Bayesian tomography of high-dimensional on-chip biphoton frequency combs with randomized measurements

Owing in large part to the advent of integrated biphoton frequency combs, recent years have witnessed increased attention to quantum information processing in the frequency domain for its inherent high dimensionality and entanglement compatible with fiber-optic networks. Quantum state tomography of such states, however, has required complex and precise engineering of active frequency mixing operations, which are difficult to scale. To address these limitations, we propose a solution that employs a pulse shaper and electro-optic phase modulator to perform random operations instead of mixing in a prescribed manner. We successfully verify the entanglement and reconstruct the full density matrix of biphoton frequency combs generated from an on-chip Si$_3$N$_4$ microring resonator in up to an 8 × 8-dimensional two-qudit Hilbert space, the highest dimension to date for frequency bins. More generally, our employed Bayesian statistical model can be tailored to a variety of quantum systems with restricted measurement capabilities, forming an opportunistic tomographic framework that utilizes all available data in an optimal way.

Encoding $d$ levels of quantum information on single photons, known as photonic qudits\(^1\), offers crucial advantages for quantum communication and networking applications\(^2\), such as higher information capacities\(^3\), increased noise tolerance\(^4\), and stronger violations of Bell's inequalities\(^5\). Generation and manipulation of photonic qudits have been explored in many degrees of freedom, including path\(^6\), orbital angular momentum\(^7\), frequency bins\(^8\)–\(^11\), and time bins\(^12\)–\(^14\). Integrated photonics plays a pivotal role in scaling the complexity of quantum states\(^15\)–\(^17\) and quantum operations\(^18\), and the frequency degree of freedom is particularly attractive as on-chip biphoton frequency combs (BFCs) can produce a large number of spectrally entangled bins in a compact fashion.

Joint spectral intensity (JSI) measurements are commonly used to characterize BFCs, but such measurements are insensitive to phase coherence (and hence entanglement) across frequency-bin pairs. Thus, reconstruction of BFC density matrices has been realized through active mixing of frequency bins\(^11\)–\(^13\), such that measurements in multiple bases can be realized. In one method, by properly setting...
the amplitude and phase on a pulse shaper and the modulation voltage on a subsequent electro-optic phase modulator (EOM), one can filter out overlapping sidebands to perform projective frequency-bin measurements. Alternatively, a quantum frequency processor\textsuperscript{21} can be used to synthesize full quantum gates for tomography. Nevertheless, both methods face roadblocks en route to higher dimensions: aggressive amplitude filtering of the input state is inevitable in the first approach, while the number of elements required for arbitrary frequency qudit operations\textsuperscript{22} limits the maximum dimensionality possible with current technology. Accordingly, these existing methods are ill-suited to single-frequency electro-optic modulation: for them, the infinite Fourier series of Bessel functions produced by sinewave electro-optic modulation—a far cry from standard quantum bases—present a challenge to be overcome.

In this work, we instead leverage the complex mixing behavior of an EOM to our advantage as built-in randomized measurements for BFC characterization. By applying Bayesian tomographic techniques\textsuperscript{23,24}, we obtain complete state estimates for any dataset, including uncertainties commensurate with the data gathered, thus bolstering all results obtained from our measurement technique with a principled foundation. Importantly, these Bayesian features extend beyond the specific nuances of frequency-bin encoding to any quantum system, offering the promise of meaningful inference irrespective of whatever experimental constraints may have limited the measurements performed.

**Results**

**Experimental scheme and Bayesian analysis**

Figure 1a illustrates the experimental setup and concept behind our proposed scheme. The states of interest are BFCs with mode spacing \(\Delta\omega\) for the signal (idler); \(\omega_m\) is the phase of each pair. The integer \(B\) here denotes the number of signal (idler) bins at the center of the biphoton spectrum that are blocked by bandstop filters (omitted in Fig. 1a). Details regarding BFC state preparation can be found in the “Methods.”

The generated state is then directed to a pulse shaper and an EOM for the implementation of \(R_{\text{tot}}\) randomly chosen operations. For each operation, we apply a set of \(d\) random spectral phases \(\theta_m\) onto the signal (idler) bins with the pulse shaper. The spectral phases \((\theta_m, \phi_m)\) are uniformly sampled between 0 and \(2\pi\). The EOM is driven by a sinusoidal voltage with amplitude \(\delta\) and frequency equal to the mode spacing \(\Delta\omega\), imposing a temporal phase \(\exp[-i\delta \sin \Delta \omega t]\) onto each photon—equivalently introducing coupling between distinct frequency bins with weights given by Bessel functions of the first kind \(J_n(\delta)\). The strength of the imposed phase modulation \(\delta\) is selected from a set of \(R_{\text{rot}}\) values equispaced between \(\delta = 0\) and \(\delta = \delta_{\text{max}}\), with \(\delta_{\text{max}}\) set by the maximum radio-frequency (RF) power attainable at \(\Delta\omega\). Among these \(R_{\text{rot}}\) operations, we designate \(\delta = 0\) for the first measurement (i.e., the EOM turned off, making it a conventional JSI), while the modulation indices for the remaining \(R_{\text{rot}} - 1\) operations are chosen in a random order from the equispaced set, without replacement.

The photons are then passed to a wavelength-selective switch and we scan the filters to collect coincidences over the original \(d \times d\) frequency mode grid for each implemented operation, omitting any photons scattered outside of this computational space. Figure 1d shows examples of such measurements, simulated using multinomial statistics, corresponding to two different random operations for the cases of a classically correlated separable state row is used experimentally. c JSI of the output generated by the MRR. d Examples of JSI measurements simulated for the cases of both an entangled state and a classically correlated separable state for two different operations with \(\delta = 1.5\) rad (left) and \(\delta = 2\) rad (right) and randomly chosen spectral phases.
Ideal and retrieved density matrices for \(d = 4\) and \(d = 5\) are shown in Fig. 3. For \(R = 21\) measurements, \(d = 5\) is the optimal choice, as the density matrix resembles that of the ideal state. Even for \(R = 1\), the density matrix still has near-zero off-diagonal elements, indicating that the Bayesian estimation is still effective, albeit with more noise in the measurements.

In conclusion, Bayesian tomography provides a powerful framework for estimating quantum states, even in the presence of noise. The results show that the fidelity of the retrieved state increases with the number of random operations, reaching nearly ideal values at \(R = 21\). This demonstrates the effectiveness of Bayesian tomography in recovering quantum states from noisy measurements.
These results showcase the highest dimension of a fully reconstructed maximally entangled state and white noise lower- and upper-bound by CARs of 17 and 30, respectively. Square markers correspond to log-negativities of maximally entangled states, given by \( \log \delta \). Error bars denote the standard deviation computed from 1024 Bayesian samples.

Discussion

Technologically speaking, our approach aligns closely with recent ideas presented in EOM-based frequency-bin quantum random walks\(^\text{29}\), where here we preceede the walk with random spectral phases and consider varying circuit depths. Yet beyond the confines of frequency-bin encoding, our statistical treatment hints at the much wider value of Bayesian models in quantum information. Neither the choice of measurement settings nor number of datapoints has any bearing on the legitimacy of Bayesian tomography\(^\text{23,24}\); hence, Bayesian estimation will return a reasonable result for any dataset, with automatic uncertainty quantification indicating the confidence warranted from the data. This feature imparts Bayesian inference with unique flexibility compared to other advanced random measurement approaches, in that it does not assume, e.g., low-rank states\(^\text{30}\) or rely on unitary operations drawn from specific distributions\(^\text{31,32}\). Therefore, any quantum system for which an appropriate physical model can be constructed is ripe to potentially benefit from Bayesian models like the one presented here. Although not required conceptually, well-chosen measurements are practically valuable for obtaining final estimates accompanied by low uncertainties. Our experimental results show through example that the datasets from random modulation are more than sufficient to converge from an initially uniform (Bures) prior to density matrices with small error bars and in good agreement with the expected ground truth. Indeed, arguments from a simple theoretical model suggest that random EOM measurements with \( \delta_{\text{max}} \sim \mathcal{O}(d) \) efficiently cover the entire Hilbert space of a \( d \)-dimensional quantum system (see “Methods” for details), so that our measurement approach offers a straightforward path for high-dimensional frequency-bin characterization and should widen opportunities for BFCs in quantum information processing.

Methods

PPLN source setup

Our test source is a 2.1-cm-long fiber-pigtailed PPLN ridge waveguide (AdVR), possessing an internal efficiency of 150 \%/W for second-
harmonic generation and fiber-coupling efficiency of ~75% per facet. We couple a continuous-wave laser (Toptica) operated at ~5 mW and ~780 nm into the PPLN waveguide, temperature-controlled to 51.6 °C for SPDC under type-0 phase matching. Broadband spectrally entangled photon pairs spanning ~5 THz are generated and subsequently filtered by a tunable fiber-pigtailed Fabry-Perot etalon (Luna Innovations) with 20 GHz mode spacing and a full-width at half-maximum linewidth of 1 GHz. We carefully tune the etalon’s temperature to align its transmission peaks with the generated entangled photons, i.e., to maximize the coincidences between the symmetric, spectrally filtered mode pairs.

To minimize crosstalk in our 20 GHz-resolution demultiplexer (Finisar Waveshaper 4000S/X), we utilize the first pulse shaper (Finisar Waveshaper 1000S) to perform amplitude filtering (on top of the phase masks programmed for the QST) and block every other frequency bin, resulting in a 40 GHz-spaced BFC with a measured CAR of 90. A 300 GHz bandstop filter (corresponding to the case of B = 3) is programmed on the same pulse shaper in the middle of the BFC spectrum, which allows us to apply strong modulation on the EOM without the possibility of the signal photon jumping over into the idler’s modes, and vice versa. For coincidence measurements, we use a window of 128 ps and integration times of 20 s for all dimensions.

MRR source set up

The SiN$_x$ MRR used in our experiment is fabricated using the optimized photonic Damascene reflow process$^{35}$ with a cross-section of 2 μm × 0.95 μm. Such a process has enabled ultralow loss waveguides that have paved the way for dissipative Kerr solitons with free spectral ranges (FSRs) as low as 10 GHz$^{36}$. In the current work, the radius of the ring is 561 μm, corresponding to an FSR of 40.5 GHz—within the range of commercial EOMs—allowing us to drive the EOM at a frequency equal to the FSR for random operations. The gap between the ring and the bus waveguide is 0.3 μm, resulting in strong overcoupling with an intrinsic Q-factor of ~10$^7$ and a loaded Q-factor of ~10$^6$. We pump the ring using a tunable continuous-wave laser (New Focus) operating in the optical C-band, with an on-chip power of ~10 mW—well below the classical comb threshold of ~80 mW. The pump is amplified using an erbium-doped fiber amplifier (EDFA), which is subsequently filtered using a set of two 100 GHz-wide dense wavelength division multiplexing (DWDM) filters to suppress amplified spontaneous emission from the EDFA. Lensed fibers are used to couple the pump into the MRR, which is positioned on a temperature-controlled stage maintained at ~23 °C. The fiber-to-fiber coupling loss of the ring is around ~4 dB. Such a low loss was realized using engineered inverse waveguide tapers$^{34}$. When the pump is tuned into the ring resonance (~1550.5 nm) and operated below threshold for classical comb generation, it gives rise to BFC states. Since the pump and newly generated biphotons are in the same wavelength band, it is essential to suppress the residual pump after the ring to reduce accidents in coincidence measurements. We use a set of three 200 GHz-wide DWDM filters, which, when combined with bandstop filters in the pulse shaper and demultiplexer gives a net pump suppression of ~100 dB. We also tap a portion of the pump power to track its wavelength using a wavelength meter, setting up a computer-based feedback loop to ensure that the pump is operated at the intended resonance. For coincidence measurements, we use a window of 2048 ps (roughly equivalent to the inverse resonance linewidth) and integration time of 5 s for $d \in [2, 3, ..., 6]$. For $d \in [7, 8]$, we reduce the integration time to 3 s in the interest of minimizing the total experimental duration.

It is necessary to experimentally characterize the FSR of the MRR to determine the precise modulation frequency needed. In addition to this, we also characterize the residual spectral phase accumulated by the biphotons, likely due to fiber dispersion, to precompensate for it. For characterization of these quantities, the experimental setup remains the same as shown in Fig. 1a. We select two adjacent signal-idler bin pairs in the first pulse shaper and drive the EOM at a frequency $\omega_{RF}$ with amplitude $\delta = 1.43$ such that $|y_0(\delta)|=|y_1(\delta)|$ for equal mixing. We then pass one energy-matched pair of signal-idler frequency bins through the demultiplexer, now consisting of equal contributions from the adjacent bin due to phase modulation, and measure the biphoton time-correlation function. The theoretical expression for the coincidence rate, assuming both bins have identical Lorentzian lineshape and equal probability amplitude, is given by$^{35}$

$$R(t) \propto e^{-\gamma t^2} \left[ 1 - \cos \left( \phi \pm \phi_0 - (\omega_{FSR} - \omega_{RF}) \frac{t}{\tau} \right) \right]$$  

(1)

where $\gamma$ represents the Lorentzian linewidth, $\tau$ the signal-idler delay, $\phi$ the joint spectral phase between the bins applied by the pulse shaper, $\phi_0$ the residual biphoton phase, and $\omega_{FSR}$ the FSR. In Fig. 5a, we plot the coincidences at $\tau = 0$ as a function of $\phi$ for a set of two adjoining signal-idler bin pairs, where $\tau = 0$ is defined as the peak of the histogram in the unmodulated case. The offset of the experimentally obtained cosine function from the origin can be used to deduced the residual spectral phase between the bins. This process is repeated for all sets of adjoining signal-idler bin pairs used in our experiment to deduce the residual spectral phase of the entire biphoton state. For finding $\omega_{FSR}$, we again work with a set of two adjoining signal-idler bin pairs and operate at a spectral phase that minimizes the coincidences at $\tau = 0$. In Fig. 5b, we plot the coincidences integrated over $\tau$ as we sweep the applied RF frequency $\omega_{RF}$. When $\omega_{RF} = \omega_{FSR}$, the coincidences are minimized, so that Fig. 5b implies $\omega_{FSR} \Delta 2 \pi = 40.5$ GHz. The solid lines in Fig. 5 are theoretical estimates, scaled and vertically offset to match the datapoints via least squares. The vertical offset and the scaling account for both the nonzero accidents and unequal bin probability amplitudes. The width of the trace in Fig. 5b can be expressed as $2\gamma$.
from Eq. (1), from which $y/2\pi \approx 200$ MHz is inferred. These obtained resonator linewidth and FSR values are consistent with linear spectroscopic measurements.

**Bayesian inference model**

The ultimate goal of the inference process is to estimate the full BFC state, which can be represented as $d^6 \times d^6$ density matrix $\rho$. The frequency bins of the signal (idler) qudit possess annihilation operators $\hat{a}_k^{(s)} (\hat{a}_l^{(i)})$ where $k, l \in [1, \ldots, d]$ correspond to center frequencies of $\omega_k = \omega_0 + (k + B)\Delta\omega$ and $\omega_l = \omega_0 - (l + B)\Delta\omega$. The density matrix of interest can be formally written as

$$\rho = \sum_{k,l,\ell = 1}^{d} \rho_{kl\ell} (\hat{a}_k^{(s)}) \left(\hat{a}_l^{(i)}\right)^\dagger \text{vac} \langle\text{vac}|\hat{a}^{(s)}_{\ell}\hat{a}^{(i)}_{\ell}\rangle \text{vac},$$

(2)

where $|\text{vac}\rangle$ is the vacuum state.

We use the index $s$ to denote a specific measurement, which consists of a particular EOM and pulse shaper setting $r(s) \in [1, \ldots, R_{\text{tot}}]$, the signal frequency bin measured $m(s) \in [1, \ldots, d]$, and the idler bin measured $n(s) \in [1, \ldots, d]$. For notational convenience, the explicit dependence is suppressed in many of the formulas below, but remains implied for $r, m, n$. For a given $r$, the frequency bins of the signal photon undergo unitary transformations into the output modes via

$$b_m^{(s)} = \sum_{k = 1}^{d} W_{mk}^{(s)} \hat{a}_k^{(s)}$$

(3)

and similarly for idler photon we have

$$b_n^{(i)} = \sum_{l = 1}^{d} W_{nl}^{(i)} \hat{a}_l^{(i)}.$$  

(4)

The unitary operations $V^{(s)}$ and $W^{(i)}$ consist of line-by-line phase shifts on each of the input modes, followed by sinewave electro-optic modulation. The index $r$ defines the phase shifts and modulation index for a specific setting: $\theta_{m,l}^{(s,r)}$, $\phi_{m,l}^{(i,r)}$, $\eta_{m,l}^{(i,r)}$. We can therefore write

$$V_{mk}^{(s)} = J_{m-k}(\theta_{m,k}^{(s)}) e^{i\theta_{m,k}^{(s)}}$$

(5)

and

$$W_{nl}^{(i)} = J_{n-l}(\phi_{n,l}^{(i)}) e^{i\phi_{n,l}^{(i)}},$$

(6)

where $J_n(\cdot)$ is the Bessel function of the first kind, and for definiteness we have assumed a modulation function of the form $\exp[-i\theta^{(i)} \sin\Delta\omega t]$. The only major difference between the signal and idler equations is the input/output index reversal in the Bessel function order, which results from our definition of signal frequencies that increase with index and idler frequencies that decrease with index.

After these operations, we look for coincidences between bins $m$ and $n$, the probability of which can be computed for the input density matrix as

$$p_{r} = \langle m, n|\rho|m, n\rangle = \langle \text{vac}|b_m^{(s)} b_n^{(i)} \rho \left(b_m^{(s)}\right)^\dagger \left(b_n^{(i)}\right)^\dagger \text{vac}\rangle$$

(7)

using Eqs. (2)–(6) to simplify. This expression provides a linear mapping from the density matrix elements $\rho_{kl\ell}$ to each output probability $p_r$.

We then need to relate these probabilities to the observed coincidence counts $N_r$ through an appropriate likelihood function. In typical quantum tomographic contexts where the full Hilbert space is detected at each setting, a multinomial model is the most conceptually straightforward\cite{22} in our case, this would consist of products of the factors $p_{r}$. However, such a model does not readily apply to situations in which some of the outcomes are unmonitored. In our particular experiment, many bins outside of the original $d^6$-dimensional computational space can be populated, owing to the nonzero-Bessel function weights in Eq. (7). As a rule of thumb, the success probability (i.e., the possibility of a single photon staying in the original $d^6$-mode computational space) is roughly 1/2 when the modulation depth $\theta - 1/d\omega$ increases with $r$. Rather than attempting to measure all output mode combinations—which in principle involves an infinite-dimensional Hilbert space and in practice means the addition of many measurements with few counts—we focus only on the central $d \times d$ space here.

On the model side, we can account for unobserved outcomes by introducing an additional flux parameter $K$, defined as the average number of total coincidences that would be measured if all $(m, n)$ combinations were tested. Since the EOM and pulse shaper operations are unitary when considered over all modes—apart from an overall insertion loss that does not vary with setting $r$—this scale factor is fixed for all measurement settings. It also automatically accounts for efficiency; explicitly, it can be written as $K = N_{fl}/\Phi\Delta T$, where $N_{fl} (\Phi)$ is the total system efficiency from generation through detection for the signal (idler), $\Phi$ the photon pair generation flux, and $\Delta T$ the integration time. Although these quantities themselves could be inferred by considering both singles counts as well as coincidences\cite{22,23,24,25}, they are not of interest in the present investigation, and so through $K$ we are able to reduce to their combined effect only.

Thus, the mean number of coincidences for a specific setting $s$ becomes $Kp_r$, which we can model with a Poissonian distribution, the product of which produces the full likelihood $L_{\gamma}(\rho, K) \propto \text{Pr}(D|\rho, K)$

$$L_{\gamma}(\rho, K) = \prod_{r=1}^{R d^2} e^{-Kp_r} (Kp_r)^{N_r},$$

(8)

where $D = \{N_1, \ldots, N_{R d^2}\}$ denotes the set of measured coincidences for all $R d^2$ settings. This likelihood matches the form we adopted previously in the construction of JSIs (not the full quantum state) from random measurements\cite{22}. And not only does it readily handle probabilities that do not encompass the full Hilbert space, but it also more accurately reflects the physical situation of high-dimensional measurements with single-outcome detectors. For example, our use of raster scanning with two single-photon detectors means that, for any given configuration, we measure events only for a specific $(m, n)$ frequency-bin pair. Although one can pool the results for all pairs $(m, n)$ and view them synthetically as resulting from a single $\Delta T$ integration time of true $d^6$-outcome measurements—the situation assumed by a multinomial distribution—this ultimately does not align with the actual measurement procedure. Accordingly, the Poissonian likelihood provides both a practical and conceptually satisfying model for our tomographic scenario.

With the likelihood defining the relationship between a given density matrix and the observed data, Bayesian inference next requires specification of a suitable prior distribution for $\rho$ and $K$. In the case of QST, uniform priors are generally preferred, as these apply appreciable weights to all possible states and thereby minimally bias the final results. Although a variety of reasonable uninformative priors can be posited for density matrices, we select the Bures distribution, which enjoys a unique position as the single monotonic metric which reduces to both the Fisher and Fubini–Study metrics in the classical and pure
To work with the Bures ensemble, it is convenient to express any $d^2 \times d^2$ density matrix $\rho$ in the computational basis as

$$
\rho = \frac{(I_d + U)G G^\dagger (I_d + U^\dagger)}{\text{Tr}[(I_d + U)G G^\dagger (I_d + U^\dagger)]}, \quad (9)
$$

where the $d^2 \times d^2$ matrices $I_d$, $U$, and $G$ are the identity, a unitary matrix, and a general complex matrix, respectively. This expression automatically satisfies all physical conditions (unit-trace, hermiticity, and positive semidefiniteness); by sampling $U$ from the Haar distribution and $G$ from the Ginibre ensemble, the $\rho$ thus formed represents a single draw from the Bures distribution. We can in turn parameterize $\rho$ by the complex vector $y = (y_1, \ldots, y_{d^2})$, with each $y_k$ observing a complex standard normal distribution $y_k \sim \mathcal{C}(O, 1)$. $K$ of the components comprise the $d^2 \times d^2$ elements of the Ginibre matrix $G$ directly, while the remaining $d^2$ parameters form a second Ginibre matrix which is converted to the unitary $U$ through the Mezzadri algorithm, thereby ensuring Haar randomness.

In addition to the parameters forming $\rho$, the scale factor $K$ must also be suitably parameterized. Following ref. 39, we find it convenient to write $K = K_0(1 + \sigma)$, where $K_0$ and $\sigma$ are hyperparameters defined separate of the inference process, and $z$ is taken to follow a standard normal distribution $\mathcal{N}(O, 1)$, leading to a normal prior on $K$ of mean $K_0$ and standard deviation $K_0\sigma$. We take $\sigma = 0.1$ and $K_0$ equal to the sum of the counts in all $d^2$ bins for the first JSI measurement ($r = 1$), where the absence of modulation ensures that all initial photon flux remains in measured bins, i.e., $K_0 = \sum_{i=1}^{d^2} N_i$. This provides an effectively uniform prior, since a fractional deviation of 0.1 is much larger than the maximum amount of fractional uncertainty $1/\sqrt{K_0} = 0.02$ expected from statistical noise at our total count numbers; the use of a normal distribution simplifies the sampling process.

The total parameter set can therefore be expressed as the vector $x = (y, z) = (y_1, \ldots, y_{d^2}, z)$, with the prior distribution

$$
\pi_0(x) \propto \prod_{k=1}^{d^2} e^{-|y_k|^2} e^{-z^2}. \quad (10)
$$

We note that this parameterization entails a total of $4d^2 + 1$ independent real numbers ($2d^2$ complex parameters for $\rho$, one real parameter for $K$)—noticeably higher than the minimum of $d^2 - 1$ required to uniquely describe a density matrix. Nevertheless, this $\rho(y)$ parameterization is to our knowledge the only existing constructive method to produce Bures-distributed states, and is straightforward to implement given its reliance on independent normal parameters only.

Following Bayes’s rule, the posterior distribution becomes

$$
\pi(x) = \frac{1}{Z} L(x) \pi_0(x) \propto \left(\prod_{k=1}^{d^2} e^{-|y_k|^2} |K(z)|^2 \mathcal{N}(y_k)|^2 \prod_{k=1}^{d^2} e^{-|y_k|^2} e^{-z^2}. \quad (11)
$$

where $Z$ is a constant such that $\int dx \, \pi(x) = 1$. We have adopted this notation for Bayes’s theorem—rather than the more traditional $P(x|D) = P(D|x) P(x)/P(D)$—to emphasize the functional dependencies on $x$, which are all that must be accounted for in the sampling algorithm below. From $\pi(x)$, the Bayesian mean estimator $f_B$ of any quantity (scalar, vector, or matrix) expressible as a function of $x$ can be estimated as

$$
\hat{f}_B = \int dx \, \pi(x) f(x) = \frac{1}{S} \sum_{j=1}^{S} f(x^{(j)}), \quad (12)
$$

where, in lieu of direct integration, $S$ samples $\{x^{(1)}, \ldots, x^{(S)}\}$ are obtained from the distribution $\pi(x)$ through Markov chain Monte Carlo (MCMC) techniques, as described below.

**MCMC sampling**

Acquiring the samples necessary for computation of high-dimensional integrals of the form in Eq. (12) forms the primary bottleneck in Bayesian inference. The most common family of solutions to address this challenge fall under the general umbrella of MCMC, in which samples from a well-chosen Markov chain are designed to approach the statistics of $\pi(x)$ asymptotically. Recently, we applied a particularly efficient MCMC algorithm—known as preconditioned Crank-Nicolson (pCN) to the problem of Bayesian QST, finding significant computational improvements over previous implementations. We utilize the same pCN approach here; the only difference from the algorithm in ref. 24 is a simpler acceptance probability depending exclusively on the likelihood ratio, a consequence of the fact that all parameters here have normally distributed priors.

For each collection of measurement results considered in the main text, we retain $S = 2^9$ samples $\{x^{(0)}, \ldots, x^{(S-1)}\}$ from an MCMC chain of total length $S T$, where $T$ is a thinning factor that is successively doubled until convergence is obtained. From these samples, we then compute the Bayesian mean estimate of the density matrix

$$
\rho_B = \frac{1}{S} \sum_{j=1}^{S} \rho(x^{(j)}), \quad (13)
$$

each of which is plotted in Figs. 2–4. The mean and standard deviation of fidelity with respect to some ideal state $|\Psi_d\rangle$ can also be computed as

$$
\bar{F}_d = \frac{1}{S} \sum_{j=1}^{S} \langle \Psi_d | \rho(x^{(j)}) | \Psi_d \rangle, \quad (14)
$$

and

$$
\Delta F_d = \frac{1}{S} \sum_{j=1}^{S} \left( \langle \Psi_d | \rho(x^{(j)}) | \Psi_d \rangle \right)^2 \frac{1}{\bar{F}_d}, \quad (15)
$$

respectively. For the PPLN states specifically, which are not dispersion-compensated, we take the state for comparison as

$$
|\Psi_d\rangle = \frac{1}{\sqrt{d}} \sum_{m=1}^{d} |\psi_{12d^2}(m, m)\rangle, \quad (16)
$$

with $\beta_2 = 2.06 \times 10^{-3}$ ps$^2$ m$^{-1}$ for standard single-mode fiber and $L = 20$ m. The phases of the MRR states are precompensated, so for computing fidelities for them we use the uniform-phase version

$$
|\Psi_d\rangle = \frac{1}{\sqrt{d}} \sum_{m=1}^{d} |m, m\rangle. \quad (17)
$$

In order to monitor convergence with $T$, we consider the sequential fidelity, defined for a given $T = 2^n$ as

$$
F_{n,n-1} = \left( \text{Tr} \sqrt{\rho_B^{(n-1)} \rho_B^{(n)} \rho_B^{(n-1)}} \right)^2, \quad (18)
$$

where we use the notation $\rho_B^{(n)}$ to denote the Bayesian mean estimate [Eq. (12)] for a chain of length $S T = 5 \times 2^n$. For a sufficiently large thinning value $F_{n,n-1}$ will converge to unity and remain for all subsequent $n$. We note that this metric contains no reference to the ideal state $|\Psi_d\rangle$, but checks for consistency between subsequent Bayesian estimates only.
Figure 6 plots these sequential fidelities as a function of thinning \((n = \log_2 T)\) for all BFCs characterized in the main text. In all cases, we consider the full measurement sets, \(R = 21\) for the PPLN results and \(R = 30\) for the MRR findings, as these cases generally possess the slowest MCMC convergence. As expected, the sequential fidelity converges more rapidly with \(T\) for lower \(d\); nonetheless, all examples ultimately reach \(F_{n,n-1} > 0.99\) at their respective maximum of \(n\), indicating high MCMC convergence for all reported \(\rho\) matrices. Continuing to even larger values of \(T\) would certainly be desirable, particularly for \(d = 7\) and \(d = 8\), yet we are currently limited by computational power. For example, the \(d = 8\) MCMC chain with \(T = 2^{14}\) required almost a week to complete; with a total of 16,385 parameters to infer \((4d^2 + 1)\), it is no surprise that we are pushing the limits of our desktop computer. Indeed, to our knowledge the present results with \(d = 64\) represent a record-high Hilbert-space dimension for complete fidelity vs. thinning factor \(n\).

**Theoretical fidelities and ebits**

In our quantum system, several sources of noise, such as multipair emission and dark counts, are expected to be uniform across the BFC bins. Accordingly, as a simple model, we theoretically anticipate a ground truth quantum state of the form

\[
\rho_k = \lambda |\Psi_d\rangle \langle \Psi_d| + \frac{1 - \lambda}{d^2} I_d^d,
\]

where \(|\Psi_d\rangle\) is the ideal maximally entangled state [Eq. (16) or (17)], \(I_d^d\) is the \(d^2 \times d^2\) identity operator, and \(\lambda \in [0, 1]\) determines the noise level. In this section, we provide quantities of interest for \(\rho_k\) in order to compare against those found in Bayesian estimation, including CAR

\[
\text{CAR} = \frac{\max_{m,n} \langle m| \rho_k |m, n\rangle}{\min_{m,n} \langle m| \rho_k |m, n\rangle} = 1 + \frac{d \lambda}{(1 - \lambda)},
\]

fidelity \(F_d\)

\[
F_d = \frac{\langle \Psi_d | \rho_k | \Psi_d \rangle}{\langle \Psi_d | \Psi_d \rangle} = \frac{(d^2 - 1)\lambda + 1}{d^2},
\]

and log-negativity \(E_d\) (adapted from ref. 28)

\[
E_d = \begin{cases} 
0 & : F_d < 1/d \\
\log_2 d F_d & : F_d > 1/d
\end{cases}
\]

The fidelities and log-negativities reported for \(\rho_k\) in the main text are calculated using these equations. In general, both fidelity and \(E_d\) deviate more strongly from their respective ideals as \(d\) increases for a fixed CAR. The good agreement in the main text between the Bayesian-estimated states and this simple white noise model suggests that our understanding of noise processes in the system is well justified.

**Theoretical analysis of measurement efficiency**

In this section, we provide a basic theoretical analysis of the effectiveness of our proposed tomographic method. We first consider a highly simplified model that captures the key aspects of how phase modulation and mode mixing enable tomographic reconstruction, concentrating in this case on a single frequency-bin qubit occupying \(d\) modes \(x \in [1, \ldots, d]\). While the experiments in the main text examine the joint state of two qudits instead, this simpler single-qubit case reveals the basic principles of the tomography method with minimal distractions. The only type of measurement considered is projection onto the photon’s output mode, reflecting the condition of frequency-resolved detection; the probability of obtaining bin \(x\) is \(p_{x}^{(0)} = p_{xx}\), and the qudit assumption ensures that \(p_x = 0\) for \(x < 1\) or \(x > d\).

Consider the mode mixing operator \(S_k\) which weakly mixes each mode \(x\) with modes \(x + k\) and \(x - k\):

\[
S_k |x\rangle = -\epsilon |x - k\rangle + |x\rangle + \epsilon |x + k\rangle.
\]

For \(k = 1\), this approximates the action of an EOM with small modulation depth \(\delta = 2\epsilon\) and modulation frequency equal to the fundamental mode spacing. More generally, for small \(\epsilon\), \(S_k\) is approximately unitary and describes a weak, translation-invariant mode mixing operation (Fig. 7a).

Suppose we apply \(S_k\) to the photon and measure its bin. The probability of observing bin \(x\) is

\[
p_{x}^{(k)} = \langle x| S_k \rho S_k^\dagger |x\rangle.
\]
Fig. 7 | Theoretical analysis of measurement approach. a Effect of the mode-mixing operator $S_k$ [Eq. (23)] on the density operator $\rho$. b Behavior of the Bessel functions, which describe the modulation operators $T_k$ [Eq. (31)]. c The modulation operator $T_\delta$ with $\delta = k$ probes a $k$-band of $\rho$, i.e., the elements at most $k$ above or below the diagonal. d Distribution of singular values for random measurement matrices $O$ for $d = 8$, $\delta \in [0, \delta_{\text{max}}]$, and $\mathcal{R} = 2d$: $\delta_{\text{max}} = 4$ (gold); $\delta_{\text{max}} = 8$ (blue); $\delta_{\text{max}} = 16$ (violet).

Using $S_k^j|x\rangle = \epsilon|x\rangle + |x\rangle - \epsilon|x+k\rangle$ and the fact that $\rho_{xx} = \rho_{xy}^*$ we obtain

$$p^{(k)}_x = (\epsilon|x-k\rangle + |x\rangle - \epsilon|x+k\rangle)\rho (|x-k\rangle + |x\rangle - \epsilon|x+k\rangle)$$

$$= \rho_{xx} + \epsilon (\rho_{x-k,x} + \rho_{x-k,x} - \rho_{x+k,x} - \rho_{x+k,x} + C(\epsilon^2)$$

$$= \rho_{xx} + 2\epsilon (\rho_{x-k,x} - \rho_{x+k,x} + C(\epsilon^2))$$  \hspace{1cm} (25)

Now, from the fact that $|p^{(k)}_x|^2 \leq \rho_{xx}\rho_{yy}$ and the fact that the photon was initially restricted to modes $[1, \ldots, d]$ we have that the support of $\rho_{x+k}$ is $x \in [1, \ldots, d-k]$ and the support of $\rho_{x-k}$ is $x \in [k+1, \ldots, d]$. Considering just the outcomes in bins $[1, \ldots, d]$ yields the approximate system of equations

$$-\epsilon \rho_{1+k} = \frac{1}{2\epsilon} (p^{(k)}_1 - p^{(0)}_1)$$

$$-\epsilon \rho_{2+k} = \frac{1}{2\epsilon} (p^{(k)}_2 - p^{(0)}_2)$$

$$\rho_{1+k+1} - \epsilon \rho_{1+k+2} = \frac{1}{2\epsilon} (p^{(k)}_{1+k+1} - p^{(0)}_{1+k+1})$$

$$\rho_{2+k+1} - \epsilon \rho_{2+k+2} = \frac{1}{2\epsilon} (p^{(k)}_{2+k+1} - p^{(0)}_{2+k+1})$$

$$\rho_{d-2+k,d-k} - \epsilon \rho_{d-2+k,d-k} = \frac{1}{2\epsilon} (p^{(k)}_{d-2+k,d-k} - p^{(0)}_{d-2+k,d-k})$$

$$\rho_{d-2+k+1,d-k+1} = \frac{1}{2\epsilon} (p^{(k)}_{d-2+k+1,d-k+1} - p^{(0)}_{d-2+k+1,d-k+1})$$

$$\rho_{d-2+k,d-k+2} = \frac{1}{2\epsilon} (p^{(k)}_{d-2+k,d-k+2} - p^{(0)}_{d-2+k,d-k+2})$$  \hspace{1cm} (26)

valid for $k \in [1, \ldots, d-1]$ and small $\epsilon$. In principle, this (generally overdetermined) system of $d$ equations enables one to determine the real parts of $\rho_{1,k+1}, \ldots, \rho_{d-k,d}$, the nonzero part of the $k$th diagonal of $\rho$. (Note that the $-k$th diagonal of $\rho$ is just the conjugate of the $k$th diagonal.) The elements on diagonals $[-k, \ldots, 0, \ldots, k]$ will be collectively referred to as the $k$-band.

To determine the imaginary parts of the $k$th diagonal, each pair of components $\rho_{xx}$ and $\rho_{x+k,x}$ must be mixed with a relative phase of $\pi/2$. Let $\Phi_k$ be the operation

$$\Phi_k|x\rangle = \omega_k|x\rangle$$  \hspace{1cm} (27)

where $\omega_k = e^{i\pi k}$ is the $k$th principle root of $i$. Suppose we apply $\Phi_k$ and then $S_k$ to the photon and measure its mode. The probability of observing bin $x$ in this case is

$$p_x^{(k)} = \langle x|S_k\Phi_k\rho \Phi_k^\dagger S_k^\dagger|x\rangle$$

$$= \langle x|S_k\rho S_k^\dagger|x\rangle$$  \hspace{1cm} (28)

where $\rho = \Phi_k \rho \Phi_k^\dagger$. Using $\rho_{xy} = \omega_k^{-1} \rho_{xy}$ we have

$$p_x^{(k)} = (\epsilon|x-k\rangle + |x\rangle - \epsilon|x+k\rangle)\rho (|x-k\rangle + |x\rangle - \epsilon|x+k\rangle)$$

$$= \rho_{xx} + \epsilon (\rho_{x-k,x} + \rho_{x-k,x} - \rho_{x+k,x} - \rho_{x+k,x} + C(\epsilon^2)$$

$$= \rho_{xx} + \epsilon (\rho_{x-k,x} + \rho_{x+k,x} - \rho_{x+k,x} + \rho_{x+k,x}) + C(\epsilon^2).$$  \hspace{1cm} (29)

Since $\omega_k^k = \pm i$ we have

$$p_x^{(k)} = \rho_{xx} + \epsilon (\rho_{x-k,x} + \rho_{x+k,x} - \rho_{x+k,x} + \rho_{x+k,x}) + C(\epsilon^2)$$

$$= \rho_{xx} + 2\epsilon (\rho_{x-k,x} - \rho_{x+k,x}) + C(\epsilon^2).$$  \hspace{1cm} (30)

Using the probabilities of bins $1$ through $d$ we obtain a system of $d$ equations which, together with $p^{(0)}$, in principle enables determination of the imaginary parts of the $k$th diagonal.

To summarize: $p^{(0)}$ determines the diagonal elements of $\rho$, then $p^{(k)}$ and $p^{(k)}$ (obtained by applying $S_k$ and $S_k \Phi_k$, respectively) determine the off-diagonal elements of $\rho$ that lie $k$ positions above and below the diagonal. Thus, all $d^2$ elements of the single-qudit density matrix can be
probled by \(1 + 2(d-1) = 2d-1\) experimental settings with \(d\) outcomes each.

To relate this to the experimental approach, we note that the EOM effects a mode mixing operation \(T_\delta\) of the form [cf. Eqs. (5) and (6)]

\[
T_\delta |k\rangle = \sum_{k=-\infty}^{\infty} f_k(\delta) |k + k\rangle
\]  

(31)

As shown in Fig. 7b, \(f_k(\delta)\) is oscillatory in \(k\) with approximate support \(k \in [-[\delta], \ldots, [\delta]]\). That is, \(T_\delta\) is approximately a linear combination of \([S_\delta, \ldots, S_{[\delta]}]\), mixing each \(x\) with multiple modes ranging from roughly \(x - [\delta]\) to \(x + [\delta]\) and yielding information about the elements in the \([\delta]\)-band of \(\rho\) (Fig. 7c). If the elements in the \((-\delta, -\delta)\)-band have already been determined, \(T_\delta\) provides new information primarily concerning the \([\delta]\)-band. In our specific experiments, we applied a set of random phases with the pulse shaper in addition to the modulation, allowing us to probe linear combinations of both the real and imaginary parts of the operators.

To relate this to the experimental approach, we note that the EOM provides new information primarily concerning the \([\delta]\)-band of \(\rho\) (Fig. 7c) without any improvement to mixing within the space. Our numerical findings therefore join the simple theory above in suggesting a maximum modulation index of \(\delta_{\text{max}}\). Nevertheless, we note that all cases in Fig. 7c sample the Hilbert space comprehensively, so that any \(\delta_{\text{max}} - \Omega(d)\) is likely to prove sufficient in a tomographic context.

As examples, we compute singular values for a variety of settings at \(d = 8\): the maximum single-qudit dimension characterized experimentally. We consider \(R = 2d = 16\) settings, which we found sufficient to obtain robust and repeatable distributions. Each measurement configuration involves a modulation index \(\delta\) chosen uniformly at random in the interval \([0, \delta_{\text{max}}]\), preceded by a random phase vector \(\phi \in [0, 2\pi]^d\) applied by the pulse shaper. Figure 7d plots histograms of singular values from 2,000 measurement matrices \(\delta\) each for \(\delta_{\text{max}} = 4, 8, 16\), corresponding to \(d/2, d,\) and \(2d = 8\). A narrow peak around \(\log_{10}s = 0\) indicates high and comparable sensitivity for all elements in the density matrix; indeed, a complete set of mutually unbiased bases—an ideal choice for tomography—possesses \(d^2\) equal singular values at \(s = 1\). By this criterion, \(\delta_{\text{max}} = d = 8\) is seen to provide the most efficient measurement distribution of the three examples. For the smaller index of \(\delta_{\text{max}} = 4\), the main peak is accompanied by a strong tail indicating small sensitivity to an appreciable percentage of the matrix elements. For the larger index \(\delta_{\text{max}} = 16\), the main peak shifts to lower values, which reflects “over-modulating” of the quantum state; taking \(\delta_{\text{max}}\) beyond \(d\) increases the probability of scattering the input outside of the computational space (see Fig. 7c) without any improvement to mixing within the space. Our numerical findings therefore join the simple theory above in suggesting a maximum modulation index of \(\delta_{\text{max}}\). Nevertheless, we note that all cases in Fig. 7c sample the Hilbert space comprehensively, so that any \(\delta_{\text{max}} - \Omega(d)\) is likely to prove sufficient in a tomographic context. Ultimately, we emphasize that the theory developed here is meant to provide heuristic guidelines for implementing EOM-based frequency-bin tomography, not to imply the optimality of our specific measurements. For example, it is quite possible that alternative distributions for \(\delta\)—e.g., other than the uniform draw \(\delta \in [0, \delta_{\text{max}}]\)—may show more favorable properties with further research. Nevertheless, our reliance on Bayesian estimation ensures that these questions need not be answered for useful inference.

Extending to the experimental task of two-qudit characterization, our model thus indicates one should consider joint measurements of the form \(T_\delta \otimes T_\delta\), with \(\theta_1, \theta_2 \in [0, d]\). Considering our experimental values of \(\delta_{\text{max}} = 2.5\) for the PPLN tests and 3.4 for the MRR, \(\delta_{\text{max}}\) falls in the range of 0.425\(d\) and 1.7\(d\) for all qudit dimensions examined, aligning well with the \(\Omega(d)\) desideratum. Since the entire density matrix of a \(d^2\)-dimensional two-qudit state requires specification of \(d^2 - 1\) real parameters and each experimental pulse shaper/EOM setting \(r\) provides \(d^2\) outcomes, \(R = \Omega(d^4)\) would be expected to be required to fully probe the two-qudit Hilbert space. Empirically, however, we were able to attain low-error state reconstruction with fewer measurements: e.g., \(R = 10\) instead of 25 for the \(d = 5\) PPLN tests (Fig. 2a) and \(R = 30\) instead of 64 for the \(d = 8\) MRR tests (Fig. 2b). Several aspects are likely responsible for this reduction. From a broad perspective, the positive semidefiniteness of the density matrix imposes additional constraints that are not reflected in a linear system analysis but automatically accounted for in the Bayesian inference procedure. Thus, simply comparing the number of measurements with the number of parameters generally yields an overly pessimistic assessment of the information required for tomography.

Moreover, we suspect that specific features of our quantum state also contribute to more efficient reconstruction. The highly correlated nature revealed in the first JSI measurement implies that only off-diagonal elements of the form \(p_{xxy}(x/k, x/k, x/k)\) are determined by biphoton energy conservation—can be significantly different than zero. This reduces the number of appreciable elements in our density matrix from \(O(d^4)\) to \(O(d^2)\). Of course, the other off-diagonal density matrix elements are not strictly zero in practice, and the Bayesian model neither requires nor assumes such a simplification. However, the strong frequency correlations do eliminate a large portion of the Hilbert space from consideration. Consequently, future experiments with more general states may
require more experimental settings for low-uncertainty estimation than we have currently used.

Data availability
The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Code availability
Analysis code used in this study is available from the corresponding author on request.

References
1. Erhard, M., Krenn, M. & Zeilinger, A. Advances in high-dimensional quantum communication: benefits, progress, and future challenges. Adv. Quantum Technol. 2, 1900038 (2019).
2. Barreiro, J. T., Wei, T.-C. & Kwiat, P. G. Beating the channel capacity limit for linear photonic superdense coding. Nat. Phys. 4, 282 (2008).
3. Cerf, N. J., Bourennane, M., Karlsson, A. & Gisin, N. Security of quantum key distribution using d-level systems. Phys. Rev. Lett. 88, 127902 (2002).
4. Ecker, S. et al. Overcoming noise in entanglement distribution. Phys. Rev. X 9, 041042 (2019).
5. Vértesi, T., Pirion, S. & Brunner, N. Closing the detection loophole in Bell experiments using qudits. Phys. Rev. Lett. 104, 060401 (2010).
6. Wang, J. et al. Multidimensional quantum entanglement with large-scale integrated optics. Science 360, 285 (2018).
7. Qiang, X. et al. Large-scale silicon quantum photonics implementing arbitrary two-qubit processing. Nat. Photonics 12, 534 (2018).
8. Kues, M. et al. Quantum optical microcombs. Nat. Photonics 13, 170 (2019).
9. Lu, H.-H., Simmerman, E. M., Lougovski, P., Weiner, A. M. & Lukens, J. M. Fully arbitrary control of frequency-bin qudits. Phys. Rev. Lett. 125, 120503 (2020).
10. Martin, A. et al. Quantifying photonic high-dimensional entanglement. Phys. Rev. Lett. 118, 110501 (2017).
11. Ikuta, T. & Takesue, H. Implementation of quantum state tomography for time-bin qudits. New J. Phys. 19, 013039 (2017).
12. Moody, G., Chang, L., Steiner, T. J. & Bowers, J. E. Chip-scale nonlinear photonics for quantum light generation. AVS Quantum Sci. 2, 041702 (2020).
13. Yang, Z. et al. A squeezed quantum microcomb on a chip. Nat. Commun. 12, 4781 (2021).
14. Elshaari, A. W., Pernice, W., Sriwirawan, K., Benso, O. & Zwiller, V. Hybrid integrated quantum photonic circuits. Nat. Photonics 14, 285 (2020).
15. Lu, H.-H. et al. Quantum interference and correlation control of frequency-bin qudits. Optica 5, 1455 (2018).
16. Ikuta, T. & Takesue, H. Implementation of quantum state tomography for time-bin qudits. New J. Phys. 19, 013039 (2017).
17. Moody, G., Chang, L., Steiner, T. J. & Bowers, J. E. Chip-scale nonlinear photonics for quantum light generation. AVS Quantum Sci. 2, 041702 (2020).
18. Yang, Z. et al. A squeezed quantum microcomb on a chip. Nat. Commun. 12, 4781 (2021).
19. Elshaari, A. W., Pernice, W., Sriwirawan, K., Benso, O. & Zwiller, V. Hybrid integrated quantum photonic circuits. Nat. Photonics 14, 285 (2020).
20. Lu, H.-H. et al. Quantum interference and correlation control of frequency-bin qudits. Optica 5, 1455 (2018).
21. Lu, H.-H., Weiner, A. M., Lougovski, P. & Lukens, J. M. Quantum information processing with frequency-comb qudits. IEEE Photon. Technol. Lett. 31, 1858 (2019).
22. Lukens, J. M. et al. All-optical frequency processor for networking applications. J. Light. Technol. 38, 1678 (2020).
23. Blume-Kohout, R. Optimal, reliable estimation of quantum states. New J. Phys. 12, 043034 (2010).
24. Lukens, J. M., Law, K. J. H., Jasra, A. & Lougovski, P. A practical and efficient approach for Bayesian quantum state estimation. New J. Phys. 22, 063038 (2020).
25. Liu, J. et al. Photonic microwave generation in the x- and k-band using integrated soliton microcombs. Nat. Photonics 14, 486 (2020).
26. Robert, C. P. & Casella, G. in Monte Carlo Statistical Methods Ch. 1 (Springer, 1999).
27. Peres, A. Separability criterion for density matrices. Phys. Rev. Lett. 77, 1413 (1996).
28. Vidal, G. & Werner, R. F. Computable measure of entanglement. Phys. Rev. A 65, 032314 (2002).
29. I many, P., Lingaraju, N. B., Alshaykh, M. S., Leaird, D. E. & Weiner, A. M. Probing quantum walks through coherent control of high-dimensionally entangled photons. Sci. Adv. 6, eaba8066 (2020).
30. Gross, D., Lu, Y.-K., Flammia, S. T., Becker, S. & Eisert, J. Quantum state tomography via compressed sensing. Phys. Rev. Lett. 105, 010504 (2010).
31. Brydges, T. et al. Probing Rényi entanglement entropy via randomized measurements. Science 364, 280 (2019).
32. Huang, H.-Y., Kueng, R. & Preskill, J. Predicting many properties of a quantum system from very few measurements. Nat. Phys. 16, 1050 (2020).
33. Liu, J. et al. High-yield, wafer-scale fabrication of ultralow-loss, dispersion-engineered silicon nitride photonic circuits. Nat. Commun. 12, 1 (2021).
34. Liu, J. et al. Double inverse nanotapers for efficient light coupling to integrated photonic devices. Opt. Lett. 43, 3200 (2018).
35. Myllysamy, K. V. et al. Time-resolved detection of phase-coherent biphon frequency combs from SiN₄ microring, in CLEO JMF3.5 (Optical Society of America, 2021).
36. Lu, H.-H. et al. Electro-optic frequency beam splitters and tritters for high-fidelity quantum information processing. Phys. Rev. Lett. 120, 030502 (2018).
37. Williams, B. P. & Lougovski, P. Quantum state estimation when qubits are lost: a no-data-left-behind approach. New J. Phys. 19, 043003 (2017).
38. Lu, H.-H. et al. A controlled-NOT gate for frequency-bin qudits. npj Quantum Inf. 5, 24 (2019).
39. Simmerman, E. M., Lu, H.-H., Weiner, A. M. & Lukens, J. M. Efficient compressive and Bayesian characterization of biphon frequency spectra. Opt. Lett. 45, 2886 (2020).
40. Sommers, H.-J. & Życzkowski, K. Bures volume of the set of mixed quantum states. J. Phys. A: Math. Gen. 36, 10083 (2003).
41. Osipov, V. A., Sommers, H.-J. & Życzkowski, K. Random Bures mixed states and the distribution of their purity. J. Phys. A: Math. Theor. 43, 055302 (2010).
42. Mezzadri, F. How to generate random matrices from the classical compact groups. Not. Am. Math. Soc. 54, 592 (2007).
43. MacKay, D. J. C. in Information Theory, Inference, and Learning Algorithms Ch. 4 (Cambridge University Press, 2003).
44. Cotter, S. L., Roberts, G. O., Stuart, A. M. & White, D. MCMC methods for functions: modifying old algorithms to make them faster. Statist. Sci. 28, 424 (2013).
45. Jacob, P. E., O’Leary, J. & Atchadé, Y. F. Unbiased Markov chain Monte Carlo methods with couplings. J. R. Statist. Soc. B 82, 543 (2020).
46. Wootters, W. K. & Fields, B. D. Optimal state-determination by mutually unbiased measurements. Ann. Phys. 191, 363 (1989).

Acknowledgements
We thank AdR for loaning the PPLN ridge waveguide; P. Imany, N.B. Lingaraju, and A.J. Moore for valuable discussions; A.A.N. Ovi for laboratory help; and B.T. Kirby for introducing us to the Bures distribution. This work was performed in part at Oak Ridge National Laboratory, operated by UT-Battelle for the U.S. Department of Energy under contract no. DE-AC05-00OR22725. Funding was provided by the U.S. Department of Energy, Office of Advanced Scientific Computing Research, Early Career Research Program (Field Work Proposal ERKJ353), the National Science Foundation (1839191-ECCS, 2034019-ECCS), the Air Force Office of Scientific Research (Award no. FA9550-19-1-0250), and the Swiss National Science Foundation under Grant no. 176563 (BRIDGE). K.V.M. acknowledges support from the QISE-NET fellowship program of the National Science Foundation (DMR-1747426). M.S.A. acknowledges support from the College of Engineering Research Center at King Saud University. The Si3N4 samples were fabricated in the EPFL Center of MicroNanoTechnology (CMI).

Author contributions
H.H.L. and K.V.M. contributed equally to this work. H.H.L. initiated the concepts, led the experiments, and contributed to the theoretical analysis. K.V.M. performed experiments with MRR sources and led writing of the paper. S.S., M.S.A., and D.E.L. assisted the experiments. R.S.B. contributed to the theoretical analysis. J.L. designed and fabricated the integrated device, supervised by T.J.K. A.M.W. supervised the work at Purdue and assisted with system design and analysis. J.M.L. managed the theoretical analysis and numerical simulation, and supervised the project. All authors reviewed the results and contributed to the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-022-31639-z.

Correspondence and requests for materials should be addressed to Hsuan-Hao Lu, Karthik V. Myilswamy or Joseph M. Lukens.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2022