

Page curves and typical entanglement in linear optics

Joseph T. Iosue^{1,2}, Adam Ehrenberg^{1,2}, Dominik Hangleiter^{2,1}, Abhinav Deshpande³, and Alexey V. Gorshkov^{1,2}

¹Joint Quantum Institute, NIST/University of Maryland, College Park, Maryland 20742, USA

²Joint Center for Quantum Information and Computer Science, NIST/University of Maryland, College Park, Maryland 20742, USA

³Institute for Quantum Information and Matter, California Institute of Technology, Pasadena, CA 91125, USA

May 17, 2023

Bosonic Gaussian states are a special class of quantum states in an infinite dimensional Hilbert space that are relevant to universal continuous-variable quantum computation as well as to near-term quantum sampling tasks such as Gaussian Boson Sampling. In this work, we study entanglement within a set of squeezed modes that have been evolved by a random linear optical unitary. We first derive formulas that are asymptotically exact in the number of modes for the Rényi-2 Page curve (the average Rényi-2 entropy of a subsystem of a pure bosonic Gaussian state) and the corresponding Page correction (the average information of the subsystem) in certain squeezing regimes. We then prove various results on the typicality of entanglement as measured by the Rényi-2 entropy by studying its variance. Using the aforementioned results for the Rényi-2 entropy, we upper and lower bound the von Neumann entropy Page curve and prove certain regimes of entanglement typicality as measured by the von Neumann entropy. Our main proofs make use of a symmetry property obeyed by the average and the variance of the entropy that dramatically simplifies the averaging over unitaries. In this light, we propose future research directions where this symmetry might also be exploited. We conclude by discussing potential applications of our results and their generalizations to Gaussian Boson Sampling and to illuminating the relationship between entanglement and computational complexity.

1 Introduction

In his pioneering paper [1], Page considered a Haar-random pure state in an N -dimensional Hilbert space. He conjectured an exact formula for the average entanglement entropy of the reduced density matrix on an rN -dimensional subspace, where $r \in [0, 1]$. His conjecture was proven soon after [2–4]. This average,

Joseph T. Iosue: jtiosue@umd.edu

which is itself a function of r and N , is now known as the *Page curve*. Page also defined the *average information of the subsystem* as the difference between the maximum of the entropy and the average entropy. This quantity has since come to be known as the *Page correction*. Since then, the Page curve and correction have found applications in a variety of areas, such as black holes [5–7], quantum information theory [8, 9], and statistical mechanics [10–17], among others. As a next step, many works considered the typical deviation of the entanglement entropy from its average. In particular, a system is said to exhibit *typical entanglement* if there is a vanishingly small probability that a random state has entanglement bounded away from the average. This phenomenon was introduced and studied in a variety of systems [18–27].

Entanglement is a key feature of quantum physics and can be used as a resource to complete various tasks, such as teleportation, key distribution, dense coding, and many others [28–32]. Furthermore, entanglement is a necessary ingredient for quantum advantage, since quantum computations with little entanglement can be simulated efficiently classically [33–35]. One can define quantum advantage using the language of complexity as using quantum resources to perform a task that is classically hard but quantumly easy. One such task is sampling from the output probability distribution of a Gaussian Boson Sampling experiment [36–40]. The general relationship between entanglement and complexity is largely unknown, but at least some entanglement is necessary for classical hardness. Similarly, in the setting of Boson Sampling, at least some amount of non-Gaussianity is necessary for sampling hardness [41]. On the other hand, too much entanglement can cause a state to be useless for computation. In particular, the typical (over the Haar measure) finite-dimensional quantum state is too entangled to be useful for computation [23, 24]. Thus, studying average and typical entanglement is necessary for learning about the useful part of entanglement and what utility random states have.

Entanglement in infinite-dimensional quantum states generated by bosonic Gaussian inputs and Gaussian operations has found direct application in areas such as quantum sensing [42–46] and quantum

communication [47, Ch. 12]. Furthermore, the original motivation for studying the finite-dimensional Page curve was to study the black hole information paradox [1, 5]. Since the degrees of freedom around black holes are inherently infinite-dimensional bosonic (e.g. photonic) modes, an infinite-dimensional bosonic Page curve may help to understand black hole information dynamics [47, Ch. 14].

Our contributions. In this work, we study Page curves and the typicality of entanglement in continuous-variable bosonic Gaussian states. Specifically, we compute average and typical entanglement quantities averaged over passive (energy-conserving) Gaussian unitaries, also called linear optical unitaries, with a fixed initial product state of squeezed vacuum states on n modes. Indeed, this setup is exactly that of a Gaussian Boson Sampling experiment, which is of great recent interest due to experimental claims of quantum advantage via Gaussian Boson Sampling [48–50].

We describe this setup in more detail in Section 2. Then in Section 3, we begin by studying the regime when all the initial squeezing strengths are equal, and denote this squeezing strength by s . We derive an analytic expression for the average Rényi-2 entropy of a subsystem of k modes as a function of s , n , and k that is exact asymptotically in n for arbitrary values of s and k . Using this expression, we exactly compute the corresponding Page correction. These results are summarized in Fig. 1. In Section 4, we then study the presence of typical entanglement for various scalings of k with n . When the distance between the entanglement of a random state and the average entanglement value vanishes additively (resp. multiplicatively), we say that entanglement is strongly (resp. weakly) typical. We prove that entanglement as measured by the Rényi-2 entropy is weakly typical for any k and strongly typical whenever $k \in o(n)$. We further show that entanglement as measured by the von Neumann entropy is weakly typical whenever $k \in o(n)$. Finally, in Section 5, we generalize our discussion to the regime when the initial squeezing strengths are not necessarily equal. We show that, if a certain conjecture is true, then the Rényi-2 and von Neumann entanglement entropies are both weakly typical whenever $k \in o(n)$.

Prior work. Refs. [26, 51] studied Page curves and entanglement in fermionic Gaussian states. Serafini et al. have considered typical entanglement in bosonic Gaussian states [52, 53], where they defined two measures, namely the microcanonical and canonical measures, on the set of all n -mode bosonic Gaussian states and averaged with respect to these measures. Roughly, averaging over the microcanonical measure corresponds to integrating over all bosonic Gaussian states up to a fixed bounded total energy, and averaging over the canonical measure corresponds to integrating over all bosonic Gaussian states with

a Boltzmann weight factor decaying with the energy of the state. Fukuda and Koenig generalized these results by studying entanglement averaged over passive Gaussian unitaries with a fixed initial product state of squeezed vacuum states on n modes [54]. Indeed, their setup is exactly the one we consider in this work. As noted in Ref. [54], the measure defined by fixing squeezing strengths and then applying a random passive Gaussian unitary generalizes both the microcanonical and canonical measures, as the latter two measures can be expressed as convolutions of the former with certain distributions on the set of all squeezing configurations.

Serafini et al. and Fukuda and Koenig consider entanglement in a subsystem of k bosonic modes when $k \in o(n)$; roughly, Serafini et al. allow $k \in \mathcal{O}(1)$ and Fukuda and Koenig allow $k \in o(n^{1/3})$. To the best of our knowledge, there are currently no results on average entanglement or typical entanglement in the regime when $k \in \Omega(n^{1/3})$. This is what we address in our work. Concerning typical entanglement, we emphasize that our results do not supplant the results of Ref. [54] in general, but rather only in certain regimes. In particular, we primarily consider the situation of equal squeezing strengths, whereas their results pertain to the situation of arbitrary squeezing. Similarly, we primarily consider the Rényi-2 entropy, whereas their results apply to both the Rényi-2 and von Neumann entropies. We summarize our work and that of Ref. [54] on typical entanglement in bosonic Gaussian states in Table 1.

For general quantum states, the von Neumann entropy has certain properties, such as *strong subadditivity*, that the Rényi- α entropies do not. For other such properties, we refer to Ref. [32]. Because of this, the von Neumann entropy is generally considered a better measure of entanglement than the Rényi- α entropies. Notably, however, it has been shown that the Rényi-2 entropy is special when restricting to bosonic Gaussian states. For example, it was recently proven that for bosonic Gaussian states, the Rényi-2 entropy also obeys strong subadditivity [55, 56]. The Rényi-2 entropy is also equal, up to a constant, to the phase-space Shannon sampling entropy $-\int W(\mathbf{x}, \mathbf{p}) \log W(\mathbf{x}, \mathbf{p}) d^n \mathbf{x} d^n \mathbf{p}$ of the Wigner distribution $W(x, p)$ of the state for Gaussian states [55]. Here the position vector \mathbf{x} and momentum vector \mathbf{p} parameterize the phase space of n oscillator modes. The phase-space Shannon sampling entropy has an operational meaning in terms of sampling via homodyne detections [55, 57]. Furthermore, it has been shown that for pure tripartite Gaussian states, the Rényi-2 entropy obeys a strong subadditivity inequality that is stronger than that for the von Neumann entropy [58]. Finally, correlation measures for Gaussian states based on the Rényi-2 entropy have been found that have no counterpart when using the von Neumann entropy [59]. As noted in Ref. [55], the

aforementioned results are “planting the seeds for a full Gaussian quantum information theory based on the Rényi-2 entropy.” Our work is focused primarily on entanglement in bosonic Gaussian states as measured by the Rényi-2 entropy and can thus be viewed as planting a few more seeds.

2 Setup

In this work, we consider a linear optical system of n modes, where each mode is an independent quantum harmonic oscillator. We will restrict our attention to bosonic Gaussian states. We refer to Refs. [60, 54] for background on the theory of Gaussian states, and we provide the necessary details to understand this paper in Appendix A.1.

Consider an n -mode mixed state ρ in the Hilbert space of square integrable wavefunctions $\mathcal{H} = L^2(\mathbb{R})^{\otimes n}$. For each $i \in \{1, \dots, n\}$, let \hat{x}_i and \hat{p}_i be the position and momentum quadrature operators on the i^{th} mode. For each i , define $\hat{r}_i := \hat{x}_i$ and $\hat{r}_{n+i} := \hat{p}_i$. ρ is a Gaussian state if there is a $\beta > 0$ and a Hamiltonian \hat{H} that is at most quadratic in the quadrature operators such that ρ is the thermal state $\rho \propto e^{-\beta \hat{H}}$. Since \hat{H} contains only linear and quadratic quadrature terms in \hat{r}_i , ρ is fully characterized by its first and second moments, $\text{Tr}(\rho \hat{r}_i)$ and $\sigma_{ij} = \frac{1}{2} \text{Tr}[\rho(\hat{r}_i \hat{r}_j + \hat{r}_j \hat{r}_i)] - \text{Tr}(\rho \hat{r}_i) \text{Tr}(\rho \hat{r}_j)$. σ is called the *covariance matrix* of the state ρ .

A unitary U is Gaussian if there exists a Hamiltonian \hat{H} that is at most quadratic in the quadrature operators such that $U = e^{-i\hat{H}}$. One can show that a Gaussian unitary maps Gaussian states to Gaussian states. Any Gaussian state can be generated by acting on the vacuum state with a Gaussian unitary. The set of all Gaussian unitaries is isomorphic to the real symplectic group of $2n \times 2n$ matrices $\text{Sp}(2n)$. By the Euler decomposition theorem of a symplectic matrix, it follows that any pure Gaussian state can be generated by acting on an initial product state of squeezed vacuum modes with a passive (energy conserving) Gaussian unitary. The set of all passive Gaussian unitaries is isomorphic to $\text{Sp}(2n) \cap \text{O}(2n) \cong \text{U}(n)$, where $\text{O}(2n)$ is the orthogonal group of $2n \times 2n$ matrices and $\text{U}(n)$ is the unitary group of $n \times n$ matrices. $\text{Sp}(2n)$ is not compact and thus does not have a finite Haar measure to average over. Physically, this is due to the fact that the Gaussian operation of squeezing can take on unbounded values in \mathbb{R} . However, $\text{U}(n)$ is compact, and it is hence well-defined to consider uniformly sampling from $\text{U}(n)$ according to the finite Haar measure.

We are therefore motivated to consider the following notion of a random pure Gaussian state. Namely, we initialize the i^{th} mode to be in a squeezed vacuum state with fixed squeezing parameter $s_i \in \mathbb{R}$ for each $i \in \{1, \dots, n\}$. We then randomly sample a passive Gaussian unitary from $\text{U}(n)$ and apply it to the n modes. This random state is thus character-

ized by a fixed choice of the squeezing parameters s_i and a Haar-random choice of a passive Gaussian unitary U . Understanding the properties of such random states is of great interest, particularly because Gaussian Boson Sampling experiments rely on precisely those state preparations. For squeezing parameters s_i for $i \in \{1, \dots, n\}$, the total expected number of bosons of the state on the n modes is $\sum_{i=1}^n \sinh^2(s_i)$. For simplicity, we will begin by considering the case when all the squeezing parameters are equal; $s_i = s$ for each i for some fixed s . In the case of equal squeezings, the average total boson number per mode is $\sinh^2 s$. The general case of unequal squeezings will be discussed in Section 5.

The n modes are then partitioned into two groups — one group of $k = rn$ modes for some $0 \leq r \leq 1$, and one group of $n - k = (1-r)n$ modes. We then compute the entropy of the reduced state of the k modes, or equivalently, since we are considering pure states, the entropy of the reduced state of the $n - k$ modes [28]. Let the density matrix of the reduced state be ρ . For the entropy function, we will be primarily focused on the Rényi-2 entropy $S_2 = -\log \text{Tr} \rho^2$, although we will also prove various statements on the von Neumann entropy $S_1 = -\text{Tr} \rho \log \rho$ as well. The Rényi-2 entropy takes the elegant form $S_2 = \frac{1}{2} \log \det \sigma = \frac{1}{2} \text{Tr} \log \sigma$, where σ is the covariance matrix of ρ [60]. For the Gaussian state generated by the unitary $U \in \text{U}(n)$ acting on the squeezed product state with squeezing strength s on each mode, the Rényi-2 entropy of the subsystem of k modes is denoted by $S_2(U)$, and the von Neumann entropy by $S_1(U)$, where the dependence of $S_1(U)$ and $S_2(U)$ on r , n , and s is implicit. We will derive statistical properties of $S_2(U)$ and $S_1(U)$ in the asymptotic limit $n \rightarrow \infty$.

Our main results involve the Rényi-2 entropy, and the following proposition allows us to use many of the results on the Rényi-2 entropy to bound the von Neumann entropy. Furthermore, we will also make use of the maximum Rényi-2 entropy to prove our later results on the Page correction.

Proposition 1. *Let $s \in \mathbb{R}$ and $r \in [0, 1]$. Then for each $j \in \{1, 2\}$,*

$$\max_{U \in \text{U}(n)} S_j(U) = n \min(r, 1 - r) h_j(\cosh(2s)), \quad (1)$$

where $h_1(x) = \frac{x+1}{2} \log \frac{x+1}{2} - \frac{x-1}{2} \log \frac{x-1}{2}$ and $h_2(x) = \log x$. Furthermore, for any fixed $U \in \text{U}(n)$, $S_1(U) \geq S_2(U)$ and

$$S_1(U) < S_2(U) + n \min(r, 1 - r)(1 - \log 2). \quad (2)$$

The full proof of Proposition 1 is given in Appendix B. A tighter version of Eq. (2) was originally derived in Ref. [58, Eq. 15], but we will only need this weaker version. For completeness, we provide a different proof of Eq. (2). Eq. (1) is perhaps implicit in various results in Refs. [52, 53, 61, 60, 62], but we have not

found it directly stated anywhere. The lower bound $S_1(U) \geq S_2(U)$ is a general property of the Rényi entropies [28, 32], whereas the upper bound holds only for Gaussian states. While trivial upper bounds also exist in the general case, Eq. (2) is tighter. Intuitively one may view this as an extension of the fact that for Gaussian states, the Rényi-2 and von Neumann entropies share many useful properties, such as strong subadditivity and others mentioned in Section 1.

3 Expectation value

Our first results concern the Rényi-2 Page curve, which is the expectation value of the Rényi-2 entan-

glement entropy as a function of the partition size ratio $r = k/n$ and the squeezing strength s . Recall that the dependence of $S_2(U)$ on r , n , and s is implicit. We find an exact formula as an infinite series for the Page curve in the limit $n \rightarrow \infty$.

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{E}_{U \in U(n)} \frac{1}{n} S_2(U) &= r \log \cosh(2s) - \sum_{\ell=1}^{\infty} r^{\ell+1} \tanh^{2\ell}(2s) \frac{C_{\ell}}{2\ell} {}_2F_1(1-\ell, \ell; \ell+2; r) \quad (3) \\ &= r \log \cosh(2s) + \sum_{d=2}^{\infty} r^d \sum_{\ell=\lceil d/2 \rceil}^{d-1} (-1)^{d-\ell} \tanh^{2\ell}(2s) \frac{(2\ell-1)!}{d(d-1)(d-\ell-1)!(2\ell-d)!\ell!}. \quad (4) \end{aligned}$$

This function is symmetric under $r \mapsto 1-r$, and hence the formula holds when r is replaced with $\min(r, 1-r)$. Furthermore, asymptotically in n , $\mathbb{E}_{U \in U(n)} S_2(U) = n\alpha(s, r) - \lambda(s, r) + o(1)$, where $\alpha(s, r)$ is precisely $\lim_{n \rightarrow \infty} \mathbb{E}_{U \in U(n)} \frac{1}{n} S_2(U)$ given above, $\lambda(s, r)$ is independent of n , and $o(1)$ denotes terms that go to zero as $n \rightarrow \infty$.

We plot the analytic Page curve for $s = 3/4$ in Fig. 1(a), and we confirm our results numerically in Fig. 1(b). The proof of Theorem 2, given in Appendix C.1, primarily uses two ingredients. The first ingredient is the asymptotic form of the Weingarten calculus for integrating over the unitary group with respect to the Haar measure [69, 70]. From this we get an equation for the Page curve that is initially daunting. The second main ingredient is the fact that the Page curve must be symmetric under $r \mapsto 1-r$ since the global state on the n modes is pure [28]. Quite miraculously, this fact is enough to simplify the equation for the Page curve and arrive at Theorem 2.

The Page curve derived in Theorem 2 can be written as

$$\lim_{n \rightarrow \infty} \mathbb{E}_{U \in U(n)} \frac{1}{n} S_2(U) = \sum_{\ell=1}^{\infty} \frac{\tanh^{2\ell}(2s)}{2\ell} G_{\ell}(r), \quad (5)$$

where $G_{\ell}(r) := r - f_{\ell}(r)$ and $f_{\ell}(r)$ is a polynomial of degrees $\ell+1$ through 2ℓ in r (given in Eq. (C20) in Appendix C.1). Polynomials $G_{\ell}(r)$ of this form are uniquely determined by the requirement that $G_{\ell}(r) = G_{\ell}(1-r)$, which ensures that the Rényi-2 entropy of

Theorem 2 (Rényi-2 Page curve). *Fix $s \in \mathbb{R}$ and $r \in [0, 1]$. Let $C_{\ell} := \frac{1}{\ell+1} \binom{2\ell}{\ell}$ be the ℓ^{th} Catalan number, and let ${}_2F_1$ be the hypergeometric function [63–67]. Then*

a subsystem is equal to that of its complement since we are considering pure states. It is from this requirement that we ultimately derive the Page curve. We show that the resulting $G_{\ell}(r)$ can be understood as a good approximation to $m(r) := \min(r, 1-r)$ from below, which we will call the ℓ^{th} approximation. Indeed, the approximation is especially good near the endpoints $r = 0$ and $r = 1$, where the first ℓ derivatives of $G_{\ell}(r)$ match those of $m(r)$. As $\ell \rightarrow \infty$, the approximation becomes better and better such that $\lim_{\ell \rightarrow \infty} G_{\ell}(r) = m(r)$.

This provides an interpretation of the derived form of the Page curve. The strength of the squeezing s determines the weight that the Page curve has on the ℓ^{th} approximation to $m(r)$. For small squeezing, only low order approximations contribute, with the most dominant contribution being the parabolic shape $G_1(r) = r(1-r)$. When the squeezing is increased, there is more contribution from higher order approximations, giving the Page curve more of the triangle shape of $m(r)$. We see a manifestation of this interpretation as

$$\lim_{s \rightarrow 0} \lim_{n \rightarrow \infty} \mathbb{E}_{U \in U(n)} \frac{1}{s^2 n} S_2(U) = 2r(1-r), \quad (6)$$

$$\lim_{s \rightarrow \infty} \lim_{n \rightarrow \infty} \mathbb{E}_{U \in U(n)} \frac{1}{s n} S_2(U) = 2 \min(r, 1-r). \quad (7)$$

Meanwhile, from Eq. (1), the maximal Rényi-2 entropy is $\max_U \frac{1}{n} S_2(U) = m(r) \log \cosh(2s)$. As stated, near the endpoints $r = 0$ and $r = 1$, $G_{\ell}(r)$ is a very good approximation to $m(r)$. Thus, regardless of the squeezing strength, when the subsystem size

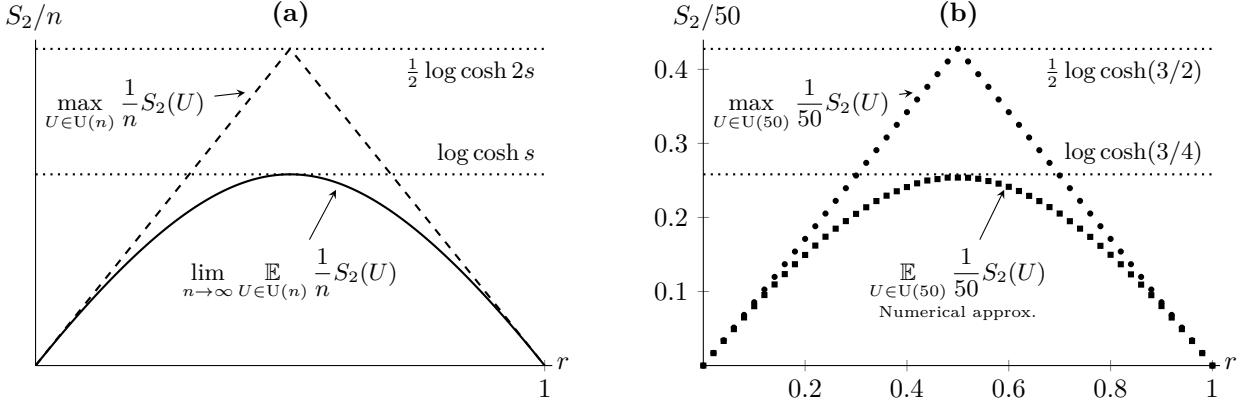


Figure 1: (a) Exact results for the Rényi-2 Page curve from Proposition 1, Theorem 2, and Corollary 3. (b) Numerical simulations of the Rényi-2 Page curve for $n = 50$ modes and squeezing $s = 3/4$. We plot the values for $r = k/50$ for each $k \in \{0, 1, \dots, 50\}$. We perform the simulation by generating random unitary matrices and doing the matrix multiplication described in Appendix A.1. We provide the code for this simulation on GitHub [68].

$k = rn$ is small (or when its complement is small), the average entanglement is very close to maximal.

Unfortunately, we are unable to simplify the infinite sum for general r in Theorem 2 further. However, the Page curve can be fully simplified at $r = 1/2$ — which is where the maximum for a fixed s occurs — to $\log \cosh s$. Indeed, we find that

$$G_\ell(1/2) = \frac{1}{2} \left(1 - 4^{-\ell} \binom{2\ell}{\ell} \right). \quad (8)$$

From this and the maximum Rényi-2 entropy from Eq. (1), we also find the exact expression for the Page correction at $r = 1/2$ to be $\frac{1}{2} \log(1 + \tanh^2 s)$. Let us formally state these results as follows.

Corollary 3. *For a fixed $s \in \mathbb{R}$ when $r = 1/2$,*

$$\lim_{n \rightarrow \infty} \mathbb{E}_{U \in \mathbb{U}(n)} \frac{1}{n} S_2(U) = \log \cosh s, \quad (9)$$

and the Page correction is

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \left(\max_{U \in \mathbb{U}(n)} S_2(U) - \mathbb{E}_{U \in \mathbb{U}(n)} S_2(U) \right) \\ = \frac{1}{2} \log(1 + \tanh^2 s). \end{aligned} \quad (10)$$

The proof of Corollary 3, given in Appendix C.2, is a straightforward consequence of a simplification of the hypergeometric function ${}_2F_1(a, 1-a; c; 1/2)$ in terms of gamma functions due to Bailey's theorem [63–67]. Altogether, we have derived the exact formula [Eqs. (3) and (4)] for the Rényi-2 page curve in the regime of equal squeezers as a series in $\tanh^2(2s)$ and r . In the special case when $r = 1/2$, we simplified the series to obtain an exact value of $\log \cosh s$. From this, we derived the Page correction, or information of the subsystem, at $r = 1/2$ to be exactly $\frac{1}{2} \log(1 + \tanh^2 s)$. Furthermore, since the Page curve

is concave in r while the maximum entropy is linear, this correction is maximized at $r = 1/2$. A summary of the Page curve results thus far is provided in Fig. 1.

We now shift our attention to the constant term in the Page curve. Theorem 2 states that asymptotically in n , $\mathbb{E}_{U \in \mathbb{U}(n)} S_2(U) = n\alpha(s, r) - \lambda(s, r) + o(1)$, and it provides the exact expression for $\alpha(s, r)$. We further find the following result for λ , which is proven in Appendix C.3.

Proposition 4. *Fix $s \in \mathbb{R}$ and $r \in [0, 1]$. Then,*

$$\lambda(s, r) = -\frac{1}{8} \log(1 - 4r(1-r) \tanh^2(2s)). \quad (11)$$

Note that $\lambda(s, r) \geq 0$ for all s and r , and therefore $\mathbb{E}_U S_2(U) \leq n\alpha(s, r)$ asymptotically. At $r = 1/2$, this simplifies to $\lambda(s, 1/2) = \frac{1}{4} \log \cosh(2s)$.

The proof of Proposition 4 is very similar to the proof of Theorem 2, again crucially using the symmetry $r \mapsto 1 - r$ to simplify the expression coming from the Weingarten function. However, there is one substantial difficulty in the proof of Proposition 4 that does not occur in that of Theorem 2. In this case, the symmetry $r \mapsto 1 - r$ is *almost* enough to fully determine λ ; however, there are constants that cannot be determined by the symmetry alone. We derive expressions for these constants that are complicated sums involving permutations and Catalan numbers. Using objects that arise in bioinformatics when studying gene orders, called *breakpoint graphs* [71], Ref. [72] evaluates these sums, hence completing the proof of Proposition 4. These constants are discussed more in Appendix C.3.

In summary, we have derived an explicit form of the asymptotic Rényi-2 entropy Page curve for all r and s up to corrections that vanish as $n \rightarrow \infty$.

4 Variance and typicality

Next, we shift our attention to the variance of the Rényi-2 entropy so as to make statements about typicality of entanglement; that is, how different a random state's entanglement is from the average. Using the results for the Rényi-2 entropy, we will also be able to prove some weaker results for the von Neumann entropy. The typicality results presented below are summarized in Table 1.

Typicality is of interest because it characterizes the applicability of statistical averages. Indeed, statistical mechanics often relies on quantities being typical so that thermodynamic average quantities, such as average energy and average pressure, can accurately represent their true values. In order to quantify the deviation from average, we consider two measures of deviation corresponding to multiplicative and additive distance. If the multiplicative distance between a quantity and its average vanishes in the thermodynamic limit, then that quantity is called *weakly typical*. If the additive distance vanishes in this limit, then that quantity is called *strongly typical*. With this intuition, we now formally define weak and strong typicality following Ref. [27].

Definition 5 (Typicality). *Let S be a nonnegative random variable on the unitary group $U(n)$, and denote its value at $U \in U(n)$ by $S(U)$. S is called weakly typical if for any constant $\epsilon > 0$,*

$$\lim_{n \rightarrow \infty} \Pr_{U \in U(n)} \left[\left| \frac{S(U)}{\mathbb{E}_{V \in U(n)} S(V)} - 1 \right| < \epsilon \right] = 1. \quad (12)$$

S is called strongly typical if for any constant $\epsilon > 0$,

$$\lim_{n \rightarrow \infty} \Pr_{U \in U(n)} \left[\left| S(U) - \mathbb{E}_{V \in U(n)} S(V) \right| < \epsilon \right] = 1. \quad (13)$$

If $\mathbb{E}_{U \in U(n)} S(U)$ does not decay as $n \rightarrow \infty$, then strong typicality clearly implies weak typicality. We will be concerned with typicality of S_1 and S_2 , the von Neumann and Rényi-2 entropy respectively. One can compute the variance of the entropy over the Haar measure and then apply Chebyshev's inequality to obtain typicality results. Therefore, we now focus on the variance. We first find the general form of the asymptotic Rényi-2 variance in the equal squeezing regime and show that it is independent of n . Recall that the dependence of $S_2(U)$ on r , n , and s is implicit.

Theorem 6 (Rényi-2 variance). *Let $s \in \mathbb{R}$ and $r \in [0, 1]$. Then*

$$\lim_{n \rightarrow \infty} \text{Var}_{U \in U(n)} S_2(U) = \sum_{d=2}^{\infty} \omega^{(d)} \tanh^{2d}(2s) (r(1-r))^d, \quad (14)$$

where $\omega^{(d)} \in \mathbb{Q}$ is some number that depends only on d . In particular, $\omega^{(2)} = 1/2$.

The proof of this theorem, given in Appendix D, is very similar to the proof of Theorem 2. To prove this theorem, we again crucially use that $\text{Var}_U S_2(U)$ must be symmetric under $r \mapsto 1-r$ since the full state on the n modes is pure. Interestingly, in contrast to Theorem 2 where we found that the Page curve grows linearly with n , the asymptotic variance is independent of n . Indeed, Theorem 6 is the bosonic analogue of the result that the variance for fermionic Gaussian states is asymptotically constant [26].

From Theorem 6, we find typicality in certain regimes as an immediate corollary. In particular, since the variance is independent of n while the average grows with n , the entanglement is always weakly typical. Furthermore, the variance is asymptotically zero if $s \in o(1)$ and/or $r \in o(1)$, and typicality is therefore strong in those regimes. Altogether, Theorem 6 and Chebyshev's inequality lead to the following corollary.

Corollary 7. *The Rényi-2 entropy is weakly typical for any $s \in \mathbb{R}$ and $r \in [0, 1]$. Furthermore, if the subsystem size $k = rn$ scales as $k \in o(n)$, or if the squeezing scales as $s \in o(1)$, then the Rényi-2 entropy is strongly typical, and $\Pr_{U \in U(n)} [|S_2(U) - \mathbb{E}_V S_2(V)| < \epsilon] \geq 1 - \frac{\Omega(1/n)}{\epsilon^2}$. On the other hand, if r and s are constant in n , then $\lim_{n \rightarrow \infty} \Pr_{U \in U(n)} [|S_2(U) - \mathbb{E}_V S_2(V)| < \epsilon] \geq 1 - \frac{c(r,s)}{\epsilon^2}$, where $c(r,s)$ is a constant independent of n (but depends on r and s).*

In summary, we have shown weak typicality in the Rényi-2 entropy for all s and k and strong typicality whenever $k \in o(n)$. On the other hand, if s and r do not tend to zero with increasing n , the variance converges to a constant value independent of n . Hence, we cannot use Chebyshev's inequality to prove strong typicality in this case. Notably, an asymptotically constant variance does not necessarily imply an absence of strong typicality either — that is, the probability that the entropy deviates from its average can scale as $1/n^2$, but the entropy can deviate by an amount proportional to n thus resulting in a constant variance. We are therefore unable to make a definitive statement about strong typicality in the $r \in \Theta(1)$ and $s \in \Omega(1)$ case.

In the next corollary, we will use these results to address typicality as measured by the von Neumann entropy. Specifically, we will use Proposition 1 to show weak typicality of the von Neumann entropy as long as the subsystem size scales sublinearly with the system size.

Corollary 8. *Let $s \in \mathbb{R}$ and $r \in [0, 1]$. If the subsystem size $k = rn$ scales as $k \in o(n)$, then the von Neumann entropy is weakly typical.*

Proof. Using Proposition 1 to upper bound $\mathbb{E}_U \frac{1}{n^2} S_1(U)^2$ and to lower bound $(\mathbb{E}_U \frac{1}{n} S_1(U))^2$, we

		$k \in \Theta(n)$	$k \in o(n)$	$k \in o(n^{1/3})$ [54]
Equal squeezing	<i>Rényi-2 von Neumann</i>	weak ?	strong weak	strong strong
Unequal squeezing	<i>Rényi-2 von Neumann</i>	?	weak* weak*	strong strong

Table 1: A summary of the current status of rigorous results on typical entanglement in Gaussian bosonic systems, where in this figure we assume single-mode squeezing parameters that are independent of n . Strong and weak typicality are defined in Definition 5. Note that “weak*” indicates that the result is not fully proven, but depends on Conjecture 10. Where we say “weak”, we have not ruled out the possibility that the typicality is also strong. The total number of modes is denoted by n , and $0 \leq k \leq n$ is the number of modes in the subsystem. “Equal squeezing” refers to the case when each mode is initially squeezed with the same strength, whereas “unequal squeezing” refers to the general case when each mode can be squeezed independently. The two leftmost columns come from Corollaries 7 and 8 and Remark 11. The rightmost column all follows from the results of Ref. [54]. Prior to Ref. [54], Refs. [52, 53] proved strong typicality in the regime $k \in \mathcal{O}(1)$.

find that

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n^2} \operatorname{Var}_{U \in \mathcal{U}(n)} S_1(U) &< \lim_{n \rightarrow \infty} \frac{1}{n^2} \operatorname{Var}_{U \in \mathcal{U}(n)} S_2(U) \\ &+ 2r \log(e/2) \lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E}_{U \in \mathcal{U}(n)} S_2(U) \\ &+ r^2 \log^2(e/2). \end{aligned} \quad (15)$$

From Theorem 2, the first term in the right-hand side is always zero. Furthermore, from Theorem 6, the second term is $\mathcal{O}(r^2)$. Hence, $\operatorname{Var}_U S_1(U)/n^2 \in \mathcal{O}(r^2)$, which is zero as $n \rightarrow \infty$ when $r \in o(1)$. \square

We again emphasize that our typicality results thus far, which are summarized in Table 1, only apply to the case when the initial squeezing strength on each mode is the same. On the other hand, Ref. [54] proves strong typicality when $k \in o(n^{1/3})$ for both the von Neumann and Rényi-2 entropies in the general case when squeezing strengths can be different on different modes. It is for this reason that our results do not supplant those of Ref. [54] in general, but rather only in certain regimes. To the best of our knowledge, our results are the first to address the $k \in \Omega(n^{1/3})$ regime.

Ultimately, we would like to determine in exactly which regimes strong and weak typicality occur and do not occur. Our results thus far almost complete the story for the regime of equal squeezing when typicality is measured with the Rényi-2 entropy, since we have proven strong typicality when $k \in o(n)$ and weak when $k \in \Theta(n)$; the missing piece is whether or not typicality is strong when $k \in \Theta(n)$. However, the story is even more incomplete for the regime of equal squeezing when typicality is measured with the von Neumann entropy, though we made some progress by proving that typicality is at least weak whenever $k \in o(n)$. In this regime, the best known result for strong typicality is when $k \in o(n^{1/3})$ as proven in Ref. [54]. Indeed, this is also currently the best known result in the regime of unequal squeezing. In Section 5, we will use our results on equal squeezing typicality to add to the story for unequal squeezing.

5 Generalizing to unequal squeezing

So far, we have considered the restricted setting where each mode i is initially squeezed with strength $s_i = s$ for some $s \in \mathbb{R}$. We now generalize by allowing the squeezing strengths to be different on each mode. As such, for the remainder of this section, the squeezings will be (s_1, \dots, s_n) , where each $s_i \in \mathbb{R}$, and s_{\max} and s_{\min} are defined as $s_{\max} := \max_i |s_i|$ and $s_{\min} := \min_i |s_i|$.

To begin, we focus on the Rényi-2 Page curve in the regime of unequal squeezing. When squeezing is small, we can utilize the equal squeezing Page curve in Theorem 2 to compute the Page curve for unequal squeezing. Recall that the dependence of $S_2(U)$ on r , n , and s is implicit.

Corollary 9. *Let each $s_i \in \mathbb{R}$ and $r \in [0, 1]$. Then, as $n \rightarrow \infty$,*

$$\mathbb{E}_{U \in \mathcal{U}(n)} S_2(U) = 2r(1-r) \sum_{i=1}^n s_i^2 + \mathcal{O}(rns_{\max}^4). \quad (16)$$

Proof. The Rényi-2 entropy of a Gaussian state with covariance matrix σ is proportional to $\log \det \sigma$. \log is real analytic and $\det \sigma$ is analytic in the parameters of the initial covariance matrix σ_0 . Hence, S_2 can be written as a power series in s_i . There exists a passive Gaussian unitary, specifically a product of one-mode phase shifters, that acts on σ_0 via the transformation $s_i \rightarrow -s_i$. Hence, by the translational invariance of the Haar measure — that is, the invariance of the Haar measure under the application of any fixed unitary — the power series must be an expansion in s_i^2 . Furthermore, there exists a passive Gaussian unitary, specifically a product of beamsplitters, that acts on σ_0 via the transformation $s_i \rightarrow s_{\tau(i)}$ for some permutation τ of n elements. Therefore, again by the translational invariance of the Haar measure, the power series must be symmetric under $s_i \rightarrow s_{\tau(i)}$. It follows that the $\mathcal{O}(s_i^2)$ term must be of the form $g(r, n) \sum_i s_i^2$ for some function $g(r, n)$. When all s_i are equal, the

power series must reduce to Theorem 2, which fixes $g(r, n)$ to be $2r(1 - r)$.

The next term is $\mathcal{O}(s_{\max}^4)$, but what is the k dependence? We will show that the k dependence is at most linear, proving that the remaining terms in the power series are $\mathcal{O}(ks_{\max}^4)$. Recall that $S_2(U) = \frac{1}{2} \text{Tr} \log \sigma(U)$ where $\sigma(U)$ is the covariance matrix for the state generated by U from the initial product squeezed state σ_0 . Since the log function is concave, Jensen's inequality implies that $\mathbb{E}_U S_2(U) \leq \frac{1}{2} \text{Tr} \log \mathbb{E}_U \sigma(U)$. $\mathbb{E}_U \sigma(U)$ is calculated in Eq. (A5) in Appendix A.1 to be $\nu \mathbb{I}_{2k \times 2k}$, where $\nu = \frac{1}{n} \sum_i \cosh(2s_i) \leq \cosh(2s_{\max})$. Therefore, $\mathbb{E}_U S_2(U) \in \mathcal{O}(k \log \nu)$. \square

One particularly interesting application of this corollary is when each $s_i \in \mathcal{O}(1/\sqrt{n})$. In this case, the average total number of bosons in the n modes is $N = \sum_i \sinh^2(s_i) \in \mathcal{O}(1)$. Thus, when one considers a constant number N of bosons in the system as the number of modes is taken to infinity, one finds the Page curve to be $2r(1 - r)N$.

We now shift our focus to entanglement typicality in the regime of unequal squeezing. The results described in the remainder of this section are summarized in Table 1. Ideally, we would like to make further statements about entanglement in the regime of unequal squeezing by using our previous results on equal squeezing. One way to potentially proceed is to use the equal squeezing results to bound the unequal squeezing quantities. Intuitively, we expect that $\mathbb{E}_U S_2(U)$ for arbitrary unequal squeezing is upper bounded by $\mathbb{E}_U S_2(U)|_{\text{all } s_i=s_{\max}}$ and lower bounded by $\mathbb{E}_U S_2(U)|_{\text{all } s_i=s_{\min}}$. In other words, by *increasing* all of the squeezing strengths until they all are equal, the average entanglement will *increase*. In this spirit, we make the following conjecture.

Conjecture 10. *Let $r \in [0, 1]$, $s_i \in \mathbb{R}$, and $n \in \mathbb{N}$. Then, for any i ,*

$$0 \leq \frac{\partial}{\partial(s_i^2)} \left(\mathbb{E}_{U \in \mathbb{U}(n)} \frac{1}{n} S_2(U) \right), \text{ and} \quad (17)$$

$$0 \leq \frac{\partial}{\partial(s_i^2)} \left(\mathbb{E}_{U \in \mathbb{U}(n)} \frac{1}{n^2} S_2(U)^2 \right). \quad (18)$$

Conjecture 10 seems intuitive — by increasing the magnitude of any individual squeezing strength, the number of bosons in the system increases, and therefore it would seem surprising for the average entanglement to decrease. We note that, somewhat counterintuitively, one can find explicit unitaries and squeezing configurations for which $\frac{\partial}{\partial(s_i^2)} S_2(U) < 0$ (we provide an example in our code repository [68]), and hence the presence of the \mathbb{E}_U is necessary for the conjecture. Despite its intuitiveness, we have been unable to rigorously prove Conjecture 10. The derivative of

$S_2(U) = \frac{1}{2} \log \det \sigma(U)$ is $\frac{1}{2} \text{Tr} \left(\sigma(U)^{-1} \frac{\partial \sigma(U)}{\partial(s_i^2)} \right)$. The difficulty in computing the expectation value over $U \in \mathbb{U}(n)$ arises due to the presence of the inverse $\sigma(U)^{-1}$.

Nonetheless, under the assumption that Conjecture 10 is true, we can immediately upper and lower bound the Rényi-2 Page curve for an arbitrary squeezing configuration (s_1, \dots, s_n) by using Theorem 2 with $s = s_{\max}$ and $s = s_{\min}$ respectively,

$$\lim_{n \rightarrow \infty} \mathbb{E}_{U \in \mathbb{U}(n)} \frac{1}{n} S_2(U) \geq \lim_{n \rightarrow \infty} \mathbb{E}_{U \in \mathbb{U}(n)} \frac{1}{n} S_2(U)|_{s_i=s_{\min}}, \quad (19)$$

$$\lim_{n \rightarrow \infty} \mathbb{E}_{U \in \mathbb{U}(n)} \frac{1}{n} S_2(U) \leq \lim_{n \rightarrow \infty} \mathbb{E}_{U \in \mathbb{U}(n)} \frac{1}{n} S_2(U)|_{s_i=s_{\max}}. \quad (20)$$

Furthermore, the conjecture implies that

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{E}_{U \in \mathbb{U}(n)} \frac{1}{n^2} S_2(U)^2 \\ \leq \lim_{n \rightarrow \infty} \mathbb{E}_{U \in \mathbb{U}(n)} \frac{1}{n^2} S_2(U)^2|_{s_i=s_{\max}}. \end{aligned} \quad (21)$$

From this, we can also make statements on weak typicality for unequal squeezers by bounding the variance. The variance is $\mathbb{E}_U S_2(U)^2 - (\mathbb{E}_U S_2(U))^2$. We can therefore upper bound the variance by upper bounding the first term with Eq. (21) and lower bounding the second term with Eq. (19).

Remark 11. *Let $r \in [0, 1]$ and $s_i \in \mathbb{R}$. If Conjecture 10 is true and if the subsystem size $k = rn \in o(n)$ or $s_{\max} \in o(1)$, then the Rényi-2 entropy is weakly typical. Similarly, if $k \in o(n)$, then the von Neumann entropy is weakly typical.*

Proof. We can bound the variance of the Rényi-2 entropy as

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n^2} \text{Var}_U S_2(U) &\leq \lim_{n \rightarrow \infty} \frac{1}{n^2} \\ &\times \left[\mathbb{E}_U \left(\frac{1}{n} S_2(U)|_{\text{all } s_i=s_{\max}} \right)^2 \right. \\ &\left. - \left(\mathbb{E}_U \frac{1}{n} S_2(U)|_{\text{all } s_i=s_{\min}} \right)^2 \right]. \end{aligned} \quad (22)$$

From Theorem 6, the \mathbb{E}_U can be brought inside the parentheses in the first term. Then, the right hand side can be computed using Theorem 2, which gives $\lim_{n \rightarrow \infty} \frac{1}{n^2} \text{Var}_U S_2(U) \in \mathcal{O}(r^2 s_{\max}^4)$. Hence S_2 is weakly typical if $r \in o(1)$ or $s_{\max} \in o(1)$.

For the von Neumann entropy, we again make use of Eq. (15), which gives $\lim_{n \rightarrow \infty} \frac{1}{n^2} \text{Var}_U S_1(U) \in \mathcal{O}(r^2 s_{\max}^2)$, which goes to zero if $r \in o(1)$. \square

In summary, we have used our results from Sections 3 and 4 on entanglement in the equal squeezing regime to prove various statements in the unequal

squeezing regime. In particular, under the assumption that Conjecture 10 is true, we prove both the Rényi-2 and von Neumann entropies are weakly typical whenever the subsystem size k scales as $k \in o(n)$. The best known result for the presence of strong typicality in the unequal squeezing regime is when $k \in o(n^{1/3})$ as proven in Ref. [54]. The current status of rigorous results on typical entanglement in Gaussian bosonic systems is summarized in Table 1.

6 Conclusion

In this work, we studied the average and variance of the Rényi-2 and von Neumann entropies in random bosonic Gaussian systems. We computed the Rényi-2 Page curve and Page correction when all the initial squeezing strengths are equal, and we proved various results on the typicality of the Rényi-2 and von Neumann measures of entanglement. Given that the Rényi-2 entropy is a function of only the purity, it is often tractable to measure experimentally. It would be interesting to compare the analytic formula in Theorem 2 to an experimental Gaussian Boson Sampling device to determine how well it is generating and maintaining bipartite entanglement.

We have identified several open problems that would generalize and expand our results. One such open problem is to prove Conjecture 10, which would allow our results on the Page curve to apply more generally. Perhaps the most important remaining task is to complete Table 1 by proving typicality of entanglement in the remaining regimes, such as the regime of unequal squeezing and the von Neumann entropy.

For the latter, we note two potentially fruitful avenues. The first comes from the formula for the von Neumann entropy of a Gaussian state with covariance matrix σ given in Ref. [73]. Let $D_\sigma := \sqrt{\det \sigma} = e^{S_2}$, where S_2 is the Rényi-2 entropy of σ . Ref. [73] derives expressions for all the Rényi- α entropies, including the von Neumann entropy, as functions of D_σ . In our work, we found an expression for $\mathbb{E}_U S_2(U)$ in the regime of equal squeezers by expanding $S_2(U)$ in a power series and exactly computing asymptotic expectation values over the unitary Haar measure. To find the Rényi- α entropy $\mathbb{E}_U S_\alpha(U)$, one could similarly attempt to expand in powers of D_σ , and therefore compute $\mathbb{E}_U S_\alpha(U)$ by computing $\mathbb{E}_U e^{jS_2(U)}$ for various values of j . A second potential way of computing $\mathbb{E}_U S_1(U)$ is similar, where one could use the formula

$$S_1 = \frac{1}{2} \log \det \left[\frac{\sigma + i\Omega}{2} \right] + \frac{1}{2} \text{Tr}[\text{arccoth}(i\Omega\sigma)i\Omega\sigma], \quad (23)$$

where Ω is the $2n \times 2n$ symplectic form given in Eq. (A1) [74, 60, 75]. The equations resulting from using these methods with the Weingarten calculus may potentially be too difficult to simplify at first glance,

as was the case in this work. However, it would be interesting to see if the presence of the $r \mapsto 1 - r$ symmetry is enough, as it was in this work, to reduce the complicated Weingarten expressions to something much more tractable and simple.

In this paper, we have only considered truly Haar-random unitaries. One important question concerns how these results translate to the case where one uses random local passive Gaussian gates to generate random unitary circuits of finite depth. Indeed, an interesting open problem is to determine the depth dependence of entanglement, sampling complexity, and gate complexity in linear optical circuits. Sampling complexity refers to the classical complexity of generating samples from the output probability distribution defined by a fixed depth linear optical circuit, and gate complexity refers to the minimum number of nearest-neighbor beamsplitters required to generate the probability distribution. Currently, the precise relationship between entanglement and complexity is largely unknown. Numerical analyses of entanglement dynamics in linear optical circuits have been reported in Refs. [76, 77]. Partial analytical work was done in Ref. [78], but only in the regime where one or a small number of modes are not initially vacuum. On the contrary, a typical Gaussian Boson Sampling experiment initially squeezes many or all of the modes. Information on such entanglement growth could yield insights on implementations of Gaussian Boson Sampling experiments as well as the complexity of computing output probabilities from such experiments. On the complexity side, many recent works have studied classical simulation and classical sampling complexity of linear optical circuits in certain regimes of low depth and the phase transition at which the complexity passes from easy to hard [79–83]. Indeed, both entanglement and complexity are expected to grow with depth, and further study may reveal that the relationship is even more intimate.

In this work, we have characterized the entanglement properties of Gaussian states such as they arise in Gaussian Boson Sampling. In this setting, we also know that sampling from Fock basis measurements of the Gaussian state is computationally intractable. It remains an exciting question to better understand the role that entanglement plays in this context. An important aspect of this direction is to understand how entanglement and measurement bases interact. After all, some form of non-Gaussianity is crucial to generate complexity in bosonic computations [41].

Acknowledgments

We thank Marcel Hinsche, Marios Ioannou, Kunal Sharma, and Brayden Ware for discussions related to the topic of this paper. JTI, AE, and AVG acknowledge support from the DoE ASCR Accelerated Research in Quantum Computing program (award

No. DE-SC0020312), DARPA SAVaNT ADVENT, AFOSR, DoE QSA, NSF QLCI (award No. OMA-2120757), DoE ASCR Quantum Testbed Pathfinder program (award No. DE-SC0019040), NSF PFCQC program, ARO MURI, and AFOSR MURI. JTI thanks the Joint Quantum Institute at the University of Maryland for support through a JQI fellowship. DH acknowledges funding from the US Department of Defense through a QuICS Hartree fellowship. AD acknowledges funding provided by the Institute for Quantum Information and Matter, an NSF Physics Frontiers Center (NSF Grant PHY-1733907), the National Science Foundation RAISE-TAQS 1839204, and Amazon Web Services, AWS Quantum Program. Specific product citations are for the purpose of clarification only, and are not an endorsement by the authors or NIST.

Appendices

A Preliminaries	11
A.1 Bosonic Gaussian states	11
A.2 Weingarten calculus	12
A.3 Series formula for the Rényi-2 entropy	12
B Rényi-2 and von Neumann entropies — Proof of Proposition 1	14
C Rényi-2 Page curve	15
C.1 Linear term — Proof of Theorem 2	15
C.2 Maximum value — Proof of Corollary 3	20
C.3 Constant term — Proof of Proposition 4	20
D Variance of the Rényi-2 entropy — Proof of Theorem 6	23
References	25

A Preliminaries

In this preliminary appendix, we will establish some notation and equations that will be used throughout the rest of the appendices. In particular, in Appendix A.1, we review bosonic Gaussian states and describe our setup. In Appendix A.2, we describe integration over the unitary group with the Weingarten calculus. Finally, in Appendix A.3, we restrict our attention to the case when all initial squeezing strengths are equal and derive a series formula for the Rényi-2 entropy that is used in many of our proofs.

A.1 Bosonic Gaussian states

Here, we describe the setup and fix the notation required for the proofs of our main results. We consider a very similar setup as the one described in Ref. [54] and use much of the same notation as them. For a review of bosonic Gaussian states, we recommend Ref. [60]. Since we are only interested in entanglement properties, the first moments — displacements — of the Gaussian states will be irrelevant, and we will ignore them.

We consider a system of n bosonic modes. Each mode $1 \leq i \leq n$ is initially in a squeezed state with squeezing strength $s_i \in \mathbb{R}$. Define the diagonal matrix $Z = \text{diag}(e^{2s_1}, \dots, e^{2s_n})$. The initial state can be represented by the covariance matrix $\sigma_0 = Z \oplus Z^{-1}$. Define $A = \frac{1}{2}(Z - Z^{-1})$ and $B = \frac{1}{2}(Z + Z^{-1})$.

The set of all passive Gaussian unitaries — that is, energy-conserving unitaries — acting on n modes is $\text{Sp}(2n) \cap \text{O}(2n)$ which acts on the covariance matrix by conjugation. Here $\text{O}(2n)$ is the orthogonal group of $2n \times 2n$ matrices, and $\text{Sp}(2n)$ is the real symplectic group of $2n \times 2n$ matrices defined with respect to the symplectic form

$$\Omega := \begin{pmatrix} 0_{n \times n} & \mathbb{I}_{n \times n} \\ -\mathbb{I}_{n \times n} & 0_{n \times n} \end{pmatrix}. \quad (\text{A1})$$

Let $\eta: \text{U}(n) \rightarrow \text{Sp}(2n) \cap \text{O}(2n)$ be the isomorphism

$$\eta(U) = \begin{pmatrix} \text{Re}(U) & \text{Im}(U) \\ -\text{Im}(U) & \text{Re}(U) \end{pmatrix}. \quad (\text{A2})$$

We evolve the initial state with covariance matrix σ_0 by a passive Gaussian unitary, which corresponds to a $U \in \text{U}(n)$. The resulting state is $\tilde{\sigma}(U) := \eta(U)\sigma_0\eta(U^\dagger) = \eta(U)\sigma_0\eta(U)^T$.

Define the $k \times n$ matrix P and the $n \times n$ projector Π as

$$P := (\mathbb{I}_{k \times k} \ 0_{k \times (n-k)}), \quad \Pi := \begin{pmatrix} \mathbb{I}_{k \times k} & 0_{k \times (n-k)} \\ 0_{(n-k) \times k} & 0_{(n-k) \times (n-k)} \end{pmatrix} = \mathbb{I}_{k \times k} \oplus 0_{(n-k) \times (n-k)}. \quad (\text{A3})$$

Then let $\hat{P} := P \oplus P$ and $\hat{\Pi} = \Pi \oplus \Pi$. The covariance matrix corresponding to the reduced state on the first $k \leq n$ modes is $\sigma(U) := \hat{P}\tilde{\sigma}(U)\hat{P}^T$. Denote the element-wise complex conjugate of the unitary U by \bar{U} , and the conjugate transpose by U^\dagger . By simply doing the matrix multiplication, one finds that

$$\sigma(U) = \frac{1}{2} \begin{pmatrix} P [UBU^\dagger + \bar{U}B\bar{U}^\dagger + UA\bar{U}^\dagger + \bar{U}AU^\dagger] P^T & -iP [UBU^\dagger - \bar{U}B\bar{U}^\dagger - UA\bar{U}^\dagger + \bar{U}AU^\dagger] P^T \\ iP [UBU^\dagger - \bar{U}B\bar{U}^\dagger + UA\bar{U}^\dagger - \bar{U}AU^\dagger] P^T & P [\bar{U}BU^\dagger + \bar{U}B\bar{U}^\dagger - UA\bar{U}^\dagger - \bar{U}AU^\dagger] P^T \end{pmatrix}. \quad (\text{A4})$$

Note that $\sigma(U)$ is a covariance matrix on k modes and is correspondingly a positive $2k \times 2k$ matrix. Throughout this paper, we define $r := k/n$. One can derive from Eq. (A6) that the average over the Haar measure is $\mathbb{E}_U UBU^\dagger = \mathbb{E}_U \bar{U}B\bar{U}^\dagger = \frac{\text{Tr } B}{n} \mathbb{I}_{n \times n}$, whereas all the other terms have expectation value 0 since they do not contain an even number of U 's and \bar{U} 's. Therefore,

$$\mathbb{E}_{U \in \text{U}(n)} \sigma(U) = \frac{\text{Tr } B}{n} \mathbb{I}_{2k \times 2k}. \quad (\text{A5})$$

The *symplectic eigenvalues* of $\sigma(U)$ are the positive eigenvalues of $i\Omega\sigma(U)$. There are k symplectic eigenvalues labeled as ν_i for $1 \leq i \leq k$. Let the von Neumann entropy of the reduced state be $S_1(U)$, and the Rényi-2 entropy of the reduced state be $S_2(U)$. Then $S_j(U) = \sum_{i=1}^k h_j(\nu_i)$, where $h_1(x) = \frac{x+1}{2} \log \frac{x+1}{2} - \frac{x-1}{2} \log \frac{x-1}{2}$ and $h_2(x) = \log x$ [60]. The Rényi-2 entropy of the reduced state takes a particularly nice form in terms of the standard eigenvalues of $\sigma(U)$, namely $S_2(U) = \frac{1}{2} \log \det \sigma(U) = \frac{1}{2} \text{Tr} \log \sigma(U)$. The dependence of $S_1(U)$ and $S_2(U)$ on r , n , and s is implicit.

A.2 Weingarten calculus

Since we are interested in average entanglement, we will be averaging over the unitary group $\text{U}(n)$ with respect to the unique unit normalized Haar measure. To do so, we will use the *Weingarten calculus* [69, 70]. For a matrix U , let U_{ij} denote the entry in row i and column j . Then,

$$\mathbb{E}_{U \in \text{U}(n)} U_{i_1 j_1} \dots U_{i_q j_q} \bar{U}_{i'_1 j'_1} \dots \bar{U}_{i'_q j'_q} = \sum_{\sigma, \tau \in S_q} \delta_{i_1 i'_{\sigma(1)}} \dots \delta_{i_q i'_{\sigma(q)}} \delta_{j_1 j'_{\tau(1)}} \dots \delta_{j_q j'_{\tau(q)}} \text{Wg}(\sigma \tau^{-1}, n), \quad (\text{A6})$$

where S_q denotes the permutation group on q elements. Wg is called the *Weingarten function*. In our proofs, we will need the asymptotic form of the Weingarten function, which is given by

$$\text{Wg}(\sigma, n) = \frac{1}{n^{q+|\sigma|}} \prod_i (-1)^{|c_i|-1} C_{|c_i|-1} + \mathcal{O}(n^{-q-|\sigma|-2}), \quad (\text{A7})$$

where $|\sigma|$ denotes the minimum number of transpositions needed to generate the permutation σ , $C_m = (2m)!/m!(m+1)!$ is the m^{th} Catalan number, and σ is a product of cycles c_i of length $|c_i|$.

A.3 Series formula for the Rényi-2 entropy

In this subappendix, we Taylor expand $S_2(U) = \frac{1}{2} \text{Tr} \log \sigma(U)$ and derive a series formula for the Rényi-2 entropy when all the initial squeezing values are equal. Hence, for each $1 \leq i \leq n$, we set $s_i = s$. Crucially, the resulting formula is a series in the squeezing strength s , and the effect of the unitary U is separated from that of the squeezing strength s .

We would like to apply the Taylor series for the matrix logarithm, and hence must first consider its convergence. We find that

$$\|\sigma(U) - \mathbb{I}_{2k \times 2k}\| = \left\| \hat{P} \eta(U) (\sigma_0 - \mathbb{I}_{2n \times 2n}) \eta(U)^T \hat{P}^T \right\| \quad (\text{A8})$$

$$= \left\| \hat{\Pi} \eta(U) (\sigma_0 - \mathbb{I}_{2n \times 2n}) \eta(U)^T \hat{\Pi}^T \right\| \quad (\text{A9})$$

$$\leq \left\| \hat{\Pi} \right\|^2 \left\| \eta(U) (\sigma_0 - \mathbb{I}_{2n \times 2n}) \eta(U)^T \right\| \quad (\text{A10})$$

$$= \left\| \eta(U) (\sigma_0 - \mathbb{I}_{2n \times 2n}) \eta(U)^T \right\| \quad (\text{A11})$$

$$= \left\| \sigma_0 - \mathbb{I}_{2n \times 2n} \right\| \quad (\text{A12})$$

$$= \max\{|\text{e}^{2s} - 1|, |\text{e}^{-2s} - 1|\} \quad (\text{A13})$$

$$= \text{e}^{2|s|} - 1, \quad (\text{A14})$$

and therefore the Taylor series for \log ,

$$\log \sigma(U) = - \sum_{j=1}^{\infty} \frac{(-1)^j}{j} (\sigma(U) - \mathbb{I}_{2k \times 2k})^j, \quad (\text{A15})$$

converges for all $|s| < R := \frac{1}{2} \log 2$. Since $\sigma(U)$ is a positive, real symmetric matrix, this expression is indeed real and nonnegative. To make this work for all $s \in \mathbb{R}$, we can let $N \geq 1$ and consider

$$S_2(U) = \frac{1}{2} \log \det \sigma(U) \quad (\text{A16})$$

$$= \frac{1}{2} \log \left(N^{2k} \det \frac{\sigma(U)}{N} \right) \quad (\text{A17})$$

$$= k \log N + \frac{1}{2} \text{Tr} \log \left(\frac{1}{N} \sigma(U) \right), \quad (\text{A18})$$

which follows from the fact that the determinant of products is equal to the product of determinants. For any given $s \in \mathbb{R}$, we can choose N large enough such that $\left\| \frac{1}{N} \sigma(U) - \mathbb{I} \right\| < 1$ and therefore the Taylor series for \log can be used. We therefore find that for large enough N ,

$$S_2(U) = rn \log N - \frac{1}{2} \text{Tr} \sum_{j=1}^{\infty} \frac{(-1)^j}{j} \left(\frac{1}{N} \sigma(U) - \mathbb{I}_{2k \times 2k} \right)^j. \quad (\text{A19})$$

When all $s_i = s$, A simplifies to $A = \sinh(2s) \mathbb{I}_{n \times n}$ and B to $B = \cosh(2s) \mathbb{I}_{n \times n}$, and therefore $\sigma(U)$ simplifies to $\sigma(U) = \cosh(2s) \mathbb{I}_{2k \times 2k} + \sinh(2s) M$, where

$$M := \begin{pmatrix} P \text{Re}(\bar{U}U^\dagger)P^T & P \text{Im}(\bar{U}U^\dagger)P^T \\ P \text{Im}(\bar{U}U^\dagger)P^T & -P \text{Re}(\bar{U}U^\dagger)P^T \end{pmatrix}. \quad (\text{A20})$$

With Eq. (A19),

$$S_2(U) = nr \log N - \frac{1}{2} \text{Tr} \sum_{j=1}^{\infty} \frac{(-1)^j}{j} \left(\left[\frac{1}{N} \cosh(2s) - 1 \right] \mathbb{I}_{2k \times 2k} + \frac{1}{N} \sinh(2s) M \right)^j \quad (\text{A21})$$

$$= nr \log N - \frac{1}{2} \text{Tr} \sum_{j=1}^{\infty} \frac{(-1)^j}{j} \sum_{\ell=0}^j \binom{j}{\ell} \frac{1}{N^\ell} \sinh^\ell(2s) \left[\frac{1}{N} \cosh(2s) - 1 \right]^{j-\ell} M^\ell \quad (\text{A22})$$

$$= nr \log N - rn \sum_{j=1}^{\infty} \frac{(-1)^j}{j} \left[\frac{1}{N} \cosh(2s) - 1 \right]^j \quad (\text{A23})$$

$$- \frac{1}{2} \text{Tr} \sum_{j=1}^{\infty} \frac{(-1)^j}{j} \sum_{\ell=1}^j \binom{j}{\ell} \frac{1}{N^\ell} \sinh^\ell(2s) \left[\frac{1}{N} \cosh(2s) - 1 \right]^{j-\ell} \text{Tr} M^\ell$$

$$= nr \log N + nr \log(\cosh(2s)/N) - \frac{1}{2} \sum_{j=1}^{\infty} \frac{(-1)^j}{j} \sum_{\ell=1}^j \binom{j}{\ell} \frac{1}{N^\ell} \sinh^\ell(2s) \left[\frac{1}{N} \cosh(2s) - 1 \right]^{j-\ell} \text{Tr} M^\ell \quad (\text{A24})$$

$$= nr \log \cosh(2s) - \frac{1}{2} \sum_{\ell=1}^{\infty} \frac{1}{N^\ell} \sinh^\ell(2s) \text{Tr} M^\ell \sum_{j=\ell}^{\infty} \frac{(-1)^j}{j} \binom{j}{\ell} \left[\frac{1}{N} \cosh(2s) - 1 \right]^{j-\ell} \quad (\text{A25})$$

$$= nr \log \cosh(2s) - \frac{1}{2} \sum_{\ell=1}^{\infty} \frac{1}{N^\ell} \sinh^\ell(2s) \text{Tr} M^\ell \frac{(-1)^\ell}{\ell} (N \operatorname{sech}(2s))^\ell \quad (\text{A26})$$

$$= nr \log \cosh(2s) - \frac{1}{2} \sum_{\ell=1}^{\infty} \frac{(-1)^\ell}{\ell} \tanh^\ell(2s) \text{Tr} M^\ell. \quad (\text{A27})$$

Of course, $S_2(U)$ is independent of the choice of N , and hence the N dependence has dropped out.

The only thing left to compute is $\text{Tr} M^\ell$. Since we are now only dealing with traces, we can replace P with Π in M . This nicely simplifies some formulas, since Π is a square matrix, and indeed a projector, whereas P is a rectangular matrix. Henceforth, we will therefore let

$$M = \begin{pmatrix} \Pi R \Pi & \Pi I \Pi \\ \Pi I \Pi & -\Pi R \Pi \end{pmatrix}, \quad (\text{A28})$$

where $R = \text{Re}(\bar{U}U^\dagger)$ and $I = \text{Im}(\bar{U}U^\dagger)$. Then, doing the matrix multiplication,

$$M^2 = \begin{pmatrix} \Pi R \Pi R \Pi + \Pi I \Pi I \Pi & \Pi R \Pi I \Pi - \Pi I \Pi R \Pi \\ \Pi I \Pi R \Pi - \Pi R \Pi I \Pi & \Pi R \Pi R \Pi + \Pi I \Pi I \Pi \end{pmatrix}. \quad (\text{A29})$$

We then notice that $M^2 = \begin{pmatrix} D & E \\ -E & D \end{pmatrix}$, where $D = \text{Re}(W)$, $E = \text{Im}(W)$, and $W = \Pi U \bar{U}^\dagger \Pi \bar{U} U^\dagger \Pi$. It is then easy to verify that $M^{2j} = \begin{pmatrix} D_j & E_j \\ -E_j & D_j \end{pmatrix}$, where $D_{j+1} = D_j D - E_j E$, $E_{j+1} = D_j E + E_j D$, $D_1 = D$, and $E_1 = E$.

It is then also easy to verify that this recurrence relation is solved by $D_j = \text{Re}(W^j)$ and $E_j = \text{Im}(W^j)$. Finally, $\text{Tr } M^{2j} = 2 \text{Tr } D_j$.

Using this recurrence, we can do the matrix multiplication $M^{2j+1} = M^{2j}M$ to find that $\text{Tr } M^{2j+1} = 0$. Hence, we only need to worry about even powers of M , giving

$$S_2(U) = nr \log \cosh(2s) - \frac{1}{2} \sum_{\ell=1}^{\infty} \frac{1}{2\ell} \tanh^{2\ell}(2s) \text{Tr } M^{2\ell}. \quad (\text{A30})$$

We then use that $\text{Tr } M^{2j} = 2 \text{Tr } D_j$ to find that

$$S_2(U) = nr \log \cosh(2s) - \sum_{\ell=1}^{\infty} \frac{1}{2\ell} \tanh^{2\ell}(2s) \text{Tr } \text{Re } W^{\ell}. \quad (\text{A31})$$

Furthermore, since W is Hermitian, $\text{Tr } W^j = \text{Tr } \bar{W}^j$. Therefore, we arrive at

$$S_2(U) = nr \log \cosh(2s) - \sum_{\ell=1}^{\infty} \frac{1}{2\ell} \tanh^{2\ell}(2s) \text{Tr } W^{\ell} \quad (\text{A32})$$

$$= n \sum_{\ell=1}^{\infty} \frac{\tanh^{2\ell}(2s)}{2\ell} \left(r - \frac{1}{n} \text{Tr } W^{\ell} \right), \quad (\text{A33})$$

where in the last step we used the Taylor expansion of $\log \cosh(\text{arctanh } t)$ in the variable $t = \tanh(2s)$.

Eqs. (A32) and (A33) hint at why equal initial squeezings simplify the problem of studying averaged entanglement properties. Specifically, the contribution from the squeezing strength and the contribution from the unitary are separated. Thus, to determine averaged entanglement properties, we only need to deal with the matrix $W = \Pi U U^T \Pi \bar{U} \bar{U}^T \Pi$.

B Rényi-2 and von Neumann entropies — Proof of Proposition 1

In this appendix, we prove Proposition 1. We derive the maximum of the Rényi-2 and von Neumann entropies, and the former will be of use later when we derive the Rényi-2 Page correction. Furthermore, we prove that the von Neumann entropy can be bounded by the Rényi-2 entropy. In this way, our results on the Rényi-2 Page curve can be used to bound the von Neumann Page curve.

We begin by proving that the maximum of the Rényi-2 entropy is

$$\max_{U \in \text{U}(n)} S_2(U) = n \min(r, 1-r) \log \cosh(2s). \quad (\text{B1})$$

Proof. Recall that $W = F^{\dagger}F$, where $F = \Pi \bar{U} U^{\dagger} \Pi$. Therefore, W is a nonnegative operator, and $\text{Tr } W^{\ell} \geq 0$. The proposition then immediately follows from Eq. (A32) if we can show that for every $r \leq 1/2$, there exists a unitary such that $W = 0$. The case when $r > 1/2$ is taken care of by the fact that the Rényi-2 entropy is symmetric under $r \mapsto 1-r$ since the global state on the n modes is pure. Therefore, we now assume that $r \leq 1/2$, and we show that there are unitaries $U \in \text{U}(n)$ such that $W = 0$.

Since $F^{\dagger}F$ is nonnegative, we need to prove that there exists a unitary such that $F = \Pi \bar{U} U^{\dagger} \Pi = 0$. Hence we must prove that there exists a U such that $(\bar{U} U^{\dagger})_{ji} = 0$ for all $1 \leq i, j \leq k = rn$. Equivalently, we can conjugate the expression, giving $(\bar{U} U^{\dagger})_{ji} = \sum_{a=1}^n U_{ia} U_{ja} = 0$. Therefore, we just need to find a set of $k = rn$ orthonormal vectors $\mathcal{S} = \{|\psi_1\rangle, \dots, |\psi_k\rangle\}$ in \mathbb{C}^n such that $\langle \bar{\psi}_i | \psi_j \rangle = 0$. Let $|\psi_i\rangle = |R_i\rangle + i|I_i\rangle$ for real vectors $|R_i\rangle$ and $|I_i\rangle$. Since \mathcal{S} is orthonormal, we find that

$$\delta_{ij} = \langle \psi_i | \psi_j \rangle \quad (\text{B2})$$

$$= (\langle R_i | - i \langle I_i |)(|R_j\rangle + i|I_j\rangle) \quad (\text{B3})$$

$$= \langle R_i | R_j \rangle + \langle I_i | I_j \rangle + i(\langle R_i | I_j \rangle - \langle I_i | R_j \rangle). \quad (\text{B4})$$

The condition that $\langle \bar{\psi}_i | \psi_j \rangle = 0$ implies that

$$0 = \langle \bar{\psi}_i | \psi_j \rangle \quad (\text{B5})$$

$$= (\langle R_i | + i \langle I_i |)(|R_j\rangle + i|I_j\rangle) \quad (\text{B6})$$

$$= \langle R_i | R_j \rangle - \langle I_i | I_j \rangle + i(\langle R_i | I_j \rangle + \langle I_i | R_j \rangle). \quad (\text{B7})$$

Hence, we just need to find k vectors $|R_i\rangle \in \mathbb{R}^n$ and k vectors $|I_i\rangle \in \mathbb{R}^n$ satisfying

1. $\langle R_i | R_j \rangle + \langle I_i | I_j \rangle = \delta_{ij}$,
2. $\langle R_i | R_j \rangle - \langle I_i | I_j \rangle = 0$,
3. $\langle R_i | I_j \rangle + \langle I_i | R_j \rangle = 0$, and
4. $\langle R_i | I_j \rangle - \langle I_i | R_j \rangle = 0$,

for all $i, j \in \{1, \dots, k\}$. This is trivial since $k \leq n/2$, and we give one construction here. Let $\mathcal{R} = \{|R_1\rangle, \dots, |R_k\rangle\}$ be an orthogonal set satisfying $\langle R_i | R_j \rangle = \delta_{ij}/2$. Since $\dim[\text{span}(\mathcal{R})^\perp] = n - k \geq k$, we can choose another orthogonal set $\mathcal{I} = \{|I_1\rangle, \dots, |I_k\rangle\}$ satisfying $\langle I_i | I_j \rangle = \delta_{ij}/2$ such that $\text{span}(\mathcal{I}) \cap \text{span}(\mathcal{R}) = \{0\}$. This immediately means that condition 3 and 4 are satisfied, because $\langle R_i | I_j \rangle = 0$ for all i and j . Furthermore, condition 1 is satisfied because $\langle R_i | R_j \rangle + \langle I_i | I_j \rangle = \delta_{ij}/2 + \delta_{ij}/2 = \delta_{ij}$. Finally, condition 2 is satisfied because $\langle R_i | R_j \rangle - \langle I_i | I_j \rangle = \delta_{ij}/2 - \delta_{ij}/2 = 0$. \square

Next, we use Eq. (B1) to prove that the maximum of the von Neumann entropy is

$$\max_{U \in \text{U}(n)} S_1(U) = n \min(r, 1-r) h_1(\cosh(2s)). \quad (\text{B8})$$

Proof. From Ref. [73], the Rényi- α entropies for $\alpha \geq 1$ are all increasing with increasing $\det \sigma$. Therefore, when S_2 increases, so does S_α . It follows that the unitary that maximizes S_2 also maximizes S_1 . Then, when $r \leq 1/2$, Eqs. (A20) and (B1) together imply that S_1 is maximized when $\sigma = \cosh(2s) \mathbb{I}_{2k \times 2k}$ and thus when $\nu_i = \cosh(2s)$ for each i , yielding $\max_U S_1(U) = nrh_1(\cosh(2s))$. The case when $r > 1/2$ is taken care of by the symmetry of the von Neumann entropy when $r \mapsto 1-r$ since we are dealing with a bipartite system [28]. \square

Finally, we prove the bound

$$S_1(U) < S_2(U) + n \min(r, 1-r)(1 - \log 2), \quad (\text{B9})$$

which was first derived in Ref. [58, Eq. 15].

Proof. Recall that $S_j = \sum_{i=1}^k h_j(\nu_i)$, where $\nu_i \geq 1$ are the symplectic eigenvalues of the covariance matrix σ , $h_2(x) = \log x$, and $h_1(x) = \frac{x+1}{2} \log \frac{x+1}{2} - \frac{x-1}{2} \log \frac{x-1}{2}$ (see Appendix A.1). We will prove that $h_1(x) < h_2(x) + 1 - \log 2$, which proves the claim. We do this by noting that $h_1(x) - h_2(x)$ is monotonically increasing, so that for any x , $h_1(x) - h_2(x) < \lim_{y \rightarrow \infty} (h_1(y) - h_2(y))$. A simple calculation shows that this limit is equal to $1 - \log 2$. \square

C Rényi-2 Page curve

In this appendix, we prove our main results on the Rényi-2 Page curve. In Appendix C.1, we prove that asymptotically in n , $\mathbb{E}_{U \in \text{U}(n)} S_2(U) = n\alpha(s, r) - \lambda(s, r) + o(1)$. We then derive the exact formula for the linear term $\alpha(s, r)$. Then, in Appendix C.2, we simplify $\alpha(s, 1/2)$ and use the result to prove a simple expression for the Rényi-2 Page correction at $r = 1/2$. Finally, in Appendix C.3, we derive a formula for the constant term $\lambda(s, r)$ up to constant factors $a^{(\ell)}$. We then derive the formula for $a^{(\ell)}$, simplify it for $\ell \in \{1, 2, 3, 4, 5\}$, and conjecture the exact value of $a^{(\ell)}$ for all ℓ .

C.1 Linear term — Proof of Theorem 2

Theorem 2 concerns the expectation value of $S_2(U)$ over $\text{U}(n)$ when all the initial squeezing values are equal. From Eq. (A32), we see that it only remains to compute $\mathbb{E}_U \text{Tr } W^\ell$. Writing the matrix multiplication of $W^\ell = (\Pi U U^T \Pi \bar{U} \bar{U}^T \Pi)^\ell$ in terms of the matrix entries of U and Π and simplifying, we find that

$$\begin{aligned} \text{Tr } W^\ell = & \sum_{i_1, \dots, i_{2\ell}=1}^k \sum_{i'_1, \dots, i'_{2\ell}=1}^k \sum_{j_1, \dots, j_{2\ell}=1}^n \sum_{j'_1, \dots, j'_{2\ell}=1}^n \\ & \delta_{i'_{2\ell}, i_1} \delta_{i'_1, i_2} \delta_{i'_2, i_3} \delta_{i'_3, i_4} \dots \delta_{i'_{2\ell-1}, i_{2\ell}} \\ & \times \delta_{j_1, j_2} \delta_{j'_1, j'_2} \delta_{j_3, j_4} \delta_{j'_3, j'_4} \dots \delta_{j_{2\ell-1}, j_{2\ell}} \delta_{j'_{2\ell-1}, j'_{2\ell}} \\ & \times U_{i_1, j_1} \dots U_{i_{2\ell}, j_{2\ell}} \bar{U}_{i'_1, j'_1} \dots \bar{U}_{i'_{2\ell}, j'_{2\ell}}. \end{aligned} \quad (\text{C1})$$

We note that this is a simple result of matrix multiplication. The restriction on the i and i' indices to $\{1, \dots, k\}$ is a result of the fact that $\Pi_{a,a} = 0$ for all $a > k$. Applying Eq. (A6), we immediately find that

$$\begin{aligned} \mathbb{E}_{U \in \mathbb{U}(n)} \text{Tr } W^\ell &= \sum_{i_1, \dots, i_{2\ell}=1}^k \sum_{i'_1, \dots, i'_{2\ell}=1}^k \sum_{j_1, \dots, j_{2\ell}=1}^n \sum_{j'_1, \dots, j'_{2\ell}=1}^n \sum_{\sigma, \tau \in S_{2\ell}} \text{Wg}(\sigma\tau^{-1}, n) \\ &\quad \times \delta_{i'_{2\ell}, i_1} \delta_{i'_1, i_2} \delta_{i'_2, i_3} \dots \delta_{i'_{2\ell-1}, i_{2\ell}} \\ &\quad \times \delta_{j_1, j_2} \delta_{j'_1, j'_2} \dots \delta_{j_{2\ell-1}, j_{2\ell}} \delta_{j'_{2\ell-1}, j'_{2\ell}} \\ &\quad \times \delta_{i_1, i'_{\sigma(1)}} \dots \delta_{i_{2\ell}, i'_{\sigma(2\ell)}} \\ &\quad \times \delta_{j_1, j'_{\tau(1)}} \dots \delta_{j_{2\ell}, j'_{\tau(2\ell)}}. \end{aligned} \tag{C2}$$

Simplifying Eq. (C2) at first seems impossible, but it will turn out that we do not need to. All we need to learn from it is the following lemma.

Lemma C.1. *Fix a positive integer ℓ . There exist coefficients $\alpha_d^{(\ell)}$ for $d \in \{\ell+1, \ell+2, \dots, 2\ell\}$ such that*

$$f_\ell(r) := \lim_{n \rightarrow \infty} \mathbb{E}_{U \in \mathbb{U}(n)} \frac{1}{n} \text{Tr } W^\ell = \sum_{d=\ell+1}^{2\ell} \alpha_d^{(\ell)} r^d. \tag{C3}$$

Proof. The proof will proceed as follows. First, we will prove that $\text{Tr } W^\ell$ contains a term proportional to n and no terms proportional to n^a for any $a > 1$. Therefore, $f_\ell(r)$ is indeed independent of n . Next, we will prove that $f_\ell(r)$ has no terms r^a for $a > 2\ell$ and $a \leq \ell$. Throughout this proof, we interpret the delta functions in Eq. (C2) as constraints on the summations. Different permutations on the indices result in a different number of constraints and hence terms with different powers of n and k .

Recall that $W = F^\dagger F$ where $F = \Pi \bar{U} U^\dagger \Pi$. Therefore, $\text{Tr } W$ is equal to the Frobenius norm $\|F\|_F^2$, which is the sum of the square absolute values of the entries of F . Thus, by removing the projector Π from W (i.e. setting it to \mathbb{I}), the trace cannot decrease. It follows that the presence of Π cannot increase the trace $\text{Tr } W^\ell$. Getting rid of the Π from W and using the cyclic nature of the trace, we see that $\text{Tr } W^\ell \leq n$. Furthermore, consider Eq. (C2) with $\sigma = \tau$ defined by $\sigma(1) = 2\ell, \sigma(2) = 1, \sigma(3) = 2, \dots, \sigma(2\ell) = 2\ell - 1$. Then, $\sigma\tau^{-1}$ is the identity, $\text{Wg}(\sigma\tau^{-1}, n)$ contributes a factor of $n^{-2\ell}$. With this σ , the sum over the i and i' yields a factor of $k^{2\ell}$. Finally, with this chosen τ , the sum over j will yield $\delta_{j'_1, j'_2} \dots \delta_{j'_{2\ell-1}, j'_{2\ell}} \delta_{j'_1, j'_{2\ell}} \delta_{j'_2, j'_1} \dots \delta_{j'_{2\ell}, j'_{2\ell-1}}$. Then summing over j' , we get a single factor of n . Hence, the term with the specific permutation described above yields a term of the form $nk^{2\ell}n^{-2\ell} = nr^{2\ell}$.

We have shown that there is a term proportional to n and that there are no terms proportional to n^2, n^3 , etc. Since we are working asymptotically in n , we can therefore ignore all terms proportional to $\frac{1}{n^a}$ for every nonnegative a . This proves that $\lim_{n \rightarrow \infty} \mathbb{E}_{U \in \mathbb{U}(n)} \frac{1}{n} \text{Tr } W^\ell$ is independent of n and only depends on r , which justifies the definition of the function $f_\ell(r)$. The only thing left to show is that $f_\ell(r)$ has only terms $r^{\ell+1}$ through $r^{2\ell}$. So we only need to show that there are no terms r^a for $a > 2\ell$ and $a \leq \ell$. We begin with the former.

To look at powers of r , it is sufficient to look at the powers of k . We therefore restrict our attention to the sum over i and i' in Eq. (C2). The sum over $i_1, \sum_{i_1=1}^k$, will give either a factor of 1 or a factor of k depending on how the index i_1 is constrained by the Kronecker delta functions, and similarly for i_2 through $i_{2\ell}$. In order to get the highest power of k , we require the fewest constraints on i and i' (i.e. the fewest *distinct* Kronecker deltas). Hence, we require σ to be the permutation satisfying

$$\delta_{i'_{2\ell}, i_1} \delta_{i'_1, i_2} \delta_{i'_2, i_3} \dots \delta_{i'_{2\ell-1}, i_{2\ell}} = \delta_{i_1, i'_{\sigma(1)}} \dots \delta_{i_{2\ell}, i'_{\sigma(2\ell)}}. \tag{C4}$$

This permutation is $\sigma(1) = 2\ell, \sigma(2) = 1, \sigma(3) = 2, \dots, \sigma(2\ell) = 2\ell - 1$. With this σ , we see that a sum over i and i' will give a factor of $k^{2\ell}$. Hence, 2ℓ is the highest power of k that can be achieved.

Next we need to show that $\ell+1$ is the lowest power of k that can be achieved. The sum over j and j' can give at most a factor of n^ℓ . This is because the first line of delta functions, $\delta_{j_1, j_2} \delta_{j'_1, j'_2} \dots \delta_{j_{2\ell-1}, j_{2\ell}} \delta_{j'_{2\ell-1}, j'_{2\ell}}$, reduces the sum over 2ℓ indices j and 2ℓ indices j' down to just a sum over ℓ indices j and ℓ indices j' . The second line of delta functions, $\delta_{j_1, j'_{\tau(1)}} \dots \delta_{j_{2\ell}, j'_{\tau(2\ell)}}$, cannot be made equivalent to the first line by any choice of τ ; in fact, the second line imposes all new constraints. Therefore, this line further reduces the sum over ℓ indices j and ℓ indices j' down to just a sum over ℓ indices j (or ℓ indices j' , but not both). Hence, the highest power of n that we get from the summations over j and j' is n^ℓ . Putting this together with the fact that asymptotically $\text{Wg}(\pi, n)$ is at most $n^{-2\ell}$, we find that any term coming from Eq. (C2) is at most $n^{-\ell} \times (\text{dependence on } k)$.

Therefore, any powers of k that are less than $\ell + 1$ can be ignored; if the sum over i and i' yields a term that is k^a for some $a \leq \ell$, then that term will be constant or decreasing with n . But from above, we already have terms that are proportional to n resulting from Eq. (C2), and so terms that are constant or decreasing can be ignored. \square

As alluded to in the proof of the lemma, the asymptotic form of $\mathbb{E}_U \text{Tr } W^\ell$ will be a term linear in n times a function of r , plus a term constant in n times a function of r , plus terms that decay to zero asymptotically with n . Hence, $\mathbb{E}_U S_2(U) = n\alpha(s, r) - \lambda(s, r) + o(1)$. In this section, we are interested only in the linear term $\alpha(s, r)$ because we are computing $\mathbb{E}_U S_2(U)/n$. However, in the next section, we prove Proposition 4 which provides the form of $\lambda(s, r)$.

Actually computing $\alpha_d^{(\ell)}$ from Eq. (C2) seems challenging. However, there is a nice workaround that uses what we know about the Rényi-2 entropy being symmetric under $r \mapsto 1 - r$, as the total state on the n modes is pure.

Lemma C.2. *Let $\alpha_d^{(\ell)}$ be as in Lemma C.1. Then*

$$\alpha_d^{(\ell)} = 2(-1)^{d-\ell-1} \binom{2\ell-1}{\ell-1} \binom{\ell}{d-\ell-1} \frac{2\ell-d+1}{(d-1)d}. \quad (\text{C5})$$

Note that this corresponds to the sequence [A062991](#) on OEIS [84].

Proof. From Lemma C.1, asymptotically $\mathbb{E}_U \text{Tr } W^\ell = n f_\ell(r)$ for a polynomial f_ℓ that is a sum over terms of degree $\ell + 1$ through 2ℓ in r . From Eq. (A33), we know that

$$\lim_{n \rightarrow \infty} \mathbb{E}_U \frac{1}{n} S_2(U) = \sum_{\ell=1}^{\infty} \frac{t^{2\ell}}{2\ell} (r - f_\ell(r)), \quad (\text{C6})$$

where $-1 < t = \tanh(2s) < 1$. This whole function must be symmetric under $r \rightarrow 1 - r$ at every order in t . The reason it must be symmetric at every order is as follows. Suppose we choose s such that $t \in \mathcal{O}(n^{-1/2})$. Then $\mathbb{E}_U S_2(U) = \frac{1}{2}nt^2(r - f_1(r)) + o(1)$. Hence, $r - f_1(r)$ must be symmetric since S_2 is symmetric. We then choose s such that $t \in \mathcal{O}(n^{-1/4})$ and consider $\mathbb{E}_U [S_2(U) - \frac{1}{2}nt^2(r - f_1(r))]$, which will be equal to $\frac{1}{4}nt^4(r - f_2(r)) + o(1)$. Since S_2 is symmetric and $r - f_1(r)$ is symmetric, it follows that $r - f_2(r)$ is symmetric. We then continue inductively like this to prove that $r - f_\ell(r)$ must be symmetric for all ℓ .

We therefore find that for all ℓ , $f_\ell(r)$ must satisfy $r - f_\ell(r) = 1 - r - f_\ell(1 - r)$, or

$$f_\ell(r) - f_\ell(1 - r) = 2r - 1. \quad (\text{C7})$$

We then plug in the form of $f_\ell(r) = \sum_{d=\ell+1}^{2\ell} \alpha_d^{(\ell)} r^d$. Eq. (C7) must hold for every r , and therefore we can equate the coefficients in front of each r^d term to get a system of equations that can be solved for the values of $\alpha_d^{(\ell)}$. We then find that

$$2r - 1 = \sum_{d=\ell+1}^{2\ell} \alpha_d^{(\ell)} (r^d - (1 - r)^d) \quad (\text{C8})$$

$$= \sum_{d=\ell+1}^{2\ell} \alpha_d^{(\ell)} r^d - \sum_{d=\ell+1}^{2\ell} \alpha_d^{(\ell)} \sum_{j=0}^d \binom{d}{j} (-r)^j \quad (\text{C9})$$

$$= \sum_{d=\ell+1}^{2\ell} \alpha_d^{(\ell)} r^d - \sum_{j=0}^{2\ell} (-r)^j \sum_{d=\max(j, \ell+1)}^{2\ell} \alpha_d^{(\ell)} \binom{d}{j} \quad (\text{C10})$$

$$= \sum_{d=\ell+1}^{2\ell} \alpha_d^{(\ell)} r^d - \sum_{d=0}^{2\ell} (-r)^d \sum_{j=\max(d, \ell+1)}^{2\ell} \alpha_j^{(\ell)} \binom{j}{d} \quad (\text{C11})$$

$$= - \sum_{d=0}^{\ell} (-r)^d \sum_{j=\ell+1}^{2\ell} \alpha_j^{(\ell)} \binom{j}{d} + \sum_{d=\ell+1}^{2\ell} r^d \left(\alpha_d^{(\ell)} - (-1)^d \sum_{j=d}^{2\ell} \alpha_j^{(\ell)} \binom{j}{d} \right). \quad (\text{C12})$$

Equating the degrees in r , we see the following conditions:

1. $1 = \sum_{j=\ell+1}^{2\ell} \alpha_j^{(\ell)},$
2. $2 = \sum_{j=\ell+1}^{2\ell} j \alpha_j^{(\ell)},$
3. For $2 \leq d \leq \ell$, $0 = \sum_{j=\ell+1}^{2\ell} \binom{j}{d} \alpha_j^{(\ell)},$
4. For $\ell+1 \leq d \leq 2\ell$, $\alpha_d^{(\ell)} = (-1)^d \sum_{j=d}^{2\ell} \binom{j}{d} \alpha_j^{(\ell)}.$

Note that one can derive equivalent conditions by requiring that the polynomial $(x + 1/2) - f_\ell(x + 1/2)$ is even in x . Condition 2 and condition 3 together constitute a linear system of ℓ linearly independent equations. To verify this, one must show that $\det C \neq 0$ where C is the $\ell \times \ell$ matrix with entries $C_{ij} = \binom{\ell+i}{j}$. As noted by Benoit Cloitre in OEIS sequence [A000984](#) [84], $\det C = \binom{2\ell}{\ell}$. For a proof, see Ref. [85].

Since we have ℓ linearly independent equations for ℓ variables α , if there is a solution to the four conditions then there is a *unique* solution. One can then verify that a solution to the four conditions, and therefore *the* solution, is given by Eq. (C5).

When verifying the four conditions, one finds that the right hand side of all four of the conditions can be simplified in terms of the hypergeometric function ${}_2F_1$ defined as

$${}_2F_1(-m, b; c; r) = \sum_{a=0}^m (-1)^a \binom{m}{a} \frac{(c-1)!(a+b-1)!}{(b-1)!(a+c-1)!} r^a \quad (\text{C13})$$

when m is nonnegative [63–67]. For example, condition 4, written as $(-1)^d = \sum_{j=d}^{2\ell} \binom{j}{d} (\alpha_j^{(\ell)} / \alpha_d^{(\ell)})$, reduces to $(-1)^d = {}_2F_1(d-2\ell, d-1; d-\ell; 1)$. Define the Pochhammer symbol as $(x)_a := \frac{\Gamma(x+a)}{\Gamma(x)}$. When $a \in \mathbb{N}$, ${}_2F_1(-a, b; c; 1) = \frac{(c-b)_a}{(c)_a}$ [63–67]. We therefore find that

$${}_2F_1(d-2\ell, d-1; d-\ell; 1) = \frac{(1-\ell)_{2\ell-d}}{(d-\ell)_{2\ell-d}} \quad (\text{C14})$$

$$= \frac{\Gamma(d-\ell)\Gamma(1-d+\ell)}{\Gamma(\ell)\Gamma(1-\ell)} \quad (\text{C15})$$

$$= \frac{\Gamma(d-\ell)\Gamma(1-\ell)(-d+\ell)(-d+\ell-1)\dots(1-\ell)}{\Gamma(\ell)\Gamma(1-\ell)} \quad (\text{C16})$$

$$= (-1)^{1+(1-\ell)-(-d+\ell)} \frac{\Gamma(d-\ell)(d-\ell)(d-\ell+1)\dots(\ell-1)}{\Gamma(\ell)} \quad (\text{C17})$$

$$= (-1)^d \frac{\Gamma(d-\ell)(\ell-1)!}{\Gamma(\ell)(d-\ell-1)!} \quad (\text{C18})$$

$$= (-1)^d, \quad (\text{C19})$$

hence proving condition 4. Conditions 1, 2, and 3 are similar. \square

Plugging in this result, we find that

$$f_\ell(r) := \frac{1}{n} \mathbb{E}_U \text{Tr } W^\ell = 2 \binom{2\ell-1}{\ell-1} \sum_{i=0}^{\ell-1} (-1)^i r^{i+\ell+1} \binom{\ell}{i} \frac{\ell-i}{(\ell+i)(\ell+i+1)} \quad (\text{C20})$$

$$= r^{\ell+1} C_\ell {}_2F_1(1-\ell, \ell; \ell+2; r), \quad (\text{C21})$$

where ${}_2F_1$ is the hypergeometric function and $C_\ell := \frac{1}{\ell+1} \binom{2\ell}{\ell}$ is the ℓ^{th} Catalan number. The simplification from the first to second line follows from the definition of the hypergeometric function given in Eq. (C13). Plugging this exact formula for $f_\ell(r) = \lim_{n \rightarrow \infty} \mathbb{E}_U \frac{1}{n} \text{Tr } W^\ell$ into Eq. (A32), and swapping sums, we arrive precisely at Theorem 2. We have therefore completed the proof.

To get a sense for f_ℓ , we list a few here.

$$\begin{aligned} f_1(r) &= r^2 \\ f_2(r) &= -r^4 + 2r^3 \\ f_3(r) &= 2r^6 - 6r^5 + 5r^4 \\ f_4(r) &= -5r^8 + 20r^7 - 28r^6 + 14r^5 \end{aligned}$$

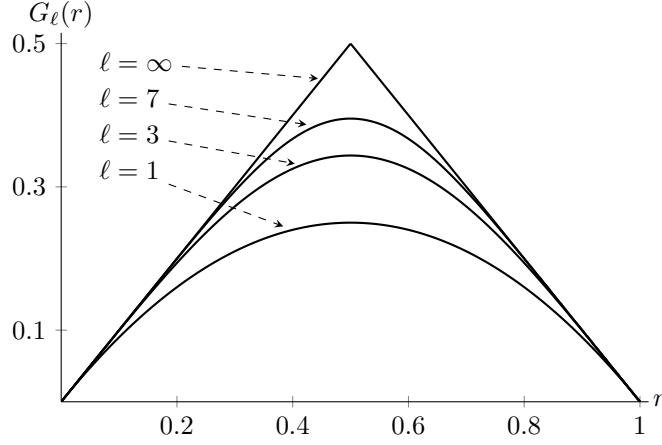


Figure C.1: The plots of $G_\ell(r) := r - f_\ell(r)$ for various values of ℓ . $f_\ell(r)$ is given in Eq. (C20). In our proof, we crucially used that $G_\ell(r)$ is symmetric under $r \mapsto 1 - r$.

$$f_5(r) = 14r^{10} - 70r^9 + 135r^8 - 120r^7 + 42r^6$$

$$f_6(r) = -42r^{12} + 252r^{11} - 616r^{10} + 770r^9 - 495r^8 + 132r^7$$

$$f_7(r) = 132r^{14} - 924r^{13} + 2730r^{12} - 4368r^{11} + 4004r^{10} - 2002r^9 + 429r^8$$

$$f_8(r) = -429r^{16} + 3432r^{15} - 11880r^{14} + 23100r^{13} - 27300r^{12} + 19656r^{11} - 8008r^{10} + 1430r^9.$$

Recall that we required $G_\ell(r) := r - f_\ell(r)$ to be symmetric under $r \mapsto 1 - r$. We show some of the resulting plots in Fig. C.1. Intuitively, one can understand $G_\ell(r)$ as being an approximation to the function $m(r) := \min(r, 1 - r)$; as ℓ increases, this approximation gets better and better.

More specifically, we can interpret the polynomials $G_\ell(r)$ in multiple ways. Once the form of $f_\ell(r)$ is fixed by Lemma C.1, we showed in Lemma C.2 that G_ℓ is uniquely determined by the symmetry requirement that $G_\ell(r) = G_\ell(1 - r)$. In the proof of Lemma C.2, we showed that G_ℓ is uniquely determined by conditions 1, 2, and 3. Each of these conditions has a simple interpretation. Condition 1, $1 = \sum_{j=\ell+1}^{2\ell} \alpha_j^{(\ell)}$, is the condition that $G_\ell(1) = 0$. Condition 2, $2 = \sum_{j=\ell+1}^{2\ell} j \alpha_j^{(\ell)}$, is the condition that the derivative $G'_\ell(1) = -1$. Condition 3, $0 = \sum_{j=\ell+1}^{2\ell} \binom{j}{d} \alpha_j^{(\ell)}$ for each $2 \leq d \leq \ell$, is the condition that the d^{th} derivative $G_\ell^{(d)}(1) = 0$. Hence, condition 1, 2, and 3 are imposing that $G_\ell^{(d)}(0) = m^{(d)}(0)$ and $G_\ell^{(d)}(1) = m^{(d)}(1)$ for all $0 \leq d \leq \ell$. Indeed, from the derived form of f_ℓ , we find that $G_\ell(r)$ is an approximation from below to $m(r)$, and it is an especially good approximation near the endpoints $r = 0$ and $r = 1$. We show some examples of this in Fig. C.1.

We will call $G_\ell(r)$ the ℓ^{th} order approximation to $m(r)$. From Eq. (A33), $\lim_{n \rightarrow \infty} \mathbb{E}_U \frac{1}{n} S_2(U) = \sum_{\ell=1}^{\infty} \frac{t^{2\ell}}{2\ell} G_\ell(r)$. Thus, $t = \tanh(2s)$ is weighting how relevant each approximation is. For small squeezing, most of the weight is concentrated on low-order approximations. The lowest order approximation is $G_1(r) = r(1 - r)$ resulting in a parabolic shaped Page curve. When the squeezing is very large, more and more weight is placed on high-order approximations so that the Page curve begins to resemble the triangle $G_\infty(r) = m(r)$. We see a manifestation of this interpretation as

$$\lim_{s \rightarrow 0} \lim_{n \rightarrow \infty} \mathbb{E}_{U \in \mathbb{U}(n)} \frac{1}{s^2 n} S_2(U) = 2r(1 - r), \quad (\text{C22})$$

$$\lim_{s \rightarrow \infty} \lim_{n \rightarrow \infty} \mathbb{E}_{U \in \mathbb{U}(n)} \frac{1}{s n} S_2(U) = 2 \min(r, 1 - r), \quad (\text{C23})$$

where the latter comes from the full expression in Theorem 2. Meanwhile, the maximal Rényi-2 entropy is $\max_U \frac{1}{n} S_2(U) = m(r) \log \cosh(2s)$ from Eq. (B1). As stated, near the endpoints $r = 0$ and $r = 1$, $G_\ell(r)$ is a very good approximation to $m(r)$. Thus, regardless of the squeezing strength, when the subsystem size $k = rn$ is small (or when its complement is small), the average entanglement is very close to maximal.

C.2 Maximum value — Proof of Corollary 3

In Appendix C.1, we derived the exact formula for the Rényi-2 Page curve as an infinite series. Here we will show that the series can be completely simplified when $r = 1/2$. Bailey's theorem says that [63–67]

$${}_2F_1(a, 1-a; c; 1/2) = \frac{\Gamma(c/2)\Gamma((1+c)/2)}{\Gamma((a+c)/2)\Gamma((1+c-a)/2)}. \quad (\text{C24})$$

Plugging this into the Page curve in Theorem 2 at $r = 1/2$ and simplifying with the duplication formula $\Gamma(\ell + 1/2) = \frac{\sqrt{\pi}(2\ell-1)!}{(\ell-1)!2^{2\ell-1}}$ [86],

$$\lim_{n \rightarrow \infty} \mathbb{E}_{U \in \mathbb{U}(n)} \frac{1}{n} S_2(U) = \frac{1}{2} \log \cosh(2s) - \frac{1}{4} \sum_{\ell=1}^{\infty} (\tanh^2(2s)/4)^{\ell} \frac{1}{\ell} \binom{2\ell}{\ell}. \quad (\text{C25})$$

We find that the second term is $\frac{1}{4} \int_0^{t/4} \frac{f(x)}{x} dx$, where $t = \tanh^2(2s)$ and $f(x)$ is the generating function of the central binomial coefficients $f(x) = \sum_{\ell=1}^{\infty} \binom{2\ell}{\ell} x^{\ell}$. Via the generalized binomial theorem, one finds the generating function evaluates to $f(x) = (1-4x)^{-1/2} - 1$ [84, A000984]. Performing the integral, we find the second term to be $\frac{1}{4} \log\left(\frac{4-4\sqrt{1-t}}{t+t\sqrt{1-t}}\right)$, which simplifies to $\frac{1}{2} \log \cosh(2s) - \log \cosh s$. Subtracting the second term from the first yields $\log \cosh s$ as desired.

Finally, using Eq. (B1) at $r = 1/2$, we find the Page correction to be $\frac{1}{2} \log \cosh(2s) - \log \cosh s$, which simplifies to $\frac{1}{2} \log(1 + \tanh^2 s)$.

C.3 Constant term — Proof of Proposition 4

In Appendix C.1, we found that when all the initial squeezers are equal to s , $\mathbb{E}_U S_2(U) = n\alpha(s, r) - \lambda(s, r) + o(1)$, and we explicitly computed $\alpha(s, r)$. In this section, we determine $\lambda(s, r)$ up to a set of constants and conjecture an explicit value of the constants.

In Appendix C.1, we found that, asymptotically in n , $\mathbb{E}_U \text{Tr } W^{\ell} = n f_{\ell}(r) + g_{\ell}(r) + o(1)$, where f_{ℓ} and g_{ℓ} are functions of r . From Eq. (A32), it follows that $\lambda(s, r) = \sum_{\ell=1}^{\infty} \frac{1}{2\ell} \tanh^{2\ell}(2s) g_{\ell}(r)$. Furthermore, in Lemma C.1, we found that f_{ℓ} takes the form $f_{\ell}(r) = \sum_{d=\ell+1}^{2\ell} \alpha_d^{(\ell)} r^d$. Indeed, the proof of Lemma C.1 applies almost identically to $g_{\ell}(r)$, except that because we are interested now in the constant term in $\mathbb{E}_U \text{Tr } W^{\ell}$ instead of the term linear in n , g_{ℓ} takes the form $g_{\ell}(r) = \sum_{d=\ell}^{2\ell} \beta_d^{(\ell)} r^d$. The extra term $\propto r^{\ell}$ in the polynomial g_{ℓ} does not occur in f_{ℓ} .

The Rényi-2 entropy must be symmetric under $r \mapsto 1-r$ at every order in s and n , since the full state on the n modes is pure. Therefore, $\lambda(s, r)$ must be symmetric under $r \mapsto 1-r$ at every order in $\tanh(2s)$, meaning that $g_{\ell}(r) = g_{\ell}(1-r)$. The following lemma will therefore be useful here, and we will also find use of it in Appendix D.

Lemma C.3. *Fix a polynomial $g_{\ell}(r) = \sum_{d=\ell}^{2\ell} \beta_d^{(\ell)} r^d$. If $g_{\ell}(r) = g_{\ell}(1-r)$, then $g_{\ell}(r) = \beta_{2\ell}^{(\ell)} (-1)^{\ell} (r(1-r))^{\ell}$.*

Proof. Simplifying $g_{\ell}(r) - g_{\ell}(1-r) = 0$ with the binomial theorem, we find

$$0 = \sum_{d=\ell}^{2\ell} \beta_d^{(\ell)} (r^d - (1-r)^d) \quad (\text{C26})$$

$$= - \sum_{d=0}^{\ell-1} (-r)^d \sum_{j=\ell}^{2\ell} \beta_j^{(\ell)} \binom{j}{d} + \sum_{d=\ell}^{2\ell} r^d \left(\beta_d^{(\ell)} - (-1)^d \sum_{j=d}^{2\ell} \beta_j^{(\ell)} \binom{j}{d} \right). \quad (\text{C27})$$

Equating all degrees of r , we find

1. For $0 \leq d \leq \ell-1$, $0 = \sum_{j=\ell}^{2\ell} \beta_j^{(\ell)} \binom{j}{d}$;
2. For $\ell \leq d \leq 2\ell$, $\beta_d^{(\ell)} = (-1)^d \sum_{j=d}^{2\ell} \beta_j^{(\ell)} \binom{j}{d}$.

Condition 1 is a system of ℓ linearly independent equations. To verify this, one must show that $\det D \neq 0$, where D is the $(\ell+1) \times (\ell+1)$ matrix with entries $D_{ij} = \binom{i+\ell-1}{j-1}$ for $1 \leq i \leq \ell+1$ and $1 \leq j \leq \ell$, and entries $D_{i,\ell+1} = \delta_{i,\ell+1}$. In other words, the rightmost column is all zeros except for the entry on the diagonal. Inserting this rightmost column is equivalent to adding one new equation to the linear system, where this new equation

simply fixes the value of one of the variables. Since the rightmost column is all zeros except for the $(\ell+1, \ell+1)$ entry, we use Laplace's expansion to find that $\det D = \det D'$, where D' is the $\ell \times \ell$ upper left submatrix of D . One can prove that $\det D' \neq 0$ in a similar way as shown in Ref. [85]. Alternatively, one can use Corollary 11 from Ref. [87], which shows that the matrix with entries $\binom{a_i}{j-1}$ has zero determinant if and only if there are indices $i \neq j$ such that $a_i = a_j$. For the matrix D' , $a_i = \ell + i - 1$, and therefore the determinant is nonzero.

In summary, we have a system of ℓ linearly independent equations for $\ell + 1$ variables. We can therefore uniquely express the solution by fixing one of the variables. Suppose we know the value of $\beta_{2\ell}^{(\ell)}$. Then one can verify that conditions 1 and 2 are satisfied by $\beta_{\ell+i}^{(\ell)} = (-1)^{\ell+i} \binom{\ell}{i} \beta_{2\ell}^{(\ell)}$, for $0 \leq i \leq \ell$. With the binomial theorem, this simply becomes $g_\ell(r) = \beta_{2\ell}^{(\ell)} (-1)^\ell (r(1-r))^\ell$.

When verifying that the two conditions are satisfied by $\beta_{\ell+i}^{(\ell)} = (-1)^{\ell+i} \binom{\ell}{i} \beta_{2\ell}^{(\ell)}$, one finds that the right hand side of both conditions can be simplified in terms of the hypergeometric function via Eq. (C13). For example, condition 2, written as $(-1)^d = \sum_{j=d}^{2\ell} \binom{j}{d} (\beta_j^{(\ell)} / \beta_d^{(\ell)})$, reduces to $(-1)^d = {}_2F_1(d-2\ell, d+1; d-\ell+1; 1)$. As in the proof of Lemma C.2, we find that

$${}_2F_1(d-2\ell, d+1; d-\ell+1; 1) = \frac{(-\ell)_{2\ell-d}}{(d-\ell+1)_{2\ell-d}} \quad (C28)$$

$$= \frac{\Gamma(d-\ell+1)\Gamma(\ell-d)}{\Gamma(-\ell)\Gamma(\ell+1)} \quad (C29)$$

$$= \frac{\Gamma(d-\ell+1)\Gamma(-\ell)(\ell-d-1)(\ell-d-2)\dots(-\ell)}{\Gamma(-\ell)\Gamma(\ell+1)} \quad (C30)$$

$$= (-1)^{1+(\ell-d+1)-(-\ell)} \frac{\Gamma(d-\ell+1)(d+1-\ell)(d+2-\ell)\dots(\ell)}{\Gamma(\ell+1)} \quad (C31)$$

$$= (-1)^d \frac{\Gamma(d-\ell+1)\ell!}{\Gamma(\ell+1)(d-\ell)!} \quad (C32)$$

$$= (-1)^d, \quad (C33)$$

which proves condition 2. The proof of condition 1 is similar. \square

For convenience, we define $a^{(\ell)} := \beta_{2\ell}^{(\ell)}$. It follows from Lemma C.3 that

$$\lambda(s, r) = \sum_{\ell=1}^{\infty} \frac{(-1)^\ell}{2\ell} a^{(\ell)} \tanh^{2\ell}(2s) (r(1-r))^\ell. \quad (C34)$$

Note that if $a^{(\ell)} = (-1)^\ell 4^{\ell-1}$, then the Taylor series of \log implies $\lambda(s, r) = -\frac{1}{8} \log(1 - 4r(1-r) \tanh^2(2s))$, which is precisely Proposition 4. Hence, the only thing left to prove is the following lemma.

Lemma C.4. *For all $\ell \in \mathbb{N}$, $a^{(\ell)} = (-1)^\ell 4^{\ell-1}$.*

To make progress on Lemma C.4, we write the formula for $a^{(\ell)}$. Recall that $\beta_{2\ell}^{(\ell)}$ is the coefficient in front of the $r^{2\ell}$ term in $g_\ell(r)$. Looking at Eq. (C2), we see that the only way to get an $r^{2\ell}$ term (i.e. a $k^{2\ell}$ term) is if σ is the permutation $\sigma(1) = 2\ell, \sigma(2) = 1, \sigma(3) = 2, \dots, \sigma(2\ell) = 2\ell - 1$. Therefore, we are interested in the constant term (the term asymptotically independent of n) in the following term of Eq. (C2):

$$\begin{aligned} & n^{2\ell} \sum_{j_1, \dots, j_{2\ell}=1}^n \sum_{j'_1, \dots, j'_{2\ell}=1}^n \sum_{\tau \in S_{2\ell}} \text{Wg}(\sigma\tau^{-1}, n) \\ & \times \delta_{j_1, j_2} \delta_{j'_1, j'_2} \dots \delta_{j_{2\ell-1}, j_{2\ell}} \delta_{j'_{2\ell-1}, j'_{2\ell}} \\ & \times \delta_{j_1, j'_{\tau(1)}} \dots \delta_{j_{2\ell}, j'_{\tau(2\ell)}}, \end{aligned} \quad (C35)$$

which simplifies to

$$\begin{aligned} & = n^{2\ell} \sum_{j_1, \dots, j_{2\ell}=1}^n \sum_{j'_1, \dots, j'_{2\ell}=1}^n \sum_{\tau \in S_{2\ell}} \text{Wg}(\tau^{-1}, n) \\ & \times \delta_{j_1, j_2} \delta_{j'_1, j'_2} \dots \delta_{j_{2\ell-1}, j_{2\ell}} \delta_{j'_{2\ell-1}, j'_{2\ell}} \\ & \times \delta_{j_1, j'_{\tau(2\ell)}} \delta_{j_2, j'_{\tau(1)}} \delta_{j_3, j'_{\tau(2)}} \dots \delta_{j_{2\ell}, j'_{\tau(2\ell-1)}} \end{aligned} \quad (C36)$$

ℓ	$a^{(\ell)}$	
1	$-C_1$	$= -1$
2	$-4C_0^2C_1 + 4C_0C_2$	$= 4$
3	$-9C_0^4C_1 + 15C_0^2C_1^2 - C_1^3 + 18C_0^3C_2 - 6C_0C_1C_2 - 9C_0^2C_3$	$= -16$
4	$-16C_0^6C_1 + 80C_0^4C_1^2 - 40C_0^2C_1^3 + 48C_0^5C_2 - 112C_0^3C_1C_2 + 16C_0^2C_2 + 8C_0^2C_2^2 - 48C_0^4C_3 + 24C_0^2C_1C_3 + 16C_0^3C_4$	$= 64$
5	$-25C_0^8C_1 + 250C_0^6C_1^2 - 380C_0^4C_1^3 + 80C_0^2C_1^4 - C_1^5 + 100C_0^7C_2 - 600C_0^5C_1C_2 + 440C_0^3C_1^2C_2 - 20C_0C_1^3C_2 + 130C_0^4C_2^2 - 50C_0^2C_1C_2^2 - 150C_0^6C_3 + 320C_0^4C_1C_3 - 60C_0^2C_1^2C_3 - 40C_0^3C_2C_3 + 100C_0^5C_4 - 60C_0^3C_1C_4 - 25C_0^4C_5$	$= -256$

Table C.1: A table showing the first five values of $a^{(\ell)}$ from Eq. (C45), which matches Lemma C.4.

$$= n^{2\ell} \sum_{j_2, j_4, \dots, j_{2\ell}=1}^n \sum_{j'_1, \dots, j'_{2\ell}=1}^n \sum_{\tau \in S_{2\ell}} \text{Wg}(\tau, n) \times \delta_{j'_1, j'_2} \dots \delta_{j'_{2\ell-1}, j'_{2\ell}} \times \delta_{j_2, j'_{\tau(2\ell)}} \delta_{j_2, j'_{\tau(1)}} \delta_{j_4, j'_{\tau(2)}} \delta_{j_4, j'_{\tau(3)}} \dots \delta_{j_{2\ell}, j'_{\tau(2\ell-2)}} \delta_{j_{2\ell}, j'_{\tau(2\ell-1)}} \quad (\text{C37})$$

$$= n^{2\ell} \sum_{j'_1, \dots, j'_{2\ell}=1}^n \sum_{\tau \in S_{2\ell}} \text{Wg}(\tau, n) \times \delta_{j'_1, j'_2} \dots \delta_{j'_{2\ell-1}, j'_{2\ell}} \times \delta_{j'_{\tau(2\ell)}, j'_{\tau(1)}} \delta_{j'_{\tau(2)}, j'_{\tau(3)}} \dots \delta_{j'_{\tau(2\ell-2)}, j'_{\tau(2\ell-1)}}. \quad (\text{C38})$$

Note we used that $\text{Wg}(\tau, n) = \text{Wg}(\tau^{-1}, n)$ from Eq. (A7). Let $\#(\tau)$ be the number of cycles in the disjoint cycle decomposition of τ and let $\{c_i^{(\tau)} \mid i \in \{1, \dots, \#(\tau)\}\}$ be the cycle decomposition. Then, from Eq. (A7), asymptotically,

$$\text{Wg}(\tau, n) = \frac{1}{n^{2\ell+|\tau|}} \prod_{i=1}^{\#(\tau)} (-1)^{|c_i^{(\tau)}|-1} C_{|c_i^{(\tau)}|-1}, \quad (\text{C39})$$

where recall C_m is the m^{th} Catalan number. Therefore, $a^{(\ell)}$ is the constant term independent of n in

$$\sum_{j_1, \dots, j_{2\ell}=1}^n \sum_{\tau \in S_{2\ell}} \frac{1}{n^{|\tau|}} \prod_{i=1}^{\#(\tau)} (-1)^{|c_i^{(\tau)}|-1} C_{|c_i^{(\tau)}|-1} \times \delta_{j_1, j_2} \dots \delta_{j_{2\ell-1}, j_{2\ell}} \times \delta_{j_{\tau(2\ell)}, j_{\tau(1)}} \delta_{j_{\tau(2)}, j_{\tau(3)}} \dots \delta_{j_{\tau(2\ell-2)}, j_{\tau(2\ell-1)}}. \quad (\text{C40})$$

Define $\xi: S_{2\ell} \rightarrow \mathbb{N}$ such that

$$n^{\xi(\tau)} = \sum_{j_1, \dots, j_{2\ell}=1}^n \delta_{j_1, j_2} \delta_{j_3, j_4} \dots \delta_{j_{2\ell-1}, j_{2\ell}} \times \delta_{j_{\tau(2\ell)}, j_{\tau(1)}} \delta_{j_{\tau(2)}, j_{\tau(3)}} \dots \delta_{j_{\tau(2\ell-2)}, j_{\tau(2\ell-1)}} \quad (\text{C41})$$

$$= \sum_{j_1, \dots, j_{\ell}=1}^n \delta_{j_{\lceil \tau(2\ell)/2 \rceil}, j_{\lceil \tau(1)/2 \rceil}} \delta_{j_{\lceil \tau(2)/2 \rceil}, j_{\lceil \tau(3)/2 \rceil}} \dots \delta_{j_{\lceil \tau(2\ell-2)/2 \rceil}, j_{\lceil \tau(2\ell-1)/2 \rceil}}. \quad (\text{C42})$$

Note that the definition of ξ is independent of the value of n , and so we can equivalently define ξ as

$$\xi(\tau) = \log_2 \sum_{j_1, \dots, j_{\ell}=1}^2 \delta_{j_{\lceil \tau(2\ell)/2 \rceil}, j_{\lceil \tau(1)/2 \rceil}} \delta_{j_{\lceil \tau(2)/2 \rceil}, j_{\lceil \tau(3)/2 \rceil}} \dots \delta_{j_{\lceil \tau(2\ell-2)/2 \rceil}, j_{\lceil \tau(2\ell-1)/2 \rceil}}. \quad (\text{C43})$$

To get the constant term, we need $\frac{n^{\xi(\tau)}}{n^{|\tau|}} = 1$. Therefore,

$$a^{(\ell)} = \sum_{\substack{\tau \in S_{2\ell} \text{ s.t.} \\ \xi(\tau) = |\tau|}} \prod_{i=1}^{\#(\tau)} (-1)^{|c_i^{(\tau)}|-1} C_{|c_i^{(\tau)}|-1}. \quad (\text{C44})$$

Finally, since the sum of the lengths of the cycles of a permutation $\tau \in S_{2\ell}$ is always 2ℓ , we find

$$a^{(\ell)} = \sum_{\substack{\tau \in S_{2\ell} \text{ s.t.} \\ \xi(\tau) = |\tau|}} (-1)^{\#(\tau)} \prod_{i=1}^{\#(\tau)} C_{|c_i^{(\tau)}|-1}. \quad (\text{C45})$$

From this equation, we can exactly compute $a^{(\ell)}$ on a computer for small values of ℓ . Table C.1 shows the first five of these, which all match Lemma C.4 saying that $a^{(\ell)} = (-1)^\ell 4^{\ell-1}$. Note that if we change the condition of $\xi(\tau) = |\tau|$ in Eq. (C45) to $\xi(\tau) = |\tau| + 1$, then this gives us the term that is linear in n and hence is the equation for $\alpha_{2\ell}^{(\ell)}$ from Lemma C.2, which is $\frac{(-1)^{\ell+1}}{2\ell-1} \binom{2\ell-1}{\ell-1}$.

To complete the proof of Lemma C.4, we need to evaluate Eq. (C45) for all $\ell \in \mathbb{N}$. This is done in Ref. [72] using objects called *breakpoint graphs* that arise in the study of gene orders in bioinformatics [71]. Roughly, ξ has an interpretation in terms of cycles of breakpoint graphs.

D Variance of the Rényi-2 entropy — Proof of Theorem 6

In this section, we shift our attention away from $\mathbb{E}_U S_2(U)$ and instead to $\text{Var}_U S_2(U) = \mathbb{E}_U S_2(U)^2 - (\mathbb{E}_U S_2(U))^2$, and we will prove Theorem 6. We are again interested in the case where all the initial squeezers are equal to s . Using Eq. (A32), this becomes

$$\text{Var}_{U \in \mathbb{U}(n)} S_2(U) = \sum_{\ell, \ell'=1}^{\infty} \frac{1}{(2\ell)(2\ell')} \tanh^{2\ell+2\ell'}(2s) \left(\mathbb{E}_U (\text{Tr } W^\ell)(\text{Tr } W^{\ell'}) - \left(\mathbb{E}_U \text{Tr } W^\ell \right) \left(\mathbb{E}_U \text{Tr } W^{\ell'} \right) \right) \quad (\text{D1})$$

$$= \sum_{d=2}^{\infty} \tanh^{2d}(2s) \sum_{\ell=1}^{d-1} \frac{1}{4\ell(d-\ell)} \left(\mathbb{E}_U (\text{Tr } W^\ell)(\text{Tr } W^{d-\ell}) - \left(\mathbb{E}_U \text{Tr } W^\ell \right) \left(\mathbb{E}_U \text{Tr } W^{d-\ell} \right) \right). \quad (\text{D2})$$

As a direct consequence of Lemma C.1, we find that asymptotically

$$\left(\mathbb{E}_U \text{Tr } W^\ell \right) \left(\mathbb{E}_U \text{Tr } W^{\ell'} \right) = n^2 p_{\ell, \ell'}(r) + n q_{\ell, \ell'}(r) + t_{\ell, \ell'}(r) + o(1), \quad (\text{D3})$$

where $p_{\ell, \ell'}$ is a polynomial of degrees $\ell + \ell' + 2$ through $2\ell + 2\ell'$ in r , $q_{\ell, \ell'}$ is a polynomial of degrees $\ell + \ell' + 1$ through $2\ell + 2\ell'$ in r , and $t_{\ell, \ell'}$ is a polynomial of degrees $\ell + \ell'$ through $2\ell + 2\ell'$ in r .

Furthermore, we find an analogous result for $\mathbb{E}_U (\text{Tr } W^\ell)(\text{Tr } W^{\ell'})$. Let $L := 2\ell + 2\ell'$. Using Eq. (A6), we find that

$$\begin{aligned} \mathbb{E}_{U \in \mathbb{U}(n)} (\text{Tr } W^\ell)(\text{Tr } W^{\ell'}) &= \sum_{i_1, \dots, i_L=1}^k \sum_{i'_1, \dots, i'_L=1}^k \sum_{j_1, \dots, j_L=1}^n \sum_{j'_1, \dots, j'_L=1}^n \sum_{\sigma, \tau \in S_L} \text{Wg}(\sigma \tau^{-1}, n) \\ &\quad \times \delta_{i'_{2\ell}, i_1} \delta_{i'_1, i_2} \delta_{i'_2, i_3} \dots \delta_{i'_{2\ell-1}, i_{2\ell}} \\ &\quad \times \delta_{i'_L, i_{2\ell+1}} \delta_{i'_{2\ell+1}, i_{2\ell+2}} \delta_{i'_{2\ell+2}, i_{2\ell+3}} \dots \delta_{i'_{L-1}, i_L} \\ &\quad \times \delta_{j_1, j_2} \delta_{j'_1, j'_2} \dots \delta_{j_{2\ell-1}, j_{2\ell}} \delta_{j'_{2\ell-1}, j'_{2\ell}} \\ &\quad \times \delta_{j_{2\ell+1}, j_{2\ell+2}} \delta_{j'_{2\ell+1}, j'_{2\ell+2}} \dots \delta_{j_{L-1}, j_L} \delta_{j'_{L-1}, j'_L} \\ &\quad \times \delta_{i_1, i'_{\sigma(1)}} \dots \delta_{i_L, i'_{\sigma(L)}} \\ &\quad \times \delta_{j_1, j'_{\tau(1)}} \dots \delta_{j_L, j'_{\tau(L)}}, \end{aligned} \quad (\text{D4})$$

and the asymptotic form of Wg function is given in Eq. (A7). In a similar proof to Lemma C.1 but with Eq. (D4) instead of Eq. (C2), we analogously find that asymptotically

$$\mathbb{E}_{U \in \mathbb{U}(n)} (\text{Tr } W^\ell)(\text{Tr } W^{\ell'}) = n^2 P_{\ell, \ell'}(r) + n Q_{\ell, \ell'}(r) + T_{\ell, \ell'}(r) + o(1), \quad (\text{D5})$$

where $P_{\ell, \ell'}$ is a polynomial of degrees $\ell + \ell' + 2$ through $2\ell + 2\ell'$ in r , $Q_{\ell, \ell'}$ is a polynomial of degrees $\ell + \ell' + 1$ through $2\ell + 2\ell'$ in r , and $T_{\ell, \ell'}$ is a polynomial of degrees $\ell + \ell'$ through $2\ell + 2\ell'$ in r . For completeness, we prove this result for the $P_{\ell, \ell'}$ term in the following lemma. The proofs for the $Q_{\ell, \ell'}$ and $T_{\ell, \ell'}$ terms follow from trivially tweaking the final part of the proof.

Lemma D.1. Fix positive integers ℓ and ℓ' . There exist coefficients $\beta_d^{(\ell, \ell')}$ for $d \in \{\ell + \ell' + 2, \dots, 2\ell + 2\ell'\}$ such that

$$P_{\ell, \ell'}(r) := \lim_{n \rightarrow \infty} \frac{1}{n^2} \mathbb{E}_U (\text{Tr } W^\ell) (\text{Tr } W^{\ell'}) = \sum_{d=\ell+\ell'+2}^{2\ell+2\ell'} \beta_d^{(\ell, \ell')} r^d. \quad (\text{D6})$$

Proof. Much of the details of this proof are the same as in the proof of Lemma C.1. We will use the asymptotic form of the Wg function, which is written in Eq. (A7). The proof will proceed as follows. First, we will prove that $(\text{Tr } W^\ell)(\text{Tr } W^{\ell'})$ contains a term proportional to n^2 and no terms proportional to n^a for any $a > 2$. Therefore, $P_{\ell, \ell'}(r)$ is indeed independent of n . Next, we will prove that $P_{\ell, \ell'}(r)$ has no terms r^a for $a > L$ and $a \leq \ell + \ell' + 1$. Throughout this proof, we interpret the delta functions in Eq. (D4) as constraints on the summations. Different permutations on the indices result in a different number of constraints and hence terms with different powers of n and k .

Recall that Π can only decrease the trace. Getting rid of the Π in W and using the cyclic nature of the trace, we see that $(\text{Tr } W^\ell)(\text{Tr } W^{\ell'}) \leq n^2$. Furthermore, consider Eq. (D4) with $\sigma = \tau$ defined by $\sigma(1) = 2\ell, \sigma(2) = 1, \dots, \sigma(2\ell) = 2\ell - 1$ and $\sigma(2\ell + 1) = L, \sigma(2\ell + 2) = 2\ell + 1, \dots, \sigma(L) = L - 1$. Then, $\sigma\tau^{-1}$ is the identity, and so $\text{Wg}(\sigma\tau^{-1}, n)$ contributes a factor of n^{-L} . With this σ , the sum over the i and i' yields a factor of k^L . Finally, the sum over j yields

$$\delta_{j'_1, j'_2} \delta_{j'_3, j'_4} \dots \delta_{j'_{L-1}, j'_L} \delta_{j'_1, j'_2} \delta_{j'_2, j'_1} \dots \delta_{j'_{2\ell}, j'_{2\ell-1}} \delta_{j'_{2\ell+1}, j'_L} \delta_{j'_{2\ell+2}, j'_{2\ell+1}} \dots \delta_{j'_L, j'_{L-1}}. \quad (\text{D7})$$

Then summing over j' , we get two factors of n . Hence, the term with the specific permutation described above yields a term of the form $n^2 k^L n^{-L} = n^2 r^L$.

We have shown that there is a term proportional to n^2 and that there are no terms proportional to n^a for $a > 2$. Since we are working asymptotically in n , we can therefore ignore all terms proportional to n^a for every $a < 2$. This proves that $\lim_{n \rightarrow \infty} \frac{1}{n^2} (\text{Tr } W^\ell)(\text{Tr } W^{\ell'})$ is independent of n and only depends on r , which justifies the definition of the function $P_{\ell, \ell'}(r)$. The only thing left to show is that $P_{\ell, \ell'}(r)$ has only terms $r^{\ell+\ell'+2}$ through $r^{2\ell+2\ell'} = r^L$. So we only need to show that there are no terms r^a for $a > L$ and $a \leq \ell + \ell' + 1$. We begin with the former.

To look at powers of r , it is sufficient to look at powers of k . We therefore restrict our attention to the sum over i and i' in Eq. (D4). In order to get the highest power of k , we require the least constraints on i and i' (i.e. the least distinct Kronecker deltas). We therefore require σ to be the permutation so that

$$\delta_{i'_{2\ell}, i_1} \delta_{i'_1, i_2} \delta_{i'_2, i_3} \dots \delta_{i'_{2\ell-1}, i_{2\ell}} \delta_{i'_L, i_{2\ell+1}} \delta_{i'_{2\ell+1}, i_{2\ell+2}} \delta_{i'_{2\ell+2}, i_{2\ell+3}} \dots \delta_{i'_{L-1}, i_L} = \delta_{i_1, i'_{\sigma(1)}} \dots \delta_{i_L, i'_{\sigma(L)}}. \quad (\text{D8})$$

This permutation σ is exactly the σ described above that gave the term proportional to $n^2 r^L$. Hence, L is the highest power of r that can be achieved.

Next we need to show that $\ell + \ell' + 2$ is the lowest power of k that can be achieved. The sum over j and j' can give at most a factor of $n^{\ell+\ell'}$. This is because the first line of delta functions, $\delta_{j_1, j_2} \delta_{j'_1, j'_2} \dots \delta_{j_{L-1}, j_L} \delta_{j'_{L-1}, j'_L}$, reduces the sum over the L indices j and the L indices j' down to just a sum over $\ell + \ell'$ indices j and $\ell + \ell'$ indices j' . The second line of delta functions, $\delta_{j_1, j'_{\tau(1)}} \dots \delta_{j_L, j'_{\tau(L)}}$, cannot be made equivalent to the first line by any choice of τ ; in fact, the second line imposes all new constraints. Therefore, this line further reduces the sum over the $\ell + \ell'$ indices j and the $\ell + \ell'$ indices j' to just a sum over the $\ell + \ell'$ indices j (or the $\ell + \ell'$ indices j' , but not both). Hence the highest power of n that we can get from the summations over the j and j' is $n^{\ell+\ell'}$. Putting this together with the fact that asymptotically $\text{Wg}(\pi, n)$ is at most $n^{-L} = n^{-2\ell-2\ell'}$, we find that any term coming from Eq. (D4) is at most $n^{-\ell-\ell'} \times (\text{dependence on } k)$. Therefore, any powers of k that are less than $\ell + \ell' + 2$ can be ignored; if the sum over i and i' yields a term that is k^a for some $a \leq \ell + \ell' + 1$, then that term will scale linearly or less with n . But from above, we already have terms that are proportional to n^2 , and so terms that are linear or less in n can be ignored. \square

Therefore, from Eq. (D2), we find that asymptotically

$$\text{Var}_{U \in \mathbb{U}(n)} S_2(U) = \sum_{d=2}^{\infty} \tanh^{2d}(2s) (n^2 p_d(r) + n q_d(r) + t_d(r)) + o(1), \quad (\text{D9})$$

where p_d is a polynomial of degrees $d + 2$ through $2d$ in r , q_d is a polynomial of degrees $d + 1$ through $2d$ in r , and t_d is a polynomial of degrees d through $2d$ in r . The variance must be symmetric under $r \mapsto 1 - r$ at every order in s and n . Therefore each p_d , q_d , and t_d must themselves be symmetric under $r \mapsto 1 - r$. It then immediately follows from Lemma C.3 that $t_d(r) = \omega^{(d)} (r(1 - r))^d$ for some constants $\omega^{(d)}$. In the following lemma, we show that p_d and q_d must be the zero polynomial.

Lemma D.2. Let d be a positive integer, and $f(r) = \sum_{j=d+1}^{2d} \gamma_j r^j$. If $f(r) = f(1-r)$, then $f(r) = 0$.

Proof. Expanding with the binomial theorem, we find

$$0 = f(r) - f(1-r) \quad (\text{D10})$$

$$= \sum_{j=d+1}^{2d} \gamma_j (r^j - (1-r)^j) \quad (\text{D11})$$

$$= \sum_{j=d+1}^{2d} \gamma_j r^j - \sum_{j=d+1}^{2d} \gamma_j \sum_{i=0}^j (-r)^i \binom{j}{i} \quad (\text{D12})$$

$$= \sum_{j=d+1}^{2d} \gamma_j r^j - \sum_{i=0}^{2d} (-r)^i \sum_{j=\max(i, d+1)}^{2d} \gamma_j \binom{j}{i} \quad (\text{D13})$$

$$= - \sum_{j=0}^d (-r)^j \sum_{i=d+1}^{2d} \gamma_i \binom{i}{j} + \sum_{j=d+1}^{2d} r^j \left(\gamma_j - (-1)^j \sum_{i=j}^{2d} \gamma_i \binom{i}{j} \right). \quad (\text{D14})$$

This must be true to all orders in r , and hence we equate each degree of r to zero. This gives us the conditions:

1. For $0 \leq j \leq d$, $0 = \sum_{i=d+1}^{2d} \gamma_i \binom{i}{j}$;
2. For $d+1 \leq j \leq 2d$, $\gamma_j = (-1)^j \sum_{i=d+1}^{2d} \gamma_i \binom{i}{j}$.

Choosing any d of the equations from the first condition gives a linear system of d linearly independent equations. To verify this, one must show that $\det C \neq 0$, where C is the $d \times d$ matrix with entries $C_{ij} = \binom{d+i}{j}$. This was shown in the proof of Lemma C.2.

Therefore, we have d linearly independent equations for d variables. Hence, if there is a solution, then there is one *unique* solution. We easily see that $\gamma_j = 0$ is a solution, and therefore it is the solution. This gives $f(r) = 0$, completing the proof. \square

Since p_d and q_d are zero, we have therefore found that

$$\lim_{n \rightarrow \infty} \operatorname{Var}_{U \in \mathbb{U}(n)} S_2(U) = \sum_{d=2}^{\infty} \omega^{(d)} \tanh^{2d}(2s) (r(1-r))^d, \quad (\text{D15})$$

proving the first part of Theorem 6. Equating this equation to Eq. (D2), we find that

$$\omega^{(d)} = \lim_{n \rightarrow \infty} (r(1-r))^{-d} \sum_{\ell=1}^{d-1} \frac{1}{4\ell(d-\ell)} \left(\mathbb{E}_U (\operatorname{Tr} W^\ell) (\operatorname{Tr} W^{d-\ell}) - \left(\mathbb{E}_U \operatorname{Tr} W^\ell \right) \left(\mathbb{E}_U \operatorname{Tr} W^{d-\ell} \right) \right) \quad (\text{D16})$$

$$= \lim_{n \rightarrow \infty} (r(1-r))^{-d} \sum_{\ell=1}^{d-1} \frac{1}{4\ell(d-\ell)} \operatorname{Cov}_{U \in \mathbb{U}(n)} (\operatorname{Tr} W^\ell, \operatorname{Tr} W^{d-\ell}), \quad (\text{D17})$$

where Cov is the covariance, and we know that $\omega^{(d)}$ is independent of r and n . From our expressions for $\operatorname{Tr} W^\ell$ in terms of the Weingarten calculus, it follows that $\omega^{(d)} \in \mathbb{Q}$. Recall that W has two factors of U and two factors of \bar{U} . Hence, we can exactly compute $\omega^{(2)}$ by integrating over fourth moments of the Haar measure on the unitary group. To do this, we use the Mathematica package RTNI that symbolically computes expressions over the Haar measure [88]. This Mathematica package precomputes the symbolic expressions for $\operatorname{Wg}(\sigma, n)$ for $\sigma \in S_4$. From this, $\omega^{(2)}$ is a sum over powers of $\operatorname{Tr} \Pi = k$ with coefficients depending on n . One can then simplify this expression and take the limit $n \rightarrow \infty$ to find that $\omega^{(2)} = 1/2$. Our Mathematica code for this calculation is provided on GitHub [68].

References

- [1] Don N. Page. “Average entropy of a subsystem”. *Physical Review Letters* **71**, 1291–1294 (1993).
- [2] S. K. Foong and S. Kanno. “Proof of Page’s conjecture on the average entropy of a subsystem”. *Physical Review Letters* **72**, 1148–1151 (1994).

- [3] Jorge Sánchez-Ruiz. “Simple proof of Page’s conjecture on the average entropy of a subsystem”. *Physical Review E* **52**, 5653–5655 (1995).
- [4] Siddhartha Sen. “Average Entropy of a Quantum Subsystem”. *Physical Review Letters* **77**, 1–3 (1996).
- [5] Don N. Page. “Information in black hole radiation”. *Physical Review Letters* **71**, 3743–3746 (1993).
- [6] Patrick Hayden and John Preskill. “Black holes as mirrors: quantum information in random subsystems”. *Journal of High Energy Physics* **2007**, 120–120 (2007).
- [7] Eugenio Bianchi, Tommaso De Lorenzo, and Matteo Smerlak. “Entanglement entropy production in gravitational collapse: covariant regularization and solvable models”. *Journal of High Energy Physics* **2015**, 180 (2015).
- [8] Patrick Hayden, Debbie W. Leung, and Andreas Winter. “Aspects of generic entanglement”. *Communications in Mathematical Physics* **265**, 95–117 (2006).
- [9] Pavan Hosur, Xiao-Liang Qi, Daniel A. Roberts, and Beni Yoshida. “Chaos in quantum channels”. *Journal of High Energy Physics* **2016**, 4 (2016).
- [10] Hiroyuki Fujita, Yuya O. Nakagawa, Sho Sugiura, and Masataka Watanabe. “Page curves for general interacting systems”. *Journal of High Energy Physics* **2018**, 112 (2018).
- [11] Tsung-Cheng Lu and Tarun Grover. “Renyi entropy of chaotic eigenstates”. *Physical Review E* **99**, 032111 (2019).
- [12] Yuya O. Nakagawa, Masataka Watanabe, Hiroyuki Fujita, and Sho Sugiura. “Universality in volume-law entanglement of scrambled pure quantum states”. *Nature Communications* **9**, 1635 (2018).
- [13] Lev Vidmar and Marcos Rigol. “Entanglement entropy of eigenstates of quantum chaotic hamiltonians”. *Physical Review Letters* **119**, 220603 (2017).
- [14] Lev Vidmar, Lucas Hackl, Eugenio Bianchi, and Marcos Rigol. “Entanglement entropy of eigenstates of quadratic fermionic hamiltonians”. *Physical Review Letters* **119**, 020601 (2017).
- [15] Lucas Hackl, Lev Vidmar, Marcos Rigol, and Eugenio Bianchi. “Average eigenstate entanglement entropy of the XY chain in a transverse field and its universality for translationally invariant quadratic fermionic models”. *Physical Review B* **99**, 075123 (2019).
- [16] Sheldon Goldstein, Joel L. Lebowitz, Roderich Tumulka, and Nino Zanghi. “Canonical typicality”. *Physical Review Letters* **96**, 050403 (2006).
- [17] Luca D’Alessio, Yariv Kafri, Anatoli Polkovnikov, and Marcos Rigol. “From quantum chaos and eigenstate thermalization to statistical mechanics and thermodynamics”. *Advances in Physics* **65**, 239–362 (2016).
- [18] Patrick Hayden, Debbie Leung, Peter W. Shor, and Andreas Winter. “Randomizing Quantum States: Constructions and Applications”. *Communications in Mathematical Physics* **250**, 371–391 (2004).
- [19] Benoit Collins, Carlos E. Gonzalez-Guillen, and David Pérez-Garcia. “Matrix Product States, Random Matrix Theory and the Principle of Maximum Entropy”. *Communications in Mathematical Physics* **320**, 663–677 (2013).
- [20] M. B. Hastings. “Random MERA States and the Tightness of the Brandao-Horodecki Entropy Bound”. *arXiv.1505.06468* (2015).
- [21] Silvano Garnerone, Thiago R. de Oliveira, and Paolo Zanardi. “Typicality in random matrix product states”. *Physical Review A* **81**, 032336 (2010).
- [22] Sandu Popescu, Anthony J. Short, and Andreas Winter. “Entanglement and the foundations of statistical mechanics”. *Nature Physics* **2**, 754–758 (2006).
- [23] D. Gross, S. T. Flammia, and J. Eisert. “Most Quantum States Are Too Entangled To Be Useful As Computational Resources”. *Physical Review Letters* **102**, 190501 (2009).
- [24] Michael J. Bremner, Caterina Mora, and Andreas Winter. “Are Random Pure States Useful for Quantum Computation?”. *Physical Review Letters* **102**, 190502 (2009).
- [25] Eugenio Bianchi and Pietro Donà. “Typical entanglement entropy in the presence of a center: Page curve and its variance”. *Physical Review D* **100**, 105010 (2019).
- [26] Eugenio Bianchi, Lucas Hackl, and Mario Kieburg. “Page curve for fermionic Gaussian states”. *Physical Review B* **103**, L241118 (2021).
- [27] Oscar CO Dahlsten, Cosmo Lupo, Stefano Mancini, and Alessio Serafini. “Entanglement typicality”. *Journal of Physics A: Mathematical and Theoretical* **47**, 363001 (2014).
- [28] Michael A. Nielsen and Isaac L. Chuang. “Quantum computation and quantum information”. Cambridge University Press (2010). 10th anniversary edition.
- [29] Ingemar Bengtsson and Karol Życzkowski. “Geometry of quantum states: An introduction to quantum entanglement”. Cambridge University Press (2008).
- [30] Luigi Amico, Rosario Fazio, Andreas Osterloh, and Vlatko Vedral. “Entanglement in many-body systems”. *Reviews of Modern Physics* **80**, 517–576 (2008).

[31] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki. “Quantum entanglement”. *Reviews of Modern Physics* **81**, 865–942 (2009).

[32] Mark Wilde. “Quantum information theory”. Cambridge University Press (2017). Second edition.

[33] Richard Jozsa and Noah Linden. “On the role of entanglement in quantum-computational speed-up”. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* **459**, 2011–2032 (2003).

[34] Guifré Vidal. “Efficient Classical Simulation of Slightly Entangled Quantum Computations”. *Physical Review Letters* **91**, 147902 (2003).

[35] F. Verstraete, J. J. Garcia-Ripoll, and J. I. Cirac. “Matrix Product Density Operators: Simulation of Finite-Temperature and Dissipative Systems”. *Physical Review Letters* **93**, 207204 (2004).

[36] A. P. Lund, A. Laing, S. Rahimi-Keshari, T. Rudolph, J. L. O’Brien, and T. C. Ralph. “Boson Sampling from a Gaussian State”. *Physical Review Letters* **113**, 100502 (2014).

[37] Craig S. Hamilton, Regina Kruse, Linda Sansoni, Sonja Barkhofen, Christine Silberhorn, and Igor Jex. “Gaussian Boson Sampling”. *Physical Review Letters* **119**, 170501 (2017).

[38] Regina Kruse, Craig S. Hamilton, Linda Sansoni, Sonja Barkhofen, Christine Silberhorn, and Igor Jex. “Detailed study of Gaussian boson sampling”. *Physical Review A* **100**, 032326 (2019).

[39] Abhinav Deshpande, Arthur Mehta, Trevor Vincent, Nicolás Quesada, Marcel Hinsche, Marios Ioannou, Lars Madsen, Jonathan Lavoie, Haoyu Qi, Jens Eisert, Dominik Hangleiter, Bill Fefferman, and Ish Dhand. “Quantum computational advantage via high-dimensional Gaussian boson sampling”. *Science Advances* **8**, eabi7894 (2022).

[40] Daniel Grier, Daniel J. Brod, Juan Miguel Arrazola, Marcos Benicio De Andrade Alonso, and Nicolás Quesada. “The Complexity of Bipartite Gaussian Boson Sampling”. *Quantum* **6**, 863 (2022).

[41] Ulysse Chabaud and Mattia Walschaers. “Resources for Bosonic Quantum Computational Advantage”. *Physical Review Letters* **130**, 090602 (2023).

[42] Quntao Zhuang, Zheshen Zhang, and Jeffrey H. Shapiro. “Distributed quantum sensing using continuous-variable multipartite entanglement”. *Physical Review A* **97**, 032329 (2018).

[43] Emanuele Polino, Mauro Valeri, Nicolò Spagnolo, and Fabio Sciarrino. “Photonic quantum metrology”. *AVS Quantum Science* **2**, 024703 (2020).

[44] Changhun Oh, Changhyoup Lee, Seok Hyung Lie, and Hyunseok Jeong. “Optimal distributed quantum sensing using Gaussian states”. *Physical Review Research* **2**, 023030 (2020).

[45] Marco Malitesta, Augusto Smerzi, and Luca Pezzè. “Distributed Quantum Sensing with Squeezed-Vacuum Light in a Configurable Network of Mach-Zehnder Interferometers”. *arXiv:2109.09178* (2021).

[46] Marco Barbieri. “Optical Quantum Metrology”. *PRX Quantum* **3**, 010202 (2022).

[47] Gerardo Adesso. “Entanglement of Gaussian states”. *arXiv:0702069 [quant-ph]* (2007).

[48] Han-Sen Zhong, Hui Wang, Yu-Hao Deng, Ming-Cheng Chen, Li-Chao Peng, Yi-Han Luo, Jian Qin, Dian Wu, Xing Ding, Yi Hu, Peng Hu, Xiao-Yan Yang, Wei-Jun Zhang, Hao Li, Yuxuan Li, Xiao Jiang, Lin Gan, Guangwen Yang, Lixing You, Zhen Wang, Li Li, Nai-Le Liu, Chao-Yang Lu, and Jian-Wei Pan. “Quantum computational advantage using photons”. *Science* **370**, 1460–1463 (2020).

[49] Han-Sen Zhong, Yu-Hao Deng, Jian Qin, Hui Wang, Ming-Cheng Chen, Li-Chao Peng, Yi-Han Luo, Dian Wu, Si-Qiu Gong, Hao Su, Yi Hu, Peng Hu, Xiao-Yan Yang, Wei-Jun Zhang, Hao Li, Yuxuan Li, Xiao Jiang, Lin Gan, Guangwen Yang, Lixing You, Zhen Wang, Li Li, Nai-Le Liu, Jelmer J. Renema, Chao-Yang Lu, and Jian-Wei Pan. “Phase-Programmable Gaussian Boson Sampling Using Stimulated Squeezed Light”. *Physical Review Letters* **127**, 180502 (2021).

[50] Lars S. Madsen, Fabian Laudenbach, Mohsen Falamarzi, Askarani, Fabien Rortais, Trevor Vincent, Jacob F. F. Bulmer, Filippo M. Miatto, Leonhard Neuhaus, Lukas G. Helt, Matthew J. Collins, Adriana E. Lita, Thomas Gerrits, Sae Woo Nam, Varun D. Vaidya, Matteo Menotti, Ish Dhand, Zachary Vernon, Nicolás Quesada, and Jonathan Lavoie. “Quantum computational advantage with a programmable photonic processor”. *Nature* **606**, 75–81 (2022).

[51] Budhaditya Bhattacharjee, Pratik Nandy, and Tanay Pathak. “Eigenstate capacity and Page curve in fermionic Gaussian states”. *Physical Review B* **104**, 214306 (2021).

[52] A Serafini, OCO Dahlsten, D Gross, and MB Plenio. “Canonical and micro-canonical typical entanglement of continuous variable systems”. *Journal of Physics A: Mathematical and Theoretical* **40**, 9551 (2007).

[53] Alessio Serafini, Oscar C. O. Dahlsten, and Martin B. Plenio. “Teleportation Fidelities of Squeezed States from Thermodynamical State Space Measures”. *Physical Review Letters* **98**, 170501 (2007).

[54] Motohisa Fukuda and Robert Koenig. “Typical entanglement for Gaussian states”. *Journal of Mathematical Physics* **60**, 112203 (2019).

[55] Gerardo Adesso, Davide Girolami, and Alessio Serafini. “Measuring Gaussian Quantum Information and Correlations Using the Rényi Entropy of Order 2”. *Physical Review Letters* **109**, 190502 (2012).

[56] Giancarlo Camilo, Gabriel T. Landi, and Sebas Eliëns. “Strong subadditivity of the Rényi entropies for bosonic and fermionic Gaussian states”. *Physical Review B* **99**, 045155 (2019).

[57] V. Bužek, C. H. Keitel, and P. L. Knight. “Sampling entropies and operational phase-space measurement. I. General formalism”. *Physical Review A* **51**, 2575–2593 (1995).

[58] Gerardo Adesso and R Simon. “Strong subadditivity for log-determinant of covariance matrices and its applications”. *Journal of Physics A: Mathematical and Theoretical* **49**, 34LT02 (2016).

[59] Ludovico Lami, Christoph Hirche, Gerardo Adesso, and Andreas Winter. “Schur Complement Inequalities for Covariance Matrices and Monogamy of Quantum Correlations”. *Physical Review Letters* **117**, 220502 (2016).

[60] Alessio Serafini. “Quantum continuous variables: a primer of theoretical methods”. CRC Press (2017).

[61] F. C. Khanna, J. M. C. Malbouisson, A. E. Santana, and E. S. Santos. “Maximum entanglement in squeezed boson and fermion states”. *Physical Review A* **76**, 022109 (2007).

[62] Stasja Stanisic, Noah Linden, Ashley Montanaro, and Peter S. Turner. “Generating entanglement with linear optics”. *Physical Review A* **96**, 043861 (2017).

[63] Marko Petkovšek, Herbert S. Wilf, and Doron Zeilberger. “A=B”. A K Peters (1996).

[64] W.N. Bailey. “Generalized Hypergeometric Series, By W.N. Bailey”. Cambridge Tracts in Mathematics and Mathematical Physics, No. 32. Cambridge University Press (1964). url: books.google.com/books?id=TVyswgEACAAJ.

[65] Wadim Zudilin. “Hypergeometric heritage of W. N. Bailey”. *Notices of the International Congress of Chinese Mathematicians* **7**, 32–46 (2019).

[66] Lucy Joan Slater. “Generalized hypergeometric functions”. Cambridge Univ. PressCambridge (1966).

[67] “Hypergeometric2F1”. WolframResearch (2001). <https://functions.wolfram.com/HypergeometricFunctions/Hypergeometric2F1/03/02/0002/>.

[68] Joseph T. Iosue. “GLO”. GitHub (2022). <https://github.com/jtiosue/GLO>.

[69] Don Weingarten. “Asymptotic behavior of group integrals in the limit of infinite rank”. *Journal of Mathematical Physics* **19**, 999–1001 (1978).

[70] Benoit Collins. “Moments and Cumulants of Polynomial random variables on unitary groups, the Itzykson-Zuber integral and free probability”. [arXiv.math-ph/0205010](https://arxiv.org/abs/math-ph/0205010) (2002).

[71] N. Alexeev, A. Pologova, and M. A. Alekseyev. “Generalized Hultman Numbers and Cycle Structures of Breakpoint Graphs”. *Journal of Computational Biology* **24**, 93–105 (2017).

[72] Max Alekseyev, Adam Ehrenberg, Joseph T. Iosue, and Alexey V. Gorshkov. “Calculating entanglement in linear optics via breakpoint graphs”. In preparation.

[73] Ilki Kim. “Rényi- α entropies of quantum states in closed form: Gaussian states and a class of non-Gaussian states”. *Physical Review E* **97**, 062141 (2018).

[74] Mark Wilde and Kunal Sharma. “PHYS 7895: Gaussian Quantum Information, Lecture 10” (2019). <https://markwilde.com/teaching/2019-spring-gqi/scribe-notes/lecture-10-scribed.pdf>.

[75] Lucas Hackl and Eugenio Bianchi. “Bosonic and fermionic Gaussian states from Kähler structures”. *SciPost Physics Core* **4**, 025 (2021).

[76] Quntao Zhuang, Thomas Schuster, Beni Yoshida, and Norman Y Yao. “Scrambling and complexity in phase space”. *Physical Review A* **99**, 062334 (2019).

[77] Tianci Zhou and Xiao Chen. “Non-unitary Entanglement Dynamics in Continuous Variable Systems”. *Physical Review B* **104**, L180301 (2021). [arXiv:2103.06507](https://arxiv.org/abs/2103.06507).

[78] Bingzhi Zhang and Quntao Zhuang. “Entanglement formation in continuous-variable random quantum networks”. *npj Quantum Information* **7**, 1–12 (2021).

[79] Abhinav Deshpande, Bill Fefferman, Minh C. Tran, Michael Foss-Feig, and Alexey V. Gorshkov. “Dynamical Phase Transitions in Sampling Complexity”. *Physical Review Letters* **121**, 030501 (2018).

[80] Gopikrishnan Muraleedharan, Akimasa Miyake, and Ivan H. Deutsch. “Quantum computational supremacy in the sampling of bosonic random walkers on a one-dimensional lattice”. *New Journal of Physics* **21**, 055003 (2019).

[81] Changhun Oh, Youngrong Lim, Bill Fefferman, and Liang Jiang. “Classical simulation of bosonic linear-optical random circuits beyond linear light cone”. [arXiv:2102.10083](https://arxiv.org/abs/2102.10083) (2021).

[82] Nishad Maskara, Abhinav Deshpande, Adam Ehrenberg, Minh C. Tran, Bill Fefferman, and Alexey V. Gorshkov. “Complexity Phase Diagram for Interacting and Long-Range Bosonic Hamiltonians”. *Physical Review Letters* **129**, 150604 (2022).

[83] Changhun Oh, Youngrong Lim, Bill Fefferman, and Liang Jiang. “Classical Simulation of Boson Sampling Based on Graph Structure”. *Physical Review Letters* **128**, 190501 (2022).

[84] OEIS Foundation Inc. “The On-Line Encyclopedia of Integer Sequences” (2022). Published electronically at <http://oeis.org>.

- [85] Ewan Delanoy. “Determinant of matrix defined by binomial coefficient”. Mathematics Stack Exchange (2017). <https://math.stackexchange.com/q/2277633>.
- [86] Milton Abramowitz and Irene A. Stegun, editors. “Handbook of mathematical functions: with formulas, graphs, and mathematical tables”. Dover books on mathematics. Dover Publ (2013).
- [87] Darij Grinberg. “A hyperfactorial divisibility”. [Darij Grinberg homepage](#).
- [88] Motohisa Fukuda, Robert König, and Ion Nechita. “RTNI—A symbolic integrator for Haar-random tensor networks”. *Journal of Physics A: Mathematical and Theoretical* **52**, 425303 (2019).