nature portfolio

Peer Review File

Visualizing Dynamics of Charges and Strings in (2+1)D Lattice Gauge Theories

Corresponding Author: Dr Pedram Roushan

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Referee #1

(Remarks to the Author)
The paper simulated the dynamics of 2+1d Z2 gauge theory on a 2d lattice of superconducting qubits. Very rich physics were found, including the confined/deconfined phases of the gauge field, propagation of excited gauge charges in both phases, dynamics of strings, and string breaking in the confined phase. It is quite amazing that so many results can be observed in a relatively small system size.

Overall speaking this is a very good paper, and I enjoyed reading it. I believe it warrants publication on a top journal like Nature.

I just have some minor questions/comments:

I would recommend the authors provide some theoretical insights, or numerical simulation for the results presented in Fig.4d. For example, what would these plots look like in the ideal situation, i.e. when there is no decoherence at all and the system size is infinite?

In particular, why should the Z(t)Z(0) correlation decay but increase again over time for small h_M (if my reading of the figure is correct)? Do we expect the correlation to oscillate with time in the ideal situation?

I recommend publication of the paper after these comments are addressed.

Referee #2

(Remarks to the Author)
The authors of the manuscript present a fascinating experiment demonstrating the ability to study the time evolution of a particular two-dimensional lattice gauge theory, specifically the dynamics of a string excitation in two different regimes, probing the confinement-deconfinement regimes. The manuscript presents protocols to prepare the initial ground state at different system parameters, the initial excitations, the time evolution, and -very importantly- the necessary measurements to probe the physics under study.

The manuscript is clearly written and the reported experiment demonstrates a clear advance in the path pursued by different groups of theoretical and experimental physicists in the last decade, struggling to bring to the lab the original vision put forward by Feynman forty years ago. This experiment, which first proves a digital quantum simulation of an excitation string dynamics in a two-dimensional LGT, could be a landmark in this direction. However, there are a number of shortcomings that prevent me from recommending this manuscript for publication. I would encourage the authors to consider them:

1. The main objection is that, despite the beauty of this experiment, nothing is learned from the quantum computing part of it. In other words, all that is shown can be simulated (as done by the authors) straightforwardly by classical computers. I do not underestimate the complexity of this experiment, nor the theoretical analysis and protocol developments needed to reach this result. I fully appreciate their beauty. However, here we are discussing if this route can, at least eventually, lead to some strong demonstration of quantum computation usefulness.

Here it is completely unclear: there are no statements on the difficulties of classical simulations while scaling the system size, on the continuum limit that will give effective useful physical results, etc. How is the entanglement scaling in these simulations? is it possible to simulate them classically, e.g., with tensor network? There have been recent results in higher-dimension on ground states and dynamics. My feeling is that the system remains pretty much low-entangled all the time here. There are instead unsupported statements here and there "prohibitively hard to simulate" in the conclusions and ""generally beyond the perturbative limits" in the abstract.
Without extensively discussing this point, I do not think that this is a Nature paper.

2. Related to the previous point: how does the entanglement scale? The numerical simulations already made by the authors contain this information. If this is very low entangled, then the difficulty of this experiment is actually overestimated.

3. The authors refer to the initial excitations as particles in the abstract. However, given that they create them by applying an X operator, they are similar to classical charge, not to any particles (system excitations). The authors should clarify this throughout the text.

4. Another possible strong point of the paper is to compare their circuit with already reported digital quantum computations: they state that each trotter step counts for about 117 CZ gates, it seems to em they perform O(10) trotter steps, that is O(1000) two-qubit gates. How does this compares with previous experiments?

5. The preparation of the ground state is discussed in great detail. However, I cannot find a theoretical prediction of the protocol in terms of fidelity (again possible in the numerical simulations). I see that the final energy is near to the expected one, but how well is the preparation, theoretically? A possible test is to prepare the state and see if this evolves in time. Has this test been performed in the QPU?

Minor points:

1. The legend of Fig. 3.d is incomplete in my pdf.
2. Despite the fact that I can probably guess the meaning, it is not clear to me what is represented in Fig. 3e in the four different columns: are these different results of measurement? What is the probability written in a given position with the "measurement symbol". More explanations are needed there.

Referee #3

(Remarks to the Author)
This is the first work presenting a quantum simulation of the dynamics of a full-fledged lattice gauge theory on a system of a size sufficient to study the confinement of matter particles. For this reason, it will have a major impact on the community of quantum simulations and it sets a fundamental milestone for the growing community studying quantum theories inspired by high-energy physics through quantum technologies.

Several experimental works already addressed 2D lattice gauge theories on quantum processors, by using ion traps (see, for instance, arXiv:2310.12110), superconducting quantum processors (see, for instance, Phys. Rev. D 103, 094501 (2021); Phys. Rev. D 106, 074502 (2022)) or quantum annealers (for instance Phys. Rev. D 104, 034501 (2021)). To the best of my knowledge, however, only arXiv:2310.12110 addressed both matter and gauge degrees of freedom, and it did it on very small lattices, such that I would consider that result based on ion setups only a proof of the efficient realization of the building blocks for a 2D LGT, rather than a more complete simulation.

The current manuscript from the Google collaboration, instead, demonstrates a quantum simulation of the full Z_2 lattice gauge theory on a truly 2D systems with a clearly defined and visualizable geometry with open boundary conditions. I believe that the combination of the relative theoretical simplicity of the obtained results and their clear visualization makes this manuscript outstanding and of doubtless value from both the scientific point of view and as a showcase for quantum computation applications.

For these reasons, I believe that the results presented in this manuscript meet the impact and innovation criteria of Nature. There are, however, a few parts in the presentation of lattice gauge theories that I find not so rigorous. Additionally, several aspects of the presentation and the description of the data should be improved. Moreover, in the following, I will also suggest providing some additional data and materials in the supplementary information. Consequently, a revision of the manuscript is necessary.

Let me first address the main semantic issues about the general presentation of the analyzed lattice gauge theory.

Since the abstract, the authors refer to the coefficient h_M in the Hamiltonian (1) as "the magnetic field". I do not find this kind of nomenclature rigorous from the point of view of lattice gauge theory and I believe it can cause major confusion for the readers, since the magnetic field of the theory is related to the plaquette operators. The coefficient h_M is related instead

to the coupling constant of the theory and I suggest changing its denomination. I also disagree with the statement "The h_M terms denote a magnetic field on each link that creates magnetic flux excitations.": these terms describe the energy density of the electric field, and they are commonly referred to as electric field terms. Therefore, it is very confusing to refer to h_M as magnetic field. Moreover, I also find very confusing the following remark "The electric field terms h_E generate hopping of matter fields located at adjacent vertices, as mediated by a gauge field on their connecting link.". I agree that these terms describe the matter hopping, but they commute with the magnetic plaquette term and it is unacceptable and contrary to the standard characterization of lattice gauge theory to define them as "electric field terms" (which, as I wrote, are instead the h_M terms). In general, the authors must adopt a more conventional and rigorous nomenclature, following the standard set by lattice gauge theories. Fig.1 needs to be changed accordingly, as well as a few other parts of the text (for instance lines 339, 370).

Again concerning theterminology, in the caption of Fig. 1 the authors indicate as a "pure LGT" the theory after removing the matter degrees of freedom via the gauge constraints. This is wrong. A pure LGT typically refers to models in which $h_E =0$ and J_E diverges, irrespective of the choice made to represent it. The fact that a model is a pure lattice gauge theory is a physical statement (matter particles have infinite mass), not a statement that depends on the choice of the adopted degrees of freedom. The representation of the Hamiltonian obtained by gauging out the matter degrees of freedom is usually called "unitary gauge" (see, for instance, Fradkin's textbook "FIELD THEORIES OF CONDENSED MATTER PHYSICS" (second edition), Chapter 9.10).

Besides these general aspects about the description of the Z_2 lattice gauge theory, in the following I list several points that the authors should address concerning the presented data.

Concerning Fig. 2c, I find it curious that the yellow data (product state) display a simulated curve with a worse energy than the experimental results. Can the authors comment about it? Are their simulations to obtain the three solid lines including any sort of noise or errors in the computation? or do they assume an exact implementation of all gates without errors?
Sec. IVA of the supplemental information mentions that the experimental results include a correction for readout errors, but I do not understand how correcting the data may lead to lower values than the theoretical predictions.
The result is counterintuitive because, for the yellow curve, it seems that the experimental errors after the initialization to the trivial product state would be beneficial to approach the ground state. Do the authors confirm that the yellow data points are obtained measuring plaquettes and links directly on the state |0>^N irrespective of the value of h_M?

Again concerning Fig. 2c, it may help the reader to provide an estimate of the critical value of h_M in the thermodynamic limit from the literature. For $h_E=0$, this value would be slightly below 1/3 from the known critical field of the Ising model (see, for instance Phys. Rev. E 66, 066110 (2002).). It should not change much for $h_E=0.25$, and it may be interesting to compare it to the crossing between the yellow and blue data. This indication is also useful for the reader when comparing the three regimes of the string dynamics in Fig. 4c.

Regarding the data displayed in Fig 2d, the color scale does not allow for a clear quantitative understanding. I think it would be important to add in the supplementary information a graph that displays the average expectation values of the plaquettes, the Z operators on bulk and boundaries and the A vertex terms as a function of h_M (the latter unless postselection automatically discards all <A>=-1 instances).

Concerning Fig. 4c there is a detail that I do not fully understand. In the panel at $h_M=0.1$ and $t=0$, $S_{ZZ}(t=0)$ is sizeably different from zero not only on the sites affected by the applied X string, but also on neighboring sites. At t=0, though, S_{ZZ} is just the expectation value of Z(0), and, for sites that are not involved in the X string, it must match the expectation value of Z after the WALA initialization. Did the authors check whether the values plotted in the panel $h_M=0.1$ and $t=0$ for the sites that do not belong to the X string match the expectation value of the WALA? Similar data are plotted in Fig. 2d, but $h_M=0.1$ is missing there and I cannot make a comparison. Therefore I would ask the authors to show them in their resubmission letter. Indeed, I find quite surprising the difference of the <S_ZZ(t=0)> value between the boundaries and the central site for $h_M=0.1$.

Regarding the data of Fig. 4d and the dynamics of the string at low h_M, the authors write "In the deconfined phase, $h_M = 0.1$, applying an X-string on top of the WALA sequence does not create a string excitation. Thus, the correlation function $S_{ZZ}(t)$ quickly trends towards zero.". I must admit that I find this description too simplistic. The behavior of S_{ZZ} remains close to zero in this case simply because the sort of normalization <Z(0)> is very low close to the toric code limit. However, the first two panels of Fig. 4d reveal that the correlation function <Z(t)Z(0)> has a marked coherent oscillation which is stronger than the cases with large h_M (this is even more evident in Fig. S15b). Therefore I agree that in this deconfined phase there is no creation of an electric flux string. However, I disagree on the fact that "the correlation function $S_{ZZ}(t)$ quickly 'trends' towards zero". It seems to me that it acquires lower values since the beginning just because of the adopted <Z(0)> normalization, but, otherwise, the dynamics trend seems to me more oscillating and coherent than the large h_M case. This is why I dislike its description as 'quickly trends towards zero'.

In line 330, the authors write: "Surprisingly, the string is still moving freely around the top qubits despite the large string tension." As the authors mention afterwards, this is mostly due to length-preserving deformations of the string and I do not find it so surprising. I would avoid writing "Surprisingly".

In line 354, the authors write "the string initial state possesses considerably more charge particles". I disagree with this sentence. It is the state obtained after the initial insertion of a string that at time t=2.7 possesses considerably more charge particles. The initial string state, if I understand correctly, should have no charge particles instead (<A>=1 from the WALA

initialization and the boundary to boundary string commutes with all A's in the bulk). Am I wrong?

From the caption of Fig. 5c, one understands that $h_M=1.4$, however it is not explicitly stated (it is written that the grey region corresponds to this value). I advise to revise the caption and add this indication more explicitly.

Concerning the discussion in the supplemental information III, when discussing Eq. S8 the authors write: "Intuitively, such a suppression of large loops is what one might expect when adding an onsite Z-field to the toric code Hamiltonian".
This intuitive picture I think can be misleading for a couple of reasons:
A) Let us take a Z-field with the same sign convention of Hamiltonian (1). If $h_M<0$, then the states |0> are favored and long loops are certainly suppressed when $h_M<-1$. Therefore theta must decrease for negative h_M. If this is what the authors had in mind, one needs to specify that the Z-field would correspond to $h_M<0$.
B) For $h_M>0$, the |1> states are favored. This does not automatically correspond to the fact that theta must increase beyond Pi/2. On the contrary, Fig. 2b and Sec. III C of the SI show the other way round. The point is that in order to polarize the spins along |1>, we should apply the B operator only to either even or odd plaquettes [in the thermodynamic limit], not to all of them. By applying all the plaquette operators (theta-> Pi in Eq. S8) one would flip twice all the spins but the ones on the boundaries, resulting in a single large loop around the boundary.
For these reasons I think that the description of Eq. S8 should be slightly modified to avoid intuitive but potentially misleading statements.

In the caption of Fig. S4 b, replace "(dotted line)" with "(dashed line)"

In Sec. IVA of the supplemental information, the authors use the ket notation |\phi> "to be the probabilities of each computational basis state...". I find it a misleading notation, since the ket is used for pure quantum states. Do the authors mean that |\phi> constitute a diagonal density matrix of the system instead? Or maybe they could represent \phi simply as a vector of probabilities.

In Sec. IVA of the supplemental information, the authors refer to the Loschmidt echo of the state preparation circuits. I find this part very technical and not self-contained. I invite the authors to include additional references or a more detailed explanation of the procedure. A brief definition of the energy of the Loschmidt circuit would be useful.

In Fig. S11e, the color scale (blue vs black) is not specified [in Fig. 3 it is ok instead].

Concerning the analysis of the dynamics of the single particle in Sec. IVD1, the authors interpret their data in Fig. S12d in terms of the motion of the charge excitation, that, in their opinion gets slower for strong confinement. I tend to think, instead, that a more rigorous interpretation would be based on screening rather than confinement and particle velocity. The initial conditions defined in Fig. S12a with a pinned qubit Q_0 are consistent with an electric flux line entering the system through Q_0. If we consider the ground state with this boundary condition, depending on the value of J_E and h_M (for $h_E>0$) we would experience a crossover from weak screening at $J_E >> h_M$ to a strong screening when $h_M >> J_E$. When screening is weak, an electric flux line propagates on average over many links in the lattice before being screened by matter excitations. On the contrary, when screening is strong, electric charges localize close to the top left corner to screen the incoming electric flux line. Now, the data presented concerns the dynamics of the system rather than its ground state properties. However it is likely that the long-time behavior (bottom row of Fig. S12d) is displaying the formation of the screening cloud associated to the incoming electric field line. For small h_M the screening cloud would be larger than the system size, thus the average electric field percolated on the whole system. For large h_M, instead, the electric field is screened and remains localized on the top left corner. In this respect, it may be instructive to plot also the expectation value of Z over all the sites of the system (as done in Fig. 2d). In this way a reader could easily spot whether and how the electric field percolates across the system. As a more general note, the authors focus mostly on small values of h_E, where the distinction between confined and deconfined phases is sound. However, usually, confinement [meant as an exponential decay of the Wilson loop with the area of the loop] is a property of the pure lattice gauge theory [$h_E=0$]. For $h_E>0$, as the authors demonstrate, there are string breaking mechanisms that are a manifestation of screening. This is why, at large h_M and $h_E>0$ I think that the concept of screening would be more appropriate than confinement.

In line 424, the following sentence is grammatically wrong "In the topological phase, the motion of the electric excitations at the end points of the string are not confined.".

In conclusion, the results presented in this work constitute a decisive step forward in the field of quantum simulations of lattice gauge theories. This work constitutes indeed a milestone for quantum simulations as it convincingly demonstrates the possibility of performing the simulation of the dynamics of interacting 2D quantum many-body problems on a system size going beyond the single building blocks that have been investigated in previous works. However, some parts of the exposition of this work are written in an unsatisfactory way (especially concerning the description of the elements of the lattice gauge theory) and, in general, a rigorous review is necessary before a final evaluation of the manuscript.

Version 1:

Reviewer comments:

Referee #2

(Remarks to the Author)

The authors have adequately addressed my previous comments, thus I recommend the publication of the manuscript in its present form.
However, let me mention that I do not agree with some of the comments that are present in their reply (not in the manuscript though),
let me mention two major points:

1. The authors seem unaware of recent developments in tensor network simulations. To date, there are simulations of 2+1 and even 3+1D ground states of LGT and in the dynamics of two-dimensional systems of pretty large (16x16) spin systems, mostly based on Tree Tensor Networks. Thus, the presented simulations are far from being state-of-the-art. However, they show that the entanglement grows in the initial state in such a way that simulations will be in any case very hard. The fact that what is measured is sensitive to this entanglement is still not completely clear (see next point), but what is shown is enough for me to justify the importance of this experiment.

2. The statement "On the experimental side, entanglement is not the limiting factor. The experimental difficulty rather lies in maintaining coherence for sufficiently long times." is kind of misleading: exactly as for tensor networks, the ability to maintain entanglement is the limiting factor, if not as it is a form of coherence to be maintained. And that is more difficult to maintain than that of single qubits. If this is not evident in this experiment, it means that it can be reproduced with tensor network methods efficiently with a very low bond dimension.

Referee #3

(Remarks to the Author)
The authors answered in full detail to the points I raised. I believe that the new version of the manuscript is now rigorous and its presentation is now consistent with the standard terminology of lattice gauge theory.

This makes the text readable and enjoyable for a broad readership.

I confirm my first impression: I believe that this work is a milestone for the field of quantum simulations of many-body systems. It showcases for the first time the possibility of investigating in detail the dynamics of an intricate two-dimensional quantum many-body problem on a quantum processor beyond previous proofs of principle. The model addressed, namely the Z2 lattice gauge theory, is a crucial model for different areas of physics, but the results presented here have a significance that goes beyond this specific model. They show a way forward for remarkable advances in many fields of research including the study of models of interest for condensed matter physics, particle physics and quantum chemistry.

For these reasons I fully support the publication of this work in Nature.

At the level of minor details, I suggest the authors to consider these final adjustments:

1) The authors solved the issues related to the lattice gauge theory terminology. However, I suggest that they improve even further the rigorousness of their abstract. In line 10 they write "As the electric field is increased,...". The meaning of this sentence, however, should be "As the coupling constant associated with the electric field energy density is increased,...", since the electric field is the operator Z and not the constant h_E. Therefore, I would suggest to rephrase that sentence. A possible and relatively short correction could be "As the electric field coupling constant is increased".

2) In line 450 of the supplemental material there is a typo: "...is is threaded in."

Sincerely yours,

Michele Burrello
(University of Pisa)

Dear Editor

Please find our point-by-point response to the referees' comments below.

— — —

Note to all reviewers: Taking up the suggestion of reviewer #3, we have updated the notation. The Hamiltonian now reads:

$$\mathcal{H} = -J_E \sum_v A_v - J_M \sum_p B_p - h_E \sum_{\text{links}} Z_l - \lambda \sum_{\text{links}} X_l.$$

The Z_l terms are now referred to as the electric field and the X_l terms are deemed the matter-gauge coupling. We have used this new scheme to respond to all questions (which we quote as they were formulated in the old notation).

The formatting for this reply is as follows:

- Direct quotes from reviewers are green
- Our responses are black
- *Quotes from the manuscript are italicsized*
- *Additional text added in response to reviewer comments is highlighted.*

Referee #1 (Remarks to the Author):

The paper simulated the dynamics of 2+1d Z2 gauge theory on a 2d lattice of superconducting qubits. Very rich physics were found, including the confined/deconfined phases of the gauge field, propagation of excited gauge charges in both phases, dynamics of strings, and string breaking in the confined phase. It is quite amazing that so many results can be observed in a relatively small system size. Overall speaking this is a very good paper, and I enjoyed reading it. I believe it warrants publication on a top journal like Nature.

I just have some minor questions/comments:

I would recommend the authors provide some theoretical insights, or numerical simulation for the results presented in Fig. 4d. For example, what would these plots look like in the ideal situation, i.e. when there is no decoherence at all and the system size is infinite? In particular, why should the Z(t)Z(0) correlation decay but increase again over time for small h_M (if my reading of the figure is correct)? Do we expect the correlation to oscillate with time in the ideal situation?

The oscillatory behavior of the Z(t)Z(0) correlation function for small h_E can be best understood by looking at this quantity directly in the toric code limit ($h_E, \lambda \to 0$), where it can be analytically computed. There, the Z(0) operator creates two magnetic excitations, leading to an energy difference of 4J compared to the ground state and hence the correlation function oscillates with this frequency,

$\langle Z(t)Z(0)\rangle = \exp[-4iJt]$. In our setting, we create the excitation at boundary qubits. As a consequence, one of the magnetic excitations is outside the system and does not contribute to the energy difference. Therefore, for the boundary qubits Q_1 and Q_2 considered in Fig. 4d, the correlation function oscillates with half the frequency: $\langle Z(t)Z(0)\rangle = \exp[-2iJt]$. The amplitude of the oscillations is 1, however, in our experiment decoherence leads to decay in time, gray shaded area in Fig 4d. Note that the real part of the measured correlation function $\mathrm{Re}[Z(t)Z(0)]$ is shown in the main text (Fig. 4d) and the imaginary part $\mathrm{Im}[Z(t)Z(0)]$ is shown in the supplement (Fig. S17b). Another limiting behavior, which is directly understood, is the limit of $h_E \to$ Infinity. There the oscillations disappear since the ground state is an eigenstate of Z and thus the correlations remain constant over time.

Action 1: To clarify this issue, we have added the following text to Supplementary Information IV.F : "*These oscillations can be understood in the toric code limit (λ , $h_E = 0$), where Z couples the ground state to an excited state with +4J higher energy. This leads to $\langle Z(t)Z(0)\rangle = e^{-4iJt}$ for a bulk qubit. However, for qubits on the boundary, such as Q_1 and Q_2, one of the magnetic excitations is outside the system and does not contribute to the energy difference, therefore $\langle Z(t)Z(0)\rangle_{edge} = e^{-2iJt}$. The imaginary part of $\langle Z(t)Z(0)\rangle$ is suppressed as h_E is increased to $h_E = 1.4$. This is also expected, since as $h_E \to \infty$ the ground state becomes an eigenstate of Z.*"

Action 2: Additionally we have amended the main text to read, *"The correlation function $Re[\langle Z(t)Z(0)\rangle]$ experiences coherent oscillations from the energy difference between the WALA state and the first excited state, while the small expectation value of $\langle Z(0)\rangle$ suppresses $S_{ZZ}(t)$."*

I recommend publication of the paper after these comments are addressed.

We thank the reviewer for their time and effort considering our manuscript. We are delighted to hear their recommendation.

Referee #2 (Remarks to the Author):

The authors of the manuscript present a fascinating experiment demonstrating the ability to study the time evolution of a particular two-dimensional lattice gauge theory, specifically the dynamics of a string excitation in two different regimes, probing the confinement-deconfinement regimes. The manuscript presents protocols to prepare the initial ground state at different system parameters, the initial excitations, the time evolution, and -very importantly- the necessary measurements to probe the physics under study.

The manuscript is clearly written and the reported experiment demonstrates a clear advance in the path pursued by different groups of theoretical and experimental physicists in the last decade, struggling to bring to the lab the original vision put forward by Feynman forty years ago. This experiment, which first proves a digital quantum simulation of an excitation string dynamics in a two-dimensional LGT, could be a landmark in this direction. However, there are a number of shortcomings that prevent me from recommending this manuscript for publication. I would encourage the authors to consider them:

We thank the reviewer for recognizing that our work constitutes a "landmark" result in the field of quantum simulation of LGTs. We also appreciate the constructive feedback that they have provided. Below we address the comments, point by point.

1. The main objection is that, despite the beauty of this experiment, nothing is learned from the quantum computing part of it. In other words, all that is shown can be simulated (as done by the authors) straightforwardly by classical computers. I do not underestimate the complexity of this experiment, nor the theoretical analysis and protocol developments needed to reach this result. I fully appreciate their beauty. However, here we are discussing if this route can, at least eventually, lead to some strong demonstration of quantum computation usefulness.

Here it is completely unclear: there are no statements on the difficulties of classical simulations while scaling the system size, on the continuum limit that will give effective useful physical results, etc. How is the entanglement scaling in these simulations? is it possible to simulate them classically, e.g., with tensor network? There have been recent results in higher-dimension on ground states and dynamics. My feeling is that the system remains pretty much low-entangled all the time here. There are instead unsupported statements here and there "prohibitively hard to simulate" in the conclusions and ""generally beyond the perturbative limits" in the abstract.

Without extensively discussing this point, I do not think that this is a Nature paper.

Solving the quantum many-body problem (for example the 2D lattice gauge theory considered here) is one of the main open challenges in natural sciences. We will first list a couple of reasons why, at least eventually, when scaling to late times and large systems, the problem we consider is indeed prohibitively hard to simulate with classical resources and after that we will show numerical simulation results.

- Even *ground states* of 2D lattice gauge theories are generically difficult to access with classical algorithms. In the limit of large system sizes, numerical quasi-1D MPS ansätze fail, due to the area law scaling of the entanglement entropy. Generalised 2D TNS ansätze capture the area

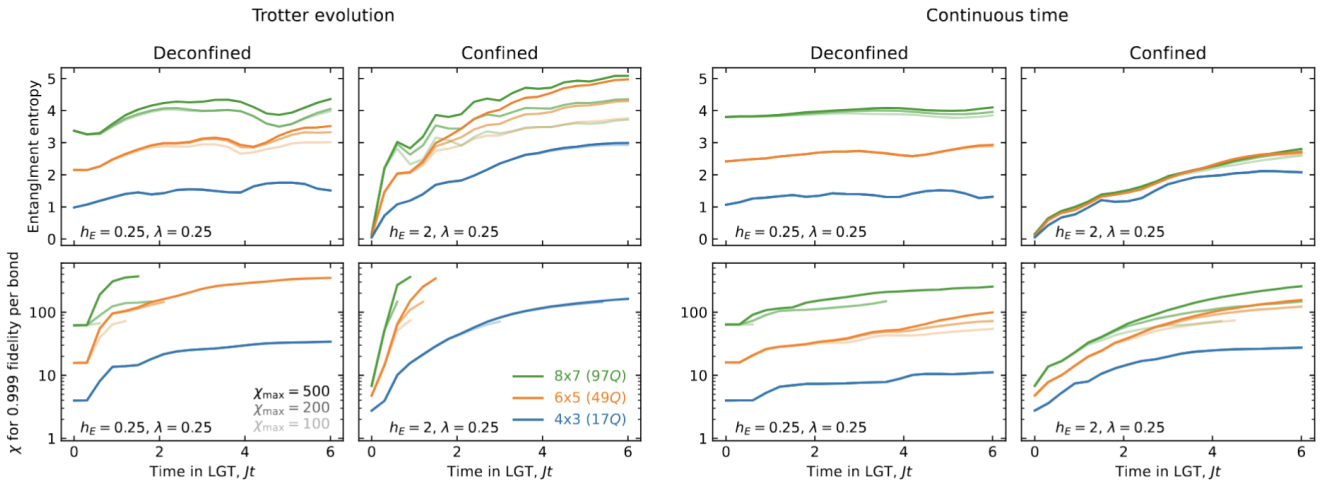
law, however, an exact contraction is #P hard and thus further approximation are required to avoid an exponential scaling of complexity. These approximations cause numerous instabilities and in particular fail at least in the vicinity of quantum phase transitions. In our model, this concerns in particular the transition from the confined to the deconfined phase.

- Characterizing the *dynamics of excitations* on top of ground states is even more difficult, as additional entanglement is generated in the dynamics. Especially near the phase transition, the intricate dynamics of large bound states are expected to yield large entanglement in the scaling limit and thus make TNS simulations prohibitively expensive.

Due to these reasons, studying the dynamics of 2D lattice gauge theory is currently not possible with classical resources in the scaling limit. Having said this, of course, there has been a lot of progress in pushing numerical simulations to their limits to gain some understanding of finite size systems. In order to assess where the limits of such classical simulations are, we have numerically computed the dynamics of excitations using a quasi-1D MPS ansatz, which we believe is currently the best available classical approach to study dynamics of excitations in our system.

In the new Section VI of the Supplementary Information, we simulated the dynamics of the experiment shown in Fig. 4a in the main text, where we have a string stretched across the system with a single bump in the middle. The resulting entanglement entropy of the MPS during the time evolution and the minimum bond dimension required to achieve 0.999 fidelity on each bond (relative to the state with the indicated χ_{max}) is shown in the figure below. We considered both the case where we start from the ground state as obtained from DMRG and simulate the exact time evolution (right panels), and the case where we start from the WALA state and evolve with a Trotter step dt=0.3 (left panels). We chose one set of parameters in the deconfined phase (h_E=0.25, λ =0.25, first and third columns), and another set of parameters in the confined phase (h_E=2.00, λ =0.25, second and fourth columns). We also considered different system sizes, starting from a system with 4x3 vertices as in the experiment (blue) and increasing the system size to 6x5 vertices (orange) and 8x7 vertices (green). In the deconfined phase, we can see that the entanglement entropy grows linearly with the system size, since the 1D MPS ansatz with fixed bond dimension cannot capture the area law entanglement in two dimensions, and slowly with time. As expected, the minimum required bond dimension scales exponentially with system size. In the confined phase, the initial entanglement entropy is smaller but quickly grows with time because the string excitation costs more energy than in the deconfined phase. The simulation of the experimental circuit shows even faster growth of entanglement, stemming from the WALA state not being an exact eigenstate and the Trotter approximation of the time evolution. This leads to the creation of further excitations on top of the initial state, which increase the entanglement.

The numerical simulations thus underline the preceding discussion, that generally simulating the dynamics of 2D lattice gauge theories for large systems and late times is eventually beyond the capabilities of current tensor network methods. In the continuous-time case, in the deconfined regime, the 8x7 system took 26 hours to simulate using 180 CPUs. Because the algorithm scales as χ^3 and we see that the required bond dimension grows exponentially with system size, we expect this to quickly become intractable with these methods as we scale up the system size.



2. Related to the previous point: how does the entanglement scale? The numerical simulations already made by the authors contain this information. If this is very low entangled, then the difficulty of this experiment is actually overestimated.

In gapped two-dimensional ground states the bipartite entanglement entropy scales as the area law $S_{vN} = c\, L_y$, where L_y is the length of the cut. States with a few excitations still have zero energy density, and hence they are captured by area law states as well, $S_{vN} = c'\, L_y$ but with a larger prefactor $c' > c$, which renders numerical simulations difficult on practical grounds, see above.

On the experimental side, entanglement is not the limiting factor. The experimental difficulty rather lies in maintaining coherence for sufficiently long times.

Action: In response to the referee, we have added numerical results to the Supplemental Information VI showing the entanglement dynamics of the time evolved state and discuss the general difficulty of simulating the time evolution of these systems with classical resources along the lines of the response above. Furthermore, we extended the discussion after introducing the WALA state: "*The WALA circuit is equivalent to a mean-field ansatz for the dual Ising model and has the advantage of being very easy to optimize classically even for larger system sizes in contrast to the actual ground state (see Supplementary Information III A). The resulting quantum state, $|\psi_0\rangle$, is then used as a*

low-energy-density initial state for simulating the dynamics. In classical simulations, the dynamics are hard to capture due to the fast entanglement growth (see Supplementary Information VI)."

3. The authors refer to the initial excitations as particles in the abstract. However, given that they create them by applying an X operator, they are similar to classical charge, not to any particles (system excitations). The authors should clarify this throughout the text.

In the infinite-field limit ($h_E \to \infty$) the string operator which creates the electric charge excitations is given by $W(\ell) = \Pi_{e\in\ell} X_e$, where ℓ is an open string. In that limit, the string is an eigenstate of the system and has no dynamics. For finite electric fields ($h_E \neq \infty$), the initial state created by applying a string $W(\ell) = \Pi_{e\in\ell} X_e$, on the ground state is not an eigenstate of the system. However, this operator will still have a significant weight on the elementary excitations of the system, as also demonstrated by the observations of our experiments. Therefore, we created a non-equilibrium initial state that allows us to probe the string fluctuations dynamically.

In order to clarify this point, we emphasize that our initial state creates a non-equilibrium configuration that has a large overlap with the elementary excitations.

Action: To avoid confusion, we have modified the text any time flips of the A_v operators were referred to as "particles". Opting instead for "electric charge excitations" or simply "charges."

We have also added the following sentence to the main text, "*We prepare a pair of electric excitations on neighboring sites in the center of the system by applying a single Pauli-X on top of the WALA state (Fig. 3a). This operator creates a non-equilibrium initial state which has a significant overlap with the elementary excitations and allows us to probe deconfinement-confinement dynamics of the charges.*"

4. Another possible strong point of the paper is to compare their circuit with already reported digital quantum computations: they state that each trotter step counts for about 117 CZ gates, it seems to me they perform O(10) trotter steps, that is O(1000) two-qubit gates. How does this compares with previous experiments?

We thank the reviewer for suggesting that we highlight the cutting edge technical achievement of our experiments. Recently, superconducting qubit chips have allowed for proof-of-principle measurements for quantum error correction [arXiv:2408.13687] and have been compared to tensor network contraction methods in random circuit sampling [Morvan *et al.* Nature 634, 328 (2024)]. The high fidelity of the experimental operations was key for the success of this project (Fig. S1).

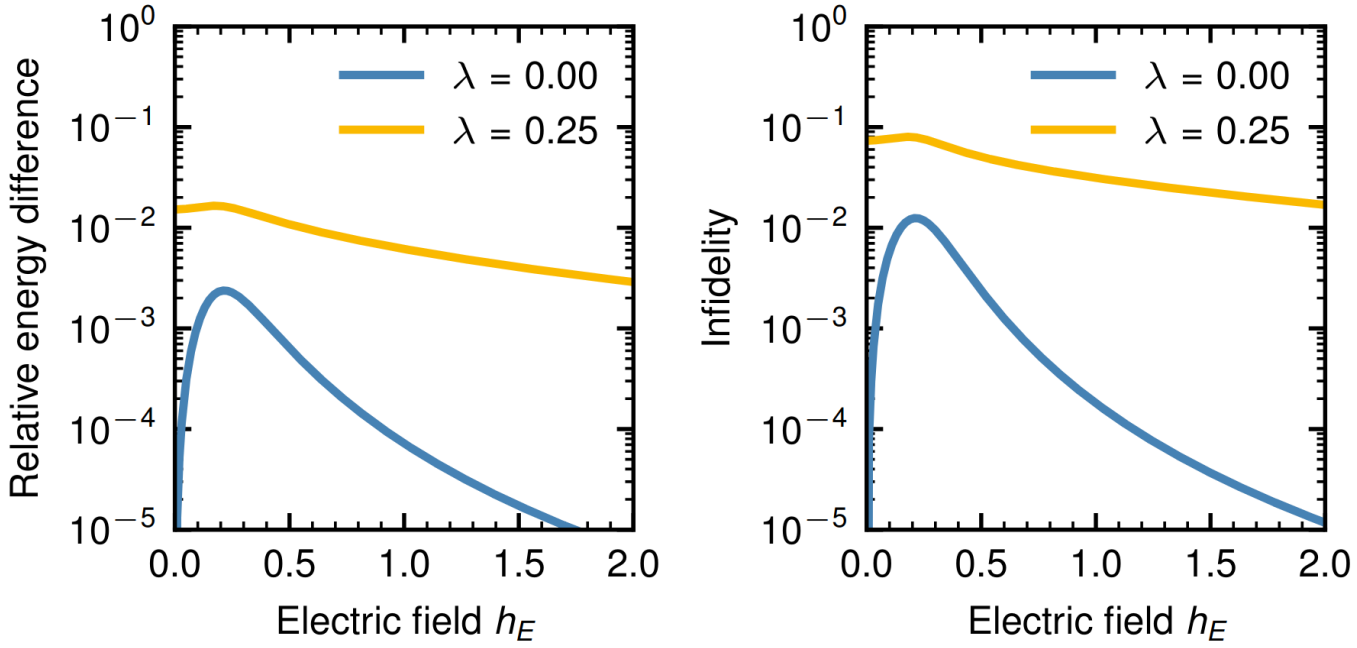
Let us compare our experiment to circuit measurements, where Morvan *et al.* report results after 880 entangling gates, spread over 32 layers focusing on cross-entropy benchmarking approaches. Since we are considering physical observables instead of fidelity measurements, we can even reach slightly deeper circuits. In particular, we measure appreciable signals after ~1075 entangling gates, spread over 77 layers (9 Trotter steps). Further, we note that signal damping from depolarizing noise does not limit us from going to longer times. Rather, we set a cap of 10^6 total shots per randomized compiling instance as shown in Fig. S3c. The trends of the effective depolarization probabilities indicate the total gate count and number of layers could be increased. However, we expect significant finite size effects for times $\gtrsim 4$ (in the units used in the work). So any increase in evolution time should be paired with scaling the system up on a larger chip.

Action: In Supplementary Information Section II.B. we add the following text, "*Executing state preparation and 9 Trotter steps, our experiment uses ~ 1075 total entangling gates. Since we are focusing on local observables we measure appreciable signals in these circuits, whose depth is slightly larger than state-of-the-art random circuit sampling experiments, which focus on cross entropy benchmarking [13]*.

5. The preparation of the ground state is discussed in great detail. However, I cannot find a theoretical prediction of the protocol in terms of fidelity (again possible in the numerical simulations). I see that the final energy is near to the expected one, but how well is the preparation, theoretically? A possible test is to prepare the state and see if this evolves in time. Has this test been performed in the QPU?

While the WALA state approximates the ground state, it is not optimized to provide optimal overlap. Since we are principally interested in the dynamics of charges and strings, it suffices to start with a suitably low energy initial state on top of which we can prepare excitations. However, we agree with the reviewer that including the fidelity of the WALA state is helpful context for the reader.

Action 1: We have added the following figure (Fig. S9) and relevant discussion to the Supplemental Material, that shows the theoretical fidelity of the WALA state. For finite coupling, λ, the infidelity of the state approaches 10% near the transition for our system size. However, this ansatz does a good job in preparing a low-energy initial state for the study of dynamics across the entire parameter regime that we explore.



Regarding the evolution of the WALA initial state on the QPU, Fig. S18a shows the behavior for h_E=1.4 and λ ={0, 0.25, 0.5}. In the raw data, it appears that the electric excitation density increases linearly (Fig. S18c top panel). However, much of this density comes from decoherence on the device; we see that beige crosses (λ = 0) in the top panel Fig. S18c increases linearly. However when λ=0, all the A_v terms commute with the Hamiltonian, so they cannot have any intrinsic dynamics. After rescaling accounting for depolarizing errors, we see the probability of electric vertex excitations is nonzero, but small (~2% for λ=0.25 and ~5% for λ=0.5). The experimental values agree well with numerical simulations.

Action 2: To highlight this point, we have added the following sentence to the Supplemental Information, "*At the same time, the low probability of electric charge creation in the WALA state by time t = 2.7 (~ 2% for λ = 0.25 and ~ 5% for λ = 0.5 after depolarization mitigation) confirms that this ansatz is a good low-energy initial state with only a small amount of intrinsic dynamics.*"

Minor points:

The legend of Fig. 3.d is incomplete in my pdf.

We thank the referee for their careful review of our figures.

Action: The legend for Fig. 3d has been amended.

Despite the fact that I can probably guess the meaning, it is not clear to me what is represented in Fig. 3e in the four different columns: are these different results of measurement? What is the probability written in a given position with the "measurement symbol". More explanations are needed there.

We thank the reader for identifying the ambiguity.

Action: We have made our description in the figure caption more specific:

($|\psi_+\rangle$) and destructively ($|\psi_-\rangle$) at short distances, respectively. **d**, The separation of the excitations as a function of time for different magnetic fields h_M. **e**, Spatial maps of the probability that an electric vertex $A_{i,j}$ is excited, conditioned on the electric vertex $A_{\mathscr{A}}$ being excited at time $t = 3.5$: $P(i,j|\mathscr{A})$. The different columns correspond to different $A_{\mathscr{A}}$, indicated by the boxed $\mathscr{A}$ symbols. The unconditioned probability that $A_{\mathscr{A}}$ is excited is written below the $\mathscr{A}$ symbol. The color scale represents $P(i,j|\mathscr{A})$ for the all $A_{i,j}$. The top and bottom rows shows results for $h_M = 0$ and 2, respectively. We implement dynamical decoupling and randomized compiling to mitigate control errors as well as idle dephasing (see Supplementary Information II A).

Referee #3 (Remarks to the Author):

This is the first work presenting a quantum simulation of the dynamics of a full-fledged lattice gauge theory on a system of a size sufficient to study the confinement of matter particles. For this reason, it will have a major impact on the community of quantum simulations and it sets a fundamental milestone for the growing community studying quantum theories inspired by high-energy physics through quantum technologies.

Several experimental works already addressed 2D lattice gauge theories on quantum processors, by using ion traps (see, for instance, arXiv:2310.12110), superconducting quantum processors (see, for instance, Phys. Rev. D 103, 094501 (2021); Phys. Rev. D 106, 074502 (2022)) or quantum annealers (for instance Phys. Rev. D 104, 034501 (2021)). To the best of my knowledge, however, only arXiv:2310.12110 addressed both matter and gauge degrees of freedom, and it did it on very small lattices, such that I would consider that result based on ion setups only a proof of the efficient realization of the building blocks for a 2D LGT, rather than a more complete simulation.

The current manuscript from the Google collaboration, instead, demonstrates a quantum simulation of the full Z_2 lattice gauge theory on a truly 2D system with a clearly defined and visualizable geometry with open boundary conditions. I believe that the combination of the relative theoretical simplicity of the obtained results and their clear visualization makes this manuscript outstanding and of doubtless value from both the scientific point of view and as a showcase for quantum computation applications.

For these reasons, I believe that the results presented in this manuscript meet the impact and innovation criteria of Nature. There are, however, a few parts in the presentation of lattice gauge

theories that I find not so rigorous. Additionally, several aspects of the presentation and the description of the data should be improved. Moreover, in the following, I will also suggest providing some additional data and materials in the supplementary information. Consequently, a revision of the manuscript is necessary.

We thank the reviewer for contextualizing our work so well.

Let me first address the main semantic issues about the general presentation of the analyzed lattice gauge theory.

1. Since the abstract, the authors refer to the coefficient h_M in the Hamiltonian (1) as "the magnetic field". I do not find this kind of nomenclature rigorous from the point of view of lattice gauge theory and I believe it can cause major confusion for the readers, since the magnetic field of the theory is related to the plaquette operators. The coefficient h_M is related instead to the coupling constant of the theory and I suggest changing its denomination. I also disagree with the statement "The h_M terms denote a magnetic field on each link that creates magnetic flux excitations.": these terms describe the energy density of the electric field, and they are commonly referred to as electric field terms. Therefore, it is very confusing to refer to h_M as magnetic field. Moreover, I also find very confusing the following remark "The electric field terms h_E generate hopping of matter fields located at adjacent vertices, as mediated by a gauge field on their connecting link.". I agree that these terms describe the matter hopping, but they commute with the magnetic plaquette term and it is unacceptable and contrary to the standard characterization of lattice gauge theory to define them as "electric field terms" (which, as I wrote, are instead the h_M terms). In general, the authors must adopt a more conventional and rigorous nomenclature, following the standard set by lattice gauge theories. Fig.1 needs to be changed accordingly, as well as a few other parts of the text (for instance lines 339, 370).

We thank the reviewer for pointing out this issue. We agree that the naming convention that we had chosen is unconventional and in a sense incorrect.

Action: Following reviewer's suggestion, we have changed the coefficient of the σ^Z term from $h_M \rightarrow h_E$, since as pointed out above this term denotes the 'external electric field' at the gauge qubits. We also agree with the referee's second point, and we have changed the gauge-matter coupling to λ, which is often used in the literature. The text and figures are adopted accordingly. We thank the referee again for pointing this inconsistency out.

2. Again concerning the terminology, in the caption of Fig. 1 the authors indicate as a "pure LGT" the theory after removing the matter degrees of freedom via the gauge constraints. This is wrong. A pure LGT typically refers to models in which $h_E =0$ and J_E diverges, irrespective of the choice made to represent it. The fact that a model is a pure lattice gauge theory is a physical statement (matter particles have infinite mass), not a statement that depends on the choice of the adopted degrees of freedom. The representation of the Hamiltonian obtained by gauging out the matter degrees of freedom is usually called "unitary gauge" (see, for instance, Fradkin's textbook "FIELD THEORIES OF CONDENSED MATTER PHYSICS" (second edition), Chapter 9.10).

Action: In order to be consistent with the conventional notation, we have rephrased the figure caption to remove any references to a pure LGT. The relevant sentence now reads, "*The local gauge structure can be leveraged to eliminate the matter field and arrive at an effective theory involving gauge fields only (right).*"

Besides these general aspects about the description of the Z_2 lattice gauge theory, in the following I list several points that the authors should address concerning the presented data.

3. Concerning Fig. 2c, I find it curious that the yellow data (product state) display a simulated curve with a worse energy than the experimental results. Can the authors comment about it? Are their simulations to obtain the three solid lines including any sort of noise or errors in the computation? or do they assume an exact implementation of all gates without errors?

Sec. IVA of the supplemental information mentions that the experimental results include a correction for readout errors, but I do not understand how correcting the data may lead to lower values than the theoretical predictions.

We thank the reviewer for catching the initially perplexing result of a lower energy from the measured state prepared on the quantum processor, compared to the expectation from circuit simulation.

Firstly, the circuit simulation is not assuming any noise on the device. Secondly, data presented in Fig. 2c includes uncorrelated readout error mitigation (as described in Supplementary Information IV.A). There we explain that "*To mitigate readout errors, we construct individual readout confusion matrices, … as determined by sampling random bitstrings*." Therefore, the mitigated data (Fig. 2c) is calculated using error from random bitstrings, which can cause a small bias when measuring the polarized state. This is because the measurement of random bitstrings allows for the possibility of some crosstalk error from neighboring qubits that are excited, which is not present when measuring the all |0> state. In other words, a small bias is introduced from the slight deviation from our assumption of uncorrelated readout errors.

To verify that the lower energy of the mitigated data is an artifact of the readout-mitigation procedure, we can examine the "*Raw*" energy achieved from simply summing expectation values of each term in the Hamiltonian before any readout mitigation. This was plotted in Fig. S10a. There we see that indeed the measurement of the energy from the "*Polarized*" initial state is slightly higher than the ideal simulation (as expected).

Action: We have added "*after readout error mitigation (Supplementary Information IV A)*" to the Figure caption of Fig. 2 as well as the following sentences in the supplement, "*We note a bias for the polarized state, which arises because of a small deviation from perfectly uncorrelated readout errors of the all |0> state. This results in the readout-mitigated values being slightly lower in energy than the noiseless simulation.*"

4. The result is counterintuitive because, for the yellow curve, it seems that the experimental errors after the initialization to the trivial product state would be beneficial to approach the ground state.

To emphasize the point, the primary error for the "*Polarized*" state is not from state preparation. The preparation of this state is a null circuit, since |0> is the ground state of each qubit, i.e. the initialized state that results from passive qubit reset with very high fidelity. Rather the dominant error in this case comes from the measurement itself. The all |0> state (or the all |1> state) is in some sense pathological, in that it is expected to maximize any deviation from uncorrelated readout errors.

5. Do the authors confirm that the yellow data points are obtained measuring plaquettes and links directly on the state |0>^N irrespective of the value of h_M?

Since there is only one polarized initial state, all of the yellow points in Figs. 2c and S9 are drawn from the same dataset and the energy is simply calculated with the different Hamiltonian values. The same holds for the toric code initial state.

Action: The following sentences were added to the Supplementary Information to clarify this issue, "*Since only the WALA state depends on h_E, the data for the polarized and toric code initial states was collected only once. The different points correspond to the total energy calculated from the measured observables using different values of h_E.*"

6. Again concerning Fig. 2c, it may help the reader to provide an estimate of the critical value of h_M in the thermodynamic limit from the literature. For $h_E=0$, this value would be slightly below 1/3 from the known critical field of the Ising model (see, for instance Phys. Rev. E 66, 066110 (2002).). It should not change much for $h_E=0.25$, and it may be interesting to compare it to the crossing between the yellow and blue data. This indication is also useful for the reader when comparing the three regimes of the string dynamics in Fig. 4c.

In the limit of λ=0, the critical point is located at h_E = 0.328474(3), see Phys. Rev. E 66, 066110 (2002); obtained from the mapping to the transverse field Ising model. We do not know the precise critical point for λ=0.25, however, for λ=0.3, we have recently estimated the phase transition point using the FM string order parameter and have found it to be at h_E = 0.335(1); see arXiv:2402.00127. Hence, in the thermodynamic limit the phase transition point is indeed located at approximately h_E ~ ⅓. For our finite size system, the crossing of the energies of the polarized and the toric code are at approximately ⅓ in FIg. 2c. This crossover field is easily determined analytically as the ratio between the number of magnetic plaquettes and the number of gauge sites. For our finite system, this turns out to be h_E = 6/17 ~ 0.35. However, this crossing shifts in the thermodynamic limit to h_E ~ ½, which is significantly different from the true transition point. Hence a direct comparison seems not pertinent.

Action: We have added the expected critical field in the thermodynamic limit to the main text, "*These dynamical signatures support the onset of a confining potential near the Ising critical point, which in the thermodynamic limit is located between hE ~ 0.33 and 0.34 depending on λ as obtained from numerical ground state studies [40, 41, 47, 48].*"

7. Regarding the data displayed in Fig 2d, the color scale does not allow for a clear quantitative understanding. I think it would be important to add in the supplementary information a graph that displays the average expectation values of the plaquettes, the Z operators on bulk and boundaries and the A vertex terms as a function of h_M (the latter unless postselection automatically discards all =-1 instances).

We thank the reviewer for making this suggestion to characterize the initial state preparation more quantitatively. We note that we do not post-select on the vertex operators, but their average expectation values lie between 0.953 and 0.957 for all values of h_E.

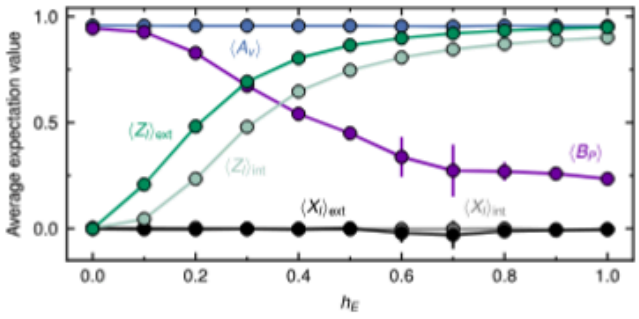


FIG. S11. **Average expectation values of terms in the $\mathcal{H}$ for WALA.** Expectation values of each term in the Hamiltonian, averaged over all equivalent vertices/plaquettes/links. For the expectation values of links, $\langle Z_l \rangle$ is expected to behave differently for qubits on the edge and those in the bulk (Supplementary Information III B). Therefore, we plot the average expectation values for these two sets of qubits separately. The expectation values are plotted for electric vertices $\langle A_v \rangle$ (blue markers), magnetic plaquettes $\langle B_p \rangle$ (purple markers), external edge links $\langle X_l \rangle_{\text{ext}}$ and $\langle Z_l \rangle_{\text{ext}}$ (dark green and black markers respectively), and internal bulk links $\langle X_l \rangle_{\text{int}}$ and $\langle Z_l \rangle_{\text{int}}$ (light green and gray markers respectively). Error bars correspond to the standard deviation over all vertices, plaquettes, or links.

Action 1: We have added a figure to the supplementary information that shows the average expectation values of all of the equivalent terms in the Hamiltonian (Fig. S11):

Action 2: In addition, the following sentences have been added to the supplementary information, "*After carrying out this uncorrelated readout error mitigation procedure, the corrected values for each term in the Hamiltonian are extracted and plotted in Fig.~2d. Averaging equivalent vertices, plaquettes and links allows for a detailed understanding of how each term in ⌘ depends on h_E (Fig. S11).*"

8. Concerning Fig. 4c there is a detail that I do not fully understand. In the panel at $h_M=0.1$ and $t=0$, $S_{ZZ}(t=0)$ is sizeably different from zero not only on the sites affected by the applied X string, but also on neighboring sites. At t=0, though, S_{ZZ} is just the expectation value of Z(0), and, for sites that are not involved in the X string, it must match the expectation value of Z after the WALA initialization. Did the authors check whether the values plotted in the panel $h_M=0.1$ and $t=0$ for the sites that do not belong to the X string match the expectation value of the WALA? Similar data are plotted in Fig. 2d, but $h_M=0.1$ is missing there and I cannot make a comparison. Therefore I would ask the authors to show them in their resubmission letter. Indeed, I find quite surprising the difference of the <S_ZZ(t=0)> value between the boundaries and the central site for $h_M=0.1$.

We thank the reviewer for carefully evaluating the S_{ZZ} data when h_E=0.1. The results are understandable by analyzing the difference between the expectation values of Z(0) for interior and exterior link qubits, presented in the new Fig. S11. When h_E=0.1, the exterior edge qubits have a significantly higher Z expectation value (0.208) compared to the interior bulk qubits (0.045). This is expected since $Z_{edge} \sim \cos(\theta)$, while $Z_{bulk} \sim \cos^2(\theta)$ and for h_E=0.1, $\cos(\theta)$ is still close to zero (Supplementary Information III.B.).

For the exterior edge qubits that are not along the trajectory of the initial string excitation, we would expect their initial expectation values to be ~0.2, which is what we observe in Fig. S16c for qubit Q_2 (h_E = 0.1). We would then expect qubit Q_1 to have an expectation value ~-0.2, since this qubit is acted on by an X gate as part of the string excitation. This is again consistent with the data in Fig S16c. For interior bulk qubits, we would expect their expectation values to be ± 0.05, which is also consistent with Fig. 4c.

9. Regarding the data of Fig. 4d and the dynamics of the string at low h_M, the authors write "In the deconfined phase, $h_M = 0.1$, applying an X-string on top of the WALA sequence does not create a string excitation. Thus, the correlation function $S_{ZZ}(t)$ quickly trends towards zero.". I must admit that I find this description too simplistic. The behavior of S_{ZZ} remains close to zero in this case simply because the sort of normalization <Z(0)> is very low close to the toric code limit. However, the first two panels of Fig. 4d reveal that the correlation function <Z(t)Z(0)> has a marked coherent oscillation which is stronger than the cases with large h_M (this is even more evident in Fig. S15b).

Therefore I agree that in this deconfined phase there is no creation of an electric flux string. However, I disagree on the fact that "the correlation function $S_{ZZ}(t)$ quickly 'trends' towards zero". It seems to me that it acquires lower values since the beginning just because of the adopted <Z(0)> normalization, but, otherwise, the dynamics trend seems to me more oscillating and coherent than the large h_M case. This is why I dislike its description as 'quickly trends towards zero'.

We thank the reviewer for identifying this confusing sentence. They are correct that S_{ZZ} is suppressed by the <Z(0)> term.

Action 1: We have changed this sentence to read, "*The correlation function $Re[\langle Z(t)Z(0)\rangle]$ experiences coherent oscillations from the energy difference between the WALA state and the first excited state, while the small expectation value of $\langle Z(0)\rangle$ suppresses $S_{ZZ}(t)$.*"

Action 2: Additionally, we have added the following text to to Supplementary Information IV.F. to describe these coherent oscillations.: "*These oscillations can be understood in the toric code limit (λ, $h_E = 0$), where Z couples the ground state to an excited state with +4J higher energy. This leads to $\langle Z(t)Z(0)\rangle = e^{-4iJt}$ for a bulk qubit. However, for qubits on the boundary, such as Q_1 and Q_2, one of the magnetic excitations is outside the system and does not contribute to the energy difference, therefore $\langle Z(t)Z(0)\rangle_{edge} = e^{-2iJt}$. The imaginary part of $\langle Z(t)Z(0)\rangle$ is suppressed as h_M is increased to h_E = 1.4. This is also expected, since as $h_E \to \infty$ the ground state becomes an eigenstate of Z.*"

10. In line 330, the authors write: "Surprisingly, the string is still moving freely around the top qubits despite the large string tension." As the authors mention afterwards, this is mostly due to length-preserving deformations of the string and I do not find it so surprising. I would avoid writing "Surprisingly".

We thank the reviewer for noticing this slightly misleading use of the work "surprisingly."

Action: We have changed this sentence to " *However, the string…* "

11. In line 354, the authors write "the string initial state possesses considerably more charge particles". I disagree with this sentence. It is the state obtained after the initial insertion of a string that at time t=2.7 possesses considerably more charge particles. The initial string state, if I understand correctly, should have no charge particles instead (=1 from the WALA initialization and the boundary to boundary string commutes with all A's in the bulk). Am I wrong?

We thank the reviewer for pointing out this ambiguity in our text. Their analysis is correct.

Action: We have changed the sentence to, "*However, for finite h_E values, the time-evolved string initial state possesses considerably more charges*."

12. From the caption of Fig. 5c, one understands that $h_M=1.4$, however it is not explicitly stated (it is written that the grey region corresponds to this value). I advise to revise the caption and add this indication more explicitly.

Action: We have added the h_E value to the figure caption.

13. Concerning the discussion in the supplemental information III, when discussing Eq. S8 the authors write: "Intuitively, such a suppression of large loops is what one might expect when adding an onsite Z-field to the toric code Hamiltonian".

This intuitive picture I think can be misleading for a couple of reasons:

A) Let us take a Z-field with the same sign convention of Hamiltonian (1). If $h_M<0$, then the states |0> are favored and long loops are certainly suppressed when $h_M<-1$. Therefore theta must decrease for negative h_M. If this is what the authors had in mind, one needs to specify that the Z-field would correspond to $h_M<0$.

B) For $h_M>0$, the |1> states are favored. This does not automatically correspond to the fact that theta must increase beyond Pi/2. On the contrary, Fig. 2b and Sec. III C of the SI show the other way round. The point is that in order to polarize the spins along |1>, we should apply the B operator only to either even or odd plaquettes [in the thermodynamic limit], not to all of them. By applying all the plaquette operators (theta-> Pi in Eq. S8) one would flip twice all the spins but the ones on the boundaries, resulting in a single large loop around the boundary.

For these reasons I think that the description of Eq. S8 should be slightly modified to avoid intuitive but potentially misleading statements.

We thank the referee for pointing out the subtleties of adding an onsite field with different signs to the Hamiltonian. Indeed, we only had in mind the case where we add the term $-h_M \Sigma_{links} Z_i$ with $h_M > 0$, so that this term favors qubits in the |0> state and suppresses qubits in the |1> state.

Action: We have clarified this restriction in the description of Eq. (S8).

14. In the caption of Fig. S4 b, replace "(dotted line)" with "(dashed line)"

Action: We have implemented this suggestion..

15. In Sec. IVA of the supplemental information, the authors use the ket notation |\phi> "to be the probabilities of each computational basis state...". I find it a misleading notation, since the ket is used for pure quantum states. Do the authors mean that |\phi> constitute a diagonal density matrix of the system instead? Or maybe they could represent \phi simply as a vector of probabilities.

We have indeed used a non-standard notation. We intend to only describe a vector of probabilities.

Action: For clarity we have updated the notation:

such that $\mathcal{R}_{\mathcal{O}}$ is a $2^n \times 2^n$ matrix. Then, taking $\vec{P}$ to be the probability vector in the computational basis of the qubits in the support of $\mathcal{O}$, we perform the readout error mitigation:

$$\vec{P}_{\text{mitigated}} = \mathcal{R}_{\mathcal{O}}^{-1} \vec{P}_{\text{measured}} \tag{S21}$$

16. In Sec. IVA of the supplemental information, the authors refer to the Loschmidt echo of the state preparation circuits. I find this part very technical and not self-contained. I invite the authors to include additional references or a more detailed explanation of the procedure. A brief definition of the energy of the Loschmidt circuit would be useful.

We emphasize that the Loschmidt echo and global depolarization mitigation are not used in the main text (Fig. 2) and that the primary novelty of our work is the visualization of dynamics of a LGT system in two spatial dimensions. For completeness, though, we included these error-mitigated results of the WALA energy in the supplement.

Action 1: To better describe the error mitigation using the Loschmidt echo in Fig. S10, we have added the following references [Mi *et al.* Science 374, 1479 (2021), O'Brian *et al.* Nat Phys 19, 1787 (2023)].

Action 2: We also have amended the text in Supplementary Information IV.A. to clarify the point that the measurement of energy requires measurements in the Z-basis and X-basis, which yields two different Loschmidt echo circuits. The connection between the Loschmidt circuit measurements and the global depolarization mitigation is also discussed in more detail.

17. In Fig. S11e, the color scale (blue vs black) is not specified [in Fig. 3 it is ok instead].

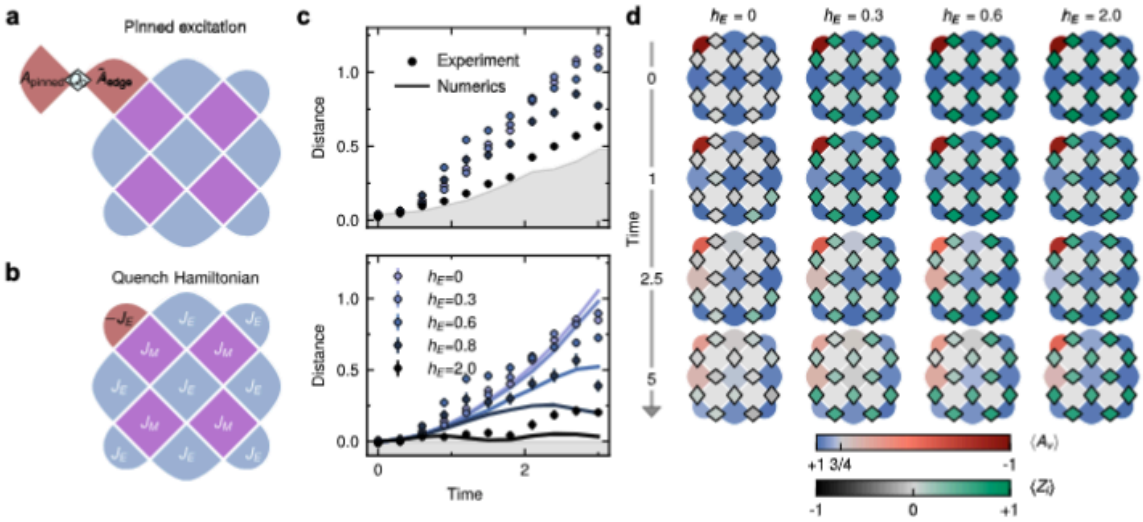
We have added the h_E correspondence to the blue/black traces in Fig. S13e (previously Fig. S11e).

18. Concerning the analysis of the dynamics of the single particle in Sec. IVD1, the authors interpret their data in Fig. S12d in terms of the motion of the charge excitation, that, in their opinion gets slower for strong confinement. I tend to think, instead, that a more rigorous interpretation would be based on screening rather than confinement and particle velocity. The initial conditions defined in Fig. S12a with a pinned qubit Q_0 are consistent with an electric flux line entering the system through Q_0. If we consider the ground state with this boundary condition, depending on the value of J_E and h_M (for $h_E>0$) we would experience a crossover from weak screening at $J_E >> h_M$ to a strong screening when $h_M >> J_E$. When screening is weak, an electric flux line propagates on average over many links in the lattice before being screened by matter excitations. On the contrary, when

screening is strong, electric charges localize close to the top left corner to screen the incoming electric flux line. Now, the data presented concerns the dynamics of the system rather than its ground state properties. However it is likely that the long-time behavior (bottom row of Fig. S12d) is displaying the formation of the screening cloud associated to the incoming electric field line. For small h_M the screening cloud would be larger than the system size, thus the average electric field percolated on the whole system. For large h_M, instead, the electric field is screened and remains localized on the top left corner. In this respect, it may be instructive to plot also the expectation value of Z over all the sites of the system (as done in Fig. 2d). In this way a reader could easily spot whether and how the electric field percolates across the system. As a more general note, the authors focus mostly on small values of h_E, where the distinction between confined and deconfined phases is sound. However, usually, confinement [meant as an exponential decay of the Wilson loop with the area of the loop] is a property of the pure lattice gauge theory [$h_E=0$]. For $h_E>0$, as the authors demonstrate, there are string breaking mechanisms that are a manifestation of screening. This is why, at large h_M and $h_E>0$ I think that the concept of screening would be more appropriate than confinement.

We thank the reviewer for offering this very physical interpretation of the data acquired with a pinned external charge in Supplementary Information IV.D.

Action: Following their suggestion, we include new data showing $\langle Z_l \rangle$ for all links in the spatio-temporal maps in Fig. 14d (below). This data is in agreement with the notion that charges act as a screening cloud of the externally threaded flux in the confined phase. We have added a discussion of this complimentary perspective in the Supplementary Text.



19. In line 424, the following sentence is grammatically wrong "In the topological phase, the motion of the electric excitations at the end points of the string are not confined."

We thank the reader for catching this grammatical error. We have removed "*the motion of*" to make the sentence correct.

In conclusion, the results presented in this work constitute a decisive step forward in the field of quantum simulations of lattice gauge theories. This work constitutes indeed a milestone for quantum simulations as it convincingly demonstrates the possibility of performing the simulation of the dynamics of interacting 2D quantum many-body problems on a system size going beyond the single building blocks that have been investigated in previous works. However, some parts of the exposition of this work are written in an unsatisfactory way (especially concerning the description of the elements of the lattice gauge theory) and, in general, a rigorous review is necessary before a final evaluation of the manuscript.

We thank the reviewer for their detailed comments which helped us to significantly improve and clarify the presentation of our work. We are thrilled that they consider our work "a decisive step forward in the field."

We thank the referees for considering our augmented manuscript and are delighted that they have both recommended publication. Below we provide a short discussion of their final "Remarks to Author."

Legend
Authors: Black.
Referees: Green.

Referee #2:

The authors have adequately addressed my previous comments, thus I recommend the publication of the manuscript in its present form.

We are very thankful to the referee for their comments that have allowed us to improve our manuscript. We are excited that they now recommend publication in the present form.

However, let me mention that I do not agree with some of the comments that are present in their reply (not in the manuscript though),
let me mention two major points:

1. The authors seem unaware of recent developments in tensor network simulations. To date, there are simulations of 2+1 and even 3+1D ground states of LGT and in the dynamics of two-dimensional systems of pretty large (16x16) spin systems, mostly based on Tree Tensor Networks. Thus, the presented simulations are far from being state-of-the-art. However, they show that the entanglement grows in the initial state in such a way that simulations will be in any case very hard. The fact that what is measured is sensitive to this entanglement is still not completely clear (see next point), but what is shown is enough for me to justify the importance of this experiment.

We thank the referee for pointing out the tree tensor network simulations that have recently been applied to large systems in both 2+1D and 3+1D. We emphasize that in our work we have used matrix product state simulations to precisely study the non-equilibrium dynamics of string excitations on top of the ground ground state and WALA state respectively. We could not find related simulations in the existing literature.

Different tensor network algorithms may lead to slightly different performance, which needs to be tested in detail for each case. The general trend for scaling of the matrix product state simulations, reported in the supplement, will help the reader to assess the complexity of the studied problem, as suggested by the referee.

We thank the referee again for drawing our attention to the above-mentioned works.

2. The statement "On the experimental side, entanglement is not the limiting factor. The experimental difficulty rather lies in maintaining coherence for sufficiently long times." is kind of

misleading: exactly as for tensor networks, the ability to maintain entanglement is the limiting factor, if not as it is a form of coherence to be maintained. And that is more difficult to maintain than that of single qubits. If this is not evident in this experiment, it means that it can be reproduced with tensor network methods efficiently with a very low bond dimension.

We thank the referee for raising the distinction between maintaining coherence and entanglement. Our experimental results point to the strong effect that errors have on the observables of interest, which are well described by a global depolarizing model after implementing randomized compiling. We interpret these errors as a source of decoherence of the quantum state, which in turn will affect the entanglement properties. We leave to future work the detailed analysis of how well reduced bond dimensions in tensor network simulations recreate the coupling of our quantum processor to the environment in this particular simulation of Z_2 lattice gauge theory.

Referee #3 (Remarks to the Author):

The authors answered in full detail to the points I raised. I believe that the new version of the manuscript is now rigorous and its presentation is now consistent with the standard terminology of lattice gauge theory.

This makes the text readable and enjoyable for a broad readership.

We thank Professor Burrello for their comments, which greatly improved the readability of our manuscript.

I confirm my first impression: I believe that this work is a milestone for the field of quantum simulations of many-body systems. It showcases for the first time the possibility of investigating in detail the dynamics of an intricate two-dimensional quantum many-body problem on a quantum processor beyond previous proofs of principle. The model addressed, namely the Z2 lattice gauge theory, is a crucial model for different areas of physics, but the results presented here have a significance that goes beyond this specific model. They show a way forward for remarkable advances in many fields of research including the study of models of interest for condensed matter physics, particle physics and quantum chemistry.

For these reasons I fully support the publication of this work in Nature.

We are thrilled that the reviewer recognizes the broad impact our work will have and are very heartened to hear that they fully support the publication of this work in Nature.

At the level of minor details, I suggest the authors to consider these final adjustments:

1) The authors solved the issues related to the lattice gauge theory terminology. However, I suggest that they improve even further the rigorousness of their abstract. In line 10 they write "As the electric field is increased,...". The meaning of this sentence, however, should be "As the

coupling constant associated with the electric field energy density is increased,...", since the electric field is the operator Z and not the constant h_E. Therefore, I would suggest to rephrase that sentence. A possible and relatively short correction could be "As the electric field coupling constant is increased".

We have implemented the suggested change to the abstract.

2) In line 450 of the supplemental material there is a typo: "...is is threaded in."

We have corrected this minor typo.