A Robust and Computationally Efficient Subspace-based Fundamental Frequency Estimator

Johan Xi Zhang, *Student Member, IEEE,* Mads Græsbøll Christensen, *Member, IEEE,* Søren Holdt Jensen *Senior Member, IEEE,* and Marc Moonen, *Fellow, IEEE*

Abstract

This paper presents a method for high-resolution fundamental frequency (F_0) estimation based on subspaces decomposed from a frequency-selective data model, by effectively splitting the signal into a number of subbands. The resulting estimator is termed frequency-selective harmonic MUSIC (F-HMUSIC). The subband based approach is expected to ensure computational savings and robustness. Additionally, a method for automatic subband signal activity detection is proposed, which is based on information-theoretic criterion where no subjective judgment is needed. The F-HMUSIC algorithm exhibits good statistical performance when evaluated with synthetic signals for both white and colored noises, while its evaluation on real-life audio signal shows the algorithm to be competitive with other estimators. Finally, F-HMUSIC is found to be computationally more efficient and robust than other subspace-based F_0 estimators, besides being robust against recorded data with inharmonicities.

Part of this work has been presented at IEEE International Conference on Acoustics, Speech, and Signal Processing 2009.

The work of J. X. Zhang is supported by the Marie Curie EST-SIGNAL Fellowship, contract no. MEST-CT-2005-021175. J. X. Zhang, S. H. Jensen are with Dept. of Electronic Systems, Aalborg University, Aalborg, Denmark. Emails: {jxz, shj}@es.aau.dk. M. G. Christensen is with Dept. of Media Technology, Aalborg University. Email: mgc@imi.aau.dk

M. Moonen is with Dept. of Electrical Engineering (ESAT-SCD), Katholieke Universiteit Leuven, Leuven, Belgium. Email: marc.moonen@esat.kuleuven.be

A Robust and Computationally Efficient Subspace-based Fundamental Frequency Estimator

***Index Terms*—fundamental frequency estimation, pitch estimation, subband processing, subspace methods.**

I. Introduction

The problem of estimating the fundamental frequency (F_0) or pitch in a recorded signal has been of interest to the signal processing community for many years. Many sophisticated algorithms have been proposed where the motivation for the intensive research in F_0 estimators is its wide usability, both within and outside the field of engineering. The non-ideal characteristics of recorded data make the estimators particularly challenging to design. For more details about F_0 properties of musical instruments, one can refer to, [1], [2]. In signal processing, the F_0 estimator is often a key component in speech and audio applications, such as linear prediction based speech coding, coding of speech and audio, using a harmonic sinusoidal model, and musical information retrieval. Even in the field of linguistics, F_0 estimators can be applied when the analysis of tones (pitch) is an important part of understanding and classifying the language, such as tonal languages [3], [4].

Most existing methods suffer from a degraded performance owing due to non-ideal characteristics of the recorded data, such as low signal-to-noise ratio (SNR), missing partials, inharmonicity, signal transients, and reverberations. Estimators are often time-domain techniques based on the auto-correlation function, cross-correlation function, averaged magnitude difference function, or average squared difference function. Other methods are based mainly on spectral extraction of the spectrogram. In most of the cases, only a "rough" estimate of F_0 can be obtained. For a historical review of F_0 estimation methods, one can refer to [5]–[7].

The harmonic structure of speech and audio signals can be modeled as follows: considering a set of harmonic signals with frequencies ω_l for $l = 1, ..., L$ embedded in noise:

$$y(t) = \sum_{l=1}^{L} \beta_l e^{j\omega_l t} + e(t), \qquad \beta_l = \alpha_l e^{j\theta_l}, \qquad (1)$$

where $t = 0, ..., N-1$ is the time index, L is the model order, α_l is the real-valued amplitude of lth complex exponential, θ_l is its phase, and $e(t)$ is complex symmetric white Gaussian noise. For perfect harmonic signals, the frequencies of the complex exponentials are exact integer multiples of ω_0 with unit [rad/s]. This perfect harmonic model is not always valid. Depending on the instrument, different parametric models of the inharmonicity of the harmonic can be derived from physical models [1], [2]. A common model used for stiff-stringed instruments is $\omega_l = \omega_0 l\sqrt{1 + Bl^2}$ for $B \ll 1$ where B is normally referred to as the inharmonicity coefficient, which is dependent on physical parameters of the string. The problem considered here is the estimation of ω_0 with or without estimation of the model order L in a time frame of N measured samples. The estimation problem associated with real valued signals can be cast as (1), by using analytic signals, which is valid when there is little or no spectral content of interest near 0 and π. To simplify the sinusoidal model as well as the algorithm only complex-valued signals are considered here.

Recently, F_0 estimation algorithms based on subspace techniques have shown good estimation performance with a high accuracy in low SNR conditions; they also provide flexibility for robust estimation on inharmonic signals [8]–[10], and for multi-pitch signals of known orders in [9] and for unknown orders in [11]. Currently, the main disadvantages of subspace-based F_0 estimators are the high computational complexity of the subspace decomposition process, and their sensitivity to colored noise on the estimation of signal and noise subspaces.

This paper presents an algorithm for high-resolution fundamental frequency estimation, based on subspaces decomposed from a frequency-selective (FS) data matrix model using inputs from a discrete Fourier transform (DFT). The resulting algorithm, termed Frequency-selective Harmonic MUSIC (F-HMUSIC), represents a frequency domain extension of HMUSIC [8]. F-HMUSIC adopts a subband-based approach, following which the signal spectrum is divided into Q equally spaced subbands where in each band an individual estimation problem is considered. This approach leads to a computationally more efficient algorithm, when compared with that of HMUSIC where the subspace decomposition is applied directly on the fullband covariance matrix. Besides, the averaging of estimated F_0 from the different subbands is expected to lead to

more robustness to colored noise. Moreover, the signal model order estimation used in HMUSIC is limited to model orders $L \geq 1$, and therefore automatic signal presence detection for the case $L = 0$ in subbands is not possible [12], [13]. Estimating fundamental frequency on subband without complex exponentials will give erroneous estimates, which will strongly reduce the estimation accuracy of F-HMUSIC. Therefore, a new method for automatic signal activity detection in subbands is proposed, which is based on information-theoretic criterion [14]. The main advantage of this method is that no subjective judgment is required in the decision process. Based on this knowledge of the subband activity, additional computational savings can be achieved, e.g. by simplifications on the order estimation stage, and by estimating the fundamental frequency only in active subbands. For a more complete discussion on order detection and estimation based on information-theoretical criteria, one can refer to [14]–[18]. Besides these, one can refer to [19] for a general overview of subspace-based estimation techniques, and [20], [21] for an overview of frequency estimation algorithm using FS-data model.

The performance of the automatic detection method was evaluated using Monte Carlo simulations where different parameters where examined. Using this automatic detection in the subbands, F-HMUSIC was evaluated on recorded musical signals [22] and its performance compared with that of HMUSIC and YIN [7], [8]. Parameter selections and the problems encountered during the evaluations are discussed. Additionally, the statistical properties of F-HMUSIC were evaluated using Monte Carlo simulations for synthetic signals.

The remaining part of the paper is organized as follows. In Section II, the development of F-HMUSIC is introduced where the frequency-selective data matrix model is reviewed, and an automatic subband detection method is proposed. The evaluation results from both recorded and synthetic signals are presented in Section III, and conclusions are drawn in Section IV.

II. Proposed Methods

A. Frequency-Selective Data Matrix Model

For a given signal sequence (1), the vector form is given as

$$\mathbf{y} = \left[\begin{array}{cccc} y(0) & \dots & y(N-1) \end{array} \right]^T, \tag{2}$$

which is first Fourier transformed using an N point DFT, T is the vector transpose. It is then assumed that the components of interest lie in a prespecified subband composed of the following Fourier frequencies:

$$\left\{ \begin{array}{cccc} \frac{2\pi}{N}k_1^q & \frac{2\pi}{N}k_2^q & \dots & \frac{2\pi}{N}k_M^q \end{array} \right\}, \tag{3}$$

where q denotes the subband index of $Q = N/(2M)$ equally divided subbands, and $\left\{k_1^q ... k_M^q\right\}$ the M given consecutive integers. The number of components L_q lying in the qth subband specified by (3) is assumed to be $L_q \leq L$. FS data model was first formulated with the DFT of vector (2) at frequency index k denoted as

$$Y_k = \mathbf{v}_k^* \mathbf{y}, \qquad k = 0, 1, \dots, N-1 \tag{4}$$

with $\mathbf{v}_k^*$ denoting the Fourier vector given as

$$\mathbf{v}_k = \left[\begin{array}{cccc} 1 & z_k & \dots & z_k^{N-1} \end{array} \right]^T, \tag{5}$$

where $z_k = e^{j\frac{2\pi}{N}k}$ and $*$ is the conjugate transpose.

Let $\left\{\omega_l^q\right\}_{l=1}^{L_q}$ denote the components of interest lying in the qth subband, then the FS data model is formulated by stacking phase-shifted Y_k and multiplying it with shift vector $\mathbf{u}_k$,

$$\mathbf{u}_k = \left[\begin{array}{cccc} z_k & \dots & z_k^s \end{array} \right]^T, \tag{6}$$

where s is a user-defined parameter. In the ideal case the phase shift is equivalent to time-delay if no noise is present in the signal [21]. The choice of s will be discussed in II-B. The stacked DFT element is then formulated as

$$\mathbf{u}_k Y_k = \sum_{l=1}^{L_q} \beta_l \mathbf{u}_k \left[\mathbf{v}_k^* \mathbf{b}(\omega_l^q)\right] + \mathbf{u}_k E_k, \tag{7}$$

where the scalar E_k is the DFT of the noise at frequency index k,

$$E_k = \mathbf{v}_k^* \mathbf{e}, \qquad k = 0, 1, \dots, N-1, \tag{8}$$

where $\mathbf{e} = \left[\begin{array}{cccc} e(0) & \dots & e(N-1) \end{array} \right]^T$ is the noise vector. With tedious manipulations [23], for a general ω the term $\mathbf{u}_k \left[\mathbf{v}_k^* \mathbf{b}(\omega)\right]$ can be rewritten as

$$\mathbf{u}_k \left[\mathbf{v}_k^* \mathbf{b}(\omega)\right] = \mathbf{a}(\omega) \left[\mathbf{v}_k^* \mathbf{b}(\omega)\right] + \left[\begin{array}{c} \boldsymbol{\gamma}_1^*(\omega) \\ \vdots \\ \boldsymbol{\gamma}_s^*(\omega) \end{array} \right] \mathbf{u}_k, \tag{9}$$

where vectors $\mathbf{a}(\omega_l^q)$ and $\mathbf{b}(\omega_l^q)$ are specified as

$$\mathbf{a}(\omega_l^q) = \left[\begin{array}{cccc} e^{j\omega_l^q} & \dots & e^{js\omega_l^q} \end{array} \right]^T \tag{10}$$

$$\mathbf{b}(\omega_l^q) = \left[\begin{array}{cccc} 1 & e^{j\omega_l^q} & \dots & e^{j(N-1)\omega_l^q} \end{array} \right]^T, \tag{11}$$

which express the harmonic components of the signal, and $\boldsymbol{\gamma}_p$ for $p = 1, ..., s$,

$$\begin{aligned} \boldsymbol{\gamma}_p = (1 - e^{i\omega N})[& e^{j\omega(p-1)} \quad e^{j\omega(p-2)} \\ & \cdots \quad 1 \quad 0 \quad \cdots \quad 0 \ ], (1 \times s). \end{aligned} \tag{12}$$

The first term $\mathbf{a}(\omega)\,[\mathbf{v}_k^*\mathbf{b}(\omega)]$ in (9) contains the parameter of interest. With the result in (7) and (9), the key equation of the FS data matrix model involving the DFT sequence Y_k is then given as [23]

$$\begin{aligned}\mathbf{u}_k Y_k &= \left[\mathbf{a}(\omega_1^q)\ldots\mathbf{a}(\omega_{L_q}^q)\right]\begin{bmatrix}\beta_1\mathbf{v}_k^*\mathbf{b}(\omega_1^q)\\ \vdots\\ \beta_{L_q}\mathbf{v}_k^*\mathbf{b}(\omega_{L_q}^q)\end{bmatrix}\\ &+ \mathbf{\Gamma}_q\mathbf{u}_k + \mathbf{u}_k E_k. \end{aligned} \quad (13)$$

where $\mathbf{\Gamma}_q \in \mathbb{C}^{s\times s}$ is defined as

$$\mathbf{\Gamma}_q = \left\{\sum_{l=1}^{L_q}\beta_l\begin{bmatrix}\gamma_1^*(\omega_l^q)\\ \vdots\\ \gamma_s^*(\omega_l^q)\end{bmatrix}\right\}, \quad (14)$$

this matrix has no importance in what follows.

For the in-band components of interest in (13), the following notation is used:

$$\mathbf{A}_q = \begin{bmatrix}\mathbf{a}(\omega_1^q) & \ldots & \mathbf{a}(\omega_{L_q}^q)\end{bmatrix} \quad (15)$$

$$\mathbf{x}_k = \begin{bmatrix}\beta_1\mathbf{v}_k^*\mathbf{b}(\omega_1^q)\\ \vdots\\ \beta_{L_q}\mathbf{v}_k^*\mathbf{b}(\omega_{L_q}^q)\end{bmatrix}. \quad (16)$$

It has been proved in [20], [21] that the out-of-band components are significantly smaller than the in-band components; therefore, in this paper leakage signals in the subband is assumed to be zero. A compact matrix form of (13), for DFT frequencies of interest in the qth subband, is given as

$$\mathbf{Y}_q = \mathbf{A}_q\mathbf{X}_q + \mathbf{\Gamma}_q\mathbf{U}_q + \mathbf{E}_q, \quad (17)$$

where matrices in (17) are defined as

$$\mathbf{Y}_q = \begin{bmatrix}\mathbf{u}_{k_1^q}Y_{k_1^q} & \ldots & \mathbf{u}_{k_M^q}Y_{k_M^q}\end{bmatrix} \quad (18)$$

$$\mathbf{E}_q = \begin{bmatrix}\mathbf{u}_{k_1^q}E_{k_1^q} & \ldots & \mathbf{u}_{k_M^q}E_{k_M^q}\end{bmatrix} \quad (19)$$

$$\mathbf{U}_q = \begin{bmatrix}\mathbf{u}_{k_1^q} & \ldots & \mathbf{u}_{k_M^q}\end{bmatrix} \quad (20)$$

$$\mathbf{X}_q = \begin{bmatrix}\mathbf{x}_{k_1^q} & \ldots & \mathbf{x}_{k_M^q}\end{bmatrix}, \quad (21)$$

with $\mathbf{Y}_q \in \mathbb{C}^{s\times M}$. The last term in (17) is the noise term. Term $\mathbf{\Gamma}_q\mathbf{U}_q$ in (17) is eliminated by postmultiplying (17) with the projection matrix,

$$\mathbf{\Pi}_q^\perp = \mathbf{I} - \mathbf{U}_q^*(\mathbf{U}_q\mathbf{U}_q^*)^{-1}\mathbf{U}_q, \quad (22)$$

which is the orthogonal projection matrix onto the null space of $\mathbf{U}_q$. This, in turn, is a $s \times M$ matrix, where s is chosen such that $M > s$. The resulting expression is written as

$$\mathbf{Y}_q\mathbf{\Pi}_q^\perp = \mathbf{A}_q\mathbf{X}_q\mathbf{\Pi}_q^\perp + \mathbf{E}_q\mathbf{\Pi}_q^\perp, \quad (23)$$

The matrix $\mathbf{Y}_q\mathbf{\Pi}_q^\perp$ obtained for qth subband can be decomposed using singular value decomposition (SVD) [8], [23]:

$$\mathbf{Y}_q\mathbf{\Pi}_q^\perp = \mathbf{H}_q\mathbf{\Lambda}_q\mathbf{V}_q^*. \quad (24)$$

The matrices $\mathbf{H}_q \in \mathbb{C}^{s\times s}$ and $\mathbf{V}_q \in \mathbb{C}^{M\times M}$ are both orthonormal. The matrix $\mathbf{H}_q$ in (24) is written as

$$\mathbf{H}_q = \begin{bmatrix}\mathbf{h}_1 & \mathbf{h}_2 & \ldots & \mathbf{h}_s\end{bmatrix}, \quad (25)$$

where the columns of $\mathbf{H}_q$ contain the singular vectors defining the signal and noise subspace, and $\mathbf{\Lambda}_q$ is a diagonal matrix with the corresponding non-negative singular values sorted in a decreasing order. Furthermore, $\mathbf{S}_q$ and $\mathbf{G}_q$ are denoted as follows:

$$\mathbf{S}_q = \begin{bmatrix}\mathbf{h}_1 & \mathbf{h}_2 & \ldots & \mathbf{h}_{L_q}\end{bmatrix} \quad (26)$$

$$\mathbf{G}_q = \begin{bmatrix}\mathbf{h}_{L_q+1} & \mathbf{h}_{L_q+2} & \ldots & \mathbf{h}_s\end{bmatrix}, \quad (27)$$

with $\mathbf{S}_q$ denoting the signal subspace associated with the L_q principal singular values, and $\mathbf{G}_q$ denoting the noise subspace associated with $s - L_q$ singular values. The noise subspace spanned by $\mathbf{G}_q$ is then orthogonal to the Vandermonde matrix $\mathbf{A}_q$ defined in (15), i.e.,

$$\mathbf{A}_q^*\mathbf{G}_q = \mathbf{0}. \quad (28)$$

for frequencies ω_l^q where $l = 1, ..., L_q$.

B. F-HMUSIC

In this part, F-HMUSIC is formulated with a subband-based approach for estimating both ω_0 and the model order L for complex exponentials with frequencies $\omega_l = \omega_0 l\sqrt{1+Bl^2}$, $l = 1, ..., L$. The spectrum support from 0 to π is divided into Q equally spaced subbands where subband containing complex exponentials is referred as active subband with Q'. For the sake of simplicity, Q' is assumed to be known. In the next subsection, the proposed subband activity detection method is described.

The harmonic model order L of (1) is given as

$$L = \sum_{q=1}^{Q} L_q, \quad (29)$$

where L_q denotes the number of complex exponentials in subband q. The number of complex exponentials in each subband is further derived from the laws of inharmonicity written as

$$L_q = \left\lfloor L_q' - \sum_{i=1}^{q-1}\lfloor L_i'\rfloor \right\rfloor, \quad (30)$$

where

$$L_q' = \frac{\left(\frac{2\pi}{N}k_M^q\right)}{\omega_0\sqrt{\frac{1}{2}+\sqrt{\frac{1}{4}+B\left(\frac{\frac{2\pi}{N}k_M^q}{\omega_0}\right)^2}}} \quad (31)$$

is derived from $\frac{2\pi}{N}k_M^q > \omega_0 L_q\sqrt{1+BL_q^2}$. In this paper, B is assumed to be known. In F_0 estimations on recorded piano notes, the average B measured from various pianos can often be used, as an example in [24] good estimation results have been obtained by using the average B in estimations on recorded piano notes. If B is unknown, it can be estimated as a parameter of interest in the extended cost function [10].

The Vandermonde matrix $\mathbf{A}_q$ in (15) has been derived without taking into account the harmonic structure of the signal. When taken into account, (15) can be parameterized with fundamental frequency as

$$\mathbf{A}_q = \left[\begin{array}{ccc} \mathbf{a}\left(\omega_0(L'_{q-1}+1)\right) & \ldots & \mathbf{a}\left(\omega_0 L_q\right) \end{array} \right], \quad (32)$$

for notational simplicity $B=0$. The joint order and fundamental frequency estimation cost function for subband q are obtained by projecting the Vandermonde matrix $\mathbf{A}_q$ onto the noise subspace $\mathbf{G}_q$ where the matrix $\mathbf{A}_q$ is a matrix function dependent on both ω_0 and the order L_q where $\mathbf{G}_q$ is dependent on only L_q. The complete cost function for all active subbands is given as

$$J(\omega_0, L) = \frac{1}{Q'}\sum_{q=1}^{Q'} \frac{\left\|\mathbf{A}_q^*\mathbf{G}_q\right\|_F^2}{s\min(L_q, s-L_q)}, \quad (33)$$

where the order of individual subband L_q is calculated for each given L using (30). Furthermore, the denominator is a scaling factor that makes the noise floor of the cost function invariant to the changing matrix dimensions of $\mathbf{A}_q$ and $\mathbf{G}_q$ because of the angle between subspaces defined in [12]. More specifically, the measure is the average over cosine of all the non-trivial angles between the subspaces spanned by the column of $\mathbf{A}_q$ and $\mathbf{G}_q$. The estimates for the order L and the fundamental frequency ω_0 are obtained by minimizing (33) thus,

$$\hat{\omega}_0 = \arg\min_{\omega_0\in\Omega}\min_{L\in\mathcal{L}} J(\omega_0, L), \quad (34)$$

where Ω is the searching space for the fundamental frequency, and $\mathcal{L}$ for the order estimation.

The performance of the proposed method depends on many parameters, such as the data length N, the number of subbands Q, the user parameter s, and the search space Ω. Obviously, increased N improves resolution while increasing Q reduces the resolution and increases the algorithm sensitivity to noise, because the resulting reduction in the number of consecutive DFT data samples in each subband M leads to reduced maximum possible value of s. Previous experience with similar approaches shows that selecting as large a user parameter s as possible increases the number of linearly independent vectors in the noise subspace. Nevertheless, the parameter s must be less than M to achieve the correct estimate of the FS data matrix model [23], [25]. The advantage of having a large number of subbands Q is to reduce the computational complexity, which is explained by the cubic increase of computational complexity of SVD algorithm. In full-band processing of HMUSIC, the computational complexity resulting from using SVD on full-band covariance matrix is of the order $\mathcal{O}\left(N^3\right)$, and by splitting up the estimation problem into sub-problems as in F-HMUSIC, the computational load will be $\mathcal{O}\left(\frac{N^3}{(2Q)^3}\right)$ when frequency samples from region 0 to π are used. This is a reduction in computational complexity of $\frac{1}{(2Q)^3}$. Another important factor is the search space Ω: in the cost function, the linearly independent vectors in $\mathbf{G}_q$ are dependent on L_q which in turn is dependent on ω_0. It is stressed here that L_q increases with reduced ω_0. So, the lower bound of the search space should be so selected as to include sufficient number of linearly independent vectors in $\mathbf{G}_q$ to retain satisfactory levels of noise robustness. Additionally, when very low ω_0 is selected, vectors in the Vandermonde matrix $\mathbf{A}_q$ will be rank deficient.

The cost function in (33) can be minimized using either an FFT-based method or a gradient based method. Both methods are described in [8]. The FFT-based method gives only a coarser estimate. Therefore, for applications that require accurate estimates for a given model order, the gradient search algorithm method described in [8] can be used, but with minor modifications.

C. Subband detection

As the cost function (33) is defined only for $L>1$, it is impossible to distinguish between one complex exponential and no complex exponential. Therefore, automatic subband detection algorithm is proposed to estimate the active subbands. This detection method is formulated using the information-theoretic criterion for model selections described in [14]. It is known from [26] that the absolute square magnitude of DFT elements is asymptotically equal to the eigenvalues of the covariance matrix. The squared DFT elements are written as Y_k^2, where the elements are sorted in the descending order, with the new sorted index denoted as k'. The sorted magnitude $Y_{k'}^2$ is then inserted into the cost function derived in [14], which is given as

$$\begin{aligned} MDL(k') &= -\log\left(\frac{\prod_{n=k'+1}^{M}|Y_n|^{2/(M-k')}}{\frac{1}{M-k'}\sum_{n=k'+1}^{M}|Y_n|^2}\right)^{(M-k')N} \\ &+ \frac{1}{2}k'(2M-k')\log N, \end{aligned} \quad (35)$$

The first term in (35) is in fact the log-likelihood of the maximum likelihood estimator of the model parameters

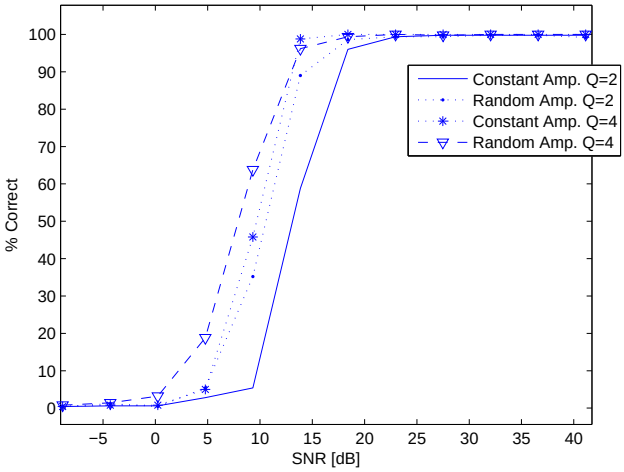


Fig. 1: *Percentages of correctly estimated subband activity on the proposed method evaluated on different SNR under the white noise conditions.*

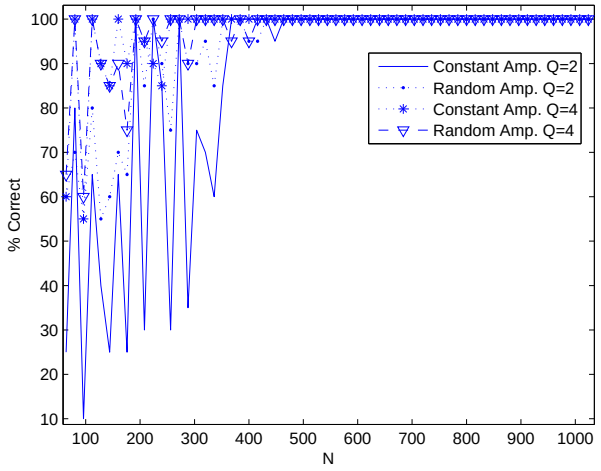


Fig. 2: *Percentages of correctly estimated subband activity on the proposed method evaluated on varying frame length N with fixed SNR=25dB.*

and the second term is a penalty term [14], [16]–[18]. In the proposed method only the activity of the band is of interest therefore when the minimum of (35) is $k' > 0$ the subband is decided as active. In general this subband detection method is used as a pre-processor to detect active subband. For computational simplicity, (33) is then only evaluated on active subbands.

Algorithm outline:

1) Extract the DFT element in the specified subband q on Y_k, with index defined in (3)
2) Sort Y_k^2 in the descending order. The new sorted index is denoted as k'.
3) Insert sorted magnitudes into (35). Find the argument that gives minimum value of (35).
4) Detect subband activities using the following rules:
   $k' > 0$, subband is active.
   $k' = 0$, subband is not active.

In this paper, when subband signal detection method is used, the active subband is assumed to have full model order L. The search range for $\mathcal{L}$ is bounded by (30). Therefore, the simplified cost function is denoted as

$$\hat{\omega}_0 = \arg \min_{\omega_0 \in \Omega} J(\omega_0, L), \tag{36}$$

where L is fixed and founded using (29).

III. Experimental Results

A. Statistical Evaluation of subband detection algorithm

Before evaluating F-HMUSIC on real recorded signals, the proposed subband signal detection method was evaluated with Monte Carlo simulations where errors were measured as correctness in detection. The test

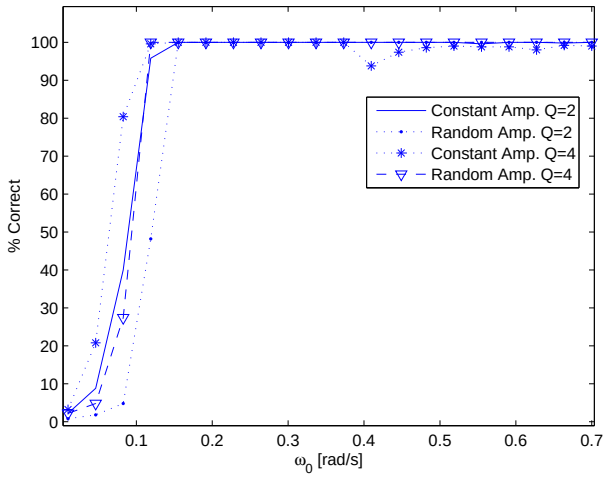


Fig. 3: *Percentages of correctly estimated subband activity on the proposed method evaluated on varying ω_0 with fixed SNR=30dB and $N = 1024$*

signal was generated according to (1), where the signal was perfectly harmonic (i.e. $B = 0$). In all simulations, $\omega_0 \in [0, 2\pi]$ with unit [rad/s]. Two types of signal amplitudes were evaluated: one with constant amplitudes and the other with random amplitudes generated according to Rayleigh distribution. For the sake of simplicity, the active subband detection errors were measured on the first subband with index $q = 1$, and the subband configurations were selected for $Q = 2$ and $Q = 4$, respectively. Four different scenarios will be evaluated. Signal model was assumed to have full model order set to $L = \left\lfloor \frac{\pi}{\omega_0} \right\rfloor$, except in the last test where various model orders were evaluated.

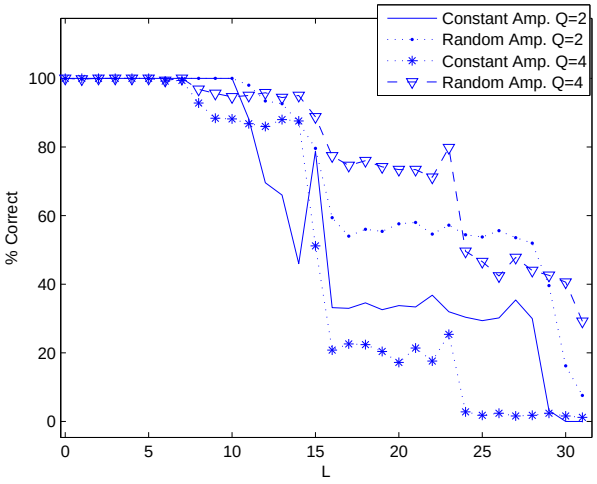


Fig. 4: *Percentages of correctly estimated subband activity on the proposed method evaluated on varying L with fixed SNR=30dB and $N = 1024$*

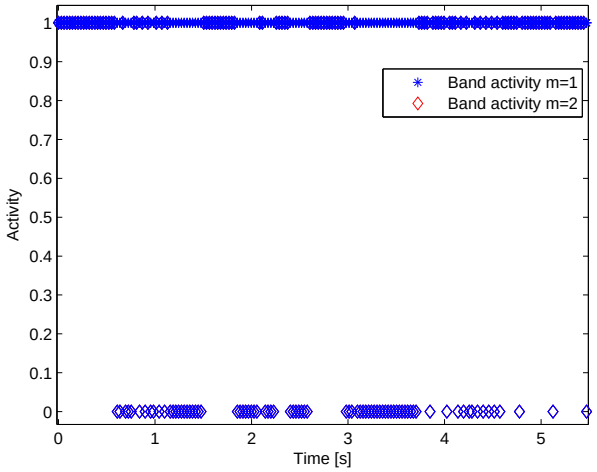
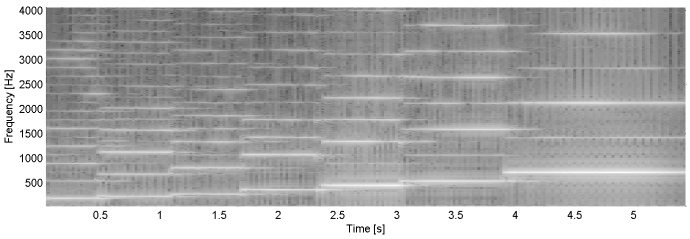
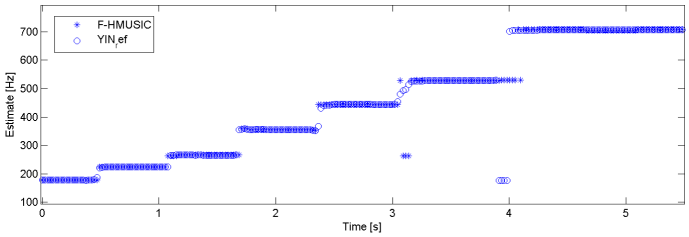


Fig. 6: *Band detection performance on individual bands of the clarinet signal in Fig. 5.*



(a)



(b)

Fig. 5: *a) Spectrogram of the clarinet signal. b) Fundamental frequencies estimated using F-HMUSIC and YIN.*

First, an experiment of detection performance versus signal-to-noise ratio (SNR) was carried out where the sample length was fixed at $N = 512$. For each SNR, 500 Monte Carlo simulations were evaluated. The signals had the fundamental frequency $\omega_0 = 0.23$ and the performance is shown in Fig. 1. From the simulations, it can be seen that almost 100% accuracy can be achieved when the SNR is above 25dB. The performances with randomly generated signal amplitudes are better than those with constant amplitudes. This can be explained by the limited sample length where DFT magnitudes are far from being equal to the asymptotically equal eigenvalues. In the case of asymptotically equality between Fourier power magnitudes and eigenvalues, every harmonic should have only one element representation in the Fourier spectrum. This is usually never the case when sample length is limited, and frequency smearing of the complex exponentials in the frequency domain will be obtained. The smearing effect is not crucial to on white noise, because perfect white noise has a flat spectrum distribution. Therefore, the proposed method performs better with randomly generated amplitudes, because interfering elements might be treated as noise elements when the power of the amplitudes is close to the noise variance. Furthermore, the same signal setup was used with SNR fixed at 30dB and applied on different sample lengths evaluated from $N = 64$ to $N = 1024$. The results are illustrated in Fig. 2. The detection algorithm performance seems to stabilize when the sample length is above $N = 512$. By increasing the sample length, a better approximation of DFT magnitudes to the eigenvalues is achieved.

The next test was carried out to evaluate the performance when ω_0 is varied from 0.01 to 0.7, with the frame length fixed at $N = 1024$ and SNR at 30dB. The simulation results are shown in Fig. 3. The difficulty in this test is mainly on the lower ω_0 where the complex exponentials are so closely spaced as to contain proportionally more interfering elements than those at higher ω_0. Finally, the performance was evaluated by varying the model order from $L = 0$ to $L = \left\lfloor \frac{\pi}{\omega_0} \right\rfloor$, from the corresponding results are shown in Fig. 4, it can be seen that the performance of band detection increases with reduced model order.

The simulation results clearly confirm that the proposed subband detection method can satisfactorily detect subband activity under various conditions. In almost all cases, no significant difference in the performance of the subband setup with $Q = 2$ and $Q = 4$ could be noticed.

B. Signal Examples

Under this section the proposed method on a recorded sequence of clarinet playing an upgoing arpeggio is first explained. The clarinet signals were assumed to be perfectly harmonic with $B = 0$. Spectrogram of the signal is shown in Fig. 5a and the estimates of F_0 using F-HMUSIC and YIN are shown in Fig. 5b. From Fig. 5b it can be seen that the proposed algorithm can successfully estimate the fundamental frequency except in the boarder region where the signal is not well defined owing to non-ideal circumstances, such as reverberation in the room that may cause a multi-pitch scenario where our model in (1) is insufficient. The setup used in F-HMUSIC in this example was on a signal with sampling frequency $f_s = 11025$Hz processed with a frame length of $N = 512$, and 50% overlaps. The user parameter was selected as $s = \lfloor 0.9M \rfloor$. The cost function was evaluated from 100Hz to 1000Hz with a search grid of 0.5Hz. The method is generally sensitive to the choice of s. For short frames, a bigger number of s is preferred. Two subbands were selected where the active subbands were automatically detected using the proposed detection algorithm. The detection results are shown in Fig. 6. In all frames, the first band was detected to be active, but according to the signal spectrogram in Fig. 5a both the subbands should be active. The subband $q = 2$ was not detected active in all the frames because its signal power was significantly less than that in the subband $q = 1$.

What follows now is the evaluation of F-HMUSIC on recordings from a database consisting of transcribed notes played by pianos [22]. The database was created in an environment reverberating at three loudness levels (piano, mezzo-forte and forte). For each note, a test set consisting of recordings played with six different pianos was selected. On average, 1000 frames of data were processed for each note. Getting information on the onset and offset timings of the note is challenging, because the test data include both metallic thumps of hammers against strings, and constantly degradation of SNR during the release of the note. Fig. 7 illustrates a spectrogram example of one note at 466Hz, which clearly shows the non-ideal signal conditions during onset and release of the note. At the onset stage, the sound produced by the piano hammer appears in the form of noise between sinusoidal tones, strongest between 0 to 0.4 second. During the release of the note, the number of sinusoidal tones decreases progressively with time. As a result, the entire evaluation had to be done under different circumstances with a general fixed parameter setup. Furthermore, the subband signal detection performance on low F_0 was more sensitive to frames with low SNR. The statistical performance on low F_0 is shown in Fig. 3. The main objective here is to evaluate the robustness of subspace based methods on real-recorded data. The method proposed here are compared with both HMUSIC[1] and YIN[2]. Most previous studies of YIN refer to it as a very robust single pitch estimator, while the performance of HMUSIC has been good on synthetic signals and on small sets of recorded signals.

During the evaluations, each estimated F_0 was quantized to the nearest note in the musical scale with A4 tuned to 440 Hz. Errors were then measured as incorrect MIDI note estimates. The evaluated signals were analyzed with a window length of $N = 1024$ and sliding forward in time with 50% overlap. For the sake of computational simplicity, the signal was downsampled to $f_s = 11025$Hz. Two types of subband configurations were evaluated in F-HMUSIC, with $Q = 2$ and $Q = 4$, respectively. The remaining parameters of F-HMUSIC and HMUSIC were empirically selected as follows: $\Omega \in [103.83, 4310]$ Hz with a search grid of 0.1 Hz, and $s = \lfloor 0.85M \rfloor$, $s' = \lfloor 0.60N \rfloor$ where s' was the user parameter for HMUSIC. Piano notes were evaluated on MIDI notes ranging from 45 to 108. In this evaluation, the lowest possible ω_0 was selected to be a little higher than the lowest note that can be produced by a piano. This is to ensure that the linearly independent vectors in noise subspace were sufficient enough to avoid rank deficiency in Vandermond matrix (32), which otherwise could give inaccurate estimations. Also, the cost function evaluated on those regions will degrade the overall performance of F_0 [25]. Therefore, one should carefully select Ω to stabilize the overall performance. It is to be noted that F-HMUSIC uses fewer data samples, which render it more sensitive to noise and rank deficiency problem than HMUSIC. Owing to the limited data length inserted into the FS data model, s needs to be selected close to M to reduce the noise influence of the data. The stiffness parameters B for different F_0 are from the results presented in [2, page 365]. The estimation errors evaluated on MIDI notes are presented in Fig. 8, which clearly show that

[1]The HMUSIC used in evaluation on recorded data is based on fixed order where the order $L = \left\lfloor \frac{\pi}{\omega_0} \right\rfloor$

[2]Parameters used in YIN is the standard parameter found on author's webpage

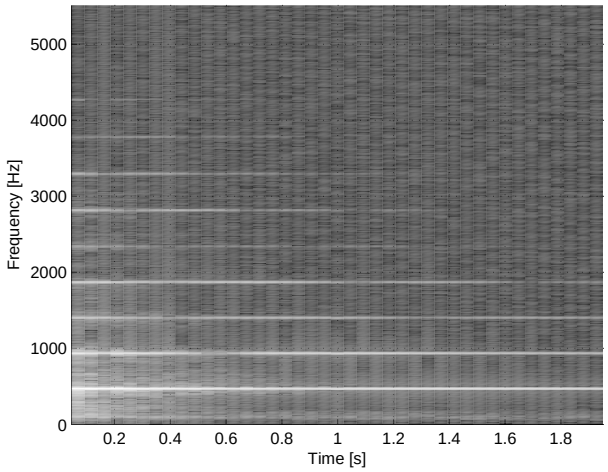


Fig. 7: *Spectrogram example of one note on MIDI note* 70 *with fundamental frequency* 466.2*Hz.*

both subspace-based methods suffer from degradation in estimating robustness on recorded signals. This can be explained by model mismatches where subspace based F_0 estimators do not make additional adaptation to the model changes. Model mismatch situations arises most probably during onset and release of the notes.

The performance on higher MIDI notes in $Q = 2$ gets significantly reduced, because as F_0 increases, model order L decreases. This reduces the estimation performance according to the asymptotic CRLB described in [8]. Regarding temporal aspects of the test signal, the reduced detection performance is attributed to physical properties of the piano sound whose amplitudes of sinusoids decay[3] rapidly for frequencies above 2800Hz [2, page 384], leading to a significant increase in the number of frames with low SNR.

To make the comparison fair between the methods involved, the errors in MIDI notes–the operating region of the algorithms concerned–will be discussed $[45, 95]$. Table I summarizes the errors, which show that the performance of HMUSIC and F-HMUSIC estimators with $Q = 2$ is comparable with that of YIN. The bad performance of F-HMUSIC with $Q = 4$ is due to the reduction in the number of subband samples when compared to $Q = 2$ case. The main disadvantage of selecting many subbands is the increases in its sensitivity to noise, while the advantage is reduction in computational complexity. In the case of $Q = 4$ the large errors above MIDI note 87 are due to the detection errors in subband $q = 1$. As this band has no complex exponentials, all erroneously detection will cause large estimation error.

[3]The speed at which amplitude decreases is often referred as decay time [2, page 384]

TABLE I: *Summarized errors of MIDI notes* $[45, 95]$

MIDI Notes Errors	Mean	Sub-octave	Octave
F-HMUSIC:Q=4	31%	2.1%	1.1%
F-HMUSIC:Q=2	6.4 %	6.3%	17.6%
HMUSIC	7.4 %	18.7 %	45.6 %
YIN	6.2 %	9.83 %	27.2%

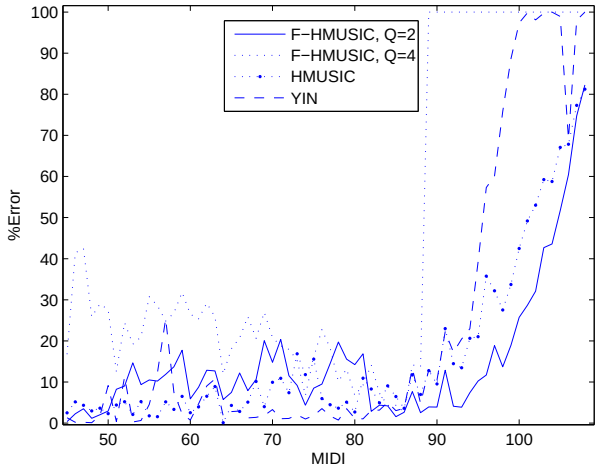


Fig. 8: *Percentage errors of the quantized MIDI notes evaluated from* $[45, 108]$.

The bad detection ability in $q = 1$ for $Q = 4$ is mainly due to the noise from the thumb of hammer, where the automatic subband detection algorithm fails to operate. Furthermore, HMUSIC is more sensitive to octave errors[4] than F-HMUSIC, but no significant differences have been observed between the performances. Even though F-HMUSIC makes a few octave and sub-octave errors, they are hard to avoid. Nevertheless, the estimator proposed here does not significantly improve the robustness of F_0 estimations on real-recorded signals. However, high-resolution estimates can be obtained, which is not possible with YIN or similar time-domain methods. Another advantage of F-HMUSIC is its lower computational complexity when compared with that of other subspace-based F_0 estimators.

C. Statistical Evaluation of F-HMUSIC

Next, the statistical properties of the proposed method were evaluated using Monte Carlo simulations under both white and colored noise conditions. For this, only the statistical properties of the algorithm were of interest and errors introduced by automatic subband signal detection were not considered. Therefore, the subbands

[4]The octave errors is defined as the ratio between the number of octave errors and the total number of frames.

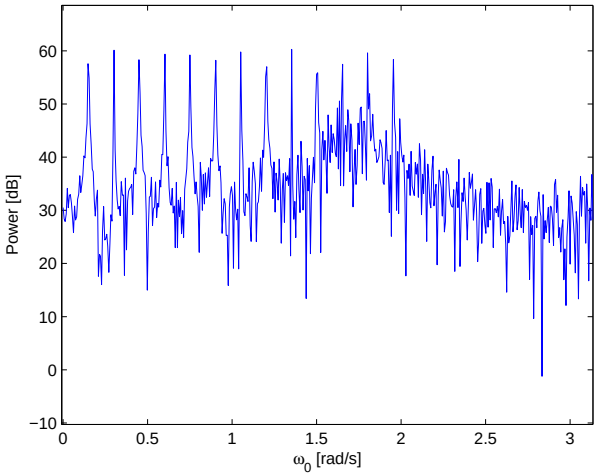


Fig. 9: *Frequency domain representation of one realization on the harmonic signal embedded in colored noise. The SNR is at* $11dB$ *with the constant amplitudes* $\alpha_l = 1 \forall l$.

containing the signal was assumed to be known. The signal is perfect harmonic with $B = 0$.

In each trial, a signal segment was generated according to the model in (1), with randomized noise. The estimators were evaluated in terms of the root mean square error (RMSE) defined as

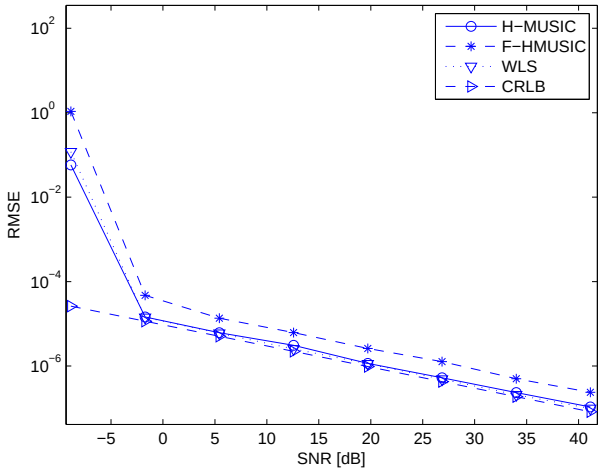
$$RMSE = \sqrt{\frac{1}{D}\sum_{d=1}^{D}(\hat{\omega}_0 - \omega_0)^2}, \tag{37}$$

where ω_0 was the true fundamental frequency, $\hat{\omega}_0$ the estimate, and D the number of Monte Carlo simulations. In this paper $D = 200$. This is done for various SNRs defined as

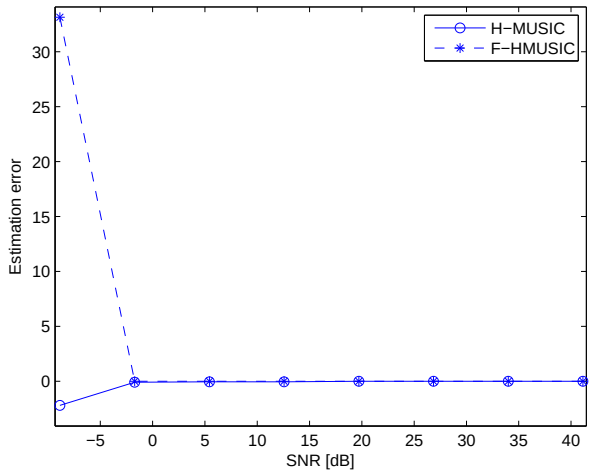
$$SNR = 10\log_{10}\sum_{l=1}^{L}\frac{\alpha_l^2}{\phi(\omega_l)}, \tag{38}$$

where the function $\phi(\omega_l)$ is the power spectrum of the noise at frequency $\omega_l = \omega_0 l$. In the case of white noise, the power spectrum is equal to the variance of noise, and in colored noise it is white noise filtered with AR process. Furthermore, model order error on the harmonic signals is defined as the difference between the estimated order and the true order. The results are compared with those of the exact CRLB for both white and colored noise cases using the equations in [27], [28].

Following are the signal and noise setup used in the experiments described below. The signal consisted of $L = 13$ complex exponentials, embedded in noise with a fundamental frequency of $\omega_0 = 0.15$. Both white and colored noises were evaluated. The constant amplitudes



(a)



(b)

Fig. 10: *a) The RMSE with a varying SNR in the case of constant amplitudes embedded in white noise. b) Corresponding model order estimation errors*

of $\alpha_l = 1 \ \forall \ l$ were considered. For both F-HMUSIC and HMUSIC, common parameters were set where the searching candidates of ω_0 were set to $\Omega \in [0.06, 0.4]$. The model order considered was $\mathcal{L} \in [5, \lfloor \pi/\omega_0 \rfloor - 1]$. It is to be noted that the interval for ω_0 included both $2\omega_0$ and $\frac{1}{2}\omega_0$, which are normally referred to as octave errors. The user parameter $s = \lfloor 0.5M \rfloor$ was selected for F-HMUSIC, and $s' = \lfloor 0.5N \rfloor$ for HMUSIC. As a second reference, the WLS estimator proposed in [29] was used. This method is computationally efficient with good statistical performance. It operates in a two-step approach where the unconstrained frequencies of complex exponentials were estimated first. Then, from these fundamental frequencies, the weighted fundamen-

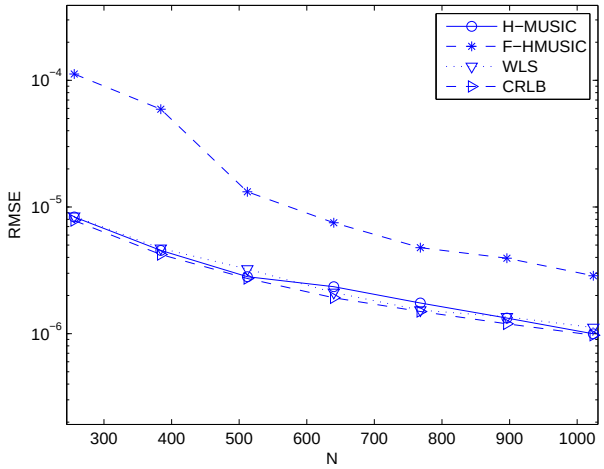


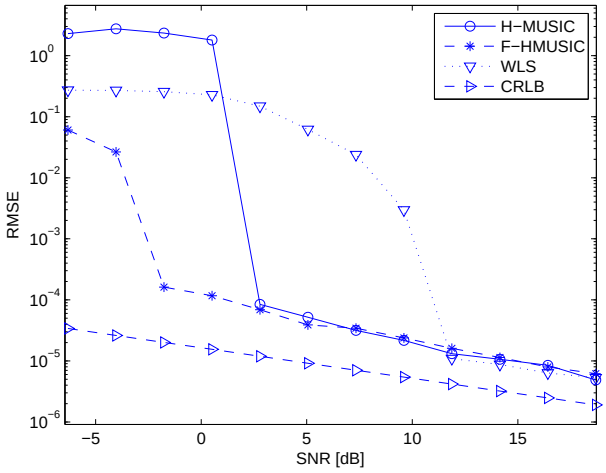
Fig. 11: *The RMSE performance with a varying frame length N where amplitudes is constant distributed and SNR fixed at $25dB$*

tal frequency was estimated. Here, WLS algorithm was assumed to have full knowledge of the signal amplitude and model order L. The implementation of WLS was taken from the toolbox in [30].

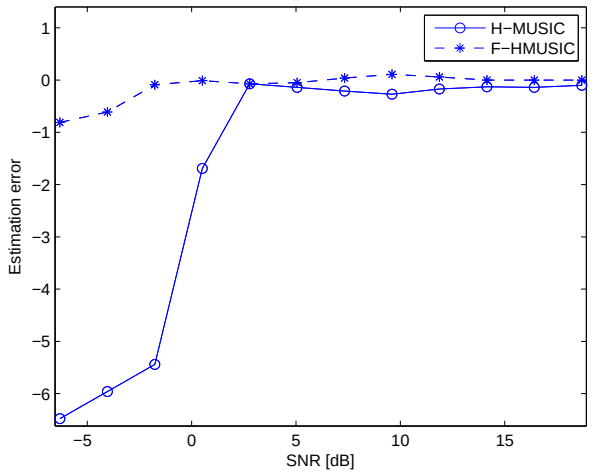
In the first example, F-HMUSIC was evaluated in white noise scenario where the amplitudes of complex exponentials were constant. The corresponding results of estimated RMSE versus a varying SNR are shown in Fig. 10a and the plots of the associated model order estimation errors in Fig.10b. The performance curve of F-HMUSIC closely corresponds to CRLB for the region above the break down[5] region of the algorithm. Even though the performance of F-HMUSIC is slightly worse than that of the reference methods, considered the computational savings, the performance of F-HMUSIC is still comparable with that of both HMUSIC and WLS as shown in Fig. 10a. Overall, F-HMUSIC has shown good statistical performance, close to that of CRLB for harmonic signals embedded in white noise. In the operation region above the breakdown point, the order estimates have also shown good accuracy, and so have the remaining experiments. Fig. 10b shows that the estimate of L is close to the true value when the estimate of ω_0 is close to CRLB.

In the next example, the RMSE performance was evaluated on a varying window length with SNR fixed at 25dB and the amplitudes at $\alpha_l = 1, \forall l$. For various frame lengths, $\lfloor 0.85M \rfloor$ was selected as the user parameter

[5]The algorithm performance break down problem is a common problem at low SNR region which is also referred as subspace swapping problem where the high noise level cause part of the signal subspace erroneously determined as noise subspace.



(a)



(b)

Fig. 12: *a) The RMSE with a varying SNR in the case of constant amplitudes embedded in colored noise. b) Corresponding model order estimation error*

for F-HMUSIC and $\lfloor 0.85N \rfloor$ for HMUSIC. The performance, illustrated in Fig. 11 shows that both algorithms can operate on any frame length N, between 256 to 1024, with judiciously selected user parameters. As expected the performance of WLS still closely follows the CRLB.

In the final example, the performance of F-HMUSIC was evaluated in colored noise scenario. The signal setup was similar to that of the previous examples, excepting that the embedded noise was filtered here with a second-order AR process $(\frac{1}{1+0.3z^{-1}+0.8z^{-2}})$, where the power of the noise was concentrated mainly on subband $q = 2$. One realization of the signal, embedded in colored noise, is shown in Fig. 9. To enhance the performance of both the methods, a slightly different setup was used for both

F-HMUSIC and HMUSIC, where the searching space for ω_0 was $\Omega \in [0.1, 0.8]$. All the other parameters remained the same as in the earlier examples. The evaluation results of colored noise are shown in Fig. 12. From these results it can be seen that the method proposed here is more robust against colored noise than HMUSIC and WLS. The algorithm breakdown region for F-HMUSIC is significantly lower than both HMUSIC and WLS. By averaging the estimates from both the subbands, a more robust estimation can be achieved. In the WLS algorithm, the poor performance is mainly due to incorrect decomposition of the signal subspace which introduces errors in the unconstrained frequency estimates.

IV. Conclusions

In this paper, a high-resolution fundamental frequency estimator termed F-HMUSIC with automatic subband signal detection has been proposed. This algorithm is a frequency-domain-based estimator using subspaces decomposed from FS data matrix model to efficiently estimate the fundamental frequency, where a subband-based approach is adopted to reduce the sensitivity to the colored noise and increase the computational efficiency. Additionally, an automatic subband signal detection method has been proposed, which is based on information-theoretic criterion where no subjective judgment is needed. The performance of F-HMUSIC has been evaluated on both synthetic and recorded signals. The simulations on synthetic data show that F-HMUSIC is more robust against colored noise than HMUSIC and WLS. The price that is paid for computational efficiency and robustness to colored noise is reduction in estimation accuracy. Furthermore, the robustness of the method has been demonstrated by evaluation on recorded signals where the performance of F-HMUSIC for $Q = 2$ approximate that of YIN for MIDI notes between $[45, 95]$, and exceeds for MIDI note above 95. Overall, with appropriate selected subband Q the performance of F-HMUSIC is considered accurate and robust for the operating region.

References

[1] T. D. Rossing, *The Science of Sound, 2nd Edition*, Addison-Wesley Publishing Company, 1990.

[2] N. H. Fletcher and T. D. Rossing, *The Physics of MUSICAL instruments*, Springer, 1998.

[3] P. J. Rose, "On the non-equivalence of fundamental frequency and pitch in tonal description," *In Prosodic Analysis and Asian Linguistics*, 1988.

[4] R. Herman, M. Beckman, and K. Honda, "Linguistic models of F0 use, physiological models of F0 control, and the issue of "mean response time"," *Language and Speech*, vol. 42, pp. 373–399, 1999.

[5] W. Hess, *Pitch Detemination of Speech Signals*, Springer-Verlag, Berlin, 1983.

[6] W. Hess, "Pitch and voicing determination," *Advances in Speech Signal Processing*, pp. 3–48, 1992.

[7] A. de Cheveigne and H. Kawahara, "YIN, a fundamental frequency estimator for speech and music," *J. Acoust. Soc. of Am.*, vol. 111(4), pp. 1917–1930, 2002.

[8] M.G. Christensen, A. Jakobsson, and S.H. Jensen, "Joint high-resolution fundamental frequency and order estimation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 15, no. 5, pp. 1635–1644, 2007.

[9] M. G. Christensen, P. Stoica, A. Jakobsson, and S. H. Jensen, "Multi-pitch estimation," *Elsevier Signal Processing*, vol. 88(4), pp. 972–983, 2008.

[10] M.G. Christensen, P. Vera-Candeas, S.D. Somasundaram, and A. Jakobsson, "Robust subspace-based fundamental frequency estimation," *IEEE Int. Conf. Acoust., Speech and Signal Processing*, pp. 101–104, 2008.

[11] M.G. Christensen, A. Jakobsson, and S.H. Jensen, "Multi-pitch estimation using Harmonic MUSIC," *in Rec. Asilomar Conf. Signals, Systems, and Computers*, pp. 521–525, 2006.

[12] M.G. Christensen, A. Jakobsson, and S.H. Jensen, "Sinusoidal order estimation using angles between subspaces," *Submitted to: IEEE Transaction in Signal Processing*, 2008.

[13] A. Jakobsson, M. G. Christensen, and S. H. Jensen, "Frequency selective sinusoidal order estimation," *IEEE Electronic Letters*, vol. 43(21), pp. 1164–1165, 2007.

[14] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 33, pp. 387– 392, 1985.

[15] M. S. Bartlett, "A note on the multiplying factors for various x^2 approximations," *J. Roy. Stat. Soc.*, vol. Ser. B,16, pp. 296–298, 1954.

[16] H. Akaike, "Information theory and an extenstion of the maximum likelihood principle," in *in Proc. 2nd Int. Symp. Inform. Theory*, 1973.

[17] J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 6, pp. 461–464, 1978.

[18] G. Schwartz, "Estimating the dimention of a model," *Ann. Stat.*, vol. 6, pp. 461–464, 1978.

[19] H. Krim and M. Viberg, "Two decades of array processing research-the parametric approach," *IEEE SP. Mag.*, July 1996.

[20] J. Gunnarsson and T. McKelvey, "High SNR performance analysis of F-ESPRIT," in *Rec. of 38th Asilomar Conference on Signals, Systems and Computers*, 2004.

[21] T. McKelvey and M. Viberg, "A robust frequency domain subspace algorithm for multi-component harmonic retrieval," in *Rec. of 34th Asilomar Conference on Signals, Systems and Computers*, 2001.

[22] V. Emiya, *Transcription automatique de la musique de piano*, Ph.D. thesis, École Nationale Supérieure des Télécommunications, 2008.

[23] P. Stoica and R. Moses, *Spectral Analysis of Signals*, Prentice Hall, 2005.

[24] V. Emiya, B. David, and R. Badeau, "A parametric method for pitch estimation of piano tones," *IEEE Int. Conf. Acoust., Speech and Signal Processing.*, 2007.

[25] P. Stoica and T. Soderstrom, "Statistical analysis of music and subspace rotation estimates of sinusoidal frequencies," *IEEE Trans. on Signal Processing,*, vol. 39, pp. 1836–1847, 1991.

[26] R. M. Gray, "Toeplitz and circulant matrices: A review,"," *Foundations and Trends in Communications and Information Theory*, vol. 2(3), pp. 155–239, 2006.

[27] J. M. Francos and B. Friedlander, "Bounds for estimation of complex exponentials in unknown colored noise," *IEEE Trans. Signal Process.*, vol. 43, pp. 2176–2185, 1995.

[28] P. Stoica, A. Jakobsson, and J. Li, "Cisoid parameter estimation in the colored noise case: asymptotic cramer-rao bound, maximum likelihood, and nonlinear least-squares," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 45, no. 8, pp. 2048–2059, 1997.
[29] H. Li, P. Stoica, and J. Li, "Compuationally efficient parameter estimation for harmonic sunusoidal signals," *Signal Processing 80*, pp. 1937–1944, 2000.
[30] M. G. Christensen and A. Jakobsson, *Multi-Pitch Estimation*, Synthesis Lectures on Speech and Audio Processing, 2009.



Johan Xi Zhang (S'06) received the M.Sc. degree in electrical engineering from Chalmers University of technology, Gothenburg, Sweden, in 2006, and a former Ph.D. student with the Department of Electronic Systems, Aalborg University, Denmark with expected graduation date 2010. He has been a visiting researcher at Electrical Engineering Department of Katholieke Universiteit Leuven, Leuven, Belgium. Currently, he is employed as researcher at Swedish Defence Research Agency, with the division Defence and Security Systems and Technology. His research interests include digital signal processing theory and methods with an application to acoustic signals, in particular parametric analysis, modeling, and source separation.



Mads Græsbøll Christensen (S'00–M'06) was born in Copenhagen, Denmark in March 1977. He received the M.Sc. and Ph.D. degrees in 2002 and 2005, respectively, from Aalborg University in Denmark, where he is also currently employed at the Department of Media Technology as Associate Professor. He was previously with the Department of Electronic Systems, Aalborg University and has been a visiting researcher at Philips Research Labs, Ecole Nationale Superieure des Telecommunications (ENST), and Columbia University. He has received several awards, namely an IEEE Int. Conf. Acoust. Speech, and Signal Proc. Student Paper Contest Award, the Spar Nord Foundation's Research Prize awarded anually for an excellent Ph.D. thesis, and a Danish Independent Research Council's Young Researcher's Award. He is author (with A. Jakobsson) of the book "Multi-Pitch Estimation", Morgan & Claypool Publishers, 2009. His research interests include digital signal processing theory and methods with a application to speech and audio, in particular parametric analysis, modeling, and coding of speech and audio signals.



Søren Holdt Jensen (S'87-M'88-SM'00) received the M.Sc. degree in electrical engineering from Aalborg University, Aalborg, Denmark, in 1988, and the Ph.D. degree in signal processing from the Technical University of Denmark, Lyngby, Denmark, in 1995. Before joining the Department of Electronic Systems of Aalborg University, he was with the Telecommunications Laboratory of Telecom Denmark, Ltd, Copenhagen, Denmark; the Electronics Institute of the Technical University of Denmark; the Scientific Computing Group of Danish Computing Center for Research and Education, Lyngby; the Electrical Engineering Department of Katholieke Universiteit Leuven, Leuven, Belgium; and the Center for PersonKommunikation (CPK) of Aalborg University. He is Full Professor and is currently heading a research team working in the area of numerical algorithms and signal processing for speech and audio processing, image and video processing, multimedia technologies, and digital communications.

Dr. Jensen was an Associate Editor for the IEEE Transactions on Signal Processing and is currently Member of the Editorial Board of Elsevier Signal Processing and the EURASIP Journal on Advances in Signal Processing. He is a recipient of an European Community Marie Curie Fellowship, former Chairman of the IEEE Denmark Section, and Founder and Chairman of the IEEE Denmark Section's Signal Processing Chapter.



Marc Moonen (M'94, SM'06, F'07) received the electrical engineering degree and the PhD degree in applied sciences from Katholieke Universiteit Leuven, Belgium, in 1986 and 1990 respectively.

Since 2004 he is a Full Professor at the Electrical Engineering Department of Katholieke Universiteit Leuven, where he is heading a research team working in the area of numerical algorithms and signal processing for digital communications, wireless communications, DSL and audio signal processing.

He received the 1994 K.U.Leuven Research Council Award, the 1997 Alcatel Bell (Belgium) Award (with Piet Vandaele), the 2004 Alcatel Bell (Belgium) Award (with Raphael Cendrillon), and was a 1997 "Laureate of the Belgium Royal Academy of Science". He received a journal best paper award from the IEEE Transactions on Signal Processing (with Geert Leus) and from Elsevier Signal Processing (with Simon Doclo).

He was chairman of the IEEE Benelux Signal Processing Chapter (1998-2002), and is currently Past-President of EURASIP (European Association for Signal Processing) and a member of the IEEE Signal Processing Society Technical Committee on Signal Processing for Communications.

He has served as Editor-in-Chief for the "EURASIP Journal on Applied Signal Processing" (2003-2005), and has been a member of the editorial board of "IEEE Transactions on Circuits and Systems II" (2002-2003) and "IEEE Signal Processing Magazine" (2003-2005) and "Integration, the VLSI Journal". He is currently a member of the editorial board of "EURASIP Journal on Applied Signal Processing", "EURASIP Journal on Wireless Communications and Networking", and "Signal Processing".