $\hat{f}_k^0(t_{l+1}) = \hat{f}_k^0(t_l)$
- FOR adaptation $i = 1, 2, \cdots$

  FOR frame $l = 1, 2, \cdots, L$

  1. Compute $a_k^l, b_k^l$ using $\hat{\phi}_k^{i-1}(t)$ and (10)

END

# Adaptation algorithm for aQHM

- Let $\hat{A}_k(t_l), \hat{f}_k(t_l), \hat{\phi}_k(t_l)$ denote the inst. amplitude, frequency, and phase at time instant $t_l$ of the $k^{th}$ component, with $l = 1, \cdots, L$, where $L$ is the number of frames:
- Initialization: (QHM)
  1) $\hat{f}_k^0(t_l) = \hat{f}_k^0(t_{l-1}) + \rho_{2,k}^0/2\pi$
  2) $\hat{A}_k^0(t_l) = |a_k^0|, \ \ \hat{\phi}_k^0(t_l) = \angle a_k^0$
  3) $\hat{f}_k^0(t_{l+1}) = \hat{f}_k^0(t_l)$
- FOR adaptation $i = 1, 2, \cdots$
  FOR frame $l = 1, 2, \cdots, L$
  1. Compute $a_k^l, b_k^l$ using $\hat{\phi}_k^{i-1}(t)$ and (10)
  2. Update $\hat{f}_k^i(t_l)$ using (9)

END

# Adaptation algorithm for aQHM

- Let $\hat{A}_k(t_l), \hat{f}_k(t_l), \hat{\phi}_k(t_l)$ denote the inst. amplitude, frequency, and phase at time instant $t_l$ of the $k^{th}$ component, with $l = 1, \cdots, L$, where $L$ is the number of frames:
- Initialization: (QHM)
  1) $\hat{f}_k^0(t_l) = \hat{f}_k^0(t_{l-1}) + \rho_{2,k}^0/2\pi$
  2) $\hat{A}_k^0(t_l) = |a_k^0|, \;\; \hat{\phi}_k^0(t_l) = \angle a_k^0$
  3) $\hat{f}_k^0(t_{l+1}) = \hat{f}_k^0(t_l)$
- FOR adaptation $i = 1, 2, \cdots$

  FOR frame $l = 1, 2, \cdots, L$
  1. Compute $a_k^l, b_k^l$ using $\hat{\phi}_k^{i-1}(t)$ and (10)
  2. Update $\hat{f}_k^i(t_l)$ using (9)
  3. Compute $\hat{A}_k^i(t_l)$ and $\hat{\phi}_k^i(t_l)$

  END

END

# Adaptation algorithm for aQHM

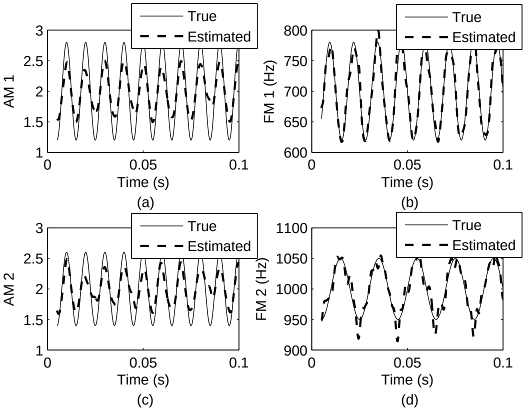
- Let $\hat{A}_k(t_l), \hat{f}_k(t_l), \hat{\phi}_k(t_l)$ denote the inst. amplitude, frequency, and phase at time instant $t_l$ of the $k^{th}$ component, with $l = 1, \cdots, L$, where $L$ is the number of frames:
- Initialization: (QHM)
  1) $\hat{f}_k^0(t_l) = \hat{f}_k^0(t_{l-1}) + \rho_{2,k}^0/2\pi$
  2) $\hat{A}_k^0(t_l) = |a_k^0|, \ \ \hat{\phi}_k^0(t_l) = \angle a_k^0$
  3) $\hat{f}_k^0(t_{l+1}) = \hat{f}_k^0(t_l)$
- FOR adaptation $i = 1, 2, \cdots$
  FOR frame $l = 1, 2, \cdots, L$
  1. Compute $a_k^l, b_k^l$ using $\hat{\phi}_k^{i-1}(t)$ and (10)
  2. Update $\hat{f}_k^i(t_l)$ using (9)
  3. Compute $\hat{A}_k^i(t_l)$ and $\hat{\phi}_k^i(t_l)$

  END
  4. Interpolation of the parameters $\{\hat{A}_k^i(t), \hat{f}_k^i(t), \hat{\phi}_k^i(t)\}$

  END

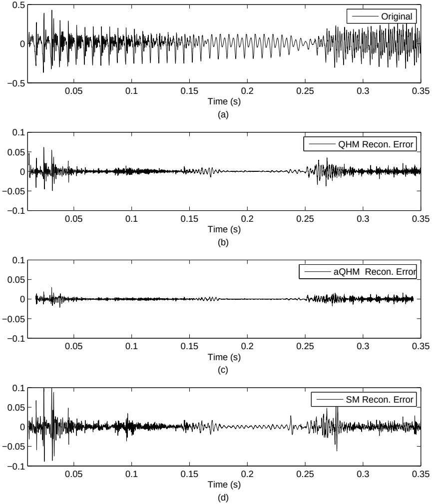
# Adaptive Quasi-Harmonic model

Synthetic signal:



# Adaptive Quasi-Harmonic model

Real Signal:



# aQHM

- Pros:

# aQHM

- Pros:
  - Phase adaptation, non-parametric approach

# aQHM

- Pros:
  - Phase adaptation, non-parametric approach
  - Local nonstationarity is partially solved

# aQHM

- Pros:
  - Phase adaptation, non-parametric approach
  - Local nonstationarity is partially solved
  - High signal-to-reconstruction error ratio (SRER):
    $SRER = 20 \log_{10} \frac{\sigma_{\hat{x}(t)}}{\sigma_{\hat{x}(t)-x(t)}}$

# aQHM

- Pros:
  - Phase adaptation, non-parametric approach
  - Local nonstationarity is partially solved
  - High signal-to-reconstruction error ratio (SRER):

    $$SRER = 20 \log_{10} \frac{\sigma_{\hat{x}(t)}}{\sigma_{\hat{x}(t)-x(t)}}$$
- Cons:

# aQHM

- Pros:
  - Phase adaptation, non-parametric approach
  - Local nonstationarity is partially solved
  - High signal-to-reconstruction error ratio (SRER):

    $$SRER = 20\log_{10}\frac{\sigma_{\hat{x}(t)}}{\sigma_{\hat{x}(t)-x(t)}}$$
- Cons:
  - Needs larger analysis window (as iQHM)

# aQHM

- Pros:
  - Phase adaptation, non-parametric approach
  - Local nonstationarity is partially solved
  - High signal-to-reconstruction error ratio (SRER):
    $SRER = 20 \log_{10} \frac{\sigma_{\hat{x}(t)}}{\sigma_{\hat{x}(t)-x(t)}}$
- Cons:
  - Needs larger analysis window (as iQHM)
  - Amplitudes are not adapted to the signal

# Outline

# Extended Adaptive Quasi-Harmonic model - Kafentzis, Pantazis, Stylianou, Rosec, 2011

- Adaptation is allowed for amplitude as well as phase

# Extended Adaptive Quasi-Harmonic model - Kafentzis, Pantazis, Stylianou, Rosec, 2011

- Adaptation is allowed for amplitude as well as phase
  - Projection of the signal onto non-stationary basis functions $\alpha_k(t)e^{j\phi_k(t)}$ [6]!

# Extended Adaptive Quasi-Harmonic model - Kafentzis, Pantazis, Stylianou, Rosec, 2011

- Adaptation is allowed for amplitude as well as phase
  - Projection of the signal onto non-stationary basis functions $\alpha_k(t)e^{j\phi_k(t)}$ [6]!
- Model:

$$s_d(t) = \sum_{k=-K}^{K} (a_k + tb_k)\left(\alpha_k(t)e^{j\phi_k(t)}\right)w(t)$$

# Extended Adaptive Quasi-Harmonic model - Kafentzis, Pantazis, Stylianou, Rosec, 2011

- Adaptation is allowed for amplitude as well as phase
  - Projection of the signal onto non-stationary basis functions $\alpha_k(t)e^{j\phi_k(t)}$ [6]!
- Model:

$$s_d(t) = \sum_{k=-K}^{K} (a_k + tb_k)\left(\alpha_k(t)e^{j\phi_k(t)}\right)w(t)$$

- complex amplitudes $a_k, b_k$: estimated via Least Squares

# Extended Adaptive Quasi-Harmonic model - Kafentzis, Pantazis, Stylianou, Rosec, 2011

- Adaptation is allowed for amplitude as well as phase
  - Projection of the signal onto non-stationary basis functions $\alpha_k(t)e^{j\phi_k(t)}$ [6]!
- Model:

$$s_d(t) = \sum_{k=-K}^{K} (a_k + tb_k)\left(\alpha_k(t)e^{j\phi_k(t)}\right)w(t)$$

- complex amplitudes $a_k, b_k$: estimated via Least Squares
  -

$$\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = (E^H W^H W E)^{-1} E^H W^H W s \qquad (11)$$

# Extended Adaptive Quasi-Harmonic model - Kafentzis, Pantazis, Stylianou, Rosec, 2011

- Adaptation is allowed for amplitude as well as phase
  - Projection of the signal onto non-stationary basis functions $\alpha_k(t)e^{j\phi_k(t)}$ [6]!
- Model:

$$s_d(t) = \sum_{k=-K}^{K} (a_k + tb_k)\left(\alpha_k(t)e^{j\phi_k(t)}\right)w(t)$$

- complex amplitudes $a_k, b_k$: estimated via Least Squares
  - $$\begin{bmatrix}\hat{a} \\ \hat{b}\end{bmatrix} = (E^H W^H W E)^{-1} E^H W^H W s \qquad (11)$$
  - where $E = [E_0|E_1]$:

# Extended Adaptive Quasi-Harmonic model - Kafentzis, Pantazis, Stylianou, Rosec, 2011

- Adaptation is allowed for amplitude as well as phase
  - Projection of the signal onto non-stationary basis functions $\alpha_k(t)e^{j\phi_k(t)}$ [6]!
- Model:

$$s_d(t) = \sum_{k=-K}^{K} (a_k + tb_k)\left(\alpha_k(t)e^{j\phi_k(t)}\right)w(t)$$

- complex amplitudes $a_k, b_k$: estimated via Least Squares
  - $$\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = (E^H W^H W E)^{-1} E^H W^H W s \qquad (11)$$
  - where $E = [E_0|E_1]$:
  - $(E_0)_{n,k} = \alpha_k(t_n)e^{j\phi_k(t_n)}$

# Extended Adaptive Quasi-Harmonic model - Kafentzis, Pantazis, Stylianou, Rosec, 2011

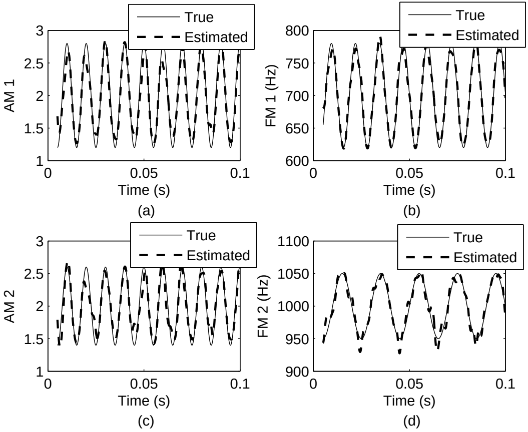
- Adaptation is allowed for amplitude as well as phase
  - Projection of the signal onto non-stationary basis functions $\alpha_k(t)e^{j\phi_k(t)}$ [6]!
- Model:

$$s_d(t) = \sum_{k=-K}^{K} (a_k + tb_k)\left(\alpha_k(t)e^{j\phi_k(t)}\right)w(t)$$

- complex amplitudes $a_k, b_k$: estimated via Least Squares
  - $$\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = (E^H W^H W E)^{-1} E^H W^H W s \quad (11)$$
  - where $E = [E_0|E_1]$:
  - $(E_0)_{n,k} = \alpha_k(t_n)e^{j\phi_k(t_n)}$
  - $(E_1)_{n,k} = t_n\alpha_k(t_n)e^{j\phi_k(t_n)} = t_n(E_0)_{n,k}$

# Extended Adaptive Quasi-Harmonic model

Synthetic signal:



# Extended Adaptive Quasi-Harmonic model

Synthetic signal:

- Robustness in noise is demonstrated:

# Extended Adaptive Quasi-Harmonic model

Synthetic signal:

- Robustness in noise is demonstrated:
-
$$MAE\{\hat{\theta}\} = \frac{1}{M}\sum_{i=1}^{M} |\hat{\theta}^{(i)} - \theta|$$

  where $\theta^{(i)}$ is the estimated parameter at the $i^{th}$ simulation, and $M$ is the number of Monte Carlo simulations.

# Extended Adaptive Quasi-Harmonic model

Synthetic signal:

- Robustness in noise is demonstrated:
- $$MAE\{\hat{\theta}\} = \frac{1}{M}\sum_{i=1}^{M} |\hat{\theta}^{(i)} - \theta|$$

  where $\theta^{(i)}$ is the estimated parameter at the $i^{th}$ simulation, and $M$ is the number of Monte Carlo simulations.

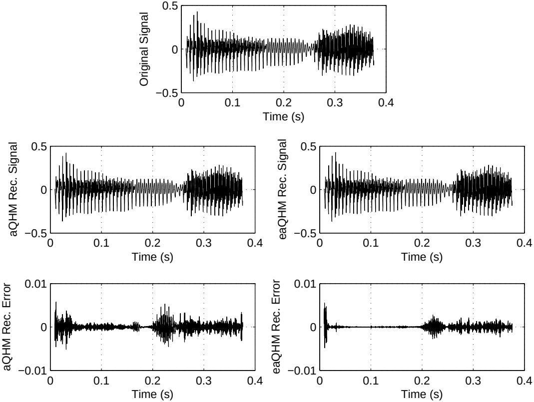
- 

| MAE scores and SRER | | | | | | |
|---|---|---|---|---|---|---|
| SNR | Model | $a_1(t)$ | $a_2(t)$ | $F_1(t)$ | $F_2(t)$ | SRER(dB) |
| $\infty$ | aQHM | 0.2380 | 0.1842 | 7.6105 | 9.1731 | 22.6 |
| | eaQHM | **0.0889** | **0.0949** | **5.9217** | **7.0505** | **42.0** |
| 10 dB | aQHM | 0.2317 | 0.1860 | 8.6071 | 9.0302 | 10.7 |
| | eaQHM | **0.1490** | **0.1476** | **8.0513** | **8.1022** | **10.9** |

Table: MAE scores and SRER for aQHM and eaQHM for $10^4$ Monte Carlo simulations.

# Extended Adaptive Quasi-Harmonic model

Real Signal:



# Adaptation algorithm for eaQHM

- Let $\hat{A}_k(t_l), \hat{f}_k(t_l), \hat{\phi}_k(t_l)$ denote the inst. amplitude, frequency, and phase at time instant $t_l$ of the $k^{th}$ component, with $l = 1, \cdots, L$, where $L$ is the number of frames:
- Initialization: (QHM)
  1) $\hat{f}_k^0(t_l) = \hat{f}_k^0(t_{l-1}) + \rho_{2,k}^0/2\pi$
  2) $\hat{A}_k^0(t_l) = |a_k^0|, \ \hat{\phi}_k^0(t_l) = \angle a_k^0$
  3) $\hat{f}_k^0(t_{l+1}) = \hat{f}_k^0(t_l)$
- FOR adaptation $i = 1, 2, \cdots$
  FOR frame $l = 1, 2, \cdots, L$
  1. Compute $a_k^l, b_k^l$ using $\hat{\phi}_k^{i-1}(t)$ and (11)
  2. Update $\hat{f}_k^i(t_l)$ using **(9)**
  3. Compute $\hat{A}_k^i(t_l)$ and $\hat{\phi}_k^i(t_l)$

  END
  4. Interpolation of the parameters $\{\hat{A}_k^i(t), \hat{f}_k^i(t), \hat{\phi}_k^i(t)\}$

  END

# Extended Adaptive Quasi-Harmonic model

Analysis-Synthesis System

- Seperate speech into two parts: deterministic and stochastic

# Extended Adaptive Quasi-Harmonic model

Analysis-Synthesis System

- Seperate speech into two parts: deterministic and stochastic
- Deterministic part: a sum of non-stationary sinusoids (ea/aQHM)

# Extended Adaptive Quasi-Harmonic model

Analysis-Synthesis System

- Seperate speech into two parts: deterministic and stochastic
- Deterministic part: a sum of non-stationary sinusoids (ea/aQHM)
- Stochastic part: time and frequency modulated (energy-based envelope and AR modeling)

# Extended Adaptive Quasi-Harmonic model

Analysis-Synthesis System

- Seperate speech into two parts: deterministic and stochastic
- Deterministic part: a sum of non-stationary sinusoids (ea/aQHM)
- Stochastic part: time and frequency modulated (energy-based envelope and AR modeling)
- Very high quality of speech signal reconstruction

# Outline

# References I


R. J. McAulay and T. F. Quatieri,
Speech Analysis/Synthesis based on a Sinusoidal Representation
IEEE Trans. on Acoust., Speech and Sig. Proc., vol. 34, pp. 744754, 1986.

X. Serra,
A system for sound analysis/transformation/synthsis based on a determistic plus stochastic decomposition
Ph.D. dissertation, Stanford University, 1989.

Y. Stylianou,
Harmonic plus noise models for speech, combined with statistical methods, for speech and speaker modification
Ph.D. dissertation, E.N.S.T - Paris, 1996.

M. Macon,
Speech synthesis based on sinusoidal modeling
Ph.D. dissertation, Georgia Institute of Technology, 1996.

Y. Pantazis, O. Rosec, and Y. Stylianou,
Adaptive AMFM signal decomposition with application to speech analysis,
IEEE Trans. on Audio, Speech, and Lang. Proc., vol. 19, pp. 290300, 2011.

G. P. Kafentzis, Y. Pantazis, O. Rosec, and Y. Stylianou,
An Extension of the Adaptive Quasi-Harmonic Model
in Proc. IEEE ICASSP, Kyoto, March 2012.

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

# Time for Questions!

Thank you for your attention!
Any questions? :-)

# Τέλος Ενότητας

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

# Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Κρήτης**» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

# Σημειώματα

.

# Σημείωμα Αναφοράς

Copyright Πανεπιστήμιο Κρήτης, Στυλιανού Ιωάννης. «Ψηφιακή Επεξεργασία Φωνής. Προσαρμόσιμα Ημιτονοειδή Μοντέλα». Έκδοση: 1.0. Ηράκλειο/Ρέθυμνο 2015. Διαθέσιμο από τη δικτυακή διεύθυνση: http://www.csd.uoc.gr/~hy578