Success Probabilities for Linear Optics Gates

Institute for Mathematical Sciences

Koenraad M.R. Audenaert

Quantum Information at Imperial College London

06/06/06

Impressum

Extended Club Remix of work by Scheel, Lütkenhaus, Audenaert, Eisert:

- S. Scheel, *quant-ph/0406127*.
- S. Scheel and N. Lütkenhaus, NJP **6**, 51 (2004).
- S. Scheel and K. Audenaert, NJP **7**, 149 (2005).
- J. Eisert, PRL **95**, 040502 (2005).

Unified presentation, shorter proofs, more general results, and a small ToDo-list...

A basic component in linear-optics quantum computing

- One of the promising routes to implementing small-scale quantum networks is by using linear-optical networks.
- Certain two-mode gates (e.g. controlled-σ_z) can be generated by acting separately on both modes within a Mach–Zehnder interferometric setup.
- In that way the complexity of a two-mode gate is reduced to a single-mode gate: the nonlinear sign-shift gate

$$c_0|0\rangle + c_1|1\rangle + c_2|2\rangle \longmapsto c_0|0\rangle + c_1|1\rangle - c_2|2\rangle$$

- A number of theoretical works (Ralph, Scheel, Lapaire, KLM, Lufthansa) give a variety of networks capable of implementing this gate.

Subject of the talk

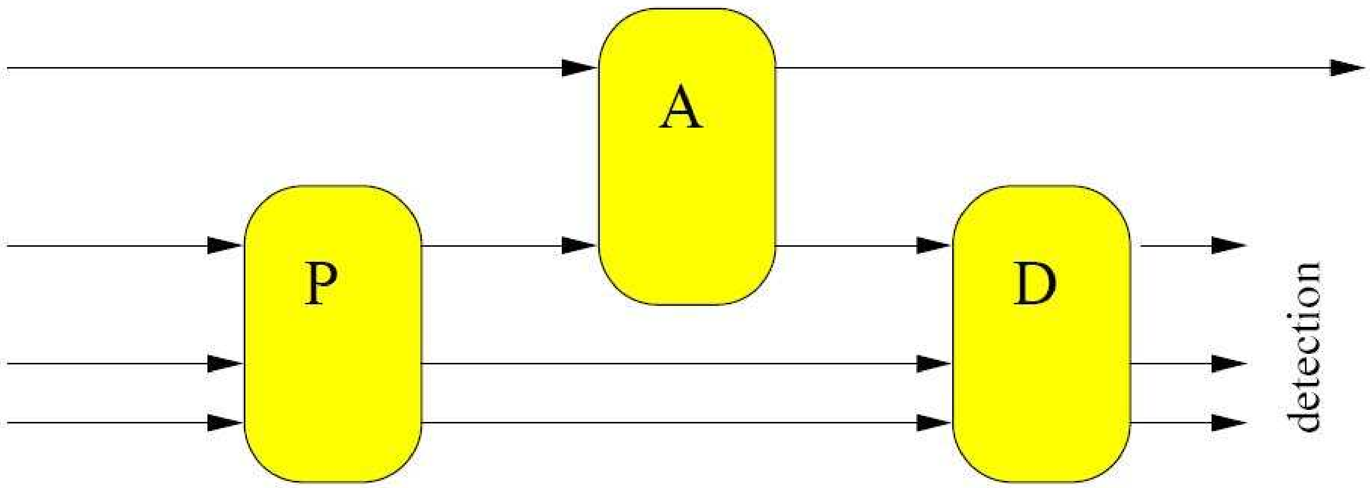
- The crucial question for the ability to concatenate linear-optical gates are their probabilities of success.
- In these works we studied the question how the probability of success scales as the signal state becomes higher-dimensional, thereby giving a hint on the possible scaling law for multi-mode quantum gates.
- We considered the $N+1$-dimensional sign-shift gate

$$c_0|0\rangle + \ldots + c_{N-1}|N-1\rangle + c_N|N\rangle \mapsto c_0|0\rangle + \ldots + c_{N-1}|N-1\rangle - c_N|N\rangle.$$

- We limit ourselves to single-shot implementations without conditional feed-forward.
- We consider an abstract three-stage type of network that includes all possible networks.

Basic Circuit

- A: active beam splitter
- P: ancilla preparation
- D: ancilla projection



Basic Circuit

- The generality of this abstraction follows from the observation that each U(N) network can be decomposed into a triangle-shaped network of at most $N(N-1)/2$ beam splitters (and some additional phase shifters) [Reck et al, 94].
- Then, the signal mode can be chosen to impinge only on a single beam splitter 'A', with the rest of the network divided into the ancilla preparation and detection stages 'P' and 'D' [Scheel and Lütkenhaus, 04].
- The 'D' stage is a two-outcome POVM:
  - outcome '1' = "accept";
  - outcome '0' = "bad luck; try again"
- We can show that having just one "accept" outcome is optimal.

Basic Circuit

- A: active beam splitter, $\hat{U}$, transmittivity T, reflectivity R; $|T|^2 + |R|^2 = 1$
- P: ancilla preparation, $|\phi\rangle = \sum_{l=0}^{n} \gamma_l \, |l\rangle|n-l\rangle$
- D: ancilla projection on $|\psi\rangle = \sum_{l=0}^{n} \alpha_l^* \, |l\rangle|n-l\rangle$
- The vectors α and γ are, of course, normalised.
- The states $|n-l\rangle$ can be multimode, and play no further role.
- The matrix elements of $\hat{U}$ in the n-photon sector are

$$\langle m_1, m_2|\hat{U}|n_1, n_2\rangle = \frac{1}{\sqrt{m_1! m_2! n_1! n_2!}} \operatorname{Per}\begin{pmatrix} T E_{m_1,n_1} & R E_{m_1,n_2} \\ -R^* E_{m_2,n_1} & T^* E_{m_2,n_2} \end{pmatrix},$$

with $m_1 + m_2 = n_1 + n_2 = n$, and $E_{j,k}$ denotes a $j \times k$ matrix all of whose entries are 1.

Basic Circuit Behaviour

- Let the input state be the basis vector $|k\rangle$.
- By orthogonality of the states $|n-l\rangle$, the post-measurement state $|\psi_k\rangle$ is

$$|\psi_k\rangle = \sum_{j=0}^{N} \sum_{l=0}^{n} \alpha_l \gamma_l \langle j,l|\hat{U}|k,l\rangle \, |j\rangle.$$

- Since a lossless beam splitter is photon-number preserving, there are only contributions from the term $j = k$.
- Denote $a_{kl} := \langle k,l|\hat{U}|k,l\rangle$, then

$$|\psi_k\rangle = (\sum_{l=0}^{n} \alpha_l \gamma_l a_{kl})|k\rangle.$$

- As the states $|n-l\rangle$ have left the picture, we can henceforth let n be as large as imaginable, say ∞.

A Formula for the Success Probability

- Let A be the matrix with entries a_{jk}, $j = 0 \ldots N$ and $k = 0 \ldots \infty$.
- Theorem [Scheel and Audenaert]: The maximal success probability for obtaining from A the non-linear gate $c_k \mapsto y_k c_k$, for $k = 0 \ldots N$, is given by

$$P_{\max} = \max_x \left\{ \frac{1}{|x|_1^2} : Ax = y \right\}$$

- $|.|_1$ is the ℓ_1 vector norm.
- The minimisation is over all infinite-dimensional complex vectors $x = (x_j)_{j=0}^{\infty}$.

Proof

- The transformation realised using ancilla γ_l and measurement α_l^* is

$$c_k \mapsto y'_k c_k \quad (k = 0 \ldots N), \text{ with } y'_k = \sum_{l=0}^{\infty} a_{kl}(\alpha_l \gamma_l).$$

- More concisely, I write this as $y' = A(\alpha\gamma)$, where $A = (a_{kl})_{k=0}^{N}{}_{,l=0}^{\infty}$.
- The target transformation is realised, with probability P, if and only if

$$y' = \sqrt{P} y.$$

- To find a realisation with maximal P, first fix a vector x s.t. $y = Ax$.
- Find normalised α and γ s.t. $\alpha\gamma = \sqrt{P}x$ and P is as large as possible.

Proof

- Find normalised α and γ s.t. $\alpha\gamma = \sqrt{P}x$ and P is as large as possible.
- Fix γ. Then $\alpha = \sqrt{P}x/\gamma$. Since $|\alpha|_2$ must be 1, this fixes $P = 1/|x/\gamma|_2^2$.
- Now maximise this P by varying over all normalised γ.
- We find $\min_\gamma\{|x/\gamma|_2 : |\gamma|_2 = 1\} = |x|_1$.
  Proof: using Cauchy-Schwarzy
  $$|x/\gamma|_2\,|\gamma|_2 \geq \sum_l |(x_l/\gamma_l)\gamma_l| = |x|_1,$$
  and noting that equality can be obtained (with $|\gamma_l|^2 = c|x_l|$).
- Hence the maximum P for given x (s.t. $Ax = y$) is $P = 1/|x|_1^2$.
- Finally, we maximise this by varying over all x s.t. $Ax = y$. $\square$

A Formula for the Success Probability

- Corollary [Scheel and Audenaert]: The maximal success probability for obtaining from A the non-linear gate $c_k \mapsto y_k c_k$, for $k = 0 \ldots N$, is given by

$$P_{\max} = \max_x \left\{ \frac{1}{|x|_1^2} : Ax = y \right\}$$

where the x can be restricted to those that have at most $N+1$ non-zero entries.

- Geometrical proof

Numerical method

- Corollary reduces the optimisation over $\mathbb{C}^\infty$ to an optimisation over $\mathbb{N}^{N+1}$.
- We can replace x by a list of $N+1$ positions, $0 \leq n_1 < n_2 < \ldots < n_l$, and corresponding values x_l.
- The optimisation over x_l can be solved directly by matrix inversion: only one x is feasible, namely $x^0 := \tilde{A}^{-1} y$. Here, $\tilde{A}$ is an $(N+1) \times (N+1)$ matrix containing the $N+1$ columns of A designated by the positions n_l.
- While this observation dramatically simplifies numerical approaches, the remaining optimisation over the n_l still makes for tedious business if we are after analytical results.
- For that reason we will consider an alternate expression for the maximal success probability, first obtained by Jens Eisert.

Upper bound on Success Probability

- Let A again be the matrix with entries a_{jk}, $j = 0 \ldots N$ and $k = 0 \ldots \infty$,
- Theorem [Eisert]: The maximal success probability for obtaining from A the non-linear phase-shift gate $c_k \mapsto y_k c_k$, with $|y_k| = 1$, for $k = 0 \ldots N$, is upper bounded by

$$P_{\max} \le \min_s \left(\frac{|A^\dagger (ys)|_\infty}{|\sum_{j=0}^N s_j|} \right)^2 .$$

- $|.|_\infty$ is the ℓ_∞ vector norm,
- The minimisation is over all $(N+1)$-dimensional (finite!) complex vectors $s = (s_j)_{j=0}^N$.

Proof

- Hölder's inequality for vectors: $|\langle a|b\rangle| \leq |a|_1 \, |b|_\infty$.

Proof

- Hölder's inequality for vectors: $|\langle a|b\rangle| \leq |a|_1 \, |b|_\infty$.
- In our case: put $a = x$ and $b = A^\dagger(ys)$

$$|A^\dagger(ys)|_\infty \, |x|_1 \geq |x^\dagger A^\dagger(ys)|$$

Proof

- Hölder's inequality for vectors: $|\langle a|b\rangle| \le |a|_1\,|b|_\infty$.
- In our case: $Ax = y$

$$|A^\dagger(ys)|_\infty\,|x|_1 \ge |x^\dagger A^\dagger(ys)| = |y^\dagger(ys)|$$

Proof

- Hölder's inequality for vectors: $|\langle a|b\rangle| \leq |a|_1\, |b|_\infty$.
- In our case:

$$|A^\dagger(ys)|_\infty\, |x|_1 \geq |x^\dagger A^\dagger(ys)| = |y^\dagger(ys)| = |\sum_j s_j |y_j|^2|$$

Proof

- Hölder's inequality for vectors: $|\langle a|b\rangle| \leq |a|_1\,|b|_\infty$.
- In our case: $|y_j| = 1$

$$|A^\dagger(ys)|_\infty\,|x|_1 \geq |x^\dagger A^\dagger(ys)| = |y^\dagger(ys)| = |\sum_j s_j|y_j|^2| = |\sum_j s_j|$$

Proof

- Hölder's inequality for vectors: $|\langle a|b\rangle| \leq |a|_1\,|b|_\infty$.
- In our case:

$$|A^\dagger(ys)|_\infty\, |x|_1 \geq |x^\dagger A^\dagger(ys)| = |y^\dagger(ys)| = |\sum_j s_j |y_j|^2| = |\sum_j s_j|$$

- Thus, for all s and for all x satisfying $Ax = y$:

$$\frac{|A^\dagger(ys)|_\infty}{|\sum_{j=0}^N s_j|} \geq \frac{1}{|x|_1},$$

Proof

- Hölder's inequality for vectors: $|\langle a|b\rangle| \leq |a|_1\,|b|_\infty$.
- In our case:

$$|A^\dagger(ys)|_\infty\,|x|_1 \geq |x^\dagger A^\dagger(ys)| = |y^\dagger(ys)| = |\sum_j s_j|y_j|^2| = |\sum_j s_j|$$

- Thus, for all s and for all x satisfying $Ax = y$:

$$\frac{|A^\dagger(ys)|_\infty}{|\sum_{j=0}^N s_j|} \geq \frac{1}{|x|_1},$$

- Minimise LHS over all s, and maximise RHS over all x s.t. $Ax = y$:

$$\min_s \frac{|A^\dagger(ys)|_\infty}{|\sum_{j=0}^N s_j|} \geq \max_{x:Ax=y} \frac{1}{|x|_1} .$$

Proof

- Hölder's inequality for vectors: $|\langle a|b\rangle| \leq |a|_1 \, |b|_\infty$.
- In our case:

$$|A^\dagger(ys)|_\infty \, |x|_1 \geq |x^\dagger A^\dagger(ys)| = |y^\dagger(ys)| = |\sum_j s_j |y_j|^2| = |\sum_j s_j|$$

- Thus, for all s and for all x satisfying $Ax = y$:

$$\frac{|A^\dagger(ys)|_\infty}{|\sum_{j=0}^N s_j|} \geq \frac{1}{|x|_1},$$

- Minimise LHS over all s, and maximise RHS over all x s.t. $Ax = y$:

$$\min_s \frac{|A^\dagger(ys)|_\infty}{|\sum_{j=0}^N s_j|} \geq \max_{x:Ax=y} \frac{1}{|x|_1} = \sqrt{P_{\max}}.$$

□

Maximal Success Probability

- We have seen that the expression for the a_{jk} depends on the transmittivity T, which is also a tunable parameter for maximising P. Hence we write $A = A(T)$.
- The claim is that the phase-shift gate $c_k \mapsto y_k c_k$, with $y_k = 1$, for $k = 0 \ldots N-1$, and $y_N = -1$, can be obtained with probability at most $P_{\max} = 1/N^2$.
- This can be proven using the Eisert bound.
- Need to find $s(T)$ such that

$$\frac{|A(T)^\dagger (ys(T))|_\infty}{|\sum_{j=0}^{N} s(T)_j|} = 1/N.$$

Maximal Success Probability

- For $N = 2$, and restricting to real T, the following s does the trick:

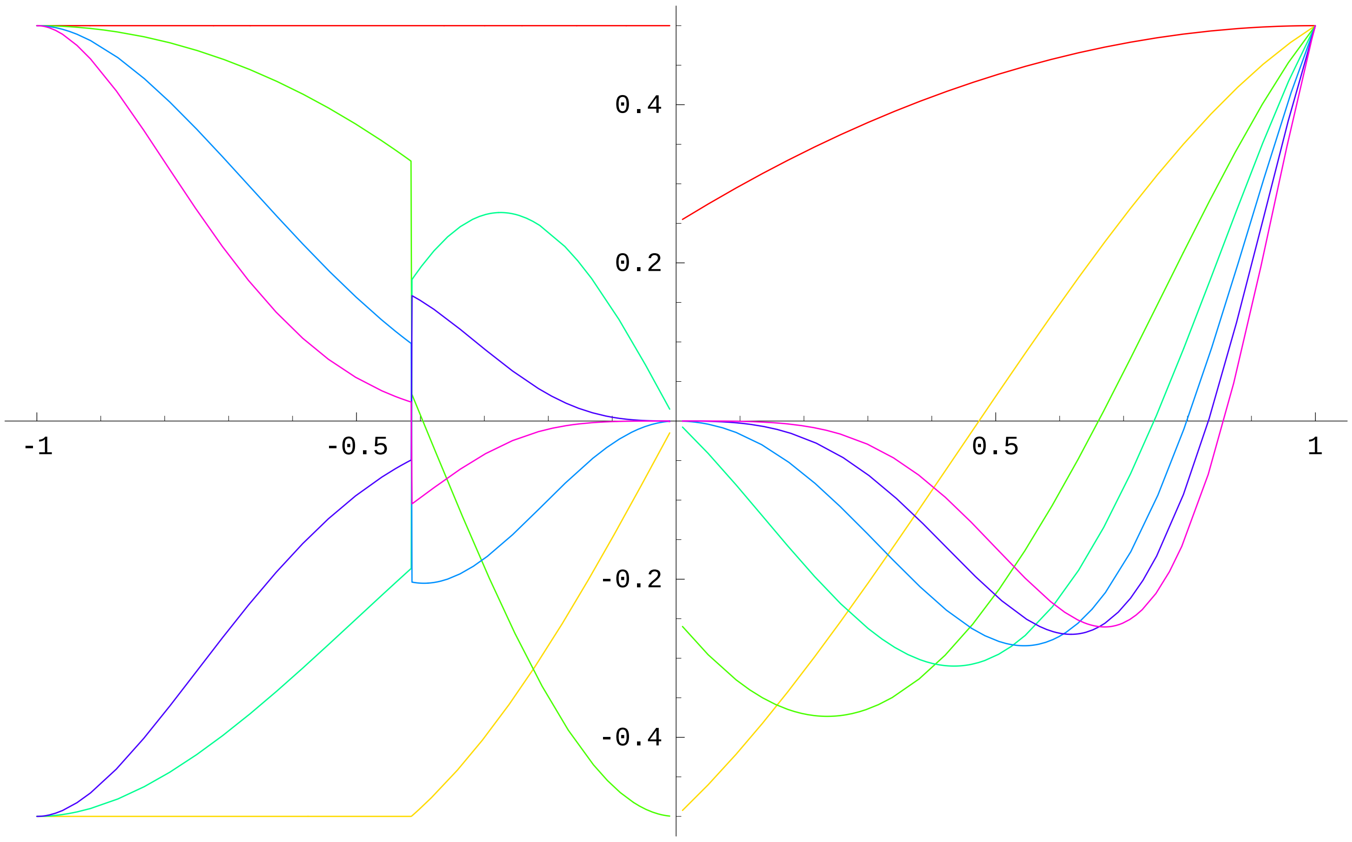
$$s_1, s_2 = \frac{1}{2} \begin{cases} 1/(1-T), \; 0, & \text{if } T \in [-1, 1-\sqrt{2}), \\ 0, \; 1/(1+T^2), & \text{if } T \in [1-\sqrt{2}, 0), \\ 1, \; 1/2, & \text{if } T \in [0, 1), \end{cases}$$

  and $s_0 = 1 - s_1 - s_2$.

- For general N, similar choices can most likely be found (To Do!)
- On the next page I plot the first 7 entries of $A(T)^\dagger(ys(T))$, showing that they are all bounded by 1/2, which suggests that P is bounded by 1/4.
- A real proof is of course needed that this is true for all entries (Need More Chocolate!)

Maximal Success Probability

First 7 entries of $A^{\dagger}(ys)$ versus T



Is this bound achievable?

- To prove that the upper bound $P \leq 1/N^2$ is achievable, consider the original problem formulation (modified with an additional maximisation over T)

$$P_{\max} = \max_T \max_x \left\{ \frac{1}{|x|_1^2} : A(T)x = y \right\}$$

- Claim **(a)**: $P_{\max} = 1/N^2$ is achieved for $T = 1 - 2^{1/N}$ and $x = A_N(T)^{-1}y$ where A_N is the leftmost $(N+1) \times (N+1)$ submatrix of A.
- Claim **(b)**: for this x, $x_l = 0,\ l \geq N$.
- We already knew that $N+1$ non-zero entries were enough; now this says that N non-zero entries suffice. Moreover, its the entries with lowest photon number.
- Hence, maximum P is achieved with the lowest possible ancilla dimension!

Entries of A and its Properties

- Some preliminaries...
- First we need an explicit formula for the a_{jk}.
- The permanents can be calculated explicitly, using the Lagrange expansion formula for the permanent. After a long calculation we find

$$a_{kl} = (T^*)^{l-k} \sum_{m=0}^{k} \binom{k}{m} \binom{l+m}{m} t^m,$$

where $t = |T|^2 - 1$.

- This expression does not explicitly contain N. This means that the $(N+1)\times\infty$ matrix A is just a submatrix of the infinite-dimensional $\mathscr{A} := (a_{kl})_{k,l=0}^{\infty}$.

Entries of A and its Properties

- This matrix $\mathscr{A}$ can be decomposed as a product of a number of matrices with a very simple structure.
- Define the matrices

$$
\begin{aligned}
D(x) &:= \operatorname{diag}(1, x, x^2, \ldots, x^\infty) \\
\Delta &:= (\binom{k}{l})_{k,l=0}^{\infty}.
\end{aligned}
$$

  Δ is a representation of Pascal's triangle by a lower triangular matrix.
- A straightforward calculation leads to the decomposition

$$\mathscr{A} = D(T^*)\ \Delta\ D(\frac{|T|^2 - 1}{|T|^2})\ \Delta^T\ D(T).$$

Entries of A and its Properties

- We will also need the matrix A_N, which is the upper left $(N+1)\times(N+1)$ submatrix of $\mathscr{A}$. Define $D_N(x)$ and Δ_N accordingly.
- Since $\det(\Delta_N)=1$, $\det(A_N)=\det(D_N(t))=t^{N(N+1)/2}$. Thus A_N is invertible provided $t\neq 0$, i.e. $|T|<1$.
- The inverse of Δ_N is Δ_N itself, but with additional negative signs arranged in a checkerboard pattern:

$$\sum_{k=0}^{N}(-1)^{k-m}\binom{p}{k}\binom{k}{m}=\delta_{pm}.$$

- This easily gives the inverse of A_N as

$$(A_N^{-1})_{jk}=(-T^*)^{k-j}\sum_{l=0}^{N}t^{-l}\binom{l}{k}\sum_{p=0}^{N}\binom{p}{j}\binom{p}{l}.$$

Equations: buy one, get one free

- We're claiming that x with $x_l = 0$ for $l > N$ are optimal.
- The only such x that satisfies $Ax = y$ is $x = A_N^{-1} y$.
- With $t = |T|^2 - 1$,

$$
\begin{aligned}
x_j &= \sum_{k=0}^{N} (A_N^{-1})_{jk} y_k = \sum_{k=0}^{N} (A_N^{-1})_{jk} - 2(A_N^{-1})_{jN} \\
&= (-T^*)^{-j} \left(\sum_{p=0}^{N-1} \binom{p}{j} \left(1 + \frac{1-T^*}{t}\right)^p - \left(\frac{-T^*}{t}\right)^N \binom{N}{j} \right).
\end{aligned}
$$

Equations: buy one, get one free

- In $T = 1 - 2^{1/N}$, x_j simplifies to

$$x_j = (-1)^j \frac{1}{(T+1)^{N-1}} \sum_{l=0}^{N-j-1} \binom{N}{l} T^l.$$

- Thus, clearly, $x_N = 0$ in that point: claim **(b)**.
- For $T = 1 - 2^{1/N}$ the sum over l always yields positive values. Thus

$$|x_j| = \frac{1}{(T+1)^{N-1}} \sum_{l=0}^{N-j-1} \binom{N}{l} T^l.$$

Equations: buy one, get one free

- Summing over all j finally yields

$$\begin{aligned}\sum_{j=0}^{N-1}|x_j| &= \frac{1}{(T+1)^{N-1}}\sum_{j=0}^{N-1}\sum_{l=0}^{N-j-1}\binom{N}{l}T^l\\ &= \frac{1}{(T+1)^{N-1}}\sum_{l=0}^{N-1}(N-l)\binom{N}{l}T^l\\ &= N.\end{aligned}$$

- Thus, for the success probability we obtain claim **(a)**:

$$P = \frac{1}{N^2}.$$

□

Conclusion

- We have proven that the $N+1$ dimensional phase-gate can be implemented using a P-A-D circuit with success probability $1/N^2$.
- We have shown that, for $N = 2$, this is the maximum.
- It remains to find trial functions $s_1(T), \ldots, s_N(T)$ to show that this is the maximum for any $N > 2$.
- We have shown that this success probability, taken as a function of the chosen ancilla photon numbers, can be obtained if the ancilla contains as few photons as possible (namely $N-1$).
- This is fortunate because in this case the least quantum-state engineering is needed to generate the ancilla in the preparation stage 'P' and, furthermore, decoherence affects the ancilla only minimally.