**Statistics and Computing manuscript No.**
(will be inserted by the editor)

# Long memory estimation for complex-valued time series

**Marina I. Knight · Matthew A. Nunes**

Received: date / Accepted: date

**Abstract** Long memory has been observed for time series across a multitude of fields and the accurate estimation of such dependence, e.g. via the Hurst exponent, is crucial for the modelling and prediction of many dynamic systems of interest. Many physical processes (such as wind data), are more naturally expressed as a complex-valued time series to represent magnitude and phase information (wind speed and direction). With data collection ubiquitously unreliable, irregular sampling or missingness is also commonplace and can cause bias in a range of analysis tasks, including Hurst estimation. This article proposes a new Hurst exponent estimation technique for complex-valued persistent data sampled with potential irregularity. Our approach is justified through establishing attractive theoretical properties of a new complex-valued wavelet lifting transform, also introduced in this paper. We demonstrate the accuracy of the proposed estimation method through simulations across a range of sampling scenarios and complex- and real-valued persistent processes. For wind data, our method highlights that inclusion of the intrinsic correlations between the real and imaginary data, inherent in our complex-valued approach, can produce different persistence estimates than when using real-valued analysis. Such analysis could then support alternative modelling or policy decisions compared with conclusions based on real-valued estimation.

M. I. Knight
Department of Mathematics, University of York, Heslington, York, YO10 5DD, UK
E-mail: Marina.Knight@york.ac.uk

M. A. Nunes
Department of Mathematics and Statistics, Fylde College, Lancaster University, Lancaster, LA1 4YF, UK
E-mail: m.nunes@lancaster.ac.uk

**Keywords** complex-valued time series · Hurst exponent · irregular sampling · long-range dependence · wavelets

## 1 Introduction

Complex-valued time series arise in many scientific fields of interest, for example digital communication and signal processing (Curtis, 1985; Martin, 2004), environmental series (Gonella, 1972; Lilly and Gascard, 2006; Adali et al, 2011) and physiology (Rowe, 2005). Modelling and analysis of such series in the complex domain is not only natural, but also convenient. In addition, complex-valued time series models are often able to represent more realistic behaviour in observed physical processes, see e.g. Mandic and Goh (2009); Sykulski et al (2017). A particular modelling aspect which has received recent attention is the property of *impropriety* or *noncircularity*, describing series whose statistics are not rotationally invariant in the complex plane (for a precise definition, the reader is directed to Sykulski and Percival (2016)). Such models of improper processes have seen growing interest in the statistics community, see e.g. Schreier and Scharf (2003); Rubin-Delanchy and Walden (2008); Mohammadi and Plataniotis (2015). Furthermore, complex-valued analysis of *real-valued* data has been shown to be beneficial in a number of settings, see e.g. Olhede and Walden (2005); Hamilton et al (2017). For a comprehensive introduction to complex-valued signals, we refer the reader to Schreier and Scharf (2010); see Adali et al (2011) and Walden (2013) for recent advances in modelling complex-valued signals.

Recently, there has been an increased interest in models for complex-valued stochastic processes exhibiting long-range dependence (i.e. persistent) behaviour,

which has seen extensions of real-valued process modelling frameworks for the complex-valued fractional Brownian motion (fBM) and Matérn processes, see respectively Coeurjolly and Porcu (2017b) and Lilly et al (2017), as well as for (improper) fractional Gaussian noise (Sykulski and Percival, 2016). For these constructions, just as for real-valued processes (Hurst, 1951; Mandelbrot and Van Ness, 1968), the degree of memory can still be quantified by means of a single parameter, the Hurst exponent parameter (Amblard et al, 2012; Sykulski and Percival, 2016). Accurate estimation of the Hurst parameter offers valuable insight into a multitude of modelling and analysis tasks, such as model calibration and prediction (Beran et al, 2013; Rehman and Siddiqi, 2009; Knight et al, 2017).

Complex-valued processes, both proper (circular) and improper (noncircular), are relevant across fields such as oceanography and geophysics (Adali et al, 2011; Sykulski et al, 2017), where data are typically difficult to acquire and will frequently suffer from omissions/ missingness or be irregularly sampled (see e.g. Figure 1). In the next section, we describe datasets arising in environmental science that feature missing observations, which can be examined for long memory with a complex-valued representation. However, we note here that data from other scientific areas may benefit from analysis with our proposed methodology, see Section 6 for further discussion.

### 1.1 Persistence in wind series

Our motivating data example in this article arises from climatology. More specifically, wind series have been analysed extensively in the literature for modelling local weather patterns and spread of pollutants, as well as global climate dynamics. Long memory in wind series has been established by a number of authors, see e.g. Haslett and Raftery (1989); Chang et al (2012); Piacquadio and de la Barra (2014) and references therein. Specifically, Hurst exponent estimates for wind speed series on a range of sampling resolutions, including the five minute scale considered here, have been shown to be in the range $0.7 - 0.9$, indicating strong long-range dependence, see e.g. Fortuna et al (2014). Accurate Hurst exponent estimation is used for accurate forecasting of wind speed, for example to assess future power yields (Haslett and Raftery, 1989; Bakker and van den Hurk, 2012).

Wind speed analysis in the literature is predominantly performed using real-valued data, such as (magnitude) wind speed series. However, more recently a number of authors have advocated modelling wind measurements as *complex-valued*, developing analysis tools which exploit both speed *and* directional information of wind time series, see e.g. Goh et al (2006); Tanaka and Mandic (2007). These complex-valued modelling approaches have resulted in methodology for improved prediction for series such as those considered in this article (Mandic et al, 2009; Dowell et al, 2014). To our knowledge, *long memory estimation* for stationary time series is exclusively performed using real-valued time series. In this article, we analyse the degree of persistence (long memory intensity) exhibited by complex-valued wind measurements, i.e. series which have both wind speed and direction, using new *complex-valued* Hurst estimation methodology we propose here.

The wind series we consider in this article consists of two datasets measured at a five minute resolution from the Iowa Department of Transport's Automated Weather Observing System (AWOS). The (speed and angular) measurements for both datasets are available at `http://mesonet.agron.iastate.edu/AWOS/`. We firstly analyse data obtained from the Atlantic Municipal Airport (AIO) monitoring site over a period from 15th April 2017 until 30th April 2017. Whilst the sampling interval for the measurements is reported as five minutes, due to a number of reasons, for example faulty recording devices, the data in fact feature missingness which results in a mix of sampling intervals – our first dataset has intervals ranging from 5 to 15 minutes.

Since we have both speed and directional information for the dataset, we shall view the series using a complex-valued representation. The real and imaginary components of the series are shown in Figure 1(a) and Figure 1(b), together with the locations of the missing data (depicted by triangles). The length of the first series is $n = 3131$ with an overall rate of missingness of 12%. Similar datasets from the Iowa monitoring system have been previously studied in the literature for the non-missing case but not in the context of Hurst estimation, see e.g. Tanaka and Mandic (2007); Adali et al (2011).

To explore the potential persistence in wind series, we examine the autocorrelation in the real and imaginary parts of the series, shown in Figure 2(a) and Figure 2(b) for the `Wind A` series. For these data, both components show highly significant autocorrelation over a range of lags, indicating long memory.

To further illustrate potential benefits of a more considered analysis approach for such data, we also investigate a dataset from the same monitoring site but for a different time period, specifically, 30th April 2017 until 14th May 2017. For this dataset, the majority of the data are observed at a spacing of 5 minutes, but a significant amount have intra-measurement sampling

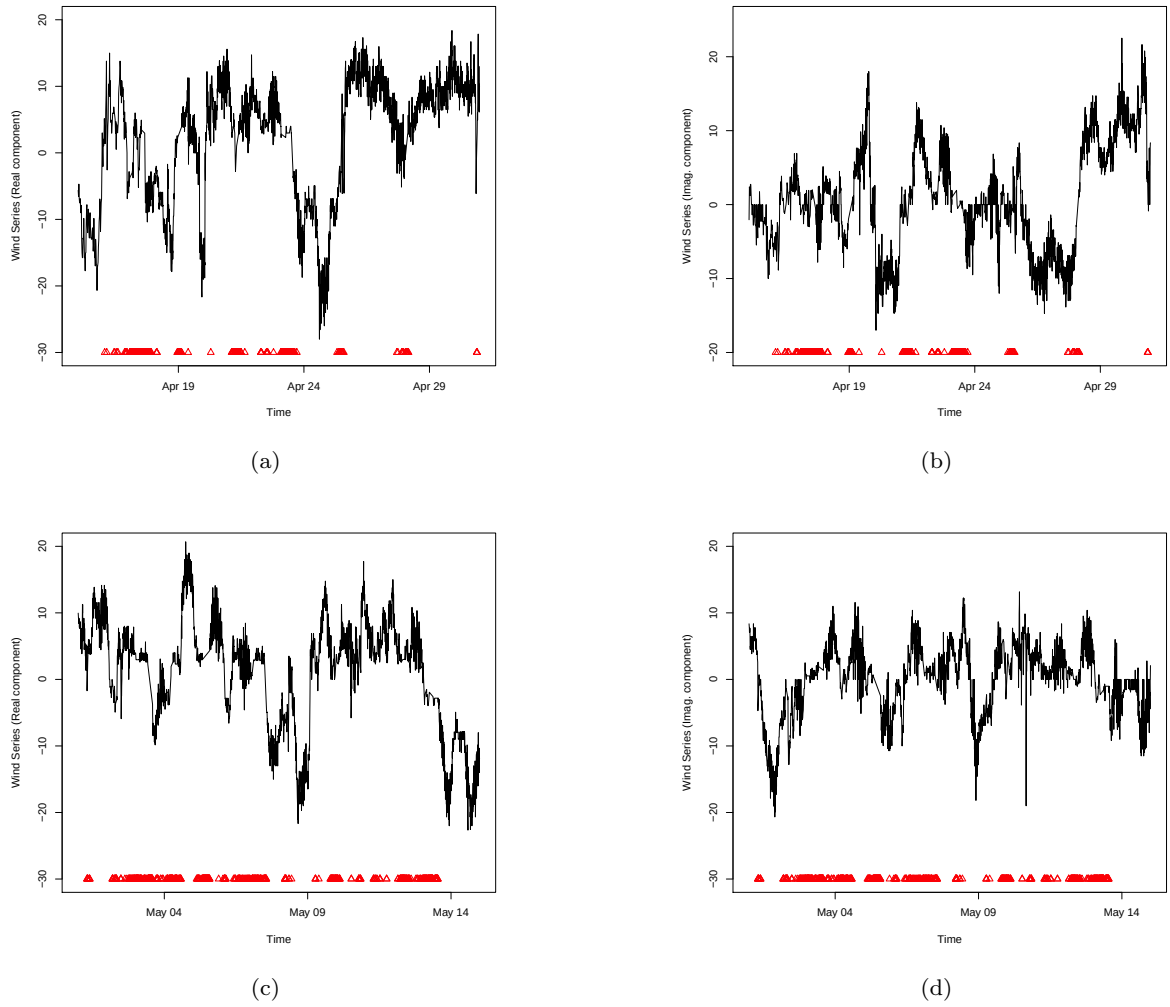



Fig. 1: (a) Real component of the `Wind A` data series; (b) Imaginary component of the `Wind A` data series; (c) Real component of the `Wind B` data series; (d) Imaginary component of the `Wind B` data series. Red triangles indicate missing data locations.

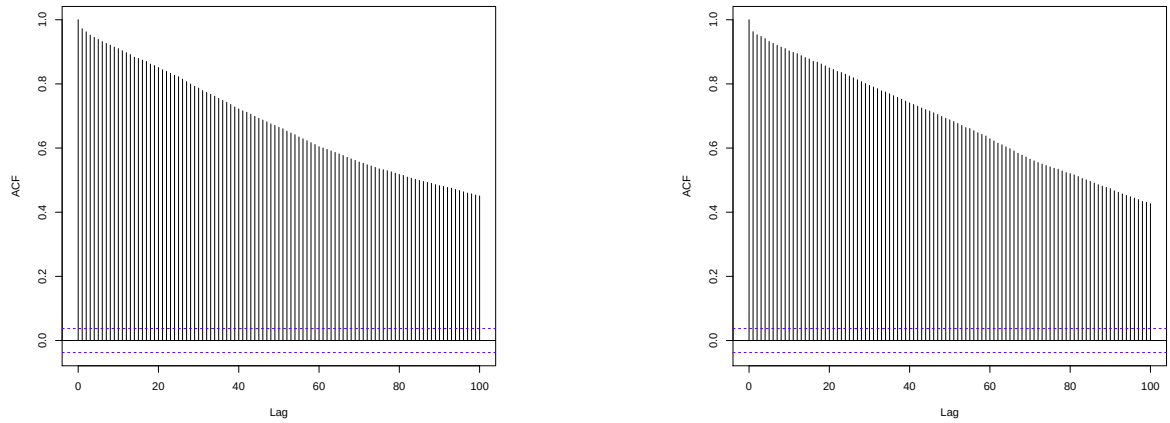
between 10 and 20 minutes resulting from a missingness proportion of 20%; the series is of length $n = 2942$. We have specifically chosen to examine this second time period due to its high degree of missingness. The two components of the complex-valued series can be seen in Figure 1(c) and Figure 1(d) (triangles indicate missing series values).

Similar observations about potential long memory characteristics can be made for the second complex-valued wind series. In particular, both real and imaginary components of the series show considerable autocorrelation over a large range of lags (Figure 2(c) and Figure 2(d)).

In addition, plotting the series in the complex plane, we see that both datasets exhibit a rotational behaviour, due to the angular component of the series (Figure 3). The series are not symmetric, exhibiting clear noncircularity, suggesting a model which allows for impropriety is appropriate for analysis (for an in-depth discussion of these properties the reader is directed to e.g. Sykulski and Percival (2016)). This reflects similar observations on impropriety shown for other Iowa AWOS data in Adali et al (2011), as well as other wind series (Mandic and Goh, 2009).

### 1.2 Aim and structure of the paper

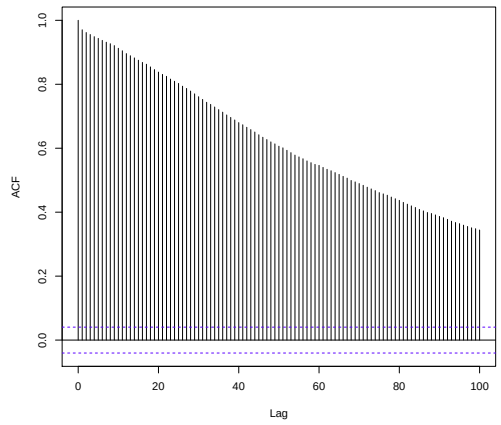
A feature of many geophysical series, such as described in Section 1.1, is that there is a need to jointly analyse both components of a bivariate signal in order to reveal a common behaviour. Due to the natural representation in the complex plane, one mathematical so-



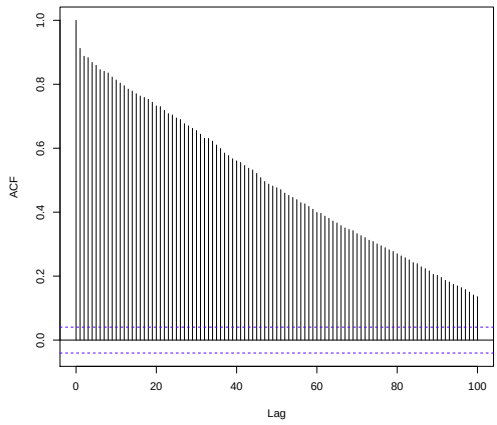

(a)




(b)




(c)



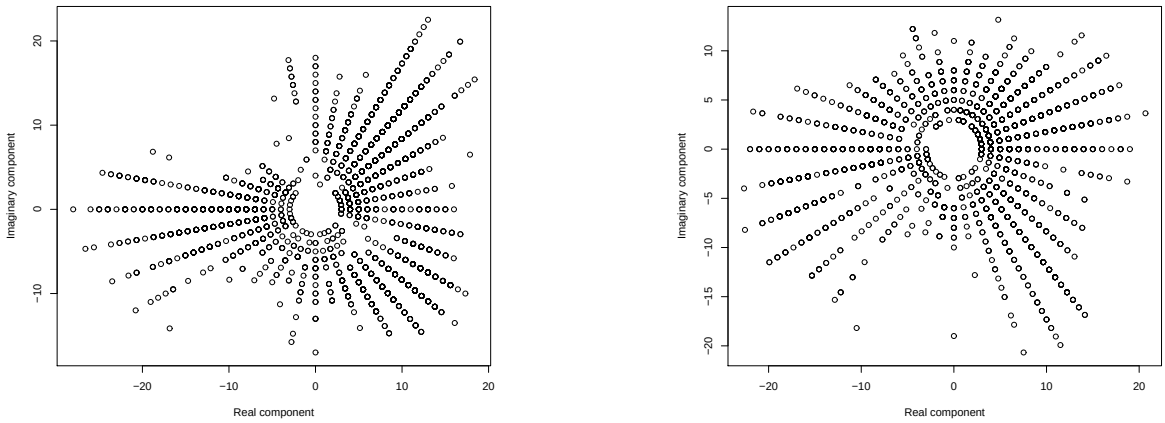

(d)

Fig. 2: (a) Autocorrelation for (a) the real component of the `Wind A` series from Figure 1; (b) the imaginary component of the `Wind A` series; (c) the real component of the `Wind B` series from Figure 1; (d) the imaginary component of the `Wind B` series (all treated as regularly spaced). Both components of the two datasets show autocorrelation at large lags, indicating persistent behaviour.

lution is to combine the two pieces of information into a single, *complex-valued* series and analyse its properties (Mandic and Goh, 2009). Adopting this approach thus calls for analysis techniques capable of dealing with complex-valued data. Additionally, for many applications the process sampling structure is inherently irregular, as the two components may be measured at irregular times, or the data may be blighted by missingness due to measurement device failures. In the real-valued case, the common practice of preprocessing the data to mitigate against irregular or missing observations, results in inaccuracies in long memory estimation by traditional methods. More specifically, there is now well-documented evidence that preprocessing by imputation or interpolation, as well as data aggregation leads to overestimation of persistence, see for example, Beran et al (2013), Zhang et al (2014) or Knight et al (2017).

In practice, to the authors' best knowledge, the only technique that permits Hurst exponent estimation for complex-valued processes is that of Coeurjolly and Porcu (2017b) which tackles the setting of regularly sampled (proper) complex-valued fractional Brownian motion. Motivated by the serious implications of inaccurate estimation in the real-valued setting, in this work we propose the first methodological approach that answers the timely challenge of accurate assessment of long memory persistence for complex-valued processes featuring regular *or* irregular sampling (including missingness).

At the heart of our methodology is a second generation wavelet-based approach. The reasoning behind this




(a) (b)

Fig. 3: Scatter plot of real and imaginary series values for (a) the `Wind A` data and (b) the `Wind B` series shown in Figure 1. Both series exhibit noncircular (improper) characteristics.

choice is two-fold: (i) (classical) wavelets have proved to be very successful in the context of regularly sampled (real-valued) time series with long memory and are considered the 'right domain' of analysis (Flandrin, 1998); and (ii) for irregularly sampled (real-valued) processes, or those featuring missingness, the wavelet lifting algorithm of Knight et al (2017) has provided a first long memory estimation solution and was shown to yield competitive results even for regularly sampled data.

The main contributions of the work in this paper are as follows. We propose (1) a novel lifting algorithm designed to work on complex-valued data with a potentially irregular sampling structure and (2) a Hurst parameter estimator for complex-valued processes sampled with a regular or irregular structure. Our method will be shown to improve on real-valued Hurst estimation results, including for regularly spaced data.

The remainder of this article is organised as follows. We begin, in Section 2, by reviewing (complex-valued) long memory processes and giving an overview of wavelet lifting transforms. Section 3 introduces our novel complex-valued lifting transform, establishes its iterative bases construction and theoretical results on its decorrelation properties. Section 4 demonstrates how these properties can be exploited to design our proposed lifting-based Hurst exponent estimation procedure for complex-valued data sampled with irregularity/ missingness. Section 5.1 contains a simulation study evaluating the performance of our new method using synthetic data. In Section 5.2, we consider the application of our approach to the wind series datasets introduced in Section 1.1, discussing the potential consequences of our analysis. Finally, Section 6 outlines some avenues of future work and discusses other potential applications.

## 2 Review of complex-valued processes, long-range dependence and wavelet lifting

### 2.1 Complex-valued processes

Let us denote a (complex-valued) second-order stationary time series by $\{X_t\}$ and its autocovariance function as $\gamma_X(t_i - t_j) = \mathbb{E}(X_{t_i}\overline{X}_{t_j})$, under the assumption that $\mathbb{E}(X_t) = 0$ and denoting by $\bar{\cdot}$ complex conjugation. As the autocovariance function $\gamma_X$ does not completely characterise a complex-valued time series, we also make use of its complementary or pseudo-covariance, $r_X(t_i - t_j) = \mathbb{E}(X_{t_i}X_{t_j})$, again assuming $\mathbb{E}(X_t) = 0$. In general, both autocovariances are complex-valued and have the properties of Hermitian symmetry and symmetry, respectively (see e.g. Sykulski and Percival (2016)).

In many applications, such as radar and communications, processes are assumed to have the property that $r_X(\cdot) = 0$ (Neeser and Massey, 1993; Picinbono, 1994; Adali et al, 2011); such processes are known as proper or circularly-symmetric and are completely determined by their autocovariance $\gamma_X$. In contrast, applications such as those described in Schreier and Scharf (2010); Adali et al (2011); Chandna and Walden (2017) deal with improper processes, whereby there exists a lag $\tau$ such that $r_X(\tau) \neq 0$. Another often encountered property is that of time-reversibility; for complex-valued processes Didier and Pipiras (2011) have shown that time-reversibility results in complex-valued processes with

real-valued autocovariances, which is precisely the setting under which Sykulski and Percival (2016) develop their exact simulation method for improper stationary Gaussian processes.

### 2.2 Long memory and its estimation

Classical literature for long-range behaviour of real-valued processes shows that persistence is often characterized by a parameter, such as the Hurst exponent, $H$, introduced to the literature by Hurst (1951) in hydrology and its estimation is treated across a large body of established literature, e.g. Beran et al (2013). Mandelbrot and Van Ness (1968) introduced self-similar and related processes with long memory, along with the associated statistical inference. Extensions of fractional Brownian motion to the complex-valued case, defined as a self-similar Gaussian process with stationary increments, are dealt with in e.g. Coeurjolly and Porcu (2017b); Lilly et al (2017). Put simply, the property of self-similarity amounts to the preservation of the process' statistical properties in the face of rescaling, thus naturally fostering the definition of the Hurst exponent.

Just as in the real-valued case, a complex-valued self-similar process $\{X_t\}$ with parameter $H$ satisfies $X(at) \stackrel{d}{=} a^H X(t)$ for $a > 0$, $H \in (0,1)$ and where $\stackrel{d}{=}$ means equal in distribution (Coeurjolly and Porcu, 2017b). Note that the self-similarity definition implies that both the real and imaginary strands of the complex-valued process $\{X_t\}$ evolve according to the same exponent $H$. The property of self-similarity results into the fBM spectrum to behave as $f_X(\omega) = A^2|\omega|^{-2\delta}$ for frequencies $\omega$, a constant $A$ and $\delta \in (1/2, 3/2)$. The spectral slope parameter $\delta$ is linked to the aspect ratio of process rescaling for self-similar behaviour as $H = \delta - 1/2 \in (0,1)$ and also determines the degree of persistence in the differenced version of the process, the fractional Gaussian noise (Lilly et al, 2017). An example of such a process is the improper fractional Gaussian noise with the pseudo-covariance proportional to the autocovariance (both real-valued), both proportional to $\tau^{2\delta-3}$ (Sykulski and Percival, 2016; Lilly et al, 2017).

**Definition 1 (Lilly et al (2017))** A stationary (finite variance) complex-valued process $\{X_t\}$ with real-valued autocovariance $\gamma_X$ is said to have long memory if $\gamma_X(\tau) \sim c_\gamma|\tau|^{-\beta}$ as $|\tau| \to \infty$ and $\beta \in (0,1)$, where $\sim$ means asymptotic equality. In other words, the process autocovariance displays long term decay.

Equivalently, the autocovariance Fourier pair, namely the spectral density, has the property that $f_X(\omega) \sim c_f|\omega|^{-\alpha}$ for frequencies $\omega \to 0$ and $\alpha \in (0,1)$ with $\alpha = 1 - \beta = 2H - 1$. In general, if $0.5 < H < 1$ the process exhibits long memory, with higher $H$ values indicating stronger dependence, whilst if $0 < H < 0.5$ the process has short memory. An improper fractional Gaussian noise constructed as outlined above (Sykulski and Percival, 2016) with $1 < \delta < 3/2$ thus has long memory ($-\beta = 2\delta - 3 = 2H - 2 \in (-1, 0)$ hence $1/2 < H < 1$).

For real-valued time series, estimation of the Hurst exponent $H$ traditionally takes place in the time domain (Mandelbrot and Taqqu, 1979; Bhattacharya et al, 1983; Taqqu et al, 1995; Giraitis et al, 1999; Higuchi, 1990; Peng et al, 1994) and/ or in the frequency domain by means of connections to Fourier or wavelet spectrum decay e.g. Lobato and Robinson (1996), McCoy and Walden (1996), Whitcher and Jensen (2000) and Abry et al (2013). Recent works that deal with long memory estimation in various settings are Vidakovic et al (2000), Shi et al (2005), Hsu (2006), Jung et al (2010), Coeurjolly et al (2014). Some authors have recently considered Hurst estimation using complex-valued wavelets in the regularly spaced real-valued image context, see Nelson and Kingsbury (2010); Jeon et al (2014); Nafornita et al (2014). Reviews comparing several techniques for Hurst exponent estimation (for real-valued series) can be found in e.g. Taqqu et al (1995). Even when only considering real-valued data, Knight et al (2017) show that methods designed for regularly spaced data often fail to deliver a robust estimate if the time series is subject to missing observations or has been sampled irregularly, and in this context they propose a lifting-based approach for Hurst estimation. While this approach serves well when the process is real-valued, it cannot cope with complex-valued processes. Coeurjolly and Porcu (2017b) propose a method of estimation in the setting of (circular) complex-valued fractional Brownian motion assuming a regular sampling structure, but cannot readily cope with sampling irregularity or measurement dropout/ missingness.

### 2.3 Wavelet lifting paradigm for irregularly sampled real-valued data

The *lifting algorithm*, first introduced by Sweldens (1995), constructs 'second-generation' wavelets adapted for non-standard data settings, such as intervals, surfaces, as well as irregularly spaced data. Lifting has since been used successfully for a variety of statistical problems dealing with real-valued signals, including nonparametric regression, spectral estimation and long memory estimation see e.g. Trappe and Liu (2000); Nunes et al (2006); Knight et al (2012, 2017); Hamilton et al (2017).

For a recent review of lifting, the reader is directed to Jansen and Oonincx (2005).

As our proposed lifting transform and subsequent long memory estimation method both make use of a recently developed lifting transform, the *lifting one coefficient at a time* (LOCAAT) transform of Jansen et al (2001, 2009), we shall briefly introduce it next.

Suppose a real-valued function $f(\cdot)$ is observed at a set of $n$, possibly irregular, locations or time points, $\underline{x} = (x_1, \ldots, x_n)$ and is represented by $\{(x_i, f(x_i) = f_i)\}_{i=1}^n$. The lifting algorithm of Jansen et al (2001) begins with the $\underline{f} = (f_1, \ldots, f_n)$ values, known as *scaling function* values, together with an interval associated to each location, $x_i$, which represents the 'span' of that point. By performing LOCAAT, we aim to transform the initial $\underline{f}$ into a set of, say, $L$ coarser scaling coefficients and $(n-L)$ wavelet or detail coefficients, where $L$ is a desired 'primary resolution' scale. This is achieved by repeating three steps: split, predict and update. In the algorithm of Jansen et al (2001), the *split* step is performed by choosing a point to be removed ('lifted'), $j_n$, say. We denote this point by $(x_{j_n}, f_{j_n})$, and identify its set of neighbouring observations, $\mathscr{I}_n$. The *predict* step estimates $f_{j_n}$ by using regression over the neighbouring locations $\mathscr{I}_n$. The prediction error (the difference between the true and predicted function values), $d_{j_n}$ or detail coefficient, is then computed by

$$d_{j_n} = f_{j_n} - \sum_{i \in \mathscr{I}_n} a_i^n f_i, \tag{1}$$

where $(a_i^n)_{i \in \mathscr{I}_n}$ are the weights resulting from the regression procedure. For points with only one neighbour, the prediction is simply $d_{j_n} = f_{j_n} - f_i$. This prediction via regression can of course be carried out using a variety of weights. Notably, Hamilton et al (2017) proposed to use two (rather than just one) prediction filters and encompassed the detail information into complex-valued wavelet coefficients. As more information was extracted from the signal, this approach was shown to improve results for nonparametric regression and spectral/ coherence estimation settings, but nevertheless is limited to real-valued signals. The *update* step consists of updating the $f$-values of the neighbours of $j_n$ used in the predict step using a weighted proportion of the detail coefficient:

$$f_i^{(updated)} := f_i + b_i^n d_{j_n}, \quad i \in \mathscr{I}_n, \tag{2}$$

where the weights $(b_i^n)_{i \in \mathscr{I}_n}$ are subject to the constraint that the algorithm preserves the signal mean value (Jansen et al, 2001, 2009). The interval lengths associated with the neighbouring points are also updated to account for the effect of the removal of $j_n$. In effect, this attributes a portion of the interval associated to the removed point to each neighbour.

These *split*, *predict* and *update* steps are then repeated on the updated signal, and after each iteration a new wavelet coefficient is produced. Hence, after say $(n-L)$ removals, the original data is transformed into $L$ scaling and $(n-L)$ wavelet coefficients. This is similar in spirit to the classical discrete wavelet transform (DWT) step which takes a signal vector of length $2^\ell$ and through filtering operations produces $2^{\ell-1}$ scaling and $2^{\ell-1}$ wavelet coefficients.

An attractive feature of lifting schemes, including the LOCAAT algorithm, is that the transform can be inverted easily by reversing the split, predict and update steps.

The current scarcity of Hurst estimation techniques for complex-valued processes, both in a uniform, but even more so in a non-uniform sampling setting, as well as the effectiveness of the lifting transform in representing irregularly sampled information, jointly motivate our proposed approach to tackle this analysis problem: firstly we propose a novel lifting transform able to cope with irregularly sampled complex-valued processes, and secondly we construct a long memory estimator using the corresponding complex-valued lifting coefficients. Notably, the proposed method is suitable for regularly or irregularly sampled processes, both real- and complex-valued; in particular, Hurst estimation is addressed for improper complex-valued processes that have real-valued covariances, as introduced in Sykulski and Percival (2016), as well as for proper complex-valued series, as described in Coeurjolly and Porcu (2017b).

## 3 A new lifting algorithm for complex-valued signals and its properties

In this section, we introduce our proposed lifting algorithm for a complex-valued function and establish its decorrelation properties.

### 3.1 Proposed $\mathbb{C}^2$-LOCAAT algorithm for complex-valued signals

Suppose now a *complex-valued* function $f(\cdot)$ is observed at a set of $n$, possibly irregular, locations or time points, $\underline{x} = (x_1, \ldots, x_n)$ and is represented by $\{(x_i, f(x_i) = f_i)\}_{i=1}^n$. Our proposed algorithm builds a redundant transform that starts with the complex-valued signal $\underline{f} = (f_1, \ldots, f_n) \in \mathbb{C}^n$ and transforms it into a set of, say, $R$ coarse (complex-valued) scaling coefficients and $2 \times (n-R)$ (complex-valued) detail coefficients, where $R$ is the desired primary resolution scale. As is usual in

lifting, our algorithm re-iterates the three steps—split, predict and update—in a modified version, as described below.

At the first stage ($n$) of the algorithm, denote the smooth coefficients as $c_{n,k} = f_k$, the set of indices of smooth coefficients by $S_n = \{1, \ldots, n\}$ and the set of indices of detail coefficients by $D_n = \emptyset$. The sampling structure is accounted for using the distance between neighbouring observations, and at stage $n$ we define the *span* of $x_k$ as $s_{n,k} = \frac{x_{k+1} - x_{k-1}}{2}$.

At the next stage $(n-1)$, the proposed algorithm proceeds as follows:

*Split:* Choose a point to be removed and denote its index by $j_n$. Typically, points from the densest sampled regions are removed first, but other predefined removal choices are also possible, as we shall discuss below. We shall often refer to the removal order as a *trajectory*, following Knight and Nason (2009).

*Predict:* The set of neighbours ($J_n$) of the point $j_n$ are identified. Note that the set of neighbours is indexed by $n$ as the choice will depend on the removal stage (via the points remaining at that stage). The *predict* step estimates $c_{n,j_n} = f_{j_n}$ by using regression over the neighbouring locations $J_n$ and *two* prediction schemes, a strategy first suggested by Hamilton et al (2017) for real-valued signals. Each prediction scheme is defined by its respective filter, $\mathbf{L}$ and $\mathbf{M}$, orthogonal on each other. The filter $\mathbf{L}$ corresponds to the (possibly) linear regression choice as is usual in LOCAAT. The filter $\mathbf{M}$ is linked to $\mathbf{L}$ through a specific set of properties, discussed in detail in Hamilton et al (2017) and described in step 2 of Algorithm 1. Both filters are constructed such that the corresponding wavelet coefficients of any constant polynomial are 0 (known in the wavelet literature, as possessing (at least) one vanishing moment).

The prediction residuals following the use of each filter are given by

$$\lambda_{j_n} = l^n_{j_n} c_{n,j_n} - \sum_{i \in J_n} l^n_i c_{n,i}, \tag{3}$$

$$\mu_{j_n} = m^n_{j_n} c_{n,j_n} - \sum_{i \in J_n} m^n_i c_{n,i}, \tag{4}$$

where $\{l^n_i\}_{i \in J_n \cup \{j_n\}}$ and $\{m^n_i\}_{i \in J_n \cup \{j_n\}}$ are the prediction weights associated with filters $\mathbf{L}$ and $\mathbf{M}$; as is typical in LOCAAT, we take $l^n_{j_n} = 1$.

Our proposal is to obtain *two* complex-valued detail (wavelet) coefficients by combining the two prediction residuals as follows

$$d^{(1)}_{j_n} = \lambda_{j_n} + \mathrm{i}\,\mu_{j_n}, \tag{5}$$

$$d^{(2)}_{j_n} = \lambda_{j_n} - \mathrm{i}\,\mu_{j_n}. \tag{6}$$

Note that if the original signal is real-valued, then $\underline{d}^{(2)} = \overline{\underline{d}}^{(1)}$ and all we need is $\underline{d}^{(1)}$. However, when the process is complex-valued as is the case here, $\underline{d}^{(2)} \neq \overline{\underline{d}}^{(1)}$ and we need both $\underline{d}^{(1)}$ and $\underline{d}^{(2)}$. This is in contrast to Hamilton et al (2017), where the information from the two prediction schemes is corroborated into just *one* complex-valued wavelet coefficient, and although its naive implementation on the real and imaginary process strands would yield two sets of complex-valued wavelet coefficients, it would not be obvious how to best combine their information.

*Update:* In the update step, both the (complex-valued) smooth coefficients $\{c_{n,i}\}$ and (real-valued) spans of the neighbours $\{s_{n,i}\}$ are updated according to filter $\mathbf{L}$:

$$\begin{aligned} c_{n-1,i} &= c_{n,i} + b^n_i \lambda_{j_n}, \\ s_{n-1,i} &= s_{n,i} + l^n_i s_{n,j_n} \quad \forall i \in J_n, \end{aligned} \tag{7}$$

where $b^n_i = (s_{n,j_n} s_{n-1,i}) / (\sum_{i \in J_n} s^2_{n-1,i})$ are the update weights, again computed so that the mean of the signal is preserved (Jansen et al, 2009). Updating the neighbours' spans accounts for the modification to the sampling grid induced by removing one of the observations, and using just one filter for update (akin to the approach of Hamilton et al (2017)) ensures the use of a common scale across both $\underline{d}^{(1)}$ and $\underline{d}^{(2)}$.

The observation $j_n$ is then removed from the set of smooth coefficients, hence after the first algorithm iteration, the index set of smooth coefficients is $S_{n-1} = \{1, ..., n\} \backslash \{j_n\}$ and the index set of detail coefficients is $D_{n-1} = \{j_n\}$. The algorithm is then reiterated until the desired primary resolution level $R$ has been achieved. In practice, the choice of the primary level $R$ in LOCAAT lifting schemes is not crucial provided it is sufficiently low (Jansen et al, 2009), with $R = 2$ recommended by Nunes et al (2006).

The three steps are then repeated on the updated signal, and each repetition yields two new wavelet coefficients. After points $j_n, j_{n-1}, \ldots, j_{R+1}$ have been removed, the function can be represented as a set of $2 \times (n-R)$ detail coefficients, $\{d^{(1)}_{j_k}\}_{k \in D_{n-R}}$ and $\{d^{(2)}_{j_k}\}_{k \in D_{n-R}}$, and $R$ smooth coefficients, $\{c_{r-1,i}\}_{i \in S_{n-R}}$, thus resulting in a redundant transform. An algorithmic description of $\mathbb{C}^2$-LOCAAT appears in Algorithm 1.

The proposed algorithm can then be easily inverted by recursively 'undoing' the update, predict and split steps described above for the *first* filter ($\mathbf{L}$). More specifically, the inverse transform can be performed by the steps

*Undo Update:* $c_{n,i} = c_{n-1,i} - b^n_i \lambda_{j_n}, \ \forall i \in J_n$

`Proposed` $\mathbb{C}^2$`-LOCAAT using two symmetrical neighbours:` Choose a removal order (trajectory), either dictated by the sampling sequence or following a random permutation.

1. Split: Choose the first/next point to be removed from the set of smooth coefficients $S_n = \{1, ..., n\}$ and denote its index by $j_n$.
2. Predict:
   (a) Determine the set of neighbours $J_n$ (one each side of $j_n$) and use linear regression over the neighbourhood in order obtain a prediction at $j_n$.
   Calculate the prediction residual, $\lambda_{j_n}$, as the difference between the observed and predicted values at $j_n$ (see equation (3)). This coupled with the requirement of achieving at least one vanishing moment amounts to obtaining a filter $\mathbf{L} = (l_1, 1, l_3)$ with $l_1 + l_3 = 1$.
   (b) Construct a new filter $\mathbf{M} = (Am, (1 + A)m, m)$ with $A = \frac{l_1-2}{l_1+1}$ and $m = \frac{l_1+1}{\sqrt{3}}$. By construction, $\mathbf{M}$ is orthogonal on $\mathbf{L}$, has at least one vanishing moment and $\|\mathbf{L}\| = \|\mathbf{M}\|$. Using $\mathbf{M}$, obtain a new prediction residual, $\mu_{j_n}$ (see equation (4)).
   (c) The complex-valued wavelet (detail) coefficients at $j_n$ are $d^{(1)}_{j_n} = \lambda_{j_n} + \mathrm{i}\,\mu_{j_n}$ and $d^{(2)}_{j_n} = \lambda_{j_n} - \mathrm{i}\,\mu_{j_n}$.
3. Update: the smooth coefficients and their associated scales using the filter $\mathbf{L}$ (see equations (7)).
   Update the index sets of smooth and detail coefficients as $S_{n-1} = S_n \backslash \{j_n\}$ and $D_{n-1} = \{j_n\}$ respectively.
4. **Iterate** steps 1–3 for $j_{n-1}, \ldots, j_{R+1}$ with a typical primary resolution level $R = 2$, hence obtain a set of complex-valued wavelet coefficients indexed by $D_R = \{j_n, ..., j_{R+1}\}$.

Alg. 1: The complex-valued lifting scheme ($\mathbb{C}^2$-LOCAAT) on a complex-valued signal.

*Undo Predict:*

$$c_{n,j_n} = \frac{\lambda_{j_n} - \sum_{i \in J_n} l^n_i c_{n,i}}{l^n_{j_n}} \qquad \text{or} \tag{8}$$

$$c_{n,j_n} = \frac{\mu_{j_n} - \sum_{i \in J_n} m^n_i c_{n,i}}{m^n_{j_n}}. \tag{9}$$

Undoing either predict (8) or (9) step is sufficient for inversion.

A few remarks on our proposed $\mathbb{C}^2$-LOCAAT lifting algorithm are now in order.

*Transform matrix representation.* As with any linear transform, the algorithm that determines one set of detail coefficients, say $\underline{d}^{(1)}$, can also be represented using a matrix transform, i.e. $\underline{d}^{(1)} = W^{(c)}\underline{f}$, where $W^{(c)}$ is a $n \times n$ matrix with complex-valued entries. When expressed as a matrix transform, our proposed $\mathbb{C}^2$-LOCAAT algorithm for a complex-valued process ($\underline{f}$) can be expressed as

$$\underline{d} = \begin{pmatrix} W^{(c)} \\ \overline{W}^{(c)} \end{pmatrix} \underline{f} \tag{10}$$

$$= \begin{pmatrix} \underline{d}^{(1)} \\ \underline{d}^{(2)} \end{pmatrix}, \tag{11}$$

with $\underline{d}^{(1)} = W^{(c)}\underline{f}$ and $\underline{d}^{(2)} = \overline{W}^{(c)}\underline{f}$.

*Wavelet lifting scales and artificial levels.* The ($\log_2$) span associated with an observation at the last stage before its removal, say $\log_2(s_{k,j_k})$ for the detail coefficient $d_{j_k}$ obtained at stage $k$, is used as a (continuous) measure of scale – this indirectly stems from the fact the wavelets are not dyadically scaled versions of a single mother wavelet. As the notion of scale of lifting wavelets is continuous, Jansen et al (2009) group wavelet functions of similar (continuous) scales into 'artificial' levels, to mimic the dyadic levels of classical wavelets (see Jansen et al (2001, 2009) for more details). We also adopt this strategy to group the complex-valued wavelet coefficients produced using our $\mathbb{C}^2$-LOCAAT algorithm. An alternative is to group the coefficients via their interval lengths into ranges $(2^{j-1}\alpha_0, 2^j\alpha_0]$, where $j \geq 1$ and $\alpha_0$ is the minimum scale. This construction more closely resembles classical wavelet dyadic scales, but both produce similar results. Note that by construction, the $\mathbb{C}^2$-LOCAAT transform crucially uses a common scale for both real and imaginary parts, and it is this feature that ensures that information is obtained on the same scale at every step.

*Choice of removal order.* The lifting algorithms in Section 2.3 and Section 3.1 are inherently dependent on the order in which points are removed as the algorithm progresses. Jansen et al (2009) remove points in order from the finest continuous scale to the coarsest, to mimic the DWT, which produces coefficients at the finest scale first, then at progressively coarser scales. However, in our proposed $\mathbb{C}^2$-LOCAAT scheme, we can choose to remove points according to a predefined *path* (or *trajectory*) $T = (x_{o_1}, \ldots, x_{o_n})$, where $(o_1, o_2, \ldots, o_n)$ is a permutation of the set $\{1, \ldots, n\}$. Knight and Nason (2009) introduced the *nondecimated* lifting transform, which proposes examining data using $P$ bootstrapped paths from the space of $n!$ possible trajectories. Aggregating the information obtained via this approach typically improves estimator variance and accuracy, not only in the long memory estimation context (Knight et al, 2017), but also for e.g. nonparametric regression (Knight and Nason, 2009). This strategy will be embedded in our proposed methodology in Section 4.

### 3.2 Refinement equations for the scaling and wavelet functions under $\mathbb{C}^2$-LOCAAT

Although not explicitly apparent, the wavelet lifting construction induces a biorthogonal (second generation) wavelet basis construction, see e.g. Sweldens (1995). In the *real-valued* lifting one coefficient at a time paradigm, as the algorithm progresses, scaling and wavelet functions decomposing the frequency content of the signal are built recursively according to the predict and update equations (1) and (2) (Jansen et al, 2009). Also, the (dual) scaling functions are defined recursively as linear combinations of (dual) scaling functions at the previous stage.

Let us now investigate the basis decomposition afforded by our proposed $\mathbb{C}^2$-LOCAAT transform, as a result of performing the split, predict and update steps. As our construction involves two prediction filters, we decompose $f$ on two biorthogonal bases. Our construction is reminiscent of the dual tree complex wavelet transform (ℂWT) (Kingsbury, 2001; Selesnick et al, 2005) which employs two separate classical wavelet transforms, but fundamentally differs through the construction of linked orthogonal filters.

In our proposed construction, let us denote the two scaling function and wavelet biorthogonal bases by $\left\{\underline{\varphi}^{(1)}, \underline{\tilde{\varphi}}^{(1)}, \underline{\psi}^{(1)}, \underline{\tilde{\psi}}^{(1)}\right\}$ and $\left\{\underline{\varphi}^{(2)}, \underline{\tilde{\varphi}}^{(2)}, \underline{\psi}^{(2)}, \underline{\tilde{\psi}}^{(2)}\right\}$ respectively. We now explore their relationships and recursive construction.

At stage $r$, the complex-valued signal $f$ can be decomposed on each basis as

$$f(x) = \sum_{\ell \in D_r} d_\ell^{(i)} \psi_\ell^{(i)}(x) + \sum_{k \in S_r} c_{r,k}^{(i)} \varphi_{r,k}^{(i)}(x), \quad i = 1, 2, \tag{12}$$

with $d_\ell^{(i)} =< f, \tilde{\psi}_\ell^{(i)} >$ and $c_{r,k}^{(i)} =< f, \tilde{\varphi}_{r,k}^{(i)} >$ for both bases $i = 1, 2$, where the inner product is as usual defined on $L^2(\mathbb{C})$. As the update step is the same for both bases, it follows that $c_{r,k}^{(1)} = c_{r,k}^{(2)}$. Hence denote $c_{r,k} =< f, \tilde{\varphi}_{r,k}^{(1)} >=< f, \tilde{\varphi}_{r,k}^{(2)} >$, for all $r, k$ and thus the dual scaling functions coincide under both bases. In what follows we shall denote these by $\tilde{\varphi}_{r,k}$.

**Proposition 1** *Suppose we are at stage $r-1$ of the $\mathbb{C}^2$-LOCAAT algorithm. The recursive construction of the primal scaling and wavelet functions corresponding to the coefficients $\underline{d}^{(1)}$, in terms of the functions at the previous stage $r$, is given by*

$$\varphi_{r-1,j}^{(1)}(x) = \varphi_{r,j}^{(1)}(x) + \tilde{a}_j^r \varphi_{r,j_r}^{(1)}(x), \ \textit{if } j \in J_r, \tag{13}$$

$$\varphi_{r-1,j}^{(1)}(x) = \varphi_{r,j}^{(1)}(x), \ \textit{if } j \notin J_r, \tag{14}$$

$$\psi_{j_r}^{(1)}(x) = \frac{\overline{a}_{j_r}^r}{|a_{j_r}^r|^2} \varphi_{r,j_r}^{(1)}(x) - \sum_{j \in J_r} b_j^r \varphi_{r-1,j}^{(1)}(x), \tag{15}$$

*where $a_j^r = \ell_j^r + \mathrm{i}\, m_j^r$ and $\tilde{a}_j^r = \frac{\overline{a}_{j_r}^r a_j^r}{|a_{j_r}^r|^2}$.*

*Similarly, the recursive construction for the primal scaling and wavelet functions corresponding to the coefficients $\underline{d}^{(2)}$, in terms of the functions at the previous stage $r$, is given by*

$$\varphi_{r-1,j}^{(2)}(x) = \varphi_{r,j}^{(2)}(x) + \overline{\tilde{a}}_j^r \varphi_{r,j_r}^{(2)}(x), \ \textit{if } j \in J_r, \tag{16}$$

$$\varphi_{r-1,j}^{(2)}(x) = \varphi_{r,j}^{(2)}(x), \ \textit{if } j \notin J_r, \tag{17}$$

$$\psi_{j_r}^{(2)}(x) = \frac{a_{j_r}^r}{|a_{j_r}^r|^2} \varphi_{r,j_r}^{(2)}(x) - \sum_{j \in J_r} b_j^r \varphi_{r-1,j}^{(2)}(x). \tag{18}$$

*For the corresponding dual bases the recursive constructions are given by*

$$\tilde{\varphi}_{r-1,j}(x) = \tilde{\varphi}_{r,j}(x) + b_j^r \tilde{\psi}_{j_r}^L(x), \forall j \in J_r, \tag{19}$$

$$\tilde{\varphi}_{r-1,j}(x) = \tilde{\varphi}_{r,j}(x), \forall j \notin J_r, \tag{20}$$

$$\tilde{\psi}_{j_r}^{(1)}(x) = a_{j_r}^r \tilde{\varphi}_{r,j_r}(x) - \sum_{j \in J_r} a_j^r \tilde{\varphi}_{r,j}(x), \tag{21}$$

$$\tilde{\psi}_{j_r}^{(2)}(x) = \overline{a}_{j_r}^r \tilde{\varphi}_{r,j_r}(x) - \sum_{j \in J_r} \overline{a}_j^r \tilde{\varphi}_{r,j}(x), \tag{22}$$

*where $\tilde{\psi}^L$ denotes the dual wavelet function corresponding to the* **L***-filter only.*

The proof can be found in Appendix A, Section A.1.

Summarizing, the two bases can be represented as $\{\underline{\varphi}^{(1)}, \underline{\tilde{\varphi}}, \underline{\psi}^{(1)}, \underline{\tilde{\psi}}^{(1)}\}$ and $\{\underline{\overline{\varphi}}^{(1)}, \underline{\tilde{\varphi}}, \underline{\overline{\psi}}^{(1)}, \underline{\tilde{\psi}}^{(2)}\}$ and their recursive construction established above will be used in obtaining the formal properties required to justify our proposed long memory estimation approach.

### 3.3 Decorrelation properties of the $\mathbb{C}^2$-LOCAAT algorithm

Wavelet transforms are known to possess good decorrelation properties, see in the context of long memory processes e.g. Abry et al (2000); Jensen (1999); Craigmile et al (2001) for classical wavelets, and Knight et al (2017) for lifting wavelets constructed by means of LOCAAT. The decorrelation property amounts to the consequent removal of the long memory in the wavelet domain, and thus estimation of the Hurst exponent can be carried out in this simplified context. Therefore, we next provide mathematical evidence for the decorrelation properties of the $\mathbb{C}^2$-LOCAAT algorithm and these will subsequently benefit our proposed long memory estimation procedure (see Section 4). The statement of Proposition 2 (next) aims to establish decorrelation results similar to earlier ones concerning regular wavelets (see e.g. Abry et al (2000, p.51) for fractional Gaussian noise, Jensen (1999, Theorem 2) for fractionally integrated processes or Theorem 5.1 of Craigmile and

Percival (2005) for fractionally differenced processes) and lifting wavelets (see Proposition 1 in Knight et al (2017)). In what follows, we establish the decorrelation properties for the proposed complex-valued lifting transform $\mathbb{C}^2$-LOCAAT in a more general data setting than previously considered for lifting wavelets, involving complex-valued stationary processes with real-valued autocovariances, that may be proper or improper in nature.

**Proposition 2** *Let $X = \{X_{t_i}\}_{i=0}^{N-1}$ denote a (zero-mean) stationary long memory complex-valued time series with Lipschitz continuous spectral density $f_X$. Assume the process is observed at irregularly spaced times $\{t_i\}_{i=0}^{N-1}$ and let $\{\{c_{R,i}\}_{i\in\{0,\ldots,N-1\}\setminus\{j_{N-1},\ldots,j_{R-1}\}}, \{\underline{d}_{j_r}\}_{r=R-1}^{N-1}\}$ be the $\mathbb{C}^2$-LOCAAT transform of $X$, where $\underline{d}_{j_r} = \begin{pmatrix} d_{j_r}^{(1)} & d_{j_r}^{(2)} \end{pmatrix}^T$. Then both sets of detail coefficients $\{d_{j_r}^{(1)}\}_r$ and $\{d_{j_r}^{(2)}\}_r$ have autocorrelation and pseudo-autocorrelation whose magnitudes decay at a faster rate than for the original process.*

The proof can be found in Appendix A, Section A.2 and uses similar arguments to the proof of Proposition 1 in Knight et al (2017), adapted for the $\mathbb{C}^2$-LOCAAT algorithm and complex-valued setting we address here. Just as for LOCAAT (Knight et al, 2017), Proposition 2 above assumes no specific lifting wavelet and we conjecture that if smoother lifting wavelets were employed, it might be possible to obtain even better rates of decay.

## 4 Long memory parameter estimation using complex wavelet lifting (ℂLoMPE)

As the newly constructed wavelet domain through $\mathbb{C}^2$-LOCAAT displays small magnitude autocorrelations, we now focus on the wavelet coefficient variance and show that the $\log_2$-variance of each of the complex-valued lifting coefficients $d^{(1)}$ and $d^{(2)}$ is linearly related to their corresponding artificial scale level, a result paralleling classical and real-valued lifting wavelet results. This result suggests a Hurst parameter estimation method for potentially irregularly sampled long memory processes that take values in the complex ($\mathbb{C}$) domain.

Proposition 3 next establishes a result similar to that in Proposition 2 of Knight et al (2017) by taking into account the specific $\mathbb{C}^2$-LOCAAT construction and thus extends the scope of Hurst estimation methodology to irregularly sampled complex-valued processes.

**Proposition 3** *Let $X = \{X_{t_i}\}_{i=0}^{N-1}$ denote a (zero-mean) complex-valued long memory stationary time series with finite variance and spectral density $f_X(\omega) \sim c_f|\omega|^{-\alpha}$ as $\omega \to 0$, for some $\alpha \in (0,1)$. Assume the series is observed at irregularly spaced times $\{t_i\}_{i=0}^{N-1}$ and transform the observed data $X$ into a collection of lifting coefficients, $\{d_{j_r}^{(1)}\}_r$ and $\{d_{j_r}^{(2)}\}_r$, via application of $\mathbb{C}^2$-LOCAAT from Section 3.1.*

*Let $r$ denote the stage of $\mathbb{C}^2$-LOCAAT at which we obtain the wavelet coefficients $d_{j_r}^{(\ell)}$ (with $\ell = 1, 2$) and let its corresponding artificial level be $j^\star$. Then, denoting by $|\cdot|$ the $\mathbb{C}$-modulus, we have for some constant $K$*

$$(\sigma_{j^\star}^{(\ell)})^2 = \mathbb{E}(|d_{j_r}^{(\ell)}|^2) \sim 2^{j^\star(\alpha-1)} \times K. \tag{23}$$

The proof can be found in Appendix A, Section A.3. This result suggests a long memory parameter estimation method for an irregularly sampled, complex-valued time series, described in Algorithm 2 below, which we shall refer to as ℂLoMPE (Complex-valued Long Memory Parameter Estimation Algorithm). Section 5.1, next, will show that our proposed ℂLoMPE methodology below not only adds a new much needed tool in the estimation of long memory for complex-valued processes, but also improves Hurst exponent estimation for real-valued processes, sampled both regularly and irregularly.

## 5 Simulated performance of ℂLoMPE and real data analysis

### 5.1 Simulated performance of ℂLoMPE

In what follows we investigate the performance of our Hurst parameter estimation technique for complex-valued series. We simulated realisations of two types of long memory processes, namely circularly symmetric complex fractional Brownian motion, as introduced in Coeurjolly and Porcu (2017a), and improper complex fractional Gaussian noise (with real-valued covariances) as described in Sykulski and Percival (2016)[1], investigating series of lengths of 256, 512 and 1024. These lengths were chosen to reflect realistic data collection scenarios – long enough for the Hurst parameter (a low-frequency asymptotic quantity) to be reasonably estimated, whilst reflecting lengths of datasets encountered in practice.

To investigate the effect of sampling irregularity on the performance of our method, we simulated datasets with different levels of random missingness (5% to 20%), which are representative of degrees of missingness reported in many application areas, for example in paleoclimatology and environmental series (Broersen, 2007; Junger and Ponce de Leon, 2015).

[1] We would like to thank Adam Sykulski for supplying the `Matlab` code to simulate the improper complex fractional Gaussian noise processes.

`Complex-valued Long Memory Parameter Estimation Algorithm (`$\mathbb{C}$`LoMPE):`
Assume that $\{X_{t_i}\}_{i=0}^{N-1}$ is as in Proposition 3. We estimate $\alpha$ as follows.

1. Apply $\mathbb{C}^2$-LOCAAT to the complex-valued observed process $\{X_{t_i}\}_{i=0}^{N-1}$ using a particular lifting trajectory to obtain the coefficients $\{\underline{d}_{j_r} = \begin{pmatrix} d_{j_r}^{(1)} & d_{j_r}^{(2)} \end{pmatrix}^T\}_r$, see equation (10).
2. Normalize both sets of (complex-valued) detail coefficients by their corresponding $\mathbb{C}$-modulus: divide each squared ($\mathbb{C}$) modulus by the corresponding diagonal entry of $W^{(c)}\overline{W}^{(c),T}$, where $W^{(c)}$ is the complex-valued lifting transform matrix corresponding to $\underline{d}^{(1)}$.
3. Group the coefficients into a set of artificial scales as described in Section 2.3. Estimate the wavelet energy within the artificial level $j^\star$ by

$$\left(\hat{\sigma}_{j^\star}^{(\ell)}\right)^2 := (n_{j^\star}-1)^{-1} \sum_{r=1}^{n_{j^\star}} |d_{j_r}^{(\ell)}|^2, \text{ for each } \ell = 1, 2, \quad (24)$$

   where $n_{j^\star}$ is the number of observations in artificial level $j^\star$. Note that the $\mathbb{C}^2$-LOCAAT construction, by its use of an unique update step, ensures that the number of observations in each $j^\star$ artificial level coincide for both $\ell = 1$ and $\ell = 2$.
4. Fit a weighted linear regression to all points $\log_2\left(\hat{\sigma}_{j^\star}^{(\ell)}\right)^2$ with $\ell = 1, 2$ versus $j^\star$; use its slope to estimate $\alpha$ as suggested by the results in Proposition 3. Note that equation (23) allows us to pull the information across both $\underline{d}^{(1)}$ and $\underline{d}^{(2)}$.
5. **Iterate** steps A-1 to A-4 for $P$ bootstrapped trajectories, obtaining an estimate $\hat{\alpha}_p$ for each trajectory $p \in \overline{1, P}$. The final estimator is $\hat{\alpha} = P^{-1}\sum_{p=1}^{P} \hat{\alpha}_p$, from which an appropriate estimate for $H$ can be obtained.

Alg. 2: The long memory parameter estimation procedure ($\mathbb{C}$LoMPE) for a complex-valued process $\{X_{t_i}\}_{i=0}^{N-1}$, sampled at potentially irregularly spaced times.

We compared results across the range of Hurst parameters $H = 0.6, \ldots, 0.9$. Each set of results is taken over $K = 100$ realizations and $P = 50$ lifting trajectories. Our $\mathbb{C}$LoMPE technique was implemented using modifications to the code from the *liftLRD* package (Knight and Nunes, 2016) and *CNLTreg* package (Nunes and Knight, 2017) for the `R` statistical programming language (R Core Team, 2013), both available on CRAN. The measure we use to assess the performance of the methods is the mean squared error (MSE) defined by

$$\text{MSE} = K^{-1}\sum_{k=1}^{K}(H - \hat{H}^k)^2. \quad (25)$$

In the case of *regularly spaced* circularly symmetric fractional Brownian motion (i.e. 0% missingness), we compare our $\mathbb{C}$LoMPE estimation technique with the recent estimation method in Coeurjolly and Porcu (2017b) (denoted "CP")[2].

Table 1 reports the mean squared error for our $\mathbb{C}$LoMPE estimator on the complex-valued fractional Brownian motion series for different degrees of missingness (0% up to 20%). In the case of regularly spaced series, our estimation method works well when compared to the "CP" method. This is pleasing since the "CP" method is designed for regularly spaced series, whereas $\mathbb{C}$LoMPE is specifically designed for irregularly spaced series. The tables also show that the $\mathbb{C}$LoMPE technique is robust to the presence of missingness, attaining good performance even for high degrees of missingness (20%).

For the complex-valued fractional Gaussian noise, Table 2 demonstrates that our $\mathbb{C}$LoMPE estimation technique performs well for regular and irregular settings, with only a slight degradation in performance for increasing missingness.

We also studied the empirical bias of our estimator for both types of long memory process. For reasons of brevity we do not report these results here, but these can be found in Appendix B in the supplementary material. As for the mean squared error results above, there is a small drop in performance with increasing missingness, and our estimator performs only slightly worse in terms of bias when compared to the "CP" method.

*Real-valued processes.* To assess whether our complex-valued approach achieves performance gains for real-valued processes, we repeated the simulation study from Knight et al (2017) for a number of long memory processes. In particular, we studied the performance of our estimator for real-valued fractional Brownian motion, fractional Gaussian noise and fractionally integrated series, for a range of Hurst parameters and levels of missingness. The processes were simulated via the *fArma* add-on package (Wuertz et al, 2013). We compare our method with the real-valued lifting technique of Knight et al (2017), shown to perform well in a number of settings. Again, for brevity, we do not report these bias results here, but they can be found in Appendix B in the supplementary material. The results show that our method is competitive with the real-valued estimation method in Knight et al (2017), achieving bet-

[2] The authors would like to thank Jean-François Coeurjolly for providing the `R` code for simulating the circular fractional Brownian motion series, as well as for the implementation of the estimation technique of Coeurjolly and Porcu (2017b).

Table 1: Mean squared error ($\times 10^3$) for fractional Brownian motion series featuring different degrees of missing observations for a range of Hurst parameters for the ℂLoMPE estimation procedure. Boxed numbers indicate best result for the regularly spaced setting. Numbers in brackets are the estimation errors' standard deviation.

| | $n = 256$ | | | | | $n = 512$ | | | | | $n = 1024$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Missingness proportion, $p$ | | | | | Missingness proportion, $p$ | | | | | Missingness proportion, $p$ | | | | |
| | CP | ℂLoMPE | | | | CP | ℂLoMPE | | | | CP | ℂLoMPE | | | |
| $H$ | 0% | 0% | 5% | 10% | 20% | 0% | 0% | 5% | 10% | 20% | 0% | 0% | 5% | 10% | 20% |
| 0.6 | 2 (3) | **1** (2) | 1 (2) | 1 (1) | 2 (3) | **1** (2) | **1** (1) | 0 (0) | 0 (1) | 1 (1) | **1** (1) | **1** (1) | 0 (0) | 0 (0) | 0 (0) |
| 0.7 | 2 (3) | **1** (2) | 1 (1) | 1 (2) | 2 (3) | **1** (1) | **1** (1) | 1 (1) | 1 (1) | 1 (1) | **0** (1) | 2 (1) | 1 (1) | 1 (1) | 0 (0) |
| 0.8 | 3 (3) | **2** (2) | 2 (2) | 1 (2) | 2 (2) | **1** (2) | 2 (2) | 1 (2) | 1 (2) | 1 (2) | **1** (1) | 3 (2) | 2 (2) | 2 (1) | 1 (1) |
| 0.9 | **2** (3) | 3 (4) | 2 (3) | 2 (3) | 2 (2) | **1** (2) | 2 (2) | 2 (3) | 2 (2) | 2 (2) | **2** (2) | **2** (2) | 3 (2) | 3 (2) | 2 (2) |

Table 2: Mean squared error ($\times 10^3$) for fractional Gaussian noise featuring different degrees of missing observations for a range of Hurst parameters for the ℂLoMPE estimation procedure. Numbers in brackets are the estimation errors' standard deviation.

| | $n = 256$ | | | | $n = 512$ | | | | $n = 1024$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Missingness proportion, $p$ | | | | Missingness proportion, $p$ | | | | Missingness proportion, $p$ | | | |
| $H$ | 0% | 5% | 10% | 20% | 0% | 5% | 10% | 20% | 0% | 5% | 10% | 20% |
| 0.6 | 1 (2) | 1 (2) | 1 (2) | 2 (2) | 1 (1) | 1 (1) | 1 (1) | 1 (1) | 1 (1) | 1 (1) | 1 (1) | 1 (1) |
| 0.7 | 1 (2) | 2 (2) | 2 (2) | 2 (3) | 1 (1) | 2 (2) | 2 (2) | 3 (2) | 2 (1) | 2 (1) | 2 (1) | 3 (2) |
| 0.8 | 2 (2) | 2 (3) | 2 (3) | 3 (5) | 2 (2) | 3 (3) | 3 (3) | 4 (4) | 2 (2) | 3 (2) | 3 (2) | 5 (3) |
| 0.9 | 3 (4) | 3 (3) | 3 (3) | 3 (5) | 2 (2) | 2 (3) | 3 (3) | 3 (3) | 2 (2) | 3 (2) | 3 (2) | 4 (3) |

ter results (in terms of MSE and bias) in the majority of cases for fractional Gaussian noise and fractionally integrated series. For fractional Brownian motion, we observe that our method achieves gains in mean square error, albeit at a cost of a decrease in bias performance. These results agree with other studies using complex-valued wavelet methodology, which is shown to outperform its real-valued counterpart in a variety of applications, from denoising (Barber and Nason, 2004) to Hurst estimation in the (real-valued) image context (Nelson and Kingsbury, 2010; Jeon et al, 2014; Nafornita et al, 2014). This is due to the use of two rather than just one filter, thus eliciting more information from the signal under analysis.

## 5.2 Analysis of complex-valued wind series with ℂLoMPE

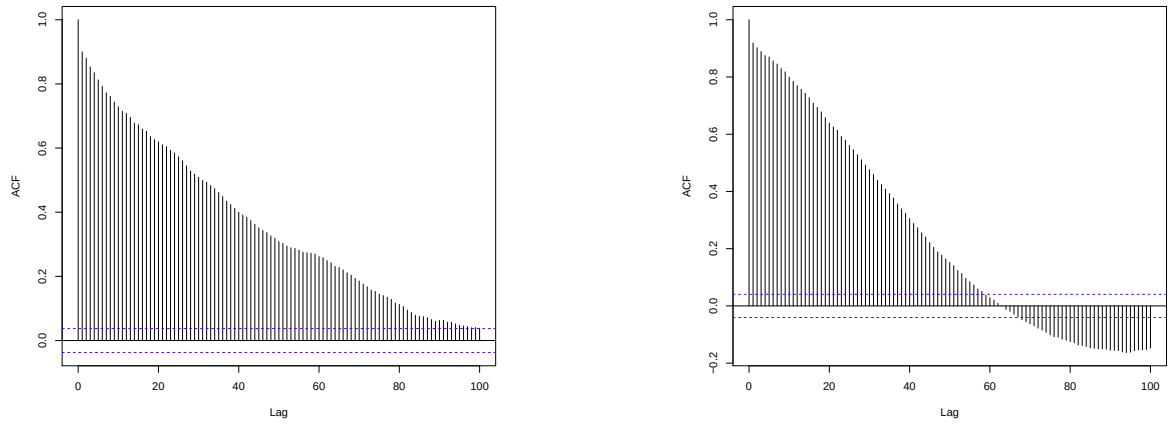
In this section we provide a more detailed long memory analysis of the complex-valued wind series described in Section 1.1. More specifically, we applied our ℂLoMPE Hurst estimation method to the (detrended) irregularly sampled wind series to assess its persistence properties. The estimated Hurst parameter was $\hat{H}_{\mathbb{C}} = 0.86$ for the `Wind A` series and $\hat{H}_{\mathbb{C}} = 0.8$ for the `Wind B` series, based on $P = 50$ lifting trajectories. Both of these estimates indicate moderate long memory.

To highlight potential differences with other approaches, we also performed the LoMPE technique of Knight et al (2017) to each of the real and imaginary components of the two series. In addition, we also estimated the Hurst exponent using the Knight et al (2017) method for the two *magnitude* series, since such series (i.e. data without directional information) are most commonly analysed in the literature. The Hurst exponent estimates are denoted $\hat{H}_{\mathbb{R}}$ and $\hat{H}_{\mathscr{I}}$ for the real and imaginary component series, and $\hat{H}_{Mod}$ for the magnitude series. The estimates are summarized in Table 3.

Table 3: Hurst parameter estimates for the `Wind A` and `Wind B` data from complex-valued series using ℂLoMPE and from real-valued component and magnitude series using LoMPE.

| Dataset | $\mathbb{R}$ | $\mathscr{I}$ | $Mod$ | $\mathbb{C}$ |
|---|---|---|---|---|
| `Wind A` | 0.90 | 0.82 | 0.80 | 0.86 |
| `Wind B` | 0.85 | 0.75 | 0.80 | 0.80 |

For the `Wind A` dataset, our ℂLoMPE technique estimates the persistence as between those of the real and imaginary components, and higher than that of the magnitude series. In contrast, for the `Wind B` dataset, the estimate from our complex-valued approach coincides with the result for the series derived from the ℂ-modulus. This analysis highlights that ignoring the dependence structure between the real and imaginary components of the series may result in misestimation. Hence we recommend an approach that uses the complex-




(a) (b)

Fig. 4: (a) Autocorrelation for the *magnitude* wind series for the `Wind A` series from Figure 1 (treated as regularly spaced); (b) autocorrelation for the *magnitude* `Wind B` dataset from Figure 1 (treated as regularly spaced). The dependence structure is markedly different to that shown for the real and imaginary series components shown in Figure 2.

valued structure of the data, thus accounting for its intrinsic rotary structure and dependence, not visible by only using the traditional magnitude series or individual real and imaginary strands.

It could also be argued that these differences in estimates are unsurprising, since the dependence structure for the magnitude series, shown in Figure 4, is visibly different to that of the real and imaginary component series shown in Figure 2. We argue that our estimation of the long memory parameter for this series is more reliable than that in the current existing literature, as our proposed algorithm naturally encompasses both the complex-valued and improper features of wind series. A complex-valued analysis using our approach could hence provide more accurate long memory information, reducing miscalibration of predictive climate models. We further suggest that this precision would provide more certainty when assessing renewable energy resource potential, as discussed in e.g. Bakker and van den Hurk (2012).

## 6 Discussion

Hurst exponent estimation is a recurrent topic in many scientific applications, with significant implications for modelling and data analysis. One important aspect of real-world datasets is that their collection and monitoring are often not straightforward, leading to missingness, or to the use of proxies with naturally irregular sampling structures. In parallel, in many applications of interest there is a natural complex-valued representation of data. To this end, this article has proposed the first Hurst estimation technique for complex-valued processes with sampling missingness or irregularity, and in doing so it has also constructed a novel lifting algorithm able to work on complex-valued data sampled with irregularity. Until the work in this article, Hurst estimation methods have not been able to exploit the wealth of signal information in such data, whilst also coping with irregular sampling regimes. Our $\mathbb{C}$LoMPE wavelet lifting methodology was shown to give accurate Hurst estimation for a variety of complex-valued fractional processes, and is suitable for both proper and improper complex-valued processes. Simulations demonstrate that the technique is robust to estimation with significant degrees of missingness, as well as in the non-missing (regular) setting.

We have demonstrated the use of our $\mathbb{C}$LoMPE technique in an application arising in environmental science. Through our analysis of wind speed data, we have shown that embedding directional wind information in the analysis can lead to significantly different Hurst exponent estimates when compared to only considering real-valued information, such as magnitude series. This highlights that not exploiting a complex-valued data representation in this setting can potentially result in misleading conclusions being drawn about wind persistence. This in turn has a subsequent impact on parameters in climate models and inefficiencies in resource management decisions.

Whilst the development of our proposed complex-valued Hurst estimator was motivated by an application in climatology, we believe that the work in this article has sufficient generality to have appeal in other settings. We thus conclude this article with outlining some example applications in which our methodology is potentially beneficial.

***Data from neuroimaging studies.*** Functional magnetic resonance imaging (fMRI) data continues to enjoy popularity in the neuroscience community due to its non-invasive acquisition and data richness, see e.g. Aston and Kirch (2012) for an accessible introduction to the area from the statistical perspective. In particular, fMRI studies often measure information on blood flow in the brain; these voxel-level data are used to investigate neuronal activity of participants during task-based experiments, and many authors have asserted that such time courses possess fractional noise structure, see e.g. Bullmore et al (2003). Evaluation of the Hurst exponent in this context has been shown to be important in characterising brain activity under a range of conditions, indicating different levels of cognitive effort (Park et al, 2010; Ciuciu et al, 2012; Churchill et al, 2016). Despite data collection being performed in a controlled setup, recent work has highlighted the need for tailored statistical methodology to cope with both unbalanced designs, as well as missingness, which can feature in fMRI data for a number of reasons (Lindquist, 2008; Ferdowsi and Abolghasemi, 2017). In actuality, fMRI scanners record both phase and magnitude information, though most studies only use the magnitude image for analysis. As a result, there has been a recent body of work dedicated to *complex-valued* analysis of fMRI data, most notably by Rowe and collaborators (see e.g. Rowe (2005, 2009); Adrian et al (2017)). Such an approach has shown improvements over real-valued methods for a range of analysis tasks, see also the work by Adali and collaborators (Calhoun et al, 2002; Li et al, 2011; Rodriguez et al, 2012). Thus our methodology has the potential of taking advantage of the full complex-valued image information whilst also coping with the inherent non-uniform sampling.

***Ocean surface measurement devices.*** There is a long-standing history of studying ocean circulation using GPS-tracked ocean buoy drifters, see e.g. Osborne et al (1989). Since these trajectories are measured in the longitude-latitude plane, they are often converted to complex-valued vector series, see e.g. Sykulski et al (2017). It has long been observed that, due to the buffeting motion of ocean currents, positional drifter trajectories often exhibit fBM-like behaviour, whilst their velocity over time resemble fGn characteristics (Sanderson and Booth, 1991; Summers, 2002; Qu and Addison, 2010; Lilly et al, 2017). In this context, accurate Hurst exponent estimation is useful in indicating the intensity of ocean turbulence, giving evidence towards particular theorized dynamical regimes (Osborne et al, 1989). These in turn, can provide insight into initial conditions and origin of ocean circulation. Moreover, the trajectories often display rotary characteristics (Elipot and Lumpkin, 2008; Elipot et al, 2016). Due to the interrupted nature of satellite coverage and the possibility of measurements from multiple satellite orbits, the temporal sampling of the trajectories are typically highly nonuniform. In addition, due to the irregular sampling structure, the data are often interpolated prior to analysis (Elipot et al, 2016). One aspect of exploration in this setting could be to contrast Hurst estimation using our proposed methodology with/ without data interpolation to investigate its effect, since previous work substantiates that such processing can produce bias (in the context of Hurst exponent estimation) for real-valued series (Knight et al, 2017). It would also be interesting to investigate modifications to our technique to parameter estimation for Matérn processes discussed in Lilly et al (2017).

**Acknowledgements** The R package `CliftLRD` implementing the ℂLoMPE technique will be released via CRAN in due course.

## A Proofs and theoretical results

This appendix gives the theoretical justification of the results from Sections 3 and 4, following the notation outlined in the text.

### A.1 Proof of Proposition 1

To obtain the recursive construction for each basis, we start with the basis indexed by $i = 1$. At stage $n$, we have $f(x) = \sum_{k\in S_n} c_{n,k}\varphi^{(1)}_{n,k}(x)$ with $\varphi^{(1)}_{n,k}(x) = \chi_{I_{n,k}}(x)$ as proposed in the LOCAAT construction (Jansen et al, 2009).

Let us now suppose $f(x) := \varphi^{(1)}_{n-1,j}(x)$, thus $\varphi^{(1)}_{n-1,j}(x) = d^{(1)}_{j_n}\psi^{(1)}_{j_n}(x) + \sum_{k\in S_{n-1}} c_{n-1,k}\varphi^{(1)}_{n-1,k}(x)$. Hence $d^{(1)}_{j_n} = 0$, $c_{n-1,k} = 0, \forall k \neq j$ and $c_{n-1,j} = 1$. From the update relationship $c_{n-1,k} = c_{n,k} + b^n_k\lambda_{j_n}$ from (7), we have $c_{n-1,k} = c_{n,k}, \forall k \in J_n$ (as $\lambda_{j_n} = 0$ from $d^{(1)}_{j_n} = 0$) and also $c_{n-1,k} = c_{n,k}, \forall k \notin J_n$.

From equations (5) we have

$$d^{(1)}_{j_n} = \lambda_{j_n} + \mathrm{i}\,\mu_{j_n} = c_{n,j_n}\left(\ell^n_{j_n} + \mathrm{i}\,m^n_{j_n}\right) + \sum_{k\in J_n} c_{n,k}\left(\ell^n_k + \mathrm{i}\,m^n_k\right). \tag{26}$$

By denoting $a^n_k = \ell^n_k + \mathrm{i}\,m^n_k$, we obtain $d^{(1)}_{j_n} = c_{n,j_n}a^n_{j_n} - \sum_{k\in J_n} a^n_k c_{n,k}$. Using also the fact that $d^{(1)}_{j_n} = 0$, we have

$c_{n,j_n} = \frac{\overline{a}^n_{j_n}}{|a^n_{j_n}|^2} \sum_{k\in J_n} a^n_k c_{n,k}$. If $j \in J_n$ then $c_{n,j} = 1$ and all others are zero, so $c_{n,j_n} = \frac{\overline{a}^n_{j_n} a^n_j}{|a^n_{j_n}|^2} := \tilde{a}^n_j$. Thus

$$\varphi^{(1)}_{n-1,j}(x) = \varphi^{(1)}_{n,j}(x) + \tilde{a}^n_j \varphi^{(1)}_{n,j_n}(x), \text{ if } j \in J_n, \quad (27)$$

$$\varphi^{(1)}_{n-1,j}(x) = \varphi^{(1)}_{n,j}(x), \text{ if } j \notin J_n. \quad (28)$$

For the *primal wavelet function* construction, we can similarly take $f(x) := \psi^{(1)}_{j_n}(x)$, and obtain the corresponding wavelet decomposition with coefficients $d^{(1)}_{j_n} = 1$ (thus $\lambda_{j_n} = 1$ and $\mu_{j_n} = 0$) and $c_{n-1,k} = 0, \forall k \neq j_n$. From the update equations, we have $c_{n,j} = -b^n_j, \forall j \in J_n$ and $c_{n,j} = 0, \forall j \notin J_n$.

Using $d^{(1)}_{j_n} = c_{n,j_n} a^n_{j_n} - \sum_{j\in J_n} a^n_j c_{n,j}$ (as above) and $d^{(1)}_{j_n} = 1$, we have $c_{n,j_n} a^n_{j_n} = 1 - \sum_{j\in J_n} a^n_j b^n_j$ and $c_{n,j_n} = \frac{\overline{a}^n_{j_n}}{|a^n_{j_n}|^2}\left(1 - \sum_{j\in J_n} a^n_j b^n_j\right)$. Since $f(x) := \psi^{(1)}_{j_n}(x)$, we then have

$$\begin{aligned}\psi^{(1)}_{j_n}(x) &= \frac{\overline{a}^n_{j_n}}{|a^n_{j_n}|^2}\left(1 - \sum_{j\in J_n} a^n_j b^n_j\right)\varphi^{(1)}_{n,j_n}(x) - \sum_{j\in J_n} b^n_j \varphi^{(1)}_{n,j}(x) \\ &= \frac{\overline{a}^n_{j_n}}{|a^n_{j_n}|^2}\varphi^{(1)}_{n,j_n}(x) - \sum_{j\in J_n} b^n_j \left(\varphi^{(1)}_{n,j}(x) + \tilde{a}^n_j \varphi^{(1)}_{n,j_n}(x)\right).\end{aligned}$$

Using the primal scaling function construction in equation (27), we obtain an expression for the primal wavelet function

$$\psi^{(1)}_{j_n}(x) = \frac{\overline{a}^n_{j_n}}{|a^n_{j_n}|^2}\varphi^{(1)}_{n,j_n}(x) - \sum_{j\in J_n} b^n_j \varphi^{(1)}_{n-1,j}(x),$$

which demonstrates the recursive construction from stage $n$ to $n-1$ and concludes the proof for the primal wavelet and scaling function construction.

For the *dual* scaling functions, we use the update equations and the fact that $c_{r,j} =< f, \tilde{\varphi}_{r,j} >$ for any $r$, hence we have, at stage $n$,

$$< f, \tilde{\varphi}_{n-1,j} > = < f, \tilde{\varphi}_{n,j} > + b^n_j < f, \tilde{\psi}^L_{n,j} >, \forall j \in J_n$$
$$< f, \tilde{\varphi}_{n-1,j} > = < f, \tilde{\varphi}_{n,j} > \forall j \notin J_n,$$

where $\tilde{\psi}^L$ denotes the dual wavelet function corresponding to the **L**-filter only.

Thus the recursive relations for the dual scaling functions are

$$\tilde{\varphi}_{n-1,j}(x) = \tilde{\varphi}_{n,j}(x) + b^n_j \tilde{\psi}^L_{n,j}(x), \forall j \in J_n$$
$$\tilde{\varphi}_{n-1,j}(x) = \tilde{\varphi}_{n,j}(x), \forall j \notin J_n.$$

Similarly, since $d^{(1)}_{j_n} = c_{n,j_n} a^n_{j_n} - \sum_{j\in J_n} a^n_j c_{n,j}$, we have $< f, \tilde{\psi}^{(1)}_{j_n} > =< f, a^n_{j_n}\tilde{\varphi}_{n,j_n} - \sum_{j\in J_n} a^n_j \tilde{\varphi}_{n,j} >$ and we obtain the dual wavelet construction

$$\tilde{\psi}^{(1)}_{j_n} = a^n_{j_n}\tilde{\varphi}_{n,j_n}(x) - \sum_{j\in J_n} a^n_j \tilde{\varphi}_{n,j}(x).$$

These steps are subsequently re-iterated, and hence the same also holds for stage $r$.

In order to obtain the primal scaling function recursive construction corresponding to the *second* basis, we proceed in the same way as for the first basis and similarly obtain

$$\varphi^{(2)}_{n-1,j}(x) = \varphi^{(2)}_{n,j}(x) + \overline{\tilde{a}}^n_j \varphi^{(2)}_{n,j_n}(x), \text{ if } j \in J_n,$$
$$\varphi^{(2)}_{n-1,j}(x) = \varphi^{(2)}_{n,j}(x), \text{ if } j \notin J_n.$$

We obtain the primal wavelet equations in a similar manner to the previous development

$$\psi^{(2)}_{j_n}(x) = \frac{a^n_{j_n}}{|a^n_{j_n}|^2}\varphi^{(2)}_{n,j_n}(x) - \sum_{j\in J_n} b^n_j \varphi^{(2)}_{n-1,j}(x).$$

The above equations show that the primal scaling and wavelet functions corresponding to the second basis are the conjugates of the corresponding primal and wavelet functions under the first basis, respectively.

As already explained, the update step is the same for both bases and $c_{r,k} =< f, \tilde{\varphi}^{(1)}_{r,k} >=< f, \tilde{\varphi}^{(2)}_{r,k} >$, for all $r, k$ thus the dual scaling functions coincide under both bases ($\tilde{\varphi}^{(1)}_{r,k} = \tilde{\varphi}^{(2)}_{r,k}$).

For the dual wavelet function, following the same approach as above, we obtain

$$\tilde{\psi}^{(2)}_{j_n}(x) = \overline{a}^n_{j_n}\tilde{\varphi}_{n,j_n}(x) - \sum_{j\in J_n} \overline{a}^n_j \tilde{\varphi}_{n,j}(x).$$

This concludes the proof for the second basis. □

## A.2 Proof of Proposition 2

Let $\{X_t\}$ be a zero-mean complex-valued stationary long memory series with autocovariance $\gamma_X(\tau) \sim c_\gamma|\tau|^{-\beta}$. We note here that for improper processes of the type considered in Sykulski and Percival (2016), the pseudo-autocovariance has the same decay rate as the autocovariance ($r_X(\tau) \sim c_r|\tau|^{-\beta}$) while for proper processes, $r_X(\tau) = 0, \forall\tau$, hence we shall concentrate on the lifting decorrelation properties for improper processes.

The autocovariance of $\{X_t\}$ can be written as $\gamma_X(t_i - t_j) = \mathbb{E}(X_{t_i}\overline{X}_{t_j})$ and $r_X(t_i - t_j) = \mathbb{E}(X_{t_i}X_{t_j})$, assuming $\mathbb{E}(X_t) = 0$, where 0 is to be understood as the complex number $0 = 0 + \mathrm{i}\,0$. Hence, $\mathbb{E}(d^{(\ell)}_j) = 0$ for $\ell = 1, 2$. In what follows we drop the superscript $(\ell)$ in order to avoid notational clutter.

Using the assumption that $\mathbb{E}(d_j) = 0$ it follows that

$$\mathbb{E}(d_{j_r}\overline{d}_{j_k}) = \int_{\mathbb{R}} \overline{\tilde{\psi}}_{j_r}(t)\left\{\int_{\mathbb{R}} \tilde{\psi}_{j_k}(s)\gamma_X(t-s)\,ds\right\}dt, \quad (29)$$

where we have used $d_{j_r} =< X, \tilde{\psi}_{j_r} >$, and the timepoints $j_r$ and $j_k$ are distinct. In what follows, denote the interval length (i.e. continuous scale) of detail $d_{j_r}$ by $I_{r,j_r}$.

Since from (15) and (22), regardless of whether we work with the basis indexed by $\ell = 1$ or $\ell = 2$, the (dual) wavelet functions are linear combinations of the (same) dual scaling functions, hence equation (29) can be re-written as

$$\begin{aligned}\mathbb{E}(d_{j_r}\overline{d}_{j_k}) = \int_{\mathbb{R}} \overline{\left\{\tilde{\varphi}_{r,j_r}(t) - \sum_{i\in J_r} A^r_i \tilde{\varphi}_{r,i}(t)\right\}} \times \\ \int_{\mathbb{R}} \left\{\tilde{\varphi}_{k,j_k}(s) - \sum_{j\in J_k} A^k_j \tilde{\varphi}_{k,j}(s)\right\}\gamma_X(t-s)\,ds\,dt, \quad (30)\end{aligned}$$

where $A$ generically denotes the appropriate coefficient that corresponds to basis $\ell = 1$ or $\ell = 2$, but $\tilde{\varphi}$ is the same for both bases.

As $\mathbb{C}^2$-LOCAAT progresses, the (dual) scaling functions are defined recursively as linear combinations of (dual) scaling

functions at the previous stage, see e.g. equation (19). Hence the scaling functions in the above equation can be written as linear combinations of scaling functions at the first stage (i.e. $r = n$). Due to the linearity of the integral operator, (30) can be written as a linear combination with complex-valued coefficients of terms like

$$B_{n,i,j} := \int_{\mathbb{R}} \tilde{\varphi}_{n,i}(t) \left\{ \int_{\mathbb{R}} \tilde{\varphi}_{n,j}(s)\gamma_X(t-s)\, ds \right\} dt = \int_{\mathbb{R}} \tilde{\varphi}_{n,i}(t) \left(\tilde{\varphi}_{n,j} \star \gamma_X\right)(t)\, dt, \quad (31)$$

where $\star$ is the convolution operator, and $i$ and $j$ refer to time locations that were involved in obtaining $d_{j_r}$ and $d_{j_k}$. Note that at this stage we do not use complex conjugation as the (dual) scaling functions are initially defined (at stage $r = n$) as scaled characteristic functions of the intervals associated with the observed times, i.e. $\tilde{\varphi}_{n,i}(t) = I_{n,i}^{-1}\chi_{I_{n,i}}(t)$ (thus real-valued).

Using Parseval's theorem in equation (31) gives

$$B_{n,i,j} = (2\pi)^{-1} \int_{\mathbb{R}} \hat{\tilde{\varphi}}_{n,i}(\omega) \left(\widehat{\tilde{\varphi}_{n,j} \star \gamma_X}\right)(\omega)\, d\omega = (2\pi)^{-1} \int_{\mathbb{R}} \hat{\tilde{\varphi}}_{n,i}(\omega)\overline{\hat{\tilde{\varphi}}_{n,j}}(\omega)\overline{f_X}(\omega)\, d\omega, \quad (32)$$

where in general $\hat{g}$ denotes the Fourier transform of $g$. As the Fourier transform of an initial (dual) scaling function (scaled characteristic function on an interval, $(b-a)^{-1}\chi_{[a,b]}$) is

$$\left\{\widehat{(b-a)^{-1}\chi_{[a,b]}}\right\}(\omega) = \operatorname{sinc}\left\{\omega(b-a)/2\right\} \exp\left\{-\mathrm{i}\,\omega(b+a)/2\right\},$$

where $\operatorname{sinc}(x) = x^{-1}\sin(x)$ for $x \neq 0$ and $\operatorname{sinc}(0) = 1$ is the (unnormalized) sinc function, we can write (32) as

$$\int_{\mathbb{R}} \operatorname{sinc}\left(\omega I_{n,i}/2\right) \operatorname{sinc}\left(\omega I_{n,j}/2\right) \exp\left\{-\mathrm{i}\,\omega\delta(I_{n,i}, I_{n,j})\right\} f_X(\omega) d\omega, \quad (33)$$

where $\delta(I_{n,i}, I_{n,j})$ is the distance between the midpoints of intervals $I_{n,i}$ and $I_{n,j}$ at the initial stage $n$.

Equation (33) can be interpreted as the Fourier transform of $u(x) = f_X(x)\operatorname{sinc}\left(xI_{n,i}/2\right)\operatorname{sinc}\left(xI_{n,j}/2\right)$ evaluated at $\delta(I_{n,i}, I_{n,j})$.

Since the sinc function is infinitely differentiable and the spectrum is Lipschitz continuous, results on the decay properties of Fourier transforms (Shibata and Shimizu, 2001, Theorem 2.2) imply that, for $i \neq j$, terms of the form $B_{n,i,j}$ decay as $O\left\{\delta(I_{n,i}, I_{n,j})^{-1}\right\}$. Hence as in Knight et al (2017), the further away the time points are, the less autocorrelation is present in the detail coefficients and as the rate of autocorrelation decay is of reciprocal order, it is faster than that of the original process assumed to have long memory (hence $O(|\tau|^{-\beta})$ with $\beta \in (0,1)$).

A similar argument as above applies for the pseudo-covariance $r_X(t_i - t_j) = \mathbb{E}(X_{t_i}X_{t_j})$, as

$$\mathbb{E}(d_{j_r}d_{j_k}) = \int_{\mathbb{R}} \overline{\tilde{\psi}}_{j_r}(t) \left\{ \int_{\mathbb{R}} \overline{\tilde{\psi}}_{j_k}(s) r_X(t-s)\, ds \right\} dt, \quad (34)$$

and concludes the proof. □

## A.3 Proof of Proposition 3

As $\operatorname{Cov}(X_{t_i}, X_{t_j}) = \gamma_X(t_i - t_j)$ and $d_{j_r} =< X, \tilde{\psi}_{j_r} >$, it follows that $d_{j_r}$ has mean zero (as the original process is zero-mean) and in a similar manner to (29) we have

$$\mathbb{E}(|d_{j_r}|^2) = \int_{\mathbb{R}} \overline{\tilde{\psi}}_{j_r}(t) \left\{ \int_{\mathbb{R}} \tilde{\psi}_{j_r}(s)\gamma_X(t-s)\, ds \right\} dt, \quad (35)$$

where again we have dropped the basis index $\ell = 1, 2$ for notational brevity and we remind the reader that $|\cdot|$ denotes the $\mathbb{C}$-modulus. As before, we denote the associated interval length of the detail $d_{j_r}$ by $I_{r,j_r}$.

Using the recursiveness in the dual wavelet construction (equations (15) and (22)), it follows that the (dual) wavelet functions are linear combinations of the (same) scaling functions. For the first basis, equation (35) can be re-written as

$$\mathbb{E}(|d_{j_r}^{(1)}|^2) = \int_{\mathbb{R}} \left\{ \overline{a}_{j_r}^r \overline{\tilde{\varphi}}_{r,j_r}(t) - \sum_{j \in J_r} \overline{a}_j^r \overline{\tilde{\varphi}}_{r,j}(t) \right\} \times \int_{\mathbb{R}} \left\{ a_{j_r}^r \tilde{\varphi}_{r,j_r}(s) - \sum_{j' \in J_r} a_{j'}^r \tilde{\varphi}_{r,j'}(s) \right\} \gamma_X(t-s)\, ds\, dt. \quad (36)$$

This can be expanded as

$$\begin{aligned}\mathbb{E}(|d_{j_r}^{(1)}|^2) = &\int_{\mathbb{R}}\int_{\mathbb{R}} a_{j_r}^r \overline{a}_{j_r}^r \overline{\tilde{\varphi}}_{r,j_r}(t)\tilde{\varphi}_{r,j_r}(s)\gamma_X(t-s)\, ds\, dt \\ &- \sum_{j \in J_r} \int_{\mathbb{R}}\int_{\mathbb{R}} \overline{a}_j^r a_{j_r}^r \overline{\tilde{\varphi}}_{r,j}(t)\tilde{\varphi}_{r,j_r}(s)\gamma_X(t-s)\, ds\, dt \\ &- \sum_{j' \in J_r} \int_{\mathbb{R}}\int_{\mathbb{R}} \overline{a}_{j_r}^r a_{j'}^r \overline{\tilde{\varphi}}_{r,j_r}(t)\tilde{\varphi}_{r,j'}(s)\gamma_X(t-s)\, ds\, dt \\ &+ \sum_{j \in J_r}\sum_{j' \in J_r} \int_{\mathbb{R}}\int_{\mathbb{R}} \overline{a}_j^r a_{j'}^r \overline{\tilde{\varphi}}_{r,j}(t)\tilde{\varphi}_{r,j'}(s)\gamma_X(t-s)\, ds\, dt. \end{aligned} \quad (37)$$

As in Proposition 1, using Parseval's theorem we obtain that the above is a linear combination of terms of the form

$$B_{r,i,j} = \int_{\mathbb{R}} \tilde{\chi}_{r,i}(t) \left\{ \int_{\mathbb{R}} \tilde{\chi}_{r,j}(s)\gamma_X(t-s) ds \right\} dt = \int_{\mathbb{R}} \operatorname{sinc}(\omega I_{r,i}/2)\operatorname{sinc}(\omega I_{r,j}/2) e^{-\mathrm{i}\,\omega\delta(I_{r,i}, I_{r,j})} f_X(\omega) d\omega, \quad (38)$$

where recall that the hat notation denotes the Fourier transform of a function and $\delta(I_{r,i}, I_{r,j})$ denotes the distance between the midpoints of intervals $I_{r,i}$ and $I_{r,j}$.

Due to the artificial level construction, the sequence of lifting integrals is approximately log-linear in the artificial level (see Knight et al (2017) for details), i.e. for those points $j_r$ in the $j^\star$th artificial level, we have $\log_2(I_{r,j_r}) = j^\star + \Delta$ where $\Delta \in \{-1 + \log_2(\alpha_0), \log_2(\alpha_0)\}$ for some $\alpha_0$, thus $I_{r,j_r} = R2^{j^\star}$ for some constant $R > 0$.

Now suppose $i = j$ and both points belong to the $j^\star$th artificial level. In equation (38) we make a change of variable $\eta = \omega R 2^{j^\star}$ to obtain

$$\begin{aligned} B_{r,i,i} &= \int_{\mathbb{R}} \operatorname{sinc}^2(\eta/2) f_X(\eta/R2^{j^\star}) \left(R2^{j^\star}\right)^{-1} d\eta \\ &\sim \int_{\mathbb{R}} \operatorname{sinc}^2(\eta/2) c_f |\eta|^{-\alpha} \left(R2^{j^\star}\right)^{\alpha-1} d\eta, \qquad (j^\star \to \infty) \\ &= 2^{j^\star(\alpha-1)} \int_{\mathbb{R}} c_f R^{\alpha-1} \operatorname{sinc}^2(\eta/2)|\eta|^{-\alpha} d\eta \\ &= 2^{j^\star(\alpha-1)} R^{\alpha-1} 4 c_f \Gamma(-1-\alpha)\sin(\pi\alpha/2) \quad (39) \\ &= 2^{j^\star(\alpha-1)} R^{\alpha-1} M, \quad (40) \end{aligned}$$

where $\alpha \in (0,1)$, $\Gamma$ is the Gamma function and $M = 4c_f\Gamma(-1-\alpha)\sin(\pi\alpha/2)$.

If $i \neq j$ are points from the same neighbourhood $J_r$ and both belong to the same artificial level $j^\star$, then their artificial scale measure will be the same. Performing the same change of variable as above, we obtain (as $(j^\star \to \infty)$

$$\begin{aligned} B_{r,i,j} &\sim \int_{\mathbb{R}} \mathrm{sinc}^2(\eta/2) e^{-\mathrm{i}\,\eta} \left(R2^{j^\star}\right)^{-1} c_f |\eta|^{-\alpha} \left(R2^{j^\star}\right)^{\alpha} d\eta, \\ &= 2^{j^\star(\alpha-1)} c_f R^{\alpha-1} \int_{\mathbb{R}} \mathrm{sinc}^2(\eta/2) e^{-\mathrm{i}\,\eta} |\eta|^{-\alpha} d\eta \\ &= 2^{j^\star(\alpha-1)} R^{\alpha-1} 4c_f (2^\alpha - 1)\sin(\pi\alpha/2)\Gamma(1-\alpha) \quad (41) \\ &= 2^{j^\star(\alpha-1)} R^{\alpha-1} N, \quad (42) \end{aligned}$$

where $N = 4c_f(2^\alpha - 1)\sin(\pi\alpha/2)\Gamma(1-\alpha)$.

All terms in (37) involve points from the same neighbourhood $J_r$, and thus using (40) and (42) together with the linearity of the integral operator, we have that

$$\begin{aligned} &\mathbb{E}(|d^{(1)}_{j_r}|^2) \sim 2^{j^\star(\alpha-1)} R^{\alpha-1} \times \\ &\left( M^2 |a^r_{j_r}|^2 + N^2 \sum_j \sum_{j'} \overline{a}^r_j a^r_{j'} - MN \sum_j \overline{a}^r_j a^r_{j_r} - MN \sum_{j'} \overline{a}^r_{j_r} a^r_{j'} \right) \\ &\quad = 2^{j^\star(\alpha-1)} R^{\alpha-1} |M\, a^r_{j_r} - N \sum_{j \in J_r} a^r_j|^2 \\ &\quad = C\, 2^{j^\star(\alpha-1)}, \end{aligned}$$

where $C$ is a constant depending on $c_f, R$ and $\alpha$.

A similar argument applies to the the second basis and completes the proof. □

## References


Abry P, Flandrin P, Taqqu MS, Veitch D (2000) Wavelets for the analysis, estimation and synthesis of scaling data. In: Park K, Willinger W (eds) Self-similar Network Traffic and Performance Evaluation, Wiley, Chichester, pp 39–88

Abry P, Goncalves P, Véhel JL (2013) Scaling, Fractals and Wavelets. Wiley, New York

Adali T, Schreier PJ, Scharf LL (2011) Complex-valued signal processing: The proper way to deal with impropriety. IEEE Trans Sig Proc 59(11):5101–5125

Adrian DW, Maitra R, Rowe DB (2017) Complex-valued time series modeling for improved activation detection in fMRI studies. Ann Appl Stat (to appear)

Amblard PO, Coeurjolly JF, Lavancier F, Philippe A (2012) Basic properties of the multivariate fractional Brownian motion. Sém & Congr 28:65–87

Aston JAD, Kirch C (2012) Evaluating stationarity via change-point alternatives with applications to fMRI data. Ann of Appl Stat 6(4):1906–1948

Bakker AMR, van den Hurk BJJM (2012) Estimation of persistence and trends in geostrophic wind speed for the assessment of wind energy yields in Northwest Europe. Clim Dynam 39(3-4):767–782

Barber S, Nason GP (2004) Real nonparametric regression using complex wavelets. Journal of the Royal Statistical Society B 66(4):927–939

Beran J, Feng Y, Ghosh S, Kulik R (2013) Long-memory Processes. Springer, New York

Bhattacharya RN, Gupta VK, Waymire E (1983) The Hurst effect under trends. J Appl Prob 20:649–662

Broersen PMT (2007) Time series models for spectral analysis of irregular data far beyond the mean data rate. Meas Sci Technol 19(1):1–13

Bullmore E, Fadili J, Breakspear M, Salvador R, Suckling J, Brammer M (2003) Wavelets and statistical analysis of functional magnetic resonance images of the human brain. Stat Meth Med Res 12(5):375–399

Calhoun VD, Adalı T, Pearlson GD, Van Zijl PCM, Pekar JJ (2002) Independent component analysis of fMRI data in the complex domain. Magn Reson Med 48(1):180–192

Chandna S, Walden AT (2017) A frequency domain test for propriety of complex-valued vector time series. IEEE Trans Sig Proc 65(6):1425–1436

Chang TP, Ko HH, Liu FJ, Chen PH, Chang YP, Liang YH, Jang HY, Lin TC, Chen YH (2012) Fractal dimension of wind speed time series. Applied Energy 93:742–749

Churchill NW, Spring R, Grady C, Cimprich B, Askren MK, Reuter-Lorenz PA, Jung MS, Peltier S, Strother SC, Berman MG (2016) The suppression of scale-free fMRI brain dynamics across three different sources of effort: aging, task novelty and task difficulty. Nature Sci Rep 6:30,895

Ciuciu P, Varoquaux G, Abry P, Sadaghiani S, Kleinschmidt A (2012) Scale-free and multifractal time dynamics of fMRI signals during rest and task. Front Physiol 3

Coeurjolly JF, Porcu E (2017a) Fast and exact simulation of complex-valued stationary Gaussian processes through embedding circulant matrix. J Comput Graph Stat (to appear)

Coeurjolly JF, Porcu E (2017b) Properties and Hurst exponent estimation of the circularly-symmetric fractional Brownian motion. Stat Prob Lett 128:21 – 27

Coeurjolly JF, Lee K, Vidakovic B (2014) Variance estimation for fractional Brownian motions with fixed Hurst parameters. Commun Stat – Theory and Meth 43(8):1845–1858

Craigmile PF, Percival DB (2005) Asymptotic decorrelation of between-scale wavelet coefficients. Trans Im Proc 51(3):1039–1048

Craigmile PF, Percival DB, Guttorp P (2001) The impact of wavelet coefficient correlations on fractionally differenced process estimation. In: Casacuberta C, Miró-Roig RM, Verdera J, Xambó-Descamps S (eds) European Congress of Mathematics, Birkhäuser, Basel, pp 591–599

Curtis TE (1985) Digital signal processing for sonar. In: Adaptive Methods in Underwater Acoustics, Springer, pp 583–605

Didier G, Pipiras V (2011) Integral representations and properties of operator fractional Brownian motions. Bernoulli 17(1):1–33

Dowell J, Weiss S, Infield D, Chandna S (2014) A widely linear multichannel wiener filter for wind prediction. In: IEEE Work. Stat. Sig. Proc. (SSP) 2014, IEEE, pp 29–32

Elipot S, Lumpkin R (2008) Spectral description of oceanic near-surface variability. Geophys Res Lett 35(5)

Elipot S, Lumpkin R, Perez RC, Lilly JM, Early JJ, Sykulski AM (2016) A global surface drifter data set at hourly resolution. J Geophys Res: Oceans 121(5):2937–2966

Ferdowsi S, Abolghasemi V (2017) Simultaneous BOLD detection and incomplete fMRI data reconstruction. Med & Biol Eng & Comput (to appear)

Flandrin P (1998) Time-frequency/Time-scale Analysis. Academic Press, San Diego

Fortuna L, Nunnari S, Guariso G (2014) Fractal order evidences in wind speed time series. In: Intern. Conf. Fract. Diff Appl. (ICFDA) 2014, IEEE, pp 1–6

Giraitis L, Robinson PM, Surgailis D (1999) Variance-type estimation of long memory. Stoch Proc Appl 80(1):1–24

Goh SL, Chen M, Popović DH, Aihara K, Obradovic D, Mandic DP (2006) Complex-valued forecasting of wind profile. Renew Energy 31(11):1733–1750

Gonella J (1972) A rotary-component method for analysing meteorological and oceanographic vector time series. In: Deep Sea Research and Oceanographic Abstracts, Elsevier, vol 19, pp 833–846

Hamilton J, Nunes MA, Knight MI, Fryzlewicz P (2017) Complex-valued wavelet lifting and applications. Technom 60(1):46–60

Haslett J, Raftery AE (1989) Space-time modelling with long-memory dependence: Assessing Ireland's wind power resource. Appl Stat pp 1–50

Higuchi T (1990) Relationship between the fractal dimension and the power law index for a time series: a numerical investigation. Physica D 46(2):254–264

Hsu NJ (2006) Long-memory wavelet models. Stat Sin 16:1255–1271

Hurst HE (1951) Long-term storage capacity of reservoirs. Trans Amer Soc Civ Eng 116:770–808

Jansen M, Oonincx P (2005) Second Generation Wavelets and Applications. Springer Verlag

Jansen M, Nason GP, Silverman BW (2001) Scattered data smoothing by empirical Bayesian shrinkage of second generation wavelet coefficients. In: Unser M, Aldroubi A (eds) Wavelet Applications in Signal and Image Processing IX, SPIE, vol 4478, pp 87–97

Jansen M, Nason GP, Silverman BW (2009) Multiscale methods for data on graphs and irregular multidimensional situations. J Roy Stat Soc B 71(1):97–125

Jensen MJ (1999) Using wavelets to obtain a consistent ordinary least squares estimator of the long-memory parameter. J Forecast 18(1):17–32

Jeon S, Nicolis O, Vidakovic B (2014) Mammogram diagnostics via 2-D complex wavelet-based self-similarity measures. The São Paulo J Math Sci 8(2):265–284

Jung YY, Park Y, Jones DP, Ziegler TR, Vidakovic B (2010) Self-similarity in NMR spectra: an application in assessing the level of cysteine. J Data Sci 8(1):1

Junger WL, Ponce de Leon A (2015) Imputation of missing data in time series for air pollutants. Atmos Envir 102:96–104

Kingsbury N (2001) Complex wavelets for shift invariant analysis and filtering of signals. Appl Comput Harm Anal 10(3):234–253

Knight MI, Nason GP (2009) A nondecimated lifting transform. Stat Comput 19(1):1–16

Knight MI, Nunes MA (2016) liftLRD: Wavelet lifting estimators for the Hurst exponent for regular and irregular time series. R package version 1.0-5

Knight MI, Nunes MA, Nason GP (2012) Spectral estimation for locally stationary time series with missing observations. Statistics and Computing 22(4):877–8951

Knight MI, Nason GP, Nunes MA (2017) A wavelet lifting approach to long-memory estimation. Stat Comput 27(6):1453–1471

Li H, Correa NM, Rodriguez PA, Calhoun VD, Adali T (2011) Application of independent component analysis with adaptive density model to complex-valued fMRI data. IEEE Trans Biomed Eng 58(10):2794–2803

Lilly JM, Gascard JC (2006) Wavelet ridge diagnosis of time-varying elliptical signals with application to an oceanic eddy. Nonlin Proc Geophys 13(5):467

Lilly JM, Sykulski AM, Early JJ, Olhede SC (2017) Fractional Brownian motion, the Matérn process, and stochastic modeling of turbulent dispersion. Nonlin Proc Geophys 24:481–514

Lindquist MA (2008) The statistical analysis of fMRI data. Stat Sci pp 439–464

Lobato I, Robinson PM (1996) Averaged periodogram estimation of long memory. J Econometr 73(1):303–324

Mandelbrot BB, Taqqu MS (1979) Robust R/S analysis of long-run serial correlation. Bull Intern Stat Inst 48(2):59–104

Mandelbrot BB, Van Ness JW (1968) Fractional Brownian motions, fractional noises and applications. SIAM Rev 10(4):422–437

Mandic DP, Goh VSL (2009) Complex valued nonlinear adaptive filters: noncircularity, widely linear and neural models, vol 59. John Wiley & Sons

Mandic DP, Javidi S, Goh SL, Kuh A, Aihara K (2009) Complex-valued prediction of wind profile using augmented complex statistics. Renew Energy 34(1):196–201

Martin KW (2004) Complex signal processing is not complex. IEEE Trans Circ Sys I: Regular Papers 51(9):1823–1836

McCoy EJ, Walden AT (1996) Wavelet analysis and synthesis of stationary long-memory processes. J Graph Comput Stat 5(1):26–56

Mohammadi A, Plataniotis KN (2015) Improper complex-valued multiple-model adaptive estimation. IEEE Trans Sig Proc 63(6):1528–1542

Nafornita C, Isar A, Nelson JDB (2014) Regularised, semi-local hurst estimation via generalised lasso and dual-tree complex wavelets. In: IEEE Intern. Conf. Im. Proc. (ICIP) 2014, IEEE, pp 2689–2693

Neeser FD, Massey JL (1993) Proper complex random processes with applications to information theory. IEEE Transactions on Information Theory 39(4):1293–1302

Nelson JDB, Kingsbury NG (2010) Dual-tree wavelets for estimation of locally varying and anisotropic fractal dimension. In: 17th IEEE Intern. Conf. Im. Proc. (ICIP) 2010, IEEE, pp 341–344

Nunes MA, Knight MI (2017) CNLTreg: Complex-Valued Wavelet Lifting for Signal Denoising. R package version 0.1

Nunes MA, Knight MI, Nason GP (2006) Adaptive lifting for nonparametric regression. Stat Comput 16(2):143–159

Olhede SC, Walden AT (2005) Local directional denoising. IEEE Trans Sig Proc 53(12):4725–4730

Osborne AR, Kirwan Jr AD, Provenzale A, Bergamasco L (1989) Fractal drifter trajectories in the Kuroshio extension. Tellus A: Dyn Meteo Ocean 41(5):416–435

Park C, Lazar NA, Ahn J, Sornborger A (2010) A multiscale analysis of the temporal characteristics of resting-state fMRI data. J Neurosci Meth 193(2):334–342

Peng CK, Buldyrev SV, Havlin S, Simons M, Stanley HE, Goldberger AL (1994) Mosaic organization of DNA nucleotides. Phys Rev E 49(2):1685

Piacquadio M, de la Barra A (2014) Multifractal analysis of wind velocity data. Energy Sustain Devel 22:48–56

Picinbono B (1994) On circularity. IEEE Trans Sig Proc 42(12):3473–3482

Qu B, Addison PS (2010) Modelling flow trajectories using fractional brownian motion. In: Intern. Work. Chaos-Fractals Theories Appl. (IWCFTA), 2010, IEEE, pp 420–424

R Core Team (2013) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, URL `http://www.R-project.org`

Rehman S, Siddiqi AH (2009) Wavelet based Hurst exponent and fractal dimensional analysis of Saudi climatic dynamics. Chaos Solitons & Fractals 40(3):1081–1090

Rodriguez PA, Calhoun VD, Adalı T (2012) De-noising, phase ambiguity correction and visualization techniques for complex-valued ICA of group fMRI data. Patt Recogn 45(6):2050–2063

Rowe DB (2005) Modeling both the magnitude and phase of complex-valued fMRI data. Neuroim 25(4):1310–1324

Rowe DB (2009) Magnitude and phase signal detection in complex-valued fMRI data. Magn Reson Med 62(5):1356–1357

Rubin-Delanchy P, Walden AT (2008) Kinematics of complex-valued time series. IEEE Trans Sig Proc 56(9):4189–4198

Sanderson BG, Booth DA (1991) The fractal dimension of drifter trajectories and estimates of horizontal eddy-diffusivity. Tellus A: Dyn Meteo Ocean 43(5):334–349

Schreier PJ, Scharf LL (2003) Second-order analysis of improper complex random vectors and processes. IEEE Trans Sig Proc 51(3):714–725

Schreier PJ, Scharf LL (2010) Statistical signal processing of complex-valued data: the theory of improper and noncircular signals. Cambridge University Press

Selesnick I, Baraniuk R, Kingsbury N (2005) The dual-tree complex wavelet transform. IEEE Sig Proc Mag 22(6):123–151

Shi B, Vidakovic B, Katul GG, Albertson JD (2005) Assessing the effects of atmospheric stability on the fine structure of surface layer turbulence using local and global multiscale approaches. Phys Fluids 17(5):055,104

Shibata Y, Shimizu S (2001) A decay property of the fourier transform and its application to the stokes problem. J Math Fluid Mech 3(3):213–230

Summers DM (2002) Impulse exchange at the surface of the ocean and the fractal dimension of drifter trajectories. Nonlin Proc Geophys 9(1):11–23

Sweldens W (1995) The lifting scheme: A new philosophy in biorthogonal wavelet construction. In: Laine A, Unser M (eds) Wavelet Applications in Signal and Image Processing III, Proc. SPIE 2569, pp 68–79

Sykulski AM, Percival DB (2016) Exact simulation of noncircular or improper complex-valued stationary Gaussian processes using circulant embedding. In: IEEE 26th Intern. Work. Mach. Learn. Sig. Proc. (MLSP) 2016, pp 1–6

Sykulski AM, Olhede SC, Lilly JM, Early JJ (2017) Frequency-domain stochastic modeling of stationary bivariate or complex-valued signals. IEEE Trans Sig Proc 65(12):3136–3151

Tanaka T, Mandic DP (2007) Complex empirical mode decomposition. IEEE Sig Proc Lett 14(2):101–104

Taqqu MS, Teverovsky V, Willinger W (1995) Estimators for long-range dependence: an empirical study. Fractals 3(04):785–798

Trappe W, Liu K (2000) Denoising via adaptive lifting schemes. In: Proceedings of SPIE, Wavelet applications in signal and image processing VIII, A, Aldroubi, A., Laine, M.A. and Unser, M.A. (eds), vol 4119, pp 302–312

Vidakovic BD, Katul GG, Albertson JD (2000) Multiscale denoising of self-similar processes. J Geophys Res 105(D22):27,049–27,058

Walden AT (2013) Rotary components, random ellipses and polarization: a statistical perspective. Phil Trans R Soc A 371(1984):20110,554

Whitcher B, Jensen MJ (2000) Wavelet estimation of a local long memory parameter. Explor Geophys 31:94–103

Wuertz D, et al (2013) fARMA: ARMA Time Series Modelling. URL `http://CRAN.R-project.org/package=fArma`, R package version 3010.79

Zhang Q, Harman CJ, Ball WP (2014) Evaluation of methods for estimating long-range dependence (LRD) in water quality time series with missing data and irregular sampling. In: Proc. Amer. Geophys. Union Fall Meet. 2014, San Francisco, CA