Estimation of frequency for AM/FM models using the phase vocoder framework : experimental and theoretical complements

Estimation de fréquence pour les modèles AM/FM dans le cadre du vocodeur de phase : compléments expérimentaux et théoriques

Michaël Betser
Patrice Collen
Bertrand David
Gaël Richard

2007D010

2007

Département Traitement du Signal et des Images
Groupe Audio, Acoustique et Ondes

Estimation de fréquence pour les modèles AM/FM dans le cadre du vocodeur de phase: compléments expérimentaux et théoriques

Michaël Betser, Patrice Collen, Bertrand David and Gaël Richard

Abstract

Ce document technique est un complément expérimental et théorique à l'article intitulé "Estimation of frequency for AM/FM models using the phase vocoder framework" (IEEE Trans. on Signal Processing, 2007). Plusieurs propriétés concernant certains estimateurs de fréquences basés sur la transformée de Fourier sont démontrés. La première démonstration concerne le réassignement spectral, et prouve que la méthode est toujours valide pour le modèle AM/FM d'ordre 1. Les autres démonstrations concernent les estimateurs de fréquences basés sur le vocodeur de phase, dans le cas AM/FM, et développés dans l'article cité ci-dessus. Il s'agit principalement des développements des biais et des variances théoriques. Des compléments expérimentaux sont également présentés et apportent un éclairage sur l'influence de certains paramètres (notamment le zéro padding).

Estimation of frequency for AM/FM models using the phase vocoder framework: Experimental and theoretical complements

Abstract

This technical report is a theoretical and experimental complement to the IEEE Transactions on Signal Processing paper [5]. Several demonstrations concerning Fourier-based estimators are presented. The first demonstration concerns the reassignment, and shows that this method is still valid for the AM/FM case. The other demonstrations concern the phase-vocoder-based frequency estimation, in the AM/FM case.

Contents

1 Introduction

This technical report is a theoretical and experimental complement to the IEEE Transactions on Signal Processing paper [5].

The model under study in this report is the first-order AM/FM model, defined as:

$$x(\tau) \triangleq \mathrm{e}^{\lambda+\mu\tau} \cdot \mathrm{e}^{j(\alpha+\beta\tau+\gamma\tau^2/2)} \tag{1.1}$$

where α is the phase, β is the frequency, γ is the Frequency Change Rate (FCR), λ is the log-amplitude and μ is the Log-Amplitude Change rate (ACR). τ is the local time. The time of the n^{th} sample is $\tau_n = n/F$ where F is the sampling frequency.

2 Reassignment and the AM/FM model

The frequency reassignment is known to perfectly localize chirp signals [2]. A simple demonstration for the continuous Fourier Transform in the FM case is presented in [3]. Using the same method, it can be shown easily that the time-frequency Reassignment is also perfectly valid for the AM/FM model.

Let's define the continuous Fourier Transform as:

$$FT(x;\omega) \triangleq \int_{-\infty}^{+\infty} x(\tau)\,\mathrm{e}^{-j\omega\tau}\,d\tau \tag{2.1}$$

Let g be $g(\tau) \triangleq x(\tau)h(\tau)$, where h is the Fourier Transform window. Using (1.1), we have:

$$\begin{aligned}\frac{dg}{d\tau}(\tau) =& \frac{dx}{d\tau}(\tau)h(\tau) + \frac{dh}{d\tau}(\tau)x(\tau)\\ =& (j(\gamma\tau+\beta)+\mu)x(\tau)h(\tau) + \frac{dh}{d\tau}(\tau)x(\tau)\end{aligned} \tag{2.2}$$

The FT is applied to the relation (2.2) for a frequency ω such that the sinusoid x has a non-zero energy for this particular frequency ($FT(g;\omega) \neq 0$), on the definition interval of the window h:

$$\begin{aligned} j\omega FT(g;\omega) &= j\gamma FT(\tau g;\omega) + (j\beta+\mu)FT(g;\omega)\\ &\quad + FT(\frac{dh}{d\tau}x;\omega)\\ \Leftrightarrow \beta - j\mu + \gamma\frac{FT(\tau g;\omega)}{FT(g;\omega)} &= \omega + j\frac{FT(\frac{dh}{d\tau}x;\omega)}{FT(g;\omega)} \end{aligned} \tag{2.3}$$

$$\Rightarrow \beta + \gamma\Re\Big(\frac{FT(\tau g;\omega)}{FT(g;\omega)}\Big) = \omega - \Im\Big(\frac{FT(\frac{dh}{d\tau}x;\omega)}{FT(g;\omega)}\Big) \tag{2.4}$$

This formula is exactly the same as the one obtained in [3] for the FM model. The influence of the ACR has been removed by taking the real part of equation (2.3). Accordingly to the usual formulation of the reassignment, let's define $\dot{h}(\tau) \triangleq \frac{dh}{d\tau}(\tau)$. The first term of equation (2.4) is the frequency of the partial for the time

$$\hat{t} = t + \Re(\frac{FT(thx;\omega)}{FT(hx;\omega)}) \tag{2.5}$$

which is the time reassignment operator. The second part of the equation corresponds to the frequency reassignment operator

$$\hat{\beta} = \omega - \Im(\frac{FT(\dot{h}x;\omega)}{FT(hx;\omega)}) \tag{2.6}$$

3 The phase vocoder frequency estimator and the AM/FM model

In this part the model is supposed valid on an interval W, and λ, α, β corresponds to the log-amplitude, phase and frequency for the time t_M, *i.e.* the time corresponding to the middle of the window W).

$$x(\tau) = \mathrm{e}^{\lambda_M + j\alpha_M}\,\mathrm{e}^{\mu\tau + j(\beta_M\tau+\gamma\tau^2/2)} \tag{3.1}$$

Let's define Γ as:

$$\Gamma(\beta, \mu, \gamma; h) = \sum_{i=-(N-1)/2}^{(N-1)/2} h(\tau_i)\, \mathrm{e}^{\mu\tau_i}\, \mathrm{e}^{j(\beta\tau_i + \gamma\frac{\tau_i^2}{2})} \tag{3.2}$$

The STFT of x for the time t_{m_i} and the frequency ω_{k_i} is:

$$X(t_{m_i}, \omega_{k_i}; h) = \mathrm{e}^{\lambda_i + j\alpha_i}\, \Gamma(\beta_i - \omega_{k_i}, \mu, \gamma; h) \tag{3.3}$$

$$X(t_{m_i}, \omega_{k_i}; h) = \mathrm{e}^{\lambda_i + j\alpha_i}\, \Gamma_i \tag{3.4}$$

where $\alpha_i = \alpha_M + \beta_M(t_{m_i} - t_M) + \gamma(t_{m_i} - t_M)^2/2$, $\beta_i = \beta_M + \gamma(t_{m_i} - t_M)$, ω_{k_i} is the frequency of the closest maximal bin k_i to β_i for the time t_{m_i}. Finally, h is the window.

If we take the argument of this last equation:

$$\psi_i \triangleq \arg(X_i) = \alpha_i + \arg(\Gamma_i) \tag{3.5}$$

If we consider the phase difference between two time-frequency points, X_1 and X_2, such that $t_{k_2} - t_M = t_M - t_{k_1} = T/2$:

$$\Delta\psi = \arg(X_2) - \arg(X_1) \tag{3.6}$$

$$= T\beta_M + \arg(\Gamma_2\bar{\Gamma}_1) + 2\pi n \tag{3.7}$$

The first two subsections will recall how work the two methods presented in the article [5], namely the phase corrected vocoder and the reassigned vocoder. Then, in the next three subsections, the property of these estimators, concerning the unwrapping factor n, their biases and their variances, will be derived.

3.1 Phase corrected vocoder (PCV)

This method is derived for the FM model ($\mu = 0$). The Fourier transform is not a direct estimator of the phase for chirp signals. An improvement to the phase vocoder consists in correcting the Fourier phase estimation, as in [4], using the error function $\Gamma(\Delta\beta, 0, \gamma; h)$. The estimation scheme proposed involves two steps:

1. Estimation of the corrected phases (modulo 2π) $\hat{\alpha}_1$ and $\hat{\alpha}_2$, and the unwrapping factor $\hat{n}$.

2. Estimation of β_M using the phase vocoder formula[1]

$$\hat{\beta}_M = \frac{\mathrm{mod}(\hat{\alpha}_2) - \mathrm{mod}(\hat{\alpha}_1) + 2\pi\hat{n}}{T} \tag{3.8}$$

The function Γ requires the knowledge of the frequencies corresponding to t_{m_1} and t_{m_2} (namely β_1 and β_2). Therefore the first step of the estimation scheme will involve a first frequency estimation for β_1 and β_2. As there is no knowledge about the FCR in this step, it is proposed to use one of the frequency estimators based on the classical sinusoidal model. Although these estimators are biased for the FM model, it is shown in [4] that this scheme can greatly improve the precision on the phase estimates.

The parameter γ, and the unwrapping factor n can be deduced from the frequencies β_1 and β_2, using the formulas:

$$\hat{\gamma} = \frac{\hat{\beta}_2 - \hat{\beta}_1}{T}$$

$$\hat{n} = \mathrm{round}\left(\frac{1}{2\pi}\left(\mathrm{mod}(\hat{\alpha}_1) - \mathrm{mod}(\hat{\alpha}_2) + \frac{\hat{\beta}_1 + \hat{\beta}_2}{2}T\right)\right)$$

[1] mod() is the modulo 2π function

3.2 Reassigned vocoder (RV)

For this method, an accurate approximation of $\arg(\Gamma_2\bar{\Gamma}_1)$ will be derived. The first step is to express $\arg(\Gamma_1\bar{\Gamma}_2)$ as a function of β_M, the frequency to be estimated. In this process, $\Delta\beta_1$ and $\Delta\beta_2$ can be decomposed into two bounded terms B and G, described below. Let's define ω_M as the mean bin frequency and $\Delta\omega$ as half the frequency variation in bins:

$$\omega_M = \frac{\omega_{k_1} + \omega_{k_2}}{2}, \qquad \Delta\omega = \frac{\omega_{k_2} - \omega_{k_1}}{2} \tag{3.9}$$

In addition, from the definition of the quadratic phase model (3.1), the FCR γ follows this relation:

$$\gamma\frac{T}{2} = \frac{\beta_2 - \beta_1}{2} \tag{3.10}$$

Let $B = \beta_M - \omega_M$ and $G = \Delta\omega - \gamma\frac{T}{2}$. From the previous definitions, $\Delta\beta_i$ can be expressed as:

$$\Delta\beta_1 = B + G, \qquad \Delta\beta_2 = B - G \tag{3.11}$$

And since ω_{k_1} and ω_{k_2} are respectively the closest maximal bins to β_1 and β_2, then $|G| < R$, where R is half the FT precision. In the pure FM case the relation $|B| < R$ is also verified. In the AM case the maximal bin ω_i are not anymore the closest bin to the β_i, and B will be bounded by a constant C, depending on ACR, and such that $C > R$. The theoretical value of C is difficult to obtain, but numerical analysis shows that C increases very slowly when the ACR increases.

The first-order Taylor expansion in $G = 0$ of Γ_1 and Γ_2 is given by:

$$\begin{aligned}\Gamma_1 &= \Gamma(B,\mu,\gamma;h) + G\Gamma'(B,\mu,\gamma;h) + \epsilon_1\\ \Gamma_2 &= \Gamma(B,\mu,\gamma;h) - G\Gamma'(B,\mu,\gamma;h) + \epsilon_2\end{aligned}$$

Where ϵ_1 and ϵ_2 are the Lagrange remainders. The frequency derivation property of the STFT leads to:

$$\begin{aligned}\Gamma_1 &= \Gamma(B,\mu,\gamma;h) + jG\Gamma(B,\mu,\gamma;th) + \epsilon_1\\ \Gamma_2 &= \Gamma(B,\mu,\gamma;h) - jG\Gamma(B,\mu,\gamma;th) + \epsilon_2\end{aligned}$$

For an order 1 Taylor expansion in 0 of the argument function, we obtain:

$$\arg(\Gamma_1\bar{\Gamma}_2) = 2G\Re(\frac{\Gamma(B,\mu,\gamma;th)}{\Gamma(B,\mu,\gamma;h)}) + \epsilon \tag{3.12}$$

This approximation has proven to be quite accurate for the intervals of parameter considered. Indeed the deterministic bias has a magnitude of 10^{-3}Hz in average for $\mu \in [0,100]$ and $\gamma \in [0,8000]$ (cf. section 3.4).

$\Re(\frac{\Gamma(B,\mu,\gamma;th)}{\Gamma(B,\mu,\gamma;h)})$ is in fact equivalent to the discrete version of the reassigned time. Indeed, the STFT can be rewritten as a function of Γ:

$$\begin{aligned}X(t_M,\omega_M;th) &= \mathrm{e}^{\lambda_M + j\alpha_M}\,\Gamma(B,\mu,\gamma;th)\\ X(t_M,\omega_M;h) &= \mathrm{e}^{\lambda_M + j\alpha_M}\,\Gamma(B,\mu,\gamma;h)\end{aligned}$$

where $th(\tau) \triangleq \tau h(\tau)$ and $\alpha_M = \alpha + \beta\tau_M + \frac{\gamma}{2}\tau_M^2$. We can therefore conclude that:

$$\Re(\frac{\Gamma(B,\mu,\gamma;th)}{\Gamma(B,\mu,\gamma;h)}) = \Re(\frac{X(t_M,\omega_M;th)}{X(t_M,\omega_M;h)})$$

and

$$\arg(\Gamma_1\bar{\Gamma}_2) = 2G\Re(\frac{X(t_M,\omega_M;th)}{X(t_M,\omega_M;h)}) + \epsilon$$

Using the previous expression in equation (3.7):

$$T\beta_M = \Delta\psi + 2\pi n + 2G\Re(\frac{X(t_M,\omega_M;th)}{X(t_M,\omega_M;h)}) + \epsilon$$

Replacing G by its definition leads to:

$$\begin{aligned}\beta_M + \gamma\Re(\frac{X(t_M,\omega_M;th)}{X(t_M,\omega_M;h)}) &= \frac{\Delta\psi + 2\pi n}{T} \\ &\quad + 2\frac{\Delta\omega}{T}\Re(\frac{X(t_M,\omega_M;th)}{X(t_M,\omega_M;h)}) + \frac{\epsilon}{T}\end{aligned} \tag{3.13}$$

The left part of this expression is the frequency for the time: $\hat{t} = t_M + \Re(\frac{X(t_M,\omega_M;th)}{X(t_M,\omega_M;h)})$. The right part is the vocoder estimator corrected by a term depending on the reassigned time and on $\frac{2\Delta\omega}{T}$, which can be interpreted as a first FCR estimate using frequency bins.

3.3 Study of the unwrapping factor

The last problem to solve is the computation of the unwrapping factor n. It will be achieved using the estimator

$$\hat{n} = \text{round}\big((\Omega T - \Delta\psi)/(2\pi)\big) \tag{3.14}$$

where Ω is a chosen reference frequency $\Omega = \omega_M$ [8]. This choice imposes a theoretical limit on the hop-size length of the phase vocoder, which is now discussed. From (3.7), n verifies this relation:

$$n = \frac{1}{2\pi}\left[\omega_M T - \Delta\psi + \Delta\beta_M T - \arg(\Gamma_1\bar{\Gamma}_2)\right] \tag{3.15}$$

where $\Delta\beta_M = \beta_M - \omega_M$. The chosen estimator of n is:

$$\hat{n} = \text{round}\left(\frac{\omega_M T - \Delta\psi}{2\pi}\right) \tag{3.16}$$

The condition for identity between n and $\hat{n}$ is:

$$n = \hat{n} \Leftrightarrow |T\Delta\beta_M - \arg(\bar{\Gamma}_1\Gamma_2)| \leq \pi \tag{3.17}$$

But from the triangular inequality, we have:

$$|T\Delta\beta_M - \arg(\bar{\Gamma}_1\Gamma_2)| \leq T\Delta_m + \Gamma_m(\Delta_m, \mu_m, \gamma_m; h) \tag{3.18}$$

where Δ_m is the largest difference between β_M and the maximal bin ω_M. Γ_m is the maximal value of the corrective term for the system parameters considered:

$$\Gamma_m(\Delta_m, \mu_m, \gamma_m; h) = \max_{|\Delta\beta_i|\leq\Delta_m, |\mu|\leq\mu_m, |\gamma|\leq\gamma_m} |\arg(\Gamma_1\bar{\Gamma}_2)| \tag{3.19}$$

Therefore, a sufficient condition for (3.17) to be verified is:

$$T\Delta_m \leq \pi - \Gamma_m(\Delta_m, \mu_m, \gamma_m; h) \tag{3.20}$$

As $T = H/F$ (H is the hop-size in samples), we finally get:

$$H \leq \frac{\pi F}{\Delta_m}\left(1 - \frac{\Gamma_m(\Delta_m, \mu_m, \gamma_m; h)}{\pi}\right) \tag{3.21}$$

In the classical and AM cases ($\gamma_m = 0$), $\arg(\Gamma_1\bar{\Gamma}_2) = 0$ and we find the classical unwrapping condition $H \leq N$.

In the FM case, Δ_m is equal to R, half the Fourier precision:

$$H \leq N\left(1 - \frac{\Gamma_m(R, \mu_m, \gamma_m; h)}{\pi}\right) \tag{3.22}$$

Table 1 presents a numerical evaluation of the maximal hop-size values for various system parameters. It can be seen that the maximal theoretical hop-sizes decrease very slowly when the FCR increases. For usual applications, which use much lower hop-sizes than this limit, this means that the FCR will have no

Table 1: Maximal hop-size values in samples for the RV method ($N = 512$, $F = 16000$)

	Hann	Hamming	Blackman	Gaussian
$\gamma_m = 1000$	508	505	509	506
$\gamma_m = 8000$	495	492	500	494

Table 2: Evaluation of ϵ_{RV} for different values of μ_m, and γ_m: maximal bias and absolute mean in Hz.

	Hann	Hamming	Blackman	Gaussian
$\mu_m = 10, \gamma_m = 2\pi 1000$	2E-2;1.6E-3	3E-2;2.6E-3	6.6E-3;5.3E-4	1.9E-2;1.8E-3
$\mu_m = 100, \gamma_m = 2\pi 8000$	0.39;2.8E-3	0.49;3.2E-2	0.19;9.5E-3	0.37;2.3E-2

impact on the unwrapping estimation. The rectangular window cannot be used with this method (the time reassignment requires smooth functions) and is therefore not present in the table.

When the amplitude varies, the energy attributed to each frequency will be shifted depending on the ACR. The maximum of energy will no longer correspond to β_M, the sinusoid frequency for the middle of the window. Therefore, Δ_m will be superior to R, and will increase as the ACR increases. In the AM/FM rate model, the maximal theoretical hop-sizes for the unwrapping estimation are more difficult to compute in this case, as Δ_m is not explicitly known. They should be lower than the values presented in table 1. Nevertheless, this problem can be minimized by using intermediate phases.

3.4 Study of the bias

All experiments are done for a window size of 32ms and a hop-size of 8ms.

For the RV method

For high FCR, a bias will appear, caused by the approximation (3.12). Although the bias does not have a simple expression, it can be easily evaluated numerically as

$$\epsilon_{RV} = \frac{1}{T} \left| (\arg(\Gamma_1 \bar{\Gamma}_2) - 2G\Re(\frac{\Gamma(B, \mu, \gamma; th)}{\Gamma(B, \mu, \gamma; h)}) \right| \tag{3.23}$$

Table 2 shows values of this bias for different windows and two different ACR and FCR intervals. In each case, the first figure corresponds to the maximal bias, and the second figure is the mean value. The RV method is applied to the maximal bins of the Fourier Transform and is used without padding. The mean value is an average of 10000 experiments. It can be seen that the biases are kept within 1 Hz even for strong AM/FM modulations. If $\gamma_m = 0$, all the biases disappear. The bias only slightly increases when μ_m increases and is lower when the window is more concentrated in time, as for the Blackman window. The bias will be reduced even more if a padding factor is used (cf. Figure 1).

For the PCV method

As for the RV method, the bias in the PCV does not have a simple mathematical expression and is evaluated numerically:

$$\epsilon_{PCV} = \frac{1}{T} \left| (\arg(\Gamma_1 \bar{\Gamma}_2) - \arg(\hat{\Gamma}_1 \bar{\hat{\Gamma}}_2)) \right| \tag{3.24}$$

In the first step, the frequency estimation method chosen is the interpolator described in [6]. The method is applied to maximum bins, without padding, and the results are based on an average of 10000 experiments. In the PCV case, Table 3 shows that the bias is within 1Hz for high FCR, if the ACR stays relatively small. For high ACRs, the method no longer works, as the PCV does not take into account this parameter. The use of a padding factor greatly decreases the biases (cf. Figure 2).

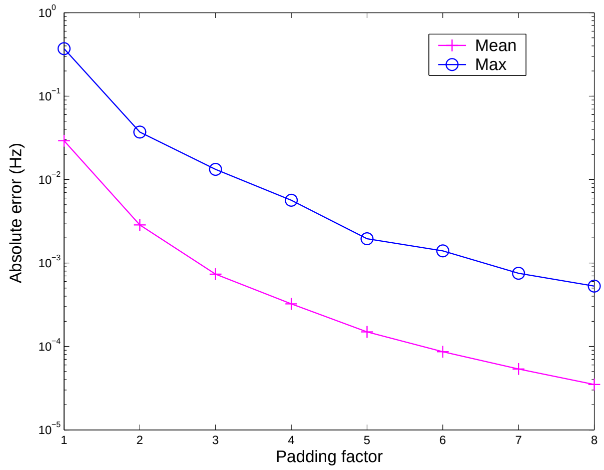


Figure 1: Influence of the padding factor on the RV method ($\mu_m = 100$,$\gamma_m = 2\pi 8000$, Hann window)

Table 3: Evaluation of ϵ_{PCV} and for different values of μ_m, and γ_m: maximal bias and absolute mean in Hz.

	Hann	Hamming	Blackman	Gaussian
$\mu_m = 0, \gamma_m = 2\pi 8000$	0.5;5.6E-2	0.65;9.7E-2	0.27;2.3E-2	0.56;8.6E-2
$\mu_m = 10, \gamma_m = 2\pi 1000$	0.37;8.6E-2	0.47;0.1	0.26;6.5E-2	0.43;9.6E-2
$\mu_m = 10, \gamma_m = 2\pi 8000$	1.1;0.14	1.2;0.19	0.64;9.6E-2	1.0;1.7E-2

3.5 Study of the variance

The influence of a white noise on the classical phase vocoder is studied in [9]. A more recent reference [1] uses the same method, but presents a simpler formula applicable to any window. The results presented in [9, 1] are generalized here for the AM/FM model.

It is well known that when the frequency is constant, the Fourier Transform asymptotically resolves the sinusoid. In the FM case, the Fourier Transform will no longer resolve the chirp when $N \to \infty$[2]. Instead of an asymptotic property of the estimator, we will suppose that the sinusoid is well resolved. This has two consequences. First the energy of the noise is negligible compared to the energy of the sinusoids within the chosen bins, $X_i \gg N_i$. Secondly, for a chirp to be resolved, the width of the main lobe window must be superior to the maximum frequency variation of the chirp inside the window. If not, multiple peaks will appear on the main lobe. For example, if γ_m is the largest FCR possible for the signal considered, and if the main lobe of the window frequency response has a width of K bins (independently of the window size N), this condition can be expressed as: $\gamma_m \frac{N}{F} < 2\pi \frac{KF}{N} \Leftrightarrow \frac{N}{F} < \sqrt{\frac{2\pi K}{\gamma_m}}$. For strong chirps ($\gamma_m \gg 1\pi K$), this condition can be expressed $\tau_N = N/F \ll 1s$.

If $S_i = S(t_{m_i}, \omega_{k_i}; h)$ and $N_i = N(t_{m_i}, \omega_{k_i}; h)$ are the Fourier Transform of s and n respectively, then:

$$S_i = X_i + N_i$$
$$S_i = X_i(1 + N_i^{'})$$

where $N_i^{'} \triangleq N_i/X_i$. The conjugate product $S_2\bar{S}_1$ can be written as:

$$S_2\bar{S}_1 = X_2\bar{X}_1(1 + Z) \tag{3.25}$$

where $Z \triangleq 1 + N_2^{'} + \bar{N}_1^{'} + N_2^{'}\bar{N}_1^{'}$.

[2]Indeed,for a pure FM signal, when N tends to infinity, the chirp covers all the frequency range, with equal energy, and will no longer be resolved by the Fourier Transform.

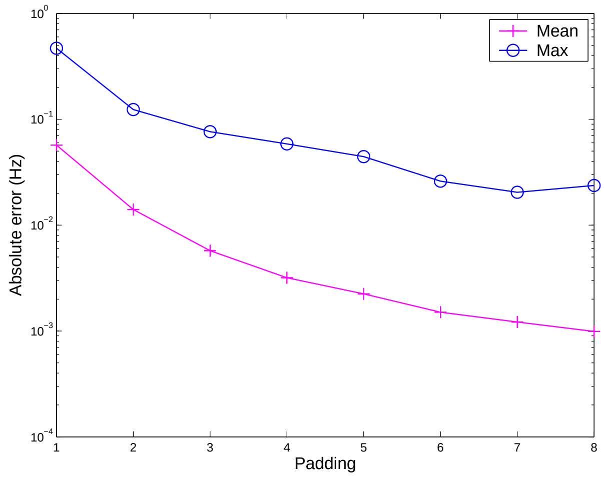


Figure 2: Influence of the padding factor on the PCV method ($\mu = 0$,$\gamma \in [0, 8000]$, Hann window)

As it is assumed that the STFT resolves the sinusoid x from the disturbance n, $X_i \gg N_i$ for the bins close to the maximum, and $\arg(1 + Z) \approx \Im(Z) \approx \Im(N_2^{'}) - \Im(N_1^{'})$. $\arg(S_2\bar{S}_1)$ can be written as

$$\arg(S_2\bar{S}_1) \approx \arg(X_2\bar{X}_1) + \Im(Z)$$

From equation (3.7), this relation becomes:

$$\arg(S_2\bar{S}_1) \approx T\beta_M + \arg(\Gamma_2\bar{\Gamma}_1) + 2\pi n + \Im(Z)$$

And the expression of β_M is:

$$\beta_M \approx \frac{\arg(S_2\bar{S}_1) - \arg(\Gamma_2\bar{\Gamma}_1) + 2\pi n - \Im(Z)}{T}$$

Let's note $\epsilon_{N,voc} = \Im(Z)$. The PCV and RV methods both use an estimate of $\arg(\Gamma_2\bar{\Gamma}_1)$ and n to compute the frequency. As the sinusoid are supposed resolved from the noise, there will be no error in the estimation of n. It will now be proved that the stochastic error resulting from the estimation of $\arg(\Gamma_2\bar{\Gamma}_1)$ is negligible compared to $\epsilon_{N,voc}$.

PCV case

In the PCV case, Γ_i will be replaced by an estimate $\hat{\Gamma}_i = \Gamma(\hat{\beta}_i - \omega_i, 0, \hat{\gamma}; h)$, where $\hat{\beta}_i$ and $\hat{\gamma}$ are estimates computed using other Fourier-based estimators. The PCV estimation scheme has been derived for the FM model, *i.e.* $\mu = 0$. As there is no knowledge on γ in a first step, the frequency estimator used will be based on the classical model, and will be biased when the slope is present. It will now be proved that the influence of the stochastic error from this first step is negligible.

Let $\epsilon^{\beta}_{D_i}$ and $\epsilon^{\beta}_{N_i}$ be respectively the deterministic and the stochastic error of the first step estimator for the frequency β_i. It is supposed that this estimator verifies the following assumptions:

1. $\hat{\beta}_i = \beta_i + \epsilon^{\beta}_{D_i} + \epsilon^{\beta}_{N_i}$
2. $\tau_N \epsilon^{\beta}_{N_i} \ll 1$
3. $E(\epsilon^{\beta}_{N_i}) = 0$
4. $\mathrm{var}(\epsilon^{\beta}_{N_i}) \leq \mathrm{var}(\Im(N_i^{'}))$

In the classical case, many Fourier-based estimators verify these assumptions asymptotically, in particular the discrete Fourier spectrum interpolators using phase, such as the methods described in [6, 10, 7]. If we suppose that the sinusoids are well resolved within the bins used, assumptions 1-4 will remain true.

The FCR estimate is defined as:

$$\hat{\gamma} = \frac{\hat{\beta}_2 - \hat{\beta}_1}{T} \tag{3.26}$$

Therefore, from the first assumption, the stochastic error of $\hat{\gamma}$ will also verify:

$$\hat{\gamma} = \gamma + \epsilon_D^{\gamma} + \epsilon_N^{\gamma} \tag{3.27}$$

$$\epsilon_N^{\gamma} \triangleq \frac{\epsilon_{N_2}^{\beta} - \epsilon_{N_1}^{\beta}}{T} \tag{3.28}$$

Let's define:

$$\Gamma_{k,i} = \Gamma(\beta_i + \epsilon_{D_i}^{\beta}\omega_i, 0, \gamma + \epsilon_D^{\gamma}; h\tau^k) \tag{3.29}$$

From the second assumption, the following approximation will hold:

$$\begin{aligned}
\arg(\hat{\Gamma}_i) &\approx \arg(\Gamma_{0,i}) + \arg(1 + j(\epsilon_N^{\beta}\frac{\Gamma_{1,i}}{\Gamma_{0,i}} + \epsilon_N^{\gamma}\frac{\Gamma_{2,i}}{2\Gamma_{0,i}})) && (3.30)\\
&\approx \arg(\Gamma_i) + \arg(\frac{\Gamma_{0,i}}{\Gamma_i}) + \epsilon_{N_i}^{\beta}C_{1,i} + \epsilon_N^{\gamma}C_{2,i} && (3.31)\\
&\approx \arg(\Gamma_i) + \epsilon_{D_i} + \epsilon_{N_i} && (3.32)
\end{aligned}$$

where,

$$C_{1,i} = \Im\left(\frac{\Gamma_{1,i}}{\Gamma_{0,i}}\right), \qquad C_{2,i} = \Im\left(\frac{\Gamma_{2,i}}{2\Gamma_{0,i}}\right) \tag{3.33}$$

Using equation (3.28), ϵ_{N_1} can be written as:

$$\epsilon_{N_1} = (C_{1,i} - \frac{C_{2,i}}{T})\epsilon_{N_1}^{\beta} + \frac{C_{2,i}}{T}\epsilon_{N_2}^{\beta} \tag{3.34}$$

As explained earlier, for a system aimed at analyzing chirps with high FCR, the relation $\tau_N \ll 1$ holds. From this relation, it can be proved that $C_{1,i} \ll 1$ and $C_{2,i}/T \ll 1$. From assumption 4, we can conclude that $\mathrm{var}(\epsilon_{N_1}^{\beta}) \ll \mathrm{var}(\Im(N_1^{'}))$. Similarly, it can be shown that $\mathrm{var}(\epsilon_{N_2}^{\beta}) \ll \mathrm{var}(\Im(N_1^{'}))$. Combining these results, we therefore have $\mathrm{var}(\epsilon_{N_1}^{\beta} + \epsilon_{N_2}^{\beta}) \ll \mathrm{var}(\epsilon_{N,voc})$.

In summary, if the frequency estimator of the first step verifies the assumption 1-4 and if the sinusoid is well resolved, the stochastic error due to the first estimates will be negligible, and the noised PCV equation can be written as:

$$\beta_M \approx \frac{1}{T}\arg(S_2\bar{S}_1) - \frac{1}{T}\arg(\hat{\Gamma}_2\bar{\hat{\Gamma}}_1) + 2\pi n + \epsilon_{PCV} - \frac{\Im(Z)}{T} \tag{3.35}$$

where ϵ_{PCV} is the deterministic error of the PCV method.

RV case

In the RV case, $\arg(\Gamma_2\bar{\Gamma}_1)$ is replaced by the approximation (3.12), with a noise perturbation. In keeping with the previous notations, $S_M = S(t_M, \omega_M; h)$ (idem for X_M and N_M) and $S_{M,1} = S(t_M, \omega_M; th)$ (idem for $X_{M,1}$ and $N_{M,1}$).

$$\begin{aligned}
\widehat{\arg(\Gamma_1\bar{\Gamma}_2)} &= 2G\Re(\frac{S_{M,1}}{S_M})\\
&= 2G\Re(\frac{X_{M,1} + N_{M,1}}{X_M + N_M})
\end{aligned}$$

Table 4: Value of Q for usual windows

	Hann	Hamming	Blackman	Gaussian
$Q(h)$	0.02	0.02	0.01	0.02

As in the PCV case, the sinusoid is supposed well resolved from the noise, $X(t_M, \omega_M; h) \gg N(t_M, \omega_M; h)$. The previous equation can be approximated as:

$$\widehat{\arg(\Gamma_1 \bar{\Gamma}_2)} \approx 2G\Re(\frac{X_{M,1}}{X_M}) + \epsilon_{N,RV}$$
$$\epsilon_{N,RV} \triangleq 2G\Re(\frac{N_M}{X_M}\frac{\bar{X}_{M,1}}{\bar{X}_M} + \frac{N_{M,1}}{X_M}) \tag{3.36}$$

As in the PCV case, it will now be proved that $\epsilon_{N,RV}$ is negligible compared to $\epsilon_{N,voc}$.

We know that $\mathrm{var}(N_1) = \mathrm{var}(N_2) = \mathrm{var}(N_M)$. As k_1, k_2 and k_M are maximum bins, if $\mu = 0$ then $|X_1| \approx |X_M| \approx |X_2|$. Recall that $N_i^{'} = N_i/X_i$, therefore if $\mu \neq 0$ we have:

$$\mathrm{e}^{\mu T}\,|X_1|^2 \approx |X_M|^2 \approx \mathrm{e}^{-\mu T}\,|X_2|^2$$
$$\mathrm{e}^{-\mu T}\,\mathrm{var}(N_1^{'}) \approx \mathrm{var}(N_M^{'}) \approx \mathrm{e}^{\mu T}\,\mathrm{var}(N_2^{'})$$
$$\mathrm{var}(N_M^{'}) \leq \max\left(\mathrm{var}(N_1^{'}), \mathrm{var}(N_2^{'})\right)$$

But $\epsilon_{N,voc}$ also verifies:

$$\mathrm{var}(\epsilon_{N,voc}) = \mathrm{var}(\Im(N_1^{'})) + \mathrm{var}(\Im(N_2^{'})) - E(\Im(N_1^{'})\Im(N_2^{'}))$$
$$\mathrm{var}(\epsilon_{N,voc}) = \frac{1}{2}(\mathrm{var}(N_1^{'}) + \mathrm{var}(N_2^{'})) - E(\Im(N_1^{'})\Im(N_2^{'}))$$
$$\mathrm{var}(\epsilon_{N,voc}) \approx \frac{1}{2}\max\left(\mathrm{var}(N_1^{'}), \mathrm{var}(N_2^{'})\right)$$

Therefore the following relation is approximately verified:

$$\mathrm{var}(\epsilon_{N,voc}) \geq \frac{1}{2}\,\mathrm{var}(N_M^{'})$$
$$\mathrm{var}(\epsilon_{N,voc}) \geq \mathrm{var}(\Re(N_M^{'})) \tag{3.37}$$

From the definition of G, we have $|G| \leq R$ and $R = \pi\frac{F}{P_c N} = \frac{\pi}{P_c \tau_N}$, where P_c is the padding factor used. As $\mathrm{var}(N_{M,1}) = \sigma^2 \sum_i h_i^2 \tau_i^2$ and $\mathrm{var}(N_M) = \sigma^2 \sum_i h_i^2$, the variance of $2G\Re(\frac{N_{M,1}}{X_M})$ verifies these inequalities:

$$\mathrm{var}(2G\Re(\frac{N_{M,1}}{X_M})) \leq \frac{2\pi^2}{P_c^2 \tau_N^2}\frac{\sigma^2}{|X_M|^2}\sum_i h_i^2 \tau_i^2$$
$$\mathrm{var}(2G\Re(\frac{N_{M,1}}{X_M})) \leq \frac{4\pi^2}{P_c^2 \tau_N^2}\frac{\sum_i h_i^2 \tau_i^2}{\sum_i h_i^2}\,\mathrm{var}(\Re(N_M^{'}))$$
$$\mathrm{var}(2G\Re(\frac{N_{M,1}}{X_M})) \leq \frac{4\pi^2}{P_c^2} Q(h,N)\,\mathrm{var}(\Re(N_M^{'}))$$
$$Q(h,N) \triangleq \frac{\sum_i h_i^2 \tau_i^2}{\tau_N^2 \sum_i h_i^2}$$

As $\frac{\sum_i h_i^2 \tau_i^2}{\sum_i h_i^2}$ is $O(N^2)$, $Q(h,N)$ will have a finite limit $Q(h)$ as N tends to infinity. For a rectangular window, $Q(h)$ is equal to $1/12$. For the other windows, a numerical evaluation of $Q(h)$ has been done in Table 4. From the value of Table 4, we can see that $Q(h) \ll 1$. Given that the padding factor P_c is large enough, we have:

$$\mathrm{var}(2G\Re(\frac{N_{M,1}}{X_M})) \ll \mathrm{var}(\Re(N_M^{'})) \tag{3.38}$$

Table 5: Value of $Q^{'}$ for usual windows, for $N = 512$ and $\mu = 100$

	Hann	Hamming	Blackman	Gaussian
$Q^{'}(h, 512, 100)$	0.03	0.04	0.02	0.04

The variance of $2G\Re(\frac{X_{M,1}}{X_M} N^{'}_M))$ verifies this relation:

$$\mathrm{var}(2G\Re(\frac{X_{M,1}}{X_M} N^{'}_M)) \leq \frac{4\pi^2}{P_c^2 \tau_N^2} \left| \frac{X_{M,1}}{X_M} \right|^2 \mathrm{var}(\Re(N^{'}_M))$$

Parseval's theorem states that for any signal y with a DFT equal to Y_k for the bin k:

$$\sum_{i=-(N-1)/2}^{(N-1)/2} |y_i|^2 = \frac{1}{N} \sum_{k=0}^{N-1} |Y_k|^2 \tag{3.39}$$

Applying this formula to $y_i = h(\tau_i)\tau_i x(\tau_i)$, we can conclude that:

$$|X_{M,1}|^2 < \sum_{k=0}^{N-1} |X(t_{m_M}; \omega_{k_M}; th)|^2 \tag{3.40}$$

$$|X_{M,1}|^2 < N\, \mathrm{e}^{2\lambda_M} \sum_{i=-(N-1)/2}^{(N-1)/2} h_i^2 \tau_i^2\, \mathrm{e}^{2\mu\tau_i} \tag{3.41}$$

Consider now $y_i = h(\tau_i)x(\tau_i)$. From the hypothesis that the signal is well resolved, the Fourier Transform of y_i, which corresponds to $X(t_{m_M}; \omega_{k_M}; h)$, has its energy concentrated near the maximum k_M. Therefore, for the closest bin k_M to the maximum, we have: $|X(t_{m_M}; \omega_{k_M}; h)|^2 \approx \sum_{k=0}^{N-1} |X(t_{m_M}; \omega_{k_M}; h)|^2$. From Parseval's theorem, we can conclude that:

$$|X_M|^2 \approx N\, \mathrm{e}^{2\lambda_M} \sum_{i=-(N-1)/2}^{(N-1)/2} h_i^2\, \mathrm{e}^{2\mu\tau_i} \tag{3.42}$$

From equations (3.41) and (3.42), the following relation holds:

$$\mathrm{var}(2G\Re(\frac{X_{M,1}}{X_M} N^{'}_M)) \leq \frac{4\pi^2}{P_c^2} Q^{'}(h, N, \mu)\, \mathrm{var}(\Re(N^{'}_M))$$

$$Q^{'}(h, N, \mu) \triangleq \frac{\sum_{i=-(N-1)/2}^{(N-1)/2} h_i^2 \tau_i^2\, \mathrm{e}^{2\mu\tau_i}}{\tau_N^2 \sum_{i=-(N-1)/2}^{(N-1)/2} h_i^2\, \mathrm{e}^{2\mu\tau_i}}$$

$Q^{'}$ is an increasing function of μ, and has a finite limit when N tends to infinity. When $\mu = 0$, we find that $Q^{'}(h, N, 0) = Q(h, N)$. In the case that $\mu = 100$ and $N = 512$, Table 5 shows that $Q^{'}(h, N) \ll 1$. Given that the padding factor P_c is large enough, we can say that:

$$\mathrm{var}(2G\Re(\frac{X_{M,1}}{X_M} N^{'}_M)) \ll \mathrm{var}(\Re(N^{'}_M)) \tag{3.43}$$

From (3.43) and (3.38), it holds that $\mathrm{var}(\epsilon_{N,RV}) \ll \mathrm{var}(\Re(N^{'}_M))$, and, using equation (3.37), that $\mathrm{var}(\epsilon_{N,RV}) \ll \mathrm{var}(\epsilon_{N,voc})$. Therefore, the RV noised estimation formula can be written as:

$$\begin{aligned}\beta_M + \gamma\Re(\frac{X(t_M, \omega_M; th)}{X(t_M, \omega_M; h)}) \approx \frac{1}{T} \arg(S_2 \bar{S}_1) + 2\frac{\Delta\omega}{T} \Re(\frac{X(t_M, \omega_M; th)}{X(t_M, \omega_M; h)}) \\ + 2\pi n + \epsilon_{RV} - \frac{\Im(Z)}{T}\end{aligned}$$

where ϵ_{RV} is the deterministic error of the RV method.

Expression of the variance for both methods

In summary, if the sinusoids are well resolved, the noised expression of the PCV and RV estimators can both be written as:

$$\hat{\beta} \approx \beta + \epsilon + \frac{\Im(Z)}{T} \tag{3.44}$$

where β is the frequency to be estimated, which is β_M for the PCV and $\beta_M + \gamma\Re(\frac{X(t_M,\omega_M;th)}{X(t_M,\omega_M;h)})$ for the RV estimator. $\hat{\beta}$ is the estimator for β and is given by equation (3.8) for the PCV and equation (3.13) for the RV. ϵ is the deterministic bias. For both methods, the stochastic error is approximately $\Im(Z)/T$.

The expectation of the estimators is $\beta + \epsilon$, and their variance is given by:

$$\begin{aligned}
\mathrm{var}(\hat{\beta}) &= \frac{E(\Im(Z)^2)}{T^2} && (3.45)\\
&= \frac{E((\Im(N_2^{'}) - \Im(N_1^{'}))^2)}{T^2} && (3.46)\\
&= \frac{E\left(\Im(N_1^{'})^2 + \Im(N_2^{'})^2 - 2\Im(N_2^{'})\Im(N_1^{'})\right)}{T^2} && (3.47)
\end{aligned}$$

From the definition of $N_k^{'}$, we have:

$$\Im(N_k^{'}) = \mathrm{e}^{-L_k}\left(\sum_i h_i n^I_{m_k+i}\cos(\omega_k\tau_i + \Theta_k) - \sum_i h_i n^R_{m_k+i}\sin(\omega_k\tau_i + \Theta_k))\right) \tag{3.48}$$

where $L_k = \lambda_k + \log(|\Gamma_k|)$, $\Theta_k = \alpha_k + \arg(\Gamma_k)$. m_k is the sample corresponding to the middle of the STFT number k. $n^I_{m_k+i}$ (resp. $n^R_{m_k+i}$) is the imaginary part (resp. real part) of the noise for the sample $m_k + i$. $\Im(N_k^{'})$ is a linear combination of independent, zero-mean random variables n_i, with the same variance and with real coefficients a_i: $\Im(N_k^{'}) = \sum_i a_i n_i$. Using the property $E(\Im(N_k^{'})^2) = E(n_i^2)\sum_i a_i^2$, we get:

$$E(\Im(N_k^{'})^2) = \frac{\sigma^2}{2}\,\mathrm{e}^{-2L_k}\,H_0 \tag{3.49}$$

where $H_0 = \sum_i h_i^2$.

Let's define the following variables:

$$\begin{aligned}
\Delta\lambda &\triangleq L_2 - L_1 = \tau_H\mu + \log(|\frac{\Gamma_2}{\Gamma_1}|)\\
\Delta\Theta &\triangleq \Theta_2 - \Theta_1 = \tau_H\beta_M + \arg(\Gamma_2\bar{\Gamma}_1)\\
\Delta\Phi &\triangleq \Delta\Theta - \tau_H\omega_M\\
H_1 &\triangleq \sum_{i=-(N-1-H)/2}^{(N-1-H)/2} h_{i+H/2}h_{i-H/2}\cos(\tau_i(\omega_1 - \omega_2))
\end{aligned}$$

Now, the expectation of the cross term $\Im(N_2^{'})\Im(N_1^{'})$ will be derived. From the definition of $N_k^{'}$, we have:

$$\begin{aligned}
N_2^{'}\bar{N}_1^{'} &= \mathrm{e}^{-L_1-L_2-j\Delta\Theta}\sum_i h_i n_{m_2+i}\,\mathrm{e}^{-j\omega_{k_2}\tau_i}\sum_i h_i\bar{n}_{m_1+i}\,\mathrm{e}^{j\omega_{k_1}\tau_i}\\
E(N_2^{'}\bar{N}_1^{'}) &= \mathrm{e}^{-L_1-L_2-j\Delta\Theta}\sum_{i=-(N-1-H)/2}^{(N-1-H)/2} h_{i+H/2}h_{i-H/2}E(|n_i|^2)\,\mathrm{e}^{j(\omega_1\tau_{i+H/2}-\omega_2\tau_{i-H/2})}
\end{aligned}$$

In the second equation, we have used the fact that $E(n_k\bar{n}_l) = 0$ if $k \neq l$.

$$\begin{aligned}
E(N_2^{'}\bar{N}_1^{'}) &= \sigma^2\,\mathrm{e}^{-L_1-L_2-j\Delta\Phi}\sum_{i=-(N-1-H)/2}^{(N-1-H)/2} h_{i+H/2}h_{i-H/2}\,\mathrm{e}^{j\tau_i(\omega_1-\omega_2)}\\
E(N_2^{'}\bar{N}_1^{'}) &= \sigma^2\,\mathrm{e}^{-L_1-L_2-j\Delta\Phi}\,H_1
\end{aligned}$$

The last equality comes from the parity hypothesis on h. In a similar way, it can be proved that $E(N_2^{'}N_1^{'}) = 0$ because $E(n_i^2) = E({n_i^R}^2) - E({n_i^I}^2) + 2E(n_i^R n_i^I) = 0$. Let's remark that

$$\Im(N_1^{'})\Im(N_2^{'}) = \frac{1}{2}\Re(\bar{N}_1^{'}N_2^{'} - N_1^{'}N_2^{'}) \tag{3.50}$$

$$E(\Im(N_1^{'})\Im(N_2^{'})) = \frac{1}{2}(\Re(E(\bar{N}_1^{'}N_2^{'})) - \Re(E(N_1^{'}N_2^{'}))) \tag{3.51}$$

We can therefore conclude that

$$E(\Im(N_1^{'})\Im(N_2^{'})) = \sigma^2\,\mathrm{e}^{-\Delta\lambda}\cos(\Delta\Phi)H_1 \tag{3.52}$$

Using equations (3.49) and (3.52), the variance of the estimators is finally obtained:

$$\mathrm{var}(\hat{\beta}) = \frac{\sigma^2}{\mathrm{e}^{L_1+L_2}}\frac{\Big[\cosh(\Delta\lambda)H_0 - \cos(\Delta\Phi)H_1\Big]}{T^2} \tag{3.53}$$

$$\mathrm{var}(\hat{\beta}) = \frac{\sinh(\mu\tau_W)}{\mu\tau_W}\frac{\Big[\cosh(\Delta\lambda)H_0 - \cos(\Delta\Phi)H_1\Big]}{\eta T^2|\Gamma_2\Gamma_1|} \tag{3.54}$$

where η is the Signal to Noise Ratio (SNR),

$$\eta \triangleq \frac{\mathrm{e}^{\lambda_1+\lambda_2}}{\sigma^2}\frac{\sinh(\mu\tau_W)}{\mu\tau_W}$$

From equation (3.54), the variance for the AM, FM and classical models can be deduced directly.

For the classical model, $\mu = 0$, $\gamma = 0$, $\Delta\lambda = 0$ and $\omega_{k_1} = \omega_{k_2} = \omega$, $\beta_1 = \beta_2 = \beta$. Equation (3.54) simplifies to:

$$\mathrm{var}(\hat{\beta}) = \frac{\Big[H_0 - \cos(\Delta\Phi)H_1\Big]}{\eta T^2|\Gamma|^2} \tag{3.55}$$

where,

$$\eta = \frac{\mathrm{e}^{2\lambda}}{\sigma^2}, \qquad H_1 = \sum_{i=-(N-H)/2}^{(N-H)/2} h_{i+\frac{H}{2}}\, h_{i-\frac{H}{2}},$$

$$\Delta\Phi = T(\beta_M - \omega_M), \qquad \Gamma = \sum_{i=-N/2}^{N/2} h_i$$

This last equation is the same as in [1].

Four examples are given on Figure 3(a) and 3(b). On Figure 3(a) the three upper curves correspond to the FM model and three lower one to the AM/FM model. In areas where the stochastic errors dominate, the theoretical variance matches the experimental MSE of the estimators. For the AM/FM model (upper curves), biases appear at high SNRs and low SNRs. In the former case, it is caused by the deterministic error of the estimator, and in the latter case, by the tracking scheme.

On Figure 3(a) the three upper curves corresponds to the FM model and three lower one to the AM/FM model. The theoretical variance matches the experimental curves. For the AM model, in the low SNRs case, the error due to the tracking scheme is slightly visible.

4 Conclusion

In this paper, it has been proved that the Fourier-based reassignment method is valid for an AM/FM model, using an original method.

The phase-vocoder frequency estimator has also been studied in the case of an AM/FM model. Two modified phase-vocoder-based schemes have been proposed: the Phase Corrected Vocoder (PCV) which aims at correcting the biased Fourier phases, and the Reassigned Vocoder (RV) which is an accurate estimator involving time reassignment. For both methods, the theoretical variance has been derived for a white-Gaussian-noise perturbation, and an experimental study of the biases has been done.

References

[1] S. Abeysekera and K. Padhi. An investigation of window effects on the frequency estimation using the phase vocoder. *IEEE Transactions on Audio Speech and Signal Processing*, 14(4):1432–1439, Jul 2006.

[2] François Auger and Patrick Flandrin. Improving the readability of time-frequency and time-scale representation by the reassignment method. *IEEE Transaction on Signal Processing*, 43(5):1068–1088, May 1995.

[3] M. Betser, P. Collen, B. David, and G. Richard. Review and discussion on STFT-based frequency estimation methods. *In the proceedings of the 120th Audio Engineering Society Convention*, Paris, France, 2006.

[4] M. Betser, P. Collen, and J.-B. Rault. Accurate FFT-based phase estimation for chirp-like signals. *In the proceedings of the 120th Audio Engineering Society convention*, Paris, France, 2006.

[5] M. Betser, P. Collen, G. Richard, and B. David. Estimation of frequency for AM/FM models using the phase vocoder framework. *IEEE Transactions on Signal Processing (in press)*, 2007.

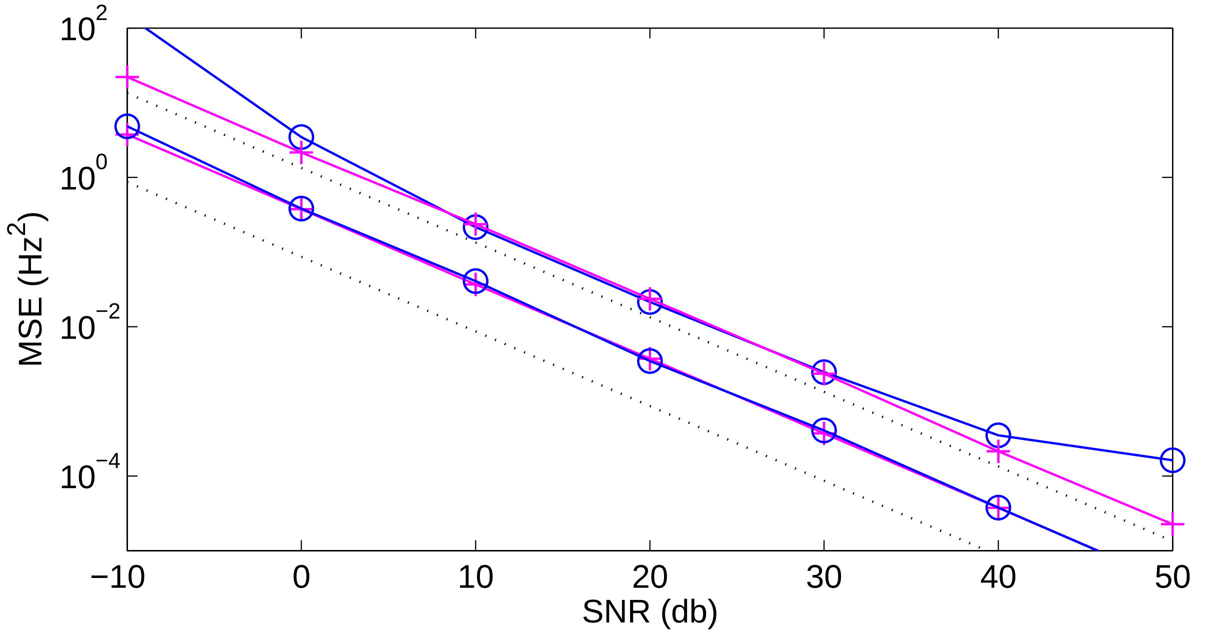
[6] M. Betser, P. Collen, and G. Richard. Frequency estimation based on adjacent DFT bins. *In Proceedings of EUSIPCO*, Florence, Italy, 2006.

[7] M. Macleod. Fast nearly ML estimation of the parameters of real or complex single tones or resolved multiple tones. *IEEE Transaction on Signal Processing*, 46(1):141–148, Jan 1998.

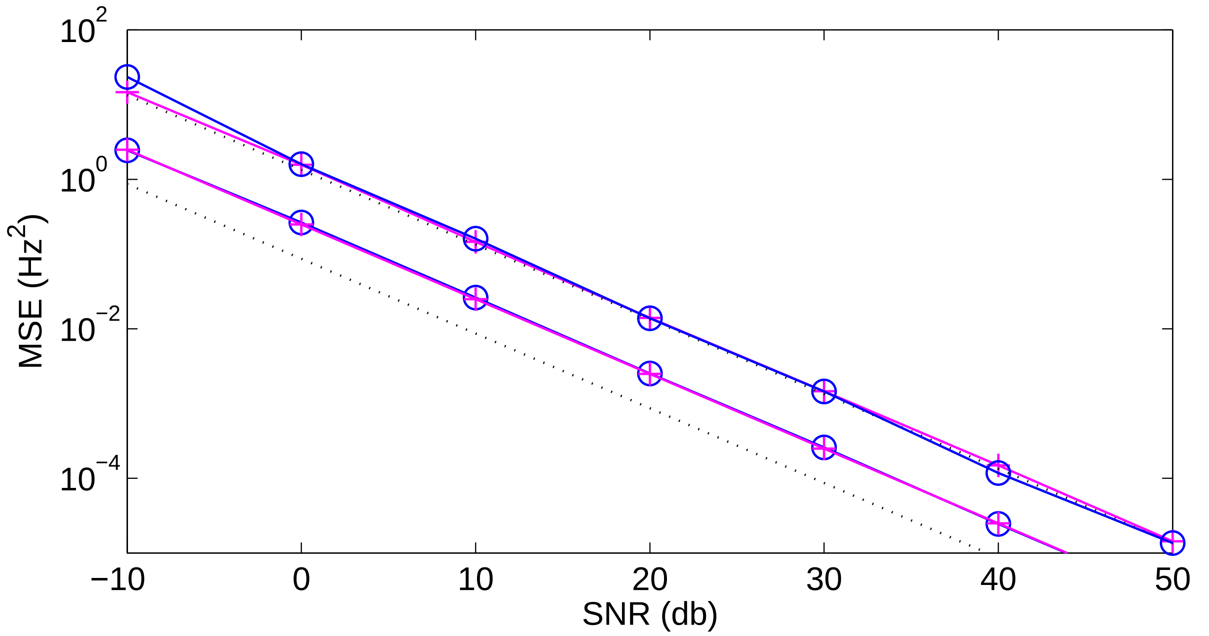
[8] R.J. McAulay and T.F. Quatieri. Speech analysis/synthesis based on a sinusoidal representation. *IEEE transactions on Acoustics, Speech and Signal Processing*, ASSP-34(4):744–754, Aug 1986.

[9] Miller S. Puckette and Judith C. Brown. Accuracy of frequency estimate using the phase vocoder. *IEEE Transaction on Speech and Audio Processing*, 6(2):166–176, Mar 1998.

[10] Barry G. Quinn. Estimating frequency by interpolation using Fourier coefficients. *IEEE Transactions on Signal Processing*, 42(5):1264–1268, May 1994.



(a) Upper curves correspond to the AM/FM model with $\mu \in [0, 100]$ and $\gamma \in [0, 8000]$, and lower curves to the FM model.



(b) Upper curves correspond to the AM model with $\mu \in [0, 100]$ and lower curves to the classical model.

Figure 3: Comparison of the theoretical vocoder variance ('+' markers) to the CRB (doted lines) and to the MSE of the RV method ('o' markers).

Dépôt légal : 2007 – 3ème[r] trimestre
Imprimé à l'Ecole Nationale Supérieure des Télécommunications – Paris
ISSN 0751-1345 ENST D (Paris) (France 1983-9999)

Ecole Nationale Supérieure des Télécommunications

Groupe des Ecoles des Télécommunications - membre de ParisTech

46, rue Barrault - 75634 Paris Cedex 13 - Tél. + 33 (0)1 45 81 77 77 - www.enst.fr

Département TSI