Non-parametric estimation of intraday spot volatility: disentangling instantaneous trend and seasonality

Thibault Vatter*[1], Hau-Tieng Wu†[2], Valérie Chavez-Demoulin‡[1], and Bin Yu§[3]

[1]Faculty of Business and Economics (HEC), University of Lausanne, Switzerland.
[2]Department of Mathematics, Stanford University, USA.
[3]Department of Statistics, University of California, Berkeley, USA.

November 22, 2013

Abstract

We provide a new framework to model trends and periodic patterns in high-frequency financial data. Seeking adaptivity to ever changing market conditions, we enlarge the popular Fourier flexible form into a richer (time-varying) functional class and relax the assumptions from usual (static) intraday models in two directions. First, the realized volatility (i.e. the slowly time-varying component) is modelled as an instantaneous trend which evolves in real time. Second, the seasonality is no longer assumed constant over the sample. We provide the estimators associated with our class models and show that they have low variance and are essentially unbiased. As an application, we analyze the trajectories of the spot volatility in the foreign exchange market. Our results suggest that failing to factor in the seasonality's dynamic properties may lead to severe underestimation/overestimation of the intraday spot volatility.

Keywords. list of keywords High frequency market returns, Diffusion model, Volatility trajectories, Time-frequency analysis, Synchrosqueezing
JEL: C14, C22, C51, C52, G12, G17

*thibault.vatter@unil.ch
†hauwu@stanford.edu
‡valerie.chavez@unil.ch
§binyu@stat.berkeley.edu

1 Introduction

Over the last two decades, increasingly easier access to high frequency data has offered a magnifying glass to study financial markets. Analysis of these data poses unprecedented challenges to econometric modeling and statistical analysis. As for traditional financial time series (e.g. daily closing prices), the most critical variable at higher frequencies is arguably the asset return: from the pricing of derivatives to portfolio allocation and risk-management, it is a cornerstone both for academic research and practical applications. Because its properties are of such utmost importance, a vast literature sparked to find high-frequency models consistent with the new observed features of the data.

Among all characteristics of the return, its second moment structure is empirically found as preponderant, partly because of its power to asses market risk. Therefore, its time-varying nature has received a lot of attention in the literature (see e.g. Andersen et al. (2003)). While most low frequency stylized features find themselves at higher frequencies, an empirical regularity referring to seasonality is left unaddressed by most models designed to capture heteroskedasticity in financial time series. It is now well documented that seasonality, namely intraday patterns due to the cyclical nature of market activity, represents one of the main sources of model misspecification for GARCH-like or many other stochastic volatility models (see e.g. Guillaume et al. (1994); Andersen and Bollerslev (1997, 1998)).

More specifically, some periodicity in the volatility is inevitable, due, for instance, to market openings and closings around the world, but is unaccounted for in the vast majority of econometrics models. To avoid bias due to misspecified models, one possibility is to explicitly take into account the seasonality in traditional models (e.g. with the periodic-GARCH of Bollerslev and Ghysels (1996)). Alternatively, a pre-filtering step combined with a non-periodic model can be used as in Andersen and Bollerslev (1997, 1998); Boudt et al. (2011); Müller et al. (2011); Engle and Sokalska (2012). Even though an explicit inclusion of seasonality is advocated as more efficient (see Bollerslev and Ghysels (1996)), the resulting models are at the same time less flexible and computationally more expensive. While they may potentially propagate errors, two-step procedures are more convenient and often consistent if each step is consistent (see Newey and McFadden (1994)).

As a result, pre-filtering of high-frequency financial time series has received a considerable amount of attention in the literature dedicated to high-frequency financial time series. In this context, an attractive approach is to generalize the removal of weekends and holidays from daily data and use the "business clock" (see Dacorogna et al. (1993)); a new time scale where time goes faster when the market is inactive and conversely during "power hours", at the cost of synchronicity between assets. The other approach, which allows researchers to work in physical time (removing only closed market periods), is to model the periodic patterns directly, either non-parametrically with estimators of scale (e.g. in Martens et al. (2002); Boudt et al. (2011); Engle and Sokalska (2012)) or using smoothing methods (e.g. splines in Giot (2005)) or parametrically with the

Fourier Flexible Form (FFF) (see Gallant (1981)), introduced by Andersen and Bollerslev (1997, 1998) in the context of intraday volatility in financial markets.

Until now, the standard assumption in the literature is to consider a constant seasonality over the sample period. However, it is seldom tested in practice and few empirical studies acknowledge this issue. Notably, Deo et al. (2006) uses frequency leakage as evidence in favor of a slowly varying seasonality. Furthermore, Laakkonen (2007) observes that smaller sample periods yield improved seasonality estimators. In this paper, we argue that the constant assumption can only hold for arbitrary small sample periods, because the entire shape of the market (and its periodic patterns) evolves over time. Contrasting with seasonal adjustments usually considered in the literature, we relax the assumption of constant seasonality and suggest a non-parametric framework to obtain "instantaneous" estimates of trend and seasonality.

From a time-frequency decomposition, the proposed method extracts the instantaneous trend (estimated from the lowest frequencies) and seasonality (estimated from the highest frequencies). In fact, it is comparable to a "dynamic" combination of realized-volatility (RV) and Fourier Flexible Form (FFF). In this case, the instantaneous trend and seasonality can be estimated respectively using rolling moving averages (for the RV part) or rolling regressions (for the FFF part). Nonetheless, our model differs in three aspects. First, it disentangles in a single step the trend from the seasonality. Second, it yields naturally smooth pointwise estimates. Third, it is data-adaptive in the sense that there are less parameters to fine tune.

There is a major reason why dynamic seasonality models are important when modeling intraday returns, concerning practitioners and academics alike: non-dynamic models may lead to severe underestimation/overestimation of intraday spot volatility. Hence there are possible implications whenever seasonality pre-filtering is used as an intermediate step. From a risk management viewpoint, intraday measures such as intraday Value-at-Risk (VaR) and Expected-shortfall (ES) under the assumption of constant seasonality may suffer from inappropriate high quantile estimation. In other words, the assumption may lead to alternate periods of underestimated (respectively overestimated) VaR and ES when the seasonality is higher (respectively lower) than suggested by a constant model. In the context of jump detection, the assumption may induce an underestimation (respectively overestimation) of the number and size of jumps when the seasonality is higher (respectively lower).

The rest of the paper is organized as follows. In Section 2, we describe our model for the intraday return, from a continuous-time point of view to the discretized process. To relax the constant seasonality assumption, we define a class of seasonality models which includes the Fourier Flexible Form (FFF) as a special case. Indeed, a non-dynamic model such as the combined FFF and GARCH can usually be solved parametrically and we use the generalized method of moments (GMM) framework to derive asymptotic properties of a two-step estimation procedure. However, the conditions required to obtain consistent estimates highlight the technical difficulties encountered by traditional econometrics tools such as the GMM when applied to less trivial cases with time-

varying periodicities. In Section 3, we provide a detailed exposition of a method aimed at studying the class of models defined in Section 2. We start by recalling the link between the FFF and the Fourier transform and we conclude that neither the FFF nor the Fourier transform helps understanding the dynamics of the seasonality. We then sketch the theoretical basics of time-frequency analysis, that was introduced to overcome this limitation (see Daubechies (1992); Flandrin (1999)). Finally, we introduce the Synchrosqueezing transform, which is specially aimed at studying the class of models defined in Section 2, and we provide associated estimators of instantaneous trend and seasonality. In Section 4, we conduct a realistic simulation to study the properties of the estimators from Section 3. We show that they are essentially unbiased, have low variance and high signal-to-noise ratio. In Section 5, we estimate the model using 8 months of high-frequency foreign exchange data. To obtain confidence intervals for the estimated trend and seasonality, we develop tailor-made resampling procedures. We conclude in Section 6.

2 Intraday seasonality

Suppose the log-price follows the usual Brownian Semimartingale (BSM) diffusion model:

$$\mathrm{d}p(t) = \mu(t)\mathrm{d}t + \sigma(t)\mathrm{d}W(t)\,, \tag{1}$$

where $p(t)$ is the log-price, $\mathrm{d}p(t)$ consists of conditionally normal random process with mean $\mu(t)\mathrm{d}t$ and variance $\sigma(t)^2\mathrm{d}t$ and $W(t)$ the standard Brownian motion which is independent of $\sigma(t)$. It is common (see Boudt et al. (2011) and references therein) to assume a multiplicative decomposition of the spot volatility $\sigma(t)$ such as

$$\sigma(t) = \sigma_1(t)\sigma_2(t), \tag{2}$$

where $\sigma_1(t)$ is slowly varying and $\sigma_2(t)$ is fast varying. For example, defining a window of length w_1 around time t, a popular choice of $\sigma_1(t)$ is the square root of the quadratic covariation over the window around t,

$$\sigma_1^2(t) = \int_{t-w_1/2}^{t+w_1/2} \sigma^2(s)\mathrm{d}s \tag{3}$$

which is the so-called integrated volatility in Barndorff-Nielsen and Shephard (2002). In this case, and provided that $\sigma(t)$ is smooth enough, we obtain an approximation of the multiplicative decomposition of σ as

$$\sigma(t) = \sigma_1(t)\sigma_2(t) + O(w_1),$$

However, this specification for $\sigma_1(t)$ is only one of the many possibilities for the "trend" part of the volatility, which is intuitively a very "low-frequency" function.

As for the second component, it is can be specified as an exponential form

$$\sigma_2^2(t) = e^{s(t)}, \tag{4}$$

where $s(t)$ is a periodic function over an interval w_2 (i.e. $s(t + w_2) = s(t)$) and a normalization condition is then needed for $\sigma_2(t)$ to make this decomposition unique. As noted for example in Deo et al. (2006); Müller et al. (2011), the exponential frees $s(t)$ from being restricted to $\mathbb{R}_+$, allowing additional modelling possibilities. To be more precise, s can be modeled as

$$s(t) = Af(\xi_0 t), \tag{5}$$

where $f \in C([0, 1])$ is a 1 periodic continuous function called the *wave shape function* (Wu, 2013), A is the *amplitude* and $\xi_0 > 0$ the *frequency* of the periodic function. One particular example of f is the cosine function. Essentially, under this model, the signal $s(t)$ oscillates ξ_0 times every unit of time, and each oscillation follows the wave-shape f. Then assuming a constant daily pattern implies $\xi_0 = 1$ and a natural estimator of $s(t)$ can be defined as the time-of-day conditional expectation of the log-squared and demeaned return (up to a normalization condition).

2.1 The adaptive harmonic model

As useful as the periodicity model (5) may be, both the amplitude and the shape function itself might vary according to time, due to the inherited dynamics of the natural system. In this case, (5) is mis-characterized and model bias may be inevitable. To alleviate this limitation, we consider a different approach, based on a time-frequency decomposition of the signal.

Intuitively, to capture the dynamics, we may write

$$s(t) = \sum_{k=1}^{K} a_k(t) \cos(2\pi\phi_k(t)), \tag{6}$$

where $K \in \mathbb{N}$, $a_k(t) > 0$ and $\phi_k'(t) > 0$. The function $a_k(t)$ is called the *amplitude modulation* (AM), $\phi_k(t)$ the *phase function* and $\phi_k'(t)$ the *instantaneous frequencies* (IF) of the k-th component. The model is therefore a sum of Amplitude Modulated-Frequency Modulated (AM-FM) periodic components. If the phases are linear ($2\pi xi_k t + \eta_k$ with $xi_k > 0$ and $\eta_k \in \mathbb{R}$) and amplitude modulations are constant ($a_k > 0$), then the model reduces to the well-known harmonic model (see section 3). Furthermore, model (5) is a special case of (6) with $f(t) = \sum_{k=1}^{K} a_k \cos(2\pi k\xi_0 t + \eta_k)$.

When the phases are non-linear, IF generalizes the concept of frequency, capturing the number of oscillations one observes during an infinitesimal time period. The same reasoning applies to the amplitude modulation, which represents the "instantaneous" magnitude of the oscillation. While those time-varying quantities allow to capture momentary behavior, there is in general no unique representation for an arbitrary s satisfying (6), even if $K = 1$.

Indeed, there exists an infinity of smooth pairs of functions $\alpha(t)$ and $\beta(t)$ so that $\cos(t) = (1+\alpha(t))\cos(t+\beta(t))$, $1+\alpha(t) > 0$ and $1+\beta'(t) > 0$. This is known as the identifiability problem studied in Chen et al. (2013). To resolve this issue, it is necessary to restrict the functional class and we borrow the following definition from Chen et al. (2013):

Definition 2.1 (Intrinsic Mode Function Class $\mathcal{A}_\epsilon^{c_1,c_2}$)**.** For fixed choices of $0 < \epsilon \ll 1$ and $\epsilon \ll c_1 < c_2 < \infty$, the space $\mathcal{A}_\epsilon^{c_1,c_2}$ of Intrinsic Mode Functions (IMFs) consists of functions $f : \mathbb{R} \to \mathbb{R}$, $f \in C^1(\mathbb{R}) \cap L_\infty(\mathbb{R})$ having the form

$$f(t) = A(t)\cos(2\pi\phi(t)), \tag{7}$$

where $A : \mathbb{R} \to \mathbb{R}$ and $\phi : \mathbb{R} \to \mathbb{R}$ satisfy the following conditions for all $t \in \mathbb{R}$:

$$\begin{aligned} A \in C^1(\mathbb{R}) \cap L_\infty(\mathbb{R}), \quad & \inf_{t\in\mathbb{R}} A(t) > c_1, \quad \sup_{t\in\mathbb{R}} A(t) < c_2, \\ \phi \in C^2(\mathbb{R}), \quad & \inf_{t\in\mathbb{R}} \phi'(t) > c_1, \quad \sup_{t\in\mathbb{R}} \phi'(t) < c_2, \\ |A'(t)| \le \epsilon\,|\phi'(t)|, \quad & |\phi''(t)| \le \epsilon\,|\phi'(t)|. \end{aligned}$$

This definition describes functions composed of a single AM-FM function, referred to as IMF, mainly satisfying two requirements. First, the AM and IF are both continuously differentiable and bounded from above and below. Second, the rate of variation of both the AM and IF are small compared to the IF itself. As these conditions exclude jumps in AM and IF, we make the working assumption that the market evolves slowly over time and we do not try to model abrupt structural changes. In fact, because usual tests for structural breaks are aimed at detecting if and where a break occurs, they do not help for the class $\mathcal{A}_\epsilon^{c_1,c_2}$ that does not feature any. To asses its validity when modelling financial data, we will assume the model and resort to confidence interval estimation.

The identifiability theorem for $\mathcal{A}_\epsilon^{c_1,c_2}$ is stated in the following way (Chen et al., 2013, Theorem 2.1). Suppose $f(t) = a(t)\cos\phi(t) \in \mathcal{A}_\epsilon^{c_1,c_2}$ can be represented in a different form, that is, $f(t) = A(t)\cos\varphi(t)$, which also satisfies the conditions of $\mathcal{A}_\epsilon^{c_1,c_2}$. Define $\alpha(t) = \varphi(t) - \phi(t)$ and $\beta(t) = A(t) - a(t)$, then $\alpha \in C^2(\mathbb{R})$, $\beta \in C^1(\mathbb{R})$ and $|\alpha'(t)| \le C\epsilon$, $|\alpha(t)| \le C\epsilon$ and $|\beta(t)| < C\epsilon$ for all $t \in \mathbb{R}$ for some universal constant C depending only on c_1. In a nutshell, if a member of $\mathcal{A}_\epsilon^{c_1,c_2}$ can be represented in two different forms, then the differences in its phase function, AM and FM between the two forms are controllable by the small model constant ϵ.

Note that the dynamic model from equation (6) was a sum of more than one IMF. To resolve the identifiability problem in this case, further conditions are necessary and we borrow another definition from Chen et al. (2013):

Definition 2.2 (Superposition of IMFs $\mathcal{A}_{\epsilon,d}^{c_1,c_2}$)**.** Fix $0 < d < 1$. The space $\mathcal{A}_{\epsilon,d}^{c_1,c_2}$ of superpositions of IMFs consists of functions f having the form

$$f(t) = \sum_{k=1}^{K} f_k(t) \tag{8}$$

for some finite $K > 0$ such that for each $k = 1, \ldots, K$, $f_k(t) = A_k(t)\cos(2\pi\phi_k(t)) \in \mathcal{A}_\epsilon^{c_1,c_2}$ and $\phi_k(t)$ satisfies

$$\phi_k'(t) > \phi_{k-1}'(t) \quad \text{and} \quad \phi_k'(t) - \phi_{k-1}'(t) \geq d\left[\phi_k'(t) + \phi_{k-1}'(t)\right] \tag{9}$$

for all $t \in \mathbb{R}$.

Functions in the class $\mathcal{A}_{\epsilon,d}^{c_1,c_2}$ are composed of more than one IMF satisfying the condition (9), which has the following natural interpretation: to be resolved simultaneously, different IMFs of the superposition must be well-enough separated in their IFs, but no further restriction on the AM is required. In this case, an identifiability theorem similar to the one for $\mathcal{A}_\epsilon^{c_1,c_2}$ exists (Chen et al., 2013, Theorem 2.2): if a member of $\mathcal{A}_{\epsilon,d}^{c_1,c_2}$ can be represented in two different forms, then the two forms have the same number of IMFs, and the differences in their phase function, AM and FM for each IMF are small. To summarize, with the $\mathcal{A}_{\epsilon,d}^{c_1,c_2}$ model and its identifiability theorem, the IFs and AMs are well defined up to an uncertainty of order ϵ and we refer the reader to Chen et al. (2013) for precise statements of the theorems and more discussions.

Using $s(t) \in \mathcal{A}_{\epsilon,d}^{c_1,c_2}$ and the multiplicative decomposition from equation (2), we consider the *adaptive harmonic model*

$$\log\sigma^2(t) = 2\log\sigma_1(t) + s(t), \tag{10}$$

to capture the dynamics of the spot volatility $\sigma^2(t)$, where

$$T(t) = 2\log\sigma_1(t) \tag{11}$$

is called the *trend function* and satisfies the following technical condition:

(T1) Fix a smooth function $\psi \in C^\infty(\mathbb{R})$ which decays faster than any polynomial at infinity, as do all its derivatives. Denote by F_ψ its Fourier transform and suppose that $\text{supp}F_\psi \subset [1-\Delta, 1+\Delta]$, where $0 < \Delta < 1$. Assume that $T(t) : \mathbb{R} \to \mathbb{R}$ is in $C^1(\mathbb{R})$ so that its Fourier transform exists, and $\left|\int T(t)\frac{1}{\sqrt{a}}\psi(\frac{t-b}{a})\mathrm{d}t\right| \leq C_T\epsilon$, $\left|\int T'(t)\frac{1}{\sqrt{a}}\psi(\frac{t-b}{a})\mathrm{d}t\right| \leq C_T\epsilon$ for all $b \in \mathbb{R}$ and $a \in (0, \frac{1+\Delta}{c_1}]$, for some $C_T \geq 0$.

As we ideally want a trend to be slowly time-varying, the intuition behind the technical condition (T1) is as a bound on how fast a function oscillates locally. A special case satisfying (T1) is a continuous function T for which its Fourier transform F_T exists and is compactly supported in $\left(-\frac{1-\Delta}{1+\Delta}c_1, \frac{1-\Delta}{1+\Delta}c_1\right)$.

There are two well-known examples of such trend functions; the first is the polynomial function $s(x) = \sum_{l=1}^{L}\alpha_l x^l$, where $\alpha_l \in \mathbb{R}$, which is commonly applied to model trends. By a direct calculation, its Fourier transform is supported at 0. The second is a harmonic function $s(x) = \cos(\xi_0 x)$ with "very low" frequency (i.e. with $|\xi_0| < \frac{1-\Delta}{1+\Delta}c_1$), as its Fourier transform is supported at ξ_0.

More generally, one can verify using Plancheral theorem that $\left(-\frac{1-\Delta}{1+\Delta}c_1, \frac{1-\Delta}{1+\Delta}c_1\right)$ implies $\int T(t)\frac{1}{\sqrt{a}}\psi(\frac{t-b}{a})(t)\mathrm{d}t = 0$ for all $a \in (0, \frac{1+\Delta}{c_1}]$ and all $b \in \mathbb{R}$. In our case

however, the trend is non-parametric by assumption and its Fourier transform might in all generality be supported everywhere in the Fourier domain. Then the conditions in (T1) describe a trend that essentially captures the slowly varying features of those models (i.e. its local behavior is similar to that of a polynomial or a very low frequency periodic function) but is more general.

2.2 A special case: the Fourier Flexible Form

A special case of the adaptive harmonic model (10) is the popular Fourier Flexible Form (FFF) (see Gallant (1981)), introduced by Andersen and Bollerslev (1997, 1998) in this context. Using Fourier series to decompose any periodic function into simple oscillating building blocks with an arbitrarily small prediction error, a simplified version of the FFF without additional dummy variables reads

$$\log \sigma_1 = s_0/2, \quad s(t) = \sum_{k=1}^{K} \left[a_k \cos(2\pi k t) + b_k \sin(2\pi k t) \right]. \tag{12}$$

The model (12) contains two terms: a constant trend s_0 and a sum of sines and cosines with integer frequencies, designed to capture the daily oscillations around the base level. For $d \leq 1/(2K-1)$, we can view $s(t)$ as a superposition of IMFs (i.e. a member of $\mathcal{A}_{\epsilon,d}^{c_1,c_2}$) with $\phi_k(t) = 2\pi k t + \eta_k$ and $a_k(t) = a_k$ for $k = 1, \ldots, K$.

It should be noted that in Andersen and Bollerslev (1997, 1998), the authors consider the addition of dummy variables to capture weekday effects or particular events such as holidays in particular markets, but also unemployment reports, retail sales figures, etc[1]. While the periodic model captures most of the seasonal patterns, their dummy variables allow to quantify the relative importance of calendar effects and announcement events.

In the following, we build on FFF for several reasons. First, it is empirically found to be a better in-sample estimator of seasonality than scale estimators and smoothing methods (Laakkonen (2007)) and also provides better out-of-sample forecasts (Martens et al. (2002)). Second, the parameters can be conveniently interpreted and/or used in tests for structural breaks (i.e. violation of the constant seasonality assumption). Third, as $\mathcal{A}_{\epsilon,d}^{c_1,c_2}$ is its natural generalization, it can be readily extended to more complicated continuous-time models $\mathcal{A}_{\epsilon,d}^{c_1,c_2}$ of seasonality. Before we move on to illustrate how such a model can be estimated using traditional statistical methods and how those are inadequate for more involved models from the $\mathcal{A}_{\epsilon,d}^{c_1,c_2}$ class, it is necessary to fill the gap between this continuous-time construction and discrete-time processes.

[1] In fact, it is also possible to add a dependence on the daily volatility level but it is discarded here because empirically insignificant both in equity and spot foreign exchange data (see Andersen and Bollerslev (1997, 1998).

2.3 Discrete-time approximation

Although trading happens in continuous time, the price process is only collected at discrete points in time. Hence, it is natural to consider a sampling of the BSM model (1) at discrete time stamps. Excluding overnights and week-ends, we assume that our sample contains D trading days of P equally-spaced intraday observations. With a sampling interval defined as $\tau = 1/P$ and letting $p \in \{1, \cdots, P\}$ and $d \in \{1, \cdots, D\}$, there is a unique correspondence between each index $n = P(d-1) + p \in \{1, \cdots, N\}$ with $N = PD$, and the sampling time $t = n\tau$ (in daily units).

At the sampling time $t = n\tau$, we therefore denote the return, mean return, spot volatility and diffusion term by

$$\begin{aligned} r_n &= \log p(n\tau) - \log p((n-1)\tau), \quad \mu_n = \mu(n\tau), \\ \sigma_n &= \sigma(n\tau) \quad \text{and} \quad w_n = W(n\tau) - W((n-1)\tau), \end{aligned} \tag{13}$$

where $n = 1, \ldots, N$. Clearly, the sampling effect leads to

$$r_n \neq \tau\mu_n + \sigma_n w_n.$$

Ideally, if the full information of the process was observable, we might discretize it by

$$\widetilde{\mu}_n = \int_{(n-1)\tau}^{n\tau} \mu(u)\mathrm{d}u \quad \text{and} \quad \widetilde{\sigma\mathrm{d}W}_n = \int_{(n-1)\tau}^{n\tau} \sigma(u)\mathrm{d}W(u), \tag{14}$$

where $n = 2, \ldots, N$. By direct calculation, we obtain

$$r_n = \widetilde{\mu}_n + \widetilde{\sigma\mathrm{d}W}_n,$$

for all $n = 2, \ldots, N$. To fill the gap between the sampling scheme (13) and the discretization scheme (14), we show in Appendix A that

$$r_n \approx \mu_n\tau + \sigma_n w_n \tag{15}$$

holds uniformly in probability.

In the following, we further assume a constant mean return[2], that is $\mu(t) = \mu$ for all $t \in \mathbb{R}$. We mention that under this assumption, when τ is small enough, the term $\mu_n\tau$ is negligible. Putting all the above facts together, we define the (spot) *log-volatility process* y_n as

$$y_n = \log(r_n - \mu\tau)^2 \; = \; T_n + s_n + \log\left(w_n^2\right), \tag{16}$$

where $T(t) = 2\log\sigma_1(t)$ for all $t \in \mathbb{R}$, $T_n = T(n\tau)$ and $s_n = s(n\tau)$. In other words, y_n is the discretized version of a noisy adaptive harmonic model (10),

[2]This assumption for the mean return is reasonable as both its value and variability are usually small for financial data sampled from a few minutes to hours.

that is, a sum of an well separated IMFs and a trend satisfying (T1), plus a disturbance term[3].

Within this framework, every specific choice of trend T, seasonality s and noise w yields a different model. Due to their non-parametric nature, estimating T, s or the AM and IF inside s for a generic σ satisfying the adaptive harmonic model (10) is non-trivial. In the next subsection, we solve the special case where s is an FFF and w is a GARCH(1,1), which we can reduce to a parametric problem. Finishing with a critique of the models with constant seasonality, we motivate the need for a generalized approach that we expose in the remainder of this paper.

2.4 Fourier Flexible Form-GARCH asymptotics and limitations

Without loss of generality, we assume a window length of one day ($w_1 = 1$). From the daily GARCH in Andersen and Bollerslev (1997, 1998) to commercial factor models in Engle and Sokalska (2012), various alternatives have been proposed for the trend T_n. Referring to the growing literature on realized volatility measures (see e.g. Andersen et al. (2007) and Barndorff-Nielsen and Shephard (2007) for surveys), another natural estimator for T_n is obtained by discretizing of equation (3).

As for any linear model with known predictors, the Fourier Flexible Form (FFF) from (12) can be easily estimated by regressing the predictors on $y_n - \widehat{T_n}$. In the BSM model, the diffusion process is assumed to be the standard Brownian motion. Then although w_n is i.i.d. Gaussian, the disturbance $\log w_n^2$ that appears in the regression is not. While Boudt et al. (2011) propose a truncated maximum likelihood estimator (MLE) of the FFF under the assumption that w_n is Gaussian, empirical evidence suggests that the residual noise process is heteroskedastic (see Andersen and Bollerslev (1997, 1998); Martens et al. (2002); Engle and Sokalska (2012)). Hence we propose a two-step approach that combines the ordinary least-squares (OLS) to estimate the FFF and the maximum likelihood estimator of a common model to capture heteroskedasticity, namely the Gaussian GARCH(p,q) (Engle (1982); Bollerslev (1986) or see Bollerslev et al. (1992, 1994) for surveys). As usual standard errors at each step are asymptotically inconsistent and must be revised, we explicitly take this issue into account.

Instead of an i.i.d. Gaussian process for the underlying noise w_n, we write

$$w_n = v_n \eta_n \text{ with } \begin{cases} v_n^2 = \omega + \sum_{i=1}^{p} a\, w_{n-i}^2 + \sum_{j=1}^{q} b\, v_{n-j}^2 \\ \eta_n \sim N(0,1), \end{cases} \tag{17}$$

where $\omega > 0$, $a_i, b_j \geq 0$ for all $i \in \{1, \cdots, p\}$ and $j \in \{1, \cdots, q\}$. While the GARCH(p,q) has been generalized in various ways to capture different proper-

[3]Under the BSM model assumption and the discretization (13), w_n is an i.i.d. Gaussian process, but it can be modified to further capture the properties of financial data and is discussed in the following sections.

ties of financial returns (see Bollerslev et al. (1992, 1994) for surveys), we restrict ourselves to the GARCH(1,1), parametrized only by (ω, a, b). The reason is that it is parcimonious, easy to estimate and that it also performs reasonably well both at lower (Lunde and Hansen (2005)) and higher (see Giot (2005); Engle and Sokalska (2012)) frequencies; the latter after correcting for seasonality. However, our derivation of the consistency and asymptotic normality of this two-step estimation holds for more complex GARCH models.

In the remainder of this Subsection, we use

$$z_n = (r_n - \widehat{\mu}\tau)e^{-\widehat{T}_n/2}, \tag{18}$$

as the "data", where $\widehat{\mu}$ is the sample mean and $\widehat{T}_n$ a previously estimated trend. In all generality, the estimation of those two quantities also influences the the asymptotic standard errors that we derive below, which are in fact lower bounds. For the seasonality, we use $\boldsymbol{\beta}$ and $\boldsymbol{x_n}$ as the vectors of parameters, respectively predictors, that is,

$$\begin{aligned} \boldsymbol{\beta} &= (a_1, b_1, \ldots, a_K, b_K)^\top \in \mathbb{R}^{2K} \\ \boldsymbol{x_n} &= (\cos(2\pi n\tau), \sin(2\pi n\tau), \ldots, \cos(2\pi Kn\tau), \sin(2\pi Kn\tau))^\top \in [-1,1]^{2K}, \end{aligned}$$

where $n = 1, \ldots, N$. Then the FFF-GARCH(1,1) likelihood is

$$L(z_1, \cdots, z_N; \boldsymbol{\beta}, \boldsymbol{\theta}) = \Pi_{n=1}^N \frac{e^{-\frac{w_n^2}{2v_n^2}}}{\sqrt{2\pi v_n^2}} \text{ with } \begin{cases} w_n = z_n e^{-\frac{\boldsymbol{x_n}^\top \boldsymbol{\beta}}{2}} \\ v_n^2 = \omega + a\,\epsilon_{n-1}^2 + b\,v_{n-1}^2, \end{cases} \tag{19}$$

with a new vector of parameters $\boldsymbol{\theta} = (\omega, a, b)$. In the two-step case, $\boldsymbol{\theta}$ is estimated by replacing the true parameter $\boldsymbol{\beta}$ by a first step estimate and the GARCH MLE standard errors also need to be revised.

To derive asymptotically valid standard errors for the two-step procedure described above, the generalized method of moments (GMM) is a natural framework (see Chapter 6 of Newey and McFadden (1994)). We closely follow Engle and Sokalska (2012), where the same approach is carried out for the time-of-day seasonality estimator. Stacking all $2K$ FFF and 3 GARCH parameters in $\boldsymbol{\Theta} = (\boldsymbol{\beta}, \boldsymbol{\theta})$, we define the vector of moments conditions as

$$g(\boldsymbol{\Theta}) = \begin{pmatrix} \mathbb{E}[g_1] \\ \mathbb{E}[g_2] \end{pmatrix} \text{ with } \begin{cases} g_1 = \boldsymbol{x_n} \circ (y_n\,\mathbf{1} - \boldsymbol{x_n} \circ \boldsymbol{\beta}) \\ g_2 = \nabla_{\boldsymbol{\theta}} \left(\log v_n^2 + \frac{z_n^2}{v_n^2} e^{-x_n^\top \boldsymbol{\beta}} \right), \end{cases} \tag{20}$$

where $\circ$ represents an element-wise multiplication and $\mathbf{1}$ is a $2K$ vector of ones. As for g_1 and g_2, they represent respectively estimates at the first and second optimization steps cast in terms of moment conditions, namely predictors and residuals that are orthogonal for the OLS and the log-likelihood's gradient that equals zero for the MLE. Including the dummy variable part of the FFF does not change the derivation that follows, because non-linear least squares can also be cast in term of moments conditions.

To obtain $\widehat{\boldsymbol{\Theta}}$, we replace $g(\boldsymbol{\Theta})$ by its sample sum counterpart $\widehat{g}(\boldsymbol{\Theta})$. Furthermore, we use the identity as GMM weighting because the system is just-identified (i.e. the number of parameters and estimation equations is equal). In this case, the estimator is

$$\widehat{\boldsymbol{\Theta}} = \operatorname{argmin} \widehat{g}(\boldsymbol{\Theta})^{\top} \widehat{g}(\boldsymbol{\Theta}). \tag{21}$$

In other words, $\widehat{\boldsymbol{\Theta}}$ minimizes the Euclidean distance between the sample moments and their population counterpart, that is zero. Under some mild regularity conditions for $g(\boldsymbol{\Theta})$ and with $\boldsymbol{\Theta}_0 = (\boldsymbol{\beta}_0, \boldsymbol{\theta}_0)$ the true vector of parameters, Theorem 6.1 of Newey and McFadden (1994) establishes the consistency and asymptotic normality of (21) as a special case of Theorem 3.4 in Newey and McFadden (1994):

$$\sqrt{N}\left(\widehat{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_0\right) \xrightarrow{d} N\left(0, G^{-1}\Omega G^{-1\top}\right), \tag{22}$$
$$\text{with } G = \mathbb{E}\begin{pmatrix} \nabla_{\boldsymbol{\beta}} g_1 & 0 \\ \nabla_{\boldsymbol{\beta}} g_2 & \nabla_{\boldsymbol{\theta}} g_2 \end{pmatrix}\Bigg|_{\boldsymbol{\Theta}_0} \text{ and } \Omega = \mathbb{E}\begin{pmatrix} g_1^2 & g_1 g_2 \\ g_1 g_2 & g_2^2 \end{pmatrix}\Bigg|_{\boldsymbol{\Theta}_0}$$

For this theorem to apply, the main requirement is the consistency of both $\widehat{\boldsymbol{\beta}}$ and $\widehat{\boldsymbol{\theta}}$, inherited in this case from the OLS and MLE, respectively. Because the predictors $\boldsymbol{x_n}$ are exogeneous and orthogonal to each other (no collinearity), consistency of $\widehat{\boldsymbol{\beta}}$ is inherited from the OLS. As for the GARCH part, compactness of the parameter space and stationarity of the process (Bollerslev (1986)) imply consistency of the MLE. Finally, It should be noted that the matrices G and Ω can be estimated consistently using the sample sum and $\widehat{\boldsymbol{\Theta}}$ (see Hansen (1982)).

In practice, this two-step construction is mainly useful if all its assumptions hold. While the heteroskedasticity of the data is undeniable, the constant seasonality part is more puzzling, as it is arguably challenged by empirical evidence. In Laakkonen (2007), the author observes that an FFF estimated yearly, then quarterly and finally weekly performs increasingly better at capturing periodicities in the data. Whereas this sub-sampling is interpretable as a varying seasonality, it is still piecewise constant. Because a piecewise constant function implies that the system under study undergoes structural changes (i.e. shocks) at every break point, it is an unlikely candidate for the market's seasonality. We take the point of view of Deo et al. (2006), where the authors use frequency leakage in the power spectrum of the absolute return to argue that the seasonality is not (piecewise) constant but slowly varying. While they allow non-integer frequencies, their "generalized FFF" parameters are still constant for the whole sample. In other words, they correct for frequency leakage, but do not allow for potential slowly varying frequencies and amplitude modulations (i.e. time-varying FFF parameters).

Dynamic alternatives include either arbitrarily small intervals or a rolling version of their "generalized FFF". In the first case, an arbitrarily increase in the estimator's variance (as in the issue of testing for a general smooth member

of $\mathcal{A}_{\epsilon,d}^{c_1,c_2}$) is expected. In the second case (that we compare to our method in Section 5), it is not clear a priori how one would choose the optimal window size.

In the next Section, we propose an alternative framework to study the time-frequency properties of the volatility. Allowing the trend and seasonality to evolve dynamically over time, we are able to obtain T and s in the adaptive harmonic model in a single step. In fact, if the log-volatility process $\{y_n\}$ is stationary or close to stationary in some sense, a transform called the Synchrosqueezing transform (SST) theoretically helps to obtain accurate pointwise estimates of T satisfying (T1) and $s \in \mathcal{A}_{\epsilon,d}^{c_1,c_2}$ (Daubechies et al. (2011); Chen et al. (2013)).

3 From Fourier to Synchrosqueezing

In this Section, we start by recalling the relationship between the harmonic model and the Fourier transform. Then we discuss the adaptive harmonic model and sketch the theoretical basics of time-frequency analysis. Finally, we present the Synchrosqueezing transform (SST), which is specially aimed at studying the adaptive harmonic model. Its theoretical properties are discussed in the end.

3.1 Fourier transform and the harmonic model

Oscillatory signals are ubiquitous in many fields. Seasonality is a synonym of oscillation broadly used in finance, econometrics, public health, biomedicine, etc. There are several "characteristics" one can use to describe an oscillatory signal (see Flandrin (1999)): for example, how often an oscillation repeats itself *on average*, how large an oscillation is *on average*, etc. To capture these features, the following *harmonic model* is widely used:

$$x(t) = \sum_{k=1}^{K} a_k \cos(2\pi\xi_k t + \eta_k), \tag{23}$$

where $\sum a_k^2 < \infty$, $a_k > 0$, $\xi_k > 0$ and $\eta_k \in \mathbb{R}$. The quantities a_k, ξ_k and η_k are respectively called the *amplitude*, the *frequency* and the *initial phase* of the k-th component of x. As we stated in Section 2, both the FFF and its generalized version from Deo et al. (2006) are based on this model. When a given signal satisfies model (23), the Fourier transform is helpful in determining the oscillatory features a_k and ξ_k. For $x(t)$ defined in (23), we recall (see (Folland, 1999, Chapter 9)) that its Fourier transform is

$$F_x = \frac{1}{2}\sum_{k=1}^{K} a_k e^{i\eta_k}\delta_{\xi_k} + \frac{1}{2}\sum_{k=1}^{K} a_k e^{-i\eta_k}\delta_{-\xi_k},$$

where δ is the Dirac measure. Thus, if a signal satisfies the harmonic model, the frequency ξ_k and the amplitude a_k can be read directly from its Fourier

transform. For more details on the property of the Fourier transform $F_x(\omega)$ of a function $x(t)$ with $t, \omega \in \mathbb{R}$, we refer the reader to (Folland, 1999, Chapter 9).

As useful as the harmonic model can be, it does not fit well many real signals. Indeed, due to the nonlinear nature of many systems, the oscillatory pattern might change over time. For instance, the signal may oscillate more often in the beginning but less in the end, or oscillate stronger in the beginning but weaker in the end. We call these time-varying patterns the dynamics of the seasonality, and it is well known that capturing the dynamics is beyond the scope of the harmonic model. Unfortunately, when departing from static models such as (23) to consider dynamic models such as the whole class $\mathcal{A}_{\epsilon,d}^{c_1,c_2}$, the usefulness of the Fourier transform is also limited due to the loss of time localization of events. In other words, because of its incapacity to capture momentary behavior, we cannot use the Fourier transform to analyze the dynamics of seasonality.

3.2 Time-frequency analysis

To overcome the limitation of the Fourier transform, time-frequency analysis was introduced (see Daubechies (1992); Flandrin (1999) and the references inside for the historical discussion). Heuristically, to capture the dynamics of the oscillatory signal, we take a small subset of the given signal $x(t)$ around a given time t_0, and analyze its power spectrum. Ideally, this provides information about the local behavior of $x(t)$ around t_0. We repeat the analysis at all times, and the result is referred to as the *time-frequency (TF) representation* of the signal. Mathematically, this idea is implemented by the short-time Fourier transform (STFT) and continuous wavelet transform (CWT). The STFT is defined by

$$G_x(t,\omega) = \int x(s)\, g(s-t)\, e^{-i2\pi\omega s} \mathrm{d}s \ , \tag{24}$$

where $\omega > 0$ is frequency, and g is the *window function* associated with the STFT which extracts the local signal for analysis. g has to be smooth enough and such that $\int g(t)\mathrm{d}t = 1$. The CWT is defined by

$$W_x(t,a) = \int x(s)\, a^{-1/2}\, \overline{\psi}\left(\frac{s-t}{a}\right) \mathrm{d}s \ , \tag{25}$$

where $a > 0$. We call a the wavelet's *scale* and ψ the *mother wavelet* which is smooth enough and such that $\int \psi(t)\mathrm{d}t = 0$. In a nutshell, we select a subset of the whole signal by the window function or the mother wavelet, and study how the signal oscillates locally. From the viewpoint of frame analysis, the STFT and CWT are defined as the inner products between the signal to be analyzed and a pre-assigned family of "templates", or atoms – which are $\{g(\cdot - t)e^{-i2\pi\omega\cdot}\}_{\omega\in\mathbb{R}^+,t\in\mathbb{R}}$ in STFT and $\{a^{-1/2}\,\overline{\psi}\left(\frac{\cdot - t}{a}\right)\}_{a\in\mathbb{R}^+,t\in\mathbb{R}}$ in CWT. Furthermore, this inner product provides a way to resolve events simultaneously in time (by translation) and frequency (by modulation and dilatation respectively), but the properties of their respective templates yield different trade-offs

between temporal and frequency resolution due to the Heisenberg uncertainty principle. Take the high frequency region for example. The STFT offers a better frequency resolution in frequency events but an inferior resolution in time. In contrast, the CWT provides an inferior resolution in frequency but a better resolution in time.

To further understand the time-frequency (TF) analysis, we consider the continuous wavelet transform (CWT) of a purely harmonic signal $x(t) = A\cos(2\pi\omega t)$, and the same reasoning follows for the short-time Fourier transform (STFT). From the Plancheral theorem, the CWT of x becomes

$$W_x(t,a) = Aa^{1/2}\overline{F_\psi}(\omega a)e^{i2\pi\omega t}. \tag{26}$$

We observe that if $F_\psi(\omega)$ is concentrated around $\omega = 1$, then $W_x(t,a)$ is concentrated around $a = 1/\omega$.

In the remainder of this section, we focus on the CWT, and assume $x = T+s$, where T satisfies (T1) and $s \in \mathcal{A}_{\epsilon,d}^{c_1;c_2}$, and take its discretization $\boldsymbol{x} = \{x_n\}_{n\in\mathbb{Z}}$, where $x_n = x(n\tau)$ and $\tau > 0$ is the sampling interval. In other words, we assume the model (16) without $\log\left(w_n^2\right)$. As discussed in Section 2.3, we only have access to the discretized sample data. We thus discretize the CWT for $a > 0$ and $n \in \mathbb{Z}$ by

$$W_{\boldsymbol{x}}(n\tau, a) := \tau \sum_{m\in\mathbb{Z}} \frac{x_m}{a^{1/2}}\bar{\psi}\Big(\frac{m\tau - n\tau}{a}\Big).$$

In Figure 1, we show a simulated dataset to be used in the remainder of this section. The contaminated signal is the sum of a trend, an amplitude modulated-frequency modulated-periodic component and an additive GARCH(1,1) noise. More precisely, we consider the following functions

$$\begin{aligned} T(t) &= 1 + 0.2t + 2\exp\left[-(t-20)^2\right], \\ s(t) &= \left(1 + \cos(t/(2\pi))^2\right)\cos(2\pi\phi(t)), \\ \phi(t) &= t + t^2/40, \end{aligned}$$

and obtain the sampled time series

$$x_n = T(n\tau) + s(n\tau) + w_n,$$

where $n \in \{1, \cdots, N\}$, $N = 2000$, $\tau = 1/100$ and w_n is the GARCH(1,1) equation (17) with parameters $a = 0.1$, $b = 0.85$ and $\omega = 1/(1-a-b)$. Note that the instantaneous frequency is $\phi'(t) = 1 + t/20$ which increases linearly with t (i.e. in this case, the number of oscillations per unit of time is twice higher at the end).

Although one may suspect the existence of a trend and of some periodic component in the right column of Figure 1, both the amplitude modulations and the increasing frequency are rather difficult (if not impossible) to see. Furthermore, although the power spectrum of the contaminated signal that we display in the right column of Figure 1 suggests events localized both at very low frequencies

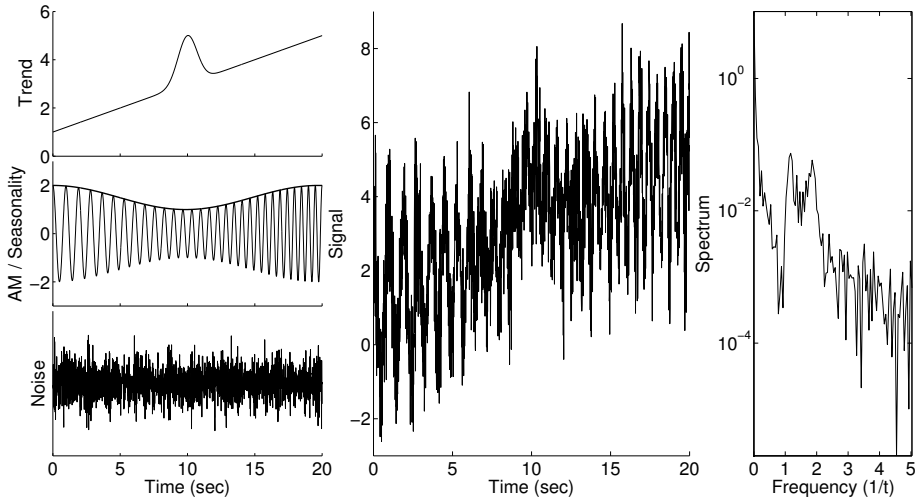

Figure 1: **Toy data.** Left column, from top to bottom: the trend, the seasonality (thin line) with the amplitude modulation (thick line) and the noise; Middle column: the contaminated signal, that is the sum of the three components. Right column: the power spectrum of the contaminated signal.

and in the $[1, 2]$ interval, the information displayed is insufficient to understand the dynamics of the system.

In Figure 2, we show the CWT[4] of $\boldsymbol{x}$ based on two different mother wavelets: the left panel is based on the Morlet wavelet and the right panel is based on the Meyer wavelet. Unlike with the Fourier transform, it is now possible to observe the time-varying frequency properties of the signal: the large increasing band on both CWT corresponds to the periodic component. Nonetheless, both pictures are blurred because the resolution in the time-frequency plane is limited by Heisenberg's uncertainty principle. Furthermore, Figure 2 illustrates the fact that different mother wavelets lead to different results. As we observe by comparing the left and right panels, the choice of template "colors" the representation and therefore influences the interpretation from which properties of the signal are deduced.

3.3 The Synchrosqueezing transform

In Kodera et al. (1978); Auger and Flandrin (1995); Chassande-Mottin et al. (2003, 1997); Flandrin (1999), the reassignment technique was proposed to improve the time-frequency resolution and alleviate the "coloration" effect of the time-frequency analysis. The synchrosqueezing transform (SST) is a special case of the reassignment technique originally introduced in the context of audio

[4]For more details about the matlab implementation of the CWT, available upon request from the authors, we refer to Appendix B.

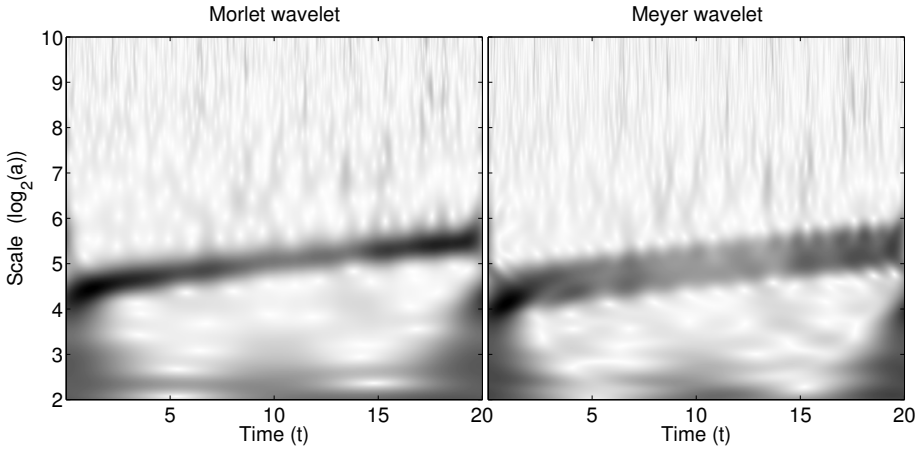

Figure 2: **Continuous wavelet transform of the toy data.** On the left: Morlet wavelet. On the right: Meyer wavelet.

signal analysis in Daubechies and Maes (1996). Its theoretical properties are analyzed in Daubechies et al. (2011); Thakur et al. (2013); Chen et al. (2013), where it is shown that the SST allows to determine $T(t)$, $a_k(t)$ and $\phi_k'(t)$ in the adaptive harmonic model (10) uniquely and up to some pre-assigned accuracy. In brief, by *reallocating* the CWT coefficients, the resolution of the time-frequency plane is increased so that we are able to accurately extract the trend, the instantaneous frequencies and the amplitude modulations.

The SST for x satisfying the adaptive harmonic model (10) plus some noise can be summarized as follows. Suppose the mother wavelet in equation (25) is such that $\text{supp} F_\psi \subset [1-\Delta, 1+\Delta]$, where $\Delta \ll 1$ and evaluate the CWT of x. According to Daubechies et al. (2011), we extract information about the IF $\phi'(t)$ by

$$\omega_x(t,a) = \begin{cases} -i\frac{\partial_t W_x(t,a)}{2\pi W_x(t,a)} & W_x(t,a) \neq 0 \\ -\infty & \text{otherwise} \end{cases} . \tag{27}$$

The motivation for (27) can be easily seen from the following example. For the purely harmonic signal $x(t) = \cos(2\pi\omega t)$, where $\omega > 0$, by (26) and the derivative of (26) with respect to time, one obtains

$$\omega_x(t,a) = \begin{cases} \omega & a \in \left[\frac{1-\Delta}{\omega}, \frac{1+\Delta}{\omega}\right] \\ -\infty & \text{otherwise} \end{cases} ,$$

which is the desired information about its IF. In this case, the SST is defined by

$$S_x(t,\omega) = \int_{a\,;\,|W_x(t,a)|>\Gamma} W_x(t,a)\, a^{-3/2}\, \delta(\,\omega_x(t,a) - \omega\,)\mathrm{d}a, \tag{28}$$

where $\Gamma > 0$ is the threshold chosen by the user, $\omega > 0$ is the frequency and $\omega_x(t,a)$ is defined in (27). Keeping in mind that the frequency is reciprocally

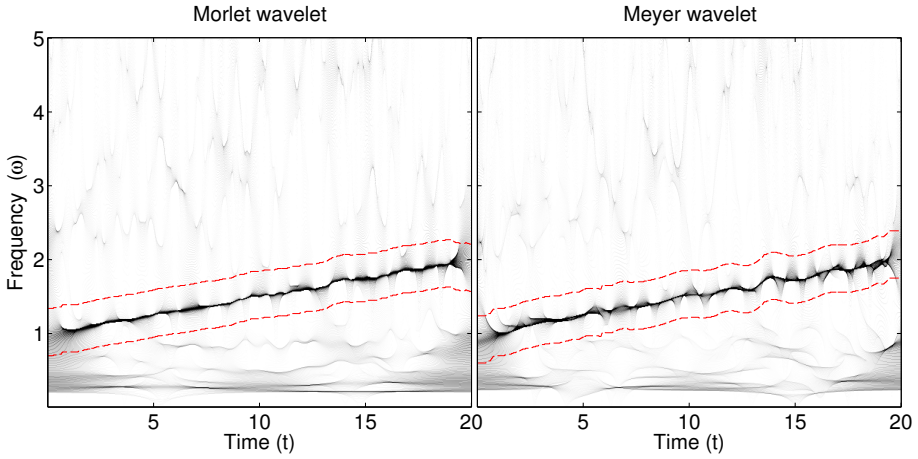

Figure 3: **Synchrosqueezed transform of the toy data.** On the left: Morlet wavelet. On the right: Meyer wavelet.

related to the scale, then (28) can be understood as follows: based on the IF information ω_x, at each time t, we collect the CWT coefficients at scale a where a seasonal component with frequency close to ω is detected. This additional step allows to mitigate the coloration due to the choice of template.

In this case, we discretize the SST. First, ω_x is discretized by

$$\omega_{\boldsymbol{x}}(n\tau, a) := \begin{cases} \dfrac{-i\partial_t W_{\boldsymbol{x}}(n\tau, a)}{2\pi W_{\boldsymbol{x}}(n\tau, a)} & \text{when} \quad |W_{\boldsymbol{x}}(n\tau, a)| \neq 0; \\ -\infty & \text{when} \quad |W_{\boldsymbol{x}}(n\tau, a)| = 0, \end{cases}$$

where $\partial_t W_{\boldsymbol{x}}(n\tau, a)$ is the discretization of $\partial_t W_x(t, a)$, which is defined as

$$\partial_t W_{\boldsymbol{x}}(n\tau, a) := \tau \sum_{m \in \mathbb{Z}} \frac{x_m}{a^{3/2}} \bar{\psi}'\big(\frac{m\tau - n\tau}{a}\big).$$

Similarly, $S_x(t, \omega)$ is discretized as

$$S_{\boldsymbol{x}}(n\tau, \omega) := \int_{a:\ |W_{\boldsymbol{x}}(n\tau, a)| \geq \Gamma} \delta(|\omega_{\boldsymbol{x}}(n\tau, a) - \omega|) W_{\boldsymbol{x}}(n\tau, a) a^{-3/2} \mathrm{d}a.$$

We now illustrate the analysis result of the toy data $\boldsymbol{x}$ by SST in Figure 3[5]. Whether we use the Morlet wavelet or the Meyer wavelet, it is visually obvious that the Synchrosqueezed transform of the noisy signal, with a darker area that seems to form a straight line from 1 to 2 in both cases, is close to the instantaneous frequency $\phi'(t) = 1 + t/20$. Compared with the CWT, it is

[5]For more details about the matlab implementation of the SST, available upon request from the authors, we refer again to Appendix B.

visually obvious that $S_{\boldsymbol{x}}$ is concentrated in narrower bands around the curves in the TF plane defined by $(t, \phi_k'(t))$ (here with $k = 1$).

Using any curve extraction technique, it is possible to estimate IF with high accuracy. We denote the estimated IF of the k-th component as $\widehat{\phi'}_k$. Furthermore, the restriction of $S_{\boldsymbol{x}}$ to the k-th narrow band suffices to reconstruct the k-th component of x. More formally, the k-th (complex) component at time $n\tau$ is estimated by

$$\widehat{f}_k^{\mathbb{C}}(n\tau) := \frac{1}{\mathcal{R}^{\psi}} \int_{\left|\phi_k'(n\tau)-\omega\right| \leq \Delta} S_x(n\tau, \omega) \mathrm{d}\omega, \tag{29}$$

where $\mathcal{R}^{\psi} := \int \frac{F_{\psi}(\omega)}{\omega} \mathrm{d}\omega$. The estimator of the k-th component f_k at time $n\tau$ is $\widehat{f}_k(n\tau) := \Re \widehat{f}_k^{\mathbb{C}}(n\tau)$, where $\Re z$ denotes the real part of $z \in \mathbb{C}$. We then estimate the amplitude modulations a_k at time $n\tau$ by

$$\widehat{a}_k(n\tau) := \left|\widehat{f}_k^{\mathbb{C}}(n\tau)\right|. \tag{30}$$

Finally, the trend at time $n\tau$ can be estimated by setting a low-frequency threshold ω_l and writing

$$\widehat{T}(n\tau) := x(t) - \frac{1}{\mathcal{R}^{\psi}} \int_{\omega_l}^{\infty} S_x(n\tau, \omega) \mathrm{d}\omega. \tag{31}$$

The theoretical properties of the above estimators have been shown in (Chen et al., 2013, Theorem 3.2), and will be summarized later.

In Figure 4, we show the reconstructed component for the toy data using a Morlet mother wavelet, where integration bands in (29) are the dashed red lines from Figure 3. Because the reconstruction is not dependent on the chosen mother wavelets[6], it is not shown here for the Meyer wavelet. Usually, while the instantaneous frequency is recovered precisely, the estimated amplitude modulations are less ideal. Although this is important when the goal is forecasting, it can be dampened by further smoothing the amplitude (here with a simple moving average filter of length $1/\tau$).

Before closing the section, we summarize the theoretical properties of the SST when it is applied to analyze the adaptive harmonic model (16) for the sampled log-volatility process $\boldsymbol{y} = \{y_l\}_{l \in \mathbb{Z}}$. Suppose y satisfies the following additional condition – $A_k(t) \in C^2(\mathbb{R})$ and $\sup_{t \in \mathbb{R}} |A_k''(t)| \leq \epsilon c_2$ for all $k = 1, \ldots, K$, and $T \in C^2(\mathbb{R})$ so that $|T''(\psi_{a,b})| \leq C_T \epsilon$ for all $b \in \mathbb{R}$ and $a \in (0, \frac{1+\Delta}{c_1}]$. If the sampling interval τ satisfies $0 < \tau \leq \frac{1-\Delta}{(1+\Delta)c_2}$ and $\mathrm{var}\left[\log w_n^2\right] < \infty$, then it is highly probably that the estimators $\widehat{f}_k$, $\widehat{\phi'}_k$ and $\widehat{a}_k$ evaluated from $\boldsymbol{y}$ are accurate up to $C_\tau(\sigma + \epsilon)$, where C_τ is a constant depending on C' and τ which decreases when the sampling rate increases, that is, τ decreases. Moreover, $\widehat{T}$

[6]This statement is correct up to the moments of the chosen mother wavelet ψ. The reconstruction error depends on the first three moments of the chosen mother wavelet and its derivative but not on the profile of ψ (Daubechies et al. (2011); Chen et al. (2013))

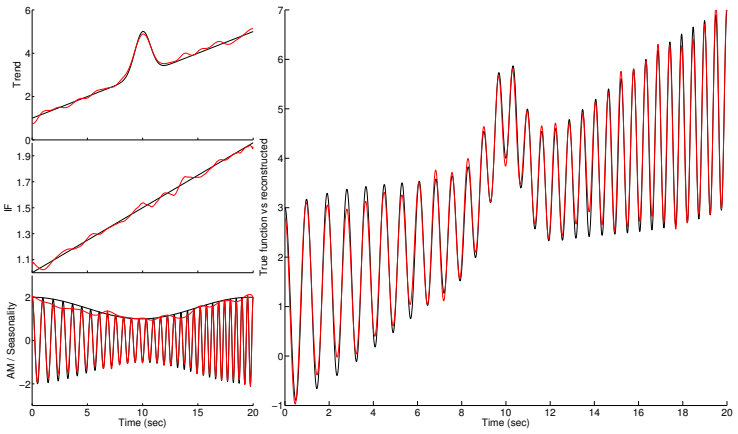

Figure 4: **Reconstruction of the components.** Left column, from top to bottom: the trend, the instantaneous frequency and the seasonality (thin line) with the amplitude modulation (thick line) with the true underlying function (black) and reconstructed result (red); Right column: the true signal, that is the sum between the trend and the seasonality, (black) and the reconstructed function (red).

is accurate up to $C_{\tau,T}(\sigma + \epsilon)$, where $C_{\tau,T}$ is a constant depending on C' τ and C_T which decreases when τ decreases. We refer the reader to Chen et al. (2013) for the precise statement of the theorem and more discussions.

In summary, for the harmonic model (23), the Fourier transform allows us to extract the features of interest about the oscillation. When it comes to the adaptive harmonic model (10), we count on the synchrosqueezing transform to extract the oscillatory features of a given signal.

4 Simulation study

In this section, we present a simulation study that uses a realistic setting relevant to the empirical application in Section 5. To study the properties of the estimators proposed in Section 3, we directly sample 1000 times 170 days of data (i.e. 48960 observations) from the discretized log-volatility from equation (16), that is

$$y_n := T_n + s_n + \log\left(w_n^2\right),$$

where T_n and s_n are deterministic functions and w_n is noise. Using the EUR/USD sample (cf. Section 5, for more details), we estimate trend and seasonality used in the simulation study:

- T_n is the trend obtained using the SST estimate from equation (31),
- s_n is the seasonality with 4 periodic components obtained either using an FFF or the SST estimator described in Section 5,

- w_n is either a student-t or a GARCH(1,1) with student-t innovations.

The noise processes is generated from processes estimated with the residuals obtained using the corresponding s_n: for example, we use the FFF version of s_n to remove seasonality from the EUR/USD sample, we fit a student-t for the residuals and use the estimated parameters to generate the noise.

In Figures 5 and 6, we show the results of the simulation study. We use f as a quantity to be estimated (either the trend or one of the periodic components) and $\widehat{f}$ as the SST estimator. The left panels of Figures 5 and 6 show the true f. The first line corresponds to the trend. In the lines below, "Am i" denotes a periodic component whose frequency equals i. In each sub-figure the plain and dashed lines correspond respectively to student and GARCH(1,1) (with student innovations) noises. In both cases of constant seasonality (Figure 5) and amplitude modulated seasonality (Figure 6), the resulting estimated biases (middle left panels) for either student or GARCH(1,1) noises are close to 0 (even though slightly positive for the trend but below 0.2), suggesting unbiased SST estimators of trend and amplitudes.

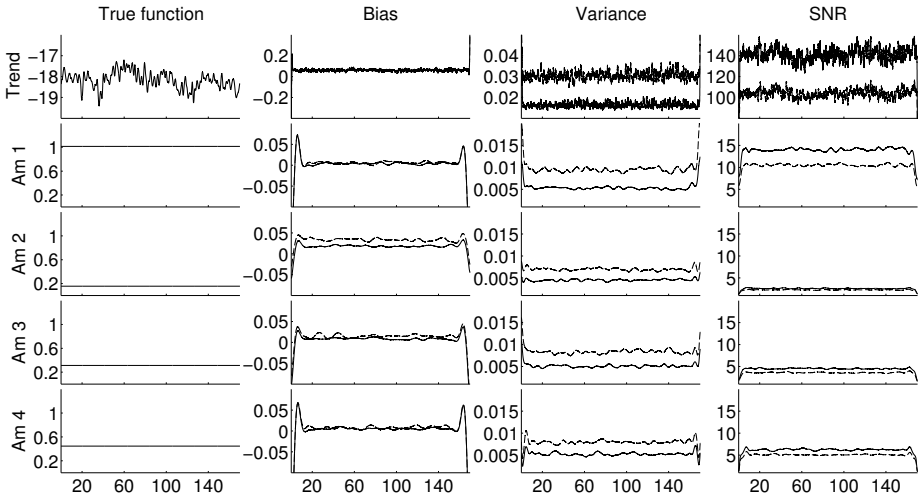

Figure 5: **Simulation using constant seasonality.** We use f as a quantity to be estimated (either the trend or one of the periodic components) and $\widehat{f}$ as the SST estimator. Left: the true f; Middle left: the estimated bias of the estimator, i.e. $\widehat{\mathbb{E}}\left[\widehat{f}\right] - f$; Middle right: the estimated variance of the estimator, i.e. $\widehat{\text{var}}\left[\widehat{f}\right]$; Right: the estimated signal-to-noise ratio, i.e. $\left|\widehat{\mathbb{E}}\left[\widehat{f}\right]\right| / \sqrt{\widehat{\text{var}}\left[\widehat{f}\right]}$. The first line corresponds to the trend. In the lines below, "Am i" denotes a periodic component whose frequency equals i and in each sub-figure the plain and dashed lines correspond respectively to student and GARCH(1,1) noises.

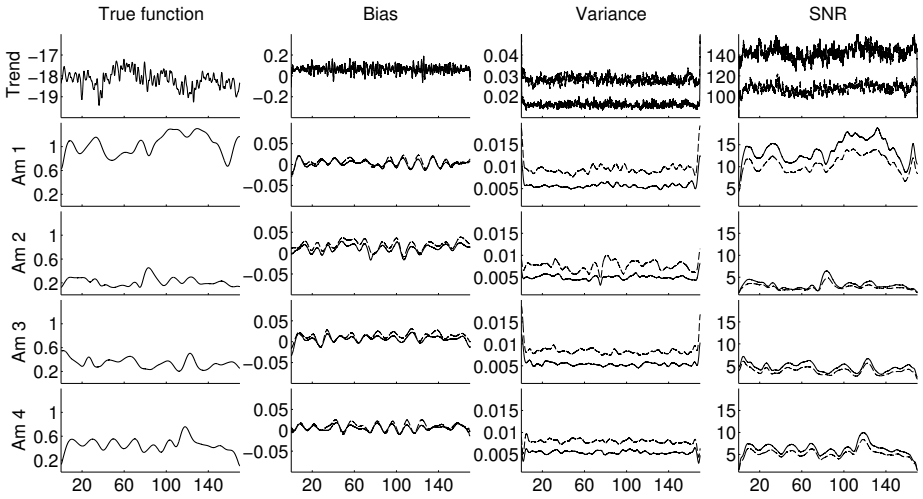

Figure 6: **Simulation using amplitude modulated seasonality.** We use f as a quantity to be estimated (either the trend or one of the periodic components) and $\widehat{f}$ as the SST estimator. Left: the true f; Middle left: the estimated bias of the estimator, i.e. $\widehat{\mathbb{E}}\left[\widehat{f}\right] - f$; Middle right: the estimated variance of the estimator, i.e. $\widehat{\text{var}}\left[\widehat{f}\right]$; Right: the estimated signal-to-noise ratio, i.e. $\left|\widehat{\mathbb{E}}\left[\widehat{f}\right]\right| / \sqrt{\widehat{\text{var}}\left[\widehat{f}\right]}$. The first line corresponds to the trend. In the lines below, "Am i" denotes a periodic component whose frequency equals i and in each sub-figure the plain and dashed lines correspond respectively to student and GARCH(1,1) noises.

Overall, the estimated variances (middle right panels) for the student noises are lower than those for the GARCH(1,1) cases. This can be explained by the fact that the heteroskedasticity represents an additional disturbance for the SST estimator. For student noises, the estimated variances are almost constant (around 0.005) for all the different amplitudes considered, both in the case of constant seasonality (Figure 5) or amplitude modulated seasonality (Figure 6).

The right panels show the estimated signal-to-noise ratio ($\widehat{\text{SNR}}$), i.e. $\left|\widehat{\mathbb{E}}\left[\widehat{f}\right]\right| / \sqrt{\widehat{\text{var}}\left[\widehat{f}\right]}$. Obviously, the higher the true amplitude the higher the SNR. Furthermore, the SNR for the trend is undoubtedly very high. Finally, because the GARCH(1,1) variances are slightly higher than the student variances but estimators are essentially unbiased, we consistently find a higher SNR for the student.

Note that, as seen in Figure 5 using the constant FFF, the bias and the variance strongly increase near the boundaries of the data. To alleviate this effect, simulations from Figure 6 are performed by "padding" the equivalent of 50 days of data on both sides. In other words, we take the 50 times 288 first/last

simulated observations, flip them and add them back at the beginning/end of the data before running the SST estimation.

5 Application to foreign exchange data

As an empirical application, we study the foreign exchange (FX) market. Because it is open continuously from Sunday 10:00 pm GMT to Friday 10:00 pm GMT, our instantaneous trend and seasonality model constitutes a convenient framework for FX market study. Although they operate over-the-counter (OTC), market-makers post bid and ask quotes $\{p_{\text{BID}}(t), p_{\text{ASK}}(t)\}$ in real time to data providers for all major currency pairs.

Because activity is not homogeneous during trading days/weeks, quotes are usually asynchronous and preprocessing is required to obtain regularly spaced observations. From $\{p_{\text{BID}}(t), p_{\text{ASK}}(t)\}_{n=1}^{N}$, we define the (log)-price

$$p(t) = \frac{\log\left(p_{\text{BID}}(t)\right) + \log\left(p_{\text{ASK}}(t)\right)}{2}.$$

Furthermore, several quotes may share the same time stamp and tick-by-tick data users often need to take this into account; for example by taking an average.

In what follows, we avoid the preprocessing work by using data graciously made available by Dukascopy Bank SA[7], an electronic broker holding a Securities Dealer License issued by the FINMA. It contains 5 minute spaced bid-ask pairs for the EUR/USD and the USD/JPY from March 11, 2012 to November 2, 2012. This specific time-period was chosen to avoid dealing with daily saving time in the US, as the activity pattern in FX markets has been shown to differ in winter and in summer (see e.g. Andersen and Bollerslev (1998)). Hence in a total of 34 trading weeks, we obtain 48960 observations (170 days) excluding weekends.

In the top row of Figure 7, we show the return r_n for the two considered exchange rates for the first week of trading, that is from March 11, 2012 (Sunday) at 21:05 pm GMT to March 16, 2012 (Friday) at 20:55 pm GMT (1439 observations). As usual with financial data, the heteroskedasticity is obvious. However, it is arguably easier to discern the periodicity by looking at the log-volatility $y_n := 2\log|r_n - \widehat{\mu}|$ in the bottom row of Figure 7; the smallest observations correspond to the case where the price does not move during the 5 minute interval (i.e. $r_n = 0$ and $y_n = 2\log|\widehat{\mu}|$).

In Figure 8, we show the autocorrelation and the power spectrum of the log-volatility y_n. While the autocorrelation and power spectrum of the return process are usually flat for all lags and frequencies (white noise property), the two curves indicate strong periodicities in the log-volatility. Although the autocorrelation and power spectrum are different for the two exchange rates, we observe that they yield essentially two versions of the same information: on average, the seasonality is composed of periodic components of integer frequency.

[7] http://www.dukascopy.com/

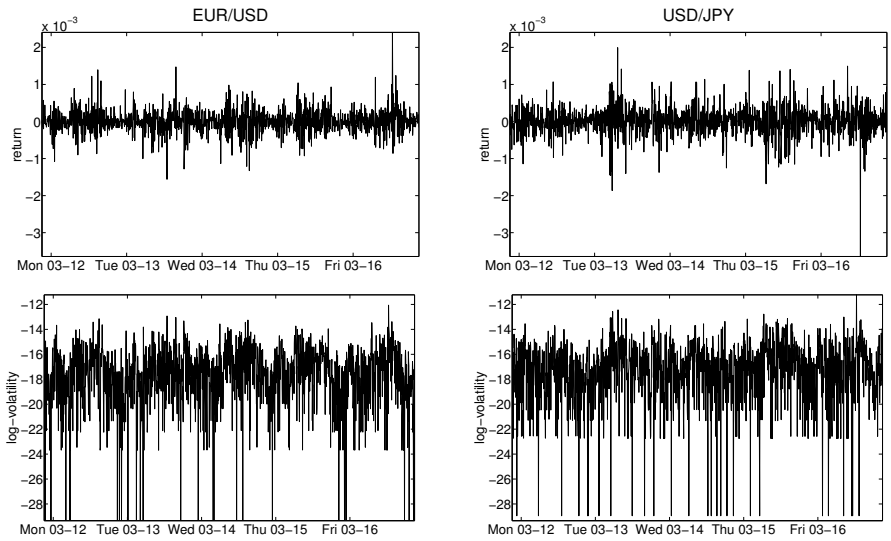

Figure 7: **Return and log-volatility.** Top: the EUR/USD (left) and USD/JPY (right) return r_n (left) sampled from March 11, 2012 (Sunday) at 21:05 pm GMT to March 16, 2012 (Friday) at 20:55 pm GMT (1439 observations); Bottom: the EUR/USD (left) and USD/JPY (right) log-volatility $y_n := 2 \log |r_n - \widehat{\mu}|$ for the same period;

Note that we used "on average" to emphasize the fact that both representations are only part of the picture. This representation misses the "dynamics" of the signal: for example, some of the underlying periodic components may strengthen, weaken or even completely disappear at some point, as illustrated in Section 3.

In Figure 9, we display $|S_y(t, \omega)|$, the absolute value of the Synchrosqueezed transform, in shades of gray for frequencies $0.5 \leq \omega \leq 4.5$. For both exchange rates, we observe much darker curves (indicating higher values of $|S_x(t, \omega)|$) at integer frequencies (although the curve at 2 daily oscillations for the EUR/USD is very thin and sometimes fading). The apparent linearity of the instantaneous frequency at integer values lends support for an amplitude-modulated version of the FFF or, in other words, a member of $\mathcal{A}_{\epsilon,d}^{c_1,c_2}$ such that $d \leq 1/(2K-1)$ and $|\phi_k'(t)| = k$ for all $k \in \{1, \ldots, K\}$ and $t \in \mathbb{R}$. Although this kind of constant frequency/amplitude modulated model can be estimated in different ways, we carry the analysis within the SST framework. For the reconstruction using the estimator from equation (29), we display the dashed red lines, which are the selected integration bands at $k \pm \Delta$ with $k \in \{1, 2, 3, 4\}$ and $\Delta = 0.05$.

In Figure 10, we show the reconstruction results and compare them to more traditional estimators. In the first row, we show the reconstructed trend (plain dark line) with bootstrapped 95% confidence intervals (dotted lines) and a rolling version of the realized log-volatility (in gray). In the second to fifth

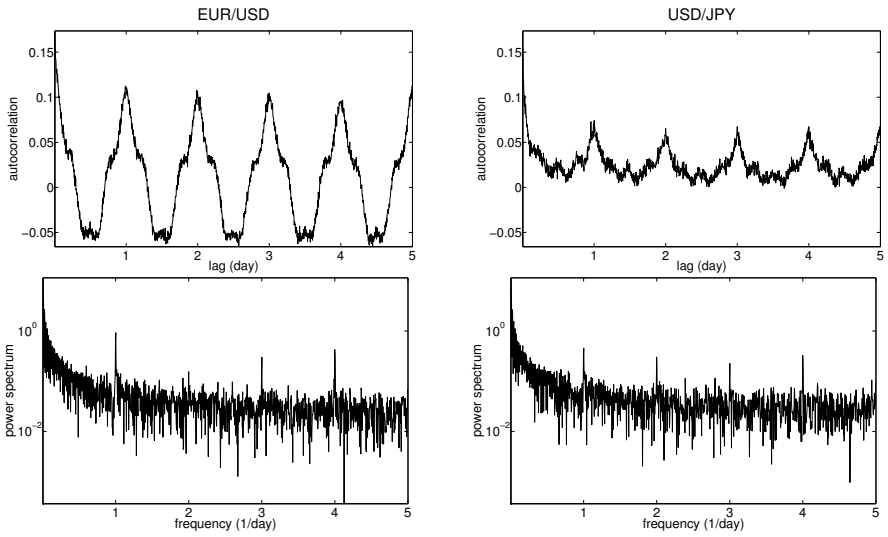

Figure 8: **Autocorrelation and power spectrum.** Top: autocorrelation function $\gamma_y(l) = \widehat{\mathbb{E}}((y_n - \widehat{\mathbb{E}}(y_n))(y_{n-l} - \widehat{\mathbb{E}}(y_n)))$ with lags l up to 1440 (the x axis ticks are divided by 288) for the EUR/USD (left) and USD/JPY (right); Bottom: power spectrum $P_y(\omega)$, with frequency ω up to 5 for the EUR/USD (left) and USD/JPY (right).

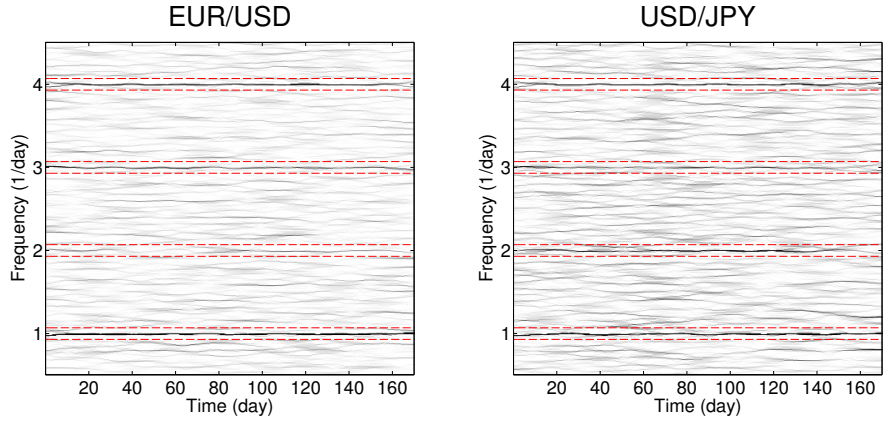

Figure 9: **Synchrosqueezed transform of the EUR/USD and USD/JPY.** The Synchrosqueezed transform of the log-volatility's absolute value, that is $|S_y(t,\omega)|$ with $y_n := 2\log|r_n - \mu|$, is displayed in shades of gray for the EUR/USD (left) and and the USD/JPY (right). The dashed red lines are the integration bands selected for the reconstruction.

rows, we show the reconstructed amplitude modulations (plain dark line) with 95% confidence intervals (dotted lines) and a rolling version of the FFF (in gray). We use "Amplitude k" to denote the amplitude of a component whose frequency equals k (i.e. the main daily oscillation corresponds to "Amplitude 1").

In what follows, we give more details on how those results are obtained and comment on them:

- We reconstruct the trend with a frequency threshold chosen at 0.95 daily oscillations, that is using a discretized version of $\widehat{T}(n\tau) := y_n - \mathfrak{Re}\frac{1}{\mathcal{R}^\psi}\int_{0.95}^{144} S_y(n\tau,\omega)\mathrm{d}\omega$ where the upper integration bound corresponds to the highest frequency that can be resolved according to the Nyquist theorem (i.e. half of the sampling frequency which is equal to 288 for 5 minute spaced observations). As for the realized log-volatility, it is simply computed as a rolling moving average of y_n with 288 observations included (except near the borders) and maximal overlap. Our trend estimator quite visibly constitutes a smoothed version of the usual volatility measure, which we find well inside the 95% confidence interval except for rare and short lived upward/downward spikes.

- We reconstruct the periodic components using the discretized version of $\widehat{f}_k(n\tau) := \frac{1}{\mathcal{R}^\psi}\int_{|k-\omega|\leq\Delta} S_y(n\tau,\omega)\mathrm{d}\omega$ described in B. It is noteworthy that when a unique FFF is estimated for the whole sample, the corresponding amplitude is very close to the mean amplitude from the SST. In fact, we extend this analogy between the FFF and the SST by dividing the sample into (possibly overlapping) smaller intervals (e.g. monthly or weekly), estimate an FFF for each of them and show that the corresponding amplitudes get closer and closer to the SST, which is essentially an instantaneous version of the FFF.

  Here, the rolling FFF parameters are computed using the OLS estimator using a two-week rolling window (say $w = 288 \times 5 \times 2$). The corresponding amplitude, which closely "tracks" the SST amplitudes, is found using the relation $\sqrt{a_k^2(n\tau) + b_k^2(n\tau)}$ where the observations y_j for $j \in \{n - \lfloor w/2 \rfloor, \ldots, n + \lfloor w/2 \rfloor\}$ are used in the estimation.

- In Chen et al. (2013), the authors give theoretical bounds on the error committed when reconstructing the trend and seasonality with the SST. However, those bounds have two disadvantages: firstly they are very rough and secondly they are fairly impractical to implement. To obtain the confidence intervals displayed in the Figure, we use the following procedure:

  (a) Use $\widehat{T}(n\tau)$ and $\widehat{f}_k(n\tau)$ to recover the residuals

  $$z_n := r_n \exp\left[-\frac{\widehat{T}(n\tau) + \sum_{k=1}^{K}\widehat{f}_k(n\tau)}{2}\right]$$

  for $n \in \{1, \cdots, N\}$.

(b) Bootstrap B samples, say z_n^b for $b \in \{1, \cdots, B\}$ and $n \in \{1, \cdots, N\}$.

(c) Add back $\widehat{T}(n\tau)$ and $\widehat{f}_k(n\tau)$ to z_n^b in order to obtain

$$r_n^b := z_n^b \exp\left[\frac{\widehat{T}(n\tau) + \sum_{k=1}^K \widehat{f}_k(n\tau)}{2}\right]$$

for $b \in \{1, \cdots, B\}$ and $n \in \{1, \cdots, N\}$.

(d) Define $y_n^b := 2\log|r_n^b - \mu|$, estimate $\widehat{T}^b(n\tau)$ and $\widehat{f}_k^b(n\tau)$ for $b \in \{1, \cdots, B\}$ and $n \in \{1, \cdots, N\}$ to compute pointwise confidence intervals.

As the residuals z_n are found to be heteroskedastic, a naive bootstrapping scheme is not appropriate. To deal with this dependency issue, we use automatic block-length selection procedure from Politis and White (2004) along with circular block-bootstrap.

Figure 10 suggests that, for all components and the two exchange rates, both the FFF parameters and SST amplitude (and hence the seasonality) evolve dynamically over time: in this example using only an eight months sample, the amplitude of the dominant component (i.e. the main daily oscillation) for the EUR/USD is roughly comprised between 0.6 and 1.4 whereas the constant FFF estimate is 1.01. In more concrete terms, using 1.01 instead of 0.6, respectively 1.4, corresponds to a 30.7% overestimation, 35.7% underestimation, of the periodic component of volatility. In a risk-management context where the volatility is the determinant time-varying feature, the implications are serious.

Figure 11 represents two diagnostic checks: the residuals autocorrelation and the estimated seasonality for a whole trading week. On the left of Figure 11, we show the log-volatility autocorrelation (dark line), the autocorrelation of $y_n - \sum_{k=1}^K \widehat{f}_k(n\tau)$ (dark gray line), the autocorrelation of $y_n - \sum_{k=1}^K \widehat{f}_k(n\tau) - \widehat{T}(n\tau)$ (light gray line) and approximate 95% confidence intervals at $\pm 2/\sqrt{N}$ (dashed line) for the EUR/USD (top) and USD/JPY (bottom). In the autocorrelation of $y_n - \sum_{k=1}^K \widehat{f}_k(n\tau)$ (dark gray line), we observe that the periodicity is removed but that the process still seems to exhibit long-memory properties (arguably easier to spot on the USD/JPY data). As it becomes negligible beyond the first 30-50 lags (or equivalently between 2.5 and 4.2 hours), the autocorrelation of $y_n - \sum_{k=1}^K \widehat{f}_k(n\tau) - \widehat{T}(n\tau)$ (light gray line), i.e. the residual intraday heteroskedasticity, suggests that the persistence is captured in the trend component.

On the right of Figure 11, we show the seasonality estimated by a constant FFF (plain line), the rolling FFF (dashed line) and the SST (dotted line) for the EUR/USD (top) and USD/JPY (bottom). Although they are very similar, the Figure helps to further explain our earlier statement that the SST allows to extract an instantaneous version of the FFF: the plain line represents the average daily oscillation, the dashed line is a wiggly estimate of the dynamics and the dotted line is our smooth instantaneous seasonality estimate.

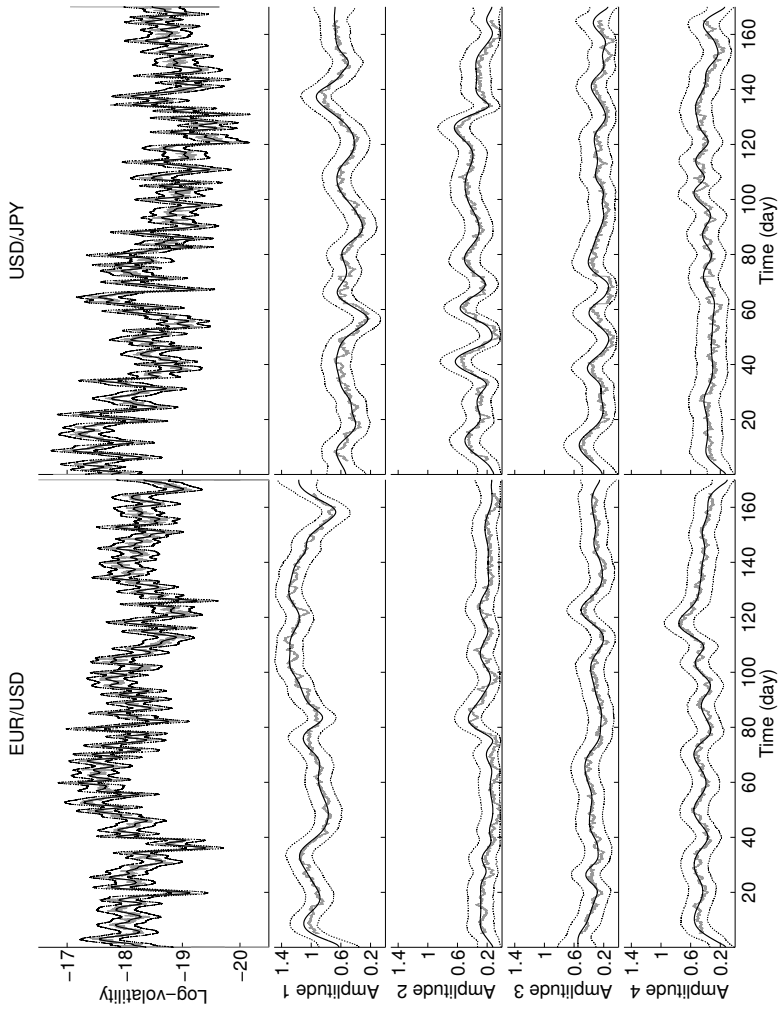

Figure 10: **Reconstruction results.** Top row: reconstructed trend (plain dark line) with 95% confidence intervals (dotted lines) and rolling realized log-volatility (in gray) for the EUR/USD (left) and USD/JPY (right). Bottom four rows: reconstructed amplitude modulations (plain dark line) with 95% confidence intervals (dotted lines) and rolling FFF (in gray) for the EUR/USD (left) and USD/JPY (right), with "Amplitude k" corresponding to the amplitude of a component of frequency k.

In summary, those results show that the SST is adaptive to the data. Although similar estimates of trend and seasonality can be obtained with moving averages and rolling regressions respectively, the SST provides smooth estimates and does not require an arbitrary choice of the length of the window or the rolling overlap. One could argue that the choice of mother wavelet is also an arbitrary choice, but as we showed in Section 2, the influence of this choice on the SST estimators is negligible. In any case, the main message is that dynamic methods

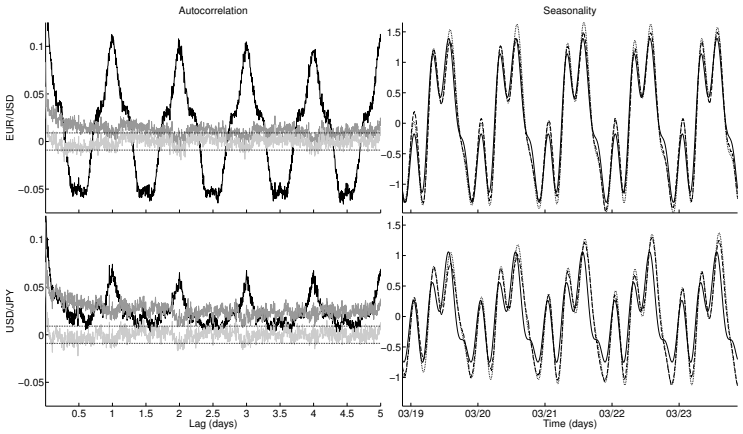

Figure 11: **Diagnostic checks.** Left: the log-volatility autocorrelation (dark line), the autocorrelation of $y_n - \sum_{k=1}^{K} \widehat{f}_k(n\tau)$ (dark gray line), the autocorrelation of $y_n - \sum_{k=1}^{K} \widehat{f}_k(n\tau) - \widehat{T}(n\tau)$ (light gray line) and approximate 95% confidence intervals at $\pm 2/\sqrt{N}$ (dashed line) for the EUR/USD (top) and USD/JPY (bottom). Right: the constant FFF (plain line), the rolling FFF (dashed line) and the SST (dotted line) for the EUR/USD (top) and USD/JPY (bottom).

are necessary to properly take the influence of seasonality into account, as static models can lead to severe underestimation/overestimation of the intraday spot volatility.

6 Discussion

The non-parametric estimation of intraday spot volatility by disentangling instantaneous trend and seasonality is an important extension of classical intraday models for situations where adaptivity to ever changing markets may be essential. The proposed method provides a realistic framework for the high frequency time series behavior. First, it lets the realized volatility be modelled as an "instantaneous" trend which evolves in real time within the day. Second, it allows the seasonality to be non-constant over the sample. In a simulation study using a realistic setting, numerical results confirm that the proposed estimators for the trend and seasonality components are unbiased and have a low variance (or equivalently a high signal-to-noise ratio). We show that this result holds even in the presence of heteroskedastic and heavy tailed noise.

Using 170 days of EUR/USD and USD/JPY exchange rates sampled every 5 minutes, we confirm empirically that the oscillation frequency is constant, as suggested originally in the FFF model from Andersen and Bollerslev (1997, 1998). In the two exchange rates, we show that the amplitude modulations of the periodic components evolve dynamically over time. In the EUR/USD,

neglecting those modulations in the periodic part of the volatility would imply either a 30.7% overestimation (when the periodic components are lower than their mean value) or a 35.7% underestimation (when the periodic components are higher than their mean value). Furthermore, we show that the SST estimator gives results comparable to a smooth version of a rolling OLS using the FFF predictors. However, the adaptivity of SST should still be emphasized, as it does not require ad-hoc choices neither of the rolling overlapping ratio nor of the length of each window as for a parametric regression.

While we illustrated our model by simultaneously disentangling the low and high frequency components of the intraday spot volatility in the FX market, it is possible to embed the new methodology into a forecasting exercise. Although this is beyond the scope of the present paper, which is rather aimed at presenting the modelling framework, the exercise may be useful in the context of an investor's optimal portfolio choice or for risk-management purposes. Further research directions include the extension of the framework to both the non-homogeneously sampled ("tick-by-tick") as well as multivariate data. We will return to these questions and related issues in future works.

Acknowledgements

Most of this research was conducted while the corresponding author was visiting the Berkeley Statistics Department. He is grateful for its support and helpful discussions with the members both from Bin Yu's research group and the Coleman Fung Risk Management Research Center. This research is also supported in part by the Center for Science of Information (CSoI), a US NSF Science and Technology Center, under grant agreement CCF-0939370, and by NSF grants DMS-1160319 and CDS&E-MSS 1228246.

References

Andersen, T. G. and T. Bollerslev (1997): "Intraday periodicity and volatility persistence in financial markets," Journal of Empirical Finance, 4, 115–158.

——— (1998): "Deutsche Mark-Dollar Volatility: Intraday Activity Patterns, Macroeconomic Announcements, and Longer Run Dependencies," The Journal of Finance, 53, 219–265.

Andersen, T. G., T. Bollerslev, F. Diebold, and P. Labys (2003): "Modeling and Forecasting Realized Volatility," Econometrica, 71, 529–626.

Andersen, T. G., T. Bollerslev, and F. X. Diebold (2007): "Roughing It Up: Including Jump Components in the Measurement, Modeling, and Forecasting of Return Volatility," The Review of Economics and Statistics, 89, 701–720.

Auger, F. and P. Flandrin (1995): "Improving the readability of time-frequency and time-scale representations by the reassignment method," IEEE Trans. Signal Process., 43, 1068–1089.

Barndorff-Nielsen, O. and N. Shephard (2002): "Econometric analysis of realized volatility and its use in estimating stochastic volatility models," Journal of Royal Statistical Society: Series B, 253–280.

——— (2007): Variation, jumps, market frictions and high frequency data in financial econometrics, Cambridge University Press, Econometric Society Monographs.

Bollerslev, T. (1986): "Generalized Autoregressive Conditional Heteroskedasticity," Journal of Econometrics, 31, 307–327.

Bollerslev, T., R. Y. Chou, and K. F. Kroner (1992): "ARCH modeling in finance : A review of the theory and empirical evidence," Journal of Econometrics, 52, 5–59.

Bollerslev, T., R. F. Engle, and D. B. Nelson (1994): "Arch models," in Handbook of Econometrics, ed. by R. F. Engle and D. McFadden, Elsevier, vol. 4 of Handbook of Econometrics, chap. 49, 2959–3038.

Bollerslev, T. and E. Ghysels (1996): "Periodic Autoregressive Conditional Heteroscedasticity," Journal of Business & Economic Statistics, 14, 139–51.

Boudt, K., C. Croux, and S. Laurent (2011): "Robust estimation of intraweek periodicity in volatility and jump detection," Journal of Empirical Finance, 18, 353–367.

Chassande-Mottin, E., F. Auger, and P. Flandrin (2003): "Time-frequency/time-scale reassignment," in Wavelets and signal processing, Boston, MA: Birkhäuser Boston, Appl. Numer. Harmon. Anal., 233–267.

Chassande-Mottin, E., I. Daubechies, F. Auger, and P. Flandrin (1997): "Differential reassignment," Signal Processing Letters, IEEE, 4, 293–294.

Chen, Y.-C., M.-Y. Cheng, and H.-T. Wu (2013): "Nonparametric and adaptive modeling of dynamic periodicity and trend with heteroscedastic and dependent errors," Journal of the Royal Statistical Society: Series B, To appear.

Dacorogna, M. M., U. A. Mueller, R. J. Nagler, R. B. Olsen, and O. V. Pictet (1993): "A geographical model for the daily and weekly seasonal volatility in the foreign exchange market," Journal of International Money and Finance, 12, 413–438.

Daubechies, I. (1992): Ten Lectures on Wavelets, Philadelphia: Society for Industrial and Applied Mathematics.

Daubechies, I., J. Lu, and H.-T. Wu (2011): "Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool," Appl. Comp. Harmonic Anal., 30, 243–261.

Daubechies, I. and S. Maes (1996): "A nonlinear squeezing of the continuous wavelet transform based on auditory nerve models," in Wavelets in Medicine and Biology, CRC-Press, 527–546.

Deo, R., C. Hurvich, and Y. Lu (2006): "Forecasting realized volatility using a long-memory stochastic volatility model: estimation, prediction and seasonal adjustment," Journal of Econometrics, 131, 29–58.

Engle, R. and M. Sokalska (2012): "Forecasting intraday volatility in the US equity market. Multiplicative component GARCH," Journal of Financial Econometrics, 10, 54–83.

Engle, R. F. (1982): "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," Econometrica, 50, 987–1007.

Flandrin, P. (1999): Time-frequency/time-scale Analysis, Wavelet Analysis and Its Applications, Academic Press Inc.

Folland, G. (1999): Real analysis: modern techniques and their applications, Pure and applied mathematics, Wiley.

Gallant, A. R. (1981): "On the bias in flexible functional forms and an essentially unbiased form: The fourier flexible form," Journal of Econometrics, 15, 211–245.

Giot, P. (2005): "Market risk models for intraday data," European Journal of Finance, 11, 309–324.

Guillaume, D. M., O. Pictet, and M. M. Dacorogna (1994): "On the intra-daily performance of GARCH processes," Working papers, Olsen and Associates.

Hansen, L. P. (1982): "Large Sample Properties of Generalized Method of Moments Estimators," Econometrica, 50, 1029–54.

Kodera, K., R. Gendrin, and C. Villedary (1978): "Analysis of time-varying signals with small BT values," Acoustics, Speech and Signal Processing, IEEE Transactions on, 26, 64 – 76.

Laakkonen, H. (2007): "Exchange rate volatility, macro announcements and the choice of intraday seasonality filtering method," Research Discussion Papers 23/2007, Bank of Finland.

Lunde, A. and P. R. Hansen (2005): "A forecast comparison of volatility models: does anything beat a GARCH(1,1)?" Journal of Applied Econometrics, 20, 873–889.

Martens, M., Y.-C. Chang, and S. J. Taylor (2002): "A comparison of seasonal adjustment methods when forecasting intraday volatility," The Journal of Financial Research, 25, 283–299.

Müller, H.-G., R. Sen, and U. Stadtmüller (2011): "Functional data analysis for volatility," Journal of Econometrics, 165, 233–245.

Newey, W. K. and D. McFadden (1994): "Large sample estimation and hypothesis testing," in Handbook of Econometrics, ed. by R. F. Engle and D. McFadden, Elsevier, vol. 4, chap. 36, 2111–2245.

Politis, D. N. and H. White (2004): "Automatic Block-Length Selection for the Dependent Bootstrap," Econometric Reviews, 23, 53–70.

Thakur, G., E. Brevdo, N. S. Fuckar, and H.-T. Wu. (2013): "The Synchrosqueezing algorithm for time-varying spectral analysis: robustness properties and new paleoclimate applications," Signal Processing, 93, 1079–1094.

Wu, H.-T. (2013): "Instantaneous frequency and wave shape functions (I)," Appl. Comp. Harmonic Anal., 35, 181–199.

A Proof of convergence for the discretization scheme

Lemma A.1. Using the notations in (13) and (14), $\mu \in C^1(\mathbb{R})$ and σ satisfying the adaptive harmonic model (10), the error terms $R_{1,n,\tau} := \widetilde{\mu}_n - \mu_n$ and $R_{2,n,\tau} := \widetilde{\sigma \mathrm{d}W}_n - \sigma_n w_n$ are such that

$$|R_{1,n,\tau}| \leq \frac{\|\mu'\|_{L^\infty}}{2}\tau$$
$$|\mathbb{E}R_{2,n,\tau}| = 0, \quad |\mathrm{var}R_{2,n,\tau}| \leq \frac{\|\sigma'\|^2_{L^\infty}}{3}\tau.$$

Furthermore, since $N\tau$ is finite and σ and μ are assumed to be continuous over $\mathbb{R}$, we know that $\|\mu'\|_{L^\infty([0,N\tau])} < \infty$ and $\|\sigma'\|^2_{L^\infty([0,N\tau])} < \infty$. Hence we obtain that $r_n \approx \mu_n\tau + \sigma_n w_n$ holds uniformly in probability.

Proof. Since $\mu(t) \in C^1(\mathbb{R})$, By Taylor's expansion, we directly have

$$R_{1,n,\tau} = \left|\frac{1}{\tau}\int_{t-\tau}^{t} \mu(u)\mathrm{d}u - \mu(t)\right| \leq \frac{1}{\tau}\int_{t-\tau}^{t} |\mu(u) - \mu(t)|\mathrm{d}u \leq \frac{\|\mu'\|_{L^\infty}}{2}\tau.$$

Similarly, we have

$$\begin{aligned} R_{2,n,\tau} &= \left|\frac{1}{\tau}\int_{t-\tau}^{t} \sigma(u)\mathrm{d}W - \sigma(t)(W(t) - W(t-\tau))\right| \\ &\leq \frac{1}{\tau}\int_{t-\tau}^{t} |\sigma(u) - \sigma(t)|\mathrm{d}W. \end{aligned}$$

Since σ is smooth, the integration and expectation are interchangeable, and hence we have

$$\mathbb{E}R_{2,n,\tau} = 0$$

and

$$\mathrm{var}(R_{2,n,\tau}) = \frac{1}{\tau^2}\int_{t-\tau}^{t} |\sigma(u) - \sigma(t)|^2\mathrm{d}u \leq \frac{\|\sigma'\|^2_{L^\infty}}{3}\tau,$$

where the first equality holds due to the definition of the stochastic integration. We thus conclude the proof. □

B Implementations details

In this section we provide the numerical SST implementation detail. The Matlab code is available from the authors upon request and we refer the readers to Thakur et al. (2013) for more implementation details.

Let us recall that each of the D days in the data set is divided into P equal periods for a total of $N = DP$ observations. Equivalently, a unique index $n = P(d-1) + p \in \{1, \cdots, N\}$ identifies observation $p \in \{1, \cdots, P\}$ of day

$d \in \{1, \cdots, D\}$. This index corresponds to time $t = n\,\tau$ (in days) where $\tau = 1/P$ is the sampling interval. We fix a discretely sampled time series $\boldsymbol{x} := \{x_n\}_{n=1}^N$, where $x_n \equiv x(t)$, x satisfies (6) and $N = 2^L$ for $L > 0$. Note that we use the bold notation to indicate the numerical implementation (or the discrete sampling) of an otherwise continuous quantity.

Step 1: numerically implement $W_x(t, a)$. We discretize the scale axis a by $a_j = 2^{j/n_v}\tau$, $j = 1, \ldots, Ln_v$, where n_v is the "voice number" chosen by the user. In practice we choose $n_v = 32$. We denote the numerical CWT as a $N \times n_a$ matrix $\boldsymbol{W}^x$. This is a well studied step and our CWT implementation is modified from that of wavelab[8].

Step 2: numerically implement $\omega_x(a, t)$.

The next step is to calculate the IF information function $\omega_x(a, t)$ (27). The $\partial_t W_x(a, t)$ term is implemented directly by finite difference at t axis and we denote the result as a $N \times n_a$ matrix $\partial_t \boldsymbol{W}^x$. The $\omega_x(a, t)$ is implemented as a $N \times n_a$ matrix $\boldsymbol{w}_x$ by the following entry-wise calculation:

$$\boldsymbol{w}_x(i,j) = \begin{cases} \frac{-i\partial_t \boldsymbol{W}^x(i,j)}{2\pi \boldsymbol{W}^x(i,j)} & \text{when } \boldsymbol{W}^x(i,j) \neq 0 \\ \text{NaN} & \text{when } \boldsymbol{W}^x(i,j) = 0. \end{cases},$$

where NaN is the IEEE arithmetic representation for Not-a-Number.

Step 3: numerically implement $S_x(t, \omega)$.

We now compute the Synchrosqueezing transform S_x (28). We discretize the frequency domain $[\frac{1}{N\tau}, \frac{1}{2\tau}]$ by equally spaced intervals of length $\Delta_\omega = \frac{1}{N\tau}$. Here $\frac{1}{N\tau}$ and $\frac{1}{2\tau}$ are the minimal and maximal frequencies detectable by the Fourier transform theorem. Denote $n_\omega = \lfloor \frac{\frac{1}{2\tau} - \frac{1}{N\tau}}{\Delta_\omega} \rfloor$, which is the number of the discretization of the frequency axis. Fix $\gamma > 0$, S_x is discretized as a $N \times n_\omega$ matrix $\boldsymbol{S}^x$ by the following evaluation

$$\boldsymbol{S}^x(i,j) = \sum_{k:|\boldsymbol{w}_x(i,k) - j\Delta_\omega| \leq \Delta_\omega/2,\, |\boldsymbol{W}^x(i,j)| \geq \gamma} \frac{\log(2)\sqrt{a_j}}{\Delta_\omega n_v} \boldsymbol{W}^x(i,k),$$

where $i = 1, \ldots, N$ and $j = 1, \ldots, n_\omega$. Notice that the number γ is a hard thresholding parameter, which is chosen to reduce the influence of noise and numerical error. In practice we simply choose $\gamma = 0.1$. If the error is Gaussian white noise, the choice of γ is suggested in Thakur et al. (2013). In general, determining how to adaptively choose γ is an open problem.

Step 4: Estimate IF, AM and trend from $\boldsymbol{S}^x$.

We fit a discretized curve $\boldsymbol{c}^* \in Z_{n_\omega}^N$, where $Z_{n_\omega} = \{1, \ldots, n_\omega\}$ is the index set of the discretized frequency axis, to the dominant area of $\boldsymbol{S}^x$ by maximizing the following functional over $\boldsymbol{c} \in Z_{n_\omega}^N$:

$$\Big[\sum_{m=1}^N \log \left(\frac{|\boldsymbol{S}^x(m, \boldsymbol{c}_m)|}{\sum_{i=1}^{n_\omega} \sum_{j=1}^N |\boldsymbol{S}^x(j,i)|} \right) - \lambda \sum_{m=2}^N |\boldsymbol{c}_m - \boldsymbol{c}_{m-1}|^2 \Big],$$

[8] http://www-stat.stanford.edu/~wavelab/

where $\lambda > 0$. The first term is used to capture the maximal value of $\boldsymbol{S}^x$ at each time and the second term is used to impose regularity of the extracted curve. In other words, the user-defined parameter λ determines the "smoothness" of the resulting curve estimate. In practice we simply choose $\lambda = 10$. Denote the maximizer of the functional as $\boldsymbol{c}^* \in \mathbb{R}^N$. In that case, the estimator of the IF of the k-th component at time $t = n\tau$ is defined as

$$\boldsymbol{\phi'}_k(n) := \boldsymbol{c}^*(n)\Delta_\omega,$$

where $\boldsymbol{\phi'}_k \in \mathbb{R}^N$. With $\boldsymbol{c}^*$, the k-th component $a_k(t)\cos(2\pi\phi_k(t))$ and its AM, $a_k(t)$ at time $t = n\tau$ are estimated by:

$$\boldsymbol{f}_k(n) := \Re\frac{2}{\mathcal{R}^\psi}\Delta_\omega \sum_{i=\boldsymbol{c}^*(n)-\lfloor\Delta/\Delta_\omega\rfloor}^{\boldsymbol{c}^*(n)+\lfloor\Delta/\Delta_\omega\rfloor} \boldsymbol{S}^x(n,i),$$

$$\boldsymbol{a}_k(n) := \left|\frac{2}{\mathcal{R}^\psi}\Delta_\omega \sum_{i=\boldsymbol{c}^*(n)-\lfloor\Delta/\Delta_\omega\rfloor}^{\boldsymbol{c}^*(n)+\lfloor\Delta/\Delta_\omega\rfloor} \boldsymbol{S}^x(n,i)\right|,$$

where $\Re$ is the real part, $\boldsymbol{f}_k \in \mathbb{R}^N$ and $\boldsymbol{a}_k \in \mathbb{R}^N$. Lastly, we estimate the trend $T(t)$ at time $t = n\tau$ by

$$\boldsymbol{T}(n) := \boldsymbol{x}_n - \Re\frac{2}{\mathcal{R}^\psi}\Delta_\omega \sum_{i=\lfloor\omega_\mathrm{l}/\Delta\omega\rfloor}^{n_\omega} \boldsymbol{S}^x(n,i),$$

where $\boldsymbol{T} \in \mathbb{R}^N$ and $\omega_\mathrm{l} > 0$ is determined by the model.