This is the accepted manuscript made available via CHORUS. The article has been published as:

**Optimal signal recovery for pulsed balanced detection**

Yannick A. de Icaza Astiz, Vito Giovanni Lucivero, R. de J. León-Montiel, and Morgan W. Mitchell

Phys. Rev. A **90**, 033814 — Published 10 September 2014

DOI: [10.1103/PhysRevA.90.033814](https://doi.org/10.1103/PhysRevA.90.033814)
Optimal Signal Recovery for Pulsed Balanced Detection

Yannick A. de Icaza Astiz,1, * Vito Giovanni Lucivero,1 R. de J. León-Montiel,1 and Morgan W. Mitchell1,2
1ICFO-Institut de Ciencies Fotoniques, Mediterranean Technology Park, 08860 Castelldefels (Barcelona), Spain
2ICREA-Institució Catalana de Recerca i Estudis Avançats, 08015 Barcelona, Spain

We demonstrate a new tool for filtering technical and electronic noises from pulses of light, especially relevant for signal processing methods in quantum optics experiments as a means to achieve the shot-noise level and reduce strong technical noise by means of a pattern function. We provide the theory of this pattern-function filtering based on balance detection. Moreover, we implement an experimental demonstration where 10 dB of technical noise is filtered after balance detection. Such filter can readily be used for probing magnetic atomic ensembles in environments with strong technical noise.

PACS numbers: 42.62.Eh, 42.50.Lc, 42.50.Dv, 07.05.Kf

I. INTRODUCTION

Balanced detection provides a unique tool for many physical, biological and chemical applications. In particular, it has proven useful for improving the coherent detection in telecommunication systems [1, 2], in the measurement of polarization squeezing [3–7], for the detection of polarization states of weak signals via homodyne detection [8, 9], and in the study of light-atom interactions [10]. Interestingly, balanced detection has proved to be useful when performing highly sensitive magnetometry [11, 12], even at the shot-noise level, in the continuous-wave [13, 14] and pulsed regimes [15, 16].

The detection of light pulses at the shot-noise level with low or negligible noise contributions, namely from detection electronics (electronic noise) and from intensity fluctuations (technical noise), is of paramount importance in many quantum optics experiments. While electronic noise can be overcome by making use of better electronic equipment, technical noise requires special techniques to filter it, such as balanced detection and spectral filtering.

Even though several schemes have been implemented to overcome these noise sources [17–19], an optimal shot-noise signal recovery technique that can deal with both technical and electronic noises, has not been presented yet. In this paper, we provide a new tool based on the operation of polarization-rotation angles, where technical noise cannot be completely removed by means of balanced detectors [20]. Furthermore, we show that our scheme outperforms the Wiener filter, a widely used method in signal processing [21].

The paper is organized as follows. In section II we present the theoretical model of the proposed technique, in section III we show the operation of this tool by designing and implementing an experiment, where high amount of noise (technical and electronic) is filtered. Finally in section IV we present the conclusions.

II. THEORETICAL MODEL

To optimally recover a pulsed signal in a balanced detection scheme, it is necessary to characterize the detector response, as well as the “electronic” and “technical” noise contributions [22]. We now introduce the theoretical framework of the filtering technique and show how optimal pulsed signal recovery can be achieved.

A. Model for a balanced detector

To model a balanced detector, see Fig. 1, we assume that it consists of 1) a polarizing beam splitter (PBS), which splits the $H$ and $V$ polarization components to two different detectors 2) the two detectors $PD_H$ and $PD_V$, whose output currents are directly subtracted, and 3) a linear amplifier

Because the amplification is linear and stationary, we can describe the response of the detector by impulse response functions $h(\tau)$. If the photon flux at detector $X$ is $\phi_X(t)$, the electronic output can be defined as

$$v_{\text{out}}(t) \equiv h_H \ast \phi_H + h_V \ast \phi_V + v_N(t),$$

where $v_N$ is the electronic noise of the photodiodes, including amplification. Here, $h \ast \phi$ stands for the convolution of $h$ and $\phi$, i.e., $(h \ast \phi)(t) \equiv \int_{-\infty}^{\infty} h(t-\tau)\phi(\tau)d\tau$. For clarity, the time dependence will be suppressed when possible. It is convenient to introduce the following notation: $\phi_S \equiv \phi_H + \phi_V$, $\phi_D \equiv \phi_H - \phi_V$, $h_S \equiv h_H + h_V$ and
h_D ≡ h_H - h_V. Using these new variables, Eq. (1) takes the form

\[ v_{out}(t) = \frac{1}{2} (h_S * \phi_S + h_D * \phi_D) + v_N(t). \] (2)

From this signal, we are interested in recovering the differential photon number \( S \equiv \int_T \phi_H(t) dt - \int_T \phi_V(t) dt \), where \( T \) is the time interval of the desired pulse, with minimal uncertainty. More precisely, we want to find an estimator \( \hat{S}[v_{out}(t)] \), that is unbiased (\( \hat{S} = \langle S \rangle \)), and has minimal variance \( \text{var}(\hat{S}) \).

### B. Signal recovery estimator

In order to make \( \hat{S} \) unbiased, we realize that it must linearly depend on \( v_{out} \). This because \( S \) and \( v_{out} \) are linear in both \( \phi_H \) and \( \phi_V \). Therefore, the estimator must have the form

\[ \hat{S} = \int_{-\infty}^{\infty} v_{out}(t) \gamma(t) dt. \] (3)

In Eq. (3), \( \gamma(t) \) refers to as pattern function, which describes the most general linear estimator. In this work, we will consider three cases: 1) a raw estimator, \( \gamma(t) = 1 \) for \( t \in T \) and 0 otherwise; 2) a Wiener estimator, which makes use of a Wiener-filter-like pattern function, \( \gamma(t) = w(t) \), where \( w(t) \) represents the Wiener filter in the time domain [21], and 3) a model-based pattern function estimator \( \gamma(t) = g(t) \). Notice that both \( w(t) \) and \( g(t) \) are defined in \(( -\infty, \infty )\), allowing to properly choose a desired pulse. In what follows, we explicitly show how to calculate the model-based pattern function estimator \( g(t) \).

### C. Conditions of the pattern function

We assume that \( \phi_S, \phi_D \) have known averages (over many pulses) \( \bar{\phi}_S(t), \bar{\phi}_D(t) \), and similarly the response functions \( h_S(\tau), h_D(\tau) \) have averages \( \bar{h}_S(\tau), \bar{h}_D(\tau) \). Then the average of the electronic output reads as

\[ \bar{v}_{out}(t) = \frac{\bar{h}_S * \bar{\phi}_S + \bar{h}_D * \bar{\phi}_D}{2}, \] (4)

and \( \langle \hat{S} \rangle = \int_{-\infty}^{\infty} dt \langle \bar{h}_S * \bar{\phi}_S + \bar{h}_D * \bar{\phi}_D \rangle / 2 \). In writing Eq. (4), we have assumed that the noise sources are uncorrelated.

From this we observe that if a balanced optical signal is introduced, i.e. \( \bar{\phi}_D = 0 \), the mean electronic signal \( \bar{v}_{out}(t) \) is entirely due to \( \bar{h}_S * \bar{\phi}_S \). In order that \( \hat{S} \) correctly detects this null signal, \( g(t) \) must be orthogonal to \( \bar{h}_S * \bar{\phi}_S \), i.e.

\[ \int_{-\infty}^{\infty} g(t) \cdot (\bar{h}_S * \bar{\phi}_S)(t) dt = 0. \] (5)

Our second condition derives from

\[ \int_{-\infty}^{\infty} g(t) \cdot (\bar{T}_D * \bar{\phi}_D)(t) dt = \int_T \bar{\phi}_D(t) dt, \] (6)

which is in effect a calibration condition: the right-hand side is a uniform-weight integral of \( \bar{\phi}_D \), while the left-hand side is a non-uniform-weight integral, giving preference to some parts of the signal. If the total weights are the same, the above gives \( \langle \hat{S} \rangle = \langle S \rangle \). We note that this condition is not very restrictive. For example, given \( \bar{h}, \bar{\phi} \), and given \( g(t) \) up to a normalization, the equation simply specifies the normalization of \( g(t) \).

Notice that the condition given by Eq. (6) may still be somewhat ambiguous. If we want this to apply for all possible shapes \( \phi_D(t) \), it would imply \( g(t) = \text{const.} \), and would make the whole exercise trivial. Instead, we make the physically reasonably assumption that the input pulse, with shape \( \phi_S \) is uniformly rotated to give \( \bar{\phi}_H(t), \bar{\phi}_V(t) \propto \bar{\phi}_S \). Similarly, it follows that \( \bar{\phi}_D(t) \propto \bar{\phi}_S \). We note that this assumption is not strictly obeyed in our experiment and is a matter of mathematical convenience: a path difference from the PBS to the two detectors will introduce an arrival-time difference giving rise to opposite-polarity features at the start and end of the pulse, as seen in Fig. 3(a). A delay in the corresponding response functions \( h \) is, however, equivalent, and we opt to absorb all path delays into the response functions. In our experiment the path difference is \( \approx 5 \text{ cm} \), implying a time difference of less than 0.2 ns, much below the smallest features in Fig. 3(a). Absorbing the constant of proportionality into \( g(t) \), we find

\[ \int_{-\infty}^{\infty} g(t) \cdot (\bar{T}_D * \bar{\phi}_S)(t) dt = \int_T \bar{\phi}_S(t) dt, \] (7)

which is our calibration condition.
D. Noise model

We consider two kinds of technical noise: fluctuating detector response and fluctuating input pulses. We write the response functions in the form \( h_X = h_X + \delta h_X \), for a given detector \( X \), where the fluctuating term \( \delta h_X \) is a stochastic variable. Similarly, we write \( \varphi_Y = \bar{\varphi}_Y + \delta \varphi_Y \), where \( Y \) is \( H, V, S \) or \( D \). By substituting the corresponding fluctuating response functions into Eq. \( (2) \), the electronic output signal becomes

\[
v_{out}(t) = \frac{1}{2} \left( \bar{h}_S \star \bar{\varphi}_S + \bar{h}_D \star \bar{\varphi}_D \right) + v_N(t) \\
+ \frac{1}{2} \left( \delta h_S \star \bar{\varphi}_S + \delta h_D \star \bar{\varphi}_D \right) \\
+ \frac{1}{2} \left( \bar{h}_S \star \delta \varphi_S + \bar{h}_D \star \delta \varphi_D \right) + O(\delta h \delta \varphi) \tag{8}
\]

where \( v_T(t) = \frac{1}{2} (\delta h_S \star \bar{\varphi}_S + \delta h_D \star \bar{\varphi}_D + \bar{h}_S \star \delta \varphi_S + \bar{h}_D \star \delta \varphi_D) \) is the summed technical noise from both \( \delta h \) and \( \delta \varphi \) sources. We note that the optical technical noise, in contrast to optical quantum noise, scales as \( \text{var}(\delta \varphi) \propto \phi^2 \), so that \( \text{var}(v_T) \propto \phi^2 \). In passing to the last line we neglect terms \( O(\delta h \delta \varphi) \) on the assumption \( \delta h \ll h, \delta \varphi \ll \phi \). We further assume that \( v_N \) and \( v_T \) are uncorrelated.

We find the variance of the model-based estimator, \( N_\sigma \equiv \text{var}(S_{\text{opt}}) \), is

\[
N_\sigma = \left\langle \left| \int_{-\infty}^{\infty} g(t)v_T(t)dt \right|^2 \right\rangle + \left\langle \left| \int_{-\infty}^{\infty} g(t)v_N(t)dt \right|^2 \right\rangle, \tag{10}
\]

with the first term describing technical noise, and the second one electronic noise.

To compare against noise measurements, we transform Eq. \( (10) \) to the frequency domain. Using Parseval’s theorem, see Eq. \( (A2) \), we can write the noise power as

\[
N_\sigma = \int_{-\infty}^{\infty} |G(\omega)|^2 (|V_T(\omega)|^2 + |V_N(\omega)|^2) d\omega. \tag{11}
\]

Our goal is now to find the \( G(\omega) \) that minimizes \( N_\sigma \) satisfying the conditions in Eqs. \( (5) \) and \( (7) \), which in the frequency space are

\[
I_{\text{or}} \equiv \int_{-\infty}^{\infty} d\omega G^*(\omega) \overline{\Phi}_S(\omega) \Phi_S(\omega) = 0, \tag{12}
\]

\[
I_{\text{cal}} \equiv \int_{-\infty}^{\infty} d\omega G^*(\omega) \overline{\Phi}_D(\omega) \Phi_S(\omega) = \Phi_S(0). \tag{13}
\]

The specific form of the solution is given in Appendix B.
III. EXPERIMENT

A. Pulse detection and detector characterization

In our experimental setup, pulsed signals are produced using an external cavity diode laser at 795 nm (Topica DL100), modulated by two acousto-optic modulators (AOMs) in series. We have used two AOMs to prevent a shift in the optical frequency of the pulses, and also to ensure a high extinction ratio ($r_e > 10^5$).

Balanced detection is performed by using a Thorlabs PDB150A detector [23] that contains two matched photodiodes wired back-to-back for direct current subtraction, amplified by a switchable-gain transimpedance amplifier. We use the gain settings $10^3$ V/A and $10^5$ V/A, with nominal bandwidths of 150 MHz and 5 MHz, respectively. Figure 2(a) shows the average pulse shapes $p(t)$ and $p'(t)$, observed with bandwidth settings 150 MHz and 5 MHz, respectively. These shapes are obtained by blocking one detector and averaging over 1000 pulse traces (280 ns width).

In this way, to determine the impulse response functions $h_H(t)$, $h_V(t)$ of the photodiodes PD$_H$ and PD$_V$, respectively, we first assume the form

$$h_X(t) = \frac{e^{-t/\tau_{\text{TIA}}} - e^{-t/\tau_X}}{\tau_{\text{TIA}} - \tau_X},$$

where $X \in \{H, V\}$ indicates the photodiode. This describes a single-pole filter with time constant $\tau_X$ for the photodiode [24, 25] followed by a single-pole filter with time-constant $\tau_{\text{TIA}}$ for the transimpedance amplifier. We choose the parameters $\tau_{\text{TIA}}, \tau_X$ by a least-squares fit of

$$p'(t) \approx \int_{-\infty}^{\infty} p(\tau) h_X(t-\tau) d\tau.$$

As seen in Fig. 3(a), a small difference in the speeds of the two detectors leads to electronic pulses with a negative leading edge and a positive trailing edge, even when the optical signal is balanced, i.e. even when the average electronic output is zero.

B. Producing technical noise in a controlled manner

In order to prove that it is possible to remove technical noise, first we need to produce it in a controlled manner. To this end, we introduce technical noise in our system perturbing the main frequency of the AOMs using the circuit described in Fig. 4. The main frequency is produced by a voltage controlled oscillator (VCO) set to 80 MHz. Then, it is split with a power splitter, one of the arms is mixed with a signal from an arbitrary wave generator (AWG) and attenuated, whereas in the other arm the signal is passed by a phase shifter. Finally, both signals are put back together with a power combiner. In this way, we have a main frequency of 80 MHz and sidebands at the frequency of the signal introduced with the AWG. We can then programe the AWG with technical noise for a particular frequency and bandwidth, as illustrated in Fig. 5.

C. Calculating the optimal pattern function for different optical powers

To measure the noise spectra upon which the pattern function will be based, we use an oscilloscope (Lecroy Wavejet-324), rather than a spectrum analyzer. This allows us to use the same instrument for noise characterization and optimization as we will later use to acquire signals to process by digital filtering.

We collect $5 \times 10^4$ samples in a 1000 µs acquisition time.
containing a total of 800 pulses ~400 ns duration, with a duty cycle of 1/3. For this train of pulses we compute the power spectral density (PSD) for three cases: 1) signal without added technical noise, 2) signal with added technical noise, and 3) the electronic noise. Figure 6 shows an example of PSD calculated for these cases. From these PSDs we can then extract the parameters necessary for computing the optimal pattern function, namely electronic background, technical noise power and shot-noise power. Using these parameters, and following the method explained in section II, we have calculated the optimal pattern function \( g(t) \) for different average powers of the beam, from 0 to 400 \( \mu W \) in steps of 20 \( \mu W \).

D. Shot-noise limited detection with pulses and measurement of the technical noise with pulses

Because the pulses are non-overlapping, as seen in Fig. 3, we can isolate any single pulse by keeping only the signal in a finite window containing the pulse, to get a waveform as illustrated in Fig. 7. Also shown there is the optimal pattern function. This illustrates some qualitative features of the optimal pattern function, which is 1) orthogonal to the residual common-mode signal \( h_S \ast \phi_S \), which first goes negative and then positive, 2) well overlapped with the differential-mode signal \( h_D \ast \phi_D \), which is positive, and 3) smooth with some ringing, to suppress both high-frequency and low-frequency noise.

For each pulse we compute the estimators \( \hat{S}_{\text{raw}}, \hat{S}_W \) and \( \hat{S}_{\text{opt}} \), using pattern functions \( \gamma(t) = 1 \) (raw estimator), \( \gamma(t) = w(t) \) (Wiener estimator) and \( \gamma(t) = g(t) \) (optimal model-based estimator), respectively. The Wiener filter \( w(t) \) can be defined as the Fourier transform of the frequency domain representation of the Wiener filter \( W(\omega) \), given by the ratio of the cross-power spectrum of the noisy signal with the desired signal over the power spectrum of the noisy signal [21]. For more details see the appendix C.

1. Shot-noise limited detection with pulses

We first show that the system is shot-noise limited in the absence of added technical noise. For this, we compute the variance of \( \hat{S}_{\text{raw}} \), this variance is a noise estimation, computed from a pulse train without technical noise, as a function of optical power \( P \). We fit the resulting variances with the quadratic \( \text{var}(\hat{S}_{\text{raw}}) = A + BP + CP^2 \), and obtain \( A = 4.5 \pm 0.3 \times 10^{-20} J^2, B = 2.4 \pm 0.1 \times 10^{-22} J^2/\mu W \) and \( C = 6.7 \pm 0.6 \times 10^{-26} J^2/\mu W^2 \). The data and fit are shown in Fig. 8(a), and clearly show a linear dependence on \( P \), a hallmark of shot-noise limited performance.

2. Measuring technical noise with pulses

Now, we proceed as before with the exception that in this case we introduce technical noise to the signal. We obtain the following fitting parameters: \( A = 4.5 \pm 0.3 \times 10^{-20} J^2, B = 1.9 \pm 0.1 \times 10^{-22} J^2/\mu W \) and \( C = 4.12 \pm 0.05 \times 10^{-24} J^2/\mu W^2 \).

We observe from Fig. 8(b) that the noise estimation for the data that has technical noise exhibits a clearly quadratic trend, in contrast to the linear behavior where no technical noise is introduced. The results shown in Figs. 8(a) and 8(b) prove that, with our designed system, it is possible to introduce technical noise in a controlled way.

E. Filtering 10 dB of technical noise using an optimal pattern function

To illustrate the performance of our technique when filtering technical noise, we introduce a high amount of noise —about 60 dB above the shot noise level at the maximum optical power— to the light pulses produced by the AOMs. After balancing a maximum of 10 dB

![Image of power spectral density](image-url)
remains in the electronic output, which is then filtered by means of the optimal pattern function technique.

We have verified the correct noise filtering by comparing the results with shot-noise limited pulses. For this purpose, we compute \( \text{var}(S_{\text{opt}}) \), the variance of the optimal estimator for each power, and for each data set, the shot-noise limited and the noisy one. Figure 9 shows the computed noise estimation as function of the optical power for both. Notice that the two noise estimations are linear with the optical power. Moreover, we observe that both curves agree at \( \sim 91 \pm 5\% \), using the ratio of the slopes, which allows us to conclude that, by using this technique, we can retrieve shot-noise limited pulses from signals bearing high amount of technical noise.

F. Optimal estimation of the polarization-rotation angle.

The experimental setup that we have implemented, see Fig. 1, can perform also as a pulsed signal polarimeter. For instance, it is possible to determine a small polarization-rotation angle \( \varphi \) from a 45° linear polarized light pulse. Along these lines, we make use of three estimators \( \hat{S}_{\text{raw}}, \hat{S}_{\text{W}} \) and \( \hat{S}_{\text{opt}} \) to determine the amount of noise on the estimation of the polarization-rotation angle. From the obtained results, we show that the model-based estimator outperforms the other two.

We proceed to calculate the noise on the polarization-rotation angle \( \varphi \) estimation, for this determination we calculate the variance of \( \varphi \). We notice that the Taylor approximation of the variance of \( \hat{S}(\varphi) \) is

\[
\text{var}(\hat{S}) \approx \left( \frac{d\hat{S}}{d\varphi} \right)^2 \text{var}(\varphi). \tag{16}
\]

For small angles \( \varphi \), the function \( \hat{S}(\varphi) \) is approximately linear on \( \varphi \), so the contribution from higher order terms can be disregarded.

Therefore, the noise on the angle estimation is

\[
\text{var}(\varphi) = \frac{\text{var}(\hat{S})}{\left( \frac{d\hat{S}}{d\varphi} \right)^2}. \tag{17}
\]

We can then compute this expression using the three before mentioned estimators. For such task we use the experimental data together with an analytical approximation of the derivative, that takes as input the measured data. Figure 10 depicts the noise angle estimation, showing that the optimal pattern function performs better than the other estimators when eliminating the technical noise and reducing the electronic noise. In particular, the model-based estimator surpasses the Wiener estimator, which is a widely used method in signal processing [21].

IV. CONCLUSIONS

We have studied in theory and with an experimental demonstration, the optimal recovery of light pulses via balanced detection. We developed a theoretical model for a balanced detector and the noise related to the detection of optical pulses. We minimized the technical and electronic noise contributions obtaining the optimal (model-based) pattern function. We designed and implemented...
an experimental setup to test the introduced theoretical model. In this experimental setup, we produced technical noise in a controlled way, and retrieved shot-noise limited signals from signals bearing about 10 dB of technical noise after balanced detection. Finally, we compare against naive and Wiener filter estimation for measuring rotation angles, and confirm superior performance of the model-based estimator.

The results presented here might lead to a better polarization-rotation angle estimations when using pulses leading to probe magnetic atomic ensembles in environments with technical noise [15, 27]. This possibility is especially attractive for balanced detection of sub-shot-noise pulses [6, 13], for which the acceptable noise levels are still lower.

Appendix A: Parseval

We note the inner-product form of Parseval’s theorem

$$\int_{-\infty}^{\infty} G^*(t)x(t)dt = \int_{-\infty}^{\infty} G^*(\omega)X(\omega)d\omega, \tag{A1}$$

where the functions $G(\omega), X(\omega)$ are the Fourier transforms of $g(t), x(t)$, respectively. For any stationary random variable $x(t)$, $(X(\omega)X(\omega')) = \delta(\omega - \omega')$ (if this were not the case, there would be a phase relation between different frequency components, which contradicts the assumption of stationarity). From this, it follows that

$$\left(\int_{-\infty}^{\infty} g(t)x(t)dt \right)^2 = \int_{-\infty}^{\infty} |G(\omega)|^2 |X(\omega)|^2 d\omega. \tag{A2}$$

Appendix B: Formal derivation of the pattern function

We will minimize the noise power $N_\sigma$ (see Eq. (11)) with respect to the pattern function $G(\omega)$ using the two conditions (see Eq. (12) and Eq. (13)). We solve this by the method of Lagrange multipliers. For this, we write

$$L(G, \lambda_1, \lambda_2) = N_\sigma + \lambda_1(I_{or} - 0) + \lambda_2(I_{cal} - \overline{\Phi_S}(\omega = 0)), \tag{B1}$$

and then solve the equations

$$\partial_{G^*}L = 0,$$
$$\partial_{\lambda_1}L = 0,$$
$$\partial_{\lambda_2}L = 0. \tag{B2}$$

The first equation reads

$$\partial_{G^*}L = G(\omega)(|V_T(\omega)|^2 + |V_N(\omega)|^2) + \lambda_1\overline{H_S}(\omega)\overline{\Phi_S}(\omega) + \lambda_2\overline{H_D}(\omega)\overline{\Phi_D}(\omega) = 0,$$

with formal solution

$$G(\omega) = \frac{\lambda_1\overline{H_S}(\omega)\overline{\Phi_S}(\omega) + \lambda_2\overline{H_D}(\omega)\overline{\Phi_D}(\omega)}{|V_T(\omega)|^2 + |V_N(\omega)|^2}. \tag{B3}$$

The second and third equations from Eq. (B2) are the same as Eq. (12) and Eq. (13) above. The problem is then reduced to finding $\lambda_1, \lambda_2$ which (through the above), make $G(\omega)$ satisfy the two constraints.

Substituting Eq. (B4) into Eq. (12) and Eq. (13), we find

$$O_1\lambda_1 + O_2\lambda_2 = 0, \tag{B5}$$

and

$$C_1\lambda_1 + C_2\lambda_2 = \Phi_0. \tag{B6}$$

where

$$O_1 \equiv \int_{-\infty}^{\infty} \frac{\overline{H_S}(\omega)^2|\overline{\Phi_S}(\omega)|^2}{(|V_T(\omega)|^2 + |V_N(\omega)|^2)} d\omega, \tag{B7}$$
$$O_2 \equiv \int_{-\infty}^{\infty} \frac{\overline{H_D}(\omega)|\overline{\Phi_S}(\omega) - \overline{H_S}(\omega)|\overline{\Phi_D}(\omega)}{|V_T(\omega)|^2 + |V_N(\omega)|^2} d\omega, \tag{B8}$$
$$C_1 \equiv \int_{-\infty}^{\infty} \frac{\overline{H_S}(\omega)|\overline{\Phi_S}(\omega)|^2}{(|V_T(\omega)|^2 + |V_N(\omega)|^2)} d\omega, \tag{B9}$$
$$C_2 \equiv \int_{-\infty}^{\infty} \frac{|\overline{H_D}(\omega)|^2|\overline{\Phi_S}(\omega)|^2}{(|V_T(\omega)|^2 + |V_N(\omega)|^2)} d\omega. \tag{B10}$$
with $\Phi_0 \equiv \overline{\Phi}_S(\omega = 0)$. The solution to the set of Eqs. (B5) and (B6) is then given by

$$\lambda_1 = \frac{\Phi_0 O_2}{C_1 O_2 - C_2 O_1}, \quad \lambda_2 = \frac{\Phi_0 O_1}{C_2 O_1 - C_1 O_2}. \quad (B11)$$

It should be noted that quantum noise is not explicitly considered in the model. Rather, it is implicitly present in $\phi_H, \phi_V$ which may differ from their average values $\bar{\phi}_H, \bar{\phi}_V$ due to quantum noise. Note that the point of this measurement design is to optimize the measurement $\bar{\phi}_H(t) - \bar{\phi}_V(t)$, including the quantum noise in that variable. For this reason, it is sufficient to describe, and minimize, the other contributions.

**Appendix C: Wiener filter estimator**

The Wiener filter estimator $\hat{S}_W$ can be derived from the frequency domain Wiener filter output $\hat{X}(\omega)$ [21] define as

$$\hat{X}(\omega) \equiv W(\omega)V(\omega), \quad (C1)$$

where $W(\omega)$ and $V(\omega)$ are the Wiener filter and the electronic output in frequency domain, respectively.

We define $W'(\omega) \equiv W^*(\omega)$ and $w'(t) \equiv w^*(t)$ and make use of the inner product of the Parseval’s theorem, see Eq. (A1).

$$\int_{-\infty}^{\infty} W'(\omega)V_{\text{out}}(\omega)d\omega = \int_{-\infty}^{\infty} w'(t)v_{\text{out}}(t)dt. \quad (C2)$$

Then the Wiener filter estimator $\hat{S}_W$ is

$$\int_{-\infty}^{\infty} w'(t)v_{\text{out}}(t)dt$$

corresponding to Eq. (3) for $\gamma(t) = w'(t)$.

The Wiener filter $W(\omega)$ is

$$W(\omega) = \frac{|V_{\text{ideal}}^{*}(\omega)V_{\text{out}}(\omega)|}{\langle |V_{\text{out}}(\omega)|^2 \rangle}. \quad (C3)$$

In order to compute the Wiener filter it is necessary to construct the ideal signal $V_{\text{ideal}}(\omega)$, a signal without all noise contributions.

**ACKNOWLEDGMENTS**

We thank F. Wolfgramm, F. Martín Ciurana, J. P. Torres, F. Beduini and J. Zielińska for helpful discussions. This work was supported by the European Research Council project “AQUOMET”, the Spanish MINECO project “MAGO” (Ref. FIS2011-23520), and by Fundación Prvada CELLEX Barcelona. Y. A. de I. A. was supported by the scholarship BES-2009-017461, under project FIS2007-60179.


