Research Article | Vol. 26, No. 26 | 24 Dec 2018 | OPTICS EXPRESS 33970
Optics EXPRESS

Coherent detection with an incoherent local oscillator

ANTONIO MECOZZI,[1,*] AND MARK SHTAIF[2]

[1]Department of Physical and Chemical Sciences, University of L'Aquila, L'Aquila 67100, Italy
[2]School of Electrical Engineering, Tel Aviv University, Tel Aviv 69978, Israel
[*]antonio.mecozzi@univaq.it

Abstract: We demonstrate the feasibility of fully coherent reconstruction of the complex envelope of arbitrary optical fields while using an incoherent source as a local oscillator (LO). The reconstruction relies on a signal processing procedure that we describe, and the only requirement from the system is that the receiver's electrical bandwidth and sampling rate are at least twice as high as the bandwidth of the received signal and of the LO. The proposed scheme is particularly attractive in spectral regions where no high-quality lasers are available.

1. Introduction

Coherent receivers have a longstanding story. The "beat receptor," the first heterodyne receiver, was invented by Nikola Tesla in 1896 [1], and from that time on, various forms of coherent receivers, heterodyne, superheterodyne and homodyne have been widely used in radio communications. Yet, since all currently known forms of coherent detection [2] rely on the existence of a highly coherent reference signal, known as a local-oscillator (LO), the penetration of coherent receivers to the optical domain was delayed until after the invention of the laser in the nineteen sixties [3]. Even today, when coherent receivers became ubiquitous in modern optical communications [4], coherent detection remains difficult in spectral regions in which lasers of adequate coherence are not widely available. A relevant example is the terahertz region, with applications to spectroscopy, astronomy, and medical imaging [5, 6].

In this paper, we propose a novel detection method that allows the accurate reconstruction of complex waveforms without relying on the availability of a coherent source. In particular, we show that with the proposed procedure, the local oscillator can be any band-limited signal, including the signal obtained from a filtered thermal source. The method is implemented by means of digital signal processing and hence the only practical requirement is that the bandwidths of both the signal and the LO remain smaller than half of the available sampling rate so as to allow lossless sampling.

2. Receiver description and principle of operation

A standard homodyne coherent receiver admits two optical fields at its input and produces two electrical signals at its output. The optical inputs are the signals $a(t)\exp(-i\omega_0 t)$, which is to be measured, and $b(t)\exp(-i\omega_0 t)$, which functions as a local oscillator. The electric outputs are $\mathrm{Re}\{I_{ab}(t)\}$ and $\mathrm{Im}\{I_{ab}(t)\}$, where $I_{ab}(t) = a(t)b^*(t)$, and the asterisk * denotes complex conjugation. In a traditional balanced homodyne receiver, the local oscillator is coherent to the extent that its complex envelope $b(t)$ can be treated as a known constant b, whose phase is set to 0 for reference. In this case the electric outputs reduce to $b\mathrm{Re}\{a(t)\}$ and $b\mathrm{Im}\{a(t)\}$, thereby yielding the real and imaginary parts of the unknown signal.

In this work we consider a scheme where both complex envelopes $a(t)$ and $b(t)$ are reconstructed without any of them being known in advance. Our scheme relies on a standard homodyne receiver that is supplemented by a direct measurement of the intensity of the incoming optical signal $I_{aa}(t) = |a(t)|^2$, as illustrated in Fig. 1. The most general situation with regard to the spectra of

#338773 https://doi.org/10.1364/OE.26.033970
Journal © 2018 Received 19 Jul 2018; revised 27 Nov 2018; accepted 2 Dec 2018; published 13 Dec 2018

Optics EXPRESS

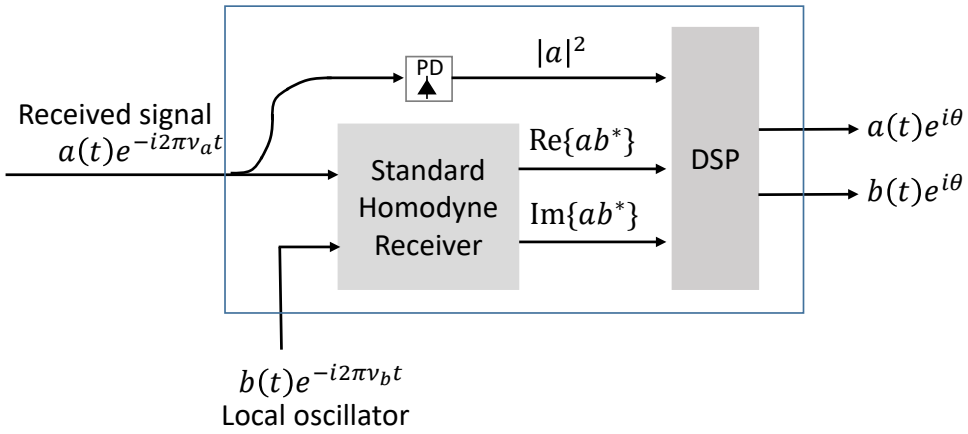

Fig. 1. Schematic of the proposed receiver. The received signal $a(t)\exp(-i2\pi\nu_a t)$ and the local oscillator $b(t)\exp(-i2\pi\nu_b t)$ are mixed in a standard homodyne receiver whose electric outputs are the real and imaginary parts of the complex signal $I_{ab}(t) = a(t)b^*(t)$. This signal together with the intensity $I_{aa}(t) = |a(t)|^2$, directly measured with a photo-diode (PD), are fed into a digital signal processing (DSP) unit whose outputs are the digitized versions of the complex envelopes $a(t)$ and $b(t)$ up to an arbitrary, time independent phase offset θ. Both $a(t)$ and $b(t)$ are arbitrary complex band-limited signals.

the two signals is that the spectrum of $a(t)$ is contained between ν_a and $\nu_a + B_a$, whereas the spectrum of $b(t)$ is contained between ν_b and $\nu_b + B_b$, where $\nu_{a,b}$ are reference frequencies and $B_{a,b}$ are bandwidths. In practice, since photo-detection is not sensitive to the absolute frequency, we can artificially set $\nu_a = 0$, and replace ν_b with $\Delta\nu$ – the difference between the reference frequencies. The minimum required sampling rate of the real-valued intensity signal $I_{aa}(t)$ is equal to its electric bandwidth $B_s = 2B_a$, whereas the minimum sampling rate needed in order to accommodate the real and imaginary parts of $I_{ab}(t)$ is $B_s = 2\max(|B_b + \Delta\nu|, |B_a - \Delta\nu|)$ [7]. It is assumed that the receiver electronics is sufficiently fast to accommodate this bandwidth. After photo-detection and sampling, the effect of the frequency difference $\Delta\nu$ can be removed by digitally multiplying $I_{ab}(t)$ by $\exp(-2\pi i\Delta\nu t)$, which allows resampling $I_{ab}(t)$ at a reduced sampling rate of $B_s = B_a + B_b$. Hence, in what follows, without any loss of generality, we will assume that $\Delta\nu = 0$, so that the spectra of $I_{aa}(t)$ and $I_{ab}(t)$ are contained in $\nu \in (-B_a, B_a)$ and $\nu \in (-B_a, B_b)$, respectively.

In the subsection that follows we describe the principle of operation in an idealized case where the signals $a(t)$ and $b(t)$ are periodic with the period being equal to the processed time window. Subsequently, in Sec. 4 we evaluate the performance of this scheme numerically in a more realistic setting, which includes the effects of detection noise and removes the periodicity assumption. The performance is found to be impressively good in a broad range of cases.

2.1. Properties of periodic band-limited waveforms

In this section we will assume that $a(t)$ is periodic with a period T and that the product $B_a T = M_a$ is an integer. In this case $a(t)$ is rigorously represented by a Fourier series with a finite number of elements

$$a(t) = \sum_{k=0}^{M_a-1} F_a(k)e^{-ik\Omega t}, \tag{1}$$

Optics EXPRESS

where $\Omega = 2\pi/T$ and where the Fourier coefficients are given by

$$F_a(k) = \frac{1}{T}\int_0^T a(t)e^{ik\Omega t}\mathrm{d}t. \tag{2}$$

Equation (1) assumes that the highest frequency in $a(t)$ is $B_a(1 - {M_a}^{-1})$, and it approaches B_a in the practically relevant cases, when $M_a >> 1$. We now define the *Z-extension* of the function $a(t)$ as

$$\mathcal{Z}_a(Z) = \sum_{k=0}^{M_a-1} F_a(k)Z^k. \tag{3}$$

The Z-extension is a tool that we use repeatedly in what follows, and it is equal to the familiar Z-transform of the function's Fourier coefficients $\{F_a(k)\}_{k=0}^{M_a-1}$. Clearly, $a(t) = \mathcal{Z}_a\big(\exp(-i\Omega t)\big)$, and hence special attention needs to be paid to cases in which the value of Z is on the unit circle. When $F_a(M_a - 1) \neq 0$, $\mathcal{Z}_a(Z)$ is an $M_a - 1$ degree polynomial which has $M_a - 1$ roots $\{Z_{a,k}\}_{k=1}^{M_a-1}$, and therefore it can be expressed as

$$\mathcal{Z}_a(Z) = \frac{F_a(0)}{\prod_{k=1}^{M_a-1}(-Z_{a,k})}\prod_{k=1}^{M_a-1}(Z - Z_{a,k}). \tag{4}$$

In what follows we will relate to the roots of $\mathcal{Z}_a(Z)$ as the *zeros of* $a(t)$. Equation (4) together with the relation $a(t) = \mathcal{Z}_a\big(\exp(-i\Omega t)\big)$, demonstrates that up to an immaterial multiplicative constant, $a(t)$ *is uniquely determined by its zeros*.

At this point we present three simple properties of the Z-extension that are instrumental for the purpose of waveform reconstruction. The proofs of these properties are fairly simple and they are outlined in Appendix 7.1.

- The zeros of $a^*(t)$ are $\{1/Z_{a,k}^*\}_{k=1}^{M_a-1}$. Namely, they are the inverse conjugates of the zeros of $a(t)$.
- The intensity $I_{aa}(t) = |a(t)|^2$ has $2(M_a - 1)$ zeros, which are the union of the zeros of $a(t)$ and the zeros of $a^*(t)$. Namely, $\{Z_{a,k}, 1/Z_{a,k}^*\}_{k=1}^{M_a-1}$.
- Let $b(t)$ be a second periodic band-limited signal, whose spectrum is contained in $\omega \in [0, 2\pi B_b]$, and whose period is an integer number of B_b^{-1}. If one further assumes that the ratio B_b/B_a is a rational number, then T can be set to be the joint period of both $a(t)$ and $b(t)$ with $TB_a = M_a$ and $TB_b = M_b$ being integers. In this case the set of zeros of $I_{ab}(t) = a(t)b^*(t)$ is the union of the zeros of $a(t)$ and the inverse conjugate zeros of $b(t)$. Namely, $\{Z_{a,k}\}_{k=1}^{M_a-1} \bigcup \{1/Z_{b,k}^*\}_{k=1}^{M_b-1}$, so that there are $M_a + M_b - 2$ zeros in total.

The above properties imply that knowledge of the zeros of $I_{aa}(t)$ and $I_{ab}(t)$ is sufficient for determining the zeros of $a(t)$ and $b(t)$, and hence the waveforms $a(t)$ and $b(t)$ themselves, up to an immaterial constant phase offset [8]. The idea behind doing that is the following: The zeros of $a(t)$ are those that are common to $I_{aa}(t)$ and $I_{ab}(t)$, whereas the zeros of $b(t)$ are the inverse conjugates of the zeros that are present in $I_{ab}(t)$, but not in $I_{aa}(t)$ [9].

2.2. Finding the zeros of $I_{aa}(t)$ and $I_{ab}(t)$

As noted earlier, the spectra of $I_{aa}(t)$ and $I_{ab}(t)$ occupy the frequency ranges $\omega \in (-2\pi B_a, 2\pi B_a)$ and $\omega \in (-2\pi B_b, 2\pi B_a)$, respectively. Therefore

$$I_{ax}(t) = \sum_{k=-M_x+1}^{M_a-1} F_{ax}(k)e^{-ik\Omega t}, \tag{5}$$

Optics EXPRESS

where x stands for either a or b. The Z-extension of $I_{ax}(t)$ is given by

$$\mathcal{Z}_{I_{ax}}(Z) = \frac{1}{Z^{M_x-1}} \sum_{k=0}^{M_a+M_x-2} F_{ax}(k - M_x + 1)Z^k, \tag{6}$$

and its relevant zeros are the roots of the polynomial appearing in the summation. Note that since the highest power of Z in the summation is higher than $M_x - 1$, the term $Z^{-(M_x-1)}$ does not add zeros at infinity, and the pole singularity that it produces at $Z = 0$ is immaterial to our study. The polynomial coefficients $F_{ax}(k)$ can be obtained from Eq. (2) with $a(t)$ replaced by $I_{ax}(t)$, but a more practical approach follows from noting that $F_{ax}(k)$ is identical to the k-th inverse digital Fourier-transform coefficient of the series $\{I_{ax}(nT/N)\}_{n=0}^{N-1}$, where $N = M_a + M_x$ [10]. This relation allows the trivial extraction of $F_{ax}(k)$ from the standard fast-Fourier-transform of the sampled $I_{ax}(t)$.

A simple illustration of the method is provided in Fig. 2. Two complex periodic signals $a(t)$ and $b(t)$ were generated by inverse Fourier transforming two independent realizations of white noise that is filtered to one half of the sampling frequency. Their real and imaginary parts are shown in Fig. 2a and 2b, the detected field with solid blue lines, and the reconstructed fields with dashed red lines. Fig. 2c shows the power spectra of the two signals. Fig. 2d shows the zeros of $I_{aa}(t)$ by blue stars and crosses, and the zeros of $I_{ab}(t)$ by red circles and squares. The stars and the circles show the zeros that are common to both $I_{aa}(t)$ and $I_{ab}(t)$, whereas the crosses show the zeros of $I_{aa}(t)$ that are the inverse conjugate zeros of $a(t)$ and the squares show the zeros of $I_{ab}(t)$ that are the inverse conjugates of the zeros of $b(t)$. The reconstruction ideally reproduces the waveforms in Figs. 2a and 2b, provided that the constant phase offset is removed [11].

3. The case of practical signals

All of the above theory was derived for the idealized case of periodic signals, which does not coincide with the situation in practice. In the case of non-periodic signals, the consideration of a finite time window affects the spectrum in a way that prevents the treatment of the processed signals as band limited. The reason for this is that when extracting a finite interval of a band-limited function, the effect in the Fourier domain is that of convolving the spectrum of the original waveform with the spectrum of a square time-window, which is a sinc function of infinite support. In order to reduce the impact of this bandwidth expansion we introduce two additional elements into our field reconstruction scheme. These are smooth time windowing and oversampling. The time windowing stage involves multiplication of both signals by a function $w(t)$ that gradually decays to zero towards its edges. Time windowing produces a new pair of signals $a'(t) = a(t)w(t)$ and $b'(t) = b(t)w(t)$. Owing to the smooth decay of $w(t)$, these two waveforms and their low-order derivatives, are nearly zero on both sides of the measured time-frame, and hence they can be approximately viewed as single periods of periodic functions whose period T coincides with the measured time-frame. Oversampling needs to be introduced in order to account for the fact that the bandwidths of $a'(t)$ and $b'(t)$ are larger than B_a and B_b because of the time windowing function $w(t)$. In principle, multiplication by $w(t)$ can be performed by means of a pair of optical modulators, but a more practical approach would be to multiply $I_{aa}(t)$ and $I_{ab}(t)$ by $w^2(t)$ in the electrical domain (where, without loss of generality, $w(t)$ is taken to be real-valued). This operation is particularly convenient because it can be implemented digitally, after sampling the received waveforms. The effect of time windowing is undone after reconstruction, by means of dividing the reconstructed waveforms by $w(t)$.

4. Numerical examples

In this section we describe the performance of the field reconstruction algorithm in cases where both the signal $a(t)$ and the local oscillator $b(t)$ are random Gaussian processes characterized

Optics EXPRESS

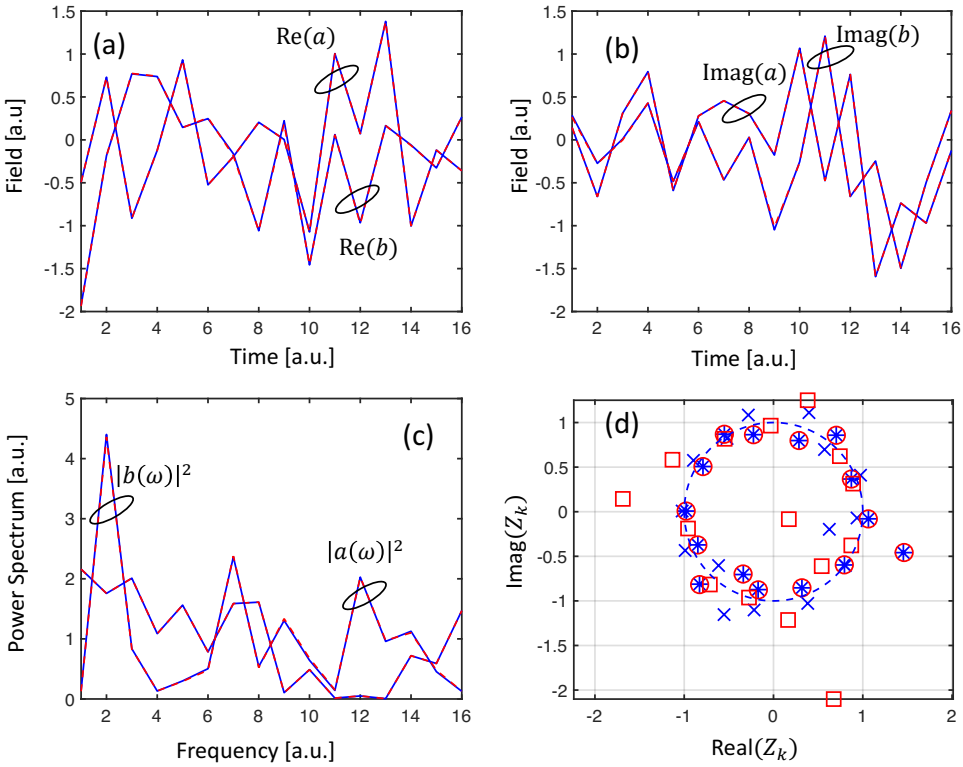

Fig. 2. The real (a) and imaginary (b) parts of $a(t)$ and $b(t)$, the detected field shown with solid blue lines, the reconstructed fields with dashed red lines. (c) Shows the power spectra of the two waveforms and (d) shows the zeros of $I_{aa}(t)$ (crosses and stars) and of $I_{ab}(t)$ (circles and squares). The blue stars and the red circles are the zeros of $a(t)$ in $I_{aa}(t)$ and $I_{ab}(t)$ respectively, the blue crosses are the inverse conjugate of the zeros of $a(t)$ in $I_{aa}(t)$ and the red squares are the inverse conjugates of the zeros of $b(t)$ in $I_{ab}(t)$. The reconstruction of $a(t)$ and $b(t)$ on the basis of the their zeros as found in (d) is perfect, as long as the intensities are correctly normalized and the constant phase-offset is eliminated (as in traditional homodyne). The signals where generated with $B_a = B_b = 16/T$.

Optics EXPRESS

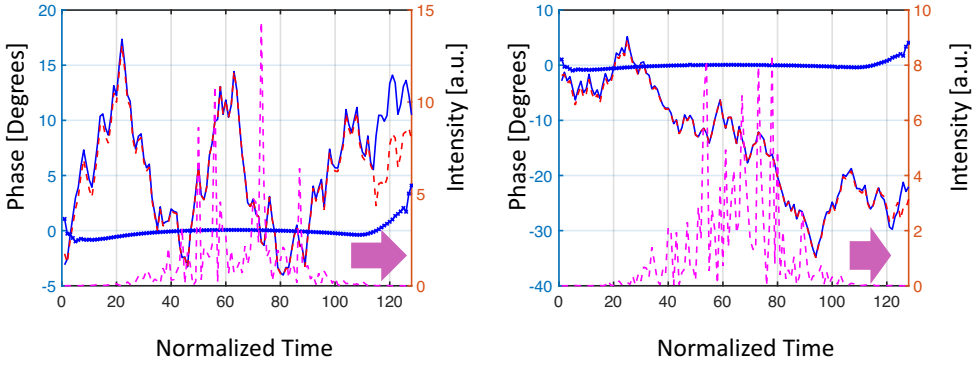

Fig. 3. The phase of the original signal (solid blue lines), the reconstructed phase (dashed red lines), and the phase reconstruction error (blue "+" signs). The intensity waveform multiplied by the time windowing function $w^2(t)$ is shown for reference by the purple dashed curve. The left (a) and right (b) panels relate to the reconstruction of $a(t)$ and $b(t)$, respectively.

by square-shaped spectral densities. For the reader's convenience, the code implementing the algorithm is provided in [12]. In the results shown in this paper we used $B_a = B_b$, and a sampling rate $B_s = 2.064B_a$. In general, for good performance, the sampling rates must be high enough in order to adequately accommodate the signal's bandwidth (including the broadening caused by the time-windowing), but not much higher than that. The reason is that a higher sampling rate implies an artificial increase of the order of the polynomials and as a result it leads to a higher sensitivity to errors in the estimated values of the obtained roots. The width of the processed time window was $T = 128B_s^{-1}$, and the time windowing function was Gaussian $w(t) = \exp[-(t - T/2)^2/(2T_0^2)]$, with $T_0 = 32B_s^{-1}$. In order to account for detection noise and imperfect balance between detectors, we used $I_{ab}(t) = (a(t) + n_a(t))b^*(t)$, where $n_a(t)$ is Gaussian noise which is independent of $a(t)$ and $b(t)$ and whose power spectrum is square and identical in shape to the spectrum of $a(t)$. The average power of $n_a(t)$ is taken to be 40 dB below the average power of $a(t)$. Notice that $n_a(t)$ does not represent the noise that accompanies the received signal, but rather it is only meant to account for the difference between the way in which $a(t)$ affects the measurements of $I_{aa}(t)$ and of $I_{ab}(t)$ as a result of detector noise and detector imbalance. As the amplitude of $a(t)$ can be directly obtained from the measurement of the intensity $|a(t)| = \sqrt{I_{aa}(t)}$, the task of reconstructing the waveform $a(t)$ is equivalent to reconstructing its phase. In Fig. 3 we show an example for the accuracy with which the proposed scheme reproduces the optical phase of the signal. The solid blue curve shows the optical phase of the original signal, whereas the dashed red curve shows the phase of the reconstructed signal. We remove the constant (time independent) offset between the original and reconstructed phases, and show the phase reconstruction error in the curve marked by the blue "+" signs. The dashed purple curve shows the input signal's intensity multiplied by the windowing function $w^2(t)$ and unlike all of the other curves in the figure, it corresponds to the right-side axis. The left and right panels correspond to the reconstruction of $a(t)$ and $b(t)$, respectively. The accuracy of the phase reconstruction is seen to be excellent everywhere, except near the edges where $|w(t)|$ is small. The quality of the final reconstruction of the complex fields can be seen in Fig. 4. The real and imaginary parts of the same $a(t)$ and $b(t)$ that were used in Fig. 3 are shown before and after the reconstruction. Here the effect of the Gaussian time windowing function was removed by dividing by $w(t)$, and the temporal edges (i.e. $t \le T/2 - T_0$ and $t > T/2 + T_0$) are discarded. The excellent accuracy of the reconstruction is self-evident. The problem of finding the zeros of the sampled waveforms $I_{aa}(t)$ and $I_{ab}(t)$ in the non-idealized case is not trivial, and our method of solving it (which relies on an adaptation of a standard root

Optics EXPRESS

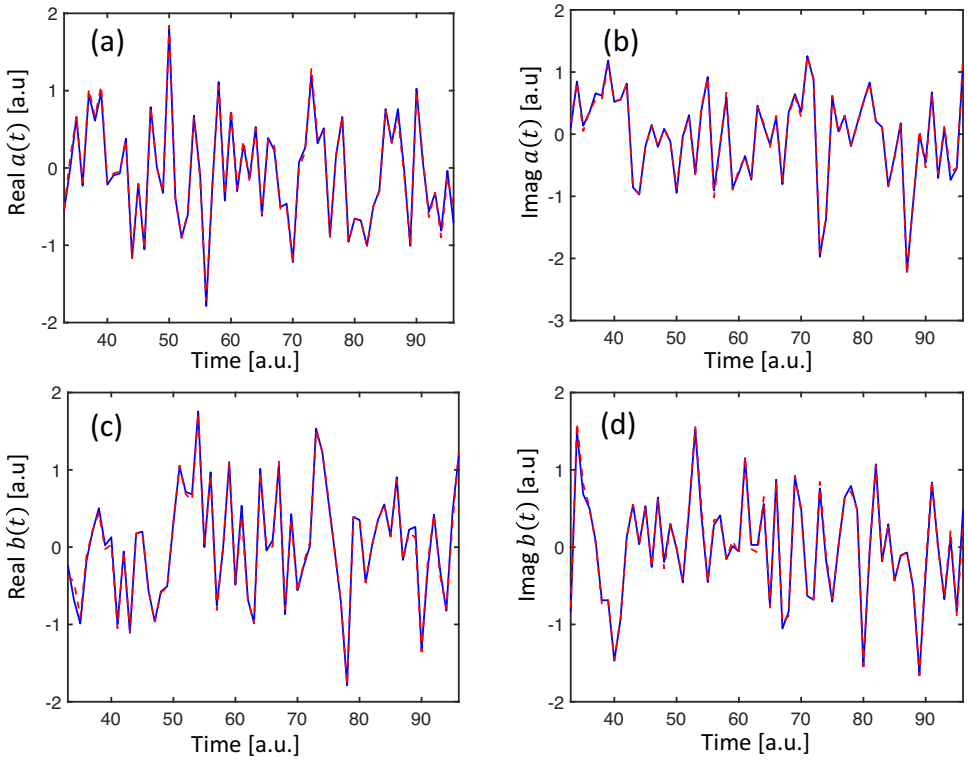

Fig. 4. Time domain profiles of the real and imaginary parts of the signals $a(t)$ and $b(t)$ after removing the effect of the time-windowing and discarding the temporal tails outside the range $t \in [T/2 - T_0, T/2 + T_0]$. The original and reconstructed waveforms are marked by the solid blue and dashed red curves, respectively, and they are almost indistinguishable in the figures (demonstrating the quality of the reconstruction).

finding algorithm [13]) is described in Appendix 7.2. Similarly challenging is the problem of determining which of the zeros of $I_{aa}(t)$ and $I_{ab}(t)$ should be identified as overlapping with each other. To solve it we formulate the zero-classification problem in a way that allows us to make use of the rectangular assignment algorithm described in [14–16]. The details of this procedure are also provided in Appendix 7.2. An exemplary result of these two procedures is shown in Fig. 5. The blue symbols (stars and crosses) show the zeros of $I_{aa}(t)$, whereas the red symbols (squares and circles) are the zeros of $I_{ab}(t)$. Unlike in the ideal case, the overlap between zeros of $I_{aa}(t)$ and $I_{ab}(t)$ is never perfect. Whenever the classification algorithm determines that a given zero of $I_{aa}(t)$ and a given zero of $I_{ab}(t)$ represent the same zero (which belongs to $a(t)$), they appear as a blue star and a red circle. A blue cross, on the other hand, indicates a zero of $I_{aa}(t)$ for which no matching zero of $I_{ab}(t)$ was found, and similarly a red square indicates a zero of $I_{ab}(t)$ for which there is no matching zero in $I_{aa}(t)$ (and hence its inverse conjugate is a zero of $b(t)$). In order to quantify the quality of the reconstruction, we look at the correlation function between the original waveform $a(t)$ and the reconstructed waveform $\hat{a}(t)$, which is defined as follows

$$\rho = \left| \frac{\int a^*(t)\hat{a}(t)\mathrm{d}t}{\sqrt{\int |a(t)|^2\mathrm{d}t \int |\hat{a}(t)|^2\mathrm{d}t}} \right|, \tag{7}$$

Optics EXPRESS

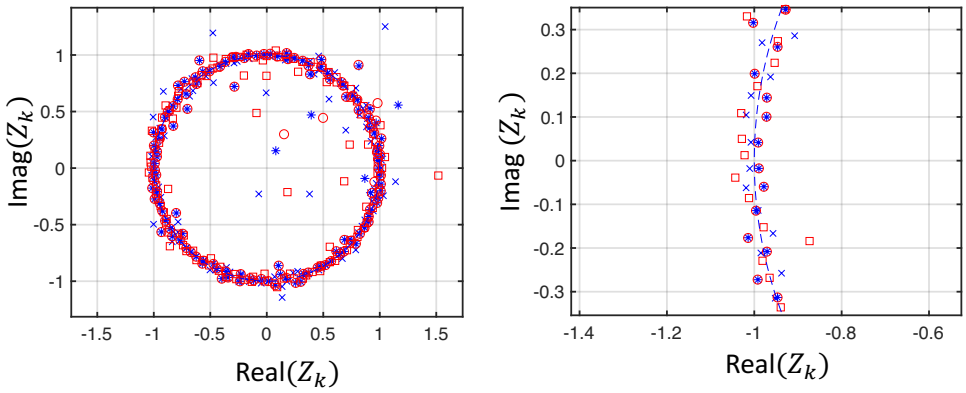

Fig. 5. The zeros of $I_{aa}(t)$ and of $I_{ab}(t)$. Blue star: zeros of $a(t)$ in $I_{aa}(t)$; blue crosses: zeros of $a^*(t)$ in $I_{aa}(t)$; red circles: zeros of $a(t)$ in $I_{ab}(t)$; red squares: zeros of $b^*(t)$ in $I_{ab}(t)$. The figure in (b) is a zoomed in version of (a).

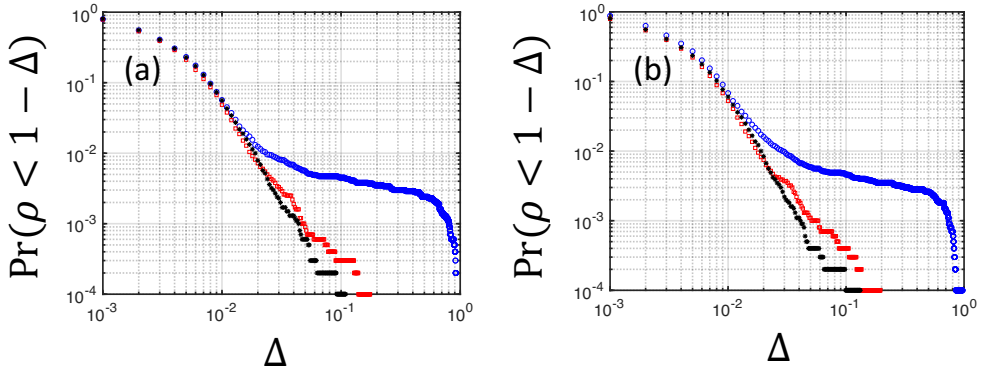

Fig. 6. The probability that $\rho < 1 - \Delta$ as a function of Δ with for $a(t)$ (left panel) and $b(t)$ right panel. The curves represent different noise to signal ratios (i.e. the ratio between the variance of the imbalance noise $n_a(t)$ and the average power of $a(t)$). Blue circles show $P_n/P_a = 10^{-4}$, red squares show $P_n/P_a = 10^{-6}$, and black stars show $P_n/P_a = 0$.

and whose allowed values range between 0 and 1 (with $\rho = 1$ indicating perfect reconstruction). In order to avoid the noisy regions near the tails of the windowing function $w(t)$ (see Fig. 3), all integrals are performed between $T/2 - T_0$ and $T/2 + T_0$, whereas outside this range (i.e. under the tails of the windowing function), the reconstructed signal is discarded. Figure 6 shows the probability that $\rho < 1 - \Delta$ for both $a(t)$ (left panel) and $b(t)$ (right panel). The displayed curves correspond to $P_n/P_a = 10^{-4}$, $P_n/P_a = 10^{-6}$, and $P_n/P_a = 0$, where P_n is the average power of the imbalance noise $n_a(t)$, and P_a is the average power of $a(t)$. Notice that al three curves overlap initially, but then the slope of the curves that include the effect of imbalance noise changes abruptly. In all cases the probability that ρ is less than 0.97 (i.e. $\Delta > 0.03$) is smaller than 1%.

5. Discussion

One relevant aspect of the proposed technique is its comparison with the differential receiver of the kind considered in [17–20]. The differential receiver beats two slightly delayed versions of the received field against one another and thereby recovers their phase difference. The idea is that

Optics EXPRESS

knowledge of the initial phase at $t = 0$ allows reconstruction of the phase at $t = \Delta t$, from which the phase at $t = 2\Delta t$ can be recovered and so on. There are two fundamental difficulties with this approach that are avoided in the scheme described in this manuscript. The first is that differential phase reconstruction fails when the sampled field's amplitude is equal (or close) to 0. The other is the accumulation of phase reconstruction errors from one sample to the next. Nonetheless, one can find an interesting relation between the differential approach and the scheme described in this paper if one notes that the local oscillator in our scheme can be set to be a delayed version of the signal, $b(t) = a(t - T_d)$. The only requirement is that the zeros of $a(t - T_d)$ do not overlap with the zeros of $a(t)$, which can be ensured if T_d is set to be greater than both the width of the processing time window and of the coherence time characterizing $a(t)$. However in practice much shorter values of T_d should be perfectly appropriate [21]. Another somewhat related scheme is that of the Stokes receiver which has been proposed in [22] and recently revived (see e.g. the recent review in [23]). It turns out that the Stokes receiver's scheme is equivalent to the scheme illustrated in Fig. 1 if $a(t)$ and $b(t)$ are set to be the two orthogonal polarizations of the same electric field [24]. Our analysis suggests that the fields can be reconstructed almost ideally, in which case the information capacity of an optical communications system using a Stokes receiver should approach the capacity of a fully coherent system. This conclusion differs from the one made in [25], where it was found that the information capacity of a system using a Stokes receiver is only three quarters of the fully coherent system's capacity. The difference may be attributed to the fact that in our implementation, the receiver bandwidth must be large enough to accommodate the spectral width of $I_a(t)$ and $I_{ab}(t)$, whereas in [25] the receiver's bandwidth is assumed to be equal to that of the incoming optical signals in the two polarizations $a(t)$ and $b(t)$.

Finally, it should be pointed out that a fundamental limitation of our proposed scheme is the inability to accurately determine the central frequency of the measured signal. As in standard coherent receivers, the frequency reference is set by the LO, except that in our case the LO is incoherent. Its bandwidth sets the maximum accuracy with which the central frequency of the measured signal can be determined.

6. Conclusions

We have proposed a method of performing coherent detection of complex-valued signals while relying on an incoherent local oscillator such as a thermal field. The scheme is entirely symmetric with respect to the roles of the measured signal and the local oscillator and in this sense it can be viewed as a method of coherently reconstructing the complex fields of two unknown waveforms that can be mixed with each other and detected using standard optical intensity detectors. The method requires well balanced low-noise photo-detection as well as an analog to digital converter using a sampling rate that is at least twice larger than the spectral width of each of the two optical signals.

7. Appendixes

7.1. Proof of the properties of the Z-extension

Here we outline the proof to the properties of the Z-extension that were listed in Sec. 2.1. The first property is that the zeros of $a^*(t)$ are the inverse conjugates of the zeros of $a(t)$. This property is seen by conjugating the expression for $a(t)$ in Eq. (1) and identifying Z with $\exp(-i\Omega t)$. Then it is evident from the definition in (3) that

$$\mathcal{Z}_{a^*}(Z) = \left[\sum_{k=0}^{M_a-1} F_a(k)\left(\frac{1}{Z^*}\right)^k\right]^* = [\mathcal{Z}_a(1/Z^*)]^* . \tag{8}$$

Optics EXPRESS

Since $\{Z_{a,k}\}_{k=1}^{M_a-1}$ are the only roots of $\mathcal{Z}_a(Z)$, the roots of $\mathcal{Z}_{a^*}(Z)$ must obey the condition $1/Z^* = Z_{a,k}$, implying that $\{1/Z_{a,k}^*\}_{k=1}^{M_a-1}$ are the only roots of $\mathcal{Z}_{a^*}(Z)$. Similarly, writing the product $|a^*(t)a(t)|$ using Eq. (1), and identifying Z with $\exp(-i\Omega t)$, one easily concludes that $\mathcal{Z}_{aa}(Z) = \mathcal{Z}_a(Z)\mathcal{Z}_{a^*}(Z)$, and hence each of the zeros of either $\mathcal{Z}_a(Z)$ or $\mathcal{Z}_{a^*}(Z)$ is a zero of $\mathcal{Z}_{aa}(Z)$ as well. The last property, relating to the zeros of $a(t)b^*(t)$, is a trivial generalization of the above.

7.2. The procedure of finding and classifying the zeros of $I_{aa}(t)$ and $I_{ab}(t)$

We start by describing the extraction of the zeros of $I_{aa}(t)$ and $I_{ab}(t)$, which relies on a standard root-finding algorithm [13]. This algorithm can readily handle polynomials of very high orders, but it has an inherent difficulty in situations where roots are very close to one another to the extent that their effective multiplicity becomes higher than 1. In the case of $I_{ab}(t)$ the probability of such proximity between zeros is negligible. For this to happen, $a(t)$ or $b(t)$ themselves must have degenerate zeros, or zeros of $a(t)$ must coincide with zeros of $b^*(t)$ — fairly uncommon situations in the case of practical independent complex waveforms. However some of the zeros of $I_{aa}(t)$ do tend to be extremely close to one another, thereby causing the standard root-finding algorithm to falter. The reason for this proximity between zeros of $I_{aa}(t)$ is that they arrive in inverse conjugate pairs. Therefore, whenever a zero happens to be close to the unit circle, it practically coincides with its inverse conjugate counterpart. This situation is further exacerbated by the fact that the zeros of broad-band Gaussian processes tend to accumulate in the vicinity of the unit circle [26]. In order to overcome this situation we first determine the minimum phase waveform whose intensity coincides with $I_{aa}(t)$. This waveform is given by $\alpha(t) = \sqrt{I_{aa}(t)}\exp(i\phi)$, where ϕ is the Hilbert transform of the logarithm of $\sqrt{I_{aa}(t)}$ [27]. All of the zeros of the minimum phase waveform are located outside [28] the unit circle including its boundaries [27] and they can be readily determined by means of the standard root-finding procedure [13]. The zeros of $I_{aa}(t)$ are then established by supplementing the zeros of the minimum-phase function with their inverse conjugates.

Next we describe the root classification algorithm that we used. Our goal is to find which $M_a - 1$ of the $2(M_a - 1)$ zeros of $I_{aa}(t)$ have matching zeros among the zeros of $I_{ab}(t)$. A convenient method of doing that relies on the so called *rectangular assignment algorithm* [14–16]. Given an $L \times P$ matrix with non-negative elements, this algorithm efficiently finds the L elements whose sum is minimal under the constraint that no two elements can be in the same row or in the same column. We formulate our zero assignment problem so as to fit the definition of the rectangular assignment algorithm. In order to do that we combine all the zeros of $I_{ab}(t)$ into a vector $\underline{Z}_{ab}$, such that its first M_+ elements are zeros that fall inside the unit circle, whereas the remaining M_- zeros are outside of it so that $M_+ + M_- = M_a + M_b - 2$. Zeros that fall exactly on the unit circle can go in either category. Then we construct a vector $\underline{Z}_{aa}$ containing only the $M_a - 1$ zeros of $I_{aa}(t)$ that fall inside the unit circle [29]. Then the matrix $\mathbf{\Delta}$ to which the rectangular assignment algorithm is applied, is defined as follows. Its elements $\Delta_{j,k}$ in the first M_+ columns (i.e. $k \le M_+$) are the distances between the j-th element of $\underline{Z}_{aa}$ and the k-th element of $\underline{Z}_{ab}$. The elements $\Delta_{j,k}$ in the remaining M_- columns ($k > M_+$) contain the distances between the inverse conjugate of the j-th element of $\underline{Z}_{aa}$ and the k-th element of $\underline{Z}_{ab}$. In other words, the first M_+ columns of $\mathbf{\Delta}$ contain the distances between all pairs of zeros of $I_{aa}(t)$ and $I_{ab}(t)$ that are inside the unit circle, whereas the remaining M_- columns contain the distances between all pairs of zeros that are outside of it. Once the matrix $\mathbf{\Delta}$ is constructed, the rectangular assignment algorithm is applied to it. As explained earlier, for each row j it gives the value of k that ensures that $\sum_{j=1}^{M_a-1}\Delta_{j,k(j)}$ is minimal. Thus for each j, if $k \le M_+$ then the j-th element of $\underline{Z}_{aa}$ is a zero of $a(t)$. Otherwise, if $k > M_+$, the inverse conjugate of the j-th element of $\underline{Z}_{aa}$ is a zero of $a(t)$.

Finally, to complete the description of the algorithm, we define the distance between two zeros

Optics EXPRESS

Z_1 and Z_2 to be $|\ln(Z_1/Z_2)|$. This definition was found to perform better than some of its obvious alternatives (such as $|Z_1 - Z_2|$). It is interesting to note that this definition is invariant not only to switching the roles of $a(t)$ and $b(t)$ in the algorithm, but also to replacing these waveforms by their complex conjugates. This is an appealing property of the algorithm, given the symmetry of the physical problem.

Funding

Israel Science Foundation (1401/16); Italian Government CIPE resolution project INnovating City Planning through Information and Communication Technologies (INCIPICT).

References

1. N. Tesla, *Lecture Before the New York Academy of Sciences - April 6, 1897*, Leland I. Anderson, ed., Twenty-First Century Books, pp. 73–74 (1994).
2. The only relevant exception is that of the differential receiver, which we discuss later in the paper.
3. M. Bertolotti, *The history of laser* Institute of Physics Publishing, London (2005).
4. K. Kikuchi, "Fundamentals of Coherent Optical Fiber Communications," J. Lightwave Technol. **34**, 157–179 (2016).
5. H.-W. Hübers, S. G. Pavlov, A. D. Semenov, R. Köhler, L. Mahler, A. Tredicucci, H. E. Beere, D. A. Ritchie, and E. H. Linfield, "Terahertz quantum cascade laser as local oscillator in a heterodyne receiver," Opt. Express **13**, 5890–5896 (2005).
6. H.-W. Hübers, "Terahertz Heterodyne Receivers," IEEE Journal of Selected Topics in Quantum Electronics **14**, 378–391 (2008).
7. This implies that the lowest sampling rate of $B_s = (B_a + B_b)/2$ is achieved when $\Delta\nu = (B_a - B_b)/2$.
8. Obviously, the roles of $a(t)$ and $b(t)$ are symmetric, meaning that the two waveforms can also be obtained from the zeros of $I_{ab}(t)$ and $I_{bb}(t) = |b(t)|^2$.
9. This procedure will rigorously reproduce the correct waveforms only as long as none of the zeros are common to both $a(t)$ and $b(t)$ (or to $a(t)$ and $b^*(t)$). Yet the probability of this happening with unrelated waveforms is too low to be of any practical relevance.
10. The reason that it is the inverse and not the direct transform has to do with opposite sign conventions between our definition of the Fourier transform (adopted from the optics literature) and the definition used in digital signal processing applications.
11. Notice that the inability to reconstruct the constant phase offset is not unique to the proposed reconstruction method. The constant phase is not recoverable also in standard homodyne and heterodyne receivers.
12. A. Mecozzi and M. Shtaif, "Coherent detection with an incoherent local oscillator: supplementary material," figshare (2018), http://dx.doi.org/10.6084/m9.figshare.7064774
13. M. Petkovic, "Iterative Methods for Simultaneous Inclusion of Polynomial Zeros," *Lecture Notes in Mathematics* Volume 1387, Springer-Verlag Berlin Heidelberg (1989).
14. H. W. Kuhn, "The Hungarian Method for the assignment problem," Naval Research Logistics Quarterly **2**, 83–97 (1955).
15. J. Munkres, "Algorithms for the Assignment and Transportation Problems," Journal of the Society for Industrial and Applied Mathematics **5**, 32–38 (1957).
16. J. Bijsterbosch and A. Volgenan, "Solving the Rectangular assignment problem and applications," Annals of Operations Research **181**, 443–462 (2010).
17. N. Kikuchi, K. Mandai, S. Sasaki and K. Sekine, "Proposal and first experimental demonstration of digital incoherent optical field detector for chromatic dispersion compensation," in Proceedings of European Conference on Optical Communications 2006, Post-deadline Paper Th4.4.4.
18. M. Nazarathy, X. Liu, L. Christen, Y. Lize, and A. Wilner, "Self-coherent optical detection of multisymbol differential phase-shift-keyed transmission ," J. Lightwave Technol. **26**, 1921–1934 (2008).
19. Xiang. Liu, S. Chandrasekhar, A. Leven, "Digital self-coherent detection," Opt. Express **16**, 792–803 (2008)
20. A. Magen and O. Amrani, "Approaching coherent performance in differential detection via diversity," Opt. Express **23**, 4529-4538 (2015).
21. Some insight can be obtained by considering the example where $a(t)$ is periodic with period T. Then the zeros of $a(t - T_d)$ are obtained by rotating the zeros of $a(t)$ by an angle of $2\pi T_d/T$, and the pairs of zeros are perfectly distinguishable even when T_d/T is a small fraction of unity (e.g. $T_d \sim 0.1T$).
22. S. Betti, G. De Marchis, and E. Iannone, "Polarization modulated direct detection optical transmission systems," J. Lightwave Technol. **10**, 1985–1997 (1992).
23. Di Che, An Li, Xi Chen, Qian Hu, and W. Shieh, "Rejuvenating direct modulation and direct detection for modern optical communications," Opt. Commun. **409**, 86–93 (2017).
24. It should be stressed that this equivalence is only in terms of the measured quantities and not in terms of the processing that we apply to them.

Optics EXPRESS

25. H. Khodakarami, Di Che, and W. Shieh, "Information Capacity of Polarization-Modulated and Directly Detected Optical Systems Dominated by Amplified Spontaneous Emission Noise," J. Lightwave Technol. **35**, 2797–2802 (2017).
26. J. B. Hough, M. Krishnapur, Y. Peres, B. Virag, *Zeros of Gaussian analytic functions and determinantal point processes* (University Lecture Series, (2009)n vol. 51.
27. A. V. Oppenheim, R. W. Schafer, and J. R. Buck, *Discrete-Time Signal Processing* (Prentice Hall, 1999).
28. This is because we define the Z-extension as in Eq. (3), in a way consistent with a definition of Z-transform as $\sum_k a_k z^k$. With the most conventional definition of Z-transform $\sum_k a_k z^{-k}$ the zeros of the minimum phase waveform would be all inside the unit circle.
29. When a zero of $I_{aa}(t)$ falls exactly on the unit circle, it must be a double zero because it coincides with its inverse conjugate. One such zero can be included in $\underline{Z}_{aa}$.