Entanglement compression in scale space: from the multiscale entanglement renormalization ansatz to matrix product operators

Karel Van Acoleyen,1 Andrew Hallam,1 Matthias Bal,1 Markus Hauru,1 Jutho Haegeman,1 and Frank Verstraete1

1Department of Physics and Astronomy, Ghent University, Krijgslaan 281, S9, B-9000 Ghent, Belgium

The multiscale entanglement renormalization ansatz (MERA) provides a constructive algorithm for realizing wavefunctions that are inherently scale invariant. Unlike conformally invariant partition functions however, the finite bond dimension $\chi$ of the MERA provides a cut-off in the fields that can be realized. In this letter, we demonstrate that this cut-off is equivalent to the one obtained when approximating a thermal state of a critical Hamiltonian with a matrix product operator (MPO) of finite bond dimension $\chi$. This is achieved by constructing an explicit mapping between the isometries of a MERA and the local tensors of the MPO.

Introduction. Arguably the central problem in the study of extended quantum many-body systems - whether it is in the context of high energy physics or condensed matter physics - is the identification of the relevant infrared (IR) degrees of freedom. A new perspective on this problem arose with the advent of tensor network states (TNS) [1, 2]. The crucial feature of the TNS approach is that it directly targets the quantum many-body state and in particular its entanglement structure. In this sense TNS realise a compression or truncation of the exact quantum state in terms of the most relevant entanglement degrees of freedom. In particular, as the virtual (contracted) indices of the tensors encode the entanglement structure, the bond dimension $\chi$ of these indices can be thought of as the systematic control parameter of the entanglement compression achieved by a particular TNS. Given their success thus far it is desirable to better understand the precise nature of this compression. In this letter we focus on systems with one spatial dimension.

For matrix product states (MPS), this picture of entanglement compression was recently made more concrete in terms of the Euclidean path integral representation of the state [3–5]. In particular it was argued that MPS approximations can be understood as efficient compressions of the strip path integral over a finite imaginary time interval $L$, with $L$ scaling polynomially in $\chi$. This explains why MPS approximations for ground-states of gapped systems are essentially exact for sufficiently large bond dimension, while for critical systems this motivates the use of finite entanglement scaling methods in direct correspondence with the finite length scaling methods of Monte-Carlo simulations [6–10].

In this letter we address the role of the finite bond dimension and the type of compression in the case of the multi-scale entanglement renormalization ansatz (MERA) [11–13]. This is a different type of TNS, that is tailored specifically towards the description of critical systems. A remarkable feature of this ansatz is that it realises an explicit renormalization group (RG) flow directly at the level of the quantum state: the network has an extra scale dimension with subsequent layers of local isometries and unitaries (disentanglers) effecting the RG flow. In contrast to MPS, this structure allows to capture the power-law decay of correlations and the logarithmic growth of the finite interval entanglement entropy, already at finite bond dimension. Furthermore, the conformal data of the theory - the different scaling dimensions and operator product expansion (OPE) coefficients - can be extracted directly from local scaling operators that are constructed in terms of the MERA isometries and unitaries [14–16]. Notice that it has been shown already that one can construct MERA states directly from the application of the tensor network renormalization (TNR) procedure on the half-finite Euclidean path integral with open physical indices [17]. But regarding the role of the bond dimension this merely shifts the question to the role of the bond dimension in TNR.

Here we will offer a different view: we will argue that MERA approximations for critical systems can be understood naturally in scale space. For this we revisit and further develop the results of [18] on the MERA realisation of the conformal plane to strip map, which for the MERA is a map from physical to scale space; see also [19, 20] for related approaches. In particular we will argue and show numerically that the half-infinite subsystem density matrix reduces to a thermal (strip) state in scale space, $\rho_{\text{scale}} = e^{-\beta H}$, for a Hamiltonian $H$ belonging to the same universality class as the original critical Hamiltonian $H$ for which the MERA was optimised. The picture that is emerging is one where a finite $\chi$ MERA yields a particularly efficient matrix product operator (MPO) approximation for thermal states. Intriguingly, our numerics indicate that $\chi$ is controlling both an IR and ultraviolet (UV) cut-off of the CFT Hamiltonian $H$. Furthermore, the (single-site) MERA scaling operator appears as the transfer matrix of this strip state, which relates the MERA procedure for extracting scaling dimensions to Cardy’s finite-size scaling [21–23]. Finally, motivated by the MPO compression view, in the last part of this letter we propose and implement a new algorithm for the construction of isometries and corresponding scaling operators from direct MPO approximations of the thermal state in terms of the original Hamiltonian $H$.

From real space to scale space, from the plane to the strip. It is useful to start with a brief review of some CFT results in the continuum as this will guide our interpretation of the discretised MERA states. We take a CFT on the infinite line and consider the vacuum density matrix...
for the infinite right half-line subsystem. First, for any relativistic QFT (including CFTs) a foliation of the Euclidean path integral in the angular direction leads to the general expression (see e.g. [24]):

$$\rho = \exp -2\pi \int_0^{\infty} dx T_{00}(x),$$

where $T_{00}(x)$ is the energy density operator. Here and in the following we drop the normalisation constant ensuring $\text{Tr}\rho = 1$. For CFTs then, the conformal map from the plane to the strip with width $\beta$,

$$z \rightarrow w = \frac{\beta}{2\pi} \log z,$$

maps this density matrix to a Gibbs state that is translation invariant:

$$\rho_{\text{strip}} = \exp -\beta \int_{-\infty}^{\infty} ds T_{00}(s).$$

Here $z = x + it$ and $w = s + i\tau$. Notice that at the level of Hilbert space, this conformal map boils down to a canonical (isometric) map $W$ of the original CFT Hilbert space on the half-line subsystem $x \in [0, +\infty]$ to a new CFT Hilbert space on the full line $s \in [-\infty, +\infty]$, with e.g. the transformation for the energy density operator: $W^{-1} T_{00}(x)W = (\frac{d}{ds})^2 T_{00}(s) = (\frac{\beta}{2\pi})^2 T_{00}(s)$, where $T_{00}(s)$ is the energy density operator on the new Hilbert space and we drop the Schwarzian derivative term proportional to the unit operator [25]. It is readily checked then that this map $W$ indeed transforms $\rho$ into $\rho_{\text{strip}}$: $W^{-1} \rho W = \rho_{\text{strip}}$.

We now consider the discretised version of a half-line CFT density matrix as realised by a numerically optimised ternary MERA with bond dimension $\chi$. As illustrated in fig. 1a, by contracting the ket and bra to the left of the origin, we obtain the reduced density matrix $\rho$ defined on the real space interval $x \in [0, +\infty]$.

All tensors to the left of the causal cone of the origin site drop out due to the unitary and isometric restrictions on the tensors, resulting in an explicit expression $\rho = W \rho_{\text{scale}} W^{-1}$. Here $\rho_{\text{scale}}$ is the matrix product operator (MPO) of fig. 1b, with the MPO tensors constructed out of the MERA isometries and their conjugates at the different layers labeled by the RG log-scale $s$. $W$ is the full isometry from the physical real space indices on the right half-line, to the virtual indices in scale space that are cut by the right causal cone. For a critical MERA $\rho_{\text{scale}}$ will be translation invariant up to the first few $s$-dependent isometries that capture the non-universal UV behaviour of the state. The previous continuum CFT discussion then strongly suggests that $W$ gives us a discretised version of the conformal plane to strip map, in particular with $z = 3^s$, as local operators at position $x$ in physical space get mapped to local operators at position $s \sim \log_3(x)$ in scale space. Notice that different from [18], we consider here both the bra and ket, and as a consequence the map $W$ is exactly isometric.

Taking $W$ as the discretized plane to strip map we get the prediction that the bulk density matrix in scale space is a thermal (strip) density matrix with inverse temperature $\beta = \frac{2\pi}{\log(3)}$ of a Hamiltonian $\tilde{H}$ with the same IR CFT fixed point as the original Hamiltonian for which the MERA was optimised:

$$\rho_{\text{scale}} = \exp -\frac{2\pi}{\log(3)} \beta.$$

Before discussing the numerics in light of this prediction, let us point out that this ties in nicely with Cardy’s seminal results on finite-size scaling for CFTs. From the conformal plane to strip map it was shown that the correlation functions on the strip should decay exponentially, in particular the transfer matrix $T$ on the strip (with periodic boundary conditions) should have the spectrum

$$T = \exp \frac{2\pi}{\log(3)} \Delta,$$

with $\Delta$ the scaling dimensions of the local primary operators and their descendants [21–23]. As shown explicitly in fig. 1c, upon a proper identification of the indices the transfer matrix of $\rho_{\text{scale}}$ indeed precisely reduces to the single-site MERA scaling (super-)operator $S$ that is used to extract the scaling dimensions from the MERA simulation according to $S = \exp -\log(3) \Delta$.

We have studied the properties of $\rho_{\text{scale}}$ at different MERA bond dimensions $\chi$, for the transverse-field Ising model at its critical point. First we have numerically optimised all the MERA tensors with the standard MERA algorithm [13] - including two non-scale invariant UV layers, imposing the explicit $Z_2$ symmetry of the Ising model [26, 27]. Then we have considered the MPO expression for $\rho_{\text{scale}}$ constructed out of the (scale-independent) bulk isometries (fig. 1b). Using standard matrix product state (MPS) techniques, we then have constructed an MPS approximation of the fixed point of the MPO in the thermodynamic limit, and have used the MPS particle momentum eigenstate ansatz [28], to approximate the one-particle excitations on top of the fixed point.
In fig.2 we show the resulting log spectrum, $\omega = -\log \rho_{\text{scale}}$ for MERA optimizations on the Ising model for different $\chi$. Notice a subtlety involving a MERA gauge choice: depending on this choice one ends up with an isometry that gives rise to a positive or non-positive scaling operator $S$, and a dispersion relation of the type of fig.2a or fig.2b [29].

From the prediction (4) we would expect a momentum regime with a linear spectrum, $\omega(k) = \beta k$, with $\beta = 2\pi/\log(3)$. For the different $\chi$ values we indeed find a linear regime in very good agreement with this prediction. To discuss the role of $\chi$ we focus on the dispersion relations of the first type (fig.2a). It turns out there is a small mass gap for $k \to 0$, which for growing $\chi$ decreases as an approximate power law, as can be seen in fig.3a (red crosses). Notice that we have made sure that this is not an artefact of our MPS approximation of the fixed point: our results were well converged in the MPS bond dimension. Furthermore $\chi$ also has its effect at the other side of the spectrum, where we again see a deviation from the linear result. To associate a UV energy scale $\Lambda$ to this effect, we have identified $\omega(\pi) = \beta \Lambda \sin(\pi/\Lambda)$. Notice that there does not seem to be a unique preferred way to identify this UV scale. The rationale for our particular choice is that the right-hand side corresponds to an Ising dispersion relation on a coarse grained lattice with $O(\Lambda)$ sites blocked into one effective site. In fig.3b we see that $1/\Lambda$ again appears to obey an approximate power law in $\chi$. So, taken at face value the bond dimension $\chi$ seems to control both a relevant (IR) and irrelevant (UV) perturbation to the CFT. In the proper scaling regime both type of operators give power corrections to the finite size-scaling (5) [22, 30, 31] which in our case should therefore translate to power corrections for the scaling dimensions obtained from the MERA scaling operator $S$. This is reasonably consistent with our numerical results, see fig.3c and fig.3d.

From thermal states to isometries and scaling operators. The results above suggest that a MERA optimisation is in effect performing an indirect compression of a thermal state for a critical Hamiltonian in terms of an MPO of the particular form fig.1b. Our numerics indicate that this compression is efficient, with both the IR and UV cut-off scale, together with the corrections on the exact CFT scaling dimensions decaying according to a power law in $\chi$. As such this is no surprise as in general, both on theoretical grounds [33, 34] and based on many previous numerical applications, it is known that thermal states of local Hamiltonians can be efficiently approximated by an MPO.

To make contact with the MERA results of the previous section we have considered a specific direct MPO compression scheme. In particular we have followed the approach of [35], by evolving the infinite-temperature state with the Hamiltonian in imaginary time using time-evolving block decimation in a second order Trotter approximation of the specific form fig.1b. Our numerics indicate that this compression is efficient, with both the IR and UV cut-off scale, together with the corrections on the exact CFT scaling dimensions decaying according to a power law in $\chi$. As such this is no surprise as in general, both on theoretical grounds [33, 34] and based on many previous numerical applications, it is known that thermal states of local Hamiltonians can be efficiently approximated by an MPO.
Figure 4: (a): The log spectra of the MPOs $\omega = -\log \rho$, for MPO representations of the Ising thermal state at temperature $\beta = 2\pi/\log 3$ for different $D$. (b): The velocity of the free boson describing the XXZ model for $H = \sigma_x \sigma_x + \sigma_y \sigma_y - g \sigma_z \sigma_z$, along the full critical line $|g| < 1$. Notice that for this model, in general the velocity $v(g) \neq 1$. This implies a linear dispersion $\omega(k) = \beta v(g) k/2$ on our (2 site) blocked lattice. From our results on the log spectrum of $\rho$ we can fit a value for $v(g)$ that agrees nicely with the exact result [32], as can be seen in fig.4b. The velocity $v(g)$ also shows up in the scaling dimension relation, which now reads: $S = \exp -\frac{\pi}{\log \Lambda}$. Fig.4d shows the $g = -0.5$ results from this scaling operator. The XXZ model is significantly more challenging than the Ising model, this is reflected in the fact that a significantly higher bond dimension is needed to accurately determine its scaling dimensions.

Outlook. Our results demonstrate that the finite bond dimension $\chi$ approximation of a MERA is of the same nature as the finite bond dimension $\chi$ MPO approximation of a thermal state of a critical Hamiltonian. This points to a new algorithm for creating critical MERA states. We have shown how to obtain the isometries, the complete algorithm of course also requires a similar construction for obtaining the disentanglers. But notice that the MPO expression of fig.1b corresponds to a special cut. One can verify that arbitrary cuts lead to MPO expressions for $\rho_{\text{scale}}$ that also involve the disentanglers, giving rise to non-translation invariant transfer matrices that depending on $s$ take the form of the single-site scaling operator or the left, centre or right double-site scaling operators. This shows how to obtain the disentanglers: by minimising the breaking of translation invariance for MPOs corresponding to such general cuts [37]. From the (bulk) disentanglers, in combination with the isometries, one can then also obtain the OPE coefficients, overcoming the main limitation of the thermal MPO approach that we presented here.

Acknowledgements. We would like to thank Laurens Vanderstraeten for inspiring discussions. This work is supported by an Odysseus grant from the FWO and has received funding from the European Research Council (ERC) under the European Union Horizon 2020 research and innovation programme (grant agreement No 647905 (QUITE), No 715861 (ERQUAF) and SIQS).
