**THE EUROPEAN**
**PHYSICAL JOURNAL B**

Regular Article

# Entanglement entropy of random partitioning[*]

Gergő Roósz[1,2,a], István A. Kovács[2,3,4], and Ferenc Iglói[2,5]

[1] Institute of Theoretical Physics, Technische Universität Dresden, 01062 Dresden, Germany
[2] Wigner Research Centre for Physics, Institute for Solid State Physics and Optics, P.O. Box 49, 1525 Budapest, Hungary
[3] Department of Physics and Astronomy, Northwestern University, Evanston, IL 60208-3112, USA
[4] Department of Network and Data Science, Central European University, Budapest 1051, Hungary
[5] Institute of Theoretical Physics, Szeged University, 6720 Szeged, Hungary

**Abstract.** We study the entanglement entropy of random partitions in one- and two-dimensional critical fermionic systems. In an infinite system we consider a finite, connected (hypercubic) domain of linear extent $L$, the points of which with probability $p$ belong to the subsystem. The leading contribution to the average entanglement entropy is found to scale with the volume as $a(p)L^D$, where $a(p)$ is a non-universal function, to which there is a logarithmic correction term, $b(p)L^{D-1} \ln L$. In $1D$ the prefactor is given by $b(p) = \frac{c}{3} f(p)$, where $c$ is the central charge of the model and $f(p)$ is a universal function. In $2D$ the prefactor has a different functional form of $p$ below and above the percolation threshold.

## 1 Introduction

The entanglement properties of many-body quantum systems are subjects of recent intensive theoretical studies [1–4]. The entanglement between two partitions $A$ and $B$ of a system being in a pure state $|\Psi\rangle$ can be measured by the entanglement entropy [5]: $S = -\text{Tr}_B \rho_B \ln \rho_B = -\text{Tr}_A \rho_A \ln \rho_A$. Here $\rho_A = \text{Tr}_B \rho$ and $\rho_B = \text{Tr}_A \rho$ are the reduced density matrices of the subsystems $A$ and $B$, respectively, and $\rho = |\Psi\rangle\langle\Psi|$.

Most of the studies are restricted to bipartitions with a smooth, regular boundary between $A$ and $B$, for example in one dimension ($1D$) the subsystem $A$ contains the successive sites $i = 1, 2, \ldots, L$, and $B$ is represented by the rest of the sites. If the system is gapped the entanglement entropy generally satisfies the so-called area law [3]: $S \sim L^{D-1}$. In one-dimensional critical systems, with algebraically decaying correlations, the area law is supplemented by a logarithmic correction, which for conformally invariant systems is given by [6–14]

$$S(L) \simeq \frac{c}{3} \ln L + c_1 \qquad (1)$$

where $c$ is the central charge of the conformal algebra.

Multidimensional ($D > 1$) free fermion systems satisfy an area law if the spectrum is gapped [15] or the Fermi surface has high codimension [16]. In the case of a sharp $D-1$ dimensional Fermi surface, there is a logarithmic correction to the are law [15,17], and the entanglement entropy is given by the following expression [18–20]

$$S(L) = \frac{L^{d-1}}{(2\pi)^{d-1}} \frac{\ln L}{12} \int \int |n_x \cdot n_k| dA_x dA_k , \qquad (2)$$

where the integral is over a scaled version of the spatial and the Fermi surface, in such a way that the volume of the (scaled) Fermi sea is 1, and $n_x$ and $n_k$ are unit normals to the real space boundary and the Fermi surface, respectively.

In disordered quantum spin chains the average entanglement entropy at the critical point has a logarithmic size dependence [21–27], too, which can be calculated by the strong disorder RG method [28,29]. In higher dimensional, critical random quantum systems there is an additive logarithmic correction due to corners, the prefactor of which is universal, i.e. independent of the form of disorder [30].

If the subsystem $A$ is not a singly connected domain, much less (analytical) results are available. Here we mention that if $A$ and $B$ contain the sites of two sublattices the contact points between them scale with the volume of the system and so behaves the entanglement entropy, too [31]. The entanglement entropy of irregular subsystems with non-continuous border is also subject of research in the recent years. General upper and lower bounds have been set for fractal boundary in real space and fractal like Fermi-surface in reference [20]. Fractal bipartition in

---

[*] Contribution to the Topical Issue "Recent Advances in the Theory of Disordered Systems", edited by Ferenc Iglói and Heiko Rieger.

[a] e-mail: gergoe.roosz@tu-dresden.de

the topologically ordered phase of the toric code with a magnetic field was also investigated in reference [32].

In the present paper we study the entanglement entropy when $A$ is selected by a random partition. This type of setting has already been used in reference [33], in which points of a subsystem were selected with a probability $p$ and the low-lying part of the entanglement Hamiltonian is calculated for the non-critical Kitaev-chain [34]. Here we consider a somewhat different selection of the points of the subsystem. The complete system has an infinite extent in which a finite domain of linear size, $L$ is defined and points of this domain belong to $A$ with probability $p$. Due to a large number of interfaces between $A$ and $B$ the average entanglement entropy, $\langle S(L) \rangle$, is expected to have a leading volume term (for $p < 1$), to which singular subleading corrections could appear in critical systems, like for $p = 1$. Here we consider critical fermionic models, hopping models in $1D$ and $2D$, as well as the critical Kitaev-chain.

The rest of the paper is organised as follows. Models and the methods of calculations are presented in Section 2. Lower and upper bounds for the entanglement entropy are calculated in Section 3, while small $p$ and small $1-p$ expansions are performed in Section 4. This is followed by extensive numerical calculations in Section 5, for different values of the occupation probability $p$ and the linear size of the (hypercubic) domain, $L$. We discuss our results in Section 6 while detailed calculations are put to the Appendices.

## 2 Models and methods

We consider fermionic hopping models with half filling defined by the Hamiltonian

$$H = -t \sum_{\langle i,j \rangle} c_i^\dagger c_j \,, \tag{3}$$

in terms of the fermion creation, $c_i^\dagger$, and annihilation, $c_j$, operators at lattice sites $i$ and $j$, respectively and the summation runs over nearest neighbour lattice sites. The lattice is either an infinite chain ($1D$) or an infinite square lattice ($2D$), in the latter case the components of the positions are $i = (i_x, i_y)$ and $j = (j_x, j_y)$.

Having the two-point correlation function, $C(i,j) = \langle c_i^\dagger c_j \rangle$ for $i,j \in A$, we can calculate the entanglement entropy of the system as

$$S = -\operatorname{Tr}_A \left[ C \ln C + (1-C) \ln(1-C) \right]$$
$$= -\sum_{i=1}^{N_A} [\zeta_i \ln \zeta_i + (1 - \zeta_i) \ln(1 - \zeta_i)] = -\sum_{i=1}^{N_A} s(\zeta_i) \,, \tag{4}$$

where $N_A$ is the dimension of the correlation matrix (number of sites in the subsystem), and $\zeta_i$ are the eigenvalues of the correlation matrix. As discussed in the introduction we consider finite domains of linear extent, $L$ (subsequent points in $1D$ and a square in $2D$) the points of which

belong to the subsystem $A$ with probability $p$. We have $2^L$ ($2^{L^2}$) different subsystems in $1D$ ($2D$), for which the entanglement entropy needs to be averaged, while the average value of $N_A$ is $pL$ ($pL^2$) in $1D$ ($2D$).

For the hopping model $C(i,i) = 1/2$, whereas for $i \neq j$ we have

$$C(i,j) = \frac{1}{\pi(i-j)} \sin \frac{\pi(i-j)}{2} \tag{5}$$

in $1D$ and

$$C(i,j) = \begin{cases} 0, \text{ if } \quad (i_x - j_x)^2 - (i_y - j_y)^2 = 0 \\ \\ -\dfrac{(-1)^{i_x - j_x} - (-1)^{i_y - j_y}}{\pi^2 \left[ (i_x - j_x)^2 - (i_y - j_y)^2 \right]}, \text{ otherwise} \end{cases}, \tag{6}$$

in $2D$.

For the non-random partition (with $p = 1$) the bipartite entanglement entropy in $1D$ is given by equation (1) with the central charge $c^{\text{hop}} = 1$ and the constant is $c_1^{\text{hop}} \approx 0.726$ [11–13]. In $2D$ for an $L \times L$ subsystem the prefactor of $L \ln L$ in equation (2) is given by $2/3$, which has been verified by numerical calculations [19,38].

Our second fermionic model is the critical Kitaev chain defined by the Hamiltonian [34]

$$H_{Kit} = -\sum_{l=-\infty}^{l=+\infty} \left[ (c_{l+1}^\dagger c_l + c_l^\dagger c_{l+1}) - (c_{l+1}^\dagger c_l^\dagger + c_l c_{l+1}) \right. $$
$$\left. + (c_l^\dagger c_l - 1/2) \right] \,. \tag{7}$$

This model corresponds to the fermionic form of the critical quantum Ising chain, what is obtained after performing the standard Jordan-Wigner transformation [35]. The relevant correlation function for this model is

$$C(i,j) = \langle (c_i^\dagger - c_i)(c_j^\dagger + c_j) \rangle \,, \tag{8}$$

given by [26]

$$C(i,j) = \frac{2}{\pi} \frac{(-1)^{i-j}}{2(i-j)+1} \,. \tag{9}$$

For the non-random partition (with $p = 1$) the entanglement entropy of the Kitaev chain (or the critical quantum Ising chain) is again given by equation (1), with the central charge $c^{\text{Kit}} = 1/2$ and the constant is $c_1^{\text{Kit}} = c_1^{\text{hop}}/2 + c^{\text{Kit}}/3 \approx 0.528$. Note, however, that if the subsystem $A$ is not a single connected domain, as is the case for random partitions, then the entanglement entropy of the critical Kitaev chain and that of the critical quantum Ising chain is different, due to the non-local nature of the Jordan-Wigner transformation [31].
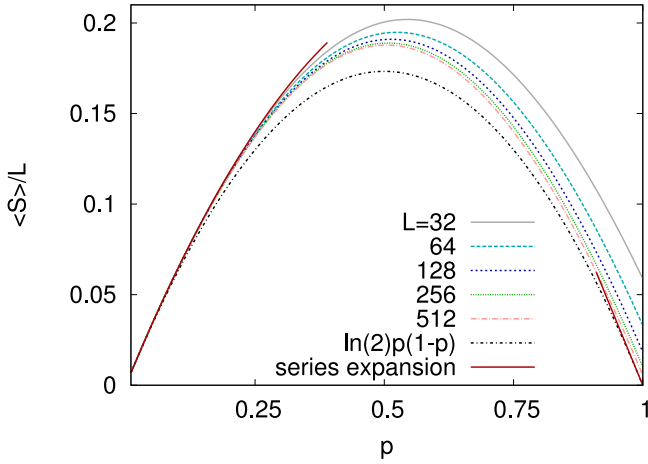
**Fig. 1.** Average entanglement entropy per domain size of the $1D$ hopping model calculated by the indirect method. The series expansion results for $p \ll 1$ and for $1 - p \ll 1$ are presented by red lines, whereas the obtained lower bound is drawn by a dotted line.

## 3 Lower and upper bounds from particle number fluctuations

Following standard techniques [36,37], we can calculate a lower bound to the entanglement entropy by using the inequality

$$s(x) \geq 4\ln(2)x(1-x) \qquad (10)$$

where $s(x)$ is defined in equation (4). Then, for the entropy we obtain

$$S \geq 4\ln 2 \operatorname{Tr}\left[C - C^2\right] = 4\ln 2 \left[\langle N^2 \rangle - \langle N \rangle^2\right] , \qquad (11)$$

where $N = \sum_{i=1}^{N_A} c_i^\dagger c_i$ is the particle number operator in the subsystem. This lower bound was used to prove the behaviour of the entanglement entropy of multidimensional free fermions [17] as well as of fractal-shaped partitions [32]. The particle number fluctuations in the case of a random partition of a uniform probability are presented in Appendix A in one and two dimensions. From these we obtain the lower bounds

$$\langle S \rangle_{1D} \geq \ln(2)Lp(1-p) + 4\ln(2)\frac{p^2}{\pi^2}\ln(L-1) , \qquad (12)$$

$$\langle S \rangle_{2D} \geq \ln(2)L^2 p(1-p) + 4\ln(2)\frac{2p^2}{\pi^2}L\left(\ln L - \frac{4}{\pi^2}\right) . \qquad (13)$$

These bounds are plotted in Figures 1 and 6. The leading term of the $2D$ lower bound in equation (13) is proportional to the area of $2D$ percolation clusters [39], given by $E_s^{\text{tot}} = 4L^2 p(1-p)$.

We note, that by shifting the parabola $x(1-x)$ upwards, one can also obtain upper bounds, for example with a shift of 0.08 it holds as

$$s(x) \leq 4\ln(2)x(1-x) + 0.08 . \qquad (14)$$

This leads to an upper bound for the prefactor of the volume term.

## 4 Limiting behaviours for $p \ll 1$ and for $1 - p \ll 1$

The average entanglement entropy can be calculated as a series expansion in $p$, performed in Appendix B for $1D$, leading to the following result up to $\mathcal{O}(p^3)$

$$\langle S \rangle_{1D} = (p\ln 2 - \alpha p^2 + \dots)L + \left(\frac{2p^2}{\pi^2} + \dots\right)\ln L , \qquad (15)$$

where $\alpha = 0.5335$ is defined in equation (B.7).

Similarly in two dimensions, the leading terms are

$$\langle S \rangle_{2D} = \left(p\ln 2 + \mathcal{O}(p^2)\right)L^2 + \left(\frac{2p^2}{\pi^2} + \mathcal{O}(p^4)\right)L\ln L . \qquad (16)$$

The other limiting case, $1 - p \ll 1$ can be treated as follows. Here, the logarithmic corrections approach the clean system's results, but due to dilution a volume term will appear. Let us now concentrate on the infinite subsystem (denoted by $B$), which – in the limit $1 - p \ll 1$ – consists of two half lines (in $1D$) or the whole plane without the square (in $2D$), as well as some isolated points from the interval (square). By neglecting correlations between isolated points, the correlation matrix of the infinite subsystem becomes block diagonal. One block contains the correlation matrix of the infinite subsystem without the isolated sites, i.e. with $p = 1$. The other block corresponds to the isolated sites, and has the dimension of the number of isolated points, $\tilde{N}_D = (1-p)L^D$, containing $1/2$ in its diagonal as

$$C_{\text{B}} = \left[\begin{array}{cc} C(p=1) & 0 \\ 0 & \frac{1}{2}\mathbb{1} \end{array}\right] . \qquad (17)$$

From this follows that the leading correction to the entanglement entropy is

$$\langle S \rangle = \operatorname{Tr} s(C_{\text{B}}) \approx S(p=1) + \ln(2)(1-p)L^d + \mathcal{O}\left[(1-p)^2\right]. \qquad (18)$$

Note, that the series expansion results for the volume term agree with the lower bound for $p \ll 1$ and $1 - p \ll 1$.

## 5 Numerical results

For finite subsystems of linear size, $L$, we have calculated the average entanglement entropy numerically. For the averaging process over different samples we have used three different methods. In the *direct method* a large number of samples are generated at a fixed value of $p$ and the entanglement entropy is calculated for each sample, and this calculation is repeated for several values of $p$. This

method generally gives an accurate average value at a given $p$, if the number samples is large enough ($10^6 \ldots 10^8$) for each $p$, but comparing the results for different values of $p$ leads to large errors since the samples are different at each $p$.

In the so called *indirect method* we use the relation

$$\langle S \rangle(p) = \sum_{n=0}^{N} \langle S_n \rangle \begin{pmatrix} N \\ n \end{pmatrix} p^n (1-p)^{N-n} , \qquad (19)$$

where $N = L$ ($N = L^2$) in $1D$ ($2D$), and $\langle S_n \rangle$ denotes the average entanglement entropy of a subsystem of $n$ sites, where the averaging is performed over all possible partitions. In practice, we have generated a large number of random samples with uniform probability for all values of $n$. These samples and their entanglement entropy are stored and used to calculate $\langle S \rangle(p)$ for different values of $p$. The advantage of the indirect method is that we need comparatively less samples ($\sim 10^6$) and the numerical derivation with respect to $p$ is more smooth, compared to the direct method.

In our third, *replica method* we generated for each random subsystem sample one (four) replicas in $1D$ ($2D$) and fused them together. By comparing the entanglement entropy of the original and the replicated sample one can cancel the leading volume term and gain direct access to the more interesting, subleading corrections. For the best results, our calculations have generally combined the replica method with the indirect method.

## 5.1 $1D$ hopping model

For the $1D$ hopping model we used the indirect method to calculate the average entanglement entropy for domain sizes $L = 16, 32, \ldots, 1024$, with $10^6$ realizations in each case. In agreement with the analytical results on the lower bound in equation (12) and the perturbation expansions in equations (15) and (18), the leading contribution scales linearly with the number of sites in the domain, $L$. This is illustrated in Figure 1, where the average numerical vale of the entanglement entropy per domain size $\langle S \rangle / L$ is plotted together with the analytical results.

The asymptotic behavior of the average entanglement entropy is expected to contain sub-leading terms in the form

$$\langle S \rangle(L) = a(p)L + b(p)\ln L + c_1(p) . \qquad (20)$$

The prefactor of the volume contribution, $a(p)$, which can be represented by extrapolating the curves in Figure 1, is close to symmetric, $a(p) \approx \tilde{a}p(1-p)$. More interesting are the subleading terms in equation (20), which are conveniently analysed by the replica method using the difference

$$2\langle S \rangle(L,p) - \langle S \rangle_{\mathrm{repl}}(2L,p) = b(p)(\ln L - \ln 2) + c_1(p) . \qquad (21)$$

Here $\langle S \rangle_{\mathrm{repl}}(2L,p)$ denotes the average entanglement entropy in the replicated samples, which are obtained by
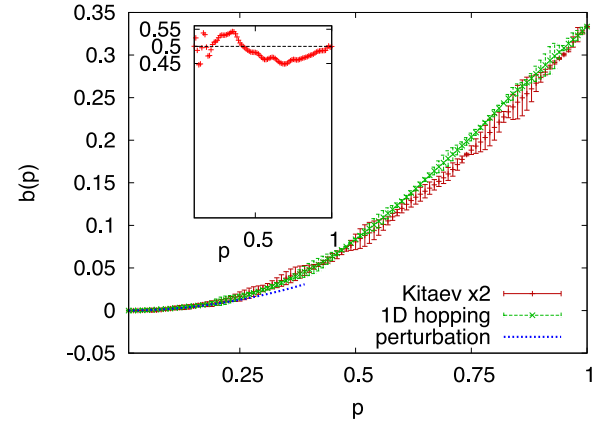


**Fig. 2.** Prefactor of the logarithmic term of the average entanglement entropy of the $1D$ hopping chain (green) and twice the same the $1D$ critical Kitaev chain (red), calculated by the replica method, see text. The series expansion result for the $1D$ hopping chain at small $p$ is shown by a dotted line. In the inset the ratio of the two prefactors are shown and the dashed line represents the ratio of the conformal charges: $c^{\mathrm{Kit}}/c^{\mathrm{hop}} = 1/2$.
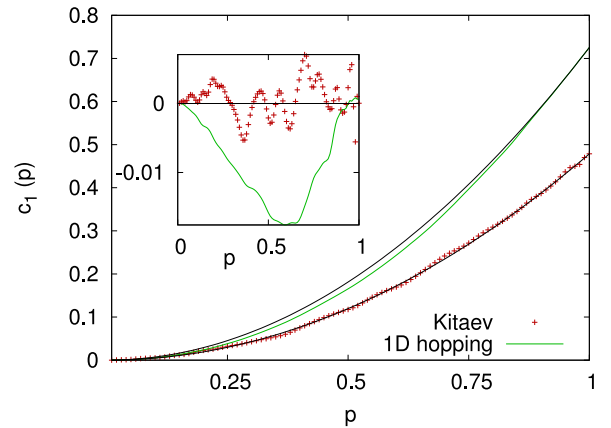


**Fig. 3.** The constant of the average entanglement entropy of the $1D$ hopping chain (green) and that of the $1D$ critical Kitaev chain (red), calculated by the replica method, see text. The dashed curves on the main panel corresponds to $c_1(1)p^2$ In the inset the difference $c_1(p) - c_1(1)p^2$ are shown for the two models.

joining the same sample behind another copy. By comparing results at sizes $L$ and $2L$, finite-size estimates are calculated for the prefactor, $b(p)$, and the constant, $c_1(p)$, which are then extrapolated. These are plotted in Figures 2 and 3, respectively. The prefactor, $b(p)$, starts quadratically for small $p$, in agreement with the series expansion in equation (15), while at $p = 1$ reaches the conformal result: $b(1) = c^{\mathrm{hop}}/3 = 1/3$, see in equation (1). The constant, $c_1(p)$, also appears to start quadratically at small $p$, while at $p = 1$ reaches the known result, as quoted below equation (6). Interestingly, we have found an overall quadratic dependence: $c_1(p) \approx p^2 c_1(1)$, as illustrated in the inset of Figure 3.

We have also studied the distribution of the entanglement entropy, shown in Figure 4 for different sizes at $p = 0.25$. These distributions are well represented by
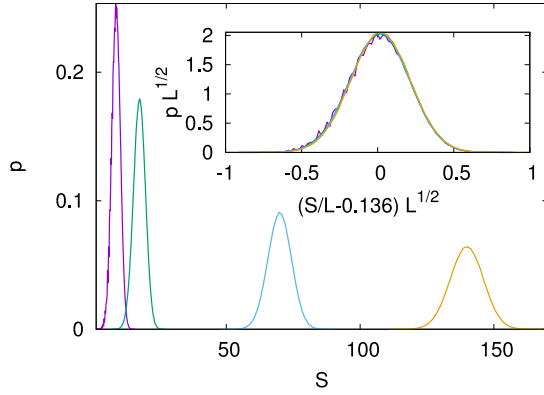
**Fig. 4.** Probability distribution of the entanglement entropy of the $1D$ hopping chain at $p = 0.25$ for different sizes: $L = 64, 128, 512, 1024$, from left to right. In the inset the scaled curves are shown assuming Gaussian behaviour.
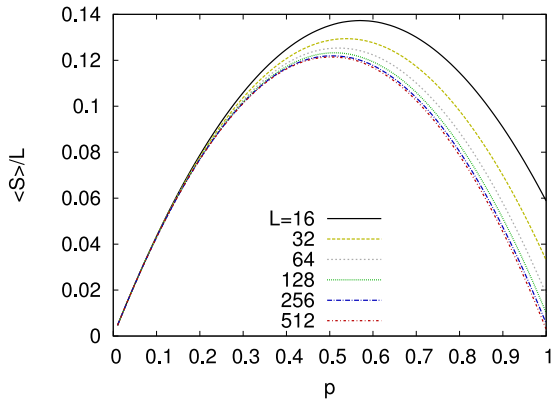


**Fig. 5.** Average entanglement entropy per domain size of the critical Kitaev chain calculated by the indirect method.

Gaussians, as illustrated in terms of scaled distributions in the inset. Similar, Gaussian distributions are observed for other values of $p$ as well, but close to $p = 1$ there is a cross-over regime, where the volume contribution, $\ln(2)(1-p)L$, and the logarithmic term, $c/3 \ln L$, compete, see in equation (18).

## 5.2 Critical Kitaev chain

For the critical Kitaev chain we have calculated the entanglement entropy of random partitions, as described in the previous subsection. Here we have used finite domains of size: $L = 16, 32, \ldots, 512$ and the averages are calculated by the indirect method over $10^6$ samples. As for the $1D$ hopping chain, the dominant contribution to the average entanglement entropy is the volume term, as illustrated in Figure 5 where the average entanglement entropy per domain size is shown for different values of $L$. The shape of the extrapolated curve is similar to that of the $1D$ hopping chain in Figure 1 and it is again approximately symmetric with respect to $p \to (1-p)$.

The subleading correction terms are found to be in the same form as given in equation (20). Using the replica method and equation (21) we have calculated estimates
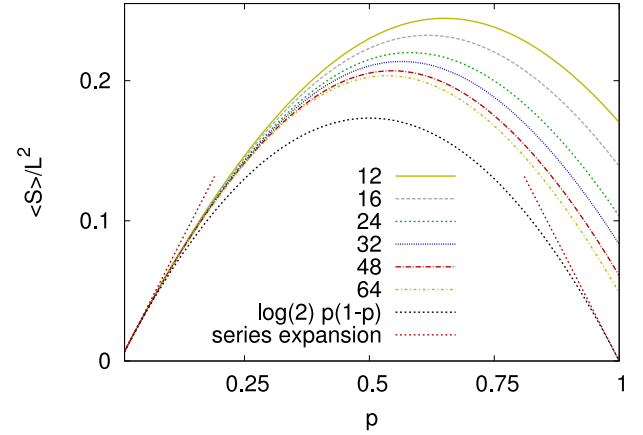


**Fig. 6.** Average entanglement entropy per domain volume of the $2D$ hopping model calculated by the indirect method. The series expansion for $p \ll 1$ and $1 - p \ll 1$ are presented by red dotted line, whereas the lower bound is drawn by a black dotted line.

for the prefactor of the logarithmic term, $b(p)$, as well as of the constant, $c_1(p)$, and their extrapolated values are plotted in Figures 2 and 3, respectively. Considering the prefactor, $b(p)$, its form is very similar to that found for the $1D$ hopping chain: their ratio is given by $b(p)^{\text{Kit}}/b(p)^{\text{hop}} \approx 1/2 = c_{\text{Kit}}/c^{\text{hop}}$. This is illustrated in the inset of Figure 2. As seen in Figure 3 the constant term, $c_1^{\text{Kit}}(p)$, has also an approximately quadratic dependence: $c_1^{\text{Kit}}(p) \approx p^2 c_1^{\text{Kit}}(1)$, as illustrated in the inset of Figure 3.

## 5.3 $2D$ hopping model

Here we consider the hopping model in a square lattice, in which the domain is an $L \times L$ square. We have calculated the entanglement entropy of samples having finite subsystems with linear extension $L = 8, 12, 16, 24, 32, 48$ and $64$, while averages are obtained through the indirect method over $10^5$ samples. According to the analytical results in equations (13) and (16) the average entanglement entropy is expected to be dominated by the surface term to which the first correction is logarithmic:

$$\langle S \rangle(L) = a(p)L^2 + b(p)L \ln L + \ldots \qquad (22)$$

This is in agreement with our numerical results in Figure 6, showing the average entanglement entropy per domain surface. For increasing $L$, the curves approach the prefactor, $a(p)$, which is approximately symmetric, $a(p) \approx \tilde{a}p(1-p)$. Comparing this figure with the one-dimensional results in Figures 1 and 5 the convergence is here slower, due to considerably smaller linear size of the domains in $2D$.

The prefactor of the logarithmic term is estimated through the replica method: comparing the (four times) entanglement entropy of each sample of size $L$, with those composed of four joint identical samples, thus having a linear size $2L$

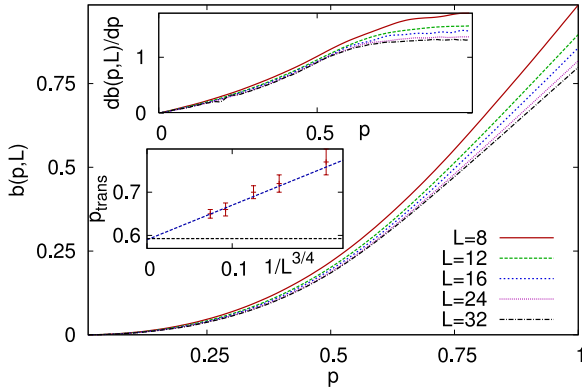$$4\langle S \rangle(L, p) - \langle S \rangle_{\text{repl}}(2L, p) = b(p, L)2L \ln(L/2) . \qquad (23)$$

**Fig. 7.** Effective, size-dependent prefactor of the logarithmic correction term in the average entanglement entropy of the 2$D$ hopping model calculated through the replica method in equation (23). In the first inset the derivative $\frac{\partial b(p,L)}{\partial p}$ is shown. In the second inset the finite-size transition points are plotted as a function of $L^{-3/4}$, see the text. The dashed blue line is guide to the eye, the horizontal black dashed line represents the transition point for site percolation.

Equation (23) defines an effective, size-dependent prefactor, $b(p,L)$, which is plotted in Figure 7. As seen in this figure $b(p,L)$ starts quadratically for small $p$ and becomes approximately linear for larger values of the probability. To study this behaviour further we have calculated the derivative of $b(p,L)$ with respect to $p$, which is shown in the first inset of Figure 7. We note, that in the indirect method the differentiation of equation (19) can be performed at each value of $p$, which reduces the error of the calculation. Inspecting the behaviour of $\frac{\partial b(p,L)}{\partial p}$ we can identify two regions. In the first regime the derivative continuously increases, while in the second regime it becomes approximately constant. In finite subsystems there is an extended cross-over region between the two regimes, which, however, shrinks with increasing $L$.

We summarize these findings in the conjecture that the change in the behaviour of $b(p,L)$ is related to the percolation transition [40], which takes place in the random partitioning at a critical value $p_c = 0.592$, if $L \to \infty$. To further check this hypothesis, we have defined finite-size transition points between the two regions as the crossing point, where the linear continuation of the curve starting from $p = .5$ for $p > 0.5$ reaches the value of the constant measured at $p \lesssim 1$. These finite-size transition points are plotted in the second inset of Figure 7 as a function of $L^{-1/\nu}$, with $\nu = 4/3$ being the correlation-length critical exponent of 2$D$ percolation, governing finite-size effects [40]. Indeed the extrapolated value of the finite-size transition point agrees with $p_c$, within the error of the calculation. We have also checked that the volume term with $a(p)$ shows no sign of a singularity at any value of $p$.

# 6 Discussion

We have studied the entanglement entropy of critical free-fermion models in one and two dimensions, when the sites

of the subsystem were taken from a hypercubic domain of linear size $L$ randomly, with probability $p$. We have investigated the average entanglement entropy by calculating lower bounds, by series expansions and performing extensive numerical calculations. When the entire system has infinite extent, the average entanglement entropy for $0 < p < 1$ is found to be dominated by the volume term $a(p)L^D$, which is supplemented by logarithmic corrections as $b(p)L^{D-1} \ln L$. The volume term is non-universal, which is connected to the fact, that the distribution of the entanglement entropy is Gaussian. On the contrary, the logarithmic correction is found to contain information about the universal, critical characteristics of the system. In 1$D$, comparing the results of the hopping chain and that of the Kitaev chain the prefactor of the logarithm is found to scale as: $b(p) = cf(p)$, where $f(p)$ is a universal, model independent function and $c$ is the central charge of the critical model. Interestingly, for both models the constant term is obtained in a pure quadratic form: $c_1(p) \approx p^2 c_1(1)$. In 2$D$, for the hopping model $b(p)$ is shown to change its behaviour at the percolation transition point, $p_c$, where the random subsystem develops an infinite cluster. According to our numerical results the derivative, $\frac{\partial b(p,L)}{\partial p}$, is increasing with $p$ for $p < p_c$, but it saturates to a constant for $p > p_c$. This conjectured behaviour would be interesting to justify independently by physical arguments, perhaps even with some rigorous method.

Our study can be extended to several further directions as discussed next.

## 6.1 Finite environment

We studied the case when the size of the entire system is $L_{\text{tot}} \to \infty$. For a finite value of $L_{\text{tot}}$, the results should depend on the ratio $L/L_{\text{tot}}$. In 1$D$, for non-random partitions with $p = 1$, the functional form of the entanglement entropy as a function of $L/L_{\text{tot}}$ is known from conformal invariance [7]. For the 1$D$ hopping model with random partitions, we have checked that the prefactor $b(p)$ vanishes in the case $L/L_{\text{tot}} = 1$. This is illustrated in Figure 8, where the ratio

$$\frac{\langle S \rangle(2L) - \langle S \rangle(L)}{L \ln 2} = \frac{a}{\ln 2} + b \frac{1}{L} \qquad (24)$$

is plotted against $1/L$ at $p = 1/2$. The slope of the points, which defines $b$, indeed tends to zero for $L \to \infty$, as seen in the inset of Figure 8.

## 6.2 Position dependent selection probability

Another potential extension of our study is to consider a different type of probability distribution for selecting the points of the subsystem. For the 1$D$ hopping model, we have also checked a position dependent probability

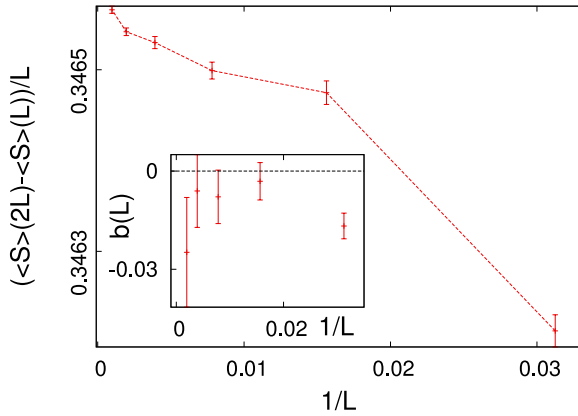$$p_i = 1 - \frac{1}{2} \frac{1}{l_i^{\kappa}} \,, \qquad (25)$$

**Fig. 8.** The ratio in equation (24) versus $1/L$ in the case with $L/L_{\mathrm{tot}} = 1$ and $p = 0.5$. In the inset the slope of the curve calculated by two point fit is presented, which defines finite-size estimates for the prefactor $b$.
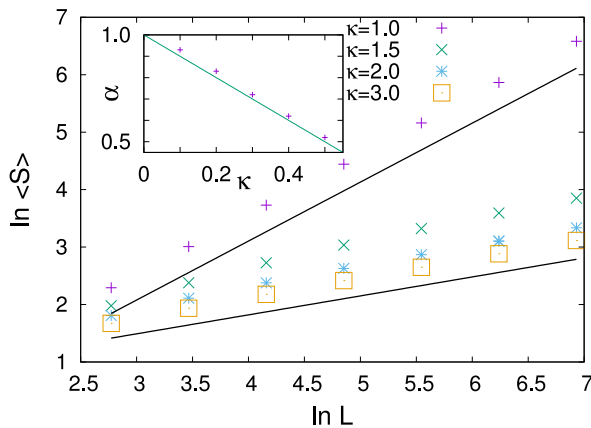


**Fig. 9.** Average entanglement entropy with site dependent probability in one dimension. The logarithm of the average entanglement entropy $\ln\langle S\rangle$ as a function of $\ln L$, for $1 \le \kappa \le 3$. The black lines are guides to the eye, the slope of the lower and upper lines are $1/3$ and $1/3 + \ln(2)$, respectively. In the inset the $\alpha$ exponent of the power-law increase of the average entanglement entropy is shown as a function of $0.1 \le \kappa \le 0.5$, see text.

where $l_i = \min(i, L - i)$ is the distance of the point $i$ from the nearest edge of the domain. By varying the decay exponent $\kappa \ge 0$ one can interpolate between the non-random partitioning with $p_i = 1$ for $\kappa \to \infty$ and the uniform probability partitioning with $p_i = 1/2$ for $\kappa = 0$. The number of internal contact points between the subsystem and the environment scales as $\int_0^L l^{-\kappa}\mathrm{d}l$, which is finite for $\kappa > 1$, it scales as $\ln L$ for $\kappa = 1$ and behaves as $L^{1-\kappa}$ for $0 \le \kappa < 1$.

We have calculated the average entanglement entropy for different values of $\kappa$, ranging between 0.1 and 3.0 and the results are presented in Figure 9. For $\kappa > 1$ the dominant contribution is $\langle S\rangle = 1/3 \ln L$ as for the non-random partitioning case with $\kappa \to \infty$. This is due to the fact, that the "volume term", which scales with the number of contact points is now $\mathcal{O}(1)$, thus it is subleading. In the borderline case, $\kappa = 1$, the size-dependence of

$\langle S\rangle$ is still logarithmic, however with a different prefactor: $\langle S\rangle \approx (1/3 + \ln 2)\ln L$. The increase of the prefactor now is due to the "volume contribution", which scales also logarithmically. Finally, for $0 \le \kappa < 1$ the average entanglement entropy scales as: $\langle S\rangle \sim L^\alpha$, with $\alpha = 1 - \kappa$, which involves the number of contact points and now it is the dominant contribution. These results are illustrated in the inset of Figure 9.

### 6.3 Some open problems

It would be interesting to try to determine the function $f(p)$ by a different method, perhaps even analytically. It could also be interesting to check, if the observed universality holds for non-fermionic models, too. In addition, one can consider models with random couplings and pose similar questions treated in this paper.

## Author contribution statement

All authors contributed equally.

## Appendix A: Particle number fluctuations

In $1D$, the average particle number fluctuations are given by

$$\left[\langle N^2\rangle - \langle N\rangle^2\right]_{1D} = \langle \mathrm{Tr}\, C\rangle - \langle \mathrm{Tr}\, C^2\rangle$$

$$= \left\langle \sum_{i\in A} c_i^\dagger c_i\right\rangle - \left\langle \sum_{i\in A}\sum_{j\in A} |C(i,j)|^2\right\rangle$$

$$= \frac{1}{2}pL - p\sum_{i=1}^{L}|C(i,i)|^2 - p^2\sum_{i\neq j}|C(i,j)|^2$$

$$= \frac{1}{2}pL - \frac{1}{4}pL - 2p^2\sum_{k=1}^{L/2}(L-2k+1)\frac{1}{\pi^2(2k-1)^2}$$

$$\approx \frac{1}{4}Lp(1-p) + \frac{p^2}{\pi^2}\ln L\,. \tag{A.1}$$

In $2D$, we obtain by a similar calculation

$$\left[\langle N^2\rangle - \langle N\rangle^2\right]_{2D} \approx \frac{L^2}{4}p(1-p) + \frac{2p^2}{\pi^2}L\left(\ln(L) - \frac{4}{\pi^2}\right) . \tag{A.2}$$

## Appendix B: Series expansion

In $1D$, the average entanglement entropy is written as

$$\langle S\rangle = \sum_{N=1}^{L} p^N (1-p)^{L-N} \sum_{A, |A|=N} \mathrm{Tr}[s(C_A)] \tag{B.1}$$

where the second sum goes for every subsystems $A$ including $N$ sites, and having the correlation matrix $C_A$. For small $p$, we keep in equation (B.1) the terms with $N=1$ and $N=2$ and omit terms with $\mathcal{O}(p^3)$, leading to

$$\langle S\rangle = pL\ln 2 - p^2(L-1)L\ln 2 + p^2 \sum_{A,|A|=2} \mathrm{Tr}[s(C_A)] + \mathcal{O}(p^3).\tag{B.2}$$

The correlation matrix of the two site subsystem is

$$C = \left[\begin{array}{cc} 1/2 & C(m,n) \\ C(m,n) & 1/2 \end{array}\right] \tag{B.3}$$

where $m$ and $n$ are indices of the points included in the subsystem and $C(m,n)$ is given in equation (5). The eigenvalues of $C$ are

$$\zeta_{n,m} = \begin{cases} 1/2 & \text{if} \quad |n-m| = \text{even} \\ 1/2 \pm \dfrac{1}{\pi(n-m)} & \text{if} \quad |n-m| = \text{odd} . \end{cases} \tag{B.4}$$

The entropy contribution from a term with $|n-m| = $ even is $2\ln 2$, whereas from a term with $|n-m| = $ odd is given by $2\ln 2 + \Delta(n,m)$, with

$$\Delta(n,m) = \left[2s\left(1/2 + \frac{1}{\pi(n-m)}\right) - 2\ln 2\right] . \tag{B.5}$$

Substituting this into equation (B.2) leads to

$$\begin{aligned} \langle S\rangle &= pL\ln 2 + p^2 \sum_{n,m=1;\ |n-m|\,\mathrm{odd}}^{L} \Delta(n,m) \\ &= pL\ln 2 + np^2 \sum_{k=1}^{L/2} (L - 2k + 1)\Delta(2k,1) \\ &\approx L(p\ln 2 - \alpha p^2) + p^2\frac{2}{\pi^2}\ln L , \end{aligned} \tag{B.6}$$

with

$$\alpha = -\sum_{k=1}^{\infty} \left[2s\left(1/2 + \frac{1}{\pi(2k-1)}\right) - 2\ln 2\right] \approx 0.5335 . \tag{B.7}$$

## References

1. L. Amico, R. Fazio, A. Osterloh, V. Vedral, Rev. Mod. Phys. **80**, 517 (2008)
2. P. Calabrese, J. Cardy, B. Doyon, J. Phys. A **42**, 500301 (2009)
3. J. Eisert, M. Cramer, M.B. Plenio, Rev. Mod. Phys. **82**, 277 (2010)
4. N. Laflorencie, Phys. Rep. **643**, 1 (2016)
5. C.H. Bennett, H.J. Bernstein, S. Popescu, B. Schumacher, Phys. Rev. A **53**, 2046 (1996)
6. C. Holzhey, F. Larsen, F. Wilczek, Nucl. Phys. B **424**, 443 (1994)
7. P. Calabrese, J. Cardy, J. Stat. Mech. **2004**, P06002 (2004)
8. G. Vidal, J.I. Latorre, E. Rico, A. Kitaev, Phys. Rev. Lett. **90**, 227902 (2003)
9. J.I. Latorre, E. Rico, G. Vidal, Quantum Inf. Comput. **4**, 048 (2004)
10. I. Peschel, J. Phys. A: Math. Gen. **36**, L205 (2003)
11. B.-Q. Jin, V.E. Korepin, J. Stat. Phys. **116**, 79 (2004)
12. A.R. Its, B.-Q. Jin, V.E. Korepin, J. Phys. A **38**, 2975 (2005)
13. R. Its, B.-Q. Jin, V.E. Korepin, in *Fields Institute Communications, Universality and Renormalization*, edited by I. Bender and D. Kreimer (2007), Vol. 50, p. 151
14. F. Iglói, R. Juhász, Europhys. Lett. **81**, 57003 (2008)
15. M.M. Wolf, Phys. Rev. Lett. **96**, 010404 (2006)
16. W. Li, L. Ding, R. Yu, S. Haas, Phys. Rev. B **74**, 073103 (2006)
17. S. Farkas, Z. Zimborás, J. Math. Phys. **48**, 102110 (2007)
18. B. Swingle, Phys. Rev. Lett. **105**, 050502 (2010)
19. T. Barthel, M.-C. Chung, U. Schollwöck, Phys. Rev. A **74**, 022329 (2006)
20. D. Gioev, I. Klich, Phys. Rev. Lett. **96**, 100503 (2006)
21. G. Refael, J.E. Moore, Phys. Rev. Lett. **93**, 260602 (2004)
22. R. Santachiara, J. Stat. Mech. Theor. Exp. **2006**, L06002 (2006)
23. N.E. Bonesteel, K. Yang, Phys. Rev. Lett. **99**, 140405 (2007)
24. G. Refael, J.E. Moore, Phys. Rev. B **76**, 024419 (2007)
25. N. Laflorencie, Phys. Rev. B **72** 140408 (R) (2005)
26. F. Iglói, Y.-Ch. Lin, J. Stat. Mech. **2008**, P06004 (2008)
27. G. De Chiara, S. Montangero, P. Calabrese, R. Fazio, J. Stat. Mech. **2006**, L03001 (2006)
28. For a review, see: F. Iglói, C. Monthus, Phys. Rep. **412**, 277 (2005)
29. F. Iglói, C. Monthus, Eur. Phys. J. B **91**, 290 (2018)
30. I.A. Kovács, F. Iglói, Europhys. Lett. **97**, 67009 (2012)
31. F. Iglói, I. Peschel, Europhys. Lett. **89**, 40001 (2010)
32. A. Hamma, D.A. Lidar, S. Severini, Phys. Rev. A **81**, 010102(R) (2010)
33. S. Vijay, L. Fu, Phys. Rev. B **91**, 22 (2015)
34. A.Y. Kitaev, Phys. Usp. **44**, 131 (2001)
35. P. Pfeuty, Phys. Lett. A **72**, 245 (1979)
36. I. Peschel, Braz. J. Phys. **42**, 267 (2012)
37. V. Eisler, I. Peschel, J. Stat. Mech. **2018**, 104001 (2018)
38. W. Li, L. Ding, R. Yu, T. Roscilde, S. Haas Phys. Rev. B **74**, 073103 (2006)
39. P.S. Grinchuk, O.S. Rabinovich, J. Exp. Theor. Phys. **96**, 301 (2003)
40. D. Stauffer, A. Aharony, *Introduction To Percolation Theory* (Taylor and Francis, London, 1992)