Coherency of Lightning Sferics

X. Bai1 and M. Füllekrug1

1Centre for Space, Atmospheric and Oceanic Science, University of Bath, Bath, UK

Abstract The amplitudes of lightning sferics are commonly used for data analyses, for example, in lightning location networks. This contribution investigates the coherency of lightning sferics, which is calculated from the phase of the analytic complex signal. The complex signal is obtained from the Hilbert transform with the recorded signal as the input signal. This analytic signal is composed of amplitudes with corresponding phases. Rigorous selection criteria are applied to ensure that consistent events are used to calculate distance-dependent average sferics waveforms which are summarized in waveform banks for the amplitude and coherency. The impulse response to lightning flashes is derived from an averaged waveform which is used to detect individual lightning events. It is shown that the coherency waveform bank is in general agreement with the amplitude waveform bank, that is, the relative arrival times of the ground and sky waves. The coherency waveform bank exhibits the same number of skywaves on the logarithmic scale, but a less attenuated ground wave when compared to the amplitude waveform bank. The results of this study show that the phase of sferics offers additional information to amplitudes, which could potentially be used for the interferometric lightning location with long-range lightning detection networks.

1. Introduction

Lightning discharges are natural phenomena that are able to generate energies up to several GJ (Rakov & Uman, 2006). The transient radio waves emitted by the lightning discharges are called “atmospherics”, or “sferics” for short. These electromagnetic waves widely cover the frequency spectrum from ∼1 Hz to ∼300 MHz, with relatively large spectral amplitudes in the extremely low frequency (ELF) range and very low frequency (VLF) range and a relative maximum at ∼10 kHz (e.g., Burke & Jones, 1992; Taylor, 1960; Weidman & Krider, 1986). Another source of VLF electromagnetic waves is radio transmitters (Barr et al., 2000).

Lightning discharges fall into two major categories: cloud-to-ground discharges (CGs) and in-cloud discharges (ICs), where the CGs exhibit much larger peak currents compared to the ICs (e.g., Betz et al., 2009; Fiser et al., 2010). Cloud-to-ground discharges can lead to serious hazards and are more relevant to human life than ICs. The peak current of CGs has been studied for the purpose of lightning protection (e.g., Chowdhuri et al., 2005; Schulz et al., 2016; Takami & Okabe, 2007; Visacro, 2004), and existing lightning location networks mainly aim at detecting cloud-to-ground discharges with the time-of-arrival (TOA) technique.

The ionosphere (σ ≈ 10^{-4} Sm^{-1} – 10^{-2} Sm^{-1}) and the earth’s ground (σ ≈ 10^{-3} Sm^{-1}) form a natural wave guide with large conductivity boundaries, while the atmosphere in between exhibits a much lower conductivity. This wave-guide is able to guide the propagation of sferics (e.g., Crombie, 1965; Magunia, 1996; Schlegel & Füllekrug, 1999; Silber et al., 2015; Stuchly, 2009). The attenuation of the propagating VLF electromagnetic waves is a function of the propagation distances and the ionospheric conductivity (Burkholder et al., 2013). The ground conductivity also contributes to the attenuation of the propagating waves (Wait, 1962).

The earth-ionosphere wave guide exhibits a selective ionospheric absorption toward the various frequency components of the VLF signal (Macario & Chapman, 1956). According to the Austin Cohen Law, there is a positive correlation between the attenuation rate and frequencies above ∼10 kHz. The attenuation effects on ELF waves and VLF waves are relatively low which allows them to propagate globally (e.g., Barr et al., 2000; Macario & Chapman, 1956; Stuchly, 2009).

In both, the earth-ionosphere cavity and the lower ionosphere (D region and lower E region), the profiles of the conductivity exhibit an exponential increase with altitude (Galejs, 1961). When electromagnetic waves enter the ionosphere, a small portion travels through the ionospheric D region which results in energy loss (Burkholder et al., 2013). These ionospheric D region reflection losses are caused by the large electron-neutral collision rate, which decreases the amplitude of the received electromagnetic waves (Jacobson et al., 2008). There is evidence...
showing that the spectrum of sferics is closely related to ionospheric height and slope (Cummer et al., 1998). Furthermore, the sferics spectrum contains information about the ionosphere along the propagation path. In the daytime, the VLF waves are reflected from a wide range of heights with the average reflection height of ~60 km, while at the nighttime, the reflection height is much more stable at ~90 km (e.g., Deeks, 1966; Smith et al., 2004; Stuchly, 2009). The electromagnetic waves exhibit lower attenuation during nighttime when compared during daytime, which means nighttime lightning sferics are more likely to exhibit skywaves (waves reflected back and forth between the ground and the ionosphere) (e.g., Cummer et al., 1998; Macario & Chapman, 1956).

The ground conductivity effects on the electromagnetic wave propagation are frequency dependent, and the ground wave (wave propagating along the ground) amplitude attenuation caused by different ground conductivities are studied in the time and frequency domain (e.g., Caligaris et al., 2008; Cooray, 2009; Rachidi et al., 1996). For long-distance propagation of sferics from lightning, the time delay of the ground wave peak is related to the propagation distance (e.g., Honma et al., 1998; Pessi et al., 2009; Shao & Jacobson, 2009; Zhou et al., 2021). Besides the ground conductivity, the Earth curvature is also an important factor that causes larger ground wave attenuation over long distances (Hou et al., 2018).

The observed electromagnetic waves are a mixture of ground waves and skywaves. The reflection process can largely reduce the skywave amplitude during the propagation and the skywave amplitude decreases as the number of reflections increases. Therefore, the number of detectable skywaves can be used to evaluate the merit of using the amplitude and phase of the analytic sferics signal.

The analytic signal of lightning sferics is obtained from the Hilbert transform by treating the recorded signal as the input signal and rotating the phase by 90° (Liu et al., 2018; Taner et al., 1979). The amplitude and phase can then be obtained from the analytic signal. To compare the amplitude and phase, the waveform bank is used here to increase the signal-to-noise ratio (SNRs) of lightning sferics. The waveform bank consists of distance-dependent waveforms that characterize lightning sferics (Said et al., 2010).

The main purpose of this contribution is to investigate a novel technique, that is, complex analysis, which has the potential to expand the use of the collected lightning sferics when compared to the traditional TOA technique. Zhu et al. (2021) studied the interferometric method using the peak amplitude of the lightning sferic, which increases the detection rate and is capable of detecting multiple events even when their pulses are overlapping at some of the receivers. It is interesting to investigate the interferometric method using the phase coherency with several receivers working synchronously while observing the same area. This paper introduces the preliminary work required to study such complex interferometry for the potential use in long-range lightning location networks. The paper is organized as follows:

The event selection process for the recorded data is described in Section 2, the amplitude waveform bank is introduced in Section 3, the complex analysis and coherency waveform bank is developed in Section 4, one application of the waveform bank, lightning detection with the impulse response, is introduced in Section 5, the comparison between the amplitude and coherency waveform bank is summarized in Section 6, and the final conclusions are presented in Section 7.

2. Event Selection

A summer measurement campaign was carried out in southern France to collect lightning sferics from 18th to 31 August 2019. A flat plate antenna was set up in Rustrel (France) at a latitude and longitude of 43.94°N, 5.48°E. The receiver was set up on a remotely located mountain top with a relatively small level of local noise interference, and it measures the electromagnetic fields within the frequency range from ~4 Hz to ~400 kHz at a sampling frequency of 1 MHz (Füllekrug, 2010). The digital filter used in this work is a band pass filter from 1 to 400 kHz. The <1 kHz frequency components are removed because they are contaminated by power line harmonic radiation. Lightning information for each sferic is provided by Meteorage (Pédeboy, 2015) which includes the time of occurrence, location, polarity, peak current and lightning type (CG or IC). Negative cloud-to-ground discharges account for 93.96% of all the collected CGs, and the median peak current of negative CGs is −16.4 kA. From all the data reported by Meteorage, only negative CGs are used because they are most common. However, further quality control is necessary because there are a few misidentified +CGs.
The recorded signals have a lower attenuation during the nighttime compared to the daytime, which means the waveforms of nighttime events are more likely to exhibit skywaves (Macario & Chapman, 1956). Thus, all the data were recorded during astronomical nighttime when the sun is 18° below the horizon. The recording times for each of the 13 nights were rigorously selected to ensure that the storm and the propagation paths are all immersed in the nighttime environment.

Events are selected by evaluating the measured maximum electric field and the reported peak current. The recorded sferics traveled from different directions with various propagation distances, which makes their maximum electric fields not comparable to each other. Hence, the electric field coefficient, that is, the electric field at 100 km from the source, for each individual event is determined by applying a distance correction. The distance correction is achieved by calculating the attenuation coefficient along the propagation path. As aforementioned, the attenuation coefficient is frequency dependent. For the frequency range above ∼10 kHz, the attenuation increases with increasing frequency. Thus, the high frequency components are less important in the distance correction as they have larger attenuation (Macario & Chapman, 1956).

The simple wave propagation model used here is adapted from Kolmašová et al. (2016) and Kašpar et al. (2017) and defined in Equation 1

$$E = \frac{A}{D} 10^{-\alpha D},$$

where $E$ is the received maximum electric field, $A$ is the corresponding electric field coefficient, $D$ is the propagation distance normalized to 100 km, and $\alpha$ is an experimental attenuation coefficient. The electric field coefficient $A$ is linearly related to the peak current $I$ of the corresponding lightning event with a constant ratio $k$, as illustrated in Equation 2 (e.g., Kašpar et al., 2017; Kolmašová et al., 2016; Uman & McLain, 1970)

$$A = kI.$$  

Equation 3 is obtained by rearranging Equation 1 and Equation 2

$$\frac{E}{I} = \frac{k}{D} 10^{\alpha D}.$$  

Equation 3 shows that the ratio $E/I$ is distance-dependent. Thus, the experimental attenuation coefficient $\alpha$ can be obtained by fitting the $E/I$ ratio using the distance $D$ to the model as described in Equation 3. With a large amount of data collected, the average $E/I$ ratio is determined from grouped events which are classified by their propagation distances. The distance resolution is 10 km and each event group is required to include more than 100 events to produce a reliable average $E/I$ ratio. For each distance bin, the maximum electric field $E$ and peak current $I$ of grouped events are used to fit a representative $E/I$ ratio using linear regression analysis (Figure 1a). In this example, the $E/I$ ratio is calculated for the event group with a propagation distance of 650 km. Strong events with large peak currents deviate more from the regression line, but if these deviations are calculated as a percentage of the corresponding peak currents, they are in an acceptable range. After the distance-dependent $E/I$ ratios are calculated for all available distances, the attenuation coefficient can be determined using Equation 3 (Figure 1b). The experimental attenuation coefficient calculated for the receiver in Rustrel is 5.11 dB/Mm. The attenuation coefficient investigated by Kolmašová et al. (2016) varies from 17.4 dB/Mm to 23 dB/Mm with a mean value of 21.5 dB/Mm. Compared to the reported attenuation value of 21.5 dB/Mm, the attenuation coefficient in this work is relatively low. One possible explanation is that all the events in this study are collected during the nighttime with a lower attenuation coefficient than during the day. In addition, the receiver used in Kolmašová et al. (2016) is a broadband antenna measuring from 5 kHz to 37 MHz, which has a much higher frequency range compared to the receiver used in this work. The higher frequency components are more strongly attenuated which contributes to the larger attenuation reported by Kolmašová et al. (2016).

The observed peak current can be calculated using Equation 3 based on the measured maximum electric field $E$. Consistent events are selected with an observed peak current which is within ±30% of the peak current reported by Meteorage.
3. Amplitude Waveform Bank

Lightning sferics in the VLF band exhibit extremely low attenuation effects which allow them to be detected almost globally. However, the received radio waves still contain noise, which may arise from wave propagation and interference at the receiver location. To have a better understanding of the lightning waveforms, a waveform bank is calculated here which consists of distance-dependent lightning sferics by averaging the received electromagnetic waves. The averaging process thereby cancels out noise and removes part of the natural lightning variability such that average lightning characteristics are more prominent.

The event selection process ensures that the determined electric field coefficients $A$ of the selected events are proportional to their reported peak currents $I$. However, this step cannot guarantee that all the selected events are −CGs. One possible reason for this is for example, that the waveforms of return strokes from some +CGs can exhibit untypical shapes which are not accounted for during the automated lightning detection process. Hence, the misidentified strong +CG events need to be excluded before the averaging process. The waveforms of selected events are referenced to $t = 0$ ms by taking out the propagation time under the assumption that the propagation velocity equals the speed of light. The waveform length is 6 ms which is from −1 to 5 ms with respect to the lightning occurrence time. For each individual lightning waveform, the local maximum is defined as the maximum electric field within the time range of 0–20 μs, which is attributed to the arrival of the ground wave reported by Meteorage. Even though the receiver is precise to ∼1 μs, the ground wave peak can be elongated up to 20 μs as a result of the rise time of the ground wave. The global maximum is defined as the maximum electric field within the entire waveform (i.e., −1–5 ms), which can have different sources. The +CGs are excluded by comparing the ratio $R_E$ between the global maximum and the local maximum. For the waveform of a −CG at relatively short distances, both, the local maximum and global maximum are identical and associated with the ground wave, which leads to a ratio $R_E$ of 1. For the waveform of a −CG at large distances, the global maximum is the peak of the first skywave while the local maximum is the ground wave peak, therefore, the ratio $R_E$ is distance-dependent. The ratio $R_E$ increases with distance. The analysis shows that the ratio $R_E$ can increase to ∼4.17 for a distance of 1,220 km (more details in Section 6). For some rare events, the local maximum is associated with the preceding small amplitude of the ground wave (inset figure in Figure 4c), and the global maximum is the peak of the ground wave, which would lead to a very large ratio $R_E$. Therefore, in this contribution, a threshold value of ratio $R_E$ is set to distinguish such +CG events. The most distant event collected in this campaign has a propagation distance of 1,341 km, which means the ratio $R_E$ for this event is larger than 4.17. By analyzing the waveforms of the rare +CGs, it is found that the ratio $R_E$ of these +CGs is usually much higher than 10. Therefore, the threshold value of the ratio $R_E$ is set as 10. That is, if the ratio $R_E$ is larger than 10, the event would be marked as +CG and is excluded from the averaging process.

Figure 1. Illustration of event selection. (a) An example of $E/I$ ratio determination for the distance bin of 650 km. The black dots are the events whose propagation distances are 650 km. The red line is the determined $E/I$ ratio using linear regression. (b) Average $E/I$ ratios versus corresponding distances with the best matched experimental attenuation coefficient. The red curve is calculated with the best fitting experimental attenuation coefficient.
After the rigorous event selection process, each lightning waveform is subsequently scaled to its maximum electric field of the ground wave such that the averaged waveform is not dominated by exceptionally strong lightning events. Note that the ground wave peak is defined as the maximum electric field within the time interval of 0–20 μs because the ground wave peak is delayed by the rise time as explained at the beginning of this section. The individual waveforms are then grouped by their propagation distances with a distance resolution of 10 km. According to Said et al. (2010), lightning sferics of 50 events are sufficient to define the average lightning waveform shape. Due to the large amount of data collected, that is, ∼24,000 events selected for Rustrel data, the event number threshold for each distance bin is chosen to be 100 events to achieve overall improved accuracy. The waveform bank is shown as an example in Figures 2a and 2b. This amplitude waveform bank covers the distance bins from 190 to 1,220 km and the black gaps shown in the waveform banks are caused by an insufficient number of events at some distance bins.

Figure 2a shows the amplitude waveform bank in the time domain. After averaging the selected events, the received random noise is canceled out so that the ground wave and skywaves are clearly shown here. Up to fourth order skywaves can readily be distinguished at long distances on the linear scale (Figure 2a). Along the propagation path, the incident electromagnetic waves which travel toward the ionosphere are partially reflected while a small portion could penetrate into a higher region of the ionosphere (i.e., E and F region), and further travel into the magnetosphere as lightning whistlers. This explains the amplitude reduction of the skywaves. As for waveforms with distances larger than ∼540 km, the amplitude of the first skywave appears to be larger than

---

**Figure 2.** Amplitude waveform bank in the time domain (a) and the frequency domain (b). Coherency waveform bank in the time domain (c) and the frequency domain (d).
the ground wave. While the propagation distance increases, the time difference between the ground wave and the skywave decreases, which appears like “wave crowding” in the waveform bank.

The waveform bank in the frequency domain is shown in Figure 2b, which shows the main energy lies below ~50 kHz and is centered at ~10 kHz. The patterns shown here are the result of the combination of ground waves and skywaves. With increasing distance, the patterns tend to be more dispersed. The blue patterns are spectral minima that separate consecutive relative maxima. The spectral minima show the characteristic of the distance along the propagation path while the relative maxima are caused by the superposition of various wave propagation modes (Liu et al., 2018). Some transmitters consistently appear in the frequency waveform bank, such as Europe 1 (closed down on 31 December 2019) at 183 kHz, BBC Radio 4 (24 hr) at 198 kHz, Radio Monte Carlo (closed down from 29 Mar 2020) at 216 kHz, RTL (24 hr) at 234 kHz, and Radio Algérie, Chaîne 3 (24 hr) at 252 kHz.

4. The Complex Analysis and Coherency Waveform Bank

Traditional lightning detection and location networks only use the amplitude of the received electromagnetic field. To expand the use of the collected lightning sferics, the complex analysis is investigated. In this section, the concept of an analytic signal, phase and coherency will be introduced, and the coherency waveform bank will be illustrated.

Received lightning sferics can be transferred to complex time series by obtaining the imaginary part using the Hilbert transform (Liu et al., 2018; Taner et al., 1979) which is represented by Equation 4

\[ y_n(t) = E(t)e^{j(\omega t - k kn)} \quad n = 1, 2, 3 \ldots N \]

where \( n \) is the receiver number of \( N \) receivers, \( y_n(t) \) is the received analytic complex signal at receiver \( n \), \( E(t) \) is the source lightning electric field and \( e^{j(\omega t - k kn)} \) indicates the plane wave propagation mode. The source electric field \( E(t) \) is composed of amplitude \( A(t) \) and phase \( \varphi(t) \) which is \( E(t) = A(t)e^{j\varphi(t)} \). The phase of the received signal is the angle between the imaginary part and real part of the analytic signal \( y_n(t) \), and the unit phasor can be determined using the analytic signal \( y_n(t) \) divided by its absolute amplitude value \( |y_n(t)| \). Usually, the phase progression \( e^{j\omega t} \) is removed by multiplication with the down converting factor \( e^{-j\omega t} \) where \( \omega \) is the center frequency. In this work, however, the phase progression is kept since the collected data are broadband measurements. The phase delay factor \( e^{-j(\omega t + k kn)} \) is removed by taking out the propagation time under the assumption that the propagation velocity equals the speed of light. Coherency is a statistic that measures the similarity of the phase in data sets. The coherence of the analytic signal phase, named coherency for short in the following section, is defined in Equation 5 (Füllekrug et al., 2016), where \( \varphi_n \) is the event source phase at receiver \( n \). According to the definition of the coherency, the maximum coherency value is one

\[ coh = \frac{1}{N} \sum_{n=1}^{N} \frac{|E(t)e^{j\varphi_n}|}{|E(t)|} = \frac{1}{N} \sum_{n=1}^{N} e^{j(\omega t + \varphi_n)} \]

Unlike the amplitude waveform data, the <1 kHz frequency components are associated with power line harmonic radiation and could be kept for the coherency calculation, because these components do not significantly distort the lightning phase. In terms of the selected lightning events, the coherency waveform bank is determined by the event groups classified by their propagation distance with a 10 km distance resolution. The time domain coherency waveform bank (Figure 2c) is produced by calculating the coherency from the phase information of the grouped events, and the frequency domain coherency waveform bank (Figure 2d) is calculated from the spectra of the grouped events. The coherency waveform bank exhibits similar lightning characteristics as the amplitude waveform bank, and it exhibits up to nine skywaves on the linear scale. More details on the comparison between the amplitude waveform bank and the coherency waveform bank will be given in Section 6.

The coherency is also calculated for lightning sferics filtered for different frequency ranges which indicate the different lightning components. The average waveform of the event group at 190 km distance is shown here as an example. Figure 3 shows the coherency over broadband lightning sferics (0–400 kHz), return stroke component waveforms (5–15 kHz), continuing current component waveforms (0–2 kHz), and intra-cloud flash component waveforms (150–400 kHz) from top to bottom. The return stroke is the most powerful component of the
lightning discharge process, and its coherency waveform exhibits the largest value and a relatively wide pulse compared to the broadband results. The continuing current component exists for a relatively long time but with lower energy, which is reflected in the coherency waveform as a wider pulse with a lower peak value. The coherency over the intra-cloud flashes is a sharp pulse whose peak has the smallest value. The same processes are repeated for long distance bins which lead to the same conclusions.

5. Impulse Detection

The averaged waveforms can be used as a transfer function that represents the characteristics of lightning sferics at the corresponding distance. The transfer function $T_d$ is defined in Equation 6 (note: the index $d$ indicates different distances such that each transfer function is related to a different propagation distance)

$$T_d = F^{-1}[F(y_{imp})F(y)]$$

where $F$ is the Fourier transform operator, the factor $y_{imp}$ is the analytic signal of the averaged waveform of $N$ negative CG events whose propagation distance is $d$, and $y_{imp}$ is the Kronecker delta function, that is, a digital impulse.

The transfer function is able to detect lightning events when it is applied to the received electromagnetic waves as shown in Equation 7. When the measured analytic signal $y_m$ is multiplied by the transfer function at the corresponding distance, the output analytic signal $y_{out}$ is supposed to be an impulse

$$y_{out} = F^{-1}[F(y_m)T_d].$$

The events selected in Section 2 are used to simulate the impulse detection, while some strong positive CGs are kept in this step to test the sensitivity of the methodology toward the rarely misidentified +CGs events. Note that the test events do not include the events used to calculate the transfer function. All the test events are classified by their propagation distances into event groups. This subsequent analysis is based on these grouped events. The format of the test event waveform is the same as the transfer function that is referenced to $t = 0$ with a duration that covers 1 ms before and 5 ms after the lightning occurrence time. The test waveforms are also scaled to the ground wave maximum to mitigate the impact of exceptionally strong intensities. To remove the interference from high frequency contributions, a low pass filter with a cutoff frequency of 50 kHz is applied to calculate the analytic output signal $y_{out}$. This low pass filter frequency was chosen because larger frequency components are less coherent as shown in Figure 2.

A numerical algorithm is needed here to evaluate whether the output signal is an impulse, given a large number of test events. The amplitude distribution of the output analytic signal $y_{out}$ is calculated to identify a threshold value $E_{thr}$, which is defined as the value that is larger than the 97th percentile of the waveform $y_{out}$. The factor $R$ is defined in Equation 8

$$R = E_{max}/E_{thr}$$

where the $E_{max}$ is the maximum amplitude of the output analytic signal $y_{out}$. This factor $R$ measures the prominence of the impulse in the analytic output signal $y_{out}$. The output signal can only be regarded as an impulse signal under the conditions that: (a) The factor $R$ is larger than 2; (b) the time stamp of the maximum amplitude $E_{max}$ is within the $t = \pm 10\mu s$ range. When the analytic signal $y_{out}$ is identified as an impulse signal, the impulse detection process of the test event is classified as a success.

For individual events, only the amplitude of the impulse detection result is discussed here. Both, the successful cases and failure cases can be illustrated as follows: A standard successful impulse detection example is shown in Figure 4a. This is an event with a peak current of $-20.6\, kA$ at a propagation distance of 610 km. In the time domain, despite the test waveform having more noise compared to the averaged waveform from the waveform
bank, the matched output signal is an outstanding impulse signal. The maximum amplitude of the output analytic signal $y_{out}$ is much larger than the threshold amplitude value $E_{thr}$. This method is also able to detect relatively weak events, which are hidden in the noise. An example is shown in Figure 4b with a peak current of $-6.3$ kA and a propagation distance of 610 km. The propagation distance is 540 km. The inset figure shows the small preceding amplitude change between −0.1 and 0.1 ms. (d) The test event with a peak current of $-3.8$ kA and a propagation distance of 610 km.

Figure 4. Impulse detection examples. For each figure, the top signal is the averaged waveform that is used to produce the transfer function $T_d$, the middle signal is the test event waveform $y_{in}$, the bottom signal is the matched output signal $y_{out}$ with the threshold electric field $E_{thr}$ marked as the red line. All the waveforms illustrated here are the real parts of the analytic signals. (a) The test event with a peak current of $-20.3$ kA and a propagation distance of 610 km. (b) The test event with a peak current of $-6.3$ kA and a propagation distance of 610 km. (c) The test event was classified to be a −CG with a peak current of $-57.9$ kA while the waveform exhibits positive polarity. The propagation distance is 540 km. (d) The test event with a peak current of $-3.8$ kA and a propagation distance of 610 km.
neighboring events is assumed to be relatively small, the use of the transfer function decided by the first event works quite well in these cases.

The failure cases can be attributed to two main reasons. A small number of strong +CGs with unusual waveforms are misidentified as −CGs, therefore, the +CGs cannot be detected using the transfer function calculated by −CG events. These untypical shapes of some +CGs can cause misidentification of lightning polarity. The impulse detection procedure can easily distinguish the misidentified strong +CGs and an example is shown in Figure 4c. This event is identified as −57.9 kA with a propagation distance of 540 km. According to the polarity of the event waveform, this event is certainly a positive CG event as it exhibits an extremely large negative ground wave pulse when using the atmospheric electricity polarity convention. The arrow marked in the lightning waveform is pointing at a very small preceding amplitude change with an inverted polarity that occurs at the lightning occurrence time reported by Meteorage. There are several +CGs waveforms which have similar patterns. This means that these small amplitude changes warrant further investigation in future work. However, at the current stage, the source and cause of these pulses are unknown. The second and more common failure mode occurs when the impulse detection process is applied to weak test events with low SNRs. An example event is shown in Figure 4d which has a peak current of −3.8 kA and a propagation distance of 610 km. The lightning sferics of weak events are largely distorted along the long propagation distance and end up with comparable amplitudes as the local noise, which makes them difficult to detect with the impulse detection method.

The amplitude of the impulse detection has been discussed, and the phase information will be discussed in this paragraph. The large amplitudes of the recorded sferics indicate the existence of lightning events. Similarly, the phase of the lightning analytic signal is also able to indicate the lightning occurrence. A Gaussian impulse is used as an idealized lightning signal to gain a first insight into the general behavior of phase. The phase of the Gaussian impulse remains constant when the amplitude of the Gaussian pulse is very small. The phase starts rising from negative to positive when the Gaussian signal amplitude starts rising, and the phase exhibits a zero crossing when the pulse amplitude reaches maximum. For the individual events simulated in the impulse detection process, the phase has the trend to increase from negative to positive and cross 0 at the lightning occurrence time while for other times the phase remains erratic, such that this result matches well with the phase simulation using a Gaussian pulse. Based on this property, it was decided to check the phase of the output impulse response $y_{out}$. However, it is found that this property is not clearly shown when the phase progression factor $e^{j\omega t}$ is dominating the phase of an individual signal. Therefore, it requires further investigation.

The performance of the impulse detection is evaluated here. The relation between the factor $R$ and the peak current is investigated. The larger the $R$ value is, the larger the certainty of lightning existence. The analysis shows that the factor $R$ and the peak current $I$ are positively correlated (not shown). This finding coincides with the event selection that the source electric field coefficient $A$ is linearly proportional to the peak current as the impulse detection process also takes out the distance factor. For all the tested events, the peak currents are mostly within the range from 0 to −40 kA, while the unmatched cases are mostly weak events with peak currents below −20 kA. In addition, the time stamp of the maximum amplitude $E_{max}$ of the matched impulse signal $y_{stim}$ tends to occur from ~0–5 μs.

The detection efficiency of the impulse detection method using amplitudes was initially based on $N > 100$ different events used to calculate the transfer function $T_{d}$. It was initially assumed that the detection efficiency would increase with the increase of $N$ until the detection is stable. However, the processed results show that the transfer function calculated based on $N = 10$ events is almost as effective as the transfer function calculated based on $N = 100$ events. As a result, it was decided to vary $N$ from 10 to 100 in steps of 10 to estimate an experimental uncertainty for the detection efficiency. To make sure that at least 100 events are available as test events, each distance group needs to have more than 200 events for a rigorous test. Therefore, there are 47 distance groups that are processed using the complex analytic signal. In total, 45 of the distance groups exhibit an average impulse detection efficiency of >80% and 33 groups exhibit an average impulse detection efficiency of >90%. The detection efficiencies >90% are found for the distance groups <240 km and >920 km. The distance groups from 510 to 830 km exhibit average detection efficiencies ~88% with a relatively large of standard deviation ~6% when compared to the uncertainties of detection efficiencies at other distances. This uncertainty may be caused by the ground and first sky wave ambiguity for two reasons: (a) The time delay between the ground wave and first sky wave decreases which results in an apparent “wave crowding”. (b) The ground wave exhibits a larger attenuation than the first sky wave such that the ground and sky wave in this distance range exhibit comparable amplitudes.
For the distance groups >920 km, the detection efficiency increases again, because the first skywave dominates the waveform of the sferic.

This novel impulse detection methodology is able to (a) detect the weak event with high accuracy (b) detect multiple events within a short time range (c) avoid polarity misidentification compared with the traditional detecting method and distinguish the +CG event even if it only rarely happens. This method potentially can contribute to distinguishing IC events as the averaged waveforms in this work only uses −CGs.

6. Comparison Between Amplitude Waveform Bank and Coherency Waveform Bank

Section 3 and Section 4 introduced two types of waveform banks, based on the lightning amplitude and the coherency that was derived from the phase of analytic lightning signal. This section will investigate the differences between the amplitude and phase by evaluating the waveforms from the amplitude waveform bank and coherency waveform bank. The comparison between the waveforms at a distance of 1,210 km are used for an exemplary illustration (Figure 5). On the linear scale, the peak amplitudes of consecutive sky waves drop off faster with time when compared to the peak coherencies of the sky waves (Figure 5a). As a result, the sky waves appear to be more readily apparent in the coherencies when compared to the amplitudes. However, on the logarithmic scale, the peak coherencies of the sky waves appear to be more similar to the peak amplitudes (Figure 5b). It is interesting to note that the noise level of the coherency can be calculated directly from the number of averaged waveforms N because the coherency of normally distributed number is proportional to $1/\sqrt{N}$. For example, for the event number threshold N = 100 used in this work, the theoretical noise level for the coherency is $\sim 1/\sqrt{100} = 0.1$. This noise level is $\sim 10\%$ of the the theoretical maximum coherency. For the amplitude noise level, the receiver used in this work is remotely located with minimum noise level. In addition, the major noise source, which is power line harmonic radiation, is filtered out before average process. This leads to a relatively low noise level in the amplitude waveform. Therefore, the noise level is not used as the reference when comparing amplitude and coherency. The ground wave is also not suitable to be taken as reference, because the ground wave exhibits more attenuation effects caused by Earth curvature and conductivity. The first skywave is used as the reference for comparison such that both figures use the peak amplitude of the first sky wave as the upper limit for the y-axis. The envelope is used to represent the amplitude waveform for a better comparison. The ground wave is marked with black dot and eight skywaves are marked with red dots. The upper signal is the amplitude waveform and the bottom signal is the coherency waveform. Both waveforms exhibit the large ground wave and a few consecutive skywaves.

The envelope of the skywaves exhibits an exponential decay for the amplitude waveform such that the ratios between the consecutive sky waves is a fixed value which is named the “common ratio”. This is a common feature...

---

**Figure 5.** Comparison between amplitude waveform bank (upper) and coherency waveform bank (bottom) at 1,210 km on the (a) linear scale and (b) log scale. Black dots are ground waves, red dots are skywaves.
for the amplitude waveform bank. The common ratios are calculated based on different distances. For the amplitude waveform of each distance, the amplitudes of the first five skywaves are used to calculate amplitude ratios $R_i$:

$$R_i = \frac{A_{i+1}}{A_i} \quad i = 1, 2, 3, 4$$

(9)

where the $A_i$ is the amplitude of $i$th skywave. Then the common ratio $R_c$ is determined by minimizing the root mean square value $S_r$:

$$S_r = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (R_i - R_c)^2}$$

(10)

The common ratio between the skywaves amplitude for the amplitude waveform is positively correlated to the distance. Except for four distance groups, the common ratio increases linearly from $\sim-6.06$ dB to $\sim-2.88$ dB with a mean value of $-5.12$ dB.

The coherency waveform indicates that the coherency value of the skywaves has a linear decay. This means that there is a constant difference between the successive terms and this constant difference is named the “common difference”. The coherency differences $D_i$ are calculated using the first five skywaves of each coherency waveform:

$$D_i = \text{coh}_{i+1} - \text{coh}_i \quad i = 1, 2, 3, 4$$

(11)

where the $\text{coh}_i$ is the coherency of $i$th skywave. The common difference $D_c$ can be calculated using the same minimum root mean square value technique:

$$S_d = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (D_i - D_c)^2}$$

(12)

The coherency common difference $D_c$ is not distance-dependent in that it fluctuates between $\sim-0.21$ and $\sim-0.11$ with an average value of $-0.17$. Compared to the exponential decay of the skywaves amplitude in the amplitude waveform, the coherency waveform has a near linear decay of the skywaves coherency value which leads to a better prominence for higher order skywaves on the linear scale. The coherency waveform bank shows up to nine skywaves for large distances in Figure 2c.

This study shows that the skywaves of amplitude waveforms have an exponential decay, which is the reason why the amplitude waveform does not exhibit pronounced high order skywaves on the linear scale. However, the amplitude waveform has the same ability to exhibit as many skywaves as the coherency waveform on the log scale. An example of this comparison at the same distance of 1,210 km is shown in Figure 5b. In this case, both waveforms clearly exhibit eight skywaves.

In Figure 5a, the amplitude of the first skywave is significantly larger than the ground wave such that there is a possibility that the ground wave is mistakenly considered as a small fluctuation on the rising edge of the ground wave while the first skywave in this figure is treated as the “ground wave”. In contrast, the coherency of the first skywave is slightly larger than the coherency of the ground wave. The amplitude ratio and coherency ratio between the first skywave and the ground wave on the linear scale are studied for the amplitude waveform bank and coherency waveform bank in detail (Figure 6a). The amplitude ratio of the ground wave to the first skywave, which is $1/R_p$, is inversely proportional to the distance in that it decreases from $\sim3.43$ to $\sim0.24$ ($\sim1/4.17$) with the distance increasing from 190 km to 1,220 km. The amplitude of the first skywave exceeds the ground wave amplitude after 540 km. The coherency ratio of the ground wave to the first skywave is also negatively related to the distance that it decreases from $\sim1.26$ to $\sim0.92$ for the same distance range.

When calculating the coherency, the waveforms are not scaled to the ground wave maximum which means the coherency only measures the phase stability without considering the peak current. The phase can therefore be a better measurement than the amplitude for weak events.

Some conclusions can be drawn from the above analysis: (a) The phase is well preserved over long propagation distances, (b) the phase is less distorted by the ionosphere reflection when compared to the amplitude and (c) the
The measured time delay between the ground wave and skywaves can be used to determine the ionospheric height (e.g., Schonland et al., 1940; Smith et al., 2004). As the skywaves are reflected at different locations of the ionosphere, this method estimates the average ionospheric height during the entire propagation path. For each available distance, both the amplitude waveform and the coherency waveform are used to extract time stamps of the ground wave and the first five skywaves. The measured time differences $\Delta t_{di}$ between the $i$th skywave time stamp $t_{di}$ and the ground wave time stamp $t_{dG}$ can be calculated in Equation 13

$$\Delta t_{di} = t_{di} - t_{dG} \quad i = 1, 2, 3, 4, 5$$

For a simulated ionospheric height $h$, the theoretical time delays $\Delta t_{di}'$ between each skywave and the ground wave can be derived using the spherical earth model in Equation 14 (Schonland et al., 1940)

$$\Delta t_{di}' = \sqrt{\frac{4 \times i^2 \times h^2 + d^2 \times (1 + h/r) - d}{c}} \quad i = 1, 2, 3, 4, 5.$$  \hspace{1cm} (14)

$d$ represents the propagation distance in km, $r$ refers to the earth radius in kilometer and $c$ is the speed of light under the assumption that lightning propagation velocity equals the speed of light.

The timing uncertainty $\Delta T$ measures the differences between the measured time differences $\Delta t_{di}$ and the theoretical time differences $\Delta t_{di}'$ of the simulated ionospheric height $h$. As each time the emitted electromagnetic waves reflect from the bottom area of the ionosphere leads to an energy loss, the time stamps of lower order skywaves are supposedly more accurate when compared to the higher order skywaves for amplitude waveforms. Therefore, when it comes to calculating the time uncertainty $\Delta T$, a weight factor $W_i$ is applied. As mentioned previously, the mean common ratio between skywave amplitudes is 0.31 on the linear scale, that is, −5.12 dB, the weight factor $W_i$ for the amplitude waveform is defined in Equation 15

$$W_i = (0.31)^{i-1} \quad i = 1, 2, 3, 4, 5.$$  \hspace{1cm} (15)

The corresponding weight factor for the coherency waveform is taken to be 1.

The timing uncertainty $\Delta T$ is therefore defined in Equation 16

$$\Delta T = \frac{1}{n} \sum_{i=1}^{5} \left( \frac{\Delta t_{di}' - \Delta t_{di}}{W_i} \right)^2 \times W_i \quad i = 1, 2, 3, 4, 5.$$  \hspace{1cm} (16)

Figure 6. (a) The ratio between the ground wave and the first skywave. Black dots are from coherency waveforms, red dots are from amplitude waveforms. (b) Ionospheric height determination. The black dots are the ionospheric heights determined by the coherency waveform bank and the red dots are the ionospheric heights determined by the amplitude waveform bank.
Different ionospheric heights \( h \) are simulated to find the minimum timing uncertainty \( \Delta T \). The least square method is used to derive the best fit ionospheric height for each distance. The resulting ionospheric heights are illustrated in Figure 6b. The amplitude and coherency determined average ionospheric height \( h \) over all available distance groups are 90.8 and 90.7 km and both methods give a standard deviation of ±1.2 km. These results match well with the mean ionospheric heights derived by Leal et al. (2017), who also used the data collected during summer nighttime. Their mean ionospheric heights calculated by their CG waveforms are 90–92 km. The ionospheric heights derived by the amplitude waveform bank and coherency waveform bank show comparable results. While the method using the amplitude waveform bank needs to take priority for low order skywaves, the method using the coherency waveform bank contributes equally by different order skywaves.

7. Conclusions

This contribution investigated a novel technique, that is, complex analysis, which will be used to study complex interferometry of long-range lightning location networks in future work. The complex analytic signal is determined using the Hilbert transform, and its phase information is used subsequently to calculate the coherency.

The event selection process is carried out on data collected as part of field work during the summer in 2019. The experimentally determined attenuation coefficient of the peak electric field recorded with a receiver in Rustrel is 5.11 dB/Mm. This value is lower than the attenuation reported by Kolmašová et al. (2016) which varies from 17.4 dB/Mm to 23 dB/Mm with a mean value of 21.5 dB/Mm. The most likely reasons for this are (a) the data we are using was collected during nighttime when the attenuation is smaller than during the daytime, (b) the receiver used in their work has a much higher frequency range, and high frequency components are more strongly attenuated.

Both, an amplitude waveform bank and a coherency waveform bank are investigated. Before calculating the waveform bank, the ratio \( R_E \) is determined for each lightning waveform to exclude rarely misidentified +CGs events. A complex sferics analysis is introduced. Averaged waveforms are used to detect lightning events by use of an impulse detection method. In total, 45 and 33 of the available 47 distance groups exhibit an average impulse detection efficiency of >80% and >90% respectively. This novel impulse detection methodology is able to detect weak events and multiple events in a short time range, and assist to avoid the rare polarity misidentification of CGs. Potentially, the method may help to distinguish ICs from −CGs events as the averaged waveforms only uses −CGs. In the future, the impulse detection methodology could be extended to detect lightning waveforms without a priory knowledge of the distance which might enable an estimation of the distance range solely based on the lightning waveform.

The amplitudes of skywaves exhibit exponential decay with a mean common ratio of −5.12 dB for the amplitude waveform on the linear scale. The coherency waveform on the linear scale indicates that the coherency of skywaves has a linear decay with an average common difference of −0.17. Therefore, the coherency waveform bank exhibits more pronounced higher order skywaves when compared to the amplitude waveform bank on the linear scale, that is, up to nine skywaves can be clearly distinguished in Figure 2c. However, both, the amplitude waveform bank and coherency waveform bank have the same ability to exhibit high order skywaves on the log scale (Figure 5b).

In long-range lightning detection networks, the skywaves can result in significant interference for the detection of lightning return strokes because (a) for long-distance propagation, the time differences between the ground wave and skywaves can be quite low, and this “wave crowding” effect can cause difficulties to distinguish the ground wave from skywaves in the waveform (b) the ground wave attenuates fast when the propagation distance increases as a result of the Earth curvature and ground conductivity. Therefore, the first skywave can be a significant interference when compared to the ground wave (Figure 5a). This characteristic is quantitatively analyzed. The ratios between the ground wave and the first skywave are inversely proportional to the propagation distances for both, the amplitude waveform bank and coherency waveform bank. The ratios decrease from ~3.43 to ~0.24 for the amplitude waveform bank, and range from ~1.26 to ~0.92 for the coherency waveform bank for the same distance range, that is 190 km to 1,220 km. This result shows that the coherency is a better parameter to detect the ground wave as the coherency is preserved over long propagation distances when compared to the amplitude waveform bank.
Acknowledgments

The work of X.B. was sponsored by URSA under the project Grant No. EA-E1250. The work of M.F. was sponsored by the Royal Society (UK) grant NMG/R1/180252 and the Natural Environment Research Council (UK) under grants NE/L012669/1 and NE/H024921/1. This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement 722337. X.B. conducted the analysis and wrote the paper, and M.F. assisted with conceiving the concept of the paper. The authors wish to thank Stéphane Pedebey at Meteorage for access to selected lightning flash data, the team of the Laboratory Souterrain a Bas Bruit (LSBB) in Rustrel for their generous hospitality, Adam Peverell for extracting and communicating the data, Yuying Liu, Mert Yircemiz, Serge Soula, Jean-Louis Pinçon, Franck Elle, and Matthieu Garnung for assisting with the acquisition of the electric field data used in this study, and the reviewers for their assistance to improve the quality of the manuscript.

References

Barr, R., Jones, D. L., & Rodger, C. J. (2000). ELF and VLF radio waves. Journal of Atmospheric and Solar-Terrestrial Physics, 62(17–18), 1689–1718. https://doi.org/10.1016/s1364-6826(00)00121-8

Betz, H. D., Schmidt, K., Laroche, P., Blanchet, P., Oettinger, W. P., Defer, E., et al. (2009). LINET—An international lightning detection network in Europe. Atmospheric Research, 91(2–4), 564–573. https://doi.org/10.1016/j.atmosres.2008.06.012

Burke, C. P., & Jones, D. (1992). An experimental investigation of ELF attenuation rates in the Earth-ionosphere duct. Journal of Atmospheric and Terrestrial Physics, 54(3–4), 243–250. https://doi.org/10.1016/0021-9169(92)90005-6

Burkholder, B. S., Hutchins, M. L., McCarthy, M. P., Pfaff, R. F., & Holworth, R. H. (2013). Attenuation of lightning-produced sferics in the Earth-ionosphere waveguide and low-latitude ionosphere. Journal of Geophysical Research: Space Physics, 118(6), 3692–3699. https://doi.org/10.1002/jgra.50351

Caligaris, C., Delfino, F., & Procopio, R. (2008). Cooray-rubenstein formula for the evaluation of lightning radial electric fields: Derivation and implementation in the time domain. IEEE Transactions on Electromagnetic Compatibility, 50(1), 194–197. https://doi.org/10.1109/t EMC.2007.913226

Chowdhuri, P., Anderson, J., Chisholm, W., Field, T., Ishii, M., Martinez, J., et al. (2005). Parameters of lightning strokes: A review. IEEE Transactions on Power Delivery, 20(1), 346–358. https://doi.org/10.1109/tpwrd.2004.835039

Cooray, V. (2009). Propagation effects due to finitely conducting ground on lightning-generated magnetic fields evaluated using Sommerfeld integrals. IEEE Transactions on Electromagnetic Compatibility, 51(3), 526–531. https://doi.org/10.1109/t EMC.2009.2019759

Crombie, D. D. (1965). On the use of VLF measurements for obtaining information on the lower ionosphere (especially during solar flares). Proceedings of the IEEE, 53(12), 2027–2034. https://doi.org/10.1109/PROC.1965.4479

The calculation of the amplitude waveform bank requires scaling the waveform to the ground wave maximum. This step makes sure that the averaged waveform is not dominated by strong events. The calculation of the coherency waveform bank does not need this extra step and this means that the coherency calculation has the potential to detect weak events, which will be studied in future work. However, the coherency calculation is based on averaged events data while the amplitude waveform can be extracted for an individual event.

Both of the waveform banks are used to calculate ionospheric heights. The average ionospheric heights over the entire propagation path determined by the amplitude waveform bank and the coherency waveform bank are 90.8 and 90.7 km with the same standard deviation of 1.2 km. These results match well with the mean ionospheric heights calculated by Leal et al. (2017), which are 90–92 km. The methodology used here shows that to achieve these comparable results, the amplitude waveform bank needs to prioritize lower order skywaves and the method using the coherency waveform bank contributes equally to different order skywaves.

This work studied the amplitude waveform bank and coherency waveform bank in detail. In the future, the coherency will be used to produce a coherency map by shifting the recorded waveforms from a set of receivers to plot the coherency at different pixels of a map. This method will use more phase information of lightning sferics compared with the traditional method that only picks a single point for each event. Such simulation can be time dependent such that it is possible to produce a dynamic coherency map which potentially can offer more detailed lightning information than currently available.

8. Summary

The complex analysis of sferic waveforms is studied in this contribution using the coherency. Amplitude and coherency waveform banks are calculated based on distant-dependent measurements collected during the summer 2019. A lightning detection method is investigated using the averaged waveforms in the amplitude waveform bank and the corresponding lightning detection efficiencies are quantified. The amplitude waveform bank is compared in detail to the coherency waveform bank. Both waveform banks exhibit distant-dependent ground waves and numerous consecutive sky waves. Higher order sky waves are readily apparent in the coherency waveform bank, in particular at relatively large distances, similar to the logarithmic amplitudes in the amplitude waveform bank. The ground wave and first sky waves exhibit similar amplitudes at distances ranging around ~540 km. At larger distances the ground wave is more attenuated than the sky waves and the time delay between the ground wave and sky waves decreases such that interference occurs. This study lays the foundation for using the coherency of sferic waveforms toward two-dimensional interferometric methods based on the complex analytic signal for use in long-range lightning detection networks.

Data Availability Statement

The data used for this publication will be available from https://doi.org/10.15125/BATH-01057.
Weidman, C. D., & Krider, E. P. (1986). The amplitude spectra of lightning radiation fields in the interval from 1 to 20 MHz. *Radio Science, 21*(6), 964–970. https://doi.org/10.1029/RS021i006p00964

Zhou, X., Wang, J., Ma, Q., Huang, Q., & Xiao, F. (2021). A method for determining D region ionosphere reflection height from lightning skywaves. *Journal of Atmospheric and Solar-Terrestrial Physics, 221*, 105692. https://doi.org/10.1016/j.jastp.2021.105692

Zhu, Y., Stock, M., & Bitzer, P. (2021). A new approach to map lightning channels based on low-frequency interferometry. *Atmospheric Research, 247*, 105139. https://doi.org/10.1016/j.atmosres.2020.105139