Quark-Hadron Transition and Entanglement

Berndt Müller
Department of Physics, Duke University, Durham, NC 27708-0305, USA

Andreas Schäfer
Institut für Theoretische Physik, Universität Regensburg, D-Regensburg, Germany

The dual holographic description has enjoyed many successes in explaining fundamental properties of the early stages of relativistic heavy ion collisions up to the formation of a minimal-viscosity quark-gluon fluid. However, there have been few attempts to extend its application beyond this stage. Here we explore the prospects for such an extension beyond the time of hadronization. Our discussion makes use of recent insights into the duality of entanglement properties of field theory states in the edge of Anti-de Sitter space and non-trivial topologies of horizons in the bulk, often referred to as ER = EPR duality. We discuss this topic from the point of view of heavy-ion phenomenology, review several relevant concepts, and map out a path toward combining them into a comprehensive, at least semiquantitative description of relativistic heavy ion collisions. We outline possible next steps in this direction.

I. INTRODUCTION

One of the fundamental questions of high energy physics is how apparently thermal behavior emerges in relativistic heavy ion collisions. A large number of measurements show such thermal behavior and are best described when one assumes that the produced transient Quark-Gluon Plasma (QGP) is in a fully thermalized state. Yet the nuclear reaction, which occurs in isolation, cannot produce a state of high von Neumann entropy because the reaction is governed by the laws of quantum chromodynamics (QCD) which assures unitarity of the S-matrix. The transient QGP and all the hadrons measured in the final state must exist in a highly entangled quantum state that mimics the properties of a thermal ensemble that is in interaction with an external heat bath.

One is immediately led to the question whether there are any observable differences between a thermal ensemble and such a highly entangled excited state. The answer is not as simple as one might think as local observables and, more generally, observables that depend on only a fraction of the complete many-particle final state have difficulty differentiating between these two alternatives.

Whether entropy can be generated in an isolated system in a process that is governed by time reversal invariant laws or not, depends on the definition of generalized, non-equilibrium “entropy”. Several proposals for such a notion exist and, unsurprisingly, the answer depends on the entropy definition used. In addition, as the deviation from equilibrium is time dependent, the most suitable notion of entropy may yield an answer that depends on the time elapsed after the start of the collision.

The complexity of this issue is illustrated by the fact that the Anti-de Sitter space-conformal field theory (AdS/CFT) duality allows to map the production of a QGP to the formation of a black hole in five-dimensional Anti-de Sitter (AdS5) space. Any description of entropy production will share the subtleties of the black hole information paradox expressed most succinctly by the Page curve [1,2]. For any pure quantum state of an isolated many-particle system the von Neumann entropy is zero and remains zero under unitary time evolution. If, however, only part of the system is observed, i.e. if the information residing in the rest of the system is discarded, the reduced system appears thermal with the corresponding thermal entropy. Although the Page curve was specifically introduced to resolve the information problem of black holes, it is understood to be the characteristic property of any generic many-body quantum system [3]. Its existence has been confirmed experimentally in various experiments, see, e.g., Fig. 4 in Kaufman et al. [4].

Our goal here is to outline a program which aims at describing essential information theoretical properties of high energy nuclear collisions, including those properties resulting in a Page curve, using concepts from AdS/CFT duality. Obviously this is an ambitious goal, and in the present article we can only outline the first steps. We will provide a list of open issues at the end of this work that may help define a systematic pathway towards a more precise understanding of how the apparent thermal properties of the final state of a relativistic heavy ion collision can be reconciled with the unitarity of the S-matrix in QCD.

There are many quantum field theories for which holographic duals are known to exist. In the realm of gauge theories the original prototype, the $N = 4$ supersymmetric conformal Yang-Mills theory, is by far the most widely explored example, because for large $N_c$ and large $’t$ Hooft coupling $\lambda$ it has a tractable holographic dual, i. e. classical supergravity in AdS space. The exact holographic dual for QCD is not known, although models that share important features with QCD have been constructed [5,6]. Therefore, any attempt to apply AdS/CFT phenomenology to QCD is problematic from the start and can only be successful when combined with other field theory and string theory techniques which allow to correct for some of the differences. For example, the violation of conformal symmetry can be treated per-
turbatively in QCD, see e.g. [7,9], the finiteness of the \'{t}Hooft coupling can be corrected by string perturbation theory, see e.g. [10], and the calculation of $1/N_c$ expansions on the field theory side has evolved into a broad research field of its own [11].

However, QCD has also certain properties, like the confinement/deconfinement transition, which cannot be treated perturbatively and thus require a more incisive approach. As we are interested in the hadronization of the quark-gluon plasma we cannot avoid studying holographic models that incorporate a confinement/deconfinement transition into their basic framework through some versions of the Hawking-Page transition [12]. Examples of such models include the $\mathcal{N}=4$ super-Yang-Mills theory on a compactified spatial volume, such as $S^3$, models including a compactified additional dimension of AdS space [13], and models that break conformal invariance through the introduction of a dilaton field [14,15]. Although there are reasons to believe that such holographic models can describe many relevant properties of QCD qualitatively, it is far from clear whether they can be refined to describe these relevant features with sufficient precision to allow for a quantitative comparison with experimental data.

In view of these caveats the success of holographic methods in modeling certain aspects of relativistic heavy ion collisions came as a pleasant surprise. They have been used extensively to understand rapid thermalization [16–22], the fast transition to hydrodynamic flow [23,24], the small value of the specific shear viscosity [25], and much more. In fact, the last two decades have witnessed such continuous progress that today a much larger assortment of powerful techniques and insights exists than ever before. For example, semiclassical quantum gravity and applications to black hole physics have recently conceptually resolved the information paradox [26,30]. All these encouraging successes motivate us to try to understand aspects of the breakup of the QGP into hadrons, which have only rarely been explored using holographic methods.

Our effort to expand the use of holographic duality to a description of the complete heavy ion collision is also motivated by the recognition that a field theoretical description of hadronization on the quantum level is far too complex to be tractable. Without tractable holographic descriptions that capture salient features of this transition, even if they cannot reproduce all aspects quantitatively, we simply lack a perspective to fully understand the many-body quantum mechanics of a heavy ion collision.

Let us start by reminding the reader of a few salient phenomenological properties of heavy ion collisions that any comprehensive description must address:

- There exist two distinct ways in which the particles that are ultimately detected, hadrons, are produced from the evaporating QGP. Beginning immediately after formation of the QGP “fireball”, hadrons are emitted from its surface. As the expanding QGP cools down to the pseudocritical temperature $T_c$, hydrodynamics ceases to be valid, and the fireball converts into hadrons throughout the remaining volume. The thermal model [35] describes the production rates of all hadrons and nuclei remarkably well with a single universal temperature parameter (the chemical freeze-out temperature $T_{\text{CF}} = 156.0 \pm 1.7$ MeV [36]) that is determined to percent-level accuracy. Because of the closeness of this value to the central temperature of the QCD crossover transition determined by lattice QCD ($T_c = 157.8 \pm 0.6$ MeV [37] or $T_c = 156.5 \pm 1.5$ MeV [38]) $T_{\text{CF}}$ is interpreted as QCD deconfinement temperature. The precision of the coincidence between $T_{\text{CF}}$ and $T_c$ itself is remarkable because the QCD crossover transition, as measured by the chiral susceptibility, has an intrinsic width of $15 \pm 1$ MeV [37].

- The fact that the thermal model works so well is quite astonishing in view of the strong interactions among the constituents of the system before and after the hadronization transition. A focus of the discussion on this point is provided by the hypertriton $\Lambda^\Lambda$, whose yield in Pb-Pb collisions at LHC agrees with the prediction of the thermal model for the universal temperature $T_{\text{CF}}$ despite the fact that its binding energy is only 0.4 MeV, i.e. much smaller than $T_{\text{CF}}$, and its size is comparable to the size of the collision system. Recently, however, it was found [39] that this is no longer true for a smaller collision system (p-Pb) with earlier hadronization time, adding to the conundrum of hadronization in heavy ion collisions.

- Heavy ion collisions create thermal conditions that resemble those prevailing in the early universe, but the time scales are vastly different. In the cosmos, different particle species fell out of thermal equilibrium at different times due to cosmic expansion and cooling. For example, photons and neutrinos have different cosmic background radiation temperatures. The same could be expected for heavy ion collisions but is not observed for Pb-Pb collisions. As we argued in [40] one framework that can potentially provide an explanation for the lack of thermal differention is the Eigenstate Thermalization Hypothesis (ETH) [41,42]. The ETH may explain not only the success of the thermal model but also the difference between Pb-Pb and p-Pb collisions as the lifetime of the QGP in p-Pb collisions may be too short to apply the full ETH formalism, which is based on energy eigenstates (see Section [43]).

Our manuscript is organized as follows. Section [44] contains some general remarks about entanglement entropy in the context of relativistic heavy ion collisions, and we introduce the concept of the Page curve. Because
the ETH plays an important role in our arguments we present a brief review in Section III of those aspects that are relevant for us. In Sections IV and VII we will further discuss some relevant aspects of heavy ion collisions and provide more details on the properties mentioned above. In particular, we will argue in Section VII that the usual numerical AdS calculations in the Poincaré patch are insufficient to obtain the Page curve. In Section V we will review these holographic model calculations of the early stages of heavy ion collisions up until apparent thermalization. In this time interval they contain the relevant physics and give phenomenologically satisfactory results.

To extend the holographic description to later times we propose two distinct hadronization mechanisms: surface radiation and bulk hadronization. In Section VII we adapt recent ideas [13, 44, 45] for the Hawking-Page phase transition to bulk hadronization. To describe the quantum entanglement of hadrons emitted from the fireball surface requires a very different approach. As hadrons cannot exist within the QGP fireball, its surface acts as a boundary for hadron states. In Section VIII we discuss how hadron emission can be treated as the production of particle-hole states at the QGP surface in analogy to Hawking radiation from the black hole horizon. This allows us to adapt the arguments for the formation of "islands" within black holes which are entangled with the outgoing Hawking radiation to the QGP fireball. The increasing entanglement of the remaining QGP with the emitted hadrons is holographically mapped onto the black hole evaporation process. This will make it possible to transcribe the arguments for the occurrence of a Page curve to the QGP hadronization process. We synthesize these considerations into a schematic holographic description of the hadronization process in Section X.

Our arguments are admittedly speculative, but they have the advantage that their underlying assumptions can be confirmed or refuted by detailed calculations. In Section X we present a limited list of possible future studies.

II. ENTANGLEMENT ENTROPY

In high-energy heavy ion collisions one never measures the complete final state but detects only a subset of all emitted particles. Detector constraints typically limit the detection to hadrons in a certain rapidity interval $\Delta y \sim 1$ that is much smaller than the full beam rapidity range $-2y_{\text{beam}} \leq y \leq 2y_{\text{beam}}$. Because the subsystem $\Phi_{\Delta y}$ within the kinematic range $\Delta y$ is entangled with the rest of the final state $\Phi_{\Delta y}^{'\prime} = \Phi \setminus \Phi_{\Delta y}$ outside this rapidity window, it carries a certain entanglement entropy:

$$S(\Phi_{\Delta y}) = \text{Tr}[\rho(\Phi_{\Delta y}) \ln \rho(\Phi_{\Delta y})].$$

(1)

Here

$$\rho(\Phi_{\Delta y}) = \text{Tr}_{\Phi^{'\prime}}[\rho(\Phi)],$$

(2)

denotes the reduced density matrix of the final state constrained to the rapidity window $\Delta y$. $\Phi$ denotes the complete final state, and $\rho(\Phi)$ is the complete final-state density matrix. Unitarity of the S-matrix in QCD ensures that $S(\Phi) = 0$, if the initial state is a pure quantum state, which is satisfied to a high degree of precision in the collision of two heavy ions.

For $\Delta y \ll 2y_{\text{beam}}$ the entanglement entropy $S(\Phi_{\Delta y})$ grows with the size of the rapidity window. Boost invariance at midrapidity, realized at high collision energy, dictates that $S(\Phi_{\Delta y})/\Delta y$ is approximately independent of $\Delta y$ for small rapidity windows. The complementarity law of quantum information also dictates that $S(\Phi_{\Delta y}) = S(\Phi_{\Delta y}^{'\prime})$. Accordingly, when $\Delta y$ exceeds half the total rapidity range of the final state, $\Delta y > y_{\text{beam}}$, the entanglement entropy has to decrease again and approach zero when $\Delta y \rightarrow 2y_{\text{beam}}$. The entanglement entropy associated with a rapidity window thus follows a Page curve [1, 3], as sketched by the red solid line in Fig. 1.

![Fig. 1. Page curve (red solid line) for the entanglement entropy associated with a rapidity window $\Delta y$.](image_url)

Particle detectors are not capable of measuring all information contained within $\Delta y$, e.g., they generally do not record the relative phases among all emitted particles. Although many modern detectors are capable of recording the full set of kinematic correlations among hundreds or thousands of detected particles, a large part of this information is usually discarded in the data analysis and only selected few-body correlations are retained [17]. The reduction to a subset of data within the rapidity window $\Delta y$ further changes the concept of entropy into a notion that has been aptly called "entropy of ignorance" about the full details of the final state [18]. What is commonly called "entropy" of the final state of a relativistic heavy ion collision is derived from the single-particle distribution in momentum space; here we denote this single-particle entropy as $S_{sp}(\Delta y)$. The single-particle entropy grows at a different rate than the entanglement entropy (see Duan et al. [18] for a detailed analysis in a slightly different context) and continues to grow as $\Delta y$ exceeds $y_{\text{beam}}$, as illustrated by the blue dashed line in Fig. 2.

A number of estimates of the single-particle entropy have been presented in the literature. The estimates are based on the measured spectra of produced particles and calculate the entropy of an equilibrated Boltz-
mamm (or Bose/Fermi) gas with the same momentum and mass distribution. This is usually done per unit of rapidity (at midrapidity), resulting in a measured value of $dS_{sp}/dy = S_{sp}(\Delta y)/\Delta y$, which corresponds to the slope of the dashed line in Fig. 2. At top RHIC energy, $dS_{sp}/dy \approx 5,000$ \cite{49,50}; at top LHC energy $dS_{sp}/dy \approx 11,500$ \cite{51}.

III. THE EIGENSTATE THERMALIZATION HYPOTHESIS

ETH \cite{41,43} posits that even a single energy eigenstate can appear like a thermal system for most, if not all, observables of practical interest. It is thought to apply generally to systems that exhibit chaos at the quantum level. The ETH represents a generalization of Random Matrix Theory (RMT) \cite{52} and, like RMT, is formulated in terms of the matrix elements of an observable $A$ in the energy basis:

$$ A_{\alpha\beta} = \langle E_\alpha | A | E_\beta \rangle. $$

The ETH goes beyond RMT by assuming that the off-diagonal elements of the matrix $A_{\alpha\beta}$ can be written as normalized elements $R_{\alpha\beta}$ a random matrix modified by factors involving the microcanonical entropy $S(E)$ and the spectral function $f(E,\omega)$ of the system:

$$ A_{\alpha\beta} = A(E)\delta_{\alpha\beta} + e^{-S(E)/2}f(E,\omega)R_{\alpha\beta}, \quad (4) $$

where $E = (E_\alpha + E_\beta)/2$ and $\omega = E_\alpha - E_\beta$. The exponential factor implies that the individual off-diagonal elements are suppressed relative to the diagonal elements for a system with a high energy level density, i.e., with $S(E) \gg 1$. Then the diagonal matrix elements alone determine the thermal average of the observable:

$$ \langle A \rangle_T = Z(T)^{-1} \int \frac{dE}{E} e^{S(E)-E/T} A(E), $$

$$ Z(T) = \int \frac{dE}{E} e^{S(E)-E/T}, \quad (5) $$

with corrections that are exponentially suppressed.

On the basis of the assumption (4) it is possible to show that the system prepared in an energy eigenstate behaves like a thermal system. A few particularly noteworthy properties are \cite{43}:

- The long-time average of an observable equals the thermal average: $\overline{A} = \langle A \rangle_T$.
- The quantum fluctuations of $A$ are equal to the thermal fluctuations with corrections of $O(1/N)$, where $N$ is the number of degrees of freedom of the system.
- The time correlation function of finite-time expectation values $\langle A \rangle_t$ obeys a Kubo relation with the function $f(E,\omega)$ as the spectral density.

These properties assure that the system, when monitored through the observable $A$, is indistinguishable from a thermal system.

The Thouless energy $E_{Th}$ is defined as the energy difference $\omega$ below which the factor $f(E,\omega)$ in (4) can be replaced with a constant for most observables, and ETH reaches the RMT limit. In a dynamically evolving system, the relevant range of energy differences $\omega$ is inversely related to the evolution time by the uncertainty relation. It is often assumed that the behavior of a chaotic quantum system for times longer than the Thouless time $t_{Th} = \hbar/E_{Th}$ is described by RMT. However, it was found in numerical studies of discrete quantum systems that the time until full ETH behavior is established, $t_{ETH}$, can be much longer, depending on the operator under investigation \cite{53,54}. In fact, can be parametrically longer than $t_{Th}$ by a factor proportional to the system size.

The intuitive interpretation of this latter time scale is that the establishment of quantum entanglement requires causal connection across the whole system, and full entanglement is only reached asymptotically. The most important prediction of ETH for a fully entangled system is that if only a small part of the density operator enters an observable, i.e., if the trace is taken over more than half of the states, the observable behaves as for a system in contact with a heat bath. It is tempting to assume that this property explains the experimental success of the thermal model in relativistic heavy ion collisions mentioned in the Introduction.

The operator dependence of the onset of ETH behavior might actually allow us to understand why the measured yield of hypertritons in p+Pb collisions differs from the
The various time scales pose a challenge to any holographic description spanning the entire collision.

However, upon closer inspection one realizes that all these time and length scales can be traced back to three primary scales related to the underlying field theory (QCD), and the initial and spatial boundary conditions:

- The energy density $\varepsilon_0$ initially deposited in the collision; in the glasma model $\varepsilon_0 \sim Q_s^4$, where $Q_s$ is the gluon saturation scale \[^{[57]}\].
- The initial transverse extent $R_T$ of the collision region, which depends on the nuclear radii $R$ and the impact parameter $b$.
- The QCD confinement scale $\Lambda_{\text{QCD}}$ reflected in the nucleon radius $r_N$ and the pseudocritical temperature $T_c$.

There are three analogous independent scales that can be present in the AdS/CFT dual description:

- The energy density $\varepsilon_0$ initially deposited in the shock wave collision.
- The initial transverse extension $R_T$ of the collision region, which depends on the transverse width of the shock waves and the impact parameter $b$.
- In holographic models, which contain an extraneous length scale, this can set the scale for a transition of the Hawking-Page type \[^{[58]}\] that marks the boundary between a confined and a deconfined phase. For example, in the strong coupling limit of the simplest holographic model dual to $\mathcal{N} = 4$ super-Yang-Mills theory \[^{[59]}\], such a transition occurs when the dual theory is considered in global AdS space \[^{[12]}\]. Similar transitions exist in more realistic holographic models of QCD even when considered on the Poincaré patch.

When the dual description is constrained to the Poincaré patch of AdS space \[^{[60]}\], as it is necessary to describe a heavy ion collision, the third scale must be introduced by some modification of the holographic dual that breaks the conformal symmetry, if hadronization is to be described. Examples include models with imbedded D-branes \[^{[6,61]}\] and the Scherk-Schwarz compactification of a dual theory with an additional dimension considered by Witten \[^{[5]}\] and used in the study of Aharony et al. \[^{[13]}\].

In recognition of these primary scales we can divide the progression of a relativistic heavy ion collision into four stages, which are summarized in Table IV:

| Stage | Description |
|-------|-------------|
| 1     | Initial state very far from thermal equilibrium, but the time available to reach it is limited by the rapid longitudinal expansion of the created quark-gluon system. Hadronization, a confinement-deconfinement transition, takes place during this short time span and drastically changes the properties of the system. This transition is usually modeled as an early "chemical freezout", to be distinguished from complete kinematic freeze-out, which is assumed to occur later. The very large momenta of the colliding nuclei imply large Lorentz factors $\gamma$ in the center-of-mass (CM) system, up to $\gamma \sim O(1000)$. Accordingly, the transverse extent of the colliding nuclei at the initial collision instant in the CM system is much larger (by a factor $\gamma$) than their longitudinal thickness. In addition, transverse fluctuations are substantial, of order 50% for the energy density \[^{[55]}\], on length scales between the inverse saturation scale $1/Q_s \sim 0.2$ fm and the nucleon radius (0.7 fm). This generates, in combination with the velocity of light $c$ for ballistic processes or the velocity of sound $v_s = c/\sqrt{3}$ for diffusive processes, additional time scales.
| 2     | The various time scales pose a challenge to any holographic description spanning the entire collision. The question is whether the time of hadronization, $t_H$, is larger or smaller than $t_{\text{ETH}}$. As the hypertriton wave function is much more extended than that of the proton one should expect a strong form-factor suppression of the corresponding matrix element, resulting in an especially large $t_{\text{ETH}}$. At the same time $t_H$ is especially small for the proton size fireball produced in p+Pb. Both effects suggest that ETH behavior has not yet established itself when hypertritons are formed in p+Pb collisions, which would explain the observed suppression compared to the thermal model.
| 3     | The QCD confinement scale $\Lambda_{\text{QCD}}$ is assumed to occur later. The phenomenology of thermalization in heavy ion collisions is characterized by multiple different time scales. The initial state is very far from thermal equilibrium, but the time available to reach it is limited by the rapid longitudinal expansion of the created quark-gluon system. Hadronization, a confinement-deconfinement transition, takes place during this short time span and drastically changes the properties of the system. This transition is usually modeled as an early "chemical freezout", to be distinguished from complete kinematic freeze-out, which is assumed to occur later.
| 4     | The very large momenta of the colliding nuclei imply large Lorentz factors $\gamma$ in the center-of-mass (CM) system, up to $\gamma \sim O(1000)$. Accordingly, the transverse extent of the colliding nuclei at the initial collision instant in the CM system is much larger (by a factor $\gamma$) than their longitudinal thickness. In addition, transverse fluctuations are substantial, of order 50% for the energy density \[^{[55]}\], on length scales between the inverse saturation scale $1/Q_s \sim 0.2$ fm and the nucleon radius (0.7 fm). This generates, in combination with the velocity of light $c$ for ballistic processes or the velocity of sound $v_s = c/\sqrt{3}$ for diffusive processes, additional time scales.

We now explain the reasoning behind Table IV and describe the roles the primary scales play in the different
TABLE I. The different stages of a heavy ion collision (HIC) and their proposed holographic modelling.

| Stage | Time | HIC phenomenology | Holographic dual model |
|-------|------|-------------------|------------------------|
| Stage I | \( t \leq 1 \text{ fm/c} \) | **Hydrodynamization**<br>hydrodynamic attractors lead from transient large amplitude fluctuations to viscous hydrodynamic expansion | **Numerical AdS simulations**<br>entropy production mapped by apparent horizon<br>\( t \leq 0.2 \text{ fm/c} \) equilibration without fluctuations<br>\( t \leq 1 \sim 2 \text{ fm/c} \) equilibration with fluctuations |
| Stage II | \( t \leq O(15) \text{ fm/c} \) | Expansion of the QGP fireball<br>hadronization at QGP surface | Collision of localized shocks [62]<br>(Sections [V and VII] |
| Stage III | \( t \approx O(15) \text{ fm/c} \) | **Bulk hadronization**<br>chemical freeze-out | Smoothed **HP transition**<br>(Sections [VII and IX]) |
| Stage IV | \( t \geq O(15) \text{ fm/c} \) | Expansion of HRG<br>kinetic freeze-out | Entangled hadrons<br>network of ER bridges (Section [VIII]) |

**Stage I:** At very short times neither the confinement scale nor the transverse extent of the reaction region are important. During this stage the local energy density far exceeds the critical density \( \varepsilon_c \) at the deconfinement transition, and information about the finite transverse size has not spread widely because of causality. The initial energy density \( \varepsilon_0 \) is the only relevant scale both on the QFT and holographic side of the duality. Both statements may not apply to very small collision systems, such as those produced in proton-proton or proton-nucleus collisions.

**Stage II:** The resulting hydrodynamic QGP expands during the second stage. During this expansion a small fraction of the QGP hadronizes on the surface of the expanding fireball. By construction, this surface is at the confinement-deconfinement transition temperature \( T_c \), which is of the order of \( \Lambda_{QCD} \). Thus, the expansion dynamics of the QGP fireball depends on the two scales \( \varepsilon_0 \) and \( R_T \), while the hadronization processes at its surface depend, in addition, on \( \Lambda_{QCD} \). Therefore, the latter aspect cannot be described in the standard framework of AdS modeling on the Poincaré patch, while the hydrodynamical expansion can.

**Stage III:** At the end of the second stage the remaining fireball becomes thermodynamically and hydrodynamically unstable and completely hadronizes into a quantum entangled hadron gas. The duration of this process depends on the nature of the phase transition. In models with a very large number of colors and a first-order transition the conversion from a gauge plasma to a hadron (glueball) gas takes an extended period of time because the formation rate of color singlet hadrons is small [13]. In QCD, where the transition is a rapid, but smooth crossover, the process is deemed to be so fast that it is usually described as instantaneous. Because time evolution in QCD is unitary no thermal entropy is produced, but ETH applies to this stage, explaining the validity of the “thermal” model. (For this to apply the time from the start of the collision until this third stage must be long enough for global ETH behavior to be established [53], which may not apply to very small collision systems. Numerical simulations in the AdS Poincaré patch are also unable to describe this stage adequately.)

The transition from (slow) surface hadronization to (rapid) bulk hadronization means that the Page curve for a heavy ion collision has an asymmetric shape. Before the final hadronization transition only a small fraction of the QGP (in Section [VI] we will argue for roughly 20%) has decayed, while the rest decays on a much shorter time scale during Stage III. Figure 3 shows a schematic sketch of the Page curve for this scenario.

**Stage IV:** The final stage of a heavy ion collision, which features an expanding cloud of entangled hadrons, resembles the final stage of a decaying black hole which comprises an expanding cloud of entangled photons without a black hole remnant at the center. Again, numerical simulations limited to the Poincaré patch of AdS space are insufficient to capture this entanglement.

While each stage of a heavy ion collision has a well-motivated holographic dual, it is not clear how well holography can model a heavy ion collision in quantitative terms. This is not a concern for the very early stage,
S

\[ S \]
\[ t \]

FIG. 3. Schematic representation of the Page curve for the hadronizing quark-gluon plasma in a relativistic heavy ion collision. The growth rate of the entanglement entropy of the emitted hadrons is low during Stage II because hadrons are only emitted from the surface. The growth rate then increases quickly during Stage III due to bulk hadronization before dropping equally rapidly after half of the QGP has hadronized. At the end of the hadronization transition, the entropy again equals the initial entropy because all hadrons are entangled with each other.

when the quark-gluon system is far away from the confinement scale, and a holographic description can be justified by its approximate conformal symmetry and the limited sensitivity of hydrodynamics to details of the initial state. It is unclear whether tractable holographic models can provide for a quantitative description at later times where more time scales are relevant.

Even if the holographic model cannot mirror a real heavy ion collision quantitatively, in can potentially provide important conceptual insights with parametric validity. As an example, we consider the final state when the whole system consists of hadrons that are freely streaming toward the detectors. There are two possibilities: Either the hadrons are in thermal equilibrium and described by a density matrix with high entropy, which is the standard assumption made in heavy ion physics, or the hadrons form a highly entangled, nearly pure quantum state with very little entropy, as we argue here. In the first case there exist well established techniques that permit detailed theoretical predictions, whereas any such calculation requires additional assumptions in the latter case. We will argue in the next Section, building on work by van Raamsdonk et al. [63], that the holographic dual of an entangled state of many hadrons differs only in subtle ways from the holographic dual of a QGP fireball and that, therefore, a holographic description can extend smoothly beyond the hadronization transition.

Another important question is whether the fireball reaches ETH-type behavior before hadronization. This cannot be true for all observables, as the following argument shows: As already noted, transverse variations in locally conserved quantities equilibrate by diffusion. The associated time scale \( R_s^2/D \), where \( D \) is the relevant diffusion constant in the QGP, is identical to the one introduced in the previous Section as Thouless time \( t_{Th} = 1/E_{th} \). It is useful to estimate this time constant for a typical volume of QGP created in a heavy ion collision. The typical value of a diffusion coefficient in QCD is \( D \sim (\pi T)^{-1} \) [64]. With \( R_s \approx 5 \text{ fm} \) and \( T \approx 300 \text{ MeV} \) one finds \( t_{Th} \sim 120 \text{ fm}/c \), which is much longer than the lifetime of the QGP.

The crucial question in this context is, how long it takes until few-particle observables that are amenable to experimental measurement are well approximated by their thermal value. The time it takes for entanglement to propagate throughout the fireball provides for a lower bound. Quantum information propagates with velocity \( v_E < c \) yielding this information time scale to be of order \( R_s/v_E \). At strong coupling holographic models in \( d = 4 \) dimensions give \( v_E \approx 0.62 \) [65], which means that the speed of information transport is similar to that for energy transport, which is given by the speed of sound \( c_s = c/\sqrt{3} \). This means that entanglement can spread through the entire fireball before the onset of hadronization.

In this work we discuss certain ideas that may ultimately make it possible to extend the successful holographic description of heavy ion collisions beyond the Page time. The crucial feature of any such attempt is the preservation of full quantum coherence which, in turn, requires the addition of novel concepts to the holographic description that become relevant at later times in the collision. As we discussed above, this implies that numerical simulations of AdS gravity must be expanded to cover the global AdS geometry, not just the Poincaré patch.

The underlying idea is that a heavy ion collision proceeds from the highly entangled many-body wave functions of the colliding nuclei to a highly entangled QGP which evolves continuously into a highly entangled multi-hadron state but never to a truly thermal state with large von Neumann entropy. In other words, the collision can be viewed as a unitary S-matrix mapping between the initial and final hadron states. In this language, it is completely irrelevant that in some intermediate state certain local observables look thermal. It is ultimately the interaction of the emitted hadrons with the detectors, which act as incoherent environment, that creates the thermal entropy. This view implies that the establishment of full ETH behavior and thus the success of the thermal model is far less of a mystery, because \( t_{Th} \), is much shorter than the time that elapses until final-state hadrons hit the first detectors.

Elucidating the microscopic processes which govern thermalization of many-particle quantum systems is relevant for many subfields of physics. For example, in [66] it was argued that thermodynamics can be based on the assumption of a pure state wave function of the universe
VII. is feasible in practice. This will be discussed in Section graphic description of entanglement and ETH behavior and, therefore, the experimental verification of ETH behavior and Hawking radiation for QED systems suggest that QCD systems behave in a similar manner. How- ever, while these insights are intellectually satisfying, their practical utility depends on whether or not a holo- graphic description of entanglement and ETH behavior is feasible in practice. This will be discussed in Section VII.

V. HOLOGRAPHIC SIMULATIONS OF EARLY COLLISION STAGES

Starting with the pioneering papers [20, 21, 73] numerical AdS calculations have played an ever more important role to inform our understanding of the early phase of high-energy heavy ion collisions. Over time the numerical techniques have been steadily improved, such that today realistic three-dimensional collisions can be studied instead of somewhat schematic collisions of infinite plane shock waves. Also, more and more details and special cases were studied [62] instead of somewhat schematic collisions of infinite plane shock waves. In parallel, calculations of the entropy density [82], shear viscosity [83], conductivity [84], and inverse equilibration times [85] at NLO in the string coupling were found to be in good agreement with QGP phenomenology. As these corrections can be encoded in higher derivative terms of classical gravity, there is a possible path to more quantitative holographic simulations of the early collision stages.

Numerical solutions of the classical Einstein equations with AdS boundary conditions share the difficult problem of having to deal with the diffeomorphism invari- ance of general relativity. Using a suitable metric as an ansatz one can significantly reduce the diffeomorphism freedom. Chesler and Yaffe used Eddington-Finkelstein coordinates and managed to rewrite the Einstein equa- tion as a system of nested ordinary differential equations, which they solved using functional methods (Chebychev functions). A nice review of early developments can be found in [84]. In recent years activity has somewhat abated although there is still continuous progress, see e.g. [62]. Already early on these simulations showed that AdS dynamics leads to rapid hydrodynamicization [76, 80] with sizeable transverse flow [87].

The prevailing strategy has been to merge these holographic simulations of the early stages of heavy ion colli- sions with statistical descriptions of the later stages of the collisions [86]. Early calculations focused on highly sym- metric settings in order to reduce the computational demands, but more recently asymmetric settings have been studied, and by now semi-realistic heavy ion collisions can be simulated [74, 76, 81]. In particular, these stud- ies have shown that holographic simulations reproduce relativistic viscous hydrodynamics, validating the hybrid holographic-transport approach, see [88] for a recent re- view. In the present context it is relevant that hydrody- namization occurs at a fixed proper time [80] and that AdS dynamics and viscous hydrodynamics are indistin- guishable for hydrodynamic observables close to hydrody- namization such that the whole QGP fireball shows hydrodynamic behavior at nearly the same proper time.

Holographic calculations are attractive because they allow to treat problems that are too difficult to solve using QFT techniques. The hope is that the successful application of holographic techniques, which has been re- alized for the early stages of heavy ion collisions, can be extended to the later collision stages. This hope is nur- tured by several observations documented in the litera- ture, which we will discuss now.

The insight that quantum entanglement in the field theory is holographically encoded in features of the higher dimensional geometry was most clearly expressed by Maldacena and Susskind [26] who argued, building on ideas formulated in [89, 90], that the maximally extended (eternal) AdS-BH geometry is dual to a thermofield dou- ble state in the quantum field theory, which means that the quantum states of the field theory in the two asym- ptotic regions are maximally entangled. More generally, in this picture, the entanglement between the quantum states in two different space-time regions is geometrically represented by an Einstein-Rosen (ER) bridge between two regions in AdS space. We refer the reader to [91] for a review of the many publications building on this paradigm and as well as its relation to recent progress in understanding the role of entanglement for BH decay by Hawking radiation.

In [26] the label “ER = EPR” was coined for this con- nection, where EPR stands for Einstein-Podolsky-Rosen
and, more generally, for any form of entanglement in the quantum field theory. For our discussion Fig. 13 in [26] is most stimulating as it suggests that the emission of entangled Hawking radiation can be described as creation of higher dimensional Einstein-Rosen bridges between the decaying BH and the emitted particles. This picture, however, also displays an obvious problem of this idea: In a high-energy heavy ion collision thousands of hadrons are produced allowing for a plethora of entanglement patterns, far more than the number of simple geometries available for the dual picture.

In this context the recent ideas advanced by van Raamsdonk et al. [63, 92–95] are relevant, which posit that one large connected region of conformal field theory (CFT) and several smaller entangled domains of the same CFT have a nearly indistinguishable holographic dual. If it were legitimate to interpret the large region as QGP fireball and the many small domains as individual hadrons this would imply that the crossover transition from QGP to hadron gas could be continuous in a holographic dual description. This notion instills hope that a holographic description of the hadronization process is feasible and practical. [96] The dual description would be the disintegration of one large AdS black hole into several small ones as discussed, e.g., in [25] in the context of black hole evaporation.

The analogy is now clear. Quarks and gluons are confined to the interior of hadrons. Therefore, a diluted gas of thousands of hadrons can be interpreted in analogy to many isolated conformal field theories (CFT) on the edge. As each of these covers only a small fraction of the fireball or expanding hadron cloud, in the bulk AdS direction this leads only to noticeable effects very close to the asymptotic region of AdS space. (In order to distinguish this asymptotic region from spatial boundaries of localized states, we will use the term “AdS edge”. Such localized systems are usually modeled in the framework of boundary conformal field theory (BCFT) and its holographic dual description [97].) Deeper inside the bulk, all hadronic regions will remain connected via a network of ER bridges. As a result, the dual geometry will be very similar to that of the QGP fireball which is just one large AdS-BH. This means that from the information theoretical point of view there is no marked difference between entangled quarks and gluons with large relative momentum in the early fireball phase, a mixed QGP plus hadron gas state at intermediary times, and an entangled purely hadronic state at late times and that, therefore, many properties of this system can be calculated from the dual description valid at early times. In Sections VII and VIII we will expand on this from the AdS perspective.

To summarize, we argued that a successful extension of AdS calculations to the description of hadronization requires their extension from the Poincaré patch to the global AdS geometry and the inclusion of the scale $L$. We will suggest in Sections VII and VIII simplified models as first steps towards this long-term goal.

VI. HADRONIZATION IN RELATIVISTIC HEAVY ION COLLISIONS

At very high energies, the valence quarks of two colliding nuclei effectively pass through each other and deposit some of their energy on a time scale much shorter than 1 fm/c, a process that can be modeled as a quantum quench. Following this quench, the deposited energy thermalizes and forms a quark-gluon plasma on a time scale of the order $\tau_{th} \sim O(1/T)$ where $T$ denotes the temperature after thermalization. The thermalization process can be studied in holographic models, either by energy shell collapse in 5-dimensional Anti-de Sitter space (AdS$_5$) [17, 18] or by numerically solving shock front collisions in AdS$_5$ [20, 21].

The quark-gluon plasma then expands hydrodynamically while hadronization occurs in regions where the temperature reaches $T_c \approx 150$ MeV. At high collision energies the evolution of the matter near the center of momentum of the two nuclei is approximately boost invariant and can be described in terms of Milne coordinates $\tau, \eta, x, y$, where $\tau = \sqrt{t^2 - z^2}$ is the proper time and $\eta = \frac{1}{2} \ln[(t+z)/(t-z)]$ is called space-time rapidity. Massless particles moving along lines of constant $\eta$ also have rapidity $\eta$ in momentum space. Boost invariance means that to a good approximation the evolution is a function of $\tau$ and the transverse coordinates $x, y$ only for moderate values $|\eta| \ll \eta_{\text{beam}}$, where $\eta_{\text{beam}}$ is the beam rapidity in the center-of-momentum frame.

Hydrodynamics simulations show that the hypersurface where the transition from quark-gluon plasma (deconfined matter) to hadrons (confined matter) occurs is composed of two main domains (see Fig. 4). One domain is time-like and located approximately at constant $\sqrt{x^2 + y^2} \equiv r \approx R$, where $R$ is the nuclear radius, and stretches from $\tau_{th}$ to a time $\tau_c$ when the quark-gluon plasma converts to hadrons in bulk. The bulk hadronization proceeds along a space-like hypersurface given by $r < R$ and $\tau = \tau_c$, which defines the second domain.

![FIG. 4. Contour plot of the evolution of the energy density in a midcentral Au+Au collision at the highest RHIC energy $\sqrt{s_{NN}} = 200$ GeV. The horizontal axis shows the proper time $\tau$, the vertical axis shows one of the transverse coordinates. The black dashed line delineates the hadronization hypersurface $T = T_c = 150$ MeV [98].](image)

The hadronization processes in these two domains are...
different: On the time-like boundary of the quark-gluon plasma, hadron emission can be visualized as surface radiation. The emission rate is given by

$$dE_{\text{had}}/d\tau = \int dx dy \tau d\eta \delta(r - R) T_{0i} n_{i},$$  \hspace{1cm} (6)$$

where $n_{i}$ is the outward directed (space-like) normal vector on the hadronization hypersurface. Introducing the notation $S_{c} = T_{0i} n_{i}$ for the energy flow density at the hadronization temperature $T_{c}$, one finds

$$dE_{\text{had}}/d\eta = 2\pi R \int_{\tau_{\text{th}}}^{\tau_{c}} \tau d\tau S_{c} \approx \pi R^{2} \tau_{c} S_{c},$$  \hspace{1cm} (7)$$

since $\tau_{c} \gg \tau_{\text{th}}$. The energy converted into hadrons on the space-like hadronization surface is similarly given by

$$dE_{\text{had}}/d\eta = 2\pi \int_{0}^{R} r dr \tau_{c} \varepsilon_{c} \approx \pi R^{2} \tau_{c} \varepsilon_{c},$$  \hspace{1cm} (8)$$

where $\varepsilon_{c}$ is the energy density of matter at temperature $T_{c}$.

For noninteracting hadrons with mass $m$ at temperature $T = 1/\beta$, neglecting quantum statistics, one has

$$S(T, m) = \frac{T^{4}}{4\pi^{2}} \left(3(1 + \beta m) + (\beta m)^{2}\right) e^{-\beta m}$$  \hspace{1cm} (9)$$

$$\varepsilon(T, m) = \frac{T^{4}}{2\pi^{2}} \left(3(\beta m)^{2}K_{2}(\beta m) + (\beta m)^{3}K_{1}(\beta m)\right).$$

When summed over all well established hadron species in the Particle Data Book weighted by their statistical degeneracies $d_{i}$, one finds

$$\frac{S_{c}}{\varepsilon_{c}} \equiv \frac{\sum_{i} d_{i} S(T_{c}, m_{i})}{\sum_{i} d_{i} \varepsilon(T_{c}, m_{i})} \approx 0.17.$$  \hspace{1cm} (10)$$

Putting everything together and considering that $\tau_{c} \approx R$, one obtains

$$\frac{dE_{\text{had}}/d\eta}{dE_{\text{had}}/d\eta} \approx \frac{\tau_{c} S_{c}}{R \varepsilon_{c}} \approx 0.15 - 0.20.$$  \hspace{1cm} (11)$$

This means that approximately 80 – 85% of the hadrons produced by hadronization of the quark-gluon plasma are created during the bulk transition.\cite{99}

The two hadronization mechanisms can be experimentally separated by measuring the emission of hadron pairs instead of single hadrons. Hadron pairs radiated off the surface of the outward flowing fireball carry a large outward directed total momentum $K_{T} = p_{T,1} + p_{T,2}$. On the other hand, hadron pairs formed when the QGP hadronizes in bulk have on average $K_{T} = 0$. This is nicely seen in Fig.\[1\] where the emission hypersurface of hadron pairs at midrapidity for the same Au+Au collisions at RHIC is shown in the $(x, y, t)$ space, where $x, y$ are the transverse coordinates.\cite{100}. Just as in Fig.\[1\] the hypersurface resembles a capped cylinder. The upper panel shows the emission points of hadron pairs with $K_{T} = 0$; the lower panel shows the emission points with $K_{T} = 2$ GeV/c. Red color indicates the highest rate of emission; blue color indicates the lowest emission rate.

\section{VII. Hadronization as an AdS Phase Transition}

The Einstein equations with a negative cosmological constant have two solutions with asymptotic AdS geometry. One is plain AdS space; the other is a black hole (AdS-BH) imbedded in AdS space. The black hole geometry is parametrized by the Schwarzschild horizon (we will use the notation $r_{h}$). As we shall discuss, $r_{h}$ is related to the color screening distance in the dual gauge theory. At thermal equilibrium, $r_{h}$ is uniquely determined by the temperature $T = 1/\beta$ encoded in the period of the Euclidean version of the geometry.\cite{101}.

The Euclidean action plays the role of free energy in the space of geometries; its minimum thus corresponds to the stable equilibrium state at a given temperature. At low values of $T$, plain AdS space ($r_{h} = 0$) has the lowest Euclidean action; above a certain temperature $T_{c}$,
which depends on the details of the dual gravity theory, the AdS-BH geometry has the lowest free energy. As mentioned in a footnote in Section [IV] we here refer to such transitions as “Hawking-Page type” transitions [12]. In the simplest holographic model, the pure $(d+1)$-dimensional Einstein action with negative cosmological constant $\Lambda = -d/L^2$ on global AdS space, the transition is discontinuous with $r_h(T_c) = L$, which means that the Hawking-Page transition is a first-order phase transition. For a detailed derivation and the exact relation between $r_h$ and $T$, see e. g. Ref. [3].

In the dual gauge theory, the Hawking-Page transition corresponds to the confinement-deconfinement transition [5, 102]. This can be seen in multiple ways. The most commonly used argument is that the free energy of the thermal plain AdS geometry is described by a Hagedorn spectrum of gauge-singlet excitations, whereas the free energy of the AdS-BH geometry is proportional to $N_c^2$. Maybe the most intuitive argument is obtained by considering the potential of a heavy quark-antiquark pair, i.e., static objects carrying color charge in the fundamental representation of the gauge group [103]. In the gauge theory, this potential is determined by a Wilson loop connecting the world lines of the quark-antiquark pair. In the holographic dual representation the potential is determined by the Nambu-Goto action of a string connecting the quark-antiquark pair through the bulk (see Ref. [103]).

$$F = \frac{2}{3} \pi L^2 \sqrt{\frac{28}{9} \pi n} \left( \frac{869}{T} \right),$$

where $\pi$ is the gravitational constant and $T$ is the temperature. This translates into a thermal color screening length $\ell_s \approx 0.869/(\pi T)$ in $d = 4$ space-time dimensions.

It is possible to relax the firm connection between AdS-BH radius $r_h$ and the temperature $T$ by allowing for geometries with a conical singularity [107]. Such geometries do not correspond to a minimum of the free energy but they allow for a smooth interpolation between the equilibrium AdS-BH geometry and the thermal AdS geometry. This is illustrated in Fig. [7] for AdS$_5$. The solid lines show the free energy $F(r_h, T)$ for several different values of $T$, and the dashed line traces the location of the minima. The uppermost solid (red) curve is for the lowest temperature $T_{\text{min}}$ for which an AdS-BH solution of Einstein’s equations exists, $T_{\text{min}} = \sqrt{2}/(\pi L)$, the middle solid (blue) curve represents the free energy at the temperature of the Hawking-Page phase transition, $T_c = 3/(2\pi L)$, and the lowest solid (black) curve shows the free energy for $T = 1.2T_c$.

![String configurations in the AdS bulk connecting a static quark-antiquark pair separated by distance $\ell$.](image)

**FIG. 6.** String configurations in the AdS bulk connecting a static quark-antiquark pair separated by distance $\ell$. The left panel shows the string in the plain AdS geometry; the right panel shows the two disconnected strings reaching from the quark (antiquark) to the black hole horizon $r_h$ in the AdS-BH geometry.

For the conformally invariant $\mathcal{N} = 4$, large-$N_c$ supersymmetric Yang-Mills theory and for the plain AdS geometry in the bulk, the potential is found to be [103]

$$U(\ell) = -\frac{4\pi^2\sqrt{2\lambda}}{\Gamma(1/4)^2} \ell^4,$$

where $\lambda = g^2 N_c$ is the ‘t Hooft coupling and $\ell$ is the quark-antiquark separation. (There is no confining potential because the theory is conformally invariant.) For the AdS-BH geometry one finds that the potential $U(\ell)$ vanishes for separations

$$\ell > \ell_s \approx 0.869L^2/r_h,$$

which defines the screening distance of the color force in the gauge theory [104–106]. The screening occurs, because for $\ell > \ell_s$ the lowest energy string configuration corresponds to a disconnected pair of strings stretching straight from the quark (antiquark) to the black hole horizon $r_h$ (shown in the right panel of Fig. [6]). We can thus consider $r_h$ as a geometric parameter related to the color screening length in the gauge theory. Since for $T \gg T_c$ the black hole radius is related to the temperature $T$ at thermal equilibrium by $r_h = (4\pi/n)L^2T$, this translates into a thermal color screening length $\ell_s \approx 0.869/(\pi T)$ in $d = 4$ space-time dimensions.
As is evident from Fig. 7, the Hawking-Page transition for the conformal, large-$N_c$ super-Yang-Mills theory is a first-order phase transition. This is mirrored in the behavior of the pure non-supersymmetric SU($N_c$) gauge theory, which also exhibits a first-order deconfinement transition for $N_c \geq 3$ \cite{108}. Gravity dual models that resemble pure SU($N_c$) gauge theory more quantitatively can be constructed by adding a dilaton field to 5-dimensional Einstein gravity, which describes the running of the gauge coupling constant. In general, dilaton gravity duals of confining gauge theories exhibit a first-order deconfinement phase transition similar to that of the conformal super-Yang-Mills theory \cite{14}.

On the other hand, the deconfinement transition in QCD is known to be a smooth crossover \cite{37,38}. The thermal properties of dual dilaton gravity models including fundamental matter (quarks) that mimic the running coupling and chiral properties of QCD have been studied \cite{40}, and models that can change from a first-order phase transition to a crossover transition have been constructed \cite{109}.

The dynamics of the bulk transition depends on the rate at which the temperature drops. If the cooling rate is slow compared with the microscopic times scales, the transition will proceed at or near $T_c$ through a mixed phase via bubble formation \cite{13,45,110}. In the case of rapid cooling, the transition occurs via a Gregory-Laflamme instability \cite{111} from a supercooled phase at $T_{\text{min}}$ \cite{15,112,115}. The second scenario prevails in the limit $N_c \to \infty$, but for holographic models of QCD with $N_c = 3$ it is likely that the first scenario is realized under the conditions of a heavy ion collision. For a smooth crossover certainly the first scenario is realized.

VIII. HADRON EMISSION AS ANALOGUE OF HAWKING RADIATION

Decades of work by many theorists were needed to reach the present level of understanding of the black hole information puzzle. Understanding the mechanisms ruling the generation of entanglement in hadronization is most probably a problem of comparable difficulty. Therefore, it is tempting to profit from the insights of the black hole community by constructing a model which treats hadronization in analogy to Hawking radiation. In this analogy the hadrons correspond to the Hawking radiation outside of the black hole horizon which is what we also expect for the hadronic state outside of the QGP. Complete entanglement among the hadrons is only reached at the end of hadronization.

There exist various phenomenological observations on the QCD side which could fit into our model. As explained in the Introduction the uniform temperature of all observed hadron yields, which is well described by the thermal hadron resonance gas model, is difficult to understand. A possible line of argument was suggested by us in Ref. \cite{40} based on the phenomenologically very successful quark recombination model \cite{116}. In this model quarks and gluons from the fireball coalesce at its surface to form hadrons. In doing so their energies $E_i$ add up and their probability densities multiply, generating for any Fock-state of the hadron $h$ with energy $E_h$ the common factor \cite{117}

$$\prod_i e^{-E_i/T_{ch}} = e^{-\sum_i E_i}/T_{ch} = e^{-E_h/T_{ch}}, \quad (14)$$

where the subscript “ch” indicates that the temperature parameter is derived from the chemical composition of the hadron gas.

An *ad hoc* feature of this model, dictated by phenomenology, is that the temperature of the QGP fireball and that of all produced hadrons has the same value $T_{ch}$. This feature is not easily understood without entanglement, because different hadrons scatter differently and thus are not expected to decouple at the same time from the expanding fireball. This could be naturally explained if a highly entangled QGP state and a highly entangled hadronic state are basically indistinguishable as we argued above based on the ideas of van Raamsdonk \cite{63}.

IX. A HOLOGRAPHIC PICTURE OF HADRONIZATION

As long as information about the state of a time reversal invariant quantum system is not lost by any kind of measurement and the associated (partial) collapse of the complete many-body wave function, it must be possible “to run the movie backwards”. The highly complex initial state $\Phi_1$ of a relativistic heavy ion collision has
nearly zero entropy. In the parton (quark-gluon) basis, this state is characterized by a density matrix with two blocks describing the two nuclei approaching each other. This entangled initial state evolves by a unitary transformation into another highly entangled final state \( \Phi_f \) characterized by a full density matrix in the parton basis. Eventually, this final many-parton quantum state is projected onto hadron states by experimental measurements, which identify the asymptotic eigenstates of the many-parton system. This happens at a time of order \( O(10^{-9} \, \text{s}) \), much longer than the duration of the nuclear reaction which is of order \( O(10^{-23} \, \text{s}) \). Since the detector acts as a heat bath, this leads to decoherence and thus to entropy production.

For a realistic holographic description of a relativistic heavy ion collision that includes hadronization it is crucial to describe entanglement at each time. This is not possible in numerical solutions of the classical Einstein equations, which depend only on the classical, local energy momentum tensor on the AdS edge. It is thus unclear whether an approximately valid AdS model that keeps track of entanglement during a transition of the Hawking-Page type exists. We argue in this Section that it does.

Our argument rests crucially on the fact that for a many-particle quantum state, such as the final state of a heavy ion collision, the effects of entanglement are only relevant if the complete wave function is considered, but are negligible for observables involving only a few hadrons. This is a consequence of the "monogamy of entanglement" which is well established in quantum information theory.

There exists a fundamental constraint, closely related to the no-cloning theorem \([120]\), which states that quantum entanglement cannot be freely shared among many objects \([121][122]\). This has been analyzed in detail for systems of qubits for which quantitatively precise statements can be made and formal proofs are possible. One can define a quantity \( \tau(\rho_{AB}) \), called "tangle", which quantifies entanglement between the elements of bi-partitions of multi-particle quantum states \( A, B \), described by a density matrix \( \rho_{AB} \). \( \tau \) can have values between 0 (no entanglement) and 1 (complete entanglement). For a quantum state of \( n \) subsystems \( A_1, A_2, \ldots, A_n \),
There exist many concepts in the literature which could be adapted to our situation. For example, Malda and Susskind suggested in [26] (see Fig. 13) a geometric, holographic interpretation of the entanglement of Hawking radiation in terms of ER bridges for a Minkowski space black hole, which is somewhat similar to ours. The difference between their picture and ours is that we do not study a horizon in the edge but in the bulk of AdS space and that we treat hadrons as entangled with the AdS horizon dual to the QGP fireball rather than photons as entangled with a black hole horizon.

FIG. 10. E) As the Hawking-Page-like transition for QCD at small baryon number density is a cross-over the difference between the situation just before (at $T = T_c + \epsilon$, see Fig. [C]) and after (at $T = T_c - \epsilon$, see Fig. [D]) is small. Therefore, E), which is strictly valid only above $T_c$, is expected to provide a good approximation. Note that because hadrons on the AdS edge are maximally entangled with states on the horizon, they are uncorrelated with all other edge hadrons such that all of them form a thermal ensemble with temperature $T_c$. This is just one of the crucial elements of ETH. Note also, that the difference in this respect to Fig. [D] is only of order $O(1/N_h)$ and thus negligible. For more discussions, please see the main text.

As discussed in Section I, hadronization in heavy ion collisions is usually modeled as “chemical freeze out”, which is hypothesized to explain the observation that all hadron yields agree with the predictions of the thermal model although hadronic interactions continue in the expanding hadron gas. As we argue that the success of the thermal model is a consequence of quantum coherence, e. g. in the form of ETH, we do not need any such additional mechanism. However, even our mechanism requires a rapid transition from the hydrodynamic QGP phase to the hadronic phase. In fact, microscopic simulations indicate a rapid growth of the specific viscosity at temperatures below 150 MeV signalling a rapid breakdown of hydrodynamics when the expanding fireball cools below $T_c$ [24].

Holographically, this transition corresponds to the transition between Figs. [9] B and D, where Fig. [C] is a schematic representation of an intermediate stage. We do not yet have a reliable quantitative description of these intermediate stages, neither in QCD nor in holography, but one future goal stated in Section II is to develop such a description in the holographic model of the transition.

With respect to phenomenology, however, we can argue that the mapping to a purely hadronic late stage...
should be a smooth one, i.e., a cross-over rather than a first-order phase transition, in line with the fact that the deconfinement/confinement phase transition of QCD is a cross-over.

We conclude this Section with the following remark: If a holographic dual of QCD exists, it must be possible to develop a quantitative description of the hadronization of a quark-gluon plasma. At present we do not know whether this is the case. However, even if it does not or turns out to be unachievable in practice, much could be learned about the quantum physics of the confinement transition by investigation of a holographic model that incorporates salient features of the quark-hadron transition in QCD. In the next Section we will discuss some steps that can be taken in this direction.

X. (SOME) OPEN QUESTIONS

As already emphasized our reasoning is speculative and must be consolidated or refuted by detailed investigations. Because the problem of decoherence and thermalization of many-particle quantum states is a very generic one, there exists a large and rapidly expanding literature on it, where the ETH plays a prominent role. This leads to our first set of open questions:

- Does QCD exhibit ETH behavior?

We interpret the success of the thermal model as heuristic indication that it does show ETH behavior, but it would be desirable to affirm this with rigorous methods, such that the success of the thermal model becomes a prediction rather than merely an observation. To make contact with the experimental data, it is not only necessary to confirm ETH properties in principle, but to also establish the time scale at which ETH behavior becomes manifest (see [53]).

In fact, this question is much discussed in quantum information theory and answering it could become an early success of quantum simulators (see, e.g., Ref. [127] and references therein). There are two crucial advantages of quantum computing in this context. First, for the arguments given in Section III, the extensive nature of the microcanonical entropy leads to a strong exponential suppression of all off-diagonal matrix elements except for very small systems, which are easiest to simulate. Second, the simulation of realistic hadrons is not required to test the validity of ETH as long as the Hamiltonian is well enough modeled. Also, showing ETH behavior for the gauge group SU(2), which is easier to simulate, should be sufficient.

Another promising approach is lattice gauge theory. Just as RMT behavior was established for the QCD Dirac operator by simulations on small lattices with classical computers [129, 127] the same may be possible for ETH.

- Is ETH necessary or only sufficient to explain success of the thermal model? If it is necessary the relevant question becomes: How long does it take until few particle observables look in good approximation thermal?

Recently a number of observations suggest that ETH is less universal than often assumed and may not even be necessary for a system to thermalize [128]. In Ref. [129] it was shown that non-Abelian charges reduce entropy-production rates and may enhance finite-size deviations from eigenstate thermalization. The occurrence of many-body scars may also negate ergodic behavior in certain systems [130] and slow down thermalization [131].

In general, one should be aware that ETH postulates very restrictive constraints for matrix elements of generic operators between energy eigenstates. It could well be that less restrictive requirements already imply the emergence of thermal behavior to such an extent that the differences are unobservable in realistic experiments. For example, it would be interesting to understand whether monogamy of entanglement in many-particle systems [120–123] is already sufficient to mimic thermalization as long as only small subsystems of the complete, entangled quantum states are observed. Obviously, the search for alternatives to, and generalizations of, the ETH paradigm is a vast and rapidly developing area of research.

A second set of questions revolves around the search for useful holographic models of the quark-hadron transition.

- Do recent advances in numerical solutions of gravity in AdS space make it possible to better understand the dynamics of a Hawking-Page type transition?

Over the years much effort has been invested to better understand the phenomenological implications of the original Hawking-Page transition without leading to definitive conclusions [15, 110, 112]. In the large-Nc limit, the HP transition will proceed by spinodal decomposition of a deeply supercooled phase unless the cooling process is very slow. One interesting question is whether the system re-equilibrates after the transition is complete [113]; another is whether quantum gravity corrections at finite ‘t Hooft coupling [52] help to smoothen out the HP transition. In a very slowly cooling scenario, the transition will proceed via bubble nucleation. It is an interesting question which of these scenarios is applicable to pure $\text{SU}(N_c)$ gauge theories at moderate $N_c$, where the equilibrium deconfinement transition is known to be of first order.

Progress in the numerical treatment of matrix models [132, 135], as well as AdS/CFT initial value problems [62] may make it possible to explore the real-time dynamics of Hawking-Page transitions. However, none of these efforts addresses the question how the entanglement structure of the quasi-thermal QGP is reflected in the entanglement structure of the produced hadrons. In the holographic dual this raises the challenge of computing the geometric form of an ER bridge between the asymptotic (“edge”)
regions of a holographic thermofield double state for general initial conditions.

- Would holographic models in which the transition is a smooth crossover as in QCD allow for an approximate identification of the HP transition with the hadronization transition of a quark-gluon plasma?

The possible application of the HP transition in nonconformal dilaton-gravity models \([13,136,137]\) as a holographic model of the dynamical deconfinement transition in QCD was explored in Ref. [138]. More recently, the dynamical nature of the transition including domain formation \([139]\) and domain wall propagation \([140]\) have been investigated. Even shock wave collisions were studied in a holographic model that parametrizes the evolution of the quark-hadron transition from a smooth crossover to a first-order phase transition with increasing baryon chemical potential \([141]\).

- Can hadron emission from the surface of a quark-gluon plasma be modeled as splitting quenches?

There exist low-dimensional toy models for the splitting of one domain of conformal field theory into two \([142,143]\) or the formation of an Einstein-Rosen (ER) bridge between two horizons \([34]\). Can these models be generalized to multiple black hole splittings, corresponding to the production of multiple hadrons and to higher dimensions? Is AdS/BCFT the adequate framework to describe hadronization? If it is, does it imply, in the spirit of Ref. [139], that the holographic description of the quark-gluon plasma and an interacting hadron gas are quite similar? In fact this is exactly what is observed on the field theoretical side of the duality. The properties of a hadron resonance gas and the QGP are indistinguishable close to hadronization (see, e.g., [144]).

Finally, one would like to know whether there are experimental signatures of the entanglement structure among hadrons emitted in the decay of quark-gluon plasma.

- How can the validity of the analogy between hadronization and black hole decay proposed in Section VII be tested?

As discussed in Section IX the monogamy property of entanglement \([123]\) implies that the states of any pair of hadrons among the \(N\) emitted hadrons is only entangled at the \(O(1/N)\) level. If one would know what the most accessible entanglement signature is, one could test this conclusion as a function of the QGP fireball size. One possibility is to look for quantum correlations among the spins of hadrons \([67]\).

\section{XI. Conclusions}

The application of concepts from the AdS/CFT duality to describe the early stage, spanning about \(1-2\) fm/c, of high energy heavy ion collisions as they are studied experimentally at LHC and RHIC has been remarkably successful and brought many important insights. This suggests that one should try to extend the dual holographic description to later times including the hadronization stage. As we explained, any such effort has to face the full complexity of the black hole information problem in the dual description.

The time evolution of both the AdS black hole and QCD is unitary. No information gets lost and no entropy is produced. Consequently, thermal ensemble in the strict sense cannot be produced. However, what can be produced is a system for which all realistic measurements produce results that are indistinguishable from strictly thermal ones. Recent progress on black hole evaporation has substantiated this conclusion. For QCD the corresponding field theoretical calculation, which would have to treat quantum entanglement during each step of the collision appears impractical. If, however, the AdS- and QCD-based descriptions of a heavy ion collision were holographically dual up to calculable corrections, a QCD-based calculation would be unnecessary, and one could simply adapt the insights from black holes to heavy ion collisions.

Here we focused on the question whether and how the AdS dictionary could be extended to times beyond the hadronization of the quark-gluon plasma and how the phenomenological properties of heavy ion collisions up to and beyond the Page time could be represented holographically. In doing so, we have not reached a conclusive answer, but we also did not encounter obvious roadblocks, and we formulated topics for future work that could help answer some of the many remaining questions.

\textit{Acknowledgments:} We thank M. Kaminski, L. Yaffe, and S. Waerbe for many valuable discussions and comments on a draft of this manuscript. We also thank T. Hartman for insightful discussions on Euclidean wormholes. B. M. acknowledges support from the U.S. Department of Energy Office of Science (Grant DE-FG02-05ER41367) and from Yale University during a sabbatical stay in Spring 2022.

\begin{thebibliography}{99}

[1] D. N. Page, Phys. Rev. Lett. \textbf{71}, 3743 (1993) arXiv:hep-th/9306083.

[2] D. N. Page, JCAP \textbf{09}, 028 (2013) arXiv:1301.4995 [hep-th].

[3] D. N. Page, Phys. Rev. Lett. \textbf{71}, 1291 (1993) arXiv:gr-qc/9305007.

[4] A. M. Kaufman, M. E. Tai, A. Lukin, M. Rispoli, R. Schittko, P. M. Preiss, and M. Greiner, Science \textbf{353},

\end{thebibliography}
arXiv:0907.3719 [hep-lat]

[109] M. Attems, Y. Bea, J. Casalderrey-Solana, D. Mateos, M. Triana, and M. Zilhão, Phys. Rev. Lett. 121, 261601 (2018), arXiv:1807.05175 [hep-th].

[110] R. A. Janik, M. Järvinen, and J. Sonnenschein, JHEP 09, 129 (2021), arXiv:2106.02642 [hep-th].

[111] R. Gregory and R. Laflamme, Phys. Rev. Lett. 70, 2837 (1993), arXiv:hep-th/9301052.

[112] V. E. Hubeny and M. Rangamani, JHEP 05, 027 (2002), arXiv:hep-th/0202189.

[113] A. Buchel and L. Lehner, Class. Quant. Grav. 32, 145003 (2015), arXiv:1502.01574 [hep-th].

[114] O. J. C. Dias, J. E. Santos, and B. Way, Phys. Rev. Lett. 117, 151101 (2016), arXiv:1605.04911 [hep-th].

[115] L. G. Yaffe, Phys. Rev. D 97, 026010 (2018), arXiv:1710.06455 [hep-th].

[116] R. J. Fries, V. Greco, and P. Sorensen, Ann. Rev. Nucl. Part. Sci. 58, 177 (2008) arXiv:0807.4939 [nucl-th].

[117] B. Müller, R. J. Fries, and S. A. Bass, Phys. Lett. B 618, 77 (2005), arXiv:nucl-th/0503003.

[118] The ground state of a colliding nucleus is unique, and any interaction of the nucleus with the accelerator structure is completely negligible on the scale of nuclear excitations. Finally, although the two nuclei can be Coulomb excited on their approach to each other, the excitation is coherent and does not change the fact that the nuclear quantum state is pure.

[119] We remind the reader that use the term “AdS edge” to avoid confusion of the asymptotic region of AdS space with the boundary of the QCD fireball.

[120] W. K. Wootters and W. H. Zurek, Nature 299, 802 (1982).

[121] W. K. Wootters, Phys. Rev. Lett. 80, 2245 (1998), arXiv:quant-ph/9709029.

[122] V. Coffman, J. Kundu, and W. K. Wootters, Phys. Rev. A 61, 052306 (2000), arXiv:quant-ph/9907047.

[123] T. J. Osborne and F. Verstraete, Physical Review Letters 96, 220503 (2006).

[124] Z. Yang and R. J. Fries, Phys. Rev. C 105, 014910 (2022).

[125] A. Schuckert and M. Knap, Phys. Rev. Research 2, 043315 (2020).

[126] M. E. Berbenni-Bitsch, S. Meyer, A. Schäfer, J. J. M. Verbaarschot, and T. Wettig, Phys. Rev. Lett. 80, 1146 (1998), arXiv:hep-lat/9704018.

[127] J. J. M. Verbaarschot and T. Wettig, Ann. Rev. Nucl. Part. Sci. 50, 343 (2000), arXiv:hep-ph/0003017.

[128] A. W. Harrow and Y. Huang, (2022), arXiv:2209.09826 [cond-mat.stat-mech].

[129] S. Majidy, A. Lasek, D. A. Huse, and N. Y. Halpern, (2022), arXiv:2209.14303 [quant-ph].

[130] C. J. Turner, A. A. Michailidis, D. A. Abanin, M. Serbyn, and Z. Papić, Nature Physics 14, 745 (2018).

[131] A. A. Michailidis, C. J. Turner, Z. Papić, D. A. Abanin, and M. Serbyn, Phys. Rev. X 10, 011055 (2020), arXiv:1905.08564 [quant-ph].

[132] M. Hanada and H. Watanabe, (2022), arXiv:2210.11216 [hep-th].

[133] S. Pateloudis, G. Bergner, M. Hanada, E. Rinaldi, A. Schäfer, P. Vranas, H. Watanabe, and N. Bodendorfer, (2022), arXiv:2210.04881 [hep-th].

[134] W. K. Wootters, Phys. Rev. Lett. 80, 2245 (1998), arXiv:quant-ph/9709029.

[135] E. Megias, H. J. Pirner, and K. Veschgini, Phys. Rev. D 83, 056003 (2011), arXiv:1009.2953 [hep-ph].

[136] U. Gursoy, S. Lin, and E. Shuryak, Phys. Rev. D 88, 105021 (2013), arXiv:1309.0789 [hep-th].

[137] M. Attems, Y. Bea, J. Casalderrey-Solana, D. Mateos, and M. Zilhão, JHEP 01, 106 (2020), arXiv:1905.12544 [hep-th].

[138] C. Ecker, J. Erdmenger, and W. van der Schee, SciPost Phys. 11, 047 (2021) arXiv:2103.10435 [hep-th].

[139] M. Attems, JHEP 08, 155 (2021) arXiv:2012.15687 [hep-th].

[140] T. Shimaji, T. Takayanagi, and Z. Wei, JHEP 03, 165 (2019) arXiv:1812.01176 [hep-th].

[141] P. Caputa, T. Numasawa, T. Shimaji, T. Takayanagi, and Z. Wei, JHEP 09, 018 (2019) arXiv:1905.08265 [hep-th].

[142] R. Bellwied, S. Borsanyi, Z. Fodor, J. N. Guenther, S. D. Katz, P. Parotto, A. Pasztor, D. Pesznyak, C. Ratti, and K. K. Szabo, Phys. Rev. D 104, 094508 (2021) arXiv:2102.06625 [hep-lat].

[143] K. Werner, Phys. Rev. Lett. 98, 152301 (2007) arXiv:0704.1270 [nucl-th].

[144] M. Petrovici, I. Berceanu, A. Pop, M. Târziolă, and C. Andrei, Phys. Rev. C 96, 014908 (2017) arXiv:1703.05805 [nucl-th].