Universal logarithmic correction to Rényi (Shannon) entropy in generic systems of critical quadratic fermions

Babak Tarighi,[1] Reyhaneh Khasseh,[2,3] M. N. Najafi,[1] and M. A. Rajabpour [4]

[1]*Department of Physics, University of Mohaghegh Ardabili, P.O. Box 179, Ardabil, Iran*
[2]*Theoretical Physics III, Center for Electronic Correlations and Magnetism, Institute of Physics, University of Augsburg, D-86135 Augsburg, Germany*
[3]*Max-Planck-Institut für Physik Komplexer Systeme, Nöthnitzer Strasse 38, D-01187, Dresden, Germany*
[4]*Instituto de Fisica, Universidade Federal Fluminense, Av. Gal. Milton Tavares de Souza s/n, Gragoatá, 24210-346, Niterói, RJ, Brazil*

The Rényi (Shannon) entropy, i.e., $Re_\alpha(Sh)$, of the ground state of quantum systems in local bases normally show a volume-law behavior. For a subsystem of quantum chains at a critical point there is an extra logarithmic subleading term with a coefficient which is universal. In this paper we study this coefficient for generic time-reversal translational invariant quadratic critical free fermions. These models can be parametrized by a complex function which has zeros on the unit circle. When the zeros on the unit circle do not have degeneracy and there is no zero outside of the unit circle we are able to classify the coefficient of the logarithm. In particular, we numerically calculate the Rényi (Shannon) entropy in a configuration basis for a wide variety of these models and show that there are two distinct classes. For systems with $U(1)$ symmetry the coefficient is proportional to the central charge, i.e., one half of the number of points that one can linearize the dispersion relation of the system; for all the values of α with transition point at $\alpha = 4$. For systems without this symmetry, when $\alpha > 1$, this coefficient is again proportional to the central charge. However, the coefficient for $\alpha \leqslant 1$ is a new universal number. Finally, by using the discrete version of the Bisognano-Wichmann modular Hamiltonian of the Ising chain we show that these coefficients are universal and dependent on the underlying CFT.

I. INTRODUCTION

In quantum mechanics the outcome of a measurement of an observable is one of the eigenvalues of the observable. Each outcome happens with a particular probability. These probabilities can be used to calculate Rényi (Shannon) entropy which is a representative number for the probability distribution. The number depends on the chosen observable and gives an idea about the distribution of probabilities. For many-body systems there are many possibilities to choose the observable and study its distribution and extract interesting information. In quantum chains one can look to a local observable defined on each site and find the probability of having a particular configuration for the full system in, for example, the ground state. This will lead to a set of probabilities where its size grows linearly with the size of the Hilbert space. In quantum spin chains, when one takes the ground state, the Rényi (Shannon) entropy in $\sigma^{x,y,z}$ basis present some information about the phase transition and the universality class [1–6].

Instead of calculating the Rényi (Shannon) entropy of the full system one can use marginal probabilities and calculate the same quantities for the subsystem. These quantities as their full system counterparts also show a volume-law behavior, however, at the phase transition point there is a logarithmic subleading term the coefficient of which shows interesting *universal* behavior [7–15]. Studies on many different quantum critical spin chains reveal that the coefficient of the logarithm depends on the chosen basis but shows some level of universality in some particular bases dubbed as conformal basis [11]. In these models there are infinite possibilities to choose the local observable and it seems any kind of classification is hopeless. In fermionic systems the situation seems more tractable. The most obvious local observable to take is the number operator. One can write the ground state in a configuration basis and look to the probabilities of different configurations. These probabilities are dubbed as formation probabilities and have been studied for subsystems of certain free fermions in depth [13,16–19]. For results on the full system see [3,20]. These probabilities have been also investigated in experiments [21].

Time-reversal translational invariant quadratic critical free fermions show interesting phase transitions. Depending on the couplings one can produce critical systems with integer and half-integer central charges [22,23]. They also show interesting topologically protected phases [23,24]. In addition, there are many efficient methods to calculate the formation

probabilities for extremely large systems [13,17]. These methods are also useful to work directly with subsystems embedded in the systems with infinite size. This is very useful to avoid the problem of finite size effect regarding the full system. We notice that since the number of probabilities grows exponentially with the size of the subsystem, there is an unavoidable limitation on the size of the subsystem that one can take in numerical calculations. In this paper we make a step in full classification of the coefficient of the logarithmic term in the Rényi (Shannon) entropy of generic time-reversal translational invariant quadratic critical free fermions. We calculate this quantity for various critical models and show that the coefficient is proportional to the number of points that one can linearize the dispersion relation but the proportionality constant is very much dependent on the presence (absence) of the $U(1)$ symmetry. In systems with $U(1)$ symmetry a clear picture emerges for the coefficient of the logarithm with respect to α. However, for systems without this symmetry the picture is clear just for $\alpha > 1$.

The paper is organized as follows: In Sec. II we first define the Rényi (Shannon) entropy for the subsystem. To extract the coefficient of the logarithm we define the quantity I_α for two subsystems of our original subsystem which was embedded in an infinite system. The setup used in this paper has not been considered previously. Most of the previous studies worked with a system which was a periodic finite system and partitioned the system into two parts [7,10,11]. In our setup we have a tripartite situation.

In Sec. III we introduce the kind of models that we considered in this study, i.e., time-reversal translational invariant quadratic critical free fermions. Apart from their physical appeal these models provide a series of different universality classes. They can be solved exactly and one can find the desired formation probabilities exactly and efficiently in the thermodynamic limit. We categorize these models to two types, those with and without $U(1)$ symmetry. We also show how one can find the formation probabilities out of the correlation matrices for these models. A couple of interesting dualities regarding the correlation matrices of different models will also be presented in this section.

In Sec. IV we summarize our main results. We make a few conjectures regarding the coefficient of the logarithm in the models that we considered. It seems there are two classes. Those that have $U(1)$ symmetry and models without manifest $U(1)$ symmetry. In the latter models we just consider models where the corresponding $f(z)$ function does not have zero outside of the unit circle.

In Sec. V we briefly describe our numerical and fitting procedure. Then in Sec. VI we present the details of the models that we considered and provide support for the results presented in Sec. IV. In Sec. VII we use the discrete version of the Bisognano-Wichmann modular Hamiltonian for the Ising chain and show that the results converge rapidly to the exact results. Finally, in Sec. VIII we discuss the results further and then conclude the paper in Sec. IX.

The paper is accompanied by two Appendixes. In Appendix A we provide the details of the fitting methods that we have used to extract the coefficient of the logarithm. In Appendix B we provided the exact Shannon entropy of the models that we considered for different sizes.
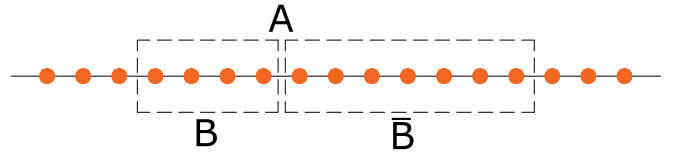

FIG. 1. The setup used to calculate $I_\alpha(\ell)$. Here B and $\bar{B}$ have sizes ℓ and $L - \ell$, respectively.

II. SETUP AND DEFINITIONS

In this section we present the basic definitions and the setup of the problem. The quantities of interest, Rényi and Shannon entropies, are defined as follows: Consider the normalized ground state of a quantum chain Hamiltonian, i.e., $|g\rangle = \sum_I a_I |I\rangle$, expressed in a particular local bases $|I\rangle = |i_1, i_2, \ldots, i_N\rangle$, where N is the system size and $i_1, i_2, \ldots, i_N$ are the eigenvalues of some local operators defined on the lattice sites. The Rényi and Shannon entropies of the total system with size N are defined as

$$Re_\alpha(N) = \frac{1}{1-\alpha} \ln \sum_I P_I^\alpha, \tag{1}$$

$$Sh(N) = -\sum_I P_I \ln P_I, \tag{2}$$

where $P_I = |a_I|^2$ is the probability of finding the system in the particular configuration given by $|I\rangle$. These probabilities are dubbed as formation probabilities in [13]. In the above definition α can be any positive real number. Note that $\alpha \to 1$ gives us just the Shannon entropy.

By considering local bases it is always possible to decompose the configurations as a combination of the configurations inside and outside of a subregion A as $|I\rangle = |I_A I_{\bar{A}}\rangle$, where I_A and $I_{\bar{A}}$ are the subconfigurations corresponding to A and $\bar{A}$. Then, one can define the marginal probabilities as $p_{I_A} = \sum_{I_{\bar{A}}} P_{I_A I_{\bar{A}}}$. Using these probabilities one can now define the Rényi and Shannon entropies of the subsystem with size L as follows:

$$Re_\alpha(L) = \frac{1}{1-\alpha} \ln \sum_{I_A} p_{I_A}^\alpha, \tag{3}$$

$$Sh(L) = -\sum_{I_A} p_{I_A} \ln p_{I_A}. \tag{4}$$

The above two quantities at the critical point normally behave as

$$Re_\alpha(L) = a_\alpha L + x_\alpha \ln L + \mathcal{O}(1), \tag{5}$$

$$Sh(L) = a_1 L + x_1 \ln L + \mathcal{O}(1). \tag{6}$$

The quantities of interest in this paper are x_α and x_1. To isolate these quantities one can divide region A into two subsystems B and $\bar{B}$ with sizes ℓ and $L - \ell$, respectively, see Fig. 1. Then one can define

$$I_\alpha(\ell) = Re_\alpha(\ell) + Re_\alpha(L - \ell) - Re_\alpha(L), \tag{7}$$

$$I_1(\ell) = Sh(\ell) + Sh(L - \ell) - Sh(L). \tag{8}$$

In the rest of paper we consider the case $\ell = \frac{L}{2}$. Then we expect

$$I_\alpha\left(\frac{L}{2}\right) = x_\alpha \ln L + \mathcal{O}(1), \tag{9}$$

$$I_1\left(\frac{L}{2}\right) = x_1 \ln L + \mathcal{O}(1). \tag{10}$$

We calculate the above two quantities for different infinite size, i.e., $N \to \infty$, critical systems. The advantage of this setup is that we are free from the finite size effects of the total system and just bounded with the limitations coming from the subsystem size itself.

III. MODELS AND METHODS OF CALCULATION

In this section we first define our Hamiltonian of interest and then present the formulas that one can use to calculate the formation probabilities and ultimately the I_α. Here we follow the notation in [23].

The Hamiltonian of the most general translational invariant (periodic) quadratic fermionic chain with time-reversal symmetry takes the form

$$H = \sum_{r=-R}^{R} \sum_{j \in \Lambda}^{N} \left[A_r c_j^\dagger c_{j+r} + \frac{B_r}{2} (c_j^\dagger c_{j+r}^\dagger - c_j c_{j+r}) \right] + \text{const.}, \tag{11}$$

with the local fermionic modes c_j, $c_j^\dagger$ and the parameters $A_r = A_{N-r}$, $B_r = -B_{N-r}$, and Λ represent the sites of the lattice. The above Hamiltonian can be exactly diagonalized after going to the Fourier space and Bogoliubov transformation as follows:

$$H = \sum_k |f(e^{ik})| \eta_k^\dagger \eta_k + \text{const.}, \tag{12}$$

where we have defined

$$\eta_k = \frac{1}{2}\left(1 + \frac{f(e^{ik})}{|f(e^{ik})|}\right) c_k^\dagger + \frac{1}{2}\left(1 - \frac{f(e^{ik})}{|f(e^{ik})|}\right) c_{-k}, \tag{13}$$

and the sum over k goes over momenta $k_n = 2\pi n/N$. All the information about the couplings are in the complex function $f(z)$ which we defined as

$$f(z) := \sum_m t_m z^m, \tag{14}$$

where we have

$$A_r = -\frac{t_r + t_{-r}}{2}, \tag{15}$$

$$B_r = -\frac{t_r - t_{-r}}{2}. \tag{16}$$

In this system the vacuum is defined as $\eta_k|0\rangle = 0$ for $\forall k$. When $B_r \neq 0$ the vacuum state is the ground state, while when $B_r = 0$ the Hamiltonian has $U(1)$ symmetry which means the particle number is conserved. In this case one needs to fill the negative modes depending on the number of particles in the system to reach the ground state.

A. Formation probabilities

Before concentrating on critical models explicitly, since all the forthcoming calculations are based on the correlation

matrix, we briefly define it here. The correlation matrix **G** for the eigenstates is defined using two Majorana fermionic operators $\gamma_j \equiv c_j^\dagger + c_j$ and $\bar{\gamma}_j \equiv i(c_j^\dagger - c_j)$ as follows:

$$iG_{jk} = \langle g|\bar{\gamma}_j \gamma_k|g\rangle. \tag{17}$$

One can use the above matrix to calculate all the observables in this system. For example, in this system the formation probabilities are defined as follows: consider the ground state of the system written in configuration basis. That means each configuration of fermions can appear with a particular probability in the ground state. These probabilities can be calculated using the following formula [13,17]:

$$p(C) = \det\left(\frac{\mathbf{I} - \mathbf{I}_c.\mathbf{G}}{2}\right), \tag{18}$$

where **I** is an identity matrix and $\mathbf{I}_c$ is a diagonal matrix made out of ± 1. We set its diagonal element to -1 when we have a fermion and $+1$ when there is no fermion at the corresponding site. When $\mathbf{I}_c = \mathbf{I}$ the corresponding probability is called emptiness formation probability. The above formula works for the full(sub) system if one takes the **G** matrix of the full(sub) system. It also works for disjoint intervals as far as one takes the **G** matrix of the subsystem. Using the determinant properties it is easy to show that the set of formation probabilities is the same for the matrices **G**, $-\mathbf{G}$, $\mathbf{G}^T$, and $-\mathbf{G}^T$. In other words, although the associated probabilities for different configurations might change, the whole set is the same. Even more generally the matrices $\mathbf{I}_c.\mathbf{G}$, $\mathbf{G}.\mathbf{I}_c$, and $\mathbf{I}_c.\mathbf{G}.\mathbf{I}_c$ have the same set of formation probabilities and consequently the same Shannon and Rényi entropies. To summarize, different models with different correlation matrices might have the same Rényi entropies.

B. Critical systems

It is known that when the complex function $f(z)$ has zeros on the unit circle the ground state is critical and depending on the number of zeros one can have different universality classes with different central charges, for a review see [25]. The reason behind this fact is that when $f(z)$ has zeros on the unit circle one can linearize the dispersion relation $|f(e^{ik})|$ around that momentum and get one gapless Majorana fermion. This Majorana fermion contributes $c = \frac{1}{2}$ to the central charge of the system so that we finally have $c = \frac{N_l}{2}$, where N_l is the number of zeros on the unit circle.

In general, one can think about two types of critical systems, those with $U(1)$ symmetry and those without this symmetry. In this work we show that the behavior of the x_α is very much dependent on the presence or absence of $U(1)$ symmetry. Because of that we will study these two cases separately.

1. Models with U(1) symmetry

In these Hamiltonians we have $B_r = 0$. Good examples of these types of Hamiltonians are the ones with the following $f(z)$ function:

$$f_z(n) = -(z^n + z^{-n}), \tag{19}$$

which corresponds to $U(1)$-symmetric n-step hopping fermions. It has the central charge $c = n$. The $n = 1$ is the

celebrated simple hopping chain. For the half-filling case the correlation matrix of the ground state is

$$\mathbf{G} = 2\mathbf{C} - \mathbf{I}, \tag{20}$$

$$C_{jk} = \frac{1}{\pi(j-k)} \sum_{m=1}^{n} (-1)^{m+n} \sin\left[\frac{\pi(2m-1)(j-k)}{2n}\right]. \tag{21}$$

The diagonal elements can be found by taking the limit. In principle, it is possible to consider more complicated models such as $f_z(\{a_n\}) = \sum_n a_n f_z(n)$. The central charge is dependent again on the number of points where one can linearize the dispersion relation and very much depend on the constants a_n. For example, consider the case $f_z(\{a_1, a_2\}) = a_1 f_z(1) + a_2 f_z(2)$. For $|a_1| \geqslant |a_2|$ we have just two points to linearize the dispersion relation and we expect $c = 1$, however, for $|a_2| > |a_1|$ we have four points to linearize so we expect $c = 2$. The **C** matrix in this case can be written as

$$C_{jk}(a_1, a_2)$$
$$= \frac{1}{\pi(j-k)} \{\sin[k_1^*(j-k)] - \sin[k_2^*(j-k)]\}; \tag{22}$$

where $k_1^* \geqslant k_2^*$ are the solutions of the equation $a_1 \cos[k] + a_2 \cos[2k] = 0$ in the range $(0, \pi)$. The diagonal elements can again be found by taking the limit.

In this work we will study $f_z(1)$, $f_z(2)$, and $f_z(\{a_1, a_2\})$ with $(a_1, a_2) = \{(1, 1), (1, 2)\}$, and make a general statement about the behavior of x_α.

2. Models without U(1) symmetry

A fairly general form of $f(z)$ with zeros on the unit circle can be written as

$$f_z(N_0, m_+, m_-, \{m_j\}; \{k_j\})$$
$$= z^{N_0} g(z)(z-1)^{m_+}(z+1)^{m_-} \prod_{j=1}^{N_c} (z - e^{ik_j})^{m_j}(z - e^{-ik_j})^{m_j}, \tag{23}$$

where $g(z)$ is a polynomial without any zeros on the unit circle or origin and $k_1 < k_2 < \cdots < k_{N_c}$. Note that since we have Hamiltonians with real couplings, all zeros are either real or come in complex conjugate pairs and all the powers are integers. For simplicity we just consider $g(z) = 1$. At this moment we assume that N_0 can be a positive or negative integer number. In addition, m_+, m_-, m_j are non-negative integer numbers. The correlation matrix of the ground state for this model is shown to be [23]

$$G_{nm} = \frac{1}{2\pi} \int_0^{2\pi} \frac{f(e^{ik})}{|f(e^{ik})|} e^{-i(m-n)k} dk. \tag{24}$$

Remarkably, the above integral can be calculated explicitly. The result for $g(z) = 1$ can be written with respect to elementary functions as follows:

$$G_{nm} = \begin{cases} \frac{4}{\pi} \frac{\frac{1}{2}G_{nm}^{Re} + G_{nm}^{Im}}{N_0 + M + Q + n - m}, & N_0 + M + Q \neq m - n, \\ \left(2\{\frac{m_+ + 1}{2}\}(-1)^{[\frac{m_+}{2}]}\right)\left((-1)^M + \frac{4}{\pi} \sum_{j=1}^{N_c} (-1)^{j-1}\{\frac{m_j}{2}\}k_j\right), & N_0 + M + Q = m - n, \end{cases} \tag{25}$$

where we have

$$G_{nm}^{Re} = (-1)^{[\frac{m_+}{2}]+1}\left(\left\{\frac{m_+}{2}\right\} - \left\{\frac{m_-}{2}\right\}(-1)^{[Q]+N_0}(-1)^{n+m}\right),$$

$$G_{nm}^{Im} = \sum_{j=1}^{N_c} (-1)^{j-1}\left\{\frac{m_j}{2}\right\} \sin\left(\frac{\pi m_+}{2} + (N_0 + M + Q + n - m)k_j\right), \tag{26}$$

and $M = \sum_{j=1}^{N_c} m_j$, $Q = \frac{m_+ + m_-}{2}$, and $\{X\}$ is defined as follows:

$$\{X\} := |X| - [|X|]. \tag{27}$$

Using the above equation we find the following duality:

$$\mathbf{G}[f_z(N_0, m_+, m_-, \{m_j\}; \{k_j\})] = (-1)^{[\frac{m_+}{2}]}\mathbf{G}[f_z(N_0 + M[m_+, m_-, \{m_j\}], h[m_+], h[m_-], \{h[m_j]\}; \{k_j\})], \tag{28}$$

where

$$M[m_+, m_-, \{m_j\}] = \left[\frac{m_+}{2}\right] + \left[\frac{m_-}{2}\right] + 2\sum_{j=1}^{N_c}\left[\frac{m_j}{2}\right] \tag{29}$$

and $h[x] = x - 2[\frac{x}{2}]$. For later use it is also useful to define

$$H[m_+, m_-, \{m_j\}] = h[m_+] + h[m_-] + 2\sum_{j=1}^{N_c} h[m_j]. \tag{30}$$

TABLE I. Coefficient of the logarithm in the Shannon entropy for different models with $U(1)$ symmetry.

		Models with $U(1)$ symmetry		
$f(z)$	$f_z(1)$	$f_z(2)$	$f_z(\{1,1\})$	$f_z(\{1,2\})$
$8x_1$	0.9968 ± 0.0001	2.00 ± 0.01	1.003 ± 0.003	1.84 ± 0.30

To the best of our knowledge, the above duality has not been discussed before in the literature. It means that when $m_+, m_-, \{m_j\}$ are bigger than one it is possible to absorb them to N_0 and remain with just one or zero powers for $m_+, m_-, \{m_j\}$. The immediate consequence of the above argument is that the even powers are noncritical and do not contribute to the central charge and the contribution of odd numbers is all the same. In other words, we have the following theorem for the central charge:

$$c[N_0, m_+, m_-, \{m_j\}; \{k_j\}] = \tfrac{1}{2}H[m_+, m_-, \{m_j\}]. \quad (31)$$

Note that we assume that the models with $c = 0$ are noncritical. In other words, all the models with $m_+, m_-, \{m_j\}$ even integer numbers are noncritical. From now on without losing any generality we consider that $m_+, m_-, \{m_j\}$ are either zero or one and not all of them are zero. One can also prove another useful duality:

$$\mathbf{G}[f_z(N_0, m_+, m_-, \{m_j\}; \{k_j\})]$$
$$= -\mathbf{G}^T[f_z(-N_0 - m_+ - m_-$$
$$-2\sum_{j=1}^{N_c} m_j, m_+, m_-, \{m_j\}; \{k_j\})]. \quad (32)$$

Combining Eqs. (28) and (32) one can conclude that without losing generality it is possible to assume that $m_+, m_-, \{m_j\}$ are either zero or one and N_0 is an integer number. To make the classification manageable and under control we will just consider the case $N_0 = 0$.

IV. SUMMARY OF RESULTS

In this section we will summarize our main results. We first discuss the case of the systems with $U(1)$ symmetry and then discuss the models without this symmetry.

A. Models with $U(1)$ symmetry

Our extensive numerical results support the following behavior for the coefficient of the logarithm:

$$x_\alpha = \begin{cases} \frac{c}{8}, & \alpha \leqslant 4, \\ \frac{\alpha}{\alpha-1} \frac{c}{8}, & \alpha > 4. \end{cases} \quad (33)$$

The case of $c = 1$ in a different geometry has been already discussed in [10]. The presence of the discontinuity at $\alpha = 4$ is attributed to the least irrelevant operator in the Luttinger liquid description of the model. As far as $\alpha < 4$ it was argued in [10] that this operator is irrelevant and one can get $x_\alpha = \frac{1}{8}$ by Luttinger model arguments. However, when $\alpha > 4$ this operator is relevant and consequently the field gets locked into one of the minima of the potential and just one of the configurations end up to have the largest contribution. Consequently, we have $x_\alpha = \frac{\alpha}{\alpha-1} \frac{c}{8}$. It seems this picture is

more general and valid for generic $f_z(n)$ models. For $n = 1$ a simple numerical investigation shows that the dominant configurations at $\alpha \to \infty$ are $|0, 1, 0, 1, \dots, 0, 1\rangle$ and $|1, 0, 1, 0, \dots, 1, 0\rangle$ consistent with the half-filling ground state. It is possible to calculate the logarithm of the probability of this configuration exactly and one finds [19] a linear term plus a logarithmic subleading term with coefficient $-\frac{1}{8}$. This result proves Eq. (33) at $\alpha \to \infty$ for $n = 1$. A simple numerical investigation shows that the largest probability for $f_z(n)$ models is attributed to the $2n$ configurations $|A_n\rangle =$

$$|\overbrace{0, 0, \dots, 0}^{n-r}\overbrace{1, 1, \dots, 1}^{n}, \dots, \overbrace{0, 0, \dots, 0}^{n}\overbrace{1, 1, \dots, 1}^{r}\rangle \quad \text{and}$$

$$|A_n\rangle = |\overbrace{1, 1, \dots, 1}^{n-r}\overbrace{0, 0, \dots, 0}^{n}, \dots, \overbrace{1, 1, \dots, 1}^{n}\overbrace{0, 0, \dots, 0}^{r}\rangle,$$

where $r = 0, 1, \dots, n-1$. We conjecture that

$$-\ln p(A_n) = a(n)L + \frac{n}{8}\ln L + \mathcal{O}(1). \quad (34)$$

It should be possible to prove the above conjecture using the methods developed in [19], however, we do not attempt to do that in this paper. We note that when the system is not half-filling a similar picture is still valid but the most relevant configuration can change. For example, for $\frac{r}{s}$ filling in $n = 1$ case the most important configuration is $|A_1(\frac{r}{s})\rangle =$

$$|\overbrace{0, 0, \dots, 0}^{s-r}\overbrace{1, 1, \dots, 1}^{r}, \dots, \overbrace{0, 0, \dots, 0}^{s-r}\rangle \text{ in which the numer-}$$

ical results show that the logarithm of the probability of this configuration has also a linear term plus logarithmic correction with coefficient $-\frac{1}{8}$, see [13].

Finally, we also found that Eq. (33) is most probably also valid for the models $f_z(\{a_n\}) = \sum a_n f_z(n)$. The numerical results in these cases have strong oscillations and consequently the estimation for x_α is poor. However, the overall behavior of the numerical results is consistent with Eq. (33).

The coefficient of the logarithm for all the considered models is summarized in Table I.

B. Models without $U(1)$ symmetry

For models with $N_0 = 0$ our numerical results done on many examples reveal the following behavior:

$$x_\alpha = \begin{cases} \frac{b(\alpha)}{8}, & \alpha \leqslant 1, \\ \frac{\alpha}{\alpha-1} \frac{c}{8}, & \alpha > 1, \end{cases} \quad (35)$$

where $b(\alpha) = \mathfrak{b}(\alpha)H[m_+, m_-, \{m_j\}]$ and $c = \frac{1}{2}H[m_+, m_-, \{m_j\}]$. The coefficient of the logarithm seems to be again increasing based on the number of gapless Majorana fermions that one can define for the model. This is reminiscent of the behavior of entanglement entropy in these systems [22]. However, for $\alpha \leqslant 1$ the coefficient is not exactly proportional to the central charge. For $\alpha = 1$ we have $\mathfrak{b}(1) = 0.480016 \pm 0.00005$ and for $0 < \alpha < 1$ the numerical results indicate a complicated but universal

TABLE II. Coefficient of the logarithm in the Shannon entropy for different models without $U(1)$ symmetry.

| | Models without $U(1)$ symmetry | | | |
$f(z)$	$z-1$	$(z-1)_{\text{BW}}$	z^2-1	z^3-1
$8x_1$	0.48008 ± 0.00002	0.48009 ± 0.00003	0.9616 ± 0.0001	1.4465 ± 0.0008

behavior, see [10] for the Ising chain. There are also regions where this coefficient is negative. For all these models the most relevant configuration is the configuration without any fermion, i.e., $|E\rangle = |0, 0, \ldots, 0\rangle$ or the one full of fermions $|E\rangle = |1, 1, \ldots, 1\rangle$. One can understand this by calculating $\langle I|H|I\rangle$ for different configurations I. An easy calculation shows that $\langle I|H|I\rangle = nA_0$, where n is the number of fermions in the configuration. It is now easy to see that depending on the sign of the A_0 just the configuration without any fermion or the one full of fermions have the lowest energies. For the subsystem configurations numerical calculations support the above argument. Note that just changing the sign of the **G** matrix interchanges the probability of the two configurations, however the set of the configurations is intact. We will be rarely concerned with this sign. The corresponding probability is called emptiness formation probability and one can calculate it explicitly using the Fisher-Hartwig formula, see [16] for the $c = \frac{1}{2}$ case. In the most general case we find

$$-\ln p(E) = a(m_+, m_-, \{m_j\}; \{k_j\})L + \frac{c}{8}\ln L + \mathcal{O}(1), \quad (36)$$

where $a(m_+, m_-, \{m_j\}; \{k_j\}) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\ln\frac{1}{2}(1\mp\frac{f(e^{ik})}{|f(e^{ik})|})$ and again we have $c = \frac{1}{2}H[m_+, m_-, \{m_j\}]$. This proves Eq. (35) for $\alpha \to \infty$. However, as it is argued already for the Ising chain in [10] it is not clear why the discontinuity in x_α should start exactly at $\alpha = 1$ in all of these models.

In Table II we summarize the coefficient of the logarithm for all the models where we did comprehensive numerical checks.

V. NUMERICAL AND FITTING PROCEDURE

In this section we briefly discuss our numerical and fitting procedures. The more comprehensive details are relegated to Appendix A.

In all of the considered models we first find the 2^L number of probabilities using Eq. (18). The largest size that we considered was $L_{\max} = 42$. After collecting all the probabilities we calculate the I_α and find the best estimate of x_α using different fitting procedures. Most importantly our fitting function is

$$I = A_0 + A_1 \log L + A_2 L^{-1}\log L + \sum_{i=3}^{m} A_i/L^{i-2}. \quad (37)$$

However, there are at least two important challenges to overcome. First of all, due to the limitation in the maximum size of L we need to use some extrapolation methods to get a good estimate of x_α. The second important hurdle is that I_α for some of the models show strong oscillations. In these cases either one needs to stick to a particular branch or average the estimated x_α over all the branches. The more sophisticated approach is to use the regularization method. We have tried all of these possibilities and in each case we report the one

with the best fit possible. In Appendix A we also explain in detail our methods to estimate the error bars in each case.

Because of the exponential nature of the calculations and the number of considered models computing all the probabilities required a quite long time, which is particularly notable for larger system sizes. As an example in the case of $L = 42$, it took about 3 days to generate 2^{42} formation probabilities using a cluster with 356 computing nodes, where each node had 16 cores. To prevent further damage to the environment in Appendix B we collected the Shannon entropy for the models that we considered so that the motivated reader can reproduce the coefficient of the logarithm by her(him)self.

VI. DETAILS OF THE ANALYSIS

In this section we will provide the details of the models that we considered. We first discuss systems with $U(1)$ symmetry and later we discuss the ones without this symmetry.

A. Models with $U(1)$ symmetry

We first considered the model $f_z(1)$ which is the simple hopping model with half-filling. The results for I_α with $\alpha = 1, 6$ are shown in Fig. 2. The results for $\alpha > 1$ have oscillations which get stronger by increasing α. To calculate the coefficient of the logarithm in these cases we first calculated the coefficient for each branch using an extrapolation method and later we averaged over the two results. The results are shown in Fig. 3 which is compatible with Eq. (33). We then considered the model $f_z(2)$. The results for I_α with $\alpha = 1, 6$ are shown in Fig. 4. There are stronger oscillations in this case. There are four visible branches for $\alpha > 1$. In these cases again we calculated the coefficient for each branch and if needed
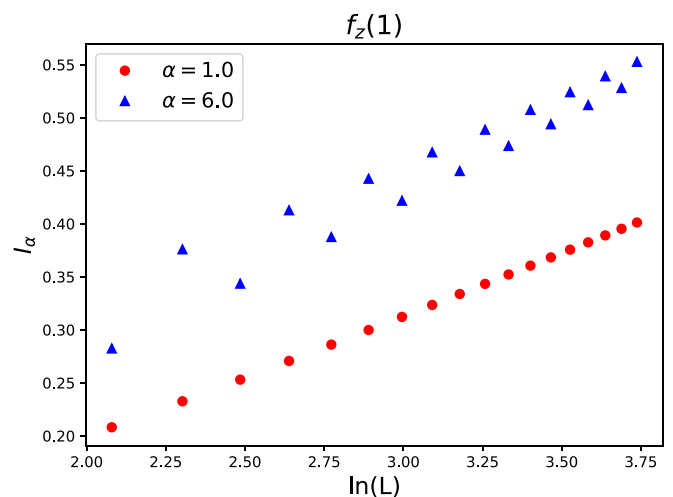

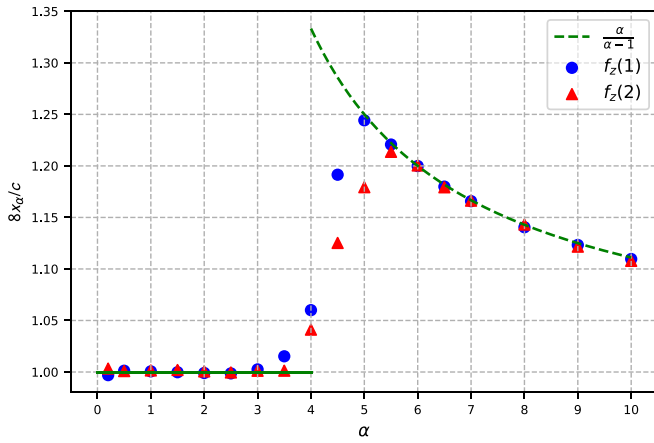FIG. 2. I_α with respect to $\ln L$ for $f_z(1)$ for two indices $\alpha = 1$ and $\alpha = 6$.

FIG. 3. The coefficient of the logarithm with respect to α for the two models $f_z(1)$ and $f_z(2)$. In all the calculations $L_{\max} = 42$.

we also used the regularization method as it is explained in Appendix A. Finally, we averaged over all the branches. The coefficient x_α with respect to α is shown again in Fig. 3.

We also considered the models $f_z(\{a_1, a_2\})$ with $(a_1, a_2) \in \{(1, 1), (1, 2)\}$. The numerical results have strong oscillations especially for the case $(a_1, a_2) = \{(1, 2)\}$. In this case for large α's it seems impossible to get a good estimate for the x_α with sizes up to $L = 42$. However, the general picture is consistent with Eq. (33). In Appendix B we just report the results for the Shannon entropy and do not show the details for the other α's.

B. Models without $U(1)$ symmetry

In this section we will provide some details regarding the models without $U(1)$ symmetry.

The first example is the famous Ising chain with $f(z) = z - 1$. The results for I_α with $\alpha = 1, 2$ are shown in Fig. 5. We do not see any oscillations for any α. To calculate the coefficient of the logarithm we used the extrapolation method explained in Appendix A. The maximum size of the subsys-
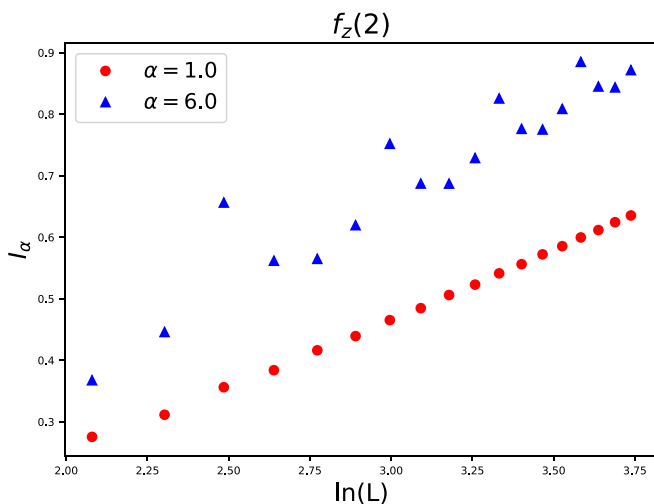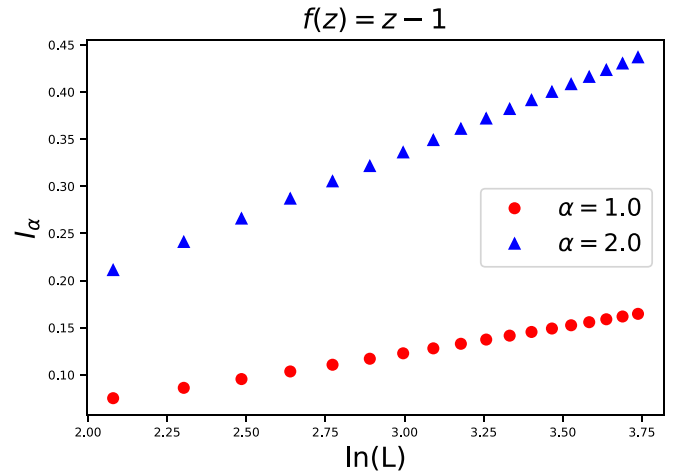

FIG. 5. I_α with respect to $\ln L$ for $f(z) = z - 1$ for two indices $\alpha = 1$ and $\alpha = 2$.

tem that we considered was $L_{\max} = 42$. The results for $\alpha \geqslant 1$ are shown in Fig. 6 which is consistent with Eq. (35). It is worth mentioning that we also analyzed the $f(z) = z + 1$ which although has a different $\mathbf{G}$ matrix the set of formation probabilities are exactly the same as the Ising chain.

The second example is $f(z) = z^2 - 1$ which is a model with central charge $c = 1$. The results for I_α with $\alpha = 1, 2$ are shown in Fig. 7. There are small oscillations for $\alpha > 1$ which are just detectable after careful numerical manipulations, see Fig. 7 inset. To calculate the coefficient of the logarithm we again separated different branches and used the extrapolation method for each branch and then finally averaged over the two branches. The results for $\alpha \geqslant 1$ are shown in Fig. 6 which is consistent with Eq. (35). Note that although the central charge here is an integer number because of lack of $U(1)$ symmetry we end up with a result which resembles the one we obtained for the Ising chain.

We also analyzed other models such as $f(z) = (z - e^{i\theta})(z - e^{-i\theta})$ with different θ's. They all have $c = 1$ and show similar structure. We realized that when θ is small or

FIG. 4. I_α with respect to $\ln L$ for $f_z(2)$ for two indices $\alpha = 1$ and $\alpha = 6$.

FIG. 6. The coefficient of the logarithm with respect to α for the three models $f(z) = z - 1$, $f(z) = z^2 - 1$, and $f(z) = z^3 - 1$. In all the calculations $L_{\max} = 42$.
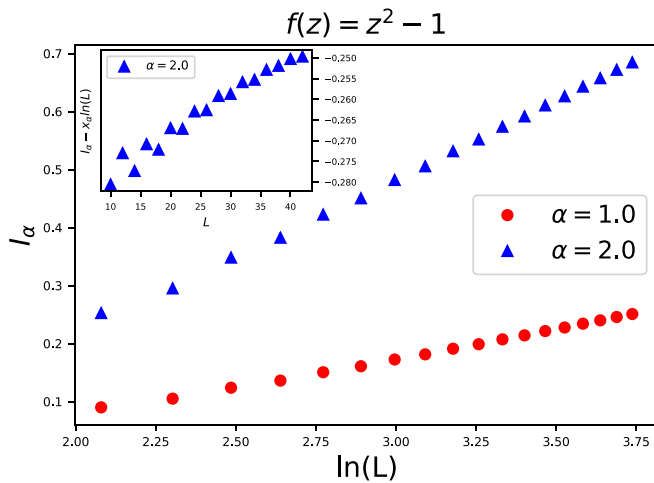
FIG. 7. I_α with respect to $\ln L$ for $f(z) = z^2 - 1$ for two indices $\alpha = 1$ and $\alpha = 2$. The inset shows there are oscillations with period two.

close to π the oscillations for $\alpha > 1$ are stronger. The fewest oscillations appear for $\theta = \frac{\pi}{2}$ which have the same set of probabilities as $f(z) = z^2 - 1$.

The third example is $f(z) = z^3 - 1$ which is a model with central charge $c = \frac{3}{2}$. The results for I_α with $\alpha = 1, 2$ are shown in Fig. 8. Similar to the previous case we have small oscillations. There are three branches and we followed the same procedure as before to estimate the coefficient of the logarithm. The results for $\alpha \geqslant 1$ are shown in Fig. 6 which is again consistent with Eq. (35).

We also considered other models with a similar central charge such as $f(z) = (z - 1)(z - e^{i\theta})(z - e^{-i\theta})$ with different θ's. The result are the same as before. However, we realized that the case $\theta = \frac{2\pi}{3}i$ has the fewest oscillations. When we decrease or increase θ the oscillations get stronger. Similar phenomena happens also for $f(z) = (z + 1)(z - e^{i\theta})(z - e^{-i\theta})$. When the zeros have the largest distance from each other the oscillations are smallest and when

FIG. 8. I_α with respect to $\ln L$ for $f(z) = z^3 - 1$ for two indices $\alpha = 1$ and $\alpha = 2$. The inset shows there are oscillations with period three.

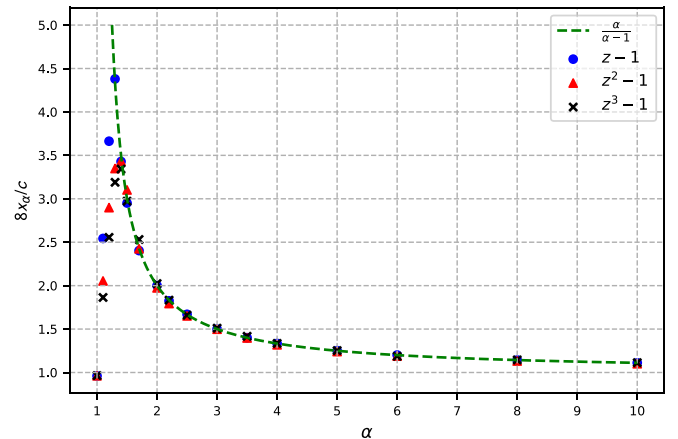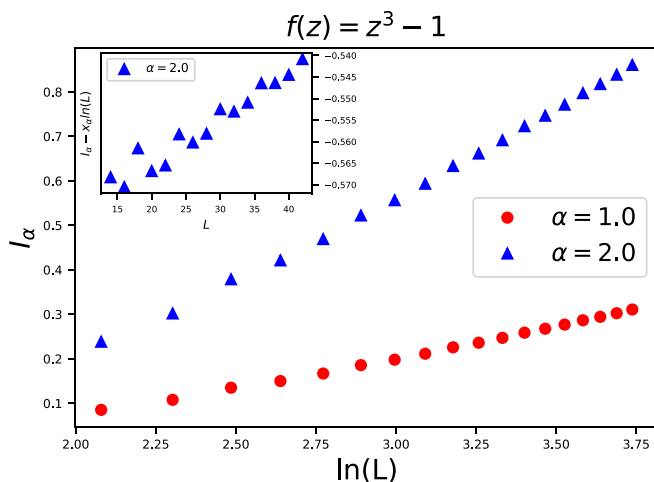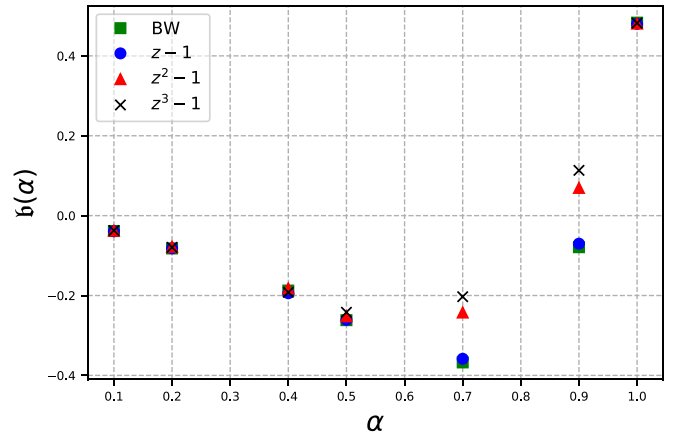
FIG. 9. The coefficient of the logarithm with respect to α for the three models $f(z) = z - 1$, $f(z) = z^2 - 1$, and $f(z) = z^3 - 1$. In all the calculations $L_{max} = 36$.

two or three of them get closer to each other we have stronger oscillations. This is a numerical observation for which we do not have a good explanation.

Apart from the above case we also considered $f(z) = (z - 1)(z + 1)(z - e^{i\theta})(z - e^{-i\theta})$ with central charge $c = 2$ with again similar conclusions. $f(z) = z^4 - 1$ has the least oscillations. The last model we considered was $f(z) = z^5 - 1$ with the central charge $c = \frac{5}{2}$. The results are consistent with Eq. (35). In most of the cases where we do not report the results here we considered $L_{max} = 36$. In some cases such as $f(z) = z^5 - 1$ we pushed the results up to $L_{max} = 42$.

In all of the above cases we also studied with the same procedure I_α with $\alpha < 1$. For some α's the coefficient of the logarithm is negative but the numerical results confirm that the behavior of the coefficient is universal and proportional to the number of points where one can linearize the dispersion relation. The results were depicted in Fig. 9. Here we considered $L_{max} = 36$. We think the visible discrepancy in the region $\alpha \in (0.6, 1)$ is due to the finite size effect which for an unknown reason to us is stronger in this interval.

VII. ANALYSIS BASED ON BISOGNANO-WICHMANN REDUCED DENSITY MATRIX

The reduced density matrix (RDM) of a quantum system ρ_A is fully encoded in the modular (or entanglement) Hamiltonian H_A defined as

$$\rho_A = \frac{e^{-H_A}}{Z_A}, \quad Z_A = \text{tr}_A e^{-H_A}. \tag{38}$$

By construction, the RDM and the modular Hamiltonian have the same eigenvectors, and their eigenvalues are simply related. The modular Hamiltonian plays a key role in quantum field theory [26]. In this context, the modular Hamiltonian of half-space partition is known to be related to the boost operator [27,28]. Its form in conformal field theory (CFT) is also known explicitly [29,30]. However, its explicit functional form in lattice models is known only in a few simple cases, see for example [31–36]. It was proposed in Refs. [37,38] to use the Bisognano-Wichmann (BW) theorem in quantum field theory and its extension in conformal field theory (CFT) to

write approximate modular Hamiltonians for lattice models. From the BW modular Hamiltonian one can construct a RDM, which has been dubbed BW RDM. The proposal has been checked extensively [36–41], showing that in many cases the BW modular Hamiltonian can reproduce to a good precision the entanglement spectrum, correlation functions and entanglement entropy. In [41] it was shown that this approximation also produces very good approximations of the formation probabilities. For a recent comprehensive review see [42]

Since the BW modular Hamiltonian is a discretization of the quantum field theory itself one might hope that the convergence of many quantities to the actual field theory result might be faster and better. Having this in mind we used the BW of the Ising model [36] to find first the $\mathbf{G}_{BW}$ matrix as follows: we first make the following T matrix:

$$e\begin{pmatrix} \mathbf{M} & \mathbf{N} \\ -\mathbf{N} & -\mathbf{M} \end{pmatrix} = \begin{pmatrix} \mathbf{T}_{11} & \mathbf{T}_{12} \\ \mathbf{T}_{21} & \mathbf{T}_{22} \end{pmatrix}, \tag{39}$$

where the $\mathbf{M}$ and $\mathbf{N}$ are $L \times L$ matrices (L is the size of subsystem) with the following elements:

$$\frac{1}{\pi} N_{lm} = \lambda(l)\delta_{l+1,m} - \lambda(m)\delta_{l,m+1}, \tag{40}$$

$$\frac{1}{\pi} M_{l,m} = \lambda(l)\delta_{l+1,m} + \lambda(m)\delta_{l,m+1} + 2\lambda\left(l - \frac{1}{2}\right)\delta_{l,m}, \tag{41}$$

where $\lambda(n) = \frac{n(L-n)}{L}$ and L is the size of the subsystem, i.e., $L = 1, 2, 3, \ldots$. Then for $\mathbf{G}$ matrix we have

$$\mathbf{G}_{BW} = \frac{\mathbf{F} - I}{\mathbf{F} + I}, \tag{42}$$

where $\mathbf{F} = \mathbf{T}_{22}^{-1} + \mathbf{T}_{12}.\mathbf{T}_{22}^{-1}$. One can use the above correlation matrix to produce formation probabilities and consequently the Rényi (Shannon) entropy. To calculate $I_\alpha(l)$ one needs to take care of a subtlety. The BW reduced density matrix is not an exact reduced density matrix. That means $\mathrm{tr}_{\bar{B}}\rho_A^{\mathrm{BW}} \neq \rho_B^{\mathrm{BW}}$. In other words, probabilities coming from $\mathrm{tr}_{\bar{B}}\rho_A^{\mathrm{BW}}$ and ρ_B^{BW} are different. We realized that the best results for $I_\alpha(l)$ can be derived by using marginal probabilities of ρ_L^{BW}. The results for the Rényi entropy is indistinguishable from the exact results when depicted in the figure so we just report the Shannon entropy in this case in Appendix B. This is an interesting demonstration of the universality of the coefficient of the logarithm and also the power of the approximate BW modular Hamiltonian.

VIII. DISCUSSION

In this paper we considered an infinite system and calculated the coefficient of the logarithm appearing in the scaling of Rényi (Shannon) entropy of the ground state of critical chains. Instead of working directly in the thermodynamic limit one could take a finite periodic system with size N and calculate $I_\alpha = Re_\alpha(\ell) + Re_\alpha(N - \ell) - Re_\alpha(N)$ which ends up to be proportional to $x'_\alpha \ln \frac{N}{\pi} \sin \frac{\pi\ell}{N}$, see [7,10,11]. We expect that in all of our models $x'_\alpha = x_\alpha$. The same is not true if one takes an open boundary condition as it was already noticed in [10]. This is because the boundary conditions can change the logarithmic subleading term drastically. Clear understanding of the coefficient in open quantum spin chains is still lacking.

All of the models that we considered in this paper can be mapped to quantum spin chains using Jordan-Wigner transformation. Based on the previous numerical calculations [7,11] it seems plausible to assume that if one calculates the Rényi (Shannon) entropy in σ^x basis the result for the coefficient of the logarithm be the same as what we found in this paper. However, as it was already noticed in [11] this might not be correct for other bases. Finally, we should mention that understanding the coefficient of the logarithm for $\alpha \leqslant 1$ in critical systems without $U(1)$ symmetry appears to be a challenge. For $\alpha = 1$ this coefficient seems almost [7] but not exactly [10] proportional to the central charge. It is an open problem to understand why this is the case.

IX. CONCLUSIONS

Rényi (Shannon) entropy of the ground state of quantum chains shows a volume-law behavior. When the system is critical these quantities for the subsystem show a subleading logarithmic term with a coefficient which is universal up to some extent. In this paper we studied these quantities at the critical point of generic time-reversal translational invariant quadratic critical free fermions. We found that there are two different classes of models. Models with $U(1)$ symmetry show a unified behavior. The coefficient is dependent on the number of points one can linearize the dispersion relation. The coefficient is constant up to $\alpha = 4$ and then there is a discontinuity and a nice decay in the form $\frac{\alpha}{\alpha-1}$. In the case of systems without $U(1)$ symmetry we have studied models where the corresponding $f(z)$ function has no zero outside of the unit circle. In these models the coefficient of the logarithm is always proportional to the number of points where one can linearize the dispersion relation. There is a discontinuity at $\alpha = 1$ and for $\alpha > 1$ we again have the $\frac{\alpha}{\alpha-1}$ kind of decay. For $\alpha \leqslant 1$ although the coefficient is still universal the exact α functionality is not known. There are also regions where this coefficient is negative. It would be interesting to generalize the above analysis to models in which $g(z) \neq 1$ and/or $N_0 \neq 0$. Due to numerous possibilities and the existence of strong oscillations in the calculation of the Rényi entropy the complete classification in these cases might not be straightforward. Finally, we also studied the same quantities using the approximate BW modular Hamiltonian and confirmed that the produced set of formation probabilities are very close to the exact ones. The derived coefficient of the logarithm was almost indistinguishable from the exact results. The biggest challenge for the future is probably to calculate analytically the $\mathfrak{b}(\alpha)$ for the second class of models to understand the nature of these numbers.

APPENDIX A: DETAILS OF THE FITTING PROCEDURE

In this Appendix we provide more details regarding fitting procedures that we followed in the main text. Let $\{(L_j, I_j)\}_{j=1}^n$
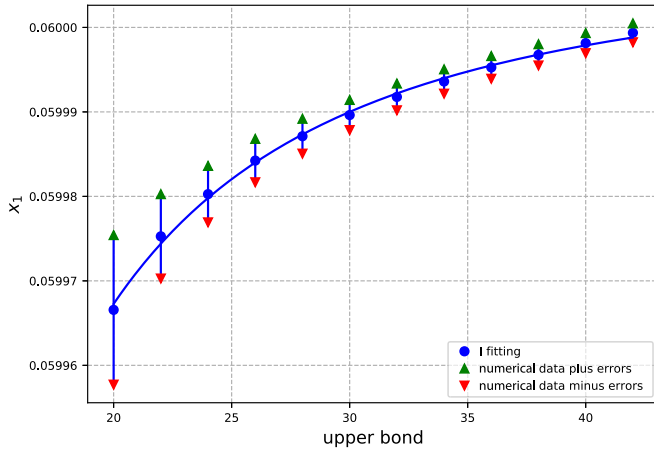
FIG. 10. Blue points are x_α estimations for each size derived using Eq. (A1). Green/red points are x_α plus/minus the error bar of the blue points. To estimate the error bar for x_α, the difference between x_α^∞ associated with the blue, red, and green points are used.

be the set of data points, in which we intend to extract the relevant physical quantities, such as the coefficient of the

logarithm, as the fitting parameters. The fitting functions of interest in this paper are of a factorized form meaning that it is a direct sum of some fitting (known) functions where the fitting parameters are the corresponding prefactors. In other words, we have

$$h(L, \{\theta\}) = \sum_{i=0}^{m} \theta_i f_i(L), \qquad (A1)$$

where m is the number of the fitting terms. In the above $f_i(L)$, $i = 0, \ldots, m$ are the fitting functions assumed to be *a priori* known $[f_0(L) \equiv 1]$ and $\{\theta\}$ shows the set of the fitting parameters, i.e., $\{\theta_0, \theta_1, \ldots, \theta_m\}$. In particular, θ_0 is called the bias. To make a connection with the paper, where the fitting formula is $I = A_0 + A_1 \log L + A_2 L^{-1} \log L + \sum_{i=3}^{m} A_i / L^{i-2}$, we have $\theta_0 \equiv A_0$, $[\theta_1, f_1(L)] = (A_1 \equiv x_\alpha, \log L)$, $[\theta_2, f_2(L)] = (A_2, L^{-1} \log L)$, and $[\theta_i, f_i(L)] = (A_i, L^{-(i-2)})$, $i = 3, \ldots, m$. The χ^2 method is an efficient approach for estimating the best fitting for a given data set, see for example [43]. One defines the χ^2 as follows:

$$\chi^2(\{\theta\}) = \frac{1}{n} \sum_{j=1}^{n} [I_j - h(L_j, \{\theta\})]^2, \qquad (A2)$$

TABLE III. Shannon entropy for different sizes for various models with $U(1)$ symmetry.

	Models with $U(1)$ symmetry			
L	$f_z(1)$	$f_z(2)$	$f_z(1, 1)$	$f_z(1, 2)$
1	0.69314718055994	0.69314718055994	0.63651416829481	0.69221884672917
2	1.30175595835581	1.38629436111989	1.20713991833245	1.38401888518543
3	1.88952473643207	1.99490313891575	1.75382909171979	2.01139838075075
4	2.46572926830832	2.60351191671162	2.29024160464855	2.63699346158892
5	3.03554130502831	3.19128069478788	2.82016064054523	3.25084607202052
6	3.60094114646480	3.77904947286415	3.34574544378459	3.86112826700399
7	4.16329822887889	4.35525400474039	3.86834033023438	4.46758265294401
8	4.72333514587764	4.93145853661664	4.38865348976192	5.07065792051820
9	5.28160174014818	5.50127057333664	4.90719393994777	5.67191793745351
10	5.83843850171108	6.07108261005664	5.42433643749537	6.27020185844070
11	6.39411934070096	6.63648245149308	5.94032544493927	6.86734966048696
12	6.94883095743287	7.20188229292957	6.45535416818048	7.46264647774241
13	7.50272914860883	7.76423937534372	6.96957656772588	8.05706979765335
14	8.05592711783444	8.32659645775783	7.48310440162153	8.65038954334950
15	8.60852182009970	8.88663337475661	7.99603073579124	9.24285216104092
16	9.16058699436123	9.44667029175474	8.50843325729873	9.83467192662273
17	9.71218703560899	10.0049368860264	9.02037209149640	10.4256806854124
18	10.2633726246811	10.5632034802959	9.53189899717722	11.0162682615565
19	10.8141886934818	11.1200402418602	10.0430585231179	11.6061362664382
20	11.3646715586883	11.6768770034228	10.5538866526276	12.1956618677633
21	11.9148538081393	12.2325578424122	11.0644150816116	12.7846447900688
22	12.4647623481134	12.7882386814019	11.5746716896468	13.3732928226244
24	13.5638519339168	13.8976619148654	12.5944598914257	14.5494962592860
26	14.6620982637734	15.0054582972159	13.6134084894882	15.7244854175891
28	15.7596246309596	16.1118542356559	14.6316397420975	16.8984152066857
30	16.8565291237941	17.2170436401148	15.6492512752802	18.0714355315951
32	17.9528910588046	18.3211739881711	16.6663221275571	19.2436688684954
34	19.0487754831082	19.4243740680976	17.6829169075352	20.4152009971075
36	20.1442363964986	20.5267452314426	18.6990893971895	21.5861034172132
38	21.2393191176783	21.6283773787716	19.7148848398415	22.7564395029841
40	22.3340608777584	22.7293430908899	20.7303414022332	23.9262777420356
42	23.4284982110165	23.8297075943315	21.7454918339490	25.0956690768337

which should be minimized with respect to all the fitting parameters $\{\theta\}$ in order to get the best fitting. In our work we mostly worked with another quantity called R value. It is defined as $R^2 \equiv 1 - \frac{\chi^2}{\sigma^2}$, where $\sigma^2 \equiv \frac{1}{n}\sum_{j=1}^{n}(I_j - \bar{I})^2$ and $\bar{I} \equiv \frac{1}{n}\sum_j I_j$. The closer this quantity is to one the better the fit is. When the number of fitting parameters are high, one can use the gradient descent method in which one updates the parameters using the equation

$$\vec{\theta}_i^{\text{new}} = \vec{\theta}_i^{\text{old}} - \eta \frac{\nabla_\theta \chi(L_j, \{\theta\})}{|\nabla_\theta \chi(L_j, \{\theta\})|}\delta\theta. \qquad (A3)$$

where $\delta\theta$ is a discretization parameter, η is the step size, and $\vec{\theta} \equiv (\theta_0, \theta_1, \ldots, \theta_m)$, and $\nabla_\theta \equiv (\frac{\partial}{\partial\theta_0}, \frac{\partial}{\partial\theta_1}, \ldots, \frac{\partial}{\partial\theta_m})$ to get the best fit after reaching the fixed point of the parameter. It is worth mentioning that one should be careful in taking an appropriate number of fitting parameters to avoid the problem of *overfitting*. Normally the sign of overfitting in numerical calculations is the huge and strongly fluctuating numbers for the fitting parameters. To avoid this problem, one adds $\lambda\sum_{i=2}^{m}\theta_i^2$ to the χ^2, where λ is a very small coefficient which prevents the coefficients to take extremely large values. This method is called the *regularization method* and was used in

this paper when necessary. Note that in this work we did not regularize the θ_0 and θ_1.

A more compact representation of Eq. (A2) can be obtained by casting the equation in a matrix form. Let $\mathbf{X}_{ji} \equiv f_i(L_j)$ be a $(n) \times (m+1)$ matrix, and $\Theta \equiv (\theta_0, \theta_1, \ldots, \theta_m)^T$ be a vector with length $m+1$, and $Y \equiv (I_1, I_2, \ldots, I_n)^T$ where $i = 0, 1, \ldots, m$ enumerates the fitting terms, and $j = 1, 2, \ldots, n$ enumerates the data points. Then the regularized χ^2 reads

$$\chi^2(\{\theta\}) = \frac{1}{n}(\mathbf{X}\Theta - Y)^T(\mathbf{X}\Theta - Y) + \lambda\Theta^T\mathbf{I}'\Theta, \qquad (A4)$$

where $\mathbf{I}'$ is a diagonal matrix with zero or one as its diagonal elements. For a fitting parameter that is not going to be regularized, the corresponding diagonal element is zero, and for the other elements, it is one. By minimizing χ^2 with respect to all θ parameters, we obtain

$$\Theta = (\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I}')^{-1}\mathbf{X}^T Y. \qquad (A5)$$

In some cases the matrix $\mathbf{X}^T\mathbf{X}$ has small eigenvalues which leads to very large θ values when λ is zero. This is the reason for introducing the regularization parameter λ.

In our analyses in this paper, we set the diagonal elements of $\mathbf{I}'$ corresponding to the bias and the x_α to zero. In addition, when overfit takes place, we consider a minimal λ value that

TABLE IV. Shannon entropy for different sizes for various models without $U(1)$ symmetry.

L	$z - 1$	$z^2 - 1$	$z^3 - 1$	$z^5 - 1$
1	0.47394663373377	0.47394663373377	0.47394663373377	0.47394663373377
2	0.92544105529219	0.94789326746755	0.94789326746755	0.94789326746755
3	1.36797061201631	1.39938768902597	1.42183990120133	1.42183990120133
4	1.80585459307135	1.85088211058439	1.87333432275975	1.89578653493511
5	2.24088987072848	2.29341166730851	2.32482874431817	2.36973316866888
6	2.67400379724519	2.73594122403263	2.77632316587659	2.82122759022730
7	3.10573474075415	3.17382520508767	3.21885272260071	3.27272201178572
8	3.53642296390859	3.61170918614271	3.66138227932483	3.72421643334414
9	3.96629704662543	4.04674446379984	4.10391183604894	4.17571085490256
10	4.39551790695338	4.48177974145697	4.54179581710399	4.62720527646098
11	4.82420308419465	4.91489366797366	4.97967979815903	5.06973483318511
12	5.25244103454533	5.34800759449037	5.41756377921407	5.51226438990921
13	5.68029998893319	5.77973853799942	5.85259905687121	5.95479394663334
14	6.10783367902435	6.21146948150821	6.28763433452831	6.39732350335750
15	6.53508517170340	6.64215770466274	6.72266961218539	6.83985306008173
16	6.96208951510668	7.07284592781729	7.15578353870211	7.27773704113673
17	7.38887561225378	7.50272001053396	7.58889746521889	7.71562102219158
18	7.81546757783182	7.93259409325053	8.02201139173568	8.15350500324665
19	8.24188574022752	8.36181495357858	8.45374233524467	8.59138898430154
20	8.66814739454094	8.79103581390823	8.88547327875337	9.02927296535642
21	9.09426737732450	9.21972099114686	9.31720422226235	9.46430824301328
22	9.52025851138501	9.64840616838662	9.74789244541657	9.89934352067105
24	10.3718974749895	10.5048820690910	10.6092688917255	10.7694140759866
26	11.2231381653332	11.3605999778672	11.4690170571602	11.6375632801578
28	12.0740383772891	12.2156673580469	12.3281120002033	12.5037911331902
30	12.9246441747402	13.0701703434076	13.1865537208623	13.3700189862121
32	13.7749928645435	13.9241790301996	14.0439240753413	14.2334808732707
34	14.6251150843050	14.7777512244745	14.9008472029367	15.0969427602233
36	15.4750363025723	15.6309351556417	15.7573231036387	15.9593619268318
38	16.3247779202053	16.4786149708125	16.6130410124114	15.8455641809791
40	17.1743580924403	17.3362947859834	17.4684336568878	16.6467809808711
42	18.0237923539904	18.1885347499710	18.3235010370748	18.5387292548529

removes the overfit. To obtain an optimal values for m, we first start with $m = 1$, and find the fittings. Then we increase m by one and repeat the fittings, and check the convergence of the fitting parameters. We continue this procedure, comparing the quality of the fittings with the previous stage, up to a stage where the fittings are optimal.

To calculate the fitting parameters in the scaling limit we used two kinds of extrapolation: the uppermost and the lowermost fixed extrapolation (UFE and LFE, respectively). Suppose that the range of fitting is $[L_{\min}, L_{\max}]$. Then in the UFE (LFE) method we fix $L_{\max}$ ($L_{\min}$) to its maximum (minimum) value and calculate the fitting parameters (especially x_α) in terms of $L_{\min}$ ($L_{\max}$) using the R^2 method. Our observations show that for all cases the resulting $x_\alpha(L)$ follow

$$x_\alpha(L) = x_\alpha^\infty + \frac{a}{L^b}, \tag{A6}$$

where x_α^∞, a, and b are some constants obtained by fitting. and x_α^∞ is the extrapolated parameter that we report in this paper. We note that the UFE is not really the usual extrapolation method because the largest size is actually fixed. However, in most of the cases, since the values for the Rényi entropy for small sizes are not very useful, we found that the UFE normally gives more stable results than the LFE.

Throughout the paper we face models where the Rényi entropy shows some oscillations. In this case we subdivide the data points to k classes where in each class the points are in the same phase of oscillations: $\{\{(L_i^{(q)}, I_i^{(q)})\}_{i=1}^n\}_{q=1}^k$. Then, using the procedure explained above we find the best fits *for each class*, with the resulting fitting parameters $\{\{\theta_p^{(q)}\}_{p=0}^m\}_{q=1}^k$, where q numerates the classes (totally k classes). Then, the average parameters are simply defined as

$$\bar{\theta}_p \equiv \frac{1}{k} \sum_{q=1}^k \theta_p^{(q)}, \tag{A7}$$

for which the corresponding fitting functions are free of oscillations. The mentioned oscillations are stronger in $U(1)$ symmetric models.

Furthermore, one can define an error of estimating coefficients of fitting, using a standard deviation method (SD) [44]. We know that for any set of data with linear behavior we have finite deviations (errors) from Eq. (A1). For n data pairs $\{(L_i, I_i), i = 1, \ldots, n\}$, the underlying relationship between I_i and L_i including this error term ϵ_i can be described as

$$I(L_j) = A_{-1} + x_\alpha \ln L_j + \cdots + \epsilon_j. \tag{A8}$$

After applying a proper fitting method and extracting fitting coefficients we can replace them in Eq. (A8) and estimate error $\epsilon_j = I_i - A_{-1} - x_\alpha \ln L_j - \cdots$. Consequently, the error of estimating coefficient x_α can be written as

$$\epsilon(x_\alpha) = \sqrt{\frac{\sum_j \epsilon_j^2}{(n-2)(\langle z^2 \rangle - \langle z \rangle^2)}}, \tag{A9}$$

where $z = \ln L$ and $\langle a \rangle = \frac{1}{n} \sum_j a_j$.

There are two sources of error in estimating the thermodynamic limit of fitting parameters, like x_α. The first one is the systematic error corresponding to Eq. (A6). The second one

is the error arising from the numerical errors in estimating the function for each system size. See for example Fig. 10, where apart from the systematic error in estimating the final value of I for large $L_{\max}$ values, there is an additional error due to the errors for each $L_{\max}$. In other words, we use Eq. (A6) for three sets of points, the fitting values of x_α coming from Eq. (A1) for different sizes and the same numbers plus/minus their error bars for each size. More precisely, in Fig. 10, although the blue bold circles converge to 0.06001052 with error 1.2×10^{-6}, the upward (downward) triangles converge to 0.06001193 (0.06000933) as a limit of maximum (minimum) estimations. The total error is then the summation of these two error bars. In the example of Fig. 10, the total error is $\pm[0.06001193 - 0.06001052 + 1.2 \times 10^{-6}]$. To summarize here we first calculate the extrapolation value of the green, blue, and red points. Then we calculate the difference between the values associated with red and green points with the blue one. Then we pick the maximum value and add it to the error bar of the fitting of the blue points.

TABLE V. Shannon entropy for different sizes for an Ising chain derived using a BW reduced density matrix. The third column is the marginal probabilities for half of the subsystem.

L	BW	
	Total system	Subsystem
1	0.45920500717509	–
2	0.92205905719825	0.47270562810661
3	1.36645271889283	–
4	1.80496685397933	0.92510102900855
5	2.24030399857181	–
6	2.67358767640636	1.36780423258969
7	3.10542377417024	–
8	3.53618169812639	1.80575515259345
9	3.96610437840695	–
10	4.39536048183681	2.24082354163218
11	4.82407203544144	–
12	5.25233024242084	2.67395633633817
13	5.68020509249127	–
14	6.10775148615426	3.10569907398650
15	6.53501329186441	–
16	6.96202612243598	3.53639516977199
17	7.38881928804390	–
18	7.81541720297786	3.96627477249654
19	8.24184042007064	–
20	8.66810640526907	4.39549965444791
21	9.09423012703542	–
22	9.52022451133198	4.82418785310530
24	10.3718688193031	5.25242813140267
26	11.2231136875357	5.68028891764912
28	12.0740172271869	6.10782407525548
30	12.9246257180375	6.53507676179803
32	13.7749766184651	6.96208208951522
34	14.6251006747952	7.38886900778397
36	15.4750234355686	7.81546166549282
38	16.3247663610014	8.24188041670934
40	17.1743476517357	8.66814257612601

APPENDIX B: LIST OF SHANNON ENTROPIES

In this Appendix we summarize the Shannon entropy for different sizes of all the models that we considered in this paper.

1. Shannon entropy of models with $U(1)$ symmetry

In this section we provide the exact values of Shannon entropy in the models with $U(1)$ symmetry. In Table III we provided the Shannon entropy for different sizes for the models we considered in the main part of the paper.

2. Shannon entropy of models without $U(1)$ symmetry

In this section we provide the exact values of Shannon entropy in the models without $U(1)$ symmetry. In Table IV we provided the Shannon entropy for different sizes for the models with $f(z) = \{z - 1, z^2 - 1, z^3 - 1, z^5 - 1, z(z - 1)\}$. Then in the Table V first column we provided the Shannon entropy coming from the BW reduced density matrix. In the second column we explicitly write the Shannon entropy for the subsystem derived using the marginal probabilities.

[1] M. M. Wolf, F. Verstraete, and M. B. Hastings, Area Laws in Quantum Systems: Mutual Information and Correlations, Phys. Rev. Lett. **100**, 070502 (2008).

[2] J.-M. Stéphan, S. Furukawa, G. Misguich, and V. Pasquier, Shannon and entanglement entropies of one- and two-dimensional critical wave functions, Phys. Rev. B **80**, 184421 (2009).

[3] J.-M. Stéphan, G. Misguich, and V. Pasquier, Rényi entropy of a line in two-dimensional Ising models, Phys. Rev. B **82**, 125455 (2010).

[4] D. J. Luitz, F. Alet, and N. Laflorencie, Universal Behavior beyond Multifractality in Quantum Many-Body Systems, Phys. Rev. Lett. **112**, 057203 (2014).

[5] D. J. Luitz, F. Alet, and N. Laflorencie, Shannon-Rényi entropies and participation spectra across three-dimensional O(3) criticality, Phys. Rev. B **89**, 165106 (2014).

[6] D. J. Luitz, F. Alet, and N. Laflorencie, Participation spectroscopy and entanglement Hamiltonian of quantum spin models, J. Stat. Mech. (2014) P08007.

[7] F. C. Alcaraz and M. A. Rajabpour, Universal Behavior of the Shannon Mutual Information of Critical Quantum Chains, Phys. Rev. Lett. **111**, 017201 (2013).

[8] J.-M. Stéphan, Emptiness formation probability, Toeplitz determinants, and conformal field theory, J. Stat. Mech. (2014) P05010.

[9] H. W. Lau and P. Grassberger, Information theoretic aspects of the two-dimensional Ising model, Phys. Rev. E **87**, 022128 (2013).

[10] J.-M. Stéphan, Shannon and Rényi mutual information in quantum critical spin chains, Phys. Rev. B **90**, 045424 (2014).

[11] F. C. Alcaraz and M. A. Rajabpour, Universal behavior of the Shannon and Rényi mutual information of quantum critical chains, Phys. Rev. B **90**, 075132 (2014).

[12] F. C. Alcaraz and M. A. Rajabpour, Generalized mutual informations of quantum critical chains, Phys. Rev. B **91**, 155122 (2015).

[13] K. Najafi and M. A. Rajabpour, Formation probabilities and Shannon information and their time evolution after quantum quench in the transverse-field XY chain, Phys. Rev. B **93**, 125139 (2016).

[14] F. C. Alcaraz, Universal behavior of the Shannon mutual information in nonintegrable self-dual quantum chains, Phys. Rev. B **94**, 115116 (2016).

[15] J. C. Getelina, F. C. Alcaraz, and J. A. Hoyos, Entanglement properties of correlated random spin chains and similarities with conformal invariant systems, Phys. Rev. B **93**, 045136 (2016).

[16] F. Franchini and A. Abanov, Asymptotics of Toeplitz determinants and the emptiness formation probability for the XY spin chain, J. Phys. A: Math. Gen. **38**, 5069 (2005).

[17] M. N. Najafi and M. A. Rajabpour, Formation probabilities and statistics of observables as defect problems in free fermions and quantum spin chains, Phys. Rev. B **101**, 165415 (2020).

[18] F. Ares and J. Viti, Emptiness formation probability and Painlevé V equation in the XY spin chain, J. Stat. Mech. (2020) 013105.

[19] F. Ares, M. A. Rajabpour, and J. Viti, Exact full counting statistics for the staggered magnetization and the domain walls in the XY spin chain, Phys. Rev. E **103**, 042107 (2021).

[20] F. Ares, M. A. Rajabpour, and J. Viti, Scaling of the formation probabilities and universal boundary entropies in the quantum XY spin chain, J. Stat. Mech. (2020) 083111.

[21] J. Zhang, G. Pagano, P. W. Hess, A. Kyprianidis, P. Becker, H. Kaplan, A. V. Gorshkov, Z.-X. Gong, and C. Monroe, Observation of a many-body dynamical phase transition with a 53-qubit quantum simulator, Nature (London) **551**, 601 (2017).

[22] A. R. Its, F. Mezzadri, and M. Y. Mo, Entanglement entropy in quantum spin chains with finite range interaction, Commun. Math. Phys. **284**, 117 (2008).

[23] N. G. Jones and R. Verresen, Asymptotic correlations in gapped and critical topological phases of 1D quantum systems, J. Stat. Phys. **175**, 1164 (2019).

[24] R. Verresen, N. G. Jones, and F. Pollmann, Topology and Edge Modes in Quantum Critical Chains, Phys. Rev. Lett. **120**, 057001 (2018).

[25] R. Shankar, Renormalization-group approach to interacting fermions, Rev. Mod. Phys. **66**, 129 (1994).

[26] E. Witten, APS medal for exceptional achievement in research: Invited article on entanglement properties of quantum field theory, Rev. Mod. Phys. **90**, 045003 (2018).

[27] J. J. Bisognano and E. H. Wichmann, On the duality condition for a Hermitian scalar field, J. Math. Phys. **16**, 985 (1975).

[28] J. J. Bisognano and E. H. Wichmann, On the duality condition for quantum fields, J. Math. Phys. **17**, 303 (1976).

[29] P. D. Hislop and R. Longo, Modular structure of the local algebras associated with the free massless scalar field theory, Commun. Math. Phys. **84**, 71 (1982).

[30] K. Najafi and M. A. Rajabpour, Entanglement entropy after selective measurements in quantum chains, J. High Energy Phys. 12 (2016) 124.

[31] H. Itoyama and H. B. Thacker, Lattice Virasoro Algebra and Corner Transfer Matrices in the Baxter Eight-Vertex Model, Phys. Rev. Lett. **58**, 1395 (1987).

[32] I. Peschel, M. Kaulke, and Ö. Legeza, Density-matrix spectra for integrable models, Ann. Phys. **511**, 153 (1999).

[33] B. Nienhuis, M. Campostrini, and P. Calabrese, Entanglement, combinatorics and finite-size effects in spin chains, J. Stat. Mech. (2009) P02063.

[34] I. Peschel and V. Eisler, Reduced density matrices and entanglement entropy in free lattice models, J. Phys. A: Math. Theor. **42**, 504003 (2009).

[35] V. Eisler and I. Peschel, Analytical results for the entanglement Hamiltonian of a free-fermion chain, J. Phys. A: Math. Theor. **50**, 284003 (2017).

[36] V. Eisler and I. Peschel, Properties of the entanglement Hamiltonian for finite free-fermion chains, J. Stat. Mech. (2018) 104001.

[37] M. Dalmonte, B. Vermersch, and P. Zoller, Quantum simulation and spectroscopy of entanglement Hamiltonians, Nat. Phys. **14**, 827 (2018).

[38] G. Giudici, T. Mendes-Santos, P. Calabrese, and M. Dalmonte, Entanglement Hamiltonians of lattice models via the Bisognano-Wichmann theorem, Phys. Rev. B **98**, 134403 (2018).

[39] T. Mendes-Santos, G. Giudici, R. Fazio, and M. Dalmonte, Measuring von Neumann entanglement entropies without wave functions, New J. Phys. **22**, 013044 (2020).

[40] T. Mendes-Santos, G. Giudici, M. Dalmonte, and M. A. Rajabpour, Entanglement Hamiltonian of quantum critical chains and conformal field theories, Phys. Rev. B **100**, 155122 (2019).

[41] J. Zhang, P. Calabrese, M. Dalmonte, and M. A. Rajabpour, Lattice Bisognano-Wichmann modular Hamiltonian in critical quantum spin chains, SciPost Phys. Core **2**, 007 (2020).

[42] M. Dalmonte, V. Eisler, M. Falconi, and B. Vermersch, Entanglement Hamiltonians: From field theory, to lattice models and experiments, arXiv:2202.05045.

[43] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, Cambridge, 2004).

[44] H. Gould, J. Tobochnik, and W. Christian, *An Introduction to Computer Simulation Methods Applications to Physical System* (Addison-Wesley, Reading, MA, 1996).