UNIVERSITY OF SOUTHAMPTON

FACULTY OF SOCIAL, HUMAN AND MATHEMATICAL SCIENCES

Mathematical Sciences

Aspects of Black Hole Instabilities and Consequences

by

Paul Rodgers

Thesis for the degree of Doctor of Philosophy

May 2021

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF SOCIAL, HUMAN AND MATHEMATICAL SCIENCES

Mathematical Sciences

Thesis for the degree of Doctor of Philosophy

ASPECTS OF BLACK HOLE INSTABILITIES AND CONSEQUENCES

by Paul Rodgers

The core topic we study in this thesis is the concept of black hole instabilities. This in essence means that if one probes a black hole background with any classical field, modes scattering the black hole horizon which are growing in time are considered unstable. There are many interesting questions which arise as a result of an unstable black hole perturbation, such as; what is the endpoint of the instability? What does the solution and phase diagram of such a black hole configuration look like? The answer in some cases is that one finds hairy black hole solutions, which turn out to be even richer in their structure as we will explore in the second part of this thesis. Sometimes however, searching for an instability within the vast structure of black hole quasi-normal modes is itself an interesting task which requires robust numerical methods. One needs to use perturbative methods to guide the numerical search, which we discuss in the first half of this thesis.

Table of Contents

Declaration of Authorship

I, Paul Rodgers, declare that the thesis entitled *Aspects of Black Hole Instabilities and Consequences* and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted form the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- parts of this work have been published as:
  - [1] O. J. C. Dias, R. Masachs, O. Papadoulaki and P. Rodgers, "Hunting for fermionic instabilities in charged AdS black holes," *Journal of High Energy Physics* ***2020***, (2020) 1-55
  - [2] O. J. C. Dias, R. Masachs and P. Rodgers, "Boson stars and solitons confined in a Minkowski box," *Journal of High Energy Physics* ***2021***, 236 (2021).
  - [3] A. Davey, O. J. C. Dias and P. Rodgers, "Phase diagram of the charged black hole bomb system," *Journal of High Energy Physics* ***2021***, (2021) 1-36.

Signed: .....

Date: .....

Acknowledgements

I would like to thank my supervisor Oscar Dias for his invaluable guidance throughout this PhD. I would also like to thank Kostas Skenderis and Marika Taylor who were always willing to discuss my work and give their opinion, which proved especially invaluable when understanding some of the finer issues from the AdS/CFT point of view. I would also like to thank my co-authors Alex, Olga, Oscar and Ramon for their collaborations on the publications presented in this thesis. I would like to particularly acknowledge Ramon Masachs for his guidance during the first two years of my PhD. Finally, I would like to thank family and friends for their support.

Part I

Introduction and background material

CHAPTER 1

Introduction

1.1 Einstein's equations and black holes.

Einstein's equations are rich and diverse in their applications. This thesis explores different applications in one of the most fascinating solution-types to Einstein's equations, namely black holes. In their most general form, Einstein's equations read

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi G_N T_{\mu\nu}, \tag{1.1.1}$$

where $R_{\mu\nu}$ is the Ricci tensor and R is the Ricci scalar associated with the metric $g_{\mu\nu}$, which describes the geometry of the spacetime manifold. The value of the cosmological constant Λ depends on the asymptotic behaviour of our spacetime, $\Lambda = 0$ for asymptotically flat spaces, $\Lambda > 0$ for de Sitter (dS), and $\Lambda < 0$ for Anti-de-Sitter (AdS) backgrounds. On the right hand side is the energy-momentum tensor, which depends on the matter content one wishes to consider. Equation (1.1.1) is normalised with Newton's constant G_N so that in the weak-field, non-relativistic limit (1.1.1) reduces to Newtonian theory. The equation (1.1.1) follows from extremising the Einstein-Hilbert action

$$S_{EH} = \frac{1}{16\pi G_N}\int \mathrm{d}^4x\sqrt{-g}\,(R - 2\Lambda) + \int \mathrm{d}^4x\sqrt{-g}\mathcal{L}_m. \tag{1.1.2}$$

In equation (1.1.2) $\mathcal{L}_m$ is the matter Lagrangian which gives rise to the stress energy tensor in equation (1.1.1). There are many interesting solutions to (1.1.1), such as neutron stars or

cosmological solutions describing the expansion of the universe. However, in this thesis we focus on black hole solutions, which I will now describe. Black holes are well known to be objects (or regions of spacetime) possessing a gravitational field that is so strong, not even light rays can enter this region and escape [1]. It is also necessary to define these objects mathematically, in order to give a precise definition we restrict our attention to asymptotically flat black holes for now. The following material can be found in much greater detail in books such as [4, 5]. In physical terms, an asymptotically flat spacetime is one which approaches the Minkowski metric at spatial infinity. Recall that future null infinity in Minkowski spacetime is denoted $\mathscr{I}^+$ and that a spacetime manifold $(\mathcal{M}, g)$ is *asymptotically flat* if there exists a spacetime $(\bar{\mathcal{M}}, \bar{g})$ and a conformal function Ω on $\mathcal{M}$ such that $\mathcal{M} \subset \bar{\mathcal{M}}$ where $\bar{\mathcal{M}}$ is an extension of $(\mathcal{M}, \Omega^2 g)$ [2]. Now suppose we have a time-orientable, asymptotically flat spacetime manifold $(\mathcal{M}, g)$. Then the set of points in $\mathcal{M}$ that can send a signal to $\mathscr{I}^+$ is $\mathcal{M} \cap \mathcal{J}^-(\mathscr{I}^+)$ where $\mathcal{J}^-(\mathscr{I}^+)$ is the causal past of the surface which is located at $\mathscr{I}^+$ and $\mathcal{J}$ is defined using the aforementioned metric $(\bar{\mathcal{M}}, \bar{g})$. The *black hole region* is therefore given by $\mathcal{B} = \mathcal{M} \setminus \{\mathcal{M} \cap \mathcal{J}^-(\mathscr{I}^+)\}$. The future event horizon is defined by $\mathcal{H}^+ = \dot{\mathcal{B}}$, which is the boundary of $\mathcal{B}$ on $\mathcal{M}$. Now, to complete the mathematical description of black holes we define the term *Killing horizon*. A null hypersurface $\mathcal{N}$ is a *Killing horizon* if there exists a Killing vector field ξ^μ in a neighbourhood of $\mathcal{N}$ such that ξ^μ is normal to $\mathcal{N}$, that is for n^μ normal to the hypersurface $\mathcal{N}$ we have $g_{\mu\nu} n^\mu \xi^\nu = 0$. This leads naturally to the following theorem by Hawking: For an asymptotically flat, stationary, analytic, vacuum black hole spacetime, the future event horizon $\mathcal{H}^+$ is a Killing horizon. To conclude our introduction to Einstein's theory and black holes it is instructive to give an example, which will also be of relevance to this thesis. Consider equation (1.1.1), for the pure vacuum one has $T_{\mu\nu} = 0$ and the unique static black hole solution is given by the famous Schwarzschild solution [6] [3]. Now suppose we wish to consider the interaction of the gravitational field in an electromagnetic vacuum, we have

$$T_{\mu\nu} = \frac{1}{4\pi}\left(F_{\mu\alpha}F_{\nu}{}^{\alpha} - \frac{1}{4} g_{\mu\nu} F_{\alpha\beta}F^{\alpha\beta}\right), \tag{1.1.3}$$

where $F = \mathrm{d}A$ is the Maxwell field strength associated with the electromagnetic potential A. The field strength obeys Maxwell's (vacuum) equation in curved spacetime, given by

$$\nabla^\alpha F_{\alpha\beta} = 0, \tag{1.1.4}$$

where ∇ is the covariant derivative associated with the torsion-free Christoffel symbols for the metric g. The static black hole solution to this theory is given by the *Reissner-*

[1] Equivalently but slightly more precisely, one cannot send a signal from inside the black hole region to spatial infinity.

[2] There are further mathematical components of this definition, which can be found in [4, 5]. For example (within $\bar{\mathcal{M}}$) $\mathcal{M}$ can be extended to $\mathcal{M} \cup \partial\mathcal{M}$ and Ω can be extended to a function on $\bar{\mathcal{M}}$ such that $\Omega = 0$ and $\mathrm{d}\,\Omega \neq 0$ on $\partial\mathcal{M}$.

[3] This is the uncharged version of Birkhoff's theorem, which we will discuss in part IV

Nordstrom (RN) solution, whose field content looks like:

$$\begin{aligned} \mathrm{d}s^2 &= -f\mathrm{d}t^2 + f^{-1}dr^2 + r^2\,\mathrm{d}\Omega_2^2, \qquad f = 1 - \frac{2M}{r} + \frac{Q^2}{2r^2}; \\ A_\mu\,dx^\mu &= A_t(r)\;\mathrm{d}t, \qquad A_t(r) = -\frac{Q}{r} + C; \end{aligned} \tag{1.1.5}$$

where $\mathrm{d}\Omega_2^2$ is the line element of a unit radius 2-sphere, M and Q are the Arnowitt-Deser-Misner (ADM) mass and charge of the black hole [7]. In (1.1.5), C is an arbitrary integration constant and fixing it amounts to choosing a particular gauge. One common gauge choice is $C = 0$ where one has $A|_\infty = 0$ and the chemical potential is $\mu = -A|_{r_+} = Q/r_+$. Another gauge that is commonly used is $C = \mu$ where we have $A|_{r_+} = 0$ and $A|_\infty = \mu$. Note that Einstein's equations hold in general d dimensions, however we focus on black holes in 4 dimensions in this thesis.

1.2 Black hole perturbations

Black holes are similar to many other physical objects in the sense that subject to small perturbations, they display a characteristic '*ringing*', these characteristic frequencies are known as *quasi-normal modes* (QNMs). The gravitational field itself has QNMs when perturbed, however fields with any spin can probe a black hole background, such as the scalar fields (spin 0)[4] or spinors (spin $\pm\frac{1}{2}$). These QNMs are intimately linked with the stability of a black hole solution. Mathematically QNMs appear as eigenmodes of dissipative systems and in practice (as we will see) the eigenvalue problems which appear in black hole perturbation theory have solutions which cannot be written in closed form, therefore requiring a numerical treatment. In general, field perturbations will inherit symmetry properties from the symmetries of the background at hand which generally leads to a separation of variables of the PDE perturbation equations so that they reduce to ODE's which can then be solved by your method of choice. These ODEs must then be supplemented by physically appropriate boundary conditions imposed at the black hole horizon and at our asymptotic infinity. For the example of black holes which are spherically symmetric (as we have in the case (1.1.5)), the associated perturbations can be decomposed in terms of spherical harmonics as $\sum_{\ell m} Y_{\ell m}(\theta)$, leaving just the radial and time fields to be found. The RN solution also has ∂_t as a Killing vector, therefore perturbations can be further Fourier decomposed as $\mathrm{e}^{-i\omega t}$, where ω are the characteristic frequencies. In empty spacetime these frequencies are real and the perturbation is oscillating in time, these are normal modes. In the presence of a black hole, which is dissipative, the frequencies are complex: $\omega = \omega_R + i\omega_I$. Therefore the time component of the perturbation looks like $\mathrm{e}^{-i\omega_R t + \omega_I t}$, hence modes with $\omega_I < 0$ decaying in time, these are QNMs, on the other hand modes with $\omega_I > 0$ are growing exponentially

[4] Scalar fields prove to be a good toy model for the more physically relevant case of the graviton (spin 2).

in time, rendering the perturbation unstable. The mathematical reasoning for having complex frequencies is down to the presence of the black hole horizon, which acts as an absorbing membrane on classical field perturbations, this forces the corresponding boundary value problem to be non Hermitian and the associated eigenvalues are complex. For backgrounds with sufficient symmetry, there usually exists a *Teukolsky master equation*[5] governing perturbations for general spin s fields. These equations can generally be written in *Schrodinger form*

$$\partial^2_{r_\star}\Psi_s + \left(\omega^2 - V_s\right)\Psi_s = 0. \tag{1.2.1}$$

In equation (1.2.1), $r_\star$ represents the *tortoise coordinate*, which for the RN background (1.1.5) is defined via the relation $\frac{\mathrm{d}r_\star}{\mathrm{d}r} = \frac{1}{f}$ and V_s is an effective potential, which depends on the spin of the field and the radial coordinate. In general the equations for field perturbations that we will encounter in this thesis can be written in the above form, such as in part II when we search for fermionic instabilities in charged AdS black holes- by computing numerically the QNMs via Dirac's equation. For a detailed review and collection of results in black hole perturbation theory see the review [9] and references therein. Now we give a detailed description of spinors and scalars in a general curved background. For more detailed discussions see [10, 11, 12, 13, 14].

To start, consider spin-s fields in Minkowski spacetime. The spin of a field can be identified by looking into how the field transforms under a Lorentz transformation, $x^\mu \to \tilde{x}^\mu = \Lambda^\mu{}_\nu x^\nu$. Let $\mathcal{M}^{\alpha\beta} = -\mathcal{M}^{\beta\alpha}$ be the generators of Lorentz transformations (*i.e.* a basis of six 4×4 antisymmetric matrices obeying the Lorentz Lie algebra). A finite Lorentz transformation is described by $\Lambda = \exp\left(\frac{1}{2}\Omega_{\alpha\beta}\mathcal{M}^{\alpha\beta}\right)$ where $\Omega_{\alpha\beta} = -\Omega_{\beta\alpha}$ are six parameters describing the particular transformation Λ (boost, rotations) of the Lorentz group $SO(3,1)$.

Under a Lorentz transformation a spin-s field $\Psi(x)$ transforms as $\Psi(x) \to \tilde{\Psi}(x) = S[\Lambda]\Psi\left(\Lambda^{-1}x\right)$. Here, the matrices $S[\Lambda]$ form a representation of the Lorentz group (i.e. $S[\Lambda_1]S[\Lambda_2] = S[\Lambda_1\Lambda_2]$, $S[\Lambda^{-1}] = S[\Lambda]^{-1}$ and $S[I] = I$) with group generators $S^{\alpha\beta} = -S^{\beta\alpha}$ such that a finite Lorentz transformation is described by $S[\Lambda] = \exp\left(\frac{1}{2}\Omega_{\alpha\beta}S^{\alpha\beta}\right)$.[6] The matrices $S[\Lambda]$ depend on the spin of the field perturbation since the generators $S^{(a)(b)}$ depend on whether we are looking into, for example, the scalar or spinor representation of the Lorentz group, we will give more details on these two cases below.

In Minkowski spacetime, under Lorentz transformations the derivative of a spin-s field transforms as $\partial_\mu\Psi \to \partial_\mu\tilde{\Psi}(x) = \Lambda_\mu{}^\nu S[\Lambda]\partial_\nu\Psi(\Lambda^{-1}x)$. If we want to couple the spin-s

[5] Named so due to the pioneering work of Teukolsky for field perturbations on the Kerr black hole background [8].

[6] Note that we are applying the same Lorentz transformation to x and Ψ; thus the coefficients $\Omega_{\alpha\beta}$ of the transformations Λ and $S[\Lambda]$ are the same although the bases of generators $\mathcal{M}^{\alpha\beta}$ and $S^{\alpha\beta}$, respectively, are different.

field to a curved background, while preserving general covariance, one needs to promote the partial derivative ∂_μ to a covariant derivative $\mathcal{D}_\mu$. This promotion is chosen such that any function of Ψ and $\mathcal{D}_\mu\Psi$ that is a scalar under Lorentz transformations in Minkowski spacetime remains a scalar – under general coordinate transformations and local changes in the *vierbein* – in the curved background. This is the case if, under an arbitrary Lorentz transformation, the covariant derivative $\mathcal{D}_\mu$ still transforms as a derivative of a spin-s field:

$$\mathcal{D}_\mu\Psi(x) \to \mathcal{D}_\mu\tilde{\Psi}(x) = \Lambda_\mu{}^\nu S[\Lambda]\mathcal{D}_\nu\Psi\left(\Lambda^{-1}x\right). \tag{1.2.2}$$

It follows that the covariant derivative of a spin-s field that preserves Lorentz invariance in a curved background is

$$\mathcal{D}_\mu = \partial_\mu - \Gamma_\mu - iqA_\mu\,, \tag{1.2.3}$$

where we took the opportunity to allow the spin-s field to have a charge q (that couples to the Maxwell background field A_μ), and Γ_μ is a covariant spin connection

$$\Gamma_\mu = -\frac{1}{2}\omega_{\mu(a)(b)}S^{(a)(b)}, \tag{1.2.4}$$

which *depends* on the spin of the field it acts on.

In more detail, for a scalar (spin-0) field $\Psi \equiv \Phi$ the Lorentz group generator is simply $S^{(a)(b)} = 0$. Therefore $S[\Lambda] = 1$ and the scalar covariant derivative (1.2.3) is simply $\mathcal{D}_\mu = \partial_\mu - iqA_\mu$. The action for a massive charged complex scalar field is given by $\mathcal{S}_\Phi = \int_\mathcal{M} d^4x\,\sqrt{-g}\left(\mathcal{D}_\mu\Phi\mathcal{D}^\mu\Phi^* + m^2\Phi\Phi^*\right)$ where * stands for complex conjugation. The factor of $\sqrt{-g}$ is introduced to ensure that the Lagrangian $\mathcal{L}_\Phi$ is a scalar density and thus the action $\mathcal{S}_\Phi = \int_\mathcal{M} d^4x\mathcal{L}_\Phi$ is a scalar. Varying this action w.r.t. Φ^* one gets the Klein-Gordon equation for the scalar field

$$\mathcal{D}_\mu\mathcal{D}^\mu\Phi - m^2\Phi = 0 \tag{1.2.5}$$

and similarly for Φ^*.

On the other hand, for a (spin-$\frac{1}{2}$) Dirac 4-spinor field[7] $\Psi \equiv \psi$, out of the gamma matrices $\gamma^{(a)}$ (2.2.8) that satisfy the covariant Clifford algebra (2.2.9) one can build the commutator ($[A, B] = AB - BA$)

$$S^{(a)(b)} = \frac{1}{4}\left[\gamma^{(a)}, \gamma^{(b)}\right], \tag{1.2.6}$$

that satisfies the Lorentz Lie algebra.[8] $S^{(a)(b)}$ is the generator of the Lorentz group in

[7] The Dirac spinor is a 4-component field ψ^A with complex components $A = 1, 2, 3, 4$. In our study we will typically omit the spinorial indices and simply write $\psi \equiv \psi^A$, $\gamma^{(a)} \equiv (\gamma^{(a)})^A{}_B$ and $S[\Lambda] \equiv S[\Lambda]^A{}_B$.

[8] More precisely, Dirac 4-spinor fields are invariant under internal local Lorentz transformations of the spinor representation of the $SU(4)$ group. There are 15 Dirac matrices that provide a 4×4 fundamental representation of the SU(4) group [13]. These are the four vectors $\gamma^{(a)}$ introduced in (2.2.8), the six tensors

the spinor representation and replacing this (1.2.6) into (1.2.4) and then the latter Γ_μ into (1.2.2) one gets the spinor covariant derivative $\mathcal{D}_\mu$, namely (1.2.3), that acts on the Dirac spinor ψ.

The action that is Lorentz invariant and describes the coupling of a spin-$\frac{1}{2}$ fermion field ψ to a curved background $\mathcal{M}$ is

$$\mathcal{S}_D = \int_{\mathcal{M}} d^4x \sqrt{-g} \left(\frac{i}{2} \left[\bar{\psi} \gamma^\mu \mathcal{D}_\mu \psi - \left(\mathcal{D}_\mu \bar{\psi} \right) \gamma^\mu \psi \right] - m \bar{\psi} \psi \right) \tag{1.2.7}$$

where we have introduced the Dirac adjoint $\bar{\psi} = \psi^\dagger \gamma^{(0)}$ with $\psi^\dagger = (\psi^*)^T$ being the Hermitian adjoint of the multi-component field ψ. One needs to work with the Dirac adjoint because the Fermi bilinears $\bar{\psi}\psi$ and $\bar{\psi}\gamma^\mu\psi$ transform covariantly (as a scalar and as a vector, respectively) under the Lorentz group (while the Hermitian partner objects do not).

Varying the action $\mathcal{S}_D$ w.r.t. $\bar{\psi}$ and ψ, respectively, one gets the Dirac equations

$$\begin{aligned} \left(i \overrightarrow{\not{\mathcal{D}}}_\mu - m \right) \psi = 0 \quad &\Leftrightarrow \quad i \gamma^\mu \mathcal{D}_\mu \psi - m \psi = 0 \,, \\ \bar{\psi} \left(i \overleftarrow{\not{\mathcal{D}}}_\mu + m \right) = 0 \quad &\Leftrightarrow \quad i \left(\mathcal{D}_\mu \bar{\psi} \right) \gamma^\mu + m \bar{\psi} = 0 \,. \end{aligned} \tag{1.2.8}$$

This concludes our introduction to black hole field perturbations, we will explicitly solve the Dirac and Klein-Gordon equations in part II. In asymptotically AdS backgrounds, there exists a special relationship between theories of gravity and conformal field theories (CFTs) living on the boundary, this is known as the *AdS CFT correspondence* [15] which we will introduce in the next section.

1.3 AdS/CFT correspondence.

The AdS/CFT correspondence is an example of a more universal physical concept called the *holographic principle*, which states that the physics describing quantum gravity in a $d+1$-dimensional *bulk* volume of spacetime is encoded in a quantum field theory (QFT) which resides on the d-dimensional boundary, or vice versa. Given that both theories describe the same physics, the degrees of freedom on both sides must match, this gives rise to the so called *holographic dictionary*. The AdS/CFT duality tells us that string theory on an asymptotically AdS (AAdS) spacetime (and a direct product of a compact manifold, such as a sphere) has the same physics as the CFT that lives on the boundary of the bulk spacetime. In the case of AdS/CFT we have a strong/weak coupling duality, that is— a strongly coupled string theory on an AAdS background is dual to a weakly coupled CFT living on the boundary and vice versa. This is particularly useful given that

$S^{(a)(b)}$ defined (1.2.6), the pseudoscalar $\gamma^{(5)}$ introduced in (2.2.10), and four axial vectors $\gamma^{(5)}\gamma^{(a)}$.

generally, strongly coupled string theories are less well understood than some weakly coupled CFTs. For example the theory of quantum chromodynamics (QCD): progress has been made in understanding the QCD phase diagram at points which are hard to access through traditional QFT techniques, through studying physics on the gravity side via the AdS/CFT correspondence. See for example [16, 17, 18, 19]. Another major milestone in the holographic programme is the method of holographic renormalisation [20] (see also [21] for a real-time prescription), whereby one can compute renormalised QFT correlation functions by performing computations exclusively on the gravity side. It is a well-known fact that string theories on AAdS spacetimes reduce to supergravity in their low energy limit, which involves classical solutions of supergravity coupled to an infinite tower of Kaluza-Klein (KK) modes resulting from the dimensional reduction of supergravity in 11 dimensions[9]. To understand what the low energy limit on the QFT side means, it is useful to come back the the example of QCD. To be precise, QCD is a gauge field theory which was formulated to describe the strong interaction of quarks and gluons, it is constructed as an $SU(3)$ gauge theory and is consistent with experimental data (that which is accessible). The theory of QCD possesses a special property: it is asymptotically free. This means that at high energies (and short distances), the constituents of QCD behave essentially as free particles, therefore allowing for perturbative methods to illuminate key features. However at large distances the quarks and gluons interact strongly (and with a low energy), therefore the theory is less compliant with traditional gauge theory perturbation techniques. Two examples of properties that have been hard to understand in the low energy regime are confinement and chiral symmetry breaking. Now consider a general $SU(N)$ gauge theory, the low energy limit corresponds to the 't Hooft large N limit [25], $\lambda \equiv g_{YM}^2 N \gg 1$ where g_{YM} is the Yang-Mills coupling. What 't Hooft showed was that if one considers the integer number N as a parameter, the $SU(N)$ gauge theories simplify in the large N limit, giving rise to a perturbative expansion in $1/N$. In taking the large N limit one must keep λ *fixed*. This is in order to keep leading terms in the beta function equation of the same order in the $N \to \infty$ limit in the perturbative expansion[10].

We now proceed by presenting an example which was also discussed in the review [26], where we derive the celebrated Breitenlöhner-Freedman (BF) bound, of crucial importance to this research. It is well known that the AdS gravitational potential well acts as a reflecting mechanism for classical fields. This is due to the fact that light rays can travel to the AdS boundary[11] and back again– in finite time– as seen by an observer moving along geodesics in AdS. Whereas massive particles moving along geodesics can never

[9] Equivalently these KK modes are the result of reducing type IIB supergravity (or type IIA supergravity, all of these supergravity theories are inextricably linked to one another via the web of string dualities. For an excellent review on these dualities see for example [22] and for detailed reviews on KK reductions see [23, 24].

[10] When the QFT is a conformal theory, other limits make sense, such as $\lambda \to \infty$ but this is beyond the scope of this work, it is discussed however in the article [26] and references therein.

[11] In the discussion surrounding equation (1.3.4) we will define precisely the AdS boundary

reach the boundary. This translates into imposing reflective boundary conditions at the AdS boundary (which equivalently imply that there is no energy flux 'leakage' through the boundary) for whatever fields are probing the AdS background.

Consider a scalar field ϕ propagating in AdS_{d+1} spacetime. The field equation is given by $(\nabla_\mu \nabla^\mu - m^2)\phi = 0$, where ∇ is the covariant derivative on AdS and m is the mass of the scalar field. The metric of AdS_{d+1} in global coordinates $x^\mu = (\tau, \rho, \Omega_i)$ is given by

$$\mathrm{d}s^2 = L^2 \left(-\cosh^2(\rho)\mathrm{d}\tau^2 + \mathrm{d}\rho^2 + \sinh^2(\rho)\mathrm{d}\Omega^2 \right), \tag{1.3.1}$$

for $\rho \geq 0$ and $0 \leq \tau < 2\pi$. The above metric represents a hyperboloid

$$X_0^2 + X_{d+1}^2 - \sum_{i=1}^{d} X_i^2 = L^2 \tag{1.3.2}$$

embedded in $d+2-$ dimensional flat space with metric $\mathrm{d}s^2 = -\mathrm{d}X_0^2 - \mathrm{d}X_{d+1}^2 + \sum_{i=1}^{d} \mathrm{d}X_i^2$. The hyperboloid equation (1.3.2) is solved by

$$\begin{aligned} &X_0 = L\cosh(\rho)\cos(\tau)\,, \quad X_{d+1} = L\cosh(\rho)\sin(\tau) \\ &X_i = L\sinh(\rho)\Omega_i\,, \quad \left(i = 1, \cdots, d; \sum_i \Omega_i^2 = 1 \right). \end{aligned} \tag{1.3.3}$$

Notice that the metric (1.3.1) is well behaved for all values of τ therefore we can extend the spacetime to $-\infty < \tau < \infty$. We wish to study the causal structure of field perturbations on AdS, therefore it is convenient to introduce the new coordinate θ which is related to ρ via $\tan(\theta) = \sinh(\rho)$ such that $0 \leq \theta < \pi/2$. The metric (1.3.1) becomes[12]

$$\mathrm{d}s'^2 = -\mathrm{d}\tau^2 + \mathrm{d}\theta^2 + \sin^2(\theta)\mathrm{d}\Omega^2. \tag{1.3.4}$$

Above is the famous metric of the Einstein static universe except $\theta \in (0, \pi/2)$ rather than the usual $\theta \in (0, \pi)$ hence we say AdS_{d+1} can be conformally mapped to *half* of the Einstein static universe. The origin of our coordinate system is located at $\theta = 0$ and the conformal boundary is at $\theta = \pi/2$. In the coordinates (1.3.4) scalar fields can be decomposed as

$$\phi(x^\mu) = \mathrm{e}^{-i\omega\tau}\tilde{F}(\theta)Y_\ell(\Omega_{d-1}), \tag{1.3.5}$$

where Y_ℓ are the $d-1$ dimensional spherical harmonics which satisfy $\triangle Y_\ell = \ell(\ell+d-2)Y_\ell$ (here $\triangle$ represents the Laplacian on S^{d-1}). The remaining ODE in terms of the radial coordinate θ resembles that of the standard hypergeometric ODE and the solution $\tilde{F}(\theta)$ is

$$\tilde{F}(\theta) = (\sin\theta)^\ell (\cos\theta)^{\lambda_\pm} {}_2F_1(a_-, a_+, c; \sin\theta), \tag{1.3.6}$$

[12] After a conformal rescaling $\times L^{-2}\cos^2(\theta)$, recall that a conformal transformation does not change the causal structure of spacetime

where ${}_2F_1$ is the standard hypergeometric function. The parameters $a_\pm, c$ and $\lambda_\pm$ are given by

$$\begin{aligned} a_\pm &= \frac{1}{2}(\ell + \lambda_\pm \pm \omega L)\,, \\ c &= \ell + \frac{d}{2} \quad \text{and} \quad \lambda_\pm = \frac{d}{2} \pm \frac{1}{2}\sqrt{d^2 + 4m^2L^2}. \end{aligned} \tag{1.3.7}$$

The total energy of the scalar field perturbation is given by $E = \int \mathrm{d}^d x \sqrt{-g} T_0^0$ and it is conserved if the energy momentum flux of the boundary term at $\theta = \pi/2$ vanishes:

$$\int_{S^{d-1}} \mathrm{d}\Omega_{d-1} \sqrt{g} n_i T_0^i|_{\theta=\pi/2} = 0. \tag{1.3.8}$$

Here n_i is the unit normal to the hypersurface located at the boundary $\theta = \pi/2$. This demonstrates how the aforementioned 'reflective' boundary conditions are equivalent to requiring vanishing flux at the boundary. The solution (1.3.6) holds iff either a_+ or a_- is an integer, as defined in (1.3.7). If we further require the energy to be real we find

$$|\omega| L = \lambda_\pm + \ell + 2n\,, \quad (n = 0, 1, 2, \cdots). \tag{1.3.9}$$

This is only possible if $\lambda_\pm$, as defined in (1.3.7), are real. By looking at the expressions, we find that this translates to requiring the term in the square root to be positive, which implies

$$m^2 L^2 \geq -\frac{d^2}{4}. \tag{1.3.10}$$

The inequality (1.3.10) is the BF bound, [27, 28]. Notice that it already implies something quite different about scalar fluctuations on AdS when compared to flat space, namely scalar fields in AdS can still have positive energy even with a negative $(\text{mass})^2$, as long as it is 'not too negative'. Whereas scalars on Minkowski space have the bound $m^2 \geq 0$.

We still have two choices for $\lambda_\pm$ which provides a good opportunity to introduce some of the holographic dictionary for scalars. For scalar fields with mass satisfying $m^2 > -(d-2)(d+1)/4L^2$ we choose λ_+ since this term is normalisable whereas the λ_- is diverging at the boundary rendering it non-normalisable. This choice is usually referred to as the *standard quantisation*. For scalar fields with mass in the range $-\frac{d^2}{4L^2} \leq m^2 \leq -(d-2)(d+1)/4L^2$ both $\lambda_\pm$ are normalisable and one can choose either. The choice λ_- is referred to as the *alternate quantisation*. Near the boundary we can expand our scalar field as

$$\phi \sim \phi_+(x)(\theta - \pi/2)^{\lambda_+} + \cdots + \phi_0(x)(\theta - \pi/2)^{\lambda_-} + \cdots\,, \tag{1.3.11}$$

where x refers to the remaining coordinates other than θ. Consider scalar field masses in the range $m^2 > -(d-2)(d+1)/4L^2$ where only the standard quantisation is possible.

Through the non-normalisable λ_- term one can define associated boundary fields via

$$\phi_0(x) \equiv \lim_{\theta\to\pi/2} \phi(\theta, x)(\theta - \pi/2)^{-\lambda_-}. \tag{1.3.12}$$

Now, the gauge/gravity duality tells us that every field Φ that propagates in the bulk spacetime has a dual gauge invariant operator $\mathcal{O}_\Phi$ living on the boundary. It was proposed in [15, 29] that the boundary values of fields (*e.g.* the scalar field boundary value from equation (1.3.12)) are identified as sources for the boundary operator which is dual to the normalisable field. In our case this translates to the following statement: the normalisable mode ϕ_+ is dual to a scalar field theory operator $\mathcal{O}_{\phi_+}$ which has conformal dimension $\triangle \equiv \lambda_+$. The source ϕ_0 can then be identified as a vacuum expectation value for $\mathcal{O}$ or a one-point correlation function, with higher point functions computed via functional differentiation with respect to the source. It should be noted however, that there are divergences on the gravity side due to the infinite volume of spacetime and QFT correlation functions diverge too, so in order for the formulae to make sense one must renormalise. As mentioned before the holographic renormalisation programme was rigorously developed in [20]. To complete our discussion I will outline briefly the dualities of other bulk-boundary fields. Suppose we have a classical gravity action $S[\phi, A_\mu, g_{\mu\nu}, \cdots]$, owing to the procedure of holographic renormalisation there exists one exact renormalised one-point function for each field

$$\begin{aligned} &\phi \to \langle \mathcal{O}(x)\rangle_S\,, \\ &A_\mu \to \langle J_i(x)\rangle_S\,, \\ &g_{\mu\nu} \to \langle T_{ij}(x)\rangle_S. \end{aligned} \tag{1.3.13}$$

Above is the scalar field which was already discussed. Then there is the boundary symmetry current J_i which is dual to the gauge field in the bulk A_μ and finally T_{ij} is the boundary energy-momentum tensor that couples to the boundary metric. The results described above hold in the saddle point approximation, where classical gravity is valid. For a detailed discussion of renormalised correlation functions in holography see the review [30].

1.4 Hairy black holes

Uniqueness theorems in General Relativity have been of interest since the pioneering work of Israel, Penrose, Wheeler and others in the 1960's, many of these results became known as *no hair theorems*. These results imply for instance that the unique asymptotically flat, stationary black hole (BH) solution to Einstein-Maxwell theory is the Kerr-Newman solution. That is, the BH solution can be characterised by three parameters, namely the mass, charge and angular momentum of the BH. In the case of static BH's,

the no hair theorems imply that the unique solution to Einstein-Maxwell is the RN solution (1.1.5) which is characterised by two parameters: mass and charge. For historical reviews on these theorems see for example [31, 32]. The latter case is known as Birkhoff's theorem [33, 34] and is the one of relevance to this thesis. Birkhoff's theorem for Einstein-Maxwell theory states that the unique spherically symmetric solution of the Einstein-Maxwell equations with non-constant area radius function r is the Reissner-Nordström solution. If r is constant then the theorem does not apply since one has the Bertotti-Robinson ($AdS_2 \times S^2$) solution.

Consider a RN BH in asymptotically flat space as given in (1.1.5), it has been established that the RN solution is linear-mode stable to scalar field and gravito-electromagnetic perturbations, intuitively this is to be expected since the field perturbation scatters off the BH horizon into the asymptotic region where it decays at the Minkowski boundary. However, if one confines the BH to a 'box' or a 'thin shell' (which we soon describe), a scalar field perturbation (for example) can drive the system unstable. This was first established as Press and Teukolsky's black hole bomb system, where a scalar perturbation (with reflecting boundary conditions imposed at the box outer layer) may extract energy continuously from the BH horizon, leading to a superradiant instability [35]. In the case of charged black holes superradiance occurs if the frequency of the scalar perturbation satisfies $\omega < e\mu$ [36]. More recently discovered was the near-horizon scalar condensation instability, whereby an (asymptotically stable) charged scalar field perturbation may (under certain conditions) violate the AdS_2 Breitenlöhner-Freedman (BF) mass bound in the near horizon geometry, which is a direct product of $AdS_2 \times S^2$. This instability was first found in the context of the AdS/CFT correspondence in the holographic superconductor programme, see for example [37, 38] where charged, planar AdS BH solutions were considered. The instability is also present in the case of charged global AdS black holes [39]. The superradiant and near horizon instabilities are in general entangled in the phase diagram, however in certain limits (as we discuss in the next sections 1.4.1 and 1.4.2) they disentangle and reveal a difference in nature. Due to the similarity between the confining box and the AdS boundary, this instability also applies to RN BH's in a box, as was shown in [40]. In the case of RN AdS, time evolutions of the aforementioned theories were carried out in [41, 42] which indicate that the endpoint of the instability should be a charged black hole with scalar hair floating above the horizon or equivalently a BH solution to Einstein-Maxwell-scalar field theory with a negative cosmological constant. It is thus expected that in the case of RN confined to a Minkowski box, the endpoint of the unstable RN configuration should be a charged black hole with charged scalar hair floating above the horizon- where the electrostatic repulsion is balanced by the tendency towards gravitational collapse. With this intuition, hairy black holes in Einstein-Maxwell-scalar field theory confined in a box have been constructed in a perturbative expansion in [43]. The goal of part IV is to extend these studies to higher energies in order to illuminate the full phase diagram of hairy BH's in a Minkowski box.

There are other interesting solutions to Einstein-Maxwell-scalar theory, namely *solitons* or *boson stars* and they will play a role in the discussion of hairy BH's. The full discussion of these soliton solutions turns out to be rather involved and we discuss it in detail in part III. It should be noted that hairy BH's in AdS spacetimes have also been constructed and they indeed display many similar properties to the BH's we present, see for example [44, 45, 46, 47].

1.4.1 Superradiance

Let us now describe the conditions required for superradiance in the case of scalar field perturbations (with electric charge q and mass m) on a RN BH confined inside a box. First, the presence of a confining box placed at $r = L$ allows the system to acquire a scaling symmetry

$$\begin{aligned} \{t, r, x, \varphi\} \to \{\lambda_1 t, \lambda_1 r, x, \varphi\}, \qquad \{f, A_t, \varphi\} \to \{f, A_t, \varphi\}, \\ \{q, L, r_+, m\} \to \left\{ \frac{q}{\lambda_1}, \lambda_1 L, \lambda_1 r_+, \frac{m}{\lambda_1} \right\} \end{aligned} \tag{1.4.1}$$

which leaves the equations of motion invariant and rescales the line element and the gauge field 1-form as $\mathrm{d}s^2 \to \lambda_1^2 \mathrm{d}s^2$ and $A_t \mathrm{d}t \to \lambda_1 A_t \mathrm{d}t$. Above $x = \cos(\theta)$ and φ is the azimuthal angular coordinate, together they parametrise S^2. We use this scaling symmetry to work with dimensionless coordinates and measure thermodynamic quantities in units of L (effectively this sets $L \equiv 1$),

$$T = \frac{t}{L}\,, \qquad R = \frac{r}{L}\,; \qquad R_+ = \frac{r_+}{L}\,, \qquad e = qL\,, \qquad m_\phi = mL\,. \tag{1.4.2}$$

The box is now located at $R = 1$. Recall that RN is a two-parameter solution, let us take these to be the horizon radius R_+ and the chemical potential μ. A charged scalar field in this background obeys the Klein-Gordon equation (1.2.5). In order to solve Klein-Gordon as a well-posed boundary value problem one must specify the boundary conditions that the scalar field must obey, both at the horizon $R = R_+$ and at the outer boundary, namely the box outer layer at $R = 1$. Outside the box $R > 1$ we are in the conditions for Birkhoff's theorem and therefore the solution must be RN, hence $\phi|_{R\geq 1} = 0$.

To explain the origin of the superradiant instability it is useful to first consider the $R_+ = 0$ limit of equation (1.4.4) and choose the gauge where $A_t|_{R_+} = 0$. This translates into the case of a scalar perturbation confined inside a box placed at $R = 1$ in Minkowski space with a constant gauge field $A = \mu \mathrm{d}t$. Naturally, the inner boundary for this case becomes the origin $R = 0$ and here we require regularity of our scalar field perturbations. The outer boundary condition remains $\phi(R = 1) = 0$. In these conditions the frequency spectrum of the scalar perturbations becomes quantised, let us now describe this in more detail. The background at hand is time independent, therefore one has ∂_t as a

Killing vector. This allows the Fourier decomposition of linear modes in the time coordinate which introduces the dimensionless frequency $\Omega = \omega L$. The background is also axisymmetric (*i.e.* has ∂_φ as a Killing vector), therefore one can also Fourier decompose in the φ coordinate. An appropriate field ansatz is then given by

$$\phi(T, R, x, \varphi) = \mathrm{e}^{-i\Omega T}\mathrm{e}^{im_\varphi \varphi} P_\ell^{m_\varphi}(x)\psi(R), \tag{1.4.3}$$

where $\mathrm{e}^{im_\varphi\varphi}P_\ell^{m_\varphi}(x)$ is a representation of the spherical harmonics $Y_{\ell m_\varphi}$, this is to be expected since the system has spherical symmetry. Note that $P_\ell^{m_\varphi}(x)$ are the usual Legendre polynomials[13]. The Klein-Gordon equation (1.2.5) in the coordinate system (1.1.5) is given by

$$\partial_R\left(R^2 f \partial_R \psi\right) + \left(\frac{R^2}{f}\left(\Omega + eA_t\right)^2 - \ell(\ell+1) - m_\phi^2 R^2\right)\psi = 0. \tag{1.4.4}$$

With the ansatz (1.4.3), in the $R_+ = 0$ limit, the Klein-Gordon equation can be solved by the function [40]

$$\psi(R) = \frac{\left[(\Omega + e\mu)^2 - m_\phi^2\right]^{-\frac{1}{4}}}{\sqrt{R}}\left[\alpha_1 J_{\ell+\frac{1}{2}}\left(R\sqrt{(\Omega+e\mu)^2 - m_\phi^2}\right) + \alpha_2 Y_{\ell+\frac{1}{2}}\left(R\sqrt{(\Omega+e\mu)^2 - m_\phi^2}\right)\right], \tag{1.4.5}$$

where $J_\nu(z)$, $Y_\nu(z)$ are Bessel functions of the first and second kind, respectively (see for example [48]), and $\alpha_{1,2}$ are arbitrary integration constants. Regularity of the field perturbation at the origin requires that we set $\alpha_2 = 0$ (in order to avoid a divergence of the order $R^{-\ell-1}$). Now, at the location of the box we require the scalar field to vanish. The Bessel function $J_\nu(z)$ has an infinite number of simple zeros for $\nu \in \mathbb{R}$, these are denoted by $j_{\nu,n}$ where $n \in \mathbb{N}$. Thus the boundary condition $\psi(1) = 0$ quantises the frequencies as $\Omega_{\ell,n}$ with

$$\Omega_{\ell,n} = \sqrt{j_{\ell+\frac{1}{2},n}^2 + m_\phi^2} - e\mu. \tag{1.4.6}$$

In parts III and IV we will search for spherically symmetry hairy solutions, therefore we set $\ell = 0$ and we also consider the perturbation with lowest energy, hence we set $n = 1$ in (1.4.6). For concreteness let us also consider massless scalar fields $m_\phi = 0$. These choices give the smallest frequency for a scalar perturbation trapped in a Minkowski box with a constant gauge field, it reads

$$\Omega_{0,1} = \pi - e\mu. \tag{1.4.7}$$

Suppose now we place our box and scalar perturbation in the RN background, we expect the frequencies to become complex therefore they will differ from (1.4.6). However if we consider small RN black holes[14] the frequency can be approximated by a series expansion in R_+. Similar expansions have been performed in AdS (which behaves similarly

[13] Here $m_\varphi = 0, \pm1, \pm2, \cdots$ and $\ell = 0, 1, 2, \cdots$ are such that $|m_\varphi| \leq \ell$ (the integer ℓ also gives the number of zeros of the eigenfunction in the polar direction $x = \cos(\theta)$).

[14] Recall $R_+ \equiv \frac{r_+}{L} \leq 1$ hence small black holes have $R_+ \ll 1$.

to our confining box), see [46, 39]. To be specific, we expand our fields as

$$\phi^{\mathcal{R}}(R) = \sum_{j\geq 0} R_+^j \phi_j^{\mathcal{R}}(R), \quad \Omega = \sum_{j\geq 0} R_+^j \Omega_j. \tag{1.4.8}$$

Above $\mathcal{R}$ refers to the region we are considering, since in these calculations we must solve the radial equation separately in the near horizon and asymptotic regions (in part II we will explain this more specifically). One must then impose (as before) that the scalar field vanishes at the box (at every order in the expansion) and that the perturbations are smooth in ingoing Eddington-Finkelstein coordinates at the RN horizon. It turns out that the leading order term in the series expansion for Ω is given by $\Omega_0 \equiv \Omega_{0,1} = \pi - e\mu$. At next to leading order, one finds that the frequency acquires an imaginary term in the expansion proportional to $\Omega_{1,0}$. To be precise, $\Omega_1 = \Re(\Omega_1) + i\Im(\Omega_1)$ where $\Im(\Omega_1) \sim \alpha_+(e\mu - \pi)$ with $\alpha_+ \geq 0$. This gives rise to a prediction, namely that these 'small BH' scalar perturbations are unstable to superradiance when $\Im(\Omega_1) \geq 0 \Leftrightarrow e \geq \pi/\mu$. The temperature of RN BHs confined to a box, in our coordinates (1.1.5) is given by $T_H L = \frac{1}{8\pi r_+}(2-\mu^2)$ hence the extremal RN BH has $\mu = \sqrt{2}$. Putting all this together we can make the statement that small extremal RN BHs confined to a box are superradiantly unstable to charged scalar perturbations when the scalar field charge e satisfies

$$e \geq \frac{\pi}{\sqrt{2}} \sim 2.221 \equiv e_S. \tag{1.4.9}$$

The above bound will be critical to our discussions in parts III and IV. This bound was found within a perturbative expansion for small horizon radius R_+, therefore at the onset of this (linear) instability we expect that a non-linear back-reaction should find hairy black hole solutions which bifurcate (continuously) from the onset of the RN instability. Moreover, since the onset occurs for $R_+ \ll 1$ we expect these hairy black holes to have a zero horizon limit. This $R_+ \to 0$ limit describes a one-parameter family of solutions that we call solitons. Solitons can therefore also be thought of as a non-linear back reaction of the boxed Minkowski scalar field modes in (1.4.6). It turns out that the above intuition only holds in certain regions of phase space when one considers the full Einstein-Maxwell-Scalar field theory trapped in a Minkowski box. Solitons exist in other regions of phase space where they are unrelated to hairy black holes, but the story is more intricate and we leave further details for later discussions in parts III and IV.

1.4.2 Near horizon scalar condensation

The following limit was first presented in [49] where the near-horizon geometry of the extremal Kerr solution in 4-dimensions was computed. This limiting procedure applies to many known black hole solutions including RN confined to a box. The near horizon geometry of the extremal RN solution (with $\mu = \sqrt{2}$) is found by making a change of

coordinate

$$T = L^2_{AdS_2}\frac{\tau}{\lambda}, \quad R = R_+ + \lambda\rho, \tag{1.4.10}$$

and taking the limit $\lambda \to 0$. At leading order, the metric becomes

$$\begin{aligned} \mathrm{d}s^2 &= L^2_{AdS_2}\left(-\rho^2\mathrm{d}\tau^2 + \frac{\mathrm{d}\rho^2}{\rho^2}\right) + R_+^2\mathrm{d}\Omega_2^2, \\ A_\mu \mathrm{d}x^\mu &= \sqrt{2}L_{AdS_2}\rho\mathrm{d}\tau, \quad \text{with} \quad L_{AdS_2} = R_+. \end{aligned} \tag{1.4.11}$$

The above geometry parametrises a direct product of AdS_2 with S^2 and it is still a solution to the full Einstein-Maxwell gravity in 4 dimensions (this is because the metric (1.4.11) that results from coordinate transformation (1.4.10) and the $\lambda \to 0$ limit, is well behaved) as well as being a solution to AdS_2 gravity. Now, we can also take the near horizon extremal limit on the Klein-Gordon equation (1.4.4)[15]. In these conditions, the Klein-Gordon equation reads

$$\partial_\rho\left(\rho^2\partial_\rho\psi\right) + \left(\frac{1}{\rho^2}\left(\tilde{\Omega} + e\sqrt{2}L_{AdS_2}\rho\right)^2 - \ell(\ell+1) - L^2_{AdS_2}m^2_\phi\right)\psi = 0. \tag{1.4.12}$$

Recall that expanding a field in AdS near the boundary gives rise to the BF bound, as detailed in the previous section. Taking the large ρ expansion of the near horizon scalar field ψ gives

$$\psi|_{\rho\to\infty} \sim \alpha_-\rho^{-\Delta_-} + \cdots + \alpha_+\rho^{-\Delta_+} + \cdots, \quad \text{with} \quad \Delta_\pm = \frac{1}{2} \pm \frac{1}{2}\sqrt{1 + 4m^2_{eff}L^2_{AdS_2}}. \tag{1.4.13}$$

The above expansion describes a massive scalar field (with an effective mass m_{eff}) propagating at the AdS_2 boundary and therefore the perturbation is unstable if the AdS_2 BF bound is violated (*i.e.* the bound (1.3.10) with $d = 1$):

$$m^2_{eff}L^2_{AdS_2} \geq -\frac{1}{4}. \tag{1.4.14}$$

The effective mass is given by the expression [40]

$$m^2_{eff}L^2_{AdS_2} \equiv m^2_\phi L^2_{AdS_2} + \ell(\ell+1) - 2e^2R_+^2. \tag{1.4.15}$$

Rearranging the inequality (1.4.14) and plugging in the expression (1.4.15) one finds that extremal RN BHs confined to a box could be unstable to scalar condensation if the scalar charge e satisfies:

$$e \geq \frac{\sqrt{(2\ell+1)^2 + 4m^2_\phi R_+^2}}{2\sqrt{2}R_+}. \tag{1.4.16}$$

[15]This is equivalent to re-deriving the Klein-Gordon equation in the near horizon coordinates (1.4.11), however one must remember to rescale the frequency in order to capture the time-dependence information correctly. The near horizon frequency is related to the frequency in the full geometry via $\Omega = \tilde{\Omega}\frac{\lambda}{L^2_{AdS_2}}$. This follows from making the identification $\mathrm{e}^{-i\Omega T} \equiv \mathrm{e}^{-i\tilde{\Omega}\tau}$.

Since our horizon radius takes the values $R_+ \in (0,1)$ we can expand the bound (1.4.14) as follows

$$e|_{R_+\sim 1} \geq \frac{\sqrt{(2\ell+1)^2+4m_\phi^2}}{2\sqrt{2}} + \frac{(2\ell+1)^2}{2\sqrt{2}\sqrt{1+4(m_\phi^2+\ell(\ell+1))}}(1-R_+) + \mathcal{O}\left[(1-R_+)^2\right]$$
$$e|_{R_+\sim 0} \geq \frac{2\ell+1}{2\sqrt{2}R_+} + \mathcal{O}(R_+). \tag{1.4.17}$$

Therefore it is clear to see that near horizon scalar condensation is suppressed for small RN black holes[16], whereas large BHs (which have $R_+ \sim 1$) can become unstable to scalar condensation. In our case we want to restrict ourselves to massless scalars, therefore in the limit where the black hole horizon approaches the box radius, scalar condensation occurs for scalar charges

$$e \gtrsim \frac{2\ell+1}{2\sqrt{2}} \tag{1.4.18}$$

Furthermore, if one restricts again to the spherically symmetric case $\ell = 0$ we find the critical charge $e \gtrsim \frac{1}{2\sqrt{2}} \approx 0.35 \equiv e_{NH}$. We will construct hairy black holes which are a result of near horizon scalar condensation for RN in a box in part IV. We have shown that in two corners of the phase diagram for RN BHs in a box, scalar perturbations can drive the system unstable to two physically different mechanisms. We find that the hairy black holes which bifurcate from the onset of these instabilities are also qualitatively very different in these corners of phase space and indeed show a mixture of properties as one ramps up the scalar field charge from e_{NH} to e_S. We will make all of this much more precise when we present our full numerical results in the main text.

1.4.3 Einstein-Maxwell-scalar gravity.

We expect hairy black holes to arise from solving the full non-linear Einstein-Maxwell-scalar system, therefore let us give the physical set-up. The action for Einstein-Maxwell gravity in four dimensions coupled to a complex scalar field with charge q is given by

$$S = \frac{1}{16\pi G_N}\int \mathrm{d}^4x\sqrt{-g}\left(\mathcal{R} - \frac{1}{2}F_{\mu\nu}F^{\mu\nu} - 2D_\mu\phi(D^\mu\phi)^\dagger + V(|\phi|)\right), \tag{1.4.19}$$

where $\mathcal{R}$ is the Ricci scalar, A is the Maxwell gauge potential, $F = \mathrm{d}A$, and $D_\mu = \nabla_\mu - iqA_\mu$ is the gauge covariant derivative of the system. We fix Newton's constant $G_N \equiv 1$.

Let us consider the quadratic potential $V(|\phi|) = m^2\phi\phi^\dagger$ where m is the scalar field mass. The equations of motion which follow from extremising the action (1.4.19) with respect

[16] This is because for small BHs $1/R_+$ blows up and one requires an unphysical, large scalar field charge.

to the gravitational field $g_{\mu\nu}$, the Maxwell gauge field A_μ and scalar field ϕ are given by

$$\begin{aligned}
&\mathcal{R}_{\mu\nu} - \frac{1}{2}\mathcal{R}g_{\mu\nu} = T_{EM} + T_\phi, \\
&\nabla_\nu F^{\nu\mu} = iqg^{\mu\nu}\left(\phi^\dagger D_\nu\phi - \phi(D^\mu\phi)^\dagger\right), \\
&D_\mu D^\mu\phi - m^2\phi = 0, \\
&\text{where} \quad T_{EM} = F_\mu{}^\rho F_{\nu\rho} - \frac{1}{4}g_{\mu\nu}F_{\rho\sigma}F^{\rho\sigma} \\
&\text{and} \quad T_\phi = D_\mu\phi\,(D_\nu\phi)^\dagger + D_\nu\phi\,(D_\mu\phi)^\dagger - g_{\mu\nu}g^{\rho\sigma}\,(D_\rho\phi)^\dagger\,D_\sigma\phi.
\end{aligned} \tag{1.4.20}$$

The instabilities that we discussed in the previous sections are present even for massless scalar fields, the electric charge that couples to the Maxwell field A is enough to drive the system unstable, therefore from here onwards and in the main text we will fix $m = 0$. Hairy solutions should exist for other values of the scalar field mass. We are interested in solutions of (3.3.1) that are static, spherically symmetric and asymptotically flat. Using reparametrisations of the time and radial coordinates, $t \to \tilde{t} = t + H(t, r)$ and $r \to \tilde{r}(r)$, we work in the 'radial/Schwarzschild gauge' where we fix the radius of a round S^2 to be the areal radius r and there is no cross term $dtdr$. An *ansatz* with the desired symmetries is

$$\mathrm{d}s^2 = -f(r)\mathrm{d}t^2 + g(r)\mathrm{d}r^2 + r^2\mathrm{d}\Omega_2^2, \qquad A_\mu\mathrm{d}x^\mu = A_t(r)\mathrm{d}t, \qquad \phi = \phi^\dagger = \phi(r), \tag{1.4.21}$$

with $\mathrm{d}\Omega_2^2$ being the metric for the unit 2-sphere. We choose to work with the static *ansatz* (1.4.21) where the scalar field is real. Horizonless solutions in this gauge are usually called solitons. However, we can also perform a $U(1)$ gauge transformation with gauge parameter $\chi = -\omega t/q$,

$$\phi = |\phi|e^{i\varphi} \to |\phi|e^{i(\varphi + q\,\chi)}\,, \qquad A_t \to A_t + \nabla_t\chi, \tag{1.4.22}$$

to rewrite the scalar field as $\phi = |\phi|e^{-i\omega t}$, in which case we would be in a frame where the scalar field oscillates in time with a frequency ω.[17] Horizonless solutions in this gauge are usually called boson stars. The solitons and boson stars of the theory are therefore the same objects since they differ only by a $U(1)$ gauge transformation.

Recall the dimensionless quantities (where thermodynamic quantities are measured in units of the box radius L):

$$T = \frac{t}{L}\,, \qquad R = \frac{r}{L}\,; \qquad R_+ = \frac{r_+}{L}\,, \qquad e = qL\,, \qquad m_\phi = mL\,. \tag{1.4.23}$$

The box is now at $R = 1$.

[17] However, since the energy-momentum tensor of the scalar field only depends on $\phi\phi^\dagger$ and $\partial\phi(\partial\phi)^\dagger$, in the new gauge the gravitational and Maxwell fields would still be invariant under the action of the Killing vector field ∂_t.

We now plug the ansatz (1.4.21) into the equations of motion (1.4.20), this gives four independent equations:

$$g'(R)+g(R)^2\left(\frac{1}{R}-\frac{e^2RA(R)^2\phi(R)^2}{f(R)}\right)-g(R)\left(\frac{RA'(R)^2}{2f(R)}+R\phi'(R)^2+\frac{1}{R}\right) \tag{1.4.24}$$

$$=f'(R)-\frac{f(R)(g(R)+R^2\phi'(R)^2-1)}{R}+\frac{R}{2}\left(A'(R)^2-2e^2A(R)^2g(R)\phi(R)^2\right) \tag{1.4.25}$$

$$=A''(R)+A'(R)\left(\frac{2}{R}-\frac{f'(R)}{2f(R)}-\frac{g'(R)}{2g(R)}\right)-2e^2A(R)g(R)\phi(R)^2 \tag{1.4.26}$$

$$=\phi''(R)+\frac{\phi'(R)}{2}\left(\frac{f'(R)}{f(R)}-\frac{g'(R)}{g(R)}+\frac{4}{R}\right)+\phi(R)\frac{e^2A(R)^2g(R)}{f(R)}=0. \tag{1.4.27}$$

Notice that the second equation above can be solved for $g(R)$ in terms of the other fields:

$$g(R)=\frac{R\left(RA'(R)^2+2f'(R)\right)+2f(R)\left(1-R^2\phi'(R)^2\right)}{2\left(e^2R^2A(R)^2\phi(R)^2+f(R)\right)} \tag{1.4.28}$$

From this we can obtain a set of three differential equations which we can solve for the fields $f(R)$, $A(R)$ and $\phi(R)$. The solution for $g(R)$ is then easily obtained via equation (1.4.28). We must supplement these ODE's with relevant boundary conditions, but we leave the discussion of this for the main text.

Recall that the RN solution (1.1.5) is a two-parameter solution to Einstein-Maxwell theory, with parameters R_+ and the chemical potential μ (say). By continuity we expect the hairy black holes that bifurcate from the unstable RN black hole to also be two-parameter solutions. Recall from the linear problem (1.4.4) that we require the scalar field to vanish at and outside the box by Birkhoff's theorem, but its derivative when approaching the box from the interior, *i.e.* as $R\to 1^-$, does not vanish (unless we have the trivial RN solution). We will call this quantity ϵ:[18]

$$\phi^{in}\big|_{R=1}=\phi^{out}\big|_{R=1}=0,\qquad \phi^{out}(R)=0,\qquad \phi'^{\,in}\big|_{R=1}\equiv\epsilon, \tag{1.4.29}$$

where ϕ^{in} refers to the solution inside the box and ϕ^{out} the solution outside. For $R\leq 1$ the scalar field is forced to have the Taylor expansion $\phi^{in}\big|_{R=1^-}=\epsilon(R-1)+\mathcal{O}(R-1)^2$. We are forcing a jump in the derivative of the scalar field normal to the timelike hypersurface Σ located at $R=1$. We will discuss the implications of this in section 3.3.2, essentially the jump in the scalar field derivative gives rise to an energy-momentum tensor at the box hypersurface layer $R=1$. This tensor gives the energy-momentum content of the box which confines the scalar hair.

Let us come back to the superradiant bound (1.4.9), suppose we have an extremal RN BH and a scalar perturbation with charge $e>e_S$. In these conditions the unstable RN BH should evolve towards a charged hairy black hole. Furthermore, we expect these

[18] Note that our theory has the symmetry $\phi\to-\phi$ so we can focus our attention only on the case $\epsilon>0$.

black holes (at least at small energies) to have a horizonless limit (*i.e.* $R_+ \to 0$) that is, the soliton described above. Indeed hairy black holes and solitons were constructed under these circumstances in [43]. This was achieved by a double expansion of the fields in the parameters R_+ and ϵ for hairy BHs and by a single expansion in ϵ for solitons. We define the chemical potential for hairy BHs by $\mu \equiv A^{in}(1) - A^{in}(R_+)$ therefore μ will also have an expansion for the black hole problem. To be precise, the fields f, A and ϕ from equations (1.4.24) are expanded as

$$\begin{aligned} f^{\mathcal{R}}(R) &= \sum_{j\geq 0} \epsilon^{2j} f^{\mathcal{R}}_{2j}(R), \quad A^{\mathcal{R}}(R) = \sum_{j\geq 0} \epsilon^{2j} A^{\mathcal{R}}_{2j}(R) \\ \phi^{\mathcal{R}}(R) &= \sum_{j\geq 0} \epsilon^{2j+1} \phi^{\mathcal{R}}_{2j+1}(R), \quad \mu \equiv A^{in}|_{R=1} \quad \text{for solitons.} \end{aligned} \tag{1.4.30}$$

Above the superscript $\mathcal{R}$ refers to fields inside and outside the box. Then for black holes we have

$$\begin{aligned} f^{in,\,near}(y) &= \sum_{j,k\geq 0} \epsilon^{2j} R_+^k f^{in,\,near}_{2j,k}(y), \quad A^{in,\,near}(y) = \sum_{j,k\geq 0} \epsilon^{2j} R_+^k A^{in,\,near}_{2j,k}(y) \\ \phi^{in,\,near}(y) &= \sum_{j,k\geq 0} \epsilon^{2j+1} R_+^k \phi^{in,\,near}_{2j+1,k}(y), \quad y = \frac{R}{R_+}, \quad \tau = \frac{T}{R_+}, \end{aligned} \tag{1.4.31}$$

$$\begin{aligned} f^{in,\,far}(R) &= \sum_{j,k\geq 0} \epsilon^{2j} R_+^k f^{in,\,far}_{2j,k}(R), \quad A^{in,\,far}(R) = \sum_{j,k\geq 0} \epsilon^{2j} R_+^k A^{in,\,far}_{2j,k}(R), \\ \phi^{in,\,far}(R) &= \sum_{j,k\geq 0} \epsilon^{2j+1} R_+^k \phi^{in,\,far}_{2j+1,k}(R), \quad \mu = \sum_{j,k\geq 0} \epsilon^{2j} R_+^k \mu_{2j,k} \end{aligned} \tag{1.4.32}$$

$$f^{out}(R) = \sum_{j,k\geq 0} \epsilon^{2j} R_+^k f^{out}_{2j,k}(R), \quad A^{out}(R) = \sum_{j,k\geq 0} \epsilon^{2j} R_+^k A^{out}_{2j,k}(R), \tag{1.4.33}$$

$$\phi^{out}(R) = 0.$$

The fields have to be solved in the near horizon and far (*i.e.* close to the cavity layer $R = 1$) regions separately and a matching procedure pieces together the solution. The coordinates $y = \frac{R}{R_+}$ and $\tau = \frac{T}{R_+}$ are introduced because the Taylor expansion of the 'in-far' fields breaks down at order $R/R_+ \sim \mathcal{O}(1)$. The junction conditions we will discuss in section 3.3.2 then join the 'in' and 'out' fields. In these conditions the hairy black holes as described above in (1.4.31) reduce to the solitons as described in (1.4.30) in the $R_+ \to 0$ limit. We confirm these expectations in parts III and IV when we solve the system numerically. Hairy black holes turn out to depend sensitively on the scalar field charge away from e_S (*i.e.* for $e < e_S$), which is only revealed once we solve the equations of motion in full. We also test the perturbative results from [43] in the main text, section 4.4.4.

This thesis is organised as follows, in part II we show that there exists a similar argument (as was presented for scalars in section 1.4.2) for spinors in a RN AdS background in 4-dimensions, to develop a near horizon instability at the linear-mode level. Motivated by this we rigorously search for unstable modes by solving the Dirac equation in RN AdS numerically, we also guide our numerics with some perturbative calculations which can be found in detail in the appendix A. However, we find no evidence of an instability, which has some interesting implications as we will discuss in the main text. In part III we discuss solitons confined in a Minkowski box, we introduced these solitons in section 1.4.3 but the phase diagram is plentiful and requires a detailed analysis. Finally in part IV we complete the discussion of the phase diagram of the charged black hole bomb by discussing in depth the hairy black holes of the theory. We also show that in certain regions of phase space hairy black holes and RN black holes co-exist, hence there is a non-uniqueness of solutions. However in these regions hairy black holes dominate the micro-canonical ensemble (they always have higher entropy) which suggests that they are the endpoint of the unstable RN configuration.

Part II

Hunting for fermionic instabilities in charged AdS black holes

CHAPTER 2

Hunting for (linear) fermionic instabilities

2.1 Introduction

It is a well known fact that bosonic waves impinging charged or rotating black holes can be amplified via superradiant scattering (see *e.g.* the review [50] and references therein). It follows that black holes perturbed by bosonic fields in the presence of a gravitational potential well – provided by, for example, an asymptotic anti-de Sitter (AdS) potential, a physical cavity at finite radius or by the mass of the field – can develop a superradiant instability. However, this is not the only instability that can be present in such systems. Indeed, the family of superradiant black holes always has a configuration with zero temperature. Typically, such extreme black holes have a near-horizon geometry that is the direct product (or a fibration) of a base space (*e.g.* a sphere) and an AdS_2 space [49]. These extreme (and near-extreme) black holes, when confined in a gravitational well, are unstable if the '*effective*' mass of the perturbation (as seen by an AdS_2 observer) violates the 2-dimensional Breitenlöhner-Freedman (BF) bound for stability [28, 27, 51, 52] (even though the asymptotic AdS_4 BF bound is obeyed). Recall that perturbations with a mass below this bound are normalizable (i.e. they have finite conserved energy) but their energy is negative and, consequently, they can trigger an instability. The superradiant and near-horizon instabilities have a different physical nature. But, in a non-extremal black hole they are usually entangled. However, if L is the typical dimension of the gravitational well (*e.g.* the radius of the cavity or the AdS radius), they disentangle for small dimensionless horizon radius r_+/L since the near-horizon instability is suppressed for

$r_+/L \ll 1$, while the superradiant instability is still present [53, 54, 55, 56, 57, 39, 40].

Not less well known is the fact that fermionic waves, unlike bosons, cannot suffer from superradiant scattering amplification [58, 59, 60, 61, 62, 50]. Perhaps less familiar is the fact that fermions, like bosons [53, 54, 55, 56], can also violate the 2-dimensional stability bound of the near-horizon geometry of near-extremal black holes, if the fermion charge is high enough (for fermions the stability bound is lower than the BF bound). In particular, this can happen in a Reissner-Nordström (RN) black hole in an asymptotically planar AdS background [55, 63, 64, 65, 66, 67, 68, 69] (see in particular section V of [55] and sections 4.4 and 4.5 of [68]) or in an asymptotically global AdS geometry (section 2.4 below). This suggests that such a RN black hole *might* be unstable to condensation of a fermionic cloud around the horizon. Moreover, *if* the features from the bosonic field extend to the fermion case, then this might be a linear instability.

Motivated by these considerations, in the present chapter we will search for *linear* mode instabilities of Dirac fields in a global AdS_4 RN black hole. We will not find any such instabilities. The absence of linear mode instabilities in the global AdS RN background is consistent with the fact that they are also not present in the planar AdS limit, $r_+/L \to \infty$, as found previously in [55, 63, 64, 65, 66, 67, 68, 69]. In view of these 'no-go' findings, in the conclusion remarks of section 2.5, we will argue that the absence of *linear* mode instabilities in the AdS RN background still leaves room (but not necessarily) for the following possibility: for a large number of fermions and in the semiclassical limit, the violation of the AdS_2 BF bound might signal a *non*-linear instability of the system.

In this chapter, we also take the opportunity to improve the understanding of the near-horizon condensation instability of scalar fields. In particular, following a similar analysis for scalar fields in a Minkowsky cavity [40], we will explicitly show that the BF bound criterion for instability in AdS-RN is *quantitatively sharp* (section 2.2.4). This scalar field analysis will fit smoothly in our presentation since reviewing its details will also allow to make a direct comparison with the Dirac field case and pinpoint major differences between them. We further take the opportunity to show that the unstable modes belong to a family of *near-extremal modes* that are connected to the *normal modes* of AdS when the dimensionless horizon radius shrinks to zero. (This is an interesting observation because in *de Sitter* black holes the 'normal mode de Sitter family' is *distinct* from the 'near-extremal family of modes', where we are using the nomenclature of [70, 71, 72]).

Necessarily, we will also clarify some misleading analyses and interpretations that were presented in previous literature. Namely, the authors of [73][1] missed that the vanishing energy flux boundary conditions at the asymptotic boundary of AdS – that they propose to be "novel" – are nothing else but the AdS/CFT correspondence no-source boundary

[1] See also Ref. [74], which appeared after the present chapter was submitted to the ArXiv.

conditions – often denoted as '*standard*' or, if allowed, '*alternative*' quantizations (discussed, for $s = 1/2$, originally in [75, 76, 77, 78, 69] and specially in [28, 27, 51, 52, 79]; this discussion applies both to bosonic and fermionic fields). The *vanishing energy flux* boundary condition is fundamental to guarantee that the energy is *conserved*. If and only if this is the case, the Schrödinger operator that describes the (bosonic or fermionic) wave equation in AdS is Hermitian[2]. It follows that *finiteness of energy* then boils down to simply require that the wavefunctions of the system have finite norm in the usual quantum mechanical sense, i.e. that the solutions are *normalizable* (square integrable). In these conditions, the Schrödinger operator of the system is self-adjoint (the associated matrix is Hermitian) and we have a well-posed initial value problem (after imposing regularity at the inner boundary). That is, the dynamical evolution of the system is deterministic.

If the energy is positive, the evolution is *stable* (this happens if we are above the stability mass bound [52, 79]); on the other hand, if the energy is negative we should have a dynamical evolution that develops an *instability* (this is the case if the mass of the perturbation is below the stability mass bound [52, 79]).[3] The aforementioned homogeneous Dirichlet (standard) or Neumann (alternative) AdS/CFT boundary conditions are special in the sense that, by construction, they yield zero energy flux normalizable modes that preserve the conformal symmetry group of AdS (and thus do not deform the boundary conformal field theory). The zero-flux boundary conditions of [73, 74] are nothing but these single-trace homogeneous boundary conditions [69, 75, 76, 77, 78, 79, 80, 81]. Besides these, the AdS/CFT correspondence literature identifies, for a certain range of the boson/fermion masses, other *normalizable* modes (finite conserved energy modes and thus with vanishing energy flux). This is the case of the inhomogeneous Dirichlet and Neumann boundary conditions but also of the often denoted mixed, Robin or multi-trace boundary conditions (see [52] for bosons and [79, 80, 81] for fermions). These zero-flux boundary conditions break the AdS conformal symmetry while still preserving its Poincaré symmetry subgroup.

The plan of this chapter is the following. In sections 2.2.1-2.2.2, we will review the Dirac equation in a AdS-RN background and we will do the necessary field redefinitions in the (physical) Dirac field that allow it to separate and even decouple. Then, in section 2.2.3 we will analyse in detail the AdS/CFT standard and alternative quantizations of a Dirac field. In particular, we will check that the requirement that the source vanishes implies that the energy flux at the conformal boundary also vanishes. In this sense, these boundary conditions can also be denoted as *reflective* boundary conditions. For a massive Dirac field these no-source boundary conditions translate into homogeneous Dirichlet or Neu-

[2] Without the zero flux or energy conservation condition, the Schröedinger operator is only symmetric [52].

[3] The stability mass bound for scalars is the BF bound $m^2 = m^2_{BF}$ [28, 27, 51, 52], while for Dirac fields it is $m^2 = 0$ [79]; see discussion of (2.2.22).

mann boundary conditions in the auxiliary decoupled Dirac radial fields. However, for a massless fermion (a Weyl field) these no-source boundary conditions – which are homogeneous Dirichlet or Neumann conditions on appropriate projections of the original *physical* Dirac field – translate into mixed (Robin) boundary conditions for the *auxiliary* decoupled Dirac radial fields. The misleading focus on the boundary condition for the *auxiliary* fields (and associated consequences) occurs recurrently. It is the case in [73] and it is similar to the one taken on the boundary conditions of the Regge-Wheeler–Zerilli master fields (aka Kodama-Ishibashi fields) in the case of *gravitational* perturbations of AdS black holes (as discussed in [82, 83]).[4] In this context it is also important to clarify that in [85, 86] massless Dirac quasinormal modes in Schwarschild-AdS were computed imposing Dirichlet boundary conditions in the auxiliary decoupled fields. These boundary conditions do not have vanishing energy flux at the asymptotic boundary, the energy of the system is thus not conserved, and it is not known what deformation they produce. Finally, in section 2.4.4 we will describe our strategy to search (unsuccessfully) for linear mode instabilities (eventually sourced by the 2-dimensional stability bound violation) of Dirac fields in global AdS_4-RN black holes. We consider both the standard and alternative quantizations and we will highlight the differences between the scalar and fermion systems. For a reader interested in a future detailed analysis of the frequency spectra, we also compute (analytically) the normal modes of massive and massless Dirac field in global AdS.

2.2 Global AdS Reissner-Nordström black hole and the Dirac equation

2.2.1 AdS-RN black holes and an orthogonal *vierbein*

The gravitational $g_{\mu\nu}$ and Maxwell A_μ fields of the AdS-RN BH are described by[5]

$$\begin{aligned} &\mathrm{d}s^2 = -f\mathrm{d}t^2 + f^{-1}dr^2 + r^2\,\mathrm{d}\Omega_2^2, \qquad f = \frac{r^2}{L^2} + 1 - \frac{2M}{r} + \frac{Q^2}{2r^2}; \\ &A_\mu\,dx^\mu = A(r)\,\mathrm{d}t, \qquad A(r) = -\frac{Q}{r} + C; \end{aligned} \tag{2.2.1}$$

where $\mathrm{d}\Omega_2^2$ is the line element of a unit radius 2-sphere, M and Q are the mass and charge parameters. We will find convenient to replace M and Q by the event horizon radius r_+ (where $f(r_+) = 0$) and chemical potential μ. The relation between these two pairs of

[4] For a discussion that AdS/CFT no-source boundary conditions for bosonic fields yield ('reflective') solutions with vanishing energy (and momentum) flux at the conformal AdS boundary see Appendix A of [84].

[5] This is a solution of the Lagrangian $\mathcal{L} = \sqrt{-g}\left(R - 2\Lambda - \frac{1}{2}F^2\right)$ with $F = dA$. Note that if we rescale $A \to \kappa A$ then the charge q of the perturbation field (to be discussed in later sections) rescales as $q \to \kappa^{-1}q$ so that qA, and thus the gauge covariant derivative, remain invariant.

parameters is

$$M = r_+\left(1+\frac{r_+^2}{L^2}\right)+\frac{1}{2}\,r_+\mu^2\,,\qquad Q=\mu\,r_+\,. \tag{2.2.2}$$

In (2.2.1), C is an arbitrary integration constant and fixing it amounts to choosing a particular gauge. One common gauge choice is $C=0$ where one has $A|_\infty=0$ and the chemical potential is $\mu=-A|_{r_+}=Q/r_+$ (we will typically use this one when presenting our results). Another gauge that is also commonly used is $C=\mu$ whereby $A|_{r_+}=0$ and $A|_\infty=\mu$.

The temperature of this black hole is

$$T_H=\frac{1}{8\pi}\frac{1}{r_+}\left(2+\frac{6r_+^2}{L^2}-\mu^2\right). \tag{2.2.3}$$

Thus, AdS-RN black holes exist for $\mu\leq\mu_{\rm ext}$ where $\mu=\mu_{\rm ext}$ with

$$\mu_{\rm ext}=\sqrt{2}\sqrt{1+3\,\frac{r_+^2}{L^2}} \tag{2.2.4}$$

describes the extremal AdS-RN black hole with zero temperature.

Later, we will consider the Dirac equation coupled to the curved spacetime (2.2.1) [10, 11, 12, 13]. For that, it will be useful to introduce the tetrad vector basis (*vierbein*) $e^{(a)}=e^{(a)}{}_\mu{\rm d}x^\mu$, with non-coordinate curved bracket Latin indices $(a)=(0),\cdots,(3)$:

$$e^{(0)}=f^{1/2}{\rm d}t\,,\quad e^{(1)}=f^{-1/2}{\rm d}r\,,\quad e^{(i)}=r\hat{e}^{(i)}\,, \tag{2.2.5}$$

where $\hat{e}^{(i)}$ is the tetrad on the S^2 manifold. This is an orthonormal basis since $g^{\mu\nu}e^{(a)}{}_\mu e^{(b)}{}_\nu=\eta^{(a)(b)}$ where $\eta^{(a)(b)}={\rm diag}(-1,1,1,1)$ is the Minkowski metric. The tetrad dual basis $e_{(a)}=e_{(a)}{}^\mu\partial_\mu$ becomes $e_{(a)}=\eta_{(a)(b)}e^{(b)}$. Latin (Greek) letters will always be used for tetrad (coordinate basis) indices.

The components of any tensor in the coordinate basis $\{{\rm d}x^\mu\}$ can be obtained from the components on the tetrad basis using the projectors $e^{(a)}$ and $e_{(a)}$. For example, $T_\mu{}^\nu=g^{\nu\sigma}e^{(b)}{}_\sigma T_{\mu(b)}=g^{\nu\sigma}e^{(a)}{}_\mu e^{(b)}{}_\sigma T_{(a)(b)}$. This example also illustrates that we can have mixed-index tensors with mixed components in the coordinate and tetrad bases.

The spin connection of non-coordinate based differential geometry can be introduced in terms of the affine connection $\Gamma^\beta{}_{\mu\nu}$ of coordinate based differential geometry as $\gamma_{(c)(a)(b)}=e_{(c)}{}^\mu\left(\partial_\nu e_{(a)\mu}-\Gamma^\beta{}_{\mu\nu}e_{(a)\beta}\right)e_{(b)}{}^\nu$. Equivalently, one can define the spin connection as

$$\omega_{\mu(a)(b)}=e^{(c)}{}_\mu\gamma_{(a)(b)(c)}=\frac{1}{2}e^{(c)}{}_\mu\left(\lambda_{(a)(b)(c)}+\lambda_{(c)(a)(b)}-\lambda_{(b)(c)(a)}\right) \tag{2.2.6}$$

with $\lambda_{(a)(b)(c)} = e_{(a)}^{\ \ \alpha}\left(\partial_\beta e_{(b)\alpha} - \partial_\alpha e_{(b)\beta}\right) e_{(c)}^{\ \ \beta}$, which allows to compute the spin connections without the use of the affine connection.

For a multi-index tensor with tetrad and coordinate indices the mixed-index covariant derivative is defined as

$$\nabla_\alpha T^{(a)\nu}_{\ \ \ \ (b)\mu} = \partial_\alpha T^{(a)\nu}_{\ \ \ \ (b)\mu} - \Gamma^\beta_{\ \mu\alpha} T^{(a)\nu}_{\ \ \ \ (b)\beta} + \Gamma^\nu_{\ \beta\alpha} T^{(a)\beta}_{\ \ \ \ (b)\mu} + \omega_\alpha^{\ (a)}_{\ (c)} T^{(c)\nu}_{\ \ \ \ (b)\mu} - \omega_\alpha^{\ (c)}_{\ (b)} T^{(a)\nu}_{\ \ \ \ (c)\mu}. \tag{2.2.7}$$

Onwards, we take the affine connection $\Gamma^\nu_{\ \mu\alpha}$ to be given by the Christoffel symbols that covariantly conserve the metric, $\nabla_\alpha g_{\mu\nu} = 0$. It follows that the spin coefficients $\omega_{\mu(a)(b)}$ defined in (2.2.6) are such that the *vierbein* is also covariantly conserved, $\nabla_\alpha e^{(a)}_{\ \ \mu} = 0$. That the latter implies the first conservation follows from $g_{\mu\nu} = \eta_{(a)(b)} e^{(a)}_{\ \ \mu} e^{(b)}_{\ \ \nu}$. Further note that the spin coefficient is anti-symmetric in the tetrad pair of indices, $\omega_{\mu(a)(b)} = -\omega_{\mu(b)(a)}$.

To discuss the Dirac equation one necessarily needs to introduce the (coordinate independent) Dirac gamma matrices $\gamma^{(a)}$. Let σ^i be the Pauli matrices and I_n the $n \times n$ identity matrix. We choose to work with the Weyl (chiral) spinor representation of the 4-dimensional Clifford algebra:

$$\gamma^{(0)} = \begin{pmatrix} 0 & iI_2 \\ iI_2 & 0 \end{pmatrix}, \quad \gamma^{(1)} = \begin{pmatrix} 0 & i\sigma^3 \\ -i\sigma^3 & 0 \end{pmatrix}, \quad \gamma^{(2)} = \begin{pmatrix} 0 & i\sigma^1 \\ -i\sigma^1 & 0 \end{pmatrix}, \quad \gamma^{(3)} = \begin{pmatrix} 0 & i\sigma^2 \\ -i\sigma^2 & 0 \end{pmatrix} \tag{2.2.8}$$

which indeed obeys the anti-commutation relations of the Clifford algebra:

$$\left\{\gamma^{(a)}, \gamma^{(b)}\right\} = 2\eta^{(a)(b)} I_4\,, \tag{2.2.9}$$

where $\{A, B\} = AB + BA$ is the usual anti-commutator, as well as the relations $(\gamma^{(0)})^2 = -I_4$ and $(\gamma^{(i)})^2 = I_4$ for $i = 1, 2, 3$.

Let us also introduce the pseudoscalar

$$\gamma^{(5)} = i\gamma^{(0)}\gamma^{(1)}\gamma^{(2)}\gamma^{(3)} = \begin{pmatrix} -I_2 & 0 \\ 0 & I_2 \end{pmatrix}, \tag{2.2.10}$$

which obeys the relations $(\gamma^{(5)})^2 = 1$ and $\{\gamma^{(5)}, \gamma^{(a)}\} = 0$.[6]

The components of the (coordinate dependent) Dirac gamma matrices in the coordinate

[6] $\gamma^{(5)}$ satisfies the 5-dimensional Clifford algebra $\left\{\tilde{\gamma}^{(A)}, \tilde{\gamma}^{(B)}\right\} = 2\eta^{(A)(B)} I_5$ if we define $\tilde{\gamma}^{(A)} = (\gamma^{(a)}, \gamma^{(5)})$, which justifies its label (from an Euclidean perspective).

basis can be obtained from the tetrad basis components (2.2.8) using

$$\gamma^{\mu} = \gamma^{(a)} e_{(a)}{}^{\mu}. \tag{2.2.11}$$

and they obey the covariant Clifford algebra $\{\gamma^{\mu}, \gamma^{\nu}\} = 2g^{\mu\nu} I_4$.

2.2.2 Dirac equation in the AdS-RN background

In the introduction 1.2 we derived the equations that scalar fields and spinors must satisfy in a curved background. Now we show explicitly how one separates variables in the RN AdS background. To find solutions of (1.2.8) it is advantageous to write the Dirac 4-spinor ψ in terms of the left-handed and right-handed 2-spinors Ψ_- and Ψ_+, respectively, as

$$\psi = \begin{pmatrix} \Psi_- \\ \Psi_+ \end{pmatrix} . \tag{2.2.12}$$

The chiral 2-spinors $\Psi_\pm$ emerge naturally when we note that the pseudoscalar $\gamma^{(5)}$ defined in (2.2.10) obeys $(\gamma^{(5)})^2 = 1$. Therefore we can introduce the Lorentz invariant projection operators $P_\pm$ that project the Dirac 4-spinor ψ into the chiral spinors:

$$P_\pm \psi = \Psi_\pm \,, \qquad \text{with} \qquad P_\pm = \frac{1}{2}\left(I_4 \pm \gamma^{(5)}\right) \tag{2.2.13}$$

and such that $P_\pm^2 = P_\pm$ and $P_+ P_- = 0$.[7] Moreover, the task of finding solutions of the Dirac equation in the AdS-RN black hole gets considerably simplified by the fact that under the separation *anstaz* [87, 88]:

$$\begin{aligned} \Psi_+(t,r,\theta,\phi) &= e^{-i\omega t} e^{i m_\phi \phi} (-gf)^{-\frac{1}{4}} \begin{pmatrix} R_1(r) S_1(\theta) \\ R_2(r) S_2(\theta) \end{pmatrix}, \\ \Psi_-(t,r,\theta,\phi) &= e^{-i\omega t} e^{i m_\phi \phi} (-gf)^{-\frac{1}{4}} \begin{pmatrix} -R_2(r) S_1(\theta) \\ -R_1(r) S_2(\theta) \end{pmatrix}, \end{aligned} \tag{2.2.14}$$

the Dirac equations (1.2.8) reduce to a set of equations where the radial and angular functions of the fermion field are decoupled. This separation *ansatz* exploits the fact that ∂_t and ∂_ϕ are Killing vectors of the background AdS-RN solution. This allows to do a Fourier decomposition in these directions which introduces the frequency ω and azimuthal angular momentum m_ϕ of the fermionic wavefunction.[8]

[7] In 4 spacetime dimensions and in the chiral representation (2.2.8) in which we work, $\Psi_\pm$ are nothing but the Weyl 2-spinors which transform in the same way under Lorentz rotations but oppositely under Lorentz boosts, and obey the Weyl equations if the fermion mass vanishes.

[8] This frequency ω is measured in the gauge $A|_\infty = 0$ – see (2.2.1) – and we will work preferentially on this gauge unless otherwise stated. In particular, all our numerical results will be using it. Note that in the alternative gauge $A|_{r_+} = 0$ (also often used) the associated frequency is $\tilde{\omega} = \omega - q\mu$.

Concretely, the radial functions $R_1(r)$, $R_2(r)$ obey the coupled system of first order ODEs

$$\begin{aligned} r\sqrt{f(r)}\Big(\frac{d}{dr} - i\,\frac{\omega + qA(r)}{f(r)}\Big)R_1(r) &= (\lambda + i\,m\,r)R_2(r), \\ r\sqrt{f(r)}\Big(\frac{d}{dr} + i\,\frac{\omega + qA(r)}{f(r)}\Big)R_2(r) &= (\lambda - i\,m\,r)R_1(r), \end{aligned} \tag{2.2.15}$$

where λ is a separation constant, while the angular functions $S_1(\theta)$, $S_2(\theta)$ satisfy the coupled system of first order ODEs

$$\begin{aligned} \left(\frac{d}{d\theta} + \frac{m_\phi}{\sin\theta} + \frac{\cot\theta}{2}\right) S_2(\theta) &= -\lambda S_1(\theta), \\ \left(\frac{d}{d\theta} - \frac{m_\phi}{\sin\theta} + \frac{\cot\theta}{2}\right) S_1(\theta) &= \lambda S_2(\theta). \end{aligned} \tag{2.2.16}$$

Furthermore, the coupled pair of first order radial equations (2.2.15) can be decoupled in a pair of second order ODEs, one for $R_1(r)$ and the other for $R_2(r)$. For that we solve the first (second) equation in (2.2.15) w.r.t. R_2 (R_1) and replace it in the second (first) equation. We end up with two decoupled second order ODEs for R_1 and R_2,

$$\begin{aligned} r\sqrt{f(r)}\frac{d}{dr}\Big(r\sqrt{f(r)}\frac{d}{dr}R_1(r)\Big) + H_1(r)\frac{d}{dr}R_1(r) + H_2(r)R_1(r) &= 0, \\ r\sqrt{f(r)}\frac{d}{dr}\Big(r\sqrt{f(r)}\frac{d}{dr}R_2(r)\Big) + H_1^*(r)\frac{d}{dr}R_2(r) + H_2^*(r)R_2(r) &= 0, \end{aligned} \tag{2.2.17}$$

where * denotes complex conjugation and we have defined

$$\begin{aligned} H_1(r) &= -\frac{m\,r^2 f(r)}{m\,r - i\lambda}, \\ H_2(r) &= -ir\sqrt{f(r)}\frac{d}{dr}K(r) + K(r)\left(K(r) - i\frac{H_1(r)}{r\sqrt{f(r)}}\right) - \lambda^2 - m^2r^2, \\ K &= \frac{r}{\sqrt{f(r)}}\Big(\omega + qA(r)\Big). \end{aligned} \tag{2.2.18}$$

Of course, we are only interested in solutions of (2.2.17) that also solve the original first order system (2.2.15). The requirement that (2.2.15) is solved imposes extra constraints on solutions of (2.2.17). This is best illustrated if we consider the Taylor expansion about the boundaries of the integration domain: the ODE pair (2.2.17) has four integration constants about each boundary but only two of them are independent when we further require that the solution solves the two first order ODEs (2.2.15); see discussion of (2.2.20) below.

Similarly, the coupled pair of first order ODEs for $S_{1,2}(\theta)$ can be written as a decoupled set of two second order ODEs for $S_1(\theta)$ and $S_2(\theta)$. They are hypergeometric equations and $S_{1,2}(\theta)$ are the spin-$\frac{1}{2}$ weighted spherical harmonics. Regularity at $\theta = 0$ and $\theta =$

π quantizes the angular separation constant as (ℓ is a harmonic number related to the number of zeros of the wavefunction)

$$\lambda = \ell + \frac{1}{2}, \qquad \ell = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \cdots \tag{2.2.19}$$

with the azimuthal number being constrained as $m_\phi \leq \ell$.

Unfortunately, the radial ODEs cannot be solved analytically[9]. We can however do a Frobenius analysis about the asymptotic boundary $r \to \infty$ to find the asymptotic behaviours of $R_1(r)$ and $R_2(r)$. One finds that (for $m \neq 0, \frac{1}{2}$)[10]

$$\begin{aligned}(-gf)^{-\frac{1}{4}} R_1\big|_{r\to\infty} &\sim r^{-\Delta_-}\left(\alpha_1 + \cdots\right) + r^{-\Delta_+}\left(\beta_1 + \cdots\right), \\ (-gf)^{-\frac{1}{4}} R_2\big|_{r\to\infty} &\sim r^{-\Delta_-}\left(-i\,\alpha_1 + \cdots\right) + r^{-\Delta_+}\left(i\,\beta_1 + \cdots\right),\end{aligned} \tag{2.2.20}$$

where we used $(-gf)^{-\frac{1}{4}}|_{r\to\infty} \sim L^{1/2} r^{-3/2}$ and, anticipating the AdS/CFT discussion below, we have introduced the conformal dimensions

$$\Delta_\pm = \frac{3}{2} \pm \sqrt{m^2 L^2}\,. \tag{2.2.21}$$

As expected for a coupled system (2.2.15) of two first order ODEs, there are two independent arbitrary constants (α_1, β_1) in the asymptotic decay (2.2.20), that is to say, the decays of R_2 are fixed by the equations of motion as a function of (α_1, β_1). The dots in (2.2.20) represent subleading terms that depend only on α_1 (in the Δ_- contribution) or β_1 (in the Δ_+ terms).

Before proceeding, one unavoidably needs to discuss the range of Dirac fermion masses that allow for normalizable solutions, i.e with conserved finite energy. We also have to distinguish the positive energy solutions (which are stable) from those negative energy states (which should trigger an instability). It was proven in section II/Appendix B of [79] (see also [80, 81, 52]) that the fermionic bound for stability (in *any* dimension) is given by

$$m^2 \geq 0 \qquad \text{(Dirac stability bound condition)}\,, \tag{2.2.22}$$

with the lower bound being the solution for which $\Delta_+ = \Delta_-$ in (2.2.21).[11] To under-

[9] For global AdS, i.e. $M = 0 = Q$ these ODEs are hypergeometric equations and can be solved analytically: see section 2.4.2.

[10] For $m = 1/2$ one of the two independent solutions decays asymptotically as a power law in r and the other as a power law multiplied by a $\log r$. For this reason (since a similar logarithmic solution appears in the scalar field case when $m^2 = m^2_{BF}$), this case is often called the BF solution of the Dirac system. We do not discuss further this special case (see [79, 63] for more details). It is however important to emphasize that for the scalar field, $m^2 = m^2_{BF}$ corresponds to $\Delta_+ = \Delta_-$ and is thus also the bound for stability while in the Dirac case, the mass stability bound is (2.2.22) not the BF mass $m = 1/2$.

[11] So, for $m^2 \geq 0$, $\Delta_\pm$ are real; otherwise they are complex numbers. Note that for a scalar field the configuration $\Delta_+ = \Delta_-$ corresponds to the BF bound where one of the independent solutions is logarithmic. However, for the Dirac field, the state $\Delta_+ = \Delta_-$ is *not* the BF logarithmic solution (which occurs instead

stand this bound it is useful to rewrite the radial Dirac equation (2.2.17) as a Schrödinger equation [52, 79]. Without further conditions, the associated Schrödinger operator is not self-adjoint (hermitian). It becomes self-adjoint if and only if we impose as boundary condition that the energy-momentum flux at the asymptotic AdS boundary vanishes. That is to say, it becomes Hermitian if and only if the energy is conserved. In these conditions looking for (conserved) finite energy solutions boils down to look for normalizable states in the standard quantum mechanical sense. That is to say, normalizable solutions are those that are square integrable.

For $m^2 < 0$ there are normalizable solutions but they have negative energy. In a mathematical language, if $m^2 < 0$, the Schrödinger operator of the Dirac equation is unbounded below and thus it does not allow for a positive self-adjoint extension [52, 79]. Alike in any other negative energy Schrödinger states, this signals the existence of an instability. We will explore further this in section 2.4.1.

On the other hand, if the mass is real, i.e. if it satisfies the bound (2.2.22), there are stable normalizable Dirac fermion solutions that are selected by a choice of boundary conditions. We will discuss in detail this issue of the boundary conditions in the next section. The upshot is that if $mL \geq 1/2$ there is an unique complete set of normalizable modes (and the non-normalizable modes must be fixed by boundary conditions; *e.g.* no-source/homogeneous boundary conditions that eliminate them) [79]. On the other hand, for $0 \leq mL < 1/2$ there is a non-unique set of normalizable modes and thus a wider band of boundary conditions that yield normalizable solutions (e.g. the no-source/homogeneous Dirichlet or Neumann boundary conditions that we will use later but also more general multi-trace boundary conditions) [79]. Further note that if we take $m \to -m$, we simply trade the role of the $\Delta_\pm$ contributions while preserving condition (2.2.22). Therefore onwards we assume, without any loss of generality, $m \geq 0$ in our discussion.

For our purposes, but without loss of generality, we will be particularly interested in the lower bound case of (2.2.22). For this $m = 0$ case and choosing the gauge $A|_\infty = 0$, a Frobenius analysis of the first order equations of motion about the asymptotic boundary

for $m = 1/2$). If follows that for the scalar field case the BF bound coincides with the bound for stability, $m^2 \geq m^2_{BF}$, but *not* in the Dirac case. Moreover, in the scalar case, there is a 1-parameter family of boundary conditions that yield stable normalizable solutions for $m^2_{BF} \leq m^2 < m^2_{BF} + 1/L^2$ and a unique boundary condition that generates stable normalizable solutions for $m^2 \geq m^2_{BF} + 1/L^2$ [28, 27, 51, 52]. However, in the Dirac case, normalizable stable states exist for: 1) a 1-parameter choice of boundary conditions for $0 \leq m^2 < m^2_{BF}$ (with $m^2_{BF} = 1/4$), and 2) a unique boundary condition for $m^2 \geq m^2_{BF}$ [80]. Further note that, unlike in the scalar case, the Dirac stability mass bound is independent of the dimension of the spacetime.

yields[12]

$$
\begin{aligned}
(-gf)^{-\frac{1}{4}} R_1\big|_{r\to\infty} &\sim r^{-\frac{3}{2}} \left(\alpha_1 + \beta_1 \frac{L}{r} + \mathcal{O}(r^{-2})\right), \\
(-gf)^{-\frac{1}{4}} R_2\big|_{r\to\infty} &\sim r^{-\frac{3}{2}} \left(-\frac{\beta_1 + i\alpha_1 \omega L}{\lambda} + \frac{\alpha_1\left(\omega^2 L^2 - \lambda^2\right) - i\beta_1 \omega L}{\lambda}\frac{L}{r} + \mathcal{O}(r^{-2})\right),
\end{aligned}
\tag{2.2.23}
$$

i.e. we can take the two independent integration constants associated to the coupled pair of first order ODEs to be α_1 and β_1 and the equations of motion then fix the decay of R_2 as a function of α_1 and β_1.

2.2.3 Boundary conditions for the Dirac spinor in AdS-RN

To find the solution of the Dirac spinor field ψ and its Dirac adjoint $\bar{\psi}$ in the AdS-RN background we have to solve a system of two equations that are first order, namely (2.2.15), subject to boundary conditions imposed at the event horizon $r = r_+$ and at the asymptotic boundary $r \to \infty$. Before imposing boundary conditions, such a system of two first order differential equations necessarily has two independent constants at the horizon boundary and another two independent constants at the asymptotic boundary (namely, α_1 and β_1 in (2.2.20)), which can be identified doing a Frobenius analysis at these two boundaries. To have a well posed formulation of the elliptic problem one should impose two boundary conditions that fix two of the independent constants and solve the equations of motion to find the other two. We certainly want the Dirac solutions to be regular at the event horizon: this boundary condition fixes one of the constants [13]. One should then fix one of the asymptotic constants α_1 or β_1 (or a relation between them) with an appropriate boundary condition [79, 80, 81]. But we certainly cannot fix both asymptotic independent constants: once the first is fixed, the second one must be found by solving the equations of motion in the bulk subject to the two aforementioned boundary conditions. This poses the question: how do we choose a boundary condition at the asymptotic boundary that is physically relevant? We should choose one that conserves the energy and thus yields a self-adoint Schrödinger operator for the system that ensures that we have a well-posed hyperbolic evolution if we let the perturbed system evolve in time. Next, we will review how two boundary conditions with these properties

[12] For a Dirac field (or scalar field) with phase φ, $\psi = |\psi| e^{i\varphi}$, $U(1)$ gauge transformations with gauge parameter χ leave the action and equations of motion invariant and transform the Dirac (scalar) and Maxwell fields as $\varphi \to \tilde{\varphi} = \varphi + q\,\chi$, $A_t \to \tilde{A}_t = A_t + \nabla_t \chi$. Thus, if in the gauge $A|_\infty = 0$ (*i.e.* $C = 0$) we denote the frequency of the Dirac (scalar) field by ω then a transformation with gauge parameter $\chi = \mu\, t$ into the gauge $\tilde{A}|_\infty = \mu$ ($C = \mu$) changes the frequency into $\tilde{\omega} = \omega - q\mu$. Thus, if we had chosen the gauge $\tilde{A}|_\infty = \mu$, then we would have to make the replacement $\omega \to \tilde{\omega} + q\mu$ in (2.2.23) (and later in the boundary conditions (2.2.36)-(2.2.37) and (2.4.22)). Further note that in (2.2.15)-(2.2.17) we are leaving the gauge choice arbitrary because we do not fix $A(r)$ introduced in (2.2.1).

[13] Alternatively, since we have a ODE system, we could use two boundary conditions to fix the two asymptotic independent constants and solving the equations of motion would yield the behaviour of the Dirac fields at the event horizon. However, this is not a good strategy because in general these solutions would not be smooth at the event horizon.

can be identified. They single out in the AdS-CFT context because they are single-trace (no-source) boundary conditions that preserve the conformal symmetry group of AdS (and thus do not deform the boundary conformal field theory) [75, 76, 77, 78, 79, 80, 69, 81] .

Dirac spinor fields ψ are intrinsically quantum fields. The dynamics of such fields can be naturally described by a path integral formulation whereby one sums over all possible field configurations in configuration space to get the transition amplitude between two states. In particular, the partition function Z (*i.e.* the generating functional of correlation functions between operators) can also be naturally computed using the path integral formulation. Schematically one has,

$$Z = \int [D\psi][D\bar{\psi}] e^{\frac{i}{\hbar}\mathcal{S}[\psi,\bar{\psi}]}, \tag{2.2.24}$$

where $[D\psi][D\bar{\psi}]$ represents the integration measure, $\hbar$ is Planck's constant and $\mathcal{S}_D[\psi,\bar{\psi}]$ is the action (1.2.7) of the Dirac field.

In the classical limit, $\hbar \to 0$, the path integral reduces simply to $Z \sim e^{\frac{i}{\hbar}\mathcal{S}_{\rm cl}[\psi,\bar{\psi}]}$, where $\mathcal{S}_{\rm cl}[\psi,\bar{\psi}]$ is the action evaluated on a solution of the classical equations of motion, that follow from the variation $\delta\mathcal{S} = 0$ subject to the boundary conditions. As emphasised in [75, 76, 77, 78, 79, 80, 69, 81] this statement that the action must be *stationary* when evaluated on a classical solution severely constrains the type of boundary conditions we can impose on the field ψ. Indeed, if $\delta\mathcal{S} = 0$ then it is not necessarily true that $\delta(\mathcal{S}+\mathcal{B}) = 0$ where $\mathcal{B}$ is the boundary term describing the desired boundary conditions (*i.e.* a total derivative term that does not change the equations of motion). That is to say, the physical choice we make for the boundary conditions must be such that $\delta\mathcal{B} = 0$. In particular, in the context of the AdS/CFT correspondence, this condition fixes the form of the boundary term that must be added to the standard Dirac action (1.2.7) to have stationary solutions. Vice-versa, this boundary term fixes the boundary field theory.

To determine the boundary term $\mathcal{B}$, one first notes that the "radial" Dirac gamma matrix $\gamma^{(1)}$ defined in (2.2.8) satisfies $(\gamma^{(1)})^2 = I_4$ and $\gamma^{(1)} = \gamma^{(1)\,\dagger}$. It follows that we can decompose the Dirac spinor as

$$\psi = \psi_+ + \psi_- \,, \qquad \bar{\psi} = \bar{\psi}_+ + \bar{\psi}_- \,, \tag{2.2.25}$$

where $\psi_\pm$ ($\bar{\psi}_\pm$) are 4-eigenspinors of $\gamma^{(1)}$ with eigenvalue ± 1 (∓ 1).[14] Using this property, including the associated properties listed in footnote 14, one finds that the terms in the

[14] In more detail, $\gamma^{(1)}\psi_\pm = \pm\psi_\pm$ and $\gamma^{(1)}\bar{\psi}_\pm = \mp\bar{\psi}_\pm$ and thus $\psi_\pm = \frac{1}{2}\left(I_4 \pm \gamma^{(1)}\right)\psi$ and $\bar{\psi}_\pm = \frac{1}{2}\bar{\psi}\left(I_4 \mp \gamma^{(1)}\right)$. A few properties follow that are useful. For example, $\bar{\psi}\gamma^{(1)} = -\bar{\psi}_+ + \bar{\psi}_-$, $\bar{\psi}_\pm\psi_\pm = \frac{1}{4}\bar{\psi}\left(I_4 - (\gamma^{(1)})^2\right)\psi = 0$, $\bar{\psi}_\pm\gamma^{(1)}\psi_\pm = \mp\bar{\psi}_\pm\psi_\pm = 0$ and $\bar{\psi}_\pm\gamma^\mu\mathcal{D}_\mu\psi_\pm = \frac{1}{4}\bar{\psi}\left(I_4 - (\gamma^{(1)})^2\right)\gamma^\mu\mathcal{D}_\mu\psi = 0$.

Dirac action (1.2.7) that contain radial derivatives of the spinor are

$$\mathcal{S}_D \supset \mathcal{S}_D\big|_{\partial_r} = i\int_{\mathcal{M}} d^4x\sqrt{-g}f^{1/2}\left(\bar{\psi}_+\partial_r\psi_- - \bar{\psi}_-\partial_r\psi_+\right), \tag{2.2.26}$$

where we used the fact that $\mathcal{D}_r = f^{1/2}\partial_r$. It follows that if we vary the Dirac action (1.2.7) w.r.t. ψ_+ and ψ_- one gets, after integration by parts,

$$\delta\mathcal{S}_D = \text{bulk terms} + \delta\mathcal{S}_{bdry}\,, \tag{2.2.27}$$

where the bulk terms describe a contribution that vanishes when the equations of motion − which are equivalent to (1.2.8) − are satisfied and $\delta\mathcal{S}_{bdry}$ is a boundary term resulting from integrating by parts the radial derivative terms (2.2.26) given by

$$\delta\mathcal{S}_{bdry} = i\int_{\partial\mathcal{M}} d^3x\sqrt{-g}f^{1/2}\left(\bar{\psi}_+\delta\psi_- - \bar{\psi}_-\delta\psi_+\right). \tag{2.2.28}$$

As discussed above, to have a well-posed boundary value problem, after requiring that the solution is regular at the event horizon we no longer have the freedom to fix both ψ_+ and ψ_- at the asymptotic boundary (these are the two independent asymptotic constants of our pair of first order ODEs). Instead, we can either fix ψ_+ at the asymptotic boundary (in which case $\delta\psi_+ = 0$) or fix the asymptotic value of ψ_- (in which case $\delta\psi_- = 0$ at the boundary).

Suppose we want to fix ψ_+ at the asymptotic boundary (a similar analysis would apply if we wanted to fix ψ_-). In order to have a well-defined variational problem one should add a boundary term that cancels the contribution $\bar{\psi}_+\delta\psi_-$ in (2.2.28). Adding the boundary term [75, 76, 77, 78, 69]

$$\mathcal{S}_\partial = -i\int_{\partial\mathcal{M}} d^3x\sqrt{-g}f^{1/2}\bar{\psi}_+\psi_-, \tag{2.2.29}$$

produces the desired effect since the total on-shell action becomes

$$\delta\mathcal{S}_{tot} = \delta\left(\mathcal{S}_D + \mathcal{S}_\partial\right) = -i\int_{\partial\mathcal{M}} d^3x\sqrt{-g}f^{1/2}\left(\bar{\psi}_-\delta\psi_+ + \delta\bar{\psi}_+\psi_-\right), \tag{2.2.30}$$

which indeed vanishes when $\delta\psi_+ = 0$ (and thus $\delta\bar{\psi}_+ = 0$). We can also compute the momentum conjugate to ψ_+ and $\bar{\psi}_+$ by varying $\mathcal{S}_{tot}$ w.r.t. ψ_+ and $\bar{\psi}_+$, respectively, yielding

$$\Pi_+ = \frac{\delta\mathcal{S}_{tot}}{\delta\psi_+} = -i\sqrt{-g}f^{1/2}\bar{\psi}_-\,, \qquad \text{and} \qquad \bar{\Pi}_+ = \frac{\delta\mathcal{S}_{tot}}{\delta\bar{\psi}_+} = -i\sqrt{-g}f^{1/2}\psi_-\,. \tag{2.2.31}$$

In terms of the functions $R_1(r)$, $R_2(r)$ and $S_1(\theta)$, $S_2(\theta)$ introduced in the separation *ansatz*

(2.2.14), the 4-spinors $\psi_\pm$ are given by

$$\psi_+ = \frac{e^{-i\omega t}e^{im_\phi\phi}}{2(-gf)^{\frac{1}{4}}}\begin{pmatrix} i(R_1+iR_2)S_1 \\ -(R_1+iR_2)S_2 \\ (R_1+iR_2)S_1 \\ -i(R_1+iR_2)S_2 \end{pmatrix}, \qquad \psi_- = \frac{e^{-i\omega t}e^{im_\phi\phi}}{2(-gf)^{\frac{1}{4}}}\begin{pmatrix} -i(R_1-iR_2)S_1 \\ -(R_1-iR_2)S_2 \\ (R_1-iR_2)S_1 \\ i(R_1-iR_2)S_2 \end{pmatrix}. \tag{2.2.32}$$

From the asymptotic decays of $R_{1,2}$ in (2.2.20) (valid for $m \neq 0, \frac{1}{2}$) or in (2.2.23) (valid for $m = 0$) one finds that $\psi_\pm$ decay as

$$\begin{cases} \psi_+\big|_{r\to\infty} \sim 2\alpha_1 r^{-\Delta_-} + a(\alpha_1)\, r^{-\Delta_+-1} + \mathcal{O}\left(r^{-\Delta_--2}\right), \\ \psi_-\big|_{r\to\infty} \sim 2\beta_1 r^{-\Delta_+} + b(\beta_1)\, r^{-\Delta_--1} + \mathcal{O}\left(r^{-\Delta_--2}\right), \qquad \text{if } 0 < mL < \frac{1}{2}; \end{cases} \tag{2.2.33}$$

$$\begin{cases} \psi_+\big|_{r\to\infty} \sim r^{-\frac{3}{2}}\left(\frac{\alpha_1(\lambda+\omega L)-i\beta_1}{\lambda} - i(\lambda+\omega L)\frac{\alpha_1(\lambda-\omega L)+i\beta_1}{\lambda}\frac{L}{r} + \mathcal{O}(r^{-2})\right), \\ \psi_-\big|_{r\to\infty} \sim r^{-\frac{3}{2}}\left(\frac{\alpha_1(\lambda-\omega L)+i\beta_1}{\lambda} + i(\lambda-\omega L)\frac{\alpha_1(\lambda+\omega L)-i\beta_1}{\lambda}\frac{L}{r} + \mathcal{O}(r^{-2})\right), \quad \text{if } m = 0; \end{cases} \tag{2.2.34}$$

$$\begin{cases} \psi_+\big|_{r\to\infty} \sim 2\alpha_1 r^{-\Delta_-} + \tilde{a}(\alpha_1)\, r^{-\Delta_--2} + \mathcal{O}\left(r^{-\Delta_+-1}\right), \\ \psi_-\big|_{r\to\infty} \sim \tilde{b}(\alpha_1)\, r^{-\Delta_--1} + 2\beta_1 r^{-\Delta_+} + \mathcal{O}\left(r^{-\Delta_--2}\right), \qquad \text{if } mL > \frac{1}{2}; \end{cases} \tag{2.2.35}$$

where α_1, β_1 are the free constants introduced in (2.2.20) or (2.2.23) and the constants $a(\alpha_1)$, $b(\beta_1)$, $\tilde{a}(\alpha_1)$ and $\tilde{b}(\alpha_1)$ are fixed as functions of α_1 or β_1 (as described by their argument) by the equations of motion (details are irrelevant for our aim). The asymptotic decays of the Dirac adjoints $\bar{\psi}_\pm$ follow straightforwardly from (2.2.33) with the exchange $\alpha_1 \to \bar{\alpha}_1$, $\beta_1 \to \bar{\beta}_1$, etc.

For $mL \geq \frac{1}{2}$ the only normalizable mode (*i.e.* with finite energy) is ψ_+ [51, 79, 80, 55, 63, 66, 81]. In the context of the AdS/CFT correspondence, the leading term of the asymptotic expansion $\lim_{r\to\infty} r^{\Delta_-}\psi_+ = 2\alpha_1$ is then identified with the source of the dual operator $\bar{\mathcal{O}}$ which has mass dimension Δ_+. We have a well-posed boundary value problem if we impose smoothness of ψ_+ at the event horizon and a Dirichlet boundary condition for α_1 at the asymptotic boundary. In particular, if we do not want to deform the boundary field theory we impose the no-source/homogeneous Dirichlet boundary condition: $\alpha_1 = 0$. We have no freedom left to fix asymptotically ψ_- *i.e.* β_-. Instead, β_- and thus $\psi_-|_\infty$ is determined by solving the Dirac equations subject to the above boundary conditions. The expectation value $\langle\bar{\mathcal{O}}\rangle$ of the dual operator is given by the conjugate momentum Π_+ defined in (2.2.31): $\langle\bar{\mathcal{O}}\rangle \propto \lim_{r\to\infty} r^{-\Delta_-}\Pi_+ \propto \bar{\beta}_1$.

On the other hand for $0 \leq mL < \frac{1}{2}$ both modes $\psi_\pm$ are normalizable [51, 79, 80, 55, 63, 66, 81]. Thus we can still impose the *standard* quantization where we identify the $\lim_{r\to\infty} r^{\Delta_-}\psi_+ \equiv \psi_+^{(0)}$ as the source of the dual operator $\bar{\mathcal{O}}$. In particular, the no-source/ho-

mogeneous standard boundary condition for all possible masses:

$$\psi_+^{(0)} = 0 \quad \Leftrightarrow \quad \begin{cases} \alpha_1 = 0, & \text{if } 0 < mL < \frac{1}{2} \text{ (or } mL \geq \frac{1}{2}); \\ \alpha_1(\lambda + \omega L) - i\beta_1 = 0, & \text{if } m = 0; \end{cases} \tag{2.2.36}$$

But, since for this range of masses both modes are normalisable,[15] we can also impose the so-called *alternative* quantization; where we identify the $\lim_{r\to\infty} r^{\Delta_+}\psi_- \equiv \psi_-^{(0)}$ as the source of the dual operator $\mathcal{O}$ with mass dimension Δ_-. In particular, if we do not want to deform the boundary field theory we impose the no-source alternative boundary condition:

$$\psi_-^{(0)} = 0 \quad \Leftrightarrow \quad \begin{cases} \beta_1 = 0, & \text{if } 0 < mL < \frac{1}{2}; \\ \alpha_1(\lambda - \omega L) + i\beta_1 = 0, & \text{if } m = 0; \end{cases} \tag{2.2.37}$$

The two quantizations (2.2.36) and (2.2.37) yield two distinct boundary conformal field theories [51, 79, 81, 55, 63]. For $m = 0$, note that the Dirichlet boundary condition on ψ_+, $\psi_+^{(0)} = 0$, implies the Neumann condition in ψ_- (*i.e.* the next-to-leading order term in the expansion for ψ_- vanishes) and vice-versa. This follows straightforwardly from an inspection of (2.2.34).

We emphasize that the no-source standard and alternative boundary conditions (2.2.36)-(2.2.37) that do not deform the boundary theory imply that the energy flux and fermion particle flux vanish at the asymptotic boundary (this is also the case for more elaborated normalizable AdS/CFT boundary conditions [79, 81]). In this sense we can regard these as 'reflective' boundary conditions. The Dirac action (1.2.7) (and (2.2.30)) is left invariant if we rotate the phase of the Dirac spinor, $\psi \to e^{-i\alpha}\psi$. The Dirac current associated to this symmetry is $j^\mu = \bar{\psi}\gamma^\mu\psi$ and one can check that it is conserved, $\nabla_\mu j^\mu = 0$ after using the first order equations of motion (1.2.8). This is an internal vector symmetry since $\psi_\pm$ transform in the same way under this symmetry. This current gives the charge flux or particle number flux of fermions. The associated conserved charge is $\mathcal{Q} = \int_V dx^3\sqrt{\gamma}j^\mu\xi_\mu = \int_V dx^3\sqrt{\gamma}\psi^\dagger\psi$ where V is the volume of a constant t hypersurface, γ_{ab} is the associated induced metric, and $\xi = \partial_t$ is the Killing vector describing time translations. In particular, $j^r|_{r\to\infty}$ gives the radial flux of particles at the asymptotic spacelike boundary Σ. One can also show that the energy flux across a spacelike boundary is proportional to the Dirac current. The energy flux across the asymptotic boundary

[15] Besides the single-trace standard/alternative boundary conditions, we can also impose multi-trace deformations which are mixed boundary conditions; see, e.g. [28, 27, 51, 52, 79, 80, 81].

$\Phi_{\partial_t}|_\infty$ is proportional to the particle flux $j^r|_\infty$ and is given by[16]

$$\Phi_{\partial_t}|_\infty \propto |R_1|^2 - |R_2|^2. \tag{2.2.38}$$

Inserting the asymptotic decays (2.2.20) for $R_{1,2}$ this yields

$$\Phi_{\partial_t}|_\infty \propto \alpha_1^* \beta_1 + \alpha_1 \beta_1^* \,, \qquad \text{if } m \neq 0, \tfrac{1}{2}. \tag{2.2.39}$$

That is, the energy flux at the asymptotic boundary vanishes if we impose the above discussed no-source Dirichlet boundary conditions $\alpha_1 = 0$ or, for the alternative quantization, $\beta_1 = 0$ which do not deform the boundary conformal field theory.

On the other hand, for $m = 0$, inserting the asymptotic decays (2.2.23) for $R_{1,2}$ into (2.2.38) yields

$$\Phi_{\partial_t}|_\infty \propto \lambda^2 \alpha_1 \alpha_1^* - (\beta_1 + i\,\alpha_1 \omega L)\,(\beta_1^* - i\,\alpha_1^* \omega L) \,, \qquad \text{if } m = 0. \tag{2.2.40}$$

Again, this flux vanishes if we impose the standard (2.2.36) or alternative (2.2.37) quantizations, $\beta_1 = -i\,(\pm\lambda + \omega L)\,\alpha_1$ (and thus $\beta_1^* = i\,(\pm\lambda + \omega L)\,\alpha_1^*$).

Here it is important to recall the clarification about AdS/CFT boundary conditions and vanishing flux conditions presented in the Introduction. The standard and alternative boundary conditions that we use have, by construction, zero energy flux at the asymptotic boundary, as reviewed above and originally discussed in [75, 76, 77, 78, 79, 80, 81]. Without noticing, these standard/alternative boundary conditions are also the boundary conditions used in [73, 74] where the "generic physical principle of zero energy flux" was used to motivate the boundary conditions originally established in [75, 76, 77, 78, 79, 80, 81] (using precisely the same rationale). But there is a broader family of zero-flux boundary conditions. The AdS/CFT standard and alternative quantizations are a special class of zero-flux boundary conditions that, additionally, preserve the conformal symmetry group of AdS [75, 76, 77, 78, 79, 80, 81]. It is this property that singles them out among other zero-flux boundary conditions that break this conformal symmetry [79, 80, 81]. Further note that zero-flux boundary conditions are sometimes denoted as 'reflective' boundary conditions in some literature and both set of words encode the familiar idea that 'AdS behaves as a confining box' (under these boundary conditions).[17]

[16] Let again $\xi = \partial_t$ be the Killing vector field conjugate to the energy. The energy-momentum tensor for the Dirac field is $T_{\mu\nu} = \frac{i}{2}\left[\bar\psi \gamma_{(\mu}\mathcal{D}_{\nu)}\psi - \left(\mathcal{D}_{(\mu}\bar\psi\right)\gamma_{\nu)}\psi\right]$ and it is conserved $\nabla_\mu T^{\mu\nu} = 0$. This conservation law together with the Killing equation, $\nabla_{(\mu}\xi_{\nu)} = 0$, imply that the 1-form $\mathcal{J}_\mu = -T_{\mu\nu}\xi^\nu$ is conserved, $d \star \mathcal{J} = 0$, where $\star$ is the Hodge dual. We can then define the energy flux across the asymptotic hypersurface Σ (like the asymptotic boundary) as $\Phi_\xi \equiv -\int_\Sigma \star\mathcal{J} = -\int_\Sigma dV_\Sigma\; T_{\mu\nu}\xi^\mu n^\nu$ where n^ν is the unit normal vector to Σ and dV_Σ is the induced volume on Σ.

[17] Note however that in AdS/CFT there are other sets of boundary conditions that yield a well-defined boundary value problem but do not correspond to zero-flux boundary conditions (*e.g.* mass deformations describe sourced solutions with important physical interpretations where gauge field(s) have a non-vanishing asymptotic flux).

It is also important to emphasise that in the AdS/CFT language the standard classification of Dirichlet/Neumann/Robin boundary conditions applies to the physical fields that obey the original differential equation (in the present $s = 1/2$ case, the Dirac equation). Often this classification does not then translate into the same type of boundary conditions on auxiliary (or even gauge invariant) fields that one might introduce. The classification should focus on the physical fields and not on auxiliary fields (we can fabricate many of these), unlike what is done for $s = 1/2$ in [73, 74]. For example, for a massless Dirac fermion, no-source Dirichlet/Neumann boundary conditions $\psi_{\pm}^{(0)} = 0$ translate into $\beta_1 = -i\,(\pm\lambda + \omega L)\,\alpha_1$ not $\alpha_1 = 0$ or $\beta_1 = 0$. Facts like this are often missed:[18] zero-flux boundary conditions that preserve conformal symmetry require $\psi_{\pm}^{(0)}$ to vanish not $R_{1,2}|_{\infty}$.[19]

2.2.4 Near-horizon geometry of the extreme AdS-RN black hole

The near-horizon geometry of the extremal AdS-RN black hole will play an important role in our discussions in sections 2.3 and 2.4. Therefore, we review it here. The limiting procedure described below was first presented in [49].

The extremal AdS-RN black hole is given by (2.2.1) with $\mu = \mu_{\rm ext}$ given by (2.2.4). To obtain the near-horizon geometry, it is convenient to work in the gauge $A|_{r_+} = 0$ ($C = \mu$; otherwise we can do a gauge transformation in the end). One first zooms in around the horizon region by making the coordinate transformations:

$$r = r_+ + \varepsilon\rho, \qquad t = L_{AdS_2}^2 \frac{\tau}{\varepsilon}\,, \tag{2.2.41}$$

where L_{AdS_2} is the AdS_2 radius (to be defined below). Now the near-horizon geometry is obtained by taking $\varepsilon \to 0$ which yields

$$\begin{aligned} ds_{NH}^2 &= L_{AdS_2}^2 \left(-\rho^2 d\tau^2 + \frac{d\rho^2}{\rho^2}\right) + r_+^2 d\Omega_2^2, \qquad L_{AdS_2} = \frac{L r_+}{\sqrt{L^2 + 6r_+^2}}; \\ A_\mu^{NH} dx^\mu &= \alpha\,\rho\,d\tau, \qquad \alpha = L_{AdS_2}\sqrt{1 + \frac{L_{AdS_2}^2}{r_+^2}}\,. \end{aligned} \tag{2.2.42}$$

This geometry is the direct product of $AdS_2 \times \mathrm{S}^2$ and has a Maxwell potential that is linear in the radial direction. As mentioned in the introduction, it is still a solution of the 4-dimensional Einstein-Maxwell-AdS theory. On the other hand, the AdS_2 metric solves the 2-dimensional Einstein-AdS equations, $R_{\mu\nu} = -L_{AdS_2}^{-2} g_{\mu\nu}$, if L_{AdS_2} is identified as a

[18] This is *e.g.* the case in [85, 86] where massless Dirac quasinormal modes of Schwarschild-AdS are computed with the Dirichlet boundary condition $\alpha_1 = 0$. This choice of boundary condition is not one of the AdS/CFT zero flux boundary conditions for a massless Dirac field.

[19] Further note that there are other boundary conditions (*e.g.* $\beta_1 = -i\,(\pm i\,\lambda + \omega L)\,\alpha_1$) that make the flux (2.2.40) vanish. These should correspond to multi-traced (*i.e.* mixed or Robin) AdS/CFT boundary conditions [79, 81] which deform the boundary theory in a way that might be interesting for other studies.

function of the AdS_4 radius L and the horizon radius r_+ as indicated in the first line of (2.2.42).

2.3 Scalar fields in a AdS-RN background and their instabilities

Scalar fields confined inside the gravitational potential (like the AdS potential or a box in an asymptotically Minkowski background) of a black hole can condense creating near-horizon linear instabilities [53, 54, 55, 56, 44, 57, 39] (for planar AdS, this instability triggered the holographic superconductor programme [53, 54, 55]). Essentially this happens because we can have scalar fields that obey the asymptotically AdS_4 UV Breitenlöhner-Freedman (BF) bound but violate the 2-dimensional BF stability bound associated to the $AdS_2 \times S^2$ near-horizon geometry of the extremal black hole of the system. As we shall discuss in section 2.4.1, a similar violation of the 2-dimensional stability bound can occur for Dirac fields. In spite of this, as we will find in section 2.4.4, it turns out that Dirac fields are not linearly unstable to the near-horizon condensation mechanism. Therefore, before we discuss the fermionic case, it is important to revisit the scalar field case. This will allow to: 1) motivate the search of linear instabilities due to Dirac fields done in this chapter, 2) eventually identify differences between the two spins that could help in understanding the opposite outcomes. We also take the opportunity to demonstrate: i) how remarkably sharp the near-horizon instability bound (2.3.7) is by comparing it with the numerical solutions of the Klein-Gordon equation, and ii) that the unstable modes are both peaked near the horizon but also connected to the AdS normal modes (that is to say, in the language of [70, 71, 72] the AdS and near-extremal families of modes coincide and describe the unstable modes).

Using the fact that the AdS-RN background (2.2.1) is static and spherically symmetric we can consider a separation ansatz for the scalar field (with mass m and charge q) with the Fourier decomposition

$$\Phi(t, r, \theta, \phi) = e^{-i\omega t} e^{i m_\phi \phi} Y_\ell(\theta) \phi(r), \tag{2.3.1}$$

where $Y_\ell(\theta)$ are the familiar (spin-0) spherical harmonics – which are regular when the separation constant of the system is quantized as $\lambda = \ell(\ell+1)$, $\ell = 0, 1, 2, \cdots$ – and $|m_\phi| \leq \ell$ is the azimuthal quantum number. The Klein-Gordon equation yields the following equation for the radial function $\phi(r)$:

$$\frac{d}{dr}\left(r^2 f \frac{d\phi}{dr}\right) + \left(\frac{r^2}{f}\left(\omega + qA\right)^2 - r^2 m^2 - \ell(\ell+1)\right)\phi = 0. \tag{2.3.2}$$

A Taylor expansion around the asymptotic boundary yields the two independent solu-

tions

$$\phi(R) \simeq r^{-\Delta_-^{(s)}}(a+\cdots) + r^{-\Delta_+^{(s)}}(b+\cdots), \quad \text{with} \quad \Delta_\pm^{(s)} = \frac{3}{2} \pm \sqrt{\frac{9}{4} + m^2 L^2} \tag{2.3.3}$$

being the conformal dimensions of the field. Such a scalar field in AdS_4 is normalizable as long as its mass obeys the AdS_4 Breitenlöhner and Freedman (BF) bound, $m^2 \geq m^2_{\rm BF} \equiv -\frac{9}{4}\,\frac{1}{L^2}$ [28, 27].

Such a scalar field that is stable in the UV region can however be unstable in the IR region. This is best understood if we take the near-horizon limit of (2.3.2). Concretely, applying the near-horizon coordinate transformation (2.2.41) together with the near-horizon frequency transformation $\tilde{\omega} \to \widehat{\omega}\,\varepsilon/L^2_{AdS_2}$ (so that $e^{-i\tilde{\omega}t} \to e^{-i\widehat{\omega}\tau}$) followed by the near-horizon limit $\varepsilon \to 0$ yields the radial Klein-Gordon equation in the near-horizon geometry (2.2.42):

$$\partial_\rho\left(\rho^2 \partial_\rho R\right) + \left(\frac{\left(\widehat{\omega} + q\,\alpha\,\rho\right)^2}{\rho^2} - m^2 L^2_{AdS_2}\right) R = 0, \tag{2.3.4}$$

This is nothing else but the Klein-Gordon equation for a scalar field around AdS_2 with an electromagnetic potential $A_\tau = \alpha\,\rho$. A Frobenius analysis of (2.3.4) yields

$$R\big|_{\rho\to\infty} \simeq \rho^{-\widehat{\Delta}_-^{(s)}}(\widehat{a}+\cdots) + + \rho^{-\widehat{\Delta}_+^{(s)}}(\widehat{b}+\cdots), \quad \text{with} \quad \widehat{\Delta}_\pm^{(s)} = \frac{1}{2} \pm \frac{1}{2}\sqrt{1 + m^2_{eff} L^2_{AdS_2}}\,, \tag{2.3.5}$$

which determines the effective mass of the scalar field from the perspective of a near-horizon observer,

$$m^2_{eff\,(s)} L^2_{AdS_2} \equiv m^2 L^2_{AdS_2} - q^2\,\alpha^2 \tag{2.3.6}$$

Now, a scalar field with mass (2.3.6) in AdS_2 has unstable modes if it violates the AdS_2 BF bound $m^2_{eff\,(s)} \geq m^2_{{\rm AdS}_2\,{\rm BF}} \equiv -\frac{1}{4}\,\frac{1}{L^2_{AdS_2}}$. It follows that extremal AdS-RN_4 black holes should be unstable whenever the charge of the scalar field obeys

$$q^2 \geq \frac{1}{L^2}\left(1 + 4m^2 L^2 \frac{r_+^2}{L^2 + 6r_+^2}\right)\frac{\left(L^2 + 6r_+^2\right)^2}{8r_+^2\left(L^2 + 3r_+^2\right)}, \qquad \text{(near-horizon instability bound)}\,. \tag{2.3.7}$$

Note that scalar fields can also induce instabilities due to another mechanism that is known as superradiance. Unlike the near-horizon instability – which is suppressed in the limit $r_+/L \to 0$; indeed (2.3.7) goes as $q^2 L^2 \geq \frac{L^2}{8r_+^2} + \mathcal{O}(1)$ – the superradiant instability is present for small $r_+/L \ll 1$ black holes. For example, for $m = 0$, from the perturbative results of [39] one finds that the superradiant instability in extremal AdS-RN_4 is present

for scalar charges[20]

$$qL \geq \frac{3}{\sqrt{2}} - \frac{9}{2\sqrt{2}}\frac{r_+^2}{L^2} + \mathcal{O}\left(\frac{r_+^4}{L^4}\right), \qquad \text{(superradiant instability bound)} \tag{2.3.8}$$

Next, we solve the Klein-Gordon equation numerically to confirm that the near-horizon and superradiant instabilities are indeed present and to find how sharp the instability bounds (2.3.7) and (2.3.8) are. We present results for scalar masses above the unitarity bound $m^2_{\rm BF}+1=-5/4$ so asymptotically we impose the Dirichlet boundary condition $a=0$; see (2.3.3).[21] On the other hand at the horizon we require that the solution is regular in the future horizon which discards outgoing modes. To present the results, note that our system has a scaling symmetry [39] which means that the physical dimensionless quantities that are relevant for the problem are (this effectively sets $L\equiv 1$)

$$\left\{\frac{r_+}{L}, \mu; mL, qL, \omega L, \ell\right\}. \tag{2.3.9}$$

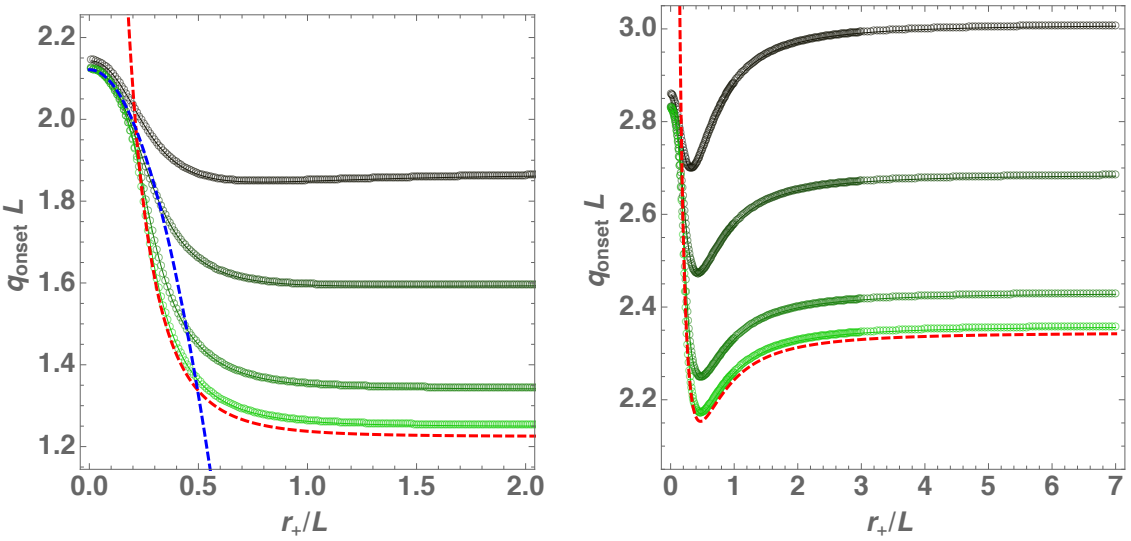

Figure 2.3.1: Onset scalar field charge as a function of the horizon radius for chemical potential $\mu=\mu_{\rm ext}(1-10^{-x})$ with $x=2,3,6,15$ (from top to bottom on the right of each panel). The left panel corresponds to massless scalar fields; the right panel to massive scalar fields with $mL=2$. The red dashed line is the near-horizon condensation analytic bound (2.3.7). In addition in the left panel we have the dashed blue line (for small horizon radius) which is the superradiant bound (2.3.8).

First, we are interested in finding the onset of the instabilities namely, the scalar field charge $q_{onset}(\mu, r_+/L; mL, \ell)$ above which the system is unstable. This onset occurs when

[20] This bound can be obtained from the expression for the frequency obtained in section III.D of [39]. Namely, the onset charge (2.3.8) of the superradiant instability is obtained by setting $\tilde{\omega}=0$ and $\mu=\mu_{\rm ext}$ in equation (55) of [39] and solving for the charge q. For further discussions between the entanglement of the superradiant and near-horizon instabilities and their different nature we ask the reader to see [39] and [40].

[21] For $m^2_{\rm BF}<m^2<m^2_{\rm BF}+1$ both modes are normalizable and thus we could also impose the Neumann boundary condition $b=0$ (the so-called alternative quantization in the context of AdS/CFT) [28, 27].

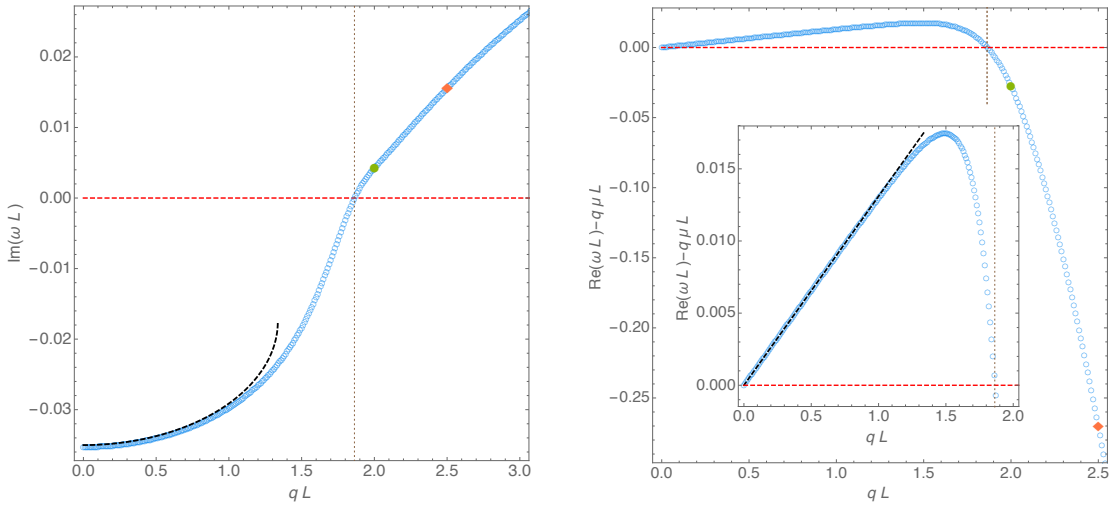

Figure 2.3.2: Scalar field frequency as a function of the dimensionless scalar charge qL for a AdS-RN black hole with $\mu = 0.99\mu_{\rm ext}$ and $r_+/L = 0.5$ ($mL = 0$ and $\ell = 1$). *Left panel*: Imaginary part of the dimensionless frequency, $\mathrm{Im}(\omega L)$. The system becomes unstable for $q > q_\star$ where $q_\star L \sim 1.863$. *Right panel*: Real part of the dimensionless frequency, $\mathrm{Re}(\omega L)$, measured with respect to $q\mu L$. This quantity changes sign at $q = q_\star$, *i.e.* when $\mathrm{Im}(\omega L)$ changes sign. In both plots, the black dashed curves describe the analytic prediction of the asymptotic matching expansion (A.2.10). We find very good agreement with the numerical results (blue curves) for $qL < 1.1$ (say). This is a further justification of the crude assumption that we should match with 0 in the overlapping region. *However* we find that these modes connect with the AdS normal modes as $r_+ \to 0$.

the frequency satisfies $\tilde{\omega} = \omega - q\mu = 0$. The Klein-Gordon equation (2.3.2) is then solved as an eigenvalue problem for $q = q_{onset}$. For concreteness, we fix $\ell = 1$ (we need $m_\phi \geq 1$ to have an instability). In the left plot of Fig. 2.3.1 we set $m = 0$ and we plot the dimensionless onset charge $q_{onset}L$ as a function of the dimensionless horizon radius r_+/L for different values of the chemical potential $\mu = \mu_{\rm ext}(1 - 10^{-x})$ that increasingly approaches the extremal value. From top to bottom, the green numerical curves describe chemical potentials with $x = 2, 3, 6, 15$. We see that as we get closer to extremality these onset curves increasingly approach (for values of r_+/L larger than ~ 0.25) the red dashed curve which describes the near-horizon bound (2.3.7). This strongly suggests that the instability, for large values of the horizon radius and near extremality, can be understood as due to the violation of the AdS_2 BF and that the associated near-horizon bound (2.3.7) is sharp (*i.e.* it is attained at extremality). On the other hand, as pointed out before, the near-horizon red dashed curve diverges as $r_+/L \to 0$. However, Fig. 2.3.1 shows that $q_{onset}L$ is finite for small r_+/L. Actually, in this regime the numerical onset curves are well described by the superradiant bound (2.3.8) (blue dashed curve with negative slope). This suggests that for small horizon radius and near extremality the instability has a superradiant nature and the superradiant bound (2.3.8) is sharp. For finite values of r_+/L, *i.e.* away from the $r_+/L \to 0$ and $r_+/L \to \infty$ regions, the superradiant and near-horizon

instabilities are entangled. These features are not unique to the massless case. For example, the onset charge plot for a scalar mass of $mL = 2$ is shown in the right panel of Fig. 2.3.1. Again, as extremality is approached the numerical green curves increasingly approach the near-horizon onset bound (2.3.7) (in this case we do not show the curve corresponding to the perturbative superradiant curve because it was not computed in [39] but we see that the behaviour of the onset curves for $r_+/L \ll 1$ is similar to the massless case).

To compare with what happens in the Dirac field case, it is enlightening to do the following exercise whose results are summarized in Fig. 2.3.2. We pick a particular AdS-RN background with chemical potential $\mu = 0.99\mu_{\rm ext}$ and horizon radius $r_+/L = 0.5$. We also fix the scalar mass to be $mL = 0$ and the scalar field harmonic number $\ell = 1$. Then we solve the Klein-Gordon equation to find the imaginary and real parts of the frequency ωL as a function of the dimensionless scalar field charge qL: these are shown in the left and right panels, respectively, of Fig. 2.3.2. From the left panel we see that, in accordance with the conclusions of Fig. 2.3.1, for small qL the system is stable (since $\mathrm{Im}(\omega L) < 0$) but there is a critical charge $q_\star L \sim 1.863$ (vertical dotted line) above which the system becomes unstable. Precisely at this critical onset charge one has $Re(\omega) - q\mu = 0$ and this quantity is negative (positive) for $q < q_\star$ ($q > q_\star$). The inset plot of Fig. 2.3.1 zooms-in the region $q < q_\star$. In section 2.4.4 we will find that the partner plot for Dirac fields is substantially different.

We take also the opportunity to understand better the frequency spectrum of scalar fields in AdS-RN. For global AdS RN black holes there are two quasinormal mode families [89, 90, 91]: one whose imaginary part grows negative without bound as the horizon radius r_+/L decreases, and another whose imaginary part vanishes as $r_+/L \to 0$ and whose real part approaches the normal modes of AdS. The unstable modes are found in this second family. This could well be the complete story. However, in *de Sitter* black holes there is a third family of quasinormal modes – called the near-extremal family – whose wavefunctions are spatially peaked near the horizon and that is distinct from the de Sitter family (as the name suggests, the latter is connected to the normal modes of de Sitter when the black hole shrinks). This naturally raises the question: could it be that in AdS one also has a near-extremal family of quasinormal modes that is not connected to the AdS family? If so, do the near-horizon unstable modes with bound (2.3.7) fit in this near-extremal family? We find a negative answer to these questions: the unstable modes belong to the AdS family of modes and the near-extremal family *coincides* with the AdS family. To arrive to this conclusion we first use a matching asymptotic expansion similar to the one used in de Sitter [70, 71, 72, 92] to find the frequency spectrum of the near-extremal family of quasinormal modes. This is done in Appendix A.2 and here we just quote the final result: near-extremality and for small scalar field charge one finds

that near-extremal modes have the frequency (for the lowest radial overtone $p = 0$)

$$\omega L \sim e\mu + \sigma\left[\frac{e\sqrt{1+3R_+^2}}{\sqrt{2}} - i\,\frac{1}{4R_+}\left((1+6R_+^2)(1+2p)\right.\right.$$
$$\left.\left.+\sqrt{(1+6R_+^2)\left[1+6R_+^2+4m^2R_+^2+4\ell(\ell+1)\right]-8q^2R_+^2(1+3R_+^2)}\right)\right]+\mathcal{O}(\sigma^2) \quad (2.3.10)$$

where $R_+ = r_+/L$, $e = qL$ and $\sigma = \frac{r_+ - r_-}{r_+}$ measures the distance away from extremality with $r_-(r_+, \mu, L)$ being the inner (Cauchy) horizon for which $f(r_-) = 0$. In Fig. 2.3.2, this analytical near-extremal frequency (with $\ell = 1$) is described by the dashed black curve. We find that it matches quite well the numerical result for small scalar charge. This indicates that the unstable modes fit into the *near-extremal family* of modes. But they also fit into the *AdS family* of normal modes. That is to say, unlike in the de Sitter case, in AdS the near-extremal and AdS family of modes coincide. To see this is indeed the case we pick two solutions in Fig. 2.3.2 that have $qL = 2.5$ (orange diamond) and $qL = 2$ and (keeping $\mu = 0.99\mu_{\rm ext}$ fixed) we follow this family of unstable modes as r_+/L decreases to zero.[22] This is done in Fig. 2.3.3 for $qL = 2.5$ and Fig. 2.3.4 for $qL = 2$. In both cases we find that, as $r_+/L \to 0$, $\mathrm{Im}(\omega L) \to 0$ and $\mathrm{Re}(\omega L) \to 3$, which is indeed the normal mode frequency of AdS with $\ell = 1$ (and lowest radial overtone).

In Fig. 2.3.3 and Fig. 2.3.4 the magenta dashed lines departing from the normal mode of AdS describe the frequency that one obtains when we consider a perturbative expansion in r_+/L and near-extremality about global AdS (and $\ell = 1$, $m = 0$). This result is taken from [39] (we already mentioned it to get the bound (2.3.8)):

$$\omega L \sim 3 - \frac{r_+}{L}\,\frac{4\left(6-4e\mu+3\mu^2\right)}{3\pi} + \frac{r_+^2}{L^2}\left[-\frac{4\left(-4e\mu+3\mu^2+6\right)\left(-52e\mu+45\mu^2+90\right)}{27\pi^2}\right.$$
$$\left.+\frac{1}{96}\left[108+\mu\Big(224e^2\mu-264e\left(\mu^2+2\right)+9\mu\left(3\mu^2+52\right)\Big)\right]-i\,\frac{16(3-e\mu)}{3\pi}\right]+\mathcal{O}\left(\frac{r_+^3}{L^3}\right) \quad (2.3.11)$$

where $e = qL$. So we see that not only the unstable modes approach the normal modes of AdS but they also do it at the expected rate in an expansion in r_+/L. The matching of our numerical results with the perturbative results (2.3.10) and (2.3.11) represents a non-trivial check of our results and illustrates the regime of validity of the perturbative results.

Now that we have highlighted the key features of the near-horizon (and superradiant)

[22] Note that $qL = 2.5$ is well above the onset curves of Fig. 2.3.1 for any r_+/L while the $qL = 2$ line is above the onset curves only above a certain horizon radius. So, for the latter charge, the system is unstable only above a critical value of r_+/L, as shown in Fig. 2.3.4.

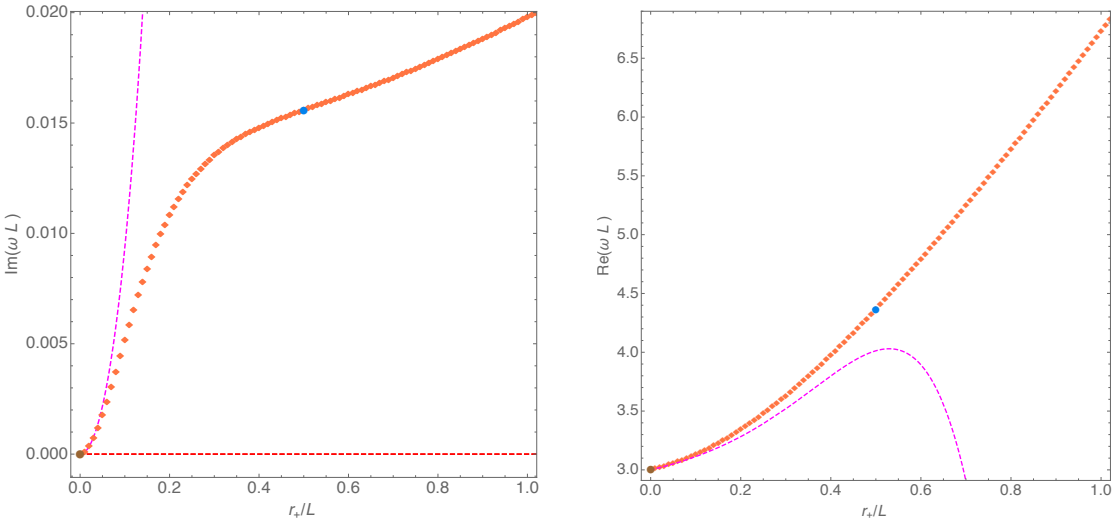

Figure 2.3.3: Scalar field frequency as a function of the dimensionless horizon radius r_+/L for a AdS-RN black hole with $\mu = 0.99\mu_{\rm ext}$ and $qL = 2.5$ that is always above the near-horizon bound (2.3.7) ($mL = 0$ and $\ell = 1$). *Left panel*: Imaginary part of the dimensionless frequency, $\mathrm{Im}(\omega L)$. *Right panel*: Real part of the dimensionless frequency, $\mathrm{Re}(\omega L)$. In both plots, the magenta dashed curves describe the analytic prediction of the perturbative expansion in r_+/L about AdS. The unstable modes are thus connected to the AdS normal modes when $r_+ \to 0$ (brown disk). For reference, the blue disk with $r_+/L = 0.5$ is shown (which makes contact with Fig. 2.3.2).

instabilities due to scalar perturbations in AdS-RN, we can proceed to the study of perturbations of Dirac fields in AdS-RN.

2.4 Searching for an instability of Dirac fields in the AdS-RN background

In section 2.3 we have seen that scalar fluctuations in the AdS-RN background give rise to the near-horizon scalar condensation instability. Moreover, we have seen that this instability is closely associated to the violation of the AdS_2 scalar BF stability bound. So much that the associated stability bound (2.3.7) for the onset of the instability is sharp. This naturally invites the questions: in the fermionic case can we also have a range of parameters where the AdS_2 fermionic stability bound is violated? If so what is the equivalent bound to (2.3.7) for the onset of the instability?

In this section we will address these questions. We will find that a near-horizon analysis of the Dirac equation indeed indicates that the AdS_2 fermionic stability bound can be violated near-extremality if the charge of the fermion is above a critical value (subsection 2.4.1). Encouraged by this result we will do a numerical analysis that will search for unstable modes in the region of parameters of interest (subsection 2.4.3). However, we

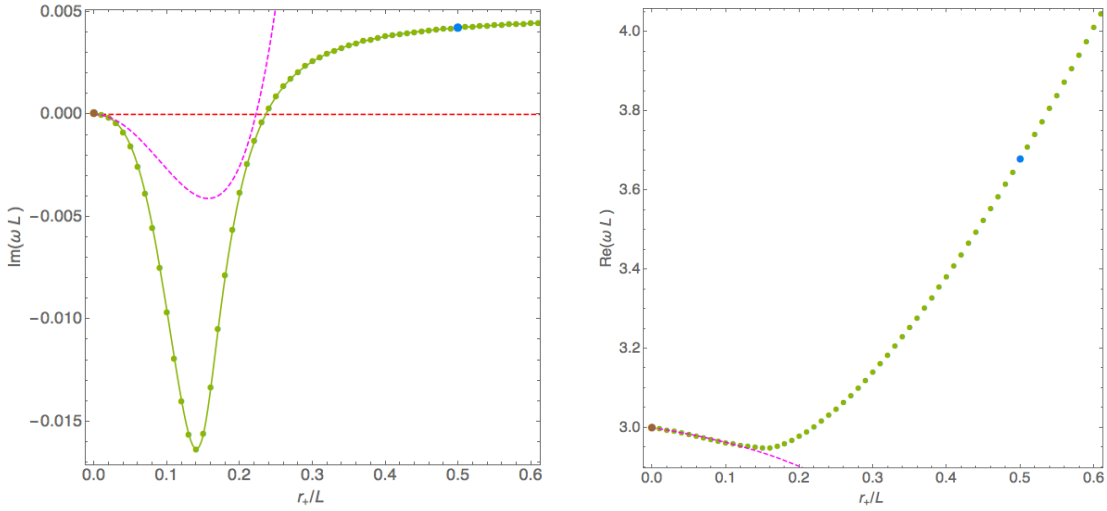

Figure 2.3.4: Similar to Fig. 2.3.3 but this time for a scalar field charge $qL = 2$ that is above the near-horizon bound (2.3.7) only for r_+/L above a certain value as seen in the left plot of Fig. 2.3.1.

will find no trace of instabilities, unlike in the scalar field case.

2.4.1 Argument for a near-horizon instability of Dirac fields

The Dirac equation in the near-horizon geometry (2.2.42) of the extreme AdS-RN black hole can be obtained taking the near-horizon limit of section 2.2.4 directly on the Dirac equation (2.2.17) for the extreme AdS-RN black hole. Concretely, applying the near-horizon coordinate transformation (2.2.41) together with the near-horizon frequency transformation $\tilde{\omega} \to \widehat{\omega}\,\varepsilon/L^2_{AdS_2}$ (so that $e^{-i\tilde{\omega}t} \to e^{-i\widehat{\omega}\tau}$) followed by the near-horizon limit $\varepsilon \to 0$ yields the Dirac equation in the near-horizon geometry (2.2.42):[23]

$$\rho\,\frac{d}{d\rho}\left(\rho\,\frac{dR_1(\rho)}{d\rho}\right) + \left[\frac{(\widehat{\omega}+q\,\alpha\,\rho)^2}{\rho^2} + i\,\frac{\widehat{\omega}}{\rho} - L^2_{AdS_2}\left(m^2+\frac{\lambda^2}{r_+^2}\right)\right] R_1(\rho) = 0 \tag{2.4.1}$$

where the AdS_2 radius L_{AdS_2} and the Maxwell near-horizon parameter α are defined in (2.2.42) and λ is the angular eigenvalue quantized as in (2.2.19). Also, recall that m and q are the mass and charge of the fermionic field.

Asymptotically, as $\rho \to \infty$, a Frobenius analysis of (2.4.1) finds that the solution $R_1(\rho)$ decays as

$$\rho^{-\frac{1}{2}} R_1\big|_{\rho\to\infty} \sim \rho^{-\widehat{\Delta}_-}\,(\widehat{\alpha}_1 + \cdots) + \rho^{-\widehat{\Delta}_+}\left(\widehat{\beta}_1 + \cdots\right), \tag{2.4.2}$$

where $\widehat{\alpha}_1, \widehat{\beta}_1$ are two arbitrary constants and we have introduced the AdS_2 conformal

[23] The field R_2 obeys a similar near-horizon Dirac equation that is just the complex conjugate of (2.4.1).

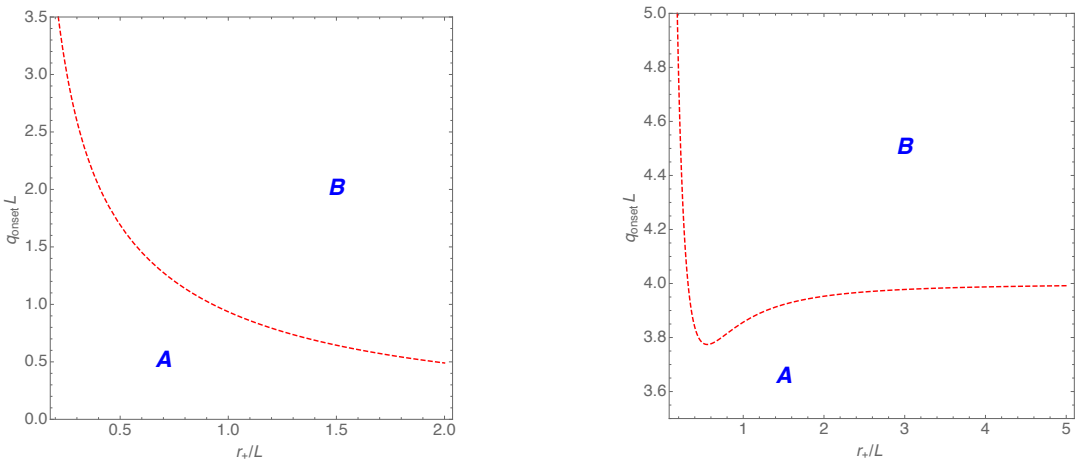

Figure 2.4.1: *Predicted* Dirac field charge for the onset of an instability as a function of the horizon radius for an extremal AdS-RN black hole ($\mu = \mu_{\rm ext}$). In both plots the red dashed curve is the lower bound of (2.4.4). The plot in the *left* (*right*) panel is for fermion mass $mL = 0$ ($mL = 4$) and harmonic number $\ell = 1/2$. The near-horizon analysis of the 2-dimensional stability bound violation leading to (2.4.4) predicts that region B should be unstable while region A should be stable (at least with respect to the stability mass bound mechanism). Note that for small r_+/L the system is not unstable because there is no superradiance for fermions and the predicted near-horizon instability is also suppressed. These Dirac figures can be (qualitatively) compared with Fig. 2.3.1 for the scalar field.

dimensions

$$\widehat{\Delta}_\pm = \frac{1}{2} \pm m_{\rm eff} L_{AdS_2} \quad \text{with} \quad m_{\rm eff} = \sqrt{m^2 + \frac{\lambda^2}{r_+^2} - \frac{q^2\alpha^2}{L^2_{AdS_2}}}\,. \tag{2.4.3}$$

The $s = 1/2$ stability bound is *independent* of the spacetime dimension and still given by (2.2.22), $m^2 \geq 0$ [79, 81]. Thus, the 2-dimensional fermionic stability bound is obeyed if $m^2_{\rm eff} \geq 0$ in (2.4.3). It follows that we can have situations where the Dirac field obeys the 4-dimensional fermionic stability bound (2.2.22), $m^2 \geq 0$, but violates the 2-dimensional stability bound. When this happens, *i.e.* when $m^2_{\rm eff} < 0$, one might expect an instability. This condition can be rewritten: the 2-dimensional stability bound is violated if the charge of the fermion is larger than

$$q \geq \frac{1}{\sqrt{2} r_+} \sqrt{\frac{L^2 + 6r_+^2}{L^2 + 3r_+^2}} \left(m^2 r_+^2 + \lambda^2\right)\,. \tag{2.4.4}$$

The equality applies strictly to the extremal case; as we move away from extremality, by continuity the instability should still be present but a higher fermion charge is needed to trigger it. Fig. 2.4.1 illustrates the regions where (2.4.4) predicts instability/stability.

At this level, we see that the near-horizon analysis of the possible violation of the AdS_2

stability bound for a Dirac field parallels very much the partner analysis done for a scalar field in section 2.3, with the minimum value for the charge (2.3.7) for the scalar case just replaced by the fermionic minimum value (2.4.4). In the scalar field case, we found (through a numerical study of linear perturbations in the AdS-RN background) that the violation of the 2-dimensional stability bound translates into the existence of a linear scalar condensation instability. Moreover, the near-horizon scalar bound (2.3.7) turns out to be very sharp, as best illustrated in Fig. 2.3.1. This scalar condensation linear instability indicates that non-linearly the AdS-RN black hole, when perturbed by a scalar field evolves towards a new configuration – a hairy black hole (with a scalar condensate floating above the horizon) – with the same UV asymptotics (since the 4-dimensional stability bound is satisfied) but with a different near-horizon geometry where the 2-dimensional stability bound is no longer violated [46, 57, 93, 39].

These considerations motivate the study done in this chapter for a Dirac field. In this case the AdS_2 stability bound can also be violated: at extremality this occurs for a fermion charge that saturates (2.4.4). From the lessons learned in the scalar field case one might well expect that the AdS-RN black hole, when perturbed by a Dirac field, is linearly unstable. To confirm whether this is the case, in the rest of this section we will solve numerically the Dirac equation in the AdS-RN background to hunt for a signature of the near-horizon linear instability. However, unlike the scalar field case, we will not find any evidence of a *linear* instability.

2.4.2 Dirac normal modes of global AdS

Before looking for potential instabilities (or frequency spectrum of damped oscillations) of Dirac modes in the global AdS-RN black hole it is convenient to first compute the normal mode spectrum of Dirac fields in global AdS. Indeed, some families of AdS-RN perturbations must reduce to these in the limit where the horizon shrinks to zero. Massive (section 2.4.2.1) and massless (section 2.4.2.2) Dirac fields require a distinct analysis.

2.4.2.1 Massive normal modes

For massive fermions in global AdS, it is not easy to solve directly the Dirac equations to get the radial functions $R_{1,2}$. There is however an appropriate combination of $R_{1,2}$ that yields equations of motion that are explicit hypergeometric equations. The linear combination for $R_{1,2}$ that we use below is motivated by a similar analysis done to compute the massive normal modes of fermions for *de Sitter* in [94].

For $m \neq 0$ and in global AdS, we introduce the new radial variable $y = -ir/L$ and make

the following field redefinitions

$$R_1(y) = (1-y^2)^{-\frac{1}{4}}(1-y)^{\frac{1}{2}}\Big[f_1(y) - f_2(y)\Big],$$
$$R_2(y) = (1-y^2)^{-\frac{1}{4}}(1+y)^{\frac{1}{2}}\Big[f_1(y) + f_2(y)\Big], \tag{2.4.5}$$

where $f_{1,2}(y)$ are functions to be determined. In these conditions, the coupled system of Dirac equations (2.2.15) yields

$$\big(1-y^2\big)\,\big(f_1' - f_2'\big) + \left(\omega L - \frac{1}{2}\right)(f_1 - f_2) + \left[mL(1+y) - \frac{1+y}{y}\left(\ell + \frac{1}{2}\right)\right](f_1 + f_2) = 0\,,$$
$$\big(1-y^2\big)\,\big(f_1' + f_2'\big) - \left(\omega L - \frac{1}{2}\right)(f_1 + f_2) - \left[mL(1-y) + \frac{1-y}{y}\left(\ell + \frac{1}{2}\right)\right](f_1 - f_2) = 0\,.$$

Adding and subtracting these two ODEs yields

$$\big(1-y^2\big)\,f_1'(y) + \left[mL\,y - \frac{1}{y}\left(\ell + \frac{1}{2}\right)\right] f_1(y) = \big(\omega L - mL + \ell\big) f_2(y)\,,$$
$$\big(1-y^2\big)\,f_2'(y) - \left[mL\,y - \frac{1}{y}\left(\ell + \frac{1}{2}\right)\right] f_2(y) = \big(\omega L + mL - \ell - 1\big) f_1(y)\,. \tag{2.4.6}$$

This pair of coupled first order ODEs can be straightforwardly rewritten as a decoupled pair of second order ODEs for f_1 and f_2. Moreover, if we introduce the new radial coordinate $z = y^2$ and the field redefinitions

$$f_1(z) = z^{\frac{\ell+1/2}{2}}\,(1-z)^{\frac{1}{4}(1-2\omega L)}\,F_1(z),$$
$$f_2(z) = z^{\frac{\ell+3/2}{2}}\,(1-z)^{\frac{1}{4}(1-2\omega L)}\,F_2(z), \tag{2.4.7}$$

each of the ODEs becomes a hypergeometric ODE with the standard form

$$z(1-z)F_i''(z) + \big[c_i - (a_i + b_i + 1)z\big]F_i'(z) - a_i b_i\,F_i(z) = 0\,, \qquad \text{for } i = 1, 2 \tag{2.4.8}$$

with parameters a_i, b_i and c_i given by

$$a_1 = \frac{1}{2}\,(1 + \ell - \omega L - mL)\,, \qquad b_1 = \frac{1}{2}\,(2 + \ell - \omega L + mL)\,, \qquad c_1 = 1 + \ell\,; \tag{2.4.9}$$
$$a_2 = \frac{1}{2}\,(3 + \ell - \omega L - mL)\,, \qquad b_2 = \frac{1}{2}\,(2 + \ell - \omega L + mL)\,, \qquad c_2 = 2 + \ell. \tag{2.4.10}$$

The most general solutions of (2.4.8) are [48]

$$\begin{aligned} F_1(z) &= A_1\,{}_2F_1\left(\frac{1}{2}(\ell-mL-\omega L+1),\frac{1}{2}(\ell+mL-L\omega+2);\ell+1;z\right) \\ &\quad +A_2 z^{-\ell}\,{}_2F_1\left(\frac{1}{2}(1-mL-\ell-\omega L),\frac{1}{2}(mL-\ell-\omega L+2);1-\ell;z\right), \\ F_2(z) &= B_1\,{}_2F_1\left(\frac{1}{2}(\ell-mL-\omega L+3),\frac{1}{2}(\ell+mL-\omega L+2);\ell+2;z\right) \\ &\quad +B_2 z^{-\ell-1}\,{}_2F_1\left(\frac{1}{2}(1-mL-\ell-\omega L),\frac{1}{2}(mL-\ell-\omega L);-\ell;z\right), \end{aligned} \tag{2.4.11}$$

where ${}_2F_1(a,b,c;z)$ is the Gaussian (ordinary) hypergeometric function and $A_{1,2}$, $B_{1,2}$ are arbitrary amplitudes. We can now plug (2.4.11) into (2.4.7) and then into (2.4.5) to get the most general solution for $R_1(r)$ and $R_2(r)$. Finally, we can insert this most general solution for $R_{1,2}(r)$ into (2.2.32) to get the most general solution for the Dirac fields $\psi_\pm(r)$. These are the physical fields that have to be regular everywhere and this constrains some of the amplitudes $A_{1,2}$ and $B_{1,2}$ and the frequencies. Namely, at the origin, $r=0$, one finds that both $\psi_\pm$ have two divergent terms of the form $B_2/r^{\ell+3/2}$ and $(2A_2-B_2)/r^{\ell+1/2}$. Regularity at the origin thus requires that one sets $A_2=0$ and $B_2=0$ and the other two amplitudes A_1 and B_1 are left arbitrary. It follows that the regular normal eigenmodes are

$$\begin{aligned} R_1 &= \sqrt{\frac{r}{L}\left(\frac{r}{L}-i\right)}\left(\frac{r^2}{L^2}+1\right)^{-\frac{\omega L}{2}}\left(-\frac{ir}{L}\right)^{\ell} \\ &\quad \left[A_1\,{}_2F_1\left(\frac{1}{2}(\ell-mL-\omega L+1),\frac{1}{2}(\ell+mL-\omega L+2);\ell+1;-\frac{r^2}{L^2}\right)\right. \\ &\quad \left. +iB_1\frac{r}{L}\,{}_2F_1\left(\frac{1}{2}(\ell-mL-\omega L+3),\frac{1}{2}(\ell+mL-\omega L+2);\ell+2;-\frac{r^2}{L^2}\right)\right], \\ R_2 &= \sqrt{1-\frac{ir}{L}}\left(\frac{r^2}{L^2}+1\right)^{-\frac{\omega L}{2}}\left(-\frac{ir}{L}\right)^{\ell+\frac{1}{2}} \\ &\quad \left[A_1\,{}_2F_1\left(\frac{1}{2}(\ell-mL-\omega L+1),\frac{1}{2}(\ell+mL-\omega L+2);\ell+1;-\frac{r^2}{L^2}\right)\right. \\ &\quad \left. -iB_1\frac{r}{L}\,{}_2F_1\left(\frac{1}{2}(\ell-mL-\omega L+3),\frac{1}{2}(\ell+mL-\omega L+2);\ell+2;-\frac{r^2}{L^2}\right)\right] \end{aligned} \tag{2.4.12}$$

We have not yet imposed the asymptotic boundary condition. A Frobenius analysis of (2.4.12) near the conformal boundary together with the use of (2.2.32) finds that $\psi_\pm$ be-

haves as (2.2.33) or (2.2.35) with

$$\begin{aligned}\alpha_1 &= \frac{A_1(-i)^\ell L^{-mL}\Gamma(\ell+1)\Gamma\left(mL+\frac{1}{2}\right)}{\Gamma\left[\frac{1}{2}(\ell+2+mL-\omega L)\right]\Gamma\left[\frac{1}{2}(\ell+1+mL+\omega L)\right]}\\ \beta_1 &= \frac{iB_1(-i)^\ell L^{mL}\Gamma(\ell+2)\Gamma\left(\frac{1}{2}-mL\right)}{\Gamma\left[\frac{1}{2}(\ell+3-mL-\omega L)\right]\Gamma\left[\frac{1}{2}(\ell+2-mL+\omega L)\right]}\end{aligned} \tag{2.4.13}$$

For $m>0$ ($m\neq 1/2$) the no-source standard boundary condition (2.2.36) requires $\alpha_1=0$. Using $\Gamma[-p]=\infty$ for $p=0,1,2,\cdots$ this quantizes the frequency as

$$\omega L=\ell+2+mL+2p \quad \text{or} \quad \omega L=-(\ell+1+mL+2p)\,, \qquad \text{(standard quantization)} \tag{2.4.14}$$

For $0<mL<\frac{1}{2}$ we can also impose the alternative quantization (2.2.37), *i.e.* $\beta_1=0$. This quantizes the frequency spectrum as (also with radial overtone $p=0,1,2,\cdots$)

$$\omega L=\ell+3-mL+2p \quad \text{or} \quad \omega L=-(\ell+2-mL+2p)\,, \qquad \text{(alternative quantization)} \tag{2.4.15}$$

2.4.2.2 Massless normal modes

In this section we find the normal modes in global AdS for a massless fermionic field. These have been previously discussed in [95, 73] but these references have not identified the full spectra of frequencies.

For $m=0$ in global AdS, introducing the change of coordinates and field redefinition

$$\begin{aligned}&z=\frac{2r}{r+iL}\,, \qquad 0\leq z\leq 2;\\ &R_1(z)=z^{\ell+\frac{1}{2}}(1-z)^{\frac{1}{2}\omega L}F(z),\end{aligned} \tag{2.4.16}$$

the radial equation (2.2.17) can be rewritten as a hypergeometric ODE in the standard form $z(1-z)F''(z)+(c-(a+b+1)z)F'(z)-abF(z)=0$, with

$$a=\ell+\frac{1}{2}\,, \qquad b=\ell+1+\omega L\,, \qquad c=2\,(\ell+1)\,. \tag{2.4.17}$$

Its most general solution is [48]

$$F(z)=C_1\,{}_2F_1\Big(\frac{1}{2}+\ell,\omega L+\ell+1,2(1+\ell),z\Big)+C_2\,z^{-1-2\ell}\,{}_2F_1\Big(-\frac{1}{2}-\ell,\omega L-\ell,-2\ell,z\Big). \tag{2.4.18}$$

Introducing this into (2.4.16) one gets $R_{1,2}(r)$ (note that $R_2=R_1^*$ as discussed in the next

section). Plugging this into (2.2.32) one finds the most general solution for the Dirac fields $\psi_\pm(r)$. At the origin, $r = 0$, these $\psi_\pm$ have a divergent term proportional to $C_2 r^{-\ell-3/2}$. Regularity at the origin thus requires that we set $C_2 = 0$. Now we need to impose the asymptotic boundary condition. One finds that asymptotically $\psi_\pm$ decays as (2.2.34) with

$$\begin{aligned}
\alpha_1 &= C_1\, i^{\omega L}\, 2^{\ell+\frac{1}{2}}\, {}_2F_1\left(\ell+\frac{1}{2}, \ell+\omega L+1; 2\ell+2; 2\right), \\
\beta_1 &= -C_1\, 2^{\ell-\frac{1}{2}}\, i^{\omega L+1}\Bigg[(2\omega L+2\ell+1)\, {}_2F_1\left(\ell+\frac{1}{2}, \ell+\omega L+1; 2\ell+2; 2\right) \\
&\quad + \frac{(2\ell+1)(\omega L+\ell+1)}{\ell+1}\, {}_2F_1\left(\ell+\frac{3}{2}, \ell+\omega L+2; 2\ell+3; 2\right)\Bigg].
\end{aligned} \tag{2.4.19}$$

As explained previously, for $m = 0$ we can impose either the standard or alternative boundary conditions. The no-source standard boundary condition (2.2.36), $\alpha_1(\lambda+\omega L) - i\beta_1 = 0$, quantizes the frequency spectrum as

$$\omega L = \ell + 2 + 2p \quad \text{and} \quad \omega L = -(\ell+1+2p), \quad p = 0, 1, 2, \ldots, \quad \text{(Standard quantization)} \tag{2.4.20}$$

On the other hand, for the no-source alternative quantization (2.2.37), $\alpha_1(\lambda-\omega L)+i\beta_1 = 0$, the normal mode frequencies of a massless Dirac field in global AdS are:

$$\omega L = \ell+1+2p \quad \text{and} \quad \omega L = -(\ell+2+2p), \quad p = 0, 1, 2, \ldots, \quad \text{(Alternative quantization)} \tag{2.4.21}$$

The positive frequencies in (2.4.20) and (2.4.21) were computed in [73] using vanishing flux boundary conditions that, as explained in the end of section 2.2.3, are exactly the AdS/CFT standard and alternative boundary conditions. However, [73] missed the existence of half of the normal mode spectrum, namely the half part that has negative frequencies. The relevance of the full spectrum (and associated relations between standard/alternative quantizations) is further analysed in the discussion of Fig. A.4.1. Further note that in RN, the four families of modes that reduce to (2.4.20)-(2.4.21) in the AdS limit become completely independent (i.e. they are not related by complex conjugation and the "degeneracy" is broken). This is further discussed in the next subsection.

2.4.3 Setup of the numerical problem

In this section we solve numerically the Dirac equation and search for linear instabilities of the Dirac solution in the AdS-RN background. Before proceeding it is important to note that: 1) the Dirac radial equation (2.2.17) for $R_2(r)$ is just the complex conjugate of the radial equation for $R_1(r)$ so if $R_1(r)$ is a solution one automatically has $R_2(r) = R_1(r)^*$, and 2) the Dirac angular equations for $S_{1,2}$ are related by the symmetry $\theta \to \pi - \theta$ so if $S_1(\theta)$ is a solution then $S_2(\theta) = S_1(\pi-\theta)$. Therefore, we just need to find

the solutions $R_1(r)$ ($S_1(\theta)$ are just the spin-weighted $s = 1/2$ spherical harmonics with quantum number ℓ).

Further note that if R_1 has charge q then $R_2 = R_1^*$ has charge $-q$, and complex conjugation maps quasinormal modes to quasinormal modes. It follows that if $\omega = \omega_r + i\omega_i$ is a linear mode frequency of R_1 then $-\omega^* = -\omega_r + i\omega_i$ is a linear mode frequency of $R_2 = R_1^*$. Thus, if we compute the frequency spectrum of R_1, we have the spectrum of R_2 too. It also follows that there is no loss of generality in assuming that $qQ > 0$ in our analysis: results for $qQ < 0$ are obtained simply by reversing the sign of the real part of the frequencies. Finally note that when we compute ω we have to allow both positive and negative values of ω_r, *i.e.* if $\omega_r = \mathrm{Re}(\omega)$ is a frequency of R_1 there is no symmetry in the system that requires $-\omega_r$ to be also a frequency of R_1. The only exception is if μ=0 (or $e = 0$) and $m = 0$ *i.e.* a massless Dirac field in Schwarzschild-AdS. In this case if $\omega = \omega_r + i\omega_i$ is an eigenvalue of R_1 so is $-\omega^* = -\omega_r + i\omega_i$, although with the opposite quantization: see (2.2.17) and (2.2.36)-(2.2.37) or (2.4.20)-(2.4.21).

For concreteness, we will set the mass of the fermion to zero, *i.e.* we solve the Dirac equation (2.2.17) for R_1 with $m = 0$ (and in the gauge $A|_\infty = 0$ where the frequency of the fermionic wave is ω; see footnote 12) subject to the physically relevant boundary conditions. The asymptotic decay of $R_1(r)$ is given in (2.2.23). For reasons discussed previously, we impose the asymptotic boundary condition (2.2.36) (standard quantization, $\psi_+^{(0)} = 0$) or (2.2.37) (alternative quantization, $\psi_-^{(0)} = 0$). At the horizon, for a non-extreme black hole, a Frobenius analysis finds that the two pairs of independent solutions are

$$R_1\big|_{r=r_+} = A_{in}\,(r-r_+)^{\frac{1}{2}-i\,\frac{\omega-q\mu}{4\pi T_H}}\Big(1+\mathcal{O}(r-r_+)\Big) + B_{out}\,(r-r_+)^{i\,\frac{\omega-q\mu}{4\pi T_H}}\Big(1+\mathcal{O}(r-r_+)\Big). \tag{2.4.22}$$

Rewriting this in ingoing Eddington-Finkelstein coordinates (v, r, θ, ϕ), with $v = t + \int f^{-1}\mathrm{d}r$, which are smooth across the future event horizon $\mathcal{H}^+$, we find that regularity of $R_1(r)$ at $\mathcal{H}^+$ requires that we impose the boundary condition $B_{out} = 0$. [24]

For the numerical solution it is convenient to redefine

$$R_1(r) = \Big(1-\frac{r_+}{r}\Big)^{\frac{1}{2}-i\,\frac{\omega-q\mu}{4\pi T_H}}\,q(r)\,, \tag{2.4.23}$$

and to work with the compact radial coordinate

$$z = \sqrt{1-\frac{r_+}{r}} \tag{2.4.24}$$

such that the horizon is now at $z = 0$ and the asymptotic infinity at $z = 1$. This has the

[24] For scalar fields, [96] used the real-time holography formalism [21] to show that imposing ingoing boundary conditions in the bulk horizon translates on the CFT side of the AdS/CFT correspondence to study *retarded* two-point functions.

advantage that analytical solutions that are smooth at the horizon simply have to obey the horizon Neumann boundary condition $q'(z=0)=0$. The asymptotic boundary condition for $q(z)$ follows straightforwardly from (2.2.36)-(2.2.37) and (2.4.23)-(2.4.24).

The numerical methods that we use are very well tested [97, 98, 99, 100, 101, 102, 56, 44, 83, 84, 103] and reviewed in [104]. To discretize the field equations we use a pseudospectral collocation grid on Gauss-Chebyshev-Lobbato points. The eigenfrequencies and associated eigenvectors are found using *Mathematica*'s built-in routine *Eigensystem*. For a given ℓ, this method has the advantage of finding several modes (i.e., from distinct families and with distinct radial overtones) simultaneously. However, to increase the accuracy of our results at a much lower computational cost we use a powerful numerical procedure which uses the Newton-Raphson root-finding algorithm discussed in detail in section III.C of the review [104]. All our results have the exponential convergence on the number of gridpoints, as expected for a code that uses pseudospectral collocation. In particular, all the results that we present are accurate at least up to the 10th decimal digit.

The Dirac equation in AdS-RN also has the scaling symmetry that determines that the physical dimensionless quantities are those listed in (2.3.9).

2.4.4 Main results

As discussed in section 2.4.3, for fermion mass $m=0$, we can have two independent homogeneous boundary conditions that yield normalizable modes: the standard ($\psi_+^{(0)}=0$) and alternative ($\psi_-^{(0)}=0$) boundary conditions. Moreover, for each of these boundary conditions, the eigenvector R_1 can have negative or positive real part of the frequency. It follows that, for a given harmonic ℓ, m_ϕ (and $m=0$) we have a total of two frequency spectra to discuss for each one of the two possible boundary conditions.

Before proceeding to the actual physical analysis of the frequency spectrum and instabilities of the system, in Appendix A.4 we first test our numerical code by comparing the associated numerical results with some analytical perturbative expansions that are derived in Appendix A.3. This confirms that our numerical code is generating physical data and we can now proceed and discuss our main physical findings.

Our aim is *not* to present the full spectrum of frequencies of a Dirac field in AdS-RN black hole. Instead, we are motivated to search for unstable modes, *i.e.* on eventually finding modes that, in some range of parameters, have $\mathrm{Im}(\omega L)>0$. There is a wide window of parameters to explore although the instability, if it exists, should appear near extremality for fermion charges q above a critical value. Thus one needs a good strategy to hunt efficiently for unstable modes. We proceed as follows. From the near-horizon bound

(2.4.4) arguing for the existence of an instability, we see that this bound is lower if we set $m = 0$ and $\ell = 1/2$. So, first, we either: 1) fixed $m = 0, \ell = 1/2$ and μ close to $\mu_{\rm ext}$, and varied $\{r_+/L, qL\}$, or 2) fixed $m = 0, \ell = 1/2$ and qL and varied $\{r_+/L, \mu\}$. In both cases, as described in the end of section 2.4.3, we solved our system as an eigenvalue problem for ωL. This finds "all" the solutions of the system (as long as the hierarchies do not grow large, *e.g.* $|\omega L| \gg 1$, which makes the numerical problem hard). This allows to eventually identify unstable modes with $\mathrm{Im}(\omega L) > 0$ or, in the worst case, to identify modes with $\mathrm{Im}(\omega L) < 0$ that are closest to the marginal case for instability ($\mathrm{Im}\,\omega = 0$). Once these interesting modes are identified we then used a Newton-Raphson root-finding algorithm to follow efficiently the modes to other values of the parameter space.

In spite of our efforts, we have found no sign of an instability. Recall that for $mL = 0$ both the standard (2.2.36) and the alternative (2.2.37) boundary conditions yield normalizable modes. In general, we do find that the stable modes with smallest $|\mathrm{Im}(\omega L)|$ are those that reduce to the alternative normal modes of AdS (2.4.21) or to the standard AdS normal modes (2.4.20), when the horizon radius shrinks to zero. Among these, we further find that the modes with smallest $|\mathrm{Im}(\omega L)|$ are, for both quantizations, the ones that reduce to the *positive* normal mode frequencies when $r_+/L \to 0$, *i.e.* $\omega L = 3/2$ (alternative boundary condition) and $\omega L = 5/2$ (standard quantization). Therefore, to avoid distraction from the main point, in the rest of this chapter we only discuss these two families of modes.

Probably the plots that best illustrate the main conclusions of our Dirac study are those of Fig. 2.4.2 (for alternative quantization) and of Fig. 2.4.5 (for standard quantization). Recall that in the best case scenario the expectation is that, close to extremality, modes should become unstable above a fermion charge q that should be higher than the near-horizon bound (2.4.4). Thus, in these figures we fix the black hole horizon to be $r_+/L = 0.5$ and choose a chemical potential close to extremality, $\mu = 0.99\mu_{\rm ext}$. Starting from $qL = 0$, where $\mathrm{Im}(\omega L) < 0$, we then increase this charge to see if there is a critical value above which $\mathrm{Im}(\omega L)$ becomes positive. (That is to say, we adopt a similar strategy as the one followed in the scalar field case to get Fig. 2.3.2).

For the alternative quantization, the left panel of Fig. 2.4.2 shows that, starting from $q = 0$, as qL grows, $\mathrm{Im}(\omega L) < 0$ increases and approaches $\mathrm{Im}(\omega L) = 0$ very closely. However, no matter how large qL is we never reach a situation where $\mathrm{Im}(\omega L) \geq 0$. Interestingly, there is a critical value of q, namely $qL = q_{\rm max}L \sim 0.9390$ (vertical brown dashed line) where $\mathrm{Im}(\omega L)$ reaches a maximum value of $\mathrm{Im}(\omega L) \sim -0.000548$ (see the inset plot which zooms-in around this maximum). But increasing qL further, $\mathrm{Im}(\omega L)$ becomes again increasingly more negative (instead of becoming positive). The Dirac field system behaves therefore substantially distinctly from the scalar field case of Fig. 2.3.2 (left panel) where there was a critical qL above which $\mathrm{Im}(\omega L)$ becomes positive. To complete the

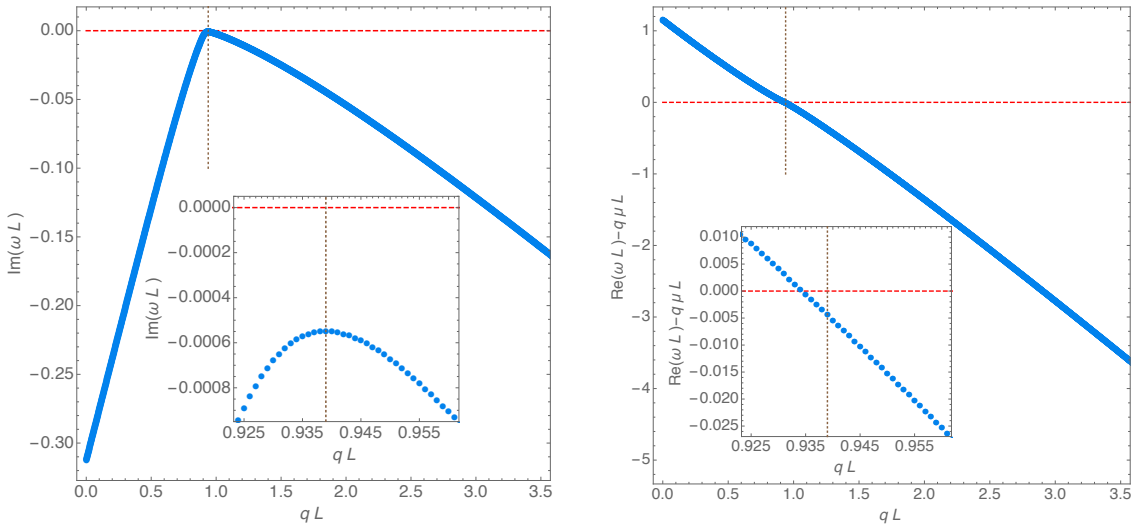

Figure 2.4.2: Dirac field frequency with alternative quantization (2.2.37) as a function of the dimensionless scalar charge qL for a AdS-RN black hole with $\mu = 0.99\mu_{\rm ext}$ and $r_+/L = 0.5$ (also, $mL = 0$ and $\ell = 1/2$). *Left panel*: Imaginary part of the dimensionless frequency, $\mathrm{Im}(\omega L)$ which attains a maximum for $q = q_{\rm max} \sim 0.9390/L$ (vertical brown dashed line). The inset plot zooms-in around this maximum and shows that $\mathrm{Im}(\omega L) < 0$ for any qL. *Right panel*: Real part of the dimensionless frequency, $\mathrm{Re}(\omega L)$, measured with respect to $q\mu L$. This quantity changes sign at $q = q_\star \sim 0.9344/L$ with $q_\star < q_{\rm max}$, *i.e.* for a smaller qL than the one where $\mathrm{Im}(\omega L)$ attains its maximum value (vertical brown dashed line): this is better seen in the inset plot which zooms-in the relevant region.

analysis, in the right panel of Fig. 2.4.2, we plot $\mathrm{Re}(\omega L) - q\mu L$. We find that for small qL this quantity is positive but becomes negative above $q = q_\star \sim 0.9344/L$. Interestingly, this occurs at a charge that is smaller than $q_{\rm max}$ where the maximum of $\mathrm{Im}(\omega L)$ is reached (vertical brown dashed line): this is better seen in the inset plot which zooms-in the relevant region. Again we note the difference to the scalar field case displayed in the right panel of Fig. 2.3.2 where $\mathrm{Re}(\omega L) - q\mu L$ changes sign precisely at the critical value of qL where $\mathrm{Im}(\omega L) = 0$. Further note that these plots also show that for a Dirac field we do not have a value of qL for which we simultaneously have $\mathrm{Re}(\omega L) - q\mu L = 0$ and $\mathrm{Im}(\omega L) = 0$. Therefore, we cannot set $\omega L = q\mu L$ in the equations of motion and solve these as an eigenvalue problem for the instability onset charge. That is to say, unlike the scalar field case, we do not have an onset charge that would produce the partner plots of the scalar field onset plots of Fig. 2.3.1. The predictions of Fig. 2.4.1 do not hold (at least at the linear mode level).

We have done similar experiments as those of Fig. 2.4.2 for other black hole parameter values μ and r_+/L. Keeping μ fixed, black holes with distinct r_+/L have plots similar to Fig. 2.4.2 with the feature that larger values of r_+/L reach the maximum of $\mathrm{Im}(\omega L)$ (but remaining negative) at smaller critical values of $q = q_{\rm max}$. On the other hand, keeping r_+/L fixed, black holes with distinct μ also have similar plots to Fig. 2.4.2 with the prop-

erty that larger values of μ reach the maximum of $\mathrm{Im}(\omega L)$ (but still negative) at smaller critical values of $q = q_{\max}$ and this maximum of $\mathrm{Im}(\omega L)$ is increasingly closer to zero as μ approaches the extremal value μ_{ext}.

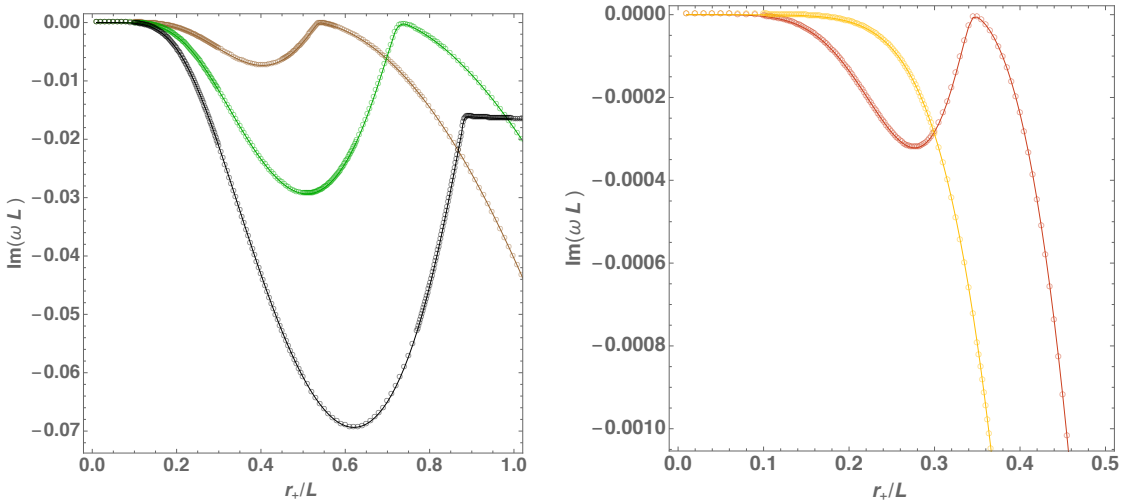

Figure 2.4.3: Imaginary part of the frequency as a function of the horizon radius for chemical potential $\mu = 0.999\mu_{\mathrm{ext}}$ for different values of the fermion field charge. In the left panel the charges are (from bottom bump to top) $qL = 0.7$ (black), $qL = 0.8$ (green), $qL = 0.9$ (brown). In the right panel the charges are $qL = 1$ (red) and $qL = 1.1$ (yellow). Notice the different regions scanned by the axes in the two plots.

To have a complementary perspective of the system's properties, in Fig. 2.4.3 and in Fig. 2.4.4 we illustrate other attempts we have made to find an instability. In Fig. 2.4.3, we keep the alternative quantization and fix the chemical potential at $\mu = 0.999\mu_{\mathrm{ext}}$, and plot $\mathrm{Im}(\omega L)$ as a function of r_+/L for five different values of qL, namely, $qL = 0.7,\ 0.8,\ 0.9$ (from bottom to top in the left panel) and $qL = 1,\ 1.1$ (right panel). (The two plots are needed for the presentation of the results because the relevant $qL = 1$ case in the right panel reaches a maximum that is approximately two orders of magnitude higher than the first three cases on the left panel). The main feature in these plots is the typical presence of a local minimum and local maximum (bump). As we increase the fermion charge from zero to a value slightly above 1, the relative minimum and relative maximum of $\mathrm{Im}(\omega L)$ raise and shift to lower values of r_+/L. But the local maximum always has $\mathrm{Im}(\omega L) < 0$, *i.e.* there is no instability. However, for charges qL above a value that is in between 1 and 1.1, the local minimum and maximum are no longer present and $\mathrm{Im}(\omega L)$ decreases monotonically with r_+/L (see *e.g.* $qL = 1.1$ displayed as the yellow curve in the right panel; higher values, $qL \geq 1.1$, have a similar monotonic behaviour).

As yet another illustration of experiments we made, in Fig. 2.4.4 we fix the fermion charge to be $qL = 1$ (which was already analysed in Fig. 2.4.3 for $\mu = 0.999\mu_{\mathrm{ext}}$) and we study the effect that changing the chemical potential has by considering a total of 5 curves with 5 different values of μ. Namely, in the left plot we consider the cases $\mu = 0.9\mu_{\mathrm{ext}}$

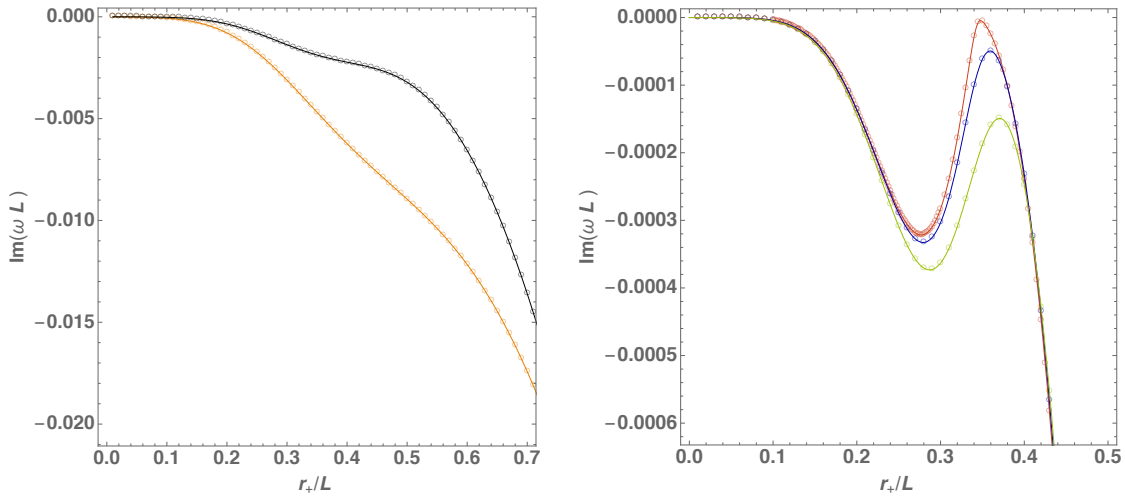

Figure 2.4.4: Imaginary part of the frequency as a function of the horizon radius for fermion field charge $qL = 1$ and different values of the chemical potential. In the left panel the chemical potentials are (from bottom to top) $\mu = 0.9\mu_{\rm ext}$ (orange) and $\mu = 0.95\mu_{\rm ext}$ (black). In the right panel the chemical potentials are closer to extremality, namely (from bottom to top curves): $\mu = 0.99\mu_{\rm ext}$ (green), $\mu = 0.995\mu_{\rm ext}$ (blue) and $\mu = 0.999\mu_{\rm ext}$ (red). Notice the different regions scanned by the axes in the two plots.

and $\mu = 0.95\mu_{\rm ext}$. These cases have no bump (no local maximum) and illustrate that it only appears close to extremality. In the right panel we show three more cases where we fix $\mu = 0.99\mu_{\rm ext}$, $\mu = 0.995\mu_{\rm ext}$ and $\mu = 0.999\mu_{\rm ext}$ (from bottom to top). The bump is now present and the local maximum increases as one approaches extremality but never becomes positive. For the case $\mu = 0.999\mu_{\rm ext}$ this local maximum is at $\mathrm{Im}(\omega L) \sim -5.93 \times 10^{-6}$.

So far we have focused our discussion of the results for the alternative quantization case because, typically, for the same values of black hole parameters this is the case where $\mathrm{Im}(\omega L)$ approaches $\mathrm{Im}(\omega L) = 0$ the most. Nevertheless, we have also tried hard to find an instability in the standard boundary condition (2.2.36) case. Again without success. To illustrate briefly this conclusion, in Fig. 2.4.5 we give the partner plot of Fig. 2.4.2 but this time for the standard quantization. Although the features of Fig. 2.4.5 are clearly more elaborated than those of Fig. 2.4.2 (*e.g.* there are several local maxima and minima), the main conclusions are still the same: *i*) one always has $\mathrm{Im}(\omega L) < 0$; *ii*) there is a $q = q_{\rm max}$ where the solution approaches $\mathrm{Im}(\omega L) = 0$ the most (vertical brown dashed line); *iii*) $Re(\omega L) - qL\mu$ changes sign at $q = q_\star < q_{\rm max}$. It follows that we find no sign of an instability and the standard boundary condition case, much like the alternative quantization case, also gives results that are very different from the scalar field case of Fig. 2.3.2.

Altogether, all our attempts – best illustrated in Figs. 2.4.2-2.4.4– to find an instability due to Dirac field perturbations with alternative boundary condition (2.2.37) failed

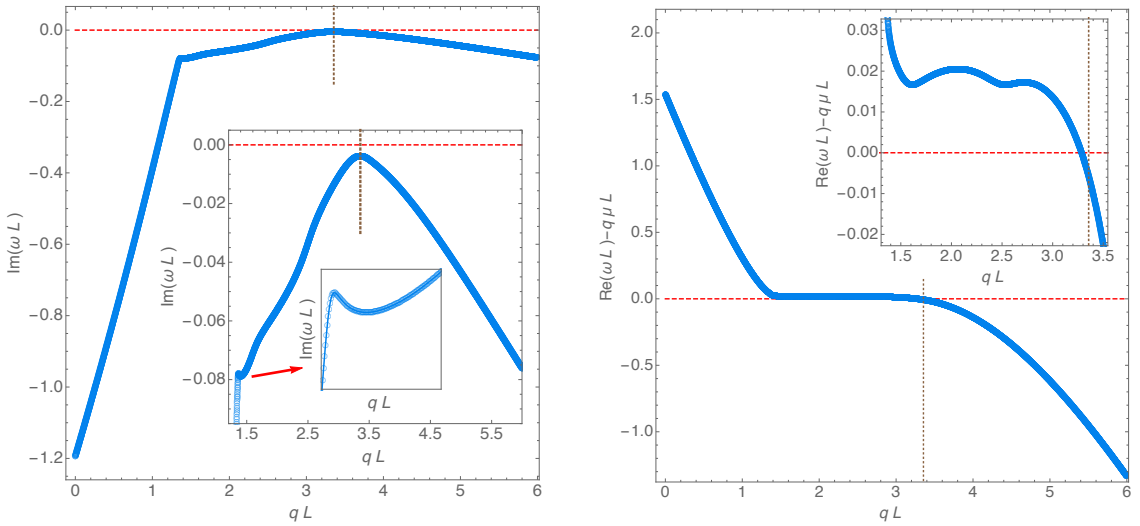

Figure 2.4.5: Dirac field frequency with standard quantization (2.2.36) as a function of the dimensionless scalar charge qL for a AdS-RN black hole with $\mu = 0.99\mu_{\rm ext}$ and $r_+/L = 0.5$ (also, $mL = 0$ and $\ell = 1/2$). *Left panel*: Imaginary part of the dimensionless frequency, $\mathrm{Im}(\omega L)$ which attains a maximum of $\mathrm{Im}(\omega L) \sim -0.0037491$ for $q = q_{\rm max} \sim 3.3555/L$ (vertical brown dashed line). The main inset plot zooms-in around this maximum and shows that $\mathrm{Im}(\omega L) < 0$ for any qL. The secondary inset plot shows the detail of the curve around $qL \sim 1.45$ to show that the apparent cusp in the main plot is smooth. *Right panel*: Real part of the dimensionless frequency, $\mathrm{Re}(\omega L)$, measured with respect to $q\mu L$. This quantity changes sign at $q = q_\star \sim 3.2873/L$ with $q_\star < q_{\rm max}$, *i.e.* for a smaller qL than the one where $\mathrm{Im}(\omega L)$ attains its maximum value (vertical brown dashed line): this is better seen in the inset plot which zooms-in the relevant region.

miserably. The outcome is similar when we consider the standard boundary condition (2.2.36). This is best summarized in Fig. 2.4.5 where we show the partner plots of the alternative quantization Fig. 2.4.2 but this time for the standard quantization. Again, and essentially, we find that $\mathrm{Im}(\omega L)$ reaches a maximum at a critical $q_{\rm max}$ but never crosses the borderline $\mathrm{Im}(\omega L) = 0$ that would signal an instability.

2.5 Discussion and conclusions

A scalar field in an asymptotically AdS_4 Reissner-Nordström black hole can satisfy the asymptotically AdS_4 UV Breitenlöhner-Freedman (BF) stability bound but violate the infrared 2-dimensional BF stability bound associated to the $AdS_2 \times S^2$ near-horizon geometry of the extremal black hole of the system, as reviewed in section 2.3. When this is the case, the AdS-RN black hole is unstable to scalar condensation and the system evolves to a new configuration in the phase diagram of solutions that preserves both the UV BF bound and the near-horizon 2-dimensional stability bound. Such a solution is a hairy black hole with a charged scalar field floating above the horizon [54, 55, 105, 46, 44, 106, 93]. Coulomb repulsion balances the gravitational force and the system is

static. There is no doubt that the violation of the AdS_2 stability bound is the physical mechanism responsible for the near-horizon scalar condensation instability since the associated minimum bound (2.3.7) on the scalar field charge that triggers the instability is sharp (at extremality) as best demonstrated by Fig. 2.3.1.

Given these considerations, the study done in this part for Dirac field perturbations in the global AdS_4 RN black hole was motivated by the following observation. Dirac fields in AdS-RN can also preserve the UV fermionic stability bound (2.2.22) [79, 81] but violate the near-horizon infrared fermionic stability bound, as seen in section 2.4.1. From the scalar field case lessons, this suggests that the system might be unstable to fermion condensation. However, in spite of our efforts to scan the relevant parameter space near extremality, we found no sign of a *linear mode* instability. The sharp distinction between the scalar and Dirac field cases is best illustrated comparing the scalar Fig. 2.3.2 with the Dirac Fig. 2.4.2 (for alternative quantization) or Fig. 2.4.5 (for standard quantization). Of course our numerical study does not prove linear stability but we did such a detailed scan that we are very confident that no linear instability is present. Our stability results are also consistent with the stability study of fermions in *planar* AdS, where no instability was found [55, 63] (see also [64, 69, 65, 66, 67, 68]).[25] Indeed, the planar AdS case is the $r_+/L \to \infty$ limit of the global AdS system.

So, the planar AdS studies [55, 63, 64, 69, 65, 66, 67, 68] and our present study in global AdS establish that the violation of the 2-dimensional stability bound of a Dirac field in AdS-RN does not lead to a *linear* mode instability. However, such solutions correspond to negative energy Schrödinger states: without a positive self-adjoint extension for the Schrödinger operator the dynamical evolution of the system should develop an instability... In particular, the system might indeed still be unstable if *non*-linear effects play a role in the discussion. That is to say, if we perturb a AdS-RN black hole with a Dirac field in a region of parameter space where the infrared stability bound is violated, it could still be the case that the system evolves non-linearly to a new configuration that has a charged Dirac field floating above the horizon and that preserves both the UV and IR stability bounds. How difficult would it be to prove whether this scenario is correct?

One must proceed with caution. To begin with one needs to first formulate more precisely the setup of the problem. It is certainly much harder to find, if they exist, the proposed Dirac hairy black holes than it was to construct the scalar hairy black holes [54, 55, 105, 46, 44, 106, 93]. There is a fundamental difference between fermionic and bosonic fields. The fermion has no classical limit: Planck's constant $\hbar$ is present in the stress tensor and associated equations of motion for a fermion. As discussed in detail in

[25] This sharp difference between the stability conclusion of scalar and Dirac fields in RN-AdS is probably due to the fact that the Dirac equation is originally a first order PDE. We acknowledge Don Marolf for this observation (private communication).

section 14.3 of Wald's textbook [5], the absence of a classical limit means that in Einstein's equation we have to promote the differential Einstein and energy-momentum operators G and T to quantum operators and the quantum version of Einstein's that gives the back-reaction of a fermion on the gravitational field is $\langle G_{\mu\nu}\rangle = 8\pi(\langle T^{\rm Max}_{\mu\nu}\rangle + \langle T_{\mu\nu}\rangle)$, where $T^{\rm Max}_{\mu\nu}$ and $T_{\mu\nu}$ stands for the Maxwell and Dirac stress tensor contributions and $\langle\cdots\rangle$ stands for the expectation value of the corresponding operator.

Thus, to find the backreaction that fermions induce on the gravitoelectromagnetic background one needs to first compute the expectation value of the fermion energy momentum tensor $\langle T_{\mu\nu}\rangle$. This is a highly non-trivial task. Even worse, once we consider the quantum backreaction of fermions one also needs to consider the quantum backreaction of gravitons and photons, *i.e.* one also needs to compute $\langle G_{\mu\nu}\rangle$ and $\langle T^{\rm Max}_{\mu\nu}\rangle$ [5]. In a best case scenario, where we have a large number N of Dirac fields, one might be able to assume that, roughly speaking, the effects of N Dirac fields are N times as relevant as that of the gravitons and photons [5]. For a 'fermionic hairy black hole', the fermionic condensate should be made of a large number of fermions. In these conditions, for large N, one might be able to neglect the quantum backreaction of gravitons and photons and work in the semi-classical limit whereby the backreaction of the Dirac field on the gravitoelectromagnetic background is simply governed by $G_{\mu\nu} = 8\pi T^{\rm Max}_{\mu\nu} + 8\pi N\langle T_{\mu\nu}\rangle$. This semi-classical system should be viewed as the leading term of a $1/N$ expansion of the full theory [5]. But this semi-classical computation still requires that one computes $\langle T_{\mu\nu}\rangle$. And this is still a formidable task. An overview on the physical and technical tools required to accomplish this task can be found in [107, 108, 68] (and references there-in) where asymptotically planar AdS quantum electron stars are discussed as semi-classical solutions of Einstein-Maxwell theory.

Finally note that in the present part we focused our attention on modes that could eventually become unstable. We have not studied in detail the full spectrum of quasinormal mode frequencies of the Dirac field in AdS-RN. Moreover, we focused on the case of a massless fermion because, as explained previously, this was enough for our purposes. However, the equations of motion and relevant boundary conditions for any fermion mass and any sector of perturbations are given in section 2.2. We have also computed the normal modes of massive fermions in AdS (previously only the massless spectrum was computed). It might be useful to have a more complete frequency spectra study for future studies/applications. It might also be interesting to look for perturbations of spin $3/2$ Rarita-Schwinger fields about AdS-RN black holes. In this case, there are also normalizable solutions that become unstable for negative square masses [80]. Probably there will be no linear near-horizon instabilities when the effective 2-dimensional mass violates the AdS_2 stability bound but, as far as we know, this was never checked.

Part III

Solitons confined in a Minkowski box

CHAPTER 3

Boson stars in a Minkowski box.

3.1 Introduction

Confining gravitational boxes became notorious in the context of superradiant instabilities when Press and Teukolsky introduced the black hole bomb system [35] (see also [109]). A scalar wave impinging on a rotating black hole may extract energy from the black hole provided its frequency ω satisfies the superradiant bound $\omega < m_\varphi \Omega_H$, where m_φ is the azimuthal quantum number and Ω_H is the angular velocity of the horizon. In standard conditions this scalar wave would then disperse to the asymptotic region and die off. However, when surrounded by a reflecting cavity, the scalar wave undergoes multiple superradiant amplifications and reflections and an instability builds up. Similar superradiant instabilities occur for charged scalar fields confined around Reissner-Nordström black holes (RN BHs), whereby the superradiant frequency bound now reads $\omega < e\mu$, where e is the charge of the scalar field and μ is the chemical potential of the black hole [36]. In both black hole bomb systems, the onset or zero-mode of the superradiant instability signals a bifurcation to a novel family of hairy black holes: the solution outside the box is described by the Kerr or RN solution but, inside the box, there is also a non-trivial scalar field floating above the horizon. The scalar field cloud is stationary because either centrifugal effects or Coulomb repulsion balance the system against gravitational collapse.

Consider now this very same gravitational box confining a charged scalar field but, this

time, it is simply placed in Minkowski background (eventually also with a radial electric field) with no event horizon. As is well known, at linear order in perturbation theory, the scalar field frequencies that can fit inside the box radius are quantized; these are the normal modes of the system. Interestingly, beyond linear order in perturbation theory, there is nothing impeding us from back-reacting a normal mode to higher orders while keeping the solution regular everywhere and still confined as a stationary configuration inside the box [43]. In this case we have an asymptotically Minkowski soliton or boson star confined in a box.

In [43], the authors constructed some of the above static charged hairy solitons (boson stars) and hairy black holes in perturbation theory (with the amplitude of the scalar field and the ratio between the horizon and box radius as expansion parameters). By construction, such solutions are valid only for small energy and charge and they are perturbatively connected to the Minkowski box solution. In the present part we complete the analysis initiated in [43] and solve the full nonlinear Einstein-Maxwell-scalar field equations to find the exact numerical solution, that describes the solitons (boson stars) of the theory confined in a box, also in the non-perturbative regime. At low energies/charges our solutions are well described by the perturbative predictions of [43]. However, at intermediate and large energies (when compared to the box lengthscale) the phase diagram of solutions develops an intricate structure that was not anticipated at all by the perturbative analysis. In particular, we will find a *main branch* of boson stars (solitons) that, in the small charge regime, is described within perturbation theory but that has (at higher charges) a Chandrashekhar mass limit and multi-branched structure that was not captured within perturbation theory. Additionally, we also find a *secondary family* of boson stars that is not captured at all by perturbation theory. Finally, we will find that the properties of boson stars have a non-trivial dependence on the electric scalar field charge e, most of which were not anticipated by the perturbative analysis.

Some key properties of asymptotically flat caged solitons turn out to be similar to those observed in asymptotically anti-de Sitter solitons [110], [46, 45, 44, 93, 47, 111, 39]. We thus identify features that seem to, or might be universal to charged scalar condensates (or other bosonic field condensates) confined in a potential well.

3.2 Summary of phase diagram

For clarity, in this section we summarize our main results since some of the plots that we will present in section 4.4 are elaborated. We find charged scalar boson stars (a.k.a. solitons) confined inside a gravitational box in an asymptotically Minkowski background. These are regular, static, horizon-free solutions to Einstein-Maxwell theory coupled to a scalar field that vanishes outside a box but not inside it (the latter has an Israel stress

tensor that supports the pressure of the scalar field and keeps the solution stationary).

As discussed in the perturbative analysis of [43], the properties of charged hairy solutions on a box depend on the charge of the scalar field e. In particular, one identifies the following four critical scalar charges (two of them, e_γ and e_c, are not captured within perturbation theory [43]):

- $e = e_{\rm NH} = \frac{1}{2\sqrt{2}} \sim 0.354$. This is the charge above which scalar fields can trigger a violation of the near horizon AdS_2 Breitenlöhner-Freedman bound of the extremal RN black hole whose horizon radius approaches, from below, the box radius. This violation renders (near-)extremal RN BHs unstable. This is the so-called near horizon scalar condensation instability first studied in the context of AdS black holes and the holographic gravity/condensed matter correspondence programme [53, 37, 38, 56]. We ask the reader to see Section III.B of [40] for a detailed analysis that gets this critical charge. As far as we could perceive, this charge does not play a relevant role in the discussion of the solitons of the theory.

- $e = e_{\rm S} = \frac{\pi}{\sqrt{2}} \sim 2.221$. This is the critical charge that saturates the superradiant bound $\omega = e\mu$ for an extremal RN BH (which has $\mu = \sqrt{2}$) when we take the frequency to be the lowest frequency, $\omega = \pi$, that can fit inside a box of (dimensionless) unit radius in Minkowski spacetime [35, 36]. We ask the reader to see the introduction for a derivation of this critical charge.

- $e = e_\gamma$ and $e = e_c$ with $e_{\rm NH} < e_\gamma < e_c < e_{\rm S}$. In the present part we find that the system has a third and a fourth critical charges, that we find within numerical error to be $e_\gamma \sim 1.13$ and $e_c \simeq 1.8545 \pm 0.0005$ (unlike the other two charges, we are not aware of a heuristic analysis that allows to capture analytically these two critical values without solving the full equations of motion). These charges are not captured by the perturbative analysis of [43]. We will find that our system has (at least) two distinct families of ground state solitons (there is then an infinite tower of excited soliton families that, in the perturbative regime, correspond to the backreaction of excited normal modes with higher radial overtones). One – that we call the *main* soliton family – can be seen as the backreaction of the charged ground state normal mode of a scalar field in a Minkowski box [43] since it exist for small energies/charges up to a Chandrashekhar limit. On the other hand, the second family – that we denote as the *secondary* soliton family – exists only for intermediate or large energies/charges and thus it is not captured in the perturbative analysis of [43]. The secondary soliton family exists only for $e_\gamma < e < e_c$. In a phase diagram of soliton solutions, as we approach the critical charge e_γ from above, the secondary soliton family ceases to exist because it no longer fits inside the box. On the other hand, as we approach the critical charge e_c from below, the main and secondary

families of solitons approach each other and they connect precisely at $e = e_c$.

In the present chapter, we solve numerically the full Einstein-Maxwell equations for a charged scalar field confined in a covariant box with dimensionless radius $R = 1$. Therefore, our analysis is now fully nonlinear and not restricted to small energies/charges. We focus our attention on soliton solutions (and leave the study of hairy BHs for the next chapter, this work was also published in [3]) since the phase diagram of these solutions is already very intricate. We will recover the main soliton family of solutions of [43] for small energies/charges (which exists for any $e > 0$), thereby confirming and setting the regime of validity of the perturbative analysis, but also solutions with intermediate and large energies with a Chandrasekhar mass limit and multi-branched structure not captured by [43]. In particular, we will find the secondary family of solitons and reveal the existence of the critical charges e_γ and e_c.[1]

By Birkhoff's theorem, outside the box (in particular, in the asymptotic region), our solutions are necessarily described by the RN solution [33, 34].[2] It follows that we cannot use the Arnowitt-Deser-Misner (ADM) mass M and charge Q (measured by a Gauss law at infinity [7]) to distinguish the several solutions of the theory. Instead, we need to resort to the Brown-York quasilocal mass $\mathcal{M}$ and charge $\mathcal{Q}$ [112], measured at the box location, to display our solutions in a phase diagram of the theory. These quantities satisfy their own first law of thermodynamics that we use to (further) check our solutions. Dimensionless quasilocal mass and charge are given in units of the box radius L, $\mathcal{M}/L$ and $\mathcal{Q}/L$, respectively. A natural reference in this quasilocal phase diagram is the extremal RN BH 1-parameter family with horizon inside the box. However, above a certain scalar charge e, the solitons have a mass/charge that can be very close to the ones of the extremal RN. For this reason we will find useful to plot $\Delta\mathcal{M}/L$ *vs* $\mathcal{Q}/L$ where $\Delta\mathcal{M} = \mathcal{M} - \mathcal{M}\big|_{\text{ext RN}}$ is the mass difference between the hairy solution and the extremal RN that has the *same* $\mathcal{Q}/L$. So, in this phase diagram the horizontal line with $\Delta\mathcal{M} = 0$ identifies the extremal RN BH solution. Its dimensionless horizon $R_+ = r_+/L$ can fit inside the box of radius L if $R_+ \leq 1$ (which corresponds to $\mathcal{Q}/L \leq 2^{-1/2}$) and non-extremal RN BHs exist above this line. But the horizon of non-extremal RN BHs fit inside the box ($R_+ \leq 1$) only if these solutions are to the left of the red dashed line that will be displayed in our plots. We will find that this line represents the maximal quasilocal charge that solutions that fit inside the box can have (with or without scalar condensate).

[1] The perturbative analysis of [43] also finds that hairy BHs that are perturbatively connected to a boxed Minkowski spacetime exist only for $e \geq e_{\text{S}}$; by construction, the zero horizon radius limit of such hairy BHs is one of our hairy solitons. Perturbative theory [43] does not capture the existence of hairy BHs with $e < e_{\text{S}}$. In the next chapter we will show that a full nonlinear analysis finds that hairy BHs also exist in the range $e_{\text{NH}} \leq e < e_{\text{S}}$ but, typically, they are no longer necessarily connected to the solitons of the theory in their zero entropy limit.

[2] Recall: Birkhoff's theorem for Einstein-Maxwell theory states that the unique spherically symmetric solution of the Einstein-Maxwell equations with non-constant area radius function r (in the gauge (4.3.1)) is the Reissner-Nordström solution. If r is constant then the theorem does not apply since one has the Bertotti-Robinson ($AdS_2 \times S^2$) solution.

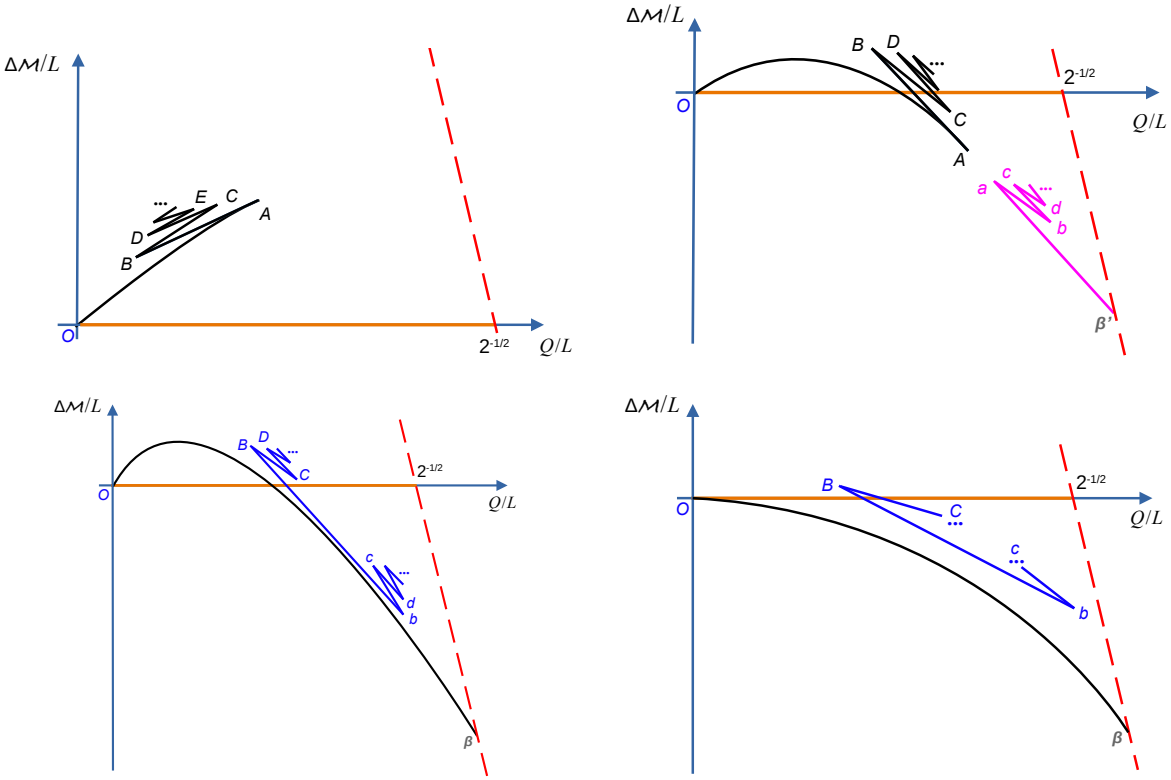

Figure 3.2.1: Sketch of the quasilocal phase diagram for solitons as we span relevant windows of scalar field charge e. The critical charges are such that $0 < e_{\rm NH} < e_c < e_{\rm S}$. The quantity $\Delta\mathcal{M}$ is the quasilocal mass difference between the soliton and an extremal RN BH that has the *same* quasilocal charge $\mathcal{Q}/L$. Hence the orange line at $\Delta\mathcal{M} = 0$ describes the extremal RN solution that must have $\mathcal{Q}/L \leq 2^{-1/2}$ to fit inside the box. The red dashed line represents the maximal quasilocal charge of solutions that can fit inside the box. It intersects the extremal RN line at $\mathcal{Q}/L = 2^{-1/2}$. Non-extremal RN BHs confined in the box have $\Delta\mathcal{M} > 0$ and fill the triangular region bounded by $\mathcal{Q} = 0$ and by the orange and red dashed lines. The main soliton family is always given by black curves that start at O. The secondary soliton family is given either by magenta or blue curves. **Top-left panel:** case $e < e_\gamma$. **Top-right panel:** case $e_\gamma < e < e_c$. **Bottom-left panel:** case $e_c < e < e_{\rm S}$. **Bottom-right panel:** case $e > e_{\rm S}$.

Using this phase diagram, a summary of our main findings is (see sketch in Fig. 3.2.1):

1. $e < e_\gamma \sim 1.13$. Later, we will give data for the case $e = 0.23$ (section 3.4.1; Figs. 3.4.1-3.4.3). Solutions with $e < e_\gamma$ are qualitatively similar and blind to the critical charge $e_{\rm NH}$. Here, we *sketch* a phase diagram that highlights the main properties of these solutions in the top-left panel of Fig. 3.2.1. We simply have the '*main soliton family*' (or '*perturbative soliton family*' represented by the black line $OABCDE\cdots$). Nearby O, the properties of the branch OA for small charge were already captured by the perturbative analysis of [43] but this soliton then develops an intricate series of cusps $A, B, C, D, E, \cdots$ that could not be anticipated by the analysis of [43]. In particular, this main soliton family has a Chandrasekhar limit at A. As we move from

the main branch OA, along the sequence of secondary zig-zagged branches AB, BC, CD, ..., we find that the Kretschmann curvature invariant at the origin of the soliton is growing without bound. For this reason, it becomes increasingly hard to follow this solution beyond a certain point (say, point E) but we gathered enough evidence to predict that this solution might well develop an infinite number of cusps as it approaches a singular limit where the curvature invariants blow up. At a physical level, nothing wrong happens at the cusps. Actually, these cusps are simply smooth turning points if we plot $\mathcal{M}/L$, $\mathcal{Q}/L$, or the values of the gravitational field $f \equiv g_{tt}$, electric potential A_t or scalar field ϕ at the origin as a function of the derivative $\epsilon \equiv \phi'(R=1)$ of the scalar field at the box radius (see Appendix). Depending on which of these functions we look at we can have a curve with damped oscillations or a spiral curve with a series of turning points $A, B, C, D, E, \cdots$ (see e.g. Figs. 3.4.2-3.4.3 for $e = 0.23$).

As e increases from $e = 0$ to $e = e_\gamma \sim 1.13$, the qualitative features of the solutions do not change significantly. These main solitons always have more quasilocal mass than the extremal RN BH with the same quasilocal charge and their slope in the $\mathcal{Q}$-$\Delta\mathcal{M}$ is positive but decreases as e grows. This discussion is best illustrated in Fig. 3.2.2 where we plot the main soliton family for different values of the electric scalar field charge e. In particular, we give three cases, $e = 0.23, 0.5$ and $e = 1$ that have $e < e_\gamma$.

2. $e_\gamma < e < e_c \simeq 1.8545 \pm 0.0005$. Later, we will give data for the case $e = 1.854$ (section 3.4.2; Figs. 3.4.4-3.4.6). Here, we summarize the analysis with a *sketch* of the phase diagram in the top-right panel of Fig. 3.2.1, that emphasizes the main properties of these solutions. As for $e < e_\gamma$, we have the '*main soliton family*' (or '*perturbative soliton family*' represented by the black line $OABCDE\cdots$). But, unlike for $e < e_\gamma$, after a gap Aa in $\mathcal{Q}/L$, we now also have a '*secondary soliton family*' $\beta' abcd\cdots$ (that we might also call the '*non-perturbative soliton family*').

The main soliton family $OABCDE\cdots$ for $e_\gamma < e < e_c$ has similar properties to the ones already found in the $e < e_\gamma$ case. The only minor difference is that, as e grows well above e_γ, at a certain point the main soliton family bends downwards as $\mathcal{Q}/L$ grows. And at a certain critical charge, $e \sim 0.5$, we can have portions of the main soliton in the neighbourhood of its Chandrasekhar point A with less quasilocal mass than the extremal RN BH with same $\mathcal{Q}/L$: see the sequence of solitons in Fig. 3.2.2 for different values of e in this range.

The top-right panel of Fig. 3.2.1 also sketches the secondary family $\beta' abcd\cdots$ of solitons (magenta line). This family only exists above a critical charge $\mathcal{Q}/L$ (point a) that depends on e and extends all the way up to the maximal charge that can

fit inside the box (represented by the point β' on the red dashed line). Thus, this family is not captured by the perturbative analysis of [43]. As we move from β' to cusp a and continue through the cusps $b, c, d, \cdots$ one finds that the Kretschmann curvature at the origin of the soliton is growing without bound and we find evidence that the solution will terminate on a singular solution where the curvature diverges, probably after going through an infinite number of cusps.

For scalar charge $e = e_\gamma$, point a coincides with point β' (see top-right panel of Fig. 3.2.1). Just above e_γ, point a is very close to β' and the gap Aa between the main and secondary solitons is very large. The secondary soliton that we found with smallest e has $e = 1.144$. We take this and the findings summarized in the right plot of Fig. 3.2.2 to extrapolate that $e_\gamma \sim 1.13$. For our purposes it is not necessary to determine e_γ with higher accuracy. As e keeps increasing away from e_γ, the gap Aa decreases very quickly and, quite importantly, it goes to zero precisely at $e = e_c$ where the main and secondary solitons merge. This evolution of the $\mathcal{Q}$-$\Delta\mathcal{M}$ phase diagram with e is best illustrated in the right panel of Fig. 3.2.2 where we display the main and secondary family of solitons for different values of e in the window $e_\gamma < e < e_c$. Note that, consistent with the description given above, secondary solitons exist only in the $\mathcal{Q}$-$\Delta\mathcal{M}$ region bounded by the closed curve $a_c\beta_c\gamma$ (see auxiliary dashed gray curve of the right panel of Fig. 3.2.2). Just above e_γ, *e.g.* for $e = 1.15$, the secondary soliton family is described by a very short segment (red diamonds) very close to point γ (*i.e.* β' is close to γ). That is to say, secondary solitons with $e < e_\gamma$ do not exist because they would not fit inside the box with radius $R = 1$ (the red dashed line $\beta_c\gamma$). On the other hand just below e_c, *e.g.* $e = 1.854$, the main soliton (black disks) family is very close to the secondary soliton family (magenta triangles that almost coincide with the limiting curve $a_c\beta_c$, *i.e.* $\beta' \to \beta_c$). Precisely at $e = e_c$, these two families merge at point $A \equiv a \equiv a_c$.

3. $e_c < e < e_s$. Later, we will give data for the case $e = 1.855$ that fits in this window (section 3.4.3; Figs. 3.4.7-3.4.9). Precisely at $e = e_c$ we find that the previously discussed main and secondary soliton families merge: point A of top-right panel of Fig. 3.2.1 merges with point a at $e = e_c$. This happens in a curious way since above e_c the branch OA merges to the branch $a\beta'$ (of the top-right panel) and they now form the main soliton family of solitons that, this time, is cusp-free: this is the black line curve $O\beta$ in the bottom-left panel of Fig. 3.2.1. On the other hand, the old (*i.e.* top-right panel) sequence of secondary branches/cusps $ABCD\cdots$ of the main soliton family is now connected to the old sequence of secondary branches/cusps $abcd\cdots$ of the secondary soliton family: this is the blue line $\cdots DCBbcd\cdots$ in the bottom-left panel of Fig. 3.2.1. So, precisely at $e = e_c$, one has 4 families of solitons "bifurcating" from the merger $A \equiv a$, *i.e.* the black and blue curves of the bottom-left panel intersect at $A \equiv a$ (equivalently, approaching e_c from below,

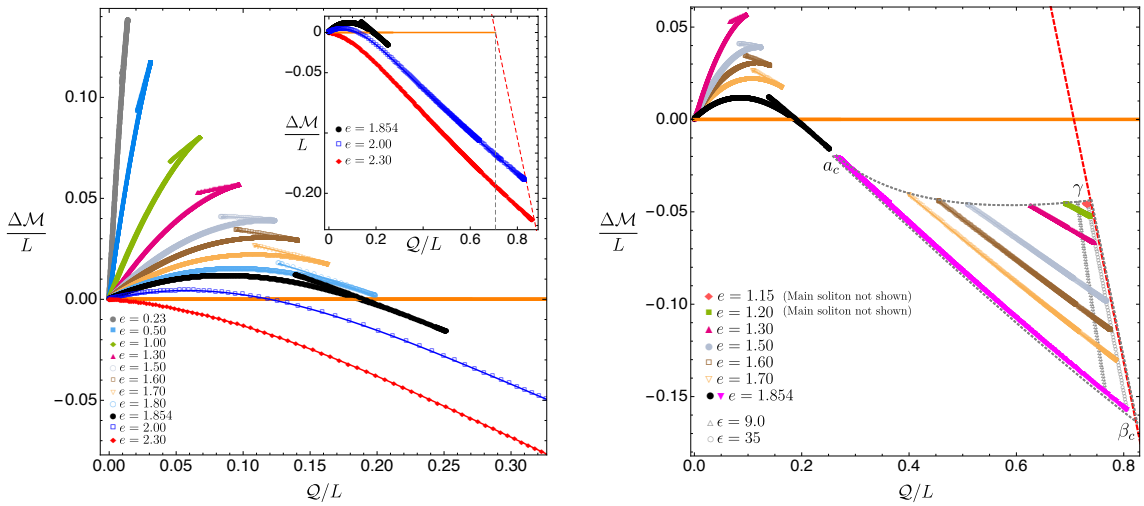

Figure 3.2.2: Quasilocal phase diagram for solitons with different scalar condensate electric charges. The quantity $\Delta\mathcal{M}$ is the mass difference between the soliton and an extremal RN BH that has the *same* charge $\mathcal{Q}/L$. Hence the orange line at $\Delta\mathcal{M} = 0$ describes the extremal RN solution. The dashed vertical line in the inset plot describes the line $\mathcal{Q} = 2^{-1/2}$ which is the maximum local charge that an extremal RN BH can have whilst fitting inside a box with radius $R = 1$. The red dashed line represents the maximal quasilocal charge of solutions that can fit inside the box. **Left panel:** here we give several examples of the main soliton family only. Notice the change in appearance of the curves as we increase the scalar field charge above $e_c \simeq 1.8545$ (see also inset plot). **Right panel:** here we concentrate our attention only in cases with $e_\gamma < e < e_c$ but, this time, we present both the main and secondary solitons. The secondary solitons only exist in the region bounded by the auxiliary gray dashed closed curve $a_c\beta_c\gamma$. The secondary soliton curve with e just above e_γ is close to point γ, while the soliton with $e = 1.854$, just below e_c, is the magenta curve (very close to $a_c\beta_c$). Note that the gap in $\mathcal{Q}/L$ between the main soliton and the secondary one starts very large at $e = e_\gamma$ but then decreases and goes to zero precisely at $e = e_c$. We also present a secondary soliton curve with constant scalar field condensate, $\epsilon = 35$, to illustrate how we can march in e to obtain different solutions and further conclude that secondary solitons are bounded by the auxiliary lines $a_c\gamma$ and $a_c\beta_c$.

the black and magenta curves of the top-right panel intersect at $A \equiv a$). Then, as we increase e above e_c the black and blue curves disconnect and their distance increases as e grows. This merging and interactions between the black/magenta and black/blue curves as we approach $e = e_c$ from bellow/above is observed not only on the quasilocal phase diagram but also undoubtedly confirmed when we analyse *e.g.* the plots for the Kretchsmann curvature invariant at the origin, $K\big|_{R=0}$, as a function of the scalar condensate $\epsilon \equiv \phi'(R = 1)$: compare latter Fig. 3.4.6 (for $e = 1.854 \lesssim e_c$) with Fig. 3.4.9 (for $e = 1.855 \gtrsim e_c$).

The sharp transition of the properties of the main soliton family when e changes from a value below e_c into one above e_c is also illustrated in Fig. 3.2.2 where we

plot the main soliton family for different values of the electric scalar field charge e. In this plot the black disk curve has $e = 1.854 < e_c$ while the blue square ($e = 2$) and red diamond ($e = 2.3$) curves have $e > e_c$. The inset plot of this figure clearly shows that the Chandrasekhar limit seen for $e < e_c$ is no longer present for $e > e_c$.

4. $e \geq e_{\rm s} = \frac{\pi}{\sqrt{2}} \sim 2.221$. Later, we will give data for the case $e = 2.3$ that fits in this window (section 3.4.4; Figs. 3.4.10-3.4.12). Solitons with charges in this range have properties that are very similar to those displayed by solitons with $e_c < e < e_{\rm s}$. In particular, as sketched in the bottom-right panel of Fig. 3.2.1, we still have similar main soliton (black curve $O\beta$) and secondary soliton (blue curve $\cdots CBbc\cdots$) families (these curves become more separated apart as e grows). But charges in this range also leave their unique footprint. Indeed, the main difference that justifies distinguishing the cases $e_c < e < e_{\rm s}$ and $e \geq e_{\rm s}$ is that, for the latter case, the main soliton family $O\beta$ always has smaller quasilocal mass than the extremal RN BH with the same charge, no matter how small $\mathcal{Q}/L$ is. That is to say, these solitons always have $\Delta\mathcal{M} < 0$, with the slope of the black curve $O\beta$ at $\mathcal{Q} = 0$ being zero precisely at $e = e_{\rm s}$ and negative for $e > e_{\rm s}$ (so, this slope at the origin for $e < e_{\rm s}$ is positive). Moreover, for $e \geq e_{\rm s}$ and only in this case, extremal RN BHs are unstable all the way down to $\mathcal{Q} = 0$ [40]. These two facts suggest that for this regime of e, as confirmed in the perturbative analysis of [43], the zero horizon radius limit of the hairy black holes of the theory should be our soliton and thus they can be constructed perturbatively for arbitrarily small scalar condensate amplitude and horizon radius.

   The secondary family of solitons (the magenta curve $\cdots CBbc\cdots$ in the bottom-right plot of Fig. 3.2.1) still exists for $e > e_{\rm s}$ but it is further separated from the main soliton family (black curve $O\beta$). It becomes increasingly much more difficult to follow this family beyond the cusps B, b for values of $e > e_c$. So it is hard to explore more energetic soliton families for values of e much larger than e_c. However, one real possibility (we found evidence for this) is that there is a sequence of critical charges $e_{c_2}, e_{c_3}\cdots$. Here, e_{c_2} would be the charge above which the cups C and c would merge (very much like the cusps A and a merged at $e = e_c \sim 1.8545$) and form a *closed* line family in the phase diagram ($CBbc$ with $c \equiv C$) together with a *third* open family of solitons that would have higher mass and that would correspond to the extension to higher e of the open curve $\cdots DCcd\cdots$ in the bottom-left plot of Fig. 3.2.1. If so, as the charge grows we could have not only two ground state families of solitons but a sequence of them: first 3 families (one of them closed), then 4 families (two of them closed), etc. For this reason we leave the curve $\cdots Bb\cdots$ incomplete a bit beyond the cusps B and b in the bottom-right panel of Fig. 3.2.1.

Once we have found the hairy solitons that are confined inside the Minkowski box we can study the properties that the box must have to support the pressure exerted by the scalar condensate that is inside it. As mentioned above, the box must have an Israel stress tensor (that accounts for the extrinsic curvature jump at the box) that prevents the scalar field from expanding all the way to the asymptotic region where it would die off. In section 3.5, we compute this Israel stress tensor and find that there are boxes with an energy density and pressure that obey several or all forms of the energy conditions [5]. We will also compute the ADM mass and charge of our solitons as measured at the asymptotic boundary [7]. These ADM charges include the contributions from the hairy soliton and the box and we will find that, quite often, the ADM mass can be negative. Recall that Schwarzschild and RN BHs with negative ADM mass are singular solutions, but our solitons with negative mass are regular everywhere (except at the box outer layer).

Many of the above physical features observed in boson stars (solitons) confined in a box in an asymptotically flat background are similar to those observed in asymptotically anti-de Sitter solitons [110] (see also [46, 45, 44, 93, 47, 111, 39]). In this case, the AdS boundary conditions act as a natural gravitational box with radius inversely proportional to the cosmological length that provide confinement. Therefore, we have good reasons to expect that many of the features that we identify in the present study are universal properties of charged scalar fields subject to some sort of confining mechanism.

3.3 Einstein-Maxwell gravity with a confined scalar field

3.3.1 Theory and setup

For completeness let us repeat the action for Einstein-Maxwell gravity in four dimensions coupled to a charged scalar field:

$$S=\frac{1}{16\pi G_N}\int \mathrm{d}^4x\sqrt{-g}\left(\mathcal{R}-\frac{1}{2}F_{\mu\nu}F^{\mu\nu}-2D_\mu\phi(D^\mu\phi)^\dagger+V(|\phi|)\right), \tag{3.3.1}$$

where $\mathcal{R}$ is the Ricci scalar, A is the Maxwell gauge potential, $F=\mathrm{d}A$, and $D_\mu=\nabla_\mu-iqA_\mu$ is the gauge covariant derivative of the system. We consider the potential $V(|\phi|)=m^2\phi\phi^\dagger$ with m the mass of the scalar field. For concreteness we will take $m=0$ but solitons with $m>0$ should also exist. We fix Newton's constant $G_N\equiv 1$.

We are interested in solitonic solutions of (3.3.1) that are static, spherically symmetric and asymptotically flat. Recall the ansatz we made for our fields in the introduction:

$$\mathrm{d}s^2=-f(r)\mathrm{d}t^2+g(r)\mathrm{d}r^2+r^2\mathrm{d}\Omega_2^2,\qquad A_\mu\mathrm{d}x^\mu=A_t(r)\mathrm{d}t,\qquad \phi=\phi^\dagger=\phi(r), \tag{3.3.2}$$

Recall also the dimensionless coordinates where thermodynamic quantities are mea-

sured in units of the box radius L,

$$T = \frac{t}{L}, \;\; R = \frac{r}{L}, \;\; R_+ = \frac{r_+}{L} \;\; e = qL, \;\; \tilde{m} = mL \tag{3.3.3}$$

In these units the box is placed at $R = 1$. The equations of motion for the fields $f(R)$, $A_t(R)$, $g(R)$ and $\phi(R)$ obtained from extremising the action (3.3.1) with $m = 0$ can be found in the introduction, equations (1.4.24). We have a system of three second order ODEs for f, A_t, ϕ and g is given by an algebraic relation of the other three functions (1.4.28). To have a well-posed boundary value problem we need to specify the boundary conditions at the origin and asymptotic boundary of our spacetime. Moreover we must also impose junction conditions at the timelike hypersurface Σ at $R = 1$ where the box is located. Our solitons have vanishing scalar field at and outside this box, $\phi(R \geq 1) = 0$.

The system is described by three second order ODEs which means that there are six arbitrary integration constants when we do a Taylor expansion around the origin, $R = 0$. Regularity, requires that we impose Dirichlet boundary conditions whereby we set three of the above integration constants to zero in order to eliminate terms that would diverge at this boundary [104]. We are thus left with only three constants f_0, A_0, ϕ_0 (say) such that the regular fields have the Taylor expansion around the origin:

$$f(0) = f_0 + \mathcal{O}(R^2), \qquad A_t(0) = A_0 + \mathcal{O}(R^2), \qquad \phi(0) = \phi_0 + \mathcal{O}(R^2). \tag{3.3.4}$$

Take now the asymptotic boundary of our spacetime, $R \to \infty$. Outside the box the scalar field vanishes, $\phi = 0$, and the solutions of the equation of motion are: $f^{out}(R) = c_f - \frac{M_0}{R} + \frac{\rho^2}{2R^2}$, $A_t^{out}(R) = c_A + \frac{\rho}{R}$ and $g^{out}(R) = c_f / f^{out}(R)$ (henceforward, the superscript ^{out} represents fields outside the box). Here, c_f, M_0, c_A and ρ are arbitrary integration constants which are not constrained, *i.e.* we have an asymptotically flat solution for any value of these constants. But the theory has a second scaling symmetry,

$$\{T, R, x, \varphi\} \to \{\lambda_2 T, R, x, \varphi\}, \quad \{f, g, A_t, \varphi\} \to \{\lambda_2^{-2} f, g, \lambda_2^{-1} A_t, \varphi\}, \quad \{e, R_+\} \to \{e, R_+\}, \tag{3.3.5}$$

that we use to set $c_f = 1$ so that $f|_{r\to\infty} = 1$ (and $g^{out} = 1/f^{out}$). Outside the box the solution to the equations of motion is then

$$f^{out}(R)\big|_{R\geq 1} = 1 - \frac{M_0}{R} + \frac{\rho^2}{2R^2}\,, \qquad A_t^{out}(R)\big|_{R\geq 1} = c_A + \frac{\rho}{R}\,, \qquad \phi^{out}(R)\big|_{R\geq 1} = 0\,, \tag{3.3.6}$$

which is the Reissner-Nordström solution as required by Birkhoff's theorem for the Einstein-Maxwell theory [33, 34]. However, (3.3.6) has three free integration constants, M_0, c_A, ρ, which will be determined only after we have the solution inside the box.

Our solutions are asymptotically flat. Therefore, some of the parameters in (3.3.6) are

related to the ADM conserved charges [7]. Namely, the adimensional ADM mass and electric charge of the system are given by (setting $G_N \equiv 1$):[3]

$$\begin{aligned} M/L &= \lim_{R\to\infty} \frac{R^2 f'(R)}{2\sqrt{f(R)g(R)}} = \frac{M_0}{2}, \\ Q/L &= \lim_{R\to\infty} \frac{R^2 A_t'(R)}{2f(R)g(R)} = -\frac{\rho}{2}. \end{aligned} \tag{3.3.7}$$

These ADM conserved charges measured by a Gauss law at the asymptotic boundary include the contribution from the energy-momentum content of the box that confines the scalar hair.

In these conditions, solitons of the theory are a 1-parameter family of solutions that we can take to be, *e.g.* f_0 as defined in (3.3.4), or the value of the (interior) derivative of the scalar field at the box, $\epsilon \equiv \phi'^{\,in}\big|_{R=1}$ as discussed in the next subsection.

As mentioned in section 3.2, it follows from Birkhoff's theorem that in the asymptotic region our solutions are necessarily described by the RN solution and we cannot use the ADM mass M and charge Q to distinguish the several solutions of the theory. Instead, we need to resort to the Brown-York quasilocal mass $\mathcal{M}$ and charge $\mathcal{Q}$, measured at the box to display our solutions in a phase diagram of the theory [112]. From section II.C of [43] (which we ask the reader to visit for details), the Brown-York quasilocal mass and charge contained inside a 2-sphere with radius $R = 1$ are ($G_N \equiv 1$)

$$\begin{aligned} \mathcal{M}/L &= R\left(1 - \frac{1}{\sqrt{g(R)}}\right)\Big|_{R=1}. \\ \mathcal{Q}/L &= \frac{R^2 A_t'(R)}{2\sqrt{g(R)}f(R)}\Big|_{R=1}. \end{aligned} \tag{3.3.8}$$

To complete the thermodynamic description of our solutions we still need to define the chemical potential of the soliton which is given by the value of the gauge potential at the box,

$$\mu = A_t\big|_{R=1}, \tag{3.3.9}$$

and the quasilocal quantities must satisfy the quasilocal first law of thermodynamics (see for example [43] and references therein):

$$\mathrm{d}\mathcal{M} = \mu\, \mathrm{d}\mathcal{Q}\,. \tag{3.3.10}$$

We will use this law as a non-trivial check of our solutions.

[3] Note that the Maxwell term in action (3.3.1) is $\frac{1}{2}F^2$, not the perhaps more common F^2 term. It follows that the extremal RN BH satisfies the ADM relation $M = \sqrt{2}|Q|$, instead of $M = |Q|$ that holds when the Maxwell term in the action is F^2.

As explained before, for reference we will often compare the soliton families of solutions against extremal RN BHs. RN BHs confined in a box can be parametrized using the dimensionless horizon radius R_+ and the chemical potential μ, and their quasilocal mass and charge are [43]

$$\mathcal{M}/L\big|_{RN} = 1 - \frac{\sqrt{2}(1-R_+)}{\sqrt{2-(2-\mu^2)R_+}}, \qquad \mathcal{Q}/L\big|_{RN} = \frac{\mu R_+}{\sqrt{2}\sqrt{2-(2-\mu^2)R_+}}. \tag{3.3.11}$$

where $0 < R_+ \leq 1$ (for the horizon to be confined inside the box) and $0 \leq \mu \leq \mu_{\rm ext}$, with extremality reached at $\mu_{\rm ext} = \sqrt{2}$. Note that at extremality one has $\mathcal{M}/L = R_+$ and $\mathcal{Q}/L = R_+/\sqrt{2}$. On the other hand, for any μ, when $R_+ = 1$ one has $\mathcal{M}/L = 1$ and $\mathcal{Q}/L = 2^{-1/2}$.

3.3.2 Junction conditions and Israel stress tensor at the box

Above, we discussed the boundary conditions at the origin and asymptotic boundaries. However, solitons are solutions that glue an interior spacetime ($R < 1$; with superscript ^{in}) with the known RN exterior background solution (3.3.6) ($R > 1$; with superscript ^{out}) . So all we need to do is to find the interior solution. But for that we must specify appropriate physical conditions at the outer boundary of our numerical integration domain, namely at $R = 1$.

Recall from the introduction that we require the scalar field to vanish at and outside the box, *i.e.* for $R \geq 1$ (by Birkhoff's theorem), hence we defined the quantity ϵ:

$$\phi^{in}\big|_{R=1} = \phi^{out}\big|_{R=1} = 0, \qquad \phi^{out}(R) = 0, \qquad \phi'^{\,in}\big|_{R=1} \equiv \epsilon, \tag{3.3.12}$$

i.e. for $R \leq 1$ the scalar field is forced to have the Taylor expansion $\phi\big|_{R=1^-} = \epsilon(R-1) + \mathcal{O}(R-1)^2$. We are forcing a jump in the derivative of the scalar field normal to the cavity time-like hypersurface Σ. The latter is defined by $\mathfrak{f}(R) = R-1 = 0$ and has outward unit normal $n_\mu = \partial_\mu \mathfrak{f}/|\partial \mathfrak{f}|$ ($n_\mu n^\mu = 1$). Naturally, this forcing condition on the scalar field has consequences: we need to impose junction conditions at Σ on the other fields. Ideally, we would like to have a smooth crossing, whereby the gravitational and gauge fields and their normal derivatives are continuous at Σ. But this is not possible when we have a non-vanishing scalar field inside the box. It follows that the Israel junction conditions require a non-vanishing jump in the extrinsic curvature across the box.

It is a good idea to set some notation to discuss this issue further. Adopting the viewpoint of an observer in the interior region, the layer surface Σ is parametrically described by $R = 1$ and $T = T^{in}(\tau) = \tau$ and the induced line element and gauge 1-form of the shell Σ

read

$$\begin{aligned} \mathrm{d}s^2|_{\Sigma^{in}} &= h^{in}_{ab}\,\mathrm{d}\xi^a\mathrm{d}\xi^b = -f^{in}|_{R=1}\mathrm{d}\tau^2 + \mathrm{d}\Omega_2^2\,, \\ A_t|_{\Sigma^{in}} &= a^{in}_a\mathrm{d}\xi^a = A^{in}_t|_{R=1}\mathrm{d}\tau\,, \end{aligned} \tag{3.3.13}$$

where ξ^a describe coordinates in Σ, h^{in}_{ab} is the induced metric in Σ and a^{in}_a is the induced gauge potential in Σ. On the other hand, from the perspective of an observer outside the cavity, Σ is parametrically described by $R = 1$ and $T = T^{out}(\tau) = N\tau$ (so, N is a reparametrization freedom parameter) so that the induced line element and gauge 1-form are

$$\begin{aligned} \mathrm{d}s^2|_{\Sigma^{out}} &= h^{out}_{ab}\,\mathrm{d}\xi^a\mathrm{d}\xi^b = -N^2 f^{out}|_{R=1}\mathrm{d}\tau^2 + \mathrm{d}\Omega_2^2\,, \\ A_t|_{\Sigma^{out}} &= a^{out}_a\mathrm{d}\xi^a = NA^{out}_t|_{R=1}\mathrm{d}\tau\,, \end{aligned} \tag{3.3.14}$$

The junction conditions required to join smoothly two backgrounds at a timelike hypersurface Σ were studied by Israel [113, 114, 115, 116] built on previous work of Lanczos and Darmois. A solution is smooth at Σ if and only if: 1) the induced metric h_{ab} and induced gauge potential a_a are continuous (*i.e.* $\mathrm{d}s^2|_{\Sigma^{in}} = \mathrm{d}s^2|_{\Sigma^{out}}$ and $A|_{\Sigma^{in}} = A|_{\Sigma^{out}}$), and 2) the extrinsic curvature K_{ab} (essentially the normal derivative of the induced metric) and the normal derivative of the induced gauge field, f_{aR}, are continuous. Denoting, as we have been doing, the solution inside (outside) Σ by the superscript ^{in} (^{out}), the Israel junction conditions are

$$a^{in}_a\big|_{R=1} = a^{out}_a\big|_{R=1}\,, \tag{3.3.15a}$$

$$h^{in}_{ab}\big|_{R=1} = h^{out}_{ab}\big|_{R=1}\,; \tag{3.3.15b}$$

$$f^{in}_{aR}\big|_{R=1} = f^{out}_{aR}\big|_{R=1}\,, \tag{3.3.15c}$$

$$K^{in}_{ab}\big|_{R=1} = K^{out}_{ab}\big|_{R=1}; \tag{3.3.15d}$$

where $h_{ab} = g_{ab} - n_a n_b$ is the induced metric at Σ and $K_{ab} = h_a{}^c\nabla_c n_b$ is the extrinsic curvature.

In the absence of the scalar condensate, we can set $N = 1$ and all the junction conditions (3.3.15) are satisfied. However, our hairy solitons are continuous but not differentiable at $R = 1$. Namely, they satisfy the 3 conditions (3.3.15a)-(3.3.15c) but not (3.3.15d). Since the latter extrinsic curvature condition is not obeyed, our hairy solitons are singular at Σ. But this singularity simply signals the presence of a Lanczos-Darmois-Israel surface stress tensor $\mathcal{S}_{ab}$ at the hypersurface layer proportional to the difference of the extrinsic curvature across the hypersurface. This Lanczos-Darmois-Israel surface stress tensor

induced in Σ is [113, 114, 115, 116]

$$\mathcal{S}_{ab} = -\frac{1}{8\pi}\Big([K_{ab}] - [K]h_{ab}\Big), \tag{3.3.16}$$

where K is the trace of the extrinsic curvature and $[K_{ab}] \equiv K^{out}_{ab}\big|_{R=1} - K^{in}_{ab}\big|_{R=1}$. This surface tensor is the pull-back of the energy-momentum tensor integrated over a small region around the hypersurface Σ *i.e.* it is obtained by integrating the appropriate Gauss-Codazzi equation [113, 114, 115, 116, 117] and it is also given by the difference between the Brown-York surface tensor just outside and inside the surface layer [112] (see also discussion in [43]). Essentially, (3.3.16) describes the energy-momentum tensor of the cavity (the "internal structure" of the box) that we have to build to confine the scalar field inside. With our explicit construction of the hairy solutions of the system we will be able to compute this Lanczos-Darmois-Israel stress tensor. Note that since the two Maxwell junction conditions (3.3.15a)-(3.3.15b) are obeyed, our solitons will have a surface layer that has no electric charge.

Our strategy is now clear. To find the soliton solution inside the box, we integrate numerically the Einstein equation in the domain $R \in [0, 1]$ subject to the boundary conditions (3.3.4) at the origin and, at the box, we impose $\phi(1^-) = 0$ and use the scaling symmetry (3.3.5) to set $f(1^-) = 1$. With this information we can already read univocally the quasilocal charges (4.3.3) of the system. We then impose the three junction conditions (3.3.15a)-(3.3.15c) at the box to match the interior solution with the outer solution described by the RN solution (3.3.6). This allows to find the parameters M_0, C_A, ρ in (3.3.6) as a function of the reparametrization freedom parameter N introduced in (3.3.14). The Israel stress tensor $\mathcal{S}^b_a$ is just a function of N and, if $\phi^{in} \neq 0$, there is no choice of N that kills all the components of $\mathcal{S}^b_a$ (there are two non-vanishing components, $\mathcal{S}^t_t$ and $\mathcal{S}^\theta_\theta = \mathcal{S}^\varphi_\varphi$). The fact that we have arbitrary freedom to select N simply reflects the freedom we have in the choice of the energy-momentum content of the box needed to contain the scalar condensate inside it. We will show that there are choices that preserve some or all the energy conditions [5]. Once we make a choice for N, we can also compute the ADM mass and charge (3.3.7) of the solution which includes the contribution from the box.

3.3.3 Numerical schemes

The solitons we search for are a 1-parameter family of solutions. We will generate these solutions numerically following one of two routes that differ on the choice made for the marching parameter along the family: 1) we march varying the value of the scalar condensate quantity $\epsilon \equiv \phi'(R = 1)$, or 2) we march changing the value of the function f at the origin, $f_0 \equiv f(0)$. In both cases, the marching parameter that we give as an input to our numerical code appears as a boundary condition.

We will use these two marching approaches because numerical convergence of each one is different at different regions of the parameter space. The first marching strategy is chosen essentially because $\epsilon \equiv \phi'(R=1)$ is the expansion parameter of the perturbative construction of [43] thus we can straightforwardly use the perturbative solution of [43] as a seed for the numerical scheme. However, we will find that away from the perturbative regime, ϵ no longer uniquely parametrizes the family of solutions (see plots in Appendix). This is because for a given ϵ there will be more than one soliton (the two or more solitons differ in their charges $\mathcal{M}$ and $\mathcal{Q}$). This happens in the neighborhood of the cusps of Fig. 3.2.1. For this reason, it is good to use an alternative parametrization where we march along the soliton branch using the value of the function $f = g_{tt}$ at the origin: we find that physical/thermodynamic quantities are a monotonic function of $f_0 \equiv f(0)$ as we move along any of the soliton families even when we cross any of the cusps sketched in Fig. 3.2.1.

When we use ϵ as a marching parameter we find convenient to introduce the field redefinitions,

$$f(R) = \tilde{q}_1(R), \qquad A(R) = \tilde{q}_2(R), \qquad \phi(R) = \left(1-R^2\right)\tilde{q}_3(R). \tag{3.3.17}$$

These redefinitions automatically impose the condition $\phi(1) = 0$ when we search for smooth functions $\tilde{q}_{1,2,3}$. Additionally, we also impose the normalization condition $\tilde{q}_1(1) = 1$. All other boundary conditions discussed in the previous two subsections follow from requiring that the equations of motion are also valid at $R = 0$ and $R = 1$.

On the other hand, when we use the marching parameter f_0, it is useful to introduce the field redefinitions

$$f(R) = \left(1-R^2\right) f_0 + R^2\left(1-\left(1-R^2\right) q_1(R)\right), \tag{3.3.18}$$

$$A(R) = q_2(R), \tag{3.3.19}$$

$$\phi(R) = \left(1-R^2\right) q_3(R).$$

which have the advantage of introducing explicitly the marching parameter f_0 in the problem and that smooth functions q_1 and q_3 automatically satisfy the boundary conditions $f(0) = f_0$, $f(1) = 1$ and $\phi(1) = 0$. The other boundary conditions for $q_{1,2,3}$ are derived boundary conditions in the sense that they follow directly from the equations of motion evaluated at the boundaries [104].

To solve numerically our boundary value problem, we use a standard Newton-Raphson algorithm and discretise the coupled system of three ODEs using pseudospectral collocation (with Chebyshev-Gauss-Lobatto nodes along the R). The resulting algebraic linear systems are solved by LU decomposition. These numerical methods are described

in detail in the review [104]. Since we are using pseudospectral collocation, our results should display exponential convergence with the number of grid points. We check this is indeed the case and the thermodynamic quantities that we display have, typically, 8 decimal digit accuracy. We further use the quasilocal first law (3.3.10) (typically, obeyed within an error smaller than $10^{-3}\%$) to check our solutions.

3.4 Phase diagram for charged solitons confined in a Minkowski box

As discussed in our summary of results (section 3.2), a system with a scalar field with charge e confined inside a box has two natural critical charges: $e_{\rm NH} = \frac{1}{2\sqrt{2}} \sim 0.354$ and $e_{\rm S} = \frac{\pi}{\sqrt{2}} \sim 2.221$. Moreover, we also find the existence of two other important critical charges, $e_\gamma \sim 1.13$ and $e_c \sim 1.8545 \pm 0.0005$, that we could not predict using heuristic or perturbative analysis. These charges satisfy the relations $e_{\rm NH} < e_\gamma < e_c < e_{\rm S}$. The charge $e_{\rm NH}$ plays no special role in the discussion of the phase diagram of solitons of the system. Thus, we do not discuss it further, and in the next subsections, we describe the properties of solitons in the following 4 windows of scalar charge: 1) $e \leq e_\gamma$, 2) $e_\gamma < e \leq e_c$, 3) $e_c < e \leq e_{\rm S}$, and 4) $e > e_{\rm S}$. For concreteness we will display results for a particular value of e for each one of these windows: 1) $e = 0.23$ (section 3.4.1), 2) $e = 1.854$ (section 3.4.2), 3) $e = 1.855$ (section 3.4.3), and 4) $e = 2.3$ (section 3.4.4). Altogether, these results (and others not presented) will allow to extract the conclusions summarized in section 3.2.

It follows from an analysis of the RN quasilocal charges (4.3.6) that, in the quasilocal $\mathcal{Q} - \mathcal{M}$ plot, the region that represents RN BHs whose horizon radius fits inside the box is the triangular surface bounded by the lines $\mathcal{Q} = 0$, $\mathcal{M}/L = 1$ and $\mathcal{M} = \sqrt{2}\mathcal{Q}$. When plotting the adimensional quasilocal mass $\mathcal{M}/L$ as a function of the dimensionless quasilocal charge $\mathcal{Q}/L$ we will often find that soliton curves are very close to the extremal RN curve that we use for reference. Therefore, we will also plot $\Delta\mathcal{M}$ as a function of $\mathcal{Q}$, where $\Delta\mathcal{M}$ is the quasilocal mass difference between the soliton and the extremal RN BH that has the *same* quasilocal charge. For reference, in our plots the orange curve describes this extremal RN BH solution which must have quasilocal charge $\mathcal{Q} \leq 2^{-1/2}$ to fit inside the covariant box with dimensionful radius L (and thus adimensional radius $R = 1$).

In the $\mathcal{Q} - \Delta\mathcal{M}$ plane, non-extremal RN BHs exist in the triangular region with boundaries $\mathcal{Q} = 0$, $\Delta\mathcal{M} = 0$ and $\Delta\mathcal{M} = 1 - \sqrt{2}\mathcal{Q}/L$. The latter curve is

$$(\mathcal{Q}/L, \Delta\mathcal{M}/L) = \left(L^{-1}\mathcal{Q}\big|_{\text{ext RN}}, 1 - L^{-1}\mathcal{M}\big|_{\text{ext RN}}\right) = \left(\frac{R_+}{\sqrt{2}}, 1 - R_+\right) \tag{3.4.1}$$

where $\mathcal{M}\big|_{\text{ext RN}}$ and $\mathcal{Q}\big|_{\text{ext RN}}$ are given by (4.3.6) with $\mu = \mu_{\rm ext} = \sqrt{2}$. In our $\mathcal{Q} - \Delta\mathcal{M}$

plots, the red dashed line is the parametric curve (4.4.1) with R_+ *allowed to take also values above 1*. It turns out that the most charged solitons that we will find seem to approach this dashed red line (in the limit where scalar condensate amplitude ϵ approaches infinity). In this sense, for a given quasilocal mass (smaller than 1), this red dashed line represents the maximal quasilocal charge that solutions that can fit inside the box can have.

To test our numerical results and to set the regime of validity of the perturbative analysis of [43], in many of our plots we will also display the perturbative prediction of [43] (which is valid only for small solitons) using green curves.

3.4.1 $e < e_\gamma \sim 1.13$

We illustrate the properties of solitons in this range using the case $e = 0.23$ (other examples can be found in Fig. 3.2.2). As discussed in our summary section 3.2, for this range of e the system only has the main family of solitons.[4]

In Fig. 3.4.1 we plot the dimensionless quasilocal mass $\mathcal{M}/L$ as a function of the dimensionless quasilocal charge $\mathcal{Q}/L$ of the main soliton branch. As a first observation, note that our full nonlinear numerical construction of solitons (black curve) agrees well with the perturbative construction of [43] at small values of the charge $\mathcal{Q}$ (green curve). However, for larger charges the elaborated zig-zagged structure that is developed is not captured by the perturbative analysis.

The zig-zagged structure in the quasilocal charge-mass plot translates into a damped oscillatory behaviour, when we plot the fields f and ϕ (or even the Kretschmann) at the origin as a function of the scalar condensate amplitude $\epsilon \equiv \phi'(1)$ (see Figs. 3.4.2-3.4.3), or into a spiral motion in the $A(0) - \epsilon$ plane (see right panel of Fig. 3.4.2); see also plots in the Appendix. Altogether, this reveals the existence of a a Chandrasekhar limit at $\mathcal{Q} = \mathcal{Q}_{Ch} \approx 0.0094L$ (point A in Figs. 3.4.1-3.4.3). We can understand this limit as follows. For very small e, the soliton is essentially an almost neutral boson star and the latter is expected to have a Chandrasekhar limit at a critical mass. Expanding in e, the soliton charge $\mathcal{Q}$ can be measured by the gauge field sourced by the above weakly charged boson star and so it should be of order e, *i.e.* the Chandrasekhar quantities $(\mathcal{M}_{Ch}, \mathcal{Q}_{Ch})$ should grow linearly with e (for small e) and then keep growing with e. We find this is indeed the case (as partially illustrated in Fig. 3.2.2) as long as $e < e_c$ (to be discussed later; as we cross e_c into $e > e_c$, $\mathcal{Q}_{Ch}$ jumps *discontinuously* to the maximum charge that can fit inside the box).

[4] We cannot exclude the existence of a second family but if it exists, it is not connected to the secondary soliton family present for $e > e_\gamma$ and that we discuss in the following subsections. Also note that beyond the ground state, there is also an infinite tower of excited solitons that are the nonlinear backreaction of the normal mode radial overtones of the system.

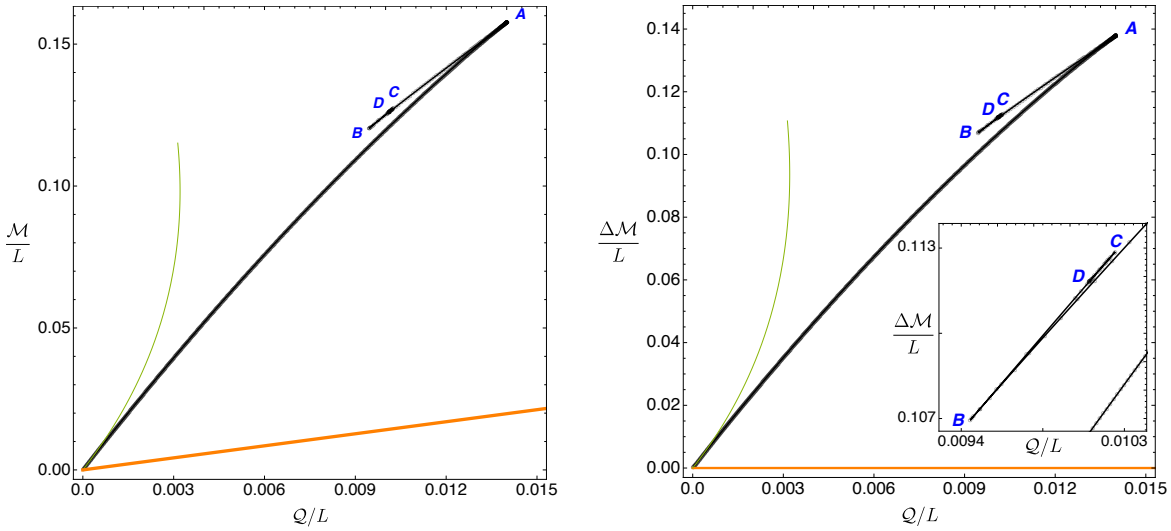

Figure 3.4.1: Main soliton family with $e = 0.23$ ($e < e_\gamma$). The orange curve describes the extremal RN BH solution, and the green curve is the perturbative prediction of [43]. **Left panel:** Quasilocal mass $\mathcal{M}/L$ as a function of the quasilocal charge $\mathcal{Q}/L$. **Right panel:** $\Delta\mathcal{M}$ is the mass difference between the soliton and an extremal RN BH that has the *same* quasilocal charge. For larger e this quantity will distinguish better some solitons from the extremal RN BH solution. The inset plots reveals a zig-zag structure along the cusps $A, B, C, D, \cdots$

At the Chandrasekhar limit (point A in Fig. 3.4.1) the family of solitons reaches a regular cusp and afterwards, it develops a sequence of zig-zagged branches AB, $BC, \cdots$ splitted by new regular cusps $B, C, D, \cdots$ As we move along these, the parameter $f_0 = f(0)$ is increasingly approaching zero, which signals that one is approaching a naked singularity (see left panel of Fig. 3.4.2). This conclusion is further corroborated by the fact that the Kretschmann curvature at the origin, $K\big|_{R=0} = R_{abcd}R^{abcd}\big|_{R=0}$, is growing unboundedly as we follow the path $OABCD\cdots$. It becomes increasingly harder to follow this soliton family beyond point D but the above results suggest that the soliton might well go through an infinite sequence of cusps (in the charge-mass plot) and an infinite sequence of damped oscillations or spirals in the plots of Figs. 3.4.2-3.4.3, before it finally reaches a naked singularity configuration where $f\big|_{R=0} \to 0$ and $K\big|_{R=0} \to \infty$.

The main soliton family of solutions has similar properties to the ones displayed in Figs. 3.4.1-3.4.3 for any $e < e_c$, although it becomes harder to capture higher order cusps as e increases.

3.4.2 $e_\gamma < e < e_c \simeq 1.8545 \pm 0.0005$

We now describe the solitons of the system for a scalar field charge in the window $e_\gamma < e < e_c \simeq 1.8545 \pm 0.0005$. We choose to discuss the particular case $e = 1.854 \lesssim e_c$ for reasons that will soon be clear (further examples all the way down to $e \to e_\gamma \sim 1.13$ can

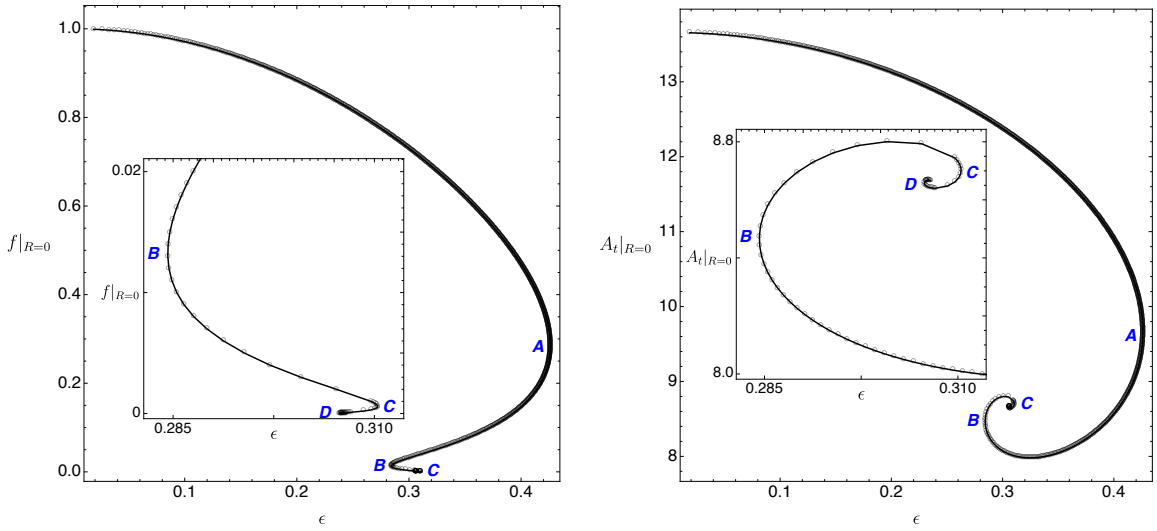

Figure 3.4.2: Main soliton family with $e = 0.23$ ($e < e_\gamma$). **Left panel:** As we ramp up the value of the amplitude $\epsilon \equiv \phi'(1)$ the core value of the metric function f is (very) slowly approaching zero. The inset plot zooms in the damped oscillations around the turning points B, C, D. **Right panel:** Contrast this with the core value of the gauge field A which is spiralling inward indefinitely as ϵ is ramped up. The inset plot zooms in the spirals around the turning points B, C, D.

be found in the right panel of Fig. 3.2.2). For the plots associated with this scalar charge, the black curve represents what we call the 'main family' (as mentioned in the introduction this family is perturbatively connected to a Minkowski box) and the magenta curve represents a distinct 'secondary family' which is *not* perturbatively connected to a Minkowski box but still describes genuine regular solitonic solutions to the theory.

The main soliton family for these values of the charge is qualitatively similar to what we had for $e = 0.23$: in the quasilocal charge-mass plane of Fig. 3.4.4, this black disk curve extends from the origin O, developing cusps $A, B, C \cdots$. The discussion presented for $e = 0.23$ applies *mutatis mutandis* to the present case. As a minor observation, note however that there is now a window of quasilocal charges (around the Chandrasekhar point A) for which the main family has a lower mass than an extremal RN BH with the same charge.

Unlike for $e < e_\gamma$, for $e = 1.854$ (and any $e_\gamma < e < e_c$) there is 'secondary family' of solitons. This is the magenta triangle curve $\beta' abc \cdots$ in Figs. 3.4.4-3.4.6. In Fig. 3.4.4, we see that the main soliton family achieves a maximum charge (the Chandrashekhar point A) which is *less than* the minimum charge (at point a) reached by the secondary soliton family. From this plot, we also infer that the secondary family of solitons extends all the way up to β' (which is a point in the red dashed "wall" that describes the maximum quasilocal charge that a solution confined in the box of radius R can have). In terms of the ϵ parameter that we use to construct these solutions, this β' configuration is only

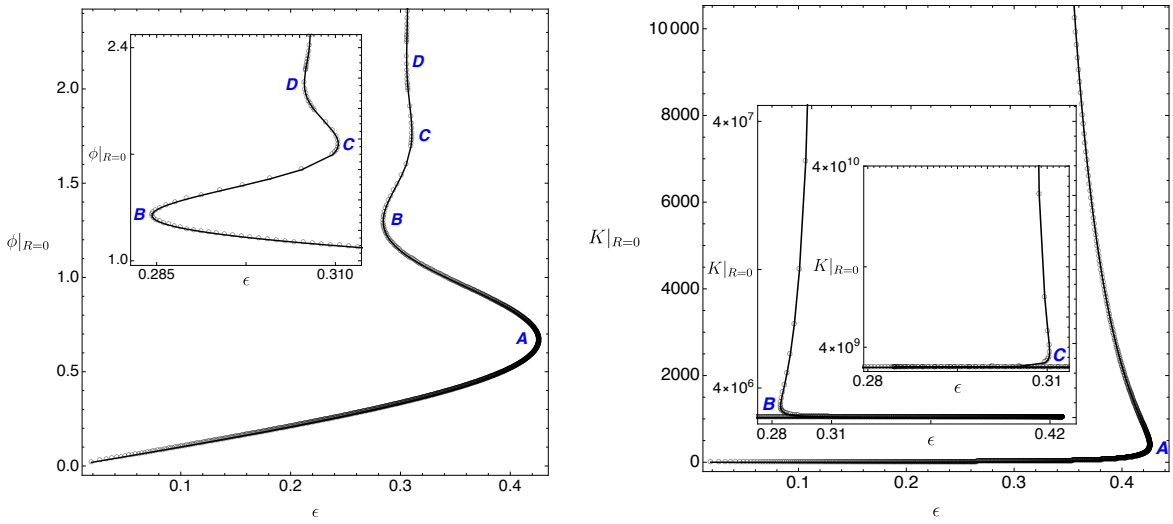

Figure 3.4.3: Main soliton family with $e = 0.23$ ($e < e_\gamma$). **Left panel:** As we ramp up the value of the amplitude $\epsilon \equiv \phi'(1)$ the core value of the scalar field ϕ is increasing seemingly without bound. The inset plot zooms in the damped oscillations around the turning points B, C, D. **Right panel:** Kretschmann invariant at the origin as a function of ϵ. The turning points A, B, C, D are seeing a substantial increase in the curvature as we go from one to the next. This is despite the fact that they lie close to each other in the phase diagram. The inset plots zoom in around the point B and then around the point C to be able to show more clearly the large growth that K experiences (note the different scales in which plot).

reached in the limit $\epsilon \to \infty$ and this is why it is difficult to extend the magenta curve all the way till it hits the dashed red line (but when we analyse the $e = 2.3$ case we will see that solitons indeed extend all the way to the dashed red line). Still in Fig. 3.4.4, we also see that the secondary family of solitons has a cusp/zig-zagged structure $abc\cdots$ that is very similar to the one found in the main family but this time around the regular cusps $a, b, c, \cdots$. In particular, this structure also translates into damped oscillations or spirals in the quantities plotted in Figs. 3.4.5-3.4.6. As we move along the cusps $a, b, c, \cdots$, $f_0 \equiv f(0)$ is approaching zero and the Kretschmann curvature is growing unbounded. This suggests that this secondary soliton family also ends in a naked singularity after, possibly, going though an infinite sequence of cusp/oscillations/spirals. On the other hand f_0 and $K\big|_{R=0}$ remain finite as we approach β'. In particular, $f_0\big|_{\beta'} \ll 1$.

The gap Aa between the main and the secondary families of solitons in the $\mathcal{Q}-\Delta\mathcal{M}$ plane of Fig. 3.4.4 is a quantity that deserves special attention. We find that this gap decreases as e increases, approaching zero when $e \to e_c$ from below. A similar gap Aa present in the several plots of Figs. 3.4.5-3.4.6 also decreases and approaches zero as $e \to e_c$. In the next subsection we will further discuss what happens for values of e at e_c and above.

Note that we can have hairy solitons with higher quasilocal charge than the maximum charge that a RN confined inside a box can have. Indeed, the maximum quasilocal

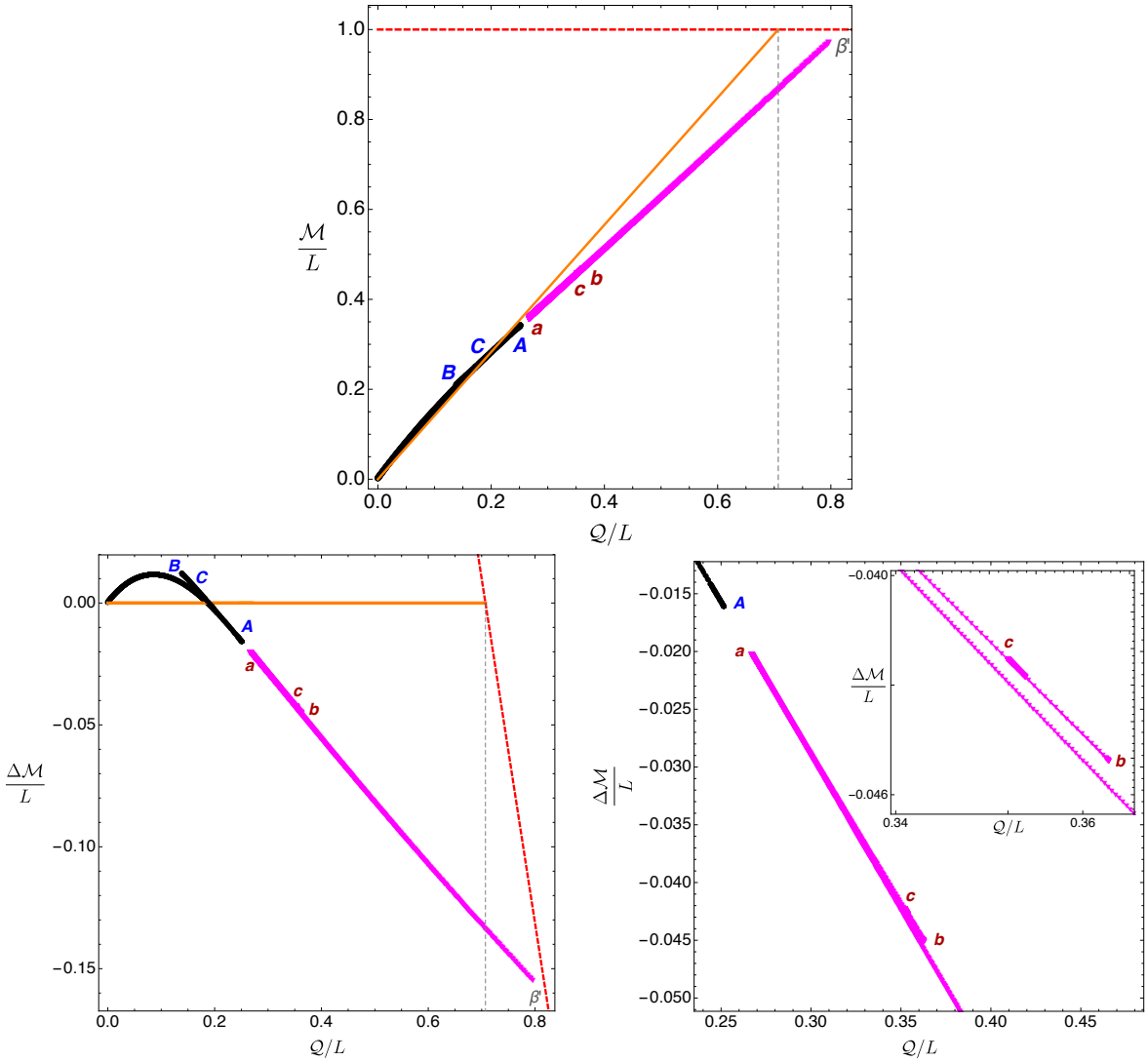

Figure 3.4.4: Soliton families with $e = 1.854$ ($e_\gamma < e < e_c$). **Top panel:** quasilocal mass *vs* charge. The main family (black curve $OABC\cdots$) looks to be on top of the extremal RN BH (orange line) which is why we find it useful to plot the quantity $\Delta\mathcal{M}$ which is shown on the bottom panel. **Bottom-left panel:** the main black soliton curve starts more massive ($\Delta\mathcal{M} > 0$) than the extremal RN but then becomes less massive ($\Delta\mathcal{M} < 0$) above a certain charge. The secondary soliton family (magenta curve) eventually hits the dashed red line at β' and extends to a finite lower charge $\mathcal{Q}$ where it develops a series of cusps $a, b, c, \cdots$. **Bottom-right panel:** zoom of the left plot to amplify the gap Aa between the main and secondary families (magenta) and the cusp structure $abc\cdots$ of the secondary family.

charge that a RN confined inside a box (*i.e.* with horizon radius $R_+ \leq 1$) can have is the one of an extremal RN BH with horizon radius $R_+ = 1$, namely $\mathcal{Q}/L = 2^{-1/2}$. This is identified by the vertical dashed grey line in Fig. 3.4.4. But this figure also demonstrates that we can have secondary solitons with a charge that is above this value.

Other examples of solitons with scalar charge in the window $e_\gamma < e < e_c$ can be found

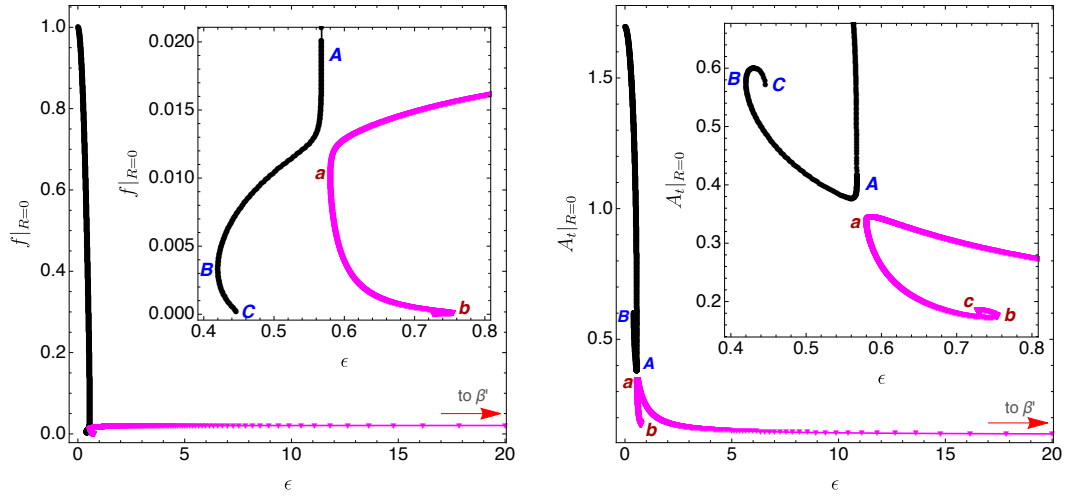

Figure 3.4.5: Soliton families with $e = 1.854$ ($e_\gamma < e < e_c$). As in previous plots, the black curve represents the main family and the magenta line represents the secondary family. In both figures, the inset plots zooms in around the region where the two soliton families approach the most. As $e \to e_c^-$, the gap between A and a decreases until these two points merge precisely at $e = e_c$. **Left panel:** The function f evaluated at the origin of the soliton is plotted as a function of the amplitude ϵ. The two families seem to be on top of each other for a small range of ϵ but the inset plot shows that this is not the case and a small gap is still present while $e \neq e_c$. **Right panel:** The gauge field A_t evaluated at the origin as a function of ϵ. Both families spiral inward towards a singularity in one of their endpoints: this corresponds to the cusp structure observed in the $\mathcal{Q} - \Delta\mathcal{M}$ plot.

in the right panel of Fig. 3.2.2. From the several cases considered, we conclude that secondary solitons (that extend to the red dashed line) exist only for $e_\gamma < e < e_c$ and inside the region bounded by the closed auxiliary dashed line $a_c\beta_c\gamma$ of Fig. 3.2.2 (these auxiliary lines are inserted to guide reader's eye). As $e \to e_c$, the secondary soliton approaches the auxiliary line $a_c\beta_c$, *i.e.* the equivalent of point a and β' of Fig. 3.4.4 approach point a_c and β_c, respectively, of Fig. 3.2.2. On the opposite end of the window, as $e \to e_\gamma$, one finds that the equivalent to points a and β' of Fig. 3.4.4 collapse to a single point denoted as γ in Fig. 3.2.2.

3.4.3 $e_c < e < e_s = \frac{\pi}{\sqrt{2}} \sim 2.221$

So, what changes when we cross $e = e_c \sim 1.8545 \pm 0.0005$? To address this question we compare the phase diagram for the charge $e = 1.854$ just below e_c (Fig. 3.4.4) with the one of Fig. 3.4.7 for a charge $e = 1.855$ just above e_c. From this comparison we can infer the following. Approaching e_c from below, precisely at $e = e_c$, the cusp A of the main soliton (black disk curve) merges with the cusp a of the secondary soliton (magenta triangle curve) in Fig. 3.4.4. That is to say, at $e = e_c$ the black and magenta curves of Fig. 3.4.4 merge at $A \equiv a$, and this becomes a bifurcation point that irradiates a total of four branches.

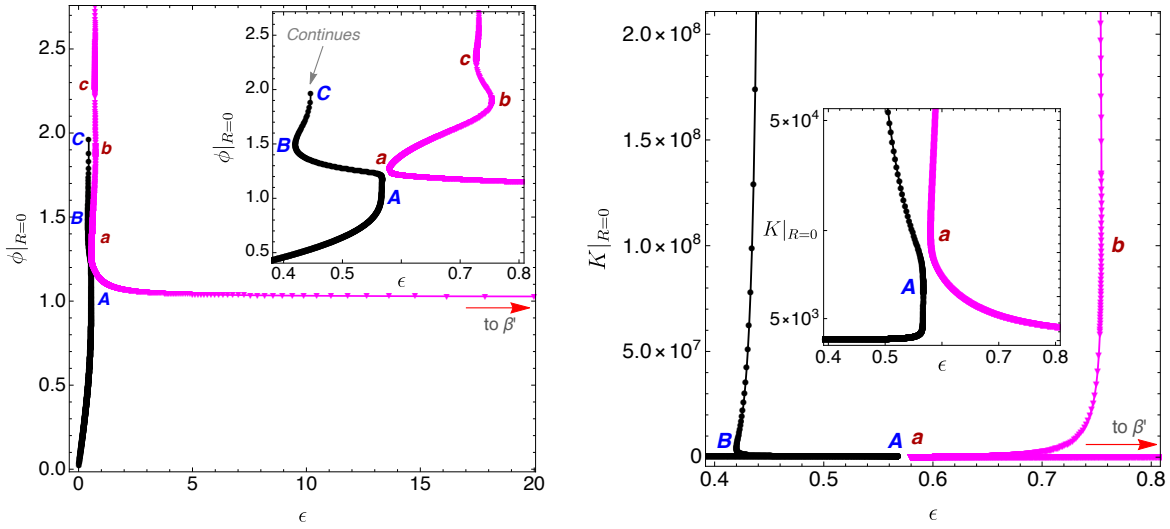

Figure 3.4.6: Soliton families with $e = 1.854$ ($e_\gamma < e < e_c$) with the main (black disks) and secondary (magenta triangles) families. In both figures, the inset plots zooms in around the region where the two soliton families approach the most. As $e \to e_c^-$, the gap between A and a decreases until these two points merge precisely at $e = e_c$. **Left panel:** The core value of the scalar field is plotted as a function of ϵ. **Right panel:** The value of the Kretschmann at the origin is plotted as a function of ϵ. Both families have diverging Kretschmann as one moves along the damped oscillation trenchs.

But as e grows above e_c, two new independent families emerge. Old point $A \equiv a$ and two of the four branches that irradiate from it at $e = e_c$ are now absorbed in the main soliton family (around the region that we vaguely identify as α in our plots of Fig. 3.4.7) that is described by the black disk curve $O\beta$ of Fig. 3.4.7. This extends from the origin all the way down to point β in the dashed red line (which is reached as $\epsilon \to \infty$). That is to say, for $e > e_c$, the main soliton family no longer has a Chandrashekhar mass limit neither cusps: the Kretschmann curvature is now always finite along this family as it extends from the origin all the way to β, where the family terminates because it reaches the maximum quasilocal charge that can fit inside the box with dimensionless radius $R = 1$. Also note that for $e > e_c$, hairy solitons within the main family can have a quasilocal charge that is higher than the maximum charge $\mathcal{Q}/L = 2^{-1/2}$ (identified by the vertical dashed grey line in Fig. 3.4.7) that a RN confined inside a box can have.

On the other hand, the system now has a novel secondary soliton family (the blue diamond curve $\cdots DCBbc\cdots$ in Fig. 3.4.7) that exists only in a window of quasilocal charge $\mathcal{Q}/L$ bounded from below by cusp B (that used to belong to the main family for $e < e_c$) and from above by the cusp b (that used to belong to the old secondary magenta triangle family for $e < e_c$). (The distinction between the main black disk family and the secondary blue diamond family is better seen in the plots of Figs. 3.4.8-3.4.9). This new secondary family absorbs the old point $A \equiv a$ and the two (out of four) branches that irradiate from it at $e = e_c$ that were not embodied in the main family. Further confirmation that this

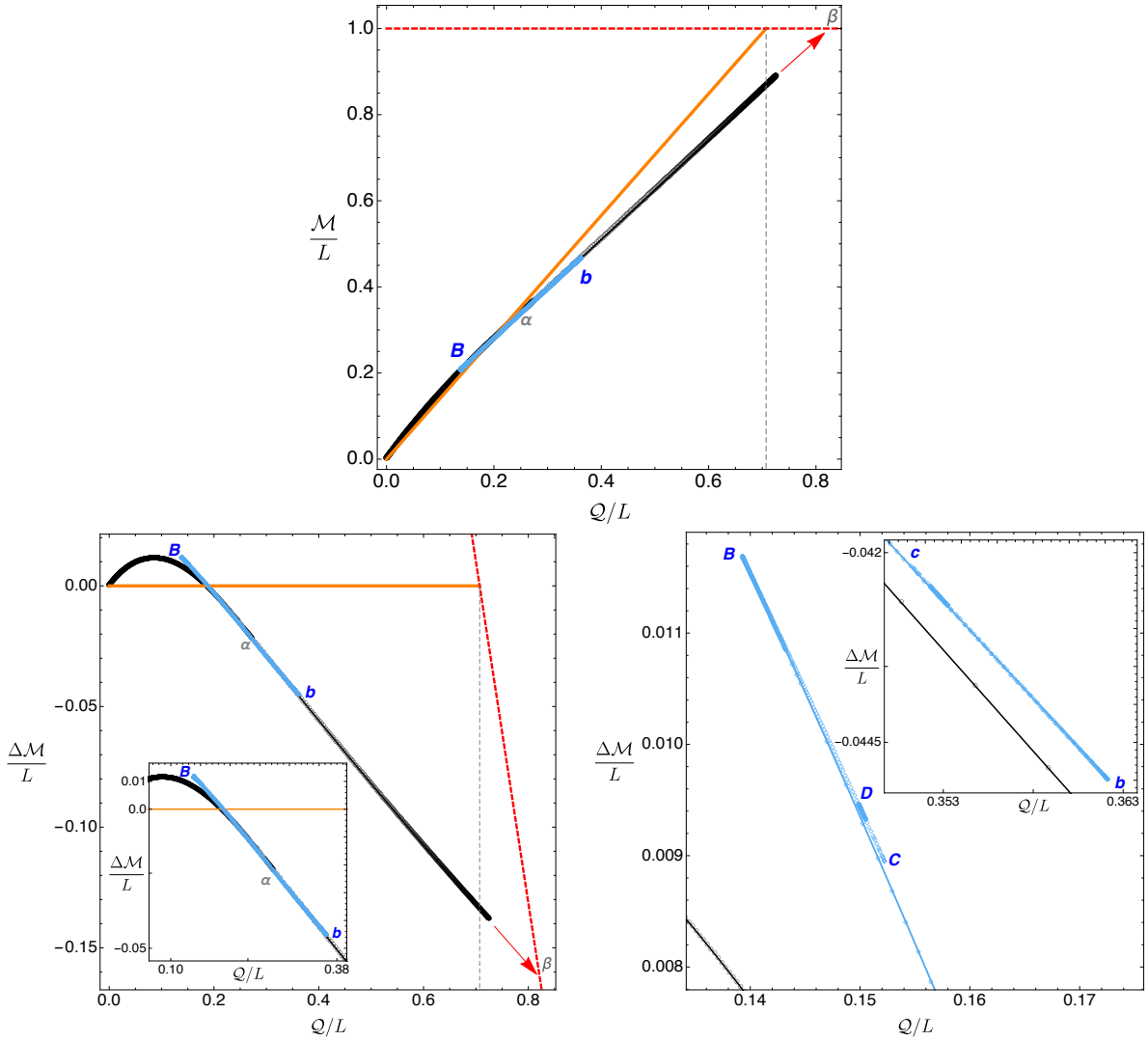

Figure 3.4.7: Soliton families with $e = 1.855$ ($e_c < e < e_s$). **Top panel:** the quasilocal charge-mass diagram is now much different from the $e < e_c$ case as better seen in the bottom plots. **Bottom-left panel:** for $e > e_c$ the main soliton curve (black disks) no longer has a Chandrasekhar limit; instead it extends from the origin all the way down to point β on the red dashed line. On the other hand, the secondary soliton family (blue diamonds curve) is now confined in a window (Bb) of quasilocal charge and with mass above the main soliton. **Bottom-right panel:** zoom of the left plot to amplify the cusp structures $BCD\cdots$ and $bc\cdots$ that appear on both sides of the secondary soliton family.

new secondary family of soliton derives from connecting the old *trench* $ABCD\cdots$ of the main black disk branch with the old *trench* $abc\cdots$ of the secondary magenta triangle branch of Fig. 3.4.4 can be obtained as follows. Starting at a point in the old trench $ABCD\cdots$ or $abc\cdots$ of Fig. 3.4.4, we can run a code where we fix ϵ or f_0 and, *this time*, we march the *electric scalar charge* e in very small steps from $e = 1.854 < e_c$ all the way up to $e = 1.855 > e_c$. Doing this exercise we find that we indeed end on solutions of the blue diamond curve of Figs. 3.4.7-3.4.9.

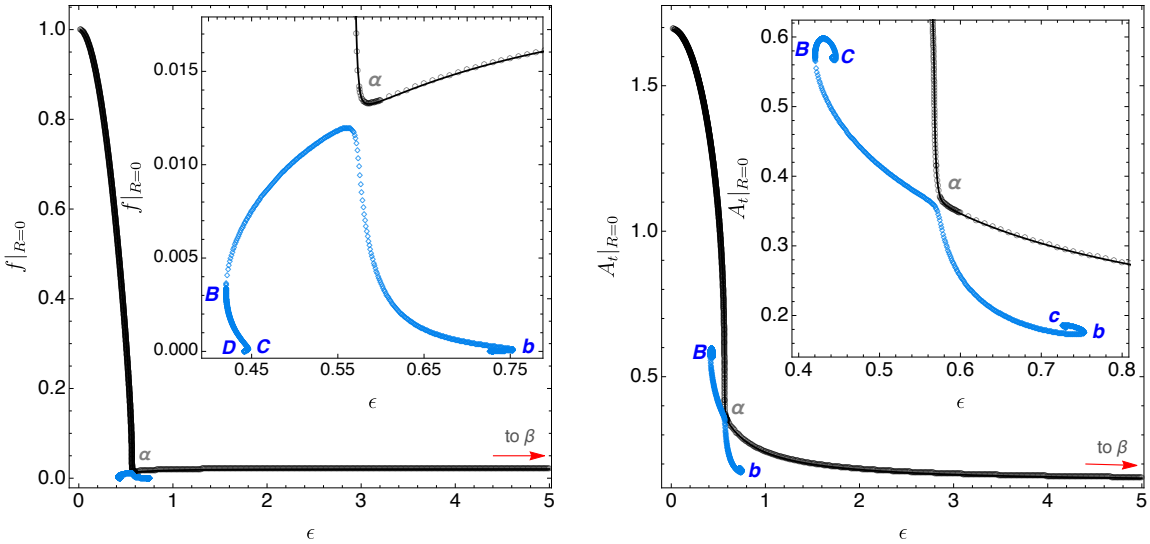

Figure 3.4.8: Soliton families with $e = 1.855$ ($e_c < e < e_s$). Comparing with the $e \lesssim e_c$ case of Fig. 3.4.5 we identify main differences: the black curve is still the main soliton family but it now extends to arbitrarily large ϵ. On the other hand, the secondary soliton family – blue diamond curve – is now confined to a small window of scalar condensate amplitude ϵ. Note that as $e \to e_c^+$, the gap between the two curves, in the region vaguely signalled as α, decreases until it vanishes precisely at $e = e_c$. **Left panel:** $f(0)$ as a function of the amplitude ϵ. **Right panel:** The gauge field evaluated at the origin $A_t(0)$ as a function of ϵ. This time the secondary family only exists for a small range of scalar field amplitudes and develops a 'double-spiral' structure, *i.e.* a cusp structure develops at either end of the curve, in the trenches $Abc\cdots$ and $BCd\cdots$.

Both around B and b, the blue diamond secondary soliton family of Fig. 3.4.7 displays a complex zig-zagged structure that we are already familiar with. In particular, on both sides of this curve, as we move along the several cusps $B, C, D, \cdots$ or $b, c, \cdots$, we find that the Kretschmann curvature is growing unbounded (see Fig. 3.4.9) and $f_0 \equiv f(0)$ is approaching zero (see Fig. 3.4.8). Altogether, this suggests that the secondary family of solitons undergoes a sequence of possibly infinite damped oscillations/spirals before it reaches a naked singularity at *both* sides of its domain.

The property that, at $e = e_c$, the black and magenta curves of Figs. 3.4.4-3.4.6 merge at $A \equiv a$ and then two distinct new soliton families (the black and blue curves of Figs. 3.4.7-3.4.9) emerge for $e > e_c$ is inferred not only from the quasilocal charge-mass plots. Indeed we arrive to the same conclusion analysing the plots for the fields f, A_t, ϕ and Kretschmann when we compare Figs. 3.4.8-3.4.9 for $e = 1.854 \lesssim e_c$ with Figs. 3.4.8-3.4.9 for $e = 1.855 \gtrsim e_c$ (see in particular their inset plots).

As we keep increasing e above e_c we find that the window of quasilocal charge $\mathcal{Q} \in [0, \mathcal{Q}_0]$ where $\Delta\mathcal{M} \geq 0$, *i.e.* where solitons have a quasilocal mass higher than the extremal RN solution with the same $\mathcal{Q}/L$, decreases and approaches zero as $e \to e_s$ (see next

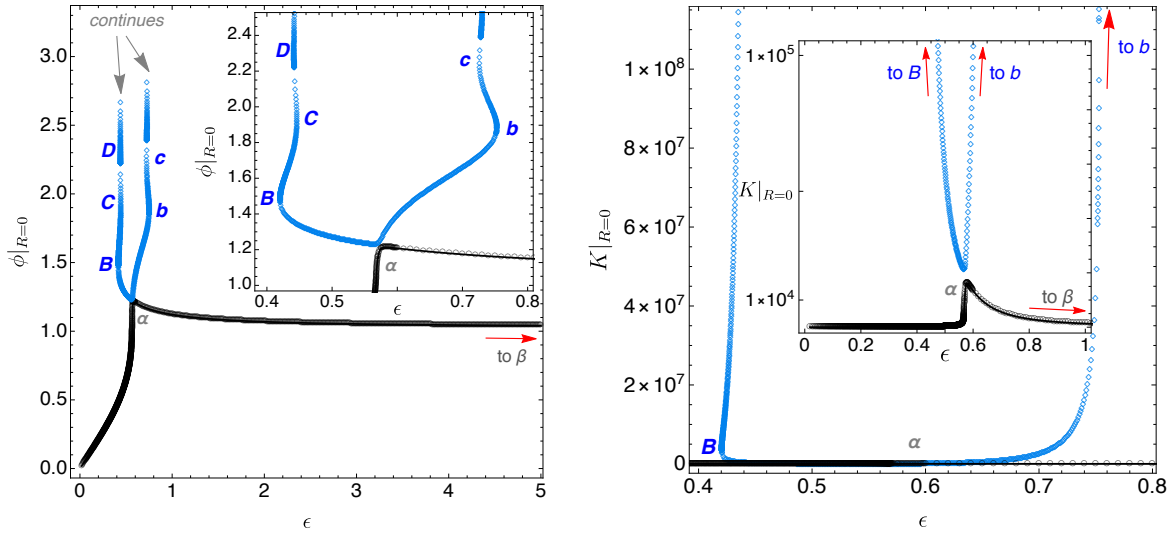

Figure 3.4.9: Soliton families with $e = 1.855$ ($e_c < e < e_s$). Alike in the previous figure, comparing the current plots with those of Fig. 3.4.6 for the $e \lesssim e_c$ case, we find major differences: the black disk curve is still the main soliton family but it now extends to arbitrarily large ϵ. On the other hand, the secondary soliton family (blue diamond curve) is now confined to a small window of ϵ. As $e \to e_c^+$, the gap between the two curves, in the region vaguely signalled as α, decreases until it vanishes precisely at $e = e_c$. **Left panel:** Core value of the scalar field as a function of ϵ with damped oscillations seen in both ends of the secondary soliton family. **Right panel:** The Kretschmann evaluated at the origin as a function of ϵ for both soliton families. Comparing with Fig. 3.4.6 for the $e \lesssim e_c$ case, we see that now the main soliton family exists for any ϵ and its curvature at the origin is finite in all its domain of existence. On the other hand, the curvature of the secondary soliton family now diverges at both of its ends.

section 3.4.4). We also find that the distance of closest approach between the secondary blue diamond family of solitons and the main black disk family increases in all the plots Figs. 3.4.7-3.4.9. There is a scenario that is highly probable: one should expect further critical charges $e_{c_i} > e_c$ ($i \geq 2$) where extra splits in the secondary blue diamond curve will occur. Indeed, at $e = e_c$ we found that the cusps A and a of Fig. 3.4.4 merge. Similarly, at a new critical charge $e_{c_2} > e_c$ we might expect that the cusps C and c in the secondary blue curve $\cdots DCBbcd\cdots$ of Fig. 3.4.7 will also merge. If so, above this new merger, *i.e.* for $e > e_{c_2}$, we should have the main family of solitons, then a *closed* curve secondary family of solitons $CBbc$ with $C \equiv c$ and then a third family of solitons with the "leftover" $\cdots Dd\cdots$ that should have an infinite sequence of cusps. We could then have a third critical charge $e_{c_3} > e_{c_2}$ where a new merger of cusps would occur. Then, above it we would end with the main family of solitons, then *two closed* families of solitons and a *fourth open* family of solitons with an infinite sequence of solitons and so on for $e_{c_i}, i > 4$ (additionally, as discussed before there is also an infinite tower of excited families of solitons besides this ground state families). It is very hard to numerically confirm this scenario. But, starting with our solutions with $e = 1.855$, fixing f_0 or ϵ and varying e to march the solutions in the blue curve of Fig. 3.4.7, we found evidence that this scenario is

very plausible. Further evidence in favour of this scenario is provided by the fact that in the context of hairy AdS solitons, closed families of solitons were found (see *e.g.* Figs. 5 and 7 of [110]).

3.4.4 $e > e_{\rm s} = \frac{\pi}{\sqrt{2}} \sim 2.221$

Strictly speaking the only critical charge that marks a substantial qualitative difference on the phase diagram of solitons is $e = e_c$. In this sense, for our discussion it would be enough to provide the illustrative cases $e = 1.854 < e_c$ (Figs. 3.4.4-3.4.6) and $e = 1.855 > e_c$ (Figs. 3.4.7-3.4.9). However, in this section we also discuss a change that occurs in the main soliton family of solitons when e increases above $e_{\rm s} = \frac{\pi}{\sqrt{2}} \sim 2.221$. Although from a pure soliton perspective this difference seems to be minor, it turns out to have a large impact when we consider the phase diagram of the theory that includes not only the solitons and RN BHs but also the hairy BHs.

We choose scalar field charge $e = 2.3$ to illustrate the properties of solitons with $e > e_{\rm s}$. The associated quasilocal charge-mass phase diagram for the main soliton is displayed in Fig. 3.4.10. Unlike for $e < e_{\rm s}$, the main solitons now have less quasilocal mass than the extremal RN BH with the same quasilocal charge, *i.e.* one always has $\Delta\mathcal{M} < 0$. For $e < e_{\rm s}$, in a neighbourhood of the origin one has $\Delta\mathcal{M} > 0$ but precisely at $e = e_{\rm s}$, one finds that the slope $\Delta\mathcal{M}/\mathcal{Q}$ vanishes at the origin and then becomes negative for $e > e_{\rm s}$. Why is this feature relevant? Well, precisely for $e \geq e_{\rm s}$ extremal RN BHs are unstable for $\mathcal{M} \geq 0$, *i.e.* the instability onset curve extends all the way down to $\mathcal{M} = 0$ (while for $e < e_{\rm s}$, this onset curve starts at the extremal RN BH at a *finite* value of $\mathcal{M}$).

These two properties (solitons have $\Delta\mathcal{M} < 0$ and RN BHs are unstable all the way down to $\mathcal{M} = 0$) strongly suggests that for $e > e_{\rm s}$, in the charge-mass phase diagram, the hairy charged BHs that emerge from the instability onset of the RN solution —which is a curve in Fig. 3.4.10 with positive $\Delta\mathcal{M}$ that approaches $\Delta\mathcal{M} = 0$ at $\mathcal{Q} = 0$ and $\mathcal{Q} = 2^{-1/2}$ (see. Fig. 3 of [40]) — should extend all the way down to the soliton (black curve $O\beta$ in Fig. 3.4.10) when the hairy BH horizon radius goes to zero. This was confirmed for small charges in the perturbative analysis of [43] and this expectation will also prove correct for higher quasilocal charges. On the other hand, for $e < e_{\rm s}$ hairy BHs should only exist above a critical quasilocal charge and the main family of hairy solitons is not the zero entropy limit of the hairy BHs of the theory, at least for small charges (and thus these hairy BHs cannot be constructed within perturbation theory for small charges [43]; one needs a full nonlinear numerical analysis that will be provided in part IV).

In Fig. 3.4.10 we also find strong evidence to a claim we made previously: as $\epsilon \to \infty$ the main soliton extends all the way to point β. This point is along the red dashed line that describes the maximal quasilocal charge that solutions confined inside a box of

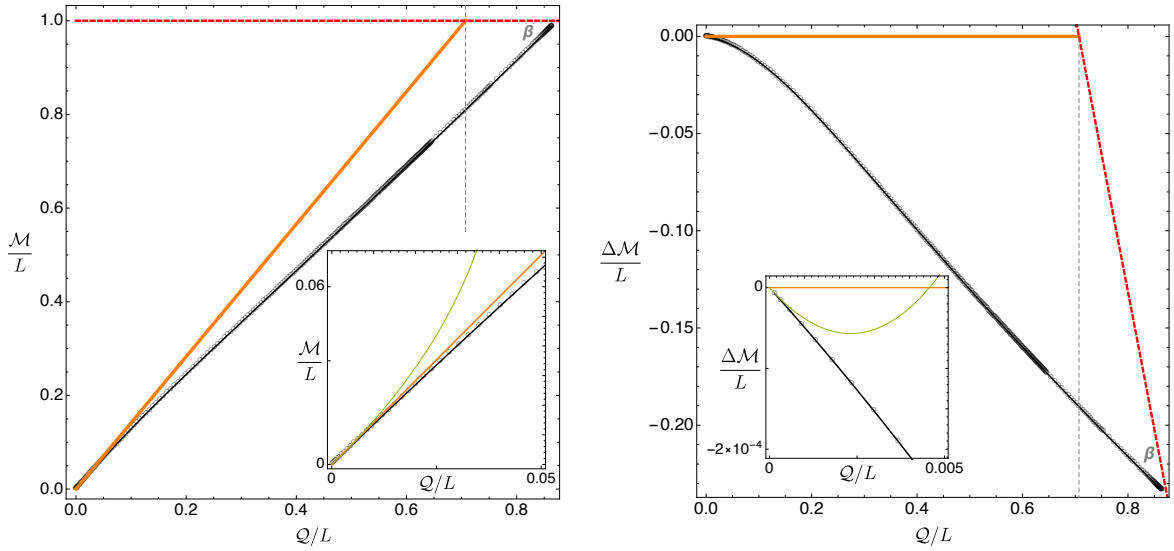

Figure 3.4.10: Main soliton family with $e = 2.3$ ($e > e_s$). The soliton extends smoothly from the origin to β and, unlike for $e < e_s$, the main soliton is now *always* less massive than its extremal RN counterpart. This case also demonstrates better a claim that we made for other cases e: as the scalar condensate amplitude grows unbounded, $\epsilon \to \infty$, the main soliton approaches point β in the red dashed line. As in other plots, the inset plots show the good agreement between the full result and the analytic prediction (green curve) at small charges.

dimensionless radius $R = 1$ can have. For smaller e, it is much harder to construct solutions with higher ϵ (compare the maximum values of ϵ in the several plots). This is why in the plots for smaller e that we have shown previously, the 'last' solution is not so close to β as in Fig. 3.4.10.

Overall, the reader can find a broad overview of the evolution (as e increases from a small value $e \ll e_s$ all the way above e_s) of the phase diagram for the main soliton family in Fig. 3.2.2.

For completeness, we also display the evolution of the fields f, A_t, ϕ and of the Kretschmann scalar with the scalar condensate amplitude in Figs. 3.4.11-3.4.12. We see that the gradients, observed in the similar plots of Figs. 3.4.8-3.4.9 for smaller e, become more pronounced.

3.5 Israel surface stress tensor and energy conditions

So far, we were able to construct the phase diagram of hairy solitons inside a box without needing to detail the matter content of the box. Indeed, to get the quasilocal phase diagrams of the previous section, we just had to integrate numerically the equations of motion in the domain $R \in [0, 1]$ subject to regular boundary conditions and vanishing

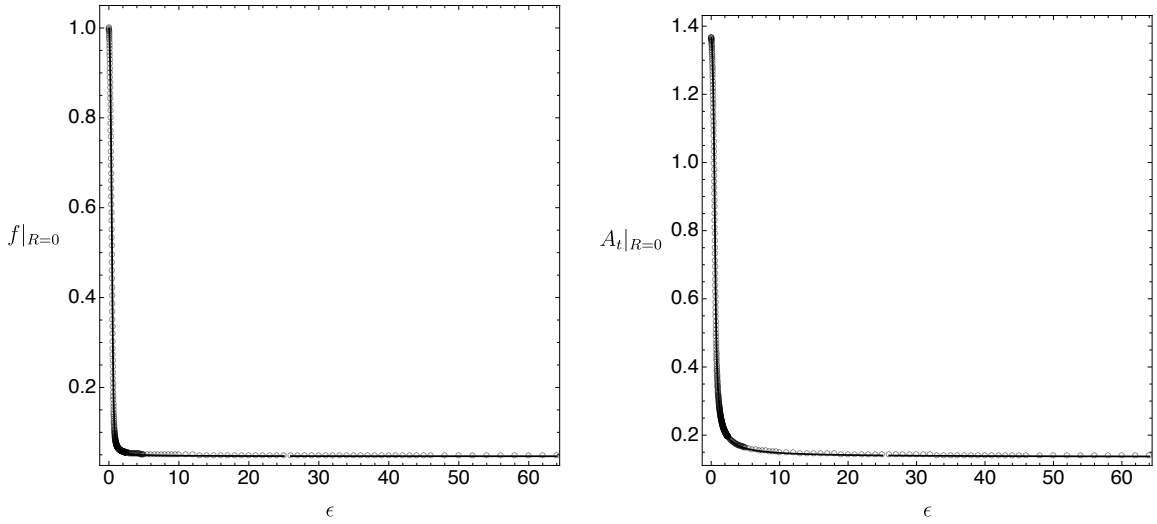

Figure 3.4.11: Main soliton family with $e = 2.3$ ($e > e_{\rm S}$).

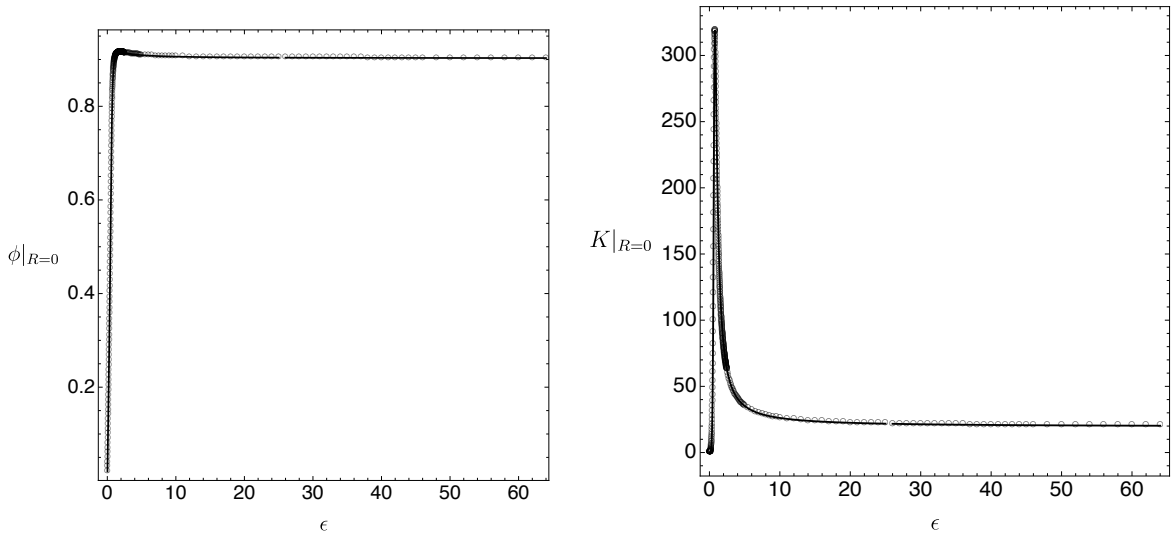

Figure 3.4.12: Main soliton family with $e = 2.3$ ($e > e_{\rm S}$).

scalar field at the box. For this, no further details are required. But the description of the solution is only complete once we discuss the full solution all the way up to the asymptotically Minkowski boundary.

In the literature one already has some studies of scalar fields confined in a Minkowski cavity: 1) at the linear level [118, 119, 41, 120, 121, 122, 123, 124, 125, 126, 127], 2) within a higher order perturbative analysis of the elliptic problem [40, 43], 3) as a nonlinear elliptic problem (although without having asymptotically flat boundary conditions [128, 129, 130] or without discussing the exterior solution [131]), and 4) as an initial-value problem [132, 133, 134]. However, except in the perturbative analysis of [43], the properties of the "internal structure" of the cavity or surface layer that is necessary to confine the scalar field and its contribution to the ADM mass and charge that ultimately describe, by Birkhoff's

theorem, the exterior RN solution are not discussed.

With the interior hairy soliton fields found in the previous section, we can now compute the Lanczos-Darmois-Israel surface stress tensor (3.3.16) which describes the energy-momentum content of the box Σ. If this satisfies relevant energy conditions, in practice, we can indeed build a physical box that confines the scalar fields and impedes that it disperses to infinity. In our construction we imposed the three Israel junction conditions (3.3.15a)-(3.3.15c) on the gravitoelectric fields on the surface layer Σ. With these conditions, the fields are continuous across Σ and the component of the electric field orthogonal to Σ is also continuous across Σ. This means that we are allowed to choose a surface layer that confines the charged scalar field without needing to have a surface electric charge density. Altogether, the three conditions (3.3.15a)-(3.3.15c) match the interior and exteriors solutions, *i.e.* they determine the parameters M_0, c_A, ρ in (3.3.6) as a function of the reparametrization freedom parameter N introduced in (3.3.14):

$$M_0 = \frac{1}{N^2}\left(1 - \frac{A_t'(1)^2}{2}\right) - 1, \qquad c_A = \frac{A_t'(1) + A_1(1)}{N}\,, \qquad \rho = -\frac{A_t'(1)}{N}\,. \tag{3.5.1}$$

That is to say, these junction conditions fix the exterior RN solution as a function of the interior field content but also as a function of the box energy-momentum content. And we have a 1-parameter freedom (N) to choose the box content that is able to confine the scalar condensate.

How can we fix N? Well, ideally we would choose it so that the gravitational field was also differentiable across the box, *i.e.* so that the fourth junction condition (3.3.15d) were also obeyed and thus the extrinsic curvature

$$K^t{}_t = -\frac{f'(R)}{2f(R)\sqrt{g(R)}}\,, \qquad K^i{}_j = \frac{1}{R\sqrt{g}}\,\delta^i{}_j\,, \quad (i,j) = (\theta,\varphi)\,, \tag{3.5.2}$$

were also continuous across the box. But for our system there is no choice of N that simultaneously makes $[K^t_t] = 0$ and $[K^i_i] = 0$, unless the scalar field vanishes everywhere. But we can fix N requiring that $[K^t_t] = 0$ (at the expense of having $[K^i_i] \neq 0$) or vice-versa, or any other combination.

Our choice of N fixes the energy density and pressure of the box since the surface tensor of the box can be written in the perfect fluid form, $\mathcal{S}_{(a)(b)} = \mathcal{E}u_{(a)}u_{(b)} + \mathcal{P}(h_{(a)(b)} + u_{(a)}u_{(b)})$, with $u = f^{-1/2}\partial_t$ and local energy density $\mathcal{E}$ and pressure $\mathcal{P}$ given by

$$\mathcal{E} = -S^t{}_t\,, \qquad \mathcal{P} = S^x{}_x = S^\phi{}_\phi\,. \tag{3.5.3}$$

Our choice should be such that relevant energy conditions are obeyed in order to be

physically acceptable. Different versions of these energy conditions read $(i = \theta, \varphi)$ [5]:

$$\text{Weak energy condition:} \qquad \mathcal{E} \geq 0 \quad \wedge \quad \mathcal{E} + \mathcal{P}_i \geq 0\,; \tag{3.5.4}$$

$$\text{Strong energy condition:} \qquad \mathcal{E} + \mathcal{P}_i \geq 0 \quad \wedge \quad \mathcal{E} + \sum_{i=1}^{2} \mathcal{P}_i \geq 0\,; \tag{3.5.5}$$

$$\text{Null energy condition:} \qquad \mathcal{E} + \mathcal{P}_i \geq 0\,; \tag{3.5.6}$$

$$\text{Dominant energy condition:} \qquad \mathcal{E} + |\mathcal{P}_i| \geq 0\,. \tag{3.5.7}$$

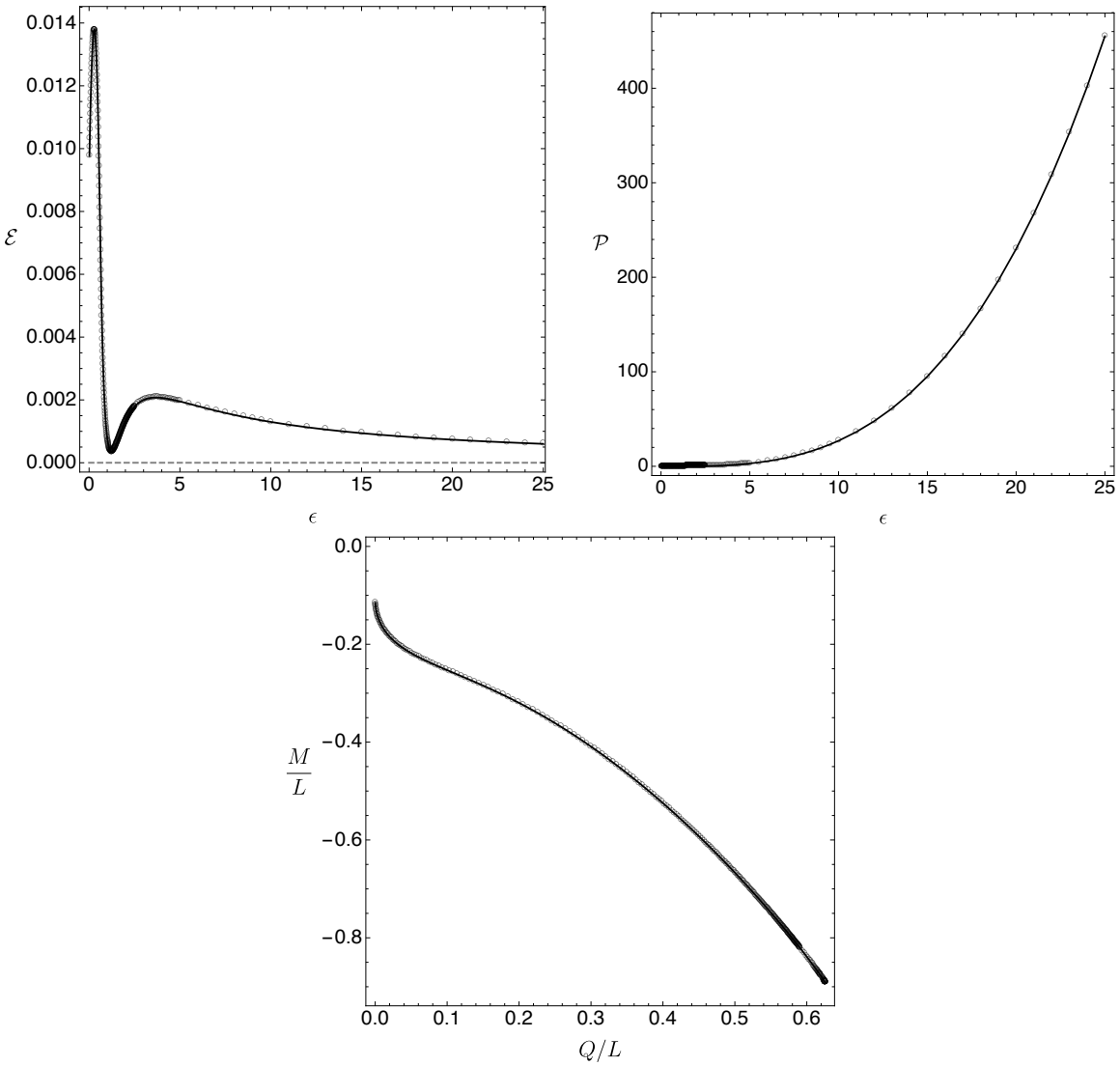

Figure 3.5.1: Properties of a box for a choice of reparametrization normalization N such that the Israel stress tensor satisfies the energy conditions (system with $e = 2.3$). **Top panel:** Energy density $\mathcal{E}$ and pressure $\mathcal{P}$. **Bottom panel:** ADM mass M *vs* ADM charge Q of the soliton solution (measured at the asymptotic region) includes the contribution from the scalar condensate confined inside the box, the energy-momentum content of the box layer and the exterior RN solution.

To illustrate that it is definitely possible to build boxes that obey these energy conditions,

let us take the main soliton family with $e = 2.3$ constructed in section 3.4.4. Let us further choose N such that one has no pressure, $N = N\big|_{\mathcal{P}=0}$. Then the evolution of box's energy density $\mathcal{E}$ as we move along the soliton family is always negative.[5] So the energy conditions are not obeyed for this choice of normalization N. Let us now insist that we want that all energy conditions (4.5.4) are obeyed. This is achieved, for example, if we choose N to be $N = N\big|_{\mathcal{P}=0} + \frac{13}{100}$. Then the energy density and pressure of the box are always positive along the whole soliton family, as displayed in the top panel of Fig. 3.5.1. Given this choice, we can also compute the ADM energy and charge of our solutions as read at the asymptotic region. These quantities are displayed in the bottom panel of Fig. 3.5.1. This case illustrates a feature that we found to be common in similar exercises we did (for other choices of N and e): the ADM mass of hairy solitons can be negative, although our solutions are regular everywhere except across the box (where the extrinsic curvature has a discontinuity). For reference, recall that Schwarzschild and RN solutions with negative ADM mass are singular solutions. This might well be a general feature of hairy solutions confined by generic gravitational potentials (e.g. hairy solutions, not necessarily with spin 0, that are confined by more realistic potentials like the one of a massive scalar field or an accretion disk wall).

[5] This pressure vanishes in the limit $\epsilon \to 0$ because in this case the interior solution is simply Minkowski space with no scalar hair and thus the box is not necessary.

Part IV

Phase diagram of charged black hole bomb system

CHAPTER 4

Hairy black holes in a Minkowski box

4.1 Introduction

The black hole bomb setup was designed by Zel'dovich [135] and Press and Teukolsky [35] (see also [109]) very much in the aftermath of having understood the mathematical theory of black hole perturbations around a Kerr black hole. It emerged naturally from the fact that the wave analogue of the Penrose-Christodoulou process [136, 137] – superradiant scattering – unavoidably occurs in rotating black holes with non-zero angular velocity Ω. If a scalar wave with frequency ω and azimuthal number m satisfying $\omega < m\Omega$ is trapped near the horizon by the potential of a box (for example), the multiple superradiant amplifications and reflections at the cavity lead to an instability. The wave keeps extracting energy and angular momentum from the black hole interior and these accumulate between the horizon and the cavity. Press and Teukolsky assumed that this build-up of radiation pressure would rise to levels that could no longer be supported by the box and the latter would eventually break apart. But the black hole bomb system does not necessarily need to have such a dreadful end. Actually, more often than not, a black hole instability is a pathway to find new solutions that are stable to the original instability, have more entropy (for given energy and angular momentum) and are thus natural candidates for the endpoint or metastable states of the instability time evolution. This is certainly the case for superradiant fields trapped by the anti-de Sitter (AdS) gravitational potential [138, 139, 140, 141, 142, 143] or massive fields in asymptotically flat black holes [144]. So we can expect the same in the original black hole bomb system.

Motivated by these considerations, we would like to find the full phase diagram of solutions that can exist in the original black hole (BH) bomb system. By this we mean to find all possible stationary solutions of the theory with boundary conditions that confine the scalar field inside the box. These would be the non-linear version of the floating solutions in equilibrium that are described in [35]. This certainly requires solving PDEs. Therefore, in this paper, we start by considering a simpler system that still has a superradiant scalar field trapped inside a box but those properties can be found solving simply ODEs. This is possible if we first place a Reissner-Nordström black hole (RN BH) with chemical potential μ inside a box and then perturb it with a scalar field with charge q and frequency ω. As long as $\omega < q\mu$, a superradiant instability will also develop leading to the charged version of the black hole bomb system [36]. We thus want to find the phase diagram of static solutions of this system, including those with a scalar condensate floating above the horizon. The latter hairy solutions might have higher entropy than the original RN BH for a given energy and charge where they coexist. If so they would be a natural candidate for the endpoint of the charged black hole bomb instability, as long as it is checked that one can in principle build boxes – with an Israel stress tensor [113, 114, 115, 116] that satisfies the relevant energy conditions [5] – that holds the internal radiation pressure without breaking apart. This will further guarantee that we can insert this boxed system in an exterior Reissner-Nordström background, as required by Birkhoff's theorem [33, 34].

Looking into the details of this programme we immediately find new physics. Indeed, a linear perturbation analysis of the Klein-Gordon equation in an RN BH finds that the system is not only unstable to superradiance but also to the near-horizon scalar condensation instability [40]. These two instabilities are typically entangled for generic RN BHs but there are two corners of the phase space where they disentangle and reveal their origin. Indeed, extremal RN BHs with arbitrarily small horizon radius only have the superradiant instability since the near-horizon instability is suppressed as inverse powers of the horizon radius. On the opposite corner, RN BHs with a horizon radius close to the box radius only have the near horizon instability. Essentially, this instability is triggered by scalar fields that violate the near horizon AdS_2 Breitenlöhner-Freedman (BF) bound [28] of the extremal RN BH. It was originally found by Gubser [53] in planar AdS backgrounds (in a study that initiated the superconductor holographic programme) but it exists in other BH backgrounds (independently of the cosmological constant sign) with an extremal (zero temperature) configuration (see *e.g.* [56, 44]).

Analysing the setup of the black hole bomb system leads to the observation that the theory also has horizonless solutions if we remove the RN BH but leave the scalar field inside a box with a Maxwell field. Indeed, we can certainly perturb a Klein-Gordon field in a cavity and the frequencies that can fit inside it will be naturally quantized and real. This suggests that, within perturbation theory, we can then back-react this linear solu-

tion to higher orders where it will source non-trivial gravitoelectric fields that are regular inside (and outside) the box [43]. These are the boson stars of the theory, also known as solitons (depending on the chosen $U(1)$ gauge; see *e.g.* [145] for a review on boson stars). This perturbative analysis is bound to capture only small mass/charge boson stars. But a full numerical nonlinear analysis can identify the whole phase space of boson stars [2]. This analysis further reveals that the phase diagram of boson stars is quite elaborate with distinct boson star families. In particular, it finds that the phase diagram of solitons depends non-trivially on a total of four critical scalar field charges. Two of them can be anticipated using simple heuristic arguments on the aforementioned superradiant and near-horizon instabilities, but the two others only emerge after solving the non-linear equations of motion.

Coming back to our main subject of study, an RN BH placed inside a box is also the starting point to discuss and find the hairy BHs of the theory. The latter have a scalar condensate floating above the horizon that is balanced against gravitational collapse by electric repulsion. A box with appropriate Israel junction conditions and stress tensor [113, 114, 115, 116] should be able to confine the scalar condensate in its interior, and it should then be possible to place the the whole boxed system in a background whose exterior solution is the RN solution. In the present paper, we confirm that this is indeed the case and we find the full phase diagram of static solutions of the charged black hole bomb system. It turns out that the aforementioned four critical scalar electric charges play a relevant role also in the phase diagram of hairy BHs. Indeed, this diagram is qualitatively distinct depending on which one of the four available windows of critical charges the scalar charge q falls into. Ultimately, the reason for this dependence follows from the fact – that we will establish– that all hairy BHs that have a zero horizon radius limit choose to terminate on the boson star of the theory (which is fully specified once q is given), in the sense that the zero entropy hairy BHs have the same (Brown-York [112] quasilocal) mass and charge as the boson star. In our system, this materializes the idea that, often, small hairy BHs can be thought of as a small BH (RN or Kerr BH) placed on top of a boson star, as long as they have the same thermodynamic potential (chemical potential or angular velocity) to have the two constituents in thermodynamic equilibrium.

One of the four hairy BH families that we find in this paper was already identified in the perturbative analysis of [43]. This is the only family of hairy BHs that extends to arbitrarily small mass and charge, thus making it prone to be captured by the perturbative analysis about an empty box with an electric field. But the other three families and their intricate properties cannot be captured by such a theory because they are not perturbatively connected to the zero mass/charge solution.

Perhaps the most important property of the hairy BHs of the charged black hole bomb is that, when both coexist, they *always* have higher entropy than the RN BH that has

the *same* mass and charge. Therefore, we will conclude that hairy BHs are always the preferred thermodynamic phase of the theory in the microcanonical ensemble.

Very much like black holes confined in a box can be a starting point to discuss certain aspects of black hole thermodynamics [146, 147, 148, 149, 150, 151, 152] they should also be useful to understand generic superradiant systems where distinct (including perhaps some astrophysical) potential barriers confine fields [35]. These two are related since the hairy solutions describe non-linear systems where the central solution is in thermodynamic equilibrium with the floating scalar radiation. In particular, we can expect that hairy solutions of the charged black hole bomb provide a toy model with *some* universal features for the phase diagram of other confined unstable systems. Actually, we find that the present phase diagram shares many common features with the phase diagram of superradiant hairy black holes in global anti-de Sitter [46, 45, 110, 44, 93, 47, 111, 39].

The plan of this part is as follows. In section 4.2 we summarize in two figures the main properties of the phase diagram of hairy black holes and boson stars. In section 4.3 we formulate the exact setup of our system. The discussion only includes aspects that guarantee that our exposition is self-contained and more details can be found in [43]. In section 4.4 we explicitly construct the hairy black hole solutions in the four relevant windows of scalar charge that, together with the boson star study of [2] (and of the previous chapter), allow us to arrive to the conclusions summarized in section 4.2. Finally, in section 4.5 we explain how data of the hairy solution inside the box can be used to find the Israel stress tensor of the cavity surface layer and be matched with the exterior Reissner-Nordstöm solution.

4.2 Summary of phase diagram of boson stars and black holes in a cavity

The Einstein−Maxwell−Klein-Gordon theory, whereby the scalar field is confined inside a box of radius L in an asymptotically flat background, is fully specified once we fix the mass and charge q of the scalar field. We consider massless scalar fields with dimensionless electric charge $e = qL$ (the system has a scaling symmetry that allows us to measure all physical, *i.e.* dimensionless, quantities in units of L). By Birkhoff's theorem [33, 34] [1], outside the cavity the hairy solutions we search for are necessarily described by the RN solution. Thus, we just need to find the hairy solutions inside the box and then confirm that the Israel junctions conditions required to confine the scalar condensate inside the cavity, while having an exterior RN solution, correspond to an Israel energy-

[1] Birkhoff's theorem for Einstein-Maxwell theory states that the unique spherically symmetric solution of the Einstein-Maxwell equations with non-constant area radius function r (in the gauge (4.3.1)) is the Reissner-Nordström solution. If r is constant then the theorem does not apply since one has the Bertotti-Robinson ($AdS_2 \times S^2$) solution.

momentum stress tensor (proportional to the extrinsic curvature jump across the box layer [113, 114, 115, 116]) that is physical, *i.e.* that satisfies relevant energy conditions [5].

Since the solution outside the box is described by the RN solution, we cannot use the Arnowitt-Deser-Misner (ADM) mass M and charge Q [7] to differentiate the several solutions of the theory. However, we can use the Brown-York quasilocal mass $\mathcal{M}$ and charge $\mathcal{Q}$ [112], computed at the box location and normalized in units of L, and associated phase diagram $\mathcal{Q}$-$\mathcal{M}$ to display and distinguish the solutions of the theory. These quasilocal quantities obey their own first law of thermodynamics that is used to (further) check the results. In the quasilocal phase diagram, the extremal RN 1-parameter family of BHs (with horizon inside the box) provides a natural reference to frame our discussions. In particular, because distinct solutions often pile up in certain regions of the phase diagram, for clarity we will find it useful to plot $\Delta\mathcal{M}/L$ *vs* $\mathcal{Q}/L$ where $\Delta\mathcal{M} = \mathcal{M} - \mathcal{M}\big|_{\text{ext RN}}$ is the mass difference between the hairy solution and the extremal RN that has the *same* $\mathcal{Q}/L$. Therefore, in this phase diagram $\mathcal{Q}$-$\Delta\mathcal{M}$, the horizontal line with $\Delta\mathcal{M} = 0$ represents the extremal RN BH solution. Its horizon at R_+ fits inside the box of radius L if $R_+ \leq 1$ (which corresponds to $\mathcal{Q}/L \leq 2^{-1/2}$) and non-extremal RN BHs exist above this line. However, horizons of non-extremal RN BHs fit inside the box ($R_+ \leq 1$) if and only if their quasilocal charges are to the left of the red dashed line that we will display in our $\mathcal{Q}$-$\Delta\mathcal{M}$ diagrams. Actually, it turns out that this line also represents the maximal quasilocal charge that hairy solutions enclosed in the box can have.

We find that the spectrum of hairy black holes and boson stars of the theory is qualitatively distinct depending on whether e is smaller or bigger than four pivotal critical scalar field charges – e_{NH}, e_γ, e_c and e_{S} – which obey the relations $0 < e_{\text{NH}} < e_\gamma < e_c < e_{\text{S}}$.

Two of these critical charges, e_{NH} and e_{S}, can be identified simply studying linear scalar field perturbations about an RN BH in a box. Such RN BHs can be parametrized by the chemical potential μ and dimensionless horizon radius $R_+ = r_+/L$. These parameters are constrained to the intervals $0 \leq \mu \leq \mu_{\text{ext}}$ (with the upper bound being the extremal configuration) and $0 < R_+ \leq 1$. Boxed RN BHs become unstable – the black hole bomb system – if e is above the instability onset charge $e_{\text{onset}}(\mu, R_+)$. Instead of displaying $e_{\text{onset}}(\mu, R_+)$, it proves to be more clear to display the 2-dimensional plot $e_{\text{onset}}(R_+)$ for fixed values of μ. A sketch of this plot is given in the left panel of Fig. 4.2.1 (which reproduces the exact results in Fig. 2 of [40]). The minimal onset charge is attained for extremal RN black holes ($\mu = \mu_{\text{ext}}$): this is the orange curve that connects points $(0, e_{\text{S}})$ and $(1, e_{\text{NH}})$. For completeness, in the left panel of Fig. 4.2.1 we also sketch the onset charge curves (green dashed) for two non-extremal RN BHs at fixed $\mu < \mu_{\text{ext}}$. Naturally, this onset charge increases as we move away from extremality. Moreover, we see that the (extremal) minimal onset curve terminates at two critical charges that, actually, can be computed analytically:

- $e = e_{\rm NH} = \frac{1}{2\sqrt{2}} \sim 0.354$. This is the charge above which scalar fields can trigger a violation of the near horizon AdS_2 Breitenlöhner-Freedman (BF) bound [28, 53, 56, 44] of the extremal RN black hole whose horizon radius approaches, from below, the box radius. For details on how to derive this critical charge please see the introduction, section 1.4.2.

- $e = e_{\rm S} = \frac{\pi}{\sqrt{2}} \sim 2.221$. This is the critical charge above which scalar fields can drive arbitrarily small RN BHs unstable via superradiance. For a detailed analysis that leads to this critical charge, please see the discussion surrounding section 1.4.1.

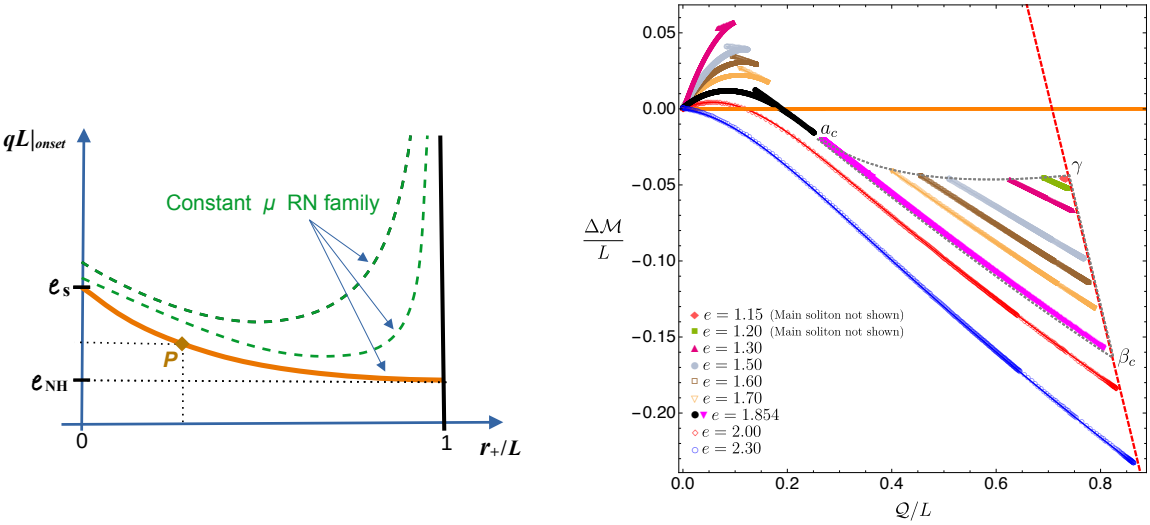

Figure 4.2.1: **Left panel:** Sketch of the scalar field electric charge $e_{\rm onset} = qL|_{\rm onset}$ as a function of the horizon radius $R_+ = r_+/L$ of RN BHs in a box (sketched from Fig. 2 of [40]). The orange curve in the bottom – that starts at $(R_+, e) = (0, e_{\rm S})$ and terminates at $(R_+, e) = (1, e_{\rm NH})$ – describes the *minimal* onset charge (which occurs at extremality, $\mu = \mu_{\rm ext}$). That is, (non-)extremal RN BHs can be unstable if an only if e is higher than the one identified by this orange onset curve. On the other hand, if we pick an RN family with constant $\mu < \mu_{\rm ext}$, for instability, we need e to be higher than the associated green dashed line $e(R_+)|_{{\rm const}\,\mu}$ also shown. In particular, we see that if we chose a charge in the range $e_{\rm NH} \leq e \leq e_{\rm S}$, RN BHs are unstable if and only if they are between the orange minimal onset curve and the horizontal line to the right of point P (gold diamond). **Right panel:** a survey of boson stars for different values of e, for $e \geq e_\gamma$ [2]. For a given $e \in [e_\gamma, e_c[$ we have the main (perturbatively connected to $(0,0)$) and secondary (non-pertubative) solitons (which only exist in the region bounded by the auxiliary grey dashed closed curve $a_c\beta_c\gamma$). The secondary soliton curve with e just above e_γ is close to the point γ, while the soliton with $e = 1.854$, just below e_c, is the magenta curve (very close to $a_c\beta_c$). Note that the gap in $\mathcal{Q}/L$ between the main soliton and the secondary one starts very large at $e = e_\gamma$ but then decreases and goes to zero precisely at $e = e_c$.

The system has two other critical charges, e_γ and e_c, that are uncovered when we do a detailed scan of the boson stars (a.k.a. solitons) of the theory. This task was completed in detail in the previous chapter. A phase diagram that summarizes the relevant properties for the present study is displayed in the right panel of Fig. 4.2.1 [2]. Note that it stores

in a single plot the solitons for *different theories*, *i.e.* for several distinct values of e. Two families of ground state boson stars (*i.e.* with smallest energy for a given charge) where found in part III. One is the *main or perturbative boson star family* which can be found within perturbation theory if we back-react a normal mode of a Minkowski cavity to higher orders. In the right panel of Fig. 4.2.1, these are the solitons that are continuously connected to $(\mathcal{Q}, \Delta\mathcal{M}) = (0, 0)$. The other one is the *secondary or non-perturbative soliton family*. In Fig. 4.2.1, these are the solitons that exist only above a critical $\mathcal{Q}$ and terminate at the red dashed line. The main/perturbative soliton family exists for any value of $e > 0$ (the system has a symmetry that allows us to consider only $e > 0$). However, the secondary/non-perturbative soliton only exists for scalar field charges that are in the window $e_\gamma \leq e < e_c$. These are the solitons enclosed in the region $a_c\beta_c\gamma$ (*i.e.* inside the auxiliary grey dashed closed line with these vertices). They only exist above $e_\gamma \sim 1.13$ (see point γ) and below $e_c \simeq 1.854 \pm 0.0005$ (see line $a_c\beta_c$ just below the magenta line). Below e_γ the non-perturbative solitons do not exist because they no longer fit inside the box. Above e_c, the gap between the two soliton families ceases to exist, *i.e.* the non-perturbative soliton merges with the perturbative soliton, and the ground state boson stars of the theory extend from the origin all the away to the red dashed line (see *e.g.* the red diamond curve with $e = 2$ or the blue circle curve with $e = 2.3$ in Fig. 4.2.1). Summarizing, the main/perturbative soliton has a Chandrasekhar limit for $0 < e < e_c$ but extends from the origin all the way to the red dashed line for $e \geq e_c$. On the other hand, the ground state secondary/non-perturbative soliton only exists in the window $e_\gamma \leq e < e_c$.

In the following sections we will find the hairy black holes of the Einstein−Maxwell−Klein-Gordon theory. We will conclude that, whenever the hairy black holes have a zero horizon radius limit, they terminate on a soliton. Accordingly, the phase diagram of solutions depends on the above four critical scalar field charges. Our main findings are summarized in the phase diagram sketches of Fig. 4.2.2 and the properties of these phase diagrams depend on the following 5 windows of scalar charge e:

1. $e < e_{\text{NH}} = \frac{1}{2\sqrt{2}} \sim 0.354$. From the left panel of Fig. 4.2.1, one concludes that RN is stable for $e < e_{\text{NH}}$ and thus no hairy BHs exist. The only non-trivial solutions of the theory are the RN BH and the main/perturbative boson star which has a Chandrasekhar limit (see details in part III).

2. $e_{\text{NH}} \leq e < e_\gamma \sim 1.13$. The phase diagram $\mathcal{Q}$-$\Delta\mathcal{M}$ of solutions for this window is sketched in the top-left panel of Fig. 4.2.2. The only boson star of the theory is the main/perturbative family $OABC\cdots$ (already present for $e < e_{\text{NH}}$) with its Chandrasekhar limit A and a series of cusps $A, B, C, \cdots$ and associated zig-zagged branches whose properties were studied in detail in part III. As the left panel of Fig. 4.2.1 indicates, RN BHs are now unstable for sufficiently large R_+ (*i.e.* $\mathcal{Q}$) if

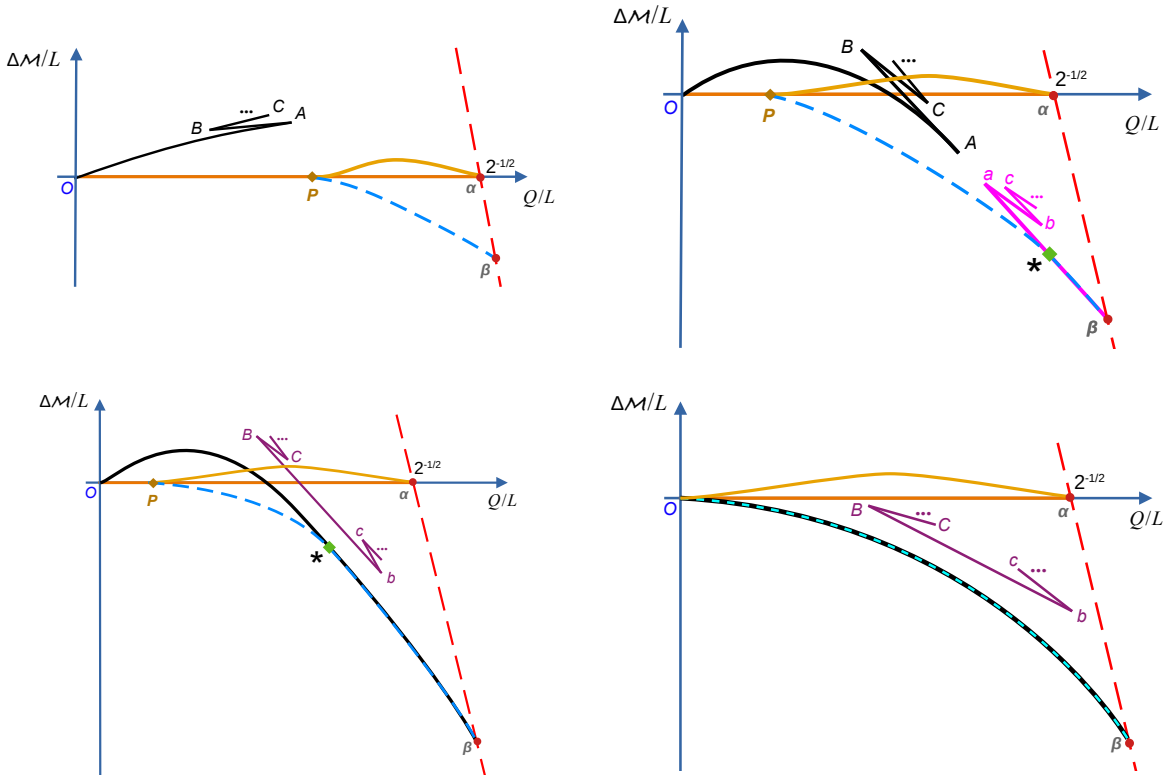

Figure 4.2.2: Sketch of the quasilocal phase diagram for black holes and solitons and as we span relevant windows of scalar field charge e. The critical charges are such that $0 < e_{\rm NH} < e_\gamma < e_c < e_{\rm S}$. The quantity $\Delta\mathcal{M}$ is the quasilocal mass difference between a given solution and an extremal RN BH that has the *same* quasilocal charge $\mathcal{Q}/L$. Hence the orange line at $\Delta\mathcal{M} = 0$ describes the extremal RN solution that must have $\mathcal{Q}/L \leq 2^{-1/2}$ to fit inside the box. The red dashed line represents the maximal quasilocal charge of solutions that can fit inside the box. It intersects the extremal RN line at $\mathcal{Q}/L = 2^{-1/2}$. Non-extremal RN BHs confined in the box have $\Delta\mathcal{M} > 0$ and fill the triangular region bounded by $\mathcal{Q} = 0$ and by the orange and red dashed lines (not shown completely). The main soliton family is always given by black curves that start at O. The secondary soliton family is given either by magenta or purple curves. Hairy black holes exist in the region $P\alpha\beta$ enclosed by the yellow merger line $P\alpha$ (between hairy and RN BHs), the blue curve $P\beta$ where the curvature grows large and the red line $\alpha\beta$. **Top-left panel:** case $e_{\rm NH} < e < e_\gamma$. **Top-right panel:** $e_\gamma \leq e < e_c$. **Bottom-left panel:** case $e_c \leq e < e_{\rm S}$. **Bottom-right panel:** case $e \geq e_{\rm S}$.

sufficiently close to extremality (see diamond point P and region bounded by the horizontal line to the right of P and the minimal onset curve). The onset of the instability translates into the yellow curve $P\alpha$ in the top-left panel of Fig. 4.2.2 and RN BHs below this onset curve $P\alpha$ (and above the extremal horizontal straight line $OP\alpha$) are unstable. Hairy BHs exist inside the region bounded by the closed curve $P\alpha\beta$. They merge with RN BHs at the onset $P\alpha$ of the instability and they extend for lower masses all the way down to the blue dashed line $P\beta$ where they terminate at *finite* entropy and *zero* temperature because the Kretschmann curvature at the

horizon blows up. They are also constrained to be to the left of the red dashed line $\alpha\beta$ because the horizon of the hairy BH must fit inside the cavity with unit normalized radius. Note that for this window of e there is no dialogue between the hairy boson star and the hairy BH families.

A representative example of a black hole bomb system with a charge $e = 1$ in this window $e_{\rm NH} \leq e < e_\gamma$ will be discussed in detail in section 4.4.1 and Fig. 4.4.1.

3. $e_\gamma \leq e < e_c \simeq 1.854 \pm 0.0005$. The phase diagram $\mathcal{Q}$-$\Delta\mathcal{M}$ of solutions for this window is sketched in the top-right panel of Fig. 4.2.2. Besides the main/perturbative family $OABC\cdots$ of boson stars (black line), the system now has the secondary/non-perturbative family $\beta abc\cdots$ of boson stars (magenta line) and there is a gap Aa between these two families. Precisely at e_γ, this gap Aa is the largest and the non-perturbative boson star family reduces to a single point β on top of the red-dashed line (*i.e.* it coincides with point γ in right panel of Fig. 4.2.1). As e grows beyond e_γ, the gap Aa decreases and it vanishes precisely at $e = e_c$ where the two ground state soliton families merge into a single one (as discussed in the previous chapter). As before, hairy BHs exist in the region enclosed by the closed line $P\alpha\beta$ with the yellow curve $P\alpha$ being again the instability onset curve where the scalar condensate vanishes and hairy BHs merge with the RN BH family. As before, the hairy BHs also extend for smaller masses all the way down to the singular blue dashed line $P\beta$ where the Kretschmann curvature at the horizon diverges. But this time, this singular curve $P\beta \equiv P\star\beta$ splits into two segments. Hairy BHs terminating in the trench $P\star$ do so at *finite entropy* and *zero temperature*, as all the hairy BHs with $e < e_c$. However, hairy BHs terminating at the trench $\star\beta$ do so at *zero entropy* and *unbounded temperature*. In such a way that in the $\mathcal{Q}$-$\Delta\mathcal{M}$ phase diagram, this hairy BH trench $\star\beta$ *coincides* with the secondary/non-perturbative soliton (magenta line between $\star$ and β). In this sense, we can say that hairy BHs with large charge ($\mathcal{Q} \geq \mathcal{Q}_\star$) terminate on the non-perturbative soliton. This point $\star$ coincides with β in the limit $e \to e_\gamma$ and it diverges away from β as e moves away from e_γ towards e_c.

A representative example of a black hole bomb system with a charge $e = 1.85$ in this window $e_\gamma \leq e < e_c$ will be discussed in detail in section 4.4.2 and Figs. 4.4.2–4.4.4.

4. $e_c \leq e < e_{\rm S} = \frac{\pi}{\sqrt{2}} \sim 2.221$. The phase diagram $\mathcal{Q}$-$\Delta\mathcal{M}$ of solutions for this window is sketched in the bottom-left panel of Fig. 4.2.2. At $e = e_c$ the perturbative and non-perturbative boson star families merge and for $e \geq e_c$ the boson star ground state is always the perturbative family $O\beta$ (black curve) that extends from the origin to the red dashed line. There is also a secondary family of boson stars $\cdots CBbc\cdots$ (purple curve) but it is not the ground state family and it plays no role

in the description of hairy BHs for this window of charges.

Hairy BHs exist inside the closed line $P\alpha\beta$. Again, the yellow curve $P\alpha$ is the instability onset curve where hairy BHs merge with the RN BH family. The hairy BHs extend for smaller masses all the way down to the singular blue dashed line $P\star\beta$ where the Kretschmann curvature at the horizon diverges. Hairy BHs terminating in the trench $P\star$ do so at *finite entropy* and *zero temperature*, while hairy BHs terminating at the trench $\star\beta$ do so at *zero entropy* and *unbounded temperature*. In the $\mathcal{Q}$-$\Delta\mathcal{M}$ phase diagram, this hairy BH trench $\star\beta$ *coincides* with the perturbative soliton (black line between $\star$ and β). In this sense, hairy BHs with large charge ($\mathcal{Q} \geq \mathcal{Q}_\star$) terminate on the perturbative soliton. As e increases above e_c, points P and $\star$ move to the left of the phase diagram, *i.e.* to lower values of $\mathcal{Q}$ and they approach the origin O as $e \to e_{\rm s}$.

A representative example of a black hole bomb system with a charge $e = 2$ in this window $e_c \leq e < e_{\rm s}$ will be discussed in detail in section 4.4.3 and Figs. 4.4.5–4.4.7.

5. $e \geq e_{\rm s}$. The phase diagram $\mathcal{Q}$-$\Delta\mathcal{M}$ of solutions for this window is sketched in the bottom-right panel of Fig. 4.2.2. Precisely at $e_{\rm s}$, the slope $\mathrm{d}\Delta\mathcal{M}/\mathrm{d}\mathcal{Q}$ of the main/perturbative boson star family (black curve $O\beta$) vanishes at the origin and for $e \geq e_{\rm s}$, perturbative boson stars always have $\Delta\mathcal{M} < 0$. Not less importantly, at and above $e_{\rm s}$, all extremal RN BHs are unstable, *i.e.* point P seen in the plots for $e < e_{\rm s}$ hits the origin O. Consequently, hairy BHs now exist for all values of $\mathcal{Q}$ (that can fit inside the cavity), *i.e.* in the 2-dimensional region with boundary $O\alpha\beta$. And, for any $\mathcal{Q}$, hairy BHs bifurcate from RN at the instability onset $O\alpha$ (yellow curve) and extend for smaller masses till they terminate – with zero entropy, and divergent temperature and Kretschmann curvature – along a curve (blue dashed line) that *coincides* with the boson star curve $O\beta$ (black curve).

A representative example of a black hole bomb system with a charge $e = 2.3$ in this window $e \geq e_{\rm s}$ will be discussed in detail in section 4.4.4 and Figs. 4.4.8–4.4.10.

In the summary above we alluded to points in the phase diagram where the solution has *zero entropy* and *unbounded temperature*. Classically, for example in Schwarzschild spacetime, when the black hole shrinks to zero entropy, it also does so with an unbounded temperature, this corresponds to a naked curvature singularity (and the spacetime becomes Minkowski). Therefore we will interpret our zero entropy, unbounded temperature solutions in the same way, as a naked curvature singularity (note that the Kretschmann curvature invariant also blows up here, as we will show in the main text). Further support for this physical interpretation is Independently of e, a universal property of hairy BHs is that, when they coexist with boxed RN BHs, they always have higher

entropy than the boxed RN BH with the same quasilocal mass and charge. That is to say, in the phase space region where they exist, hairy BHs are always the dominant phase in microcanonical ensemble. Moreover, hairy BHs are stable to the superradiant mode that drives the boxed RN BH unstable. It follows from these observations and the second law of thermodynamics that the endpoint of the superradiant/near horizon instability of the boxed RN BH, when we do a time evolution at constant mass and charge, should be a hairy BH. It would be interesting to confirm this doing time evolutions along the lines of those done in [132, 133, 134].

The present work can be seen as the final study of a sequence of works on the charged black hole bomb system. Ref. [40] started by studying the linear superradiant and near-horizon instabilities of the boxed RN BH. This identified the zero-mode and growth rates of these instabilities. The hairy boson stars and hairy BHs were found within perturbation theory in [43]. As expected, this perturbative analysis is valid only for small condensate amplitudes and small horizon radius and thus it is able to capture only small energy/charge hairy solutions. Therefore, for the solitons, the perturbative analysis can capture the main or perturbative boson star family at small mass/charge. But it misses: 1) the existence of the Chandrasekhar limit and cusps of this family, 2) the existence of the secondary or non-perturbative boson star family, and 3) it misses the existence of two important critical charges e_γ and e_c where the non-perturbative soliton starts existing and merges with the perturbative family. These properties were only identified once the Einstein−Maxwell−Klein-Gordon equation was solved fully non-linearly in [2]. On the other hand, the perturbative analysis of [43] also finds the hairy BHs that, for $e \geq e_{\rm s}$, are perturbatively connected to a Minkowski spacetime with a cavity. By construction, these perturbative hairy BHs reduce, in the zero horizon radius limit, to the boson star of the theory. However, the perturbative analysis of [43] says nothing about the hairy BHs of the theory when $e < e_{\rm s}$. In the present chapter, we fill this gap.

In the introduction we already mentioned that the potential barrier that confines the scalar condensate in our boxed or black hole bomb system might be a good toy model for other systems with potential barriers that provide confinement. In particular, we find that the phase diagram of hairy boson stars and BHs in the black hole bomb system is qualitatively similar to the one found for asymptotically anti-de Sitter solitons [46, 45, 110, 44, 93, 47, 111, 39]. In this latter case, the AdS boundary conditions act as a natural gravitational box with radius inversely proportional to the cosmological length that provides bound states. In this sense, our work also complements and completes previous AdS studies since the existence range of the secondary/non-perturbative boson star family, its merger with the main/perturbative soliton at $e = e_c$, and the fact that hairy BHs can also terminate on this soliton family for $e_\gamma \leq e < e_c$ was not established in detail in [46, 45, 110, 44, 93].

4.3 Setting up the black hole bomb boundary value problem

The setup of our problem was already discussed in the perturbative analysis of the problem in [43]. Here, to have a self-contained exposition, we discuss only the key aspects needed to formulate the problem and the strategy to compute physical quantities without ambiguities. We ask the reader to see [43] for details.

4.3.1 Einstein-Maxwell gravity with a confined scalar field

We already described Einstein-Maxwell theory with a confined scalar field in section 3.3. We want to find black hole solutions of (3.3.1) that are static, spherically symmetric and asymptotically flat. $U(1)$ gauge transformations allow us to work with a real scalar field and a gauge potential that vanishes at the horizon. Further choosing the Schwarzschild gauge, an *ansatz* with the desired symmetries is then

$$\mathrm{d}s^2 = -f(r)\mathrm{d}t^2 + g(r)\mathrm{d}r^2 + r^2\mathrm{d}\Omega_2^2, \qquad A_\mu \mathrm{d}x^\mu = A(r)\mathrm{d}t, \qquad \phi = \phi^\dagger = \phi(r), \tag{4.3.1}$$

with $\mathrm{d}\Omega_2^2$ being the metric for the unit 2-sphere (expressed in terms of the polar and azimuthal angles $x = \cos\theta$ and φ). The scalar field is forced to be confined inside a box of radius L. The system then has a scaling symmetry that allows us to normalize coordinates ($T = t/L, R = r/L$) and thermodynamic quantities in units of L, and place the box at radius $R = 1$ [43].

In these conditions, the equations of motion that follow from (3.3.1) can be found in the introduction, equations (1.4.24) after plugging in the algebraic relation for the field g (1.4.28). These are a set of three ordinary differential equations for the fields $f(R)$, $A(R)$ and $\phi(R)$, and an algebraic equation that expresses $g(R)$ as a function of the other 3 fields and their first derivatives. Well-posedness of the boundary value problem requires that we give boundary conditions at the horizon and asymptotic boundary of our spacetime. Additionally, we must specify Israel junction conditions at the timelike hypersurface $\Sigma = R - 1 = 0$ where the box is located. Again, our hairy BHs have vanishing scalar field at and outside this box, $\phi(R \geq 1) = 0$.

The horizon, with radius $R = R_+ = \frac{r_+}{L}$ is the locus $f(R_+) = 0$. We have three second order ODEs and thus there are six free parameters when we do a Taylor expansion about the horizon. Regularity demands Dirichlet boundary conditions that set three of these parameters to zero. We are thus left with only three constants f_0, A_0, ϕ_0 (say) such that the regular fields have the Taylor expansion around the origin:

$$\begin{aligned} &f(R_+) = f_0(R - R_+) + \mathcal{O}((R-R_+)^2), \\ &A(R_+) = A_0(R - R_+) + \mathcal{O}((R-R_+)^2), \qquad \phi(R_+) = \phi_0 + \mathcal{O}((R-R_+)^2). \end{aligned} \tag{4.3.2}$$

The boundary conditions at the asymptotic boundary of our spacetime are no different from those described for the soliton.

In these conditions, hairy BHs of the theory are a 2-parameter family of solutions that we can take to be the horizon radius R_+ and the value of the (interior) derivative of the scalar field at the box, $\epsilon \equiv \phi'^{\,in}\big|_{R=1^-}$.

As mentioned in section 4.2, it follows from Birkhoff's theorem that in the asymptotic region our solutions are necessarily described by the RN solution. Therefore, the ADM mass M and charge Q cannot be used to distinguish the several solutions of the theory. It is thus natural to instead use the Brown-York quasilocal mass $\mathcal{M}$ and charge $\mathcal{Q}$, measured at the box, to display our solutions in a phase diagram of the theory [112]. From section II.C of [43], the Brown-York quasilocal mass and charge contained inside a 2-sphere with radius $R = 1$ are ($G_N \equiv 1$)[2]

$$\mathcal{M}/L = R\left(1 - \frac{1}{\sqrt{g(R)}}\right)\Big|_{R=1}, \qquad \mathcal{Q}/L = \frac{R^2 A_t'(R)}{2\sqrt{g(R)}f(R)}\Big|_{R=1}. \tag{4.3.3}$$

The thermodynamic description of our solutions is complete after defining the chemical potential, temperature and entropy:

$$\mu = A(1) - A(R_+)\,, \qquad T_H L = \lim_{R\to R_+} \frac{f'(R)}{4\pi\sqrt{f(R)g(R)}}, \qquad S/L^2 = \pi R_+^2, \tag{4.3.4}$$

where we work in the gauge $A(R_+) = 0$. These quantities must satisfy the quasilocal form of the first law of thermodynamics:

$$\mathrm{d}\mathcal{M} = T_H\,\mathrm{d}S + \mu\,\mathrm{d}\mathcal{Q}, \tag{4.3.5}$$

which is used to check our solutions.

As explained before, for reference we will often compare the hairy families of solutions against extremal RN BHs. RN BHs confined in a cavity can be parametrized by the horizon radius R_+ and the chemical potential μ, and their quasilocal mass and charge are [43]

$$\mathcal{M}/L\big|_{RN} = 1 - \frac{\sqrt{2}(1-R_+)}{\sqrt{2-(2-\mu^2)R_+}}, \qquad \mathcal{Q}/L\big|_{RN} = \frac{\mu R_+}{\sqrt{2}\sqrt{2-(2-\mu^2)R_+}}. \tag{4.3.6}$$

where $0 < R_+ \leq 1$ (for the horizon to be confined inside the box) and $0 \leq \mu \leq \mu_{\rm ext}$, with extremality reached at $\mu_{\rm ext} = \sqrt{2}$. Note that at extremality one has $\mathcal{M}/L = R_+$ and $\mathcal{Q}/L = R_+/\sqrt{2}$. On the other hand, for any μ, when $R_+ = 1$ one has $\mathcal{M}/L = 1$ and $\mathcal{Q}/L = 2^{-1/2}$.

[2] The Brown-York quasilocal quantities reduce to the ADM ones when we evaluate the former at $R \to \infty$.

The strategy to find the hairy BHs of the theory can now be outlined. The hairy solution inside the box is found integrating numerically the coupled system of three ODEs in the domain $R \in [R_+, 1]$. This is done while imposing the boundary conditions (4.3.2) at the horizon and, at the box, we impose $\phi(1^-) = 0$ and use the scaling symmetry (3.3.5) to set $f(1^-) = 1$. After this task, we can compute the quasilocal charges (4.3.3) and the other thermodynamic quantities (4.3.4) of the system. To find the solution in the full domain $R \in [R_+, \infty]$ we impose the three junction conditions (3.3.15a)-(3.3.15c) at the box to match the interior solution with the outer solution (described by the RN solution (3.3.6)). This operation finds the parameters M_0, C_A, ρ in (3.3.6) as a function of the reparametrization freedom parameter N introduced in (3.3.14). The Israel stress tensor $\mathcal{S}_a^b$ is just a function of N and, if $\phi^{in} \neq 0$, we cannot choose N to kill all the components of $\mathcal{S}_a^b$ (there are two non-vanishing components, $\mathcal{S}_t^t$ and $\mathcal{S}_\theta^\theta = \mathcal{S}_\varphi^\varphi$). In this process, we have arbitrary freedom to choose N. This simply reflects the freedom we have to select the energy-momentum content of the box needed to confine the scalar condensate. We should however, make a selection that respects some or all the energy conditions [5]. Once this choice is made, we can finally compute the ADM mass and charge (3.3.7) of the hairy solution which, necessarily, includes the contribution from the box.

4.3.2 Numerical scheme

The hairy BHs we seek are a 2-parameter family of solutions, that we can take to be the horizon radius R_+ and the scalar field amplitude $\epsilon \equiv \phi'(R = 1)$ as defined in (3.3.12). In practice, we set up a two dimensional discrete grid where we march our solutions along these two parameters. In other words, we give R_+ and ϵ as inputs of our numerical code, and in the end of the day we read the horizon parameters f_0, A_0, ϕ_0 in (4.3.2), and the values of the derivative of f and the value of A and its derivative at the box, $R = 1$. Typically, we start near the merger with the RN BH where a good seed (approximation) for the Newton-Raphson method we use is the RN BH itself but with a small perturbation that also excites the scalar field.

We find it convenient to introduce a new radial coordinate

$$y = \frac{R - R_+}{1 - R_+} \tag{4.3.7}$$

so that the event horizon is at $y = 0$ and the box at $y = 1$. The equations of motion now depend explicitly on R_+.

Moreover, we also find useful to redefine the fields as

$$f = y\, q_1(y), \qquad A = y\, q_2(y), \qquad \phi = -(1 - y)\, q_3(y) \tag{4.3.8}$$

which automatically imposes the boundary conditions (4.3.2) at the horizon. We now use the scaling symmetry (3.3.5) to set $f(1^-) = 1$ and introduce the scalar amplitude (3.3.12) at the box. This can be done through imposing the boundary conditions

$$q_1(1) = 1, \qquad q_3(1) = \epsilon. \tag{4.3.9}$$

The other boundary conditions for $q_{1,2,3}$ are derived boundary conditions in the sense that they follow directly from evaluating the equations of motion at the boundaries $y = 0$ and $y = 1$ [104]. Under these conditions, the hairy BHs are described by smooth functions $q_{1,2,3}$ that we search for numerically.

To solve numerically our boundary value problem, we use a standard Newton-Raphson algorithm and discretise the coupled system of three ODEs using pseudospectral collocation (with Chebyshev-Gauss-Lobatto nodes). The resulting algebraic linear systems are solved by LU decomposition. These numerical methods are described in detail in the review [104]. Since we are using pseudospectral collocation, and our functions are analytic, our results must have exponential convergence with the number of grid points. We check this is indeed the case and the thermodynamic quantities that we display have, typically, 8 decimal digit accuracy. We further use the quasilocal first law (4.3.5) (typically, obeyed within an error smaller than $10^{-3}\%$) to check our solutions.

4.4 Phase diagram of the charged black hole bomb system

The properties of the hairy black holes of the charged black hole bomb system are closely linked to the superradiant/near-horizon instability of RN black holes[3], so we first highlight some features of this instability to provide the context needed to interpret the hairy black hole phase diagram (see [40] for details).

In the left panel of Fig. 4.2.1 we sketch (from [40]) the scalar field instability onset charge $e_{\rm onset} = q_{\rm onset} L$ as a function of R_+ for three families of RN black holes with constant chemical potential μ, *i.e.* the minimum scalar charge needed for a black hole with (R_+, μ) to be unstable. We can see that the near-horizon charge $e_{\rm NH}$ is a lower bound for an RN instability, *i.e.* caged RN BHs are always stable when $e < e_{\rm NH}$. Correspondingly, we also find no hairy black holes when $e < e_{\rm NH}$. At the other end, all extremal RN black holes, no matter their R_+, are unstable at or above the superradiant charge $e_{\rm S}$. In between these two critical charges $e_{\rm NH} < e < e_{\rm S}$ we have a window of horizon radii $R_+ \in [R_+|_P, 1]$ within which sufficiently near-extremal RN black holes are unstable. In equivalent words, for $e_{\rm NH} < e < e_{\rm S}$, extremal RN BHs are unstable for quasilocal charges in the range $\mathcal{Q}/L \in [(\mathcal{Q}/L)|_P, 2^{-\frac{1}{2}}]$. In the upcoming phase diagrams we will indicate the instability onset

[3] For a general RN black hole the superradiant and near-horizon instabilities are entangled, so we will simply refer to an RN instability, regardless of the origin.

curve as a yellow curve $P\alpha$ and, when applicable, we will also use a gold diamond point P to identify the minimum charge for instability. The onset curve starts at point P where it intersects the extremal RN curve and terminates at point α with $\mathcal{Q}/L = 2^{-\frac{1}{2}}$ (*i.e.* $R_+ = 1$) where it intersects again the extremal RN curve.

In all our plots $\mathcal{M}$ and $\mathcal{Q}$ are the Brown-York quasilocal mass and charge (4.3.3) of the system, measured at the location of the box. Different solutions tend to pile up in certain regions of the $\mathcal{Q}$-$\mathcal{M}$ diagram (as illustrated in Fig. 4.4.8). Thus, the distinction between different solutions becomes clearer if we use instead $\Delta\mathcal{M} = \mathcal{M} - \mathcal{M}\big|_{\text{ext RN}}$ which is the quasilocal mass difference of a hairy solution with an extremal RN that has the *same* quasilocal charge $\mathcal{Q}$. Thus, in our $\mathcal{Q}$-$\Delta\mathcal{M}$ plots, the horizontal orange line $O\alpha$ with $\Delta\mathcal{M} = 0$ describes the extremal RN BH. It is constrained to have $\mathcal{Q}/L \leq 2^{-\frac{1}{2}}$ (point α) in order to fit inside the box (this extremal line will be represented by a dark red line in the 3-dimensional plots $\mathcal{Q}$-$\Delta\mathcal{M}$-S).

From the RN quasilocal charges (4.3.6), in the quasilocal $\mathcal{Q} - \mathcal{M}$ plot, the region that represents RN BHs with horizon radius inside the box is the triangular surface bounded by the lines $\mathcal{Q} = 0$, $\mathcal{M} = \sqrt{2}\mathcal{Q}$ and $\mathcal{M}/L = 1$. Therefore, in the $\mathcal{Q} - \Delta\mathcal{M}$ plane, non-extremal RN BHs with $R_+ \leq 1$ are those inside the triangular region bounded by $\mathcal{Q} = 0$, $\Delta\mathcal{M} = 0$ and $\Delta\mathcal{M}/L = 1 - \sqrt{2}\mathcal{Q}/L$. The boundary $\mathcal{Q} = 0$ describes the Schwarzschild limit, $\Delta\mathcal{M} = 0$ is the extremal RN boundary and the latter curve is

$$(\mathcal{Q}/L, \Delta\mathcal{M}/L) = \left(L^{-1}\mathcal{Q}\big|_{\text{ext RN}}, 1 - L^{-1}\mathcal{M}\big|_{\text{ext RN}}\right) = \left(\frac{R_+}{\sqrt{2}}, 1 - R_+\right) \tag{4.4.1}$$

where $\mathcal{M}\big|_{\text{ext RN}}$ and $\mathcal{Q}\big|_{\text{ext RN}}$ are given by (4.3.6) with $\mu = \mu_{\text{ext}} = \sqrt{2}$. The red dashed line in the forthcoming $\mathcal{Q} - \Delta\mathcal{M}$ plots is this parametric curve (4.4.1) with R_+ *allowed to take also values above 1*. Indeed, it turns out that the most charged solutions we find approach this dashed red line (4.4.1) (in the limit where scalar condensate amplitude approaches infinity). In this sense, for a given quasilocal mass (smaller than 1), this red dashed line (4.4.1) represents the maximal quasilocal charge that confined solutions can have, with or without scalar hair.

As discussed in our summary of results (section 4.2), the charged black hole bomb system has a total of four critical scalar field charges. Besides $e_{\text{NH}} = \frac{1}{2\sqrt{2}} \sim 0.354$ and $e_{\text{S}} = \frac{\pi}{\sqrt{2}} \sim 2.221$ discussed above, the two others are $e_\gamma \sim 1.13$ and $e_c \sim 1.8545 \pm 0.0005$. Accordingly, the phase diagram of hairy boson stars and hairy black holes depends on the value of e compared to these four fundamental critical charges of the system. Thus, in the next subsections, we describe the properties of hairy solutions in the following four windows of scalar field charge: 1) $e_{\text{NH}} \leq e < e_\gamma$, 2) $e_\gamma \leq e \leq e_c$, 3) $e_c \leq e < e_{\text{S}}$, and 4) $e \geq e_{\text{S}}$. For concreteness, we will display results for a representative value of e for each one of these windows, namely: 1) $e = 1$ (section 4.4.1), 2) $e = 1.85$ (section 4.4.2), 3)

$e = 2$ (section 4.4.3), and 4) $e = 2.3$ (section 4.4.4). Altogether, these results (and others not presented) will allow us to extract the conclusions summarized in section 4.2.

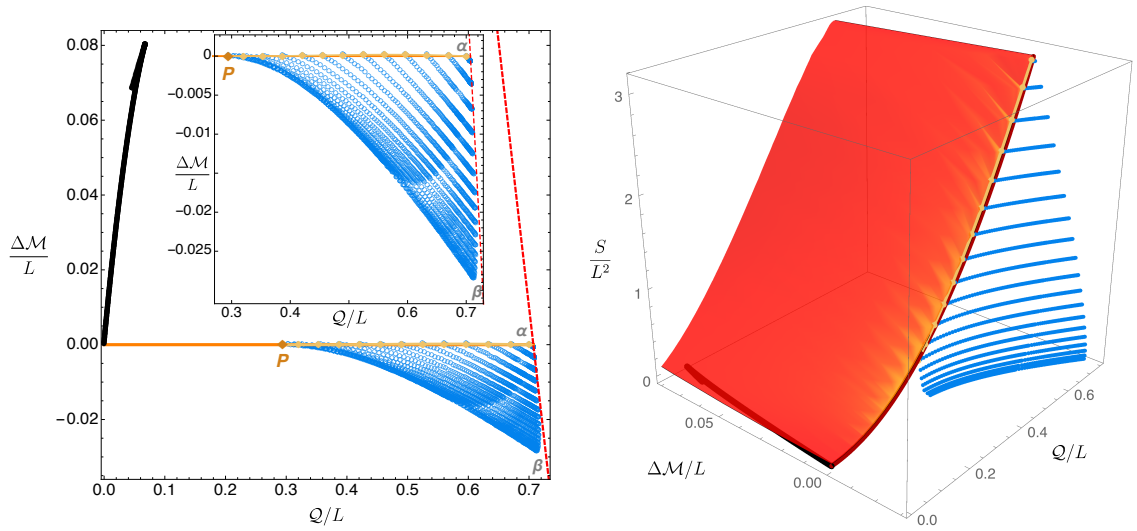

Figure 4.4.1: Phase diagrams for Einstein-Maxwell theory with a scalar field charge $e = 1$ ($e_{\rm NH} \leq e < e_\gamma$) in a Minkowski box. **Left panel:** Quasilocal mass difference $\Delta\mathcal{M}/L$ as a function of the quasilocal charge $\mathcal{Q}/L$. The black disk curve is the main/perturbative soliton family, the orange line is the extremal RN BH (RN black holes exist above it), and the blue circles describe hairy black holes. The yellow curve is the superradiant onset curve of RN (just above but very close to the extremal RN curve with the two merging at P and α). It agrees with the hairy solutions in the limit where these have $\epsilon = 0$ (no scalar condensate) and thus merge with RN family. The red dashed line with negative slope signals solutions with $\Delta\mathcal{M}/L = 1 - \sqrt{2}\mathcal{Q}/L$ *i.e.* black holes with horizon radius $R_+ = 1$ (above this value they do not fit inside the cavity). **Right panel:** Dimensionless entropy S/L^2 as a function of the quasilocal charge and mass difference. RN BHs are the two parameter red surface with extremality described by the 1-parameter curve $\Delta\mathcal{M} = 0$ (dark red). The instability onset is described by the yellow curve (very close to extremality) and RN between these two curves are unstable. When they coexist with RN BHs, for a given $(\mathcal{Q}, \mathcal{M})/L$, hairy BHs (blue dots) always have more entropy than RN, *i.e.* they dominate the microcanonical ensemble. For $e_{\rm NH} \leq e < e_\gamma$, hairy BHs terminate at an extremal BH (*i.e.* with zero temperature) and finite entropy (and divergent horizon curvature). The soliton (black dots) with zero entropy is also shown.

4.4.1 Phase diagram for $e_{\rm NH} \leq e < e_\gamma$

The left panel of Fig. 4.4.1 is the phase diagram $\mathcal{Q}$-$\Delta\mathcal{M}$ for $e = 1$, representative of the range $e_{\rm NH} \leq e < e_\gamma$. The black disk curve describes the only family of boson stars of the theory for this (range of) e which is the main/perturbative family. This corresponds to the black curve $OABC\cdots$ (already present for $e < e_{\rm NH}$) with its Chandrasekhar limit A and a series of cusps $A, B, C, \cdots$ and associated zig-zagged branches sketched in the top-left panel of Fig. 4.2.2. The properties of this boson star were already studied in much detail in the previous chapter therefore we do not expand further. Our interest here are the hairy BHs.

The horizontal orange curve $OP\alpha$ with $\Delta\mathcal{M} = 0$ is the extremal RN BH family with $\mathcal{Q}/L \leq 2^{-\frac{1}{2}}$ and boxed non-extremal BHs exist above this line and to the left of the red dashed line (4.4.1) to fit inside the cavity, as detailed above. The yellow curve $P\alpha$, that intersects and terminates at the extremal RN curve precisely at P and α, describes the instability onset curve of RN BHs as computed using linear analysis in [40]. It coincides with the merger line of hairy BHs with RN BHs, as it had to. Indeed, recall that hairy BHs can be parametrized by their horizon radius R_+ and the scalar field amplitude ϵ. When $\epsilon = 0$ we recover the 1-parameter family $P\alpha$ of RN BHs at the onset of the instability. RN BHs are unstable below this onset curve $P\alpha$ all the way down to the horizontal extremal line also labelled $P\alpha$. This region is extremely small for this value of e but it will be wider as e increases.

Hairy BHs (blue circles) exist inside the closed line $P\alpha\beta$. That is, they exist below the onset curve $P\alpha$ and to the left of the red $\alpha\beta$ dashed line (4.4.1), all the way down till they reach a line $P\beta$ where the Kretschmann curvature scalar evaluated at the horizon $K|_{\mathcal{H}} = R_{abcd}R^{abcd}\big|_{R_+}$ grows large without bound. This occurs at finite R_+ and thus at *finite entropy* $S/L^2 = \pi R_+^2$, and the *temperature vanishes* along this boundary curve $P\beta$. The entropy is however not constant along this singular extremal boundary curve (in practice, the last curve we plot has $R_+ = 0.1$ but it should extend a bit further down in the region close to $\alpha\beta$). We typically find that lines of constant R_+ extend all the way to the red $\alpha\beta$ dashed line (4.4.1), but the latter is only reached in the limit $\epsilon \to \infty$. This makes it harder to extend our solution to regions even closer to $\alpha\beta$ (a fixed step in ϵ corresponds to an increasingly smaller progression in $\mathcal{Q}$ as $\alpha\beta$ is approached). Hairy BHs do not exist for $\mathcal{Q} < \mathcal{Q}|_P$, in agreement with the linear analysis of the left panel of Fig. 4.2.1, and there is clearly no relation between the hairy BHs and the boson star of the theory when $e = 1$ and, more generically, for $e_{\text{NH}} \leq e < e_\gamma$.

Because point P does not coincide with the origin O, hairy BHs with $e_{\text{NH}} \leq e < e_\gamma$ were not found in the perturbative analysis of [43]. Indeed, this perturbative analysis only captures hairy BHs that have small mass and charge.

The right panel of Fig. 4.4.1 plots the same phase diagram as the left panel but this time with the entropy S/L^2 on the extra vertical axis. The latter is the appropriate thermodynamic potential to discuss the preferred thermal phases of the microcanonical ensemble: for a given quasilocal mass $\mathcal{M}/L$ and charge $\mathcal{Q}/L$ fixed, the dominant phase is the one with the largest entropy. The red surface represents (a subset[4]) of RN BHs, both stable and unstable with the boundary line of stability being again the yellow dotted curve, here very close to the extremal RN BH (dark red with $\Delta\mathcal{M} = 0$). In the $S = 0$ plane we find the perturbative boson star (black curve). The blue dots fill the 2-dimensional

[4]We just plot the portion of the RN surface with $\Delta\mathcal{M} < 0.085$ that covers the region where the boson star also exists.

surface that describe hairy BHs (which merge with RN along the yellow line). Again we see (not very clearly but it will be more clear for higher e) that hairy BHs coexist with RN black holes in the region between the onset and extremal RN curves. In this case, we find that hairy black holes always have a larger entropy than the corresponding RN BHs with same $\mathcal{M}/L$ and $\mathcal{Q}/L$. So they are the thermodynamically preferred phase in the microcanonical ensemble.

Hence, it follows from the second law of thermodynamics that hairy BHs with $(\mathcal{Q}, \mathcal{M})$ between the RN onset and extremality curves are natural candidates for the endpoint of the RN superradiant/near-horizon instability when we do a time evolution of the instability where we preserve the mass and charge of the system.

4.4.2 Phase diagram for $e_\gamma \leq e < e_c$

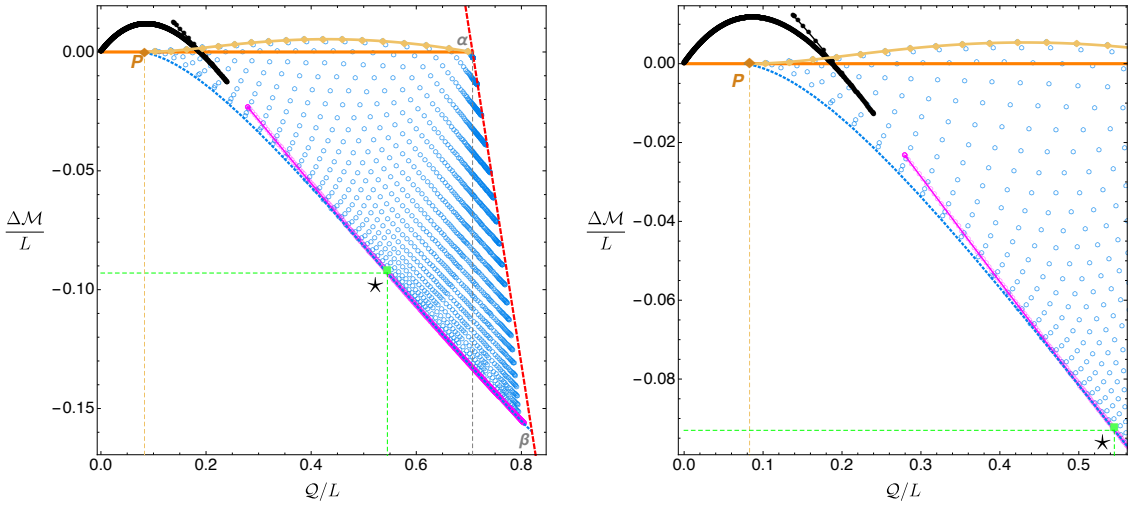

Figure 4.4.2: Phase diagram for Einstein theory with a scalar field charge $e = 1.85$ ($e_\gamma \leq e < e_c$) in a Minkowski box. The blue circles describe hairy black holes, the black disk (magenta circle) curve is the soliton main (secondary) family, and the orange line is the extremal RN BH (non-extremal RN BHs exist above it). The yellow curve is the superradiant onset curve of RN. As it could not be otherwise, it agrees with the hairy solutions in the limit where these have $\epsilon = 0$ and thus merge with RN family. The dashed vertical line is at $\mathcal{Q} = 2^{-1/2}$ which is the maximum local charge that an extremal RN BH can have while fitting inside a box with radius $R = 1$. The red dashed line (4.4.1) describes the boundary for black holes that can fit inside the cavity with radius $R = 1$. The green solid square labelled with a star ($\star$) has $(\mathcal{Q}_\star, \mathcal{M}_\star, \Delta\mathcal{M}_\star) \sim (0.545, 0.678, -0.093)$. The auxiliary blue dotted curve $P\star\beta$ in the bottom describes the line where hairy BHs terminate with unbounded horizon curvature. Hairy BHs that terminate in the trench $P\star$ of this auxiliary curve do so at finite entropy and vanishing temperature. On the other hand, hairy BHs that terminate in the trench segment $\star\beta$ (that coincides with magenta soliton line) do so at zero entropy and large (possibly unbounded) temperature.

In Fig. 4.4.2 we display the phase diagram when $e = 1.85$, which is representative of the range $e_\gamma \leq e < e_c$ that we sketched in the top-right panel of Fig. 4.2.2. As a first ob-

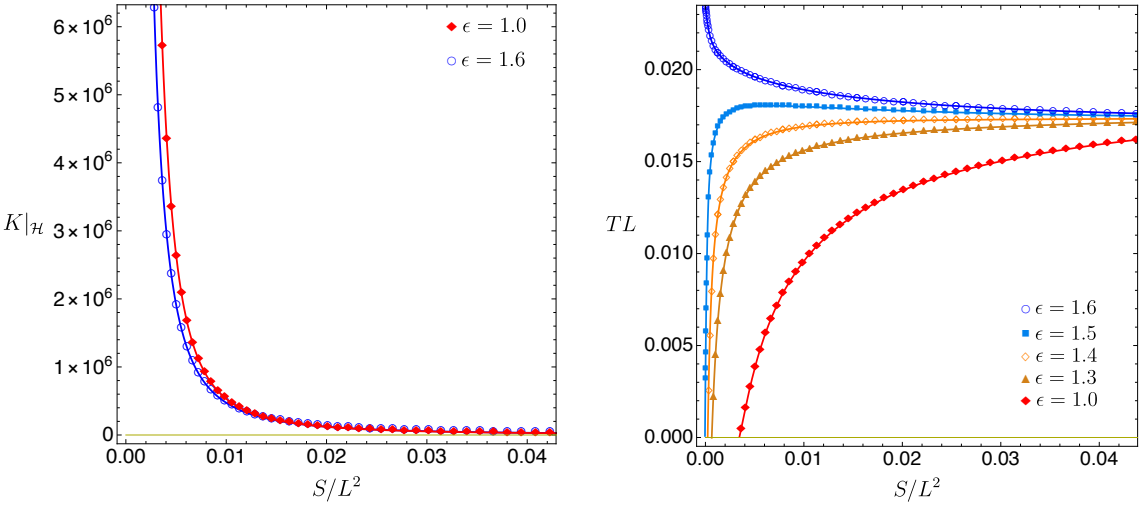

Figure 4.4.3: Kretschmann curvature at the horizon (left panel) and temperature (right panel) as a function of the entropy ($S/L^2 = \pi R_+^2$) for several hairy BH families with constant scalar amplitude ϵ and scalar field charge $e = 1.85$ ($e_\gamma \leq e < e_c$).

servation we note that, besides the main or perturbative boson star family (black curve) already present for $e < e_\gamma$, the diagram now also has the magenta line that starts at finite $\mathcal{Q}$, passes though point $\star$, and terminates at point β on the red dashed line. This is the secondary or non-perturbative family of boson stars. On its left side, this family has itself a series of cusps and zig-zagged secondary branches denoted as $B, C, \cdots$ in the sketch of the top-right panel of Fig. 4.2.2 (not displayed in Fig. 4.4.2). These details are not relevant here, and we ask the reader to see part III for an exhaustive study of boson stars' properties. It is however important to emphasize that this secondary/non-perturbative family exists (as a ground state family) only for $e_\gamma \leq e < e_c$, thus explaining the origin of the critical charges e_γ and e_c. At $e = e_c$ the magenta line of Fig. 4.4.2 merges with the black line (see section 4.4.3). On the other hand, as we decrease e below e_c one finds that the gap $\Delta\mathcal{Q}$ between the black and magenta families increases, and the "length" of the magenta line decreases because the left endpoint of this curve approaches β. It keeps doing so till it only exists on a very small neighbourhood of the red dashed line and, at $e = e_\gamma$, this line shrinks to the single point β. Below e_γ, the non-perturbative family ceases to exist (as seen in section 4.4.1). Essentially because it no longer fits inside the cavity. This discussion is better illustrated in the right panel of Fig. 4.2.1: 1) if we collect all non-perturbative solitons in a single plot, we find that they exist only in the window $e_\gamma \leq e < e_c$ and they fill the area bounded by the auxiliary dashed lines $a_c\beta_c\gamma$; 2) very close to e_c the non-perturbative soliton is almost on top of the auxiliary curve $a_c\beta_c$; and 3) on the opposite end, as $e \to e_\gamma$, the perturbative soliton line shrinks to the point γ on the red dashed line.

What are the consequences of these boson star discussions for the hairy BHs? Hairy BHs

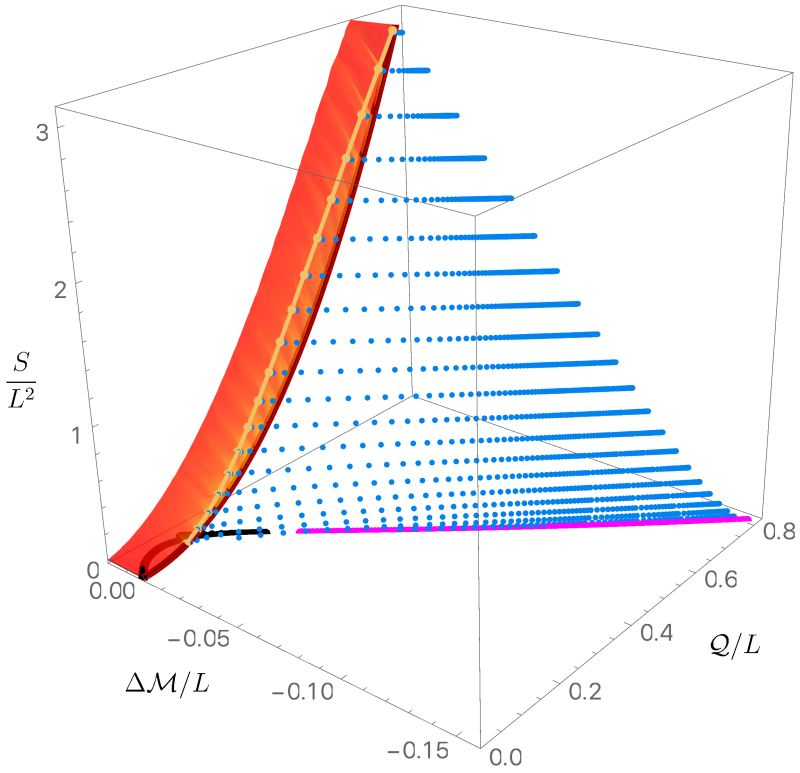

Figure 4.4.4: Entropy as a function of the quasilocal charge and mass difference for Einstein theory with a scalar field charge $e = 1.85$ ($e_\gamma \leq e < e_c$) in a Minkowski box. The red surface represents RN BHs in the range $0 \leq \Delta\mathcal{M} < 0.02$ (they extend for higher $\Delta\mathcal{M}$) with the dark red line with $\Delta\mathcal{M} = 0$ being the extremal RN BH family. The yellow line describes the merger line between RN BHs and hairy BHs, and RN BHs between this line and the dark red extremal line are unstable. The blue disks describe hairy BHs and the black (magenta) lines with $S = 0$ describe the main (secondary) soliton family. When they coexist with RN BHs, for a given $(\mathcal{Q}, \mathcal{M})/L$, hairy BHs always have more entropy than RN, *i.e.* they dominate the microcanonical ensemble.

with $e = 1.85$ are the blue circles in Fig. 4.4.2. As before, they exist in the area bounded by $P\alpha\beta$, where $P\alpha$ is the merger yellow line with RN BHs and coincides with the instability onset curve of [40], and $\alpha\beta$ is a segment of the red dashed line (4.4.1). Starting at the onset curve $P\alpha$ and moving down, *e.g.* along constant $\mathcal{Q}$ lines, we find that that hairy BHs terminate at the line $P\beta$ (or $P \star \beta$). This is the blue dashed line in Fig. 4.4.2 which describes hairy BHs with minimum entropy/horizon radius for a given charge. Along this line, the Kretschmann curvature scalar evaluated at the horizon $K|_{\mathcal{H}}$ grows very large (most probably, $K|_{\mathcal{H}} \to \infty$). To illustrate this, in the left panel of Fig. 4.4.3 we plot $K|_{\mathcal{H}}$ as a function of the entropy $S/L^2 = \pi R_+^2$ as we approach the line $P\beta$ (at small S) along curves of constant scalar amplitude ϵ (shown in the legends). Indeed, for small S/L^2 the curvature is growing very large.

So far the phase diagram of hairy BHs looks similar to the one for $e_{\rm NH} \leq e < e_\gamma$ (section 4.4.1). However, for $e_\gamma \leq e < e_c$ we now find that the way hairy BHs terminate

along the singular curve differs substantially depending on whether it ends to the left or to the right of the green square point $\star$ in Fig. 4.4.2 (with $\mathcal{Q}_\star/L \simeq 0.545$ for $e = 1.85$). When the hairy BHs terminate along $P\star$, they do so at *finite entropy* and *vanishing temperature*. On the other hand, hairy BHs that terminate along $\star\beta$ do so at *vanishing entropy* and *large (possibly unbounded) temperature*. To illustrate this, in the right panel of Fig. 4.4.3 we plot the temperature TL as a function of the entropy $S/L^2 = \pi R_+^2$ as we follow hairy BH families that approach the singular line $P\beta$ at (different; see legends) constant scalar amplitude ϵ. Point $\star$ has $(\mathcal{Q}_\star, \Delta\mathcal{M}_\star) \sim (0.545, -0.093)$ which corresponds to $(R_+, \epsilon)\big|_\star = (0, 1.55 \pm 0.05)$. Hairy BHs with $\epsilon < \epsilon_\star$ terminate at $P\star$, while hairy BHs with $\epsilon > \epsilon_\star$ end at $\star\beta$. The right panel of Fig. 4.4.3 indeed shows that hairy BHs with $\epsilon < \epsilon_\star$ approach $P\star$ at finite S/L^2 and with $TL \to 0$ (like all hairy BHs of section 4.4.1), while those with $\epsilon > \epsilon_\star$ approach $\star\beta$ with $S \to 0$ and $TL \to \infty$.

Another important conclusion that emerges from Fig. 4.4.2, is that hairy BHs which have a zero horizon radius limit terminate precisely along the segment $\star\beta$ of the secondary/non-perturbative soliton family. This means that hairy BHs terminate with the *same* $\mathcal{Q}$ and $\mathcal{M}$ as the non-perturbative soliton (but the gravitoelectric and scalar fields of the two solutions are different). On the other hand, those that end at $P\star$ do so in a manner that is very similar to the way the hairy BHs with $e_{\text{NH}} \leq e < e_\gamma$ terminate (section 4.4.1).

We find that the critical charge $\mathcal{Q}_\star(e)$ decreases as e grows from e_γ till e_c. As explained when discussing the right plot of Fig. 4.2.1, the non-perturbative soliton line shrinks to the point β when $e \to e_\gamma$. Thus, our expectation is that the critical charge $\mathcal{Q}_\star$ also reaches $\mathcal{Q}\big|_\beta$ when $e \to e_\gamma^+$. That is to say, we expect that hairy black holes are connected to the non-perturbative soliton as soon as it exists. However, determining numerically $\mathcal{Q}_\star$ in this limit is very difficult, since hairy BHs near β have very large values of ϵ.

The hairy BHs with $e_\gamma \leq e < e_c$ we find were not captured by the perturbative analysis of [43] because they do not extend to arbitrarily small mass and charge.

In Fig. 4.4.4, we plot the thermodynamic potential of the microcanonical ensemble $-$ the entropy $S/L^2 -$ as a function of $\mathcal{Q}$ and $\Delta\mathcal{M}$. In the $S = 0$ plane we find the perturbative boson star (black curve) and, for larger $\mathcal{Q}$ and after a gap, the non-perturbative boson star (magenta curve). As before, the red surface describes the RN BH family parametrized by R_+ and μ as in (4.3.6) and with $S/L^2 = \pi R_+^2$. It terminates at the dark red extremal curve with $\Delta\mathcal{M} = 0$. We only plot the portion of the RN surface with $\Delta\mathcal{M} < 0.02$ that covers the region where the perturbative boson star also exists. Unstable RN BHs are those between the instability onset (yellow dotted curve) and the extremal RN dark red curve. The blue dots fill the 2-dimensional surface that describes hairy BHs. It merges with RN BHs along the yellow dotted curve and then extends to lower $\Delta\mathcal{M}$ with an entropy that is always larger the the RN BH with the same $\mathcal{Q}$ and $\mathcal{M}$ (when they coexist). Therefore,

hairy BHs are the thermodynamically dominant phase in the microcanonical ensemble. Consequently, from the second law of thermodynamics, hairy BHs with $(\mathcal{Q}, \mathcal{M})$ between the RN onset and extremality curves are candidates for the endpoint of the RN superradiant/near-horizon instability when in a time evolution of the RN instability at constant mass and charge.

4.4.3 Phase diagram for $e_c \leq e < e_s$

In Fig. 4.4.5 we give the phase diagram for $e = 2$. This is the case we choose to illustrate the solution spectra in the range $e_\gamma \leq e < e_c$ that we sketched in the bottom-left panel of Fig. 4.2.2. Comparing with the diagram of Fig. 4.4.2 we immediately notice that the magenta line representing the non-perturbative soliton family is no longer present in Fig. 4.4.5. This is because at $e = e_c$, the perturbative and non-perturbative boson star families (*i.e.* the black and magenta lines of Fig. 4.4.2) merge and for $e \geq e_c$ the main or perturbative boson star family no longer has a Chandrasekhar mass limit and now extends from the origin O all the way to β in the red dashed line. This merger at e_c occurs in an interesting elaborated manner. In particular, going back to top-right sketch

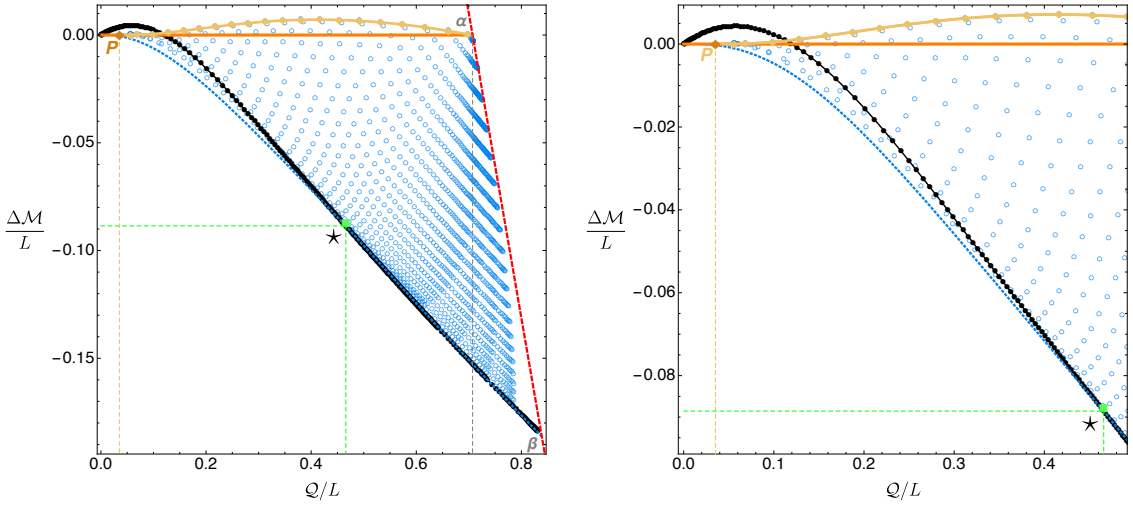

Figure 4.4.5: Phase diagram for Einstein theory with a scalar field charge $e = 2$ ($e_c \leq e < e_s$) in a Minkowski box. As before, the blue circles describe hairy black holes, the black disk curve is the soliton main family, and the orange line is the extremal RN BH (RN black holes exist above it). The gray and red dashed curves have the same interpretation as in Fig. 4.4.2. The green solid square labelled with a star ($\star$) has $(\mathcal{Q}_\star, \mathcal{M}_\star, \Delta\mathcal{M}_\star) \sim (0.466, 0.659, -0.0886)$. The auxiliary blue dotted curve $P\star\beta$ in the bottom describes the line where hairy BHs terminate with unbounded horizon curvature. Hairy BHs that terminate in the trench $P\star$ of this auxiliary curve have zero temperature ($T = 0$) and finite entropy $S/L = \pi R_+^2$. On the other hand, hairy BHs that terminate in the trench segment $\star\beta$ (that coincides with the black soliton line) have zero entropy and large (possibly unbounded) temperature. Note that these $\star\beta$ terminal hairy BHs have the same $(\mathcal{Q}, \Delta\mathcal{M})$ as the main soliton family with $\mathcal{Q} > \mathcal{Q}_\star$.

of Fig. 4.2.2, at $e = e_c$ the secondary zig-zagged branches $\cdots CBA$ of the perturbative (black) soliton also merge with the secondary zig-zagged branches $abc\cdots$ of the non-perturbative (magenta) soliton. As a consequence, for $e \geq e_c$ there is also a secondary soliton $\cdots CBbc\cdots$ (purple line in bottom-left of Fig. 4.2.2) that has higher energy than the perturbative (black) soliton $P\beta$. This secondary family is not displayed in Fig. 4.4.5 because it plays no role on the discussion of hairy BHs of the theory. The reader can find a detailed discussion of soliton's properties across the transition at $e = e_c$ in [2].

Since the colour code and associated labelling in Fig. 4.4.5 is the same as in Figs. 4.4.1 and 4.4.2 we can now immediately discuss the hairy BHs. Again they exist in the area enclosed by $P\alpha\beta$ filled with the blue circles. They merge with the RN family along the yellow dotted line $P\alpha$ when the scalar condensate vanishes, which agrees with the RN instability curve found in [40]. The hairy BHs then exist all the way down to the blue dashed line $P\beta$ (or $P\star\beta$) which, for a given charge, identifies the hairy BH that has minimum entropy/horizon radius. The Kretschmann curvature evaluated at the horizon $K|_{\mathcal{H}}$ diverges. For a given charge, $P\beta$ identifies the hairy BHs with minimum entropy/horizon radius and $K|_{\mathcal{H}}$ grows very large along it. This is confirmed in the left panel of Fig. 4.4.6: as we approach $P\beta$ (at small S) along lines of of constant scalar amplitude ϵ (identified in the legends), $K|_{\mathcal{H}}$ is growing very large.

Point $\star$ with charge $\mathcal{Q}_\star \simeq 0.466$ describes a transition point. Hairy BHs that end to the left of this point along $P\star$ do so at finite S with $T \to 0$. However, one has $S \to 0$ and $T \to \infty$ when the hairy BHs terminate along $\star\beta$ with $\mathcal{Q} > \mathcal{Q}_\star$. This is confirmed in the right panel of Fig. 4.4.6 where we plot the temperature TL as a function of the entropy $S/L^2 =$

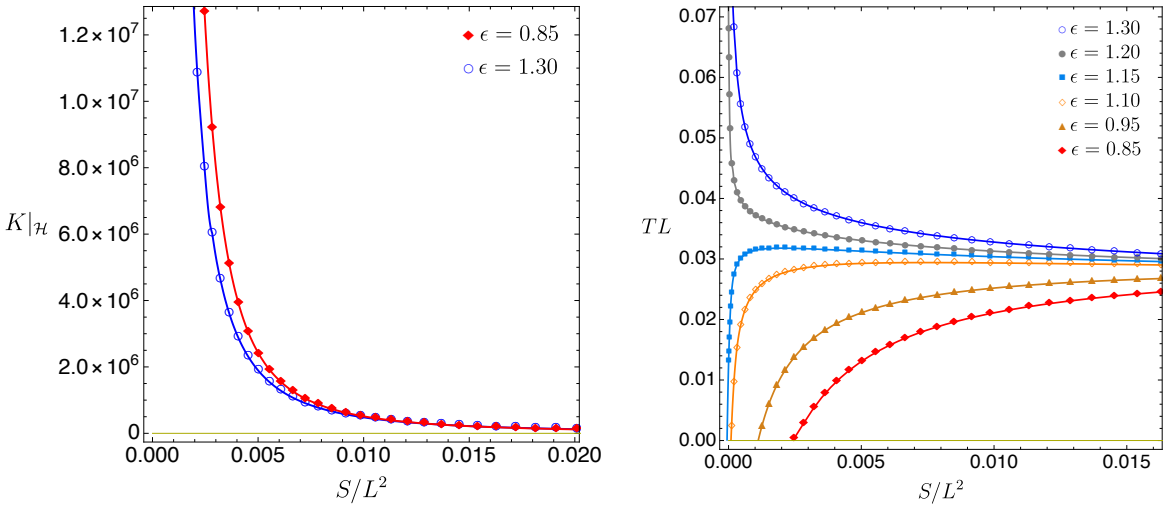

Figure 4.4.6: Kretschmann curvature at the horizon (left panel) and temperature (right panel) as a function of the entropy ($S/L^2 = \pi R_+^2$) for hairy BH families with fixed ϵ and scalar field charge $e = 2$ ($e_c \leq e < e_s$).

πR_+^2 as we follow different families of constant scalar amplitude hairy BHs that approach the singular line $P\beta$. Point $\star$ has $(\mathcal{Q}_\star, \Delta\mathcal{M}_\star) \sim (0.466, -0.0886)$ which corresponds to $(R_+, \epsilon)\big|_\star = (0, 1.175 \pm 0.005)$. Hairy BHs with $\epsilon < \epsilon_\star$ have $\mathcal{Q} < \mathcal{Q}_\star$ and terminate at $P\star$, while hairy BHs with $\epsilon > \epsilon_\star$ have $\mathcal{Q} > \mathcal{Q}_\star$ and end at $\star\beta$.

From Fig. 4.4.5 and the right panel of Fig. 4.4.6, it should not go without notice that the hairy BHs that have a zero horizon radius limit terminate along the trench $\star\beta$ of the perturbative soliton family. That is, *when* the hairy BHs have zero entropy, they have the same charge $\mathcal{Q}$ and mass $\mathcal{M}$ as the perturbative soliton. In a nutshell, hairy BHs with $e_c \leq e < e_s$ have a behaviour that is qualitatively similar to those of $e_\gamma \leq e < e_c$ (section 4.4.2). However, the zero entropy BHs now terminate on top of the perturbative soliton in the $\mathcal{Q}$-$\mathcal{M}$ phase diagram instead of ending on the non-perturbative soliton (which is now an excited solution $\cdots CBbc \cdots$ in the bottom-left panel of Fig. 4.2.2). We also find that the critical charge $\mathcal{Q}_\star(e)$ decreases and approaches $\mathcal{Q}_P$ as e grows from e_c to e_s. Moreover, we find that $\mathcal{Q}_\star \to \mathcal{Q}_P \to 0$ as $e \to e_s$.

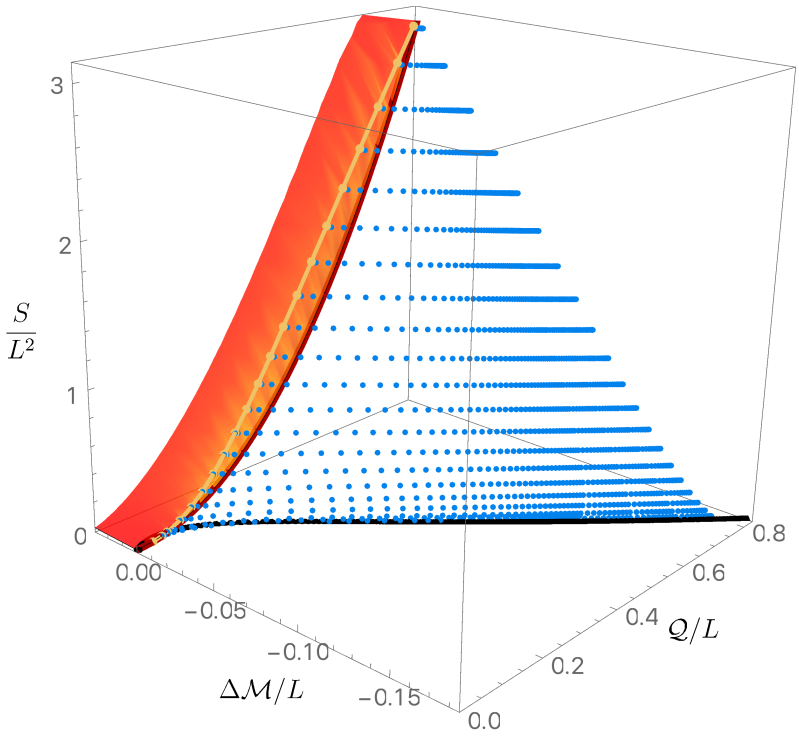

Figure 4.4.7: Entropy as a function of the quasilocal charge and mass difference for Einstein theory with a scalar field charge $e = 2$ ($e_c \leq e < e_s$) in a Minkowski box. When they coexist with RN BHs, for a given $(\mathcal{Q}, \mathcal{M})/L$, hairy BHs always have more entropy than RN, *i.e.* they dominate the microcanonical ensemble. For $e_c \leq e < e_s$, when $\mathcal{Q} > \mathcal{Q}_\star(e)$, hairy BHs have a zero entropy limit where they coincide with the soliton (black disk) curve in the sense that they have the same $(\mathcal{Q}, \mathcal{M})/L$ as the soliton (the temperature and horizon curvature diverges). However, when $0 < \mathcal{Q} < \mathcal{Q}_\star(e)$, hairy BHs terminate at an extremal BH (*i.e.* with zero temperature) and finite entropy (and divergent horizon curvature) along a line that does not coincide with the black disk one for the soliton.

Fig. 4.4.7, displays the phase diagram of the microcanonical ensemble for $e = 2$: the entropy S/L^2 as a function of $\mathcal{Q}$ and $\Delta\mathcal{M}$. The colour code of this diagram is the same as Fig. 4.4.4. Because e is bigger than the cases considered before, we see that the region between the onset yellow curve and the extremal RN dark red curve where RN BHs are unstable is now quite visible. Again we find that the hairy BHs (blue circles) that bifurcate from the yellow onset curve always have higher entropy that the RN BHs with the same $(\mathcal{Q}/L, \mathcal{M}/L)$ when they coexist. It follows that also for $e_c \leq e < e_{\rm s}$, hairy BHs are the preferred thermodynamic phase in the microcanonical ensemble. As expected from Fig. 4.4.5, for $\mathcal{Q} \geq \mathcal{Q}_\star \simeq 0.466$, the hairy BHs terminate with zero entropy on top of the perturbative boson star (black curve).

It is natural to expect that the hairy BHs we find should be the endpoint of the RN instability if we perturb an RN BH in the unstable region (where they always coexist with hairy BHs) and do a time evolution at constant charge and mass. The system would evolve to a final configuration that is stable against the original perturbation while respecting the second law of thermodynamics. Finally note that the hairy BHs with $e_c \leq e < e_{\rm s}$ described in this section were not studied in the perturbative analysis of [43] because the latter can only capture solutions that have a zero mass and charge limit.

4.4.4 Phase diagram for $e \geq e_{\rm s}$

The critical charge $e = e_{\rm s} = \frac{\pi}{\sqrt{2}} \sim 2.221$ is special for two main (related) reasons. First, it is the minimal charge above which scalar fields can drive *arbitrarily small* extremal RN BHs unstable via superradiance, as observed in the instability onset charge plot of the left panel of Fig. 4.2.1. Indeed, the extremal onset curve $e_{\rm onset}(R_+)$ reaches $e = e_{\rm s}$ as $R_+ \to 0$. The value of $e_{\rm s}$ can be predicted analytically as done in Section III.A of [43]. For $e > e_{\rm s}$, we can also have near-extremal BHs unstable for arbitrarily small R_+ or, equivalently, for arbitrarily small mass and charge.

This scalar charge $e_{\rm s}$ is also special because at $e = e_{\rm s}$ the slope of the perturbative soliton at the origin vanishes, *i.e.* $\frac{\delta\Delta\mathcal{M}}{\delta\mathcal{Q}}\big|_{\mathcal{Q}=0} = 0$. For $e < e_{\rm s}$, this slope is positive and we always have (some) perturbative solitons with higher quasilocal mass than the extremal RN (for sufficiently small $\mathcal{Q}$). On the other hand, for $e > e_{\rm s}$ the slope is always negative, $\frac{\delta\Delta\mathcal{M}}{\delta\mathcal{Q}}\big|_{\mathcal{Q}=0} < 0$, and thus perturbative solitons never coexist with RN BHs.

Ultimately as a consequence of these two properties, two important changes occur in the phase diagram of Fig. 4.4.5 as we follow its evolution across $e_{\rm s}$ and land on Fig. 4.4.8. First, the minimal charge for instability – that we have been denoting as $\mathcal{Q}_P$ – approaches zero as $e \to e_{\rm s}^-$ and $\mathcal{Q}_P = 0$ for $e \geq e_{\rm s}$. This is illustrated in Fig. 4.4.8 for the case $e = 2.3$. Second, we find that the hairy BHs (blue circles inside $O\alpha\beta$) now *always* terminate on top of the perturbative boson star (black line $O\beta$) as we move down, *e.g.* at constant $\mathcal{Q}$,

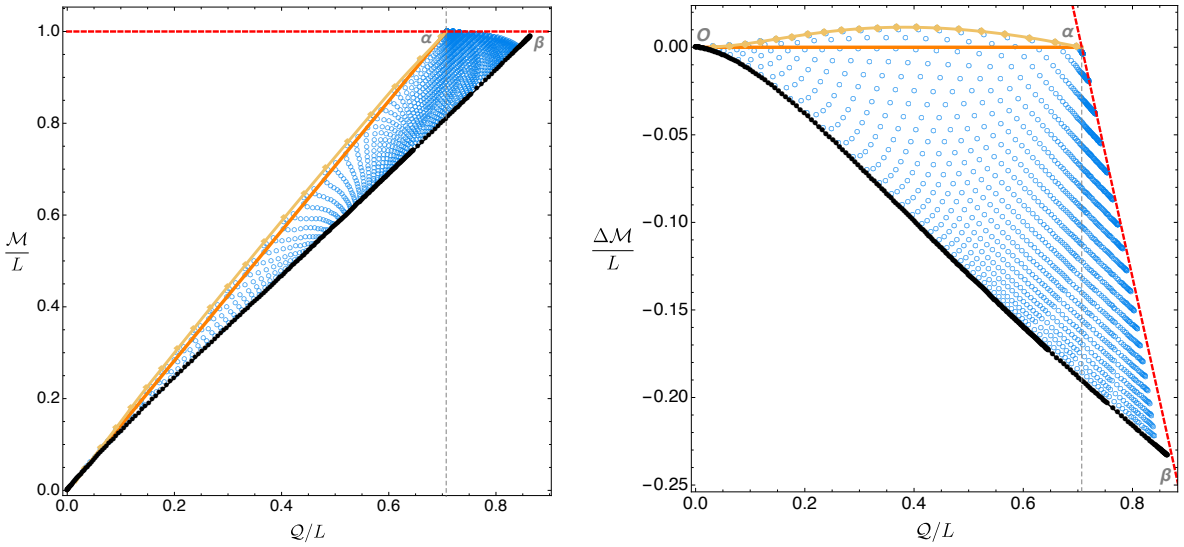

Figure 4.4.8: Phase diagram for Einstein theory with a scalar field charge $e = 2.3$ ($e > e_{\rm S}$) in a Minkowski box. In the left panel we have the $\mathcal{Q}$-$\mathcal{M}$ phase diagram: we see that the solutions pile up and this is why we have instead been plotting the phase diagram $\mathcal{Q}$-$\Delta\mathcal{M}$ (right panel). The blue circles describe hairy black holes, the black disk curve is the perturbative soliton family and the orange line is the extremal RN BH (RN black holes exist above it). The yellow curve is the superradiant onset curve of RN. As it could not be otherwise, it agrees with the hairy solutions in the limit where these have $\epsilon = 0$ and thus merge with RN family. The gray and red dashed curves have the same interpretation as in Fig. 4.4.2. For $e > e_{\rm S}$, the zero entropy limit of the hairy BH is the soliton (black disk curve) in the sense that they have the same $(\mathcal{Q}, \mathcal{M})/L$ as the soliton.

from the onset curve $O\alpha$. That is to say, one also has $\mathcal{Q}_\star = 0$ for $e \geq e_{\rm S}$. As hairy BHs approach this perturbative soliton curve, the Kretschmann curvature at the horizon, the entropy and temperature have the same behaviour as the one observed in Fig. 4.4.2 for BHs terminating along $\star\beta$: $K|_{\mathcal{H}} \to \infty$, $S \to 0$ and $T \to \infty$.

Since for $e \geq e_{\rm S}$ the hairy BHs exist all the way down to $(\mathcal{Q}, \mathcal{M}) \to (0, 0)$ one might expect that their properties can be captured by a perturbative analysis (to higher orders) around Minkowski space with gauge field in a box. This is indeed the case and such analysis was performed in [43]. This is a double expansion perturbation theory with the expansion parameters being the horizon radius R_+ and the scalar amplitude ϵ. Of course, here one assumes that $R_+ \ll 1$ and $\epsilon \ll 1$ which translates into $\mathcal{Q} \ll 1$ and $\mathcal{M} \ll 1$. The analysis of [43] culminates with explicit expansions for the thermodynamic quantities of the hairy BHs, which are listed in (5.27) of [43]. In particular, the expansion for the quasilocal mass

and charge are:

$$\begin{aligned}\mathcal{M}/L =& \left[\frac{R_+}{4}\left(\frac{\pi^2}{e^2}+2\right)+\frac{R_+^2}{32e^4}\left(\pi^4\Big(8[\mathrm{Ci}(2\pi)-\gamma-\ln(2\pi)]+5\Big)+4\left(e^2+\pi^2\right)e^2\right)+\mathcal{O}(R_+^3)\right]\\ &+\epsilon^2\left[\frac{1}{2}+\frac{R_+}{12\pi e^2}\left(9\pi^3\Big[\gamma-\mathrm{Ci}(2\pi)-2+\ln(2\pi)\Big]+\left(8\pi^2-3e^2\right)\Big[2\mathrm{Si}(2\pi)-\mathrm{Si}(4\pi)\Big]\right)\right.\\ &\left.+\mathcal{O}(R_+^2)\right]+\epsilon^4\left[\frac{15\pi^2-6e^2+16\pi\big[\mathrm{Si}(4\pi)-2\mathrm{Si}(2\pi)\big]}{24\pi^2}+\mathcal{O}(R_+)\right]+\mathcal{O}(\epsilon^6),\end{aligned} \tag{4.4.2}$$

$$\begin{aligned}\mathcal{Q}/L =& \left[\frac{\pi R_+}{2e}+\frac{R_+^2}{8e^3}\left(\pi^3\Big(2[\mathrm{Ci}(2\pi)-\gamma-\ln(2\pi)]+1\Big)+2\pi e^2\right)+\mathcal{O}(R_+^3)\right]+\epsilon^2\left[\frac{e}{2\pi}\right.\\ &\left.+\frac{R_+}{12\pi^2 e}\left(12\pi^3\Big(\gamma-\mathrm{Ci}(2\pi)+\ln(2\pi)-\frac{7}{4}\Big)+\left(8\pi^2-3e^2\right)\left[2\mathrm{Si}(2\pi)-\mathrm{Si}(4\pi)\right]\right)+\mathcal{O}(R_+^2)\right]\\ &-\left[\epsilon^4\frac{e\left(\left(8\pi^2-e^2\right)\left(2\mathrm{Si}(2\pi)-\mathrm{Si}(4\pi)\right)+4\pi e^2-8\pi^3\right)}{8\pi^4}+\mathcal{O}(R_+)\right]+\mathcal{O}(\epsilon^6),\end{aligned} \tag{4.4.3}$$

where $\mathrm{Ci}(x)=-\int_x^\infty \frac{\cos z}{z}\mathrm{d}z$ and $\mathrm{Si}(x)=\int_0^x \frac{\sin z}{z}\mathrm{d}z$ are the cosine and sine integral functions, respectively, and $\gamma\sim 0.577216$ is Euler's constant. This perturbation scheme assumes that R_+ and ϵ do not have a hierarchy of scales. When $R_+=0$, (4.4.2)-(4.4.3) reduces to the soliton thermodynamics and, when $\epsilon=0$, (4.4.2)-(4.4.3) yields the expansion of the caged RN BH thermodynamics. In [43] it was argued that (4.4.2)-(4.4.3) should

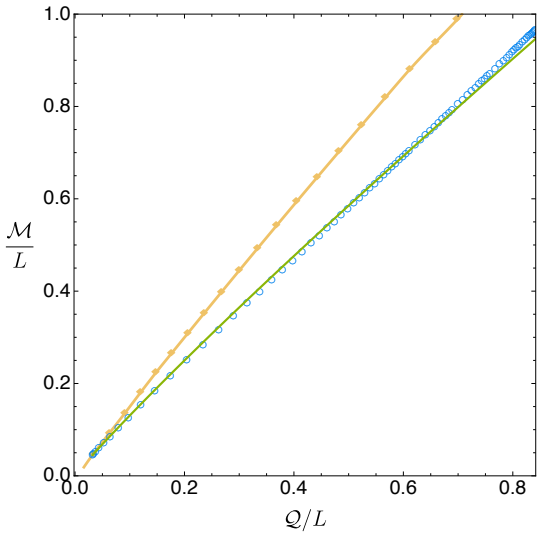

Figure 4.4.9: Comparing the exact numerical results (blue circles) with the perturbative analytical predictions (4.4.2)-(4.4.3) (green curve) for a family of black holes with constant $R_+=0.05$ and $e=2.3$. As expected, the perturbative analysis matches the exact results only for small R_+ and small ϵ (*i.e.* close to the origin and in the neighbourhood of the merger, yellow diamond, line which has $\epsilon=0$). That is to say, for the $R_+=0.05$ family shown, good agreement occurs for small $\mathcal{Q}$, say $\mathcal{Q}\lesssim 0.2$.

provide a good approximation (as monitored by the first law) for $\epsilon \lesssim 0.1, R_+ \lesssim 0.1$. Now that we have the exact (numerical) results for the hairy BHs in all their domain of existence we can use (4.4.2)-(4.4.3) to further check our numerics while, simultaneously, testing the regime of validity of (4.4.2)-(4.4.3) . As an example of this exercise, in Fig. 4.4.9 we compare the perturbative prediction (4.4.2)-(4.4.3) – the green curve– to our exact numerical results (blue circles) for the 1-parameter family of hairy BHs with $R_+ = 0.05$ parametrized by ϵ (with $\epsilon = 0$ at the merger with RN; the yellow curve). As expected, we observe good agreement for $\mathcal{Q} \lesssim 0.2$, say. Of course, the fact that the perturbative analysis does not differ much from the exact results for higher values of $\mathcal{Q}$ is to be seen as accidental; the perturbative is certainly not valid for such high charges.

As in the previous cases, we end our discussion of the $e \geq e_s$ case with the plot of Fig. 4.4.10 of the entropy as a function of the charge and mass. The colour coding is the same as in previous cases so it suffices to emphasize that again the hairy BHs (blue circles) are the preferred phase in the microcanonical ensemble. Indeed, in the region between the onset yellow curve and the extremal RN dark red curve with $\Delta\mathcal{M} = 0$ where they coexist with (unstable) RN BHs, hairy BHs always have higher dimensionless entropy for a given charge $\mathcal{Q}/L$ and mass $\mathcal{M}/L$. It further follows from the second law, that the unstable RN BHs should evolve in time towards the hairy BH we find with the same $\mathcal{Q}/L$ and $\mathcal{M}/L$.

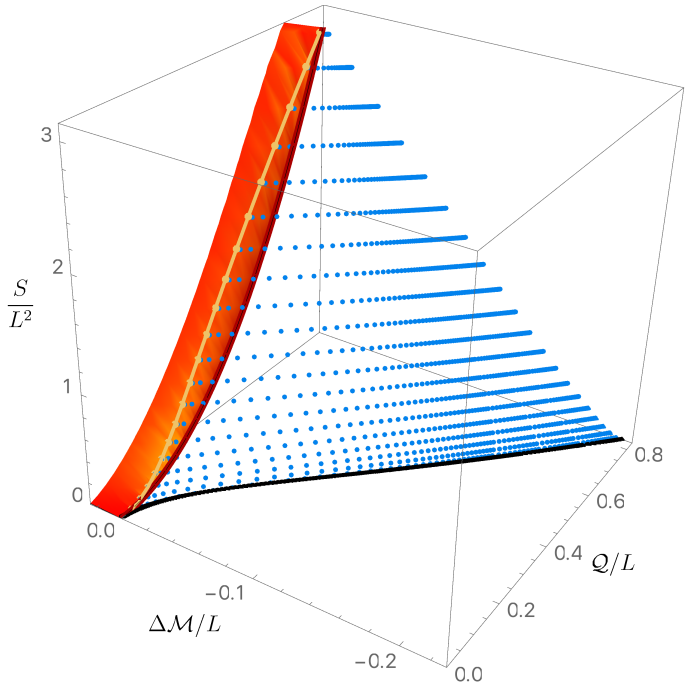

Figure 4.4.10: Entropy as a function of the quasilocal charge and mass difference for Einstein theory with a scalar field charge $e = 2.3$ ($e > e_s$) in a Minkowski box. When they coexist with RN BHs, for a given $(\mathcal{Q}, \mathcal{M})/L$, hairy BHs always have more entropy than RN, *i.e.* they dominate the microcanonical ensemble. For $e > e_s$, the zero entropy limit of the hairy BH is the soliton (black disk curve) in the sense that they have the same $(\mathcal{Q}, \mathcal{M})/L$ as the soliton (the temperature and horizon curvature diverges).

4.5 Conclusions and discussion

Recapping what we did so far, we integrated the equations of motion in the domain $R \in [R_+, 1]$ subject to regular boundary conditions at horizon and vanishing scalar field at the box. This is all we need to get the quasilocal phase diagrams of the previous section. But the description of the solution is only complete once we give the full solution all the way up to the asymptotically flat boundary.

Studies of scalar fields confined in a Minkowski cavity are already available in the literature: 1) at the linear level [118, 119, 41, 120, 121, 122, 123, 124, 125, 126, 127], 2) within a higher order perturbative analysis of the elliptic problem [40, 43], 3) as a nonlinear elliptic problem (without having asymptotically flat boundary conditions [128, 129, 130] or without matching with the exterior solution [131]), and 4) as an initial-value problem [132, 133, 134]. However, with the exception of the perturbative analysis of [43], the properties of the "internal structure" of the cavity required to confine the scalar field and its contribution to the ADM mass and charge that ultimately describe, by Birkhoff's theorem, the exterior RN solution are not discussed.

However, having the interior solution, we can compute the Lanczos-Darmois-Israel surface stress tensor (3.3.16) that describes the energy-momentum of the box Σ. We impose the three Israel junction conditions (3.3.15a)-(3.3.15c) on the gravitoelectric fields on the surface layer Σ. The fields f, A_t, ϕ are then continuous across Σ and the component of the electric field orthogonal to Σ is also C^0 across Σ. The latter means that we can confine the charged scalar condensate without needing to have a surface electric charge density on Σ. The three conditions (3.3.15a)-(3.3.15c) permit us to match the interior and exterior solutions, *i.e.* they fix the parameters M_0, c_A, ρ in (3.3.6) as a function of the reparametrization freedom parameter N introduced in (3.3.14):

$$M_0 = \frac{1}{N^2}\left(1 - \frac{A_t'(1)^2}{2}\right) - 1, \qquad c_A = \frac{A_t'(1) + A_1(1)}{N}\,, \qquad \rho = -\frac{A_t'(1)}{N}\,. \tag{4.5.1}$$

Effectively, these conditions fix the exterior RN solution as a function of the interior solution and of the box's energy-momentum. Not surprisingly, we have a 1-parameter freedom (N) to choose the box content that is able to confine the scalar condensate. Several cavities can do the job.

Ideally, we would fix N requiring that the gravitational field is not only C^0 but also differentiable across the box. That is, the fourth junction condition (3.3.15d) would also be obeyed and thus the extrinsic curvature

$$K^t{}_t = -\frac{f'(R)}{2f(R)\sqrt{g(R)}}\,, \qquad K^i{}_j = \frac{1}{R\sqrt{g}}\,\delta^i{}_j\,, \quad (i,j) = (\theta,\varphi)\,, \tag{4.5.2}$$

would also be continuous across the box. But, except when $\phi(R) = 0$, no choice of N allows us to simultaneously set $[K^t_t] = 0$ and $[K^i_i] = 0$. All we can do is to fix N requiring that $[K^t_t] = 0$ (at the expense of having $[K^i_i] \neq 0$) or vice-versa, or any other combination.

A choice of N fixes the energy density and pressure of the box since its stress tensor can be written in the perfect fluid form, $\mathcal{S}_{(a)(b)} = \mathcal{E} u_{(a)} u_{(b)} + \mathcal{P}(h_{(a)(b)} + u_{(a)} u_{(b)})$, with $u = f^{-1/2}\partial_t$ and local energy density $\mathcal{E}$ and pressure $\mathcal{P}$ given by

$$\mathcal{E} = -S^t_{\ t}\,, \qquad \mathcal{P} = S^x_{\ x} = S^\phi_{\ \phi}\,. \tag{4.5.3}$$

We are further constrained to make a choice such that relevant energy conditions are obeyed. Ultimately, failing these would mean that we cannot build the necessary box with the available materials. Different versions of these energy conditions read ($i = \theta, \varphi$) [5]:

$$\text{Weak energy condition:} \qquad \mathcal{E} \geq 0 \quad \wedge \quad \mathcal{E} + \mathcal{P}_i \geq 0\,; \tag{4.5.4}$$

$$\text{Strong energy condition:} \qquad \mathcal{E} + \mathcal{P}_i \geq 0 \quad \wedge \quad \mathcal{E} + \sum_{i=1}^{2} \mathcal{P}_i \geq 0\,; \tag{4.5.5}$$

$$\text{Null energy condition:} \qquad \mathcal{E} + \mathcal{P}_i \geq 0\,; \tag{4.5.6}$$

$$\text{Dominant energy condition:} \qquad \mathcal{E} + |\mathcal{P}_i| \geq 0\,. \tag{4.5.7}$$

We have experimented with different choices of N and found that are many selections that indeed satisfy (4.5.4) (and equally many others that don't). An example of this exercise is given in part III for the boson star case. Given that there seems to be no preferred choice, we do not do a further aleatory illustration here. Instead, we approach the problem from an experimental perspective. That is to say, in practice, we are given a cavity (that obeys the energy conditions or else it could not have been built with available materials). In principle, we can identify its stress tensor and hence compute N. We then insert this into the Israel matching conditions (4.5.1) to find the exact RN exterior solution and, in particular, the asymptotic ADM charges. We end up with an asymptotically flat static black hole solution (or boson star) that is regular everywhere except across the box (where the extrinsic curvature has a discontinuity) and that describes confined scalar radiation floating above the horizon and in thermodynamic equilibrium with it. That is to say, we have established that the configuration originally envisioned (in the rotating case) by Zel'dovich [135], Press-Teukolsky [35] and [146, 147, 148, 149, 150, 151] using linear considerations indeed exists as a non-linear equilibrium solution of the Einstein-Maxwell-scalar equations. And we further established that this is the thermal phase that dominates the microcanonical ensemble. In an ongoing programme, we are extending the current analysis to the rotating BH bomb system.

The hairy BHs we find are stable to the RN instabilities (since they merge with RN precisely at the onset of the original instability; see also [128, 129]) and have higher entropy than the RN BHs. It follows from this and the second law of thermodynamics that the charged black hole bomb does not need to break apart: in a time evolution at fixed energy and charge, the unstable RN BH should simply evolve towards the hairy BH we find. It would be interesting to confirm this doing time evolutions along the lines of those performed in [132, 133, 134] in the precise setup we described.

Not less interestingly, Minkowski space in a box (no horizon) with a scalar perturbation is itself non-linearly unstable to the formation of a BH for arbitrarily small amplitude [153], very much alike the pure global AdS spacetime [154, 155, 156, 157, 158, 159, 160, 140]. The weakly turbulent phenomenon is responsible for this instability [154, 155, 156, 157, 158, 161, 162, 163]. It would be interesting to study this non-linear instability when the scalar field is charged. Unlike the neutral case, for certain windows of charge and energy, there are now two possible families of BHs and not just the RN one. Therefore a time evolution of the non-linear instability along the lines of [156, 159, 160, 153, 164, 165, 166, 167, 140] should lead in some cases to gravitational collapse into an RN BH and in others into a hairy BH (there should also be a wide class of initial data for which no BH should form at all). Accordingly, the evolution details should differ in these different cases.

Part V

Appendices

APPENDIX A

Further properties of spinors and scalars in AdS

A.1 Dirac and Weyl Spinors

Depending on the dimensionality and the signature of spacetime one can have different types of spinors. We work in four spacetime dimensions in the bulk and three spacetime dimensions at the boundary. Let's first see the types of fermions that one has on the four dimensional space-time. The most general spinor is the Dirac spinor which in four dimensions has four complex components. Introducing the following basis for the γ - matrices:

$$\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} \tag{A.1.1}$$

one sees that the generators of the Lorentz group become block diagonal. Here indices i represent the spatial coordinates. This means that the Dirac spinor representation of the Lorentz group is reducible into two pieces and the Dirac spinor can be decomposed as

$$\Psi = \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix} \tag{A.1.2}$$

now the Lorentz group acts on each two component spinors $\psi_\pm$ independently. If one works in even dimension, one can always make a Weyl decomposition of the spinor. If moreover the spinor is massless it can be shown that the two Dirac equations are decoupled and they are reduced to the Weyl equation.

The Dirac spinor has four degrees of freedom and the Weyl spinor has two degrees of freedom.

A.1.1 Spin connection

For a metric of the form $g_{\mu\nu} = diag(-g_{tt}(r), g_{rr}(r), g_{\theta\theta}(r), g_{\phi\phi}(r,\theta))$, one can define the following tetrad basis

$$e^0 = \sqrt{g_{tt}}, \quad e^1 = \sqrt{g_{rr}}, \quad e^2 = \sqrt{g_{\theta\theta}}, \quad e^3 = \sqrt{g_{\phi\phi}} \tag{A.1.3}$$

then using $de^a = -\omega^a_b \wedge e^b = -(\omega^a_b)_c e^c \wedge e^b$, we can read off the spin connection components, where the Latin indices are referring to the orthonormal basis and the Greek ones to the spacetime metric. The non trivial spin connections are

$$\omega_{01} = -\frac{g_{tt,r}}{2\sqrt{g_{tt}}}e^0, \quad \omega_{21} = \frac{g_{\theta\theta,r}}{2g_{\theta\theta}\sqrt{g_{rr}}}e^2, \quad \omega_{31} = \frac{g_{\phi\phi,r}}{2g_{\phi\phi}\sqrt{g_{rr}}}e^3, \quad \omega_{32} = \frac{g_{\phi\phi,\theta}}{2g_{\phi\phi}\sqrt{g_{\theta\theta}}}e^3 \tag{A.1.4}$$

and using $(\omega^a_b)_c e^c_\mu = \omega^a_{b\mu}$ we have

$$\omega_{01t} = -\frac{g_{tt,r}}{2}, \quad \omega_{21\theta} = \frac{g_{\theta\theta,r}}{2\sqrt{g_{\theta\theta}}\sqrt{g_{rr}}}, \quad \omega_{31\phi} = \frac{g_{\phi\phi,r}}{2\sqrt{g_{\phi\phi}}\sqrt{g_{rr}}}, \quad \omega_{32\phi} = \frac{g_{\phi\phi,\theta}}{2\sqrt{g_{\phi\phi}}\sqrt{g_{\theta\theta}}} \tag{A.1.5}$$

as well as the ones that are related by antisymmetry $\omega_{ab\mu} = -\omega_{ba\mu}$.

A.1.2 Useful formulae

The components of $\Gamma^M D_M$ are given by

$$\Gamma^t D_t = \frac{1}{\sqrt{g_{tt}}}\Gamma^{\underline{t}}\partial_t + \Gamma^{\underline{r}}\frac{1}{4}\frac{g_{tt,r}}{\sqrt{g_{tt}}} - i\frac{1}{\sqrt{g_{tt}}}q\Gamma^{\underline{t}}A_t \tag{A.1.6}$$

$$\Gamma^r D_r = \frac{1}{\sqrt{g_{rr}}}\Gamma^{\underline{r}}\partial_r \tag{A.1.7}$$

$$\Gamma^\theta D_\theta = \Gamma^{\underline{\theta}}\frac{1}{\sqrt{g_{\theta\theta}}}\partial_\theta + \Gamma^{\underline{r}}\frac{1}{4}\frac{1}{\sqrt{g_{rr}}}\frac{g_{\theta\theta,r}}{g_{\theta\theta}} \tag{A.1.8}$$

$$\Gamma^\phi D_\phi = \Gamma^{\underline{\phi}}\frac{1}{\sqrt{g_{\phi\phi}}}\partial_\phi + \Gamma^{\underline{r}}\frac{1}{4}\frac{1}{\sqrt{g_{rr}}}\frac{g_{\phi\phi,r}}{g_{\phi\phi}} + \Gamma^{\underline{\theta}}\frac{1}{4}\frac{1}{\sqrt{g_{\theta\theta}}}\frac{g_{\phi\phi,\theta}}{g_{\phi\phi}} \tag{A.1.9}$$

The following relations are useful for the calculations involving the boundary conditions which were discussed in detail in the main text, section 2.2.3:

$$\begin{aligned} \psi_+ &= \frac{1}{2}(1+\Gamma^r)\Psi, \quad \Gamma^r\psi_+ = \psi_+ \\ \psi_- &= \frac{1}{2}(1-\Gamma^r)\Psi, \quad \Gamma^r\psi_- = -\psi_- \end{aligned} \tag{A.1.10}$$

$$\begin{aligned} \bar{\psi}_+ &= \bar{\Psi}\frac{1}{2}(1-\Gamma^r), \quad \bar{\psi}_+\Gamma^r = -\bar{\psi}_+ \\ \bar{\psi}_- &= \bar{\Psi}\frac{1}{2}(1+\Gamma^r), \quad \bar{\psi}_-\Gamma^r = \bar{\psi}_- \end{aligned} \tag{A.1.11}$$

the above follow from the anticommutation relations for the gamma matrices and the fact that $(\Gamma^r)^2 = 1$ and $\Gamma^r = \Gamma^{r\,\dagger}$.

$$\begin{aligned} \bar{\psi}_+\psi_+ &= \bar{\Psi}\frac{1}{2}(1-\Gamma^r)\frac{1}{2}(1+\Gamma^r)\Psi = \bar{\Psi}\frac{1}{4}(1-\Gamma^{r2})\Psi = 0 \\ \bar{\psi}_-\psi_- &= \bar{\Psi}\frac{1}{2}(1+\Gamma^r)\frac{1}{2}(1-\Gamma^r)\Psi = \bar{\Psi}\frac{1}{4}(1-\Gamma^{r2})\Psi = 0 \\ \bar{\psi}_+\Gamma^r\psi_+ &= -\bar{\Psi}\frac{1}{2}(1-\Gamma^r)\frac{1}{2}(1+\Gamma^r)\Psi = -\bar{\Psi}\frac{1}{4}(1-\Gamma^{r2})\Psi = 0 \\ \bar{\psi}_-\Gamma^r\psi_- &= \bar{\Psi}\frac{1}{2}(1+\Gamma^r)\frac{1}{2}(1-\Gamma^r)\Psi = \bar{\Psi}\frac{1}{4}(1-\Gamma^{r2})\Psi = 0 \end{aligned} \tag{A.1.12}$$

$$\bar{\psi}_+\Gamma^i D_i\psi_- = \bar{\Psi}\Gamma^i(1-\Gamma^{r2})D_i\Psi = 0\,. \tag{A.1.13}$$

A.2 Near-extremal scalar field modes in AdS

For a Reissner-Nordström de Sitter ($\Lambda > 0$) background, in [70, 71, 72] it was found that there is a family of quasinormal modes – denoted as the 'near-extremal' family of modes – that is distinct from the 'de Sitter' family quasinormal mode, where the latter connects to the normal modes of de Sitter when the horizon radius shrinks to zero size, $r_+ \to 0$.

This naturally raises the question of whether there is also such a 'near-extremal' family of modes in AdS and, if so, wether they do or not coincide with the AdS family of modes. In this appendix, we address this question in the simplest case, namely in the case of a (charged) scalar field that obeys the Klein-Gordon equation (2.3.2). More concretely, we arrive to the near-extremal frequency (2.3.10) which is used in the main text (see section 2.3 and the discussion there-in of the dashed curves of Fig. 2.3.2) to show that in AdS the 'near-extremal' and AdS families of modes coincide (unlike in the de Sitter case).

The 'near-extremal' modes we seek obey (2.3.2) in the background (2.2.1) (we will work with the gauge choice $C = 0$) and, at least in the near extremal limit, are expect to be highly peaked near the horizon. So we want to simultaneously zoom into the horizon and approach extremality. For that we first introduce the dimensionless quantities

$$x = 1 - \frac{r}{r_+}, \qquad \sigma = 1 - \frac{r_-}{r_+}\,. \tag{A.2.1}$$

For $x \ll 1$ one is close to the outer horizon and for $\sigma \ll 1$ the inner and outer horizon are very close, *i.e.* one is close to extremality. Next, we take the limit $\sigma \to 0$ whilst keeping $z = \frac{x}{\sigma}$ fixed. From [70, 71, 72] the 'near-extremal' modes are expected to saturate the superradiant bound $\omega = q\mu$ at extremality so onwards we measure the frequency difference $\delta\omega$ with respect to this bound via the redefinition

$$\omega = \frac{e\,Q}{r_+} + \sigma\,\delta\omega\,. \tag{A.2.2}$$

Using the condition $f(r_-) = 0$ for the location of the inner horizon one can find $r_- = r_-(r_+, Q, L)$ which is then inserted into (A.2.1) to express Q as a function of (r_+, σ, L).

In these near-extremality conditions, we are ready to find the near-horizon solution of the Klein-Gordon equation. Concretely, introducing the above redefinitions into the Klein-Gordon equation (2.3.2), to *leading order in* σ, we obtain:

$$(1-z)\,z\,\partial_z^2\phi(z) + (1-2z)\,\partial_z\phi(z) + \left[\frac{\varphi^2 - \hat{\lambda}\,z}{z(1-z)} + \eta\right]\phi(z) = 0\,, \tag{A.2.3}$$

where

$$\begin{aligned}
\varphi &= \frac{R_+\delta\tilde{\omega}}{1+6R_+^2}\,,\\
\hat{\lambda} &= \frac{2e\,R_+^2\left[\delta\tilde{\omega}\sqrt{2}\sqrt{1+3R_+^2} - e(1+3\,R_+^2)\right]}{\left(1+6R_+^2\right)^2}\,,\\
\eta &= \frac{1}{6R_+^2+1}\left(\tilde{m}^2R_+^2 + \ell(\ell+1) - \frac{2e^2\left(1+3R_+^2\right)R_+^2}{1+6R_+^2}\right)\,,
\end{aligned} \tag{A.2.4}$$

where we have introduced the dimensionless quantities $R_+ = r_+/L, e = qL, \tilde{m} = mL, \delta\tilde{\omega} = L\,\delta\omega$.

With the field redefinition

$$\phi(z) = z^{-i\varphi}\,(1-z)^{i\sqrt{\varphi^2-\hat{\lambda}}}\,\hat{f}(z)\,, \tag{A.2.5}$$

(A.2.3) is rewriten as a standard hypergeometric ODE

$$(1-z)\,z\,\hat{f}''(z)+\left[1-2i\varphi-2iz\left(-i-\varphi+\sqrt{\varphi^2-\hat{\lambda}}\right)\right]\hat{f}'(z)$$
$$+\left[\eta-\hat{\lambda}-(i+2\varphi)\left(-\varphi+\sqrt{\varphi^2-\hat{\lambda}}\right)\right]\hat{f}(z)=0\,. \qquad \text{(A.2.6)}$$

The regular (*i.e.* ingoing) solution at the future event horizon is given by

$$\phi(z)=z^{-i\varphi}\,(1-z)^{i\sqrt{\varphi^2-\hat{\lambda}}}\,{}_1F_2\,(a_-,a_+,1-2i\varphi,z)\,, \qquad \text{(A.2.7)}$$

where ${}_1F_2(a,b,c;z)$ is the standard hypergeometric function and a_-,a_+ are defined by:

$$a_\pm=\frac{1}{2}\left(1\pm\sqrt{1+4\eta}-2i\varphi+2i\sqrt{\varphi^2-\hat{\lambda}}\right). \qquad \text{(A.2.8)}$$

In the context of a matched asymptotic expansion, the near-region (near-horizon) solution (A.2.7) must now be matched with the far-region solution of (2.3.2) (in near-extremality conditions). As explained above we expect the 'near-extremal' modes we are looking into to have wavefunctions that die-off very quickly away from the black hole horizon (at least near-extremality). Therefore, as a first rude approximation we take the far-region to be described by a vanishing wavefunction. That is to say, in the overlapping region, we match the near-region solution (A.2.7) with $\phi=0$. In the end of the day, this approximation turns out to be quite good because the analytical approximation for the 'near-extremal' frequency that we obtain −see (A.2.11) − matches remarkably well the numerical solution of (2.3.2). This is best seen comparing the black dashed analytical curve of our expansion in Fig. 2.3.2 with the numerical blue dot results. For this reason, we do not try to improve further our matching asymptotic approximation.

Proceeding in these conditions, the leading order behaviour of the large $R=r/L$ series expansion ($z\to-\infty$) of ϕ, namely $\phi\approx(-z)^{\pm\sqrt{1+4\eta}}$, needs to be matched with the far-region solution $\phi=0$. Before we can do it, we still need to distinguish the cases $1+4\eta\geq0$ and $1+4\eta<0$. For our proposes (comparing with the numerical results of section 2.3), we want to consider the small scalar field charge case for which one finds that $1+4\eta\geq0$ holds as long as e $\leq e_c$ where

$$e_c^2=\frac{1+6R_+^2}{8R_+^2(1+3R_+^2)}\left[\left(6+4m^2\right)R_+^2+(1+2\ell)^2\right] \qquad \text{(A.2.9)}$$

(the reader also interested in the case $1+4\eta<0$ can follow the steps detailed in [72]). In

these conditions, from the matching condition one finds that

$$\delta\tilde{\omega} = \frac{\sqrt{2}}{2} e\sqrt{1+3R_+^2} - \frac{i}{4R_+}\Bigg((1+6R_+^2)(1+2p) + \sqrt{(1+6R_+^2)\left[1+6R_+^2+4m^2R_+^2+4\ell(\ell+1)\right]-8e^2R_+^2(1+3R_+^2)}\Bigg), \quad \text{(A.2.10)}$$

where $p = 0, 1, 2, \cdots$ is the radial overtone of the mode. Replacing this into (A.2.2) one finally finds that 'near-extremal' modes have a frequency given by:

$$\omega L \simeq e\mu + \sigma\,\delta\tilde{\omega} + \mathcal{O}(\sigma^2). \quad \text{(A.2.11)}$$

This is (2.3.10) in the main text when we set the radial overtone $p = 0$.

A.3 Perturbative results for spinors in AdS

Although we solve the Dirac equation numerically in the main text, it is very good practice to testify the numerical results against analytical predictions that can be obtained within perturbation theory in some region of the parameter space. Therefore, in this appendix we find some useful analytical perturbative approximations for the Dirac frequencies. More concretely, this Appendix is divided in two parts. In Appendix A.3.1 we use a matching asymptotic expansion approach to find the frequency approximations (A.4.1)-(A.4.2) that, in Fig. A.4.2 of Appendix A.4, are compared against our numerical results. Then, in Appendix A.3.2 we use a systematic perturbative expansion in the dimensionless horizon radius $r_+/L \ll 1$ (with no further approximations) to find the analytical frequency approximations (A.4.3)-(A.4.4) which, in Fig. A.4.3 of Appendix A.4, are also compared with our numerical results. In both cases, there is agreement between the analytical approximation predictions and the numerical results in the regime of parameter space where the former are valid.

In this appendix, as in the main text, we solve the Dirac equation (2.2.17) in a AdS-RN background (2.2.1) (gauge choice $C = 0$) with regular (ingoing) boundary conditions at the future event horizon and the standard (2.2.36) or alternative (2.2.37) boundary conditions at the conformal boundary. We will work exclusively with vanishing fermion mass, $m = 0$.

A.3.1 Matched asymptotic expansion

In this section we derive an analytical expression for the imaginary part of the Dirac frequency using the method of matched asymptotic expansion introduced in [168, 169] (see also e.g. [170, 109, 171]). We assume $r_+ \ll L$ and split our spacetime into two regions; an

asymptotic globally AdS *far region* where the effects of the black hole can be neglected and a *near region* about the black hole outer horizon where the effects of the cosmological constant can be neglected. In each region the associated perturbation equation can be solved analytically, then matching the near and far region solutions in their *overlapping region* will fix the integration constants as well as the imaginary part of the perturbation frequency ω. More concretely, the near region is defined by $r - r_+ \ll 1/\omega$ and the far region is defined by $r - r_+ \gg r_+$. It follows that the overlapping region exists for $\omega r_+ \ll 1$. A further assumption we must make is that the Coulomb interaction is weak, $Qq \ll 1$, where q (Q) is the fermion (black hole) charge.

A.3.1.1 Far region solution

Since in the far region the effect of the black hole (BH) is assumed to be negligible, we effectively have a fermion field in the global AdS background. Thus, the general solution for the massless fermionic field R_1 is given by (2.4.18) that we reproduce here:

$$\begin{aligned} F(r) &= C_1\, {}_2F_1\Big(\frac{1}{2}+\ell, \omega L+\ell+1, 2(1+\ell), \frac{2r}{r+iL}\Big) \\ &\quad + C_2 \Big(\frac{2r}{r+iL}\Big)^{-1-2\ell} {}_2F_1\Big(-\frac{1}{2}-\ell, \omega L-\ell, -2\ell, \frac{2r}{r+iL}\Big), \end{aligned} \tag{A.3.1}$$

where $C_{1,2}$ are two arbitrary amplitudes to be determined below. Asymptotically this solution decays as (2.2.23), namely $R_1\big|_{r\to\infty} \sim \alpha_1 + \beta_1 \frac{L}{r} + \mathcal{O}(r^{-2})$ with

$$\begin{aligned} \alpha_1 &= i^{\omega L}\, 2^{\ell+\frac{1}{2}} \left(C_1\, {}_2F_1\left(\ell+\frac{1}{2}, \ell+\omega L+1; 2\ell+2; 2\right) + C_2\, {}_2F_1\left(-\ell-\frac{1}{2}, -\ell+\omega L; -2\ell; 2\right)\right), \\ \beta_1 &= C_2 i^{1+L\omega} 2^{-\frac{3}{2}-\ell} \Bigg[(1+2\ell-2L\omega)\ {}_2F_1\left(-\ell-\frac{1}{2}, -\ell+\omega L; -2\ell; 2\right) + \\ &\quad \frac{1}{\ell}(1+2\ell)(\ell-L\omega)\ {}_2F_1\left(-\ell+\frac{1}{2}, -\ell+\omega L+1; -2\ell+1; 2\right)\Bigg] - \\ &\quad C_1\, 2^{\ell-\frac{1}{2}}\, i^{\omega L+1} \Bigg[(2\omega L+2\ell+1)\, {}_2F_1\left(\ell+\frac{1}{2}, \ell+\omega L+1; 2\ell+2; 2\right) \\ &\quad + \frac{(2\ell+1)(\omega L+\ell+1)}{\ell+1}\, {}_2F_1\left(\ell+\frac{3}{2}, \ell+\omega L+2; 2\ell+3; 2\right)\Bigg]. \end{aligned} \tag{A.3.2}$$

This solution has to satisfy the asymptotic boundary condition (for a massless fermion). For the standard quantization this is (2.2.36) while for the alternative quantization the boundary condition is (2.2.37) which fix β_1 as a function of α_1 or, equivalently, C_1 as a

function of C_2. This yields

$$\alpha_1(\lambda \pm \omega L) \mp i\beta_1 = 0 \quad \Leftrightarrow \quad C_1 = C_2(-1)^{-2\ell}2^{-1-2\ell}\gamma_\pm\,, \qquad \text{where}$$
$$\gamma_\pm = \frac{(1{+}2\ell{\pm}4L\omega)_2F_1\left(-\frac{1}{2}{-}\ell,{-}\ell{+}\omega L,{-}2\ell,2\right)\mp(1{+}2\ell)_2F_1\left(\frac{1}{2}{-}\ell,\omega L{-}\ell,{-}2\ell,2\right)}{(1{+}2\ell{\pm}4L\omega)_2F_1\left(\frac{1}{2}{+}\ell,1{+}\ell{+}\omega L,2(1{+}\ell),2\right)\pm(1{+}2\ell)_2F_1\left(\frac{3}{2}{+}\ell,\omega L{+}\ell{+}1,2(1{+}\ell),2\right)}\,, \tag{A.3.3}$$

where the upper sign refers to the *standard quantisation* (2.2.36) and the lower sign to the *alternate quantisation* (2.2.37).

Note that we do not impose any boundary condition at an inner boundary since this far region solution does not extend to there.

A.3.1.2 Near region solution

In the near region we can approximate $\Delta(r) = r^2 f(r)$ by:

$$\Delta(r) \approx (r - r_+)(r - r_-) \tag{A.3.4}$$

where $r_- \approx \frac{Q^2}{2r_+}$. This follows from the assumption that $r_+ \ll L$ and therefore in the near region we have $r \sim \mathcal{O}(r_+) \ll L$, so we can neglect the r^2/L^2 term in $f(r)$. Further applying the near region assumptions to the Dirac equation (2.2.17) we can neglect terms of order ωr_+ or higher powers. Other terms appear which are dominated by a $1/\Delta$ term in the small black hole (BH) approximation near the horizon; therefore we can evaluate the numerators of these terms at $r \approx r_+$.

With these approximations, and the coordinate transformation

$$z = \frac{r - r_+}{r - r_-}\,, \qquad 0 \le z \le 1, \tag{A.3.5}$$

(the horizon $r = r_+$ is at $z = 0$) the Dirac equation is approximately given by the near region equation,

$$\Big(\hat{\omega}\frac{1-z}{z} - \frac{\lambda^2}{1-z}\Big)R_1(z) + \frac{1}{2}(1-3z)R_1'(z) + (1-z)zR_1''(z) = 0, \tag{A.3.6}$$

where

$$\hat{\omega} = \frac{r_+^2\omega(r_+(i+2r_+\omega) - ir_-)}{2(r_+ - r_-)^2}. \tag{A.3.7}$$

Making the field redefinition

$$R_1(z) = z^{\tilde{\alpha}}(1-z)^{\tilde{\beta}}R(z) \tag{A.3.8}$$

where

$$\tilde{\alpha} = \frac{1}{4} + i\tilde{\sigma}, \quad \tilde{\beta} = \frac{1}{2} + \ell, \quad \tilde{\sigma} = \frac{1}{4}\sqrt{16\hat{\omega} - 1}\,, \tag{A.3.9}$$

the near horizon equation is rewritten in the standard hypergeometric form – see (2.4.8) – with parameters a, b and c given by

$$a = 1 + \ell, \quad b = 1 + \ell + 2i\tilde{\sigma}, \quad c = 1 + 2i\tilde{\sigma}\,. \tag{A.3.10}$$

The most general near horizon solution is

$$\begin{aligned} R_1^{\text{near}}(z) = \quad & \alpha z^{\frac{1}{4}-i\tilde{\sigma}}(1-z)^{\ell+\frac{1}{2}}\,{}_2F_1(1+\ell, 1+\ell-2i\tilde{\sigma}, 1-2i\tilde{\sigma}, z) \\ & +\beta z^{\frac{1}{4}+i\tilde{\sigma}}(1-z)^{\ell+\frac{1}{2}}\,{}_2F_1(1+\ell, 1+\ell+2i\tilde{\sigma}, 1+2i\tilde{\sigma}, z) \end{aligned} \tag{A.3.11}$$

Using the property of the hypergeometric function ${}_2F_1(a, b, c, 0) = 1$ we find that the $z \to 0$ behaviour of the near region solution is $R_1^{\text{near}}(z) \approx z^{\frac{1}{4}}(\alpha z^{-i\tilde{\sigma}} + \beta z^{i\tilde{\sigma}})$. Requiring regularity (only ingoing modes) at the horizon implies that we must set $\beta = 0$.

A.3.1.3 Matching

To find the large r ($z \to 1$) behaviour of the near region solution (A.3.11) with $\beta = 0$, we use the $z \to 1 - z$ transformation law of the hypergeometric function [48]:

$$\begin{aligned} {}_2F_1(a, b, c, z) = & \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)}\,{}_2F_1(a, b, a+b-c+1, 1-z) \\ & +(1-z)^{c-a-b}\frac{\Gamma(c)\Gamma(a+b-c)}{\Gamma(a)\Gamma(b)}\,{}_2F_1(c-a, c-b, c-a-b+1, 1-z). \end{aligned} \tag{A.3.12}$$

We again use that ${}_2F_1(a, b, c, 0) = 1$ as well as $1 - z \approx \frac{r_+ - r_-}{r}$ (when $r \to \infty$), to obtain:

$$R_1^{\text{near}}\Big|_{\text{large}\,r} \approx \Gamma(1-2i\tilde{\sigma})\Big[\frac{(r_+ - r_-)^{-\ell-\frac{1}{2}}\Gamma(1+2\ell)}{\Gamma(1+\ell)\Gamma(1+\ell-2i\tilde{\sigma})}r^{\ell+\frac{1}{2}} + \frac{(r_+ - r_-)^{\ell+\frac{1}{2}}\Gamma(-1-2\ell)}{\Gamma(-\ell)\Gamma(-\ell-2i\tilde{\sigma})}r^{-\ell-\frac{1}{2}}\Big]. \tag{A.3.13}$$

This needs to be matched (in the overlapping region) with the small r behaviour of the far region solution (A.3.1) subject to the asymptotic boundary conditions (A.3.3).

$$R_1^{\text{far}}\Big|_{\text{small}\,r} \approx C_1\, 2^{\frac{1}{2}+\ell}(iL)^{-\frac{1}{2}-\ell}r^{\frac{1}{2}+\ell} + C_2\, 2^{-\frac{1}{2}+\ell}(iL)^{\frac{1}{2}+\ell}r^{-\frac{1}{2}-\ell}. \tag{A.3.14}$$

In the overlapping region, one must have $R_1^{\text{near}}\big|_{\text{large}\,r} = R_1^{\text{far}}\big|_{\text{small}\,r}$. That is to say, we must match independently the $r^{\frac{1}{2}+\ell}$ and $r^{-\frac{1}{2}-\ell}$ terms of (A.3.13) with those of (A.3.14).

This matching yields:

$$\frac{C_2}{C_1}\Big(\frac{i}{2}\Big)^{2\ell+1} = \frac{\Gamma\left[1+\ell\right]\Gamma\left[-1-2\ell\right]}{\Gamma\left[1+2\ell\right]\Gamma\left[-\ell\right]}\frac{\Gamma\left[\frac{3}{2}+\ell-\frac{2ir_+^2}{r_+-r_-}\omega\right]}{\Gamma\left[\frac{1}{2}-\ell-\frac{2ir_+^2}{r_+-r_-}\omega\right]}\Big(\frac{r_+-r_-}{L}\Big)^{2\ell+1} \tag{A.3.15}$$

where we have used (A.3.9) and (A.3.7) for $\tilde{\sigma}$ to restore the explicit dependence on the frequency ω. The ratio $\frac{C_2}{C_1}$ follows straightforwardly from (A.3.3). We want to solve the transcendental equation (A.3.15) to get an analytical expression for the frequency. There is no closed form solution, unless we do some educated approximations that we now discuss. Since we are working with very small and weakly charged RN-AdS black hole, one expects that the mode frequencies in such a background are close to the massless Dirac normal modes frequencies of AdS_4, already computed in (2.4.20) (standard quantization) or (2.4.21) (alternative quantization). Denote this normal mode frequency by ω_{AdS_4}. However, since the background now has a horizon, the system becomes dissipative and with respect to the normal modes of AdS_4, the frequency of the system should acquire a small imaginary contribution. Denote it by $i\,\delta$. In (A.3.15), it is thus a good approximation to replace the frequency ω by $\omega = \omega_{AdS_4} + i\,\delta$ with $|\delta| \ll \omega_{AdS_4}$. Our target now is to solve (A.3.15) at *leading order* for $\delta \ll \omega_{AdS_4} \sim \mathcal{O}(1)$. *A posteriori*, we compare the prediction of our analytical computation with the numerical result to confirm that this approximation is valid.

Equation (A.3.15) has the additional challenges that: 1) the frequency appears in the argument of the Gamma functions and 2) $\Gamma\left[-1-2\ell\right]$ in the numerator diverges for the allowed values (2.2.19) of the harmonic number ℓ (recall that $\Gamma[-p] = \infty$ for non-negative p). To deal with these obstacles, we use the Gamma function property $\Gamma[z+1] = z\Gamma[z]$ and the assumptions of our problem, $\omega r_+ \ll 1$ and $\delta \ll \omega_{AdS_4} \sim \mathcal{O}(1)$. This allows to expand the Gamma functions whose argument depends on ω (and thus on δ) to extract δ out of the argument of the Gamma functions. In particular, this permits to find that the divergence of $\Gamma\left[-1-2\ell\right]$ in the numerator is cancelled by the Gamma function in the denominator that depends on δ.

In these conditions, one finds that the leading order solution for δ is

$$\tilde{K}\delta \approx \frac{i(-1)^{2\ell+3/2}2^{-4\ell-2}\ell!}{\sqrt{\pi}\left(\frac{2\ell-1}{2}\right)!\left(\ell+1/2\right)}\prod_{k=0}^{\ell-1/2}\frac{\ell+1/2-k}{\ell-k}\Big(\frac{r_+-r_-}{L}\Big)^{2\ell+1}, \tag{A.3.16}$$

where $\tilde{K}$ is a positive real number for each ℓ, n. For the alternative boundary condition

(2.2.37) it is given by (for overtone $p = 0$)

$$\tilde{K} = -\frac{e^{3i\pi\ell}\left({}_2F_1^{(0,1,0,0)}\left(\ell+\frac{1}{2},2\ell+2,2\ell+2,2\right)+{}_2F_1^{(0,1,0,0)}\left(\ell+\frac{3}{2},2\ell+2,2\ell+2,2\right)\right)}{{}_2F_1\left(1,-\ell-\frac{1}{2};-2\ell;2\right)-{}_2F_1\left(1,\frac{1}{2}-\ell;-2\ell;2\right)}, \tag{A.3.17}$$

whereas for the standard boundary condition (2.2.36) we have (for $p = 0$)

$$\tilde{K} = \frac{e^{3i\pi\ell}\left(-{}_2F_1^{(0,1,0,0)}\left(\ell+\frac{1}{2},2\ell+3,2\ell+2,2\right)+{}_2F_1^{(0,1,0,0)}\left(\ell+\frac{3}{2},2\ell+3,2\ell+2,2\right)\right)}{{}_2F_1\left(2,-\ell-\frac{1}{2};-2\ell;2\right)+{}_2F_1\left(2,\frac{1}{2}-\ell;-2\ell;2\right)}. \tag{A.3.18}$$

Thus the two boundary conditions (2.2.37), (2.2.36) yield different values of δ. As a concrete example (that we use to compare our numerics with), for $\ell = 1/2$, $p = 0$ we have for the alternative boundary condition (2.2.37),

$$\delta \approx -\frac{1}{4\pi}\left(\frac{r_+ - r_-}{L}\right)^2, \tag{A.3.19}$$

while for standard boundary condition (2.2.36),

$$\delta \approx -\frac{3}{4\pi}\left(\frac{r_+ - r_-}{L}\right)^2. \tag{A.3.20}$$

These are the analytical predictions we use in (A.4.1)-(A.4.2) of Appendix A.4 to compare against the numerical results: see Fig. A.4.2. We expect these analytical predictions to be valid only for small horizon radius and away from extremality and for $Qq \ll 1$, which we are able to confirm numerically in section 2.4.4. An added bonus for this method is that we have an expression for general ℓ. The method we present in the next Appendix below has to be done for each ℓ individually, but it is more systematic than this one, since it only requires an expansion in $r_+/L \ll 1$.

A.3.2 Perturbative expansion in R_+

In this section we find an analytical prediction for the frequency using a systematic perturbative expansion in r_+/L. Unlike in the previous subsection, the only approximation that will be made is that the expansion parameter of this expansion is small, $r_+/L \ll 1$. We will do this expansion up to the order that finds the first correction (in the real part of the frequency) to the global AdS normal mode frequency. Should we wish, we could go one order higher in the analysis and find also the correction to the imaginary part of the frequency (although this is computationally more demanding). For our purposes of comparing with the numerical results, it is enough to have the correction to the real part of the frequency (the results of appendix A.3.1 already allow us to test independently the imaginary part).

The systematic perturbative expansion in $r_+/L \ll 1$ used in this Appendix was first

introduced in [46] and further explored in [45, 138, 57, 39, 40] where the reader can find full details of the method (we will be very succinct in our exposition). In short, we split our spacetime into a near and far regions. We expand the frequency $\Omega = \omega L$ and the field R_1 in each region in a power series in $R_+ = r_+/L$:

$$\Omega = \sum_{k=0}^{\infty} \Omega_{(k)} R_+^k \,; \qquad R_1^{near} = \sum_{k=0}^{\infty} \psi_{(k)}^{near} R_+^k \,, \qquad R_1^{far} = \sum_{k=0}^{\infty} \psi_{(k)}^{far} R_+^k \,. \tag{A.3.21}$$

We now series expand the Dirac equation (2.2.17) in small R_+. The leading, zeroth order equation is simply the Dirac equation in global AdS_4 for a massless fermion (A.3.1) (that we already studied in sections 2.4.2.2 and A.3.1.1). Not surprisingly, the small $R = r/L$ expansion of this leading order solution breaks down at order R_+/R. This motivates splitting our spacetime into a far region, $R \gg R_+$, and a near region, $R_+ \leq R \ll 1$. In the far region we work with the radial coordinate R but in the near region we work instead with the radial coordinate $y = R/R_+$ (since the far region small R expansion breaks down at order R_+/R).

In the far region, at each order in R_+, we impose the *standard* boundary condition (2.2.36) or the *alternative* boundary condition (2.2.37). In the near region we impose boundary conditions that only allow for ingoing waves at the horizon. We then perform a matching procedure at each order in R_+, in the region where the far and near region overlap, to determine the frequency coefficients $\Omega_{(k)}$, as well as amplitudes that were not fixed by the two boundary conditions. At leading (zeroth) order, we fix $\Omega_{(0)}$ to be the normal mode frequency for a massless fermion already obtained in (2.4.21) or (2.4.20) for the boundary conditions (2.2.37) or (2.2.36), respectively. We will do this for the mode with harmonic number $\ell = 1/2$, and radial overtone to be $p = 0$. Our aim is then to find the first frequency correction $\Omega_{(1)}$ due to the presence of the black hole. For concreteness, in most of our discussion below we only explicitly present details of the case where we impose the alternative boundary condition (2.2.37). We then present the final result also for the standard boundary condition (2.2.36).

In the far region the leading order R_+^0 solution is (A.3.1) and imposing the asymptotic boundary conditions amounts to repeat *mutatis mutandis* the analysis done in (A.3.1)-(A.3.3). With our choice of $\ell = 1/2$ and $p = 0$ this fixes the frequency at order zero to be $\Omega_{(0)} = \frac{3}{2}$; see (2.4.21). To fix the normalization, we set the amplitude of the Dirac field at infinity to be 1 at all orders in R_+: $R_1^{far}|_{R\to\infty} = 1 + \mathcal{O}(1/R)$.

Introducing the near region radial coordinate $y = R/R_+$, still at leading order R_+^0, the near region Dirac equation (for $\ell = 1/2$) reads

$$\frac{1}{2}(y-1)\left(2y-\mu^2\right)\partial_y^2 \psi_{(0)}^{near} + \left(y - \frac{1}{2} - \frac{\mu^2}{4}\right)\partial_y \psi_{(0)}^{near} - \psi_{(0)}^{near} = 0. \tag{A.3.22}$$

The solution which is regular at the horizon ($y = 1$) is

$$\begin{aligned}\psi^{near}_{(0)} = \quad & \alpha_{(0)} \cosh\left(2\log\left(2\sqrt{y-1}+\sqrt{4y-2\mu^2}\right)\right) \\ & +i\beta_{(0)} \sinh\left(2\log\left(2\sqrt{y-1}+\sqrt{4y-2\mu^2}\right)\right),\end{aligned} \tag{A.3.23}$$

with

$$\beta_{(0)} = -i\alpha_{(0)} \frac{1+4\left(\mu^2-2\right)^2}{1-4\left(\mu^2-2\right)^2}. \tag{A.3.24}$$

We must now match the far and near regions solutions at order R_+^0 in their overlapping region $R_+ \ll R \ll 1$. This procedure, typically fixes all other constants of the problem that were not fixed by the boundary conditions. The large R expansion of $\psi^{near}_{(0)}$ is $\psi^{near}_{(0)}|_{\text{large } R} = \alpha_{(0)} + \beta_{(0)} R + \cdots$ whereas the small R expansion of $\psi^{far}_{(0)}$ is $\psi^{far}_{(0)}|_{\text{small } R} = (-1)^{3/4} R + \cdots$. Therefore, matching $\psi^{near}_{(0)}|_{\text{large } R} = \psi^{far}_{(0)}|_{\text{small } R}$ requires that we set $\alpha_{(0)} = 0$ and $\beta_{(0)} = (-1)^{3/4}$. Collecting the results at order 0 for the alternative quantization (2.2.37) we have:

$$\psi^{near}_{(0)} = 0\,, \qquad \psi^{far}_{(0)} = \frac{(-1)^{3/4} R\sqrt{1+\mathrm{i}R}}{(1+R^2)^{3/4}}\,; \qquad \Omega_{(0)} = \frac{3}{2}. \tag{A.3.25}$$

For the standard quantisation one has a similar result with $\Omega_{(0)} = 5/2$.

We can now proceed to the first order R_+^1 contribution that enables us to find $\Omega_{(1)}$. We repeat the procedure outlined above for the far and near regions, imposing the alternative boundary condition in the far region and ingoing boundary condition at the horizon in the near region. The near region equation for $\psi^{near}_{(1)}$ is the same as at order R_+^0 (i.e. there is no source). After imposing the horizon boundary condition, we find that the large R expansion of the near region solution gives

$$R^{near}\big|_{\text{large } R} \approx (-1)^{3/4} R + \beta_{(1)} R_+ + \cdots\,. \tag{A.3.26}$$

The far region equation at order R_+^1 can also be solved analytically for $\psi^{far}_{(1)}$. As usual, we find a solution with two integration constants, say C_1 and C_2. These constants will in general depend on μ, e and $\Omega_{(1)}$. Firstly by imposing the alternative quantisation we find an expression for C_2 in terms of μ, e and $\Omega_{(1)}$. Then we again impose that our field has amplitude 1 at infinity, yielding an expression for C_1 in terms of μ, e and $\Omega_{(1)}$. After doing this we find the small R behaviour of the far region solution to be

$$\begin{aligned}R^{far}\big|_{\text{small } R} \approx (-1)^{3/4} R + R_+ & \left(-14+2e\mu-7\mu^2-6\pi\Omega_{(1)}\right) \\ & - \frac{R_+}{R}\left(2-4e\mu+\mu^2+2\pi\Omega_{(1)}\right)+\cdots\,,\end{aligned} \tag{A.3.27}$$

In the overlapping region $R_+ \ll R \ll 1$, (A.3.27) must match (A.3.26). A straightforward matching of the terms $R_+ R^0$ fixes the constant $\beta_{(1)}$ in (A.3.26). On the other hand, there is no R_+/R term in the large R series expansion of the near region solution. It follows that $\Omega_{(1)}$ must be such that it eliminates the corresponding R_+/R term in the small R expansion of the far region solution (A.3.27). This fixes $\Omega_{(1)}$ to be

$$\Omega_{(1)} = -\frac{\mu^2 - 4\mu e + 2}{2\pi}. \tag{A.3.28}$$

Thus for the alternative boundary condition (2.2.37), the frequency up to $\mathcal{O}(R_+)$ is

$$\omega L = \frac{3}{2} - R_+ \frac{\mu^2 - 4\mu e + 2}{2\pi} + \mathcal{O}(R_+^2). \tag{A.3.29}$$

On the other hand, repeating the above analysis the standard boundary condition (2.2.36) yields the frequency:

$$\omega L = \frac{5}{2} + R_+ \frac{11\mu^2 - 20\mu e + 22}{6\pi} + \mathcal{O}(R_+^2). \tag{A.3.30}$$

The frequencies (A.3.29) and (A.3.30) are the analytical frequency approximations (A.4.3) and (A.4.4) that we reproduce in the main text and that we compare against the numerical data in Fig. A.4.3 of section A.4.

A.4 Comparing numerical results with analytical expansions for spinors in AdS

To test our numerical code we first compute the quasinormal mode frequencies for a massless Dirac field in global AdS and in Schwarzschild-AdS: see Fig. A.4.1 for the standard (2.2.36) and alternative (2.2.37) quantizations. When $r_+ = 0$ our frequencies reduce to the normal mode frequencies of global AdS (2.4.20) and (2.4.21) for the standard and alternative quantizations, respectively. Recall that ℓ is the spin-weighted harmonic number and n is the radial overtone (related to the number of nodes of the wavefunction along the radial direction). So the smallest $|\omega|$ is obtained for $\ell = 1/2$ and $n = 0$. For the alternative boundary condition these are $\omega L = 3/2$ and $\omega L = -5/2$, while for the standard quantization these are $\omega L = 5/2$ and $\omega L = -3/2$). All the numerical results we will present describe solutions with $\ell = 1/2$ and $n = 0$. If there is an instability it should already be present in this sector of perturbations (see the argument in section 2.4.1).

On the other hand, for finite r_+/L our numerical curves reproduce the values first computed in [73]. As explained in the end of section 2.2.3, this is because the AdS/CFT standard (2.2.36) and alternative (2.2.37) quantizations have vanishing energy flux and corre-

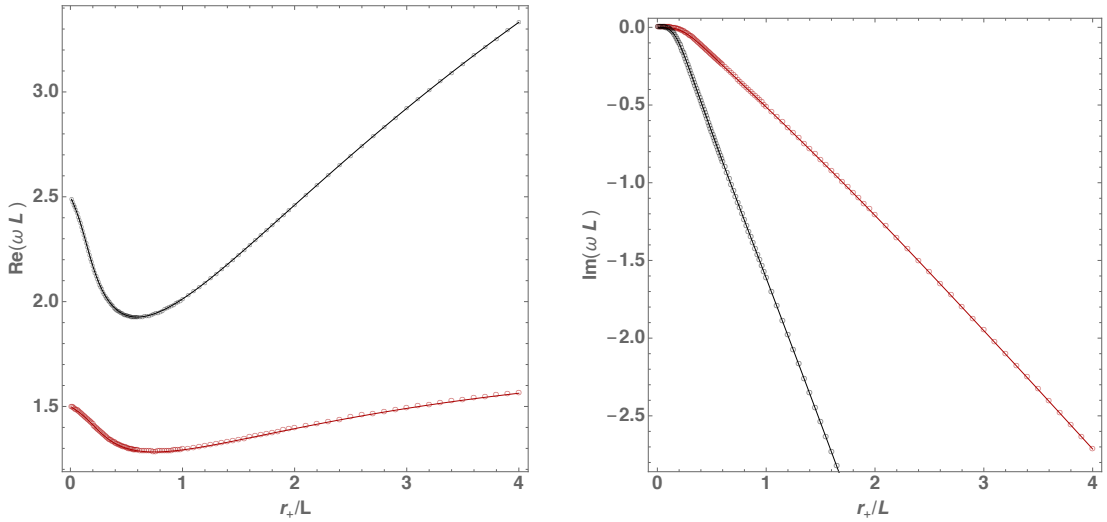

Figure A.4.1: Real (left) and imaginary (right) part of the QNM frequencies as a function of the horizon radius for the two families of boundary conditions in Schwarzschild-AdS background ($m = 0$, $\ell = 1/2$). The black curve that reduces to the normal mode of AdS $\omega L = 5/2$ has the standard quantization (2.2.36) while the red curve (that reduces to the normal mode of AdS $\omega L = 3/2$) has the alternative quantization (2.2.37). Not plotted, if we take the black curve data then $-\mathrm{Re}(\omega) + i\,\mathrm{Im}(\omega)$ is also a solution with alternative quantization (2.2.37) (which reduces to the AdS normal mode $\omega L = -5/2$ when the horizon shrinks). Similarly, if we take the red curve data then $-\mathrm{Re}(\omega) + i\,\mathrm{Im}(\omega)$ is also a solution with standard quantization (2.2.36).

spond precisely to the boundary conditions imposed in [73]. To complete the spectrum of Schwarzschild-AdS (and global AdS) note that frequencies that are the negative of the complex conjugate of the frequencies ($-\omega^*$) plotted in Fig. A.4.1 are also eigenvalues of the system (which was missed in [73]).

Next we test our numerical code for AdS-RN. As a first test, in Appendix A.3.1 we use a matching asymptotic expansion method to find that for $\omega r_+ \ll 1$ and $qQ \ll 1$ (and $m = 0, \ell = 1/2, n = 0$) the imaginary part of the frequency is approximately given by

$$\mathrm{Im}(\omega L) \approx -\frac{1}{4\pi}\left(\frac{r_+ - r_-}{L}\right)^2 + \mathcal{O}\left(\left(\frac{r_+ - r_-}{L}\right)^3\right), \qquad \text{Alternative quantization;} \quad \text{(A.4.1)}$$

$$\mathrm{Im}(\omega L) \approx -\frac{3}{4\pi}\left(\frac{r_+ - r_-}{L}\right)^2 + \mathcal{O}\left(\left(\frac{r_+ - r_-}{L}\right)^3\right), \qquad \text{Standard quantization.} \quad \text{(A.4.2)}$$

In Fig. A.4.2 we plot $\mathrm{Im}(\omega L)$ as a function of r_+/L, for $\mu = \frac{1}{2}\mu_{\mathrm{ext}}$ and a fermion charge $qL = 1$ (blue dots). We compare these numerical results with the analytical result (A.4.1). Both agree for small r_+/L, *i.e.* in the regime were the matching expansion (A.4.1) is valid.

As a second test, in Appendix A.3.2 we use a perturbative expansion in r_+/L about global

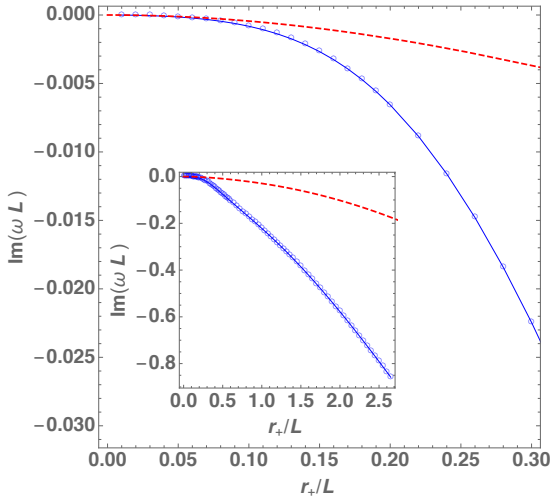

Figure A.4.2: Imaginary part of the frequency as a function of the horizon radius for fixed chemical potential $\mu = \frac{\mu_{\text{ext}}}{2}$ ($m = 0, \ell = 1/2$ and alternative quantization). The red dashed line is the analytical approximation (A.4.1). The inset plot zooms-out the main plot.

AdS to find that the real part of the frequency behaves as

$$\text{Re}(\omega L) = \frac{3}{2} - \frac{r_+}{L}\frac{\mu^2 - 4\mu e + 2}{2\pi} + \mathcal{O}\left(\frac{r_+^2}{L^2}\right), \qquad \text{Alternative quantization;} \tag{A.4.3}$$

$$\text{Re}(\omega L) = \frac{5}{2} + \frac{r_+}{L}\frac{11\mu^2 - 20\mu e + 22}{6\pi} + \mathcal{O}\left(\frac{r_+^2}{L^2}\right), \qquad \text{Standard quantization.} \tag{A.4.4}$$

The plots for the real part of the frequency are displayed in figure A.4.3. The analytical perturbative results are (A.4.3) and (A.4.4). An important feature of this result is that for a fixed chemical potential, the slope of the real part of the frequency (the term proportional to R_+) changes sign for a certain electric charge. This coincides with our numerical findings. In the top panels we show the real part of the frequency for the boundary condition (2.2.37) with $qL = 0.7,\ 0.8$ and $\mu = \mu_{\text{ext}}(1 - 10^{-3})$. In the bottom panels of figure A.4.3 we show the real part of the frequency for the boundary condition (2.2.36) as a function of the horizon radius for fixed chemical potential $\mu = \frac{\mu_{\text{ext}}}{2}$ and fermion charges $qL = 1.9,\ 2$ from left to right. The dashed red lines are the analytical results (A.4.4) and (A.4.3) respectively. Note that for these parameters the change of sign occurs for charges $qL \sim 1.95$ and $qL \sim 0.71$ respectively.

As a general comment, it is perhaps worthy to comment that we find that matching asymptotic and perturbative analysis like the ones provided in Appendices A.3.1 A.3.2 are valid for smaller windows of r_+/L in the Dirac field case when compared with similar analysis done for the scalar field.

We have confirmed that our numerical code is generating physical data. Our main phys-

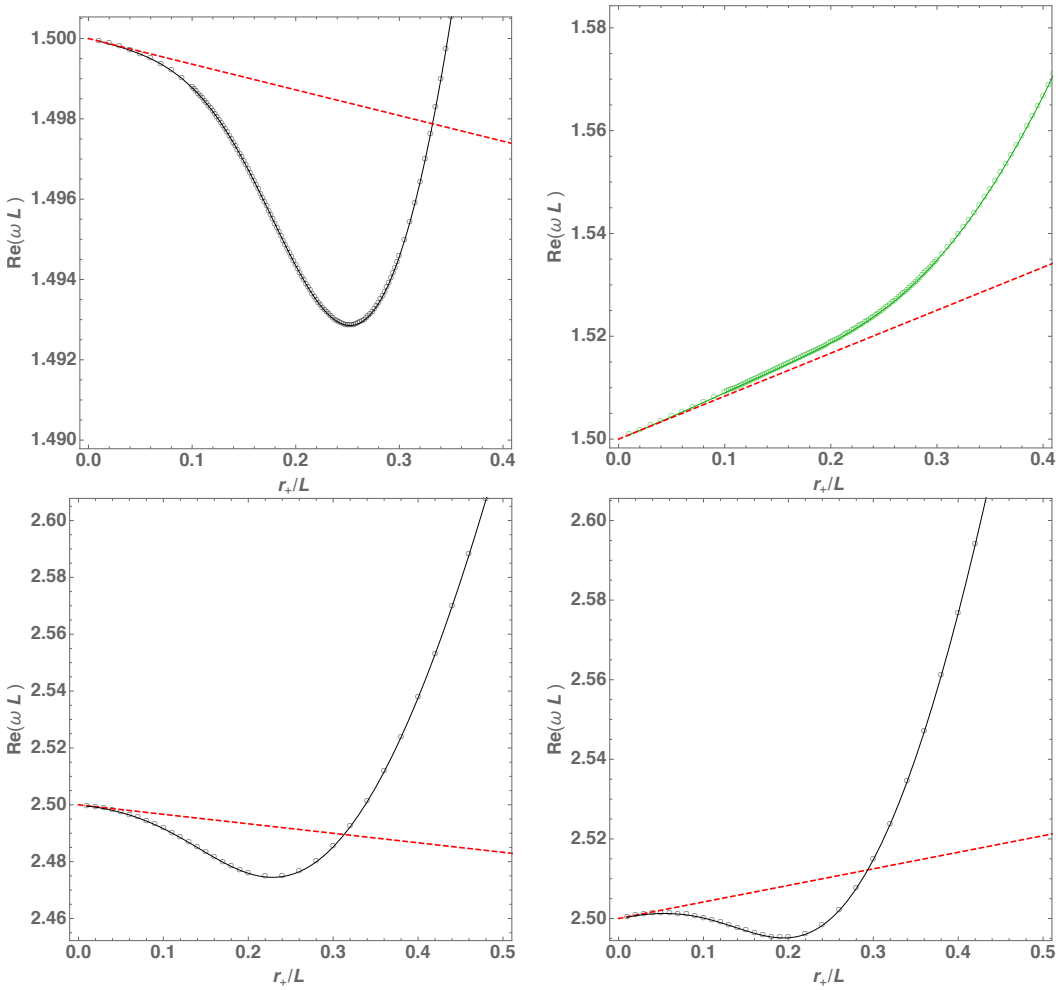

Figure A.4.3: Real part of the frequencies as a function of the horizon radius. *Top panel*: Alternative quantization (2.2.37) and fixed chemical potential $\mu = \mu_{\text{ext}}(1 - 10^{-2})$ and charges $qL = 0.7$ (left panel) and $qL = 0.8$ (right panel). The red dashed line describes the perturbative result (A.4.4). *Bottom*: Standard quantization (2.2.36) and fixed chemical potential $\mu = \frac{\mu_{\text{ext}}}{2}$ and charges $qL = 1.9$ (left panel) and $qL = 2$ (right panel). The red dashed line describes the perturbative result (A.4.3).

ical results are reported in section 2.4.4 of the main text.

APPENDIX B

Charged black hole bomb

B.1 Further properties of solitons in a Minkowski box

For completeness, in this appendix we plot the soliton quasilocal mass $\mathcal{M}$, charge $\mathcal{Q}$ and chemical potential μ as a function of the marching parameters $\epsilon = -\phi'(1)$ or $f_0 = f(0)$. We do this for the two most representative cases studied in the main text. Namely, for the solitons with $e = 1.854 \lesssim e_c$ (in Fig. B.1.1) and $e = 1.855 \gtrsim e_c$ (in Fig. B.1.2).

These quantities display the spiral behaviour already described in the main text. In particular, the turning points of these spirals translate into the regular cusps in the quasilocal charge plots: compare *e.g.* Fig. B.1.1 with Fig. 3.4.4 or Fig. B.1.2 with Fig. 3.4.7. Complementing the discussion given in the main text, these plots also illustrate how points A of the main (black) curve and point a of the secondary (magenta) curve) approach each other as e tends to e_c from below (Fig. B.1.1) and then give origin to a new branch of main (black) and secondary (blue) soliton families (Fig. B.1.2)

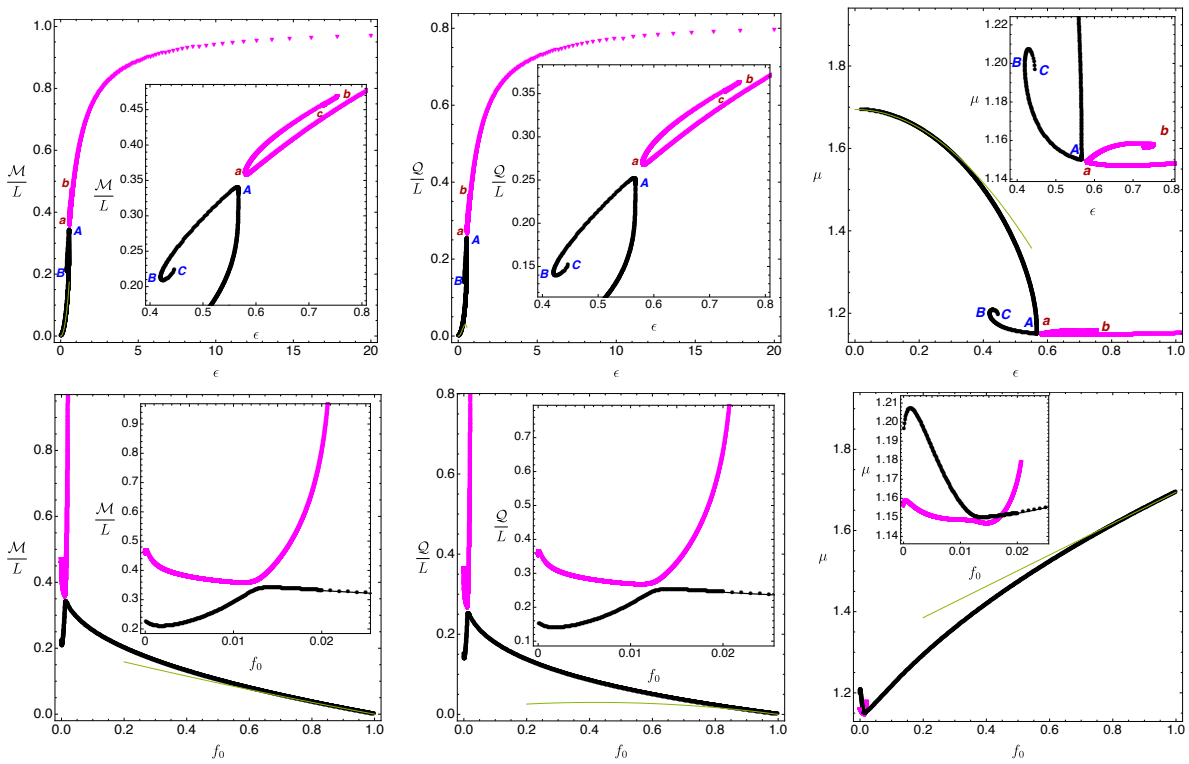

Figure B.1.1: Soliton families with $e = 1.854$ ($e_\gamma < e < e_c$). **Top panel:** The quasilocal thermodynamic quantities, namely mass, charge and chemical potential of the main (black disks) and secondary (magenta triangles) soliton families are shown as functions of the scalar field amplitude ϵ. **Bottom panel:** This time the mass, charge and chemical potential are plotted as functions of $f_0 \equiv f(0)$.

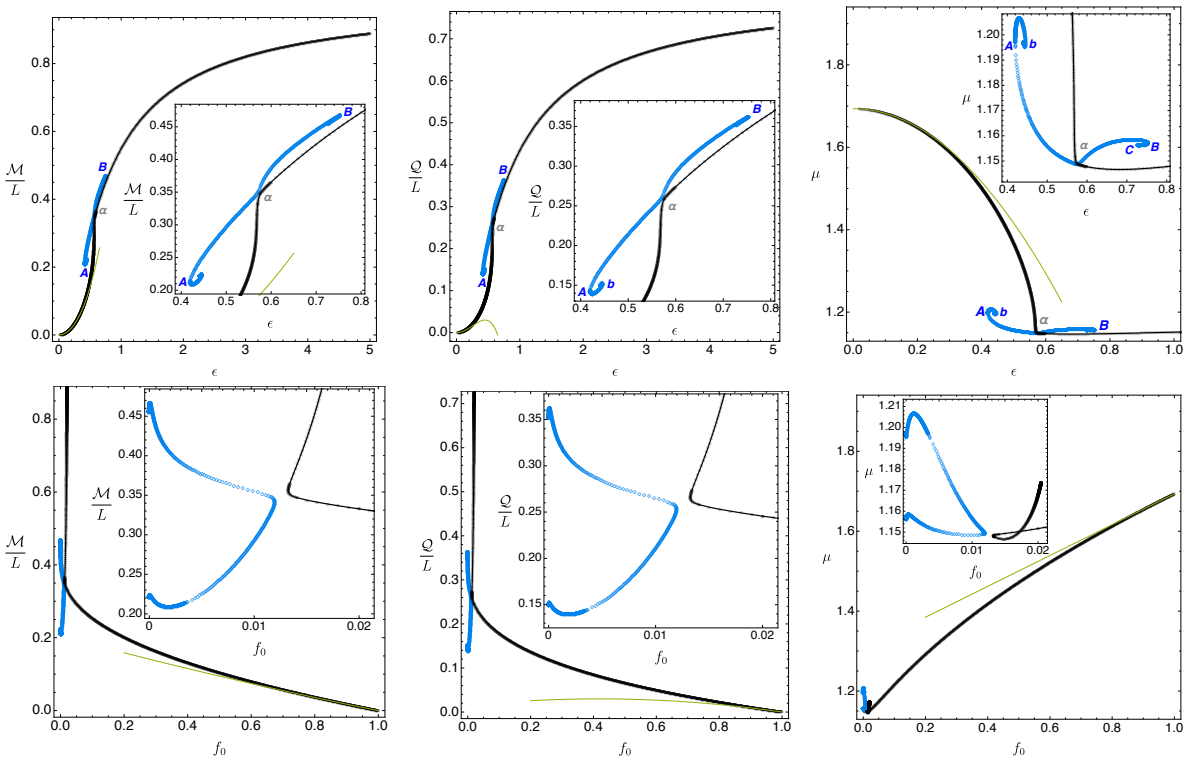

Figure B.1.2: Soliton families with $e = 1.855$ ($e_c < e < e_s$). **Top panel:** The quasilocal thermodynamic quantities, namely mass, charge and chemical potential of the main (black disks) and secondary (blue diamons) soliton families are shown as functions of the scalar field amplitude ϵ. **Bottom panel:** Quasilocal mass, charge and chemical potential plotted as functions of f_0.

Bibliography

[1] O. J. Dias, R. Masachs, O. Papadoulaki and P. Rodgers, *Hunting for fermionic instabilities in charged AdS black holes*, *Journal of High Energy Physics* **2020** (2020) 1–55.

[2] O. J. C. Dias, R. Masachs and P. Rodgers, *Boson stars and solitons confined in a Minkowski box*, *Journal of High Energy Physics* **2021** (2021) 236.

[3] A. Davey, O. J. Dias and P. Rodgers, *Phase diagram of the charged black hole bomb system*, *Journal of High Energy Physics* **2021** (2021) 1–36.

[4] S. W. Hawking and G. F. R. Ellis, *The large scale structure of space-time*, vol. 1. Cambridge university press, 1973.

[5] R. M. Wald, *General relativity*. Chicago Univ. Press, Chicago, IL, 1984.

[6] K. Schwarzschild, *Über das Gravitationsfeld einer Kugel aus inkompressibler Flüssigkeit nach der Einsteinschen Theorie*, *Sitzungsberichte der königlich preußischen Akademie der Wissenschaften zu Berlin* (1916) 424–434.

[7] R. L. Arnowitt, S. Deser and C. W. Misner, *The Dynamics of general relativity*, *Gen. Rel. Grav.* **40** (2008) 1997–2027, [gr-qc/0405109].

[8] S. A. Teukolsky, *Perturbations of a rotating black hole. I. Fundamental equations for gravitational, electromagnetic, and neutrino-field perturbations*, *The Astrophysical Journal* **185** (1973) 635–648.

[9] E. Berti, V. Cardoso and A. O. Starinets, *Quasinormal modes of black holes and black branes*, *Classical and Quantum Gravity* **26** (2009) 163001.

[10] N. D. Birrell and P. C. W. Davies, *Quantum Fields in Curved Space*. Cambridge Monographs on Mathematical Physics. Cambridge Univ. Press, Cambridge, UK, 1984, 10.1017/CBO9780511622632.

[11] D. Tong, *Lectures on Quantum Field Theory (Part III)*. Cambridge Univ., Cambridge, UK, 2007.

[12] M. D. Pollock, *On the Dirac equation in curved space-time*, *Acta Phys. Polon.* **B41** (2010) 1827–1846.

[13] J. Yepez, *Einstein's vierbein field theory of curved space*, `1106.2037`.

[14] I. I. Cotaescu, *Discrete quantum modes of the Dirac field in AdS (d+1) backgrounds*, *Int. J. Mod. Phys.* **A19** (2004) 2217–2232, [`gr-qc/0306127`].

[15] J. M. Maldacena, *The Large N limit of superconformal field theories and supergravity*, *Int. J. Theor. Phys.* **38** (1999) 1113–1133, [`hep-th/9711200`].

[16] J. Erlich, E. Katz, D. T. Son and M. A. Stephanov, *QCD and a holographic model of hadrons*, *Physical Review Letters* **95** (2005) 261602.

[17] T. Sakai and S. Sugimoto, *Low energy hadron physics in holographic QCD*, *Progress of Theoretical Physics* **113** (2005) 843–882.

[18] T. Sakai and S. Sugimoto, *More on a holographic dual of QCD*, *Progress of Theoretical Physics* **114** (2005) 1083–1118.

[19] N. R. Constable and R. C. Myers, *Exotic scalar states in the AdS/CFT correspondence*, *Journal of High Energy Physics* **1999** (1999) 020.

[20] S. de Haro, S. N. Solodukhin and K. Skenderis, *Holographic reconstruction of space-time and renormalization in the AdS / CFT correspondence*, *Commun. Math. Phys.* **217** (2001) 595–622, [`hep-th/0002230`].

[21] K. Skenderis and B. C. van Rees, *Real-time gauge/gravity duality: Prescription, Renormalization and Examples*, *Journal of High Energy Physics* **2009** (2009) 085.

[22] T. Ortín, *Gravity and strings*. Cambridge University Press, 2004.

[23] M. J. Duff, B. E. Nilsson and C. N. Pope, *Kaluza-klein supergravity*, *Physics Reports* **130** (1986) 1–142.

[24] M. Cveti, H. Lü and C. Pope, *Consistent Kaluza-Klein sphere reductions*, *Physical Review D* **62** (2000) 064028.

[25] G. 't Hooft, *A planar diagram theory for strong interactions*, in *The Large N Expansion In Quantum Field Theory And Statistical Physics: From Spin Systems to 2-Dimensional Gravity*, pp. 80–92. World Scientific, 1993.

[26] O. Aharony, S. S. Gubser, J. Maldacena, H. Ooguri and Y. Oz, *Large N field theories, string theory and gravity*, *Physics Reports* **323** (2000) 183–386.

[27] P. Breitenlohner and D. Z. Freedman, *Positive Energy in anti-De Sitter Backgrounds and Gauged Extended Supergravity*, *Phys. Lett.* **115B** (1982) 197–201.

[28] P. Breitenlohner and D. Z. Freedman, *Stability in Gauged Extended Supergravity*, *Annals Phys.* **144** (1982) 249.

[29] S. S. Gubser, I. R. Klebanov and A. M. Polyakov, *Gauge theory correlators from non-critical string theory*, *Physics Letters B* **428** (1998) 105–114.

[30] K. Skenderis, *Lecture notes on holographic renormalization*, *Classical and Quantum Gravity* **19** (2002) 5849.

[31] J. D. Bekenstein, *Black hole hair: 25 - years after*, in *Physics. Proceedings, 2nd International A.D. Sakharov Conference, Moscow, Russia, May 20-24, 1996*, pp. 216–219, 1996. `gr-qc/9605059`.

[32] P. T. Chrusciel, *'No hair' theorems: Folklore, conjectures, results*, *Contemp. Math.* **170** (1994) 23–49, [`gr-qc/9402032`].

[33] D. Wiltshire, *Spherically symmetric solutions of Einstein-Maxwell theory with a Gauss-Bonnet term*, *Physics Letters B* **169** (1986) 36 – 40.

[34] R. D'Inverno, *Introducing Einstein's Relativity*. Clarendon Press, 1992.

[35] W. H. Press and S. A. Teukolsky, *Floating Orbits, Superradiant Scattering and the Black-hole Bomb*, *Nature* **238** (1972) 211–212.

[36] G. Denardo and R. Ruffini, *On the energetics of Reissner Nordström geometries*, *Phys. Lett.* **B45** (1973) 259–262.

[37] S. A. Hartnoll, C. P. Herzog and G. T. Horowitz, *Building a Holographic Superconductor*, *Phys. Rev. Lett.* **101** (2008) 031601, [0803.3295].

[38] S. A. Hartnoll, C. P. Herzog and G. T. Horowitz, *Holographic Superconductors*, *JHEP* **12** (2008) 015, [0810.1563].

[39] O. J. C. Dias and R. Masachs, *Hairy black holes and the endpoint of AdS_4 charged superradiance*, *JHEP* **02** (2017) 128, [1610.03496].

[40] O. J. C. Dias and R. Masachs, *Charged black hole bombs in a Minkowski cavity*, *Class. Quant. Grav.* **35** (2018) 184001, [1801.10176].

[41] J. C. Degollado and C. A. R. Herdeiro, *Time evolution of superradiant instabilities for charged black holes in a cavity*, *Phys. Rev.* **D89** (2014) 063005, [1312.4579].

[42] R. Li, H. Zhang and J. Zhao, *Time evolutions of scalar field perturbations in D-dimensional Reissner-Nordström Anti-de Sitter black holes*, *Phys. Lett.* **B758** (2016) 359–364, [1604.01267].

[43] O. J. C. Dias and R. Masachs, *Evading no-hair theorems: hairy black holes in a Minkowski box*, *Phys. Rev.* **D97** (2018) 124030, [1802.01603].

[44] O. J. C. Dias, P. Figueras, S. Minwalla, P. Mitra, R. Monteiro and J. E. Santos, *Hairy black holes and solitons in global AdS_5*, *JHEP* **08** (2012) 117, [1112.4447].

[45] S. Bhattacharyya, S. Minwalla and K. Papadodimas, *Small Hairy Black Holes in $AdS_5 x S^5$*, *JHEP* **11** (2011) 035, [1005.1287].

[46] P. Basu, J. Bhattacharya, S. Bhattacharyya, R. Loganayagam, S. Minwalla and V. Umesh, *Small Hairy Black Holes in Global AdS Spacetime*, *JHEP* **10** (2010) 045, [1003.3232].

[47] J. Markeviciute and J. E. Santos, *Hairy black holes in $AdS_5 \times S^5$*, *JHEP* **06** (2016) 096, [1602.03893].

[48] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*, vol. 55. Courier Corporation, 1965.

[49] J. M. Bardeen and G. T. Horowitz, *The Extreme Kerr throat geometry: A Vacuum analog of AdS(2) x S**2*, *Phys. Rev.* **D60** (1999) 104030, [hep-th/9905099].

[50] R. Brito, V. Cardoso and P. Pani, *Superradiance*, 1501.06570.

[51] I. R. Klebanov and E. Witten, *AdS / CFT correspondence and symmetry breaking*, *Nucl. Phys.* **B556** (1999) 89–114, [hep-th/9905104].

[52] A. Ishibashi and R. M. Wald, *Dynamics in nonglobally hyperbolic static space-times. 3. Anti-de Sitter space-time*, *Class. Quant. Grav.* **21** (2004) 2981–3014, [hep-th/0402184].

[53] S. S. Gubser, *Breaking an Abelian gauge symmetry near a black hole horizon*, *Phys. Rev.* **D78** (2008) 065034, [0801.2977].

[54] S. A. Hartnoll, C. P. Herzog and G. T. Horowitz, *Building a holographic superconductor*, *Physical Review Letters* **101** (2008) 031601.

[55] T. Faulkner, H. Liu, J. McGreevy and D. Vegh, *Emergent quantum criticality, Fermi surfaces, and AdS(2)*, *Phys. Rev.* **D83** (2011) 125002, [0907.2694].

[56] O. J. C. Dias, R. Monteiro, H. S. Reall and J. E. Santos, *A Scalar field condensation instability of rotating anti-de Sitter black holes*, *JHEP* **11** (2010) 036, [1007.3745].

[57] O. J. Dias, P. Figueras, S. Minwalla, P. Mitra, R. Monteiro and J. E. Santos, *Hairy black holes and solitons in global AdS 5*, *Journal of High Energy Physics* **2012** (2012) 117.

[58] O. Klein, *Die Reflexion von Elektronen an einem Potentialsprung nach der relativistischen Dynamik von Dirac*, *Z. Phys.* **53** (1929) 157.

[59] F. Sauter, *Uber das Verhalten eines Elektrons im homogenen elektrischen Feld nach der relativistischen Theorie Diracs*, *Z. Phys.* **69** (1931) 742–764.

[60] R. G. Winter, *Klein Paradox for the Klein-Gordon Equation*, *American Journal of Physics* **27** (1959) 355.

[61] W. Unruh, *Separability of the Neutrino Equations in a Kerr Background*, *Phys. Rev. Lett.* **31** (1973) 1265–1267.

[62] B. R. Holstein, *Klein's Paradox*, *American Journal of Physics* **66** (1998) 507.

[63] N. Iqbal and H. Liu, *Real-time response in AdS/CFT with application to spinors*, *Fortsch. Phys.* **57** (2009) 367–384, [0903.2596].

[64] S.-S. Lee, *A Non-Fermi Liquid from a Charged Black Hole: A Critical Fermi Ball*, *Phys. Rev.* **D79** (2009) 086006, [0809.3402].

[65] H. Liu, J. McGreevy and D. Vegh, *Non-Fermi liquids from holography*, *Phys. Rev.* **D83** (2011) 065029, [0903.2477].

[66] D. Guarrera and J. McGreevy, *Holographic Fermi surfaces and bulk dipole couplings*, 1102.3908.

[67] N. Iqbal, H. Liu and M. Mezei, *Lectures on holographic non-Fermi liquids and quantum phase transitions*, in *Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics (TASI 2010). String Theory and Its Applications: From meV to the Planck Scale: Boulder, Colorado, USA, June 1-25, 2010*, pp. 707–816, 2011. 1110.3814. DOI.

[68] S. A. Hartnoll, A. Lucas and S. Sachdev, *Holographic quantum matter*, 1612.07324.

[69] M. Cubrovic, J. Zaanen and K. Schalm, *String Theory, Quantum Phase Transitions and the Emergent Fermi-Liquid*, *Science* **325** (2009) 439–444, [0904.1993].

[70] V. Cardoso, J. a. L. Costa, K. Destounis, P. Hintz and A. Jansen, *Quasinormal modes and Strong Cosmic Censorship*, *Phys. Rev. Lett.* **120** (2018) 031103, [1711.10502].

[71] O. J. C. Dias, F. C. Eperon, H. S. Reall and J. E. Santos, *Strong cosmic censorship in de Sitter space*, *Phys. Rev.* **D97** (2018) 104060, [1801.09694].

[72] O. J. C. Dias, H. S. Reall and J. E. Santos, *Strong cosmic censorship: taking the rough with the smooth*, *JHEP* **10** (2018) 001, [1808.02895].

[73] M. Wang, C. Herdeiro and J. Jing, *Dirac perturbations on Schwarzschild–anti–de Sitter spacetimes: Generic boundary conditions and new quasinormal modes*, *Phys. Rev.* **D96** (2017) 104035, [1710.10461].

[74] M. Wang, C. Herdeiro and J. Jing, *Charged Dirac perturbations on Reissner-Nordström-Anti-de Sitter spacetimes: quasinormal modes with Robin boundary conditions*, 1910.14305.

[75] W. Mueck and K. S. Viswanathan, *Conformal field theory correlators from classical field theory on anti-de Sitter space. 2. Vector and spinor fields*, *Phys. Rev.* **D58** (1998) 106006, [hep-th/9805145].

[76] M. Henningson and K. Sfetsos, *Spinors and the AdS / CFT correspondence*, *Phys. Lett.* **B431** (1998) 63–68, [hep-th/9803251].

[77] M. Henneaux, *Boundary terms in the AdS / CFT correspondence for spinor fields*,

in *Mathematical methods in modern theoretical physics. Proceedings, International Meeting, School and Workshop, ISPM'98, Tbilisi, Georgia, September 5-18, 1998*, pp. 161–170, 1998. hep-th/9902137.

[78] R. Contino and A. Pomarol, *Holography for fermions*, *JHEP* **11** (2004) 058, [hep-th/0406257].

[79] A. J. Amsel and D. Marolf, *Supersymmetric Multi-trace Boundary Conditions in AdS*, *Class. Quant. Grav.* **26** (2009) 025010, [0808.2184].

[80] A. J. Amsel and G. Compere, *Supergravity at the boundary of AdS supergravity*, *Phys. Rev.* **D79** (2009) 085006, [0901.3609].

[81] T. Andrade and D. Marolf, *AdS/CFT beyond the unitarity bound*, *JHEP* **01** (2012) 049, [1105.6337].

[82] G. Michalogiorgakis and S. S. Pufu, *Low-lying gravitational modes in the scalar sector of the global AdS(4) black hole*, *JHEP* **02** (2007) 023, [hep-th/0612065].

[83] O. J. C. Dias and J. E. Santos, *Boundary Conditions for Kerr-AdS Perturbations*, *JHEP* **10** (2013) 156, [1302.1580].

[84] V. Cardoso, O. J. C. Dias, G. S. Hartnett, L. Lehner and J. E. Santos, *Holographic thermalization, quasinormal modes and superradiance in Kerr-AdS*, *JHEP* **04** (2014) 183, [1312.5323].

[85] M. Giammatteo and J.-l. Jing, *Dirac quasinormal frequencies in Schwarzschild-AdS space-time*, *Phys. Rev.* **D71** (2005) 024007, [gr-qc/0403030].

[86] J.-l. Jing, *Quasinormal modes of Dirac field perturbation in Schwarzschild-anti-de Sitter black hole*, gr-qc/0502010.

[87] B. H. J. McKellar, M. J. Thomson and G. J. Stephenson, *The Dirac equation in Kerr space-time, spheroidal coordinates and the MIT bag model of hadrons*, *J. Phys.* **A26** (1993) 3649–3657.

[88] S. R. Dolan and D. Dempsey, *Bound states of the Dirac equation on Kerr spacetime*, *Class. Quant. Grav.* **32** (2015) 184001, [1504.03190].

[89] B. Wang, C.-Y. Lin and E. Abdalla, *Quasinormal modes of Reissner-Nordstrom anti-de Sitter black holes*, *Phys. Lett.* **B481** (2000) 79–88, [hep-th/0003295].

[90] E. Berti and K. D. Kokkotas, *Quasinormal modes of Reissner-Nordstrom-anti-de Sitter black holes: Scalar, electromagnetic and gravitational perturbations*, *Phys. Rev.* **D67** (2003) 064020, [gr-qc/0301052].

[91] N. Uchikata and S. Yoshida, *Quasinormal modes of a massless charged scalar field on a small Reissner-Nordstrom-anti-de Sitter black hole*, *Phys. Rev.* **D83** (2011) 064020, [1109.6737].

[92] O. J. C. Dias, H. S. Reall and J. E. Santos, *Strong cosmic censorship for charged de Sitter black holes with a charged scalar field*, *Class. Quant. Grav.* **36** (2019) 045005, [1808.04832].

[93] R. Arias, J. Mas and A. Serantes, *Stability of charged global AdS_4 spacetimes*, *JHEP* **09** (2016) 024, [1606.00830].

[94] A. Lopez-Ortega, *Absorption and quasinormal modes of classical fields propagating on 3D and 4D de Sitter spacetime*, *General Relativity and Gravitation* **38** (2006) 743.

[95] I. I. Cotaescu, *The Dirac particle on central backgrounds and the anti-de Sitter oscillator*, *Mod. Phys. Lett.* **A13** (1998) 2923–2936, [gr-qc/9803042].

[96] B. C. van Rees, *Real-time gauge/gravity duality and ingoing boundary conditions*, *Nuclear Physics B-Proceedings Supplements* **192** (2009) 193–196.

[97] O. J. C. Dias, P. Figueras, R. Monteiro, J. E. Santos and R. Emparan, *Instability and new phases of higher-dimensional rotating black holes*, *Phys. Rev.* **D80** (2009) 111701, [0907.2248].

[98] O. J. C. Dias, P. Figueras, R. Monteiro and J. E. Santos, *Ultraspinning instability of rotating black holes*, *Phys. Rev.* **D82** (2010) 104025, [1006.1904].

[99] O. J. C. Dias, P. Figueras, R. Monteiro, H. S. Reall and J. E. Santos, *An instability of higher-dimensional rotating black holes*, *JHEP* **05** (2010) 076, [1001.4527].

[100] O. J. C. Dias, P. Figueras, R. Monteiro and J. E. Santos, *Ultraspinning instability of anti-de Sitter black holes*, *JHEP* **12** (2010) 067, [1011.0996].

[101] O. J. C. Dias, R. Monteiro and J. E. Santos, *Ultraspinning instability: the missing link*, *JHEP* **08** (2011) 139, [1106.4554].

[102] O. J. C. Dias, G. S. Hartnett and J. E. Santos, *Quasinormal modes of asymptotically*

flat rotating black holes, *Class. Quant. Grav.* **31** (2014) 245011, [1402.7047].

[103] O. J. C. Dias, M. Godazgar and J. E. Santos, *Linear Mode Stability of the Kerr-Newman Black Hole and Its Quasinormal Modes*, *Phys. Rev. Lett.* **114** (2015) 151101, [1501.04625].

[104] O. J. C. Dias, J. E. Santos and B. Way, *Numerical Methods for Finding Stationary Gravitational Solutions*, *Class. Quant. Grav.* **33** (2016) 133001, [1510.02804].

[105] K. Murata, S. Kinoshita and N. Tanahashi, *Non-equilibrium Condensation Process in a Holographic Superconductor*, *JHEP* **07** (2010) 050, [1005.0633].

[106] P. Bosch, S. R. Green and L. Lehner, *Nonlinear Evolution and Final Fate of Charged Anti-de Sitter Black Hole Superradiant Instability*, *Phys. Rev. Lett.* **116** (2016) 141102, [1601.01384].

[107] S. A. Hartnoll and A. Tavanfar, *Electron stars for holographic metallic criticality*, *Phys. Rev.* **D83** (2011) 046003, [1008.2828].

[108] A. Allais and J. McGreevy, *How to construct a gravitating quantum electron star*, *Phys. Rev.* **D88** (2013) 066006, [1306.6075].

[109] V. Cardoso, O. J. C. Dias, J. P. S. Lemos and S. Yoshida, *The Black hole bomb and superradiant instabilities*, *Phys. Rev.* **D70** (2004) 044039, [hep-th/0404096].

[110] S. A. Gentle, M. Rangamani and B. Withers, *A Soliton Menagerie in AdS*, *JHEP* **05** (2012) 106, [1112.3979].

[111] J. Markeviciute, *Rotating Hairy Black Holes in $AdS_5 \times S^5$*, *JHEP* **03** (2019) 110, [1809.04084].

[112] J. D. Brown and J. W. York, Jr., *Quasilocal energy and conserved charges derived from the gravitational action*, *Phys. Rev.* **D47** (1993) 1407–1419, [gr-qc/9209012].

[113] W. Israel, *Singular hypersurfaces and thin shells in general relativity*, *Nuovo Cim.* **B44S10** (1966) 1.

[114] W. Israel, *Discontinuities in Spherically Symmetric Gravitational Fields and Shells of Radiation*, *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **248** (1958) 404–414, [http://rspa.royalsocietypublishing.org/content/royprsa/248/1254/404.full.pdf].

[115] K. Kuchar, *Charged shells in general relativity and their gravitational collapse*, *Czech. J. Phys. B* **B18** (1968) 435.

[116] C. Barrabes and W. Israel, *Thin shells in general relativity and cosmology: The Lightlike limit*, *Phys. Rev.* **D43** (1991) 1129–1142.

[117] C. W. Misner, K. S. Thorne and J. A. Wheeler, *Gravitation*. W.H. Freeman and Co., San Francisco, 1973.

[118] C. A. R. Herdeiro, J. C. Degollado and H. F. Rúnarsson, *Rapid growth of superradiant instabilities for charged black holes in a cavity*, *Phys. Rev.* **D88** (2013) 063003, [1305.5513].

[119] S. Hod, *Analytic treatment of the charged black-hole-mirror bomb in the highly explosive regime*, *Phys. Rev.* **D88** (2013) 064055, [1310.6101].

[120] S. Hod, *Resonance spectra of caged black holes*, *Eur. Phys. J.* **C74** (2014) 3137, [1410.4567].

[121] R. Li, J.-K. Zhao and Y.-M. Zhang, *Superradiant Instability of D-Dimensional Reissner-Nordstrm Black Hole Mirror System*, *Commun. Theor. Phys.* **63** (2015) 569–574, [1404.6309].

[122] S. Hod, *The charged black-hole bomb: A lower bound on the charge-to-mass ratio of the explosive scalar field*, *Phys. Lett.* **B755** (2016) 177–182, [1606.00444].

[123] O. Fierro, N. Grandi and J. Oliva, *Superradiance of charged black holes in Einstein-Gauss-Bonnet Gravity*, 1708.06037.

[124] R. Li and J. Zhao, *Superradiant instability of charged scalar field in stringy black hole mirror system*, *Eur. Phys. J.* **C74** (2014) 3051, [1403.7279].

[125] R. Li and J. Zhao, *Numerical study of superradiant instability for charged stringy black hole mirror system*, *Phys. Lett.* **B740** (2015) 317–321, [1412.1527].

[126] R. Li, Y. Tian, H.-b. Zhang and J. Zhao, *Time domain analysis of superradiant instability for the charged stringy black hole mirror system*, *Phys. Lett.* **B750** (2015) 520–527, [1506.04267].

[127] R. Li, J. Zhao, X. Wu and Y. Zhang, *Scalar clouds in charged stringy black hole-mirror system*, *Eur. Phys. J.* **C75** (2015) 142, [1501.07358].

[128] S. R. Dolan, S. Ponglertsakul and E. Winstanley, *Stability of black holes in Einstein-charged scalar field theory in a cavity*, *Phys. Rev.* **D92** (2015) 124047, [1507.02156].

[129] S. Ponglertsakul, E. Winstanley and S. R. Dolan, *Stability of gravitating charged-scalar solitons in a cavity*, *Phys. Rev.* **D94** (2016) 024031, [1604.01132].

[130] S. Ponglertsakul and E. Winstanley, *Effect of scalar field mass on gravitating charged scalar solitons and black holes in a cavity*, *Phys. Lett.* **B764** (2017) 87–93, [1610.00135].

[131] P. Basu, C. Krishnan and P. N. B. Subramanian, *Hairy Black Holes in a Box*, *JHEP* **11** (2016) 041, [1609.01208].

[132] N. Sanchis-Gual, J. C. Degollado, P. J. Montero, J. A. Font and C. Herdeiro, *Explosion and Final State of an Unstable Reissner-Nordström Black Hole*, *Phys. Rev. Lett.* **116** (2016) 141101, [1512.05358].

[133] N. Sanchis-Gual, J. C. Degollado, C. Herdeiro, J. A. Font and P. J. Montero, *Dynamical formation of a Reissner-Nordström black hole with scalar hair in a cavity*, *Phys. Rev.* **D94** (2016) 044061, [1607.06304].

[134] N. Sanchis-Gual, J. C. Degollado, J. A. Font, C. Herdeiro and E. Radu, *Dynamical formation of a hairy black hole in a cavity from the decay of unstable solitons*, *Class. Quant. Grav.* **34** (2017) 165001, [1611.02441].

[135] Y. B. Zel'dovich, *Generation of Waves by a Rotating Body*, *JETP Lett.* **14** (1971) 180.

[136] R. Penrose, *Gravitational collapse: The role of general relativity*, *Riv. Nuovo Cim.* **1** (1969) 252–276.

[137] D. Christodoulou, *Reversible and Irreversible Transformations in Black-Hole Physics*, *Phys. Rev. Lett.* **25** (Nov, 1970) 1596–1597.

[138] O. J. C. Dias, G. T. Horowitz and J. E. Santos, *Black holes with only one Killing field*, *JHEP* **07** (2011) 115, [1105.4167].

[139] O. J. C. Dias, J. E. Santos and B. Way, *Black holes with a single Killing vector field: black resonators*, *JHEP* **12** (2015) 171, [1505.04793].

[140] M. W. Choptuik, O. J. C. Dias, J. E. Santos and B. Way, *Collapse and Nonlinear Instability of AdS Space with Angular Momentum*, *Phys. Rev. Lett.* **119** (2017)

191104, [1706.06101].

[141] T. Ishii and K. Murata, *Black resonators and geons in AdS5*, *Class. Quant. Grav.* **36** (2019) 125011, [1810.11089].

[142] T. Ishii, K. Murata, J. E. Santos and B. Way, *Superradiant instability of black resonators and geons*, *JHEP* **07** (2020) 206, [2005.01201].

[143] T. Ishii, K. Murata, J. E. Santos and B. Way, *Multioscillating black holes*, 2101.06325.

[144] C. A. R. Herdeiro and E. Radu, *Kerr black holes with scalar hair*, *Phys. Rev. Lett.* **112** (2014) 221101, [1403.2757].

[145] S. L. Liebling and C. Palenzuela, *Dynamical Boson Stars*, *Living Rev. Rel.* **20** (2017) 5, [1202.5809].

[146] S. W. Hawking, *Black Holes and Thermodynamics*, *Phys. Rev. D* **13** (1976) 191–197.

[147] G. W. Gibbons and M. J. Perry, *Black Holes and Thermal Green's Functions*, *Proc. Roy. Soc. Lond.* **A358** (1978) 467–494.

[148] S. W. Hawking and W. Israel, *Penrose, R. (Chapter 12) at General Relativity: An Einstein Centenary Survey*. Univ. Pr., Cambridge, UK, 1979.

[149] D. Page, *Black hole formation in a box*, *Gen Relat Gravit.* **13** (1981) 1117–1126.

[150] S. W. Hawking and D. N. Page, *Thermodynamics of Black Holes in anti-De Sitter Space*, *Commun. Math. Phys.* **87** (1983) 577.

[151] H. W. Braden, J. D. Brown, B. F. Whiting and J. W. York, Jr., *Charged black hole in a grand canonical ensemble*, *Phys. Rev.* **D42** (1990) 3376–3385.

[152] T. Andrade, W. R. Kelly, D. Marolf and J. E. Santos, *On the stability of gravity with Dirichlet walls*, *Class. Quant. Grav.* **32** (2015) 235006, [1504.07580].

[153] M. Maliborski and A. Rostworowski, *Time-Periodic Solutions in an Einstein AdS-Massless-Scalar-Field System*, *Phys. Rev. Lett.* **111** (2013) 051102, [1303.3186].

[154] M. Dafermos, *The Black Hole Stability problem*, in *Talk at the Newton Institute*, Available at: http://www-old.newton.ac.uk/webseminars/pg+ws/2006/gmx/1010/dafermos/, University

of Cambridge, 2006.

[155] M. Dafermos and G. Holzegel, *Dynamic instability of solitons in 4+1 dimensional gravity with negative cosmological constant*, in *Seminar at DAMTP*, Available at: https://www.dpmms.cam.ac.uk/∼md384/ADSinstability.pdf, University of Cambridge, 2006.

[156] P. Bizon and A. Rostworowski, *On weakly turbulent instability of anti-de Sitter space*, *Phys. Rev. Lett.* **107** (2011) 031102, [1104.3702].

[157] O. J. C. Dias, G. T. Horowitz and J. E. Santos, *Gravitational Turbulent Instability of Anti-de Sitter Space*, *Class. Quant. Grav.* **29** (2012) 194002, [1109.1825].

[158] O. J. C. Dias, G. T. Horowitz, D. Marolf and J. E. Santos, *On the Nonlinear Stability of Asymptotically Anti-de Sitter Solutions*, *Class. Quant. Grav.* **29** (2012) 235019, [1208.5772].

[159] A. Buchel, L. Lehner and S. L. Liebling, *Scalar Collapse in AdS*, *Phys. Rev.* **D86** (2012) 123011, [1210.0890].

[160] A. Buchel, S. L. Liebling and L. Lehner, *Boson stars in AdS spacetime*, *Phys. Rev.* **D87** (2013) 123006, [1304.4166].

[161] O. J. C. Dias and J. E. Santos, *AdS nonlinear instability: moving beyond spherical symmetry*, 1602.03890.

[162] A. Rostworowski, *Comment on "AdS nonlinear instability: moving beyond spherical symmetry" [Class. Quantum Grav. 33 23LT01 (2016)]*, 1612.00042.

[163] O. J. C. Dias and J. E. Santos, *AdS nonlinear instability: breaking spherical and axial symmetries*, *Class. Quant. Grav.* **35** (2018) 185006, [1705.03065].

[164] V. Balasubramanian, A. Buchel, S. R. Green, L. Lehner and S. L. Liebling, *Holographic Thermalization, Stability of Anti-de Sitter Space, and the Fermi-Pasta-Ulam Paradox*, *Phys. Rev. Lett.* **113** (2014) 071601, [1403.6471].

[165] P. Bizon and A. Rostworowski, *Comment on Holographic Thermalization, Stability of Anti-de Sitter Space, and the Fermi-Pasta-Ulam Paradox?*, *Phys. Rev. Lett.* **115** (2015) 049101, [1410.2631].

[166] E. da Silva, E. Lopez, J. Mas and A. Serantes, *Collapse and Revival in Holographic Quenches*, *JHEP* **04** (2015) 038, [1412.6002].

[167] V. Balasubramanian, A. Buchel, S. R. Green, L. Lehner and S. L. Liebling, *Reply to Comment on Holographic Thermalization, Stability of Anti-de Sitter Space, and the Fermi-Pasta-Ulam Paradox?*, *Phys. Rev. Lett.* **115** (2015) 049102, [1506.07907].

[168] A. A. Starobinsky, *Amplification of waves reflected from a rotating "black hole".*, *Sov. Phys. JETP* **37** (1973) 28–32.

[169] W. G. Unruh, *Absorption cross section of small black holes*, *Phys. Rev. D* **14** (Dec, 1976) 3251–3259.

[170] J. Maldacena and A. Strominger, *Universal low-energy dynamics for rotating black holes*, *Physical Review D* **56** (1997) 4975.

[171] V. Cardoso and O. J. C. Dias, *Small Kerr-anti-de Sitter black holes are unstable*, *Phys. Rev.* **D70** (2004) 084011, [hep-th/0405006].