Supplementary Material: Scalability and performance metrics for boson sampling devices

Dong-Gil Im,[1,*,†] Chung-Hyun Lee,[2,†] Joonsuk Huh,[3,4,5,‡] and Yoon-Ho Kim[2,§]

[1] *Agency for Defense Development, Daejeon, 34186, Korea*
[2] *Department of Physics, Pohang University of Science and Technology (POSTECH), Pohang, 37673, Korea*
[3] *Department of Chemistry, Yonsei University, Seoul 03722, Korea*
[4] *Department of Quantum Information, Yonsei University, Incheon 21983, Korea*
[5] *Department of Computational Science and Engineering, Yonsei University, Seoul 03722, Korea*

(Dated: October 7, 2026)

I. Mode mismatch induced background and Effective internal mode model

A. Mode mismatch induced background for partially populated inputs

We consider an M-mode interferometer in which only N input modes are populated. The distinguishable fraction $1-p_{ind}$ does not contribute coherently to the intended multiphoton interference, but these photons can still propagate through the interferometer and produce detector clicks.

For output mode j, we define the total transfer weight from the populated inputs as

$$w_j = \sum_{i=1}^{N} |U_{ij}|^2, \tag{1}$$

where $\boldsymbol{U}$ is the unitary matrix describing the lossless interferometer. The mode mismatch induced click probability at output j is then given by

$$p_{mis,j} = \mu\eta_L\eta_D(1-p_{ind})w_j, \tag{2}$$

where μ is the mean photon number per populated input mode. In matrix form,

$$\boldsymbol{p}_{mis} = \mu\eta_L\eta_D(1-p_{ind})\mathrm{diag}(w_1,\cdots,w_M). \tag{3}$$

Since U is unitary, $\sum_j w_j = N$. Therefore, the total mode mismatch induced background scales with the number of populated inputs N. Within the present effective model, the fraction p_{ind} contributes to the intended quantum interference. Accordingly, the transmission appearing in the interference term of Eq. (7) in the main text is replaced by $\eta_L p_{ind}$. The distinguishable fraction contributes instead to the random count matrix,

$$\boldsymbol{p}_D = \boldsymbol{p}_{mis} + D_c\boldsymbol{I}_M, \tag{4}$$

where D_c is the intrinsic dark count probability. Substitution into Eq. (7) in the main text gives the general condition

$$\boldsymbol{\Sigma}_{mis} = 2\mu\eta_L(1-p_{ind})\mathrm{diag}(w_1,\cdots,w_M) + \frac{2D_c}{\eta_D}\boldsymbol{I}_M + \eta_L p_{ind}(\bar{t}-1)\boldsymbol{U}^\dagger\boldsymbol{J}_N\boldsymbol{U} \geq 0. \tag{5}$$

For a sufficiently mode-mixing interferometer, the contribution from the N populated inputs is approximately distributed uniformly over the M output modes. We therefore use

$$w_j \simeq \frac{N}{M} \equiv \nu, \tag{6}$$

where ν is the input filling fraction. The mode mismatch induced background then becomes

$$\boldsymbol{p}_{mis} \simeq \mu\eta_L\eta_D(1-p_{ind})\nu\boldsymbol{I}_M. \tag{7}$$

* eastgil7902@gmail.com
† These authors contributed equally to this work
‡ joonsukhuh@gmail.com
§ yoonho72@gmail.com

Under this approximation, Eq. (5) reduces to the scalar sufficient condition,

$$2\mu\eta_L(1-p_{ind})\nu + \frac{2D_c}{\eta_D} + \eta_L p_{ind}(\bar{t}-1) \geq 0. \tag{8}$$

Equation (8) is the condition used in the main text. For a fully populated interferometer, $N = M$ and therefore $\nu = 1$, recovering the expression without the filling fraction factor.

B. Effective internal mode model

Mode mismatch arises when photons injected into different spatial input modes do not occupy identical temporal, spectral, or polarization modes. To describe this effect, we decompose the internal state of a photon entering input port i as

$$|\psi_i\rangle = \sqrt{p_{ind}}\,|\psi\rangle + \sqrt{1-p_{ind}}\,\left|\psi_i^{\perp}\right\rangle \tag{9}$$

where $|\psi\rangle$ is the internal mode common to all populated inputs and $\left|\psi_i^{\perp}\right\rangle$ is orthogonal to the common component, $\langle\psi\left|\psi_i^{\perp}\right\rangle = 0$. For the equal-overlap model adopted here, we additionally assume that the residual components associated with different input ports are mutually orthogonal, $\langle\psi_i^{\perp}\left|\psi_k^{\perp}\right\rangle = \delta_{ik}$. For two photons entering different input ports $i \neq k$, it gives $\langle\psi_i\,|\psi_k\rangle = p_{ind}$, and hence their pairwise mutual fidelity is $F_{ik} = |\langle\psi_i|\psi_k\rangle|^2 = p_{ind}^2$. Under this equal-overlap model, the indistinguishability parameter used in the main text is therefore related to the mutual fidelity by

$$p_{ind} = \sqrt{\mathcal{F}}. \tag{10}$$

For a general set of internal states, the residual components need not be mutually orthogonal and the full distinguishability structure cannot, in general, be specified by a single pairwise fidelity. In that case, p_{ind} should be interpreted as an effective common-mode weight within the present model.

II. Evaluation of the boson sampling volume

To estimate the effective low-collision problem size accessible to a given device, the boson sampling volume is defined in the main text as $V_{BS}(M_L) = \max_{M \leq M_L}[\min[S_{BS}(M), \sqrt{M}]]$ where S_{BS} is the boson sampling score, M_B is the scalability bound for boson sampling, M is the number of interferometer modes, and M_L is the practical constraint on the mode size of the interferometer. To calculate $S_{BS} = \min[N_{sat}, M] \times \min[\mu p_{ind}\eta_L\eta_D, 1]$ when $M < M_B$, we consider four cases that can occur based on the saturated-input regime and the probability of detecting a photon per input mode: (i) the saturated-input regime $N_{sat} < M$ with $\mu p_{ind}\eta_L\eta_D \geq 1$, (ii) the saturated-input regime $N_{sat} < M$ with $\mu p_{ind}\eta_L\eta_D < 1$, (iii) the unsaturated-input regime $M \leq N_{sat}$ with $\mu p_{ind}\eta_L\eta_D \geq 1$, (iv) the unsaturated-input regime $M \leq N_{sat}$ with $\mu p_{ind}\eta_L\eta_D < 1$, where $N_{sat} = M/(\mu\eta_L\eta_D)$ represents the number of injection ports that saturate the coincidence measurement of the boson sampling experiment, μ is the mean photon number per port, p_{ind} is the fraction of photons that remain indistinguishable, η_L is the transmission of the interferometer, and η_D is the detector efficiency.

Among these, the fourth case—characterized by unsaturated-input regime $M \leq N_{sat}$ and a detection probability less than unity $\mu p_{ind}\eta_L\eta_D < 1$—is the most critical for calculating the boson sampling volume. This case is particularly relevant because, under high-loss conditions, the achievable boson sampling size is often limited by the scalability bound M_B. Therefore, $M\mu p_{ind}\eta_L\eta_D$ is a crucial factor for efficiently determining the boson sampling volume. To calculate the boson sampling volume V_{BS}, we need to determine how the boson sampling score S_{BS} changes with the number of interferometer modes M and at what point, if any, S_{BS} equals $\sqrt{M}$. Among the five parameters in the term $M\mu p_{ind}\eta_L\eta_D$, only the number of interferometer modes M and interferometer transmission η_L depend on M. Therefore, our method for obtaining the boson sampling volume differs based on whether $M\eta_L$ is a monotonically nondecreasing function of M. This distinction is important because when $M\eta_L$ is monotonically nondecreasing, the calculation of the boson sampling volume is significantly simplified. We consider two cases and describe the detailed evaluation procedure for each in the following subsections.

A. Boson sampling volume calculation: When $M\eta_L$ is monotonically nondecreasing

We first consider the case where $M\eta_L$ is a monotonically nondecreasing function of M, as this simplifies the calculation of the boson sampling volume V_{BS}. When $M\eta_L$ is monotonically nondecreasing, the boson sampling score S_{BS} reaches its maximum at $M_B - 1$ and becomes zero for any $M \geq M_B$. This allows us to calculate the boson sampling volume using the following equation:

$$V_{BS} = \min\left[S_{BS}(M_B - 1), \sqrt{M_B - 1}\right]. \tag{11}$$

A notable example in which $M\eta_L$ is monotonically nondecreasing is the fast Fourier transform architecture described in the main manuscript. For this architecture, $M\eta_L$ follows the relation: $M\eta_c\eta_0^{\log_2 M} = \eta_c M^{1+\log_2 \eta_0}$. Consequently, if η_0, which represents the transmission of the optical elements comprising the fast Fourier transform architecture, exceeds 0.5, $M\eta_L$ becomes a monotonically nondecreasing function.

Given this known form of the $M\eta_L = \eta_c M^{\log_2 2\eta_0}$ function with respect to M, we can determine the scalability bound M_B as follows:

$$M_B = \left\lceil 2^{\log_{1/\eta_0}\left[-\frac{\eta_c\eta_D(2\mu(1-p_{ind})+p_{ind}(\bar{t}-1))}{2D_c}\right]} \right\rceil. \tag{12}$$

By substituting this $M_B - 1$ into Eq. (11), we can readily obtain the boson sampling volume for this scenario.

B. Boson sampling volume calculation: When $M\eta_L$ is not monotonically nondecreasing

When $M\eta_L$ is not a monotonically nondecreasing function, the boson sampling score S_{BS} as a function of M increases to a maximum and then decreases, as illustrated in Fig. S1. For illustration, we generate different boson sampling score functions by varying the input coupling efficiency η_c for explanation. As shown in Fig. S1(a), if the scalability bound M_B is below the peak position of the boson sampling score function S_{BS}, we can calculate the boson sampling volume V_{BS} as in the previous case since the actual boson sampling score function S_{BS} remains monotonically nondecreasing up to the scalability bound M_B.

However, when the scalability bound M_B exceeds the peak position of the boson sampling score function S_{BS}, we consider three scenarios, as depicted in Fig. S1(b). Let M_{peak} denote the peak position of the boson sampling score function S_{BS}.

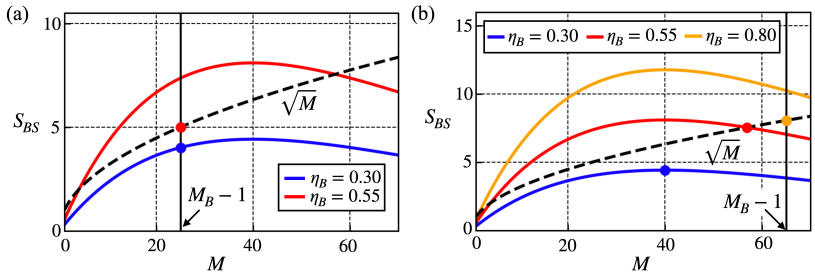

FIG. S1. S_{BS} as a function of M for evaluating V_{BS}. (a) Scenario where the scalability bound is below the peak position of the boson sampling score function ($M_B \leq M_{peak}$). (b) Scenarios where the scalability bound exceeds the peak position of the boson sampling score function ($M_B > M_{peak}$).

(i) If the peak value of the boson sampling score lies below $\sqrt{M}$ (blue line in Fig. S1(b)), the peak value $S_{BS}(M_{peak})$ itself becomes the boson sampling volume V_{BS}. Note that we define the boson sampling volume V_{BS} as the maximum value of min $\left[S_{BS}(M), \sqrt{M}\right]$.

(ii) If both the peak value of the boson sampling score and its value at $M_B - 1$ lie above $\sqrt{M}$ (orange line in Fig. S1(b)), the $\sqrt{M_B - 1}$ represents the boson sampling volume V_{BS}.

(iii) If the peak value of the boson sampling score lies above $\sqrt{M}$, but its value at $M_B - 1$ lies below $\sqrt{M}$ (red line in Fig. S1(b)), we first determine the intersection point M_{cross} between the boson sampling score function S_{BS} and the $\sqrt{M}$ function using a root-finding algorithm. We use the bisection method, as the function $f(M) = S_{BS}(M) - \sqrt{M}$ is continuous on the interval $[M_{peak}, M_B]$, and $f(M_{peak})$ and $f(M_B)$ have opposite signs. Thus, $\sqrt{M_{cross}}$ is the boson sampling volume V_{BS} in this case.

These cases are summarized as

$$V_{BS} = \begin{cases} \min\left[S_{BS}(M_B - 1), \sqrt{M_B - 1}\right] & M_{peak} \geq M_B \\ S_{BS}(M_{peak}) & M_{peak} < M_B \& S_{BS}(M_{peak}) \leq \sqrt{M_{peak}} \\ \sqrt{M_B - 1} & M_{peak} < M_B \& S_{BS}(M_{peak}) > \sqrt{M_{peak}} \text{ and } S_{BS}(M_B - 1) \geq \sqrt{M_B - 1} \\ \sqrt{M_{cross}} & M_{peak} < M_B \& S_{BS}(M_{peak}) > \sqrt{M_{peak}} \text{ and } S_{BS}(M_B - 1) < \sqrt{M_B - 1} \end{cases}, \tag{13}$$

which provides a comprehensive definition of the boson sampling volume V_{BS} under various conditions.

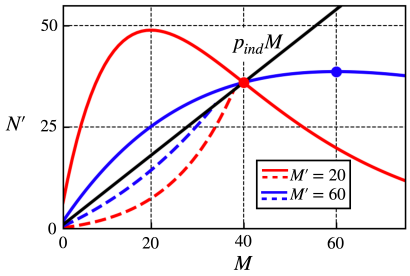

FIG. S2. Three cases for calculating the boson sampling score when the number of interferometer modes is less than the boson sampling scalability bound ($M < M_B$). N' denotes the number of input interferometer modes for each case.

As described in Eq. (13), to calculate the boson sampling volume when $M\eta_L$ as a function of M is not monotonically nondecreasing, it is crucial to determine the peak position M_{peak} of the boson sampling score function. As previously discussed, there are four cases for calculating the boson sampling score $S_{BS} = \min[N_{sat}, M] \times \min[\mu p_{ind}\eta_L\eta_D, 1]$ when the number of interferometer modes is less than the boson sampling scalability bound M_B, expressed as $S_{BS} = \min[M/(\mu\eta_L\eta_D), p_{ind}M, M, M\mu p_{ind}\eta_L\eta_D]$. Given that $p_{ind} \leq 1$, the value corresponding to the second case is always less than or equal to that corresponding to the third case ($p_{ind}M \leq M$). Consequently, we can simplify our analysis to three cases for calculating the boson sampling score S_{BS} when the number of interferometer modes is below the scalability bound ($M < M_B$). Figure S2 illustrates these three cases: (i) the colored dotted curves represent the first case ($M/(\mu\eta_L\eta_D)$), (ii) the black line represents the second case ($p_{ind}M$), (iii) the colored solid lines represent the fourth case ($M\mu p_{ind}\eta_L\eta_D$). The parameters are varied for clarity.

A critical point of interest is the intersection between the first and second cases ($M_{sat}/(\mu\eta_L\eta_D) = p_{ind}M_{sat}$), where M_{sat} denotes the intersection point of the functions $M/(\mu\eta_L\eta_D)$ and $p_{ind}M$. Another important point is the maximum position of the fourth case ($M\mu p_{ind}\eta_L\eta_D$). We can determine this by differentiating $M\eta_L$ with respect to M and solving $\partial M\eta_L/\partial M = 0$. Let M' denote this maximum position. The peak position of the boson sampling score function S_{BS} can now be determined as illustrated in Fig. S1: When $M' > M_{sat}$, M_{peak} is M'. When $M' \leq M_{sat}$, the peak position of S_{BS} is M_{sat}. Thus, $M_{peak} = \max[M', M_{sat}]$. Using these equations, we can calculate the following: M' (the maximum position of the fourth case), M_{sat} (the intersection point between the first and second cases), M_{peak} (the peak position of the boson sampling score function), M_B (the boson sampling scalability bound), $S_{BS}(M_{peak})$ (the boson sampling score at M_{peak}), and $S_{BS}(M_B - 1)$ (the boson sampling score at $M_B - 1$). These calculations allow us to obtain the boson sampling volume as described in Eq. (13).

III. Transmission scaling of Jiuzhang-type interferometer architecture

The multimode linear optical network of the Jiuzhang system incorporates phase stability, full connectivity, matrix randomness, high wavepacket overlap, and high transmission efficiency [1–3]. This optical network comprises a multi-layer rectangular interferometer followed by a multi-layer triangular interferometer [2, 3] as shown in Fig. S3(a).

In the initial stage, the rectangular interferometer allows photons within the same horizontal layer to interact, while photons in different horizontal layers remain isolated [1–3]. At this point, the modes (represented as circles) are only connected within their respective horizontal layers, as illustrated in Fig. S3(b). Subsequently, the triangular interferometer is implemented to establish connections between vertical layers as shown in Fig. S3(c). This configuration enables photons from different horizontal layers to interfere with each other, resulting in full connectivity among all modes [1–3].

As shown in Figs. S3(d) and (e), both the rectangular and triangular interferometers follow a universal architecture, typically implemented on a two-dimensional platform such as a 2D photonic chip or a free-space optical table. The depth of each interferometer is proportional to $m_{R(T)}$, where $m_{R(T)}$ denotes the number of modes in the rectangular(triangular) interferometer. Therefore, the number of interferometer modes in the complete Jiuzhang system (combining both rectangular and triangular interferometers) scale as $M \propto m_T m_R$. Consequently, the depth of the complete Jiuzhang interferometer is proportional to $\sqrt{M}$.

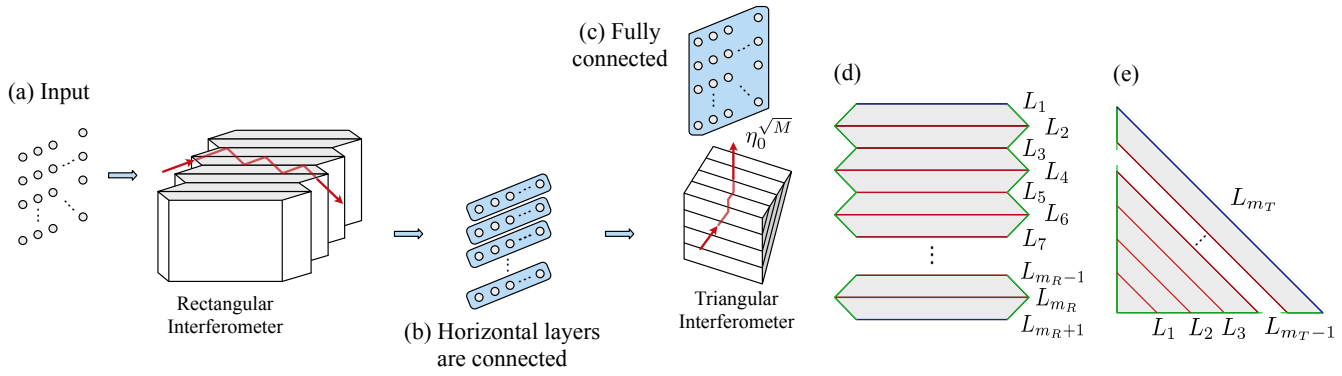

FIG. S3. Schematic design of the Jiuzhang interferometer. (a) Three-dimensional structure comprising rectangular and triangular interferometers as building blocks. (b) Configuration of the optical modes after the rectangular interferometer: horizontal layers are connected, but vertical layers remain isolated. (c) Optical modes configuration after the triangular interferometer: all modes are fully connected. (d) Cross-sectional view of the rectangular interferometer. (e) Cross-sectional view of the triangular interferometer. Red lines represent 50:50 beam-splitting films, blue lines indicate mirrors, and green lines denote antireflective coatings.

IV. Numerical evaluation of S_{BS} and V_{BS} for Jiuzhang4.0

To illustrate how the quantities introduced in the main text are evaluated from experimentally reported parameters, we consider the Jiuzhang4.0 L1024 (6 nJ) dataset as a representative example [4]. The experiment employs $N_{sq} = 1024$ squeezed input ports and $M = 8176$ output modes, with an average detected output photon number $N_{output} = 2207$ and an overall system efficiency $\eta^{sys} = 0.51$. Since the mean photon number μ is defined per occupied input port, the number of squeezed input ports N_{sq}, rather than the total output mode number M, must be used to infer μ,

$$\mu = \frac{N_{output}}{N_{sq}\eta^{sys}} = \frac{2207}{1024 \times 0.51} \simeq 4.226. \tag{14}$$

For a squeezed vacuum input, $\mu = \sinh^2 r$, and the corresponding PQD ordering threshold is $\bar{t} = e^{-2r}$. These relations give

$$r = \operatorname{arcsinh}\sqrt{\mu} \simeq 1.468, \tag{15}$$

$$\bar{t} = e^{-2r} \simeq 0.0530. \tag{16}$$

Using $p_{ind} = 0.9995$, the coefficient governing the nontrivial PQD-based transmission threshold becomes

$$C = 2\mu(1 - p_{ind})\nu + p_{ind}(\bar{t} - 1) \simeq -0.9459 < 0, \tag{17}$$

where $\nu = 1024/8176 \simeq 0.1252$. Thus, a finite transmission threshold exists. Taking the detector efficiency and dark count probability to be $\eta_D = 0.93$ and $D_c = 5 \times 10^{-6}$, respectively, we obtain

$$\eta_{th} = -\frac{2D_c}{\eta_D C} \simeq 1.14 \times 10^{-5}. \tag{18}$$

As the inferred interferometer transmission is $\eta_L = \eta^{sys}/\eta_D \simeq 0.548$, it satisfies $\eta_L \gg \eta_{th}$. Therefore, Jiuzhang4.0 is not certified as classically simulatable by the PQD-based sufficient condition at its reported scale. Note that unlike the earlier spatially encoded Jiuzhang architectures, Jiuzhang4.0 employs a hybrid spatial-temporal circuit. Consequently, the M-dependent transmission model is not uniquely implied by its physical layout. If the effective scaling model $\eta_L(M) = \eta_c \eta_0^{\sqrt{M}}$ is adopted for comparison with the earlier Jiuzhang-type architectures, it yields $M_B \simeq 3.90 \times 10^6$; this value should, however, be interpreted as model-dependent.

The boson sampling score at the implemented mode number is evaluated using $N_{sat} = M/(\mu\eta_L\eta_D) \simeq 3.793 \times 10^3$. Since $\mu p_{ind}\eta_L\eta_D > 1$, the second factor in the definition of S_{BS} is saturated at unity. Consequently,

$$S_{BS} = \min[N_{sat}, M] \simeq 3793. \tag{19}$$

The implemented-mode boson sampling volume is therefore

$$V_{BS}(8176) = \min[S_{BS}, \sqrt{8176}] = \sqrt{8176} = 90.42. \tag{20}$$

Hence, the effective problem size of Jiuzhang4.0 is limited not by the boson sampling score or the PQD-based scalability bound, but by the available mode size. Within the present metric, the experiment exceeds the $V_{BS} \simeq 50$ benchmark but remains below $V_{BS} \simeq 100$, for which at least $M = 10^4$ modes would be required.

[1] H. Wang, J. Qin, X. Ding, M.-C. Chen, S. Chen, X. You, Y.-M. He, X. Jiang, L. You, Z. Wang, C. Schneider, J. J. Renema, S. Höfling, C.-Y. Lu, and J.-W. Pan, "Boson sampling with 20 input photons and a 60-mode interferometer in a 10^{14}-dimensional Hilbert space," Phys. Rev. Lett. **123**, 250503 (2019).

[2] H.-S. Zhong, H. Wang, Y.-H. Deng, M.-C. Chen, L.-C. Peng, Y.-H. Luo, J. Qin, D. Wu, X. Ding, Y. Hu, P. Hu, X.-Y. Yang, W.-J. Zhang, H. Li, Y. Li, X. Jiang, L. Gan, G. Yang, L. You, Z. Wang, L. Li, N.-L. Liu, C.-Y. Lu, and J.-W. Pan, "Quantum computational advantage using photons," Science **370**, 1460-1463 (2020).

[3] H.-S. Zhong, Y.-H. Deng, J. Qin, H. Wang, M.-C. Chen, L.-C. Peng, Y.-H. Luo, D. Wu, S.-Q. Gong, H. Su, Y. Hu, P. Hu, X.-Y. Yang, W.-J. Zhang, H. Li, Y. Li, X. Jiang, L. Gan, G. Yang, L. You, Z. Wang, L. Li, N.-L. Liu, J. J. Renema, C.-Y. Lu, and J.-W. Pan, "Phase-programmable Gaussian boson sampling using stimulated squeezed light," Phys. Rev. Lett. **127**, 180502 (2021).

[4] H.-L. Liu *et al.*, "Gaussian boson sampling with 1,024 squeezed states in 8,176 modes," Nature **653**, 687 (2026).