nature physics

Article

https://doi.org/10.1038/s41567-023-02308-9

Quantum-inspired classical algorithms for molecular vibronic spectra

In the format provided by the authors and unedited

CONTENTS

S1. GURVITS'S ALGORITHM WITH REPEATED ROWS AND COLUMNS

We first present the original Gurvits's algorithm and generalize the algorithm to improve the bound when rows and columns are repeated. Gurvits's algorithm exploits the following equality [1, 2]:

$$\mathrm{Per}(X) = \mathbb{E}_{\boldsymbol{x}\in\{-1,1\}^n}\left[\prod_{i=1}^{n}\left(\sum_{j=1}^{n} x_i X_{ij} x_j\right)\right]. \tag{S1}$$

Thus, by sampling the random variable with uniform $\boldsymbol{x}\in\{-1,1\}^n$, the randomized algorithm gives an estimate μ of permanent of an $n\times n$ matrix X such that

$$|\mathrm{Per}(X)-\mu| < \epsilon\|X\|^n, \tag{S2}$$

with high probability $1-\delta$ with running time $T = O(\mathrm{poly}(n, 1/\epsilon, \log\delta^{-1}))$. When we consider a submatrix of a unitary matrix, this expression is sufficient because the submatrix's norm is bounded by 1. However, when we have repeated rows and columns in X from a unitary matrix, which is the case when we have multiphoton inputs, the norm $\|X\|$ is not generally bounded by 1. Therefore, the upper bound of the error can increase as $\|X\|^n$, i.e., exponentially in n. To deal with multiphoton cases with a reduced error, we need to generalize Gurvits's algorithm. Note that although Ref. [3] generalizes Gurvits's algorithm for multiphoton outputs, i.e., $\mathrm{Per}(V_{\boldsymbol{n},\boldsymbol{m}}) = \langle\boldsymbol{n}|\hat{V}|\boldsymbol{m}\rangle$, where $\boldsymbol{n}\in\{0,1\}^n$ and $\boldsymbol{m}$ is a nonnegative integer vector, it is not sufficient for our purpose where we have multiphoton input and output, i.e., $\langle\boldsymbol{n}|\hat{V}|\boldsymbol{n}\rangle = \mathrm{Per}(V_{\boldsymbol{n},\boldsymbol{n}})$.

Consider a case where the goal is to approximate a quantity

$$\frac{\mathrm{Per}(A)}{\boldsymbol{n}!}, \tag{S3}$$

where $A\in\mathbb{C}^{n\times n}$ which is obtained by repeating ith row and column of a matrix $B\in\mathbb{C}^{k\times k}$ for n_i times. Here, $\boldsymbol{n}\in\mathbb{Z}^k_{\geq 0}$ with $\sum_{i=1}^k n_i = n$. Following the method in Ref. [3], let us define a random variable

$$\mathrm{GenGly}_x(A) \equiv \frac{1}{\prod_{i=1}^k n_i^{n_i}}\prod_{i=1}^{k}\bar{y}_i^{n_i}\prod_{i=1}^{k}(y_1B_{i,1}+\cdots+y_kB_{i,k})^{n_i} = \prod_{i=1}^{k}\left(\frac{\bar{y}_i(By)_i}{n_i}\right)^{n_i}, \tag{S4}$$

where $y_i \equiv n_i^{1/2}x_i$. Here, $\boldsymbol{x}\in\mathcal{R}[n_1+1]\times\cdots\times\mathcal{R}[n_k+1]\equiv\mathcal{X}$, where $\mathcal{R}[j]$ is the set of jth roots of unity. Then, its absolute value is given as

$$|\mathrm{GenGly}_x(A)| = \left(\prod_{i=1}^{k} n_i^{n_i}\right)^{-1/2}\left|\prod_{i=1}^{k}(y_1b_{i,1}+\cdots+y_kb_{i,k})^{n_i}\right| = \left(\prod_{i=1}^{k} n_i^{n_i}\right)^{-1/2}\left|\prod_{i=1}^{k} c_i^{n_i}\right|, \tag{S5}$$

where $c_i \equiv y_1B_{i,1}+\cdots+y_kB_{i,k}$. Now, let us find its upper bound, which determines the error of the corresponding randomized algorithm by Chernoff bound. First, we have a constraint

$$\sum_{i=1}^{k}|c_i|^2 = \sum_{i=1}^{k}|(By)_i|^2 \leq \|B\|^2\|y\|^2 = n\|B\|^2. \tag{S6}$$

We then consider

$$\log\prod_{i=1}^{k}|c_i|^{n_i} = \sum_{i=1}^{k} n_i\log|c_i|. \tag{S7}$$

Using lagrange multiplier, with a constraint $\sum_{i=1}^k|c_i|^2 = c$,

$$Z \equiv \sum_{i=1}^{k} n_i\log|c_i| - \lambda\left(\sum_{i=1}^{k}|c_i|^2 - c\right), \tag{S8}$$

we obtain

$$\frac{\partial Z}{\partial|c_i|} = \frac{n_i}{|c_i|} - 2\lambda|c_i| = 0, \quad \forall i. \tag{S9}$$

Using the constraint Eq. (S6),

$$\sum_{i=1}^{k} |c_i|^2 = \sum_{i=1}^{k} \frac{n_i}{2\lambda} = \frac{n}{2\lambda} = c, \tag{S10}$$

the solution is obtained as

$$|c_i| = \sqrt{\frac{n_i}{2\lambda}} = \sqrt{\frac{cn_i}{n}}. \tag{S11}$$

Therefore, using the solution, we have an upper bound as

$$\prod_{i=1}^{k} |c_i|^{n_i} \leq \exp\left(\sum_{i=1}^{k} n_i \log \sqrt{\frac{cn_i}{n}}\right) = \prod_{i=1}^{k} n_i^{n_i/2} \left(\frac{c}{n}\right)^{n/2} \leq \prod_{i=1}^{k} n_i^{n_i/2} \|B\|^n. \tag{S12}$$

Hence, we have an upper bound of the absolute value of the random variable,

$$|\mathrm{GenGly}_x(A)| = \left(\prod_{i=1}^{k} n_i^{n_i}\right)^{-1/2} \left|\prod_{i=1}^{k} c_i^{n_i}\right| \leq \|B\|^n. \tag{S13}$$

Now, we prove that the average of the random variables gives an unbiased estimator of $\mathrm{Per}(A)/\boldsymbol{n}!$, i.e.,

$$\mathbb{E}_{x\in\mathcal{X}}[\mathrm{GenGly}_x(A)] = \frac{\mathrm{Per}(A)}{\boldsymbol{n}!}. \tag{S14}$$

First, we have

$$\mathbb{E}_{x\in\mathcal{X}}\left[\frac{1}{\prod_{i=1}^{k} n_i^{n_i}} \prod_{i=1}^{k} \bar{y}_i^{n_i} \prod_{i=1}^{k} (y_1 B_{i,1} + \cdots + y_k B_{i,k})^{n_i}\right] = \frac{1}{\prod_{i=1}^{k} n_i^{n_i}} \mathbb{E}_{x\in\mathcal{X}}\left[\prod_{i=1}^{k} \bar{y}_i^{n_i} (y_1 B_{i,1} + \cdots + y_k B_{i,k})^{n_i}\right] \tag{S15}$$

$$= \frac{1}{\prod_{i=1}^{k} n_i^{n_i}} \sum_{\sigma_1,\ldots,\sigma_n\in[k]} \left(\prod_{i=1}^{k} B_{i,\sigma_i}\right) \mathbb{E}_{x\in\mathcal{X}}\left[\prod_{i=1}^{k} \bar{y}_i^{n_i} \prod_{i=1}^{n} y_{\sigma_i}\right], \tag{S16}$$

where by the symmetry over the roots of unity,

$$\mathbb{E}_{x\in\mathcal{X}}\left[\prod_{i=1}^{k} \bar{y}_i^{n_i} \prod_{i=1}^{n} y_{\sigma_i}\right] = 0, \tag{S17}$$

unless the product insides is equal to $\prod_{i=1}^{k} |y_i|^{2n_i}$, in which case we have

$$\mathbb{E}_{x\in\mathcal{X}}\left[\prod_{i=1}^{k} \bar{y}_i^{n_i} \prod_{i=1}^{n} y_{\sigma_i}\right] = \prod_{i=1}^{k} n_i^{n_i}. \tag{S18}$$

Hence,

$$\mathbb{E}_{x\in\mathcal{X}}[\mathrm{GenGly}_x(A)] = \sum_{t_1\ldots,t_k\in[k]:|\{i:t_i=j\}|=n_j \ \forall j\in[k]} \prod_{i=1}^{k} B_{i,t_i} = \frac{\mathrm{Per}(A)}{\boldsymbol{n}!}, \tag{S19}$$

where $\boldsymbol{n}!$ appears to take into account the fact that each produce $\prod_{i=1}^{k} b_{i,t_k}$ appears $\boldsymbol{n}!$ times.

Therefore, $\mathrm{GenGly}_x(A)$ gives us an unbiased estimator of $\mathrm{Per}(A)/\boldsymbol{n}!$, and the random variables are bounded by $\|B\|^n$. Due to the Chernoff bound, the randomized algorithm with number of samples $N = O(1/\epsilon^2)$ provides an estimate μ with high probability $1-\delta$ such that

$$\left|\frac{\mathrm{Per}(A)}{\boldsymbol{n}!} - \mu\right| < \epsilon \|B\|^n. \tag{S20}$$

S2. SPARSE FAST FOURIER TRANSFORMATION FOR LARGE WEIGHT VECTOR

In the main text, we showed that we can efficiently solve the problem where $\Omega_{\max} = O(\mathrm{poly}(M))$ due to the assumption that $\omega_i = O(\mathrm{poly}(M))$ for all i's. By lifting the assumption, we now consider the case where $\Omega_{\max} = \omega(\mathrm{poly}(M))$. Let $d \equiv \Omega_{\max} + 1$ be the number of bins. In particular, we will assume that the goal is reproducing the $t = O(\mathrm{poly}(M))$ largest components $G(\Omega)$ in the spectrum, namely peaks, instead of reproducing all the components to focus on the problems that a boson sampler can solve in polynomial time. This is because estimating a superpolynomial number of quantities inevitably takes superpolynomial time. This assumption is also practically reasonable because the molecular vibronic spectra problem's typical purpose is identifying the peaks. For a boson sampler to achieve reasonable accuracy, we assume that the t peaks are as large as $1/\mathrm{poly}(M)$. We further assume that $\log \omega_i = O(\mathrm{poly}(M))$ for all $i \in [M]$ so that ω_i's can be represented in a polynomial number of bits. Consequently, $\log d = O(\mathrm{poly}(M))$.

A. Boson-sampling case

Suppose that we have obtained N number of samples by running a boson sampler. The standard Chernoff bound states that for each $\Omega \in [0, \Omega_{\max}]$ with the N samples, we obtain the estimate of $G(\Omega)$ with accuracy ϵ with high probability $1 - 2e^{-2N\epsilon^2}$. By applying the union bound for all the t peaks, we obtain the accuracy ϵ for t different Ω with probability $1 - 2te^{-2N\epsilon^2} = 1 - 2e^{-2N\epsilon^2 + \log t}$. Thus, we choose N such that $-2N\epsilon^2 + \log t = O(1)$. Then, we need at least $N = O(\log t/\epsilon^2)$. Since our goal is to estimate $G(\Omega)$ for all the t peaks, we can guarantee that the estimated peaks G'_t is $t\epsilon$-close to the true peaks:

$$\|G_t - G'_t\|_1 \leq t\epsilon, \tag{S21}$$

where G_t is the distribution only for the peaks. In other words, $N = O(t^2 \log t/\epsilon^2)$ is sufficient to guarantee with high probability that

$$\|G_t - G'_t\|_1 \leq \epsilon. \tag{S22}$$

B. Efficient classical algorithm using a sparse fast Fourier transformation

We now consider the same problem, i.e., estimating the t peaks $G(\Omega) = \Omega(1/\mathrm{poly}(M))$. For our classical algorithm, we exploit the sparse fast Fourier transformation (SFFT) [4]. For general cases, the SFFT gives the estimate $G' = G'(\Omega)$ such that

$$\|G - G'\|_2 \leq (1+\epsilon) \min_{H \in t\text{-sparse}} \|G - H\|_2 + \delta \|G\|_2, \tag{S23}$$

in $O(t/\epsilon \log(d/t) \log(d/\delta))$ time with high probability. Here, the minimization is over t-sparse distributions $H = H(\Omega)$, i.e., only t elements are non-zero. For our case, $\|G\|_2 \leq \|G\|_1 = 1$, and due to the dependence of the running time on δ, we can choose $\delta = 1/\exp(M)$. Intuitively, the multiplicative factor 1 comes from the tails, i.e., the terms that are not peaks; thus, if we focus only on the peaks, the right-hand-side should be efficiently reduced by choosing ϵ and δ to be $1/\mathrm{poly}(M)$.

More rigorously, we can show that the error on the peaks can be efficiently reduced by the fact that the final round of the algorithm in Ref. [4] gives us the estimate $G'(\Omega)$ of the distribution $G(\Omega)$ (see Eq. (6) in Ref. [4]), which satisfies

$$\|G - G'\|_\infty^2 \leq \frac{\epsilon}{t} \left(\min_{H \in t\text{-sparse}} \|G - H\|_2^2 + d\delta^2 \right). \tag{S24}$$

Therefore, the total variation distance for the peaks is given by

$$\|G_t - G'_t\|_1 \leq \sqrt{t\epsilon} \left(\min_{H \in t\text{-sparse}} \|G - H\|_2^2 + d\delta^2 \right)^{1/2}. \tag{S25}$$

Hence, choosing $\epsilon' = \sqrt{t\epsilon} \left(\min_{H \in t\text{-sparse}} \|G - H\|_2^2 + d\delta^2 \right)^{1/2}$ with $\delta = 1/\sqrt{d}$, we can achieve

$$\|G_t - G'_t\|_1 \leq \epsilon', \tag{S26}$$

in $O(t^2/\epsilon'^2 \log(d/t)\log(d/\delta))$ time with high probability. Hence, when we can compute the Fourier components $\tilde{G}(k)$ exactly, the SFFT gives an estimate G' with high probability. Thus, it already suffices to show that for the molecular vibronic spectra problem corresponding to the Gaussian boson sampling, the SFFT provides an estimate of $G(\Omega)$ as accurate as running a Gaussian boson sampler.

For the case of Fock-state boson sampling, we can only approximate the Fourier components because exactly computing them is #P-hard. Therefore, the SFFT has to be applied to the approximated Fourier components whose inverse Fourier transformation gives $\hat{G}$. If we assume that the approximation of the Fourier components does not change the position of peaks (it may change the value of the peaks but not the position of the peaks), the SFFT applied to the approximated Fourier coefficients can still recover the peaks with error produced by the approximation. Thus, under the assumption, we have the following inequality for the estimated spectrum G'_t:

$$\|G_t - G'_t\|_1 \leq \|G_t - \hat{G}_t\|_1 + \|\hat{G}_t - G'_t\|_1 \leq \sqrt{t}\epsilon'' + \epsilon', \tag{S27}$$

where ϵ' comes from the SFFT, ϵ'' comes from the approximation of the Fourier components (See Methods C.), and $\sqrt{t}$ factor comes from $\|v\|_1 \leq \sqrt{t}\|v\|_2$ for length t vector v. Thus, by setting $\epsilon'' = \epsilon/(2\sqrt{t})$ and $\epsilon' = \epsilon/2$, we obtain the accuracy $\|G_t - G'_t\|_1 \leq \epsilon$ in running time $O(\mathrm{poly}(M, 1/\epsilon))$.

Now, the remaining problem is to show that the approximation of the Fourier coefficients does not change the position of the peaks. To prove this, it suffices to show that the approximation error in the Fourier basis does not generate an error as large as $o(1/\mathrm{poly}(M))$ with high probability. To this end, we make use of the Hoeffding inequality for the sub-Gaussian distribution. This is because we approximate the permanent using the generalized Gurvits's algorithm. In particular, the Hoeffding inequality for the generalized Gurvits's algorithm from Sec. S1 gives the following inequality:

$$\Pr\left[\left|\tilde{p} - \frac{\mathrm{Per}(V_{\boldsymbol{n}})}{\boldsymbol{n}!}\right| \geq \epsilon\right] \leq 2\exp\left(-\frac{\epsilon^2 N}{2}\right), \tag{S28}$$

where N is the number of samples. On the other hand, the definition of the sub-Gaussian distribution is that

$$\Pr\left[|X| \geq t\right] \leq 2e^{-ct^2}, \tag{S29}$$

for some $c > 0$. Thus, the estimation error from the generalized Gurvits's algorithm follows a sub-Gaussian distribution for any $N > 0$.

For random sub-Gaussian variables $X_1, \ldots, X_m$, the Hoeffiding inequality gives [5]

$$\Pr_X\left[\left|\frac{\sum_{i=1}^m X_i}{m}\right| \geq \epsilon\right] \leq 2\exp\left(-\frac{cm^2\epsilon^2}{\sum_{i=1}^m \|X_i\|_{\psi_2}}\right) \equiv 2\exp\left(-Cm\epsilon^2\right) \tag{S30}$$

for an absolute constant $c > 0$, and $C = cm/\sum_{i=1}^m \|X_i\|_{\psi_2} > 0$. Here, we defined

$$\|X_i\|_{\psi_2} \equiv \inf\{c > 0 : \mathbb{E}[e^{X_i^2/c^2}] \leq 2\}, \tag{S31}$$

which is finite if and only if X_i is sub-Gaussian. Now, recall that the error of the spectra is written as

$$\Delta G(\Omega) = \frac{1}{\Omega_{\max}+1}\sum_{k=0}^{\Omega_{\max}} \Delta\tilde{G}(k)e^{ik\theta\Omega}. \tag{S32}$$

The real part of the error is written as

$$\mathrm{Re}[\Delta G(\Omega)] = \frac{1}{\Omega_{\max}+1}\sum_{k=0}^{\Omega_{\max}} \mathrm{Re}[\Delta\tilde{G}(k)e^{ik\theta\Omega}] = \frac{1}{\Omega_{\max}+1}\sum_{k=0}^{\Omega_{\max}} \left[\mathrm{Re}[\Delta\tilde{G}(k)]\cos(k\theta\Omega) - \mathrm{Im}[\Delta\tilde{G}(k)]\sin(k\theta\Omega)\right]. \tag{S33}$$

Here, we note that the random variables $\mathrm{Re}[\Delta\tilde{G}(k)]\cos(k\theta\Omega) - \mathrm{Im}[\Delta\tilde{G}(k)]\sin(k\theta\Omega)$ are not correlated for different k's because both $\mathrm{Re}[\Delta\tilde{G}(k)]$ and $\mathrm{Im}[\Delta\tilde{G}(k)]$ are not correlated for different k's. Meanwhile, to avoid the possibility that the real and imaginary errors $\mathrm{Re}[\Delta\tilde{G}(k)]$ and $\mathrm{Im}[\Delta\tilde{G}(k)]$ of $\tilde{G}(k)$ are correlated, we estimate the real part and the imaginary part independently in the generalized Gurvits's algorithm, which only doubles the computational cost. Thus, when we aim to estimate the real part, $\mathrm{Re}[\Delta\tilde{G}(k)]$, after running the generalized Gurvits's algorithm, we keep the real part while ignoring the imaginary part, and vice versa. By applying the Hoeffding bound (S30) on Eq. (S33), we obtain

$$\Pr\left[|\mathrm{Re}[\Delta G(\Omega)]| \geq \epsilon\right] \leq 2\exp\left(-C_R d\epsilon^2\right), \tag{S34}$$

where $C_R > 0$ is a constant. Using the same argument for the imaginary part and applying the union bound, we obtain

$$\Pr\left[|\Delta G(\Omega)| \geq \epsilon\right] \leq 2\exp\left(-C_R d\epsilon^2\right) + 2\exp\left(-C_I d\epsilon^2\right) \leq 4\exp\left(-Cd\epsilon^2\right), \tag{S35}$$

where C_I is the counterpart of C_R for imaginary part and $C \equiv \min(C_R, C_I)$. Now we apply the union bound for all Ω, which gives us

$$\Pr\left[|\Delta G(\Omega)| \geq \epsilon \text{ for some } \Omega \in [0, \Omega_{\max}]\right] \leq 4d\exp\left(-Cd\epsilon^2\right) = 4\exp\left(-Cd\epsilon^2 + \log d\right). \tag{S36}$$

Here, by choosing $\epsilon = 1/d^{1/4} = o(1/\mathrm{poly}(M))$, we finally prove that the approximation error in the spectrum is upper bounded by $o(1/\mathrm{poly}(M))$ with high probability.

S3. THE POSITIVE P-REPRESENTATION

In this section, we recall the generalized (positive) P-representation introduced in Ref. [6]:

$$\hat{\rho} = \int_{\mathbb{C}^{2M}} P(\boldsymbol{\alpha}, \boldsymbol{\beta})\hat{\Lambda}(\boldsymbol{\alpha}, \boldsymbol{\beta}) d\boldsymbol{\alpha} d\boldsymbol{\beta}, \quad \hat{\Lambda}(\boldsymbol{\alpha}, \boldsymbol{\beta}) \equiv \frac{|\boldsymbol{\alpha}\rangle\langle\boldsymbol{\beta}^*|}{\langle\boldsymbol{\beta}^*|\boldsymbol{\alpha}\rangle}, \tag{S37}$$

where $|\alpha\rangle$ and $|\beta^*\rangle$ are coherent states. The generalized P-representation is one of the quasi-probability distributions of a bosonic state [7]. An important property of the generalized P-representation is that the distribution $P(\boldsymbol{\alpha}, \boldsymbol{\beta})$ can always be chosen nonnegative (we call this positive P-representation.) and the expectation value of a normal-ordered operator can be readily computed [6]. The former can be proved as follows. For simplicity, we use a simpler version, where we assume that the Glauber-Sudarshan P-function exists [7], which is written as

$$\hat{\rho} = \int_{\mathbb{C}} d^2\alpha P_{GS}(\alpha)|\alpha\rangle\langle\alpha|, \tag{S38}$$

for a single-mode state. We note that although we assume the existence of the Glauber-Sudarshan P-function for simplicity, the claim is true for arbitrary density matrices [6]. The claim is that the positive P-representation is written as

$$P(\alpha, \beta) = \frac{1}{4\pi^2}\exp\left(-\frac{|\alpha - \beta^*|^2}{4}\right)\langle(\alpha + \beta^*)/2|\hat{\rho}|(\alpha + \beta^*)/2\rangle. \tag{S39}$$

If the claim is true, then it guarantees that the representation is positive and real. By directly substituting the representation, we get

$$P(\alpha, \beta) = \frac{1}{4\pi^2}\int P_{\mathrm{GS}}(\gamma)\exp\left(-\frac{1}{2}|\alpha - \gamma|^2 - \frac{1}{2}|\beta^* - \gamma|^2\right) d^2\gamma. \tag{S40}$$

Meanwhile, for an analytical function f, we have

$$f(\gamma) = \frac{1}{2\pi}\int_{\mathbb{C}} f(\alpha)\exp\left(-\frac{1}{2}|\alpha - \gamma|^2\right) d^2\alpha. \tag{S41}$$

It can be shown that for example at $\gamma = 0$,

$$\int_{\mathbb{C}} f(\alpha)\exp\left(-\frac{1}{2}|\alpha|^2\right) d^2\alpha = 2\pi\exp\left(\frac{1}{2}\frac{\partial^2}{\partial\alpha\partial\alpha^*}\right) f(\alpha)|_{\alpha=0} = 2\pi f(0), \tag{S42}$$

where the last equality holds because the function is analytic. Using this property, we obtain from Eq. (S40)

$$\int_{\mathbb{C}^2} \hat{\Lambda}(\alpha, \beta)P(\alpha, \beta)d^2\alpha d^2\beta = \frac{1}{4\pi^2}\int_{\mathbb{C}^3} d^2\alpha d^2\beta d^2\gamma\hat{\Lambda}(\alpha, \beta)P_{\mathrm{GS}}(\gamma)\exp\left(-\frac{1}{2}|\alpha - \gamma|^2 - \frac{1}{2}|\beta^* - \gamma|^2\right) \tag{S43}$$

$$= \int_{\mathbb{C}} d^2\alpha P_{\mathrm{GS}}(\alpha)|\alpha\rangle\langle\alpha| = \hat{\rho}, \tag{S44}$$

which proves the claim.

Now, we show by an example that the expectation value of a normal-ordered operator can be computed by averaging over the phase-space variable. For example, when one computes for a single-mode $:\hat{a}\hat{a}^\dagger\hat{a}\hat{a}^\dagger:=\hat{a}^{\dagger 2}\hat{a}^2$,

$$\mathrm{Tr}\big[\hat{\rho}:\hat{a}\hat{a}^\dagger\hat{a}\hat{a}^\dagger:\big]=\mathrm{Tr}\big[\hat{\rho}\hat{a}^{\dagger 2}\hat{a}^2\big]=\int_{\mathbb{C}^2}P(\alpha,\beta)\,\mathrm{Tr}\Big[\hat{\Lambda}(\alpha,\beta)\hat{a}^{\dagger 2}\hat{a}^2\Big]d\alpha d\beta \tag{S45}$$

$$=\int_{\mathbb{C}^2}P(\alpha,\beta)\frac{\langle\beta^*|\hat{a}^{\dagger 2}\hat{a}^2|\alpha\rangle}{\langle\beta^*|\alpha\rangle}d\alpha d\beta=\int_{\mathbb{C}^2}P(\alpha,\beta)\beta^2\alpha^2 d\alpha d\beta. \tag{S46}$$

Therefore, we may be able to directly integrate or estimate it by sampling (α,β) from P and averaging $\alpha^2\beta^2$ over the samples. The same procedure works for any normal-ordered operator $f(\hat{a},\hat{a}^\dagger)$ by replacing $\hat{a}\to\alpha$ and $\hat{a}^\dagger\to\beta$. Especially when the operator is written as a function of $\hat{n}$, $g(\hat{n})$, then the corresponding phase-variable is given by $g(\alpha\beta)$.

When we apply a linear-optical circuit to a given input state, the quantum state transforms as

$$\hat{\rho}\to\hat{U}\hat{\rho}\hat{U}^\dagger, \tag{S47}$$

which is written for the positive P-representation as

$$\hat{\rho}=\int_{\mathbb{C}^{2M}}P(\boldsymbol{\alpha},\boldsymbol{\beta})\hat{\Lambda}(\boldsymbol{\alpha},\boldsymbol{\beta})d\boldsymbol{\alpha}d\boldsymbol{\beta}\ \to\ \hat{\rho}=\int_{\mathbb{C}^{2M}}P(\boldsymbol{\alpha},\boldsymbol{\beta})\hat{\Lambda}(U\boldsymbol{\alpha},U^*\boldsymbol{\beta})d\boldsymbol{\alpha}d\boldsymbol{\beta}. \tag{S48}$$

Therefore, sampling from the positive P-representation of the output state of the linear-optical circuit is equivalent to sampling from the positive P-representation of the input state and transform the phase variable as $(\boldsymbol{\alpha},\boldsymbol{\beta})\to(\boldsymbol{\alpha}',\boldsymbol{\beta}')=(U\boldsymbol{\alpha},U^*\boldsymbol{\beta})$.

S4. POSITIVE P-REPRESENTATION OF A SQUEEZED VACUUM STATE

In this section, we show that a positive P-representation of a squeezed vacuum state can be chosen as [8]

$$P(x,y)=\frac{\sqrt{1+\gamma}}{\pi\gamma}\exp\left(-(x^2+y^2)(\gamma^{-1}+1/2)+xy\right), \tag{S49}$$

where x and y are real numbers. We follow the proof in Ref. [8]. We emphasize that this function is defined on the real plane, which is a distinct property of a squeezed state, and that generally the positive P-representation is defined on the complex-variable domain:

$$\hat{\rho}=\int_{\mathbb{C}^{2M}}P(\boldsymbol{\alpha},\boldsymbol{\beta})\hat{\Lambda}(\boldsymbol{\alpha},\boldsymbol{\beta})d\boldsymbol{\alpha}d\boldsymbol{\beta},\quad \hat{\Lambda}(\boldsymbol{\alpha},\boldsymbol{\beta})\equiv\frac{|\boldsymbol{\alpha}\rangle\langle\boldsymbol{\beta}^*|}{\langle\boldsymbol{\beta}^*|\boldsymbol{\alpha}\rangle}, \tag{S50}$$

where $|\alpha\rangle$ and $|\beta^*\rangle$ are coherent states. The claim that we now prove is that the quantum state $|\psi\rangle$ written as

$$|\psi\rangle=\int_{\mathbb{R}}dxF(x)|x\rangle\quad\text{with}\quad F(x)=\frac{(1+\gamma)^{1/4}}{\sqrt{\pi\gamma}}\exp\big(-x^2/\gamma\big), \tag{S51}$$

is a single-mode squeezed state. Using

$$\langle\alpha|\beta\rangle=\exp\left(\alpha^*\beta-\frac{1}{2}|\alpha|^2-\frac{1}{2}|\beta|^2\right), \tag{S52}$$

Eq. (S51) leads to Eq. (S49). Thus, it suffices to show the claim to prove that Eq. (S49) is a positive P-representation of a squeezed vacuum state.

To show that $|\psi\rangle$ is a squeezed vacuum state, we consider the characteristic function of $|\psi\rangle$,

$$\chi_\psi(\eta)=\mathrm{Tr}\left[|\psi\rangle\langle\psi|\exp\Big(\eta\hat{a}^\dagger-\eta^*\hat{b}\Big)\right]=\int_{\mathbb{R}^2}dxdyF(x)F(y)\exp\left[-\frac{1}{2}|\eta|^2+\eta y-\eta^*x-\frac{1}{2}(x-y)^2\right] \tag{S53}$$

$$=\frac{\sqrt{1+\gamma}}{\pi\gamma}\int_{\mathbb{R}^2}dxdy\exp\left[-\frac{x^2+y^2}{\gamma}-\frac{1}{2}|\eta|^2+\eta y-\eta^*x-\frac{1}{2}(x-y)^2\right] \tag{S54}$$

$$=\exp\left[-\frac{|\eta\gamma-\eta^*(\gamma+2)|^2}{8(1+\gamma)}\right]. \tag{S55}$$

By changing the variable with real numbers, $\eta \equiv q + ip$, we obtain the following expression:

$$\chi_\psi(q,p) = \exp\left[-\frac{q^2 + (1+\gamma)^2 p^2}{2(1+\gamma)}\right]. \tag{S56}$$

Thus, it is the characteristic function of a multivariate normal distribution of a covariance matrix,

$$\Sigma = \mathrm{diag}\left(\frac{1}{1+\gamma}, 1+\gamma\right). \tag{S57}$$

Since the covariance matrix of a p-squeezed vacuum state with a squeezing parameter r is $\mathrm{diag}(e^{2r}, e^{-2r})$, it shows that $|\psi\rangle$ is a p-squeezed vacuum state and the relation $\gamma + 1 = e^{2r}$.

We note that the above expression holds only if $r > 0$. Especially when $r = 0$, i.e., when the state is vacuum, the expression is not well defined, but the positive P-representation can be obtained by using Eq. (S38):

$$P_0(\alpha,\beta) = \frac{1}{4\pi^2}\exp\left[-\frac{1}{2}(|\alpha|^2 + |\beta|^2)\right], \tag{S58}$$

where α and β are complex. Since the P-representation is still Gaussian, one may easily generalize the analytical result of the Fourier components $\tilde{G}(k)$ for the vacuum, while nonzero-squeezing-parameter cases are more generic for molecular vibronic spectra. In addition, for the purpose of the Gaussian boson sampling setting, or the molecular vibronic spectra, squeezing parameters $r > 0$ is sufficient because the rotation part can be moved to the linear-optical circuit. Nonetheless, squeezed vacuum states in other directions can be obtained by simply rotating the state) as

$$\hat{\rho} \to \hat{R}(\phi)\hat{\rho}\hat{R}^\dagger(\phi) = \int_{\mathbb{R}^2} P(\alpha,\beta)\hat{\Lambda}(\alpha e^{i\phi}, \beta e^{-i\phi}) d\alpha d\beta, \tag{S59}$$

where $\hat{R}(\phi) = e^{-i\hat{a}^\dagger \hat{a}\phi}$ is the rotation operator. Thus, if we change the variables, the domain for the integral may not be the real plane.

S5. EXACT SOLUTION OF FOURIER COMPONENTS OF MOLECULAR VIBRONIC SPECTRA

Here, we show the derivation of the analytic solution of Fourier components of the vibronic spectra. Let us first write the initial positive P-representation as (see Sec. S3 for more details.) [8]

$$P_{\mathrm{in}}(\boldsymbol{x},\boldsymbol{y}) = \prod_{i=1}^{M}\left[\frac{\sqrt{1+\gamma_i}}{\pi\gamma_i}\exp\left(-(x_i^2+y_i^2)(\gamma_i^{-1}+1/2) + x_i y_i\right)\right] \tag{S60}$$

$$= \mathcal{N}\exp\left[-\boldsymbol{x}.A.\boldsymbol{x} - \boldsymbol{y}.B.\boldsymbol{y} - \boldsymbol{x}.C.\boldsymbol{y} - \boldsymbol{y}.C^{\mathrm{T}}.\boldsymbol{x}\right], \tag{S61}$$

where we define

$$A \equiv \mathrm{diag}(\gamma_i^{-1}+1/2)_{i=1}^M, \quad B \equiv \mathrm{diag}(\gamma_i^{-1}+1/2)_{i=1}^M, \quad C \equiv -\mathbb{1}/2, \quad \mathcal{N} \equiv \prod_{i=1}^{M}\frac{\sqrt{1+\gamma_i}}{\pi\gamma_i}. \tag{S62}$$

Here, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are real M-dimensional vectors. From the main text, we have found the expression of the Fourier components of molecular vibronic spectra,

$$\tilde{G}(k) = \langle e^{-ik\theta\hat{\boldsymbol{n}}\cdot\boldsymbol{\omega}}\rangle, \tag{S63}$$

where the expectation value is over the output state of Gaussian boson sampling. By using the relation $e^{-\gamma\hat{n}} =: e^{\hat{n}(e^{-\gamma}-1)} :$ [9], we can write the operators as a normal-ordered form:

$$\tilde{G}(k) = \left\langle : \exp\left[\sum_{i=1}^{M}\hat{n}_i(e^{-ik\theta\omega_i} - 1)\right] : \right\rangle. \tag{S64}$$

Now, we compute the expectation value,

$$\tilde{G}(k) = \left\langle : \exp\left[\sum_{i=1}^{M} \hat{n}_i (e^{i\phi_i} - 1)\right] : \right\rangle = \left\langle \exp\left[\sum_{i=1}^{M} x_i' y_i' (e^{i\phi_i} - 1)\right] \right\rangle_P \tag{S65}$$

$$= \left\langle \exp\left[\sum_{i=1}^{M} \left(\sum_{j=1}^{M} U_{ij} x_j + x_{0i}\right) \left(\sum_{k=1}^{M} U_{ik}^* y_k + y_{0i}\right) (e^{i\phi_i} - 1)\right] \right\rangle_P \tag{S66}$$

$$= \langle \exp\left[\boldsymbol{x}.\mathcal{U}.\boldsymbol{y} + \boldsymbol{a}.\boldsymbol{x} + \boldsymbol{b}.\boldsymbol{y} + c_0\right] \rangle_P \tag{S67}$$

$$= \mathcal{N} \int d\boldsymbol{x} d\boldsymbol{y} \exp\left[-\boldsymbol{x}.A.\boldsymbol{x} - \boldsymbol{y}.B.\boldsymbol{y} - \boldsymbol{x}.\tilde{C}.\boldsymbol{y} - \boldsymbol{y}.\tilde{C}^{\mathrm{T}}.\boldsymbol{x} + \boldsymbol{a}.\boldsymbol{x} + \boldsymbol{b}.\boldsymbol{y} + c_0\right] \tag{S68}$$

$$= \mathcal{N} \int d\boldsymbol{q} \exp\left[-\boldsymbol{q}.Q.\boldsymbol{q}/2 + \boldsymbol{c}.\boldsymbol{q} + c_0\right], \tag{S69}$$

where $\phi_i \equiv -k\theta\omega_i$,

$$\mathcal{U} \equiv U^{\mathrm{T}} \Phi U^*, \quad \Phi \equiv \mathrm{diag}(e^{i\phi_1} - 1, \ldots, e^{i\phi_M} - 1), \quad \tilde{C} \equiv C - \mathcal{U}/2, \quad Q = 2\begin{pmatrix} A & \tilde{C} \\ \tilde{C}^{\mathrm{T}} & B \end{pmatrix}, \tag{S70}$$

$$\boldsymbol{c} \equiv \begin{pmatrix} \boldsymbol{a} \\ \boldsymbol{b} \end{pmatrix}, \quad \boldsymbol{a} = U^{\mathrm{T}} \Phi \boldsymbol{y}_0, \quad \boldsymbol{b} = U^{\dagger} \Phi \boldsymbol{x}_0, \quad c_0 = \boldsymbol{x}_0 \Phi \boldsymbol{y}_0, \tag{S71}$$

and $\langle \cdot \rangle_P$ is the average over the positive P-representation. We note that the displacement occurs after the linear-optical unitary, which matches the displacement $\boldsymbol{\delta}/\sqrt{2}$ in the Doktorov transformation, i.e., $\boldsymbol{x}_0 = \boldsymbol{y}_0^* = \boldsymbol{\delta}/\sqrt{2}$ Here, we note that Q's real part is positive-definite. To see this, we decompose Q as

$$Q = 2\mathbb{1}_2 \otimes \Gamma^{-1} + \begin{pmatrix} \mathbb{1}_M & -\mathbb{1}_M \\ -\mathbb{1}_M & \mathbb{1}_M \end{pmatrix} - \begin{pmatrix} 0 & \mathcal{U} \\ \mathcal{U}^{\mathrm{T}} & 0 \end{pmatrix} = 2\mathbb{1}_2 \otimes \Gamma^{-1} + \begin{pmatrix} \mathbb{1}_M & -U^{\mathrm{T}} \mathrm{diag}(e^{i\phi_j})_{j=1}^{M} U^* \\ -U^{\dagger} \mathrm{diag}(e^{i\phi_j})_{j=1}^{M} U & \mathbb{1}_M \end{pmatrix}, \tag{S72}$$

where $\Gamma \equiv \mathrm{diag}(\gamma_1, \ldots, \gamma_M)$. Now, we can write the real part of Q as

$$\mathrm{Re}(Q) = 2\mathbb{1}_2 \otimes \Gamma^{-1} + \begin{pmatrix} \mathbb{1}_M & L \\ L^{\mathrm{T}} & \mathbb{1}_M \end{pmatrix}, \tag{S73}$$

where $L \equiv \mathrm{Re}(-U^{\mathrm{T}} \mathrm{diag}(e^{i\phi_j})_{j=1}^{M} U^*)$. Since $\|L\|_{\mathrm{op}} \leq 1$, $\mathrm{Re}(Q)$ is positive-definite [10]. Also, it is well known that the Gaussian integral in Eq. (S69) converges if the real part of the complex symmetric matrix Q is positive-definite. Therefore, the integral always converges.

Using the Gaussian integral formula,

$$\int d\boldsymbol{q} \exp[-\boldsymbol{q}.Q.\boldsymbol{q}/2 + \boldsymbol{c}.\boldsymbol{q}] = \sqrt{\frac{(2\pi)^{2M}}{|Q|}} \exp\left(\frac{1}{2}\boldsymbol{c}.Q^{-1}.\boldsymbol{c}\right), \tag{S74}$$

we obtain the analytical form of the Fourier components,

$$\tilde{G}(k) = \mathcal{N} \sqrt{\frac{(2\pi)^{2M}}{|Q|}} \exp\left(\frac{1}{2}\boldsymbol{c}.Q^{-1}.\boldsymbol{c} + c_0\right), \tag{S75}$$

where

$$Q = \begin{pmatrix} 2\Gamma^{-1} + \mathbb{1}_M & -U^{\mathrm{T}} \mathrm{diag}(e^{i\phi_j})_{j=1}^{M} U^* \\ -U^{\dagger} \mathrm{diag}(e^{i\phi_j})_{j=1}^{M} U & 2\Gamma^{-1} + \mathbb{1}_M \end{pmatrix}. \tag{S76}$$

We note that for simplicity, we have assumed nonzero squeezing, which is generic for molecular vibronic spectra, while a similar expression can be obtained when there is a zero squeezing because the positive P-function is still Gaussian (See Eq. (S58)).

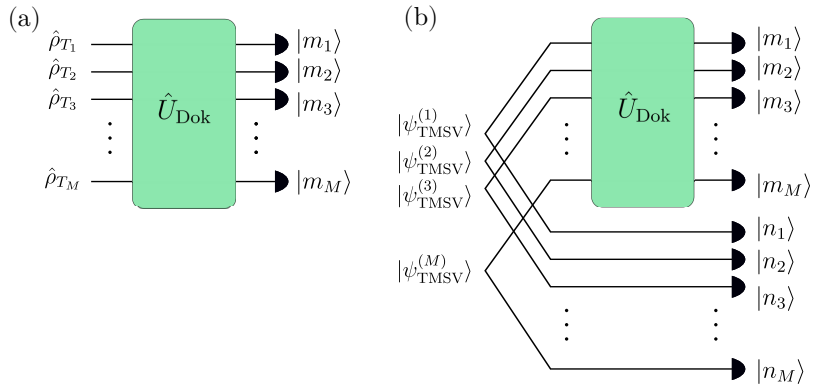

FIG. S1. Transformation of molecular vibronic spectra (a) at finite temperature to molecular vibronic spectra at zero temperature. (b) By introducing additional M modes and two-mode squeezing operations, one may treat the problem as at zero temperature. For finite temperature case, the cost is only the additional M modes and we can exploit the same solution as at zero temperature.

S6. EXACT SOLUTION FOR MOLECULAR VIBRONIC SPECTRA AT FINITE TEMPERATURE

We now show that our method is applicable for molecular vibronic spectra even at finite temperature. At finite temperature T, the molecular vibronic spectra is written as

$$\mathrm{FCP}(\Omega) = \sum_{\boldsymbol{n},\boldsymbol{m}=\boldsymbol{0}}^{\infty} p(\boldsymbol{n},\boldsymbol{m})\delta\left(\Omega - (\boldsymbol{\omega}'\cdot\boldsymbol{m} - \boldsymbol{\omega}\cdot\boldsymbol{n})\right), \tag{S77}$$

where

$$p(\boldsymbol{n},\boldsymbol{m}) = p_T(\boldsymbol{n})|\langle\boldsymbol{n}|\hat{U}_{\rm Dok}|\boldsymbol{m}\rangle|^2. \tag{S78}$$

Here, $p_T(\boldsymbol{n}) \equiv \langle\boldsymbol{n}|\hat{\rho}_T|\boldsymbol{n}\rangle$ is the probability of photon number $|\boldsymbol{n}\rangle$ from an M-mode thermal state $\hat{\rho}_T$ at temperature T (in general we do not need to assume the same temperature.). In Ref. [11], it was shown that one may construct a Gaussian boson sampling for the finite temperature case. The key idea is to add M auxiliary modes and prepare two-mode squeezed states to simulate thermal states and detect the additional modes as well, which is depicted in Fig. S1. In this case, the original M modes' output photons correspond to $\boldsymbol{m}$ and the additional M modes' output photons corresponds to $\boldsymbol{n}$. Therefore, we can again use the same framework to obtain the solution. We can rewrite the FCP as

$$\mathrm{FCP}(\Omega) = \sum_{\boldsymbol{m},\boldsymbol{n}=\boldsymbol{0}}^{\infty} |\langle\boldsymbol{m},\boldsymbol{n}|\hat{U}'_{\rm Dok}|\boldsymbol{0}\rangle|^2\delta\left(\Omega - (\boldsymbol{\omega}'\cdot\boldsymbol{m} - \boldsymbol{\omega}\cdot\boldsymbol{n})\right), \tag{S79}$$

which is the same form as the zero temperature case except for the negative part in delta function. Here, $\hat{U}'_{\rm Dok} = \hat{U}_{\rm Dok}\hat{U}_{\rm TMSV}$ is obtained by applying two-mode squeezing operators. Again, the Fourier transform of the FCP at finite temperature is given by

$$\tilde{G}(k) = \langle e^{-ik\hat{n}\cdot\boldsymbol{\omega}''\theta}\rangle, \tag{S80}$$

where $\boldsymbol{\omega}'' = (\boldsymbol{\omega}'^{\mathrm{T}}, -\boldsymbol{\omega}^{\mathrm{T}})^{\mathrm{T}}$. Therefore, the analytic solution for Fourier components can easily be found. Note that without introducing additional modes, one can directly use the positive P-representation for a squeezed thermal input state and follow the same procedure as the zero temperature case.

S7. COMPUTATION OF LOW-RANK LOOP HAFNIAN

Here, we generalize the algorithm of computing the hafnian of a low-rank symmetric matrix to loop hafnian [12]. The relevant matrix of which we compute the loop hafnian is $\tilde{\Sigma}$, which is obtained by replacing the diagonal elements

of a symmetric Σ by μ vector, and we assume that the symmetric matrix Σ has a rank r; thus, we can find $G \in \mathbb{C}^{n\times r}$ such that $GG^{\mathrm{T}} = \Sigma$. We then introduce indeterminates $x_1, \ldots, x_r$, and consider the multivariate polynomial

$$q(x_1, \ldots, x_r) = \prod_{i=1}^{n} \left(\sum_{j=1}^{r} g_{i,j} x_j + \mu_j \right). \tag{S81}$$

Note that the difference from the hafnian case is the additional term μ_j. We let $\mathcal{P}_{2s,r}$ be the set of integer r-partitions of $2s$, i.e., tuples $(p_1, p_2, \ldots, p_r)$ such that $p_i \geq 0$ for all i and $\sum_i p_i = 2s$, and let $\mathcal{E}_{2s,r}$ be the subset of $\mathcal{P}_{s,r}$ such that p_i is even for all i. We then expand the polynomial and identify the coefficients λ_p as

$$\tilde{q}(x_1, \ldots, x_r) = \sum_{s=0}^{[n/2]} \sum_{p=(p_1,\ldots,p_{2s})\in\mathcal{P}_{2s,r}} \lambda_p \prod_{i=1}^{r} x_i^{p_i}. \tag{S82}$$

where we dropped a polynomial composed of odd number of x_i's from $q(x_1, \ldots, x_r)$. Since

$$|\mathcal{P}_{2s,r}| = \binom{2s+r-1}{r-1} \quad \text{and} \quad |\mathcal{E}_{2s,r}| = \binom{s+r-1}{r-1}, \tag{S83}$$

The coefficients λ_p can be found in $|\mathcal{P}_{n,r}|\mathrm{poly}(n)$ running time. Meanwhile, the loop hafnian can be written as

$$\mathrm{lhaf}(\tilde{\Sigma}) = \sum_{s=0}^{[n/2]} \sum_{e\in\mathcal{E}_{2s,r}} \lambda_e \prod_{i=1}^{r} (e_i - 1)!!, \tag{S84}$$

where $a!! = 1$ if $a \leq 0$. For given coefficients λ_e, the loop hafnian can be computed in $|\mathcal{E}_{n,r}|\mathrm{poly}(n)$ time. Thus, when Σ is a rank r matrix, one can compute its relevant loop hafnian $\mathrm{lhaf}(\tilde{\Sigma})$ can be evaluated in $\binom{n+r-1}{r-1}\mathrm{poly}(n)$.

S8. FOURIER COMPONENTS OF MOLECULAR VIBRONIC SPECTRA

In this section, we simplify the Fourier component of molecular vibronic spectra. We focus on the displaced squeezed Fock state input $\hat{D}(\boldsymbol{\alpha})\hat{S}(\boldsymbol{r}_0)|\boldsymbol{n}\rangle$. In this case, the Fourier component can be written as

$$\tilde{G}(k) = \langle \boldsymbol{n}|\hat{S}^\dagger(\boldsymbol{r}_0)\hat{D}^\dagger(\boldsymbol{\alpha})\hat{U}^\dagger e^{-ik\theta\hat{\boldsymbol{n}}\cdot\boldsymbol{\omega}}\hat{U}\hat{D}(\boldsymbol{\alpha})\hat{S}(\boldsymbol{r}_0)|\boldsymbol{n}\rangle \equiv \langle \boldsymbol{n}|\hat{W}|\boldsymbol{n}\rangle = \langle \boldsymbol{n}|\hat{D}(\boldsymbol{\xi})\hat{U}_{\mathrm{lin}2}\hat{S}(\boldsymbol{r})\hat{U}_{\mathrm{lin}1}|\boldsymbol{n}\rangle, \tag{S85}$$

where we have used Bloch-Messiah decomposition of the Gaussian unitary matrix $\hat{W}$.

Before we derive the most general case's formula, we first consider a simpler case. First, we consider $\langle n|\hat{S}(r)|n\rangle$ for a single mode:

$$\langle n|\hat{S}(r)|n\rangle = \frac{1}{\pi^2}\int d\alpha d\beta \langle n|\alpha\rangle\langle\alpha|\hat{S}(r)|\beta\rangle\langle\beta|n\rangle = \frac{1}{n!\pi^2}\int d\alpha d\beta \alpha^n \beta^{*n} \exp\left[-\frac{1}{2}\left(|\alpha|^2 + |\beta|^2\right)\right]\langle\alpha|\hat{S}(r)|\beta\rangle. \tag{S86}$$

Note that

$$\begin{aligned}\langle\alpha|\hat{S}(r)|\beta\rangle &= \frac{1}{\sqrt{\cosh r}}\exp\Bigg[-\frac{1}{2}\Big(|\alpha - \beta\cosh r - \beta^*\sinh r|^2 - \tanh r(\alpha^* - \beta^*\cosh r - \beta\sinh r)^2 \\ &\quad + \alpha(\beta^*\cosh r + \beta\sinh r) - \alpha^*(\beta\cosh r + \beta^*\sinh r)\Big)\Bigg] \qquad \text{(S87)} \\ &= \frac{1}{\sqrt{\cosh r}}\exp\left[-\frac{1}{2}v(\alpha,\beta)^{\mathrm{T}}.\begin{pmatrix} 0 & 1/2 & 0 & 0 \\ 1/2 & -\tanh r & -\mathrm{sech}\, r & 0 \\ 0 & -\mathrm{sech}\, r & \tanh r & 1/2 \\ 0 & 0 & 1/2 & 0 \end{pmatrix}.v(\alpha,\beta)\right]. \qquad \text{(S88)}\end{aligned}$$

Here, we defined $v(\alpha,\beta) \equiv (\alpha, \alpha^*, \beta, \beta^*)^{\mathrm{T}}$. Therefore, Eq. (S86) can be rewritten as

$$\langle n|\hat{S}(r)|n\rangle = \frac{1}{n!\pi^2\sqrt{\cosh r}}\int d\alpha d\beta \alpha^n\beta^{*n}\exp\left[-\frac{1}{2}v(\alpha,\beta)^{\mathrm{T}}.\Sigma_s^{-1}.v(\alpha,\beta)\right], \tag{S89}$$

where we have defined

$$\Sigma_s^{-1} \equiv \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & -\tanh r & -\operatorname{sech} r & 0 \\ 0 & -\operatorname{sech} r & \tanh r & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad \Sigma_s = \begin{pmatrix} \tanh r & 0 & 1 & \operatorname{sech} r \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ \operatorname{sech} r & 1 & 0 & -\tanh r \end{pmatrix}. \tag{S90}$$

Finally, using Wick's theorem [13], we obtain

$$\langle n|\hat{S}(r)|n\rangle = \frac{\langle \alpha^n \beta^{*n}\rangle}{n!\sqrt{\cosh r}} = \frac{\operatorname{haf}((\Sigma_s)_n)}{n!\sqrt{\cosh r}}, \tag{S91}$$

where $(\Sigma_s)_n$ is obtained by repeating the first and fourth rows and columns n times, which accounts for α and β^*, respectively.

Using the same method, we derive a more general case:

$$\langle \boldsymbol{n}|\hat{D}(\boldsymbol{\xi})\hat{U}_{\rm lin2}\hat{S}(\boldsymbol{r})\hat{U}_{\rm lin1}|\boldsymbol{n}\rangle \tag{S92}$$

$$= \frac{1}{\pi^{4M}} \int d\boldsymbol{\alpha} d\boldsymbol{\gamma} d\boldsymbol{\omega} d\boldsymbol{\beta} \langle \boldsymbol{n}|\hat{D}(\boldsymbol{\xi})|\boldsymbol{\alpha}\rangle\langle \boldsymbol{\alpha}|\hat{U}_{\rm lin2}|\boldsymbol{\gamma}\rangle\langle \boldsymbol{\gamma}|\hat{S}(\boldsymbol{r})|\boldsymbol{\omega}\rangle\langle \boldsymbol{\omega}|\hat{U}_{\rm lin1}|\boldsymbol{\beta}\rangle\langle \boldsymbol{\beta}|\boldsymbol{n}\rangle \tag{S93}$$

$$\begin{aligned} &= \frac{1}{\boldsymbol{n}!\pi^{4M}} \int d\boldsymbol{\alpha} d\boldsymbol{\gamma} d\boldsymbol{\omega} d\boldsymbol{\beta} \prod_{i=1}^{M} [(\alpha_i + \xi_i)\beta_i^*]^{n_i} \\ &\quad \times \exp\left[-\frac{1}{2}\left(|\boldsymbol{\alpha} + \boldsymbol{\xi}|^2 + |\boldsymbol{\alpha}|^2 + 2|\boldsymbol{\beta}|^2 + |\boldsymbol{\gamma}|^2 + |\boldsymbol{\omega}|^2 \right) + \boldsymbol{\alpha}^*.U_{\rm lin2}.\boldsymbol{\gamma} + \boldsymbol{\omega}^*.U_{\rm lin1}.\boldsymbol{\beta} + \frac{1}{2}(\boldsymbol{\xi}.\boldsymbol{\alpha}^* - \boldsymbol{\xi}^*.\boldsymbol{\alpha}) \right] \langle \gamma|\hat{S}(\boldsymbol{r})|\boldsymbol{\omega}\rangle \end{aligned} \tag{S94}$$

$$\begin{aligned} &= \frac{1}{\boldsymbol{n}!\pi^{4M}\prod_{i=1}^{M}\sqrt{\cosh r_i}} \int d\boldsymbol{\alpha} d\boldsymbol{\gamma} d\boldsymbol{\omega} d\boldsymbol{\beta} \prod_{i=1}^{M} [(\alpha_i + \xi_i)\beta_i^*]^{n_i} \exp\left[-|\boldsymbol{\alpha}|^2 - |\boldsymbol{\beta}|^2 - \frac{1}{2}|\boldsymbol{\xi}|^2 - \boldsymbol{\alpha}.\boldsymbol{\xi}^* + \boldsymbol{\alpha}^*.U_{\rm lin2}.\boldsymbol{\gamma} + \boldsymbol{\omega}^*.U_{\rm lin1}.\boldsymbol{\beta} \right] \\ &\quad \times \exp\left[-\frac{1}{2} v(\boldsymbol{\gamma},\boldsymbol{\omega})^{\rm T}.\Sigma_s^{-1}.v(\boldsymbol{\gamma},\boldsymbol{\omega}) \right] \end{aligned} \tag{S95}$$

$$= \frac{1}{\boldsymbol{n}!\pi^{4M}Z} \int d\boldsymbol{\alpha} d\boldsymbol{\gamma} d\boldsymbol{\omega} d\boldsymbol{\beta} \prod_{i=1}^{M} [(\alpha_i + \xi_i)\beta_i^*]^{n_i} \exp\left(-\frac{1}{2}|\boldsymbol{\xi}|^2 - \boldsymbol{\alpha}.\boldsymbol{\xi}^* \right) \exp\left(-\frac{1}{2} v(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\omega})^{\rm T}.\Sigma^{-1}.v(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\omega}) \right) \tag{S96}$$

$$= \frac{1}{\boldsymbol{n}!\pi^{4M}Z} \int d\boldsymbol{\alpha} d\boldsymbol{\gamma} d\boldsymbol{\omega} d\boldsymbol{\beta} \prod_{i=1}^{M} [(\alpha_i + \xi_i + \zeta_i^{\alpha})(\beta_i + \zeta_i^{\beta^*})^*]^{n_i} \exp\left(-\frac{1}{2} v(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\omega})^{\rm T}.\Sigma^{-1}.v(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\omega}) \right) \tag{S97}$$

$$= \frac{\operatorname{lhaf}(\tilde{\Sigma}_{\boldsymbol{n}})}{\boldsymbol{n}!Z}. \tag{S98}$$

Here, we have introduced various parameters, which are defined as follows: First, we have defined a complex vector as

$$v(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\omega}) = (\alpha_1,\ldots,\alpha_M,\alpha_1^*,\ldots,\alpha_M^*,\beta_1,\ldots,\beta_M,\beta_1^*,\ldots,\beta_M^*,\gamma_1,\ldots,\gamma_M,\gamma_1^*,\ldots,\gamma_M^*,\omega_1,\ldots,\omega_M,\omega_1^*,\ldots,\omega_M^*), \tag{S99}$$

and the complex symmetric covariance matrix Σ as

$$\Sigma^{-1} = \begin{pmatrix} \Sigma_x & -\mathcal{W} \\ -\mathcal{W}^{\rm T} & \Sigma_s^{-1} \end{pmatrix}, \quad \mathcal{U}_{\rm lin1} = \begin{pmatrix} 0 & U_{\rm lin1}^{\rm T} \\ 0 & 0 \end{pmatrix}, \quad \mathcal{U}_{\rm lin2} = \begin{pmatrix} 0 & 0 \\ U_{\rm lin2} & 0 \end{pmatrix}, \quad \mathcal{W} = \begin{pmatrix} \mathcal{U}_{\rm lin2} & 0 \\ 0 & \mathcal{U}_{\rm lin1} \end{pmatrix}, \tag{S100}$$

$$\Sigma_x = \begin{pmatrix} \sigma_x \otimes \mathbb{1} & 0 \\ 0 & \sigma_x \otimes \mathbb{1} \end{pmatrix}, \quad \Sigma_s = \begin{pmatrix} \tanh \boldsymbol{r} & \mathbb{1}_M & 0 & \operatorname{sech} \boldsymbol{r} \\ \mathbb{1}_M & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbb{1}_M \\ \operatorname{sech} \boldsymbol{r} & 0 & \mathbb{1}_M & -\tanh \boldsymbol{r} \end{pmatrix}, \tag{S101}$$

where σ_x is the Pauli x matrix. Here, using the equality

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} + A^{-1}B(D - CA^{-1}B)^{-1}CA^{-1} & -A^{-1}B(D - CA^{-1}B)^{-1} \\ -(D - CA^{-1}B)^{-1}CA^{-1} & (D - CA^{-1}B)^{-1} \end{pmatrix}, \tag{S102}$$

we obtain

$$\Sigma = \begin{pmatrix} U_{\rm lin2}\tanh\boldsymbol{r}U_{\rm lin2}^{\rm T} & \mathbb{1} & 0 & U_{\rm lin2}\,\mathrm{sech}\,\boldsymbol{r}U_{\rm lin1} & U_{\rm lin2}\tanh\boldsymbol{r} & U_{\rm lin2} & 0 & U_{\rm lin2}\,\mathrm{sech}\,\boldsymbol{r} \\ \mathbb{1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbb{1} & 0 & 0 & 0 & 0 \\ U_{\rm lin1}^{\rm T}\,\mathrm{sech}\,\boldsymbol{r}U_{\rm lin2}^{\rm T} & 0 & \mathbb{1} & -U_{\rm lin1}^{\rm T}\tanh\boldsymbol{r}U_{\rm lin1} & U_{\rm lin1}^{\rm T}\,\mathrm{sech}\,\boldsymbol{r} & 0 & U_{\rm lin1}^{\rm T} & -U_{\rm lin1}^{\rm T}\tanh\boldsymbol{r} \\ \tanh\boldsymbol{r}U_{\rm lin2}^{\rm T} & 0 & 0 & \mathrm{sech}\,\boldsymbol{r}U_{\rm lin1} & \tanh\boldsymbol{r} & \mathbb{1} & 0 & \mathrm{sech}\,\boldsymbol{r} \\ U_{\rm lin2}^{\rm T} & 0 & 0 & 0 & \mathbb{1} & 0 & 0 & 0 \\ 0 & 0 & 0 & U_{\rm lin1} & 0 & 0 & 0 & \mathbb{1} \\ \mathrm{sech}\,\boldsymbol{r}U_{\rm lin2}^{\rm T} & 0 & 0 & -\tanh\boldsymbol{r}U_{\rm lin1} & \mathrm{sech}\,\boldsymbol{r} & 0 & \mathbb{1} & -\tanh\boldsymbol{r} \end{pmatrix}. \tag{S103}$$

Note that the relevant part for the loop hafnian is the ones that are related to $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}^*$ (for this reason, we dropped other parts in the main text), which is

$$\Sigma_{\boldsymbol{\alpha},\boldsymbol{\beta}^*} = \begin{pmatrix} U_{\rm lin2}\tanh\boldsymbol{r}U_{\rm lin2}^{\rm T} & U_{\rm lin2}\,\mathrm{sech}\,\boldsymbol{r}U_{\rm lin1} \\ U_{\rm lin1}^{\rm T}\,\mathrm{sech}\,\boldsymbol{r}U_{\rm lin2}^{\rm T} & -U_{\rm lin1}^{\rm T}\tanh\boldsymbol{r}U_{\rm lin1} \end{pmatrix}, \tag{S104}$$

and $\tilde{\Sigma}_{\boldsymbol{n}}$ is the matrix that is obtained by replacing the diagonal element of $\Sigma_{\boldsymbol{\alpha},\boldsymbol{\beta}^*}$ by $(\boldsymbol{\xi}+\boldsymbol{\zeta}^{\alpha},\boldsymbol{\zeta}^{\beta^*})$ and repeating ith row and column of each block matrix by n_i. Here,

$$\boldsymbol{\zeta} = -\Sigma.\boldsymbol{\xi}_{\rm ex}^*, \quad \begin{pmatrix} \boldsymbol{\zeta}^{\alpha} \\ \boldsymbol{\zeta}^{\beta^*} \end{pmatrix} = -\begin{pmatrix} U_{\rm lin2}\tanh\boldsymbol{r}U_{\rm lin2}^{\rm T}.\boldsymbol{\xi}^* \\ U_{\rm lin1}^{\rm T}\,\mathrm{sech}\,\boldsymbol{r}U_{\rm lin2}^{\rm T}.\boldsymbol{\xi}^* \end{pmatrix}, \tag{S105}$$

where $\boldsymbol{\xi}_{\rm ex}^*$ is interpreted as an extended vector for other parts than $\boldsymbol{\alpha}$ such as $\boldsymbol{\alpha}^*$, $\boldsymbol{\beta}$ to be zero, i.e., $\boldsymbol{\xi}_{\rm ex}^* = (\boldsymbol{\xi}^*, 0, \ldots, 0)$, and $(\boldsymbol{\zeta}^{\alpha},\boldsymbol{\zeta}^{\beta^*})$ are restriction to the space of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}^*$. Finally, the normalization factor Z can be compactly written as

$$Z^{-1} = \langle 0|\hat{D}(\boldsymbol{\xi})\hat{U}_{\rm lin2}\hat{S}(\boldsymbol{r})\hat{U}_{\rm lin1}|0\rangle = \exp\left(\frac{1}{2}\boldsymbol{\xi}_{\rm ex}^{\rm T}\Sigma\boldsymbol{\xi}_{\rm ex}\right)/\prod_{i=1}^{M}\sqrt{\cosh r_i}, \tag{S106}$$

which gives the analytic expression for the zero temperature case as we obtained in Eq. (S75).

We note that

$$\Sigma_{\boldsymbol{\alpha},\boldsymbol{\beta}^*} = \begin{pmatrix} U_{\rm lin2} & 0 \\ 0 & U_{\rm lin1}^{\rm T} \end{pmatrix}\begin{pmatrix} \tanh\boldsymbol{r} & \mathrm{sech}\,\boldsymbol{r} \\ \mathrm{sech}\,\boldsymbol{r} & -\tanh\boldsymbol{r} \end{pmatrix}\begin{pmatrix} U_{\rm lin2}^{\rm T} & 0 \\ 0 & U_{\rm lin1} \end{pmatrix} \tag{S107}$$

and that the operator norm is 1 regardless of $U_{\rm lin1}$, $U_{\rm lin2}$ and $\boldsymbol{r}$. A more general matrix element such as $\langle\boldsymbol{n}|\hat{D}(\boldsymbol{\xi})\hat{U}_{\rm lin2}\hat{S}(\boldsymbol{r})\hat{U}_{\rm lin1}|\boldsymbol{m}\rangle$ can be similarly obtained [14].

As a remark, when $\boldsymbol{r}=0$, the relevant matrix becomes

$$\begin{pmatrix} 0 & U_{\rm lin2}U_{\rm lin1} \\ U_{\rm lin2}^{\rm T}U_{\rm lin1}^{\rm T} & 0 \end{pmatrix}. \tag{S108}$$

Since for a matrix A,

$$\mathrm{haf}\begin{pmatrix} 0 & A \\ A^T & 0 \end{pmatrix} = \mathrm{Per}A \tag{S109}$$

we can see that when squeezing parameters are zero, we recover the Fock-state boson sampling case.

S9. APPROXIMATION ALGORITHM OF HAFNIAN

We provide an algorithm to approximate a hafnian. First, from Lemma 1 in Ref. [13],

$$x_1^{n_1}x_2^{n_2}\cdots x_M^{n_M} = \frac{1}{n!}\sum_{v_1=0}^{n_1}\cdots\sum_{v_M=0}^{n_M}(-1)^{\sum_{i=1}^{M}v_i}\binom{n_1}{v_1}\cdots\binom{n_M}{v_M}\left(\sum_{i=1}^{M}h_ix_i\right)^n, \tag{S110}$$

where $h_i = n_i/2 - v_i$. Using the lemma, Proposition 1 in Ref. [13] states that for $z = (z_1, \ldots, z_M)^{\mathrm{T}} \sim \mathcal{N}(0, \Sigma)$, where Σ is a real symmetric matrix and z can be a complex variable vector,

$$\mathrm{haf}(\Sigma_{\boldsymbol{n}}) = \mathbb{E}_z \left[\prod_{i=1}^{M} z_i^{n_i} \right] = \frac{1}{(n/2)!} \sum_{v_1=0}^{n_1} \cdots \sum_{v_M=0}^{n_M} (-1)^{\sum_{i=1}^{M} v_i} \binom{n_1}{v_1} \cdots \binom{n_M}{v_M} \left(\frac{h^{\mathrm{T}} \Sigma h}{2} \right)^{n/2}, \tag{S111}$$

where $n = n_1 + \cdots + n_M$ and $h = (n_1/2 - v_1, \ldots, n_M/2 - v_M)^{\mathrm{T}}$ and n is even. We note that the equality between the hafnian and the sum in Eq. (S111) holds even for general complex symmetric matrices. Especially when $n_i = 1$ for all $i \in \{1, \ldots, n\}$,

$$\mathbb{E}_z \left[\prod_{i=1}^{n} z_i \right] = \frac{1}{(n/2)!} \sum_{v_1=0}^{1} \cdots \sum_{v_M=0}^{1} (-1)^{\sum_{i=1}^{n} v_i} \left(\frac{h^{\mathrm{T}} \Sigma h}{2} \right)^{n/2} = \frac{1}{2^n} \sum_{v_1=0}^{1} \cdots \sum_{v_n=0}^{1} (-1)^{\sum_{i=1}^{n} v_i} \frac{2^n}{(n/2)!} \left(\frac{h^{\mathrm{T}} \Sigma h}{2} \right)^{n/2} \tag{S112}$$

$$= \mathbb{E}_v \left[(-1)^{\sum_{i=1}^{n} v_i} \frac{2^{n/2}}{(n/2)!} \left(h^{\mathrm{T}} \Sigma h \right)^{n/2} \right], \tag{S113}$$

where $h = (1/2 - v_1, \ldots, 1/2 - v_M)^{\mathrm{T}}$. In this case, $M = n$. Therefore,

$$(-1)^{\sum_{i=1}^{n} v_i} \frac{2^{n/2}}{(n/2)!} \left(h^{\mathrm{T}} \Sigma h \right)^{n/2} \tag{S114}$$

is an unbiased estimator for the hafnian in the sense that

$$\mathbb{E}_v \left[(-1)^{\sum_{i=1}^{n} v_i} \frac{2^{n/2}}{(n/2)!} \left(h^{\mathrm{T}} \Sigma h \right)^{n/2} \right] = \mathrm{haf}(\Sigma). \tag{S115}$$

Note that

$$\left| \frac{2^{n/2}}{(n/2)!} \left(h^{\mathrm{T}} \Sigma h \right)^{n/2} \right| = \left| \frac{2^{n/2}}{(n/2)!} \left(\frac{n}{4} \hat{h}^{\mathrm{T}} \Sigma \hat{h} \right)^{n/2} \right| = \left| \frac{n^{n/2}}{(n/2)! 2^{n/2}} \left(\hat{h}^{\mathrm{T}} \Sigma \hat{h} \right)^{n/2} \right| \leq \frac{(n \|\Sigma\|)^{n/2}}{(n/2)! 2^{n/2}}. \tag{S116}$$

Therefore, in general, the estimate by sampling $\boldsymbol{v}$ and using the estimator has an error bound given by Chernoff bound as

$$\Pr \left[|\tilde{p} - \mathrm{haf}(\Sigma)| > \epsilon \frac{(n\|\Sigma\|)^{n/2}}{(n/2)! 2^{n/2}} \right] \leq 2 \exp \left(-\frac{1}{2} \epsilon^2 T \right). \tag{S117}$$

Here, $\|\Sigma\|$ is the spectral norm, i.e., the maximum singular value.

Our goal is to estimate the Fourier components of molecular vibronic spectra, which is given by

$$\tilde{G}(k) = \frac{\mathrm{haf}(\Sigma_{\boldsymbol{n}})}{Z \boldsymbol{n}!}. \tag{S118}$$

Let us recall that the operator norm of the matrix $\Sigma_{\boldsymbol{\alpha}, \boldsymbol{\beta}^*}$ is 1 (see Eq. (S107)). Since a matrix's submatrix's operator norm is smaller than or equal to the matrix's norm, when the vector $\boldsymbol{n}$ is composed 0 or 1, $\Sigma_{\boldsymbol{n}}$'s operator norm is smaller than or equal to 1. In addition, when $\boldsymbol{n}$ has a component larger than 1, we compute hafnian for a matrix obtained by repeating the row and column. Nevertheless, when repeating the row and column the operator norm does not increase as much as $\boldsymbol{n}!$, i.e.,

$$\frac{\|\Sigma_{\boldsymbol{n}}\|}{\boldsymbol{n}!} \leq \|\Sigma\|. \tag{S119}$$

Therefore, the estimate by sampling $\boldsymbol{x}$ and using the estimator has an error bound given by Chernoff bound as

$$\Pr \left[\left| \tilde{q} - \frac{\mathrm{haf}(\Sigma)}{Z \boldsymbol{n}!} \right| \gtrsim \epsilon \frac{e^{n/2}}{\sqrt{\pi n} Z} \right] \leq 2 \exp \left(-\frac{1}{2} \epsilon^2 T \right), \tag{S120}$$

where $Z = \sqrt{\prod_{i=1}^{n} \cosh r_i}$ and we have used Stirling's formula, $N! \approx \sqrt{2\pi N} (N/e)^N$ (more precisely, $\sqrt{2\pi N} (N/e)^N e^{\frac{1}{12N+1}} < N! < \sqrt{2\pi N} (N/e)^N e^{\frac{1}{12N}}$). Thus, if $Z > e^{n/2}$, the estimation error is smaller than ϵ with exponentially small failure probability if we choose the number of samples as $T = O(1/\epsilon^2)$.

For nonzero mean, one may again employ a similar formula for $z \sim \mathcal{N}(\mu, \Sigma)$ [13]:

$$\mathrm{lhaf}(\Sigma_{\boldsymbol{n}}) = \mathbb{E}_z \left[\prod_{i=1}^{M} z_i^{n_i} \right] = \sum_{v_1=0}^{n_1} \cdots \sum_{v_M=0}^{n_M} (-1)^{\sum_{i=1}^{M} v_i} \binom{n_1}{v_1} \cdots \binom{n_M}{v_M} \sum_{r=0}^{[n/2]} \frac{\left(\frac{h^{\mathrm{T}}\Sigma h}{2}\right)^r (h \cdot \mu)^{n-2r}}{r!(n-2r)!}, \tag{S121}$$

where $n = n_1 + \cdots + n_M$ and $h = (n_1/2 - v_1, \ldots, n_M/2 - v_M)^{\mathrm{T}}$ and $\Sigma_{\boldsymbol{n}}$ is obtained by replacing the diagonal elements of Σ by μ and repeating ith row and column for n_i times. Again, let $n_i = 1$ for all i's. In this case, $M = n$, and

$$\mathrm{lhaf}(\Sigma) = \mathbb{E}_z \left[\prod_{i=1}^{n} z_i \right] = \sum_{v_1=0}^{1} \cdots \sum_{v_n=0}^{1} (-1)^{\sum_{i=1}^{n} v_i} \sum_{r=0}^{[n/2]} \frac{\left(\frac{h^{\mathrm{T}}\Sigma h}{2}\right)^r (h \cdot \mu)^{n-2r}}{r!(n-2r)!} \tag{S122}$$

$$= \frac{1}{2^n [n/2]} \mathbb{E}_{\boldsymbol{v},r} \left[(-1)^{\sum_{i=1}^{n} v_i} \frac{2^n [n/2] \left(\frac{h^{\mathrm{T}}\Sigma h}{2}\right)^r (h \cdot \mu)^{n-2r}}{r!(n-2r)!} \right]. \tag{S123}$$

Here,

$$\left| \frac{2^n [n/2] \left(\frac{h^{\mathrm{T}}\Sigma h}{2}\right)^r (h \cdot \mu)^{n-2r}}{r!(n-2r)!} \right| = \frac{[n/2] n^{n/2} \left(\hat{h}^{\mathrm{T}}\Sigma\hat{h}\right)^r (\hat{h} \cdot \mu)^{n-2r}}{2^r r!(n-2r)!} \leq \frac{[n/2] n^{n/2} \|\Sigma\|^r \|\mu\|^{n-2r}}{2^r r!(n-2r)!} \tag{S124}$$

Therefore, the Chernoff bound gives us

$$\Pr\left[|p - \mathrm{lhaf}(\Sigma)| > \epsilon \frac{[n/2] n^{n/2} \|\Sigma\|^r \|\mu\|^{n-2r}}{2^r r!(n-2r)!} \right] \leq 2 \exp\left(-\frac{1}{2} \epsilon^2 T \right). \tag{S125}$$

Especially when $\|\Sigma\| \leq 1$, which is the case that we are interested in,

$$\Pr\left[|p - \mathrm{lhaf}(\Sigma)| > \epsilon \frac{[n/2] n^{n/2} \|\mu\|^{n-2r}}{2^r r!(n-2r)!} \right] \leq 2 \exp\left(-\frac{1}{2} \epsilon^2 T \right). \tag{S126}$$

The error bound clearly depends on the mean vector μ, which makes the bound more nontrivial than the hafnian case.

S10. EMBEDDING A UNIVERSAL COMPUTING CIRCUIT INTO MOLECULAR VIBRONIC SPECTRA PROBLEM

We first show that we can construct a BQP-complete spectra problem, which is essentially equivalent to simulating a universal computing circuit, in discrete-variable (DV) system. To do that, we consider a quantum circuit $\hat{U}$ applied on n qubits which consists of arbitrary single-qubit gates and CNOT gates with a polynomially large depth. Now, we define the corresponding spectra problem as follows. For a given circuit $\hat{U}$, $\Omega \in \{0, \ldots, \Omega_{\max}\}$, and $\boldsymbol{\omega} \in \mathbb{Z}^n_{\geq 0}$ with $\omega_i \leq O(\mathrm{poly}(n))$, which are defined similarly as boson sampling cases, compute the spectra

$$G(\Omega) = \sum_{\boldsymbol{m} \in \{0,1\}^n} p(\boldsymbol{m}) \delta(\Omega - \boldsymbol{\omega} \cdot \boldsymbol{m}). \tag{S127}$$

We can easily show that one can approximate the single-qubit marginal probabilities if we can approximate the spectra. We first choose $\boldsymbol{\omega}$ as $\omega_i = 1$ for $i = 1$ and $\omega_i = 0$ otherwise. Here, without loss of generality, we will compute the first qubit's marginal probabilities. For this choice, each component of the spectra simply represents the marginal probability:

$$G(\Omega = 1) = p(1), \quad G(\Omega = 0) = p(0). \tag{S128}$$

Hence, if we can approximate $G(\Omega)$, we can approximate the single-qubit marginal probabilities with the same accuracy. Therefore, it proves that approximating the spectra within $O(\mathrm{poly}(1/\epsilon))$ error in DV system is BQP-hard. Also, since the spectra can be approximated with error in $O(\mathrm{poly}(1/\epsilon))$ by simulating the quantum circuit, measuring the qubits and repeating sampling, the problem is in BQP, which shows that it is BQP-complete.

Using this property, we will construct a bosonic circuit whose spectra problem is also BQP-complete. The key idea is to map a qubit-based circuit to a bosonic circuit using the Schwinger boson correspondence between qubit and boson states [15], which is similar to the one in Ref. [16]. Specifically, one can map a single-qubit state $|x\rangle$ as

$$|x\rangle \iff (\hat{a}^\dagger)^x(\hat{b}^\dagger)^{1-x}|0\rangle, \tag{S129}$$

where $|0\rangle$ is vacuum state. Thus, one can map a single-qubit state to a single boson in a two-mode bosonic system, which corresponds to the dual-rail encoding in the Knill-Laflamme-Milburn protocol [17]. We call the two modes as a single site. Similarly, an n-qubit state is mapped to a $2n$-mode state as

$$|\boldsymbol{x}\rangle \iff (\hat{a}_1^\dagger)^{x_1}(\hat{b}_1^\dagger)^{1-x_1}\cdots(\hat{a}_n^\dagger)^{x_n}(\hat{b}_n^\dagger)^{1-x_n}|0\rangle. \tag{S130}$$

Furthermore, one can map an arbitrary single-qubit gate to an interaction between $\hat{a}$ and $\hat{b}$ by noting that any 2×2 matrix can be decomposed by Pauli operators and using the following mapping

$$\hat{\sigma}_i^0 \iff \mathbb{1}, \quad \hat{\sigma}_i^1 \iff \hat{a}_i^\dagger\hat{b}_i + \hat{a}_i\hat{b}_i^\dagger, \quad \hat{\sigma}_i^2 \iff i(\hat{a}_i\hat{b}_i^\dagger - \hat{a}_i^\dagger\hat{b}_i), \quad \hat{\sigma}_i^3 \iff \hat{a}_i^\dagger\hat{a}_i - \hat{b}_i^\dagger\hat{b}_i. \tag{S131}$$

Here, it is worth emphasizing that all the operations mapped from the Pauli operators preserve the number of photons and restrict the number of photons to a single photon for each site. Notice that the operators in bosonic system are not unitary operators in the full Hilbert space. However, since the subspace that we need to consider is confined by the states spanned by the single-photon states per each site, within this confined subspace, one can easily verify that the mapped operators in bosonic systems are guaranteed to be unitary. As a remark, for single-qubit gates, instead of directly mapping the unitary operation, one can map the Hamiltonian of the operation using the Euler decomposition, whose unitary operator will be decomposed into beam splitters and phase shifters.

For two-qubit gates as well, we can decompose arbitrary operations using Pauli operators as

$$\hat{U} = \sum_{\mu,\nu=0}^{3} u_{\mu,\nu}\hat{\sigma}_i^\mu \otimes \hat{\sigma}_j^\nu, \tag{S132}$$

and map this to bosonic operators as single-qubit cases. Again, all the operations mapped from the Pauli operators preserve the number of photons and restrict the number of photons to a single photon for each site. Therefore, we can map arbitrary input states, single-qubit gates, and two-qubit gates in an n-qubit system to a $2n$-mode bosonic system. Hence, it shows that we can construct a BQP-complete problem in bosonic system as well.

Now, we emphasize that the constructed bosonic circuit is different from Fock-state boson sampling for various aspects. First, although the circuit preserves the number of photons, we do not expect the circuit can be simulated by linear optics unless Fock-state boson sampling constitutes universal quantum computation. Therefore, the implementation of the circuit requires non-linear effects that are beyond quadratic hamiltonians, i.e., beyond Gaussian operations. Second, the constructed circuit guarantees that the number of photons in each site composed of $\hat{a}_i$ and $\hat{b}_i$ is always occupied by a single photon, which is clearly not the case for Fock-state boson sampling. It implies not only that the constructed circuit is distinct from Fock-state boson sampling but also that this is not the most general bosonic circuit that we can consider. Therefore, general molecular vibronic spectra problems that allow non-linear operations and multiphoton cases, which include the constructed circuit as a particular example, are BQP-hard.

[1] L. Gurvits, On the complexity of mixed discriminants and related problems, in *International Symposium on Mathematical Foundations of Computer Science* (Springer, 2005) pp. 447–458.

[2] S. Aaronson and A. Arkhipov, The computational complexity of linear optics, in *Proceedings of the forty-third annual ACM symposium on Theory of computing* (2011) pp. 333–342.

[3] S. Aaronson and T. Hance, Generalizing and derandomizing gurvits's approximation algorithm for the permanent, arXiv preprint arXiv:1212.0025 (2012).

[4] H. Hassanieh, P. Indyk, D. Katabi, and E. Price, Nearly optimal sparse fourier transform, in *Proceedings of the forty-fourth annual ACM symposium on Theory of computing* (2012) pp. 563–578.

[5] R. Vershynin, *High-dimensional probability: An introduction with applications in data science*, Vol. 47 (Cambridge university press, 2018).

[6] P. Drummond and C. Gardiner, Generalised p-representations in quantum optics, Journal of Physics A: Mathematical and General **13**, 2353 (1980).

[7] D. F. Walls and G. J. Milburn, *Quantum optics* (Springer Science & Business Media, 2007).

[8] J. Janszky and A. V. Vinogradov, Squeezing via one-dimensional distribution of coherent states, Physical review letters **64**, 2771 (1990).
[9] J. Vargas-Martínez, H. Moya-Cessa, and M. Fernández Guasti, Normal and anti-normal ordered expressions for annihilation and creation operators, Revista mexicana de física E **52**, 13 (2006).
[10] R. Bhatia, *Positive definite matrices* (Princeton university press, 2009).
[11] J. Huh and M.-H. Yung, Vibronic boson sampling: generalized gaussian boson sampling for molecular vibronic spectra at finite temperature, Scientific reports **7**, 1 (2017).
[12] A. Björklund, B. Gupt, and N. Quesada, A faster hafnian formula for complex matrices and its benchmarking on a supercomputer, Journal of Experimental Algorithmics (JEA) **24**, 1 (2019).
[13] R. Kan, From moments of sum to moments of product, Journal of Multivariate Analysis **99**, 542 (2008).
[14] N. Quesada, Franck-condon factors by counting perfect matchings of graphs with loops, The Journal of chemical physics **150**, 164113 (2019).
[15] A. Auerbach, *Interacting electrons and quantum magnetism* (Springer Science & Business Media, 2012).
[16] T.-C. Wei, M. Mosca, and A. Nayak, Interacting boson problems can be qma hard, Physical review letters **104**, 040501 (2010).
[17] E. Knill, R. Laflamme, and G. J. Milburn, A scheme for efficient quantum computation with linear optics, Nature **409**, 46 (2001).