Adaptive loop-based subtraction of a single photon from a traveling beam of light

Petr Marek, Jan Provazník, Radim Filip
Department of Optics, Palacký University, 17. listopadu 1192/12, 77146 Olomouc, Czech Republic

Manipulating light by adding and subtracting individual photons is a powerful approach with a principal drawback: the operations are fundamentally probabilistic and the probability is often small. This limits not only the fundamental scalability but also the number of operations that can be applied in realistic experimental settings. We propose and analyze an adaptive technique which can significantly increase the probability of success while preserving the quality of the photon subtraction. We show the improvement both in single mode preparation and manipulation of non-Gaussian states with negative Wigner functions and in two-mode entanglement distillation protocol with Gaussian states of light.

I. INTRODUCTION

During the last decades, techniques manipulating individual quanta were used for testing many interesting and fundamental aspects of quantum physics [1–3], in particular the applications in quantum communication where they allow entanglement distillation [4, 5] or noiseless amplification of traveling beams of light [6–10]. Recently they have also expanded into the developing field of quantum thermodynamics [11, 12]. Overall, many quantum information protocols were and continuously are being tested in hybrid quantum optical experiments involving both discrete photons and continuous variable fields of light [13–15].

The best approach towards manipulating the individual photons at optical frequencies has been, so far, achieved by employing the measurement-induced-operations paradigm. Projective measurements on states of atoms, which passed through cavity containing microwave field, were used to generate highly non-classical and non-Gaussian states of the field [16, 17]. Similar results were obtained in optical domain with detectors capable of distinguishing individual photons [18–22]. The detectors themselves can operate either on separate spatial modes, but they can also filter out particular frequency modes from the full spectrum of the light field [23–25].

All of these implementations have one thing in common: the output states are always obtained by conditioning on one particular measurement result and, as a consequence, the operations are probabilistic in nature. This is a drawback both from the theoretical point of view, as it limits the protocols that can be meaningfully implemented, as well as from the experimental point of view. It is therefore quite important to work around this issue and to design operations which are deterministic or, if that cannot be done, which have as large probability of success as possible. It is worth mentioning here that the only operations that can be implemented deterministically are those emulating quantum dynamics described by positive trace preserving maps, such as [26, 27]. Some nonlinear operations, as is the case of the already mentioned photon subtraction and its compatriots, can never be performed deterministically. However, for the sake of efficient technologies and also the immediate experimental tests of noiseless amplification [6–9, 28] or entanglement distillation [4, 5], it is important and worthwhile to improve the success rate of these operations beyond what is possible today.

In this paper we take a closer look at the photon subtraction operation and propose a way of increasing its probability of success while preserving its fundamental properties. The operation is probabilistic by definition but for sufficiently intense quantum state of light we can achieve success probabilities close to one, similarly to the methods employing atoms for detection [29, 30]. We are, however, not limited to any class of states in particular. Our method employs standard avalanche photo diodes and is based on recycling the used quantum state after an unsuccessful photon subtraction and repeating the procedure until it either succeeds or until a pre-determined number of steps is achieved. This approach is adaptive; the results of the measurements shape the dynamics of the operation. In [31] a similar method was briefly proposed but it was not analyzed in detail. In this paper we thoroughly discuss the cost for increasing the success rate, which is the decoherence of the state, and show how it depends on the parameters of the operation for both single mode and two-mode scenarios. We observe that the success probability can be significantly increased while the fundamental properties of the transformation are kept intact. The experimental realization of the proposed technique requires a quantum delay element for the traveling light in order to act on the light and to release it when needed. Fortunately, the recent experimental advances show a lot of promise in this area [32, 33].

II. SINGLE PHOTON SUBTRACTION

Subtraction of a single photon from a traveling mode of light is usually done by tapping off part of the light field on a strongly unbalanced beam splitter and subjecting this stray light to a detection by an avalanche photon
FIG. 1: (color online) Schematic representation of the adaptive photon subtraction. Part of the input signal is tapped off at the mostly transmissive beam splitter (BS) and directed towards avalanche photo-diode (APD). Positive detection event results in successfully transformed signal while negative one prompts the optical switches (S) to feed the signal back for the next attempt if a predetermined number of maximal steps was not reached yet.

detector (APD) that can only distinguish between photons and no photons. When the detector registers a photon, the operation is considered successful and the remaining state of light is kept, otherwise it is discarded. Formally, the detection event transforms an arbitrary quantum state $\rho_{0,\text{in}}$ into

$$\hat{\rho}_{0,\text{out}} = \text{Tr}_1[\hat{U}_{01}\hat{\rho}_{0,\text{in}} \otimes |0\rangle_1\langle 0|\hat{U}_{0,1}^\dagger\hat{\Pi}_1^*],$$

(1)

where $\hat{U}_{01} = \exp[\kappa(\hat{a}_0\hat{a}_1^\dagger - \hat{a}_0^\dagger\hat{a}_1)]$ is the unitary operation of the beam splitter between the modes labeled by 0 and 1, with $\hat{a}_0$ and $\hat{a}_1$ being the corresponding annihilation operators for the two modes with the standard commutation relations $[\hat{a}_k, \hat{a}_j^\dagger] = \delta_{kj}$. The two possible outcomes of the measurement by APD can then be represented by its Positive Operator Value Measure (POVM) elements, which are $\hat{\Pi}_1^\circ = |0\rangle_1\langle 0|$ for the ideal detector and $\hat{\Pi}_1^* = 1 - |0\rangle_1\langle 0|$ for a realistic detector with nonunit quantum efficiency $\eta$. We neglect the dark counts as they are negligibly small for high quality superconducting detectors [34]. When the beam splitter is strongly unbalanced and the amplitude transmissivity $t = \cos\kappa$ is close to one, the unitary operator can be replaced by its Taylor expansion up to the first order and the output state can be approximated to

$$\hat{\rho}_{0,\text{out}} \approx \kappa^2 \hat{a}_0^\dagger\hat{\rho}_{0,\text{in}}\hat{a}_0^\dagger.$$  

(2)

We can see that in this limit, $\kappa \to 0$, the operation does indeed approach the perfect subtraction of a single photon, but the probability of success, given by the trace of (2), approaches zero as well.

To increase this probability, we suggest modification of the scheme as per Fig. 1. The main difference from the standard photon subtraction is that, when the detector measures no signal, we do not discard the quantum state, but instead loop it back to enter the beam splitter again. We repeat this looping procedure until a photon is detected, after which we release the field immediately. If a photon is not detected within a pre-determined number of loops $N$, we stop the procedure and discard the run. As a consequence the operation as a whole is still probabilistic. The transmission coefficient of the internal beam splitter is kept constant during the whole procedure. The architecture was inspired by the recent experimental proposals [32, 33]. If the procedure succeeds in its $n$-th step we immediately stop the looping and the output state can be expressed as

$$\hat{\rho}_{0,\text{out}}^{(n)} = \text{Tr}_{1,\ldots,n}[\hat{U}_{(n)}\hat{\rho}_{0,\text{in}} \otimes \bigotimes_{k=1}^n |0\rangle_k\langle 0|\hat{U}_{(n)}^\dagger\hat{\Pi}_{(n)}],$$

(3)

where

$$\hat{U}_{(n)} = \bigotimes_{k=1}^n \hat{U}_{0,k}, \quad \hat{\Pi}_{(n)} = \bigotimes_{k=1}^{n-1} \hat{\Pi}_k^\circ \otimes \hat{\Pi}_n^*.$$  

(4)
are the collective operators of the sequences of the beam splitters and the projective measurements. The density matrix \( \hat{\rho}_{0,\text{out}} \) is not normalized; its norm is equal to the probability of obtaining the positive detection event in that particular loop. If we do not care in which step the procedure succeeded, we need to average over the possibilities and the final output state can be expressed as

\[
\hat{\rho}_{0,\text{out}} = \frac{1}{P_S} \sum_{n=1}^{N} \hat{\rho}_{0,\text{out}}^{(n)},
\]

where \( P_S = \text{Tr}\{\sum_{n=1}^{N} \hat{\rho}_{0,\text{out}}^{(n)}\} \) is the overall probability of success. If we assume the individual beam splitter couplings to be weak we can approximate the output state as

\[
\hat{\rho}_{0,\text{out}} \approx \frac{1}{P_S} \hat{\rho}_0 \left[ (1 - t^2) \sum_{k=0}^{N} t^{k\bar{n}_1} \hat{\rho}_{0,\text{in}} t^{k\bar{n}_1} \right] \hat{n}_0^\dagger,
\]

with

\[
P_S \approx (1 - t^2) \sum_{k=0}^{N} \text{Tr} [\hat{\rho}_{0,\text{in}} t^{2k\bar{n}_0} \hat{n}_0],
\]

where \( t = \cos(\kappa) \) is the transmission coefficient of the beam splitters. The probability of success depends on the transmission coefficient \( t \) and on the maximal allowed number of steps \( N \). For each input state \( \rho_{0,\text{in}} \) and each value of \( N \), it is possible to choose the coefficient \( t \) in such way that any success rate \( P_S \) from the interval \([0, P_{S,\text{max}}]\) can be obtained. Here \( P_{S,\text{max}} = 1 - \text{Tr}[\hat{\rho}_{\text{in,0}} \hat{\Pi}] \) is the probability of finding any photons in the state. In the upcoming analysis we will compare quality of the operation directly with the success rate, keeping the implicit dependence on \( t \) hidden. However, the basic mechanism of improvement can be gleaned from (7), where we can see that if we fix the input state and the coefficient \( t \) is sufficiently close to one, the probability of success increases with \( N \) and for small values of \( N \) the increase is almost linear. The cost of this improvement is the decoherence of the output state which is now a mixture a quantum states which underwent different noiseless attenuation processes \([32]\). The output state is therefore a mixture, its decoherence being the cost for increasing the success rate.

### III. ANALYSIS OF THE TRADE-OFFS

We start with analyzing the adaptive loop-based approach for the archetypal example: subtraction of a single photon from a single mode pure squeezed state of light. In the ideal scenario this leads to creation of a squeezed single photon state. A distinctive property of this state is the negativity of its Wigner function, which also decisively marks the state as nonclassical and non-Gaussian \([30, 37]\). One of the oftentimes seen consequences of decoherence is the reduction of this negativity. Its survival under decoherence was even proposed as one of the measures of nonclassicality \([38]\). So let us start by analyzing the behavior of the negativity at the center of the phase space, represented by the value \( W(0,0) \) of the Wigner function, under the particular decoherence incurred by our operation.

Let us first consider the ideal case of a pure squeezed state with 6 dB of squeezing and 6 dB of antisqueezing. In Fig. 2 we can see the value of the Wigner function in relation to the maximal number of allowed steps \( N \) and the probability of success \( P_S \). For each particular state and a specific number of allowable steps \( N \), the probability of success \( P_S \) is in one-to-one correspondence with the transmissivity \( t \) of the tapping mirror. For the case of the ideal detection we can see that for low numbers of allowed steps there is a distinctive trade-off between the achieved negativity and the probability of success. However, with increasing numbers of steps it becomes possible to approach maximal negativity at the maximal possible success rate and the improvement is significant already for ten steps. The maximal probability of success, \( P_{S,\text{max}} = 1 - \text{Tr}[\hat{\rho}_{\text{in,0}} \hat{\Pi}] \), depends solely on the input state and the efficiency of the detectors. For sufficiently intense signals the success rate approaches one. However, unlike \([29, 30]\) who have shown this tendency for coherent states, our method is not limited for any particular class of states. In the case of currently available superconducting detectors with efficiency \( \eta = 0.8 \) the maximal achievable negativity is lower; the reduced efficiency effectively acts as loss on the initial squeezed state and disrupts its parity properties. Consequently, the trade-off between the achievable negativity and the success probability never vanishes as in the ideal case. Nevertheless, for all values of success probability and negativity the improvement by considering multiple steps is clearly visible. The method can be therefore tested with already existing detectors.

Let us comment on an interesting situation which can appear when the operation is applied to a squeezed vacuum state which is not pure. The presence of extra noise leads to larger values of \( W(0,0) \). However, unlike the previous...
FIG. 2: Central negativity of non-Gaussian quantum state prepared from a pure state with 6dB of squeezing relative to the achievable probability of success for (a) ideal and (b) contemporary detectors with quantum efficiency $\eta = 0.8$. The performance of the original single step subtraction protocol (black dotted) is significantly surpassed by taking as little as 10 steps (dashed blue) and which can be further improved by performing 100 steps (solid red). While the improvement is significant for both the ideal and the realistic detection, the losses due to quantum inefficiency make it impossible to attain maximal negativity with maximal probability of success.

FIG. 3: The central negativity of the Wigner function (solid red line) and the achievable probability of success (dashed red line) in relation to the transmission coefficient of the tapping beam splitter for using the (a) ideal and (b) realistic superconducting detectors on an impure squeezed state with 8 dB of squeezing and 10 dB of anti-squeezing. The currently available superconducting detectors have quantum efficiency $\eta = 0.8$. The black vertical lines mark the points in which the Wigner function negativity reaches the local minima even though the success probability is not saturated. The insets show the detail of these points. The effect becomes observable for sufficiently squeezed states after a higher number of steps; the present figure is obtained for $N = 100$ step procedure.

scenario, in which larger negativity was always associated with lower probability of success, in the case of impure squeezed states it is possible to simultaneously increase both the negativity and the probability of success. This behavior, which manifests for sufficiently efficient detectors and sufficiently high number of allowed steps, is illustrated in Fig. 3 on the case of subtraction of photon from a squeezed state with 8 dB of squeezing and 10 dB of anti-squeezing. The negativity of Wigner function is an important quantum feature but it does not provide a complete picture of the decoherence. Specifically, when subtracting a photon from a squeezed state by the multi-step approach, the
FIG. 4: Fidelity in respect to probability of success of the $|+\rangle \rightarrow |-\rangle$ transition facilitated by the subtraction procedure employing (a) ideal and (b) realistic superconducting detectors with detection efficiency $\eta = 0.8$. The superpositions $|\pm\rangle \propto |\alpha\rangle \pm |\alpha\rangle$ are investigated for realistic $\alpha = 3/\sqrt{2}$ in both regimes of detection. Our procedure shows a significant improvement in success probability over the original subtraction protocol (dotted black) with as little as 10 steps (dashed blue). The maximal probability relative to desired fidelity of the transition can be further increased by taking, for example, 100 steps (solid red).

final quantum state is a mixture $|\Psi\rangle$, but under the ideal conditions all the states in the mixture have Wigner function with maximal negativity at the origin. Let us therefore consider a different scenario. Instead of subtracting a photon from a squeezed state we will consider subtracting a photon from a superposition of two coherent states $|+\rangle \propto |\alpha\rangle + |\alpha\rangle$. Such states can be used for quantum information processing protocols $|39, 40\rangle$ and photon subtraction can be advantageously employed for probabilistic realization of quantum logic gates $|41, 42\rangle$. What limits the practical utilization of these states and, at the same time makes them an excellent ‘canary in a coal mine’, is their fragility. When affected by loss or noise they tend to quickly lose coherence and deteriorate from a superposition of coherent states into their mixture, which has fidelity of only 0.5 with the initial superposition. Although the fidelity is not a good quantifier of nonclassical interference aspects of the superposition, here it is a very sensitive measure of even small decoherence. Looking at the fidelity will therefore tell us more about the nature and the severity of the decoherence emerging during the adaptive operation.

The ideal subtraction transforms the state $|+\rangle$ into a different superposition $|-\rangle \propto |\alpha\rangle - |\alpha\rangle$, which is orthogonal to the original one. As a figure of merit we can then consider the fidelity of the subtracted state $\hat{\rho}_{\text{out}}$ with the ideal state $|-\rangle$, which is $F = \langle - | \hat{\rho}_{\text{out}} | - \rangle$. In Fig. 4 we can see the dependence of the fidelity on the probability of success and the number of steps. When the employed detectors are ideal, even the fragile superposed coherent states can be transformed with both fidelity and the success probability achieving the maximal values, when the number of steps is large enough. At the same time, the improvement is significant already for ten steps. In the case of imperfect detectors the decoherence can not be avoided and the maximal fidelity is achievable only in the limit of $P_S \rightarrow 0$. However, already ten steps of the adaptive method can lead to obtaining the same fidelities with significantly higher probabilities of success. It is therefore sufficient for immediate experimental test with high quality $|+\rangle$ superposition states $|43, 44\rangle$.

IV. DISTILLATION OF ENTANGLEMENT

Photon subtraction is a crucial part of the entanglement distillation protocols $|45\rangle$. There it serves the dual role of increasing the entanglement and de-Gaussifying the quantum state $|5, 44\rangle$ and even though it can be in principle substituted by other non-Gaussian operations such as photon addition $|1, 8\rangle$, photon subtraction is still the most feasible approach. For our purposes, entanglement distillation is also a suitable testing example, because the entanglement is vulnerable to decoherence differently than negativity of the Wigner function. We will show, however, that also in this scenario the improvements in success rate can overcome the costs.
In our test model we consider a pure two mode squeezed state

\[ |\psi\rangle = \frac{1}{\cosh r} \sum_{f=0}^{\infty} (\tanh r)^f |f\rangle |f\rangle. \] (8)

Such highly pure states are nearly available with almost 8 dB of generalized squeezing. In order to increase the entanglement of the state we apply the adaptive photon subtraction in both of its arms, see Fig. 5. We will consider only the first step of the distillation protocol, the degaussification. However, to reflect that we are ultimately interested in the properties of the Gaussian approximation of the final state we will quantify the entanglement of both the initial and the distilled state by Gaussian logarithmic negativity \[ N \], which can be calculated from the covariance matrix

\[ \sigma = \begin{pmatrix} \alpha & \gamma \\ \gamma^\dagger & \beta \end{pmatrix} \] (9)

of the state using the PPT transformed symplectic invariant \[ \Delta = \det \alpha + \det \beta - 2 \det \gamma \] of the variance matrix \[ \sigma \] as \[ N = \max\{0, -\log_2 \nu\} \] with

\[ \nu = \sqrt{2 \left( \Delta - \sqrt{\Delta^2 - 4 \det \sigma}\right)} . \] (10)

In Fig. 6 we show the dependence of the logarithmic negativity on the probability of success and the number of allowed steps. We have considered the symmetric scenario in which both modes of the entangled state were subjected to adaptive subtraction with the identical number of allowed steps. Overall we can say that the maximally achievable distilled Gaussian logarithmic negativity is preserved for the adaptive subtraction and that for all values of logarithmic negativity the adaptive procedure can reach higher success probability than the traditional single step technique. In all the cases, the maximal logarithmic negativity can be asymptotically reached in the limit \[ P_s \to 0 \], but the rate of convergence depends on the number of steps. For all distilled states with non-maximal logarithmic negativity the probability of obtaining them can be increased roughly by an order of magnitude when the adaptive subtraction with at most 10 steps is applied. Going to 75 steps increases the probability further, but our numerical calculations suggest the improvement begins to be saturated at that point. Interestingly enough, the qualitative improvement in success rate for all considered regimes is preserved even when the current superconducting detectors are considered.

V. CONCLUSION AND OUTLOOK

We have proposed an adaptive protocol for improved photon subtraction which can achieve both higher probability of success and higher quality of the operation. The protocol is based on a simple premise: when a standard photon
In this paper we have discussed only single photon subtraction, but the basic principle of the adaptive operation is not limited to it. In its place there could be any kind of probabilistic operation which is triggered by a certain measurement result. This could be as simple as multiple photon subtraction \[ R \], but also operations such as photon addition \[ R \] or also operations realized by coupling with discrete quantum systems \[ R \]. Further extension lies in dynamical tuning of the interaction parameters. While we have considered identical couplings in all the loops of the operation, their dynamical adjustment would likely result in improving the trade-offs between the probability of success and the quality of the operation.

**Funding**

P.M. and J. P. acknowledge grant GA18-21285S of the Czech Science Foundation. J. P. also acknowledges Palacký University project IGA PrF-2018-010. R.F. acknowledges national funding from the MEYS and the funding from European Union’s Horizon 2020 (2014-2020) research and innovation framework programme under grant agreement No 731473.

---

[1] A. Zavatta, Science 306, 660 (2004), URL [https://doi.org/10.1126/science.1103190](https://doi.org/10.1126/science.1103190)
[2] A. Ourjoumtsev, Science 312, 83 (2006), URL [https://doi.org/10.1126/science.1122858](https://doi.org/10.1126/science.1122858)
[3] V. Parigi, A. Zavatta, M. Kim, and M. Bellini, Science 317, 1890 (2007), URL [https://doi.org/10.1126/science.1146204](https://doi.org/10.1126/science.1146204)
