State Smoothing in Markov-Switching Time-Frequency GARCH Models

Ari Abramson and Israel Cohen, *Senior Member, IEEE*

***Abstract*—In this letter, we propose a state smoothing algorithm for path-dependent Markov-switching generalized autoregressive conditional heteroscedasticity (GARCH) processes. Our smoothing technique extends the *forward–backward recursions* of Chang and Hancock and the *stable backward recursion* of Lindgren, Askar and Derin. We derive two recursive steps for the evaluation of conditional densities of future observations. The first step is an upward recursion that manipulates the future observations for the evaluation of their conditional densities, and the second step is a backward recursion that integrates over the possible future paths. Experimental results demonstrate the improvement in performance, compared to using causal estimation.**

***Index Terms*—Forward–backward recursions, generalized autoregressive conditional heteroscedasticity (GARCH), stable backward recursion, state smoothing.**

I. Introduction

STATE estimation is of both theoretical and practical importance whenever the underlying statistical model switches regimes over time [1], [2]. State smoothing (i.e., noncausal state estimation) of hidden Markov processes (HMPs) has been originally introduced by Chang and Hancock [3]. Their solution for estimating the noncausal state probability, which is implemented using *forward–backward recursions*, decouples a forward recursion for the evaluation of the joint probability density of the current state and all observations up to the same time as well as a backward recursion for obtaining the future observations' density given the current state. Lindgren [4] and Askar and Derin [5] developed an alternative *stable backward recursion* for the state smoothing in HMPs. Kim [6] extended the stable backward recursion to nonmemoryless autoregressive hidden Markov processes (AR-HMPs), where both the current state (regime) and a finite set of past values are required for the conditional density evaluation (see also [2, chap. 22]).

Generalized autoregressive conditional heteroscedasticity (GARCH) models and also Markov-switching GARCH (MS-GARCH) models are widely used in the field of econometrics for volatility forecast derivation of economics rates [7]–[10], and they have recently been utilized for several signal processing applications. In [11], GARCH modeling has been applied to spatially nonuniform noise in multichannel signal processing. In [12], a regime-switching GARCH model has been utilized for speech recognition, and a complex-valued GARCH model has been proposed in [13] and [14] for modeling speech signals in the short-time Fourier transform (STFT) domain for the application of speech enhancement. Generally, when incorporating GARCH processes with switching-regimes, the volatility evaluation requires knowledge of the pertinent history of the regime-switching GARCH process, including the regime-path [7], [8]. Properties of path-dependent MS-GARCH models have been studied by Francq *et al.* [15]. In order to estimate the model parameters, they showed that the conditional likelihood depends on all the possible paths, and for a Markov-switching ARCH model (in which case, there is no dependency on past active regimes), they showed that the forward–backward recursions can be employed for the conditional likelihood evaluation. The complex-valued GARCH model has been shown to be useful in speech enhancement applications [13], [14]. Motivated by extending the dynamic formulation of the time-frequency GARCH model and enabling a better fit for a process with a more complicated time-varying statistical behavior, a Markov-switching time-frequency GARCH (MSTF-GARCH) model has been introduced [16]. However, existing smoothing solutions are inapplicable in the case of a path-dependent MS-GARCH model since both past observations and the regime path are required for the conditional variance estimation, whereas existing smoothing techniques rely on the assumption that given the current state, past active regimes are statistically independent of future densities.

In this letter, we develop a state smoothing approach for MSTF-GARCH processes. The dependency of the conditional variance on past observations and past active regimes are taken into consideration as we generalize both the forward–backward recursions of Chang and Hancock [3] and the stable backward recursion of Lindgren [4] and Askar and Derin [5]. We derive two recursive steps for the evaluation of conditional densities of future observations. The first step is an upward recursion that manipulates the future observations for the evaluation of their conditional densities, corresponding to all possible future paths. The second step is a backward recursion that integrates over these paths to evaluate the future densities required for the noncausal state probability. The computational complexity of the generalized recursions grows exponentially with the number of future observations employed for the fixed-lag smoothing. However, experimental results demonstrate that the significant part of the improvement in performance, compared to using causal estimation, is achieved by considering a few future observations.

The organization of this letter is as follows: In Section II, we introduce the MSTF-GARCH model and formulate the

Manuscript received November 23, 2005; revised December 29, 2005. This work was supported by the Israel Science Foundation under Grant 1085/05. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Alfred Hanssen.

The authors are with the Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa 32000, Israel (e-mail: aari@tx.technion.ac.il; icohen@ee.technion.ac.il).

Digital Object Identifier 10.1109/LSP.2006.871708

state smoothing problem. In Section III, we develop generalized forward–backward recursions as well as generalized stable backward recursions and derive our noncausal state probability approach. Finally, in Section IV, we provide experimental results that demonstrate state smoothing for noisy MSTF-GARCH processes.

II. Problem Formulation

Let $\mathbf{X}_t \in \mathbb{C}^K$ be a K-dimensional random vector at a discrete time t, and let X_{tk}, $k \in \{0, \ldots, K-1\}$ be its kth element. Let $\mathcal{X}_{t_1}^{t_2} = \{\mathbf{X}_t | t_1 \leq t \leq t_2\}$ represent the data set from time t_1 up to t_2, and let $\mathcal{X}^t \triangleq \mathcal{X}_0^t$. Let S_t denote the (unobserved) state at time t, and let s_t be a realization of S_t, assuming S_t is a first-order Markov chain with transition probabilities $a_{s_t s_{t+1}} \triangleq p(S_{t+1} = s_{t+1} | S_t = s_t)$. Let $\mathcal{I}^t \triangleq \{\mathcal{X}^t, \mathcal{S}^t\}$ denote all available information up to time t, where $\mathcal{S}^t \triangleq \mathcal{S}_0^t = \{s_0, \ldots, s_t\}$. We assume that X_{tk} are generated by an m-state MSTF-GARCH process of order (1,1), which follows [16]:

$$X_{tk} = \sqrt{\lambda_{tk|t-1}} V_{tk}, \quad k = 0, \ldots, K-1 \tag{1}$$

where $\{V_{tk}\}$ are iid complex-valued random variables with zero-mean, unit variance, and some known probability density. Given the state s_t, the conditional variance of X_{tk}, $\lambda_{tk|t-1,s_t} = E\{|X_{tk}|^2 | \mathcal{I}^{t-1}, s_t\}$, is a linear function of the previous conditional variance and squared absolute value

$$\lambda_{tk|t-1} \equiv \lambda_{tk|t-1,s_t} = \xi_{s_t} + \alpha_{s_t} |X_{t-1,k}|^2 + \beta_{s_t} \lambda_{t-1,k|t-2} \tag{2}$$

where $\xi_s > 0$, $\alpha_s \geq 0$, and $\beta_s \geq 0$, $s = 1, \ldots, m$ are sufficient constrains for the positivity of the conditional variance.

Let $\mathbf{Y}_t = \mathbf{X}_t + \mathbf{D}_t$ denote the observed noisy signal, where $\mathbf{D}_t$ denotes the noise process that is uncorrelated with the signal $\mathbf{X}_t$, and let $\mathbf{D}_t$ be a zero-mean complex-valued Gaussian random process with a diagonal covariance matrix $E\{\mathbf{D}_t \mathbf{D}_t^H\} = \mathrm{diag}\{\boldsymbol{\sigma}^2\}$, where $(\cdot)^H$ denotes the Hermitian transpose operation. The state conditional probability of a Markov-switching model, $p(s_t | \mathcal{Y}^\tau)$, is of considerable theoretical and practical importance for signal restoration and state sequence estimation (e.g., [1] and [16]).

Solutions of the state smoothing problem, i.e., $\tau > t$, are normally obtained for HMPs using the forward–backward recursions [3] or the stable backward recursion [4], [5]. Extensions of these recursions for nonmemoryless AR-HMPs [2, Chap. 22], [6] are based on the quality that s_t and a finite set of past clean observations give complete statistical knowledge of future densities. However, in case of a path-dependent MS-GARCH model, a recursive formulation specifies the conditional distribution of the process as dependent on both past observations and the regime path, and therefore, existing smoothing solutions are inapplicable.

III. State Probability Smoothing

In this section, we develop the noncausal state probability for the model defined in (1) and (2). The smoothed probability is derived by generalizing both the forward–backward recursions [3] and the stable backward recursion [4], [5].

A. Generalized Forward–Backward Recursions

Assume that the conditional variance of the process is recursively estimated for any given state (e.g., as proposed in [16]), and assume that the set of the recursively estimated conditional variances at time t, $\hat{\Lambda}_t \triangleq \{\hat{\boldsymbol{\lambda}}_{t|t-1,S_t} | S_t = 1, \ldots, m\}$, with the observed signal $\mathbf{Y}_t$, are sufficient statistics for the next conditional variance estimation for any given regime [14], [16]. Let $\hat{\boldsymbol{\lambda}}_{\tau_2|\tau_1,\mathcal{S}_{\tau_0}^{\tau_2}} = E\{\mathbf{X}_{\tau_2} \odot \mathbf{X}_{\tau_2}^* | \mathcal{S}_{\tau_0}^{\tau_2}, \mathcal{Y}^{\tau_1}\}$, $\tau_2 \geq \tau_1 > \tau_0$ denote the vector of estimated conditional variances at time τ_2 based on the observations up to time τ_1 and on the given set of active regimes $\mathcal{S}_{\tau_0}^{\tau_2}$, where $\odot$ denotes a term-by-term multiplication, and * denotes complex conjugation. Let $g(\boldsymbol{\lambda}_{t|t-1,s_t}, \mathbf{Y}_t) \triangleq E\{\mathbf{X}_t \odot \mathbf{X}_t^* | S_t = s_t, \boldsymbol{\lambda}_{t|t-1,s_t}, \boldsymbol{Y}_t\}$, where the function $g(\cdot)$ is determined based on the statistical model of $\{V_{tk}\}$ [14]. Define the generalized forward density by $\alpha(s_t, \mathcal{Y}^t) \triangleq f(s_t, \hat{\Lambda}_t, \mathbf{Y}_t)$ and the generalized backward density by $\beta(\mathcal{Y}_{t+l}^{t+L} | \mathcal{S}_t^{t+l-1}, \mathcal{Y}^{t+l-1}) \triangleq f(\mathcal{Y}_{t+l}^{t+L} | \mathcal{S}_t^{t+l-1}, \hat{\Lambda}_t, \mathcal{Y}_t^{t+l-1})$. Then, by substituting $l = 1$, we have $f\left(s_t, \mathcal{Y}_t^{t+L} | \hat{\Lambda}_t\right) = \alpha(s_t, \mathcal{Y}^t) \beta\left(\mathcal{Y}_{t+1}^{t+L} | s_t, \mathcal{Y}^t\right)$, and the noncausal state probability can be obtained by

$$p(s_t | \mathcal{Y}^{t+L}) = \frac{\alpha(s_t, \mathcal{Y}^t) \beta\left(\mathcal{Y}_{t+1}^{t+L} | s_t, \mathcal{Y}^t\right)}{\sum_{s_t} \alpha(s_t, \mathcal{Y}^t) \beta\left(\mathcal{Y}_{t+1}^{t+L} | s_t, \mathcal{Y}^t\right)}. \tag{3}$$

Proposition 1: The generalized forward density of an MSTF-GARCH (1,1) process, $\alpha(s_t, \mathcal{Y}^t)$, satisfies the following recursion:

$$\alpha(s_t, \mathcal{Y}^t) = f\left(\mathbf{Y}_t | s_t, \hat{\boldsymbol{\lambda}}_{t|t-1,s_t}\right) \sum_{s_{t-1}} \alpha(s_{t-1}, \mathcal{Y}^{t-1}) a_{s_{t-1}s_t} \tag{4}$$

with the initial condition $\alpha(s_0, \mathbf{Y}_0) = p(s_0) f(\mathbf{Y}_0 | s_0)$.

Proof: The generalized forward density is obtained by $\alpha(s_t, \mathcal{Y}^t) = f(\mathbf{Y}_t | s_t, \hat{\Lambda}_t) f(s_t, \hat{\Lambda}_t)$. Given the active regime, the state-dependent conditional variance is sufficient for the conditional density. Furthermore, $\hat{\Lambda}_t$ and $\{\hat{\Lambda}_{t-1}, \mathbf{Y}_{t-1}\}$ represent the same statistical information. Hence, $\alpha(s_t, \mathcal{Y}^t) = f\left(\mathbf{Y}_t | s_t, \hat{\boldsymbol{\lambda}}_{t|t-1,s_t}\right) f(s_t, \hat{\Lambda}_{t-1}, \mathbf{Y}_{t-1})$, where $f(s_t, \hat{\Lambda}_{t-1}, \mathbf{Y}_{t-1}) = \sum_{s_{t-1}} \alpha(s_{t-1}, \mathcal{Y}^{t-1}) a_{s_{t-1}s_t}$, which yields the recursive formulation for the generalized forward density.[1] ■

Proposition 2: The generalized backward density of an MSTF-GARCH (1,1) process, $\beta(\mathcal{Y}_{t+1}^{t+L} | s_t, \mathcal{Y}^t)$, satisfies the following two-step recursion.

Step 1) For $l = 1, \ldots, L$ and all $\mathcal{S}_t^{t+l}$

$$\hat{\boldsymbol{\lambda}}_{t+l|t+l-1,\mathcal{S}_t^{t+l}} = \xi_{s_{t+l}} \mathbf{1} + \alpha_{s_{t+l}} \hat{\boldsymbol{\lambda}}_{t+l-1|t+l-1,\mathcal{S}_t^{t+l-1}} + \beta_{s_{t+1}} \hat{\boldsymbol{\lambda}}_{t+l-1|t+l-2,\mathcal{S}_t^{t+l-1}} \tag{5}$$

[1]The initial conditions for the generalized forward recursion have negligible effect on the conditional densities, assuming an asymptotic stationary process that is sufficiently long. Therefore, the initial conditional variance $f(\mathbf{Y}_0 | s_0)$ can be estimated by using the state-dependent stationary density of the process.

$$\hat{\boldsymbol{\lambda}}_{t+l|t+l,\mathcal{S}_t^{t+l}} = g\left(\hat{\boldsymbol{\lambda}}_{t+l|t+l-1,\mathcal{S}_t^{t+l}}, \mathbf{Y}_{t+l}^2\right). \tag{6}$$

Step 2) For $l = L, \ldots, 1$ and all $\mathcal{S}_t^{t+l}$

$$\begin{aligned} f&\Big(\mathcal{Y}_{t+l}^{t+L}|\mathcal{S}_t^{t+l}, \hat{\boldsymbol{\lambda}}_{t|t-1,s_t}, \mathcal{Y}_t^{t+l-1}\Big) \\ &= \beta\left(\mathcal{Y}_{t+l+1}^{t+L}|\mathcal{S}_t^{t+l}, \mathcal{Y}^{t+l}\right) \\ &\quad \times f\Big(\mathbf{Y}_{t+l}|\mathcal{S}_t^{t+l}, \hat{\boldsymbol{\lambda}}_{t|t-1,s_t}, \mathcal{Y}_t^{t+l-1}\Big) \end{aligned} \tag{7}$$

$$\begin{aligned} \beta&(\mathcal{Y}_{t+l}^{t+L}|\mathcal{S}_t^{t+l-1}, \mathcal{Y}^{t+l-1}) \\ &= \sum_{s_{t+l}} f\Big(\mathcal{Y}_{t+l}^{t+L}|\mathcal{S}_t^{t+l}, \hat{\boldsymbol{\lambda}}_{t|t-1,s_t}, \mathcal{Y}_t^{t+l-1}\Big) a_{s_{t+l-1}s_{t+l}} \end{aligned} \tag{8}$$

with $\beta(\mathcal{Y}_{t+L+1}^{t+L}|\mathcal{S}_t^{t+L}, \mathcal{Y}^{t+L}) = 1$ as the initial condition for the second step, and where $\mathbf{1}$ denotes a vector of ones.

Proof: The generalized backward density can be obtained by

$$\beta\left(\mathcal{Y}_{t+1}^{t+L}|s_t, \mathcal{Y}^t\right) = \sum_{s_{t+1}} f\left(\mathcal{Y}_{t+1}^{t+L}|\mathcal{S}_t^{t+1}, \hat{\boldsymbol{\lambda}}_{t|t-1,s_t}, \mathbf{Y}_t\right) a_{s_t s_{t+1}} \tag{9}$$

where the multivariate density $f(\mathcal{Y}_{t+1}^{t+L}|\mathcal{S}_t^{t+1}, \hat{\boldsymbol{\lambda}}_{t|t-1,s_t}, \mathbf{Y}_t)$ in (13) can be obtained by

$$\begin{aligned} f\left(\mathcal{Y}_{t+1}^{t+L}|\mathcal{S}_t^{t+1}, \hat{\boldsymbol{\lambda}}_{t|t-1,s_t}, \mathbf{Y}_t\right) &= \beta\left(\mathcal{Y}_{t+2}^{t+L}|\mathcal{S}_t^{t+1}, \mathcal{Y}^{t+1}\right) \\ &\quad \times f\left(\mathbf{Y}_{t+1}|\mathcal{S}_t^{t+1}, \hat{\boldsymbol{\lambda}}_{t|t-1,s_t}, \mathbf{Y}_t\right). \end{aligned} \tag{10}$$

From (9) and (10), we recursively obtain for any $l = 1, \ldots, L$

$$\begin{aligned} \beta&\left(\mathcal{Y}_{t+l}^{t+L}|\mathcal{S}_t^{t+l-1}, \mathcal{Y}^{t+l-1}\right) \\ &= \sum_{s_{t+l}} f\left(\mathcal{Y}_{t+l}^{t+L}|\mathcal{S}_t^{t+l}, \hat{\boldsymbol{\lambda}}_{t|t-1,s_t}, \mathcal{Y}_t^{t+l-1}\right) a_{s_{t+l-1}s_{t+l}} \end{aligned} \tag{11}$$

and

$$\begin{aligned} f\left(\mathcal{Y}_{t+l}^{t+L}|\mathcal{S}_t^{t+l}, \hat{\boldsymbol{\lambda}}_{t|t-1,s_t}, \mathcal{Y}_t^{t+l-1}\right) &= \beta\left(\mathcal{Y}_{t+l+1}^{t+L}|\mathcal{S}_t^{t+l}, \mathcal{Y}^{t+l}\right) \\ &\quad \times f\left(\mathbf{Y}_{t+l}|\mathcal{S}_t^{t+l}, \hat{\boldsymbol{\lambda}}_{t|t-1,s_t}, \mathcal{Y}_t^{t+l-1}\right). \end{aligned} \tag{12}$$

The conditional density $f(\mathbf{Y}_{t+l}|\mathcal{S}_t^{t+l}, \hat{\boldsymbol{\lambda}}_{t|t-1,s_t}, \mathcal{Y}_t^{t+l-1})$ in (16) is the density of the observed data at time $t+l$ conditioned on the regime path $\mathcal{S}_t^{t+l}$, the recursively estimated conditional variance at time t given s_t, and also on all observations from time t up to time $t+l-1$. This density has a diagonal covariance matrix with the following conditional variance on its diagonal:

$$\begin{aligned} E&\Big\{\mathbf{Y}_{t+l} \odot \mathbf{Y}_{t+l}^*|\mathcal{S}_t^{t+l}, \hat{\boldsymbol{\lambda}}_{t|t-1,s_t}, \mathcal{Y}_t^{t+l-1}\Big\} \\ &= \boldsymbol{\sigma}^2 + \hat{\boldsymbol{\lambda}}_{t+l|t+l-1,\mathcal{S}_t^{t+l}} \\ &= \boldsymbol{\sigma}^2 + \xi_{s_{t+l}}\mathbf{1} \\ &\quad + \alpha_{s_{t+l}} E\Big\{\mathbf{X}_{t+l-1} \odot \mathbf{X}_{t+l-1}^*|\mathcal{S}_t^{t+l}, \hat{\boldsymbol{\lambda}}_{t|t-1,s_t}, \mathcal{Y}_t^{t+l-1}\Big\} \\ &\quad + \beta_{s_{t+l}} E\Big\{\boldsymbol{\lambda}_{t+l-1|t+l-2}|\mathcal{S}_t^{t+l-1}, \hat{\boldsymbol{\lambda}}_{t|t-1,s_t}, \mathcal{Y}_t^{t+l-2}\Big\}. \end{aligned} \tag{13}$$

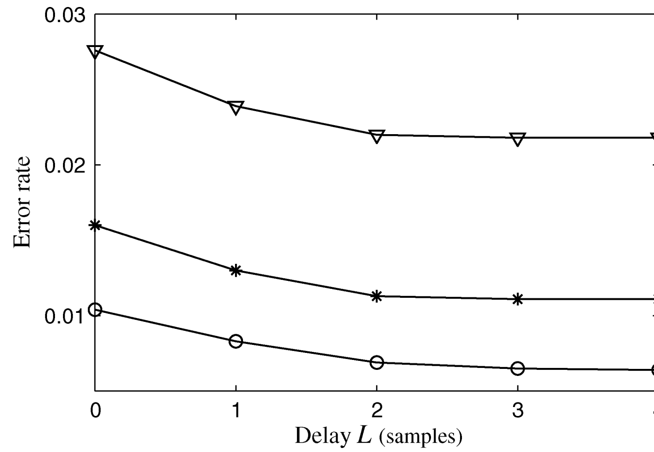

Fig. 1. State smoothing error rate for three-state MSTF-GARCH models with SNRs of 5 dB (triangle), 10 dB (asterisk), and 15 dB (circle).

The expected absolute squared value of the signal at a specific time given the active regime is independent of any future regimes; hence

$$\begin{aligned} E&\left\{\mathbf{X}_{t+l-1} \odot \mathbf{X}_{t+l-1}^*|\mathcal{S}_t^{t+l}, \hat{\boldsymbol{\lambda}}_{t|t-1,s_t}, \mathcal{Y}_t^{t+l-1}\right\} \\ &= \hat{\boldsymbol{\lambda}}_{t+l-1|t+l-1,\mathcal{S}_t^{t+l-1}} \\ &= g\left(\hat{\boldsymbol{\lambda}}_{t+l-1|t+l-2,\mathcal{S}_t^{t+l-1}}, \mathbf{Y}_{t+l-1}\right). \end{aligned} \tag{14}$$

Combining (14) with (13), we obtain Step 1) of the generalized backward recursion [see (5) and (6)], and from (11) and (12), we obtain Step 2) [see (7) and (8)]. ■

B. Generalized Stable Backward Recursion

The stable backward recursion is derived by using the smoothed probability of two sequential states, which is given by [5]

$$p\left(\mathcal{S}_t^{t+1}|\mathcal{Y}^{t+L}\right) = \frac{f\left(\mathcal{S}_t^{t+1}, \mathcal{Y}_{t+1}^{t+L}|\mathcal{Y}^t\right) f(s_{t+1}|\mathcal{Y}^{t+L})}{f\left(s_{t+1}, \mathcal{Y}_{t+1}^{t+L}|\mathcal{Y}^t\right)}. \tag{15}$$

Under the assumption that $\{\hat{\Lambda}_t, \mathbf{Y}_t\}$ are sufficient statistics for the next state-dependent conditional variance estimation, we obtain

$$\begin{aligned} f&(\mathcal{S}_t^{t+1}, \mathcal{Y}_{t+1}^{t+L}|\mathcal{Y}^t) \\ &= f(s_{t+1}, \mathcal{Y}_{t+1}^{t+L}|s_t, \mathcal{Y}^t)p(s_t|\mathcal{Y}^t) \\ &= f(\mathcal{Y}_{t+1}^{t+L}|\mathcal{S}_t^{t+1}, \mathcal{Y}^t)p(s_{t+1}|s_t, \mathcal{Y}^t)p(s_t|\mathcal{Y}^t) \\ &= f\Big(\mathcal{Y}_{t+1}^{t+L}|\mathcal{S}_t^{t+1}, \hat{\boldsymbol{\lambda}}_{t|t-1,s_t}, \mathbf{Y}_t\Big) a_{s_t s_{t+1}} p(s_t|\hat{\Lambda}_t, \mathbf{Y}_t) \end{aligned} \tag{16}$$

and

$$\begin{aligned} f&\left(s_{t+1}, \mathcal{Y}_{t+1}^{t+L}|\mathcal{Y}^t\right) \\ &= f\left(\mathcal{Y}_{t+1}^{t+L}|s_{t+1}, \mathcal{Y}^t\right) p(s_{t+1}|\mathcal{Y}^t) \\ &= f\left(\mathcal{Y}_{t+1}^{t+L}|s_{t+1}, \hat{\boldsymbol{\lambda}}_{t+1|t,s_{t+1}}\right) p(s_{t+1}|\hat{\Lambda}_t, \mathbf{Y}_t). \end{aligned} \tag{17}$$

By substituting (16) and (17) into (15) and integrating out all states at time $t+1$, we obtain the following backward recursion for the smoothed state probability:

$$p(s_t|\mathcal{Y}^{t+L}) = p(s_t|\hat{\Lambda}_t, \mathbf{Y}_t) \times \sum_{s_{t+1}} \frac{f\left(\mathcal{Y}_{t+1}^{t+L}|\mathcal{S}_t^{t+1}, \hat{\boldsymbol{\lambda}}_{t|t-1,s_t}, \mathbf{Y}_t\right) a_{s_t s_{t+1}} p(s_{t+1}|\mathcal{Y}^{t+L})}{f\left(\mathcal{Y}_{t+1}^{t+L}|s_{t+1}, \hat{\boldsymbol{\lambda}}_{t+1|t,s_{t+1}}\right) p(s_{t+1}|\hat{\Lambda}_t, \mathbf{Y}_t)} \tag{18}$$

where the conditional density $f(\mathcal{Y}_{t+1}^{t+L}|\mathcal{S}_t^{t+1}, \hat{\boldsymbol{\lambda}}_{t|t-1,s_t}, \mathbf{Y}_t)$ can be derived from the generalized backward recursion (5)–(8). However, the conditional density $f(\mathcal{Y}_{t+1}^{t+L}|s_{t+1}, \hat{\boldsymbol{\lambda}}_{t+1|t,s_{t+1}})$ in the denominator of (18) requires calculation of a similar recursion that is not informed of the regime s_t.

Although the stable backward recursion is known to be numerically more stable than the forward–backward recursions, the instability of the latter is insignificant for short delays and the former requires computation of the generalized backward recursion twice, one for evaluating $f(\mathcal{Y}_{t+1}^{t+L}|\mathcal{S}_t^{t+1}, \hat{\boldsymbol{\lambda}}_{t|t-1,s_t}, \mathbf{Y}_t)$ and one for $f(\mathcal{Y}_{t+1}^{t+L}|s_{t+1}, \hat{\boldsymbol{\lambda}}_{t+1|t,s_{t+1}})$.

IV. Experimental Results

The generalized state smoothing has been applied to state detection in noisy MSTF-GARCH (1,1) processes with three states and 5–15 dB signal-to-noise ratios (SNRs). Twenty random stationary models have been simulated with an unconditional Gaussian model and uniformly distributed parameters on the intervals (0,1/3], (1/3,2/3], and (2/3,1] for each state, respectively. For each model, 20 signals are considered, each of dimension $K = 100$ and time length $T = 100$. The conditional variances $\hat{\boldsymbol{\lambda}}_{t|t-1,s_t}$ are estimated using the recursive approach of [16]. Fig. 1 shows the detection error rate $p(\hat{s}_t \neq s_t)$ for casual estimation as well as for noncausal estimation with up to $L = 4$ samples delay. It can be seen that the state detection monotonically improves with the increase of the delay. However, the most significant improvement is achieved by using up to two future samples, and the contribution of additional future observations decays along time.

V. Conclusion

We have derived state smoothing for the MSTF-GARCH process, in which case the conditional variances depend on both past observations and the regime path. Our noncausal state probability solution generalizes both the standard forward–backward recursions and the stable backward recursion of HMP by capturing both the signal correlation along time and its conditioning on the regime path. Accordingly, the backward recursion requires two recursive steps for evaluating the conditional density of the given future observations corresponding to all optional future paths. Although the computational complexity of the generalized backward recursion grows exponentially with the delay, a small number of future observations contribute with the most significant improvement to the state estimation. Combining the generalized recursions with the recursive signal restoration algorithm of [16] facilitates a noncausal signal restoration, which is a subject for further research.

Acknowledgment

The authors would like to thank the anonymous reviewers for their helpful comments.

References

[1] Y. Ephraim and N. Merhav, "Hidden Markov processes," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1518–1569, Jun. 2002.

[2] J. D. Hamilton, *Time Series Analysis*. Princeton, NJ: Princeton Univ. Press, 1994.

[3] R. W. Chang and J. C. Hancock, "On receiver structures for channels having memory," *IEEE Trans. Inf. Theory*, vol. IT-12, no. 4, pp. 463–468, Oct. 1966.

[4] G. Lindgren, "Markov regime models for mixed distributions and switching regressions," *Scan. J. Statist.*, vol. 5, pp. 81–91, 1978.

[5] M. Askar and H. Derin, "A recursive algorithm for the Bayes solution of the smoothing problem," *IEEE Trans. Autom. Control*, vol. AC-26, no. 2, pp. 558–561, Apr. 1981.

[6] C. J. Kim, "Dynamic linear models with Markov-switching," *J. Economet.*, vol. 60, pp. 1–22, 1994.

[7] J. Cai, "A Markov model of switching-regime ARCH," *J. Bus. Econ. Statist.*, vol. 12, no. 3, pp. 309–316, Jul. 1994.

[8] J. D. Hamilton and R. Susmel, "Autoregressive conditional heteroskedasticity and changes in regime," *J. Economet.*, vol. 64, pp. 307–333, Jul. 1994.

[9] F. Klaassen, "Improving GARCH volatility forecasts with regime-switching GARCH," *Empirical Econ.*, vol. 27, no. 2, pp. 363–394, Mar. 2002.

[10] M. Haas, S. Mittnik, and M. S. Paolella, "A new approach to Markov-switching GARCH models," *J. Fin. Econometr.*, vol. 2, no. 4, pp. 493–530, Autumn 2004.

[11] H. Amiri, H. Amindavar, and R. L. Kirlin, "Array signal processing using GARCH noise modeling," in *Proc. 29th IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Montreal, QC, Canada, Mar. 18–23, 2004, pp. II-105–II-108.

[12] M. Abdolahi and H. Amindavar, "GARCH coefficients as feature for speech recognition in Persian isolated digit," in *Proc. 30th IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Philadelphia, PA, May 17–21, 2005, pp. I.957–I.960..

[13] I. Cohen, "From volatility modeling of financial time-series to stochastic modeling and enhancement of speech signals," in *Speech Enhancement*, J. Benesty, S. Makino, and J. Chen, Eds. New York: Springer-Verlag, 2005, ch. 5, pp. 97–114.

[14] ——, "Speech spectral modeling and enhancement based on autoregressive conditional heteroscedasticity models," *Signal Process.*, vol. 86, no. 4, pp. 698–709, Apr. 2006.

[15] C. Francq, M. Roussignol, and J.-M. ZakoïanZakoian, "Conditional heteroskedasticity driven by hidden Markov chains," *J. Time Series Anal.*, vol. 22, no. 2, pp. 197–220, 2001.

[16] A. Abramson and I. Cohen, "Recursive estimation of time-frequency GARCH process with hidden Markov regime," *IEEE Trans. Signal Process.*, submitted for publication.