InstaQ, or Computational Jones Decomposition: An Algorithm to Reconstruct Any Pure 2-Qubit State as a Product of Jones Matrices

Oscar Scholin
Department of Physics, Pomona College, Claremont, CA 91711

In this letter, we describe the process of decomposing any arbitrary pure 2-qubit density matrix into a series of Jones matrices representing half waveplates, quarter waveplates, quartz plates, and barium borate crystals, which provides a injective correspondence to an experimental quantum optics setup. We achieve this computationally by minimizing a loss function $L = 1 - \sqrt{\mathcal{F}}$, where $\mathcal{F}$ is the Uhlmann-Josza fidelity of our predicted state to the target. Table II presents the results of the model of test states for two primary setup configurations, with an accuracy ($\mathcal{F}$) exceeding 0.99 for states the setup is capable of creating. To our knowledge no similar model exists in the literature, so this could potentially be a valid tool in both theoretical and especially experimental uses for investigating the properties of pure 2-qubit states—once certain experimental discrepancies are addressed.

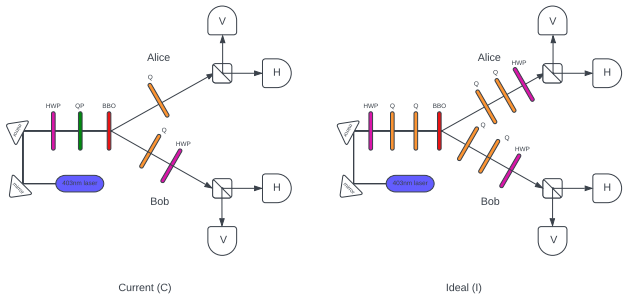

FIG. 1: Two sample experimental setup configurations: "C" on the left and "I" on the right. N.b. HWP is half waveplate, QWP is quarter waveplate, QP is quartz plate, BBO is the BBO crystal. The square with a line on its diagonal represents a polarizing beam splitter. "H" and "V" are detectors of horizontally and vertically polarized light respectively.

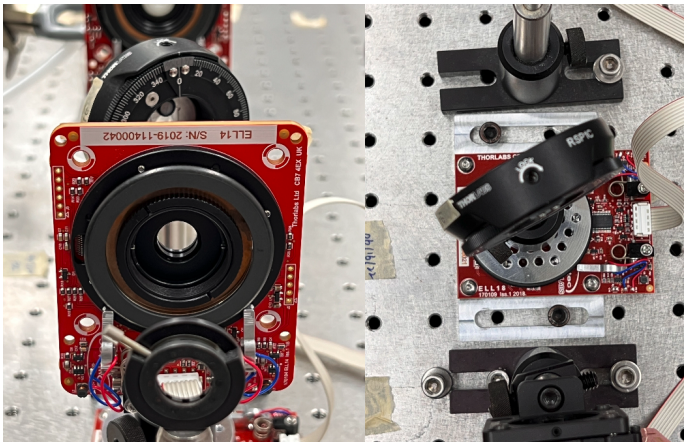

FIG. 2: Image from experimental setup. *(a–left)* A half waveplate front on, which rotates in the vertical plane (a quarter waveplate looks identical). *(b–right)* A quartz plate top down, which rotates in the horizontal plane. The red is the Thor Labs PCB that allows a computer to issue precise movement instructions for angle settings.

I. INTRODUCTION

It is known that any SU2 matrix can be decomposed as a product of three commutative matrices. These matrices can be defined as

$$H(\theta) = \begin{bmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{bmatrix},$$
$$Q(\alpha) = R(\alpha) \begin{bmatrix} e^{i\pi/4} & 0 \\ 0 & e^{-i\pi/4} \end{bmatrix} R(-\alpha), \qquad (1)$$
$$\text{for } R(\alpha) = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix},$$

which give the product $S = H(\theta)Q(\alpha_1)Q(\alpha_2)$ [1]. These matrices have strong physical significance: they are the Jones matrices for half and quarter wave plates (HWP, QWP) respectively, oriented about the vertical axis by some angle θ for the HWP and α for the QWPs.

Motivated by this theorem, we can model an arbitrary polarization state by multiplying together various Jones matrices to represent each of the optical components in our setup that permute an input state. Using the diagrams in figure 1, we can derive the expression for an arbitrary state created by changing the adjustable angle parameters in each of the matrices; this process of choosing the input parameters to the matrices is analogous to actually changing the dials on the corresponding devices in the lab, up to calibration, described in section V. To do this, we need to make some more definitions:

$$QP = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{bmatrix}, B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{bmatrix}, \qquad (2)$$
$$S_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

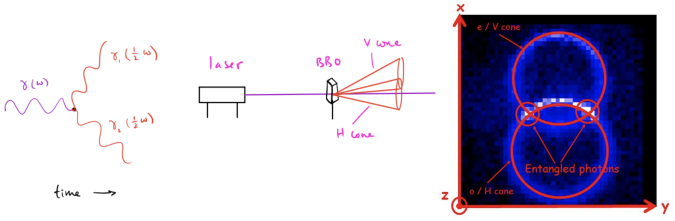

FIG. 3: Illustration of SPDC. *(a–left)* Feynmann Diagram for the creation of two 810nm photons from an incident 405nm. *(b–middle)* Sketch showing the emergence of H and V polarized cones after passing through the BBO. *(c–right)* Vertical slice experimentally showing the two cones and entangled regions (due to [2]). Interestingly, SPDC is actually quite an inefficient process: only about 1 in 1 billion photons is down converted.

where the first matrix corresponds to the action of a QP, the second a pair of orthogonally oriented barium borate crystals, which have nonlinear polarization responses that promote a process called spontaneous parametric downconversion (SPDC), by which an incident horizontally polarized photon becomes two vertically polarized photons of roughly half the energy of the original, and vice versa for an incident vertically polarized photon which becomes two horizontally polarized photons (see figure 3).

Because of this "optical meiosis", the two resulting photons are *entangled* with respect to energy, momentum, and polarization. Finally, S_0 describes the initial polarization state in matrix form, which here is oriented completely horizontally as it exits the laser. Using figure 1, we consider the two following setups:

$$\begin{aligned} P_C =& [Q(\alpha_A) \otimes Q(\alpha_B)\ H(\theta_B)]\ B\ QP(\phi)\ H(\theta_0)\ S_0 \\ P_I =& [Q(\alpha_{A_1})\ Q(\alpha_{A_2})\ H(\theta_B) \otimes Q(\alpha_{B_1})\ Q(\alpha_{B_2})\ H(\theta_B)] \\ & B\ Q(\alpha_1)Q(\alpha_0)\ H(\theta_0)\ S_0 \\ \therefore \rho_I =& P_I P_I^\dagger \\ \therefore \rho_C =& P_C P_C^\dagger, [3] \end{aligned} \tag{3}$$

where we use the tensor product to specify what happens to photon "A" and "B" separately (it is convention to call these two photons after "Alice" and "Bob"), and where ρ_I and ρ_C respectively are the density matrices given the setups described in equation (3) The result is a *2-qubit density matrix* ρ, which describes a 2-qubit quantum state and has the properties

1. $\mathrm{Tr}\,(\rho) = 1$
2. $\rho \geq 0$ (that is, all eigenvalues are nonnegative),
3. $\rho = \rho^\dagger$

II. METHODOLOGY

We are able to generate, using equation (3), an arbitrary quantum state given a set of angles. However, what if we wish to go the other way? That is, to start with some arbitrary state and *decompose* it into a product of Jones matrices? While Jones calculus is convenient in directly representing the action of a physical component on the polarization of an incoming state, these matrices are not pretty to work with if one attempts to multiply them out. Therefore, we can try a different approach, inspired by machine learning: we call this algorithm "InstaQ".

We define the closeness of two density matrices via Uhlmann-Jozsa fidelity, defined as

$$\mathcal{F}(\rho_1, \rho_2) = \left[\mathrm{Tr}\left(\sqrt{\sqrt{\rho_1}\rho_2\sqrt{\rho_1}}\right)\right]^2, \tag{4}$$

as given in [4]. We define the quantity

$$L = 1 - \sqrt{\mathcal{F}} \tag{5}$$

to be the *loss function*, since then we can take advantage of Python's `scipy.optimize.minimize` function to minimize a given scalar function. In our problem, we let $\rho_1 = \hat{\rho}$ and $\rho_2 = \rho_\mathrm{T}$, or the predicted and target density matrices respectively, where the target is the state we want to replicate. That is, this algorithm is designed to take as input some matrix, and determine the optimal angles to make L as low as possible so as to create as perfect a replica as possible. Then, one has this list of minimizing angles, which correspond to the ideal settings of the components in the lab.

However, equation (5) has either 5 or 9 dimensional domain depending on whether setup C or I is used, so running `scipy.optimize.minimize` once is in most cases not enough—it will have found a relative minimum. Therefore, we need to call this process many times in order to seed a variety of starting locations to maximize the chances of finding an optimal solution. We determine the number of times to execute the process by setting two limits: an overall limit to the number of iterations, as well as the parameter ϵ, which sets the maximum fidelity we wish to find. See section III for specific values and results.

The game then becomes how to determine the initial seed. We investigate three strategies: Random Hop, Random Fan, and Gradient Descent, outlined below.

A. Random Hop

This approach, given as PROGRAM HOP is the simplest, yet surprisingly effective.

B. Random Fan

This algorithm modifies the Random Hop by focusing its guesses around the current optimal angles, creating a

BEGIN PROGRAM HOP

while $n < N$ and $\mathcal{F}_{\max} < \epsilon$ **do**
Choose random x_0 array containing an initial guess of the angles. Minimize $L(x_0)$.
  if $\mathcal{F} > \mathcal{F}_{\max}$ **then** $\mathcal{F}_{\max} = \mathcal{F}, \mathbf{x}_{\text{best}} = x'$
  end if
  if $\mathcal{F}_{\max} > \epsilon$ **then** break.
  end if
Return $\mathcal{F}$ and x', the local optimal angles.
end while

"fan" as this window broadens if successive attempts do not further minimize loss.

BEGIN PROGRAM RANDOM FAN

while $n < N$ and $\mathcal{F}_{\max} < \epsilon$ **do**
Random guess so $\mathbf{x}_{\text{best}}$ not null. Minimize $L(x_0)$.
  Then choose random x_0 array by taking $(\mathbf{x}_{\text{best}} * (1 - \frac{mD}{fN}, \mathbf{x}_{\text{best}} * (1 + \frac{mD}{fN})$ as the range of random searching, where m is the number of attempts using this current $\mathbf{x}_{\text{best}}$ as the center of the random distribution, and f is the fraction of the total domain D on either side of $\mathbf{x}_{\text{best}}$ (default=0.1). Minimize $L(x_0)$.
  If $m > fN$, reset by choosing a new random seed.
  if $\mathcal{F} > \mathcal{F}_{\max}$ **then** $\mathcal{F}_{\max} = \mathcal{F}, \mathbf{x}_{\text{best}} = x'$
  end if
  if $\mathcal{F}_{\max} > \epsilon$ **then** break.
  end if
Return $\mathcal{F}$ and x', the local optimal angles.
end while

C. Gradient Descent

The most adaptive model architecture employs gradient descent, the steam engine of machine learning, in order to choose deliberately each new guess. Intuitively, the gradient is the measure of steepest ascent of a function, so if we swap a minus sign and move in the direction of the negative gradient, we will minimize loss most efficiently. Mathematically,

$$\mathbf{x}'_{\text{best}} = \mathbf{x}_{\text{best}} - \zeta \nabla L \tag{6}$$

where $\zeta \in (0, 1]$ is the learning rate, essentially how aggressively the model compensates at each iteration. Technically, the `scipy.optimize.minimize` algorithm mentioned earlier works by utilizing gradient descent–the idea is apply another layer of gradient descent to figure out what initial seeds should be given to then compute gradient descent all the way to the optimal loss.

We can perform estimate the gradient using `scipy.optimize.approx_fprime`, which is much more computationally efficient than trying to define a grid—essentially we take a small perturbation on the order of 10^{-6} for each of the inputs and measure the change.

We note that the algorithm also computes measurements in 12 bases (combinations of H and V, D and A, and R and L, where $|H\rangle$ refers to a horizontal state and $|V\rangle$ to a vertical, and $D = \frac{|H\rangle+|V\rangle}{\sqrt{2}}, A = \frac{|H\rangle-|V\rangle}{\sqrt{2}}$; $R = \frac{|H\rangle+i|V\rangle}{\sqrt{2}}, L = \frac{|H\rangle-i|V\rangle}{\sqrt{2}}$) via the joint-projection operator

Component	Domain
θ	$[0, \pi/4]$
α	$[0, \pi/2]$
ϕ	$[0, \pi/2]$

TABLE I: Angle domains.

$$P_{x,y} = \text{Tr}\left((\Pi_{\text{x}} \otimes \Pi_{\text{y}})\rho\right), \tag{7}$$

for $\Pi_j = |j\rangle\langle j|$—for both $\hat{\rho}$ and ρ_{T} so that the state can be verified experimentally via observed count rates with respect to different basis configurations. We follow the same structure as Hop, but replace the random angle with that updated with gradient descent.

III. RESULTS

We set $N = 1000, \epsilon = .999$. We restrict the domains of the angles to be optimized according to table I. Table II shows the results. In particular, we investigate several classes of states: the Bell states

$$|\Phi^+\rangle, |\Phi^-\rangle, |\Psi^+\rangle, |\Psi^-\rangle$$

for their ubiquity and theoretical and experimental interest as fully entangled states;

$$E_0 = \cos\eta|\Psi^+\rangle + e^{i\chi}\sin\eta|\Psi^-\rangle$$

and

$$E_1 = \frac{1}{\sqrt{2}}(\cos\eta(|\Psi^+\rangle + i|\Psi^+\rangle) + e^{i\chi}\sin\eta(|\Phi^+ + i|\Phi^-\rangle))\rangle$$

because of their experimental relevance in demonstrating the relevance of a certain class of entanglement witnesses (cite last summer's group), RS = random simplex, i.e. completely arbitrary pure 2-qubit state to probe the limits of the mode, J_C = states generated with the Jones framework on the C setup acting as a positive control for the C predictor, J_I = states generated with the Jones framework on the I setup acting as a positive control for the I predictor, and finally R = states generated according to the scheme in [5], restricted to solely mixed states (with an average purity of 0.74 and an SEM of 0.02) as a negative control for both setups—since the Jones matrix formalism can only yield pure states.

IV. DISCUSSION

Comparing the effectiveness of the Random Fan vs Random Hope methods, the results on the Bell states

State	Setup	Adapt	$\overline{n}$	$\sigma_{\overline{n}}$	$\overline{\mathcal{F}}$	$\sigma_{\overline{\mathcal{F}}}$
$\|\Phi^+\rangle$	C	1	72	8	0.99	0.01
	C	0	168	18	0.99	0.01
	I	1	85	7	0.99943	0.00003
	I	0	123	14	0.99	0.01
$\|\Phi^-\rangle$	C	1	155	14	0.99	0.01
	C	0	226	21	0.9996	0.00003
	I	1	109	10	0.99	0.01
	I	0	104	11	0.99950	0.00003
$\|\Psi^+\rangle$	C	1	118	10	0.99961	0.00003
	C	0	186	18	0.99963	0.00003
	I	1	56	7	0.99	0.01
	I	0	83	9	0.98	0.01
$\|\Psi^-\rangle$	C	1	102	10	0.99962	0.00003
	C	0	187	19	0.99	0.01
	I	1	61	6	0.98	0.01
	I	0	64	6	0.99	0.01
E_0	C	0	561	46	0.98	0.01
	I	0	627	36	0.99911	0.00006
E_1	C	0	886	27	0.98	0.003
	I	0	868	30	0.9973	0.0002
RS	C	0	996	17	0.937	0.006
	I	0	955	17	0.9963	0.0002
J_C	C	0	676	37	0.9987	0.0001
	I	0	940	21	0.9961	0.0003
J_I	C	0	992	7	0.933	0.006
	I	0	891	27	0.9966	0.0002
$R(0.74, 0.02)$	C	0	1000	0	0.78	0.01
	I	0	1000	0	0.83	0.01

TABLE II: Results of model: showing an average for 100 states per row. Note: Adapt=0 corresponds to Random Hop, 1 to Random Fan. $\overline{n}$ is the average number of iterations; $\sigma_{\overline{n}}$ is the standard error of the mean (SEM) number of iterations. $\overline{\mathcal{F}}$ is the average fidelity; $\sigma_{\overline{\mathcal{F}}}$ is the SEM for the mean fidelity. Note: full gradient descent calculations for all states have not yet been run because of the desire to tune the learning rate and random fraction values beforehand, which I have not yet had time to do.

suggest an improvement in the speed by a factor of 1.8 ± 0.3 for the Bell States, excluding$|\Phi^-\rangle$ with I setup. Both setups keep pace in the E_0 and E_1 states as well.

What's interesting is that this model can be used to test the difference between the set of possible states obtainable with the C vs I setups: we observe that while C passes its positive control with flying colors, it lowers to $\overline{\mathcal{F}} \approx 0.94$ for the I positive control, which also matches the figure for the RS states—whereas the I model *exceeds 0.99 average fidelity* without any special tricks beyond the Random Hop method. This is incredibly striking. These states that I can detect but C cannot are interesting follow-up states to try to understand what physically is unique with these states.

The negative control is also validated, as both models hover around $\overline{\mathcal{F}} \approx 0.8$ fidelity for the Roik (R) states.

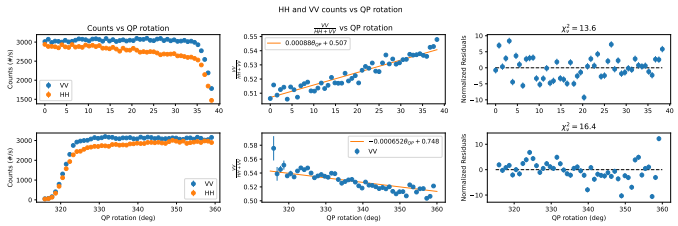

FIG. 4: Dependence of ratio of entangled photon pair production by BBO.

V. EXPERIMENTAL RESULTS

For lack of redundancy, a full discussion of experimental results in terms of making E_0 states to calculate entanglement witnesses is given in my Summer writeup. Here we briefly summarize the corrections to the theoretical argument above. First, the BBO crystals will not produce exactly equal amounts of HH and VV photons, so we replace the original matrix with one of the form

$$B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ a & 0 \end{bmatrix},$$

where $0 < a \leq 1$ is the ratio of VV to HH counts, as shown in Figure 4, which we determine by keeping the UVHWP fixed and sweeping the quartz plate.

Moreover, we need to determine the relationship between the actual angle of the quartz plate and the theoretical phase. Alec Roberson developed a theoretical expression for our actual current experimental setup, in which there is no additional quarter waveplate on Alice's Bob's path:

$$\phi = \arctan \frac{DL - DR}{RR - RL},$$

where DL is the counts in the DL basis, and so on. However, for a fully generalizable algorithm it is desirable to not have to rely on such an expression but to be able to determine it by performing sweeps. We attempted to perform this by sweeping (1) ϕ calculated vs UVHWP angle and (2) ϕ calculated vs QP angle. The idea was to vary only that component and keep the others fixed, and then use the InstaQ algorithm to determine ϕ. This approach was overall not terrible: outside a range of 10° to 40° for the UVHWP , ϕ depends greatly on the angle, which reflects the poorer performance of the a posteriori approach to determine ϕ via gradient descent. Moreover below $-22.5°$ for the QP, which was swept from $-36°$ to $-2°$ the a posteriori ϕ calculations are not very accurate. Part of this is due to the fact that, because the QP turns in the axis perpendicular to the laser beam, light will be refracted. Part of this is lalso ikely linked to the problem we presented in REFERENCE: namely, the current lack of a good way (and generalizable) to adjust the theoretical state of these generated states.

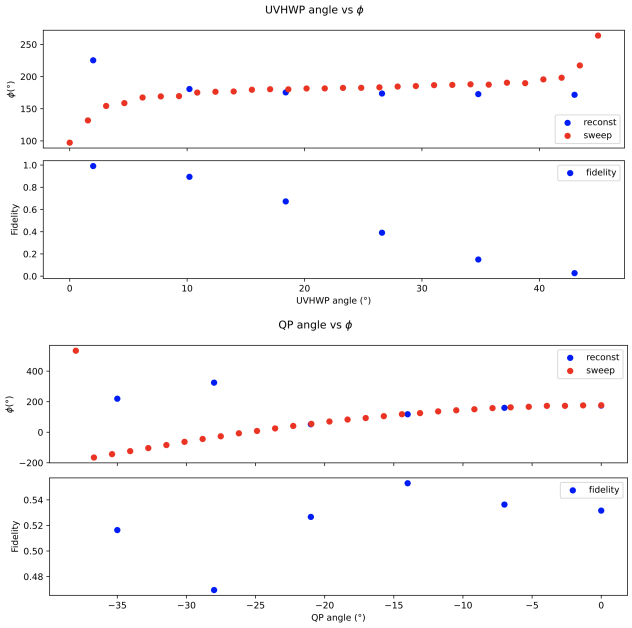

FIG. 5: Results of the effort to determine ϕ a posteriori in blue in comparison to the a priori method in red. The top figure was obtained by making $|\Psi^+\rangle$ but then sweeping UVHWP from 2 to 43° via (a) 30 iterations of 20 seconds each, measuring DL, DR, RR, and RL, and (b) by taking a full tomography and then performing gradient descent to find ϕ in the decomposition.

One clear solution would be to swap the QP with two quarter waveplates instead: we would not need to perform this sweep in that case. However, that would also cost more. Therefore, before this algorithm could be used reliably for research purposes, we would need to first effectively adjust the theoretical matrices to explain these experimental discrepancies.

VI. CONCLUSION

To our knowledge, nothing quite like this algorithm exists in the literature: [6] presents a pseudopolar decomposition of a Jones matrix into a sequence of matrix factors representing retardation and dichroic properties; the only other directly relevant work, by [7], decomposes experimentally determined Jones and Mueller matrices to obtain the polarizing and retardation characteristics of an optical system. However, we are decomposing density matrices into Jones matrices representing specific components in a quantum optics system. That said, there is still work to be done adapting to the experimental use case.

VII. ACKNOWLEDGEMENTS

We thank Professor Theresa Lynn and Alec Roberson for helpful comments, as well as Rob Fricke for discussions.

[1] R. Simon and N. Mukunda, Physics Letters A **143**, 165 (1990).
[2] C. Couteau, Contemporary Physics **59**, 291 (2018), arXiv:1809.00127 [physics, physics:quant-ph].
[3] We neglect the precompensation crystal (PCC) since in this theoretical setup the states have purity 1 ($\mathrm{Tr}\left(\rho^2\right)$).
[4] Y.-C. Liang, Y.-H. Yeh, P. E. M. F. Mendonça, R. Y. Teh, M. D. Reid, and P. D. Drummond, Reports on Progress in Physics **82**, 076001 (2019), publisher: IOP Publishing.
[5] J. Roik, K. Bartkiewicz, A. Černoch, and K. Lemr, Physical Review Applied **15**, 054006 (2021), publisher: American Physical Society.
[6] O. Arteaga and A. Canillas, JOSA A **26**, 783 (2009), publisher: Optica Publishing Group.
[7] F. L. Roy-Bréhonnet, B. L. Jeune, P. Eliès, J. Cariou, and J. Lotrian, Journal of Physics D: Applied Physics **29**, 34 (1996).