Point-to-point response to AE:

Four reports have been received, including two who reviewed the previous submission. One of the old referees did not give further comments; the other still thinks that the statistical inference needs be strengthened. One of the new reviewers does not think SST is the right approach to signal detection and believes that the test might be less powerful than the chirplet path pursuit of Cand`es et al. (2008, Appl. Comput. Harmon. Anal. pp.14–40). The other new reviewer finds this paper could have promising application in detecting the signal in astronomy, but also thinks that another transformation, the Short-time Fourier transform (STFT), is applicable to more general signals than SST and should be investigated.

I agree with the comments of the reviewers. The paper needs to consider other possible approaches: i) to make a comprehensive comparison of the proposed method with Cand`es et al. (2008), especially when the time series has large noise and high oscillation, and ii) extend the theory to STFT. In addition, I find the detail of bootstrapping is not clearly presented, for example, how to select the local neighbor B.

We thank the AE for his/her time and effort for handling this manuscript.

We found that we didn’t make it clear what SST is, and why we work hard on this paper, so the reviewers and AE are probably confused and think SST is designed specifically to detect an oscillatory component. Note that SST was introduced back in 2010, and has been applied widely for various missions. However, so far almost all theoretical works for SST are based on the deterministic model, and little effort was put on the statistical part for statistical inference. Thus, the focus of this paper is establishing the asymptotic distribution of SST, and set up a foundation for understanding various applications of SST. Among all applications of SST (Typical applications include (1) estimate the time-varying frequency and amplitude, (2) obtain the non-sinusoidal oscillatory pattern, (3) decompose the constitutional oscillatory components and their phase functions from a noisy observation, (4) determine if there is an oscillatory component and when it exists, and many others), detecting the existence of the oscillatory component is the one that can be immediately explored with the developed theoretical foundation. That is why we chose it as an application in this paper. We have improved the introduction and hope we now have made this point clear.

In addition to this major change, we have addressed all valuable comments by the reviewers and provided a point-by-point response. All the changes in the manuscript we made are marked in red. We should make clear the two points raised by the AE:

(i) Note that the SST that we analyze in this paper is designed to handle oscillatory signals with slowly varying time-varying frequency but not chirps. While it is possible to extend SST to handle chirps (like 2nd order SST or higher order SST, see citation [27]), their theoretical foundation has not been established since this paper is the foundation toward these more complicated extensions. Therefore, we do not think applying the 2nd or higher order SST without asymptotic theory for a comparison will be fair. On the other hand, the chirplet path pursuit of Cand`es et al. is a different approach that is based on the chirplet transform. While the chirplet path pursuit is certainly applicable to the blackhole signal as requested by the third reviewer, we
found a comparison in such a signal is not suitable since the SST we analyze in this paper is not designed for that purpose (but anyways, we still provide the 2nd SST result of the blackhole signal for Reviewer 3 to convince the reviewer). Thus, in the revision, to have a fair and square comparison, we consider a more complicated simulation with a more complicated time-varying frequency and contaminated by big & colored noise, but the time-varying frequency is slowly varying but not a strong chirp.

(ii) We believe this request comes from the second point of Reviewer 3. We shall mention that STFT is the base of SST. Without its properties, we cannot establish the properties of SST. So, the properties of STFT have been extensively used in the manuscript. To make this point clear, we have added a new section developing the STFT theory as the reviewer and AE requested. Moreover, in the oscillatory component detection example, we provide a comparison with a STFT-based approach.

(iii) We also thank the AE for indicating the issue about B for the local bootstrapping. Since the null model we work with in this paper is a stationary noise, to reduce the confusion, we have removed the local bootstrapping and the parameter B in this revision. Instead, a global bootstrapping procedure is used for the implementation under stationary noises. We also made the details of the bootstrap clear in the revision. Please refer to Section 6.3.

We hope the AE would understand and be satisfied by our decision of how we address your and reviewers’ comments.
Point-to-point response to Reviewer 1:

This paper establishes a statistical analysis of a tool in nonlinear-type time-frequency analysis, the synchrosqueezing transform (SST), for both the null and non-null cases. Furthermore, a block bootstrap scheme is designed based on the established SST theory to test if a given time series contains oscillatory components.

In the estimation problem where it is interesting to estimate the instantaneous frequencies of oscillatory signals, SST is useful. However, in the detection problem, especially deciding whether an oscillatory signal exists in a time series with heavy noise, SST should not be the choice for statistical inference. The reason is simple: SST requires to take the derivative of a noisy signal and the derivative operator significantly enhance the noise effect (extremely not stable against noise). Hence, from this mathematical principle, it is easy to understand that SST is not suitable for statistical inference.

We respectfully disagree with the reviewer regarding this differentiation issue. Actually the “differentiation” does not enhance the noise effect since the differentiation is over an integration of an integral kernel $g$, and such differentiation can be represented by an integration of another integral kernel $h$, which is the differentiation of $g$. The fact is that the stability of SST has been long understood and analyzed in the time-frequency literature (see the citations 10 and 35 in the revised manuscript), so we only quickly addressed it in the previous manuscript. However, we do understand that this commonly confusing point might mislead the reading, so we have added a remark on page 6 of the revised manuscript to resolve this issue:

In this sense, other time-frequency transforms without derivatives are more suitable, e.g., using Chirplet to detect oscillatory signals from noisy time series in the literature. Therefore, though the theoretical analysis in this paper is technical and can make some contribution, it is not convincing that this paper is quantified for Annals of Statistics.

Following the comment above, SST does not involve derivative. It is written & expressed in the derivative format simply to emphasize its resource and purpose – the phase information is used in the transform. In real world data, SST has been applied extensively to analyze various time series from different fields, even those with big and complicated noise. This is the reason we develop the statistical foundation of this widely applied method (note that we have also revised the Introduction to better conduct this fact and motivate this work). On the other hand, the chirplet transform over a frame does have its benefit for the oscillatory signal detection purpose, and in practice we found its
performance similar to SST. We have added a section comparing SST with the chirplet transform approach, particularly the chirplet path pursuit algorithm, in Section 7.
**Point-to-point response to Reviewer 2:**

I appreciate that the authors have made great effort to incorporate my suggestions and strengthen the inference part in the paper. My main comments for this version are as follows:

1. About the motivating example (Figure 1) in Section 1: I think that the dataset has changed from the previous version of the paper, so the figure and conclusions have changed. But I’m quite confused by the corresponding (2nd) paragraph on page 3. Where are the “blue arrows” in Figure 1? How can you tell from the figure that the “texture structure could mask the information of interest”? Do you compare them with the blue diamonds in the figure? The interpretation is not clear to me. In addition, for the blue diamonds in Figure 1, please report the corresponding significance level.

   We thank the reviewer for indicating this unclear point. Yes, the original example is changed from the peripheral venous pressure (PVP) signal to the photoplethysmogram (PPG) signal since after talking to experts in PVP signals, so far there is no standard to decide if a PVP signal is of good quality, unless under the most ideal situation. However, for PPG we know a lot more. Thus, we decided to change to PPG that we know the signal quality, so that the scientific argument is solid. And yes, the blue arrows are missing, and we have added them back. For the information, we have made it clear what we meant by revising this sentence:

   See Figure 1 for an example of SST when applied to a noisy photoplethysmography (PPG) signal. The PPG signal is non-invasive, cost-effective, and widely used in healthcare environment [1]. However, noise is inevitable in the clinic environment, which might downgrade the reliability of the signal. In this example, the second half of the signal is of low quality in the sense that the cardiac oscillation cannot be visualized. In the associated TFR determined by SST, there is a curious “texture” structure (indicated by blue arrows) in the second half, which comes from the low-quality signal, while there is a dominant curve (indicated by red arrows) in the first half, which encodes the time-varying heart rate information. If we could understand the asymptotic distribution of SST under various situations, we could further utilize information encoded in the signal.

   The p-value, the false discovery rate and the threshold are now included in the bottom subfigure in Figure 1.

2. While this version of the paper has been improved significantly, I still feel that inference is a relatively small part of the paper. Can you further strengthen it by, e.g., considering multiple testing procedures for a period of time?

   Certainly. We have extended this part in Section 6.2, which reads:
To determine when an oscillatory component exists over a period, we could repeat the above steps over a set of chosen timestamps. For the chosen kernel in Assumption 4.3, $h(c)$ is numerically zero for a sufficiently large constant $c > 0$, say, $c = 10$. Therefore, numerically the SST coefficients are independent if two timestamps are separated by $2c$. Thus, we choose a uniform grid $(t_1, t_2, \ldots, t_{r'}) \subset [-\sqrt{N}/2, \sqrt{N}/2]$ on the time axis, where $r' \in \mathbb{N}$ is chosen so that $t_i - t_{i-1} = \lfloor \sqrt{\log N} \rfloor$. Then, for each timestamp $t_i$, evaluate the Fisher-SST statistic and obtain the p-value $p_i$. Since we will run this test for $r' = \lfloor \sqrt{N}/\sqrt{\log N} \rfloor$ times and these tests are independent, we recommend to consider the false discovery rate control [4] to handle the multiple testing issue in the following way. Rank obtained p-values $p_1, \ldots, p_{r'}$ as $p(1) \leq p(2) \leq \ldots \leq p(r')$. Take $q > 0$ to be the false discovery rate and set $k^*$ to be the largest integer so that $p(k^*) \leq k^*q/r'$. At time $t_i$, the null hypothesis is rejected if $p_i \leq p(k^*)$. 


Point-to-point response to Reviewer 3:

This paper introduces a theoretical framework to investigate asymptotic distribution of a nonlinear-type TF analysis algorithm such as the synchrosqueezing transform (SST), in both the null and non-null cases. In particular, the three main challenges when investigating a nonlinear-type TF analysis algorithm were raised and dealt with very convincingly. Then, all of these analyses are used to for statistical inference such as oscillatory signal detection. The paper is really interesting including a rigorous and complex mathematical analysis and nice numerical simulations. In my opinion, it is only ready for being published on the TSP unless the following issues are addressed.

1. The main work of this paper focuses on studying asymptotic distribution of the SST and only a small part related to the statistical inference of SST for determination of an oscillatory component of the signal. Could the authors add more applications of using such an asymptotic distribution study, rather than only determining an oscillatory component. This is because that the current title is too broad compared to what is presented in the paper.

We appreciate the reviewer’s suggestion. We understand why the reviewer worries about the too broad title “inference of synchrosqueezing transform”. We fully agreed with the reviewer that no single paper can fully address all inference problems, and we obviously do not achieve so. To avoid this issue, we have decided to change the title from

“Inference of synchrosqueezing transform -- toward a unified statistical analysis of nonlinear-type time-frequency analysis”


to

“Asymptotic analysis of synchrosqueezing transform -- toward statistical inference with nonlinear-type time-frequency analysis”


to better reflect the material presented in this work.

On the other hand, note that the focus of this paper is setting the foundation of the statistical inference with SST by establishing the asymptotic distribution of SST when the input is a random process under the null and non-null setups. Among many possible applications based on the asymptotic distribution study (For example, typical applications of SST in the signal processing society include (1) estimate the time-varying frequency and amplitude, (2) obtain the non-sinusoidal oscillatory pattern, (3) decompose the constitutional oscillatory components and their phase functions from a noisy observation, (4) determine if there is an oscillatory component and when it exists, and many others), we chose to discuss the oscillatory component detection problem since the developed theory can be applied immediately. However, even setting this foundation and the oscillatory component detection has taken almost 90 pages, due to the nature of the
complicated structure in its nonlinear transform. An exploration of each of the other applications needs a whole other paper since more theoretical foundations need to be established, which we found not feasible to add in this already-long paper.

Moreover, we realized that in the introduction section we didn’t make it clear what SST is, and why we work hard on this paper, so the reviewer is confused and think SST is designed specifically to detect an oscillatory component. Note that SST was introduced back in 2010, and has been applied widely for various missions. However, so far almost all theoretical works for SST are based on the deterministic model, and little effort was put on the statistical part for statistical inference. Thus, the focus of this paper is establishing the asymptotic distribution of SST, and set up a foundation for understanding various applications of SST. We have improved the introduction and hope we now have made this point clear.

After discussing among our team, we decided to also expand the discussion section about exploring other statistical applications, particularly those that have been widely considered in the signal processing society with proper citations and postpone those exploration in our future work. Specifically, we added the following sentence in Section 8.5

SST and these TF analysis tools have been widely applied in the signal processing society. See [37] for a list of applications. Inspired by these applications, there are many important statistical inference problems remain open. For example, how to establish the inference procedure for estimating instantaneous frequency, amplitude modulation and phase function and decomposing the noisy signal into its constitutional components? How to handle the nonstationary noise (e.g. the piecewise locally stationary [43]) or study the statistical structure of a random process? How to detect the number of oscillatory components present inside a noisy signal or at which times such components exist?

We sincerely hope the reviewer agrees with our decision regarding our feedback to this valuable comment.

2. As it is well known that SST is only concentrated time-frequency (TF) version of short-time Fourier transform (STFT) through the two formulas (3) and (4). It means that all signal structures including the number of its oscillatory components are kept the same when moving from STFT domain to SST domain. My questions is whether there is oscillatory signal detection algorithm using a statistical inference via STFT. In my opinion, if that is possible, it should be more simpler while developing such an asymptotic analysis in the STFT space. If not, please explain why.

We thank the reviewer for raising this question. Note that the STFT distribution has been established in the manuscript, which is the foundation of the SST theory
(actually, the distribution of SST can be seen in Equation 18). We have added a new section to develop the necessary STFT theory for our purpose in this paper. Please see Section 5.1.

While developing the asymptotic analysis in the STFT space is simpler than SST, it is out of the scope of this paper. On the other hand, on practice, it has been well known that the performance of STFT in several signal processing missions, for example the ridge detection, is lower compared with that of SST (See S. Meignen, D.-H. Pham, and S. McLaughlin, “On demodulation, ridge detection, and synchrosqueezing for multicomponent signals,” IEEE Trans. Signal Process., vol. 65, no. 8, pp. 2093–2103, Apr. 2017.). This is because the time-frequency representation determined by STFT is blurred by the uncertainty principle, while that determined by SST is sharpened. The contrast enhancement of the time-frequency representation plays an important role here. To demonstrate this point, we have extended the numerical section with a realistic simulation (instead of the harmonic example in the previous submission) and show that the oscillatory detection based on STFT performs worse compared with SST.

3. The application of the proposed bootstrapping algorithm to the PPG signal shown in Figure 1 (page 20/25), is not clear to me. Could the authors show a rejection rate about oscillations to this signal?

The application is for the signal quality estimation of the PPG signal, which is a critical step in analyzing real world biomedical signals. We have made this part clear in Section 7 and added the p-value, false discovery rate, and other details in Figure 1. Specifically, we added:

\textit{A critical biomedical signal processing step, particularly for long-term monitoring in clinics, is determining when the signal quality is trustworthy so that the obtained information is usable for decision making. Signal quality assessment is in general challenging, and the strategy depends on the clinical problem. When the heart rate and its variability are the concern, we care if a PPG signal oscillates properly, and the TFR of a high-quality PPG encodes the time-varying heart rate as a curve with distinguishable intensity. See Figure 1 for an example, where the PPG signal in the first 50 second is of high quality. It is visually obvious to see an oscillation from 0th to 40th second, which are cardiac cycles. However, after 50th second, the signal looks chaotic and it is not clear if it provides any useful cardiac information. The signal between 40th and 50th second is also oscillatory, but the pattern is slightly less clear. This visual inspection suggests that the signal quality over each segment of predetermined length could be determined by the confidence of oscillation detection. Note that the frequency and amplitude might change slowly, so it is reasonable to assume that locally the signal oscillates with fixed}
amplitude and frequency, and the proposed local bootstrapping algorithm and the Fisher-SST statistic could be applied.

Furthermore, it should be interesting to show the applicability of this proposed algorithm to gravitational waves from a binary black hole merger.

https://www.ligo.org/science/Publication-GW150914/

Note that the gravitational waves are chirps, and the SST we study in this paper is not designed to analyze this kind of signal. The correct tool to analyze this gravitational wave is the second order SST or higher order SST (see below), which we have not established theoretical foundations yet (the SST under analysis in this paper is the first order SST). While in the numerical section we could run the higher order SST algorithm for a comparison, we found it improper to do so in AoS when a theory is missing. Thus, to avoid further problem, after discussing among the coauthors, we decided not to include it since the (first order) SST under exploration is not the correct tool for it. Instead, we offered two alternatives to answer the reviewer’s question. We hope the reviewer understands and agrees with our decision.

(a) In the simulation section, we consider a much more complicated simulation with a time-varying frequency with a chirp term (But the chirp is not as strong as the black hole signal) with a complicated noise. We show that SST performs better than the chirplet-based approach, particularly the chirplet path pursuit algorithm.

(b) Below is how the 2nd order SST of the blackhold signal looks like. In this plot, we show two gravitational signals in the first row, the first order SST (analyzed in this paper) in the second row, and the second order SST (not analyzed in this paper) in the third row. We see that over the period with a fast-varying frequency (indicated by the blue circle), the time-frequency representation determined by the first order SST is blurred and that of the second order SST is concentrated. We mention that the performance of detecting the chirp with the first order SST is similar to that with the chirplet approach (mainly due to the blurring part), while the 2nd order SST approach performs better. We will extend the current asymptotic analysis to the second order & higher order SST and its application to such signals to our future work.

Note that an application of SST to study such signal has been shown in several places. See, for example [Pham, Duong-Hung, and Sylvain Meignen. "High-order synchrosqueezing transform for multicomponent signals analysis—With an application to gravitational-wave signal." IEEE Transactions on Signal Processing 65.12 (2017): 3168-3178] and [Coifman, Ronald R., Stefan Steinerberger, and Hau-tieng Wu. "Carrier frequencies, holomorphy, and unwinding." SIAM Journal on Mathematical Analysis 49.6 (2017): 4838-4864].
ASYMPTOTIC ANALYSIS OF SYNCHROSQUEEZING TRANSFORM – TOWARD STATISTICAL INFERENCE WITH NONLINEAR-TYPE TIME-FREQUENCY ANALYSIS

BY MATT SOURISSEAU∗ HAU-TIENG WU†, ZHOU ZHOU‡

Department of Mathematics, University of Toronto, Toronto, ON, Canada∗
Department of Mathematics and Department of Statistical Science, Duke University, Durham, NC, USA†
Department of Statistics, University of Toronto, Toronto, ON, Canada‡

We provide a statistical analysis of a tool in nonlinear-type time-frequency analysis, the synchrosqueezing transform (SST), for both the null and non-null cases. The intricate nonlinear interaction of different quantities in SST is quantified by carefully analyzing relevant multivariate complex Gaussian random variables. Specifically, we provide the quotient distribution of dependent and improper complex Gaussian random variables. Then, a central limit theorem result for SST is established. As an example, we provide a block bootstrap scheme based on the established SST theory to test if a given time series contains oscillatory components.

1. Introduction. Time series contain dynamical information of a system under observation, and their ubiquity is well-known [19]. A key task in understanding and forecasting such a system is to quantify the dynamics of an associated time series according to a chosen model, a task made challenging by the fact that often the system is nonstationary. Although there is no universal consensus on how to model and analyze time series extracted from nonstationary systems, two common schools of thought are those of time series analysis [20, 7, 15] and time-frequency (TF) analysis [12, 16]. Roughly stated, the main difference between these two paradigms is the assumptions they make on the underlying random process modeling a time series.

In classical time series analysis, this random process is typically assumed to have zero first-order statistics. The focus is then on analyzing the second-order statistics, mainly for the purpose of forecasting. Seasonality of a time series (that is, an oscillatory pattern of known periodicity in its mean) is modeled separately or included in the covariance structure [7]. When a time series is modeled as a sum of a parametric periodic mean function and a...
stationary noise sequence, there is a small body of statistics literature on methods and algorithms to estimate its periodicity when unknown. The common ground with TF analysis originates in investigating “local spectral behavior” [31], a generalization of the idea of using the spectrum to capture local behavior. See Section A for more literature review. This direction has a long history, beginning with the consideration of analytic model [17, 30] and more recently progressing to the adaptive harmonic model (AHM) modeling time-varying frequency and amplitude [13, 10], or adaptive non-harmonic model (ANHM) further modeling nonsinusoidal, time-varying oscillatory patterns [37, 27]. Unlike the classical time series analysis, in this direction the oscillation is modeled in the first-order statistics.

In the TF approach, available algorithms are roughly classified into linear-type, bilinear-type and nonlinear-type. The synchrosqueezing transform (SST) and its variations, a family of nonlinear-type TF tool, were developed based on AHM/ANHM in the past decade. SST can be viewed as a special case of the reassignment technique pioneered in [23] and further explored in [2]. SST encodes the spirit of empirical mode decomposition [21]. Broadly, SST nonlinearly modifies the TF representation (TFR), and hence the spectrogram, derived from the short-time Fourier transform (STFT) by utilizing the phase information in STFT so that the TFR is sharper. Recall that spectrogram comes from dividing a time series into short segments by a chosen window (taper), and evaluate the tapered periodogram at each moment. As a result, the phase information included in STFT is lost. The key feature that distinguishes SST from traditional spectral analysis is how the phase information is used to sharpen the TFR.

From the application perspective, SST has been applied to handle diverse signal processing challenges since its development. Typical applications include (1) estimate the time-varying frequency and amplitude, (2) obtain the non-sinusoidal oscillatory pattern, (3) decompose the constitutional oscillatory components and their phase functions from a noisy observation, (4) determine if there is an oscillatory component and when it exists, and many others. See [38] for a recent review article for its successful scientific applications in medicine. Due to its flexibility, several variations of SST have been proposed. For example, taking the S-transform [22] or wave packets [42] into account, considering higher order phase information [28], combining the cepstrum tool [27], and applying multi-taper techniques [41, 14].

From the theoretical perspective, recently the theoretical analysis of SST under the AHM or ANHM when noise does not exist has been well-established. See [13, 10] for example. However, when noise (or any stochastic process) is present, the exploration has been limited to asymptotic expansion [35, 10, 43].
or only part of the algorithm [9]. To our knowledge, its statistical property, particularly the asymptotic distribution, even in the null case (that is, no oscillatory signal), is still missing.

See Figure 1 for an example of SST when applied to a noisy photoplethysmography (PPG) signal. The PPG signal is non-invasive, cost-effective, and widely used in healthcare environment [1]. However, noise is inevitable in the clinic environment, which might downgrade the reliability of the signal. In this example, the second half of the signal is of low quality in the sense that the cardiac oscillation cannot be visualized. In the associated TFR determined by SST, there is a curious “texture” structure (indicated by blue arrows) in the second half, which comes from the low-quality signal, while there is a dominant curve (indicated by red arrows) in the first half, which encodes the time-varying heart rate information. If we could understand the asymptotic distribution of SST under various situations, we could further utilize information encoded in the signal.

Motivated by the wide application of SST and the missing asymptotic analysis of SST (e.g., for a systematic study of various applications of SST) from the statistical perspective, the main focus of this work is providing a statistical analysis of SST toward the statistical inference purpose. Specifically, we write down the distribution associated with SST of a stationary colored Gaussian random process, in both null and non-null cases. We apply the developed result to design a local bootstrapping algorithm for statistical inference for testing the existence of an oscillatory component, and its theoretical justification is also provided. This algorithm is applied to determine the signal quality of the PPG signal shown in Figure 1.

The first technical challenge encountered along the way is dealing with improper multivariable complex random variables and their ratio distributions. While proper (or, “circular”) complex random variables have been widely discussed in signal processing literature [3], their improper counterparts and corresponding ratios have been mostly ignored, except for [30, 34]. In our case, impropriety arises naturally from the phase information encoded in the STFT, and handling its effect on the quotient structure forms the first part of the paper, and is of its own interest for other applications.

The second technical challenge is handling the nonlinear reassignment of STFT coefficients according to the reassignment rule. This nonlinear reassignment is challenged by handling the nonlinear change of variable by the $\mathbb{C}\mathbb{R}$-calculus computation [25] and approximating the integration of confluent hypergeometric function that naturally pops out when we evaluate moments of SST, particularly in the non-null case. The analysis is complicated in the low frequency region due to the degeneracy of the covariance struc-
ture. This step has an interesting interpretation and helps us connect time series, TF analysis and other topics; the big picture is that the reassignment rule has a natural interpretation within the kernel regression framework of time series analysis, and can be understood in the framework of diffusion geometry in the manifold learning setup.

The third technical challenge is handling the dependence structure when we show the central limit theorem (CLT) of SST. The main technique here is exploiting the $M$-dependence, and the associated critical quantity is the “effective sampling rate” – once we find a proper $M$-dependent surrogate of the original random process associated with SST, if the “effective sampling rate” is correctly specified, we show that in both null and non-null cases, SST of a stationary colored Gaussian random process follows a complex normal distribution. With the above results, we could establish a theoretical justification of a local bootstrapping algorithm for the statistical inference.

The paper is organized in the following way. In Section 2, we summarize the STFT-based SST. In Section 3, we handle quotients of two complex Gaussian random variables. These results are of interest aside from their use in the sequel. The mathematical setup for the SST analysis is given in Section 4. Section 5 includes the asymptotic analysis of SST. A local bootstrapping

![Fig 1](image-url)
algorithm for an application of SST to the oscillatory component detection problem and its theoretical validation are shown in Section 6. A numerical simulation is shown in Section 7. We conclude the paper with a discussion in Section 8. More literature review, numerical example, and all proofs are relegated to the supplementary material. In this paper, we use the following asymptotic notations. For two sets of \(\{a_u\}_{u \in U}; \{b_u\}_{u \in U} \subset \mathbb{R}_+\) indexed by a set \(U\), \(a_u \asymp b_u\) means that there exist constants \(0 < c_1 < c_2 < \infty\) so that \(c_1 b_u \leq a_u \leq c_2 b_u\) for all \(u \in U\), \(a_u = O(b_u)\) means that there exists a constant \(c\) so that \(a_u \leq cb_u\) for all \(u \in U\). For \(a, b \in \mathbb{R}\), \(a \lor b\) means \(\max\{a, b\}\).

2. A summary of the SST algorithm. Take a Schwartz function \(h\). For a tempered distribution \(f\), the STFT of \(f\) associated with the window function \(h\) is defined by the equation \(V^{(h)}_f(t, \eta) := f(h_{t,\eta})\), where \(t \in \mathbb{R}\) is the time and \(\eta > 0\) is the frequency and \(h_{t,\eta}(s) := h(s-t)e^{-2\pi i \eta (s-t)}\). We mention that this is a modification of the ordinary STFT by the phase modulation \(e^{2\pi i \eta t}\), and we choose to work with it to simplify the upcoming heavy notation. When \(f\) is represented by a function, we may abuse notation in the usual way and write

\[
V^{(h)}_f(t, \eta) = \int_{-\infty}^{\infty} f(s) h(s-t)e^{-2\pi i \eta (s-t)} \, ds,
\]

Commonly, the window function \(h\) is chosen to be a Gaussian function with mean 0 and bandwidth \(\sigma > 0\). Given the above, the STFT-based synchrosqueezing transform (SST) of \(f\) with the modified window function \(h\) with resolution \(\alpha > 0\) is defined to be

\[
S^{(h,\alpha)}_f(t, \xi) := \int V^{(h)}_f(t, \eta) g_\alpha(\xi - \Omega^{(h)}_f(t, \eta)) \, d\eta,
\]

where the reassignment rule \(\Omega^{(h)}_f(t, \eta)\) is defined by

\[
\Omega^{(h)}_f(t, \eta) := \begin{cases} 
\frac{1}{2\pi i} \frac{\partial V^{(h)}_f(t, \eta)}{V^{(h)}_f(t, \eta)} & \text{if } V^{(h)}_f(t, \eta) \neq 0 \\
-\infty & \text{otherwise}
\end{cases}
\]

and \(g_\alpha : \mathbb{C} \to \mathbb{R}\) is an approximate \(\delta\)-distribution when restricted on \(\mathbb{R}\) when \(\alpha\) is sufficiently small. For concreteness, we will take \(g_\alpha(z) = \frac{1}{\sqrt{\pi \alpha}} e^{-\frac{|z|^2}{\alpha}}\), which has the \(L^1\) norm 1. Notice that the nonlinearity of SST over signals arises from the dependence of equation (2) on the reassignment rule, which provides information about the instantaneous frequency of the signal (as made precise in [13, 36]). \(\alpha\) is interpreted as the resolution of SST in the
frequency axis. Numerically, SST is implemented by a direct discretization of (1), (2), and (3). We will come back to this part when we discuss the proposed bootstrapping algorithm in Section 6. We should mention a commonly confusing point regarding SST. At the first glance, since a differentiation is taken when we define the reassignment rule, SST is unstable when noise exists. This is not the case since \( \partial_t \Omega_f^{(h)}(t, \eta) = -\Omega_f^{(h')}(t, \eta) + 2\pi i \eta \Omega_f^{(h)}(t, \eta) \); that is, the differentiation operator is equivalent to evaluating another STFT with the differentiation of \( h \) as the window and summing \( 2\pi i \eta \Omega_f^{(h)}(t, \eta) \). Based on this fact, the stability of SST has been established in [35, 10].

To better appreciate which kind of information is utilized in SST, rewrite \( \Omega_f^{(h)}(t, \eta) = |\Omega_f^{(h)}(t, \eta)| e^{2\pi i P_f^{(h)}(t, \eta)} \), where \( P_f^{(h)}(t, \eta) \) is the “phase” of the complex value \( \Omega_f^{(h)}(t, \eta) \). By properly choosing the branch for log, we have the relationship \( \Omega_f^{(h)}(t, \eta) = \frac{1}{2\pi i} \partial_t \log(\Omega_f^{(h)}(t, \eta)) = \frac{1}{2\pi i} \partial_t \log |\Omega_f^{(h)}(t, \eta)| + \partial_t P_f^{(h)}(t, \eta) \). Suppose the magnitude \( |\Omega_f^{(h)}(t, \eta)| \) changes slowly, we see that the reassignment rule encodes the phase information.

The goal of this paper is to initiate the study of the distribution of \( S_f^{(h, \alpha)}(t, \xi) \) for \( t \in \mathbb{R} \) and \( \xi > 0 \), where \( f \) is a deterministic tempered distribution and \( \Phi \) is a generalized random process (GRP); see Section D for a summary of GRP. In this case, to understand the statistics of \( S_f^{(h, \alpha)}(t, \xi) \), we are led to consider the distribution of the ratio of the random variables

\[
\begin{align*}
(4) & \quad V_{f + \Phi}(t, \eta) = f(h_{t, \eta}) + \Phi(h_{t, \eta}) \\
(5) & \quad \partial_t V_{f + \Phi}^{(h)}(t, \eta) = \partial_t f(h_{t, \eta}) + \Phi(\partial_t h_{t, \eta}).
\end{align*}
\]

We work under the assumptions that the noise \( \Phi \) is mean-zero. Under this assumption, we have \( \mu = [\mu_1, \mu_2]^\top \) so that

\[
(6) \quad \mu := \mathbb{E}\left[V_{f + \Phi}^{(h)}(t, \eta) \right] = \left[f(h_{t, \eta}) \right] = \left[\partial_t f(h_{t, \eta}) \right].
\]

By a direct expansion and the fact that \( \partial_t [h_{t, \eta}(s)] = -(h')_{t, \eta}(s) + i2\pi \eta h_{t, \eta}(s), \)

by definition, the reassignment rule takes the form

\[
(7) \quad \Omega_{f + \Phi}^{(h)}(t, \eta) = \frac{1}{2\pi i} \left( \frac{\mu_2 + 2\pi i \eta \Phi(h_{t, \eta}) - \Phi((h')_{t, \eta})}{\mu_1 + \Phi(h_{t, \eta})} \right).
\]

Note that \( (h')_{t, \eta}(s) = h'(s - t) e^{-2\pi i \eta (s - t)}. \) If the noise is such that \( \Phi(h_{t, \eta}) \) is zero only on a set of measure zero, we may add and subtract \( 2\pi i \eta \mu_1 \) in the numerator of (7) to obtain the almost-sure equality

\[
(8) \quad \Omega_{f + \Phi}^{(h)}(t, \eta) \overset{a.s.}{=} \eta - \frac{1}{2\pi i} \left( \frac{2\pi i \eta \mu_1 - \mu_2 + \Phi((h')_{t, \eta})}{\mu_1 + \Phi(h_{t, \eta})} \right).
\]
We will treat the *null case* when \( f(t) = 0 \), and the *non-null case* when \( f(t) \) is not identically zero. The analysis depends on understanding the random variables at hand, specifically \( \frac{2\pi i n \eta_1 - \mu_2 + \Phi((h')_t, \eta)}{\mu_1 + \Phi(h_t, \eta)} \), which we address now.

### 3. Complex Gaussians and their quotients.

Suppose the complex random vector \( Z \in \mathbb{C}^n \) can be written in the form \( Z = X + iY \) for some \( n \)-dim real-valued random vectors \( X \) and \( Y \). The density of \( Z \) is then defined to be the density of \( [X^\top \ Y^\top]^\top \in \mathbb{R}^{2n} \); that is, \( f_Z(x+iy) := f_{X,Y}(x,y) \).

**Definition 3.1 (Complex Gaussian distribution [34]).** Let \( \mu \in \mathbb{C}^n \). Suppose \( \Gamma, C \in \mathbb{C}^{n \times n} \) are Hermitian positive-definite and complex symmetric, respectively, and the Hermitian matrix \( \Gamma - C^* \Gamma^{-1} C \) is positive definite. We write \( Z \sim \mathbb{C}N_n(\mu, \Gamma, C) \) and say \( Z \in \mathbb{C}^n \) follows a *complex Gaussian distribution* with mean \( \mu \), covariance \( \Gamma \), and pseudocovariance \( C \) if

\[
f_Z(z) = \pi^{-n}(\det \Sigma)^{-1/2} e^{-\frac{1}{2}(z-\mu)^* \Sigma^{-1} (z-\mu)},
\]

where \( z \in \mathbb{C}^n \), \( z := \begin{bmatrix} z & \bar{z} \end{bmatrix} \), and \( \Sigma := \begin{bmatrix} \Gamma & C \\ C^* & \Gamma^* \end{bmatrix} \) is the *augmented covariance matrix*. \( Z \) is said to be *proper* if \( C = 0 \), and *improper* otherwise.

Note that while real Gaussian vectors are completely characterized by their mean and covariance, complex Gaussian vectors are characterized by their mean and *augmented covariance*, as is clearly seen by the structure of the matrix \( \Sigma \). If a complex Gaussian vector has uncorrelated components (that is, diagonal \( \Gamma \)), it does not necessarily follow that these components are independent, as \( C \) may be nonzero. When \( Z \) is proper, commutativity of matrix inversion with conjugation and positive-definiteness of \( \Gamma \) gives us

\[
f_Z(z) = \pi^{-n}(\det \Gamma)^{-1} e^{-(z-\mu)^* \Gamma^{-1} (z-\mu)}. \tag{10}
\]

If \( \Gamma \) is also diagonal, the real and imaginary parts of \( Z \) are seen by direct calculation to be independent \( N(\Re(\mu), \Gamma/2) \) and \( N(\Im(\mu), \Gamma/2) \) random variables, respectively.

Note that positive-definiteness of \( \Sigma \) guarantees the invertibility of \( \Gamma \), and hence by [5, Proposition 2.8.3] we have \( \det \Sigma = \det \Gamma \det P \), where the so-called *Schur complement*, \( P := \Gamma - C \Gamma^{-1} C \) is also invertible. By block matrix inversion [5, (2.8.16), (2.8.18) and (2.8.20)], we have \( \Sigma^{-1} = \begin{bmatrix} \overline{P^{-1}} & -\overline{P^{-1}} \overline{R} \\ -R \overline{P^{-1}} & P^{-1} \end{bmatrix} \), where \( R := \overline{C \Gamma^{-1}} \). Observe that since \( \Sigma \) is positive definite, \( P^{-1} \) is too. Moreover, by a direct calculation, \( \overline{P^{-1}} \overline{R} \) is symmetric.
3.1. Complex Gaussian quotient density and moments. If the complex random vector $(Z_1, Z_2) \sim \mathbb{C}N(0, \Gamma, 0)$, then the density of $Z_2/Z_1$ has a simple closed-form determined in [3]. To extend this result to the most general case, we recall from [26, p. 285, (10.2.8)], [26, p. 290, (10.5.2)] and [29, (4)] that the Hermite function $H_\nu$ of order $\nu < 0$ is an analytic function of $z \in \mathbb{C}$ and satisfies the identities
\begin{align}
H_\nu(z) &= \frac{1}{\Gamma(-\nu)} \int_0^\infty t^{-(\nu+1)} e^{-t^2-2tz} \, dt \\
H_\nu(-z) + H_\nu(z) &= \frac{2^{\nu+1} \sqrt{\pi}}{\Gamma((1-\nu)/2)} F_1\left(\frac{-\nu+1}{2}; \frac{3}{2}; z^2 \right) \\
H_\nu(-z) - H_\nu(z) &= \frac{2^{\nu+2} \sqrt{\pi} z}{\Gamma(-\nu/2)} F_1\left(\frac{-\nu+1}{2}; \frac{3}{2}; z^2 \right),
\end{align}
where $F_1(\cdot; \cdot)$ is the confluent hypergeometric function [26, p. 239, (9.1.4)].

With these in mind, we have the following new result:

**Theorem 3.1 (Complex Gaussian quotient density).** If $[Z_1 \ Z_2] \sim \mathbb{CN}_2(\mu, \Gamma, 0)$ and $Q = Z_2/Z_1$, then the density of $Q$ is given by
\begin{equation}
f_Q(q) = \frac{e^{-\frac{1}{2}q^\top \Gamma^{-1} \mu} \Gamma(\frac{1}{2})}{\pi \sqrt{\text{det} \Gamma \text{det} P}} \int_0^\pi F_1\left(2, \frac{1}{2}; \frac{B_\mu(\theta, q)^2}{A(\theta, q)} \right) \frac{1}{A(\theta, q)^2} \, d\theta,
\end{equation}
where
\begin{align}
A(\theta, q) &= q^\top P^{-1} q - \Re(e^{i\theta} q^\top R^\top P^{-1} q) \\
B_\mu(\theta, q) &= \Re(e^{i\theta}(\mu^* - \mu^\top R^\top P^{-1} q))
\end{align}
and $q := [1 \ q]^\top \in \mathbb{C}^2$. In the special case when $C = 0$ and $\mu = 0$, we write $Q^0$ instead of $Q$, and we have $f_{Q^0}(q) = \frac{1}{\pi \text{det} \Gamma (q^\top P^{-1} q)^{-\frac{1}{2}}}$.

The proof of Theorem 3.1 is given in Appendix C.1. The symbols $A$ and $B_\mu$ originating from the existence of the impropriety are introduced to keep the formula (14) and the following analysis succinct. Notice that $B_\mu$ vanishes when the mean $\mu$ is zero. Moreover, $f_{Q^0}$ does not blow up at 0 because $\|q\|^2 = 1 + |q|^2$ for all $q$ and decays like $\|q\|^{-4}$ as $|q| \to \infty$.

**Theorem 3.2.** Let $[Z_1 \ Z_2] \sim \mathbb{C}N_2(\mu, \Gamma, C)$ and $Q = Z_2/Z_1$. Then
(i) $\mathbb{E}|Q|^\beta$ and $\mathbb{E}Q^\beta$ are finite when $0 \leq \beta < 2$, and infinite when $\beta = 2$.
(ii) When $C = 0$ and $\mu = 0$, we have $\mathbb{E}Q^0 = \Gamma_{21}/\Gamma_{11}$. 
When $C = 0$, $\Gamma$ is diagonal, and $\mu = (\mu_1, 0) \in \mathbb{C}^2$, we have $E Q = 0$.

The proof of Theorem 3.2 is given in C.2. Note that when $\Gamma$ is not diagonal, $Z_1$ and $Z_2$ are dependent. In the SST application we soon consider, $\Gamma$ will automatically be diagonal, and for most applications, a whitening process can achieve this condition.

**Remark.** We mention a point of potential confusion from the literature. It is well-known that the quotient of two real Gaussian random variables (independent or not) has a Cauchy tail, and its density is given explicitly in [29, Theorem 2]. One might expect a parallel statement that “a quotient of complex Gaussians should have a complex Cauchy distribution”. However, this is not the case: a distribution by the name of complex Cauchy has already appeared in the literature [34, p. 46, (2.80) with $n = 1$], its density being given by $f(z) = \frac{1}{\pi \sqrt{\det S}} \left(1 + (\bar{z} - \mu)^* S^{-1} (\bar{z} - \mu)\right)^{-3/2}$ for some location parameter $\mu$ and scatter (or dispersion) matrix $S$. This does not coincide with the distribution of the quotient of two complex Gaussians; for example, the former does not have a mean [34], whereas the latter does.

**4. Mathematical model for SST analysis.** The following mathematical model is considered to analyze SST. We follow the ideas in [18, 24] and introduce a complex version of Gaussian noise. Then, we study each step of the algorithm.

**Definition 4.1** (Stationary complex Gaussian noise). A stationary GRP $\Phi$ is complex Gaussian if for any finite collection $\psi_1, \ldots, \psi_n \in S$ we have $(\Phi(\psi_1), \ldots, \Phi(\psi_n)) \sim \mathbb{C}N_n(0, \Gamma, C)$, where $\Gamma_{i,j} = E[(\Phi(\psi_i) - E[\Phi(\psi_i)])(\Phi(\psi_j) - E[\Phi(\psi_j)])]$ and $C_{i,j} = E[(\Phi(\psi_i) - E[\Phi(\psi_i)])(\Phi(\psi_j) - E[\Phi(\psi_j)])]$ for all $1 \leq i, j \leq n$. Denote the spectral measure associated with $\Phi$ by $d\vartheta(\xi)$, where $\int (1 + |\xi|)^{-2} d\vartheta(\xi) < \infty$ for some $l > 0$; see Section D for more information.

**Assumption 4.1.** For the stationary GRP $\Phi$, we assume that the spectral measure associated with $\Phi$ is absolutely continuous related to the Lebesgue measure and there exists a smooth function $p(\xi)$, the spectral density of $\Phi$, so that $d\vartheta(\xi) = p(\xi) d\xi$ by the Radon-Nikodym theorem. Further, we assume that $p(\xi) \propto (1 + |\xi|)^\varrho$ for $\xi \in \mathbb{R}$, where $\varrho < 2l - 1$. When $\varrho = 0$, the GRP is white; otherwise it is colored.

A concrete example of the colored GRP is the continuous autoregressive and moving average (CARMA) random process [6, (2.16)]. The CARMA(1,0)
process is defined as a stationary solution of the first-order stochastic differential equation $(D + a)X(t) = bDW(t)$, where $a > 0$, $b \neq 0$, $t \geq 0$, $D$ is the differentiation with respect to $t$ in the proper sense, and $\{W(t)\}$ is the standard Brownian motion. We have $\mathbb{E}X(t) = 0$ and the power spectrum function is $p(\xi) = \frac{b^2}{2\pi |\xi + a|^2}$; that is, $\varrho = -2$.

To study the non-null case, we focus on the oscillatory signal we have interest. In practice, the frequency and amplitude of an oscillatory signal both vary with time [13, 10]. The AHM is commonly applied to model such oscillatory signals, where the frequency and amplitude are assumed to change slowly relative to its time-varying frequency. Under the slowly varying assumption, a function satisfying the AHM can be well-approximated locally by a single harmonic component [13, 10]. In light of this, we assume from now on the following:

**Assumption 4.2.** Consider $Y = f + \Phi$, where $\Phi$ is a stationary GRP and $f(t) = A e^{2\pi i (\xi_0 t + \phi_0)}$ is the oscillatory signal for some fixed frequency $\xi_0 > 0$, phase shift $\phi_0 \in [0, 1)$, and amplitude $A \geq 0$. We refer to the situation when $A = 0$ as the null case, otherwise the non-null case.

Since working with a more general kernel will not provide more insight to understanding SST but the notation will become highly intense, we make the following assumption in the following analysis.

**Assumption 4.3.** The window is $h(x) = (2\pi)^{-1/2} e^{-x^2/2}$.

5. **Statistical analysis of SST.** From now on, unless otherwise described, we always assume Assumptions 4.1, 4.2 and 4.3 hold.

5.1. **The statistical behavior of STFT.** We start from establishing the statistical behavior of STFT.

**Theorem 5.1.** Suppose Assumptions 4.1, 4.2 and 4.3 hold. For any $t \in \mathbb{R}$ and $\eta, \eta' > 0$, then $[\Phi(h_{t, \eta}) | \Phi(h_{t, \eta'})] \overset{\text{d}}{\sim} \mathcal{CN}_2(0, \Gamma_{\eta, \eta'}, C_{\eta, \eta'})$, where

$$
\Gamma_{\eta, \eta'} = \begin{bmatrix}
\gamma_0(\eta, \eta) & e^{-\pi^2(\eta' - \eta)^2} \gamma_0(\eta, \eta') \\
e^{-\pi^2(\eta' - \eta)^2} \gamma_0(\eta', \eta) & \gamma_0(\eta', \eta')
\end{bmatrix},
C_{\eta, \eta'} = e^{-\pi^2(\eta' + \eta)^2} \begin{bmatrix}
\nu_0(\eta, \eta) & \nu_0(\eta, \eta') \\
\nu_0(\eta', \eta) & \nu_0(\eta', \eta')
\end{bmatrix}.
$$

$$
\gamma_0(\eta, \eta') := \int e^{-4\pi^2(\xi + \eta + \eta')^2} d\vartheta(\xi) \text{ and } \nu_0(\eta, \eta') := \int e^{-4\pi^2(\xi + \eta + \eta')^2} d\vartheta(\xi). 
$$

In other words, we have

$$
\text{Cov}(V_{f+\Phi}^{(h)}(t, \eta), V_{f+\Phi}^{(h)}(t, \eta')) = e^{-\pi^2(\eta' - \eta)^2} \gamma_0(\eta, \eta')
$$

$$
\text{Cov}(V_{f+\Phi}^{(h)}(t, \eta), V_{f+\Phi}^{(h)}(t, \eta')) = e^{-\pi^2(\eta' + \eta)^2} \nu_0(\eta, \eta').
$$
In other words, at a fixed time $t$, the STFT coefficient at each frequency $\eta$ is a complex normal distribution, and the dependence of two coefficients of two frequencies, $\eta$ and $\eta'$, decay exponentially fast when $|\eta - \eta'|$ increases. We refer readers to [44] for more discussion of statistical inference via STFT. Next, we prepare results to study SST. By equations (4) and (5), we investigate the noise structure $W_{t,\eta} := [\Phi(h_{t,\eta}) \quad \Phi((h')_{t,\eta})]^{\top}$ The second-order statistics of $W_{t,\eta}$ are computed in the following lemma.

**Lemma 5.1.** For any $t \in \mathbb{R}$ and $\eta > 0$, $W_{t,\eta} \sim \mathbb{C}N_{2}(0, \Gamma_{\eta}, C_{\eta})$, where

$$
\Gamma_{\eta} = \begin{bmatrix}
\gamma_0(\eta) & -2\pi i\gamma_1(\eta) \\
2\pi i\gamma_1(\eta) & 4\pi^2\gamma_2(\eta)
\end{bmatrix}
$$

and

$$
C_{\eta} := e^{-4\pi^2\eta^2} \begin{bmatrix}
\gamma_0(0) & 2\pi i\eta\gamma_0(0) \\
2\pi i\eta\gamma_0(0) & 4\pi^2[\gamma_2(0) - \eta^2\gamma_0(0)]
\end{bmatrix},
$$

and $\gamma_k(s) := \int_{-\infty}^{\infty} e^{-4\pi^2(\xi+s)^2}(\xi + s)^k d\vartheta(\xi)$ for $s \in \mathbb{R}$ and $k \geq 0$.

See Section F for a proof. Clearly, $C_{\eta} \to \Gamma_{\eta}$ when $\eta \to 0$. Therefore, the eigenvalues of the augmented covariance matrix becomes more degenerate when $\eta \to 0$. On the other hand, $C_{\eta} \to 0$ when $\eta \to \infty$. Note that when the noise is white, that is, $d\vartheta(\xi) = d\xi$, the formula are simplified as

$$
\Gamma_{\eta} = \frac{1}{2\sqrt{\pi}} \begin{bmatrix} 1 & 0 \\ 0 & 1/2 \end{bmatrix}
$$

and

$$
C_{\eta} = \frac{e^{-4\pi^2\eta^2}}{2\sqrt{\pi}} \begin{bmatrix} 1 & 2\pi i\eta \\ 2\pi i\eta & 1/2 - 4\pi^2\eta^2 \end{bmatrix}.
$$

Due to the relationship between different moments, in general $\Gamma_{\eta}$ is not degenerate with a lower bound for eigenvalues, like the white noise case (see Lemma F.4).

5.2. The statistical behavior of the reassignment. Note that (4), (5) and (6) are reduced to

$$
V_{f+\Phi}^{(h)}(t, \eta) = f(t)\hat{h}(\eta - \xi_0) + \Phi(h_{t,\eta}),
$$

$$
\partial_t V_{f+\Phi}^{(h)}(t, \eta) = f'(t)\hat{h}(\eta - \xi_0) + 2\pi i\eta \Phi(h_{t,\eta}) - \Phi((h')_{t,\eta}),
$$

$$
\mu = [f(t)\hat{h}(\eta - \xi_0) - f'(t)\hat{h}(\eta - \xi_0)]^{\top} = [\mu_1 \quad 2\pi i\xi_0\mu_1]^{\top},
$$

where $\hat{h}$ is the Fourier transform of $h$, $\mu_1 = f(t)\hat{h}(\eta - \xi_0)$, and (8) becomes

$$
\Omega_{f+\Phi}^{(h)}(t, \eta) = \eta - \frac{1}{2\pi i} \left( \frac{2\pi i(\eta - \xi_0)\mu_1 + \Phi((h')_{t,\eta})}{\mu_1 + \Phi(h_{t,\eta})} \right) =: \eta - \frac{1}{2\pi i} Q_{f+\Phi}^{(h)}(t, \eta).
$$

Since $\Omega_{f+\Phi}^{(h)}(t, \eta)$ and $Q_{f+\Phi}^{(h)}(t, \eta)$ are linearly related, we only need to study one of them. As a quotient of two complex Gaussian random variables, the behavior of $Q_{f+\Phi}^{(h)}(t, \eta)$ could be immediately understood from Section 3.1 when the mean is $\mu = [\mu_1 \quad 2\pi i(\eta - \xi_0)\mu_1]^{\top}$. For example, the variance of
$Q_{f+\Phi}(t,\eta)$ does not exist. In the special case when $\eta = \xi_0$, by (18), we have

\begin{equation}
\Omega_{f+\Phi}(t,\xi_0) \overset{a.s.}{=} \xi_0 - \frac{1}{2\pi i} \left( \frac{\Phi((h')_t,\xi_0)}{f(t)} + \Phi(h_t,\xi_0) \right)
\end{equation}

since $\hat{h}(0) = 1$. When $\xi_0$ is sufficiently large so that the pseudocovariance is small by Lemma 5.1 and the noise is white so that $\Gamma_{\eta}$ is diagonal, by Theorem 3.2(iii), we know that $\Phi((h')_t,\xi_0)$ has a mean bounded by $Ce^{-4\pi^2\xi_0^2\eta_{\max}\{\eta,\eta'\}}$ for some $C > 0$. This says that when $\eta = \xi_0$ and the noise is white, the reassignment rule gives accurate frequency information. See Section G for an argument.

Since $h$ is Gaussian, the covariance between $\Phi(h_t,\eta)$ and $\Phi(h_t,\eta')$ decays exponentially when $|\eta - \eta'|$ increases. Intuitively, the covariance between $Q_{f+\Phi}(t,\eta)$ and $Q_{f+\Phi}(t,\eta')$ should also be small when $|\eta - \eta'|$ is large, but the ratio structure might obfuscate the speed of decay. The following theorem shows that this intuition is true. See Section I for the proof.

**Theorem 5.2 (Ratio covariance).** Suppose Assumptions 4.1, 4.2 and 4.3 hold. For distinct $\eta, \eta' > 0$, as $|\eta - \eta'| \to \infty$ we have

\begin{equation}
\text{Cov}(Q_{f+\Phi}(t,\eta), Q_{f+\Phi}(t,\eta')) = O((\eta + \eta')^2 e^{-|\eta - \eta'|^2}).
\end{equation}

Writing down the distribution of $Z_{\alpha,\xi,\eta}$ is not simple. We need to carry out a careful change of variable argument with the $CR$-calculus computation [25] for this purpose, which is of its own independent interest and is summarized in Section B.

The seemingly complicated random variable $Y_{f+\Phi}^{(h,\alpha,\xi)}(t,\eta)$ turns out to have nice behavior – while the reassignment rule $\Omega_{f+\Phi}(t,\eta)$ has a fat tail for every $\eta > 0$ by Proposition 3.2 and (18), after being composed with a Gaussian function this tail is “tamed”. To simplify the heavy notation, when
there is no danger of confusion, we suppress $t$, $f + \Phi$ and $\xi$ and emphasize the “bandwidth” $\alpha$ and $\eta$ by denoting

\begin{equation}
Y_{\alpha,\xi,\eta} := Y_{f+\Phi}^{(h,\alpha,\xi)}(t,\eta), \quad \Omega_{\eta} := \Omega_{f+\Phi}^{(h)}(t,\eta) \quad \text{and} \quad V_{\eta} := V_{f+\Phi}^{(h)}(t,\eta).
\end{equation}

We state below that $Y_{\alpha,\xi,\eta}$ has finite moments of all orders, one for null and one for non-null case, relegating the proof to Section H.

**Theorem 5.3** (Absolute moments and moments of $Y_{\alpha,\xi,\eta}$, null case). Suppose Assumptions 4.1, 4.2 and 4.3 hold, $A = 0$ and $k > 0$. For any $\eta, \xi, \alpha > 0$, the $k$-th (absolute) moment of $Y_{\alpha,\xi,\eta}$ is finite. Moreover,

1. (absolute moments) For $\eta \geq 1$ and $\xi > 0$, when $\alpha$ is sufficiently small, we have $\mathbb{E}|Y_{\alpha,\xi,\eta}|^k \asymp \alpha^{-k/2+1}$, where the implied constant depends on $\varrho$, $k$ and $\frac{\eta^{k+8}}{(1+4\pi^2\eta - \xi^2)(k+4)/2}$; for $\eta < 1$ and $\xi > \sqrt{\eta}$, when $\alpha$ is sufficiently small, we have

\begin{equation}
\frac{c_1 \alpha^{-k/2+1} \leq \mathbb{E}|Y_{\alpha,\xi,\eta}|^k \leq c_2 \left( \frac{1}{\sqrt{\alpha \eta}} e^{-\frac{k \eta^2}{2\pi}} \vee 1 \right) \alpha^{-k/2+1},
\end{equation}

where $c_1 > 0$ depends on $\varrho$, $k$ and $\frac{\eta^{k+8}}{(1+4\pi^2\eta - \xi^2)(k+4)/2}$, and $c_2 > 0$ depends on $\varrho$ and $k$.

2. (moments) When $k$ is odd, for any $\eta, \xi > 0$, $\mathbb{E}Y_{\alpha,\xi,\eta}^k = 0$. When $k$ is even, for $\eta \geq 1$, when $\alpha$ is sufficiently small, we have $|\mathbb{E}Y_{\alpha,\xi,\eta}^k| = O(\alpha^{-k/2+1})$, where the implied constant depends on $\varrho$, $k$ and $\frac{\eta^{k/2} e^{-4\pi^2 \eta^2}}{(1+4\pi^2\eta - \xi^2)(k+4)/2}$; for $\eta < 1$, we have a simple bound $|\mathbb{E}Y_{\alpha,\xi,\eta}^k| \leq \mathbb{E}|Y_{\alpha,\xi,\eta}|^k$.

In the theorem, we split the (absolute) moment evaluate into two cases, one is for $\eta \geq 1$ and one is for $\eta < 1$, particularly when $\eta$ is close to 0. In fact, when $\eta$ is close to 0, $\sum_{(h)}^{(h)}$ is close to being degenerate, and controlling the (absolute) moments depends on controlling the degeneracy, which needs a different approach compared with that when $\eta \geq 1$. Moreover, the spectral property of color noise is reflected in the (absolute) moment bound, and it interacts with $\xi$, the frequency we want to detect by SST. A similar fact holds for the non-null case, where the signal plays another role in the analysis. Note that the results for the non-null case is not optimal, while it is sufficient for our purpose.

**Theorem 5.4** (Absolute moments and moments of $Y_{\alpha,\xi,\eta}$, non-null case). Suppose Assumptions 4.1, 4.2 and 4.3 hold, $A > 0$ and $k \in \mathbb{N}$. For any $\eta, \xi, \alpha > 0$, the $k$-th (absolute) moment of $Y_{\alpha,\xi,\eta}$ is finite. Moreover,
1. (absolute moment) for $\eta \geq 1$, $|\eta - \xi_0| \geq 1/2$ and $\xi > 0$, when $\alpha$ is sufficiently small, we have

$$
E|Y_{\alpha,\xi,\eta}|^k \asymp \alpha^{-k/2+1},
$$

where the implied constant depends on $\varrho$, $k$ and $\frac{\eta^{k+2}}{(1+4\pi^2|\eta-\xi_0|^2)^{(k+4)/2}}$; when $|\eta - \xi_0| < 1/2$ and $\xi > 0$, when $\alpha$ is sufficiently small, we have

$$
c_1\alpha^{-k/2+1} \leq E|Y_{\alpha,\xi,\eta}|^k \leq c_2\alpha^{-k/2+1},
$$

where $c_1$ depends on $\varrho$, $k$ and $e^{-A^2(1+4\pi^2\xi_0^2)^2\xi_0^{-\varrho}}$ and $c_2$ depends on $\varrho$, $k$ and $A^{k+3}\xi_0^{(2-\varrho)(k+3)/2}$; for $\eta < 1$ (particularly when $\eta$ is close to 0), when $\alpha$ is sufficiently small so that $\alpha < \eta$, we have

$$
c_1 e^{-CA^2\tilde{h}(\xi_0)^2(\eta-\xi_0^2)^2} \alpha^{-k/2+1} \leq E|Y_{\alpha,\xi,\eta}|^k \leq c_2 \max\left\{\frac{1}{\sqrt{\eta\tilde{h}(\xi_0)^2}}, 1, \frac{1}{A\tilde{h}(\xi_0)\xi_0}\right\}\alpha^{-k/2+1}
$$

where $c_1$ depends on $\varrho$, $k$ and $\frac{\eta^{2k+8}}{(1+4\pi^2|\eta-\xi|^2)^{(k+4)/2}}$, $C > 0$ depends on $\varrho$, and $c_2$ depends on $\varrho$ and $k$.

2. (moment) When $\eta \geq 1$ and $|\xi_0 - \eta| \geq 1/2$, when $\alpha$ is sufficiently small, we have $|EY_{\alpha,\xi,\eta}^k| = o\big(e^{-2\pi^2(\xi_0^2-\eta^2)}\big)$, where the implied constant depends on $|E|Y_{\alpha,\xi,\eta}|^k|$. When $\eta < 1$ or when $\eta \geq 1$ and $|\xi_0 - \eta| < 1/2$, when $\alpha$ is sufficiently small, we have the trivial bound $|EY_{\alpha,\xi,\eta}^k| \leq E|Y_{\alpha,\xi,\eta}|^k$.

Note that $Y_{\alpha,\xi,\eta}$ is the product of two dependent random variables, $V_\eta$ and $g_\alpha(|\xi - \Omega_\eta|)$. By Lemma 5.1, we know that the covariance of $V_\eta$ and $V_{\eta'}$ decays exponentially when $|\eta - \eta'| \to \infty$, and by Theorem 5.2, the same decay is true for the covariance of $\Omega_\eta$ and $\Omega_{\eta'}$. It is thus natural to expect that the covariance of $Y_{\alpha,\xi,\eta}$ and $Y_{\alpha,\xi,\eta'}$ also decays exponentially when $|\eta - \eta'| \to \infty$. Below, we show that despite the involved nonlinear transform, the same decay rate is also true for the covariance of $Y_{\alpha,\xi,\eta}$ and $Y_{\alpha,\xi,\eta'}$.

**Theorem 5.5.** Suppose Assumptions 4.1, 4.2 and 4.3 hold. Fix $\xi > 0$. For any $\alpha > 0$, $\text{Cov}(Y_{\alpha,\xi,\eta}, Y_{\alpha,\xi,\eta'})$ and $\text{Cov}(Y_{\alpha,\xi,\eta}, \overline{Y_{\alpha,\xi,\eta'}})$ are continuous over $\eta > 0$ and $\eta' > 0$. For $\eta, \eta' > 0$ satisfying $|\eta - \eta'| \geq 1$, when $\alpha$ is sufficiently small,

$$
|\text{Cov}(Y_{\alpha,\xi,\eta}, Y_{\alpha,\xi,\eta'})| \vee |\text{Cov}(Y_{\alpha,\xi,\eta}, \overline{Y_{\alpha,\xi,\eta'}})| = O\left((\eta - \eta')^2 e^{-\pi^2(\eta-\eta')^2}\right),
$$

where the implied constants depend on $\varphi$ and $\varrho$.  


5.4. Distribution of $S_{f+\Phi}^{(h,\alpha)}(t,\xi)$. With the above preparation, we may state the main result. Without loss of generality, we focus on $t = 0$. To study the distribution of $S_{f+\Phi}^{(h,\alpha)}(0,\xi)$ for a given $\alpha > 0$ and $\xi > 0$, by viewing $Y_{\alpha,\xi,\eta}$ as a random process indexed by $\eta$, a natural approach is to discretize $Y_{\alpha,\xi,\eta}$, approximate $S_{f+\Phi}^{(h,\alpha)}(0,\xi)$ by a Riemann sum, and apply the CLT. Below we consider the following discretization in $\eta$. For each $l = 1, \ldots, n$, denote $\eta_l := l \Delta \eta$, where $\Delta \eta = n^{-1/2-\beta}$, and $\beta \geq 0$ is to be determined in the proof. Also, denote $H = n \Delta \eta = n^{1/2-\beta}$. To further simplify the notation, when there is no danger of confusion, denote $V_{l} := V_{f+\Phi}(t,\eta_l), \Omega_l := \Omega_{\eta_l}, Y_{\alpha,\xi,l} := Y_{\alpha,\xi,\eta_l}$.

For $\xi > 0$, we approximate $S_{f+\Phi}^{(h,\alpha)}(0,\xi)$ by the Riemann sum:

$$S_{\alpha,\xi,n} := \Delta \eta \sum_{l=1}^{n} Y_{\alpha,\xi,l}.$$ 

The asymptotic distribution of $S_{\alpha,\xi,n}$ when $n \to \infty$ represents the distribution of $S_{f+\Phi}^{(h,\alpha)}(0,\xi)$, which is related to integrating over a wider spectral range with a finer frequency resolution. Note that the dependence structure of $Y_{\alpha,\xi,\eta}$ generates difficulty when we evaluate (27), despite its exponential decay indicated in Theorem 5.5. To handle this difficulty, the proof heavily depends on the $M$-dependent argument. Intuitively, the behavior of the $M$-dependent random process of $Y_{\alpha,\xi,\eta}$, denoted as $Y_{\alpha,\xi,\eta}$, will be essentially the same as that of $Y_{\alpha,\xi,\eta}$ for large $M$, and we expect $\text{Var}(Y_{\alpha,\xi,\eta}) \approx \text{Var}(Y_{\alpha,\xi,\eta}) \approx \text{Cov}(Y_{\alpha,\xi,\eta}, Y_{\alpha,\xi,\eta})$ for any $(t, \eta)$. The following lemma quantifies this intuition. The main challenge toward this seemingly simple conclusion is the nonlinearity inherited from SST, which boils down to studying the relationship between the covariances of $Y_{\alpha,\xi,\eta}$ and $Y_{\alpha,\xi,\eta}$. To the best of our knowledge, this kind of problem is less considered in the Gaussian approximation literature and there is no standard approach toward it. We provide a separate section elaborating this technique in Section I, which is the basis of the proof of Theorem 5.6 shown in Section J.

**Theorem 5.6.** Suppose Assumptions 4.1, 4.2 and 4.3 hold and assume $\varrho < 5$. Fix $\xi > 0$, take a small $\delta > 0$ and set $\beta = \frac{\delta}{4(2+\delta)}$ and $\Delta \eta = n^{-1/2-\beta}$. Also assume $\alpha = \alpha(n)$ so that $\alpha \to 0$ and $n\alpha \to 1$ when $n \to \infty$. We have $S_{\alpha,\xi,n} \Rightarrow S_{\alpha,\xi,n} \to \mathbb{C}N_{1}(0,\nu,\varrho)$ weakly when $n \to \infty$, where $\nu > 0$ and $\varrho \in \mathbb{C}$ are of order $(1+\xi)^{\varrho}$, where the implied constant depends on $\varrho$ when $A = 0$ and on $A$, $\xi_0$ and $\varrho$ when $A > 0$. 
Note that the condition \( n \alpha \to 1 \) could be understood as the “sufficient sampling” condition.

6. An application – oscillatory component detection via SST.
A critical question in practice is how to determine if a given time series contains an oscillatory component. This challenging problem has attracted lots of attention [32], but so far there is no universally accepted solution, particularly when handling modern biomedical signals. Below, we propose a detection algorithm based on SST to handle this challenge.

6.1. Discretization of SST. Before introducing the algorithm, we detail the discretization of SST. First, we follow the setup in [10] to discretize \( \Phi \).

**Assumption 6.1.** Take a symmetric Schwartz function \( \psi \) so that \( \hat{\psi}(\xi) = 1 \) when \( |\xi| \leq 1/4 \) and \( \hat{\psi}(\xi) = 0 \) when \( |\xi| > 1/2 \). Set \( X_j := \Phi(\bar{N}^{-1/2} \psi(\bar{N} \cdot -j)) \), where \( \bar{N}^{1/2} \psi(\bar{N} \cdot -j) \) is of unit \( L^1 \) norm centered at \( j/\bar{N} \). \( \bar{N} > 0 \) is the sampling frequency, \( j = -\lfloor N/2 \rfloor, -\lfloor N/2 \rfloor + 1, \ldots, \lfloor N/2 \rfloor - 1, \lfloor N/2 \rfloor \), and \( N \) is the number of sampling points. Below, we assume \( \bar{N} = N^{1/2} \).

Note that \( \psi \) is the *measurement function*, which models the properties of the measurement equipment, and \( N/\bar{N} > 0 \) is the recording length. Clearly, \( X := [X_{-\lfloor N/2 \rfloor}, \ldots, X_{\lfloor N/2 \rfloor}] \) is a discretization of the stationary GRP \( \Phi \) so that \( X \) is a stationary Gaussian time series with mean 0. For \( l \in \{-\lfloor N/2 \rfloor, -\lfloor N/2 \rfloor + 1, \ldots, \lfloor N/2 \rfloor - 1, \lfloor N/2 \rfloor \} \), \( X(l) \) is the discretization at time \( l/\bar{N} \) and \( X(l) \) has a Gaussian distribution with the standard deviation \( \sigma_N := (\int |\psi(\xi/\sqrt{N})|^2 p(\xi) d\xi)^{1/2} \), which increases when \( N \) increases. Specifically, when \( \rho \geq 0 \), the high frequency noise is not negligible. Thus, when \( N \) gets larger, the “measurement period” is shorter, high frequency noise dominates, and the measured value is more uncertain.

To discretize the analysis of \( \Phi \) by SST, we need the following discretization. Without loss of generality, since \( \Phi \) is stationary, in the following analysis we fix to \( t = 0 \). First, the STFT of \( \Phi \) in (1) is discretized by \( V_{\eta,N} := \frac{1}{\bar{N}} \sum_{j=-[N/2]}^{[N/2]} X_j h\left( \frac{j}{N} \right) e^{-2\pi i \frac{\eta j}{\bar{N}}} \), where \( \eta > 0 \). Note that \( V_{\eta,N} \) is a periodic function of \( \eta \) with period \( \bar{N} \), and \( V_{\eta,N} \) is the numerical implementation of \( V_{\Phi}^{(h)}(0, \eta) = \Phi(h_{0,\eta}) \). For the reassignment rule, the \( \Phi((h')_{0,\eta}) \) in (8) can be implemented in the same way with \( h \) replaced by \( h' \). By a direct expansion, we have

\[
V_{\eta,N} = \Phi \left( \frac{1}{\bar{N}} \sum_{j=-[N/2]}^{[N/2]} \bar{N} \psi(\bar{N} \cdot -j) h\left( \frac{j}{N} \right) e^{-2\pi i \frac{\eta j}{\bar{N}}} \right).
\]
Notice that \( \frac{1}{N} \sum_{j=-\lceil N/2 \rceil}^{\lceil N/2 \rceil} \bar{N} \psi(\bar{N} y - j) h\left( \frac{j}{N} \right) e^{-2\pi i \eta j} \) could be viewed as an approximation of \( h(y) e^{-i 2\pi \eta y} \). Thus, if we define \( h_{\eta,N}(\cdot) := \frac{1}{N} \sum_{j=-\lceil N/2 \rceil}^{\lceil N/2 \rceil} \bar{N} \psi(\bar{N} \cdot - j) h\left( \frac{j}{N} \right) e^{-2\pi i \eta j} \), which is a Schwartz function, \( V_{\eta,N} = \Phi(h_{\eta,N}) \). Similarly, we can define \( h'_{\eta,N}(\cdot) := \frac{1}{N} \sum_{j=-\lceil N/2 \rceil}^{\lceil N/2 \rceil} \bar{N} \psi(\bar{N} \cdot - j) h'\left( \frac{j}{N} \right) e^{-2\pi i \eta j} \), and have \( V'_{\eta,N} := \Phi(h'_{\eta,N}) \), which is the numerical implementation of \( \Phi((h')_{0,n}) \). The reassignment rule can be implemented following (8) by a direct division. For \( \xi > 0 \), the integrand of SST is discretized in the same way as (27), which is denoted by \( \{Y_{\alpha,\xi,l,N}\}_{l=1}^{n} \). Denote \( S_{\alpha,\xi,n,N} := \Delta \eta \sum_{l=1}^{n} Y_{\alpha,\xi,l,N} \). The following corollary states the intuition that when the sampling rate \( \bar{N} \) is high, the discretization of SST approaches the continuous version of SST. The proof is postponed to Section K.

**Corollary 6.1.** Assume Assumptions 4.1, 4.2, 4.3 and 6.1 hold. Adapt notations used in Theorem 5.6. For \( \xi > 0 \) and a sufficiently large \( n \), we have \( S_{\alpha,\xi,n,N} \to S_{\alpha,\xi,n} \) in probability when \( N \to \infty \).

6.2. **Fisher-SST statistic.** Assume Assumptions 4.1, 4.2 and 4.3 hold. To test if an oscillatory component exists at a given time point, we propose to use the maximal magnitude of SST coefficients at the associated properly chosen frequency grid points as a test statistic under the null hypothesis \( H_{0} : f(t) = 0 \) against the alternative \( H_{1} : f(t) \neq 0 \). Consider the following procedure. Take a recorded time series \( x = [x_{-\lfloor N/2 \rfloor}, \ldots, x_{\lfloor N/2 \rfloor}] \in \mathbb{R}^{N} \) following the discretization scheme in Assumption 6.1; that is, the sampling period is \( 1/\sqrt{N} \). Note that \( x = X \) under \( H_{0} \). To simplify the notation, we fix to time 0. For a given \( N \), we could sample the frequency axis up to \( \sqrt{N}/2 \) Hz according to the Nyquist-Shannon theorem. Take \( n \) to be the discretization of the SST integration in (27) and suppose \( N \) is sufficiently larger than \( n \). Then choose a uniform grid \( G := (\xi_{1}, \xi_{2}, \ldots, \xi_{r}) \subset (0, \sqrt{N}/2) \) on the frequency axis, where \( r \in \mathbb{N} \) is chosen to be \( \lceil \sqrt{N}/\log(N) \rceil \) so that \( \xi_{j} = j \sqrt{N}/\lceil \sqrt{N}/\log(N) \rceil \) for \( j = 1, \ldots, r \). The **Fisher-SST statistic** is defined as

\[
m_{n,N} := \max_{\xi_{j} \in G} |S_{\alpha,\xi,j,n,N}|.
\]

For a preassigned \( a \in [0,1] \), let \( T_{a} \) be the \((100 \times a)\)-th percentile of \( m_{n,N} \) under the null. We reject the null hypothesis if the Fisher-SST statistic of \( x \) exceeds \( T_{a} \); that is, we detect a sufficiently strong oscillatory component compared with the noise around time 0.

To determine when an oscillatory component exists over a period, we repeat the above steps and evaluate Fisher-SST statistics over a set of chosen
timestamps. Note that SST on the TF domain has a dependent structure. When two consecutive time stamps are sufficiently separated, the SST coefficients would be approximately independent. Indeed, for the chosen kernel in Assumption 4.3, \( h(c) \) is numerically zero for a sufficiently large constant \( c > 0 \), say, \( c = 10 \). Therefore, numerically the SST coefficients are independent if two timestamps are separated by \( 2c \). The proof comes from the Plancheral theorem and the fact that the noise is Gaussian, and we omit details. Thus, we propose to choose a uniform grid \((t_1, t_2, \ldots, t_{r'}/N) \subset [-\sqrt{N}/2, \sqrt{N}/2] \) on the time axis, where \( r' \in \mathbb{N} \) is chosen so that \( t_i - t_{i-1} = \log(N) \). Then, for each timestamp \( t_i \), evaluate the Fisher-SST statistic and obtain the p-value \( p_i \). Since we will run this test for \( r' = \lfloor \sqrt{N}/\log(N) \rfloor \) times and these tests are numerically independent, we recommend to consider the false discovery rate control [4] to handle the multiple testing issue in the following way. Rank obtained p-values \( p_1, \ldots, p_{r'} \) as \( p^{(1)} \leq p^{(2)} \leq \ldots \leq p^{(r')} \). Take \( q > 0 \) to be the false discovery rate and set \( k^* \) to be the largest integer so that \( p^{(k^*)} \leq k^*q/r' \). At time \( t_i \), the null hypothesis is rejected if \( p_i \leq p^{(k^*)} \). See [4] for details. The result shows when an oscillatory component exists over a period.

6.3. Bootstrapping. Although we know that the discretization effect on \( S_{\alpha, \xi, n, N} \) disappears asymptotically by Corollary 6.1, the distribution of \( \lim_{N \to \infty} m_{n, N} \) is not known. We thus propose a bootstrapping algorithm to approximate the Fisher-SST statistic. Decompose \( x \) into the possibly existing oscillatory component \( y \) by the reconstruction formula provided in [13, 10] and the noise part \( n \). Based on the stationary assumption in Assumption 4.1, we could estimate the covariance structure of the noise by applying the banding covariance approximation approach [40, Section 3]. Denote the estimated covariance as \( \hat{\Sigma} \), and generate \( m \in \mathbb{N}, \text{say}, 10,000 \), pseudo-observed Gaussian noises \( n^{*{(l)}} \in \mathbb{R}^N \), where \( l = 1, \ldots, m \), with mean 0 and the covariance structure \( \hat{\Sigma} \). For each \( l \) and a timestamp, apply SST to \( n^{*{(l)}} \) and all grid points in \( G \), denoted as \( \{S_{\alpha, \xi, n, n, N}^{*{l}} | \xi \in G \} \). Define

\[
m_{n, N}^{*{(l)}} := \max_{\xi \in G} |S_{\alpha, \xi, n, n, N}^{*{l}}|.
\]

The distribution of \( m_{n, N} \) can be approximated by the empirical distribution of \( \{m_{n, N}^{*{(l)}} \}_{l=1}^m \). The justification of the proposed bootstrapping algorithm is given in the following theorem, whose proof is postponed to Section K.

**Theorem 6.2.** Assume Assumptions 4.1, 4.2, 4.3 and 6.1 hold and \( G \subset (0, \sqrt{N}/2) \) is a uniform grid with \(|G| = \lfloor \sqrt{N}/\log(N) \rfloor \). There exists a
probability space associated with the bootstrapping, called \((\Omega, \mathcal{F}, \mathbb{P})\), where we could construct a sequence of i.i.d. random variables \(\{m^*_n\}_{n=1}^m\), that follows the same distribution as that of \(m^*_{n,N}\) and \(m^\#_{n,N}\) which has the same distribution as that of \(m_{n,N}\). Then we have \(m^*_{n,N} - m^\#_{n,N} \to 0\) in probability when \(N \to \infty\).

**7. Numerical Results.** The Matlab code of SST is available in [http://hautiengwu.wordpress.com/](http://hautiengwu.wordpress.com/). More numerical results of the developed theorems could be found in Section L. We demonstrate the proposed oscillatory signal detection algorithm and show the rejection rate with a realistic simulated signal. We consider the smoothed Brownian path realizations to model an oscillation with a slowly varying amplitude and frequency [14]. Suppose \(W\) is the standard Brownian motion defined on \([0, \infty)\). A smoothed Brownian motion with the bandwidth \(B > 0\) is defined as \(W_B = W \ast K_B\), where \(K_B\) is the Gaussian function with the bandwidth \(B > 0\) and \(\ast\) denotes the convolution operator. Given \(T > 0\) and parameters \(\zeta_1, \ldots, \zeta_6 > 0\), we then define a family of random processes on \([0, T]\) by \(\Psi_{[\zeta_1, \ldots, \zeta_6]}(x) := \zeta_1 + \zeta_2 x + \zeta_3 \frac{W_{\zeta_4}(x)}{\|W_{\zeta_4}\|_{L^\infty[0,T]}} + \zeta_5 \int_0^x \frac{W_{\zeta_6}(s)}{\|W_{\zeta_6}\|_{L^\infty[0,T]}} ds\). Consider \(f(t) = A a(t) \cos(2\pi\phi(t)) + \Phi\) over \([0, T]\), where \(A \geq 0\) controls the strength of the oscillation, \(a(t)\) is a realization of \(\Psi_{[1,0,0,5,5,0,0]}(t)\), \(\phi(t)\) is a realization of \(\Psi_{[0,8,0,0,2,4]}(t) + t^{1.1}/15\), and \(\Phi\) follows an autoregressive and moving average (ARMA) process with a proper normalization so that the standard deviation is 1 at each timestamp, where the ARMA process is determined by the auto-regression polynomial \(a(z) = 0.5z + 1\) and the moving averaging polynomial \(b(z) = -0.5z + 1\), with the innovation process taken as independent and identically distributed Gaussian random variables. Set \(n = 2,048\) and \(\beta = 0.05\), and realize \(f\) with the sampling rate 64Hz and sample 4,096 points from \(f\) over \([0, 64]\) s. Take grids of the frequency and time axes as the above and and construct Fisher-SST at 32s. Note that the amplitude and frequency of the oscillation \(a(t) \cos(2\pi\phi(t))\) both change slowly, and hence locally the signal oscillates like a harmonic function. For a comparison, we also consider STFT and the chirplet path pursuit algorithm [8]. Based on Theorem 5.1, we consider STFT and define a similar statistic, called the Fisher-STFT statistic, by \(\max_{\zeta \in G} |V_{\zeta,N}|\). The bootstrapping to estimate the null distribution is carried out in the same way shown in Section 6.3, except the step of decomposing the noise out of the noisy signal. Specifically, we follow the common practice to subtract the oscillation associated with the maximal peak determined by the periodogram from the noisy signal, and view the remaining component as the noise. For Fisher-SST and
Fisher-STFT, we repeat the bootstrapping for 10,000 times to determine the threshold, where we take 0.05 as our significance level. For the chirplet path pursuit, we use the ChirpLab v1.1 package provided by the authors of [8], where we use the best path statistic with the cubic polynomial to fit the amplitude, run Monte Carlo simulation for 10,000 times, and take 0.05 as the significance level. We realize $f$ for 1,000 times with $A = 0.12(k - 1)$, where $k = 1, \ldots, 10$, and plot the simulated rejection rate in Figure 2. We see that the Fisher-SST has a higher rejection rate compared with the Fisher-STFT and chirplet path pursuit, and the chirplet path pursuit performs better than the Fisher-STFT when the signal is strong. It is expected that chirplet path pursuit performs better than Fisher-STFT since a multiscale scheme and the chirp information is captured in the chirplet path pursuit algorithm. We shall mention that chirplet coefficients used in the chirplet path pursuit, which comes from the inner product of the signal and $\exp(i2\pi(at^2/2 + bt))$ for a range of $a$ and $b$, decay at the rate $(a - a_0)^{-1/2}$ when $a_0$ is the true chirp, $b$ is fixed and $|a| \to \infty$. This slow decay could be understood as an uncertainty principle [11]. So, while chirplet could help capture an oscillatory component with a chirp, its performance might be impacted.

Finally, we come back to the PPG signal shown in Figure 1. A critical biomedical signal processing step, particularly for long-term monitoring in clinics, is determining when the signal quality is trustworthy so that the obtained information is usable for decision making. This step is usually referred to signal quality assessment. Signal quality assessment is in general challenging, and the strategy depends on the clinical problem. When the heart rate and its variability are the concern, we care if a PPG signal oscillates properly and reflects how the heart beats, so that the TFR of a high-quality PPG encodes the time-varying heart rate as a curve with distinguishable intensity. See Figure 1 for an example, where the PPG signal in the first 50 second is labeled as high quality and the remaining signal is labeled as low quality. It is visually obvious to see an oscillation from the 0th to 40th second, which are cardiac cycles. However, after the 50th second, the signal looks chaotic and it is not clear if it provides any useful cardiac information. The signal between the 40th and 50th second is also oscillatory, but the pattern is slightly distorted compared with that before the 40th second.
This visual inspection suggests that the signal quality over each segment of predetermined length could be quantified by the confidence of oscillation detection. Note that the frequency and amplitude might change slowly, so it is reasonable to assume that locally the signal oscillates with fixed amplitude and frequency, and the proposed bootstrapping algorithm and the Fisher-SST statistic could be applied. In this signal, the sampling rate is 100Hz and the signal length is 100 sec. The window satisfies Assumption 4.3, $\Delta \eta = 1/1201$, and the significance level is 0.05. We evaluate the Fisher-SST statistic every 1 second by setting the grid size to be 0.5 in $G$. By setting the desired false discovery rate to be 0.05, the rejection of the null hypothesis, marked as blue diamonds, coincides with the visually identifiable cardiac oscillations. Thus, the first 50-seconds segment are of high quality with some type I errors, which coincides with the expert’s annotation. The potential of designing a signal quality assessment algorithm based on the proposed algorithm will be further explored in our future work.

8. Discussion and Conclusion. We provide a theoretical support for the nonlinear-type TF analysis algorithm, SST, that forms a foundation for future statistical inference studies. In particular, we extend the existing quotient distribution of proper complex normal random variables to the improper case, and quantify the asymptotic distribution of SST at a given frequency entry. While there are a multitude of available nonlinear-TF analysis algorithms, to the best of our knowledge, this is the first work providing an extensive quantification of the asymptotic distribution. This result sets the stage for further analysis of SST and other nonlinear-type TF analysis algorithms. In particular, we provide several analytic tools to handle the main challenges when studying a nonlinear-type TF analysis algorithm. Specifically, in order to handle the nonlinearity involved in SST, a careful change of variables, an approximation scheme for the confluent hypergeometric function, and a construction of the associated $M$-dependent random process are given. Observe that one major challenge in nonlinear frequency domain analysis lies in the lack of systematic dependence measures, such as strong mixing conditions [33] and physical dependence measures [39]. In this article, we adopted a highly nontrivial $M$-dependent approximation scheme in the frequency domain and successfully combined it with the nonlinear kernel regression technique in time series analysis to derive the asymptotic distribution of the STFT-based SST, and construct a local bootstrap inference procedure with theoretical supports.

8.1. Relationship with kernel regression. The approximation (27) is related to the kernel regression perspective (See for instance Chapter 6 of
of time series analysis. Suppose we were able to model \( \{(\Omega_i, V_i)\}_{i=1}^n \) as a dataset sampled from a random vector \([X Y]^\top \in \mathbb{C}^2 \) so that \( Y \) and \( X \) were related by \( Y = F(X) + \mathcal{N} \), where \( \mathcal{N} \) is random noise satisfying \( \mathbb{E}[\mathcal{N}|X = \xi] = 0 \), and the response \( V_i \) and predictor \( \Omega_i \) are related by the “regression function” \( F \). If we further imagine \( f_{\Omega}(\xi) \) to model the “density of \( \Omega_i \) at \( \xi \)”, then the kernel regression \( S_{\xi,n}/f_{\Omega}(\xi) \) estimates that regression function at a fixed \( \xi > 0 \); that is, it gives the conditional expectation of \( Y \) given \( X = \xi \) so that \( F(\xi) = \mathbb{E}[Y|X = \xi] \). In our case, this model is not correct, but still we have \( S_{\xi,n}/f_{\Omega}(\xi) \to \mathbb{E}[Y|X = \xi] \) as \( n \to \infty \). Adapting this kernel regression perspective, intuitively if we view \( V_i \) as a “noisy” version of some regression function over \( \Omega_i \), with the “clean” regression function providing the “best” TF representation, then the kernel regression helps recover this representation. When the signal is only noise, we expect \( F \) to be zero and \( S_{\xi,n}/f_{\Omega}(\xi) \) to converge to 0. However, while this intuition helps us better understand how SST works, the structure of the regression function is not easy to directly identify in the non-null case.

8.2. Continuous wavelet transform based SST. The same analysis can be mimicked in the continuous wavelet transform (CWT) setup, but a significant simplification occurs regarding the pseudocovariance. In particular, let \( \psi \in \mathcal{S} \) and for \( a, b > 0 \), define \( \Psi(a, b)(t) = a^{-1/2} \psi((t-b)/a) \). Then the CWT of a tempered distribution \( f \) takes the form \( C_{f}^{(\psi)}(a, b) = f(\Psi(a, b)) \). Analogously to (4) and (5), we have \( C_{f}^{(\psi)}(a, b) = f(\Psi(a, b)) + \Phi(\Psi(a, b)), \partial_t V_{f+h}(t, \eta) = \partial_t f(\Psi(a, b)) + \Phi(\Psi'(a, b)). \) To simplify the discussion, suppose \( \Phi \) is white, and the covariance becomes \( \Gamma = \frac{1}{a^2} \left[ a^2 \int |\psi(\xi)|^2 d\xi \quad -2\piia \int \xi |\hat{\psi}(a\xi)|^2 d\xi \right] \). For the pseudocovariance, if we further assume that \( \hat{\psi} \) is analytic with \( \psi \) real and \( \text{supp}(\hat{\psi}) \subset (0, \infty) \), then the pseudocovariance matrix is manifestly zero. Hence, \( \Gamma \) and \( C \) trivially commute and are thus simultaneously diagonalizable, so there is a basis of \( \mathbb{C}^2 \) where the components of \([C_{\Phi}^{(\psi)} \partial_t C_{\Phi}^{(\psi)}]^\top \) are uncorrelated and have a zero pseudocovariance. The reassignment rule is thus made as a quotient of independent random variables, and the relevant nonlinear transform depending on the complex gaussian quotients simplify significantly when \( C = 0 \). Note that the main technical challenge in analyzing STFT-based SST is handling pseudocovariance and this challenge is not encountered in the CWT-based SST. As a result, the proof is similar to that of the STFT-based SST shown in this paper, and we omit the details.

8.3. Future work. We remark that while the bounds in the proof are sufficient for our purpose, they might not be optimal, particularly when \( \eta \to 0 \).
We need a different approach to handle the degeneracy of the covariance structure for a finer analysis. In addition to SST, there are many other nonlinear-type TF analysis algorithms, for example, reassignment [2], concentration of frequency and time [14], synchrosqueezed wave packet transform [42], synchrosqueezing S-transform [22], second-order SST [28], and bilinear TF analysis tools like Cohen and Affine classes [16]. The current work sheds light on constructing a systematic approach to the study of statistical properties of those algorithms. As mentioned in Introduction, SST and these TF analysis tools have been widely applied in the signal processing society. Inspired by these applications, there are many important statistical inference problems remain open. For example, how to generalize the proposed oscillatory component detection algorithm to the case when multiple oscillatory components exist? How to establish the inference procedure for estimating instantaneous frequency, amplitude modulation and phase function and decomposing the noisy signal into its constitutional components? How to handle the nonstationary and/or non-Gaussian noise (e.g. the piecewise locally stationary [45]) or study the statistical structure of a random process? How to detect the number of oscillatory components present inside a noisy signal or at which times such components exist? To answer these questions, we need to fully understand the distribution of SST on the TF domain (e.g. for various $\xi$ and $t$) so that an inference can be carried out on the TFR level. Note that the oscillatory signal detection algorithm proposed in Section 6 is a special inference example. More discussion can be found in Section A of the supplement.

9. Acknowledgements. The authors acknowledge Professors Almut Burchard and Mary Pugh for the fruitful discussion. They thank the authors of [8] for providing the ChirpLab 1.1 code. They also thank the associate editor and the anonymous reviewers for their valuable and constructive feedbacks and comments.

REFERENCES


