This is a repository copy of *Long memory estimation for complex-valued time series*.

Article:

Knight, Marina Iuliana orcid.org/0000-0001-9926-6092 and Nunes, Matthew A. (2018)

CrossMark

Long memory estimation for complex-valued time series

Marina I. Knight[1] · Matthew A. Nunes[2]

Abstract

Long memory has been observed for time series across a multitude of fields, and the accurate estimation of such dependence, for example 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.

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, for example, Mandic and Goh (2009) and Sykulski et al. (2017). A particular modelling

✉ Matthew A. Nunes
   m.nunes@lancaster.ac.uk

   Marina I. Knight
   Marina.Knight@york.ac.uk

[1] Department of Mathematics, University of York, Heslington, York YO10 5DD, UK

[2] Department of Mathematics and Statistics, Fylde College, Lancaster University, Lancaster LA1 4YF, UK

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, for example, Schreier and Scharf (2003), Rubin-Delanchy and Walden (2008) and Mohammadi and Plataniotis (2015). Furthermore, complex-valued analysis of *real-valued* data has been shown to be beneficial in a number of settings; see, for example, Olhede and Walden (2005) and 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 (2017) 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 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. Fig. 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 Sect. 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, for example, Haslett and Raftery (1989), Chang et al. (2012) and 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 5 min scale considered here, have been shown to be in the range 0.7–0.9, indicating strong long-range dependence; see, for example, 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, for example, Goh et al. (2006) and 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 5 min 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 15 April 2017 until 30 April 2017. Whilst the sampling interval for the measurements is reported as 5 min, 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 min.

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 Fig. 1a, 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, for example, Tanaka and Mandic (2007) and 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 Fig. 2a, 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 periods, specifically, 30 April 2017 until 14 May 2017. For this dataset, the majority of the data are observed at a spacing of 5 min, but a significant amount have intra-measurement sampling between 10 and 20 min 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 Fig. 1c, 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 (Fig. 2c, 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 (Fig. 3). The series are not symmetric, exhibiting clear noncircularity, suggesting a model which allows for impropriety is appropriate for analy-
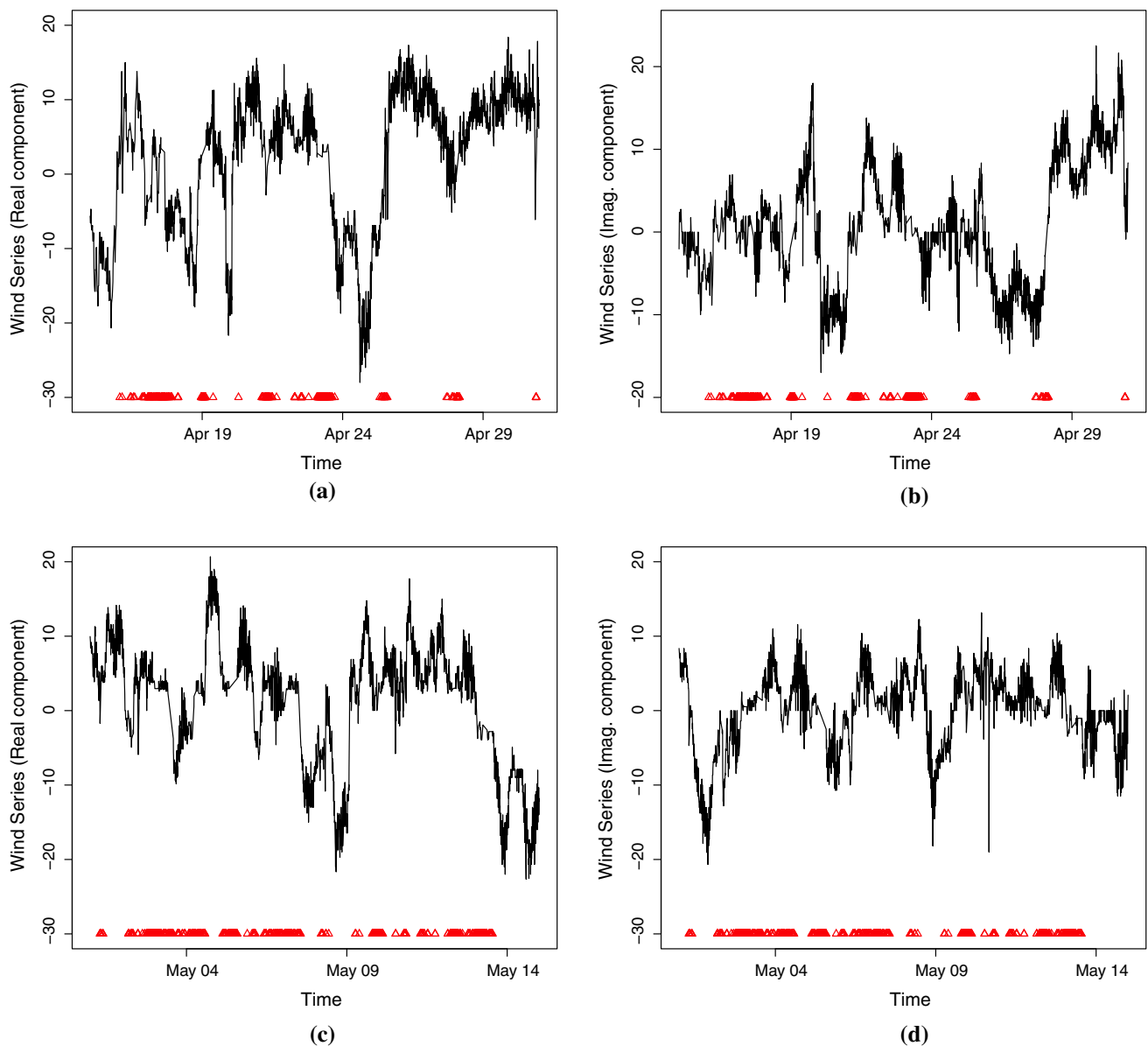
Fig. 1 **a** Real component of the `Wind A` data series; **b** imaginary component of the `Wind A` data series; **c** real component of the `Wind B` data series; **d** imaginary component of the `Wind B` data series. Red triangles indicate missing data locations. (Color figure online)

sis [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 Sect. 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 solution 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
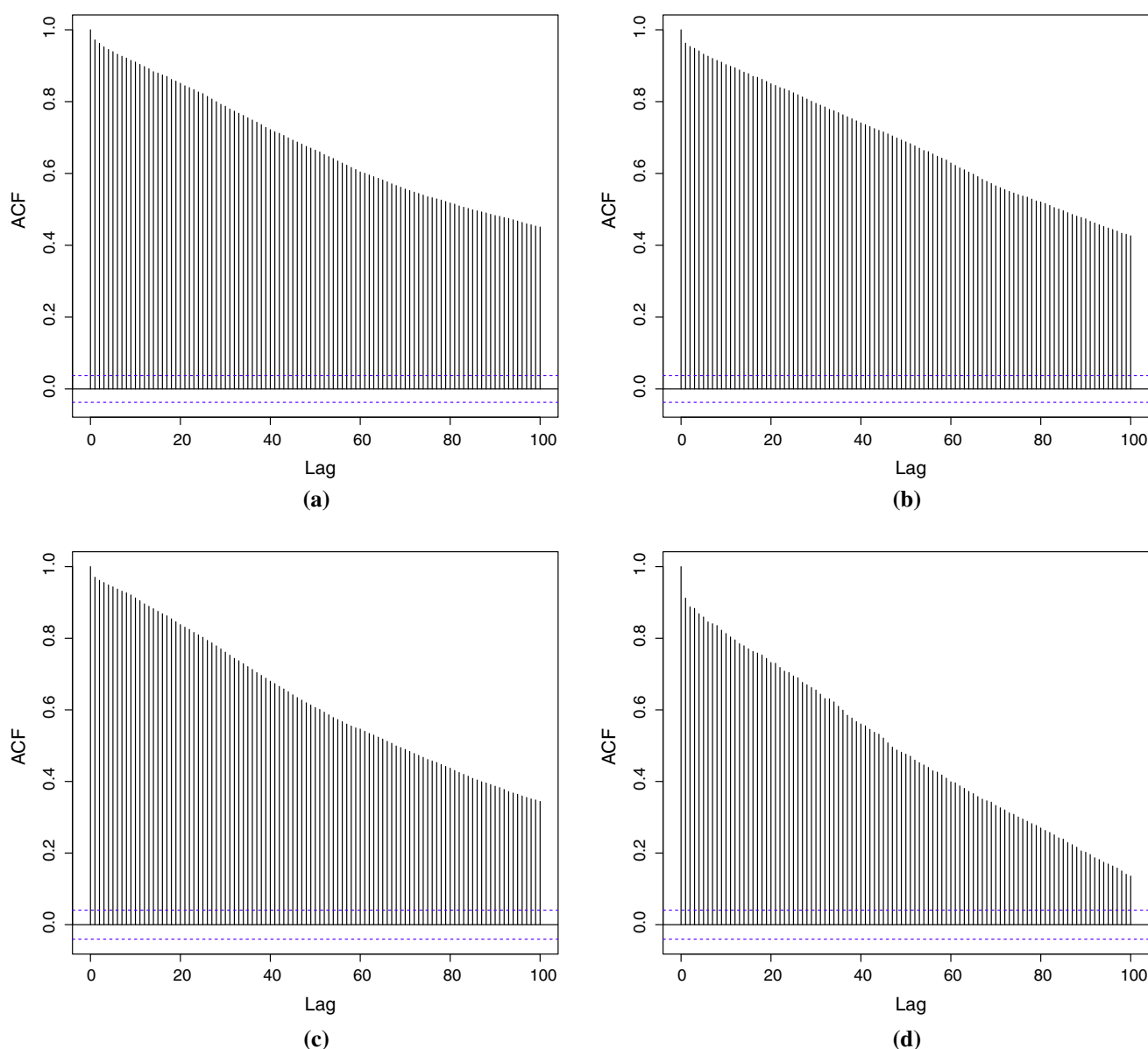
Fig. 2 **a** Autocorrelation for (**a**) the real component of the `Wind A` series from Fig. 1; **b** the imaginary component of the `Wind A` series; **c** the real component of the `Wind B` series from Fig. 1; and **d** the imagi-nary component of the `Wind B` series (all treated as regularly spaced). Both components of the two datasets show autocorrelation at large lags, indicating persistent behaviour

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 (2017) 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 choice is twofold: (1) (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 (2) for irregularly sampled (real-valued) processes, or those featuring missing-ness, the wavelet lifting algorithm of Knight et al. (2017) has provided a first long memory estimation solution and was
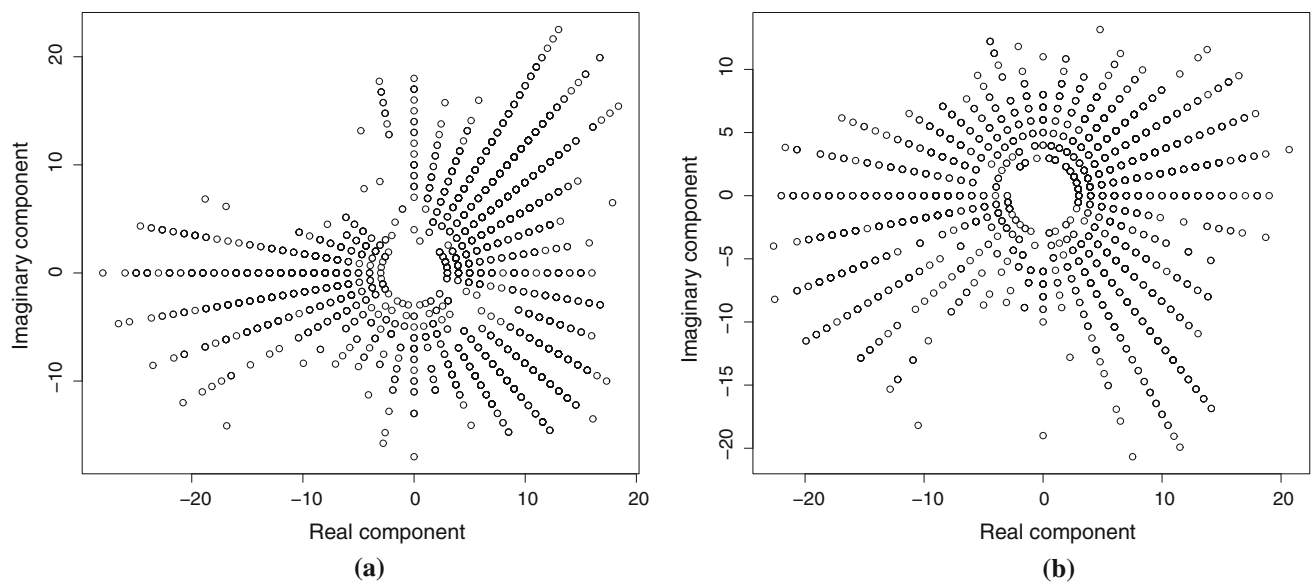
Fig. 3 Scatter plot of real and imaginary series values for **a** the `Wind A` data and **b** the `Wind B` series shown in Fig. 1. Both series exhibit noncircular (improper) characteristics

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 organized as follows. We begin, in Sect. 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 Sect. 5.2, we consider the application of our approach to the wind series datasets introduced in Sect. 1.1, discussing the potential consequences of our analysis. Finally, Sect. 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 $\overline{\cdot p}$ complex conjugation. As the autocovariance function γ_X does not completely characterize 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 $r_X(\cdot p) = 0$ (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 γ_X. In contrast, applications such as those described in Schreier and Scharf (2010), Adali et al. (2011) and Chandna and Walden (2017) deal with improper processes, whereby there exists a lag τ 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, for example Beran et al. (2013). Mandelbrot and 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, for example, Coeurjolly and Porcu (2017) and 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 2017). 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 ω, a constant A and $\delta \in (1/2, 3/2)$. The spectral slope parameter δ 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 γ_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, for example 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) and 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) and Nafornita et al. (2014). Reviews comparing several techniques for Hurst exponent estimation (for real-valued series) can be found in, for example, 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. Whilst this approach serves well when the process is real-valued, it cannot cope with complex-valued processes. Coeurjolly and Porcu (2017) 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, for example, Trappe and Liu (2000), Nunes et al. (2006), Knight et al. (2012), Knight et al. (2017) and 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), Jansen et al. (2009), we shall briefly introduce it next.

Suppose a real-valued function $f(\cdot p)$ 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 with 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:

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

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 with 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 are 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^ℓ 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, in a uniform, but even more so in a non-uniform sampling setting, and 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 (2017).

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 $f(\cdot p)$ 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 reiterates 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 is 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, **L** and **M**, orthogonal on each other. The filter **L** corresponds to the (possibly) linear regression choice as is usual in LOCAAT. The filter **M** is linked to **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

$$\lambda_{j_n} = l_{j_n}^n c_{n,j_n} - \sum_{i \in J_n} l_i^n c_{n,i}, \tag{3}$$

$$\mu_{j_n} = m_{j_n}^n c_{n,j_n} - \sum_{i \in J_n} m_i^n c_{n,i}, \tag{4}$$

where $\{l_i^n\}_{i \in J_n \cup \{j_n\}}$ and $\{m_i^n\}_{i \in J_n \cup \{j_n\}}$ are the prediction weights associated with filters **L** and **M**; as is typical in LOCAAT, we take $l_{j_n}^n = 1$.

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

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

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

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 **L**:

$$c_{n-1,i} = c_{n,i} + b_i^n \lambda_{j_n},$$
$$s_{n-1,i} = s_{n,i} + l_i^n s_{n,j_n} \quad \forall i \in J_n, \tag{7}$$

where $b_i^n = (s_{n,j_n} s_{n-1,i})/(\sum_{i \in J_n} s_{n-1,i}^2)$ 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, $\{d_{j_k}^{(1)}\}_{k \in D_{n-R}}$ and $\{d_{j_k}^{(2)}\}_{k \in D_{n-R}}$, 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 (**L**). More specifically, the inverse transform can be performed by the steps

Undo Update $c_{n,i} = c_{n-1,i} - b_i^n \lambda_{j_n}, \ \forall i \in J_n$

Undo Predict

$$c_{n,j_n} = \frac{\lambda_{j_n} - \sum_{i \in J_n} l_i^n c_{n,i}}{l_{j_n}^n} \quad \text{or} \tag{8}$$

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

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 $n \times n$

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, λ_{j_n}, as the difference between the observed and predicted values at j_n (see Eq. (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, μ_{j_n} (see Eq. (4)).

   (c) The complex-valued wavelet (detail) coefficients at j_n are $d_{j_n}^{(1)} = \lambda_{j_n} + i\,\mu_{j_n}$ and $d_{j_n}^{(2)} = \lambda_{j_n} - i\,\mu_{j_n}$.

3. Update: the smooth coefficients and their associated scales using the filter $\mathbf{L}$ (see Eq. (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

---

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), Jansen et al. (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 α_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 Sects. 2.3 and 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, for example nonparametric regression (Knight and Nason 2009). This strategy will be embedded in our proposed methodology in Sect. 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, for example 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 Eqs. (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 ($\mathbb{C}$WT) (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

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

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*

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

$$\varphi_{r-1,j}^{(1)}(x) = \varphi_{r,j}^{(1)}(x), \; \text{if } j \notin J_r, \quad (14)$$

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

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

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

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

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

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

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

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

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

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

*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 $\{\overline{\underline{\varphi}}^{(1)}, \underline{\tilde{\varphi}}, \overline{\underline{\psi}}^{(1)}, \underline{\tilde{\psi}}^{(2)}\}$ 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, for example, 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 Sect. 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 $\{\{c_{R,i}\}_{i \in \{0,...,N-1\} \setminus \{j_{N-1},...,j_{R-1}\}}, \{\underline{d}_{j_r}\}_{r=R-1}^{N-1}\}$ be the $\mathbb{C}^2$-LOCAAT transform of X, where $\underline{d}_{j_r} = \left(d_{j_r}^{(1)} \quad d_{j_r}^{(2)} \right)^T$. 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 ($\mathbb{C}$LoMPE)

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 Sect. 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* $|\cdot p|$ *the* $\mathbb{C}$-*modulus, we have for some constant* K

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

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, which we shall refer to as $\mathbb{C}$LoMPE (Complex-valued Long Memory Parameter Estimation Algorithm). Section 5.1, next, will show that our proposed $\mathbb{C}$LoMPE 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.

```
Complex-valued Long Memory Parameter
Estimation Algorithm (CLoMPE):
```
Assume that $\{X_{t_i}\}_{i=0}^{N-1}$ is as in Proposition 3. We estimate α 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 $\{\underline{d}_{j_r} = \left(d_{j_r}^{(1)} \quad d_{j_r}^{(2)}\right)^T\}_r$; see Eq. (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 Sect. 2.3. Estimate the wavelet energy within the artificial level $j^\star$ by

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

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 coincides 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 α as suggested by the results in Proposition 3. Note that Eq. (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

5 Simulated performance of $\mathbb{C}$LoMPE and real data analysis

5.1 Simulated performance of $\mathbb{C}$LoMPE

In what follows, we investigate the performance of our Hurst parameter estimation technique for complex-valued series. We simulated realizations of two types of long memory processes, namely circularly symmetric complex fractional Brownian motion, as introduced in Coeurjolly and Porcu (2018), 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 esti-

---

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

mated, 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 (5–20%), 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).

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 *CNL-Treg* 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

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

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 (2017) (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 better 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 $\mathbb{C}$**LoMPE**

In this section, we provide a more detailed long memory analysis of the complex-valued wind series described in Sect. 1.1. More specifically, we applied our $\mathbb{C}$LoMPE 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 by $\hat{H}_{\mathbb{R}}$ 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.

---

[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 (2017).

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 $\mathbb{C}$LoMPE estimation procedure. Boxed numbers indicate best result for the regularly spaced setting. Numbers in brackets are the estimation errors' standard deviation

H	$n = 256$					$n = 512$					$n = 1024$				
	Missingness proportion, p					Missingness proportion, p					Missingness proportion, p				
	CP	$\mathbb{C}$LoMPE				CP	$\mathbb{C}$LoMPE				CP	$\mathbb{C}$LoMPE			
	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 $\mathbb{C}$LoMPE estimation procedure. Numbers in brackets are the estimation errors' standard deviation

H	$n = 256$				$n = 512$				$n = 1024$			
	Missingness proportion, p				Missingness proportion, p				Missingness proportion, p			
	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)

Table 3 Hurst parameter estimates for the `Wind A` and `Wind B` data from complex-valued series using $\mathbb{C}$LoMPE 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 $\mathbb{C}$LoMPE 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 $\mathbb{C}$-modulus. This analysis highlights that ignoring the dependence structure between the real and imaginary components of the series may result in mis-estimation. Hence, we recommend an approach that uses the complex-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 Fig. 4, is visibly different to that of the real and imaginary component series shown in Fig. 2. We argue that our estimation of the long memory parameter for this series is more reliable than that in the currently 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, for example, 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
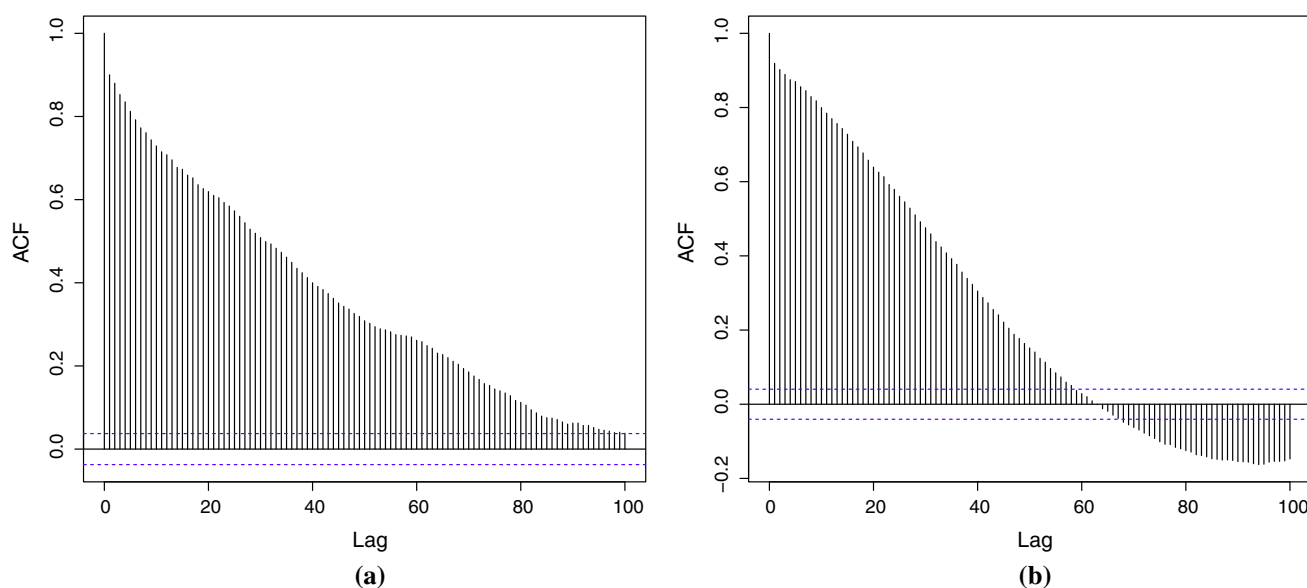
Fig. 4 **a** Autocorrelation for the *magnitude* wind series for the `Wind A` series from Fig. 1 (treated as regularly spaced); **b** autocorrelation for the *magnitude* `Wind B` dataset from Fig. 1 (treated as regularly spaced).

The dependence structure is markedly different to that shown for the real and imaginary series components shown in Fig. 2

methods have not been able to exploit the wealth of signal information in such data, whilst also coping with irregular sampling regimes. Our ℂLoMPE 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 ℂLoMPE 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 continue to enjoy popularity in the neuroscience community due to their non-invasive acqui-

sition and data richness; see, for example, 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, for example, Bullmore et al. (2003). Evaluation of the Hurst exponent in this context has been shown to be important in characterizing 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 set-up, 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 2018). 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) and Rowe (2009) and Adrian et al. (2018)]. 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, for example, 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 resembles 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 non-uniform. 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).

A Proofs and theoretical results

This appendix gives the theoretical justification of the results from Sects. 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 $f(x) =$ $\sum_{k \in S_n} c_{n,k} \varphi_{n,k}^{(1)}(x)$ with $\varphi_{n,k}^{(1)}(x) = \chi_{I_{n,k}}(x)$ as proposed in the LOCAAT construction (Jansen et al. 2009).

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

From Eq. (5), we have

$$d_{j_n}^{(1)} = \lambda_{j_n} + i\mu_{j_n} = c_{n,j_n}\left(\ell_{j_n}^n + i m_{j_n}^n\right) + \sum_{k \in J_n} c_{n,k}\left(\ell_k^n + i m_k^n\right). \quad (26)$$

By denoting $a_k^n = \ell_k^n + i m_k^n$, we obtain $d_{j_n}^{(1)} = c_{n,j_n} a_{j_n}^n - \sum_{k \in J_n} a_k^n c_{n,k}$. Using also the fact that $d_{j_n}^{(1)} = 0$, we have $c_{n,j_n} = \frac{\bar{a}_{j_n}^n}{|a_{j_n}^n|^2} \sum_{k \in J_n} a_k^n c_{n,k}$. If $j \in J_n$ then $c_{n,j} = 1$ and all others are zero, so $c_{n,j_n} = \frac{\bar{a}_{j_n}^n a_j^n}{|a_{j_n}^n|^2} := \tilde{a}_j^n$. Thus

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

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

For the *primal wavelet function* construction, we can similarly take $f(x) := \psi_{j_n}^{(1)}(x)$ and obtain the corresponding wavelet decomposition with coefficients $d_{j_n}^{(1)} = 1$ (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 $c_{n,j} = -b_j^n, \forall j \in J_n$ and $c_{n,j} = 0, \forall j \notin J_n$.

Using $d_{j_n}^{(1)} = c_{n,j_n} a_{j_n}^n - \sum_{j \in J_n} a_j^n c_{n,j}$ (as above) and $d_{j_n}^{(1)} = 1$, we have $c_{n,j_n} a_{j_n}^n = 1 - \sum_{j \in J_n} a_j^n b_j^n$ and $c_{n,j_n} = \frac{\bar{a}_{j_n}^n}{|a_{j_n}^n|^2}\left(1 - \sum_{j \in J_n} a_j^n b_j^n\right)$. Since $f(x) := \psi_{j_n}^{(1)}(x)$, we then have

$$\psi_{j_n}^{(1)}(x) = \frac{\bar{a}_{j_n}^n}{|a_{j_n}^n|^2}\left(1 - \sum_{j \in J_n} a_j^n b_j^n\right)\varphi_{n,j_n}^{(1)}(x) - \sum_{j \in J_n} b_j^n \varphi_{n,j}^{(1)}(x)$$

$$= \frac{\bar{a}_{j_n}^n}{|a_{j_n}^n|^2}\varphi_{n,j_n}^{(1)}(x) - \sum_{j \in J_n} b_j^n\left(\varphi_{n,j}^{(1)}(x) + \tilde{a}_j^n \varphi_{n,j_n}^{(1)}(x)\right).$$

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

$$\psi_{j_n}^{(1)}(x) = \frac{\bar{a}_{j_n}^n}{|a_{j_n}^n|^2}\varphi_{n,j_n}^{(1)}(x) - \sum_{j \in J_n} b_j^n \varphi_{n-1,j}^{(1)}(x),$$

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,

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

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

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

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

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

These steps are subsequently reiterated, 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

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

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

$$\psi_{j_n}^{(2)}(x) = \frac{a_{j_n}^n}{|a_{j_n}^n|^2} \varphi_{n,j_n}^{(2)}(x) - \sum_{j \in J_n} b_j^n \varphi_{n-1,j}^{(2)}(x).$$

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

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

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

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 ($r_X(\tau) \sim c_r |\tau|^{-\beta}$) whilst 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 $0 = 0 + i\,0$. Hence, $\mathbb{E}(d_j^{(\ell)}) = 0$ for $\ell = 1, 2$. In what follows, we drop the superscript (ℓ) in order to avoid notational clutter.

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

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

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 Eq. (29) can be rewritten as

$$\mathbb{E}(d_{j_r} \overline{d}_{j_k}) = \int_{\mathbb{R}} \overline{\left\{ \tilde{\varphi}_{r,j_r}(t) - \sum_{i \in J_r} A_i^r \tilde{\varphi}_{r,i}(t) \right\}}$$
$$\times \int_{\mathbb{R}} \left\{ \tilde{\varphi}_{k,j_k}(s) - \sum_{j \in J_k} A_j^k \tilde{\varphi}_{k,j}(s) \right\} \gamma_X(t-s) \, \mathrm{d}s \, \mathrm{d}t, \quad (30)$$

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, for example, Eq. (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

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

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 Eq. (31) gives

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

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

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

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

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

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 $u(x) = f_X(x) \, \mathrm{sinc}\left(x I_{n,i}/2 \right) \mathrm{sinc}\left(x I_{n,j}/2 \right)$ 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

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

and concludes the proof. $\qquad\square$

A.3 Proof of Proposition 3

As $\mathrm{Cov}(X_{t_i}, X_{t_j}) = \gamma_X(t_i - t_j)$ and $d_{j_r} = \langle X, \tilde{\psi}_{j_r} \rangle$, 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

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

where again we have dropped the basis index $\ell = 1, 2$ for notational brevity and we remind the reader that $|\cdot p|$ 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 (Eqs. (15) and (22)), it follows that the (dual) wavelet functions are linear combinations of the (same) scaling functions. For the first basis, Eq. (35) can be rewritten as

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

This can be expanded as

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

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

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

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 $\log_2\left(I_{r, j_r} \right) = j^\star + \Delta$

where $\Delta \in \{-1 + \log_2(\alpha_0), \log_2(\alpha_0)\}$ for some α_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 Eq. (38), we make a change of variable $\eta = \omega R2^{j^\star}$ to obtain

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

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

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 $j^\star \to \infty$)

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

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

All terms in (37) involve points from the same neighbourhood J_r, and thus using (40), (41) together with the linearity of the integral operator, we have that

$$
\begin{aligned}
\mathbb{E}(|d_{j_r}^{(1)}|^2) &\sim 2^{j^\star(\alpha-1)} R^{\alpha-1} \\
&\times \Bigg(M^2 |a_{j_r}^r|^2 + N^2 \sum_j \sum_{j'} \overline{a}_j^r a_{j'}^r \\
&\quad - MN \sum_j \overline{a}_j^r a_{j_r}^r - MN \sum_{j'} \overline{a}_{j_r}^r a_{j'}^r \Bigg) \\
&= 2^{j^\star(\alpha-1)} R^{\alpha-1} |M a_{j_r}^r - N \sum_{j \in J_r} a_j^r|^2 \\
&= C 2^{j^\star(\alpha-1)},
\end{aligned}
$$

where C is a constant depending on c_f, R and α.

A similar argument applies to the second basis and completes the proof. $\qquad \square$

References

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

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

Adali, T., Schreier, P.J., Scharf, L.L.: Complex-valued signal processing: the proper way to deal with impropriety. IEEE Trans. Signal Proc. **59**(11), 5101–5125 (2011)

Adrian, D.W., Maitra, R., Rowe, D.B.: Complex-valued time series modeling for improved activation detection in fMRI studies. Ann. Appl. Stat. **12**(2) (2018) (**to appear**)

Amblard, P.O., Coeurjolly, J.F., Lavancier, F., Philippe, A.: Basic properties of the multivariate fractional Brownian motion. Sém. Congr. **28**, 65–87 (2012)

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

Bakker, A.M.R., van den Hurk, B.J.J.M.: Estimation of persistence and trends in geostrophic wind speed for the assessment of wind energy yields in Northwest Europe. Clim. Dyn. **39**(3–4), 767–782 (2012)

Barber, S., Nason, G.P.: Real nonparametric regression using complex wavelets. J. R. Stat. Soc. B **66**(4), 927–939 (2004)

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

Bhattacharya, R.N., Gupta, V.K., Waymire, E.: The Hurst effect under trends. J. Appl. Prob. **20**, 649–662 (1983)

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

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

Calhoun, V.D., Adalı, T., Pearlson, G.D., Van Zijl, P.C.M., Pekar, J.J.: Independent component analysis of fMRI data in the complex domain. Magn. Reson. Med. **48**(1), 180–192 (2002)

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

Chang, T.P., Ko, H.H., Liu, F.J., Chen, P.H., Chang, Y.P., Liang, Y.H., Jang, H.Y., Lin, T.C., Chen, Y.H.: Fractal dimension of wind speed time series. Appl. Energy **93**, 742–749 (2012)

Churchill, N.W., Spring, R., Grady, C., Cimprich, B., Askren, M.K., Reuter-Lorenz, P.A., Jung, M.S., Peltier, S., Strother, S.C., Berman, M.G.: The suppression of scale-free fMRI brain dynamics across three different sources of effort: aging, task novelty and task difficulty. Nat. Sci. Rep. **6**(30), 895 (2016)

Ciuciu, P., Varoquaux, G., Abry, P., Sadaghiani, S., Kleinschmidt, A.: Scale-free and multifractal time dynamics of fMRI signals during rest and task. Front. Physiol. **3**, 186 (2012)

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

Coeurjolly, J.F., Porcu, E.: Fast and exact simulation of complex-valued stationary Gaussian processes through embedding circulant matrix. J. Comput. Gr. Stat. **27**(2), 278–290 (2018)

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

Craigmile, P.F., Percival, D.B.: Asymptotic decorrelation of between-scale wavelet coefficients. Trans. Im. Proc. **51**(3), 1039–1048 (2005)

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

Curtis, T.E.: Digital signal processing for sonar. In: Urban, H.S. (ed.) Adaptive Methods in Underwater Acoustics, pp. 583–605. Springer, Netherlands (1985)

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

Dowell, J., Weiss, S., Infield, D., Chandna, S.: A widely linear multi-channel wiener filter for wind prediction. In: IEEE Workshop on Statistical Signal Processing (SSP) 2014, pp 29–32. IEEE (2014)

Elipot, S., Lumpkin, R.: Spectral description of oceanic near-surface variability. Geophys. Res. Lett. **35**(5), L05606 (2008)

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

Ferdowsi, S., Abolghasemi, V.: Simultaneous BOLD detection and incomplete fMRI data reconstruction. Med. Biol. Eng. Comput. **56**(4), 599–610 (2018)

Flandrin, P.: Time–Frequency/Time-Scale Analysis. Academic Press, San Diego (1998)

Fortuna, L., Nunnari, S., Guariso, G.: Fractal order evidences in wind speed time series. In: International Conference on Fractional Differentiation and Its Applications (ICFDA) 2014, pp 1–6. IEEE (2014)

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

Goh, S.L., Chen, M., Popović, D.H., Aihara, K., Obradovic, D., Mandic, D.P.: Complex-valued forecasting of wind profile. Renew. Energy **31**(11), 1733–1750 (2006)

Gonella, J.: A rotary-component method for analysing meteorological and oceanographic vector time series. Deep Sea Res. Oceanogr. Abstr. **19**, 833–846 (1972)

Hamilton, J., Nunes, M.A., Knight, M.I., Fryzlewicz, P.: Complex-valued wavelet lifting and applications. Technometrics **60**(1), 46–60 (2017)

Haslett, J., Raftery, A.E.: Space-time modelling with long-memory dependence: assessing Ireland's wind power resource. Appl. Stat. **38**(1), 1–50 (1989)

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

Hsu, N.J.: Long-memory wavelet models. Stat. Sin. **16**, 1255–1271 (2006)

Hurst, H.E.: Long-term storage capacity of reservoirs. Trans. Am. Soc. Civ. Eng. **116**, 770–808 (1951)

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

Jansen, M., Nason, G.P., Silverman, B.W.: 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, vol. 4478, pp. 87–97. SPIE, Bellingham (2001)

Jansen, M., Nason, G.P., Silverman, B.W.: Multiscale methods for data on graphs and irregular multidimensional situations. J. R. Stat. Soc. B **71**(1), 97–125 (2009)

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

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

Jung, Y.Y., Park, Y., Jones, D.P., Ziegler, T.R., Vidakovic, B.: Self-similarity in NMR spectra: an application in assessing the level of cysteine. J. Data Sci. **8**(1), 1 (2010)

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

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

Knight, M.I., Nason, G.P.: A nondecimated lifting transform. Stat. Comput. **19**(1), 1–16 (2009)

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

Knight, M.I., Nunes, M.A., Nason, G.P.: Spectral estimation for locally stationary time series with missing observations. Stat. Comput. **22**(4), 877–8951 (2012)

Knight, M.I., Nason, G.P., Nunes, M.A.: A wavelet lifting approach to long-memory estimation. Stat. Comput. **27**(6), 1453–1471 (2017)

Li, H., Correa, N.M., Rodriguez, P.A., Calhoun, V.D., Adali, T.: Application of independent component analysis with adaptive density model to complex-valued fMRI data. IEEE Trans. Biomed. Eng. **58**(10), 2794–2803 (2011)

Lilly, J.M., Gascard, J.C.: Wavelet ridge diagnosis of time-varying elliptical signals with application to an oceanic eddy. Nonlinear Proc. Geophys. **23**(4), 467 (2006)

Lilly, J.M., Sykulski, A.M., Early, J.J., Olhede, S.C.: Fractional Brownian motion, the Matérn process, and stochastic modeling of turbulent dispersion. Nonlinear Proc. Geophys. **24**, 481–514 (2017)

Lindquist, M.A.: The statistical analysis of fMRI data. Stat. Sci. **23**(4), 439–464 (2008)

Lobato, I., Robinson, P.M.: Averaged periodogram estimation of long memory. J. Econom. **73**(1), 303–324 (1996)

Mandelbrot, B.B., Taqqu, M.S.: Robust R/S analysis of long-run serial correlation. Bull. Int. Stat. Inst. **48**(2), 59–104 (1979)

Mandelbrot, B.B., Van Ness, J.W.: Fractional Brownian motions, fractional noises and applications. SIAM Rev. **10**(4), 422–437 (1968)

Mandic, D.P., Goh, V.S.L.: Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models, vol. 59. Wiley, Hoboken (2009)

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

Martin, K.W.: Complex signal processing is not complex. IEEE Trans. Circ. Syst. I Reg. Pap. **51**(9), 1823–1836 (2004)

McCoy, E.J., Walden, A.T.: Wavelet analysis and synthesis of stationary long-memory processes. J. Gr. Comput. Stat. **5**(1), 26–56 (1996)

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

Nafornita, C., Isar, A., Nelson, J.D.B: Regularised, semi-local hurst estimation via generalised lasso and dual-tree complex wavelets. In: IEEE International Conference on Image Processing (ICIP) 2014, pp 2689–2693. IEEE (2014)

Neeser, F.D., Massey, J.L.: Proper complex random processes with applications to information theory. IEEE Trans. Inf. Theory **39**(4), 1293–1302 (1993)

Nelson, J.D.B., Kingsbury, N.G.: Dual-tree wavelets for estimation of locally varying and anisotropic fractal dimension. In: 17th IEEE International Conference on Image Processing (ICIP) 2010, pp 341–344, IEEE (2010)

Nunes, M.A., Knight, M.I.: CNLTreg: Complex-Valued Wavelet Lifting for Signal Denoising. R package version 1 (2017)

Nunes, M.A., Knight, M.I., Nason, G.P.: Adaptive lifting for nonparametric regression. Stat. Comput. **16**(2), 143–159 (2006)

Olhede, S.C., Walden, A.T.: Local directional denoising. IEEE Trans. Signal Proc. **53**(12), 4725–4730 (2005)

Osborne, A.R., Kirwan Jr., A.D., Provenzale, A., Bergamasco, L.: Fractal drifter trajectories in the Kuroshio extension. Tellus A Dyn. Meteorol. Oceangr. **41**(5), 416–435 (1989)

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

Peng, C.K., Buldyrev, S.V., Havlin, S., Simons, M., Stanley, H.E., Goldberger, A.L.: Mosaic organization of DNA nucleotides. Phys. Rev. E **49**(2), 1685 (1994)

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

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

Qu, B., Addison, P.S.: Modelling flow trajectories using fractional brownian motion. In: International Workshop on Chaos-Fractals Theories and Applications (IWCFTA), 2010, pp 420–424. IEEE (2010)

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

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

Rodriguez, P.A., Calhoun, V.D., Adalı, T.: De-noising, phase ambiguity correction and visualization techniques for complex-valued ICA of group fMRI data. Pattern Recognit. **45**(6), 2050–2063 (2012)

Rowe, D.B.: Modeling both the magnitude and phase of complex-valued fMRI data. Neuroimage **25**(4), 1310–1324 (2005)

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

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

Sanderson, B.G., Booth, D.A.: The fractal dimension of drifter trajectories and estimates of horizontal eddy-diffusivity. Tellus A Dyn. Meteorol. Oceangr. **43**(5), 334–349 (1991)

Schreier, P.J., Scharf, L.L.: Second-order analysis of improper complex random vectors and processes. IEEE Trans. Signal Proc. **51**(3), 714–725 (2003)

Schreier, P.J., Scharf, L.L.: Statistical Signal Processing of Complex-Valued Data: The Theory of Improper and Noncircular Signals. Cambridge University Press, Cambridge (2010)

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

Shi, B., Vidakovic, B., Katul, G.G., Albertson, J.D.: 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 (2005)

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

Summers, D.M.: Impulse exchange at the surface of the ocean and the fractal dimension of drifter trajectories. Nonlinear Proc. Geophys. **9**(1), 11–23 (2002)

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

Sykulski, A.M., Percival, D.B.: Exact simulation of noncircular or improper complex-valued stationary Gaussian processes using circulant embedding. In: IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP) 2016, pp. 1–6 (2016)

Sykulski, A.M., Olhede, S.C., Lilly, J.M., Early, J.J.: Frequency-domain stochastic modeling of stationary bivariate or complex-valued signals. IEEE Trans. Signal Proc. **65**(12), 3136–3151 (2017)

Tanaka, T., Mandic, D.P.: Complex empirical mode decomposition. IEEE Signal Proc. Lett. **14**(2), 101–104 (2007)

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

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

Vidakovic, B.D., Katul, G.G., Albertson, J.D.: Multiscale denoising of self-similar processes. J. Geophys. Res. **105**(D22), 27,049–27,058 (2000)

Walden, A.T.: Rotary components, random ellipses and polarization: a statistical perspective. Philos. Trans. R. Soc. A **371**(1984), 20110554 (2013)

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

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

Zhang, Q., Harman, C.J., Ball, W.P.: Evaluation of methods for estimating long-range dependence (LRD) in water quality time series with missing data and irregular sampling. In: Proceedings of American Geophysical Union Fall Meeting 2014, San Francisco (2014)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.