DAMPED CHIRP MIXTURE ESTIMATION VIA NONLINEAR BAYESIAN REGRESSION

Julian Neri *

SPCL, CIRMMT
McGill University, Montréal, Canada
julian.neri@mcgill.ca

Philippe Depalle *

SPCL, CIRMMT
McGill University, Montréal, Canada
philippe.depalle@mcgill.ca

Roland Badeau

LTCI, Télécom Paris
Institut Polytechnique de Paris, France
roland.badeau@telecom-paris.fr

ABSTRACT

Estimating mixtures of damped chirp sinusoids in noise is a problem that affects audio analysis, coding, and synthesis applications. Phase-based non-stationary parameter estimators assume that sinusoids can be resolved in the Fourier transform domain, whereas high-resolution methods estimate superimposed components with accuracy close to the theoretical limits, but only for sinusoids with constant frequencies. We present a new method for estimating the parameters of superimposed damped chirps that has an accuracy competitive with existing non-stationary estimators but also has a high-resolution like subspace techniques. After providing the analytical expression for a Gaussian-windowed damped chirp signal's Fourier transform, we propose an efficient variational EM algorithm for nonlinear Bayesian regression that jointly estimates the amplitudes, phases, frequencies, chirp rates, and decay rates of multiple non-stationary components that may be obfuscated under the same local maximum in the frequency spectrum. Quantitative results show that the new method not only has an estimation accuracy that is close to the Cramér-Rao bound, but also a high resolution that outperforms the state-of-the-art.

1. INTRODUCTION

Sinusoidal modeling is a primary topic in audio signal processing with many applications for audio analysis, coding, transformation, and re-synthesis. Rooted in additive synthesis and Fourier theory, sinusoidal modeling assumes that a signal is composed of a sum of sinusoidal oscillations and noise. Stationary models assume that the sinusoids have constant frequencies and amplitudes within the finite temporal window of analysis. But musical audio typically exhibits time-varying features, for example from the vibrato of a singing voice or attack of a plucked string. Estimating the parameters of non-stationary sinusoids can improve the quality and efficiency of the signal representation, enabling precise transformation and re-synthesis even when there are modulations within the short temporal analysis window. Indeed, non-stationary analysis has received much attention over the last several decades, as summarized in Section 1.1.

However, estimating mixtures of non-stationary sinusoids is still a difficult problem, especially in polyphonic music where superimposed non-stationary sinusoids with strong modulations cross in the time-frequency plane. Typically, non-stationary estimators assume that sinusoids do not overlap in the frequency domain and detect them by picking peaks one-by-one. After detection, a few spectral values around the peak are used to estimate first the non-constant parameters, and second the constant parameters. They do not usually consider the covariance between the sinusoids nor the contribution of negative frequencies to the Fourier transform that can bias their estimations. While subspace methods can detect overlapping components with high-resolution thanks to their measure of temporal covariance from an auto-correlation function, their performance is highly sensitive to the model order and they cannot estimate frequency modulations.

This paper addresses the gap between non-stationary and high-resolution estimators with a new method that jointly resolves and estimates mixtures of damped chirp sinusoids in noise. We present a variational algorithm for nonlinear Bayesian regression that fits the entire frequency spectrum to a weighted sum of spectral basis functions: the weights encode the amplitudes and phases, and the basis functions encode the frequencies, chirp rates, and decay rates of the damped chirps. As opposed to existing methods, this offers a high resolution solution because it infers a distribution that encodes the spectral covariance between the chirps and has high accuracy because it integrates information from the entire spectrum. The new method can be used with any analysis window: the spectrum of a windowed damped chirp is computed analytically when using a Gaussian window or numerically when using a non-Gaussian window. Accuracy and resolution experiments show the new method's high quality compared to the state-of-the-art and theoretical bounds.

1.1. Overview of Previous Work

Parameter estimation of a noisy chirp was proposed by Djuric et al. [1]. Zhou et al. [2] generalized the estimation problem to polynomial phase signals, which includes the damped chirp. The quadratically interpolated fast Fourier transform (QIFFT) [3] assumes that a Gaussian analysis window is used and estimates non-stationary parameters after fitting a parabola to the log-magnitude spectrum and the unwrapped phase spectrum. The reassignment method was initially proposed by Kodera et al. [4, 5], generalized to time-frequency analysis by Auger and Flandrin [6], and to estimate a non-stationary sinusoid with the generalized reassignment method (GRM) by Röbel [7] and Hainsworth [8]. Marchand and Depalle [9] developed a generalized derivative method (GDM) that uses a signal's first and second derivatives to estimate a non-stationary sinusoid's parameters. Betser [10] proposed a distribution derivative method that estimates the parameters of a general polynomial signal model by solving a set of linear equations formed from the signal and its derivatives. A comparison of these non-stationary parameter estimation methods was presented in [11]. Estimation of signal parameters via rotational invariance techniques (ESPRIT) is a high-resolution method for damped, stationary frequency sinusoids proposed by Roy and Kailath [12] and

* Thanks to the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant (RGPIN- 2018-05662) for funding.

Copyright: © 2021 Julian Neri et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 Unported License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

adapted to audio analysis by Badeau et al. [13], that has an estimation accuracy close to the theoretical bound.

1.2. Notation

- $\boldsymbol{x}^{\mathsf{T}}$: transpose of vector $\boldsymbol{x}$,
- $\bar{\boldsymbol{z}}$: complex conjugate of vector $\boldsymbol{z}$,
- $\boldsymbol{z}^{\mathsf{H}}$: conjugate (Hermitian) transpose of vector $\boldsymbol{z}$,
- $\Re\{\boldsymbol{z}\}$: real part of vector $\boldsymbol{z}$,
- $\langle z \rangle_{p(z)}$: expected value of z with respect to (w.r.t.) the probability distribution $p(z)$,
- $\boldsymbol{I}_M$: identity matrix of size $M \times M$,
- $\mathrm{Diag}(\boldsymbol{z})$: diagonal matrix formed from the elements of $\boldsymbol{z}$,
- $\mathrm{diag}(\boldsymbol{Z})$: vector formed from the diagonal entries of $\boldsymbol{Z}$,
- $\mathcal{N}_{\mathbb{F}}(\boldsymbol{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma})$: real (if $\mathbb{F} = \mathbb{R}$) or circular complex (if $\mathbb{F} = \mathbb{C}$) multivariate normal distribution over $\boldsymbol{x}$ with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$.

2. A DAMPED CHIRP AND ITS FOURIER TRANSFORM

A damped chirp is a sinusoidal oscillation with linear frequency modulation (FM) and exponential amplitude modulation (AM),

$$s(t) = \rho e^{-\alpha t} \cos\left(\theta + \omega_0 t + \frac{1}{2}\psi t^2\right) . \tag{1}$$

Its parameters include the amplitude ρ, the phase θ in radians, the decay rate α in log-amplitude per second, the angular frequency ω_0 in radians per second, and the chirp rate ψ in radians per second squared. The decay rate can be positive or negative to make an exponential decrease (damped) or increase (undamped), respectively.

A Gaussian window with scale $\beta > 0$ is defined as

$$w(t) = \exp\left(-\frac{t^2}{2\beta}\right) . \tag{2}$$

The Fourier transform $x(\omega) \in \mathbb{C}$ of a damped chirp signal has a closed-form solution when it is multiplied by the Gaussian window in equation (2):

$$x(\omega) = \int_{-\infty}^{+\infty} w(t)s(t)e^{-i\omega t}dt = f_\phi(\omega)v + \bar{f}_\phi(-\omega)\bar{v} , \tag{3}$$

where a coefficient $v \in \mathbb{C}$ encodes the amplitude and phase,

$$v = \rho e^{i\theta} , \tag{4}$$

and a spectral basis function $f_\phi(\omega) \in \mathbb{C}$ that has the parameter set $\phi = [\alpha, \omega_0, \psi]$ is defined as

$$f_\phi(\omega) = \sqrt{\pi g_\psi} \exp\left(g_\psi h^2_{\alpha,\omega_0}(\omega)\right) , \tag{5}$$

$$g_\psi = \frac{\beta}{2}\left(\frac{1 + i\beta\psi}{1 + \beta^2\psi^2}\right) , \tag{6}$$

$$h_{\alpha,\omega_0}(\omega) = \alpha + i(\omega - \omega_0) . \tag{7}$$

Since the damped chirp in equation (1) is real, it consists of an equal contribution of positive and negative frequency components, $f_\phi(\omega)v$ and $\bar{f}_\phi(-\omega)\bar{v}$, respectively.

In theory, a Gaussian window and a damped chirp both have infinite time support and are non-causal. Practically, since we process causal digital signals that have finite temporal support, we assume that the beginning of the sinusoidal component (the onset) occurs outside of the analysis window. This assumption is common to sinusoidal model estimators, which succumb to distortions in the frequency spectrum and possible estimation biases when the onset occurs within the analysis window.

This model is temporally local in two ways. First, we use a window of finite duration T that localizes the estimation around its center. Scale β may be set such that $w(\pm T/2) = \nu$, where $0 < \nu \ll 1$ is an acceptably small threshold, using the equation

$$\beta = -\frac{T^2}{8\ln(\nu)} . \tag{8}$$

Second, since real-world signals have evolving features, we are interested in using a local analysis window that slides through time, estimating parameters from a short snapshot of the signal. Perfect reconstruction is possible when the time between analysis windows, the hop length L, satisfies the following condition,

$$L \leq \frac{\sqrt{\pi\beta}}{\sqrt{2}} . \tag{9}$$

3. BAYESIAN REGRESSION MODEL

Now we turn to estimating the parameters of damped chirps given an audio signal, which we address as a nonlinear Bayesian regression problem. In nonlinear regression, the goal is to estimate the parameters of nonlinear functions, called basis functions, and regression coefficients (weights), such that the weighted sum of basis functions matches the data [14].

Data is assumed to be generated from a sum of M real-valued damped chirps plus zero-mean, normally-distributed noise. In the frequency domain, the noisy damped chirps model for complex Hermitian spectral data $x(\omega) \in \mathbb{C}$ is

$$x(\omega) = \eta(\omega) + \sum_{m=1}^{M} \left\{ f_{\boldsymbol{\phi}^{(m)}}(\omega)v_m + \bar{f}_{\boldsymbol{\phi}^{(m)}}(-\omega)\bar{v}_m \right\} , \tag{10}$$

$$\eta(\omega) \sim \mathcal{N}_{\mathbb{C}}(0, \sigma_x^2) , \tag{11}$$

where σ_x^2 is the variance of the Hermitian noise $\eta(\omega) \in \mathbb{C}$, v_m is the regression coefficient and $\boldsymbol{\phi}^{(m)} = [\alpha_m, \omega_{0,m}, \psi_m]$ is the parameter set of the mth damped chirp.

A real signal's negative and positive frequencies cause estimation bias and detection errors for existing estimators. Energy accumulation is prominent not only in lower frequencies where a component's bandwidth is likely greater than its center frequency, but also along the entire frequency range when a signal has significant chirp or decay rates. Previously, the Hilbert transform has been used to approximately remove the negative frequencies, and windows with low side-lobes were designed to reduce overlap. Instead, we enhance the estimator by explicitly modeling the contribution of both the positive and negative frequency components.

3.1. Discretization

In order to apply the concept of nonlinear Bayesian regression, we consider discretized data, basis functions, and regression coefficients, and express them in matrix form. Further, the parameters and coefficients are considered to be stochastic, random variables whose properties are described through their statistical distributions. In the following, the variables are normalized to make them independent of the sampling rate, *e.g.* ω is in radians per sample.

Data vector $\boldsymbol{x} \in \mathbb{C}^{N\times 1}$ contains N observations of a spectrum obtained by taking a zero-phased discrete Fourier transform (DFT)[1] of a truncated windowed signal with a duration of N samples, evaluated at frequencies $\omega_n = 2\pi(n-1)/N$,

$$\boldsymbol{x} = \begin{bmatrix} x(\omega_1) \\ \vdots \\ x(\omega_N) \end{bmatrix} . \tag{12}$$

Matrix $\overrightarrow{\boldsymbol{F}} \in \mathbb{C}^{N\times M}$ contains values of the basis function $f_{\boldsymbol{\phi}^{(m)}}(\omega_n)$ for each of the M damped chirps evaluated at the same N discrete frequencies of $\boldsymbol{x}$,

$$\overrightarrow{\boldsymbol{F}} = \begin{bmatrix} f_{\boldsymbol{\phi}^{(1)}}(\omega_1) & \dots & f_{\boldsymbol{\phi}^{(M)}}(\omega_1) \\ \vdots & \ddots & \vdots \\ f_{\boldsymbol{\phi}^{(1)}}(\omega_N) & \dots & f_{\boldsymbol{\phi}^{(M)}}(\omega_N) \end{bmatrix} . \tag{13}$$

Likewise, $\overleftarrow{\boldsymbol{F}} \in \mathbb{C}^{N\times M}$ contains values of $\bar{f}_{\boldsymbol{\phi}^{(m)}}(-\omega_n)$,

$$\overleftarrow{\boldsymbol{F}} = \begin{bmatrix} \bar{f}_{\boldsymbol{\phi}^{(1)}}(-\omega_1) & \dots & \bar{f}_{\boldsymbol{\phi}^{(M)}}(-\omega_1) \\ \vdots & \ddots & \vdots \\ \bar{f}_{\boldsymbol{\phi}^{(1)}}(-\omega_N) & \dots & \bar{f}_{\boldsymbol{\phi}^{(M)}}(-\omega_N) \end{bmatrix} . \tag{14}$$

Basis function parameters are concatenated in a $1\times 3M$ vector $\boldsymbol{\phi} = [\boldsymbol{\alpha}, \boldsymbol{\omega}_0, \boldsymbol{\psi}]$. The 1×3 subvector $\boldsymbol{\phi}^{(m)} = [\alpha_m, \omega_{0,m}, \psi_m]$ contains the mth component's parameter set.

Vector $\boldsymbol{v} \in \mathbb{C}^{M\times 1}$ contains M complex-valued regression coefficients,

$$\boldsymbol{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_M \end{bmatrix} = \begin{bmatrix} a_1 + ib_1 \\ \vdots \\ a_M + ib_M \end{bmatrix} = \boldsymbol{a} + i\boldsymbol{b} . \tag{15}$$

To simplify estimation and notation, we introduce the real-valued vector of variables $\boldsymbol{z} \in \mathbb{R}^{2M\times 1}$ that is formed from the vertical concatenation of $\boldsymbol{v}$'s real and imaginary parts,

$$\boldsymbol{z} = \begin{bmatrix} \boldsymbol{a} \\ \boldsymbol{b} \end{bmatrix} . \tag{16}$$

We define a $M \times 2M$ complex-valued matrix

$$\boldsymbol{C} = \begin{bmatrix} \boldsymbol{I}_M & i\boldsymbol{I}_M \end{bmatrix} , \tag{17}$$

so $\boldsymbol{v}$ and $\boldsymbol{z}$ are related through a linear transformation,

$$\boldsymbol{v} = \boldsymbol{C}\boldsymbol{z} . \tag{18}$$

Lastly, we define the composite matrix $\boldsymbol{Q} \in \mathbb{C}^{N\times 2M}$ as

$$\boldsymbol{Q} = \overrightarrow{\boldsymbol{F}}\boldsymbol{C} + \overleftarrow{\boldsymbol{F}}\bar{\boldsymbol{C}} . \tag{19}$$

3.2. Data likelihood

Following from equation (10), the likelihood of the data $\boldsymbol{x}$ given $\boldsymbol{z}$ and the entire parameter set $\boldsymbol{\phi}$ is

$$p(\boldsymbol{x}|\boldsymbol{z}, \boldsymbol{\phi}) = \mathcal{N}_{\mathbb{C}}(\boldsymbol{x}|\boldsymbol{Q}\boldsymbol{z}, \sigma_x^2\boldsymbol{I}_N) . \tag{20}$$

[1] A zero-phased DFT $\boldsymbol{x}$ is one where the center of the analysis is at time zero and is obtained by multiplying the linear DFT $\boldsymbol{y}$ by the factor $\exp(i\pi n)$: $x(\omega_n) = y(\omega_n)\exp(i\pi n)$ for $n = 0, \dots, N-1$. We center the transform at time zero to be consistent with equation (3).

3.3. Priors

The vector of regression coefficients $\boldsymbol{z}$ is given a zero-mean normal prior with diagonal $2M \times 2M$ precision (inverse covariance) matrix $\boldsymbol{\Lambda} = \mathrm{Diag}(\boldsymbol{\lambda})$ that is scaled by the data noise's level σ_x^2,

$$p(\boldsymbol{z}) = \prod_{j=1}^{2M} \mathcal{N}_{\mathbb{R}}(z_j|0, \lambda_j^{-1}\sigma_x^2) . \tag{21}$$

Estimation of $\boldsymbol{\lambda}$ is key to relevance vector regression [15], a method of finding sparse solutions to Bayesian regression problems that represents the data with only a few relevant components. An irrelevant component is trimmed from the model as its coefficient magnitude $|z_j|$ is pushed towards the prior mean of zero. Relevant components correspond to the underlying damped chirps, while irrelevant components correspond to spurious peaks from the noise-power spectral density estimate, secondary lobes of the window, and distortion.

The mth component's frequency $\omega_{0,m} > 0$ has a normal prior with variance $\tau_{\omega_0} = M^{-1}$ and mean uniformly spaced in the range $(0, \pi)$ rad/sample,

$$p(\boldsymbol{\omega}_0) = \prod_{m=1}^{M} \mathcal{N}_{\mathbb{R}}(\omega_{0,m}|\widetilde{\omega}_{0,m}, \tau_{\omega_0}) , \tag{22}$$

$$\widetilde{\omega}_{0,m} = \frac{\pi(m-1)}{M} . \tag{23}$$

Assuming a normal prior distribution (even for positive-valued variables) simplifies estimation because it is a conjugate distribution to the normal likelihood. The prior variance encourages the mth estimate to be close to $\widetilde{\omega}_{0,m}$. Further, our inclusion of both positive and negative components in the model enables accurate estimation around the zero frequency.

The mth component's chirp rate $\psi_m \in \mathbb{R}$ has a normal prior with mean $\widetilde{\psi}_m = 0$ and variance τ_{ψ_m},

$$p(\boldsymbol{\psi}) = \prod_{m=1}^{M} \mathcal{N}_{\mathbb{R}}(\psi_m|\widetilde{\psi}_m, \tau_{\psi_m}) , \tag{24}$$

$$\tau_{\psi_m} = \frac{1}{N}\left(\frac{\pi}{2} - \left|\frac{\pi}{2} - \widetilde{\omega}_{0,m}\right|\right) . \tag{25}$$

This prior encourages the chirp rate to be within $\pm\widetilde{\omega}_{0,m}/N$ because the instantaneous frequency, $\omega_{0,m} \pm \psi_m n$, should be greater than zero and less than the Nyquist frequency at $n \pm N/2$. For example, if $\widetilde{\omega}_{0,m}$ is 0 or π, then $\tau_{\psi_m} = 0$, and the estimate of ψ_m will be ≈ 0. If $\widetilde{\omega}_{0,m} = \frac{\pi}{2}$, then the chirp rate is likely in the range $-\frac{\pi}{2N} < \psi_m < \frac{\pi}{2N}$. This parametrization is opposed to directly relating the chirp rate to the frequency through a factor, which would complicate the estimation of both variables.

Lastly, the mth component's decay rate $\alpha_m \in \mathbb{R}$ has a normal prior with mean $\widetilde{\alpha}_m = 0$ and variance $\tau_\alpha = N^{-1}$,

$$p(\boldsymbol{\alpha}) = \prod_{m=1}^{M} \mathcal{N}_{\mathbb{R}}(\alpha_m|\widetilde{\alpha}_m, \tau_\alpha) . \tag{26}$$

3.4. Joint Distribution

The joint distribution is the product of the likelihood and priors,

$$p(\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{\phi}) = p(\boldsymbol{x}|\boldsymbol{z}, \boldsymbol{\phi})p(\boldsymbol{z}, \boldsymbol{\phi}) \tag{27}$$

$$= p(\boldsymbol{x}|\boldsymbol{\alpha}, \boldsymbol{\omega}_0, \boldsymbol{\psi}, \boldsymbol{z})p(\boldsymbol{\alpha})p(\boldsymbol{\omega}_0)p(\boldsymbol{\psi})p(\boldsymbol{z}) . \tag{28}$$

4. ESTIMATION

Applying Bayes' theorem gives the posterior distribution over the latent variables, $\boldsymbol{\phi}$ and $\boldsymbol{z}$, given the data,

$$p(\boldsymbol{z}, \boldsymbol{\phi}|\boldsymbol{x}) = \frac{p(\boldsymbol{x}|\boldsymbol{z}, \boldsymbol{\phi})p(\boldsymbol{z}, \boldsymbol{\phi})}{p(\boldsymbol{x})}\,, \tag{29}$$

$$p(\boldsymbol{x}) = \int\int p(\boldsymbol{x}|\boldsymbol{z}, \boldsymbol{\phi})p(\boldsymbol{z}, \boldsymbol{\phi})d\boldsymbol{z}d\boldsymbol{\phi}\,. \tag{30}$$

As with most non-trivial models, it is not possible to solve the integral for the model evidence $p(\boldsymbol{x})$ analytically.

Since exact inference is intractable for this model, we develop a variational inference algorithm for approximate inference of the posterior over the regression coefficients and the parameters,

$$p(\boldsymbol{z}, \boldsymbol{\phi}|\boldsymbol{x}) \approx q(\boldsymbol{z}, \boldsymbol{\phi})\,. \tag{31}$$

Variational inference circumvents the intractable integral involved in minimizing the Kullback-Leibler (KL) [16] divergence from q to p by instead maximizing the lower bound on model evidence [17, 18],

$$\mathcal{L}(q) = \langle \ln p(\boldsymbol{x}|\boldsymbol{z}, \boldsymbol{\phi})\rangle_{q(\boldsymbol{z},\boldsymbol{\phi})} - D_{KL}\left(q(\boldsymbol{z}, \boldsymbol{\phi})\|p(\boldsymbol{z}, \boldsymbol{\phi})\right)\,. \tag{32}$$

Approximate posterior q is factorized between the regression coefficients and basis function parameters,

$$q(\boldsymbol{z}, \boldsymbol{\phi}) = q_{\boldsymbol{z}}(\boldsymbol{z})q_{\boldsymbol{\phi}}(\boldsymbol{\phi})\,. \tag{33}$$

The optimal log distributions that maximize the lower bound are given by the calculus of variations,

$$\ln q^\star_{\boldsymbol{z}}(\boldsymbol{z}) = \langle \ln p(\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{\phi})\rangle_{q_{\boldsymbol{\phi}}(\boldsymbol{\phi})} + \text{constant}\,, \tag{34}$$

$$\ln q^\star_{\boldsymbol{\phi}}(\boldsymbol{\phi}) = \langle \ln p(\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{\phi})\rangle_{q_{\boldsymbol{z}}(\boldsymbol{z})} + \text{constant}\,. \tag{35}$$

Each distribution is updated in turn to maximize $\mathcal{L}(q)$.

4.1. Regression coefficients (amplitudes and phases)

The optimal approximate posterior $q^\star_{\boldsymbol{z}}(\boldsymbol{z})$ introduced in (34) is normal and has a closed-form solution,

$$q^\star_{\boldsymbol{z}}(\boldsymbol{z}) = \mathcal{N}_{\mathbb{R}}(\boldsymbol{z}|\boldsymbol{\mu}_z, \sigma_x^2\boldsymbol{\Sigma}_z)\,, \tag{36}$$

where the mean $\boldsymbol{\mu}_z$ and covariance matrix $\boldsymbol{\Sigma}_z$ are derived using the properties [19] of marginal and conditional normal distributions,

$$\boldsymbol{\mu}_z = \boldsymbol{\Sigma}_z\boldsymbol{Q}^{\mathsf{H}}\boldsymbol{x}\,, \tag{37}$$

$$\boldsymbol{\Sigma}_z = \left(\boldsymbol{\Lambda} + \boldsymbol{Q}^{\mathsf{H}}\boldsymbol{Q}\right)^{-1}\,. \tag{38}$$

Indeed, $\boldsymbol{\Lambda}$ regularizes the ill-posed problem when $\boldsymbol{Q}^{\mathsf{H}}\boldsymbol{Q}$ is ill-conditioned, which may happen when $|\alpha_m|$ is very large.

Amplitude and phase estimates are given by the absolute value and argument (angle) of $\boldsymbol{C}\boldsymbol{\mu}_z$, respectively,

$$\widehat{\boldsymbol{\rho}} = |\boldsymbol{C}\boldsymbol{\mu}_z|\,, \tag{39}$$

$$\widehat{\boldsymbol{\theta}} = \text{Arg}(\boldsymbol{C}\boldsymbol{\mu}_z)\,. \tag{40}$$

Maximizing the expected log-prior $\langle \ln p(\boldsymbol{z})\rangle_{q_{\boldsymbol{z}}}$ w.r.t. $\boldsymbol{\lambda}$ gives

$$\widehat{\boldsymbol{\lambda}} = \text{diag}\left(\frac{1}{\sigma_x^2}\boldsymbol{\mu}_z\boldsymbol{\mu}_z^{\mathsf{T}} + \boldsymbol{\Sigma}_z\right)^{-1}\,. \tag{41}$$

Algorithm 1 Damped Chirps Estimator

input: $\boldsymbol{x}$. **initialize:** $\boldsymbol{\phi}$, $\boldsymbol{\Lambda}$.
  repeat
    $\boldsymbol{Q} \leftarrow$ GETBASISFUNCTIONS($\boldsymbol{\phi}$)
    $\boldsymbol{\Sigma}_z \leftarrow (\boldsymbol{\Lambda} + \boldsymbol{Q}^{\mathsf{H}}\boldsymbol{Q})^{-1}$
    $\boldsymbol{\mu}_z \leftarrow \boldsymbol{\Sigma}_z\boldsymbol{Q}^{\mathsf{H}}\boldsymbol{x}$
    $\boldsymbol{\Lambda} \leftarrow \text{Diag}\left(\text{diag}(\frac{1}{\sigma_x^2}\boldsymbol{\mu}_z\boldsymbol{\mu}_z^{\mathsf{T}} + \boldsymbol{\Sigma}_z)^{-1}\right)$
    $\nabla_{\boldsymbol{\phi}}\langle \ln p(\boldsymbol{x}, \boldsymbol{\phi}, \boldsymbol{z})\rangle_{q^\star_{\boldsymbol{z}}} \leftarrow$ GETGRADIENT($\boldsymbol{\mu}_z$, $\boldsymbol{\Sigma}_z$,$\boldsymbol{\phi}$)
    $\boldsymbol{\phi} \leftarrow \boldsymbol{\phi} + \gamma\nabla_{\boldsymbol{\phi}}\langle \ln p(\boldsymbol{x}, \boldsymbol{\phi}, \boldsymbol{z})\rangle_{q^\star_{\boldsymbol{z}}}$
  until $\sum_j|\Delta_{\phi_j}| <$ threshold
output: $\boldsymbol{\phi}$, $\boldsymbol{\mu}_z$, $\boldsymbol{\Sigma}_z$.

As an element of $\boldsymbol{\mu}_z$ and its variance go towards zero, its corresponding precision will go to infinity. In turn, a large element in $\boldsymbol{\Lambda} = \text{Diag}(\widehat{\boldsymbol{\lambda}})$ causes the corresponding element in the estimate $\boldsymbol{\mu}_z$ to go to zero. Therefore, this variational estimation of the regression coefficient and precision leads to sparse solutions, since irrelevant components (ones with small magnitudes and small variances) are driven to zero, while relevant components with significant magnitudes are not altered. This is referred to as a relevance vector machine [15], or more generally, automatic relevancy determination [20].

4.2. Basis function parameters (frequencies, chirp rates, decay rates)

The optimal approximate posterior $q^\star_{\boldsymbol{\phi}}(\boldsymbol{\phi})$ introduced in equation (35) does not have a closed-form solution. However, with a Laplace approximation [21], $q_{\boldsymbol{\phi}}(\boldsymbol{\phi}) = \mathcal{N}_{\mathbb{R}}(\boldsymbol{\phi}|\widehat{\boldsymbol{\phi}}, \sigma_{\boldsymbol{\phi}}^2\boldsymbol{I}_{3M})$ for $\sigma_{\boldsymbol{\phi}}^2 \approx 0$, we can estimate the variational mode $\widehat{\boldsymbol{\phi}}$ by maximizing the right hand side of equation (35) w.r.t. $\boldsymbol{\phi}$.

Estimation of $\boldsymbol{\phi}$ is an unconstrained optimization problem [22] that can be addressed with gradient root finding. Gradient ascent performs well in this application because it is less sensitive to different initializations when compared to Newton's method but still converges quickly. The algorithm updates $\boldsymbol{\phi}$ as follows,

$$\boldsymbol{\phi} = \boldsymbol{\phi} + \gamma\nabla_{\boldsymbol{\phi}}\left(\langle \ln p(\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{\phi})\rangle_{q^\star_{\boldsymbol{z}}}\right)\,, \tag{42}$$

where $\nabla_{\boldsymbol{\phi}}$ denotes the gradient w.r.t. $\boldsymbol{\phi}$ and γ is the learning rate. Section 8.1 provides closed-form equations for the gradient of the expected log-joint, which is decomposed as

$$\nabla_{\boldsymbol{\phi}}\langle \ln p(\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{\phi})\rangle_{q^\star_{\boldsymbol{z}}} = \nabla_{\boldsymbol{\phi}}\left(\langle \ln p(\boldsymbol{x}|\boldsymbol{z}, \boldsymbol{\phi})\rangle_{q^\star_{\boldsymbol{z}}} + \ln p(\boldsymbol{\phi})\right)\,. \tag{43}$$

Pseudo-code for estimating damped chirps is presented in Algorithm 1. Its computational complexity is mainly from the $2M \times 2M$ matrix inverse in equation (38) to update $\boldsymbol{\mu}_z$, and the inner product in equation (50), Section 8.1, to update the gradient.

4.3. Expected marginal likelihood

After inference, an interesting statistic that is directly available using the sum rule of probability is the expected likelihood of the

data w.r.t. the inferred posterior over $\boldsymbol{z}$,

$$\langle p(\boldsymbol{x}|\boldsymbol{z},\boldsymbol{\phi})\rangle_{q^\star_{\boldsymbol{z}}(\boldsymbol{z})} = \int p(\boldsymbol{x}|\boldsymbol{z},\boldsymbol{\phi}) q^\star_{\boldsymbol{z}}(\boldsymbol{z}) d\boldsymbol{z}\,, \tag{44}$$

$$= \mathcal{N}_{\mathbb{C}}(\boldsymbol{x}|\boldsymbol{\mu}_x, \boldsymbol{\Sigma}_x)\,, \tag{45}$$

$$\boldsymbol{\mu}_x = \boldsymbol{Q}\boldsymbol{\mu}_z\,, \tag{46}$$

$$\boldsymbol{\Sigma}_x = \sigma_x^2\left(\boldsymbol{Q}\boldsymbol{\Sigma}_z\boldsymbol{Q}^{\mathsf{H}} + \boldsymbol{I}_N\right)\,. \tag{47}$$

The mean is a smoothed (de-noised) representation of the data, and the covariance measures uncertainty about the data.

4.4. Generalization to non-Gaussian windows

The proposed method can be used with *any* analysis window. While the Fourier transform of a damped chirp with a non-Gaussian window does not admit a closed-form solution, we can use the damped chirp's time domain basis function and derivatives given in Section 8.1, multiply them with a non-Gaussian window, then numerically compute the DFTs. This provides us with the spectral basis functions and its derivatives w.r.t. the parameters for any analysis window. The Hann window is often used for spectral analysis because it has a compact main lobe and attenuated side-lobes [23].

5. PRACTICAL EXPERIMENTS AND RESULTS

5.1. Accuracy experiments (single component)

Nonlinear Bayesian regression (NLR) was tasked with estimating the parameters of damped chirps across a range of noise levels. It was compared to the Cramér-Rao bound (CRB) defined in Section 8.2, and existing state-of-the-art non-stationary sinusoidal model estimators discussed in Section 1.1: GRM [6, 7], GDM [9], QIFFT [3], and ESPRIT [13].

For NLR, the spectral data was either obtained using a Gaussian window (NLR-G) or a Hann window (NLR-H). For NLR-G, the closed-form expressions for the spectrum and derivatives were used. The Gaussian window's scale was set using equation (8) with $\nu < 10^{-3}$. For NLR-H, the spectral basis function and its derivatives were computed numerically by taking DFTs of a Hann-windowed complex-valued damped chirp and its derivatives.

The duration of each test signal was $N = 512$ samples. The signal-to-noise ratio (SNR) in dB was

$$SNR = 10\log_{10}\left(\frac{\sum_{n=0}^{N-1} s_n^2}{\sum_{n=0}^{N-1}\zeta_n^2}\right)\,. \tag{48}$$

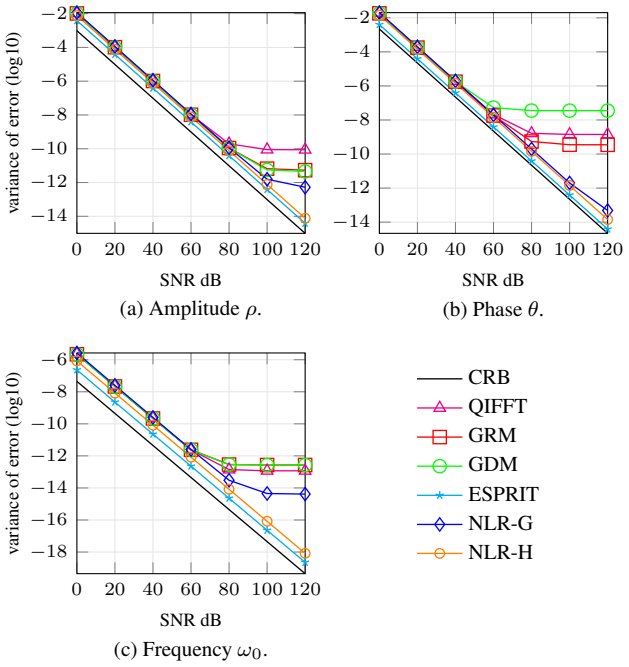
and went from 0dB to +120dB by steps of 20dB, where s_n is the signal and ζ_n is the zero-mean white noise of the nth time-domain sample. Since NLR does not have access to the true noise variance, we set $\sigma_x^2 = .1N$ constant for all tests.

For each SNR and each analysis method, the variance of error was evaluated from $R = 1000$ Monte Carlo runs. The variance of error was approximated by the equation

$$\text{var(error)} = \frac{1}{R}\sum_{r=1}^{R}\left(\xi^{(r)} - \widehat{\xi}^{(r)}\right)^2\,, \tag{49}$$

where $\xi^{(r)}$ was a target value and $\widehat{\xi}^{(r)}$ was an estimate from the rth Monte Carlo run. Monte Carlo run r involved randomly sampling from uniform distributions the phase $\theta^{(r)} \in (0, 2\pi)$, and frequency in the range $\omega_0^{(r)} \in (4\pi/N, \pi/4)$, as GRM, GDM, and QIFFT had problems with $\omega_0^{(r)} < 4\pi/N$. The chirp rate was either $\psi^{(r)} = 0$ (no FM) or randomly sampled in the range $\psi^{(r)} \in (-2\omega_0/N, 2\omega_0/N)$ (FM), so the instantaneous frequency could span $2\omega_0$ in N samples. Similarly, the decay rate was either $\alpha^{(r)} = 0$ (no AM) or randomly sampled $\alpha \in (-2/N, 2/N)$ (AM).

Figures 1 and 2 show the variance of error given stationary signals (no FM or AM), and non-stationary, damped chirp signals (FM and AM), respectively. ESPRIT is closest to the CRB for stationary signals, but does not estimate chirp rate and its quality degrades for damped chirp signals because it assumes that signals have stationary frequencies (no FM). NLR has similar accuracy to GRM, GDM, and QIFFT below 60dB and more accuracy above 60dB, remaining close to the CRB for all parameters. Compared to a Gaussian window spectrum, a Hann window's spectrum has a small main lobe and low side-lobes enabling better frequency resolution. Indeed, NLR-H was better than NLR-G across all SNRs.



(a) Amplitude ρ. (b) Phase θ. (c) Frequency ω_0.

Figure 1: Estimation of constant frequency and amplitude signals (no AM or FM).

5.2. Resolution experiments (multiple components)

Joint estimation of multiple frequency components in a noisy signal is a difficult problem. First, Figure 3 shows the result of using NLR-G and ESPRIT to estimate eight stationary components from a noisy signal of duration $N = 256$ that are close in frequency and where some share the same spectral peak. The proposed method demonstrates its high resolution, detecting obfuscated components and estimating their parameters. The number of components was set to $M = 32$, which allowed each component to be properly detected. Amplitudes of irrelevant components were automatically

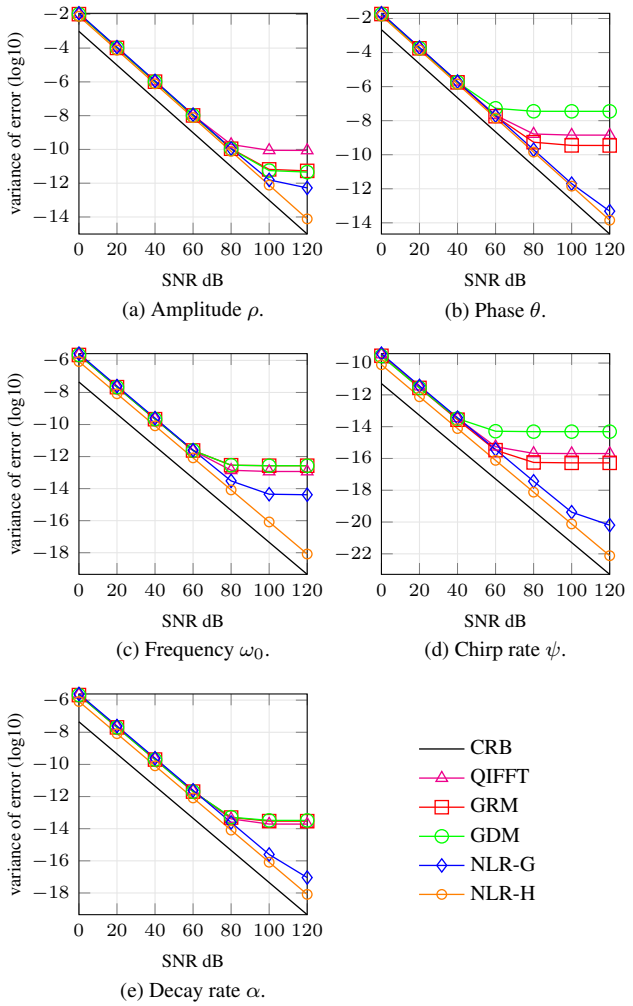


Figure 2: Estimation of damped chirp signals (AM and FM).

attenuated thanks to the sparsity inducing prior over the regression coefficients and the estimation of $\boldsymbol{\lambda}$. Irrelevant components were easily detected by checking for amplitudes lower than a threshold, *e.g.* 10^{-6}. ESPRIT with order 24 (12 components) had a similar resolution but was sensitive to the order, did not attenuate irrelevant components, and did not consider frequency modulation.

Figure 4 shows the results of estimating several damped chirp components from a noisy signal of duration $N = 256$. Fast chirps and decays created wide spectral lobes that caused significant phase cancellation artifacts in the magnitude spectrum, most noticeably at 4.3 kHz where a component's center frequency is above a valley. The proposed method resolved this difficult situation, estimating each component's chirp rate and decay rate, as shown in the bottom panels of Figure 4. GRM and GDM would not resolve these components, as they detect components by picking spectral peaks.

5.3. Short-term analysis with a sliding window

To analyze longer duration evolving sounds, we used a sliding Gaussian window of $N = 256$ duration with an $L = 256$ hop length. Since the analysis frames do not overlap in time, we can clearly show how the parameters evolve inside them. *Chirpograms* in Figures 5a and 5b show the short-term estimation of noisy (20dB SNR) signals with fast chirp rates, and two cosine FM components that cross in the time-frequency plane. When the

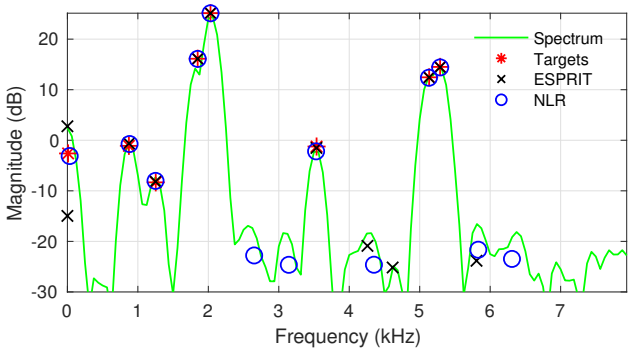


Figure 3: NLR and ESPRIT resolved close frequency components, even when the components were obfuscated under the same local maximum (peak) in the magnitude spectrum. Irrelevant components are not shown.

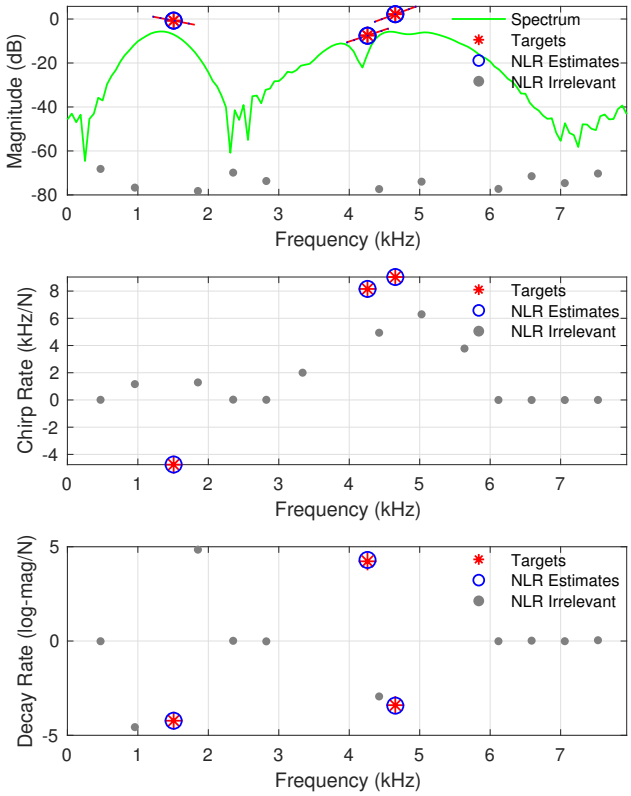
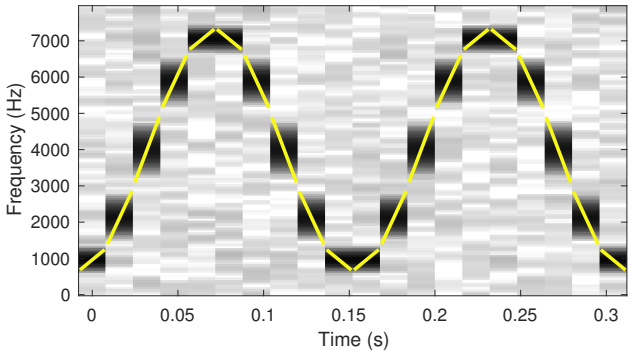
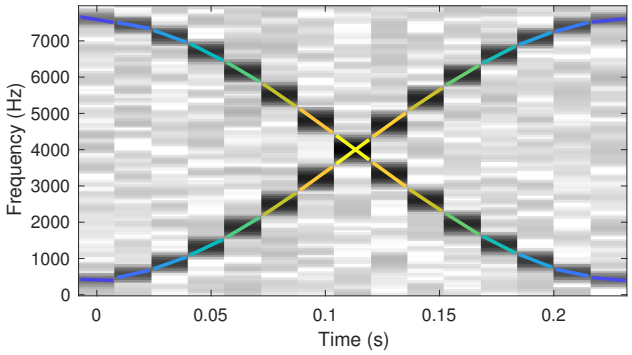


Figure 4: NLR detected multiple damped chirps from a noisy signal, even though they had fast chirp and decay rates (middle and bottom panels) and were obfuscated under the same peak in the magnitude spectrum (top panel, line slopes represent chirp rates). Components were classified as irrelevant if their estimated amplitudes were driven towards zero by the sparsity inducing prior.

components cross at time 0.12 seconds, their chirp rates have the same magnitude but are of opposite signs. NLR resolves the two components and shows their crossing even though they are under the same spectral peak at 4kHz. This quality is relevant for subsequent partial tracking [24] and additive synthesis, where incorrect detection can lead to distortions in the re-synthesized sounds.



(a) Fast chirp rate estimation from a noisy cosine FM sound.



(b) NLR accurately resolves and estimates the two cosine FM sounds from a noisy signal, even at 0.12 seconds where they share the same 4kHz spectral peak.

Figure 5: *Chirpograms*: a line segment's center, slope, and color, correspond to an estimated frequency, chirp rate, and amplitude, respectively. Chirp rate and window scale affect the sizes of the spectrogram lobes (dark parts of background image).

6. CONCLUSIONS

This paper presented a method for resolving and estimating the parameters of superimposed damped chirp sinusoids in noise. Nonlinear Bayesian regression in the frequency domain was addressed using an analytic expression for a Gaussian-windowed damped chirp's Fourier transform, and generalized to any window by computing DFTs of a damped chirp's temporal basis function and its derivatives. Explicitly modeling the additive contribution of a real-valued signal's negative and positive frequency components enables quality estimation over the entire spectrum. Its high resolution is attributed to the probabilistic modeling and inference of the non-stationary parameters and their covariances, given all the data available. It can resolve crossing partials with fast chirp rates that even share the same spectral peak, and has more accuracy than existing estimators for highly non-stationary signals. These qualities are especially relevant for analyzing, coding, and synthesizing vocal sounds and polyphonic music. Future work could accelerate the algorithm. While our structured variational approximation provided high resolution and accuracy, a mean-field factorization of the model's variables may significantly reduce the computational complexity of estimation without altering its quality. A second order optimization algorithm like Newton's method could improve convergence speed after a careful initialization. Alternatively, a variational auto-encoder could be used to map the data to approximate posterior means and variances of the model parameters.

7. REFERENCES

[1] P. M. Djurić and S. M. Kay, "Parameter estimation of chirp signals," *IEEE Trans. Acoustics, Speech, Signal Process.*, vol. 38, no. 12, pp. 2118–2126, Dec. 1990.

[2] G. Zhou, G. Giannakis, and A. Swami, "On polynomial phase signal with time-varying amplitudes," *IEEE Trans. Acoustics, Speech, Signal Process.*, vol. 44, no. 4, pp. 848–860, 1996.

[3] M. Abe and J. O. Smith, "AM/FM rate estimation for time-varying sinusoidal modeling," in *IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, 2005, vol. III, pp. 201–204.

[4] K. Kodera, C. de Villedary, and R. Gendrin, "A new method for the numerical analysis of time-varying signals with small BT values," *Phys. Earth Planet. Inter.*, vol. 12, no. 2-3, pp. 142–150, 1976.

[5] K. Kodera, R. Gendrin, and C. de Villedary, "Analysis of time-varying signals with small BT values," *IEEE Trans. Acoustics, Speech, Signal Process.*, vol. 32, no. 1, pp. 64–76, 1986.

[6] F. Auger and P. Flandrin, "Improving the readability of time-frequency and time-scale representations by the reassignment method," *IEEE Trans. Signal Process.*, vol. 43, no. 5, pp. 1068–1089, 1995.

[7] A. Röbel, "Estimating partial frequency and frequency slope using reassignment operators," in *International Computer Music Conference (ICMC)*, Sep. 2002, pp. 122–125.

[8] S. W. Hainsworth, *Techniques for the Automated Analysis of Musical Audio*, Ph.D. thesis, University of Cambridge, United Kingdom, 2003.

[9] S. Marchand and P. Depalle, "Generalization of the derivative analysis method to non-stationary sinusoidal modeling," in *Proc. of the 11th Int. Conf. on Digital Audio Effects (DAFx-08)*, September 2008.

[10] M. Betser, "Sinusoidal polynomial parameter estimation using the distribution derivative," *IEEE Transactions on Signal Processing*, vol. 57, no. 12, pp. 4633–4645, Dec. 2009.

[11] B. Hamilton and P. Depalle, "Comparisons of parameter estimation methods for an exponential polynomial sound signal model," in *AES 45th International Conference*, Mar. 2012.

[12] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoustics, Speech, Signal Process.*, vol. 37, no. 7, pp. 984–995, July 1989.

[13] R. Badeau, R. Boyer, and B. David, "EDS parametric modeling and tracking of audio signals," in *5th Int. Conf. on Digital Audio Effects (DAFx-02)*, 2002, pp. 139–144.

[14] D. MacKay, *Information Theory, Inference, and Learning Algorithms*, Cambridge University Press, 2003.

[15] M. E. Tipping, "The relevance vector machine," in *Advances in neural information processing systems*, 2000, pp. 652–658.

[16] S. Kullback and R. A. Leibler, "On information and sufficiency," *The Annals of Mathematical Statistics*, vol. 22, no. 1, pp. 79–86, Mar. 1951.

[17] D. Blei, A. Kucukelbir, and J. McAuliffe, "Variational inference: A review for statisticians," *Journal of the American Statistical Association*, vol. 112, no. 518, pp. 859–877, 2017.

[18] J. Neri, R. Badeau, and P. Depalle, "Probabilistic filter and smoother for variational inference of Bayesian linear dynamical systems," in *IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, May 2020, pp. 5885–5889.

[19] C. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.

[20] M. Beal, *Variational Algorithms for Approximate Bayesian Inference*, Ph.D. thesis, University College London, May 2003.

[21] K. Friston, J. Mattout, N. Trujillo-Barreto, J. Ashburner, and W. Penny, "Variational free energy and the Laplace approximation," *NeuroImage*, vol. 34, no. 1, pp. 220–234, 2006.

[22] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 7th edition, 2009.

[23] F. J. Harris, "On the use of windows for harmonic analysis with the discrete Fourier transform," in *Proc. IEEE*, January 1978, vol. 66, pp. 51–83.

[24] J. Neri and P. Depalle, "Fast partial tracking with real-time capability through linear programming," in *Proc. 21st Int. Conf. Digital Audio Effects (DAFx-18)*, Sep 2018, pp. 326–333.

8. APPENDIX

8.1. Gradients of the basis functions

The gradient of the log-likelihood w.r.t. $\boldsymbol{\phi}$ is

$$\nabla_{\boldsymbol{\phi}} \langle \ln p(\boldsymbol{x}|\boldsymbol{z},\boldsymbol{\phi}) \rangle_{q^\star_{\boldsymbol{z}}} = \frac{2}{\sigma_x^2} \Re \left\{ \widetilde{\boldsymbol{x}}^{\mathsf{H}} \left(\overrightarrow{\boldsymbol{F}} \odot \boldsymbol{y} + \overleftarrow{\boldsymbol{F}} \odot \bar{\boldsymbol{y}} \right) \right\} , \tag{50}$$

where $\odot$ denotes element-wise multiplication and we defined

$$\widetilde{\boldsymbol{x}} = \boldsymbol{x} - \boldsymbol{Q}\boldsymbol{\mu}_z \,, \tag{51}$$

$$\widehat{\boldsymbol{v}} = \boldsymbol{C} \left(\boldsymbol{\mu}_z + \mathrm{diag}(\boldsymbol{\Sigma}_z) \right) , \tag{52}$$

$$\boldsymbol{y} = \begin{bmatrix} \widehat{\boldsymbol{v}}^{\mathsf{T}} & \widehat{\boldsymbol{v}}^{\mathsf{T}} & \widehat{\boldsymbol{v}}^{\mathsf{T}} \end{bmatrix} \tag{53}$$

Second, the gradient of a log-prior w.r.t. $\boldsymbol{\phi}$ is

$$\nabla_{\boldsymbol{\phi}} \ln p(\boldsymbol{\phi}) = \left(\widetilde{\boldsymbol{\phi}} - \boldsymbol{\phi} \right) \mathrm{Diag}(\boldsymbol{\tau}_\phi)^{-1} , \tag{54}$$

$$\widetilde{\boldsymbol{\phi}} = \begin{bmatrix} \widetilde{\boldsymbol{\omega}}_0 & \widetilde{\boldsymbol{\psi}} & \widetilde{\boldsymbol{\alpha}} \end{bmatrix} , \tag{55}$$

$$\boldsymbol{\tau}_{\boldsymbol{\phi}} = \begin{bmatrix} \boldsymbol{\tau}_{\boldsymbol{\omega}_0} & \boldsymbol{\tau}_{\boldsymbol{\psi}} & \boldsymbol{\tau}_{\boldsymbol{\alpha}} \end{bmatrix} . \tag{56}$$

Considering equation (43), σ_x^2 gives more or less influence to the likelihood compared to the prior. We keep this value fixed regardless of the actual noise level, which we assume is unknown.

Matrix $\overrightarrow{\boldsymbol{F}} \in \mathbb{C}^{N \times 3M}$ contains the partial derivatives of the M spectral basis functions with respect to their three parameters,

$$\dot{\overrightarrow{\boldsymbol{F}}} = \begin{bmatrix} \partial \overrightarrow{\boldsymbol{F}} / \partial \phi_1 & \dots & \partial \overrightarrow{\boldsymbol{F}} / \partial \phi_{3M} \end{bmatrix} . \tag{57}$$

where $\partial \overrightarrow{\boldsymbol{F}} / \partial \phi_j$ is an $N \times 1$ vector because only one of the basis functions depends on $\phi_j, \forall j \in [1, 3M]$.

The partial derivatives of the Gaussian's spectral basis function w.r.t. each of the parameters are

$$\partial f_\phi(\omega) / \partial \alpha = 2 g_\psi h_{\alpha,\omega_0}(\omega) f_\phi(\omega) \,, \tag{58}$$

$$\partial f_\phi(\omega) / \partial \omega_0 = -i 2 g_\psi h_{\alpha,\omega_0}(\omega) f_\phi(\omega) \,, \tag{59}$$

$$\partial f_\phi(\omega) / \partial \psi = i g_\psi \left(1 + 2 g_\psi h^2_{\alpha,\omega_0}(\omega) \right) f_\phi(\omega) \,. \tag{60}$$

where the index m is not subscripted to simplify the equations.

For non-Gaussian windows, the spectral basis function and its derivatives are numerically computed with windowed DFTs of the temporal basis function $d_\phi(t)$ and its derivatives,

$$d_\phi(t) = \exp(\alpha t + i\omega_0 t + i \frac{1}{2} \psi t^2) \,, \tag{61}$$

$$\partial d_\phi(t) / \partial \alpha = t d_\phi(t) \,, \tag{62}$$

$$\partial d_\phi(t) / \partial \omega_0 = i t d_\phi(t) \,, \tag{63}$$

$$\partial d_\phi(t) / \partial \psi = i \frac{1}{2} t^2 d_\phi(t) \,. \tag{64}$$

8.2. Theoretical bounds

The Cramér-Rao bound (CRB) is defined as the best possible performance of an unbiased estimator in the presence of noise for a given dataset. For the damped chirp model in (1), the CRBs have been derived by Zhou et al. [2]. The CRB depends on the value

$$\epsilon_k = \sum_{n=0}^{N-1} \left(\frac{n - n_0}{N} \right)^k \exp\left(2\alpha \frac{n - n_0}{N} \right) , \quad k \geq 0 \,, \tag{65}$$

where $0 \leq n_0 < N$ is the time sample at which the parameters are estimated. Djurić and Kay [1] noted that the optimal choice in terms of the CRB is $n_0 = \frac{N}{2}$, *i.e.* the center of the analysis frame.

The Fischer information matrix (FIM) and CRB for the amplitude and decay rate are

$$\boldsymbol{J}_{\rho,\alpha} = \frac{2\rho^2}{\sigma_x^2} \begin{bmatrix} \epsilon_0 \frac{1}{\rho^2} & -\epsilon_1 \frac{1}{\rho} \\ -\epsilon_1 \frac{1}{\rho} & \epsilon_2 \end{bmatrix} , \tag{66}$$

$$\mathrm{CRB}\left(\begin{bmatrix} \rho \\ \alpha \end{bmatrix} \right) = \mathrm{diag}\left(\boldsymbol{J}^{-1}_{\rho,\alpha} \right) . \tag{67}$$

The FIM and CRB for the phase, frequency, and chirp rate are

$$\boldsymbol{J}_{\theta,\omega_0,\psi} = \frac{2\rho^2}{\sigma_x^2} \begin{bmatrix} \epsilon_0 & \epsilon_1 N & \epsilon_2 N^2 \\ \epsilon_1 N & \epsilon_2 N^2 & \epsilon_3 N^3 \\ \epsilon_2 N^2 & \epsilon_3 N^3 & \epsilon_4 N^4 \end{bmatrix} , \tag{68}$$

$$\mathrm{CRB}\left(\begin{bmatrix} \theta \\ \omega_0 \\ \psi \end{bmatrix} \right) = \mathrm{diag}\left(\boldsymbol{J}^{-1}_{\theta,\omega_0,\psi} \right) . \tag{69}$$