Circuit gradient

July 2022

Let us consider a photonic circuit described by some unitary matrix U.

`https://arxiv.org/abs/2102.04579v2` gives us the following formula that can be used for the probability computation:

$$\langle s| U |t\rangle = \frac{Per(U_{s,t})}{\sqrt{s!}\sqrt{t!}},$$

where $U_{s,t}$ is obtained from U by repeating s_i times the i^{th} row and t_j times its j^{th} column.

We want to compute the derivative of $Per(U_{s,t})$ relative to a parameter of some component of the circuit.

1 Derivation of the Permanent:

According to `https://books.google.fr/books?id=UHLjBwAAQBAJ&printsec=frontcover&hl=fr&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false`, page 2, for any matrices U and H:

$$Per(U+H) = \sum_{s,t;|s|=|t|;s,t\subset\{1,\ldots,n\}} Per(U_{i,j})_{i\in s,j\in t} Per(H_{i,j})_{i\in s^c,j\in t^c}.$$

Thus:

$$Per(U+H) = Per(U) + \sum_{i,j} H_{i,j} Per(U^{ij}) + o(\|H\|),$$

where: U^{ij} is obtained by taking off the i^{th} row and the j^{th} column from U. Hence:

$$\frac{\partial Per}{\partial u_{k,l}}(U) = Per(U^{kl})$$

Since the matrix $U_{s,t}$ is obtained from the circuit unitary matrix U by repeating its lines and columns, we only need to consider the following problem:

2 Unitary matrix derivation with *one photon*:

2.1 Phase shifter:

Let us consider the problem of computing the partial derivative of the term $u_{i,j}$ of the unitary circuit matrix U relative to a parameter δ of a phase shifter.

For a two mode circuit, the unitary matrix of a phase shifter applied on the first mode reads:

$$\begin{pmatrix} e^{i\delta} & 0 \\ 0 & 1 \end{pmatrix}$$

It is easy to see that $u_{k,l}$ is a polynomial function of $e^{i\delta}$. Moreover, since $e^{i\delta}$ appears only once in the different component unitary matrices, the aforementioned polynomial function has a degree of at most 1, hence the expression:

$$u_{k,l} = \alpha e^{i\delta} + \beta,$$

where α and β do not depend on δ. The sought derivative is thus:

$$\frac{\partial u_{k,l}}{\partial \delta} = i\alpha e^{i\delta}$$

Now, the remaining task is to find the expression of α. Let us suppose that U can written in the form:

$$U = A \begin{pmatrix} e^{i\delta} & 0 \\ 0 & 1 \end{pmatrix} B$$

where A and B are two unitary matrices. Let us develop the computation:

$$\begin{aligned} U &= A \begin{pmatrix} e^{i\delta} & 0 \\ 0 & 1 \end{pmatrix} B \\ &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} e^{i\delta} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x & y \\ z & w \end{pmatrix} \\ &= \begin{pmatrix} axe^{i\delta} + bz & aye^{i\delta} + bu \\ cxe^{i\delta} + dz & cye^{i\delta} + dz \end{pmatrix} \end{aligned}$$

One can remark that the $(k,l)^{th}$ coefficient in front of $e^{i\delta}$ can be written as: $\alpha = A[k,0]B[0,l]$.

More generally for a circuit with any number of modes, if the phase shifter is applied on the j^{th} mode, one can write:

$$\alpha = A[k,j]B[j,l]$$

Thus:

$$\frac{\partial u_{k,l}}{\partial \delta_j} = iA[k,j]B[j,l]e^{i\delta}$$

2.2 Beam spliter:

Let us repeat the same steps with a beam spliter. For a two mode circuit, the unitary matrix of a beam spliter reads:

$$\begin{pmatrix} \cos\theta & ie^{i\phi}\sin\theta \\ ie^{-i\phi}\sin\theta & \cos\theta \end{pmatrix}$$

Developing the computation:

$$\begin{aligned}
U &= A\begin{pmatrix} \cos\theta & e^{i\phi}\sin\theta \\ ie^{-i\phi}\sin\theta & \cos\theta \end{pmatrix} B \\
&= \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} \cos\theta & e^{i\phi}\sin\theta \\ ie^{-i\phi}\sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x & y \\ z & w \end{pmatrix} \\
&= \begin{pmatrix} ax\cos\theta + ibxe^{-i\phi}\sin\theta + iaze^{i\phi}\sin\theta + bz\cos\theta & ay\cos\theta + ibye^{-i\phi}\sin\theta + iaue^{i\phi}\sin\theta + bu\cos\theta \\ cx\cos\theta + idxe^{-i\phi}\sin\theta + icze^{i\phi}\sin\theta + dz\cos\theta & cy\cos\theta + idye^{-i\phi}\sin\theta + icue^{i\phi}\sin\theta + du\cos\theta \end{pmatrix} \\
&= \begin{pmatrix} (ax+bz)\cos\theta + i(bxe^{-i\phi} + aze^{i\phi})\sin\theta & (ay+bu)\cos\theta + i(bye^{-i\phi} + aue^{i\phi})\sin\theta \\ (cx+dz)\cos\theta + i(dxe^{-i\phi} + cze^{i\phi})\sin\theta & (cy+du)\cos\theta + i(dye^{-i\phi} + cue^{i\phi})\sin\theta \end{pmatrix}
\end{aligned}$$

We deduce from this computation that, more generally, for a beam spliter acting upon the modes j_0 and j_1:

$$u_{k,l} = (A[k,j_0]B[j_0,l] + A[k,j_1]B[j_1,l])\cos\theta + i(A[k,j_1]B[j_0,l]e^{-i\phi} + A[k,j_0]B[j_1,l]e^{i\phi})\sin\theta$$

Hence:

$$\frac{\partial u_{k,l}}{\partial\theta} = -(A[k,j_0]B[j_0,l] + A[k,j_1]B[j_1,l])\sin\theta + i(A[k,j_1]B[j_0,l]e^{-i\phi} + A[k,j_0]B[j_1,l]e^{i\phi})\cos\theta$$

and

$$\frac{\partial u_{k,l}}{\partial\phi} = (A[k,j_1]B[j_0,l]e^{-i\phi} - A[k,j_0]B[j_1,l]e^{i\phi})\sin\theta$$

3 Final formula:

We have:

$$\frac{\partial Per}{\partial\alpha}(U) = \sum_{k,l}\frac{\partial Per}{\partial u_{k,l}}(U)\frac{\partial u_{k,l}}{\partial\alpha}(\alpha),$$

where $\frac{\partial u_{k,l}}{\partial\alpha}(\alpha)$ is given by the formulas given above.

4 Stochastic Gradient Minimization:

The remaining question is how to optimize the computation of A and B. Indeed, if we apply a usual minimization algorithm, the values of A and B will be changed at each iteration, requiring a new computation of the quantities $A[., j]B[j, .]$.

The solution we propose is a stochastic gradient algorithm. We try to use a maximal (i.e. to which no one can be added) set of independent phase shifters/beam spliters as a batch (the values of A and B remain the same in each mode subspace). We represent our general interferometer as a graph with the phase shifters and the beam spliters as nodes and the modes as edges (the "lines" in the drawing of the circuit). One can see that two phase shifters/beam spliters are independent iif there is no path linking them in the graph. Our goal is to find maximal (not maximum) independent sets. We do not know if there is a more efficient algorithm (there is room for improvement), but we propose the following two step algorithm:

Step 1 - Listing: $O(N^3)$, where N is the number of phase shifters/beam shifters

1. For each phase shifter/beam shifter, do a DFS in order to find the phase shifters/beam shifters with whom it is in interaction (in one way or the other)
2. Build a map `interact_list` (a dictionary, a list...) that, for each phase shifter/beam spliter associates the list of phase shifters/beam spliters found in the first step.

Step 2 - Generation of a batch $O(m)$, `list_s` is the list of available phase shifters / beam shifters

1. Initialize `list_s` with all the phase shifters / beam shifters.
2. Draw a phase shifter/beam shifter `s` from `list_s` randomly.
3. Eliminate it and the members of `interact_list(s)` from `list_s`.
4. If `list_s` is not empty, go to step 2.

Our idea is to use the same batch a fixed number of iterations before regenerating it again. Note that the listing step needs to be performed only at the beginning of the optimisation. The batch generation step can be performed either at the beginning (resulting in constant batches, we will have to add a verification that all phase shifters/beam spliters are in at least one batch) or after each fixed number of iterations. The performances of the two solutions will have to be tested (both in final loss and in computation time).

4.1 Illustrations of the concept on independent phase shifters:

In order to give some intuition of the concept of independent phase shifters, we present two figures where some sets of independent phase shifters are represented in the same color.

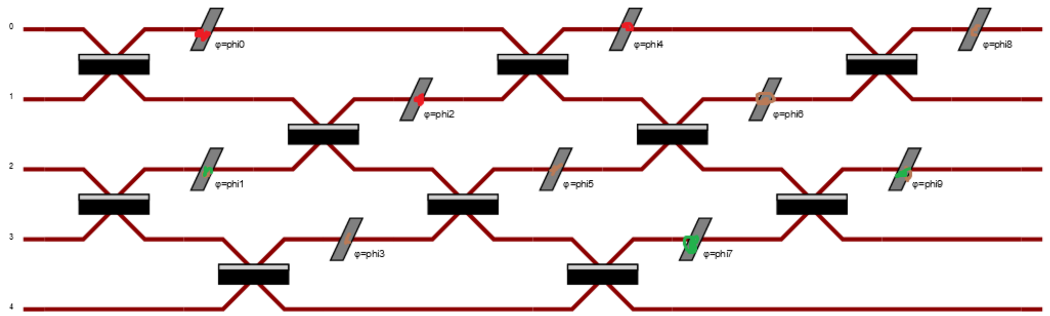

Figure 1: Some sets of independent phase shifters are represented by the same color, here for a five mode circuit.

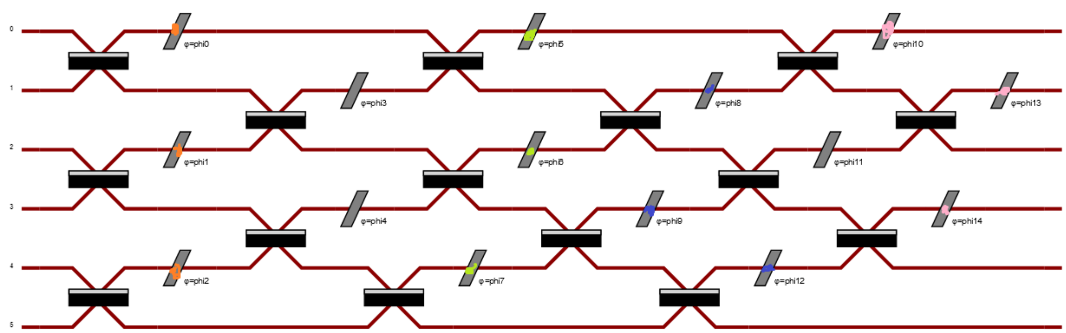

Figure 2: Some sets of independent phase shifters are represented by the same color, here for a six mode circuit.