Continuous-variable quantum information processing in real and synthetic dimensions with self-configuring optics

Aviv Karnieli[1]*, Charles Roques-Carmes[1]*, Paul-Alexis Mor[1,2], Eran Lustig[1], Jamison Sloan[1], Jelena Vučković[1], David A. B. Miller[1] and Shanhui Fan[1]

[1]E. L. Ginzton Laboratory, Stanford University, 348 Via Pueblo, Stanford, CA 94305

[2]Physics Department, Ecole polytechnique, 91128 Palaiseau cedex, France

*Author e-mail address: karnieli@stanford.edu; chrc@stanford.edu; *equal contribution*

Abstract: We design self-configuring optical network architectures for continuous-variable quantum information processing of multimode squeezed vacuum, allowing scalable implementations in real and synthetic dimensions. © 2025 The Authors

Continuous variable (CV) quantum information serves as a promising paradigm for realizing scalable photonic quantum technologies. Multimode squeezed vacuum (MSV) states are a versatile CV platform for applications ranging from quantum computing to sensing and communication. Such states can be implemented across different degrees of freedom allowing very large dimensionality, such as the spatial [1], spectral and temporal [2-4] domains, where the corresponding squeezed quadratures are encoded in supermodes [5]. To process and measure MSV states, one usually employs balanced homodyne detection, optionally with a shaped local oscillator [5]. However, in the most general case where no *a priori* knowledge of the expected supermode structure can be made, the number of required homodyne measurements scales quadratically with the number of available modes M, even if squeezing is carried by only a few principal supermodes. This unfavorable "curse of dimensionality" can hamper the usage of this important quantum optical resource.

Efficient modal decompositions can be automatically performed using variational optimization in self-configuring arrays of Mach-Zehnder interferometers for coherent [6] and partially coherent [7] classical light, as well as for quantum optical processing of photon pairs [8]. In this work, we propose architectures for quantum information processing of MSV states based on a cascade of r self-configuring (SC) layers [6], in both real and synthetic dimensions. Our scheme allows the sequential learning of the r most dominant supermodes carrying the maximally-squeezed quadratures, performing the so-called Bloch-Messiah—Williamson decomposition [5] (diagonalization of the MSV covariance matrix in terms of physical modes) using $O(rM)$ iterations and in a hardware-efficient manner. Our findings pave the way towards scalable high-dimensional entanglement processing.

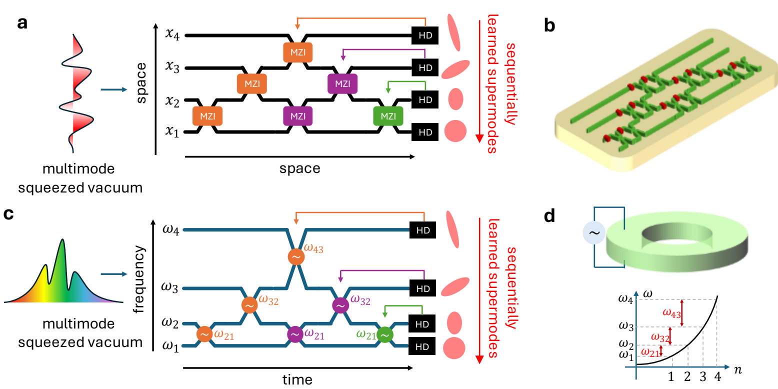

Fig 1: **CV quantum information processing with self-configuring optics in real and synthetic dimension**. The networks variationally optimize each SC layer (color-coded) using feedback from a single homodyne output measurement. The network learns the most dominant supermodes first. **a-b** Real space implementation involves an MSV state encoded in spatial bins, and the network is a MZI mesh implemented in an integrated photonics platform. **c-d** Implementation in the frequency-time synthetic dimension, involving an MSV state encoded in spectral bins and implemented using a single modulated cavity with quadratic dispersion.

Fig. 1 illustrates implementations in real space and in synthetic frequency-time dimension with a triangular mesh (other self-configuring architectures, such as a binary tree mesh [9] are also possible) of real or synthetic Mach-Zehnder interferometers (MZIs). The input MSV (a Gaussian state) is completely determined by its covariance matrix [5] $\Gamma_{\text{in}} = \langle \hat{\mathbf{q}}\hat{\mathbf{q}}^{\boldsymbol{T}} + (\hat{\mathbf{q}}\hat{\mathbf{q}}^{\boldsymbol{T}})^{\boldsymbol{T}} \rangle/2$ in terms of the quadrature operator vector $\hat{\mathbf{q}} = (\hat{x}_1, \ldots, \hat{x}_M, \hat{p}_1, \ldots, \hat{p}_M)^T$. The M modes of the MSV (spatial or frequency bins) are incident on the input ports of the network. The SC topology allows for the sequential optimization of the MZI parameters of each SC diagonal layer (color coded in Fig. 1a,c) through a single-mode homodyne measurement on its corresponding output, providing electronic feedback to the SC layer (color-coded arrows). The i-th output is optimized over the homodyne interferogram to find

$$\sigma_i^2 = \max_{||o^{(i)}||=1} o^{(i)T}\Gamma_{\text{in}} o^{(i)}, \qquad o_{\text{opt}}^{(i)} = \underset{||o^{(i)}||=1}{\text{argmax}}\, o^{(i)T}\Gamma_{\text{in}} o^{(i)}, \qquad (1)$$

where Γ_{in} is the input covariance matrix ($2M \times 2M$ symmetric real positive definite matrix; approximately-uniform losses across the modes are assumed), $o^{(i)}$ is the i-th column of O^T, and O is the real $2M \times 2M$ orthogonal matrix diagonalizing Γ_{in}, given in terms of the network complex unitary U as $O = (\text{Re}U, \text{Im}U; -\text{Im}U, \text{Re}U)$. Writing $\Gamma_{\text{in}} = O\Sigma^2 O^T$, the $i = 1, \dots, M$ ($i = M+1, \dots 2M$) -th diagonal elements of Σ^2 contain the variances of the antisqueezed (squeezed) quadratures, in descending (ascending) order of antisqueezing (squeezing), while the rows of O contain the structure of the $2M$ quadratures corresponding to the M supermodes. From the variational theorem, the sequential optimization of layer i learns the i-th most squeezed supermode $o_{\text{opt}}^{(i)}$ and its degree of (anti-)squeezing σ_i^2 .

So, this variational approach ensures that the network learns the most dominant supermodes *first*, allowing favorable scaling of both hardware and number of required iterations for the experimentally relevant scenario of sparsely encoded states of r supermodes in large Hilbert spaces ($r \ll M$). Specifically, the number of physical elements reduces from $O(M^2)$ of a full MZI mesh to $O(rM)$ (that is, one needs only $O(r)$ network layers). We note that the number of physical elements can potentially reach $O(1)$ for the synthetic dimension implementation discussed below. Further, it can be formally proven that the variational optimization of the Rayleigh quotient of Eq. (1) always converges using stochastic gradient descent [8] and requires $O(rM)$ iterations to implement the optimizing feedback. Numerical simulations of the optimization procedure, accounting for the homodyne measurement noise (modeled as Gaussian noise with the covariance matrix of the output state), are presented in Fig. 2.

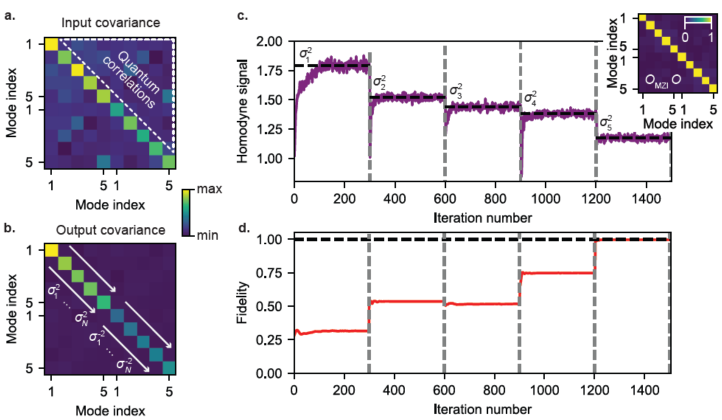

Fig 2: **Simulation of automatic Bloch-Messiah decomposition of multimode squeezed vacuum. a** Input covariance matrix for a $N = 5$ multimode squeezed vacuum (off-diagonal terms represent quantum correlations). **b** Output covariance matrix after self-configuring optics convergence. **c, d** Dynamics of self-configuring training: homodyne signal (**c**) and overall fidelity (**d**) as a function of iteration number. Inset shows that the product of the learned orthogonal matrix O_{MZI} (on x, p space) is orthogonal to the modal decomposition O corresponding to the Bloch-Messiah decomposition of the input state.

We now describe our proposed architecture using real and synthetic dimensions. In real space (Fig. 1a-b), an MZI mesh can be constructed using standard integrated photonics techniques [9] where squeezed vacuum is prepared either on-chip in a discrete-waveguide array [10] or using free-space couplers to couple onto the chip. The network's "width" and "depth" are both in the spatial domain. In contrast, one can implement such networks in a synthetic frequency dimension (Fig. 1c-d), using a single modulated resonator with a weak quadratic dispersion $\omega_n = \omega_0 + n\Omega_R + n^2\Omega_R'$ where Ω_R is the free spectral range and Ω_R' is the dispersion scale. Driving the system using uneven harmonics $\omega_{l,l-1} \equiv \Omega_R + (2l-1)\Omega_R' = \omega_l - \omega_{l-1}$, that couple adjacent frequencies ω_l, ω_{l-1}, a choice of $|\Omega_R'| \leq |\Omega_R|/3M$ ensures that the minimal detuning from any unwanted transitions in the system is $2|\Omega_R'|$. Therefore, for all $l = 1, \dots, M$, driving with Ω_l is equivalent to a frequency-domain MZI between the pair ω_l, ω_{l-1} if the modulation time satisfies $T \gg \pi/|\Omega_R'|$. The frequency-domain MZI meshes are then constructed fully in synthetic dimensions: the network "width" is the spectrum, and the "depth" is the time duration of the modulation. In this manner, the physical hardware complexity is $O(1)$ optical elements. We note that in this configuration, the squeezing is assumed to be generated in the same resonator (or an identical one coupled to it), and that the feedback homodyne measurements are done by controllably outcoupling the light after the total modulation time is over. We note that these ideas could also be extended to scattering-based architectures with the frequency domain MZIs [11]; these may still require $O(rN)$ physical elements, which is still a significant improvement to architectures proposed thus far.

References

[1] P. Sharapova et al, Phys. Rev. A **91**, 043816 (2015)
[2] Pfister, J. Phys. B: At. Mol. Opt. Phys. 53 012001 (2020)
[3] Guidry et al, Optica 10 694 (2023)
[4] Kouadou et al *APL Photonics* 8, 086113 (2023)
[5] Fabre and Treps Rev. Mod. Phys., 92, 035005 (2020)
[6] Miller, Photonics Research 1, 1 (2013).
[7] Roques-Carmes et al, *Light: Science & Applications* **13**, 260 (2024)
[8] Roques-Carmes et al, arXiv:2407.16849 (2024)
[9] Miller, Optica 7, 794, (2020).
[10] Arrazola et al *Nature* **591**, 54 (2021)
[11] Hu et al, *Nature* **599**, 587 (2021)