# Entanglement-Assisted Absorption Spectroscopy

Haowei Shi,[1] Zheshen Zhang,[2, 1] Stefano Pirandola,[3] and Quntao Zhuang[4, 1, *]

[1]*James C. Wyant College of Optical Sciences, University of Arizona, Tucson, Arizona 85721, USA*
[2]*Department of Materials Science and Engineering,*
*University of Arizona, Tucson, Arizona 85721, USA*
[3]*Department of Computer Science, University of York, York YO10 5GH, UK*
[4]*Department of Electrical and Computer Engineering,*
*University of Arizona, Tucson, Arizona 85721, USA*

Spectroscopy is an important tool for probing the properties of materials, chemicals and biological samples. We design a practical transmitter-receiver system that exploits entanglement to achieve a provable quantum advantage over all spectroscopic schemes based on classical sources. To probe the absorption spectra, modelled as pattern of transmissivities among different frequency modes, we employ broad-band signal-idler pairs in two-mode squeezed vacuum states. At the receiver side, we apply photodetection after optical parametric amplification. Finally, we perform a maximal-likehihood decision test on the measurement results, achieving orders-of-magnitude-lower error probability than the optimum classical systems in various examples, including 'wine-tasting' and 'drug-testing' where real molecules are considered. In detecting the presence of an absorption line, our quantum scheme achieves the optimum performance allowed by quantum mechanics. The quantum advantage in our system is robust against noise and loss, which makes near-term experimental demonstration possible.

Quantum sensing and metrology [1–4] harness nonclassical resources to substantially improve the performance of positioning and timing [5], target detection [6–10, 14], digital reading [4], photometry [12, 13], distributed sensing [15–20], and most prominently the Laser Interferometer Gravitational-wave Observatory (LIGO) [21–23]. There have been attempts to develop quantum-metrology protocols [24, 25] for spectroscopy [26–30], an indispensable tool for science and industry. In this regard, the entangled NOON state [31, 32] and its generalizations have been considered in interferometric estimation of loss and phase [33]. However, NOON states are hard to generate and lack robustness against imperfections. As a more experimentally accessible approach, entangled photons produced by spontaneous parametric down conversion have been utilized for loss estimation [34–39]. In particular, Refs. [37–39] reported nonlinear interferometric probing of mid-infrared absorption lines using visible photons. Despite the technical benefits, the quantum advantage over the classical schemes remains unclear.

In this work, we investigate entanglement-assisted absorption spectroscopy (EAAS), as an effective means to achieve a provable quantum advantage over all schemes using classical sources. As depicted in Fig. 1, EAAS uses a source of multichromatic entangled signal-idler mode pairs from a nonlinear media, each being in a two-mode squeezed vacuum (TMSV) state and anti-correlated in the frequency domain. The signals with different frequencies interact with the sample and experience absorption differently, while the idlers are stored locally. Then an optical parametric amplifier (OPA) is applied on the return signal-idler pairs, followed by photodetection to classify samples among a plural of possibilities.

EAAS achieves a strict quantum advantage in the discrimination of arbitrary absorption patterns. Before addressing the general case, we begin with two basic models: absorption detection—the binary testing of a single absorption line at a specific frequency, and peak positioning—pinpoint a given number of absorption lines within a frequency spectrum. Then we consider the classification of several large organic molecules, and use real spectrum data [40] to simulate the performance against the optimum classical performance in 'wine-tasting' and 'drug-testing'. Let us remark that all components in our EAAS are off-the-shelf, and the quantum advantage is robust against excess noise and idler storage loss, making experimental implementations possible in the near-term.

*Pattern recognition on absorption spectra.–* In absorption spectroscopy [26], each specific composition is associated with a unique absorption spectrum determined by measuring the transmissivities across the spectrum the input light. Therefore, the overall problem of composition identification can be formulated as a hypothesis testing between several known patterns of the frequency-dependent transmissivities, as formulated below.

The multichromatic input light is decomposed into $m$ discrete frequency modes, denoted by the annihilation operators $\{a_\ell\}_{\ell=1}^m$. The input-output relation for each mode $a_\ell$ is modelled as a thermal loss channel $\mathcal{L}^{\kappa_\ell, N_B}$ [1] described by the Bogoliubov transformation

$$a_\ell \to \sqrt{\kappa_\ell} a_\ell + \sqrt{1-\kappa_\ell} e_\ell, \qquad (1)$$

where $\kappa_\ell$ is the transmissivity and $e_\ell$ is a thermal mode with mean photon number $N_B/(1-\kappa_\ell)$ to model the environmental thermal noise, which is negligible ($N_B \sim 0$) at the optical wavelengths. However, to demonstrate the robustness of the quantum advantage, $N_B > 0$ is considered for generality.

The pattern of the transmissivity coefficient $\{\kappa_\ell\}_{\ell=1}^m$
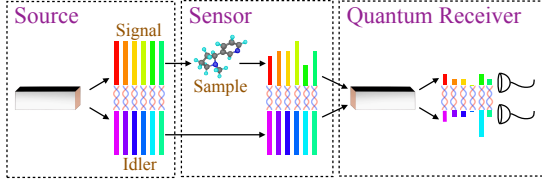
* zhuangquntao@email.arizona.edu

Figure 1. Diagram of the entanglement-assisted absorption spectroscopy. The source generates multi-chromatic entangled signal-idler pairs via a nonlinear process. The signals interact with the sample and then go through another nonlinear process jointly with the idlers at the quantum receiver. Photo detection extracts the absorption coefficients at different frequencies.
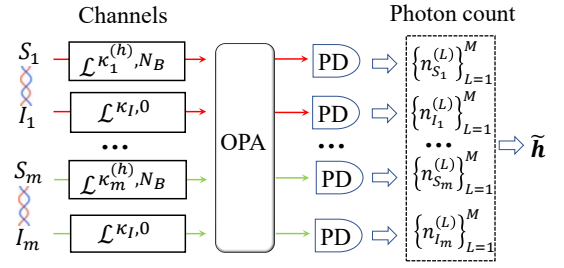


Figure 2. Schematic of receiver. Signal beams are irradiated over the sample, modeled by frequency-dependent transmissivities $\boldsymbol{\kappa}^{(h)} = \{\kappa_\ell^{(h)}\}_{\ell=1}^m$. The modulated beams go through a single optical parametric amplifier (OPA). Finally spectrally-resolving photodection (PD) offers a $2mM$-dimensional count, based on which the maximum-likelihood decision $\tilde{h}$ is made.

reveals the sample's absorption spectrum. We usually have prior information about the possible patterns, therefore the task is to discriminate between $H$ patterns, each described by a vector $\boldsymbol{\kappa}^{(h)} = \{\kappa_\ell^{(h)}\}_{\ell=1}^m$ of transmissivities, where $1 \le h \le H$ is the index of the hypotheses and $\ell$ is the frequency-mode index (see Ref. [42] for a channel formulation). In general, we allow $M$ repetitions of the probing attempt to make a decision.

Before addressing the general pattern-recognition problem described above(see Fig. 1), we consider two simplified problems: absorption detection and peak positioning.

In absorption detection, the goal is to determine whether absorption occurs at a single frequency mode ($m = 1$), therefore there are $H = 2$ hypotheses, with transmissivities $\kappa_B$ and $\kappa_T$ corresponding to the absence and the occurrence of absorption. In peak positioning, one aims to pinpoint a single absorption peak (target) within $m$ frequencies, therefore we have $H = m$ possible patterns. Each pattern $h$ has a single absorption peak with transmissivity $\kappa_T$ for frequency mode $a_h$ while all other frequency modes see a background transmissivity $\kappa_B$, i.e., $\kappa_\ell^{(h)} = \kappa_T$ if $\ell = h$, and $\kappa_B$ otherwise.

The problem of absorption detection can be generalized to finding the positions of $k$ absorption peaks in a spectrum of $m$ frequencies, which we call 'k-peak positioning'. In this more general problem, $k$ targets with transmissivity $\kappa_T$ are hidden among $m - k$ backgrounds of transmissivity $\kappa_B$, so that we have a total of $H = C_m^k$ hypotheses, where $C_m^k$ is the binomial coefficient of $m$-choose-$k$. Note that, while we consider these simple examples to introduce our results, our methodology applies to the recognition of arbitrary patterns, such as the complex molecules considered at the end of this paper.

*Classical lower bounds.—* In a classical spectroscopy scheme, one sends an arbitrary mixture of coherent states as input state. Given $mMN_S$ mean total number of photons at the input, where $N_S$ is the average mean photon number per frequency mode, the minimum error probability affecting the discrimination between the ensemble

of patterns $\{\boldsymbol{\kappa}^{(h)}\}_{h=1}^H$ is lower bounded by

$$P_{C,m,LB} = \frac{2}{(H-1)H^3} \times \tag{2}$$

$$\left[ \min_{\{X_\ell\}} \sum_{h'>h} \exp\left( -\frac{1}{2} \sum_{\ell=1}^m (\sqrt{\kappa_\ell^{(h)}} - \sqrt{\kappa_\ell^{(h')}})^2 \nu_B X_\ell \right) \right]^2,$$

where $\nu_B = 1/(1 + 2N_B)$ and the minimization is under the energy constraint $\sum_{\ell=1}^m X_\ell \le mMN_S$ (See [42] for a proof). Applying Eq. (2) to the absorption detection case, we obtain the lower bound $P_{C,1,LB} = e^{-\nu_B MN_S(\sqrt{\kappa_B} - \sqrt{\kappa_T})^2}/4$. In this case, a slightly improved bound can be obtained [4]

$$P_{C,1,LB} = \frac{1}{2}\left( 1 - \sqrt{1 - e^{-\nu_B MN_S(\sqrt{\kappa_B} - \sqrt{\kappa_T})^2}} \right). \tag{3}$$

Specifying Eq. (2) to the problem of $k$-peak positioning, one obtains [42]

$$P_{C,m,LB} = \frac{C_m^k - 1}{2C_m^k} e^{-2w_{m,k}\nu_B MN_S(\sqrt{\kappa_B} - \sqrt{\kappa_T})^2}, \tag{4}$$

where $w_{m,k} = kC_{m-1}^k/(C_m^k - 1)$. The latter term is equal to 1 for a single peak, and $w_{m,k} \simeq k(1-k/m)$ for $k$ peaks. When $N_B = 0$, the lower bound is tight in the error exponent for absorption detection and 1-peak positioning.

*Entanglement-assisted strategy.—* To achieve a quantum advantage, we exploit entanglement at the input, as given by $M$ copies of a TMSV state $\phi_{ME}$ for each signal-idler pair [42]. Each idler mode is stored locally, with imperfections modeled as a pure-loss channel $\mathcal{L}^{\kappa_I,0}$ of transmissivity $\kappa_I$ (with a mode transformation $a_I \to \sqrt{\kappa_I}a_I + \sqrt{1 - \kappa_I}v$ and $v$ being a vacuum mode); while the signal modes are sent to probe the patterns. For the special binary case of absorption detection, the error probability is bounded by the asymptotically-tight quantum Chernoff bound (QCB), which can be efficiently calculated [2, 3] from the return Gaussian states $\Xi^{(T/B)}$ composed of $M$ identical copies of $\mathcal{L}^{\kappa_{T/B},N_B} \otimes \mathcal{L}^{\kappa_I,0}(\phi_{ME})$. For the general pattern case,
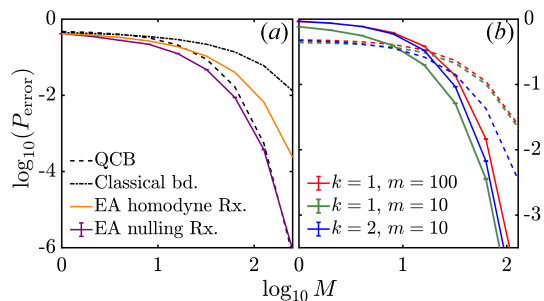
Figure 3. Error rate versus number of probing modes with practical parameters $N_S = 1$, $\kappa_T = 0.75$, and $\kappa_B = 0.95$. (a) Absorption detection. EA nulling receiver (solid orange) is compared with classical lower bound of Eq. (3) (dot-dashed black), QCB (dashed black), and the EA homodyne receiver (solid purple). (b) Peak positioning. Single-peak positioning (solid green) and double-peak positioning (solid blue) with $m = 10$ frequency slots. Single-peak positioning with $m = 100$ (solid red) provided as a reference. Classical lower bounds given by Eq. (4) (dashed, accordingly colored).

a simple tool like the QCB is missing and, for this reason, we need to design an explicit receiver that is able to show a quantum advantage.

*Entanglement-assisted receiver design.–* We begin our description of the receiver design with a simple case so as to provide its basic modus operandi. Consider the ideal case of $\kappa_B = \kappa_I = 1$ and $N_B = 0$. Then the returned state $\Xi^{(B)} = \phi_{ME}^{\otimes M}$ consists of $M$ copies of the ideal TMSV state (while $\Xi^{(T)}$ is mixed because $\kappa_T < 1$). Suppose that we perform a two-mode squeezing (TMS) operation $\mathcal{S}$ (via an OPA), that precisely anti-squeezes each TMSV state $\phi_{ME}$. Then we can 'null' $\mathcal{S}(\Xi^{(B)})$ to tensor products of vacuum, while $\mathcal{S}(\Xi^{(T)})$ is non-vacuum. Therefore, a simple photon counting on all the signals and idlers after the TMS operation can identify the input state $\Xi^{(B)}$ if there is any photon count. Errors only occur if we obtain a zero count on $\mathcal{S}(\Xi^{(T)})$: when this happens, we can only guess randomly, with an error probability $R_m$. Note that this nulling strategy has been used in classical schemes [42], whose performance is bounded by Eqs. (3) and (4). OPA has also been utilized in quantum illumination [52], however without exploiting correlations in the patterns.

Let us use a compact notation, where $m = 1$ corresponds to absorption detection, for which $R_1 = 1/2$; and $m \geq 2$ corresponds to single-peak positioning, with one copy of $\mathcal{S}(\Xi^{(T)})$ among $m - 1$ copies of $\mathcal{S}(\Xi^{(B)})$, so that $R_m = (m-1)/m$. Accounting for the zero counts, the error probability for absorption detection ($m = 1$) and single-peak positioning ($m \geq 2$) is given by [42]

$$P_{E,m} = R_m \left[ \frac{1}{1 + N_S(1 - \sqrt{\kappa_T})} \right]^{2M}. \tag{5}$$

When $N_S \ll 1$ and $M \gg 1$, we have $P_{E,m} \simeq R_m \exp\left(-2MN_S\left(1 - \sqrt{\kappa_T}\right)\right)$. Comparing this with the classical lower bounds in Eqs. (3) and (4), we see that
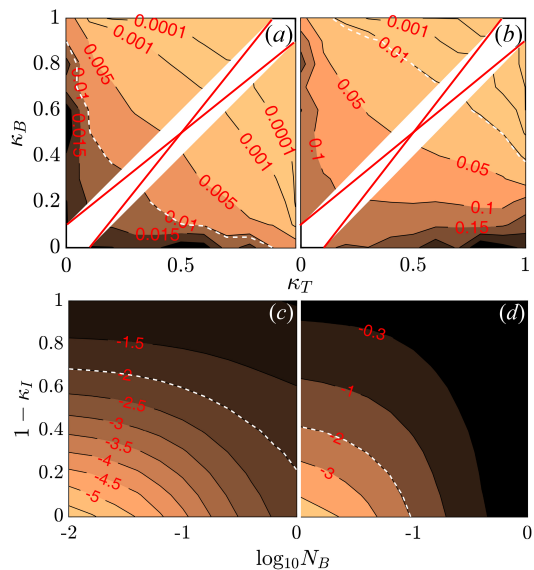


Figure 4. (a)(b) Error probability of EAAS $P_{E,m}$ versus transmissivities $\kappa_B$ and $\kappa_T$; (c)(d) $\log_{10} P_{E,m}$ versus idler loss $1 - \kappa_I$ and thermal noise $N_B$ at fixed $\kappa_T = 0.75$ and $\kappa_B = 0.95$. absorption detection ($m = 1$) in (a)(c) compared with single-peak positioning ($m = 100$) in (b)(d). $M$ is chosen to fixed the classical lower bounds to 0.01 [42]. $N_S = 1$ is assumed. Red-crossed diagonal region in (a)(b) represents the degenerate case $\kappa_B = \kappa_T$.

EAAS has an exponential advantage: $P_{E,m}/P_{C,m,LB} \simeq 2\exp\left(-MN_S\left(1 - \kappa_T\right)\right)$ for absorption detection, and $\simeq 2\exp\left(-2MN_S\left(\sqrt{\kappa_T} - \kappa_T\right)\right)$ for single-peak positioning. In fact, Eq. (5) achieves the QCB [42] and therefore it is optimal for absorption detection.

The above receiver design, and the resulting entanglement advantage, can be generalized to cope with more complex spectrum patterns and the presence of noise and idler loss ($N_B > 0$, $\kappa_I < 1$), as described by the following strategy (see Fig. 2): (i) Apply TMS operation with gain $G_\ell$ to each of the return signal-idler pairs $a''_{S\ell}, a''_{I\ell}$ to obtain new modes $a_{S\ell} = \sqrt{G_\ell}a''_{S\ell} - \sqrt{G_\ell - 1}a''^\dagger_{S\ell}$ and $a_{I\ell} = \sqrt{G}a''_{I\ell} - \sqrt{G - 1}a''^\dagger_{S\ell}$. (ii) Perform photon counting measurement on all signal and idler modes $\{a_{S\ell}, a_{I\ell}\}$'s to obtain the results as two vectors $\boldsymbol{n}_S$ and $\boldsymbol{n}_I$. (iii) Finally, apply maximum-likelihood (ML) decision rule, i.e., make the decision $\tilde{h}$ through

$$\tilde{h} = \arg\max_h P_m(\boldsymbol{n}_S, \boldsymbol{n}_I | h), \tag{6}$$

where $P_m(\boldsymbol{n}_S, \boldsymbol{n}_I | h)$ is the conditional probability of obtaining the outcomes $\boldsymbol{n}_S, \boldsymbol{n}_I$ if the true hypothesis is $h$.

To complete the description of our receiver, we need to determine the gain $G_\ell$'s and specify the conditional probabilities. Let us begin with the cases of absorption detection and peak positioning, where we adopt uniform gain $G_\ell = G$; the ideal situation is to get a quantum state close to vacuum; however, if $\kappa_B < 1$, it is only possible

to reduce the signal part of $\Xi^{(B)}$ to vacuum, by choosing $G = 1 + N_S\kappa_B/(1 + N_S(1 - \kappa_B))$. In the presence of noise $N_B > 0$ and idler loss $\sqrt{1 - \kappa_I} > 0$, 'nulling' to vacuum is not possible but the same choice of gain still provides an appreciable advantage over classical schemes. For general patterns, due to the absence of symmetry we consider optimization over the gain $G_\ell$'s at different frequency modes. Moreover, as some frequency windows may contain more essential information about the hypotheses, we also allow the optimization over the energy distribution $\{N_{S\ell}\}$ of the TMSV in different frequency modes. In these cases, although the 'nulling' decision rule does not apply, the ML decision rule in Eq. (6) still leads to an advantage [42].

Now let us compute the conditional probabilities. With $M$ identical repetitions, the probability of obtaining the $mM$-dimensional measurement results $\boldsymbol{n}_S = \{n_{S_L,k}\}_{L=1,k=1}^{M,m}$ and $\boldsymbol{n}_I = \{n_{I_L,k}\}_{L=1,k=1}^{M,m}$, conditioned on pattern $h$, is

$$P_m(\boldsymbol{n}_S, \boldsymbol{n}_I | h) = \prod_{L=1}^{M} \prod_{\ell=1}^{m} P(n_{S_L,\ell}, n_{I_L,\ell} | \kappa_\ell^{(h)}, G_\ell, N_{S\ell}), \tag{7}$$

where each term is a function of the subsystem transmissivity $\kappa_\ell^{(h)}$, the TMSV source energy $N_{S\ell}$ and the gain choice $G_\ell$ [42].

With all these theoretical elements in our hands, we can numerically evaluate the error probability $P_{E,m}$ for the problems of absorption detection, peak positioning and general spectrum recognition via Monte Carlo simulations [42]. Although we consider equal priors for simplicity, our ML decision can generally be applied to arbitrary prior probabilities for the patterns.

*Detecting and positioning absorption peaks.–* In order to investigate the problems of absorption detection and peak positioning, we assume a background transmissivity $\kappa_B = 0.95$ and a target transmissivity $\kappa_T = 0.75$. In particular, we study their error probabilities in terms of the number of modes $M$. For absorption detection, Fig. 3(a) shows that our EA nulling receiver asymptotically achieves the QCB [4], outperforming both the best known receiver, the EA homodyne receiver [3, 42], and the classical lower bound of Eq. (3). In fact, we can verify that our receiver can asymptotically saturate the QCB for absorption-detection with general choices of $\kappa_B$ and $\kappa_T$ [42]. For peak positioning, as shown in Fig. 3(b), our EA receiver is able to outperform the classical lower bound of Eq. (4) by orders of magnitude.

In a practical scenario, we are interested in how much EAAS can enhance the performance, when classical schemes fail to perform well. To showcase the advantage, in Fig. 4, we fix the classical lower bounds in Eqs. (3) and (4) to be 0.01 and plot the error probability $P_{E,m}$ of EAAS. We start with tuning the transmissivities $\kappa_B$ and $\kappa_T$ in Fig. 4(a)(b). Then we fix $\kappa_T = 0.75$ and $\kappa_B = 0.95$ and study how the quantum advantage varies with idler loss $1 - \kappa_I$ and noise $N_B$ in Fig. 4(c)(d). The white dashed
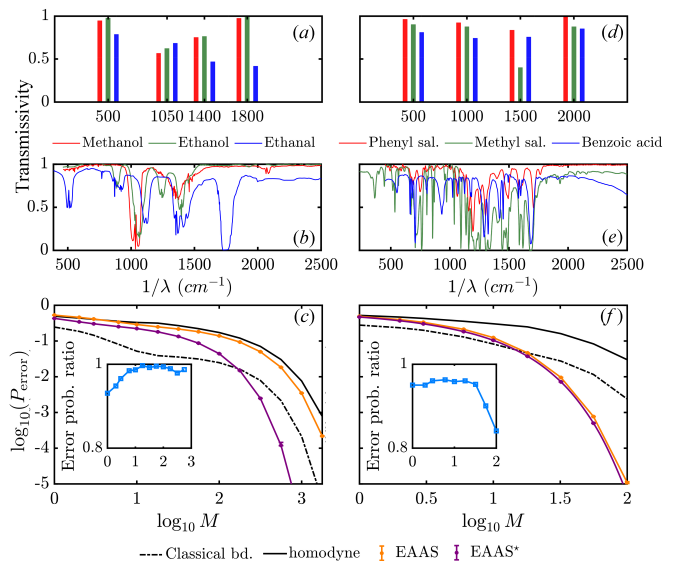


Figure 5. Identification of $H = 3$ molecules with $m = 4$ frequency slots, for wine-tasting in (a)(b)(c) and drug-testing in (d)(e)(f). (a)(d) are $m = 4$ sampled discrete spectra on the FTIR spectra in (b)(e). (c) and (f) show the logarithmic error rate of EAAS, with $N_S = 1$. 'EAAS' assumes a uniform distribution of photons at the input modes (solid orange), while numerically optimized energy distribution presented in 'EAAS⋆' (solid purple). The classical lower bound (dot-dashed black) and homodyne detection (solid black). Insets of (c) and (f): error probability ratio of EAAS with gain optimization after energy-distribution optimization and EAAS with merely energy-distribution optimization.

lines divide the parameter space with/without quantum advantage. We can see that the advantage is remarkable, and also survives for a large range of parameters, especially when $\kappa_B \simeq 1$ as in practice. The robustness of the advantages to imperfections reveals a clear possibility for a near-term experimental demonstration. See [42] for more parameter settings.

*General spectrum recognition.–* EAAS can also identify actual molecules, each of which is associated with a unique absorption spectrum. As a taste of flavor, we begin with 'wine-tasting'—where one discriminates three common alcohol-like liquids. Methanol could be lethal if mistaken for ethanol (alcohol). Meanwhile, the alcohol, as time goes by, will be dehydrogenated to ethanal, whose concentration provides the age of a vintage [53]. To consider larger molecules, the second example, 'drug-testing', involves three drugs: phenyl salicylate, methyl salicylate, and benzoic acid. In both examples, a non-destructive testing method is preferred, as we shall conduct with the extremely weak quantum light source. The transmissities are taken from real Fourier-transform infrared (FTIR) spectra [40]. These spectra are sampled by averaging them within each of $m = 4$ frequency slots [42].

As the classical benchmark, we calculate the ultimate lower bound using Eq. (2) and the performance of a ho-

modyne receiver on coherent-state input with the same energy distribution (distribution of mean photon number over frequency modes) optimized in Eq. (2). Fig. 5 shows that EAAS with uniform energy distribution and $G = 1$ (orange) outperforms the homodyne receiver (black solid) in both cases. Then, in drug-testing, EAAS beats the classical lower bound by orders of magnitude, while, in wine-tasting, this advantage is less pronounced. This is mainly due to the classical lower bound being not tight, and uniform energy being sub-optimum, as we see EAAS with energy optimization (purple) enables much better advantages. Although gain optimization only leads to a slight advantage over the energy-optimized EAAS, as evident in the inset plots of Fig. 5(c)(f). In the noisy case, it enables a much better enhancement [42].

Now we address phase noise common in experiments. Phase tracking can typically eliminate the time-invariant phase noise, so the above results directly hold; when phase tracking is not possible, we can model the phase noise by adding $a_\ell \to e^{i\theta_\ell} a_\ell$ in Eq. (1). The random phase $\theta_\ell$ clearly complicates the problem. However, if we choose uniform $G = 1$ (i.e., not applying OPA before photodetection), the same results of the orange curves in

Fig. 5(c)(f) hold, and the classical performance can only be worse than the current benchmarks (black). Thus, the quantum advantage sustains.

*Conclusion.*— We have devised a near-term feasible EAAS scheme that outperforms any classical strategy in determining the presence and position of spectral absorption peaks. The EAAS scheme saturates the QCB in binary detection of a single absorption line and offers orders-of-magnitude advantage in error probability in the discrimination of sampled spectra of molecules even in the presence of experimental nonidealities.

[1] S. Pirandola, B. R. Bardhan, T. Gehring, C. Weedbrook, and S. Lloyd, Advances in photonic quantum sensing, Nat. Photonics **12**, 724 (2018).

[2] I. Ruo Berchera and I. P. Degiovanni, Quantum imaging with sub-poissonian light: challenges and perspectives in optical metrology, Metrologia **26**, 024001 (2019).

[3] C.L. Degen, F. Reinhard, and P. Cappellaro, Quantum sensing, Rev. Mod. Phys. **89**, 035002 (2017).

[4] V. Giovannetti, S. Lloyd, and L. Maccone, Advances in quantum metrology, Nat. Photonics **5**, 222 (2011).

[5] V. Giovannetti, S. Lloyd, and L. Maccone, Quantum-enhanced positioning and clock synchronization, Nature **412**, 417 (2001).

[6] S. Lloyd, Enhanced sensitivity of photodetection via quantum illumination, Science **321**, 1463 (2008).

[7] S.-H. Tan, B. I. Erkmen, V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, S. Pirandola, and J. H. Shapiro, Quantum illumination with gaussian states, Phys. Rev. Lett. **101**, 253601 (2008).

[8] Q. Zhuang, Z. Zhang, and J. H. Shapiro, Optimum mixed-state discrimination for noisy entanglement-enhanced sensing, Phys. Rev. Lett. **118**, 040801 (2017).

[9] S. Barzanjeh, S. Guha, C. Weedbrook, D. Vitali, J. H. Shapiro, and S. Pirandola, Microwave Quantum Illumination, Phys. Rev. Lett. **114**, 080503 (2015)

[10] Zheshen Zhang, Sara Mouradian, Franco N.C. Wong, and Jeffrey H. Shapiro, Entanglement-Enhanced Sensing in a Lossy and Noisy Environment, Phys. Rev. Lett. 114, 110506 (2015)

[4] S. Pirandola, Quantum reading of a classical digital memory, Phys. Rev. Lett. **106**, 090504 (2011).

[12] G. Spedalieri, C. Lupo, S. L. Braunstein, and S. Pirandola, Thermal quantum metrology in memoryless and correlated environments, Quantum Sci. and Technol. **4**,

015008 (2018).

[13] G. Spedalieri, S. Pirandola, and S. L. Braunstein, Symmetric and asymmetric discrimination of bosonic loss: Toy applications to biological samples and photodegradable materials, Phys. Rev. A **98**, 053836 (2018).

[14] Q. Zhuang, Z. Zhang, and J. H. Shapiro, Entanglement-enhanced lidars for simultaneous range and velocity measurements, Phys. Rev. A **96**, 040304 (2017).

[15] Q. Zhuang, Z. Zhang, and J. H. Shapiro, Distributed quantum sensing using continuous-variable multipartite entanglement, Phys. Rev. A **97**, 032329 (2018).

[16] T. J. Proctor, P. A. Knott, and J. A. Dunningham, Multiparameter estimation in networked quantum sensors, Phys. Rev. Lett. **120**, 080501 (2018).

[17] W. Ge, K. Jacobs, Z. Eldredge, A. V. Gorshkov, and M. Foss-Feig, Distributed quantum metrology with linear networks and separable inputs, Phys. Rev. Lett. **121**, 043604 (2018).

[18] X. Guo, C. R. Breum, J. Borregaard, S. Izumi, M. V. Larsen, T. Gehring, M. Christandl, J. S. Neergaard-Nielsen, and U. L. Andersen, Distributed quantum sensing in a continuous-variable entangled network, Nat. Phys. **16**, 281 (2020).

[19] Y. Xia, W. Li, W. Clark, D. Hart, Q. Zhuang, and Z. Zhang, Demonstration of a Reconfigurable Entangled Radio-Frequency Photonic Sensor Network, Phys. Rev. Lett. **124**, 150502 (2020).

[20] Q. Zhuang, J. Preskill, and L. Jiang, Distributed quantum sensing enhanced by continuous-variable error correction, New J. Phys. **22**, 022001 (2020).

[21] LIGO Scientific Collaboration, Observation of gravitational waves from a binary black hole merger, Phys. Rev. Lett. **116**, 061102 (2016).

[22] LIGO Scientific Collaboration, A gravitational wave ob-

servatory operating beyond the quantum shot-noise limit, Nat. Phys. **7**, 962 (2011).

[23] LIGO Scientific Collaboration, Quantum-enhanced advanced ligo detectors in the era of gravitational-wave astronomy, Phys. Rev. Lett. **123**, 231107 (2019).

[24] K. E. Dorfman, F. Schlawin, and S. Mukamel, Nonlinear optical signals and spectroscopy with quantum light, Rev. Mod. Phys. **88**, 045008 (2016).

[25] S. Mukamel, M. Freyberger, W. P. Schleich, M. Bellini, A. Zavatta, G. Leuchs, C. Silberhorn, R. W. Boyd, L. S. Soto, A. Stefanov, *et al.*, Roadmap on quantum light spectroscopy, J. Phys. B: At., Mol. Opt. Phys. (2020).

[26] W. Demtröder, *Laser spectroscopy: basic concepts and instrumentation* (Springer Science & Business Media, 2013).

[27] M. Levenson, *Introduction to Nonlinear Laser Spectroscopy 2e* (Elsevier, 2012).

[28] W. E. Moerner, Nobel lecture: Single-molecule spectroscopy, imaging, and photocontrol: Foundations for super-resolution microscopy, Rev. Mod. Phys. **87**, 1183 (2015).

[29] O. Fischer, M. Kugler, I. Maggio-Aprile, C. Berthod, and C. Renner, Scanning tunneling spectroscopy of high-temperature superconductors, Rev. Mod. Phys. **79**, 353 (2007).

[30] T. W. Hänsch and H. Walther, Laser spectroscopy and quantum opt., Rev. Mod. Phys. **71**, S242 (1999).

[31] B. C. Sanders, Quantum dynamics of the nonlinear rotator and the effects of continual spin measurement, Phys. Rev. A **40**, 2417 (1989).

[32] H. Lee, P. Kok, and J. P. Dowling, A quantum rosetta stone for interferometry, J. Mod. Opt. **49**, 2325 (2002).

[33] H. T. Dinani, M. K. Gupta, J. P. Dowling, and D. W. Berry, Quantum-enhanced spectroscopy with entangled multiphoton states, Phys. Rev. A **93**, 063804 (2016).

[34] R. Whittaker, C. Erven, A. Neville, M. Berry, J. O'Brien, H. Cable, and J. Matthews, Absorption spectroscopy at the ultimate quantum limit from single-photon states, New J. Phys. **19**, 023013 (2017).

[35] M. Li, C.-L. Zou, D. Liu, G.-P. Guo, G.-C. Guo, and X.-F. Ren, Enhanced absorption microscopy with correlated photon pairs, Phys. Rev. A **98**, 012121 (2018).

[36] R. Nair, Quantum-limited loss sensing: Multiparameter estimation and bures distance between loss channels, Phys. Rev. Lett. **121**, 230801 (2018).

[37] D. A. Kalashnikov, A. V. Paterova, S. P. Kulik, and L. A. Krivitsky, Infrared spectroscopy with visible light, Nat. Photonics **10**, 98 (2016).

[38] A. Paterova, H. Yang, C. An, D. Kalashnikov, and L. Krivitsky, Measurement of infrared optical constants with visible photons, New J. Phys. **20**, 043015 (2018).

[39] A. V. Paterova, S. M. Maniam, H. Yang, G. Grenci, and L. A. Krivitsky, Hyperspectral infrared microscopy with visible light, arXiv:2002.05956 (2020).

[40] Coblentz Society, Inc. "Evaluated Infrared Reference Spectra" in NIST Chemistry WebBook, NIST Standard Reference Database Number 69, Eds. P.J. Linstrom and W.G. Mallard, National Institute of Standards and Technology, Gaithersburg MD, 20899, `https://doi.org/10.18434/T4D303`, (retrieved April 7, 2020).

[1] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, Gaussian quantum information, Rev. Mod. Phys. **84**, 621 (2012).

[42] See Supplemental Material, which also includes Refs. [5–11].

[43] Q. Zhuang and S. Pirandola, Entanglement-enhanced testing of multiple quantum hypotheses, Commun. Phys. **3**, 103 (2020).

[44] A. Montanaro, A lower bound on the probability of error in quantum state discrimination, in 2008 IEEE Information Theory Workshop (IEEE, 2008) pp. 378âĂŞ380.

[7] S. J. Dolinar, *Processing and Transmission of Information*, Technical Report (Research Laboratory of Electronics (RLE) at the Massachusetts Institute of Technology (MIT), 1973)

[8] S. Dolinar Jr, A near-optimum receiver structure for the detection of m-ary optical ppm signals, The Telecommunications and Data Acquisition Progress Report **42** (1982)

[9] C. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, 1976)

[48] J.C. Spall, An overview of the simultaneous perturbation method for efficient optimization, Johns Hopkins APL Technical Digest **19**, 482 (1998)

[49] P. Sadegh, Constrained optimization via stochastic approximation with a simultaneous perturbation gradient approximation, IFAC Proceedings **30** 281 (1997).

[2] S. Pirandola and S. Lloyd, Computable bounds for the discrimination of gaussian states, Phys. Rev. A **78**, 012331 (2008).

[3] S. Pirandola, C. Lupo, V. Giovannetti, S. Mancini, and S. L. Braunstein, Quantum reading capacity, New J. Phys. **13**, 113012 (2011).

[52] S. Guha and B. I. Erkmen, Gaussian-state quantum-illumination receivers for target detection, Phys. Rev. A **80**, 052310 (2009).

[53] H. Yu, H. Lin, H. Xu, Y. Ying, B. Li, and X. Pan, Prediction of enological parameters and discrimination of rice wine age using least-squares support vector machines and near infrared spectroscopy, J. Agric. Food Chem **56**, 307 (2008).

# Supplemental Materials: Entanglement-Assisted Absorption Spectroscopy

## I. CHANNEL SET-UP

In the main paper, we mainly utilize input-output relations to describe the quantum channels, here we make the quantum channel notations explicit.

The overall $H$ patterns being discriminated can each be modelled as quantum channels $\{\mathcal{J}^{(\boldsymbol{k}^{(h)})}\}_{h=1}^H$ act on $m$ subsystems. They are given by $\mathcal{J}^{(\boldsymbol{k}^{(h)})} = \otimes_{\ell=1}^m \Phi_{S_\ell}^{(h)}$, where each sub-channel $\Phi_{S_\ell}^{(h)} \equiv \left(\mathcal{L}^{\kappa_\ell^{(h)}, N_B}\right)^{\otimes M}$ is associated with $M$ probings of subsystem $S_\ell$ and $\mathcal{L}^{\kappa, N}$ is a thermal loss channel with transmissivity $\kappa$, noise $N$.

In the absorption detection case, $H = 2$, $m = 1$, and we denote the two channels as $\Phi^{(T/B)} = \left(\mathcal{L}^{\kappa_{B/T}, N_B}\right)^{\otimes M}$. The transmissivities $\kappa_B$ and $\kappa_T$ correspond to the absence (background channel $\Phi^{(B)}$) and the occurrence (target channel $\Phi^{(T)}$) of absorption.

In the peak positioning case, we have $H = m$ possible global channels, where $\mathcal{J}^{(\boldsymbol{k}^{(h)})}$ has a target channel $\Phi^{(T)}$ at subsystem $S_h$ while the rest are transparent backgrounds $\Phi^{(B)}$. The corresponding pattern is therefore described by $\kappa_\ell^{(h)} = \kappa_T$ if $\ell = h$, and $\kappa_B$ otherwise.

In the general pattern case, we consider discrimination between the ensemble of channels $\{\mathcal{J}^{(\boldsymbol{k}^{(h)})}\}_{h=1}^H$.

When we introduce entanglement assistance, each sub-channel is extended to $\Phi_{S_\ell}^{(h)} \otimes \mathcal{D}_{I_\ell}^{\kappa_I}$, where $\mathcal{D}^{\kappa_I} = \left(\mathcal{L}^{\kappa_I,0}\right)^{\otimes M}$ and $\mathcal{L}^{\kappa_I,0}$ is a noiseless lossy channel modelling the imperfections on each idler system $I_\ell$. Then the return state for each sub-channel is $\Xi_{S_\ell}^{(h)} = \Phi_{S_\ell}^{(h)} \otimes \mathcal{D}_{I_\ell}^{\kappa_I}\left(\phi_{ME}^{\otimes M}\right)$. For the case of target/background channels, the return states

$$\Xi^{(T/B)} = \Phi^{(T/B)} \otimes \mathcal{D}^{\kappa_I}\left(\phi_{ME}^{\otimes M}\right). \tag{S8}$$

## II. COVARIANCE MATRIX AND TWO-MODE PHOTON STATISTICS

### A. Covariance matrix derivation

First, we briefly introduce the notion of Gaussian states [S1], whose Wigner functions have a Gaussian shape. An $n$-mode Gaussian state $\rho$ comprising modes $a_k, 1 \le k \le n$, is fully characterized by the mean and the covariances of real quadrature field operators $q_k = a_k + a_k^\dagger, p_k = i\left(a_k^\dagger - a_k\right)$. Formally, we can define a real $2n$-dim vector of operators $\boldsymbol{x} = (q_1, p_1, \cdots, q_n, p_n)$, then the mean $\bar{\boldsymbol{x}} = \langle \boldsymbol{x} \rangle_\rho$ and the elements of the $2n$-by-$2n$ covariance matrix are given by

$$\Lambda_{ij} = \frac{1}{2}\langle\{x_i - \bar{x}_i, x_j - \bar{x}_j\}\rangle_\rho, \tag{S9}$$

where $\{,\}$ is the anticommutator and $\langle A \rangle_\rho = \text{Tr}(A\rho)$.

An important example of Gaussian state is TMSV, given by the wave-function

$$\phi_{ME} = \sum_{n=0}^\infty \sqrt{\frac{N_S^n}{(N_S + 1)^{n+1}}}|n\rangle_{S'}|n\rangle_{I'}, \tag{S10}$$

where $|n\rangle$ is the number state. From the above wavefunction, we can obtain the covariance matrix of a TMSV as

$$\Lambda_{\text{TMSV}} = \begin{pmatrix} (2N_S + 1)\mathbf{I} & 2C_0\mathbf{Z} \\ 2C_0\mathbf{Z} & (2N_S + 1)\mathbf{I} \end{pmatrix}, \tag{S11}$$

where $\mathbf{I}$, $\mathbf{Z}$ are two-by-two Pauli matrices, and $C_0 = \sqrt{N_S(N_S + 1)}$ is the amplitude of the phase-sensitive cross correlation.

We utilize multiple copies of the signal-diler pair $\{a_S', a_I'\}$ in a TMSV state to probe the sample. To begin with, we consider the case with no phase noise, where each signal goes through the channel $\mathcal{L}^{\kappa_S, N_B}$, giving the output

$$a_S'' = \sqrt{\kappa_S}a_S' + \sqrt{1 - \kappa_S}e, \tag{S12}$$

where $e$ mode is in a thermal state with mean photon number $N_B/(1 - \kappa_S)$. While the idler mode $a_I'$ goes through a pure loss channel $\mathcal{L}^{\kappa_I, 0}$.

$$a_I'' = \sqrt{\kappa_I}a_I' + \sqrt{1 - \kappa_I}v, \tag{S13}$$

where the environment mode $v$ is in vacuum state. The state $\rho_{\Lambda_0(\kappa_S)}$ of each of the signal-idler pair $\{a_S'', a_I''\}$ has the covariance matrix

$$\Lambda_0(\kappa_S) = \begin{pmatrix} (2(\kappa_S N_S + N_B) + 1)\mathbf{I} & 2C_p\mathbf{Z} \\ 2\kappa_S\kappa_I C_0\mathbf{Z} & (2\kappa_I N_S + 1)\mathbf{I} \end{pmatrix}, \tag{S14}$$

with the signal-idler cross correlation $C_p = \sqrt{\kappa_S\kappa_I N_S(1 + N_S)}$.

In the case of absorption detection, we have $\kappa_S = \kappa_B$ or $\kappa_S = \kappa_T$. With each Gaussian mode-pair of the return state specified by Eq. (S14), we can obtain the QCB on the error probability of the binary hypothesis testing through methods in Ref. [S2],

$$P_E \le P_{QCB} = \tilde{Q}_s^M/2, \tag{S15}$$

where $\tilde{Q}_s = \inf_{0 \le s \le 1}\text{Tr}\rho_{\Lambda_0(\kappa_B)}^s\rho_{\Lambda_0(\kappa_T)}^{1-s}$. The QCB is asymptotically tight as the mode number $M \to \infty$ and can be efficiently calculated in our case [S2, S3].

On the receiver side, we apply a two-mode squeezing process, parametrized by the gain $G \ge 1$, to obtain returned signal-idler mode pairs $\{a_S, a_I\}$

$$a_S = \sqrt{G}a_S'' - \sqrt{G - 1}a_I^\dagger, \tag{S16}$$

$$a_I = \sqrt{G}a_I'' - \sqrt{G - 1}a_S^\dagger. \tag{S17}$$

They are in a Gaussian state with covariance matrix

$$\mathbf{\Lambda}\left(\kappa_S\right) = \begin{pmatrix} E\mathbf{I} & C\mathbf{Z} \\ C\mathbf{Z} & S\mathbf{I} \end{pmatrix}. \tag{S18}$$

Denote $G = 1 + N'_S$ with the effective photon number $N'_S$, we have the variances of signal and idler

$$E = 1 + 2\kappa_S N_S - 4C_p\sqrt{N'_S(1 + N'_S)}$$
$$+ 2N'_S(1 + \kappa_I N_S + \kappa_S N_S) + 2N_B(1 + N'_S), \tag{S19}$$

$$S = 1 + 2\kappa_I N_S - 4C_p\sqrt{N'_S(1 + N'_S)}$$
$$+ 2N'_S(1 + \kappa_I N_S + \kappa_S N_S) + 2N_B N'_S, \tag{S20}$$

and the signal-idler correlation

$$C = 2\sqrt{N'_S(1 + N'_S)}(1 + \kappa_I N_S + \kappa_S N_S)$$
$$- 2(1 + 2N'_S)C_p + 2N_B\sqrt{(1 + N'_S)N'_S}. \tag{S21}$$

When $\kappa_I = 1, N_B = 0$, suppose we choose

$$N'_{S0} = \frac{N_S \kappa_B}{1 + N_S(1 - \kappa_B)}, \tag{S22}$$

as given in the main paper, we have

$$\mathbf{\Lambda}\left(\kappa_B\right) = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & (1 + 2(1 - \kappa_B)N_S)\mathbf{I} \end{pmatrix}. \tag{S23}$$

This means that the signal mode becomes vacuum under this choice of the gain. We will choose this gain in the absorption detection and peak positioning cases, with the exception in Sec. VII B, in which we optimize the gain. For the general pattern recognition for molecules, we will consider both uniform gain $G = 1$ and consider optimization over gain.

## B. Photon number statistics

We consider photon counting on each signal-idler pair $\{a_S, a_I\}$ in a Gaussian state with the covariance matrix in the form of Eq. (S18), which has zero phase-insensitive cross-correlation $\left\langle a_S a_I^\dagger \right\rangle = 0$ and non-zero phase sensitive cross-correlation $\langle a_S a_I \rangle \neq 0$. The corresponding joint probability of obtaining results $n_S, n_I$ is given by

$$P(n_S, n_I) = -4F_R(1 + n_S, 1 + n_I, 1, \frac{4C^2}{XY})$$
$$\times \frac{(-1 + C^2 + E + S - ES)^{1+n_S+n_I}}{X^{1+n_S}Y^{1+n_I}}, \tag{S24}$$

where $F_R$ is the regularized hypergeometric function and

$$X = 1 + C^2 + E - (1 + E)S, \tag{S25}$$
$$Y = C^2 - (E - 1)(S + 1). \tag{S26}$$

Moreover, for states with covariance matrix equal to Eq. (S18) up to a phase rotation on any mode, the photon number statistics is also given by Eq. (S24).

In our hypothesis testing protocols, $P(n_S, n_I)$ is fully determined by $N_S, N'_S, \kappa_S, \sqrt{1 - \kappa_I}, N_B$, where the last two $\sqrt{1 - \kappa_I}, N_B$ are the fixed environment parameters, idler loss and thermal noise. Given signal mean photon number $N_S$, our receiver, parametrized by the gain $G = 1 + N'_S$, is to jointly discriminate the photon count distributions per slot $\{P(\cdot, \cdot|\kappa_S, G, N_S)\}$ associated with different signal channel transmissivity patterns, conditioned on the same parameter setting $N'_S, \kappa_I, N_B$. Here we make the dependence of Eq. (S24) on $\kappa_S, G, N_S$ explicit.

## C. Conditional probabilities

In the main paper, we give the conditional probability of obtaining the $mM$-dimensional measurement results $\boldsymbol{n}_S = \{n_{S_L,k}\}_{L=1,k=1}^{M,m}$ and $\boldsymbol{n}_I = \{n_{I_L,k}\}_{L=1,k=1}^{M,m}$ conditioned on pattern $h$ in Eq. (7) of the main paper. Here we also print it for reference

$$P_m(\boldsymbol{n}_S, \boldsymbol{n}_I|h) = \prod_{L=1}^{M}\prod_{\ell=1}^{m} P(n_{S_{L,\ell}}, n_{I_{L,\ell}}|\kappa_\ell^{(h)}, G_\ell, N_{S\ell}). \tag{S27}$$

Each term $P(n_{S_{L,\ell}}, n_{I_{L,\ell}}|\kappa_\ell^{(h)}, G_\ell, N_{S\ell})$ is given in Eq. (S24), with the transmissivity $\kappa_S = \kappa_\ell^{(h)}$, also the gain distribution $G_\ell = 1 + N'_{S\ell}$ and the energy distribution $N_{S\ell}$ optimized probalistically. For the scenarios with strong symmetry, e.g. the peak positioning case, the optimal gain and energy distributions may be trivially uniform due to the symmetry.

For the absorption detection case, $m = 1$ and $H = 2$, the two hypotheses are $h = \{T, B\}$. Denote $M$-dimensional vectors of photon count $\boldsymbol{n}_S = \{n_{S_L}\}_{L=1}^{M}, \boldsymbol{n}_I = \{n_{I_L}\}_{L=1}^{M}$. Furthermore, with total energy being limited, the degree of freedom of $N_S$ is frozen. Eq. (S27) reduces to

$$P_m(\boldsymbol{n}_S, \boldsymbol{n}_I|h) = \prod_{L=1}^{M} P(n_{S_L}, n_{I_L}|\kappa_h, G), \tag{S28}$$

Similarly, for the single-peak positioning case, $H = m > 1$, each hypothesis $h \in [1, m]$ corresponds to the position of the target channel. Noting the symmetry in this scenario, we choose the uniform energy and gain distributions. Still, the degree of freedom $N_S$ is frozen and that of gain is limited to one. Eq. (S27) reduces to

$$P_m(\boldsymbol{n}_S, \boldsymbol{n}_I|h)$$
$$= \prod_{L=1}^{M}[P(n_{S_{L,h}}, n_{I_{L,h}}|\kappa_T, G) \cdot \prod_{l \neq h} P(n_{S_{L,l}}, n_{I_{L,l}}|\kappa_B, G)]. \tag{S29}$$

In general, one can tune the energy and gain distributions among the different slots to approach the optimal performance.

### D. Limiting cases of the photon statistics

To enable numerical simulation, we also need to deal with term-wise divergence in Eq. (S24), e.g. the hypergeometric function $F_R$ could diverge. Since the probability is normalized, all divergences are in fact cancelled out by pairing infinite terms with infinitesimal terms. This must be done in an analytical way before numerical calculations.

The divergence comes from special values of $z = 4C^2/XY$, as summarized in the following two cases:

#### 1. Case 1: $z \to \infty$

This can be caused by $X \to 0$ or $Y \to 0$. This only happens when all noise are zero. Then

$$
\lim_{Y \to 0} F_R(1+n_S, 1+n_I, 1, \frac{4C^2}{XY})/Y^{1+n_I} =
$$
$$
\begin{cases}
C^{-2(n_I+1)}(-4)^{-1-n_I}X^{1+n_I}(-1)^{n_S}\binom{n_I}{n_S}, & \text{if } n_I \geq n_S \\
0, & \text{otherwise}
\end{cases}
\tag{S30}
$$

Hence

$$
P(n_S, n_I) = (-1+C^2+E+S-ES)^{1+n_S+n_I}
$$
$$
(-4)^{-n_I}(-1)^{n_S}C^{-2(1+n_I)}X^{n_I-n_S}\binom{n_I}{n_S}
\tag{S31}
$$

for $n_I \geq n_S$ or 0 otherwise.

Two-mode squeezed vacuum is among this case, by taking $E, S = 1+2N_S$, $C = 2\sqrt{N_S(N_S+1)}$, we find $X = Y = 0$, which makes the above expression non-zero only when $n_I = n_S$, and indeed we have $P(n_S, n_I) = N_S^{n_S}/(1+N_S)^{n_S+1}\delta_{n_S-n_I}$.

A special scenario of case 1 happens when $z = \frac{0}{0}$. At this moment $X$ or $Y$ is zero, combined with $C = 0$. To avoid singularity $C^{-2(1+n_I)}$ in numerical calculation, we take the $C \to 0$ limit with Eq. (S31), which yields

$$
P(n_S, n_I) = \begin{cases}
2(E-1)^{n_S}(1+E)^{-1-n_S}, & \text{if } n_I = 0, \\
0, & \text{otherwise},
\end{cases}
\tag{S32}
$$

for $X = 0$ and

$$
P(n_S, n_I) = \begin{cases}
2(S-1)^{n_I}(1+S)^{-1-n_I}, & \text{if } n_S = 0 \\
0, & \text{otherwise}.
\end{cases}
\tag{S33}
$$

for $Y = 0$. Note that first case corresponds to signal mode thermal statistics

$$
P(n_S) = \frac{\overline{N}_S^{n_S}}{(1+\overline{N}_S)^{n_S+1}},
\tag{S34}
$$

with mean photon number $\overline{N}_S = (-1+E)/2$; the second case corresponds to idler mode thermal statistics

$$
P(n_I) = \frac{\overline{N}_I^{n_I}}{(1+\overline{N}_I)^{n_I+1}},
\tag{S35}
$$

with mean photon number $\overline{N}_I = (-1+S)/2$.

#### 2. Case 2: $z = 4C^2/XY \to 1$

In this case $F_R$ is infinite. Indeed it always comes with $-1+C^2+E+S-ES \to 0$ which cancels the singularity. We have

$$
P(n_S, n_I) = \left(\frac{E+S}{(E-1)(S-1)}\right)^{-1-n_S-n_I}(-1)^{-n_S-n_I}\times
$$
$$
2^{3+n_S+n_I}\frac{(2-2E)^{-1-n_I}(2-2S)^{-1-n_S}(n_S+n_I)!}{n_S!n_I!}.
\tag{S36}
$$

## III. ENTANGLE-ASSISTED HOMODYNE RECEIVER

We give a brief summary of the best known receiver design for quantum reading, a Bell-measurement receiver proposed in Ref. [S4], which we utilized to benchmark in Fig. 3(a) of the main paper. The returned signal-idler pair travels through a beamsplitter, which yields output modes $a_+ = (a_S + a_I)/\sqrt{2}$, $a_- = (a_S - a_I)/\sqrt{2}$. Then homodyne measurements are operated on the two quadratures $p_+ = (a_+ - a_+^\dagger)/i$, $q_- = (a_- + a_-^\dagger)$, which share the same variance $V_h = \sigma_h^2 = \langle p_+^2 \rangle = \langle q_-^2 \rangle = 1+N_S+N_S\kappa^{(h)}-2\sqrt{N_S(1+N_S)\kappa^{(h)}}$, $h = \{T, B\}$. With a large identical mode number $M$, the receiver design constructs $\chi^2$ test variable $\theta = \sum_{L=1}^{M} p_{+,L}^2 + q_{-,L}^2$, which comes with the probability density function

$$
P_M(\theta|h) = \frac{\theta^{M-1}\exp\left(-\frac{\theta}{2V_h}\right)}{(2V_h)^M\Gamma(M)},
\tag{S37}
$$

where $\Gamma(\cdot)$ is the Gamma function. Under maximum-likelihood decision, the error rate is $(1 - |F_T - F_B|)/2$, given the cumulative distribution function (cdf) of the chi-square distribution $F_h = \Gamma(M, \frac{t}{2V_h})/\Gamma(M)$ with the threshold $t = 2M\sigma_T^2\sigma_B^2\log(\sigma_T^2/\sigma_B^2)/(\sigma_T^2 - \sigma_B^2)$. Here $\Gamma(\cdot,\cdot)$ is the incomplete Gamma function

Generally, the Bell receiver, extracting the quantum correlations by interfering the signal and the idler with a beamsplitter, yields measurement results with a higher signal-to-noise ratio, thereby it has a significant advantage over a direct homodyne measurement in the classical scenario. Similarly, the nulling receiver discussed in the main text exploits the quantum correlation with the assistance of an OPA, which yields a significant advantage on a photon number resolving detector. In the following,

we calculate the performance of a homodyne detector assisted by OPA, where we find that the advantage remains but is undermined.

Now we consider a homodyne receiver, measuring the quadratures of the signal-idler pair nulled by an OPA with gain similarly defined by Eq. (S22). After the OPA, each of the two-mode signal-idler pair is in a Gaussian state with the covariance matrix given by Eq. (S18). For homodyne measurements, one can choose to measure the position $q_S$ or momentum $p_S$ quadrature of the signal mode and $q_I$ or $p_I$ for the idler mode. The two outcomes, defined as $x, y$, are Gaussian random variables characterized by the marginal covariance matrix. Observing the symmetry, we see that the measurement on $\{q_S, p_I\}$ has the same statistics with $\{p_S, q_I\}$, ditto for $\{q_S, q_I\}$ with $\{p_S, p_I\}$. Upon obtaining all measurement results across all mode-pairs, we utilize maximum likelihood estimation (MLE) to make the decision. In general, the error rate of this OPA-assisted homodyne receiver has no closed form.

To provide an outline of the practical performance, we give the numerical result of the ideal case $\kappa_B = 1$. In this case, the covariance matrix for the background channel is $\Lambda^{(B)} = \boldsymbol{I}$. We find it possible to eliminate the correlation between the two measurement outcomes $\{x, y\}$ per copy by a linear combination that diagonalizes the covariance matrix, mapping $\{x, y\}$ to $\{x', y'\}$. Then we have the weighted M-copy squared sum $s = \sum_{L=1}^{M} a x'^2_L + b y'^2_L$, as a sufficient statistic for the $2M$-dimensional Gaussian, thereby the $2M$-dimensional MLE reduces to 1-dimensional. Here the weights are $a = 1/\sigma^{2\,(T)}_{x'} - 1$, $b = 1/\sigma^{2\,(T)}_{y'} - 1$, where the variances $\sigma^2$'s are obtained from the diagonalized covariance matrix and we have used the fact $\sigma^{2\,(B)}_{x'} = \sigma^{2\,(B)}_{y'} = 1$ when $\kappa_B = 1$. Finally, we can express $s = a\sigma^{2\,(h)}_{x'} X_1 + b\sigma^{2\,(h)}_{y'} X_2$ as a weighted sum of two chi-square distributed random variables $X_1, X_2 \sim \chi^2(M)$. The corresponding distribution of $s$ under hypothesis $h$ is a generalized chi-square distribution, which can be numerically calculated. Then, the error probability of MLE is numerically obtained by an numerical integration.

Fig. S6 compares the Bell receiver with the OPA-assisted homodyne receivers and the classical limit, which illustrates the superiority of the Bell measurement. Note that the homodyne receiver measuring $\{q_S, p_I\}$ fails to achieve an advantage over the classical at this moment. This is due to the absence of the quantum correlation between $q_S$, $p_I$. Above all, considering its supremacy among the homodyne receivers, we compare our design only with the Bell receiver in the main text.
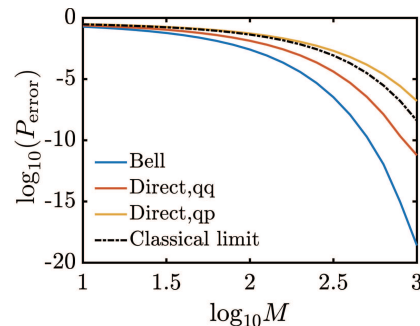


Figure S6. Error rates of the entanglement-assisted (EA) homodyne receivers with $M$ identical copies in the absorption detection scenario. We compare the performance of Bell receiver (blue solid), the OPA-assisted homodyne receivers measuring quadratures $\{q_1, q_2\}$ (red solid) and measuring quadratures $\{q_1, p_2\}$ (orange solid), with the classical limit Eq. (3) (black dashed). The mean photon number of the TMSV source is $N_S = 1$, the transmissivities of channels are $\kappa_T = 0.75$ and $\kappa_B = 1$.

## IV. ERROR PROBABILITY LOWER BOUND OF CLASSICAL PATTERN RECOGNITION

### A. Single-mode phase-insensitive Gaussian channels

As explained in Ref. [S5], the action of a single-mode (covariant) phase-insensitive Gaussian channel over input quadratures $\hat{\boldsymbol{x}} = (\hat{q}, \hat{p})^T$ can be represented by the transformation $\hat{\boldsymbol{x}} \to \sqrt{\mu}\hat{\boldsymbol{x}} + \sqrt{|1-\mu|}\hat{\boldsymbol{x}}_E + \xi$, where $\mu$ is a transmissivity ($0 \leq \mu \leq 1$) or a gain ($\mu \geq 1$), $\hat{\boldsymbol{x}}_E$ are the quadratures of an environmental mode in a thermal state with noise variance $\omega = 2N + 1$, with $N$ being the mean number of photons, and $\xi$ is additive classical noise, i.e., a random 2-D Gaussian distributed vector with covariance matrix $w_{\text{add}}\mathbf{I}$. Here we assume vacuum shot noise equal to 1.

Note that, for a coherent state at the input, the output state of the channel is generally thermal with covariance matrix $\mathbf{V} = (\mu + |1 - \mu|\omega + w_{\text{add}})\mathbf{I}$. Setting $\omega = (1 + 2E - w_{\text{add}} - \mu)/|1 - \mu|$, this matrix simply becomes $(2E + 1)\mathbf{I}$. Therefore, conditionally on a coherent state input, the channel can be described by the two parameters $\mu$ and $E$, which we denote as $\Phi_{\mu, E}$. In particular, for a thermal-loss channel, we have $1 \leq \mu \leq 1$, and $E = (\omega - 1)(1 - \mu)/2 = (1 - \mu)N$; for a noisy amplifier, we have $\mu \geq 1$, and $E = (\omega + 1)(\mu - 1)/2 = (\mu - 1)(N + 1)$; and finally, for an additive Gaussian noise channel, we have $\mu = 1$ and $E = w_{\text{add}}/2$.

## B. Ultimate lower bounds for general patterns

Consider the general case of hypothesis testing between $H$ different patterns of composite channels

$$\mathcal{E}_n = \otimes_{\ell=1}^{m} \left( \Phi_{\mu_\ell^{(n)}, E_\ell^{(n)}}^{\otimes M} \right)_{S_\ell}, 1 \leq n \leq H, \qquad (S38)$$

each acting on $m$ subsystems $\{S_\ell\}_{\ell=1}^m$. Here each subsystem consists $M$ modes and each mode goes through a single-mode phase-insensitive bosonic Gaussian channel $\Phi_{\mu_\ell^{(n)}, E_\ell^{(n)}}$. Therefore, a global channel $\mathcal{E}_n$ is specified by coefficients $\{\mu_\ell^{(n)}, E_\ell^{(n)}\}_{\ell=1}^m$.

**Lemma 1** *Consider classical states (with positive P-representation) as the input, assuming a global energetic constraint of $mMN_S$ mean photons with $M$ modes irradiated over each of the $m$ subsystems $S_k$, the error probability of equal-prior hypothesis testing between $H$ channels $\mathcal{E}_n$ in Eq. (S38) is lower bounded by*

$$P_H \geq P_{H,LB}^C = \frac{H-1}{2H} \bar{c}^2 \exp\left[-B^\star mMN_S\right], \qquad (S39)$$

*where*

$$\bar{c}^K = \prod_{n'>n} \prod_{\ell=1}^m$$

$$\left\{1 + \left[\sqrt{E_\ell^{(n)}(1+E_\ell^{(n')})} - \sqrt{E_\ell^{(n')}(1+E_\ell^{(n)})}\right]^2\right\}^{-M/2}, \qquad (S40)$$

$$B^\star = \max_\ell \frac{1}{K} \sum_{n'>n} \frac{(\sqrt{\mu_\ell^{(n)}} - \sqrt{\mu_\ell^{(n')}})^2}{1 + E_\ell^{(n)} + E_\ell^{(n')}}, \qquad (S41)$$

*with $K = H(H-1)/2$.*
*A tighter lower bound*

$$P_{H,LB}^C = \frac{K}{H^2} f_\star^2 \qquad (S42)$$

*can be achieved by solving a constrained optimization*

$$f_\star = \min_{\{X_\ell\}} : \sum_{n'>n} \frac{1}{K} C_{n,n'} \exp\left[-\frac{1}{2}\sum_{\ell=1}^m B_\ell^{(n,n')} X_\ell\right] \quad (S43)$$

$$under\ constraint\ \sum_{\ell=1}^m X_\ell \leq mMN_S. \qquad (S44)$$

*The constants*

$$B_\ell^{(n,n')} = \frac{(\sqrt{\mu_\ell^{(n)}} - \sqrt{\mu_\ell^{(n')}})^2}{1 + E_\ell^{(n)} + E_\ell^{(n')}} \qquad (S45)$$

$$C_{n,n'} =$$

$$\prod_{\ell=1}^m \left[\frac{1}{1 + \left(\sqrt{E_\ell^{(n)}(1+E_\ell^{(n')})} - \sqrt{E_\ell^{(n')}(1+E_\ell^{(n)})}\right)^2}\right]^{M/2}. \qquad (S46)$$

**Corollary 2** *(classical channel-position finding, originally derived in [S5]) When $H = m$,*

$$\mu_\ell^{(n)} = \begin{cases} \mu_T & if\ n = \ell; \\ \mu_B & if\ n \neq \ell, \end{cases} \qquad (S47)$$

*and*

$$E_\ell^{(n)} = \begin{cases} E_T & if\ n = \ell; \\ E_B & if\ n \neq \ell, \end{cases} \qquad (S48)$$

*the pattern recognition problem corresponds to channel-position finding with a target channel $\Phi_{\mu_T, E_T}^{\otimes M}$ among $(m-1)$ background channels $\Phi_{\mu_B, E_B}^{\otimes M}$, where the above bound in Ineq. (S39) reduces to*

$$P_{H,LB} = \frac{m-1}{2m} c_{E_B, E_T}^{2M} \times$$

$$\exp\left[-\frac{2MN_S(\sqrt{\mu_B} - \sqrt{\mu_T})^2}{1 + E_B + E_T}\right], \quad (S49)$$

*with $c_{E_B, E_T} = [1 + \left(\sqrt{E_B(1+E_T)} - \sqrt{E_T(1+E_B)}\right)^2]^{-1}$. In particular, for no passive signature ($E_T = E_B \equiv E$), we have the simplification*

$$P_{H,LB} = \frac{m-1}{2m} \exp\left[-\frac{2MN_S(\sqrt{\mu_B} - \sqrt{\mu_T})^2}{1 + 2E}\right], \quad (S50)$$

**Remark 3** *The corollary is easy to obtain from Lemma 1, recognizing $\bar{c}^2 = c_{E_B, E_T}^{2M}$ and $\sqrt{\mu_\ell^{(n)}} - \sqrt{\mu_\ell^{(n')}}$ is zero unless $\ell = n$ or $\ell = n'$, therefore we have*

$$B^\star = \frac{2}{m}(\sqrt{\mu_B} - \sqrt{\mu_T})^2/(1 + E_B + E_T), \qquad (S51)$$

*which easily leads to Eq. (S49).*

**Corollary 4** *(classical channel-position finding with multiple target channels) When there are $k$ target channels $\Phi_{\mu_T, E_T}^{\otimes M}$ among $(m-k)$ background channels $\Phi_{\mu_B, E_B}^{\otimes M}$, we have $H = C_m^k$ patterns. For this case, the lower bound in Eq. (S39) reduces to*

$$P_{H,LB} =$$

$$\frac{H-1}{2H} c_{E_B, E_T}^{2Mw_{m,k}} \exp\left[-\frac{2w_{m,k}MN_S(\sqrt{\mu_B} - \sqrt{\mu_T})^2}{1 + E_B + E_T}\right], \qquad (S52)$$

*with*

$$w_{m,k} \equiv m \frac{C_{m-1}^{k-1} C_{m-1}^k}{C_m^k(C_m^k - 1)} = \frac{kC_{m-1}^k}{C_m^k - 1} \qquad (S53)$$

*For no passive signature ($E_T = E_B \equiv E$) case, the above bound in Ineq. (S39) reduces to*

$$P_{H,LB} = \frac{H-1}{2H} \exp\left[-\frac{2w_{m,k}MN_S(\sqrt{\mu_B} - \sqrt{\mu_T})^2}{1 + 2E}\right]. \qquad (S54)$$

**Remark 5** *Similarly, the corollary is easy to obtain from Lemma 1. First, one realizes that due to the symmetry, the maximization in $B^\star$ is achieved by an arbitrary $\ell$, e.g. we let $\ell = 1$. And then the non-zero contribution to Eq. (S41) only comes from patterns with different channels on subsystem $\ell = 1$. Slight simplification of Eq. (S54) leads to Eq. (4) of the main paper.*

**Proof.** For the convenience of analysis, we will parameterize a coherent state $|\alpha\rangle$ with the phase and amplitude squared, i.e., $|x, \theta\rangle \equiv |\sqrt{x}e^{i\theta}\rangle$, where $x \geq 0$ and $0 \leq \theta \leq 2\pi$. In this notation, a multi-mode coherent state over the entire system takes the form $|\boldsymbol{x}, \boldsymbol{\theta}\rangle = \otimes_{\ell=1}^m (|\boldsymbol{x}_\ell, \boldsymbol{\theta}_\ell\rangle_{S_\ell})$, where each subsystem state $|\boldsymbol{x}_\ell, \boldsymbol{\theta}_\ell\rangle_{S_\ell} = \otimes_{\ell'=1}^M |x_\ell^{(\ell')}, \theta_\ell^{(\ell')}\rangle$ is again a tensor product of multiple modes with generally-different amplitudes. Here $\boldsymbol{x}_\ell$ are positive and real vectors $\boldsymbol{x}_\ell = (x_\ell^{(1)}, \cdots, x_\ell) \equiv \{x_\ell^{(\ell')}\}_{\ell'=1}^M$ and $\boldsymbol{x}$ is a simple concatenation of them, i.e., $\boldsymbol{x} = (\boldsymbol{x}_1, \cdots, \boldsymbol{x}_m)$.

In this notation, the general classical state as the input can be written as a Lebesgue integral

$$\rho = \int dP \, |\boldsymbol{x}, \boldsymbol{\theta}\rangle \langle \boldsymbol{x}, \boldsymbol{\theta}|, \tag{S55}$$

where the probability measure $P$ over $\boldsymbol{x}, \boldsymbol{\theta}$ can be arbitrary. Let us define

$$\|\boldsymbol{x}\|_1 \equiv \sum_{\ell, \ell'} |x_\ell^{(\ell')}| = \sum_{\ell, \ell'} x_\ell^{(\ell')}, \tag{S56}$$

which is the standard one-norm and equals the total mean photon number of the state $|\boldsymbol{x}, \boldsymbol{\theta}\rangle$. Then, the total energy constraint leads to the inequality

$$\int dP' \|\boldsymbol{x}\|_1 \leq mMN_S, \tag{S57}$$

where the integral has been simplified to a marginal probability measure $P'$ restricted to the non-negative variables $\boldsymbol{x}$.

The total conditional state at the output of the channel $\mathcal{E}_n$ is also a mixture, with expression

$$\rho_n^C = \mathcal{E}_n(\rho) = \int dP \rho_{\boldsymbol{x}, \boldsymbol{\theta}, n}^C, 1 \leq n \leq H, \tag{S58}$$

where each conditional state is given by

$$\rho_{\boldsymbol{x}, \boldsymbol{\theta}, n}^C = \otimes_{\ell=1}^m (\rho_{\mu_\ell^{(n)}, E_\ell^{(n)}}^C)_{S_\ell} \tag{S59}$$

The state $(\rho_{\mu_\ell^{(n)}, E_\ell^{(n)}}^C)_{S_\ell}$ is a product of $M$ displaced thermal states, each with amplitude $\sqrt{\mu_\ell^{(n)} x_\ell^{(\ell')}} e^{i\theta_\ell^{(\ell')}}$ and covariance matrix $(2E_\ell^{(n)} + 1)\boldsymbol{I}$.

We use the fidelity-based lower bound of Helstrom limit [S6],

$$P_H \geq P_{H,LB} \equiv \sum_{k'>k} p_{k'} p_k F^2(\rho_{k'}, \rho_k). \tag{S60}$$

From Eq. (S60), we can write the following lower bound to the mean error probability. Consider the equal prior case for simplicity.

$$P_{H,LB}^C = \sum_{n'>n} \frac{1}{H^2} F^2 \left[ \int dP \rho_{\boldsymbol{x}, \boldsymbol{\theta}, n}^C, \int dP \rho_{\boldsymbol{x}, \boldsymbol{\theta}, n'}^C \right]$$

$$\geq \frac{K}{H^2} \sum_{n'>n} \frac{1}{K} \left\{ \int dP F[\rho_{\boldsymbol{x}, \boldsymbol{\theta}, n}^C, \rho_{\boldsymbol{x}, \boldsymbol{\theta}, n'}^C] \right\}^2$$

$$\geq \frac{K}{H^2} \left\{ \sum_{n'>n} \frac{1}{K} \int dP F[\rho_{\boldsymbol{x}, \boldsymbol{\theta}, n}^C, \rho_{\boldsymbol{x}, \boldsymbol{\theta}, n'}^C] \right\}^2, \tag{S61}$$

where use the joint concavity of fidelity

$$F\left[ \int dp_x \rho_x, \int dp_x \sigma_x \right] \geq \int dp_x F[\rho_x, \sigma_x], \tag{S62}$$

and Jensen's inequality for the square function with $K = (H-1)H/2$.

Let us now address each fidelity term

$$F_{n,n'}^C \equiv F[\rho_{\boldsymbol{x}, \boldsymbol{\theta}, n}^C, \rho_{\boldsymbol{x}, \boldsymbol{\theta}, n'\neq n}^C]$$

$$= F[\otimes_{\ell=1}^m (\rho_{\mu_\ell^{(n)}, E_\ell^{(n)}}^C)_{S_\ell}, \otimes_{\ell=1}^m (\rho_{\mu_\ell^{(n')}, E_\ell^{(n')}}^C)_{S_\ell}] \tag{S63}$$

$$= \prod_{\ell=1}^m F[(\rho_{\mu_\ell^{(n)}, E_\ell^{(n)}}^C)_{S_\ell}, (\rho_{\mu_\ell^{(n')}, E_\ell^{(n')}}^C)_{S_\ell}] \tag{S64}$$

Using Gaussian fidelity formula [S1], we can compute

$$F[(\rho_{\mu_\ell^{(n)}, E_\ell^{(n)}}^C)_{S_\ell}, (\rho_{\mu_\ell^{(n')}, E_\ell^{(n')}}^C)_{S_\ell}]$$

$$= \left( c_\ell^{(n,n')} \right)^{M/2} \exp\left[ -\frac{1}{2} B_\ell^{(n,n')} \|\boldsymbol{x}_\ell\|_1 \right] \tag{S65}$$

where the constant

$$B_\ell^{(n,n')} = \frac{(\sqrt{\mu_\ell^{(n)}} - \sqrt{\mu_\ell^{(n')}})^2}{1 + E_\ell^{(n)} + E_\ell^{(n')}} \tag{S66}$$

$$c_\ell^{(n,n')} = \frac{1}{1 + \left( \sqrt{E_\ell^{(n)}(1 + E_\ell^{(n')})} - \sqrt{E_\ell^{(n')}(1 + E_\ell^{(n)})} \right)^2}. \tag{S67}$$

Note that $B_\ell^{(n,n')} > 0$ and $c_\ell^{(n,n')} \leq 1$. From the one-norm in the expression above, it becomes clear that the performance is exactly the same regardless how the energy is distributed among the $M$ modes impinging on a subsystem, as long as the mean total energy irradiated over the subsystem is fixed. Therefore we have

$$F_{n,n'}^C = C_{n,n'} \exp\left[ -\frac{1}{2} \sum_{\ell=1}^m B_\ell^{(n,n')} \|\boldsymbol{x}_\ell\|_1 \right], \tag{S68}$$

where we defined $C_{n,n'} = \prod_{\ell=1}^m \left( c_\ell^{(n,n')} \right)^{M/2}$. By replacing the $F_{n,n'}^C$ in Eq. (S61), and noticing that $F_{n,n'}^C$ does not depend on $\boldsymbol{\theta}$ we find the following lower bound

$$P_{H,LB}^C \geq \frac{K}{H^2} \left\{ \int dP' g(\{\|\boldsymbol{x}_\ell\|_1\}_{\ell=1}^m) \right\}^2, \tag{S69}$$

where we define the function

$$g(\{\|\boldsymbol{x}_\ell\|_1\}_{\ell=1}^m) \equiv$$

$$\sum_{n'>n} \frac{1}{K} C_{n,n'} \exp\left[-\frac{1}{2}\sum_{\ell=1}^m B_\ell^{(n,n')}\|\boldsymbol{x}_\ell\|_1\right]. \qquad (S70)$$

We can use the convexity of $e^{-cx}$ (with $c > 0$) and Jensen's inequality to move expectation value to the exponent

$$P_{H,LB}^C \geq \frac{K}{H^2}\left\{\sum_{n'>n}\frac{1}{K}C_{n,n'}\exp\left[-\frac{1}{2}\sum_{\ell=1}^m B_\ell^{(n,n')}X_\ell\right]\right\}^2, \qquad (S71)$$

where $X_\ell = \int dP'\|\boldsymbol{x}_\ell\|_1$. Note that equality is only achieved when $P'$ corresponds to a delta function. Overall we want to solve the minimization under constraint Ineq. (S57)

$$f_\star = \min : \sum_{n'>n}\frac{1}{K}C_{n,n'}\exp\left[-\frac{1}{2}\sum_{\ell=1}^m B_\ell^{(n,n')}X_\ell\right] \qquad (S72)$$

$$\text{under constraint } \sum_{\ell=1}^m X_\ell \leq mMN_S. \qquad (S73)$$

Then the lower bound would be

$$P_{H,LB}^C \geq \frac{K}{H^2}f_\star^2. \qquad (S74)$$

This gives the lower bound in Eq. (S42).

Below we obtain a further lower bound.

$$g(\{\|\boldsymbol{x}_\ell\|_1\}_{\ell=1}^m)$$

$$\geq \left\{\prod_{n'>n}\left(C_{n,n'}\exp\left[-\frac{1}{2}\sum_{\ell=1}^m B_\ell^{(n,n')}\|\boldsymbol{x}_\ell\|_1\right]\right)\right\}^{1/K} \qquad (S75)$$

$$= \bar{c}\exp\left[-\frac{1}{2K}\sum_{n'>n}\sum_{\ell=1}^m B_\ell^{(n,n')}\|\boldsymbol{x}_\ell\|_1\right] \qquad (S76)$$

where we have introduced $\bar{c} = \left\{\prod_{n'>n}C_{n,n'}\right\}^{1/K}$. We have also used $K = H(H-1)/2$ and the fact that arithmetic mean is greater than geometric mean. The equality holds if and only if $c_{n,n'}\exp\left[-\sum_{\ell=1}^m B_\ell^{(n,n')}\|\boldsymbol{x}_\ell\|_1\right]$ is equal for all $n, n'$.

Thus overall we may write

$$P_{H,LB}^C \geq \frac{\bar{c}^2 K}{H^2}\left\{\int dP'\exp\left[-\frac{1}{2K}\sum_{n'>n}\sum_{\ell=1}^m B_\ell^{(n,n')}\|\boldsymbol{x}_\ell\|_1\right]\right\}^2$$

$$\geq \frac{\bar{c}^2 K}{H^2}\left\{\exp\left[-\int dP'\frac{1}{2K}\sum_{n'>n}\sum_{\ell=1}^m B_\ell^{(n,n')}\|\boldsymbol{x}_\ell\|_1\right]\right\}^2$$

$$= \frac{H-1}{2H}\bar{c}^2\exp\left[-\frac{1}{K}\sum_{n'>n}\sum_{\ell=1}^m B_\ell^{(n,n')}X_\ell\right]. \qquad (S77)$$
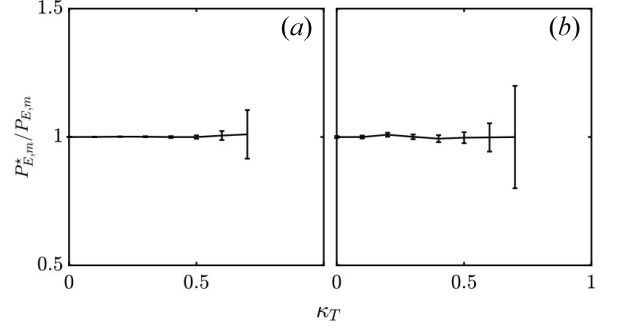


Figure S7. Verification of the analytical formulae. We compare the numerical result $P_{E,m}^\star$ with the analytical result $P_{E,m}$. $M$ is valued such that the classical lower bound $P_{c,m,LB} = 10^{-1}$. (a) the absorption detection Eq. (S91). $\kappa_B = 1, N_S = 1$. (b) Verification of the peak positioning Eq. (S93). $\kappa_B = 1, N_S = 1, m = 100$. In the simulation we take $10^8$ samples for the absorption detection case and $2\times 10^6$ samples for the peak positioning case.

For the second inequality, we use the convexity of $e^{-cx}$ (with $c > 0$) and Jensen's inequality to move expectation value to the exponent.

To minimize the lower bound in Eq. (S77), we need to solve the constrained (by constraint in Eq. (S57)) optimization

$$\max \sum_{\ell=1}^m \frac{1}{K}\sum_{n'>n} B_\ell^{(n,n')}X_\ell$$

$$\text{under constraint } \sum_{\ell=1}^m X_\ell \leq mMN_S. \qquad (S78)$$

The solution to the maximization is simple, denote $\ell^\star = \arg\max_\ell \sum_{n'>n} B_\ell^{(n,n')}$ and $B^\star = \max_\ell \frac{1}{K}\sum_{n'>n} B_\ell^{(n,n')}$, and let

$$X_{\ell^\star} = mMN_S, X_{\ell\neq\ell^\star} = 0, \qquad (S79)$$

which leads to the maximum $B^\star mMN_S$.

Overall we have the lower bound

$$P_{H,LB}^C \geq \frac{H-1}{2H}\bar{c}^2\exp\left[-B^\star mMN_S\right], \qquad (S80)$$

It is easy to check that the lower bound can be reached only if

$$c_{n,n'}\exp\left[-B_{\ell^\star}^{(n,n')}mMN_S\right] = \text{const}, \qquad (S81)$$

is independent of $n, n'$. This is not always possible and therefore the lower bound is only achievable in certain symmetric cases. In symmetric cases, when $\max_\ell \frac{1}{K}\sum_{n'>n} B_\ell^{(n,n')}$ are equal for all $\ell$, then one can evenly distribute the energy to achieve the lower bound. ∎
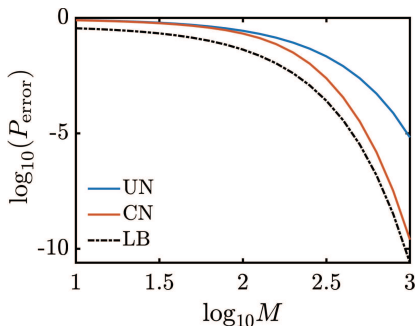
Figure S8. Error rate of the classical nuller with $M$ identical copies in the 100-slot peak positioning scenario. We compare the unconditional nuller (UN, blue solid) and the conditional nuller (CN, red solid) with the classical lower bound Eq. (4) in the main text (LB, black dashed). Source mean photon number $N_S = 1$, channel transmissivity $\kappa_T = 0.75$, $\kappa_B = 0.95$.

## V. CLASSICAL STRATEGY: NULLING RECEIVER FOR COHERENT STATE INPUT

Refs. [S7, S8] have conducted a comprehensive analysis on the classical nulling receiver in the noiseless scenario $N_B = 0$. We briefly summarize the pertinent conclusions here, and compare the performance with our lower bound. The channel setup in the classical scenario is almost identical to that in the quantum scenario, the only difference being that the source is classical, e.g. coherent states generated by a laser.

### A. Absorption Detection

In this scenario, the task is to distinguish the target channel with transmissivity $\kappa_T$ with the background channel $\kappa_B$. Given the source state being $M$ copies of coherent state $\left|\sqrt{N_S}\right\rangle$ with total mean photon number $MN_S$, the optimal error rate is tightly bounded by the Helstrom limit [S4]

$$P_H = \frac{1}{2}\left[1 - \sqrt{1 - e^{-(\sqrt{\kappa_T} - \sqrt{\kappa_B})^2 MN_S}}\right]. \quad \text{(S82)}$$

Dolinar has proposed an adaptive receiver [S7] that reaches the Helstrom bound. In this case, the improved classical ultimate lower bound Eq. (3) in the main text coincides with this bound, which results from the fact that encoding with two pure states instead of an ensembles, i.e. concentrating the source energy on a single amplitude level, has achieved the optimum.

### B. Peak Positioning

Consider single-peak position among $m$ slots. We categorize the nulling receiver into unconditional nuller and condition nuller. The unconditional nuller, which motivated our entanglement-assisted design in the main text, applies identical displacement $D(\sqrt{\kappa_B N_S})$ to all $m$ modes. In the noiseless case $N_B = 0$, the return is in coherent states. The uniform displacement nulls the returned copies through the background channels to vacuum and those through the target channel to yet a coherent state with displacement $\sqrt{\kappa_T N_S} - \sqrt{\kappa_B N_S}$. Then we apply photon counting on every mode, which immediately identifies the target channel if any click is detected. The error only occurs when the coherent state associated with the target channel yields zero photon count with the false-negative error rate $p = \exp\left[-(\sqrt{\kappa_T} - \sqrt{\kappa_B})^2 MN_S\right]$ for source mean photon energy $MN_S$. The error rate of unconditional nuller is then

$$P_{C,UN} = \frac{m-1}{m}p = \frac{m-1}{m}\exp\left[-MN_S(\sqrt{\kappa_B} - \sqrt{\kappa_T})^2\right]. \quad \text{(S83)}$$

The conditional nuller, however, applies a mode-by-mode sequence of displacements dependent on the prior measurement results. Specifically, we null the first mode and measure its photon count. If no photon is detected at the first mode, our hypothesis that the target channel be at the first mode is partially confirmed, and we forgo the nulling on the remaining modes unless any photon is detected in the subsequent measurements. Note that in this noiseless case the false positive error rate is zero, any nonzero photon count is a conclusive evidence in favor of rejecting the current hypothesis for the currently measured mode, which immediately gives the conclusion if the rejected hypothesis is 'background'. On the other hand, if any photon is detected at the first mode, the target hypothesis is conclusively rejected, we move forward to the next hypothesis that the target channel be at the second mode. In sum, the error only occurs if both measurements on the target channel and the background channel mistook by target yield false negative errors. By iteration the relation $P_{C,m} = [(1-p)P_{C,m-1} + p^2] \times (m-1)/m$, the error rate of the conditional nuller is

$$P_{C,CN} = \frac{(1-p)^m + mp - 1}{m}. \quad \text{(S84)}$$

When $M \gg 1$, we have

$$P_{C,CN} \sim \frac{m-1}{2}\exp[-2MN_S(\sqrt{\kappa_B} - \sqrt{\kappa_T})^2]. \quad \text{(S85)}$$

In this limit, the conditional nuller loses to the classical lower bound Eq. (S54) (here $E = 0$) merely by a constant factor $m$, achieving the bound in the exponent indeed.

Fig. S8 compares the two nulling receivers with the classical lower bound Eq. (4) in the main text. It is verified that the unconditional nuller is overwhelmed by the conditional nuller. Furthermore, the latter is shown close to the classical lower bound in the decaying rate, which never achieves the bound though as expected.
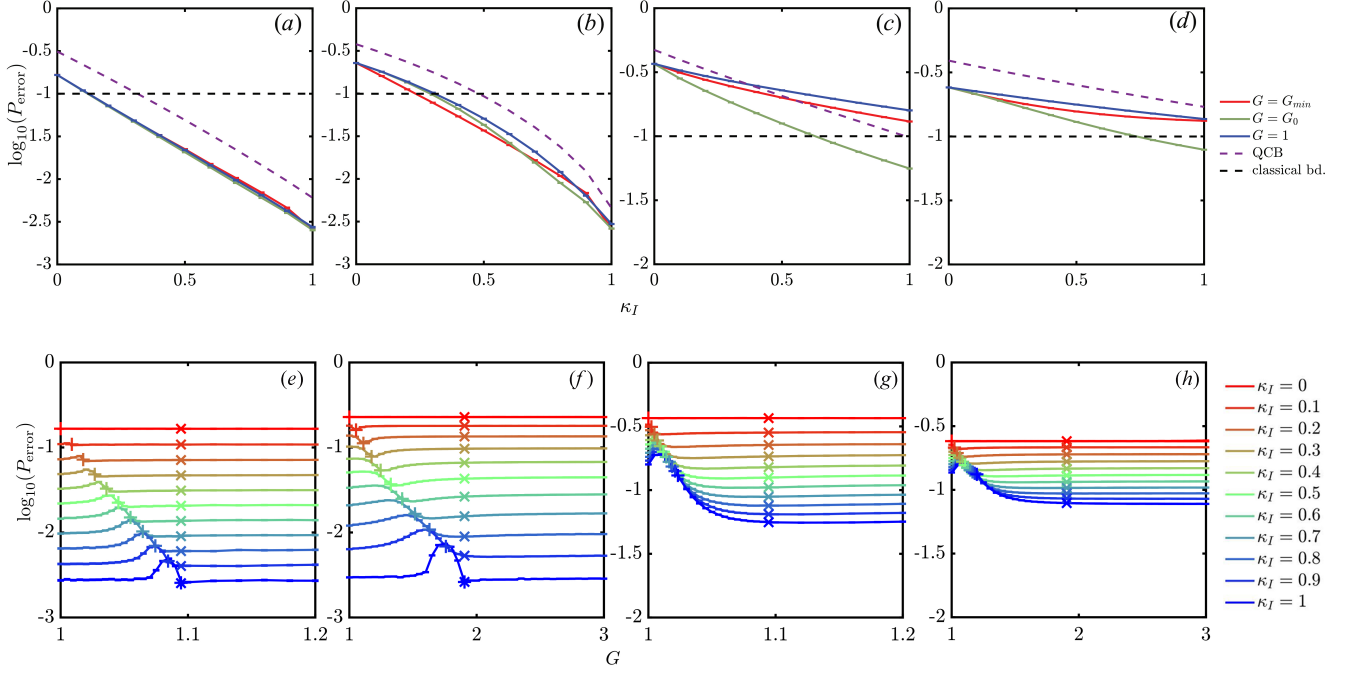
Figure S9. The effect of idler storage efficiency $\kappa_I$ (a-d) and the anti-squeezing gain $G$ (e-h) on the performance of our entanglement-assisted nulling receiver. (a)(e)$N_S = 0.1$, $N_B = 0$. (b)(f)$N_S = 1$, $N_B = 0$. (c)(g)$N_S = 0.1$, $N_B = 1$; (d)(h)$N_S = 1$, $N_B = 1$. For the gain $G$, we compare the consequential error rate between a wide range of values, including $G_0$, the default setting in the main text, and $G_{min}$, which minimizes the signal mode thermal fluctuation. With the parameter setting in this figure $G_{min} < G_0$. Identical mode number $M$ is chosen such that the classical lower bound $P_{C,1,LB}$ (black dashed line) is fixed to be 0.1. $\kappa_T = 0.75, \kappa_B = 0.95$. Note that in (c) the QCB and classical bound coincidentally get close when $\kappa_I = 1$, this is due to finite $M$ and they will deviate as $M$ further increases.

## VI. ANALYTICAL SOLUTIONS OF THE ERROR PROBABILITY

Here we present analytical solutions for the error probability in absorption detection and single-peak positioning, when $\kappa_B = \kappa_I = 1, N_B = 0$.

### A. Absorption detection

In this case, Eq. (S22) yields the gain $G = 1 + N_S$. At this moment the squeezing $\mathcal{S}$ is exactly the inverse of the two-mode squeezing operation that creates the TMSV state from vacuum. From the covariance matrix $\mathbf{\Lambda}(\kappa_B)$ in Eq. (S23), the background signal $\Xi^{(B)}$ is nulled to the vacuum state, a pure state capable to be discriminated with zero error rate. Applied on the photon statistics $P_m(\boldsymbol{n}_S, \boldsymbol{n}_I|h)$, the maximum-likelihood decision rule accepts the hypothesis $\tilde{h} = T$ for all the photon count results except for zero counts, which leads to the hypothesis to $\tilde{h} = B$. In this case, the error only happens when the nulled target-present state $\mathcal{S}(\Xi^{(T)})$ yields no photon count. Hence the error rate is $P_{E,1} = P_m(\boldsymbol{0}, \boldsymbol{0}|T)/2$. Given the transmissivities of background and target present channels $\kappa_B$, $\kappa_T$, and the mean photon number constraint on the source $N_S$, we

have the covariances $\mathbf{\Lambda}(\kappa_T)$ of each pair in $\mathcal{S}(\Xi^{(T)})$ given by Eq. (S18) with

$$E = 1 + 2N_S(1 + N_S)(1 - \sqrt{\kappa_T})^2, \qquad (S86)$$
$$S = 1 + 2N_S[(1 - \sqrt{\kappa_T})(2 + N_S(1 - \sqrt{\kappa_T}))], \qquad (S87)$$
$$C = 2[(N_S(1 - \sqrt{\kappa_T}) + 1)(1 - \sqrt{\kappa_T})\sqrt{N_S(1 + N_S)}]. \qquad (S88)$$

Thus

$$P_{E,1} = P(\boldsymbol{0}, \boldsymbol{0}|T)/2 = [P(0,0)]^M/2 \qquad (S89)$$
$$= \frac{(C^2 + E + S - ES - 1)}{C^2}, \qquad (S90)$$

where $P(n_S, n_I)$ is given by Eq. (S24). Finally

$$P_{E,1} = \frac{1}{2}\left[\frac{1}{1 + N_S(1 - \sqrt{\kappa_T})}\right]^{2M}. \qquad (S91)$$

### B. Peak positioning

The simplest case is when $\kappa_B = 1$, which is analytically solvable. First, the receiver nulls all the returned slot by uniformly applying the squeezing process $\mathcal{S}$ with the

(a) $N_S = 0.1$. $R_E/R_{QCB}$     (b) $N_S = 1$. $R_E/R_{QCB}$     (c) $N_S = 0.1$. $R_E/E_C$     (d) $N_S = 1$. $R_E/E_C$
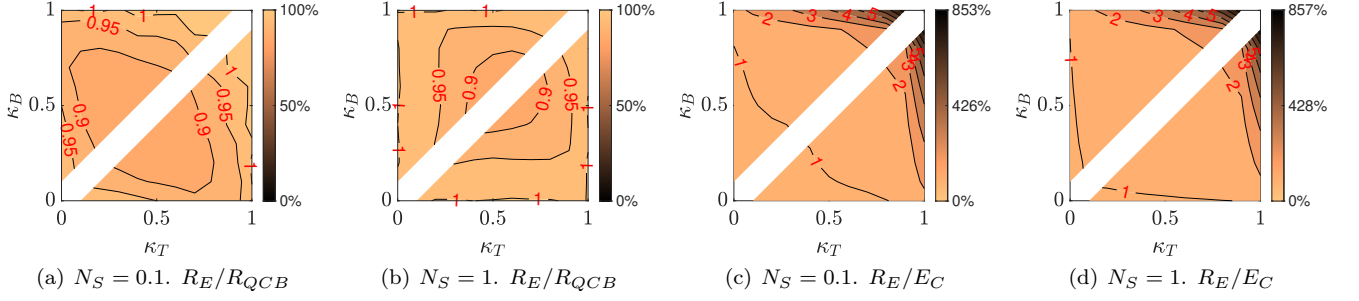
Figure S10. (a)(b) Error exponent ratio $R_E/R_{QCB}$ of noiseless absorption detection over QCB with respect to the transmissivity $\kappa_B, \kappa_T$. Based on Monte Carlo simulation with $5 \times 10^6$ ($N_S = 0.1$) and $5 \times 10^7$ ($N_S = 1$) samples. Standard deviation is (a) below 1% in all data points. (b) below 0.3% in all data points. The difference in the standard deviation results from the fact that more modes are required in (a), therefore the Monte Carlo simulation is more costly and the sample size is traded off. (c)(d) Error exponent ratio of $R_E/E_C$ in the same parameter setting.

same $N_S' = N_S$. Then the photon number per slot is measured. The slot with any click detected is the target-present slot. The errors only occur when no photon is detected among all slots. At this moment we randomly guess to make a decision. Hence the error rate is

$$P_E = \frac{m-1}{m} P_m(\mathbf{0}, \mathbf{0}|T) \tag{S92}$$

with the same $E, S, C$ as in the single slot case. Similarly we have

$$P_{E,m} = \frac{m-1}{m} \left[ \frac{1}{1 + N_S(1 - \sqrt{\kappa_T})} \right]^{2M} \tag{S93}$$

### C. Numerical verification of the formula

The solutions in Eq. (S91) and (S93) can be summarized into a unified formula, as done in Eq. (5) of the main paper.

The solvable case requires the ideal channel with $\kappa_I = 1, N_B = 0$. As shown in Fig. S7(a), the analytical formula in Eq. (S91) fits the numerical results well in the absorption detection case and achieves the QCB. As shown in Fig S7(b), the analytical formula in Eq. (S93) fits the numerical results well for the peak positioning case.

## VII. DETAILS ON NUMERICAL SIMULATION AND FULL RESULTS

### A. Monte Carlo simulation

In the main paper, we evaluate the performance of EAAS through Monte Carlo simulations. First we randomly choose the true hypothesis $h_0$ and simulate the random outcomes $\{\boldsymbol{n}\}$ measured from the quantum state corresponding to $h_0$. Based on the measurement results, the conditional probabilities $p(\boldsymbol{n}|h)$ are generated accordingly and we make the maximum likelihood decision

$\tilde{h} = \text{argmax}_h p(\boldsymbol{n}|h)$. Note that when there are multiple maximum hypotheses with equal probability, we make a random guess among them. Finally we figure out the frequency of $\tilde{h} \neq h_0$ as an estimation of the error probability $P_{E,m}$ of EAAS.

The sample size, i.e., number of total simulations, determines the precision of the estimation. According to the central limit theorem, larger sample size yields lower estimation variance $\sigma_{MC}^2$ in inverse proportion. However, we have to balance the computation cost while targeting at higher precision results. As a result, a variety of sample sizes are chosen according to the specific scenarios in the main paper: Fig. 4(a) with $10^8$ and (b)(d) with $10^6$; Figs. 4(c) and 5 with $10^7$. With these sample sets, we have checked the validity and precision of the Monte Carlo simulation by comparing with the random guess error rate $(H-1)/H$ of the H-identical-hypothesis case in the $N_S \to 0$ limit, and the analytical formulae Eqs. (S91) and (S93) for ideal cases $\kappa_B = 1$ in Fig. S7. The differences all lie in the $\pm 3\sigma_{MC}$ 99.7% confidence interval. In the supplemental material, we give the sample size directly in each figure involving Monte Carlo simulations.

### B. Optimal gain

Indeed, in all ideal cases with no noise nor idler loss, the gain $G_0 = 1 + N_{S0}'$ nulling the signal to vacuum is given by Eq. (S22). However the optimal gain that minimizes the error rate is not necessarily the same. We denote the optimal gain $G^{(opt)}$. Indeed, numerical results show that with most $\kappa_B, \kappa_T$ values under the ideal parameter $N_B = 0, \kappa_I = 1$, $G_0$ saturates QCB when $MN_S$ is relatively large, i.e. optimal gain $G^{(opt)} = G_0$. When $MN_S$ is small, $G^{(opt)}$ may deviate from $G_0$ and fluctuate in the neighbourhood.

When the idler storage efficiency $\kappa_I < 1$, although impeded from nulling the signal mode to vacuum, intuitively we expect to minimize the mean photon number of it. This requires another setting $G_{min} =$
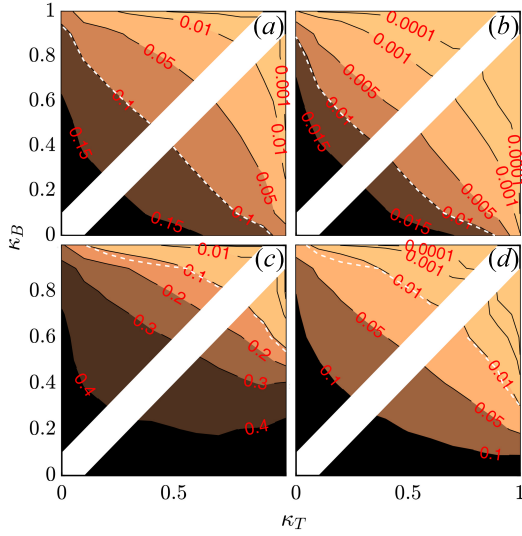
Figure S11. Error probability of EAAS versus transmissivities of the background and target channels, $\kappa_B$ and $\kappa_T$, with different $P_{C,m,LB}$ fixed. Source mean photon number $N_S = 0.1$. The white dashed line corresponds to the value of the classical benchmark $P_{C,m,LB}$, therefore indicating the boundary between regions with/without quantum advantage. (a-b) We consider absorption detection case ($m = 1$), with $M$ such that $P_{C,1,LB} = 0.1, 0.01$. (c-d) We consider single-peak positioning with $m = 100$ and $M$ such that $P_{C,100,LB} = 0.1, 0.01$. Monte Carlo simulation sample size $10^5$.
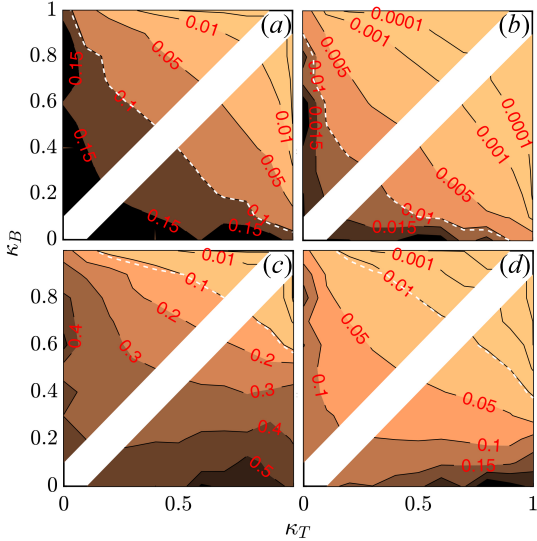


Figure S12. Error probability of EAAS versus transmissivities of the background and target channels, $\kappa_B$ and $\kappa_T$, with different $P_{C,m,LB}$ fixed. Source mean photon number $N_S = 1$. The layout is the same as Fig. S11. Monte Carlo simulation sample size $10^6$.

$\mathrm{argmin}_G \left\langle a_S^\dagger(G) a_S(G) \right\rangle$ different from $G_0$ in Eq. (S22).

We numerically compare the performances for a wide range of the gain, including $G_0$ and $G_{min}$, setting $M$ such that the classical optimum $P_{C,1,LB} = 0.1$, we see the
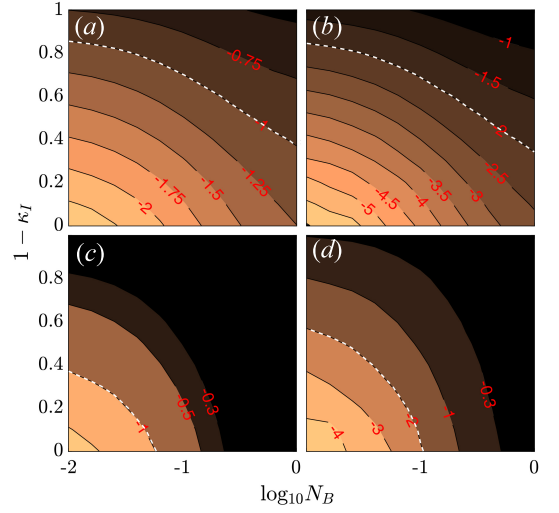


Figure S13. Logarithmic error rate $\log_{10}(P_{E,m})$ of EAAS with respect to the thermal noise $N_B$ and the idler loss $1-\kappa_I$ with different $P_{C,m,LB}$ fixed. Parameters are: $N_S = 0.1$, $\kappa_T = 0.75$, and $\kappa_B = 0.95$. The white dashed line corresponds to the value of the classical benchmark $P_{C,m,LB}$, therefore indicating the boundary between regions with/without quantum advantage. (a-b) We consider absorption detection case ($m = 1$), with $M$ such that $P_{C,1,LB} = 0.1, 0.01$. (c-d) We consider single-peak positioning with $m = 100$ and $M$ such that $P_{C,100,LB} = 0.1, 0.01$. Monte Carlo simulation sample size $10^5$.
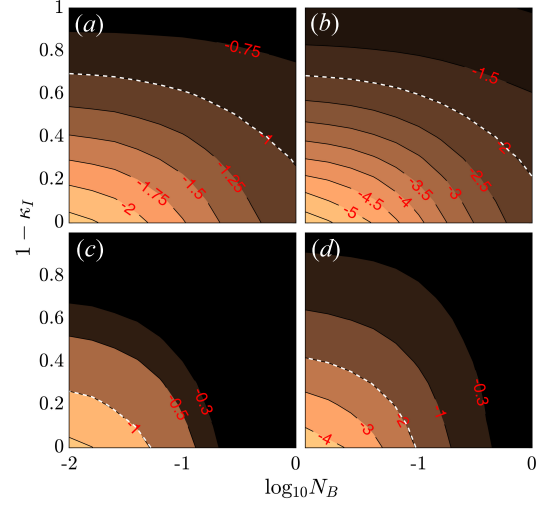


Figure S14. Logarithmic error rate $\log_{10}(P_{E,m})$ of EAAS with respect to the thermal noise $N_B$ and the idler loss $1-\kappa_I$ with different $P_{C,m,LB}$ fixed. Parameters are: $N_S = 1$, $\kappa_T = 0.75$, and $\kappa_B = 0.95$. The layout is the same as Fig. S13. Monte Carlo simulation sample size $10^6$.

decay rate of quantum advantage depends on the gain $G$ shown in Fig. S9. Contrary to the intuition, optimum sticks around $G_0$ in most of cases, as the performance saturates beyond it (in the presented cases $G_{min} < G_0$). In the $N_B = 0$ case (Fig. S9 (e)(f)), when $\kappa_I < 1$, the optimum can be between $G_{min}$ and $G_0$, however, the

optimum error probability is only slightly smaller then that of $G_0$.

## C. Saturating the QCB for absorption detection case

We further explore the saturation of QCB in the absorption detection case. Since QCB is only asymptotically tight, we compare the error exponents as

$$R_E = -\ln(P_{E,m})/M \qquad (S94)$$

and that of the QCB

$$R_{QCB} = -\ln\left(\tilde{Q}_s^M/2\right)/M \qquad (S95)$$

We also compare the classical error exponent

$$R_C = -\ln(P_{C,1,LB})/M. \qquad (S96)$$

In Fig. S10(a)(b), we plot the error exponent ratio $R_E/R_{QCB}$. To guarantee that we are considering the asymptotic region, we choose the mode number $M$ such that the QCB is around $10^{-5}$ in all data points. We see that the error exponent ratio $r = R_E/R_{QCB} \sim 1$ in all parameter region of $\kappa_B, \kappa_T$, verifying the optimality of the EA receiver in absence of noise ($N_B = 0, \kappa_I = 1$). Moreover, in Fig. S10(c)(d) we plot the error exponent ratio $R_E/R_C$ in the same parameter region, where we see advantage in almost all the parameter region of $\kappa_T, \kappa_B$. Furthermore, the entangled error exponent can have a factor of $\sim 8$ larger than the classical error exponent, showing a great advantage. Note that the expected monotonicity with respect to transmissivities breaks at few points, e.g. $(\kappa_T, \kappa_B) = (0.6, 0)$. This is due to the discreteness of $M$ when $M$ is small [even $\sim 1$ at (0.6,0)], which results in a relative sharp change. Indeed the classical lower bounds $P_{C,1,LB}$, expected to be fixed by the proper choice on $M$, fluctuates around 0.1 in these non-monotonic areas.

## VIII. CHARACTERIZING THE QUANTUM ADVANTAGE IN ABSORPTION DETECTION AND PEAK POSITIONING

Here we present more results related to Fig. 4 of the main paper.

In a practical scenario, we are interested in how much EAAS can enhance the performance, when classical schemes fail to perform well. To showcase the advantage, we fix the classical lower bound to be 0.1 or 0.01 and calculate the error probability achievable with EAAS.

In Fig. S11, we plot the error probability $P_{E,m}$ for absorption detection as a function of the transmissivities $\kappa_B$ and $\kappa_T$, while comparing it with the classical lower bounds in Eqs. (3) of the main paper. The white dashed lines divide the parameter space with/without quantum advantage. From the figure, we can see that the advantage is remarkable (several orders of magnitude for $\kappa_B, \kappa_T \simeq 1$) and also survives for a large range of parameters. In practice, when $\kappa_B \simeq 1$, we find a quantum advantage for all values of $\kappa_T$ for the problem of absorption detection.

In Fig. S12, we plot the error probability $P_{E,m}$ for 100-peak positioning as a function of the transmissivities $\kappa_B$ and $\kappa_T$, while comparing it with the classical lower bounds in Eq. (4) of the main paper. From the figures, we can see that the advantage is remarkable (several orders of magnitude for $\kappa_B, \kappa_T \simeq 1$) and also survives for a relatively large range of parameters. In practice, when $\kappa_B \simeq 1$, we find a quantum advantage for $\kappa_T \gtrsim 0.2$ for 100-peak positioning. The region permitting quantum advantage is slightly smaller than the absorption detection case, which we think is due to the fact that the classical lower bounds in Eq. (4) of the main paper is looser than Eq. (3) of the main paper.

In the analysis above, we have assumed noiseless ($N_B = 0$) and lossless ($\kappa_I = 1$) idler storage. In experimental practice, the presence of such noise and loss is inevitable. Therefore we also study how the quantum advantage varies with idler loss $\sqrt{1 - \kappa_I}$ and noise $N_B$ in Figs. S13 and S14 for $N_S = 0.1$ and $N_S = 1$. Again we fix the classical lower bound and plot the error probability of EAAS for $\kappa_T = 0.75$ and $\kappa_B = 0.95$. Regions on the left-down side of the dashed lines show quantum advantage. We see that the advantage is robust against idler loss and channel noise.

## IX. DETAILS ON GENERAL SPECTRUM RECOGNITION

Here we give the transmissivities data that are used in our calculations of wine-tasting and drug-testing.

For the wine tasting, there are $H = 3$ molecules—methanol ($h = 1$), ethanol ($h = 2$) and ethanal ($h = 3$). To discretize the spectra so that we can perform numerical simulation, samples are tested at four frequency slots 500, 1050, 1400, 1800 cm$^{-1}$ (wavelength 20, 9.5, 7.1, 5.6$\mu m$), with the transmissivities averaged in the span of $\pm 100$ cm$^{-1}$. The discrete spectrum can be described by $\kappa_\ell^{(h)} = K_{\ell h}$, where the overall data can be represented as an $m \times H$ matrix

$$K = \begin{pmatrix} 0.9460 & 0.9749 & 0.7853 \\ 0.5659 & 0.6218 & 0.6846 \\ 0.7503 & 0.7622 & 0.4683 \\ 0.9737 & 0.9891 & 0.4165 \end{pmatrix}. \qquad (S97)$$

Here each column gives the transmissivities of $\boldsymbol{\kappa}^{(h)}$ for a fixed hypothesis $h = 1, 2, 3$. The four rows in each column corresponds to the $m = 4$ spectrum slots.

For the drug-testing case, the molecules are phenyl salicylate ($h = 1$), methyl salicylate ($h = 2$), and benzoic

(a) Error rate in base 10 logarithmic scale.

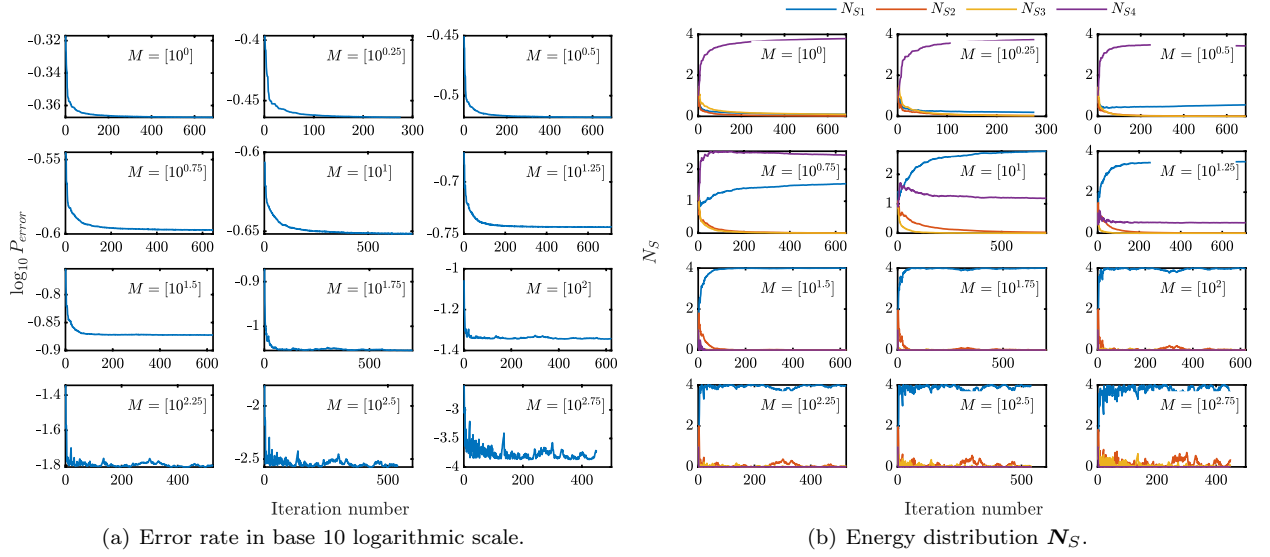(b) Energy distribution $\boldsymbol{N}_S$.

Figure S15. Logarithmic scaled learning curve of the error rate optimization over the source energy distribution in the noiseless wine tasting scenario. Average mean photon number per mode $\sum_{\ell=1}^{4} N_{S\ell}/4 = 1$, noise $N_B = 0$. We find the energy concentrates on the first frequency slot when $M$ is small, and the fourth frequency slot when $M$ goes sufficiently large. It verifies the intuition that suggests the energy be concentrated on the slot with the largest error exponent which dominates the decay of error rate.



(a) Error rate in base 10 logarithmic scale.

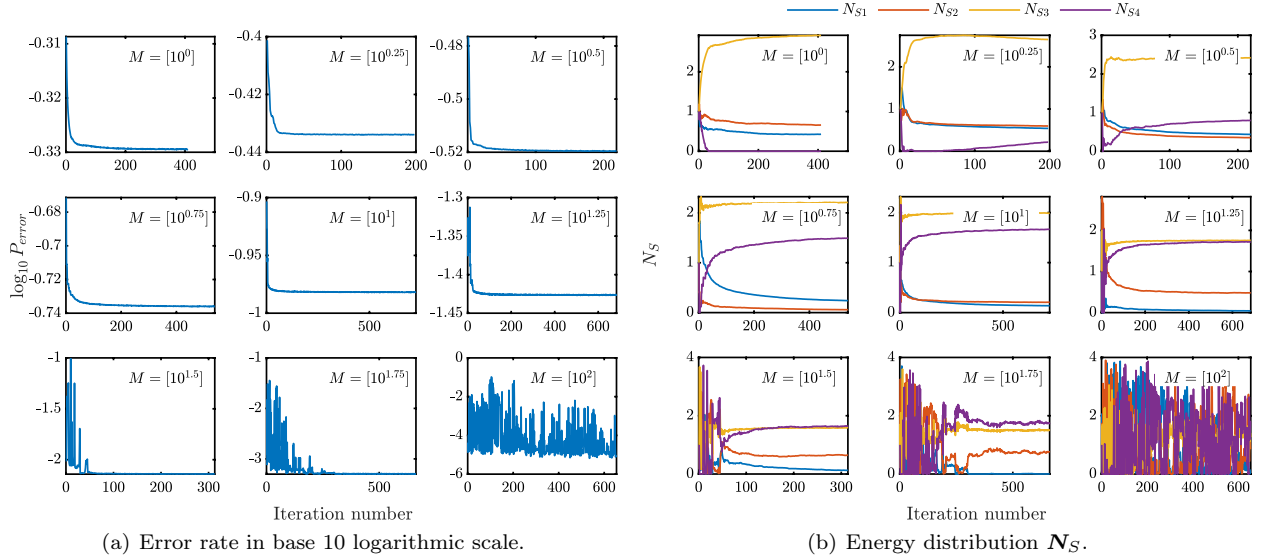(b) Energy distribution $\boldsymbol{N}_S$.

Figure S16. Learning curve of the error rate optimization over the source energy distribution in the noiseless drug testing scenario. Average mean photon number per mode $\sum_{\ell=1}^{4} N_{S\ell}/4 = 1$, noise $N_B = 0$. The large fluctuation in the last two subfigures of $M = [10^{1.75}]$ and $M = [10^{2}]$ is due to the Monte Carlo simulation error. The concentration of energy on a single frequency slot predicted by the classical bound is absent here, rather the optimum energy distribution is non-trivial on all slots.

acid ($h = 3$). Similarly, samples are tested at four frequency slots 500, 100, 1500, 2000 cm$^{-1}$ (wavelength 20, 10, 6.67, 5$\mu m$), with the transmissivities averaged in the span of $\pm 100$ cm$^{-1}$. We have $\kappa_{\ell}^{(h)} = K_{\ell h}$, with

$$K = \begin{pmatrix} 0.9613 & 0.9002 & 0.8093 \\ 0.9215 & 0.8749 & 0.7427 \\ 0.8360 & 0.4002 & 0.7556 \\ 0.9867 & 0.8749 & 0.8522 \end{pmatrix}. \qquad (S98)$$

We evaluate the performance of our receiver through numerical simulations, with optimization over energy and gain. We generate the measurement results through Eq. (S27), and perform maximum likelihood decision. And the error rates are obtained by calculating the frequencies of error happening.

Noticing that the general patterns in Eq. (S97) and (S98) break the symmetry between different (frequency) slots, symmetric strategies with uniform pa-

(a) Error rate.

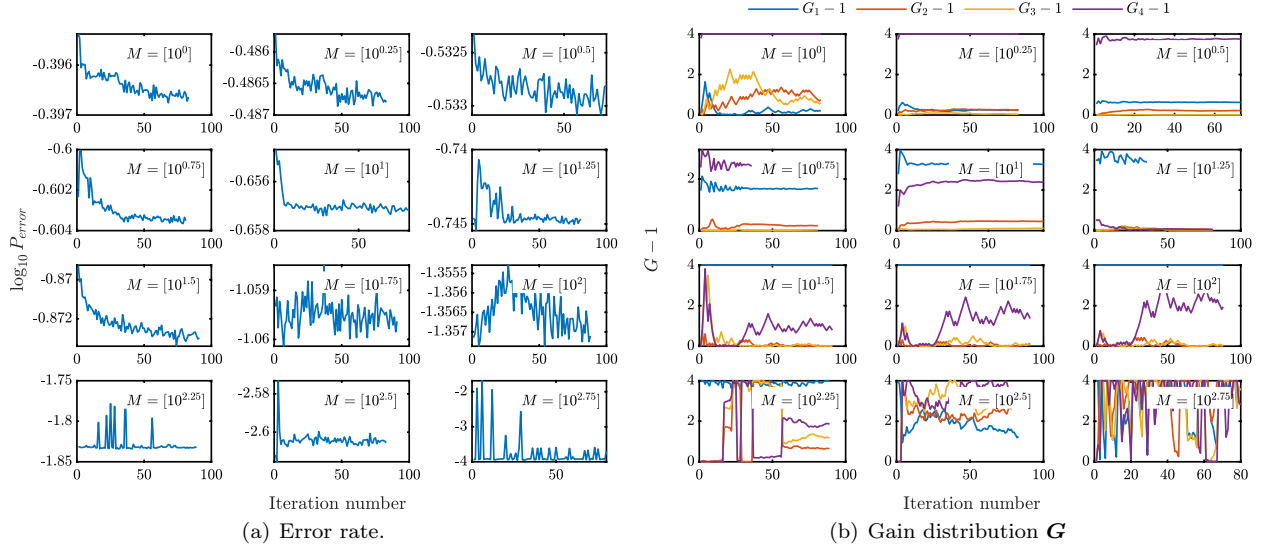(b) Gain distribution $\boldsymbol{G}$

Figure S17. Learning curve of the error rate optimization over the receiver gain in the noiseless wine tasting scenario. Average mean photon number per mode $\sum_{\ell=1}^{4} N_{S\ell}/4 = 1$, noise $N_B = 0$. The gains are restricted by $\boldsymbol{G} - 1 \leq 4$.



(a) Error rate.

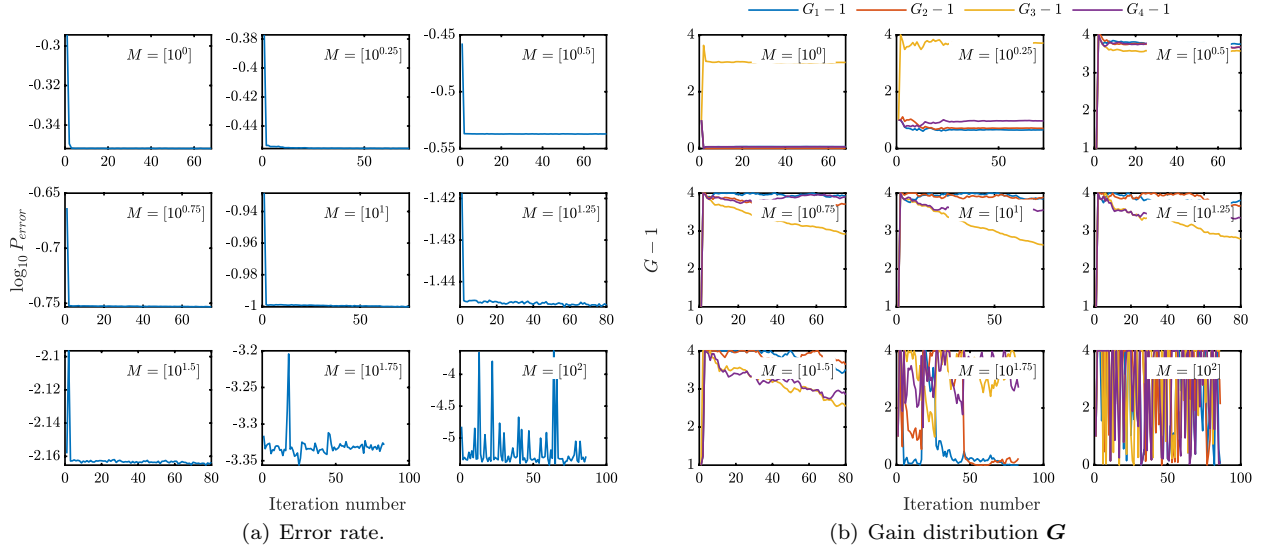(b) Gain distribution $\boldsymbol{G}$

Figure S18. Learning curve of the receiver gain optimizing the receiver gain in the noiseless drug testing scenario. Average mean photon number per mode $\sum_{\ell=1}^{4} N_{S\ell}/4 = 1$, noise $N_B = 0$. The gains are restricted by $\boldsymbol{G} - 1 \leq 4$.

rameters over all slots are unlikely to be the optimum; therefore we expect optimizing the parameters, including the source mean photon numbers $\boldsymbol{N}_S = [N_{S1}, N_{S2}, N_{S3}, N_{S4}]^T$ and the receiver gains $\boldsymbol{G} = [G_1, G_2, G_3, G_4]^T$ over the four frequency slots, to achieve further quantum advantages. Considering the large dimension of parameters in the future application with more frequency slots, we apply the constrained simultaneous perturbation stochastic approximation (SPSA) algorithm [S10, S11] to accelerate the optimization. As a variation of stochastic gradient descent methods, SPSA does not guarantee to always find the global optima. First, we will perform optimization over the energy dis-

tribution, and then we will further optimize over the gain. We find that in general energy optimization leads to improved quantum advantages, gain optimizations offers slight improvement in the noiseless case and much better improvement in the noisy case.

### A. Energy distribution of the source

In the noiseless scenario, Figs. S15 and S16 show the learning curve of the optimization over the source energy distribution, with the average mean photon number per mode fixed to $\sum_{\ell=1}^{4} N_{S\ell}/4 = 1$. From the subplots (b) in

(a) Wine tasting.

(b) Drug testing.
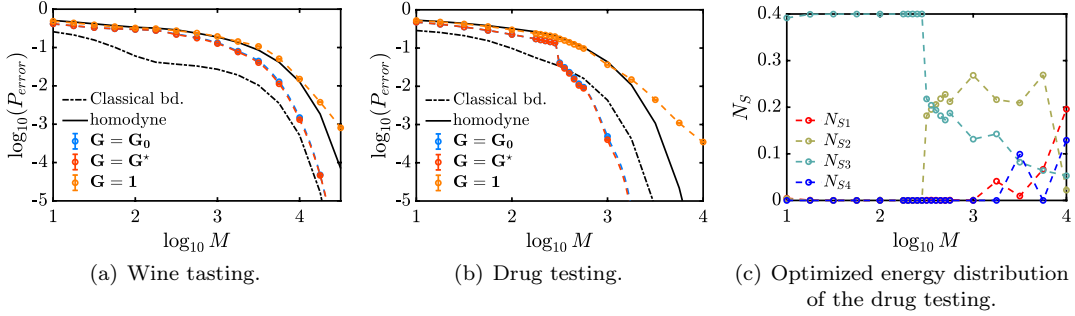
(c) Optimized energy distribution of the drug testing.

Figure S19. Error rate of the receiver in the noisy scenario with thermal photon number $N_B = 0.1$. The energy distribution of the source is optimized under the constraint on the average mean photon number per mode $\sum_{\ell=1}^{4} N_{S\ell}/4 = 1$. We compare energy-optimized EAAS under three gain settings: simply adopting Eq. (S22) in each frequency slot to obtain $\mathbf{G} = \mathbf{G}_0 = 1 + \mathbf{N}'_{S0}$ (dashed blue), numerically optimized gain $\mathbf{G} = \mathbf{G}^\star$ (dashed red), OPA-absent case $\mathbf{G} = 1$ (dashed orange), along with the classical lower bound Eq. (S42) (dot-dashed black) and homodyne detection (solid black). In the numeric optimization of $\mathbf{G}^\star$ the gains are restricted by $\mathbf{G} - 1 \leq 4$. Subplot (c) shows the optimized energy distribution which is prescribed in the optimization in subplot (b), where the leap at $M \sim 10^{2.5}$ accounts for the same leap in the performance in (b).

both figures, we see that the optimal energy distributions are far from uniform as expected. We achieve an appreciable one-order-of-magnitude improvement when $M$ is large, which guarantees the advantage over the classical lower bound in Fig. 5 of the main text.

As to the energy distribution, the classical lower bound in Eq. (2) of the main paper gives a pretty good intuition for the optimization, although the lower bound itself is likely to be loose. Specifically, the classical bound consists of a few exponential terms with respect to the difference between the transmitted energys $(\sqrt{\kappa_i} - \sqrt{\kappa_j})^2 M N_S$ through two channels $i$, $j$, and the term with the largest error exponent dominates the performance asymptotically when $M$ is sufficiently large. Hence a fair guess of the optimal strategy is to allocate most of the energy to the slot associated with the dominant term. However, in Fig. S15(b), we see that the energy tends to concentrate on the first slot when M increases, which is different from the classical optimum (the second slot). Note that the transmissivities of the first slot are relatively higher and closer to unity, this divergence from the classical prediction is likely due to the gap between the quantum scenario here and the conventional classical region, viz. that the entanglement-assisted advantage over the classical optimum surges as the transmissivity increases close to unity. This can also be seen in Fig. S16(b), the energy was scattered on the four modes without any dominant slot, instead of being concentrated on the second slot as predicted by the classical optimum. The illustrated differences from the intuition based on classical schemes demonstrate the novelty of the entanglement-assisted scenario again.

### B. Gain distribution of the receiver

With the same noiseless environment setting, we optimize the gain distribution on top of the optimal energy distribution achieved above. Figs. S17 and S18 give the learning curve of the optimization on the receiver gain distribution. The gains are restricted per slot by $\mathbf{G} - 1 \leq 4N_S = 4$, as the gain of OPA is limited by the nonlinearity of the crystal in practice. Compared with the optimization on energy distribution, we see a slight improvement by further optimizing the gains in this scenario.

The optimization of the gain distribution is more challenging than that of the energy distribution. As shown in Fig. S9 (e-h) for the binary hypothesis-testing case, the gradient of error rate with respect to the gain is close to zero almost everywhere. A large fluctuation is present in the convergence process due to the small gradient-to-noise ratio.

By contrast, gain optimization is likely to achieve a significant improvement in the noisy scenario. To demonstrate the influence of noise, we include a uniform thermal background with mean photon number $N_B = 0.1$ into the channel. Following the same procedure, we first optimize the energy distribution for the noisy case. Here the presence of noise leads to a sharp increase in the computational complexity of the photon statistics. Consequentially, we compromise the average mean photon number per mode down to $\sum_{\ell=1}^{4} N_{S\ell}/4 = 0.1$. Meanwhile, we correspondingly scale the range of copy number $M$ of interest by a factor 10. Fig. S19 illustrates the gap between the different gain settings. In both the wine-tasting and drug-testing cases, we see a substantial advantage of the gain optimization over the OPA-absent case as expected. Interestingly, a trivial choice of $\mathbf{G} = \mathbf{G}_0 = 1 + \mathbf{N}'_{S0}$ according to Eq.(S22) is sufficient to yield nearly the same improvement. This provides a neat rule-of-thumb estimation of the optimum gain in the practical implementation.

[S1] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, Gaussian quantum information, Rev. Mod. Phys. **84**, 621 (2012).

[S2] S. Pirandola and S. Lloyd, Computable bounds for the discrimination of gaussian states, Phys. Rev. A **78**, 012331 (2008).

[S3] S. Pirandola, C. Lupo, V. Giovannetti, S. Mancini, and S. L. Braunstein, Quantum reading capacity, New J. Phys. **13**, 113012 (2011).

[S4] S. Pirandola, Quantum reading of a classical digital memory, Phys. Rev. Lett. **106**, 090504 (2011).

[S5] Q. Zhuang and S. Pirandola, Entanglement-enhanced testing of multiple quantum hypotheses, Commun. Phys. **3**, 1 (2020).

[S6] A. Montanaro, A lower bound on the probability of error in quantum state discrimination, in *2008 IEEE Infor-mation Theory Workshop* (IEEE, 2008) pp. 378–380.

[S7] S. J. Dolinar, *Processing and Transmission of Information*, Technical Report (Research Laboratory of Electronics (RLE) at the Massachusetts Institute of Technology (MIT), 1973).

[S8] S. Dolinar Jr, A near-optimum receiver structure for the detection of m-ary optical ppm signals, The Telecommunications and Data Acquisition Progress Report **42** (1982).

[S9] C. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, 1976).

[S10] J.C. Spall, An overview of the simultaneous perturbation method for efficient optimization, Johns Hopkins APL Technical Digest **19**, 482 (1998)

[S11] P. Sadegh, Constrained optimization via stochastic approximation with a simultaneous perturbation gradient approximation, IFAC Proceedings **30** 281 (1997).