IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received September 1, 2019, accepted September 23, 2019, date of publication October 1, 2019, date of current version October 16, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2944788

Generalized Time-Updating Sparse Covariance-Based Spectral Estimation

YONGCHAO ZHANG[1], (Member, IEEE), ANDREAS JAKOBSSON[2], (Senior Member, IEEE), DEQING MAO[1], (Student Member, IEEE), YIN ZHANG[1], (Member, IEEE), YULIN HUANG[1], (Senior Member, IEEE), AND JIANYU YANG[1], (Member, IEEE)

[1]School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China
[2]Centre for Mathematical Sciences, Mathematical Statistics, Lund University, SE-221 00 Lund, Sweden

Corresponding author: Yin Zhang (yinzhang@uestc.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61901092 and Grant 61671117, in part by the project through the China Post-Doctoral Science Foundation, in part by the Collaborative Innovation Center of Information Sensing and Understanding, and in part by the Swedish Research Council.

ABSTRACT Recently, the time-updating q-norm sparse covariance-based estimator (q-SPICE) was developed for online spectral estimation of stationary signals. In this work, this development is furthered to deal with non-stationary signals. By introducing a weighting matrix defined by a forgetting factor, the generalized least absolute shrinkage and selection operator (LASSO) is generalized, in order to allow for changes in the spectral content. As shown here, the resulting LASSO formulation can be solved in a simple manner using cyclic minimization, enabling recursive estimation for non-stationary signals. The proposed generalized time-updating q-SPICE offers the same benefits as the original estimator, including being computationally efficient at constant computational and storage cost, but also allows for substantial improvements when dealing with non-stationary signals. The performance of the method is evaluated using both stationary and non-stationary signals, indicating the preferable performance of the generalized formulation as compared to the original time-updating SPICE algorithm.

INDEX TERMS Time-updating, sparse covariance-based spectral estimation, non-stationary signals.

I. INTRODUCTION

The problem of estimating the spectral content of periodic or quasi-periodic signals occurs in a wide variety of applications, such as in audio and speech processing, biological signal processing, and radar imaging (see e.g. [1]–[3]). In applications requiring the resulting estimate to have higher resolution than the periodogram, one often uses parametric or data-adaptive non-parametric estimation techniques. Generally, parametric approaches rely heavily on accurate a priori information of both the model structure and the model order of the assumed signal. This is a major drawback as such information is typically difficult to assess in practical situations [4]. Non-parametric approaches on the other hand are inherently robust to model assumptions, but in turn suffers from low resolution [5]. To allow for the strengths of both approaches, there has recently been significant interest in semi-parametric approaches, which makes some weak model structure assumptions, such as the vector of parameters consisting of a few dominant or nonzero elements, but does not assume detailed information of the model order. Such formulations have been found to allow for high resolution estimates by exploiting sparse reconstruction approaches [6], although typically do so by relying on one or more hyperparameters. It is often a non-trivial task to select such hyperparameters properly, although there is some work indicating that one may formulate selection algorithms for this (see e.g. [7], and the references therein).

The sparse iterative covariance-based estimator (SPICE) is a hyperparameter free semi-parametric approach that was proposed in [4]. The method is globally convergent, and has been shown to outperform several dense and sparse estimators, yielding both superior resolution and low side-lobe levels. In view of its excellent performance, several works have examined and extended upon the framework [8]–[16], and it has been shown that SPICE is a weighted version of the square-root least absolute shrinkage and selection operator (LASSO) [10]–[12], with the selected weight

The associate editor coordinating the review of this manuscript and approving it for publication was Yue Zhang.

being close to optimal. In [13], the SPICE algorithm was generalized by introducing separate penalties on the signal and noise terms, yielding the so called q-SPICE algorithm, which was found to substantially improve the resulting estimates.

Regrettably, the brute force implementation of SPICE is computationally cumbersome. To mitigate this, several works have examined efficient implementations for the SPICE framework. Inspired by previous fast implementations for the Capon estimator, the amplitude and phase estimation (APES), and the iterative adaptive approach (IAA) [18]–[26], a fast SPICE algorithm based on Gohberg-Semencul (GS) factorization was proposed in [17], exploiting the inherent Fourier structure of the dictionary and the Toeplitz structure of the covariance matrix. However, such an implementation neglects the inherent sparsity of the parameter space. In [15], a wideband SPICE algorithm was developed by introducing the use of integrated dictionary elements spanning bands of the considered parameter space, which enables a reliable sparse signal reconstruction at a much lower computational cost. Both these fast implementations consider the case of batch processing, where the entire signal is processed simultaneously.

In a recent effort, a time-updating SPICE was proposed, offering the possibility of online spectral estimation [16]. The time-updating SPICE is capable of recursively updating and refining the estimates by each obtained sample, thereby allowing for real-time processing, for instance, when the data is obtained as a stream of measurement. This allows the algorithm to scale well with the growing size of the data, resulting in an overall reduction of both the computational time and the storage cost in memory. In [27], the online implementation was extended to also allow for the q-SPICE algorithm. Both these implementations assume stationary signals. When dealing with non-stationary signals, however, they are incapable of recovering the evolution of the time-varying spectral content, as all samples are given equal importance. Different from stationary signals, the spectral content of non-stationary signals varies with time. Therefore, new samples may bring new information on the current spectral content, and it is therefore desirable to weight more recent samples more than those of the distant past. Following this train of thought, and to overcome the shortcoming of the current time-updating SPICE algorithm, this paper proposes a generalized time-updating q-SPICE algorithm that forms an online estimate of the spectrum for non-stationary signals by introducing a forgetting factor in the updating of the estimate. The resulting algorithm offers the same benefits as the original time-updating q-SPICE, including full control over the sparsity in the estimate in a data adaptive and noise independent manner, while allowing for a simple implementation using a cyclic minimizing of an equivalent LASSO problem, enabling an implementation that is computationally efficient with a constant computational and storage cost, even for cases when processing large amounts of data. The choice of the forgetting factor allows for changes in the spectral content.

The remainder of this paper is organized as follows: in section II, we reviews the q-SPICE algorithm, and formulate the problem of interest. Then, in Section III, the proposed generalized time-updating q-SPICE is detailed. Numerical examples are provided in section IV, illustrating the performance of the proposed method on both stationary and non-stationary signals. Finally, section V contains our conclusions.

Notation: We denote vectors and matrices by boldface letters. The kth component of a vector $\mathbf{u}$ is written as u_k. Further symbols are summarized in Table 1.

TABLE 1. Notation.

$(\cdot)^H$	conjugate transpose of a vector or matrix
$(\cdot)^T$	transpose of a vector or matrix
$(\cdot)^*$	conjugate of each entry of a vector or matrix
$(\cdot)^{-1}$	inverse of a square matrix
$\|\cdot\|_q$	q-norm of a vector, $q > 0$
$\|\cdot\|_{\mathrm{F}}$	Frobenius norm of a matrix
$[\cdot]_k$	the kth column of a matrix

II. BACKGROUND

Let $y(t)$ denote the signal of interest. Under the assumption that $y(t)$ contains an unknown number of sinusoidal components, one may represent the signal as

$$\mathbf{y} = \mathbf{F}\boldsymbol{\theta} + \mathbf{e} \tag{1}$$

where $\mathbf{y} = [y(t_1), y(t_2), \ldots, y(t_M)]^T$ is an $M \times 1$ vector containing a sequence of measurements sampled at (the possibly non-uniform sampling) times $t_1, t_2, \ldots, t_M$, and $\mathbf{F}$ a known $M \times K$ regressor matrix containing a dictionary of potential candidate sinusoids, such that

$$\mathbf{F} = \begin{bmatrix} \mathbf{f}_1 & \mathbf{f}_2 & \cdots & \mathbf{f}_K \end{bmatrix} = \begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \\ \vdots \\ \mathbf{h}_M \end{bmatrix} \tag{2}$$

with

$$\mathbf{f}_k = \begin{bmatrix} 1 & e^{j\omega_k} & \cdots & e^{j\omega_k(M-1)} \end{bmatrix}^T \tag{3}$$

$$\omega_k = \frac{2\pi}{K}(k-1) \tag{4}$$

for $k = 1, 2, \ldots, K$, where K denotes the number of candidate frequencies, which is assumed to be selected sufficiently large to allow some of the candidate pitches to approximate the true frequencies reasonably well. Furthermore, $\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots, \theta_K]^T$ denotes the (unknown) complex-valued amplitudes of the candidate elements, with only a few elements assumed to be non-zero, whereas $\mathbf{e} = [e(t_1), e(t_2), \ldots, e(t_M)]^T$ denotes an additive noise vector containing the background noise and any non-periodic signal components.

The problem of interest may therefore be expressed as a sparse estimation problem, wherein both the amplitudes and the number of non-zero amplitudes are sought.

A. SPICE

A now classical approach to express this problem as a sparse reconstruction problem [6]

$$\underset{\boldsymbol{\theta}}{\text{minimize}} \ \|\mathbf{y} - \mathbf{F}\boldsymbol{\theta}\|_2^2 + \mu\|\boldsymbol{\theta}\|_1 \tag{5}$$

where μ is a user set regularization parameter dictating the tradeoff between the fit of the signal and the sparsity of the solution. Although typically easier than the model order estimation problem, it is often a non-trivial task to select μ suitably. An interesting alternative was introduced in [4], where the authors proposed a novel sparse reconstruction technique based on a covariance fitting criteria that avoids the selection of any user parameters. The resulting minimization criteria is then formed as

$$\underset{p_k \geq 0}{\text{minimize}} \ \left\|\mathbf{R}^{-1/2}\left(\mathbf{y}\mathbf{y}^H - \mathbf{R}\right)\right\|_{\mathbf{F}}^2 \tag{6}$$

where

$$\mathbf{R} = \mathbf{A}\mathbf{P}\mathbf{A}^H \tag{7}$$
$$\mathbf{A} = \begin{bmatrix}\mathbf{F} & \mathbf{I}\end{bmatrix} \tag{8}$$
$$\triangleq [\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_{M+K}] \tag{9}$$
$$\mathbf{P} = \operatorname{diag}(\mathbf{p}) \tag{10}$$
$$\mathbf{p} = \left[|\theta_1|^2, |\theta_2|^2, \ldots, |\theta_K|^2, \sigma_1^2, \sigma_2^2, \ldots, \sigma_M^2\right]^T \tag{11}$$
$$\triangleq [p_1, p_2, \ldots, p_{M+K}]^T \tag{12}$$

with $\mathbf{R}^{-1/2}$ denoting the Hermitian positive definite square root of $\mathbf{R}^{-1}$, $\mathbf{I}$ the $M \times M$ identity matrix, diag $(\cdot)$ a diagonal matrix formed by the specified vector, and σ_m^2 the noise variance for sample m, for $m = 1, \ldots, M$. As shown in [4], minimizing (6) is equivalent to

$$\underset{p_k \geq 0}{\text{minimize}} \ \mathbf{y}^H\mathbf{R}^{-1}\mathbf{y} + \|\mathbf{W}\mathbf{p}\|_1 \tag{13}$$

where

$$\mathbf{W} = \operatorname{diag}([w_1, w_2, \ldots, w_{M+K}]) \tag{14}$$
$$w_k = \frac{\|\mathbf{a}_k\|^2}{\|\mathbf{y}\|^2}, \quad k = 1, 2, \ldots M + K \tag{15}$$

It is worth noting that as (13) minimizes a signal fitting criteria, which measures the distance through the inverse of the (model) covariance matrix, no explicit regularization is required.

B. q-SPICE

The constraint in (13) is a weighted 1-norm, enabling a sparse solution to be formed. However, as it is formulated in (13), both the signal and the noise components are penalized jointly. Consequently, the minimization will also impose a sparse solution on the noise variances, causing the solution to be non-sparse in order to ensure that the resulting covariance matrix is invertible. In order to remedy this drawback, the generalized SPICE modifies the criterion in (13) as [13]

$$\underset{p_k \geq 0}{\text{minimize}} \ \mathbf{y}^H\mathbf{R}^{-1}\mathbf{y} + \|\mathbf{W}_\text{s}\mathbf{p}_\text{s}\|_1 + \|\mathbf{W}_\text{n}\mathbf{p}_\text{n}\|_q \tag{16}$$

where

$$\mathbf{p}_\text{s} = [p_1, p_2, \ldots, p_K]^T \tag{17}$$
$$\mathbf{p}_\text{n} = [p_{K+1}, p_{K+2}, \ldots, p_{M+K}]^T \tag{18}$$
$$\mathbf{W}_\text{s} = \operatorname{diag}([w_1, w_2, \ldots, w_K]) \tag{19}$$
$$\mathbf{W}_\text{n} = \operatorname{diag}([w_{K+1}, w_{K+2}, \ldots, w_{M+K}]) \tag{20}$$

with the choice of q allowing for direct control of the sparseness for the dense estimate on the noise variance. Thus, the original SPICE is obtained when $q = 1$. One may increase the norm for the second term in (16) to control the robustness of the estimate, as discussed in [13].

C. PROBLEM FORMULATION

For real-time processing, the batch mode of the current implementations of the q-SPICE is not suitable, as the computational complexity and storage complexity quickly become cumbersome with increasing data size. Furthermore, for non-stationary signals, the current batch implementations cannot track frequency changes in the spectral content.

In order to do so, and to alleviate the cumbersome nature of the batch implementations, one may preferably implement the optimization recursively.

Over past decades, approaches designed for online processing, coupled with spectral analysis of non-stationary signals have been well developed, with examples including the classical recursive least squares (RLS), and the recently developed time-recursive IAA (see e.g. [18], [19], [28], [29]). Many of these methods are filterbank-based, assuming that the input signal is stationary during a short time, and then estimate the spectral content using a set of sliding-windowed data. The windowing procedure may provide the ability for tracking the varying spectral content, but results in a degraded resolution and high sidelobe levels.

In this paper, based on the inherent connection between the q-SPICE and the LASSO, we instead propose a time-updating implementation of the q-SPICE, which allowing for the forming of a recursive estimate in an online fashion.

III. GENERALIZED TIME-UPDATING q-SPICE

It has been proved that q-SPICE is equivalent to solving a weighted hyperparameter-free square-root LASSO problem [13]

$$\underset{\boldsymbol{\theta}}{\text{minimize}} \ \|\mathbf{y}-\mathbf{F}\boldsymbol{\theta}\|_2+\|\mathbf{D}\boldsymbol{\theta}\|_1 \tag{21}$$

where

$$\mathbf{D}=\operatorname{diag}\left(\sqrt{\frac{\|\mathbf{f}_1\|_2^2}{M^{1/q}}},\sqrt{\frac{\|\mathbf{f}_2\|_2^2}{M^{1/q}}},\ldots,\sqrt{\frac{\|\mathbf{f}_K\|_2^2}{M^{1/q}}}\right) \tag{22}$$

Based on this equivalence, the time-updating q-SPICE implementation has recently been developed in [16], [27], although this implementation is incapable of also tracking time-varying spectral content.

For non-stationary signals that contain time-varying frequencies, the goal is to recover the evolution of the spectral content as it varies over time. As new samples bring more information on the current spectral contents, it is desired to weight the more recent samples more than the earlier samples. Following this train of thought, to allow for changes in the spectral content, we here introduce a forgetting factor, $0 \leq \lambda \leq 1$, in the equivalent formulation in (21), such that

$$\underset{\boldsymbol{\theta}}{\text{minimize}} \ \|\Lambda(\mathbf{y}-\mathbf{F}\boldsymbol{\theta})\|_2+\|\mathbf{D}_1\boldsymbol{\theta}\|_1 \tag{23}$$

where $\Lambda=\operatorname{diag}\left(\left[\lambda^{M-1},\lambda^{M-2},\ldots,1\right]\right)$, which gives older samples less importance than newer samples, and with

$$\mathbf{D}_1=\operatorname{diag}\left(\sqrt{\frac{\|\Lambda\mathbf{f}_1\|_2^2}{M^{1/q}}},\sqrt{\frac{\|\Lambda\mathbf{f}_2\|_2^2}{M^{1/q}}},\ldots,\sqrt{\frac{\|\Lambda\mathbf{f}_K\|_2^2}{M^{1/q}}}\right) \tag{24}$$

We proceed to show that the resulting generalized LASSO problem in (23) can be solved in a simple manner using cyclic minimization.

A. CYCLIC MINIMIZATION

With respect to one component θ_k, for $k=1,\ldots,K$, the cost function in (23) may be re-written as

$$J(\theta_k)=\left[\|\Lambda(\mathbf{x}_k-\mathbf{f}_k\theta_k)\|_2^2\right]^{1/2}+d_{kk}|\theta_k|+N_k \tag{25}$$

where

$$\mathbf{x}_k=\mathbf{y}-\sum_{i=1,i\neq k}^{K}\mathbf{f}_i\theta_i \tag{26}$$

with d_{kk} denoting the kth diagonal element of $\mathbf{D}_1$, and

$$N_k=\sum_{i=1,i\neq k}^{K}d_{ii}|\theta_i| \tag{27}$$

is a constant independent of θ_k. Let $\theta_k=r_ke^{j\varphi_k}$, where $r_k>0$, and $\varphi_k\in[-\pi,\pi)$. Following the derivation in [16], [27], the quadratic term in (25) may be reformulated using polar coordinates, such that

$$\begin{aligned}\psi &= \|\Lambda(\boldsymbol{x}_k-\mathbf{f}_k\theta_k)\|_2^2\\ &= \left\|\Lambda\left(\mathbf{x}_k-\mathbf{f}_kr_ke^{j\varphi_k}\right)\right\|_2^2\\ &= \|\Lambda\mathbf{x}_k\|_2^2+\left\|\Lambda\mathbf{f}_kr_ke^{j\varphi_k}\right\|_2^2\\ &\quad -2\operatorname{Re}\left\{r_k\mathbf{f}_k^H\Lambda^H\Lambda\mathbf{x}_ke^{-j\varphi_k}\right\}\\ &= \|\Lambda\mathbf{x}_k\|_2^2+\|\Lambda\mathbf{f}_k\|_2^2r_k^2\\ &\quad -2r_k\left|\mathbf{f}_k^H\Lambda^H\Lambda\mathbf{x}_k\right|\cos\left[\arg\left(\mathbf{f}_k^H\Lambda^H\Lambda\mathbf{x}_k\right)-\varphi_k\right]\end{aligned} \tag{28}$$

yielding

$$\begin{aligned}J(\theta_k)&=J(r_k,\varphi_k)\\ &=\psi^{1/2}+d_{kk}r_i\end{aligned} \tag{29}$$

where the constant N_k has been omitted in the interest brevity. The minimizing φ_k is simply

$$\hat{\varphi}_k=\arg\left(\mathbf{f}_k^H\Lambda^H\Lambda\mathbf{x}_k\right) \tag{30}$$

Let

$$\begin{aligned}\alpha_k&=\|\Lambda\mathbf{x}_k\|_2^2\\ \beta_k&=\|\Lambda\mathbf{f}_k\|_2^2\\ \gamma_k&=\left|\mathbf{f}_k^H\Lambda^H\Lambda\mathbf{x}_k\right|\end{aligned} \tag{31}$$

such that (25) may be expressed as

$$J(r_k,\varphi_k)=\left(\alpha_k+\beta_kr_k^2-2\gamma_kr_k\right)^{1/2}+d_{ii}r_i \tag{32}$$

According to the Cauchy-Schwarz inequality, we have

$$\alpha_k\beta_k-r_k^2\geq 0 \tag{33}$$

Following the derivations in [16], [27], the cost function (25) is convex with respect to r_k. Consequently, the optimal solution to minimize (25) can be then expressed as [16]

$$\hat{\theta}_k=\begin{cases}\hat{r}_ke^{j\hat{\varphi}_k}, & \text{if } \sqrt{M^{1/q}-1}\gamma_k>\sqrt{\alpha_k\beta_k-\gamma_k^2}\\ 0, & \text{else}\end{cases} \tag{34}$$

where

$$\hat{r}_k=\frac{\gamma_k}{\beta_k}-\frac{1}{\beta_k}\left(\frac{\alpha_k\beta_k-\gamma_k^2}{M^{1/q}-1}\right)^{1/2} \tag{35}$$

Thus, the estimate of $\hat{\theta}_k$ relies on $\mathbf{x}_k$, which depends on the unknown $\hat{\theta}_i$, $i=1,\ldots,K$, $i\neq k$ that are also to be estimated. To handle this, one may set initial values on $\hat{\theta}_i$, $i=1,\ldots,K$, and then iteratively update the estimation on $\hat{\theta}_k$ while holding the remaining elements $\hat{\theta}_i$, $i=1,\ldots,K$, $i\neq k$ constant, until convergence. Because the function shown by (25) is convex, such an iterative process is global convergent. Practically, one may instead simply stop the iteration after L iterations. The behaviour of the iterative process has been investigated in [16], where only a small number of iterations was shown to be required.

TABLE 2. Generalized time-updating *q*-SPICE.

Select the forgetting factor λ and the norm q
Initialize:
    $\boldsymbol{\Gamma}(0) = \mathbf{0}$
    $\boldsymbol{\rho}(0) = \mathbf{0}$
    $\kappa(0) = 0$
    $\tilde{\boldsymbol{\theta}}(0) = \mathbf{0}$
for $m = 1, 2, \ldots, M$
    Input: $y(t_m)$, $\mathbf{h}_m$
    $\boldsymbol{\Gamma}(m) \leftarrow \lambda^2 \boldsymbol{\Gamma}(m-1) + \mathbf{h}_m^H \mathbf{h}_m$
    $\boldsymbol{\rho}(m) \leftarrow \lambda^2 \boldsymbol{\rho}(m-1) + \mathbf{h}_m^H y(t_m)$
    $\kappa(m) \leftarrow \lambda^2 \kappa(m-1) + |y(t_m)|^2$
    $\eta = \kappa(m) + \tilde{\boldsymbol{\theta}}^H \boldsymbol{\Gamma}(m) \tilde{\boldsymbol{\theta}} - 2\mathrm{Re}\left\{\tilde{\boldsymbol{\theta}}^H \boldsymbol{\rho}(m)\right\}$
    $\boldsymbol{\zeta} = \boldsymbol{\rho}(m) - \boldsymbol{\Gamma}(m)\tilde{\boldsymbol{\theta}}$
    for $l = 1, \ldots, L$ **do**
        for $k = 1, \ldots, K$ **do**
            $\alpha_k = \eta + \Gamma_{kk}\left|\tilde{\theta}_k\right|^2 + 2\mathrm{Re}\left\{\tilde{\theta}_k^* \zeta_k\right\}$
            $\beta_k = \Gamma_{kk}$
            $\gamma_k = \left|\zeta_k + \Gamma_{kk}\tilde{\theta}_k\right|$
            if $\sqrt{m^{1/q} - 1}\gamma_k > \sqrt{\alpha_k \beta_k - \gamma_k^2}$ **then**
                $\varphi_k = \arg\left(\zeta_k + \Gamma_{kk}\tilde{\theta}_k\right)$
                $\hat{\theta}_k = r_k e^{j\varphi_k}$
            else
                $\hat{\theta}_k = 0$
            end if
            $\eta \leftarrow \eta + \Gamma_{kk}\left|\tilde{\theta}_k - \hat{\theta}_k\right|^2 + 2\mathrm{Re}\left\{\left(\tilde{\theta}_k - \hat{\theta}_k\right)^* \zeta_k\right\}$
            $\boldsymbol{\zeta} \leftarrow \boldsymbol{\zeta} + [\boldsymbol{\Gamma}(m)]_k \left(\tilde{\theta}_k - \hat{\theta}_k\right)$
            $\tilde{\theta}_k \leftarrow \hat{\theta}_k$
        end for
    end for
    Output: $\hat{\boldsymbol{\theta}}$
end for

B. ONLINE FORMULATION

In the case of online processing, the length of the input data, M, is incrementally increased over the observation time. To explicitly derive the recursive estimate, unless otherwise specified, the here used symbols refer to the measurements, temporary variants, or estimates at time sample M.

Given the current estimate denoted by $\tilde{\boldsymbol{\theta}}$, let

$$\mathbf{z} = \mathbf{y} - \mathbf{F}\tilde{\boldsymbol{\theta}} \tag{36}$$

denote an auxiliary variable. Then,

$$\mathbf{x}_k = \mathbf{z} + \mathbf{f}_k \tilde{\theta}_k \tag{37}$$

Therefore, the variables in (31) may be expressed as

$$\begin{aligned} \alpha_k &= \|\Lambda \mathbf{x}_k\|_2^2 \\ &= \left\|\Lambda\left(\mathbf{z} + \mathbf{f}_k\tilde{\theta}_k\right)\right\|_2^2 \\ &= \|\Lambda \mathbf{z}\|_2^2 + \|\Lambda \mathbf{f}_k\|_2^2 \left|\tilde{\theta}_k\right|^2 + 2\mathrm{Re}\left\{\tilde{\theta}_k^* \mathbf{f}_k^* \Lambda^H \Lambda \mathbf{z}\right\} \end{aligned}$$

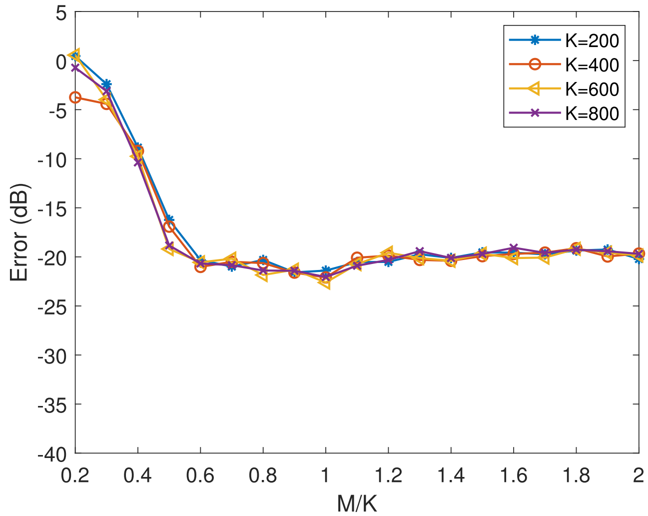


FIGURE 1. Effect of input data length, *M*, on the relative error between the batch and (generalized) time-updating *q*-SPICE estimates.

$$\begin{aligned} \beta_k &= \|\Lambda \mathbf{f}_k\|_2^2 \\ \gamma_k &= \left|\mathbf{f}_k^H \Lambda^H \Lambda \mathbf{x}_k\right| \\ &= \left|\mathbf{f}_k^H \Lambda^H \Lambda\left(\mathbf{z} + \mathbf{f}_k\tilde{\theta}_k\right)\right| \end{aligned} \tag{38}$$

Next, introduce the auxiliary variables

$$\eta = \|\Lambda \mathbf{z}\|_2^2 \tag{39}$$

$$\boldsymbol{\zeta} = \mathbf{F}^H \Lambda^H \Lambda \mathbf{z} \tag{40}$$

and the recursively computed variables

$$\Gamma(M) = \mathbf{F}^H \Lambda^H \Lambda \mathbf{F} = \lambda^2 \Gamma(M-1) + \mathbf{h}_M^H \mathbf{h}_M \tag{41}$$

$$\boldsymbol{\rho}(M) = \mathbf{F}^H \Lambda^H \Lambda \mathbf{y} = \lambda^2 \boldsymbol{\rho}(M-1) + \mathbf{h}_M^H y(t_M) \tag{42}$$

$$\kappa(M) = \mathbf{y}^H \Lambda^H \Lambda \mathbf{y} = \lambda^2 \kappa(M-1) + |y(t_M)|^2 \tag{43}$$

where $\boldsymbol{\Gamma}(M-1)$, $\boldsymbol{\rho}(M-1)$, and $\kappa(M-1)$ denote temporary variables at time sample $M-1$. Then, (38) may be simplified as

$$\begin{aligned} \alpha_k &= \eta + \Gamma_{kk}\left|\tilde{\theta}_k\right|^2 + 2\mathrm{Re}\left\{\tilde{\theta}_k^* \zeta_k\right\} \\ \beta_k &= \Gamma_{kk} \\ \gamma_k &= \left|\zeta_k + \Gamma_{kk}\tilde{\theta}_k\right| \end{aligned} \tag{44}$$

where Γ_{kk} denotes the kth diagonal entry of $\boldsymbol{\Gamma}(M)$, and ζ_k the kth entry of $\boldsymbol{\zeta}$. Similarly, (35) can be expressed as

$$\varphi_k = \arg\left(\zeta_k + \Gamma_{kk}\tilde{\theta}_k\right) \tag{45}$$

Therefore, the computation of $\hat{\boldsymbol{\theta}}$ can be expressed in terms of (44) and (45), together with the current estimate $\tilde{\boldsymbol{\theta}}$.

Once $\hat{\boldsymbol{\theta}}$ has been obtained, the current estimate, $\tilde{\boldsymbol{\theta}}$, must be updated along with the auxiliary variables, $\mathbf{z}$, to compute the subsequent coefficients of $\hat{\boldsymbol{\theta}}$. The remaining derivation then follows the one in [16], [27], and is therefore omitted here. The resulting algorithm is summarized in Table 2.

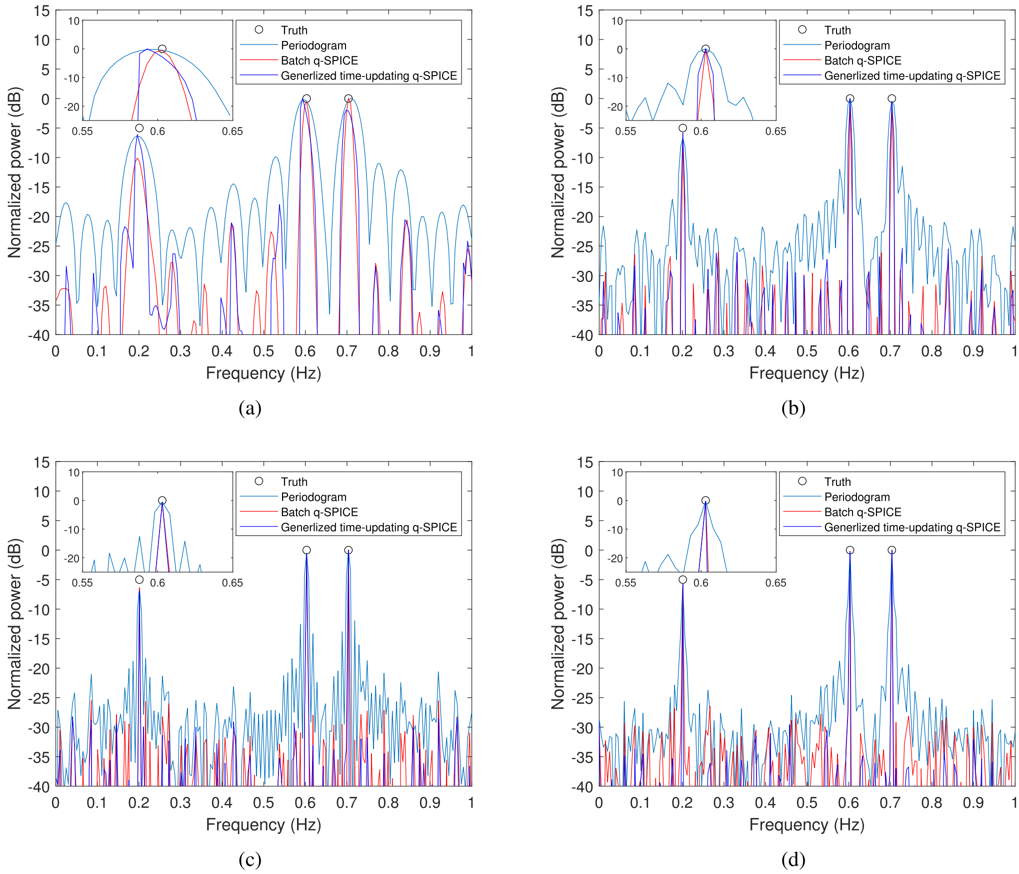


FIGURE 2. Comparison between the batch and (generalized) time-updating q-SPICE estimates under varying input data length, M, for (a) $M = 0.1K$, (b) $M = 0.3K$, (c) $M = 0.5K$, and (d) $M = 0.7K$.

C. COMPUTATIONAL COMPLEXITY ANALYSIS

As may be noted from the resulting expression, using $\lambda = 1$ yields the conventional time-updating q-SPICE algorithm. Therefore, the benefits of the conventional algorithm are well preserved by the proposed generalized time-updating q-SPICE algorithm, being computationally efficient at a constant computational cost of $\mathcal{O}\left(LK^2\right)$ operations per new sample, independent of the data length [16], [27]. In contrast, the batch SPICE implementation requires repeated inversions of $M \times M$ matrices, costing $\mathcal{O}\left(M^3\right)$ operations.

D. EFFECT OF THE CHOICE OF THE FORGETTING FACTOR λ

The choice of λ reflects assumptions on the variability of the spectral content of the signal, with $\lambda = 1$ implying a stationary signal, whereas a smaller value will allow for a quicker adaption to changes in the spectral content. However, a too large or too small value of λ will degrade the performance of the estimate. On one hand, if λ is too large, the importance of older samples is increased, which will weaken the adapting ability for changes in the spectral content. On the other hand, the estimate requires an adequate number of samples to converge to the global minimizer in (34). If λ is selected too small, more samples will in effect be discarded or partially neglected, which may result in that the estimate cannot converge to the global minimizer. According to our experience, the choice of $0.95 < \lambda < 0.97$ can well balance the tradeoff between the recoveries of stationary and non-stationary contents.

IV. NUMERICAL RESULTS

We proceed to examine the numerical performance of the presented online implementation, and illustrate the achievable performance gain in the resulting spectral estimate as compared to the conventional time-updating q-SPICE formulation proposed in [16], [27], as well as several existing online spectral analysis methods. Here, we consider additive zero-mean complex white Gaussian noise, and define the signal-to-noise ratio (SNR) as

$$\text{SNR} = 10\log_{10}\frac{P_s}{\sigma^2} \tag{46}$$

in decibels (dB), where P_s is the power of the signal component, whereas σ^2 denotes the noise variance. The presented simulations do not examine the influences of the choice of q and the iteration number, L, as they have been thoroughly investigated in previous works [13], [16], [27]. Here, we set $q = 1$ and $L = 2$.

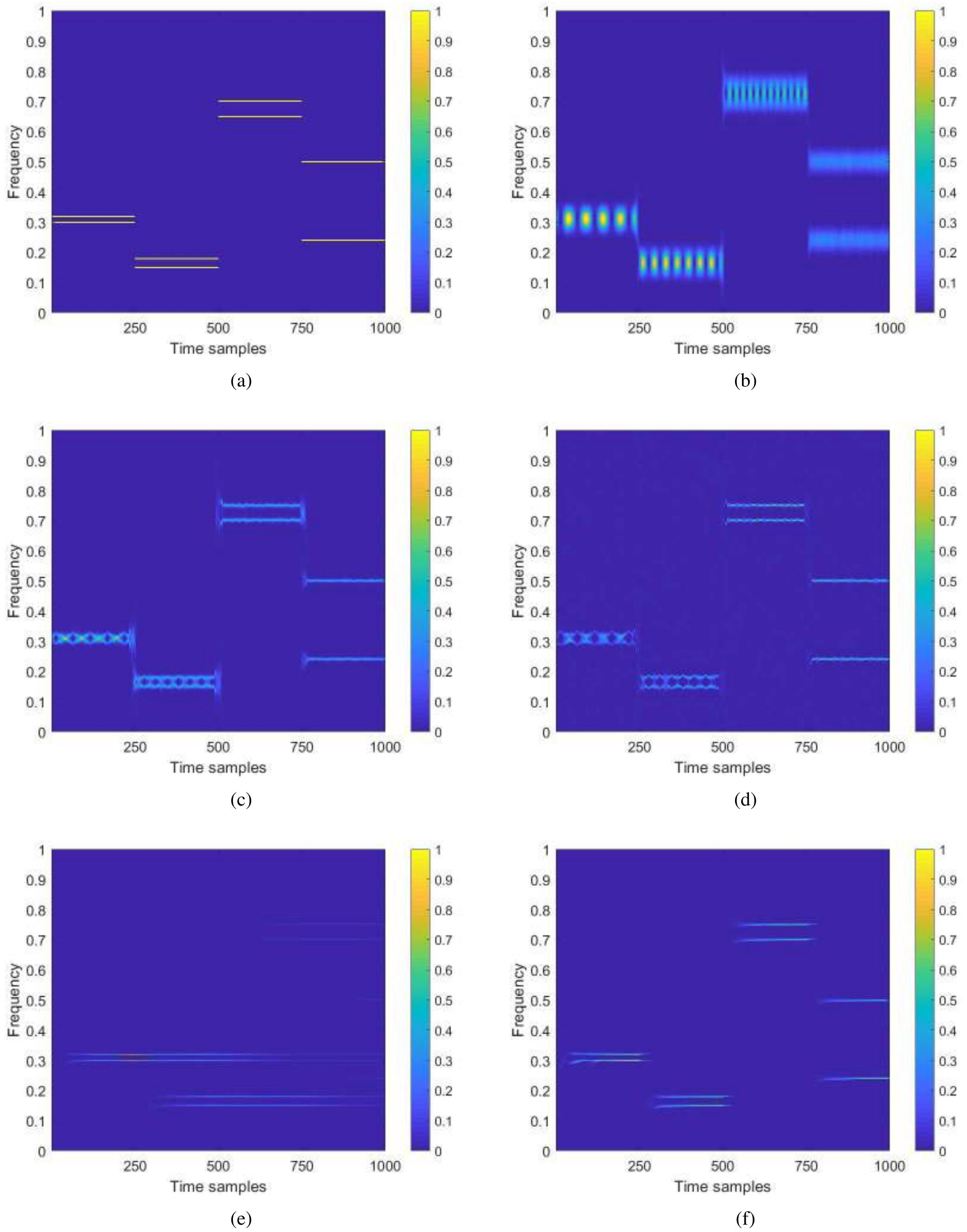


FIGURE 3. Evolution of spectral estimates for frequency hopping signals. (a) Ground-truth. (b) STFT, $W = 25$. (c) Sliding window IAA, $W = 25$. (d) Sliding window q-SPICE, $W = 25$. (e) Original time-updating q-SPICE proposed in [16], [27]. (f) Generalized time-updating q-SPICE ($\lambda = 0.95$).

A. STATIONARY SIGNALS

We begin by comparing the performance of the batch and time-updating implementations. Consider a signal containing three unit amplitude complex sinusoidal components at frequencies $f_1 = 0.2$, $f_2 = 0.6$, and $f_3 = 0.7$, with SNR = 5 dB. Initially considering stationary signals, the forgetting factor is set to $\lambda = 1$. To evaluate the difference between the batch and (generalized) time-updating q-SPICE estimates, denoted by $\hat{\boldsymbol{\theta}}_{\text{batch}}$ and $\hat{\boldsymbol{\theta}}_{\text{online}}$, respectively, we define the relative error

$$\varepsilon = \frac{\left\| \hat{\boldsymbol{\theta}}_{\text{batch}} - \hat{\boldsymbol{\theta}}_{\text{online}} \right\|_2}{\left\| \hat{\boldsymbol{\theta}}_{\text{batch}} \right\|_2} \tag{47}$$

Figure 1 shows the relative error as a function of the ratio between the length of input data and the number of potential candidate frequencies, M/K. For a given M/K ratio, the choice of K has almost no impact on the relative error,

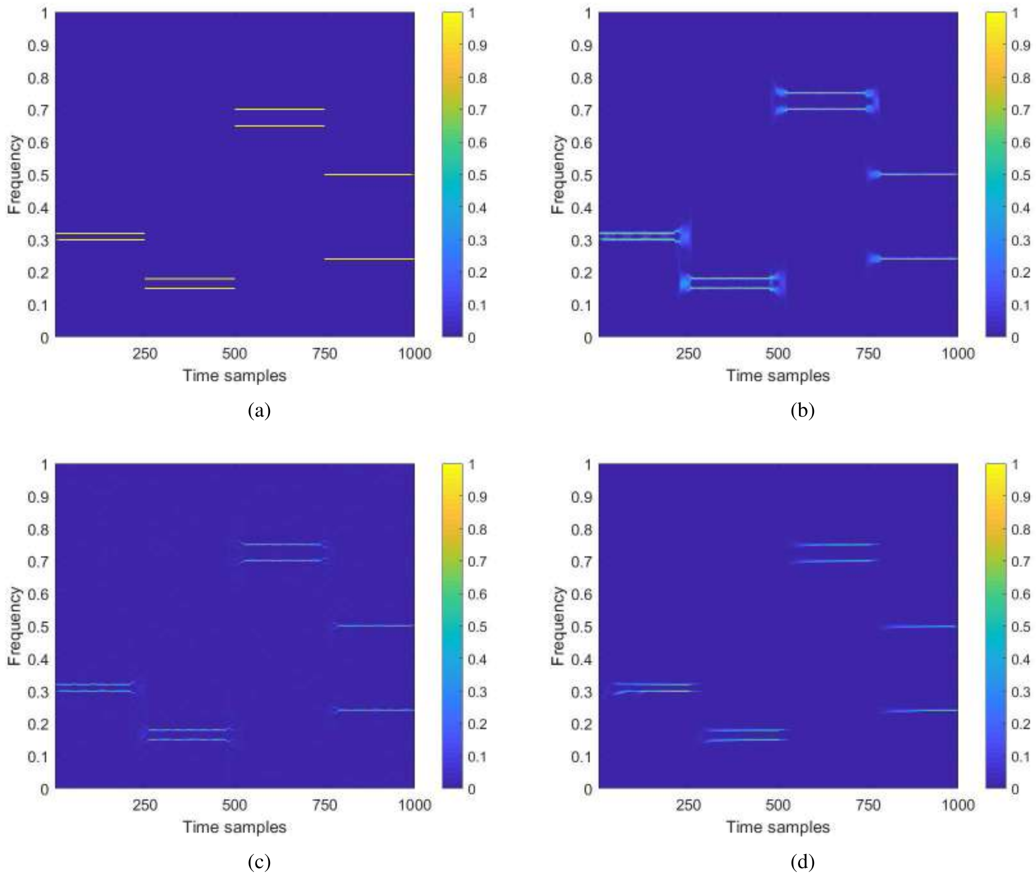


FIGURE 4. **Evolution of spectral estimates for frequency hopping signals. (a) Ground-truth. (b) Sliding window IAA, $W = 50$. (c) Sliding window q-SPICE, $W = 50$. (d) Generalized time-updating q-SPICE ($\lambda = 0.95$).**

with the results converging, as may be expected, as the ratio approach two. Figure 2 illustrates further examples, where the power and frequency estimates of each algorithm for varying M are plotted. The number of potential candidate frequencies is here set to $K = 200$. It may be seen that both the batch and time-updating q-SPICE provide much better results than the periodogram, with both higher resolution and lower sidelobes. Although there is a notable performance difference between the two kinds of q-SPICE implementation when M is relatively small, the difference decreases as M increases. Especially for the cases of moderate or large M (as compared to K), the difference between the frequency estimates becomes marginal.

B. NON-STATIONARY SIGNALS

Frequency hopping signals are widely used in applications of wireless communication and radar electron warfare. We proceed to examine such a frequency hopping signal, with length $M = 1000$, as shown in Figure 3(a), using a dictionary containing $K = 500$ potential candidate frequencies. To allow for the spectral content changes, the forgetting factor for the proposed method is set to $\lambda = 0.95$. The result of the classical short-time Fourier transform (STFT) is shown in Figure 3(b), with a window size of $W = 25$, illustrating the low resolution of the STFT and its poor sidelobe suppression. Figure 3(c) and Figure 3(d) illustrate the results of the sliding window IAA and q-SPICE, using the same window size. Although the frequency resolution has been improved significantly, some closely spaced frequencies cannot be clearly resolved. As seen from Figure 3(e), the original time-updating q-SPICE is incapable of recovering the evolution of the time-varying spectrum. In contrast, by introducing a forgetting factor, the proposed method can successfully reflect the change in the spectral content, with a higher frequency resolution as compared to the aforementioned methods, as shown in Figure 3(f). As the size of sliding window, W, balances the tradeoff between the time and frequency resolution of the sliding window IAA and q-SPICE methods, it is necessary to compare them further with the proposed generalized time-updating q-SPICE for larger W. As shown by Figure 4, when W is increased to 50, the sliding window IAA and q-SPICE methods provide similar frequency resolution as the proposed method. However, as may be seen from the figures, the sliding window implementations in these cases suffer severer leakage problems in the time dimension.

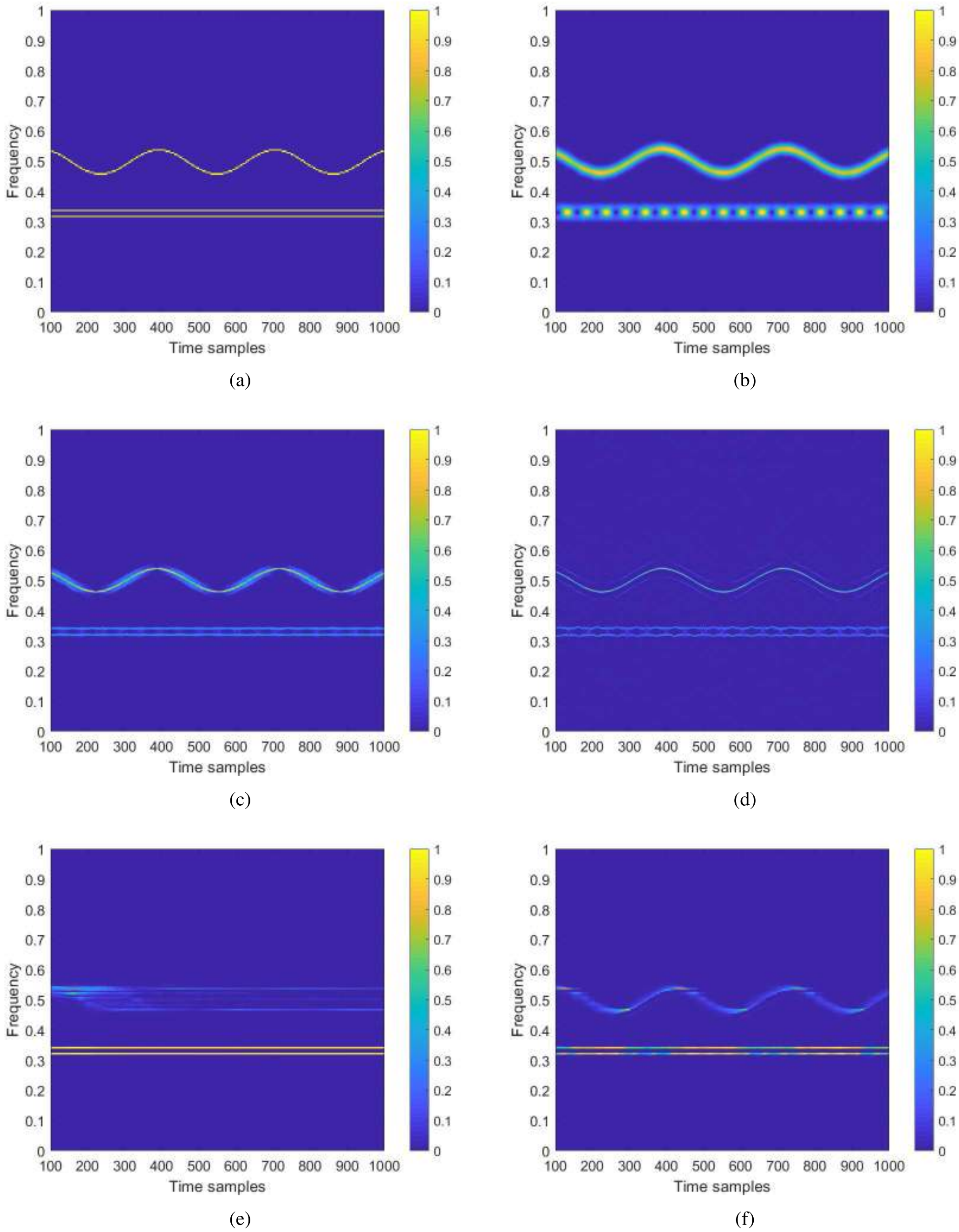
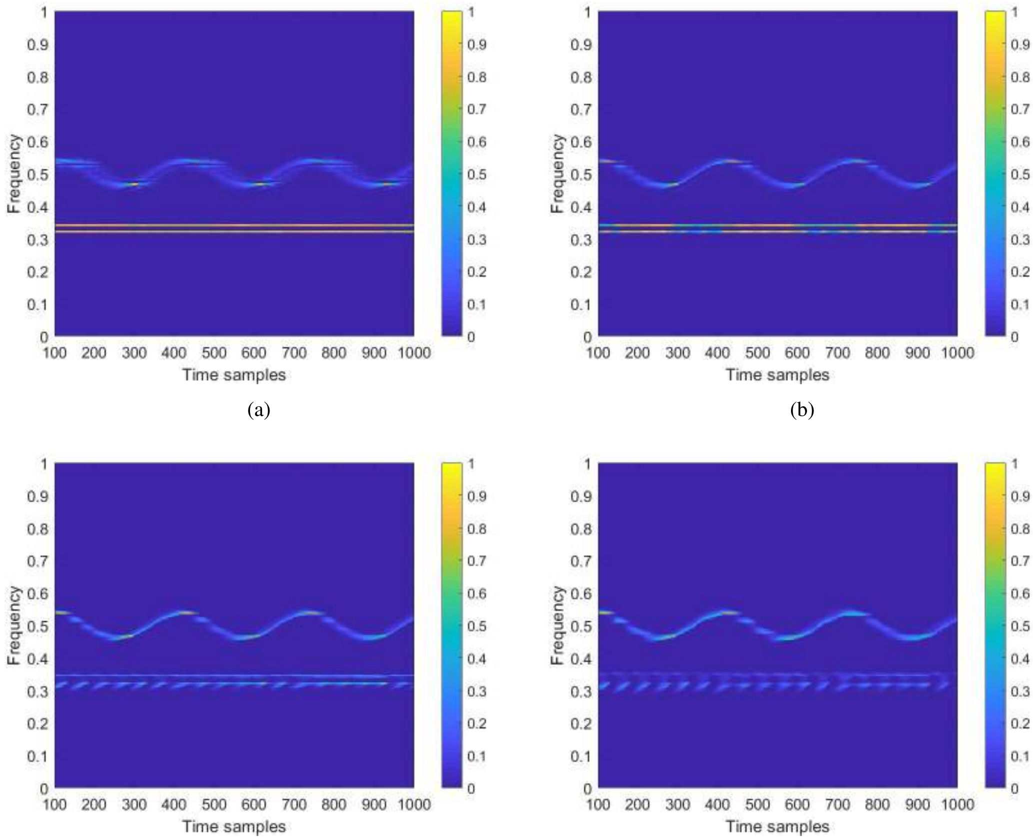


FIGURE 5. **Evolution of spectral estimates for signals containing stationary and non-stationary content. (a) Ground-truth. (b) STFT, $W = 50$. (c) Sliding window IAA, $W = 50$. (d) Sliding window q-SPICE, $W = 50$. (e) Original time-updating q-SPICE proposed in [16], [27]. (f) Generalized time-updating q-SPICE ($\lambda = 0.95$).**

Next, we consider the case of a signal containing both stationary and non-stationary components, as shown in Figure 5(a). The non-stationary signal has a continuously varying frequency content. The number of potential candidate frequencies is set to $K = 200$. Figure 5(b) shows the result of the STFT, using a window size of $W = 50$. Figure 5(c) and Figure 5(d) illustrate the results of the sliding window IAA and q-SPICE methods, using the same window size, clearly yielding preferable performance as compared to the STFT estimate. However, both contain artifacts for the estimates of the non-stationary signal. As seen from Figure 5(e), the original time-updating can only work for the stationary components. In contrast, the proposed method can successfully track changes in the spectral content change, maintaining the high resolution for the stationary signal, with a forgetting factor $\lambda = 0.95$, as shown by Figure 5(f). Although the frequency resolution for the non-stationary component is somewhat inferior to that of the sliding window IAA and



(a) (b) (c) (d)

FIGURE 6. Effect of the choice of the forgetting factor λ. (a) λ = 0.97. (b) λ = 0.95. (c) λ = 0.92. (d) λ = 0.89.

q-SPICE methods, the resulting estimate contains no spurious artifacts.

C. EFFECT OF THE CHOICE OF THE FORGETTING FACTOR λ

Figure 6 illustrates a further comparison for the degree of forgetting, for the signal shown in Figure 5(a). It may be seen that a larger λ will aggravate the blurring effect on the estimates of the non-stationary content. However, a too small λ will degrade the performance of the estimates of the stationary content. Practically, for problems similar to these, our recommendation is to choose $0.95 < \lambda < 0.97$ to well balance the tradeoff between the recoveries of stationary and non-stationary contents.

V. CONCLUSION

This paper derives a generalized time-updating q-SPICE estimator, able to allow for non-stationary signals, while still allowing for the substantial performance improvement offered by the time-updating q-SPICE as compared to the batch q-SPICE. The performance of the implementation has been demonstrated using numerical examples, clearly indicating the improved performance as compared to the original time-updating q-SPICE, as well as to other recent online spectral estimators. The user parameter λ may be selected to trade-off between tracking rapidly changing spectral content and to retain conservatism to small changes.

REFERENCES

[1] N. Bertin, R. Badeau, and E. Vincent, "Enforcing harmonicity and smoothness in Bayesian non-negative matrix factorization applied to polyphonic music transcription," *IEEE Trans. Audio, Speech, Language Process.*, vol. 18, no. 3, pp. 538–549, Mar. 2010.

[2] Y. Lu, Y. Zhou, W. Qu, M. Deng, and C. Zhang, "A Lasso regression model for the construction of microRNA-target regulatory networks," *Bioinformatics*, vol. 27, no. 17, pp. 2406–2413, 2011.

[3] M. Çetin, I. Stojanovic, O. Onhon, K. Varshney, S. Samadi, W. C. Karl, and A. S. Willsky, "Sparsity-driven synthetic aperture radar imaging: Reconstruction, autofocusing, moving targets, and compressed sensing," *IEEE Signal Process. Mag.*, vol. 31, no. 4, pp. 27–40, Jul. 2014.

[4] P. Stoica, P. Babu, and J. Li, "New method of sparse parameter estimation in separable models and its use for spectral analysis of irregularly sampled data," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 35–47, Jan. 2011.

[5] P. Stoica and R. Moses, *Spectral Analysis of Signals*. Upper Saddle River, NJ, USA: Prentice-Hall, 2005.

[6] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Roy. Statist. Soc., B (Methodol.)*, vol. 58, no. 1, pp. 267–288, 1996.

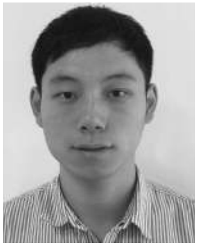
[7] T. Kronvall and A. Jakobsson, "Hyperparameter selection for group-sparse regression: A probabilistic approach," *Signal Process.*, vol. 151, pp. 107–118, Oct. 2018.
[8] P. Stoica, P. Babu, and J. Li, "SPICE: A sparse covariance-based estimation method for array processing," *IEEE Trans. Signal Process.*, vol. 59, no. 2, pp. 629–638, Feb. 2011.
[9] P. Stoica and P. Babu, "SPICE and LIKES: Two hyperparameter-free methods for sparse-parameter estimation," *Signal Process.*, vol. 92, no. 7, pp. 1580–1590, Jul. 2012.
[10] C. R. Rojas, D. Katselis, and H. Hjalmarsson, "A note on the SPICE method," *IEEE Trans. Signal Process.*, vol. 61, no. 18, pp. 4545–4551, Sep. 2013.
[11] P. Babu and P. Stoica, "Connection between SPICE and square-root LASSO for sparse parameter estimation," *Signal Process.*, vol. 95, pp. 10–14, Feb. 2014.
[12] P. Stoica, D. Zachariah, and J. Li, "Weighted SPICE: A unifying approach for hyperparameter-free sparse estimation," *Digit. Signal Process.*, vol. 33, pp. 1–12, Oct. 2014.
[13] J. Swärd, S. I. Adalbjörnsson, and A. Jakobsson, "Generalized sparse covariance-based estimation," *Signal Process.*, vol. 143, pp. 311–319, Feb. 2018.
[14] T. Kronvall, S. I. Adalbjörnsson, S. Nadig, and A. Jakobsson, "Group-sparse regression using the covariance fitting criterion," *Signal Process.*, vol. 139, pp. 116–130, Oct. 2017.
[15] Y. Zhang, A. Jakobsson, Y. Zhang, Y. Huang, and J. Yang, "Wideband sparse reconstruction for scanning radar," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 10, pp. 6055–6068, Oct. 2018.
[16] D. Zachariah and P. Stoica, "Online hyperparameter-free sparse estimation method," *IEEE Trans. Signal Process.*, vol. 63, no. 13, pp. 3348–3359, Jul. 2015.
[17] Q. Zhang, H. Abeida, M. Xue, W. Rowe, and J. Li, "Fast implementation of sparse iterative covariance-based estimation for source localization," *J. Acoust. Soc. Amer.*, vol. 131, no. 2, pp. 1249–1259, Feb. 2012.
[18] G. Glentis, "Efficient algorithms for adaptive Capon and APES spectral estimation," *IEEE Trans. Signal Process.*, vol. 58, no. 1, pp. 84–96, Jan. 2010.
[19] G. O. Glentis and A. Jakobsson, "Time-recursive IAA spectral estimation," *IEEE Signal Process. Lett.*, vol. 18, no. 2, pp. 111–114, Feb. 2011.
[20] J. Karlsson, W. Rowe, L. Xu, G. O. Glentis, and J. Li, "Fast missing-data IAA with application to notched spectrum SAR," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 50, no. 2, pp. 959–971, Apr. 2014.
[21] G. O. Glentis and A. Jakobsson, "Superfast approximative implementation of the IAA spectral estimate," *IEEE Trans. Signal Process.*, vol. 60, no. 1, pp. 472–478, Jan. 2012.
[22] G.-O. Glentis and A. Jakobsson, "Efficient implementation of iterative adaptive approach spectral estimation techniques," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4154–4167, Sep. 2011.
[23] J. R. Jensen, G. Glentis, M. G. Christensen, A. Jakobsson, and S. H. Jensen, "Fast LCMV-based methods for fundamental frequency estimation," *IEEE Trans. Signal Process.*, vol. 61, no. 12, pp. 3159–3172, Jun. 2013.
[24] A. Jakobsson, G. O. Glentis, and E. Gudmundson, "Computationally efficient time-recursive IAA-based blood velocity estimation," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3853–3858, Jul. 2012.
[25] K. Angelopoulos, G.-O. Glentis, and A. Jakobsson, "Computationally efficient sparsity-inducing coherence spectrum estimation of complete and non-complete data sets," *Signal Process.*, vol. 93, no. 5, pp. 1221–1234, 2013.
[26] Y. Zhang, Y. Zhang, Y. Huang, J. Yang, and W. Li, "Super-resolution surface mapping for scanning radar: Inverse filtering based on the fast iterative adaptive approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 1, pp. 127–144, Jan. 2018.
[27] A. Sakhnini, D. Montgomery, and A. Jakobsson, "Time-updating the generalized sparse covariance-based estimator," in *Proc. Int. Symp. Intell. Signal Process. Commun. Syst.*, Nov. 2018.
[28] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*. Englewood Cliffs, NJ, USA: Prentice-Hall, 2000.
[29] F. Elvander, J. Swärd, and A. Jakobsson, "Online estimation of multiple harmonic signals," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 25, no. 2, pp. 273–284, Feb. 2017.

YONGCHAO ZHANG (S'15–M'18) received the B.S. degree in electronic information engineering from Hainan University, Haikou, China, in 2011, and the Ph.D. degree from the School of Information and Communication Engineering, University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2018.

From 2016 to 2017, he was a Visiting Student with Lund University, Lund, Sweden. He is currently an Associate Research Fellow with the School of Information and Communication Engineering, UESTC. His research interests include array signal processing and inverse problem in radar applications.

ANDREAS JAKOBSSON (S'95–M'00–SM'06) received the M.Sc. degree from the Lund Institute of Technology, in 1993, and the Ph.D. degree in signal processing from Uppsala University, in 2000.

Since, he has held positions at Global IP Sound AB, the Swedish Royal Institute of Technology, the King's College London, and Karlstad University, as well as held an Honorary Research Fellowship at Cardiff University. He has been a Visiting Researcher with the King's College London, Brigham Young University, Stanford University, Katholieke Universiteit Leuven, and the University of California, San Diego, as well as acted as an expert for the IAEA. He is currently a Professor and the Head of the Mathematical Statistics with Lund University, Sweden. He has published his research findings in about 200 refereed journal and conference papers, and has filed five patents. He has also authored a book on *Time Series Analysis* (Studentlitteratur, 2013 and 2015), and coauthored (together with M. G. Christensen) a book on *Multi-Pitch Estimation* (Morgan & Claypool, 2009). His research interests include statistical and array signal processing, detection and estimation theory, and related application in remote sensing, telecommunication, and biomedicine. He was an Associate Editor for the IEEE Transactions on Signal Processing, from 2006 to 2010, the IEEE Signal Processing Letters, from 2007 to 2011, the *Research Letters in Signal Processing*, from 2007 to 2009, and the *Journal of Electrical and Computer Engineering*, from 2009 to 2014, and a member of the IEEE Sensor Array and Multichannel (SAM) Signal Processing Technical Committee, from 2008 to 2013. He has been a member of The Royal Swedish Physiographic Society, the EURASIP Special Area Team on Signal Processing for Multisensor Systems, since 2015, and an Associate Editor for *Signal Processing* (Elsevier).

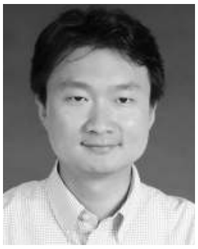


DEQING MAO (S'17) received the B.S. degree from the School of Electronic Engineering, Chengdu University of Information Technology, Chengdu, China, in 2010. He is currently pursuing the Ph.D. degree with the School of Information and Communication Engineering, University of Electronic Science and Technology of China (UESTC), Chengdu.

His research interests include radar signal processing and inverse problem in radar imaging.

YIN ZHANG (S'13–M'17) received the B.S. and Ph.D. degrees in electronic information engineering from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2008 and 2016, respectively.

From 2015 to 2016, he was a Visiting Student with the University of Delaware, USA. He is currently an Associate Research Fellow with the School of Electronic Engineering, UESTC. His research interests include radar imaging and signal processing in related radar applications.



YULIN HUANG (M'08–SM'17) received the B.S. and Ph.D. degrees in electronic engineering from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2002 and 2008, respectively.

From 2013 to 2014, he was a Visiting Researcher with the University of Houston, USA. Since 2016, he has been a Professor with the School of Electronic Engineering, UESTC. He has authored more than 80 journal and conference papers. His research interests include synthetic aperture radar, target detection and recognition, artificial intelligence, and machine Learning. He is a member of the IEEE Aerospace and Electronic Systems Society.

JIANYU YANG (M'91) received the B.S. degree from the National University of Defense Technology, Changsha, China, in 1984, and the M.S. and Ph.D. degrees from the University of Electronic Science and Technology of China (UESTC), Chengdu, in 1987 and 1991, respectively.

In 2005, he visited the Massachusetts Institute of Technology, USA. Since 1999, he has been a Professor with the School of Electronic Engineering, UESTC. He is currently a Senior Member of the Chinese Institute of Electronics. He has authored more than 120 journal and conference papers. His research interests include synthetic aperture radar and statistical signal processing.