

Local sampling of the Wigner function at telecom wavelength with loss-tolerant detection of photon statistics: Supplemental Material

G. Harder,¹ Ch. Silberhorn,^{1,2} J. Rehacek,³ Z. Hradil,³ L. Motka,³ B. Stoklasa,³ and L. L. Sánchez-Soto^{4,2}

¹*Department of Physics, University of Paderborn, Warburger Straße 100, 33098 Paderborn, Germany*

²*Max-Planck-Institut für die Physik des Lichts, Günther-Scharowsky-Straße 1, Bau 24, 91058 Erlangen, Germany*

³*Department of Optics, Palacký University, 17. listopadu 12, 77146 Olomouc, Czech Republic*

⁴*Departamento de Óptica, Facultad de Física, Universidad Complutense, 28040 Madrid, Spain*

I. FITTING DATA PATTERN

For completeness, we give here the essential details of the data pattern tomography [1, 2]. The PDC state is characterized by the two-mode photon-number distribution P_{mn} , where the first (second) index refers to the signal (idler) mode. We denote by $p_{\alpha\beta}$ the probability of simultaneous signal (α) and idler (β) detection. According to the experimental description, detections are represented by 8-digit binary numbers, where 0/1 values mean click/no click in the corresponding time bin. All in all, this gives $2^8 = 256$ distinct single-mode events and $2^{16} = 65536$ two-mode events to reckon with. This is below the limitations of the method, however 8 bins per mode are enough to reveal the subtle signal features, while preventing signal-to-noise degradation. A further splitting of the modes would increase the dark counts, decreasing the accuracy of the reconstruction. Besides, an effective reduction of the number of bins can always be done in post-processing by combining the existing detection channels.

In a linear model of the TMD detection, we have

$$p_{\alpha\beta} = \sum_{m=0}^{D-1} \sum_{n=0}^{D-1} C_{\alpha\beta, mn} P_{mn}, \quad (1)$$

where D is a cutoff dimension. The matrix C contains full information about the TMD, including losses, detector efficiencies, and afterpulsing effects. In particular, afterpulses prevent the factorization of C into signal and idler components.

We also assess single-mode and heralded events; the former (latter) are nothing but marginal (conditional) probabilities of $p_{\alpha\beta}$. For these single-mode events, we look at the total number of clicks (either in the signal or the idler), without paying attention to the particular ordering of time bins.

In a real experiment, we get the relative frequencies $f_{\alpha\beta}$ after N random samples drawn from the multinomial distribution parametrized by $p_{\alpha\beta}$. From the measured data $f_{\alpha\beta}$ we have to determine P_{mn} . The standard detector tomography proceeds in two steps: first, a detector estimation, where the measurement matrix $C_{\alpha\beta, mn} \geq 0$ is inferred from a set of calibration states. Afterwards, $P_{mn} \geq 0$ is reconstructed from the previously obtained matrix C . Unfortunately, this detector estimation is exceedingly costly, scaling as D^4 , which makes the method impractical, even for moderate values of this cutoff D .

The alternative data-pattern approach ignores the details of

the TMD and casts P_{mn} as a mixture

$$P_{mn} = \sum_{\xi=1}^M c_{\xi} P_{mn}^{(\xi)} = \sum_{\xi=1}^M c_{\xi} P_m^{(\xi)} P_n^{(\xi)} \quad (2)$$

of M linearly independent (generally, nonorthogonal) two-mode coherent probes $\{P_{mn}^{(\xi)}\}$, with positive and negative weights $\{c_{\xi}\}$. We can think of Eq. (2) as a kind of discrete P -representation and can be straightforwardly extended to any number of modes. As long as there is a match between the number of linearly independent probes and the number of free parameters, the decomposition is unique. For overcomplete set of probes, the uniqueness is not guaranteed, but all the feasible vectors c_{ξ} give the same solution.

We stress that the field of view of the setup is uniquely determined by the known set of measured probes, rather than some *ad hoc* truncation of the search space. In our experiment, the photon-number distributions of coherent states, $P_n^{(\xi)} = |\langle n | \alpha(\xi) \rangle|^2$, where $|n\rangle$ is a Fock state, serve as the probes.

The responses $f_{\alpha\beta}^{(\xi)}$ of the TMD to these coherent probes are called patterns. Then, by linearity, the data (i.e., the TMD response $f_{\alpha\beta}$ to an unknown state P_{mn}) can be modeled as

$$f_{\alpha\beta} \simeq \sum_{\xi=1}^M c_{\xi} f_{\alpha\beta}^{(\xi)}. \quad (3)$$

Once the patterns and data are acquired, the coefficients c_{ξ} can be inferred from Eq. (3) and the state reconstructed according to (2). To this end, a suitable convex measure of the distance between the left- and right-hand side of Eq. (3) has to be minimized, subject to the constraints $P_{mn} \geq 0$ and $\sum_{mn} P_{mn} = 1$: this is a quadratic program, for which stable algorithms are available.

The advantage of the inversion (2) over the standard post-processing is evident. Moreover, at difference of the standard detector tomography, the number of parameters M is independent of the cutoff dimension D and, therefore, incomplete sets of probes $M \ll D^2$ can be measured and processed much in the same way as complete ones.

II. PARTIAL TOMOGRAPHY

It happens quite often that the full distribution P_{mn} is not required for a specific purpose. For example, only a few ele-

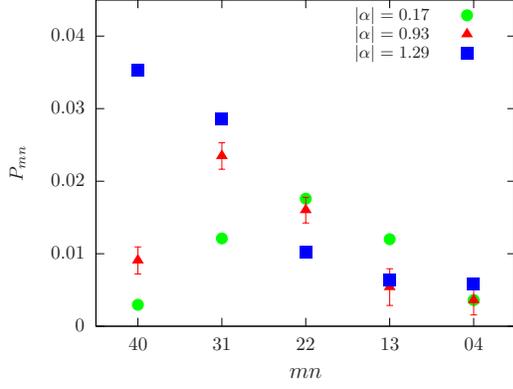


FIG. 1. Partial tomography of the anti-diagonal terms of P_{mn} for different amounts of displacement of a PDC input state with $r \approx 0.6$. In all cases, 10 two-mode coherent probes are used. Error bars are only shown for one displacement to avoid symbol clustering.

ments of P_{mn} or a linear function of P_{mn} , such as parity, might be enough. This is called partial tomography and finds applications in experiments with complex, highly-dimensional systems, where a full tomography is impractical or even impossible. A nice feature of the pattern approach is that both full and partial tomography are done in much the same way.

As a case study, we restrict P_{mn} up to four photons in each mode ($m, n = 0, \dots, 4$) and we address exclusively the anti-diagonal elements $P_{40}, P_{31}, P_{22}, P_{13},$ and P_{04} . Without displacement, the anti-diagonal should be symmetric around P_{22} ; with displacement, it is expected to become biased towards P_{40} , for displacement adds intensity in the signal mode and signal-idler correlations tend to vanish.

This is confirmed in Fig. 1, where the reconstructed anti-diagonal is shown for different displacements. Notice that a full reconstruction of a P_{mn} in this case requires using 25 linearly independent probes of size 25 each, whereas only 10 probes of size 5 each were used here.

The method has also limitations. As an example, we choose the extreme case where only one particular moment of the state distribution P_n , say parity S , will be inferred. The partial tomography

$$S = \sum_{\xi} c_{\xi} S^{(\xi)}, \quad (4)$$

where $S^{(\xi)} = \sum_n (-1)^n P_n^{(\xi)}$, should be compared with the parity resulting from the full state reconstruction,

$$S = \sum_n (-1)^n P_n, \quad (5)$$

where $P_n = \sum_{\xi} c_{\xi} P_n^{(\xi)} \geq 0$. The partial tomography does not involve physical constraints and therefore the data fitting becomes an unconstrained optimization problem. The full tomography, on the other hand, involves positivity check, making the fitting a constrained optimization. This effectively performs a regularization of the problem: unphysical estimates

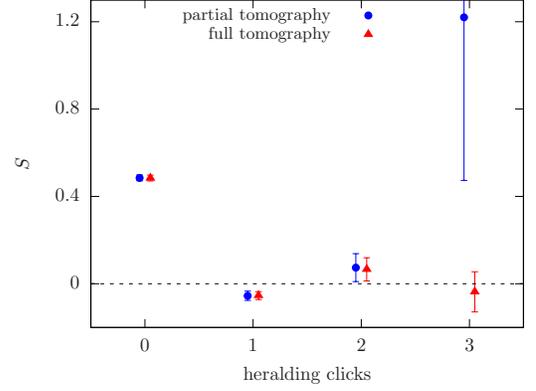


FIG. 2. Partial tomography of heralded idler parity (circles) compared to parity from full state reconstruction (squares). Heralding by up to fourfold coincidences in the signal beam is shown. PDC input state with $r = 0.93$ and $|\alpha| = 0$ was used.

are rejected. Hence, the latter technique is expected to be more stable and perform better, especially with noisy data.

To illustrate this point, an experiment has been performed, where idler-state parity was estimated by both partial and full tomography. A moderately strong PDC state was prepared without the reference beam stopped down and the idler state was heralded by different signal detection events: no click, single click, double click, and triple click. Given the intensity in the signal, multiple coincidence events are highly unlikely and heralding results in a strong reduction of idler counts. As a consequence, statistical noise grows and this effect is combined with other sources of systematic errors, such as afterpulses. Heralding by no signal click, which is the most common strategy, should not pose serious problems. On the other hand, we expect the tomography of idler heralded by triple clicks to be rather involved.

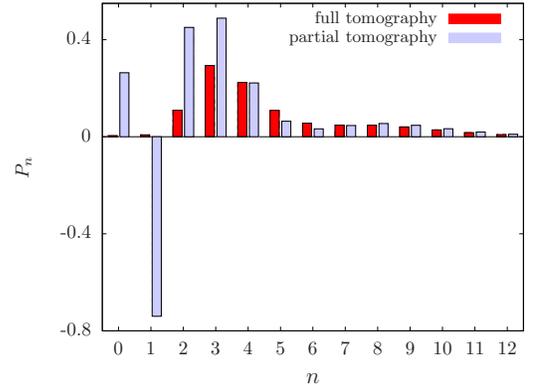


FIG. 3. Experimentally reconstructed idler photon-number distributions heralded by a signal fourfold coincidence event. Results of constrained (red bars) and unconstrained (blue bars) pattern tomography are shown. The corresponding parities are shown in Fig. 2 as the rightmost pair of symbols.

Figure 2 shows the reconstructed idler parities. Up to double-click heralding, the results are as expected: the sign of the reconstructed parity alternates and both techniques have similar accuracy. The situation changes with heralding by triple detection. In the case of full tomography, the result is still meaningful, although we observe larger reconstruction errors and the sign of the heralded Wigner function cannot be longer resolved. However, partial tomography fails completely, reporting a nonsensically large parity and absurdly large error bars.

The origin of this problem becomes apparent if we analyze the corresponding inferred idler distributions, which are shown in Fig. 3. For parity measurements, partial tomography

is equivalent to full tomography with the positivity constraint removed. Let \tilde{P}_n be the underlying idler distribution obtained by (partial) unconstrained pattern tomography. We see that triple-click heralded \tilde{P}_n is highly unphysical, with strongly negative component \tilde{P}_1 (blue bars), rendering the calculated parity useless. Taking physical constraints into account, a more reasonable reconstruction is achieved (red bars).

These results show that partial tomography can be helpful in saving experimental and computational resources. However, one has to be careful when applied to noisy data. In this case, extracting the sought after information from the full state reconstruction seems to be a more robust approach.

-
- [1] J. Řeháček, D. Mogilevtsev, and Z. Hradil, “Operational tomography: Fitting of data patterns,” *Phys. Rev. Lett.* **105**, 01040 (2010).
 [2] D. Mogilevtsev, A. Ignatenko, A. Maloshtan, B. Stoklasa, J. Re-

hacek, and Z. Hradil, “Data pattern tomography: reconstruction with an unknown apparatus,” *New J. Phys.* **15**, 025038 (2013).