Perturbations of the Almost Killing Equation and their implications

Sumanta Chakraborty\textsuperscript{1, 2} and Justin C. Feng\textsuperscript{3}

\textsuperscript{1}School of Mathematical and Computational Sciences, Indian Association for the Cultivation of Science, Kolkata-700032, India
\textsuperscript{2}School of Physical Sciences, Indian Association for the Cultivation of Science, Kolkata-700032, India
\textsuperscript{3}CENTRA, Departamento de Física, Instituto Superior Técnico – IST, Universidade de Lisboa – UL, Avenida Rovisco Pais 1, 1049 Lisboa, Portugal

(Dated: December 1, 2020)

Killing vectors play a crucial role in characterizing the symmetries of a given spacetime. However, realistic astrophysical systems are in most cases only approximately symmetric. Even in the case of an astrophysical black hole, one might expect Killing symmetries to exist only in an approximate sense due to perturbations from external matter fields. In this work, we consider the generalized notion of Killing vectors provided by the almost Killing equation, and study the perturbations induced by a perturbation of a background spacetime satisfying exact Killing symmetry. To first order, we demonstrate that for nonradiative metric perturbations (that is, metric perturbations with nonvanishing trace) of symmetric vacuum spacetimes, the perturbed almost Killing equation avoids the problem of an unbounded Hamiltonian for hyperbolic parameter choices. For traceless metric perturbations, we obtain similar results for the second-order perturbation of the almost Killing equation, with some additional caveats. Thermodynamical implications have also been explored.

1. INTRODUCTION

Symmetries are central to our understanding of the physical world and play a key role in describing a wide range of physical systems, from the determination of the Lagrangian for a mechanical system to the lattice structure of crystalline substances. This extends to general relativity and relativistic theories of gravity: symmetries and the Killing vectors that formalize them are useful for characterizing the properties of spacetime and matter. For example, the existence of a timelike Killing vector field ensures that the spacetime is time-translation invariant, leading to conserved definitions for energy for test particles and matter. Similarly, the existence of a closed spacelike Killing vector field ensures rotational invariance of the spacetime geometry, leading to a conserved definition for angular momentum. Moreover, many astrophysical systems are approximately described by spacetime geometries admitting such Killing vectors. However, the exact nature of these symmetries are lost in realistic systems due to dynamical behavior of, and irregularities in the matter configurations. This scenario can arise in various contexts, e.g., when one drops a cup of coffee into a black hole (considering the gravitational backreaction), the resulting perturbed spacetime no longer inherits the exact Killing symmetry. Fortunately, one may still construct certain generalizations of Killing vector fields in such circumstances, which are useful for understanding generalizations of conserved quantities (such as energy and momentum) for gravitating systems that lack exact symmetries.

The literature contains several approaches for defining generalized Killing vectors and symmetries. Specific examples include Matzner’s Eigenvector approach \cite{Matzner}, which has recently been of interest for studying quantum geometries in Causal Dynamical Triangulations \cite{Cai2002}, symmetry-seeking coordinates \cite{Chacko2001}, affine collineations \cite{Affine} and the almost Killing equation (henceforth AKE) \cite{Arnowitt1, Arnowitt2}. The latter approach, the generalized Killing vectors defined by the AKE, forms the subject of this paper.

The generalized Killing vector fields (henceforth GKVs) associated with the AKE may be used to define conserved charges in spacetimes with no exact Killing symmetries. Given some notion of the GKV, the generalized Komar current, as defined in \cite{Kol:2004aa, Kol:2004ab} may be used to construct generalizations of the usual Komar charges — explicit examples have been constructed and studied in \cite{Achour:2015aa} and \cite{Achour:2015ab} (see also \cite{Achour:2015ac} for further generalizations of the Komar current). For example, in \cite{Kol:2004aa}, it was shown that the generalized Komar current for solutions of the AKE, which are the GKVs, can provide a measure of the matter content of the physical system under consideration. It was also demonstrated in \cite{Kol:2004aa} that GKVs may be used with the generalized Komar current to obtain a Gauss law for systems of black holes in vacuum and matter distributions with compact support if the GKV is divergenceless or for a certain choice of parameters associated with the AKE.

Though one might hope that, for sufficiently small perturbations of symmetric spacetimes, the solutions of the AKE are close to that of an exact Killing vector field, GKVs do not necessarily approximate Killing vectors in the sense that the components of $\nabla_{(\mu} \chi_{\nu)}$ can be large compared to that of $\chi^\alpha$ [where $\chi^\alpha$ is a GKV], even in Minkowski spacetime. One might postulate that an appropriate choice for initial data for the AKE will ensure that $\nabla_{(\mu} \chi_{\nu)}$ is small in the sense that the components of $\nabla_{(\mu} \chi_{\nu)}$ are much smaller than that of $\chi^\alpha$ for some normalization. This matter was studied to some degree in \cite{Kol:2004aa}, which examines the hyperbolicity and Hamiltonian stability of the system described by the AKE. There, parameter choices were identified in which the AKE is strongly and weakly hyperbolic,
and also in which the system admits ghost modes and unbounded Hamiltonians. Ghosts and unbounded Hamiltonians are potentially dangerous, as they may correspond to runaway behavior which can potentially drive the solution far from the Killing condition \( \nabla_{(\mu} \chi_{\nu)} = 0 \), even if the initial data approximately satisfies this condition (and its time derivative). Though there is no parameter choice for which the generic AKE system is both hyperbolic and has a bounded Hamiltonian, it was shown in [12] that in vacuum \( (R_{\mu \nu} = 0) \) spacetime and for initial data satisfying \( \nabla \cdot \chi = 0 \) and its derivative, the system yields a constraint which renders it dynamically equivalent to a system with a bounded Hamiltonian and simultaneously equivalent to a strongly hyperbolic system. Moreover, it was argued that for appropriate initial data and falloff conditions, the AKE can provide a notion of an approximate Killing vector in a neighborhood of spatial infinity of asymptotically flat spacetimes.

Despite the promising results presented in [12] for the vacuum case, these do not in general extend to the non-vacuum \( (R_{\mu \nu} \neq 0) \) case. Therefore, it is not immediately apparent that the AKE can be simultaneously well-posed and equivalent to a system with a bounded Hamiltonian for spacetimes containing matter. On the other hand, for perturbations of spacetimes that admit an exact Killing vector, one might expect the AKE for the perturbed spacetime to admit solutions that approximate Killing vectors. Thus one of the primary aims of this article is to construct perturbative solutions to the AKE for perturbations of spacetimes which admit Killing vectors and to study their properties and the interpretation of the resulting generalized Komar currents and charges. Additionally, we would like to explore the connection of the perturbed Komar current and charges with the thermodynamic behavior of black hole spacetimes, e.g., the first law. As we will show there is a close correspondence between the AKE and black hole thermodynamics.

The paper is organized as follows: In Section 2 we will review the AKE and shall present a physically interesting scenario, namely that of the Vaidya spacetime, where some of the key aspects of the AKE will be demonstrated. Subsequently, the evolution of the GKV in the perturbed spacetime has been presented in Section 3 from both the action formalism and also from the perturbation of the AKE itself. The stability of the perturbed AKE, as well as its hyperbolicity, have been studied in Section 4, before discussing the nature of the solution of the AKE for both first and second-order perturbations in Section 5. Finally, the thermodynamic interpretation of the AKE has been depicted in Section 6, before presenting the concluding remarks in Section 7. Several computations performed in this work have been presented in detail in Appendix A — Appendix F.

Notations and Conventions: Throughout this paper we have used the mostly positive signature convention, such that the Minkowski metric in the Cartesian coordinates has the following form: \( \eta_{\mu \nu} = \text{diag}(-1, 1, 1, 1) \). The four-dimensional spacetime indices have been denoted by the Greek letters, \( \mu, \nu, \alpha, \ldots \). We work in units such that the fundamental constants have the values \( G = c = \hbar = 1 \). Throughout the article, indices on quantities which appear in arguments will be denoted with superscript and subscript dots, for instance, the arguments in \( A[\chi^\mu] \) and \( \mathcal{O}(\hbar^\alpha) \) represent the quantities \( \chi^\mu \) and \( h_{\mu \nu} \).

2. THE ALMOST KILLING EQUATION: A BRIEF REVIEW

In this section, we will briefly review the almost Killing equation, where the motivation for its construction and its various properties will be discussed in detail. Besides, we will also present the Vaidya geometry as an example of arriving at a solution of the almost Killing equation.

A. Motivation, construction, and properties

A Killing vector field \( \xi^\mu \) is defined as one which satisfies the Killing equation \( \mathcal{L}_\xi g_{\mu \nu} = 2 \nabla_{(\mu} \xi_{\nu)} = 0 \). The divergence of the Killing equation takes the form:

\[
\Box \xi^\alpha + R^\alpha_{\beta \gamma \delta} \, \xi^\beta = 0 . \tag{2.1} \]  

\[
\{\text{DivKillingEquation}\}
\]

As evident, Eq. (2.1) takes the form of a wave equation; on geometries that do not admit Killing vectors, one can nonetheless construct generalizations of the Killing equation by solving Eq. (2.1) for an appropriate set of initial data. The AKE is a generalization of Eq. (2.1), and is given by the following formula:

\[
\Box \chi^\alpha + R^\alpha_{\beta \gamma \delta} \, \chi^\beta + \nabla^\alpha \left[ (1 - \mu) \nabla \cdot \chi \right] = 0 , \tag{2.2} \]  

\[
\{\text{AlmostKillingEquation}\}
\]

where, \( \mu \) is a scalar, which in previous literature is assumed to be a constant—for generality, we do not assume this to be the case here. The solution of Eq. (2.2), i.e., \( \chi^\alpha \) is the GKV. It is straightforward to verify that Killing vectors satisfy the AKE — it is in this sense that solutions of the AKE may be regarded as generalizations of Killing vectors. As discussed in [12], GKV s are not necessarily approximate Killing vectors in the following sense. The vector \( \chi^\alpha \)
satisfies the AKE if the tensor \( Q_{\mu\nu} = \nabla (\mu \chi_{\nu}) \) is transverse and trace-free; however, the components of transverse and trace-free tensors need not be small.\(^1\)

It is instructive to derive any evolution equation from an action principle and the AKE is no different. The AKE, presented in Eq. (2.2), may be derived from the following action (see Appendix A for a derivation of the AKE from this action functional):

\[
A[\chi] = \int_{\mathcal{M}} d^4x \sqrt{-g} \left( -\nabla^{(\alpha} \chi^{\beta)} \nabla_{(\alpha} \chi_{\beta)} + \frac{1}{2} \mu (\nabla \cdot \chi)^2 \right). \tag{2.3} \]

Here, \( \mathcal{M} \) denotes the spacetime volume of interest and as evident, it immediately follows that the action \( A[\chi] \) vanishes if \( \chi^\alpha \) is a Killing vector field. It was shown in [12] that the AKE is strongly hyperbolic when the parameter \( \mu = 1 \), however, it fails to be hyperbolic when \( \mu \neq 2 \), and is weakly hyperbolic for all other (constant) values for \( \mu \). It was also argued that in general, the AKE may suffer from dynamical instabilities; a Hamiltonian analysis reveals the presence of ghosts for \( \mu < 2 \), and unbounded terms for \( \mu > 1/3 \); there is no parameter choice for \( \mu \) in which the AKE avoids these potential instabilities and is also hyperbolic. However, for vacuum spacetimes, at least one exception exists, which we will discuss shortly.

In any spacetime manifold, given some vector field \( V^\alpha \), it follows from differential geometry that it is possible to construct a conserved current and hence a conserved charge. This conserved current takes the following form,

\[
J^\alpha = \nabla_\beta (\nabla^\alpha V^\beta - \nabla^\beta V^\alpha) \tag{2.4}. \]

For Killing vector fields, the conserved current \( J^\alpha \) is known as the Komar current, and the associated charges are known as the Komar charges, and we have so far referred to these respectively as the generalized Komar current and generalized Komar charges. However, if the vector field \( V^\alpha \) is considered as a generator of the diffeomorphism, then \( J^\alpha \) is in fact the conserved current corresponding to the invariance under said diffeomorphism\(^2\). For this reason, it is perhaps more appropriate to call this as the Noether current since it arises out of the diffeomorphism invariance of the gravitational system [15–19]. For the remainder of this article, we shall use this terminology and shall refer to the conserved charges associated with \( J^\alpha \) as Noether charges — as [20–22] demonstrate, Noether charges defined in this manner have interesting thermodynamical interpretations when computed over certain spacelike and null surfaces.

As emphasized before, the Noether current \( J^\alpha \), defined in Eq. (2.4) is identically divergence-free, which when evaluated for solutions of the AKE takes the following form:

\[
J^\alpha = 2R^\alpha_\beta \chi^\beta + \nabla^\alpha [(2 - \mu) \nabla \cdot \chi] \tag{2.5}. \]

We note that when \( \mu = 2 \), the Noether current may be interpreted as a measure of the energy and momentum through the use of the trace-reversed Einstein’s field equations:

\[
R_{\alpha\beta} = 8\pi \left( T_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} T \right) \tag{2.6}. \]

Moreover, the divergence-free property of the Noether current yields the following expression:

\[
\square [(\mu - 2) \nabla \cdot \chi] = \chi^\beta \nabla_\beta R + 2R^\alpha_\beta \nabla_{(\alpha} \chi_{\beta)} \tag{2.7}. \]

where the contracted Bianchi identity, \( \nabla_\alpha R^\alpha_\beta = (1/2) \nabla_\beta R \) has been used. The above evolution equation for \( (\nabla \cdot \chi) \) was used in [12] to show that in a vacuum spacetime, the constraint \( \nabla \cdot \chi = 0 \) is propagated by the AKE; if the initial data satisfy the constraint \( \nabla \cdot \chi = 0 \) and its time derivative, then the time development of solution satisfies the constraint. Under this constraint, the AKE becomes strongly hyperbolic and is independent of \( \mu \), so that it is no longer subject to the instabilities associated with ghosts or unbounded terms in the Hamiltonian.

### B. Example: The Vaidya geometry

Here, we review and generalize the solution for the AKE in the Vaidya spacetime, as presented in [10], to gain some insight into the relationship between GKV, the Smarr relation, and the laws of thermodynamics. The line element associated with Vaidya spacetime takes the following form \((d\Omega^2 \) being the round metric on the 2-sphere):

\[
ds^2 = - \left[ 1 - \frac{2M(v)}{r} \right] dv^2 + 2dvdr + r^2 d\Omega^2, \tag{2.8} \]

\(^1\) Alternately, one can show that even on a Minkowskian background, the AKE admits wavelike solutions for which the components \( Q_{\mu\nu} \gg 0 \), so that they cannot be considered as approximate Killing vectors by any means.

\(^2\) For further discussion of this point, see [13] and [14].
where, $M(v)$ is an arbitrary function of the advanced null coordinate $v$. Following [10], here also we construct a solution to the AKE in the Vaidya spacetime presented in Eq. (2.8) for the $\mu = 2$ case, which has the following form:

$$\chi^\alpha = \left( \frac{M(v)}{M_0}, \frac{r M'(v) + f(v)}{M_0}, 0, 0 \right), \tag{2.9}$$

where $f(v)$ is an arbitrary function of the advanced null coordinate $v$, and $1/M_0$ is a constant factor; a natural choice for this factor is to set it equal to the ADM mass of the spacetime. The fact that the GKV $\chi^\alpha$ depends on an arbitrary function $f(v)$ is due to the fact that for $\mu = 2$, the AKE fails to be hyperbolic. Fortunately, the resulting Noether current and Noether charge are unaffected by the arbitrary function $f(v)$, so that one can regard it as a sort of “gauge” potential — there is, however, a criterion that one may use to fix this arbitrary function $f(v)$, which will be discussed later. The Noether current associated with the GKV $\chi^\alpha$ then takes the following form,

$$J_\chi^\alpha = \left( 0, \frac{4 M(v) M'(v)}{M_0 r^2}, 0, 0 \right). \tag{2.10}$$

In the Vaidya spacetime, the surface characterized by

$$r_H := 2 M(v), \tag{2.11}$$

is a special surface, since the expansion of the outgoing null generators vanishes on this surface and is the apparent horizon. Moreover, it is straightforward to demonstrate that $\chi^\alpha = (1, 0, 0, 0)$ is null on the surface $r = r_H$. Thus expressing the surface element as: $dS_{\alpha\beta} = \varepsilon_{\alpha\beta\mu
u} e^{\mu}_0 e^{\nu}_0 d\theta d\phi$, where $e^{\mu}_0 = \delta^{\mu}_\mu$ and $e^{\nu}_0 = \delta^{\nu}_\nu$ are the basis vector components on the apparent horizon, the mass within the apparent horizon is given by:

$$M_H = \frac{1}{8\pi} \oint_H \nabla^\alpha \chi^\beta dS_{\alpha\beta}. \tag{2.12}$$

The above integrand turns out to be independent of the radius of the surface on which it is being evaluated and thus one obtains the following expression for the mass enclosed by the apparent horizon,

$$M_H = \frac{M(v)^2}{M_0}. \tag{2.13}$$

In the Vaidya spacetime, the event horizon is given by $A_H = 4\pi r_H^2$ (with $r_H$ given by Eq. (2.11)), so that one may rewrite the mass within the apparent horizon, as presented in Eq. (2.13) as,

$$M_H = \frac{A_H}{16\pi M_0} = \frac{\kappa_0 A_H}{4\pi}, \quad \kappa_0 := \frac{1}{4M_0}. \tag{2.14}$$

where, $\kappa_0$ is the surface gravity associated with the surface $r = 2 M_0$, corresponding to the event horizon of the final black hole spacetime. The fact that $\kappa_0$ is a constant here is contradictory with the explicit formula: $\kappa^2 = -\frac{1}{4} (\nabla_\mu \chi^\nu) (\nabla^\mu \chi^\nu)$ for the surface gravity, since it depends on the gauge function $f(v)$. One may choose the gauge function, such that $f(v) = -2M(v)M'(v)$, which corresponds to the requirement that $\chi$ is null on the apparent horizon, in which case, one obtains:

$$\kappa = \sqrt{1 - 16 M'(v)^2} \frac{4M_0}{M_0}. \tag{2.15}$$

Even though it appears that the surface gravity is indeed dependent on the mass function, it is straightforward to verify that if the $v$ dependence is treated as a perturbation, such that $M(v) = M_0 + \epsilon \delta M(v)$, then the $v$ dependent part of $\kappa$ is a second order term in the perturbation,

$$\kappa = \frac{1}{4M_0} - \frac{2\delta M'(v)^2}{4M_0} \epsilon^2 + O(\epsilon^3). \tag{2.16}$$

Therefore, it follows that to first-order, the perturbation of the surface gravity identically vanishes, and the first law $\delta M = \kappa \delta A/8\pi$ holds identically. Even then, at first sight, Eq. (2.15) appears to be puzzling, as it seems to conflict with the expected behavior for the surface gravity, which must satisfy the exact Smarr relation $M_H = \kappa A_H/4\pi$ (neglecting angular momentum). However, upon closer inspection, one notes that since the AKE is linear in $\chi$, the constant factor $1/M_0$ is not specified by the AKE. In an asymptotically flat spacetime, a natural choice for $M_0$ is the ADM mass.
$M_{\text{ADM}}$ and hence the mass enclosed by the apparent horizon may then be written as: $M_H = (M(v)/M_{\text{ADM}}) M(v)$. This suggests that $M_H$ can be interpreted as a rescaling of $M(v)$ by the ratio of $M(v)$ to $M_{\text{ADM}}$.

To better understand this scenario, we assume that $M'(v)$ has a compact support in $v$, such that at late time, $M(v) \rightarrow M_0$ and $f(v) \rightarrow 0$, yielding: $M_0 \rightarrow M_{\text{ADM}}$, in which case one has $\chi^\alpha = (1,0,0,0)$. At early times, again assuming $M'(v) \rightarrow 0$ and $f(v) \rightarrow 0$, one has $\chi^\alpha = (M_{\text{early}}/M_{\text{ADM}},0,0,0)$, where $M_{\text{early}}$ is the mass of the spacetime before $M'(v)$ becomes non-zero and thus $\chi^\alpha$ will differ from $(\partial/\partial t)^\alpha$. Since the early-time geometry of the Vaidya spacetime approximates that of a Schwarzschild black hole, then it is appropriate to rescale $\chi^\alpha$ by a factor $M_{\text{ADM}}/M_{\text{early}}$; in doing so, one obtains an early time horizon mass consistent with the early time “Schwarzschild mass.” This is an indication that the horizon mass constructed from the Komar integral for solutions of the AKE is not identical to the “local” (in $v$) mass of the black hole.

### 3. EVOLUTION OF THE PERTURBED GKVS

In this section, we will consider the perturbation of a background spacetime with Killing symmetry (e.g., Schwarzschild or Kerr), which may or may not be vacuum. Since the perturbation need not respect the symmetry of the background spacetime, the perturbed spacetime does not admit an exact Killing vector field, but the perturbed spacetime will admit GKVs, as long as solutions to Eq. (2.2) exist in the perturbed spacetimes. The perturbed AKE associated with the perturbation of the background spacetime will be derived in two different ways, first from a variational principle, where the perturbation of the action presented in Eq. (2.3) will be considered, and then from the direct perturbation of the AKE itself. We will verify that the results arising out of these two different approaches match.

#### A. Notations and conventions

Before proceeding further, it is perhaps appropriate to settle the notations and the conventions that we will use for this section and the remainder of this article. With the exception of the GKVs, often denoted as $\chi^\mu$, ‘barred’ symbols will be used to denote exact quantities; for instance $\bar{g}_{\alpha\beta}$, $\bar{\nabla}_\alpha$ and $\bar{R}_{\alpha\beta}$ respectively denote the exact metric, connection, and Ricci tensor. Unbarred geometric quantities will be used to denote background quantities. Metric perturbations will be denoted $h_{\alpha\beta}$, and are defined by:

$$h_{\alpha\beta} := \bar{g}_{\alpha\beta} - g_{\alpha\beta} . \quad (3.1)$$

Indices are raised and lowered according to the background metric $g_{\alpha\beta}$. Killing vector fields for the background spacetime will be denoted $\xi^\mu$, and we define $\delta \xi^\mu$ to be the difference between the GKV and the exact Killing vector, in the following manner:

$$\delta \xi^\alpha := \chi^\alpha - \xi^\alpha . \quad (3.2)$$

In general, the prefix $\delta$ will denote differences between the exact and background quantities (which we will later assume to be small compared to background values), and the prefix $\Delta$ will denote first-order variations.

#### B. Perturbation of the action principle yielding almost Killing equation

As emphasized before, we assume that the background metric $g_{\mu\nu}$ admits a Killing vector field $\xi^\mu$. Since $h_{\mu\nu}$ is a perturbation over and above the background spacetime, it is legitimate to assume that $h_{\mu\nu} \ll g_{\mu\nu}$, and for some normalization of the background Killing vector field $\xi^\mu$, we also assume $\delta \xi^\mu \ll \xi^\mu$. The action presented in Eq. (2.3) may then be expanded in $h_{\mu\nu}$, $\delta \xi^\mu$ and $\delta \mu := \bar{\mu} - \mu$, keeping terms up to quadratic order in each — we treat the expansions in $h_{\mu\nu}$ and $\delta \xi^\mu$ independently, so that we keep terms of the form $h \cdot h \cdot \delta \xi \delta \xi$. For simplicity, we assume the quantity $\delta \mu$ to be independent of the spacetime coordinates. It should be emphasized that $\delta \xi^\mu \neq \bar{g}_{\mu\nu}(\chi^\nu - \xi^\nu)$.

---

3 Technically speaking, this is achieved by introducing two parameters, $\epsilon_1$ and $\epsilon_2$, with $\epsilon_{1,2} \ll 1$ and then expanding the metric as, $\bar{g}_{\mu\nu} = g_{\mu\nu} + \epsilon_1 h_{\mu\nu}$ and the GKV as, $\chi^\alpha = \xi^\alpha + \epsilon_2 \delta \xi^\alpha$. Then keeping terms linear in $\epsilon_1$ and $\epsilon_2$ will provide first order perturbation, while terms $O(\epsilon_{1,2}^2)$ yield the second order perturbations.
since indices are raised and lowered with the background metric. We begin by writing down the Lagrangian from the action functional presented in Eq. (2.3), which has the following explicit form in the spacetime with metric $\bar{g}_{\mu\nu}$:

$$L = \frac{1}{4} \left( \bar{g}_{\alpha\beta} \nabla_{\mu} \chi^\alpha + \bar{g}_{\alpha\mu} \nabla_{\nu} \chi^\alpha \right) \left( \bar{g}^{\mu\beta} \nabla_{\beta} \chi^\nu + \bar{g}^{\nu\beta} \nabla_{\beta} \chi^\mu \right) + \frac{\bar{\mu}}{2} \left( \nabla_{\mu} \chi^\mu \right)^2.$$  \hspace{1cm}  \text{(3.3)} \hspace{1cm} \{Lagrangian Form\}

We wish to express the Lagrangian presented above solely in terms of the background metric $g_{\mu\nu}$, the perturbation $h_{\mu\nu}$ and the perturbation $\delta \xi^\mu$ along with $\delta \mu$. Since the Lagrangian depicted in Eq. (3.3) consists of terms quadratic in $\chi^\mu$ and its derivatives, it can be rewritten in the following form:

$$L = \chi^\alpha \chi^\beta L_{\alpha\beta} + \chi^\mu \nabla^\alpha \chi^\beta L_{\mu\alpha\beta} + \nabla^\nu \chi^\mu \nabla^\beta \chi^\alpha L_{\mu\nu\alpha\beta},$$  \hspace{1cm}  \text{(3.4)} \hspace{1cm} \{Lagrangian Form\}

where the tensors $L_{\alpha\beta}$, $L_{\mu\alpha\beta}$, and $L_{\mu\nu\alpha\beta}$ depend on the background metric $g_{\mu\nu}$, the perturbation $h_{\mu\nu}$, the derivatives of $h_{\mu\nu}$, and $\bar{\mu}$. Note that in writing Eq. (3.4), we make no assumption about $h_{\mu\nu}$; Eq. (3.4) should hold to all orders in the metric perturbation $h_{\mu\nu}$. Using the above decomposition of the Lagrangian, the action can similarly be written down in the following form:

$$A[\delta \xi] = \int d^4 x \sqrt{-g} \left[ \chi^\alpha \chi^\beta M_{\alpha\beta} + \chi^\mu \nabla^\alpha \chi^\beta M_{\mu\alpha\beta} + \nabla^\nu \chi^\mu \nabla^\beta \chi^\alpha M_{\mu\nu\alpha\beta} \right],$$  \hspace{1cm}  \text{(3.5)} \hspace{1cm} \{Action Form\}

where we define:

$$M_{\alpha\beta} := \sqrt{\bar{g}/g} L_{\alpha\beta},$$  \hspace{1cm}  \text{(3.6)} \hspace{1cm} \{Action_Form\}

One may at this point perform the variation of the action with respect to $\delta \xi^\alpha$, the perturbed GKV, without explicit knowledge of the tensors $M_{\alpha\beta}$. Up to boundary terms, the first-order variation of the action takes the following form:

$$\Delta A = \int d^4 x \sqrt{-g} \Delta \delta \xi^\alpha E_\alpha,$$  \hspace{1cm}  \text{(3.7)} \hspace{1cm} \{PAKE-Action\}

where:

$$E_\alpha := \frac{1}{2} \left\{ 2 \nabla^\beta \chi^\nu \left[ \frac{1}{2} \left( L_{\beta\nu\gamma\alpha} + L_{\gamma\alpha\beta\nu} \right) \left( h^{\sigma\tau} \nabla^\gamma h_{\sigma\tau} - \nabla^\gamma h \right) - \left( \nabla^\nu L_{\beta\nu\gamma\alpha} + \nabla^\gamma L_{\gamma\alpha\beta\nu} \right) + L_{\alpha\beta\nu} - L_{\nu\beta\alpha} \right] \right. $$

$$+ 2 \chi^\nu \left[ \frac{1}{2} L_{\nu\beta\alpha} \left( h^{\sigma\tau} \nabla^\gamma h_{\sigma\tau} - \nabla^\gamma h \right) - \nabla^\beta L_{\nu\beta\alpha} + L_{\alpha\nu} + L_{\nu\alpha} \right] - 2 \nabla^\nu \nabla^\beta \chi^\nu \left( L_{\beta\nu\gamma\alpha} + L_{\gamma\alpha\beta\nu} \right) \right\}.$$  \hspace{1cm}  \text{(3.8)} \hspace{1cm} \{PAKE-Action\}

Keeping in mind $\chi^\alpha = \xi^\alpha + \delta \xi^\alpha$, the perturbed AKE may then be written as:

$$\frac{\bar{g}^{\mu\alpha} E_\alpha}{\sqrt{\bar{g}/g}} = 0.$$  \hspace{1cm}  \text{(3.9)} \hspace{1cm} \{PAKE-Action\}

The factor of $1/\sqrt{\bar{g}/g}$ is included because one typically factors out the volume element from the functional derivative when deriving field equations (as was done when going from Eq. (A.4) to Eq. (A.5) in Appendix A). Again, we emphasize that the analysis presented here does not require that $h_{\mu\nu}$ is small; the result in Eq. (3.9) holds to all orders in $h_{\mu\nu}$. Thus one may expand the tensors $L_{\alpha\beta}$ in various powers of the gravitational perturbation in the following manner:

$$L_{\alpha\beta} = L_0^{\alpha\beta} + L_1^{\alpha\beta} + L_2^{\alpha\beta} + O(h^3)$$

$$L_{\mu\alpha\beta} = L_0^{\mu\alpha\beta} + L_1^{\mu\alpha\beta} + L_2^{\mu\alpha\beta} + O(h^3)$$

$$L_{\mu\nu\alpha\beta} = L_0^{\mu\nu\alpha\beta} + L_1^{\mu\nu\alpha\beta} + L_2^{\mu\nu\alpha\beta} + O(h^3).$$  \hspace{1cm}  \text{(3.10)} \hspace{1cm} \{LagPerturbedTensorsDecomp\}

It turns out that to zeroth order, one has $L_0^{\alpha\beta} = 0$, $L_0^{\mu\alpha\beta} = 0$, while,

$$L_0^{\mu\nu\alpha\beta} = \frac{1}{2} \left( \bar{\mu} \ g_{\alpha\beta} g_{\mu\nu} - g_{\alpha\nu} g_{\beta\mu} - g_{\alpha\mu} g_{\beta\nu} \right).$$  \hspace{1cm}  \text{(3.11)} \hspace{1cm} \{PAKE-Action\}

To first order in $h_{\mu\nu}$, the tensors $L_1^{\alpha\beta}$, $L_1^{\mu\alpha\beta}$ and $L_1^{\mu\nu\alpha\beta}$ can be expressed as linear functions of the gravitational perturbation $h_{\mu\nu}$ as,

$$L_1^{\alpha\beta} = 0; \quad L_1^{\mu\nu\alpha\beta} = \frac{1}{2} \left\{ \bar{\mu} g_{\alpha\beta} \nabla_{\mu} h - 2 \nabla_{\mu} h_{\alpha\beta} \right\}; \quad L_1^{\mu\nu\alpha\beta} = \frac{1}{2} \left\{ h_{\alpha\nu} g_{\beta\mu} - g_{\alpha\mu} h_{\beta\nu} \right\}. $$  \hspace{1cm}  \text{(3.12)} \hspace{1cm} \{PAKE-Action\}
Finally, we present the second-order terms in the perturbation \( h_{\mu \nu} \) as,

\[
L_{\alpha \beta}^2 = \frac{1}{8} \left\{ \bar{\mu} \nabla_{\alpha} h \nabla_{\beta} h - 2 \nabla_{\alpha} h^{\rho \tau} \nabla_{\beta} h_{\rho \tau} \right\}
\]

\[
L_{\mu \alpha \beta}^2 = \frac{1}{2} \left\{ 2h_{\alpha \beta} \nabla_{\mu} h_{\sigma \tau} - \bar{\mu} g_{\alpha \beta} h^{\rho \tau} \nabla_{\mu} h_{\rho \tau} \right\}
\]

\[
L_{\mu \nu \alpha \beta}^2 = \frac{1}{2} \left\{ h_{\alpha \mu} h_{\beta \nu} - g_{\beta \nu} h_{\alpha \mu} \right\}.
\]

(3.13) \{PAKE-ActionPerturbedTensors2\}

In what follows, we consider in detail the expansion of the Lagrangian to first order in the metric perturbation \( h_{\mu \nu} \) and as we shall demonstrate the resulting perturbed AKE is consistent with the expansion coefficients determined above.

C. Explicit perturbation of the action to first order in the metric

The expression for the AKE given in Eq. (3.8) is rather complicated and somewhat opaque; it is perhaps more illustrative to show explicitly that the derivation of the perturbed AKE to first order in the metric perturbations \( h_{\mu \nu} \), simplifying the expressions along the way. To obtain the expansion of the Lagrangian to first order in \( h_{\mu \nu} \), one may either use Eq. (3.11) and Eq. (3.12) or the identities presented in Appendix B to obtain the following Lagrangian:

\[
\bar{L} = -\frac{1}{4} \left[ (g_{\mu \alpha} + h_{\mu \alpha}) \nabla_{\mu} \xi^{\alpha} + (g_{\nu \alpha} + h_{\nu \alpha}) \nabla_{\nu} \xi^{\alpha} + \chi^{\mu} \nabla_{\mu} h_{\nu \alpha} \right]
\times \left[ (g^{\beta \nu} - h^{\beta \nu}) \nabla_{\beta} \chi^{\nu} + (g^{\nu \beta} - h^{\nu \beta}) \nabla_{\nu} \chi^{\beta} + \chi^{\sigma} \nabla_{\sigma} h_{\mu \nu} \right]
\times \left[ \left( \nabla_{\alpha} \delta \xi^{\nu} - h^{\rho \beta} \nabla_{\rho} \xi^{\nu} - h^{\rho \beta} \nabla_{\rho} \delta \xi^{\nu} + \frac{1}{2} (\xi^{\nu} + \delta \xi^{\nu}) \nabla_{\nu} h_{\mu \nu} \right) \right]
\times \left[ \frac{\bar{\mu}}{2} \left( \nabla_{\mu} \delta \xi^{\nu} \right)^2 + (\xi^{\nu} + \delta \xi^{\nu}) \nabla_{\nu} h_{\mu \nu} \right] + O(h_0^2). \]

(3.14) \{lag_density_02\}

Having expressed the Lagrangian explicitly in terms of the background metric \( g_{\mu \nu} \) and the perturbation \( h_{\alpha \beta} \), we now expand the GKV field, in terms of the background Killing field \( \xi^{\mu} \) and the perturbation \( \delta \xi^{\mu} \). One can see that Eq. (3.14) has the form of Eq. (3.4), and it is not difficult to verify that the Lagrangians are equivalent for the coefficients given in Eq. (3.10), Eq. (3.11) and Eq. (3.12). Using the Killing equation for \( \xi^{\mu} \), i.e., setting \( \nabla_{\nu} \xi^{\nu} + \nabla_{\nu} \xi^{\mu} = 0 \), the Lagrangian presented in Eq. (3.14) can be further simplified. In particular, it is worth emphasizing that the on-shell value of the action for the background Killing vector field identically vanishes and thus the Lagrangian density given in Eq. (3.14) becomes,

\[
\bar{L} = -\frac{1}{2} \left[ g_{\mu \alpha} \nabla_{\mu} \delta \xi^{\alpha} + h_{\mu \alpha} \nabla_{\mu} \xi^{\alpha} + g_{\mu \alpha} \nabla_{\mu} \delta \xi^{\alpha} + h_{\mu \alpha} \nabla_{\nu} \xi^{\alpha} + h_{\mu \alpha} \nabla_{\nu} \delta \xi^{\alpha} + (\xi^{\nu} + \delta \xi^{\nu}) \nabla_{\nu} h_{\mu \nu} \right]
\times \left[ \left( \nabla_{\alpha} \delta \xi^{\nu} - h^{\rho \beta} \nabla_{\rho} \xi^{\nu} - h^{\rho \beta} \nabla_{\rho} \delta \xi^{\nu} + \frac{1}{2} (\xi^{\nu} + \delta \xi^{\nu}) \nabla_{\nu} h_{\mu \nu} \right) \right]
\times \left[ \frac{\bar{\mu}}{2} \left( \nabla_{\mu} \delta \xi^{\nu} \right)^2 + (\xi^{\nu} + \delta \xi^{\nu}) \nabla_{\nu} h_{\mu \nu} \right] + O(h_0^2). \]

(3.15) \{lag_act_simp1\}

Even though the above Lagrangian density looks sufficiently complicated, we can reduce it to a very simple form by dividing the above into three categories — (a) terms quadratic in the derivatives of \( \delta \xi^{\alpha} \), (b) terms linear in the derivatives of \( \delta \xi^{\alpha} \) and (c) terms independent of derivatives of \( \delta \xi^{\alpha} \). The terms quadratic in the derivative of \( \delta \xi^{\alpha} \) yields (see Appendix B),

\[
\text{Quadratic Terms} = \frac{1}{2} \left[ (\mu + \delta \mu) \delta \xi^{\alpha} \delta \xi^{\beta} - (\delta \alpha + h_{\alpha \beta}) (g^{\mu \nu} - h^{\mu \nu} - (\xi^{\nu} + \delta \xi^{\nu}) \left( \nabla_{\nu} \delta \xi^{\alpha} \nabla_{\nu} \delta \xi^{\beta} \right) \right], \]

(3.16) \{quad_term\}

while the terms linear in the derivative of \( \delta \xi^{\alpha} \) become (see Appendix B),

\[
\text{Linear Terms} = \frac{\bar{\mu}}{2} \left[ (\xi^{\alpha} + \delta \xi^{\alpha}) \nabla_{\alpha} h (\nabla_{\nu} \delta \xi^{\mu}) \right] - \frac{1}{2} \left[ g_{\mu \alpha} \nabla_{\mu} \xi^{\alpha} + g_{\nu \alpha} \nabla_{\nu} \delta \xi^{\alpha} \right] \times \left[ - h^{\rho \beta} \nabla_{\rho} \delta \xi^{\nu} + \frac{1}{2} (\xi^{\nu} + \delta \xi^{\nu}) \nabla_{\nu} h_{\mu \nu} \right]
- \frac{1}{2} \nabla_{\alpha} \delta \xi^{\nu} \left[ h_{\mu \alpha} \nabla_{\mu} \xi^{\alpha} + h_{\nu \alpha} \nabla_{\nu} \delta \xi^{\alpha} + (\xi^{\nu} + \delta \xi^{\nu}) \nabla_{\rho} h_{\mu \nu} \right]. \]

(3.17) \{quad_term\}

Finally, as indicated in the expression for \( L_{\alpha \beta} \) in Eq. (3.12) the terms involving no derivatives of \( \delta \xi^{\alpha} \) is \( O(h_0^2) \) and hence will not contribute in our subsequent discussion regarding the determination of the action functional of the perturbed AKE. Thus we have computed the Lagrangian of the GKV field involving linear order terms in the perturbation \( h_{\mu \nu} \) and up to quadratic order terms in the perturbed GKV \( \delta \xi^{\mu} \). However, computation of the action functional requires
the complete action for the perturbed Killing vector field $\delta \xi^\mu$ takes the following form,

$$
\mathcal{A}[\delta \xi] = \int_V d^4x \sqrt{-\bar{g}} \left\{ \left[ \frac{1}{2} \left( \mu \delta_{\alpha}^{\mu} \delta_{\beta}^{\nu} - g_{\alpha\beta} g^{\mu\nu} - \delta_{\alpha}^{\mu} \delta_{\beta}^{\nu} \right) + \frac{1}{4} h \left( \mu \delta_{\alpha}^{\mu} \delta_{\beta}^{\nu} - g_{\alpha\beta} g^{\mu\nu} - \delta_{\alpha}^{\mu} \delta_{\beta}^{\nu} \right) \right] + \frac{1}{2} \left( 1 + \frac{1}{2} h \right) \delta_{\mu}^{\alpha} \delta_{\beta}^{\nu} + \frac{1}{2} \left( g_{\alpha\beta} h^{\mu\nu} - h_{\alpha\beta} g^{\mu\nu} + \delta_{\alpha}^{\mu} h_{\beta}^{\nu} - h_{\alpha\beta} \delta_{\mu}^{\nu} \right) \right\} \left( \nabla_{\mu} \delta \xi^\alpha \nabla_{\nu} \delta \xi^\beta \right) + \left[ \frac{\mu}{2} \left( (\xi^\alpha + \delta \xi^\alpha) \nabla_{\alpha} h (\nabla_{\mu} \delta \xi^\mu) \right) - \left( \nabla_{\mu} \delta \xi^\alpha \right) h_{\alpha\beta} \nabla^\beta \xi^\beta + \left( \nabla_{\mu} \delta \xi^\alpha \right) h^{\mu\beta} \nabla_{\alpha} \xi^\beta \right] - g_{\nu\alpha} \nabla_{\mu} \delta \xi^\alpha (\xi^\sigma + \delta \xi^\sigma) \nabla_{\sigma} h^{\mu\nu} \right \}.
$$

Having derived the action to linear order in the gravitational perturbation $h_{\mu\nu}$, we can determine an arbitrary variation of the action for variation of the perturbed Killing vector field $\delta \xi^\mu$, which when set to zero should yield the corresponding perturbed AKE. The above variation of the action has been carried out in Appendix B and the final expression for the variation, ignoring any boundary contribution, takes the following form,

$$
\Delta \mathcal{A} = \int_V d^4x \sqrt{-\bar{g}} \delta (\nabla_{\nu} \delta \xi^\nu) \left\{ (1 - \mu) \nabla^{\beta} (\nabla_{\nu} \delta \xi^\nu) + \Box \delta \xi^\beta + R^\beta_{\rho\sigma\mu} \delta \xi^\rho \right\} - \delta \mu \nabla^{\beta} \nabla_{\sigma} \delta \xi^\sigma - h^{\mu\nu} \nabla_{\mu} \nabla_{\nu} \delta \xi^\beta + 2 \left( \nabla^{(\mu} \delta \xi^{\nu)} \right) \left( \nabla_{\mu} h^{\beta}_{\nu} - \frac{1}{2} \nabla^{\beta} h_{\mu\nu} \right) - R^\beta_{\rho\sigma\mu} h^{\mu\rho} \delta \xi^\sigma + \left( \frac{1 - \mu - \delta \mu}{2} \nabla^{\beta} \left[ (\xi^\rho + \delta \xi^\rho) \nabla_{\rho} h \right] - h^{\beta\mu} (1 - \mu - \delta \mu) \nabla_{\mu} \left( \nabla_{\nu} \delta \xi^\nu \right) \right\},
$$

where we have neglected all the terms quadratic in the gravitational perturbation $h_{\mu\nu}$. Setting the variation $\Delta \mathcal{A}$ to zero, for arbitrary variation of the perturbation of the Killing vector field $\delta \xi^\mu$, we obtain the following dynamical equation for the perturbed GK field $\delta \xi^\mu$,

$$
(1 - \mu) \nabla^{\beta} (\nabla_{\nu} \delta \xi^\nu) + \Box \delta \xi^\beta + R^\beta_{\rho\sigma\mu} \delta \xi^\rho = J^\beta,
$$

$$
J^\beta = \delta \mu \nabla^{\beta} \nabla_{\sigma} \delta \xi^\sigma + h^{\mu\nu} \nabla_{\mu} \nabla_{\nu} \delta \xi^\beta - 2 \left( \nabla^{(\mu} \delta \xi^{\nu)} \right) \left( \nabla_{\mu} h^{\beta}_{\nu} - \frac{1}{2} \nabla^{\beta} h_{\mu\nu} \right) + R^\beta_{\rho\sigma\mu} h^{\mu\rho} \delta \xi^\sigma - \left( \frac{1 - \mu - \delta \mu}{2} \nabla^{\beta} \left[ (\xi^\rho + \delta \xi^\rho) \nabla_{\rho} h \right] + h^{\beta\mu} (1 - \mu - \delta \mu) \nabla_{\mu} \left( \nabla_{\nu} \delta \xi^\nu \right) \right.
$$

The above provides the dynamical equation for the perturbed Killing vector field $\delta \xi^\alpha$, arising from the variation of the action. One can verify that this expression is equivalent to that obtained from Eq. (3.8); we have done this using the xAct package for Mathematica. In the subsequent discussion, we will discuss explicitly the derivation of this equation from the perturbation of the AKE itself to first order in the metric perturbations. This will depict the internal consistency of the results derived in this work.

### D. Perturbation of the almost Killing equation to first order in metric

We have derived the evolution equation for the perturbed GKV field to first order in the metric perturbations, starting from the variation of the perturbed action for the GKV field. As we will show in this section, the same equation can also be derived from direct perturbation of the AKE itself. As before, we assume that the background metric $g_{\mu\nu}$ admits a Killing vector field $\xi^\mu$ and also that $h_{\mu\nu} \ll g_{\mu\nu}$ and $\delta \xi^\mu \ll \xi^\mu$ (see also footnote 2). Thus we will expand the AKE, given in Eq. (2.2), in the perturbed spacetime with metric $g_{\mu\nu}$, to first order in $\delta \mu$ and $h_{\mu\nu}$ each, again assuming that the expansions are independent (so that we keep terms of the form $h_{..} \delta \xi^\mu$). It is convenient to first present the expansion of the following geometric quantities to linear order in the gravitational perturbation $h_{\mu\nu}$ (for the derivation of these identities, see Appendix C):

$$
\delta R_{\mu\nu} = \frac{1}{2} \left( -\Box h_{\mu\nu} - \nabla_{\mu} \nabla_{\nu} h + \nabla_{\mu} \nabla_{\alpha} h^{\alpha}_{\nu} + \nabla_{\nu} \nabla_{\alpha} h^{\alpha}_{\mu} + R_{\beta\mu} h^{\beta}_{\nu} + R_{\beta\nu} h^{\beta}_{\mu} - 2R_{\alpha\mu\beta\nu} h^{\alpha\beta} \right).
$$
\[ \nabla_\alpha \nabla_\beta V^\beta = \delta_\mu^\alpha \nabla_\alpha \nabla_\beta V^\mu = \nabla_\alpha \left( \nabla_\beta V^\beta \right) + \nabla_\alpha \left( \delta \Gamma_\beta^\rho_{\beta \rho} V^\rho \right) + \delta \Gamma_\alpha^\beta \nabla_\beta V^\rho - \delta \Gamma_\alpha^\rho \nabla_\rho V^\beta \]
\[ = \nabla_\alpha \left( \nabla_\beta V^\beta \right) + \nabla_\alpha \left( \delta \Gamma_\beta^\rho_{\beta \rho} V^\rho \right). \] (3.23) \[ \{ \text{identity} \}
\]

\[ \Box V^\mu = \left( g^{\alpha \beta} - h^{\alpha \beta} \right) \nabla_\alpha \nabla_\beta V^\mu + \left( - \frac{1}{2} \nabla^\mu h_{\alpha \rho} + \nabla_\alpha h_{\rho}^\mu \right) \left( \nabla_\alpha V^\rho + \nabla_\beta V^\alpha \right) - \nabla_\rho V^\mu \left( \nabla_\alpha h_{\alpha \rho} - \frac{1}{2} \nabla^\rho h \right) \]
\[ + \frac{1}{2} V^\rho \left( \nabla_\rho \nabla_\beta h_{\alpha \rho} - \nabla_\mu \nabla_\beta h_{\alpha \rho} + R_{\alpha \rho \beta \lambda} h_{\rho \lambda} - R_{\alpha \rho \mu} h_{\rho \beta} \right). \] (3.24) \[ \{ \text{identity} \}
\]

where \( V^\mu \) is an arbitrary vector field. In deriving the above identities we have used various properties of the Riemann tensor, e.g., \( R_{\alpha \beta \mu \nu} = R_{\nu \mu \beta \alpha} \) among others. Applying all these identities to the AKE in the perturbed spacetime and imposing the Lorenz gauge condition: \( \nabla_\alpha h_{\alpha \rho} = (1/2) \nabla_\rho h \), we obtain,

\[ \left( g^{\alpha \beta} - h^{\alpha \beta} \right) \nabla_\alpha \nabla_\beta \delta \xi^\mu + \left( - \frac{1}{2} \nabla^\mu h_{\alpha \rho} + \nabla_\alpha h_{\rho}^\mu \right) \left( \nabla_\alpha \delta \xi^\rho + \nabla_\beta \delta \xi^\alpha \right) + R^\mu_{\alpha \beta} \delta \xi^\beta \]
\[ + \left( 1 - \mu \right) \delta \xi^\alpha \nabla_\alpha \left( \nabla_\beta \delta \xi^\rho \right) + \frac{1}{2} \nabla_\alpha \left( \delta \xi^\rho \nabla_\rho h \right) \] \[ - R_{\alpha \beta \rho \sigma} h_{\alpha \rho} \delta \xi^\beta \]
\[ + \left( 1 - \mu \right) \delta \xi^\alpha \nabla_\alpha \left( \nabla_\beta \delta \xi^\rho \right) + \frac{1}{2} \left( \delta \xi^\rho \nabla_\rho h + \delta \xi^\rho \nabla_\rho h \right) = 0 , \] (3.25)

which is valid up to linear order in the gravitational perturbation \( h_{\alpha \rho} \).

At this point, we have not yet expanded in the GKV field \( \chi^\mu \); we do this now. We make use of the wave equation for the background Killing vector field \( \xi^\mu \), given in Eq. (2.1) and other properties of Killing vectors to obtain,

\[ \left( g^{\alpha \beta} - h^{\alpha \beta} \right) \nabla_\alpha \nabla_\beta \delta \xi^\mu + \left( - \frac{1}{2} \nabla^\mu h_{\alpha \rho} + \nabla_\alpha h_{\rho}^\mu \right) \left( \nabla_\alpha \delta \xi^\rho + \nabla_\beta \delta \xi^\alpha \right) + R^\mu_{\alpha \beta} \delta \xi^\beta \]
\[ + \left( 1 - \mu \right) \delta \xi^\alpha \nabla_\alpha \left( \nabla_\beta \delta \xi^\rho \right) + \frac{1}{2} \nabla_\alpha \left( \delta \xi^\rho \nabla_\rho h \right) \] \[ - R_{\alpha \beta \rho \sigma} h_{\alpha \rho} \delta \xi^\beta \]
\[ + \left( 1 - \mu \right) \delta \xi^\alpha \nabla_\alpha \left( \nabla_\beta \delta \xi^\rho \right) + \frac{1}{2} \left( \delta \xi^\rho \nabla_\rho h + \delta \xi^\rho \nabla_\rho h \right) = 0 . \] (3.26)

This is our result for the perturbed AKE. The above equation has been derived under very general conditions, without any assumptions about the nature of the perturbation. Thus it is possible to express the above equation in several different ways, under different assumptions, which we will list below. First, we rewrite the above evolution equation for the perturbed GKV field \( \delta \xi^\mu \) in the following form,

\[ \Box \delta \xi^\mu + R^\mu_{\rho \sigma} \delta \xi^\rho + g^{\mu \alpha} \nabla_\alpha \left\{ (1 - \mu) \left( \nabla_\sigma \delta \xi^\sigma \right) \right\} = j^\mu ; \] (3.27) \[ \{ \text{pake_form} \}
\]

\[ j^\mu = h^{\alpha \beta} \left( \nabla_\alpha \nabla_\beta \delta \xi^\mu + R^\mu_{\alpha \rho \beta} \delta \xi^\rho + \delta \xi^\rho \nabla_\beta \left\{ (1 - \mu - \delta \mu) \left( \nabla_\sigma \delta \xi^\sigma \right) \right\} \right) + g^{\mu \alpha} \nabla_\alpha \left( \delta \mu \nabla_\sigma \delta \xi^\sigma \right) \]
\[ - 2 \left( \nabla_\alpha h_{\rho}^\mu - \frac{1}{2} \nabla^\mu h_{\alpha \rho} \right) \nabla^{(\rho \dot{\sigma} \xi^\rho) - \frac{1}{2} (1 - \mu - \delta \mu) g^{\mu \alpha} \nabla_\alpha \left( \delta \xi^\rho \nabla_\rho h \right) \]
\[ - \frac{1}{2} \nabla_\alpha \left( 1 - \mu - \delta \mu \right) g^{\mu \alpha} \left( \delta \xi^\rho \nabla_\rho h \right) \]. (3.28)

Upon comparison, we find that this evolution equation for the perturbed GKV \( \delta \xi^\mu \) is identical to what we had derived from the action, i.e., to Eq. (3.20), except for the terms involving derivatives of \( \mu \) and \( \delta \mu \), respectively. This is because, while deriving Eq. (3.20), we have assumed for convenience that \( \mu \) and \( \delta \mu \) are constants, while that is not the case for the derivation presented above. If we assume that \( \mu \) for the background spacetime is constant, and \( \delta \mu \) to be a scalar function then the dynamics of the perturbed GKV is determined by,

\[ \left( g^{\alpha \beta} - h^{\alpha \beta} \right) \left( \nabla_\alpha \nabla_\beta \delta \xi^\mu + R^\mu_{\alpha \rho \beta} \delta \xi^\rho \right) = j^\mu ; \] (3.29) \[ \{ \text{pake_form} \}
\]

\[ j^\mu = - (1 - \mu - \delta \mu) \left( g^{\mu \alpha} - h^{\mu \alpha} \right) \nabla_\alpha \left( \nabla_\sigma \delta \xi^\sigma \right) + \frac{1}{2} g^{\mu \alpha} \nabla_\alpha \left( \delta \xi^\rho \nabla_\rho h \right) \]
\[ - 2 \left( \nabla_\alpha h_{\rho}^\mu - \frac{1}{2} \nabla^\mu h_{\alpha \rho} \right) \nabla^{(\alpha \delta \xi^\rho\beta)} \].
\[ + (\nabla_\alpha \delta \mu) \left\{ g^{\mu \sigma} \nabla_\alpha \left( \nabla_\sigma \delta \xi^\rho \right) + \frac{1}{2} g^{\mu \alpha} \left[ (\xi^\mu + \delta \xi^\mu) \nabla_\rho h \right] - \left( \nabla_\sigma \delta \xi^\sigma \right) h^{\mu \alpha} \right\} \]. \quad (3.30)

As mentioned before, it will be useful if we write down simplified versions of Eq. (3.29) for various scenarios of physical interest. These can range from the use of the transverse traceless gauge to setting \( \delta \mu = \text{constant} \). We discuss below each of these limits explicitly.

- If we choose \( \delta \mu = \text{constant} \), then the last term in the expression for \( j^\mu \) in Eq. (3.29) identically vanishes. In this context, the dynamical equation for the perturbed GKV becomes identical to that derived from the perturbation of the action, i.e., to Eq. (3.20).

- If we assume that the background spacetime is vacuum, and the perturbation involves no incoming matter fields, then use of the transverse traceless gauge (equivalently setting \( h = 0 \) in Eq. (3.29)) yields,

\[
\left( g^{\alpha \beta} - h^{\alpha \beta} \right) \left[ \nabla_\alpha \nabla_\beta \delta \xi^\mu + R^\mu_{\alpha \rho \beta} \delta \xi^\rho \right] = j^\mu , \tag{3.31}
\]

\[
j^\mu = - (1 - \mu - \delta \mu) \left[ (g^{\mu \alpha} - h^{\mu \alpha}) \nabla_\alpha \left( \nabla_\sigma \delta \xi^\sigma \right) \right] - 2 \left( \nabla_\alpha h^\mu_\rho - \frac{1}{2} \nabla^\mu h_{\alpha \rho} \right) \nabla^{(\alpha} \delta \xi^{\rho)} \]

\[ + (\nabla_\alpha \delta \mu) \left\{ g^{\mu \alpha} \nabla_\alpha \left( \nabla_\sigma \delta \xi^\sigma \right) - \left( \nabla_\sigma \delta \xi^\sigma \right) h^{\mu \alpha} \right\} \]. \quad (3.32)

Note that any term involving \( R_{\alpha \beta} \) will not contribute, since for vacuum spacetime the Ricci tensor identically vanishes. Also, if we assume \( \delta \mu \) to be constant, the above equation can be simplified even more, as the last term in \( j^\mu \) will be absent.

- If we use the fact that the perturbed GKV is really a consequence of perturbation of the spacetime geometry (we will examine this case in detail later), then we will have \( \chi^\mu = \xi^\mu + \delta \chi^\mu_1 + \delta \chi^\mu_2 \), where \( \delta \chi^\mu_1 \) is linear in the gravitational perturbation, while \( \delta \chi^\mu_2 \) is quadratic in the gravitational perturbation. An identical decomposition will work for the \( \mu \) as well. If we keep terms linear in the gravitational perturbation, we should also ignore terms \( O(\delta \chi^\mu \delta h_{\alpha \beta}) \) and so on. It follows that perturbed AKE governing the evolution of the vector \( \delta \chi^\mu_1 \), takes the relatively simple form,

\[
\Box \delta \chi^\mu_1 + R^\mu_\rho \delta \chi^\rho_1 + (1 - \mu) \left[ \nabla^\mu \left( \nabla_\sigma \delta \chi^\sigma_1 \right) + \frac{1}{2} \nabla^\mu (\xi^\rho \nabla_\rho h) \right] = 0 . \tag{3.33}
\]

Note also that if we assume the background spacetime to be vacuum with no incoming matter perturbation, then use of the transverse traceless gauge would reduce Eq. (3.33) to AKE for the background spacetime \( g_{\mu \nu} \).

The consequences of this equation with or without matter field will be discussed in a subsequent section.

Thus we have derived the evolution equation for the perturbed GKV field from the perturbed action and also from the perturbed AKE to first order in the metric perturbations. Both of these procedures yield identical equations, depicting the internal consistency of our analysis. We have also verified using computer algebra (in particular the xAct package for Mathematica) that this consistency holds to second-order as well. The AKE is rather complicated in the second-order case, so we do not present the result here — the interested reader can view the Mathematica file posted at [23]. In what follows, we will discuss the structure of the Hamiltonian associated with the dynamical equation for the perturbed AKE, leading to an understanding of the stability as well as its hyperbolicity.

4. HAMILTONIAN STABILITY AND HYPERBOLICITY

In this section, we will construct the Hamiltonian out of the Lagrangian, whose variation yields the evolution equation for the perturbed GKV. The stability of the Hamiltonian and its bounded nature will also be examined. Besides the hyperbolicity of the perturbed AKE will also be explored.

A. Hamiltonian for the perturbed AKE

We have derived the evolution equation for the perturbed GKV in the preceding section in two different ways, first by varying the perturbed action from which the AKE can be derived, and then by direct perturbation of the AKE. In this section, we will discuss the stability and the hyperbolicity of the perturbed AKE, restricting (for simplicity) to first order in the metric perturbations \( h_{\mu \nu} \). First, we will construct the Hamiltonian for the perturbed GKV field.
and discuss its stability. Subsequently, we will discuss the hyperbolicity of the perturbed AKE and its consequences. The starting point for the Hamiltonian analysis is the action for the perturbed GKV field $\delta \xi^\mu$; in particular, the zeroth component of the boundary term in the variation will provide the Hamiltonian. The action for the perturbed GKV simplifies considerably to first order in the metric perturbations. Neglecting all terms of $O(h^2)$ and using symmetry properties of the resulting expression, the structure of the action can be simplified to (see Appendix B for the derivation),

$$ A[\delta \xi] = \int_V d^4x \sqrt{-g} \left[ \frac{1}{2} \left( 1 + \frac{1}{2} \hbar \right) \left\{ \bar{\mu} \delta_\alpha^\beta \delta_\beta^\alpha - g_{\alpha\beta} g^{\mu\nu} - h_{\alpha\beta} h^{\mu\nu} + g_{\alpha\beta} h^{\mu\nu} \right\} \nabla_\mu \delta \xi^\alpha \nabla_\nu \delta \xi^\beta \right] + \frac{\hbar}{2} \left\{ (\xi^\alpha + \delta \xi^\alpha) \nabla_\alpha (\nabla_\nu \delta \xi^\nu) \right\} - (\nabla_\mu \delta \xi^\alpha) \left\{ h_{\alpha\beta} \nabla^\nu \xi^\beta + h^{\mu\beta} \nabla_\nu \xi^\beta \right\} = -g_{\nu\alpha} \nabla_\mu \delta \xi^\alpha (\xi^\sigma + \delta \xi^\sigma) \nabla_\sigma h^{\mu\nu} \right\} . \right. \] (4.1)

Note that the Lagrangian density associated with the above action is identical to Eq. (3.15), though written in a different form. The variation of the above action has been performed in detail in Appendix B; collecting all the total derivative terms that we have thrown away in the derivation of the field equation for $\delta \xi^\mu$ in the previous section we obtain (recollecting the notation $\Delta$ for the first order variation),

$$ \text{Total Derivative Terms} = \sqrt{-g} \Delta (\delta \xi^\alpha) \left\{ \left( \mu \delta_\alpha^\beta \delta_\beta^\alpha - g_{\alpha\beta} g^{\mu\nu} - \delta_\alpha^\mu \delta_\beta^\mu \right) \nabla_\nu \delta \xi^\beta \right\} + \frac{1}{2} \hbar \left( \mu \delta_\alpha^\beta \delta_\beta^\alpha - g_{\alpha\beta} g^{\mu\nu} - \delta_\alpha^\mu \delta_\beta^\mu \right) \nabla_\nu \delta \xi^\beta + \frac{1}{2} \left( 1 + \frac{1}{2} \hbar \right) \bar{\mu} \delta_\alpha^\beta \nabla_\nu \delta \xi^\beta + \left( \delta_\alpha^\mu \delta_\beta^\mu \right) \nabla_\nu \delta \xi^\beta \right\} + \frac{\hbar}{2} \left\{ (\xi^\sigma + \delta \xi^\sigma) \nabla_\sigma \hat{h} \right\} \mu \delta_\alpha^\beta \nabla^\nu \delta \xi^\nu \right\} - \left( \nabla_\mu \delta \xi^\alpha \right) \left\{ \left( h_{\alpha\beta} \nabla^\nu \xi^\beta + h^{\mu\beta} \nabla_\nu \xi^\beta \right) - g_{\nu\alpha} (\xi^\sigma + \delta \xi^\sigma) \nabla_\sigma h^{\mu\nu} \right\} \right\} = \sqrt{-g} \Delta (\delta \xi^\alpha) \ P^\mu_\alpha . \right. \(4.2)

where, the last equality defines the quantity $P^\mu_\alpha$, which is the polymomentum conjugate to the perturbed GKV field $\delta \xi^\alpha$. In particular, starting from the perturbed action presented in Eq. (4.1), one can immediately verify that $P^\mu_\alpha$ has the following expression in terms of the perturbed Killing vector field $\delta \xi^\alpha$ and its derivatives:

$$ P^\mu_\alpha = \frac{1}{\sqrt{-g} \Delta (\nabla_\mu \delta \xi^\alpha)} \Delta A \left\{ \left( 1 + \frac{1}{2} \hbar \right) \left( \mu \delta_\alpha^\beta \delta_\beta^\alpha - g_{\alpha\beta} g^{\mu\nu} - \delta_\alpha^\mu \delta_\beta^\mu \right) \nabla_\nu \delta \xi^\beta \right\} + \frac{\hbar}{2} \left\{ (\xi^\sigma + \delta \xi^\sigma) \nabla_\sigma \hat{h} \right\} \mu \delta_\alpha^\beta \nabla^\nu \delta \xi^\nu \right\} - \left( \nabla_\mu \delta \xi^\alpha \right) \left\{ \left( h_{\alpha\beta} \nabla^\nu \xi^\beta + h^{\mu\beta} \nabla_\nu \xi^\beta \right) - g_{\nu\alpha} (\xi^\sigma + \delta \xi^\sigma) \nabla_\sigma h^{\mu\nu} \right\} . \right. \] (4.3)

We write the variation of the action for the perturbed Killing vector field $\delta \xi^\alpha$ (incorporating the variation of the boundary surface) in the Weiss form [24]:

$$ \Delta A = \int_V d^4x \sqrt{-g} \ E_a \Delta (\delta \xi^\alpha) + \int_{\partial V} d\Sigma_\alpha \left[ P^\mu_\alpha \Delta (\delta \xi^\alpha) \right] + \int_{\partial V} d\Sigma_\alpha \left[ P^\mu_\alpha \Delta (\delta \xi^\alpha) \right] + \int_{\partial V} d\Sigma_\alpha \left[ P^\mu_\alpha \Delta (\delta \xi^\alpha) \right] + \int_{\partial V} d\Sigma_\alpha \left[ P^\mu_\alpha \Delta (\delta \xi^\alpha) \right] , \right. \] (4.4)

where $d\Sigma_\alpha = d^3x \sqrt{-g} \nabla_\alpha \phi$ is the volume measure on the boundary surface $\partial V$, denoting the $\phi(x^\mu) = \text{constant}$ hypersurface. If we instead use the unit normal vector $n_\alpha$, then the volume measure of the boundary hypersurface $\partial V$ becomes $d\Sigma_\alpha = \epsilon d^3x \sqrt{n_\alpha}$, where $\epsilon = -1(+1)$ for space-like (time-like) hypersurfaces respectively. Defining $d\Sigma = d^3x \sqrt{\hat{h}}$ and choosing the hypersurface to be $t = \text{constant}$ and using the $(1 + 3)$ decomposition for the metric, we find the boundary Hamiltonian to be:

$$ H = \int d\Sigma \ H ; \right. \] H = \epsilon P^\mu_\alpha n_\mu \delta \xi^\alpha - NL \] , \right. \] (4.5)

where, $N$ is the lapse function and ‘overdot’ denotes derivative with respect to time. For the space-like hypersurface we are considering, and using orthonormal coordinates, e.g., in the synchronous frame, the Hamiltonian density $\hat{H}$ takes the following form,

$$ \hat{H} = P^\mu_\alpha \delta \xi^\alpha - \mathcal{L} . \right. \] \] (4.6)

We now consider terms which are quadratic in the time derivative as well as terms which are quadratic in the spatial derivative of the perturbed GKV field $\delta \xi^\alpha$. Collecting these terms, the time derivative part of the Hamiltonian density, quadratic in the perturbed Killing vector field, becomes,

$$ \hat{H}^{(2)} \text{time} = \frac{1}{2} \left( 1 + \frac{1}{2} \hbar \right) \left\{ \hat{\mu} - 2 \right\} \left\{ \delta \xi^0 \delta \xi^0 \right\} + \frac{1}{2} \left( 1 + \frac{1}{2} \hbar \right) \left\{ g_{ij} + h_{ij} + g_{ij} h^{00} \right\} \left\{ \delta \xi^i \delta \xi^j \right\} + \left( 1 + \frac{1}{2} \hbar \right) h_{ij} \left\{ \delta \xi^0 \delta \xi^i \right\} . \right. \] (4.7)

\] \]
Here, we have performed a (1+3) decomposition of the Hamiltonian density and have collected terms quadratic in the time derivative of the perturbed GKV field. In the limit of vanishing perturbation, the above quadratic contribution to the Hamiltonian coincides with that presented in [12]. Though the kinetic term of the zeroth component of the perturbed Killing vector field, i.e., $\delta \xi^0$ in the Hamiltonian density harbours a negative sign in the unperturbed spacetime for $\mu < 2$ (see [12]), we see that in Eq. (4.7), the corresponding kinetic term in perturbed GKV has positive sign for $\mu < 2 < \bar{\mu}$. Thus ghost modes can be avoided. The other quadratic terms in the time derivative have positive sign. Similarly, terms quadratic in the space derivatives of the perturbed Killing vector field yields the following expression for the Hamiltonian density,

$$\mathcal{H}^{(2)\text{space}} = \frac{1}{2} \left( 1 + \frac{1}{2} \bar{h} \right) \left\{ 1 - \bar{\mu} \right\} \left( \partial_{\alpha} \delta \xi^i \partial_{\beta} \delta \xi^j \right) + \frac{1}{2} \left( 1 + \frac{1}{2} \bar{h} \right) \left\{ -g^{ij} + h_{00} g^{ij} + h^{ij} \right\} \left( \partial_{\alpha} \delta \xi^0 \partial_{\beta} \delta \xi^0 \right) +
\frac{1}{2} \left( 1 + \frac{1}{2} \bar{h} \right) \left\{ g_{ab} g^{ij} + h_{ab} g^{ij} - g_{ab} h^{ij} \right\} \left( \partial_{\alpha} \delta \xi^a \partial_{\beta} \delta \xi^b \right) + \left( 1 + \frac{1}{2} \bar{h} \right) h_{0a} g^{ij} \left( \partial_{\alpha} \delta \xi^0 \partial_{\beta} \delta \xi^a \right).$$  \(\text{(4.8)}\)

In the above expression involving spatial derivatives of the perturbed GKV, the first two terms can provide a negative contribution to the Hamiltonian, thereby making it unbounded from below. In the second term, even though the metric perturbations try to keep this term positive, the background metric drives it to negative values. Similarly, if we want the theory to be ghost-free, the first term will turn negative, leading to an unbounded Hamiltonian. As argued in [12] in the unperturbed case, an unbounded Hamiltonian is potentially dangerous, as it can result in runaway behavior that drives the GKV field far from the Killing condition. Thus the problems associated with the Hamiltonian density for the AKD in the exact case remain for the perturbed GKV $\delta \xi^\alpha$ as well.

\section*{B. Hyperbolicity}

We now turn our attention to the hyperbolicity of the perturbed AKD. For this purpose, we employ the methods of hyperbolicity analysis for second order systems, particularly that presented in [25, 26] (see also [12]). In this approach, we compute the principal symbol for the system of equations—if the principal symbol has real eigenvalues, the system is weakly hyperbolic, and if the principal symbol has a complete set of Eigen vectors, the system is strongly hyperbolic. Collecting all the terms involving double derivatives of $\delta \xi^\alpha$, we obtain from Eq. (3.29),

$$\left( g^{\alpha\beta} - h^{\alpha\beta} \right) \partial_{\alpha} \partial_{\beta} \delta \xi^\mu + (1 - \mu - \delta \mu) \left( g^{\mu\alpha} - h^{\mu\alpha} \right) \partial_{\alpha} \partial_{\sigma} \delta \xi^\sigma \approx 0,$$  \(\text{(4.9)}\)

where, the symbol $\approx$ denotes equality up to terms not included in the principal part. In order to express the above equation in the desired form we can decompose the metric as $g^{\alpha\beta} = q^{\alpha\beta} - n^\alpha n^\beta + s^\alpha s^\beta$, where $n^\alpha$ is a timelike unit vector and $s^\alpha$ is a spacelike unit vector. Further, defining $n^\alpha \partial_{\alpha} = \partial_n$, $s^\alpha \partial_{\alpha} = \partial_s$ and $q^{\alpha\beta} = q^{AB} \delta^\alpha_s \delta^\beta_n$, we can rewrite the above equation into three separate equations; we obtain one by contraction of Eq. (4.9) with $n^\mu$, another by contraction of Eq. (4.9) with $s^\mu$, and the last by the projection of Eq. (4.9) along transverse directions. Keeping only the principal parts of these equations, we obtain (for the derivation, see Appendix E),

$$\left( 2 - \mu - \delta \mu \right) (1 + h^{nn}) \partial^2_n \delta \xi^\mu \approx \left( 1 - h^{ss} \right) \partial^2_s \delta \xi^\mu + (3 - \mu - \delta \mu) h^{ns} \partial_n \partial_s \delta \xi^\mu +
(1 - \mu - \delta \mu) \left( 1 + h^{nn} \right) \partial_n \partial_s \delta \xi^s + (1 - \mu - \delta \mu) h^{ns} \partial^2_s \delta \xi^s. \quad \text{(4.10)}$$  \(\text{hypb_04}\)

Note that since $n^\mu$ is the time-like unit vector, $\partial^2_n \delta \xi^\mu$ corresponds to a double time derivative of the time component of the perturbed GKV. A similar analysis yields the following equation for $\partial^2_s \delta \xi^s$, i.e., for the double time derivative of the spatial component of the perturbed GKV,

$$\left( 1 + h^{nn} \right) \partial^2_n \delta \xi^s \approx (2 - \mu - \delta \mu) (1 - h^{ss}) \partial^2_s \delta \xi^s + (3 - \mu - \delta \mu) h^{ns} \partial_n \partial_s \delta \xi^s +
(1 - \mu - \delta \mu) h^{ns} \partial^2_n \delta \xi^s - (1 - \mu - \delta \mu) h^{ns} \partial^2_n \delta \xi^s. \quad \text{(4.11)}$$  \(\text{hypb_06}\)

Finally, the double time derivative of the transverse component of the perturbed GKV yields,

$$\left( 1 + h^{nn} \right) \partial^2_n \delta \xi^A \approx (1 - h^{ss}) \partial^2_s \delta \xi^A + 2 h^{ns} \partial_n \partial_s \delta \xi^A + (1 - \mu - \delta \mu) h^{As} \partial_n \partial_s \delta \xi^n +
(1 - \mu - \delta \mu) h^{As} \partial^2_n \delta \xi^n - (1 - \mu - \delta \mu) h^{As} \partial^2_n \delta \xi^n. \quad \text{(4.12)}$$  \(\text{hypb_08}\)

Therefore, we can read off the principal symbol $P^s$ for this system of second-order differential equations, which takes the following form (again see Appendix E for the derivation),

$$P^s = \begin{bmatrix} O_{4 \times 4} & I_{4 \times 4} \\ 0 & B \end{bmatrix},$$  \(\text{(4.13)}\)

where $O_{4 \times 4}$ is a 4-dimensional identity matrix and $B$ is a 4-dimensional matrix with a single non-zero entry.
where $O_{4 \times 4}$ and $I_{4 \times 4}$ are the $(4 \times 4)$ null and unit matrix respectively. The entries $A$ and $B$ are also $(4 \times 4)$ matrices with the following expressions,

\[
\begin{bmatrix}
A & B & 0 & 0 \\
C & D & 0 & 0 \\
E & F & 0 & 0 \\
G & H & 0 & 0 \\
\end{bmatrix}
\]

(4.14)

where the unknown quantities appearing in the matrices $A$ and $B$ have the following expressions,

\[
A = \frac{(1 - h_{ss})}{(2 - \mu - \delta \mu)(1 + h_{nn})}; \quad B = -\frac{(1 - \mu - \delta \mu) h_{ns}}{(2 - \mu - \delta \mu)(1 + h_{nn})}
\]

(4.15)

\[
C = \frac{(3 - \mu - \delta \mu) h_{ns}}{(2 - \mu - \delta \mu)(1 + h_{nn})}; \quad D = \frac{(1 - \mu - \delta \mu)(1 - h_{ss})}{(2 - \mu - \delta \mu)}
\]

(4.16)

\[
E = -\frac{(1 - \mu - \delta \mu) h_{nn}^2}{(2 - \mu - \delta \mu)(1 + h_{nn})}; \quad F = \frac{(2 - \mu - \delta \mu)(1 - h_{ss})}{(1 + h_{nn})}
\]

(4.17)

\[
G = -\frac{(1 - \mu - \delta \mu) h_{An}}{(2 - \mu - \delta \mu)(1 + h_{nn})}; \quad \mathcal{H} = \frac{(3 - \mu - \delta \mu) h_{ns}}{(1 + h_{nn})} - \frac{(1 - \mu - \delta \mu)^2 h_{nm}}{(2 - \mu - \delta \mu)(1 + h_{nn})}
\]

(4.18)

\[
I = -\frac{(1 - \mu - \delta \mu) h_{ss}^2}{(1 + h_{nn})}; \quad J = -\frac{(1 - \mu - \delta \mu) h_{ss}^2}{(1 + h_{nn})}
\]

(4.19)

\[
K = \frac{(1 - h_{ss})}{(1 + h_{nn})}; \quad L = \frac{(1 - \mu - \delta \mu) h_{ss}^2}{(1 + h_{nn})}
\]

(4.20)

\[
M = -\frac{(1 - \mu - \delta \mu)^2 h_{An}}{(2 - \mu - \delta \mu)(1 + h_{nn})} + \frac{(1 - \mu - \delta \mu) h_{An}}{(1 + h_{nn})}; \quad \mathcal{N} = 2\frac{h_{ss}}{(1 + h_{nn})}
\]

(4.21)

Note that as the perturbations vanish, we obtain: $A = 1/(2 - \mu)$, $B = 0 = C$, $D = (1 - \mu)/(2 - \mu)$, $E = 0$, $F = 2 - \mu$, $\mathcal{G} = -(1 - \mu)$, $\mathcal{H} = 0 = I = J$, $K = 1$ and $L = 0 = M = \mathcal{N}$. Thus our result agrees with that derived in [12]. If one chooses $\mu + \delta \mu = 1$, one has a complete set of Eigenvectors, and if one chooses $\mu + \delta \mu = 2$, the principal symbol becomes singular. Setting $\mu = 1$ and expanding the eigenvalues to first order in metric perturbations and second order in $\delta \mu$, one obtains the following expression for the eigenvalues of the principal symbol:

\[
\begin{cases}
\frac{1}{2}(h_{nn} - 2h_{ns} + h_{ss} - 2), & \frac{1}{2}(h_{nn} + 2h_{ns} + h_{ss} - 2), \\
\frac{1}{4}(-\delta \mu^2 h_{ns} + (\delta \mu^2 + 4) h_{ns} - 2h_{nn} - 2(h_{ss} - 2)), & \\
\frac{1}{4}(-\delta \mu^2 h_{ns} + (\delta \mu^2 + 4) h_{ns} - 2h_{nn} - 2(h_{ss} - 2)), & \\
\frac{1}{4}(-\delta \mu^2 h_{ns} + (\delta \mu^2 + 4) h_{ns} - 2h_{nn} - 2(h_{ss} - 2))
\end{cases}
\]

(4.22)

Since the eigenvalues, presented above, are real, this implies that to first order in the metric perturbation and second order in $\delta \mu$, the perturbed AKE is weakly hyperbolic. This is consistent with the result in [12], in which the AKE for the background spacetime was found to be weakly hyperbolic for general $\mu \neq 1, 2$.

5. **Induced Perturbations: Implications for the Almost Killing Equation**

So far, we had considered the metric perturbation and the perturbation of the GKV to be independent. However, in most of the physical scenarios of interest, e.g., perturbation of the black hole spacetime due to matter fields entering the horizon, the metric perturbation sources the perturbation of the almost Killing equation. Thus we may consider the perturbation of the GKV to be induced by the metric perturbation. In what follows we will consider such induced perturbation of the GKV and hence determine their evolution equations order by order.

---

4 These calculations were performed using the package xAct in Mathematica.
A. First order perturbation

As emphasized before, we will imagine a class of nontrivial perturbative solutions to the AKE which are induced by the metric perturbations. This may be quantified by assuming that $\delta \xi^\alpha$ and $h_{\alpha\beta}$ are implicitly proportional to the same expansion parameter $\epsilon$ (see footnote 2), and then solving the perturbed AKE order by order in $\epsilon$. In particular, for the GKV $\chi^\alpha$ we write,

$$\chi^\alpha = \xi^\alpha + \delta \chi^\alpha_1 + \delta \chi^\alpha_2 + \mathcal{O}(\epsilon^3) ,$$  \hfill (5.1) \tag{GKV$_\text{Pert}$}

where, we have assumed $\delta \chi^\alpha_1 \propto \epsilon^1$. As before, the background spacetime is assumed to admit a Killing vector field $\xi^\alpha$, which satisfies the AKE, presented in Eq. (2.2), exactly. Further, imposing the Lorenz gauge condition, $\nabla^\beta h_{\alpha\beta} = \nabla\alpha h/2$, and choosing the background value of the parameter $\mu$ to be constant, the expansion of the perturbed AKE, presented in Eq. (3.29) to $\mathcal{O}(\epsilon)$ reduces to the following equation for $\delta \chi^\alpha_1$:

$$\Box \delta \chi^\alpha_1 + R^\alpha_\beta \delta \chi^\beta_1 + (1 - \mu) \nabla^\alpha \left[ \nabla \cdot \delta \chi_1 + (\xi^\beta \nabla h_{\beta\alpha}/2) \right] = 0 .$$  \hfill (5.2) \tag{AKE$_\text{PertI}$}

Note that this reduces to the background AKE when $\mu = 1$ or, when $\xi^\beta \nabla h_{\beta\alpha}$ is a constant; in those cases, one must consider higher-order corrections to the AKE. In the cases where $\xi^\beta \nabla h_{\beta\alpha}$ is nontrivial, for instance, if the perturbations are the result of an energy-momentum tensor with nontrivial trace; one can apply the expression Eq. (2.4) for the Noether current, defined with respect to the background derivatives, directly to $\delta \chi^\alpha_1$ and substitute Eq. (5.2) to obtain the following current:

$$J^\alpha_1 = 2R^\alpha_\beta \delta \chi^\beta_1 + \nabla^\alpha \left[ (2 - \mu) \nabla \cdot \delta \chi_1 + (1 - \mu) (\xi^\alpha \nabla h_{\alpha\alpha}/2) \right] ,$$  \hfill (5.3) \tag{KomarPAKE$_\text{PertI}$}

which, satisfies the identity $\nabla \cdot J_1 = 0$. If the background spacetime is vacuum, setting $R^\alpha_\beta = 0$, this identity yields:

$$\Box \Phi_1 = 0 ,$$  \hfill (5.4) \tag{KomarPAKE$_\text{PertI}$}

where:

$$\Phi_1 := (2 - \mu) \nabla \cdot \delta \chi_1 + \frac{1}{2} (1 - \mu) (\xi^\alpha \nabla h_{\alpha\alpha}) .$$  \hfill (5.5) \tag{Constraint$_\text{PertI}$}

It follows that if the initial data satisfies the constraint $\Phi_1 = 0$ and its first time derivative also vanishes, then the constraint $\Phi_1 = 0$ is preserved by the evolution of Eq. (5.4). Under this constraint, Eq. (5.2) may be rewritten as:

$$\Box \delta \chi^\alpha_1 - \nabla^\alpha (\nabla \cdot \delta \chi_1) = 0 ,$$  \hfill (5.6) \tag{AKE$_\text{PertI}$}

which is equivalent to the AKE for $\mu = 2$ in vacuum spacetime; it was demonstrated in [12] that the $\mu = 2$ parameter choice avoids ghosts and is dynamically equivalent to the Maxwell theory, thus avoiding the problems arising from an unbounded Hamiltonian.

Furthermore, Recall that to first order in the gravitational perturbation, imposing the Lorenz gauge condition, the trace $h$ of the gravitational perturbation satisfies the following evolution equation on a vacuum background:

$$\Box h = 16\pi \delta_1 T ,$$  \hfill (5.7) \tag{EFEPERTtr$_\text{Pert}$}

where, $\delta_1 T$ is the first order perturbation of the trace of the matter energy-momentum tensor. If one assumes $\delta_1 T = 0$, i.e., the matter field perturbing the spacetime is dilute radiation and the background is vacuum, one can impose the condition that $h = 0$. In which case, from Eq. (5.5), it follows that the constraint $\Phi_1 = 0$ implies $\nabla \cdot \delta \chi_1 = 0$. Subsequently, substituting this result in Eq. (5.6) simplifies to $\Box \delta \chi^\alpha_1 = 0$, which is satisfied by the background Killing vector field $\xi^\alpha$. Thus one may expect that in this case, the perturbed spacetime will respect the symmetries of the background spacetime. On the other hand, if $\delta_1 T \neq 0$, then the background Killing vector field does not, in general, satisfy the constraint $\Phi_1 = 0$, and the resulting solutions for Eq. (5.6) differ nontrivially from the background Killing vector.

We have therefore derived that if the spacetime perturbation $h_{\alpha\beta}$ is being sourced by an energy-momentum tensor with a non-vanishing trace, then the evolution equation for the first-order perturbation $\delta \chi^\alpha_1$ must differ from the background Killing vector field in a nontrivial manner. Furthermore, we find that for the weakly hyperbolic parameter choices, i.e., $\mu \neq 1$ and $\mu \neq 2$, and initial data satisfying the constraint $\Phi_1 = 0$ and $\delta_1 \Phi_1 = 0$, the AKE propagates the constraint $\Phi_1 = 0$ and avoids the dynamical instabilities associated with ghosts and unboundedness in the Hamiltonian. First-order perturbations of the AKE are therefore different from the background Killing vector field and are well-suited for describing the perturbations of the vacuum spacetimes induced by dilute matter in which $\delta_1 T \neq 0$. 


B. Second order perturbations

As the previous section demonstrates, perturbations of the AKE to \( \mathcal{O}(\epsilon) \) are insensitive to the metric perturbations in scenarios with \( \delta_1 T = 0 \). This is the case if the perturbations consist of weak gravitational radiation, or dilute matter source encompassing radiation or null dust. In such situations, one should consider the evolution equation for \( \delta \chi_\nu^{\mu} \). Assuming constant \( \mu \) and a vacuum background spacetime, the homogeneous part of the equation is the same as that of the first-order case:

\[
\Box \delta \chi_\nu^{\mu} + (1 - \mu) \nabla^\nu (\nabla \cdot \delta \chi_\nu) = - (j_2^{\nu} + k_2^{\nu} + l_2^{\nu} + m_2^{\nu}) ,
\]

and the inhomogeneous part \( j_2^{\nu} \) takes the explicit form:

\[
j_2^{\nu} = 2 \xi^\alpha h^{\beta \sigma} \{ h^{\nu \tau} R_{\alpha \beta \tau \sigma} - h^{\beta \tau} R^{\nu} \sigma \alpha \tau \sigma - \nabla_\sigma \nabla_\alpha h^{\nu \beta} \} - \xi^\alpha \nabla_\alpha h^\beta \sigma \nabla^{\nu} h^{\beta \sigma} + 2 \nabla^\nu (\xi^\alpha h^{\beta \sigma} \nabla_\alpha h^\beta \sigma) - 2 h^{\beta \sigma} \nabla^\nu \xi^\alpha \nabla_\alpha h_{\beta \sigma} + 2 h_{\alpha \sigma} \nabla^\beta \xi^\alpha \{ \nabla_\beta h^{\nu \sigma} - \nabla^{\nu} h_\beta \sigma + \nabla_\sigma h^{\nu \beta} \} ,
\]

\[
k_2^{\nu} = 2 h^{\nu \beta} \nabla_\beta \nabla_\alpha \delta \chi_\nu^{\lambda} - 2 h_{\alpha \beta} \nabla^\beta \nabla_\alpha \delta \chi_\nu^{\lambda} + 2 \nabla^\nu \delta \chi_\nu^{\lambda} \{ \nabla_\alpha h^{\nu \beta} + \nabla_\beta h^{\nu \alpha} - \nabla^{\nu} h_{\alpha \beta} \} - 2 \delta \chi_\nu^{\lambda} h_{\beta \sigma} R^{\nu \alpha \beta \sigma} ,
\]

\[
l_2^{\nu} = h^{\nu \beta} \nabla_\beta \xi^\alpha \nabla_\alpha h - \nabla^\nu [\delta \chi_\nu^{\lambda} \nabla_\alpha h + (2 \nabla_\alpha \delta \chi_\nu^{\lambda} + \xi^\alpha \nabla_\alpha h) \delta \mu] + (\mu - 2) \{ h^{\nu \beta} \nabla_\beta (\xi^\alpha \nabla_\alpha h) - \nabla^\nu (\delta \chi_\nu^{\lambda} \nabla_\alpha h) \} ,
\]

\[
m_2^{\nu} = (\mu - 2) \{ \nabla^\nu (\xi^\alpha h^{\beta \sigma} \nabla_\alpha h_{\beta \sigma}) + 2 h^{\nu \beta} \nabla^\alpha \delta \chi_\nu^{\lambda} \} .
\]

The terms in Eq. (5.9) are organized so that \( k_2^{\nu} = 0 \) if \( \delta \chi_\nu^{\mu} \) satisfies the Killing equation, and \( l_2^{\nu} = 0 \) if \( \delta_1 T = 0 \). As pointed out in the first order case, for \( h = 0 \), the relevant solutions for \( \delta \chi_\nu^{\mu} \) are essentially those of the background Killing vector — in that case, one may choose \( \delta \chi_\nu^{\mu} = 0 \), as one can absorb it into the background Killing vector field.

Further understanding of the second-order perturbation can be achieved by computing the second-order perturbation of the Noether current assuming constant \( \mu \) and a vacuum background \( R_{\mu \nu} = 0 \):

\[
j_2^{\nu} = 4 \delta \chi_\nu^{\alpha} \delta_1 R_{\nu}^{\alpha} + 2 \xi^\alpha \left( \delta_2 R_{\nu}^{\alpha} - 2 h^{\nu \beta} \delta_1 R_{\alpha \beta} \right) + (\mu - 2) \left[ \nabla_\alpha h^{\sigma \tau} (\xi^\alpha \nabla^\nu (\xi^\sigma h^{\tau \nu} + h^{\sigma \tau} \nabla^\nu \xi^\alpha) + \xi^\alpha h^{\sigma \tau} \nabla^\nu \nabla_\alpha h^{\sigma \tau} \right]
\]

\[
+ 2 h^{\nu \beta} \nabla^\beta (\nabla \cdot \delta \chi_\nu) - \nabla^\nu (\nabla \cdot \delta \chi_\nu) - \left\{ \nabla^\nu (\delta \chi_\nu^{\lambda} \nabla_\alpha h) - h^{\nu \beta} \nabla_\beta (\xi^\alpha \nabla_\alpha h) \right\} ,
\]

where \( \delta_1 R_{\alpha \beta} \) and \( \delta_2 R_{\alpha \beta} \) are the respective first and second-order perturbations of the lowered index Ricci tensor. Note that the quantity within the curly brackets vanishes when \( h = 0 \). We further see that when \( \mu = 2 \), most of the terms in the Noether current, except for the first two terms, identically vanishes. Therefore, for \( \mu = 2 \), the Noether current depends on the perturbations of the Ricci tensor alone, as one might expect.

We then turn to the second-order perturbations of the identity presented in Eq. (2.7). Assuming vacuum spacetime along with transverse-traceless gauge \( h = 0 \), and setting \( \mu \) to be a constant, the perturbation of the Komar identity takes the following form:

\[
(\mu - 2) \Psi_L = \Psi_R ,
\]

where:

\[
\Psi_L \equiv \Box [\nabla \cdot \delta \chi_2 + \delta \chi_1 \nabla_\alpha h - h^{\sigma \tau} \nabla_\alpha (\xi^\sigma h^{\tau \nu}) + 2 h^{\sigma \tau} (\nabla_\nu \nabla_\sigma \nabla_\tau \delta \chi_2 + R_{\alpha \sigma \tau \beta} \nabla^\beta \delta \chi_2)]
\]

\[
+ \nabla_\beta [\nabla_\nu (h^{\beta \sigma} \xi^\alpha \nabla_\alpha h) + \nabla^\beta h \nabla_\beta (\xi^\alpha \nabla_\alpha h)] ,
\]

\[
\Psi_R \equiv \xi^\alpha (\nabla_\alpha \delta_2 R - 2 h^{\sigma \tau} \nabla_\alpha \delta_1 R_{\sigma \tau}) - 4 \delta_1 R_{\alpha \beta} h^{\beta \sigma} \nabla^\sigma \xi^\alpha + 2 \left( \delta \chi_2 \nabla_\alpha \delta_1 R + 2 \delta_1 R_{\alpha \beta} \nabla^\beta \delta \chi_2 \right) .
\]

Note that the quantity \( \Psi_R \) is independent of the second-order perturbation of the GKV \( \delta \chi_2 \), and when \( \mu = 2 \), it follows from Eq. (5.11) that \( \Psi_R = 0 \). For vacuum spacetime, in the transverse-traceless gauge, at first order, one may choose \( \delta \chi_1 \) to satisfy the background Killing equation and hence the terms dependent on \( \delta \chi_1 \) disappear from \( \Psi_R \), so that \( \Psi_R \) depends only on the background quantities and the perturbation \( h_{\mu \nu} \). It follows that to satisfy \( \Psi_R \) as a function of \( h_{\mu \nu} \) which satisfies \( h = \delta_1 R = 0 \) and permits a solution to the \( \mu = 2 \) AKE.

If \( \Psi_R = 0 \), the second-order perturbations should satisfy the following:

\[
\Box \Phi_2 = 0 ,
\]
where:

$$\Phi_2 := (\mu - 2) \left\{ \nabla \cdot \delta \chi_2 - h^{\sigma \tau} \nabla_\sigma (\xi^\alpha h_{\sigma \tau}) \right\} .$$  \hfill (5.15) \{Constraint\}

It follows that when $\Psi_R = 0$, the field equations propagate the constraint $\Phi_2 = 0$ for the second order perturbations (assuming initial data satisfying $\partial_t \Phi_2 = 0$). On a vacuum background, one may (assuming $\mu \neq 2$) use the constraint $\Phi_2 = 0$ to rewrite Eq. (5.8):

$$\Box \delta \chi_2 - \nabla \cdot (\delta \chi_2) = -j_2^\nu - m_2^\nu - (2 - \mu) \nabla \left\{ h^{\sigma \tau} \nabla_\alpha (\xi^\alpha h_{\sigma \tau}) \right\} ,$$  \hfill (5.16) \{AKEPertI\}

so that the equation for $\delta \chi_2^\nu$ resembles the $\mu = 2$ equation with a source. For second-order perturbations, the question of whether the perturbations suffer from dynamical instabilities due to the unboundedness of the Hamiltonian depends on the behavior of the RHS of Eq. (5.16).

The claim that $\Psi_R = 0$ for transverse-traceless metric perturbations of vacuum spacetimes that admit Killing vectors suggests an identity for such perturbations. However, the arguments we have presented so far do not yet constitute a proof of such an identity, as they depend on the existence of solutions for the exact $\mu = 2$ AKE. Though the AKE fails to admit a well-posed initial value problem for the $\mu = 2$ parameter choice, there is some reason to expect that the failure is primarily due to nonuniqueness, rather than existence. For instance, it is straightforward to show that on vacuum spacetimes, the exact AKE becomes an identity for the gradient of an arbitrary function. Furthermore, one can show that in locally flat coordinates, the time derivatives for the time component of $\chi^\mu$ disappear, so that the AKE is an underdetermined dynamical system. One might therefore expect the existence of solutions to the exact $\mu = 2$ AKE (and consequently $\Phi_2 = 0$) to hold for a large class of transverse-traceless metric perturbations of vacuum spacetimes which admit Killing vectors.

6. PERTURBATIONS OF THE NOETHER CHARGE AND ITS THERMODYNAMICAL INTERPRETATION

In this section, we will discuss how the Noether charge associated with the GKV $\chi^\alpha$ associated with the AKE is affected by the perturbation of the metric. As we will demonstrate, the perturbed Noether charge, will have an interesting thermodynamical interpretation. Applying Eq. (2.4) to the solutions of the AKE, which correspond to the GKV field $\chi^\mu$, the Noether current for the GKV in the perturbed spacetime takes the following form (see Appendix F for a derivation),

$$\bar{J}^\mu = 2 \bar{R}^\mu_\nu \chi^\nu + (2 - \bar{\mu}) \nabla^\mu \left\{ \nabla_\sigma \chi^\sigma \right\} + (\nabla_\sigma \chi^\sigma) \bar{g}^{\mu \alpha} \nabla_\alpha (2 - \bar{\mu}) ,$$  \hfill (6.1) \{Noether_\nu_\sigma_\rightarrow_\nu_\alpha_\rightarrow\}

where the AKE presented in Eq. (2.2) has been used. Let us now use the fact, following Eq. (3.1), that the spacetime metric can be expressed as $\bar{g}_{\mu \nu} = g_{\mu \nu} + h_{\mu \nu}$, where $h_{\mu \nu}$ is the perturbation, possibly due to some matter field entering the background spacetime geometry. As a consequence, we also have $\xi^\sigma \rightarrow \xi^\sigma + \delta \xi^\sigma = \chi^\sigma$, where $\chi^\sigma$ is the GKV, with the associated Noether current being given by Eq. (6.1).

Thus the Noether current associated with the GKV field can be decomposed into the Komar current for the background Killing vector field $\xi^\sigma$, and a part containing additional corrections arising out of the gravitational perturbation $h_{\mu \nu}$ and the perturbation of the Killing vector field $\delta \xi^\alpha$. We then obtain the following expression for the Noether current associated with the GKV field $\chi^\sigma$,

$$\bar{J}^\nu = 2 \bar{R}^\nu_\sigma \chi^\sigma + (-\Box h^\nu_\sigma + R^\nu_\mu h_{\sigma \mu} - R_{\sigma \mu} h^{\mu \nu} - 2 \bar{R}^\nu_\mu_\sigma_\rho_\nu h^{\mu \rho}) \chi^\sigma$$

$$+ (2 - \bar{\mu}) (g^{\nu \alpha} - h^{\nu \alpha}) \nabla_\alpha \left( \nabla_\sigma \chi^\sigma + \frac{1}{2} \chi^\sigma \nabla_\sigma h \right) + (g^{\nu \alpha} - h^{\nu \alpha}) \left( \nabla_\sigma \chi^\sigma + \frac{1}{2} \chi^\sigma \nabla_\sigma h \right) \nabla_\alpha (2 - \bar{\mu}) ,$$  \hfill (6.2) \{Noether_\nu_\sigma_\rightarrow_\nu_\alpha_\rightarrow\}

where we have used the Lorenz gauge condition to simplify the expression further. The above provides the expansion of the Noether current in terms of the gravitational perturbation, further expressing $\chi^\sigma = \xi^\sigma + \delta \xi^\sigma$, we obtain the following expression for the change in the Noether current,

$$\delta \bar{J}^\nu = 2 \bar{R}^\nu_\sigma \delta \xi^\sigma + (-\Box h^\nu_\sigma + R^\nu_\mu h_{\sigma \mu} - R_{\sigma \mu} h^{\mu \nu} - 2 \bar{R}^\nu_\mu_\sigma_\rho_\nu h^{\mu \rho}) (\xi^\sigma + \delta \xi^\sigma)$$

$$+ (2 - \bar{\mu}) (g^{\nu \alpha} - h^{\nu \alpha}) \nabla_\alpha \left( \nabla_\sigma \delta \xi^\sigma + \frac{1}{2} (2 - \bar{\mu}) g^{\nu \alpha} \nabla_\alpha (\xi^\sigma + \delta \xi^\sigma) \nabla_\sigma h \right)$$

$$- \left[ g^{\nu \alpha} \left( \nabla_\sigma \delta \xi^\sigma + \frac{1}{2} (\xi^\sigma + \delta \xi^\sigma) \nabla_\sigma h \right) - h^{\nu \alpha} (\nabla_\sigma \delta \xi^\sigma) \right] \nabla_\alpha \delta \mu .$$  \hfill (6.3) \{change_noether_current\}
Here we have assumed that the background spacetime inherits Killing symmetry and hence $\xi^\sigma$ is a Killing vector field, such that $\nabla_\sigma \xi^\sigma = 0$, which we have used in deriving the above expression. Using the expressions for the perturbations of the Ricci tensor and the Einstein tensor in the Lorenz gauge, the above change in the Noether current may be expressed in several ways, among which we quote the expression involving the Einstein tensor below (see Appendix F for the detailed derivation),

$$
\delta J^\nu = 2R^\nu_\sigma \delta \xi^\sigma + 2\delta G^\nu_\sigma (\xi^\sigma + \delta \xi^\sigma) - \left(\frac{1}{2} \nabla^2 h + R_{\mu\nu} h^{\mu\rho} \right) (\xi^\nu + \delta \xi^\nu) + (2 - \bar{\mu}) (g^{\nu\alpha} - h^{\nu\alpha}) \nabla_\alpha (\nabla_\sigma \delta \xi^\sigma) + \frac{1}{2} (2 - \bar{\mu}) g^{\nu\alpha} \nabla_\alpha [(\xi^\sigma + \delta \xi^\sigma) \nabla_\sigma h] - \left[g^{\nu\alpha} \left(\nabla_\sigma \delta \xi^\sigma + \frac{1}{2} (\xi^\sigma + \delta \xi^\sigma) \nabla_\sigma h\right) - h^{\nu\alpha} (\nabla_\sigma \delta \xi^\sigma) \right] (\nabla_\sigma \delta \mu) .
$$

(6.4) {change_noether_current_linearN}

Observe that, using the perturbed Einstein’s equations, one can replace the perturbation of the Einstein tensor appearing in the above expression with the perturbation of the matter energy-momentum tensor. In that case, the object $(\delta G^\nu_\sigma \xi^\sigma)$ will correspond to the flux of the matter energy-momentum tensor through the Killing horizon, to which the Killing vector field is orthogonal. We will come back to this point later in this section, but we first discuss a couple of interesting limits:

- Even though we have treated the gravitational perturbation and the perturbation of the Killing vector field separately, the perturbation $\delta \xi^\alpha$ we are interested in, is induced by the perturbation of the background spacetime. Thus it is natural to decompose the perturbed Killing vector field as $\delta \chi^\mu_1 + \delta \chi^\mu_2$ (cf. Eq. (5.1)), where $\delta \chi^\mu_1$ is linear in the gravitational perturbation and $\delta \chi^\mu_2$ is quadratic in the gravitational perturbation. Then to the linear order in the gravitational perturbation, the perturbed Noether current, from Eq. (6.4), takes the following form:

$$
\delta J^\nu = 2R^\nu_\sigma \delta \chi^\sigma_1 + 2\delta_1 R^\nu_\sigma \xi^\sigma + (2 - \mu) \nabla^\nu (\nabla \cdot \delta \chi_1) + \frac{1}{2} (2 - \mu) \nabla_\nu (\xi^\sigma \nabla_\sigma h) .
$$

(6.5) {change_noether_currentN}

where the Lorenz gauge condition has been used. In the first line of the above expression, we have expressed the change in the Noether current in terms of the change in the Ricci tensor. In the second line, we express the change in Noether current in terms of the change in the Einstein tensor, and apply the constraint $\Phi_1 = 0$ (with $\Phi_1$ given in (Eq. (5.5))], which follows from the fact that $\delta \chi^\mu_1$ satisfies Eq. (5.2), which as we showed earlier propagates the constraint $\Phi_1 = 0$ for an appropriate choice of initial data.

- For spacetimes which may contain a dilute amount of nongravitational radiation on vacuum backgrounds, i.e., with $R_{\mu\nu} = 0$, one can use the transverse-traceless gauge (effectively setting $h = 0$) and hence the above expression for the change in the Noether current simplifies considerably:

$$
\delta J^\nu = 2\delta_1 G^\nu_\sigma \xi^\sigma .
$$

(6.7) {reduction}

Thus using the perturbed Einstein’s equations, we find that given the constraint $\Phi_1 = 0$, the change in the Noether current to first order is simply equal to $16\pi (\delta T^\nu_\sigma \xi^\sigma)$, which corresponds to the matter field flowing into the Killing horizon. This will have thermodynamical interpretation, as we compute the associated change in the Noether charge.

- For spacetimes which contain a dilute amount of matter on vacuum backgrounds, one can instead employ the $\mu = 2$ parameter choice, in which case the Noether current also simplifies to:

$$
\delta J^\nu = 2\delta_1 R^\nu_\sigma \xi^\sigma = (2\delta_1 G^\nu_\sigma + \delta_\nu^\rho \delta_1 R) \xi^\sigma .
$$

(6.8) {reduction}

We see that for matter fields, the Noether current does not measure energy and momentum in the sense of the energy-momentum tensor, due to the term containing $\delta_1 R$. However, one may nonetheless still regard the Noether current and its associated charge as a measure of the matter content, and an analysis [27] comparing Komar integrals for radiation with that of matter in cylindrical symmetry suggests that charges constructed from Noether currents measure the effective gravitating mass.

---

5 One should be careful not to confuse this expression with Eq. (5.3), which satisfies the identity $\nabla \cdot J_1 = 0$. $\delta J^\alpha$, on the other hand, satisfies $\nabla \cdot (J + \delta J) = 0$. 
Thus we have discussed situations of physical interest and how the change in the Noether current can be simplified and interpreted in these scenarios. We will now proceed to compute the change in the Noether charge due to the gravitational perturbation and the perturbation of the Killing vector field.

In order to determine the Noether charge, we have to integrate the Noether current over a three-surface. One may define such a hypersurface as a level surface of some scalar function $\Phi = \Phi(x)$; defined in this way the surface is held fixed in the manifold and does not change under metric perturbations. However, the measure on the surface will change. In particular, for a $\Phi = \text{constant}$ surface, the integration measure is $d\Sigma_\alpha = d^3x \sqrt{h} n_\alpha$, where $n_\alpha$ is the normalized normal on this surface and $h$ is the determinant of the induced metric on the $\Phi = \text{constant}$ surface. It is also possible to express the integration measure on a $\Phi = \text{constant}$ surface as $d\Sigma_\alpha = d^3x \sqrt{-g} \nabla_\alpha \Phi$, where $\nabla_\alpha \Phi$ is the unnormalized normal to the $\Phi = \text{constant}$ surface. Under metric perturbations, the measure changes as $\sqrt{-g} \to \sqrt{-g} = \sqrt{-g}\{1 + (1/2)\hbar\}$. Thus the Noether charge associated with the GKV field $\chi^\alpha = \xi^\alpha + \delta \xi^\alpha$ becomes,

$$
\delta Q = \int d\Sigma_\alpha \delta G_{\alpha \beta} \xi^\beta,
$$

(6.11)

As we will discuss below, this expression can be understood from a physical as well as thermodynamical perspective.

We consider the case where the metric $g_{\mu \nu}$ is a vacuum black hole spacetime, with a Killing vector field $\xi^\alpha$ defining the Killing horizon, to which the Killing vector is orthogonal. Since the Killing horizon is a null surface, it follows that we can consider the surface on which the Noether charge is computed to be the Killing horizon, which is bound to have a thermodynamic interpretation. In particular, for a generic null surface, the Noether charge associated with a Killing-like vector corresponds to $Q = 16\pi T S$, where $T$ is the temperature associated with the null surface and $S$ is the associated entropy, which for general relativity is simply $\text{(Area/4)}$ [21, 28]. As evident from the Vaidya solution considered in Section B and also from the general calculation presented in Appendix F it follows that to first order, under the assumptions considered here, the temperature does not change and hence the above change in the Noether charge may be interpreted as,

$$
T \delta S = \int d\Sigma_\alpha \delta T^\alpha \beta \xi^\beta
$$

(6.12)

which is the Clausius relation. In brief, this suggests that due to radiation falling into the Killing horizon, the spacetime is perturbed and the Killing vector also ceases to be Killing, rather it becomes a GKV. This perturbation under the appropriate limit yields the Clausius relation. Thus the formalism developed here for the case of radiation provides a close correspondence with the thermodynamic nature of gravity. One may make a similar argument for the case of nonradiative matter with the $\mu = 2$ parameter choice for null surfaces, provided that $d\Sigma_\alpha \xi^\alpha = 0$, or by interpreting the Noether charge in terms of gravitating mass, as suggested in [27].

7. FINAL REMARKS

Killing vectors play a central role in characterizing spacetime symmetries, which are crucial in determining the conserved quantities that can be constructed in a given spacetime geometry. Systems of astrophysical interest are often symmetric only in a first approximation, and the spacetime geometries for such systems are often more accurately
described in terms of a symmetric spacetime background with perturbations (the latter due to gravitational radiation or inhomogeneities and highly dynamical behavior in matter fields) which explicitly break the symmetry of the background. This motivates the perturbative study of the AKE, the solutions of which provide the generalizations of Killing vectors (which we refer to as GKV) appropriate for the perturbed spacetime.

In this article, we have examined in detail the construction and behavior of these GKV as perturbative solutions to the AKE, associated with the metric perturbations of vacuum and non-vacuum spacetimes, which admit a Killing vector field. This has been achieved in two steps — (a) by considering the perturbation of the action yielding the AKE and then varying the same with respect to the perturbed GKV and (b) by perturbing the AKE and hence determining the evolution equation of the perturbed GKV. The matching of both of these equations explicitly demonstrates the internal consistency of these results.

Additionally, it turns out that the hyperbolicity and Hamiltonian stability of the perturbed GKV remains unchanged compared to its unperturbed counterpart if the GKV and the metric perturbations are kept, independent. However, we find that in the case where the perturbations of the GKV are sourced by the metric perturbation, the problem of an unbounded Hamiltonian can be avoided at first order, and at second order, the problem may also be avoided if the metric perturbations are transverse and traceless (assuming the perturbations remain well-behaved), and perturbative solutions to the $\mu = 2$ AKE exist to second order. We find that the first-order equations trivialize (they reduce to the background AKE) for traceless metric perturbations; for dilute radiation, the second-order case is necessary. We have also examined the first-order behavior of the Noether current constructed from a GKV and its associated charge.

Intriguingly, it turns out that the conservation of the Noether current introduces additional constraints in the theory, which helps significantly to simplify the evolution equation for the perturbed GKV. In particular, if the perturbed matter energy-momentum tensor is traceless, i.e., the perturbation is due to null matter field, it follows that the first-order perturbation of the GKV can be absorbed within the background Killing vector field. To second order, we find that the second-order perturbation always yields non-trivial modifications to the background Killing vector field.

Finally, the perturbation of a background spacetime respecting Killing symmetry also has interesting thermodynamic implications. In particular, as we have demonstrated, the perturbation of the Noether charge to first order can be expressed as $T\delta S$. This is because to first order in the perturbation, under these assumptions, the surface gravity does not change. This is also apparent from the example of Vaidya spacetime considered in this work, which further corroborates our claims regarding the thermodynamic interpretation for the perturbed Noether charge and currents associated with the generalized Killing vector fields.

ACKNOWLEDGMENTS

We are grateful to Edgar Gasperin and David Hilditch for helpful discussions. Some of the calculations were performed using the xAct package [29] for Mathematica. JCF acknowledges support from FCT Grant Nos. PTDC/MAT-APL/30043/2017 and UIDB/00099/2020. Research of SC is funded by the INSPIRE Faculty fellowship from DST, Government of India (Reg. No. DST/INSPIRE/04/2018/000893) and by the Start-Up Research Grant from SERB, DST, Government of India (Reg. No. SRG/2020/000409).

Appendix A: Variational principle for AKE

Here, we review the derivation of the AKE from a variational principle, following the notation and conventions of Section A. We rewrite here the action given in Eq. (2.3) in expanded form:

$$ A[\chi] = \int_{\mathcal{V}} d^4x \sqrt{-g} \left[ -\frac{1}{4} \left( \nabla_\mu \chi_\nu + \nabla_\nu \chi_\mu \right) \left( \nabla^\mu \chi^\nu + \nabla^\nu \chi^\mu \right) + \frac{\bar{\mu}}{2} \left( \nabla_\mu \chi^\mu \right)^2 \right], $$

(A.1)

where $\bar{g}_{\alpha\beta}$ is the metric of the perturbed spacetime and $\bar{\mu} \equiv \bar{\mu}(x)$ is an arbitrary function of the spacetime coordinates. Varying the above action with respect to arbitrary variations of $\chi^\mu$, including endpoint contributions, we obtain,

$$ \Delta A = \int_{\mathcal{V}} d^4x \sqrt{-\bar{g}} \left[ -\bar{g}_{\alpha\beta} \bar{g}_{\mu\nu} \left( \nabla^\mu \chi^\nu + \nabla^\nu \chi^\mu \right) \left( \nabla_\alpha \Delta_\beta + \bar{\mu} \nabla_\alpha \chi_\beta \right) + \bar{\mu} \left( \nabla_\alpha \chi_\beta \right) \left( \nabla_\alpha \Delta_\beta \right) \right] $$

$$ + \int_{\partial\mathcal{V}} d^3x \sqrt{-\bar{g}} \left[ -\frac{1}{4} \left( \nabla_\mu \chi_\nu + \nabla_\nu \chi_\mu \right) \left( \nabla^\mu \chi^\nu + \nabla^\nu \chi^\mu \right) + \frac{\bar{\mu}}{2} \left( \nabla_\mu \chi^\mu \right)^2 \right] \Delta x^\alpha \nabla_\alpha \phi , $$

(A.2)

where $\partial\mathcal{V}$ is the boundary surface of the full spacetime volume $\mathcal{V}$, described by some arbitrary scalar function, $\phi(x) = \text{constant}$. By performing integration by parts, the above expression for the variation of the action can be
further simplified and it will yield several boundary terms. Since these boundary terms will not play any significant role immediately, we will neglect all the boundary contributions and hence the above variation of the action functional can be expressed in the following manner,

\[
\Delta A = \int d^4x \sqrt{-g} \left[ \Delta \chi^\alpha \bar{g}_{\alpha\beta} \left( \Box \chi^\alpha + [\nabla_\mu, \nabla_\nu] \chi^\mu + \nabla_\nu \nabla_\mu \chi^\alpha \right) - \Delta \chi^\beta \nabla_\beta \left( \bar{\mu} \nabla_\sigma \chi^\sigma \right) \right]
\]

\[
= \int d^4x \sqrt{-\bar{g}} \bar{g}_{\alpha\beta} \Delta \chi^\beta \left[ \Box \chi^\alpha + \bar{R}^\nu_\mu \chi^\mu + \nabla_\nu \left\{ (1 - \bar{\mu}) \nabla_\sigma \chi^\sigma \right\} \right]. \tag{A.3}
\]

Here we have used the fact that the commutator of covariant derivatives acting on a vector is given by the Riemann tensor. Thus, setting the variation of the action functional \( \Delta A \), to be zero, for arbitrary variations of the GKV field \( \chi^\beta \), we obtain,

\[
\sqrt{-\bar{g}} \bar{g}_{\alpha\beta} \left[ \Box \chi^\alpha + \bar{R}^\nu_\mu \chi^\mu + \nabla_\nu \left\{ (1 - \bar{\mu}) \nabla_\sigma \chi^\sigma \right\} \right] = 0 . \tag{A.4}
\]

Since we are interested in non-degenerate spacetime, i.e., spacetimes with a metric \( \bar{g}_{\alpha\beta} \) with non-zero determinant and non-trivial inverse, the above equation can be casted in the following form,

\[
\Box \chi^\alpha + \bar{R}^\nu_\mu \chi^\mu + \nabla_\nu \left\{ (1 - \bar{\mu}) \nabla_\sigma \chi^\sigma \right\} = 0 . \tag{A.5}
\]

The above equation corresponds to the AKE satisfied by the GKV \( \chi^\alpha \) in the exact spacetime, with metric \( \bar{g}_{\mu\nu} \). Note that, in the above expression we have kept \( \bar{\mu} \) inside the derivative terms since it is a function of the spacetime coordinates.

### Appendix B: Perturbed action for the GKV

In this section, we will present the detailed derivation of the perturbed action for the GKV field. To reiterate, we have a background spacetime \( \bar{g}_{\mu\nu} \), which has a Killing vector field \( \xi^\alpha \). Due to some perturbation, possibly corresponding to some dilute matter field, the metric and GKV perturbations are respectively described according to \( \bar{g}_{\mu\nu} = g_{\mu\nu} + h_{\mu\nu} \) and \( \chi^\alpha = \xi^\alpha + \delta \xi^\alpha \), expanding to first order in \( h_{\mu\nu} \) and second order in \( \delta \xi^\alpha \). The aim here is to determine the evolution equation for \( \delta \xi^\alpha \). As a first step, we have the following expression:

\[
\nabla_\alpha \chi^\beta = \partial_\alpha \chi^\beta + \hat{\Gamma}^\beta_{\alpha\rho} \chi^\rho = \nabla_\alpha \chi^\beta + \delta \Gamma^\beta_{\alpha\rho} \chi^\rho . \tag{B.1}
\]

Thus we obtain from Eq. (3.3) the following expansion of the Lagrangian for the GKV \( \chi^\alpha \) in the perturbed metric \( h_{\mu\nu} \),

\[
\bar{L} = -\frac{1}{4} [(g_{\nu\alpha} - h_{\nu\alpha}) (\nabla_\mu \chi^\alpha + \delta \Gamma^\alpha_{\mu\rho} \chi^\rho) + (g_{\mu\alpha} + h_{\mu\alpha}) (\nabla_\nu \chi^\alpha + \delta \Gamma^\alpha_{\nu\rho} \chi^\rho)] \\
\quad \times \left[ (g^{\mu\beta} - h^{\mu\beta}) (\nabla_\beta \chi^\nu + \delta \Gamma^\nu_{\beta\rho} \chi^\rho) + (g^{\nu\beta} - h^{\nu\beta}) (\nabla_\beta \chi^\mu + \delta \Gamma^\mu_{\beta\rho} \chi^\rho) \right] \\
\quad + \bar{\mu} \left( \nabla_\mu \chi^\mu + \delta \Gamma^\mu_{\mu\rho} \chi^\rho \right)^2 \\
\quad = -\frac{1}{4} [(g_{\nu\alpha} + h_{\nu\alpha}) \nabla_\mu \chi^\alpha + g_{\nu\alpha} \delta \Gamma^\alpha_{\mu\rho} \chi^\rho + (g_{\mu\alpha} + h_{\mu\alpha}) \nabla_\nu \chi^\alpha + g_{\mu\alpha} \delta \Gamma^\alpha_{\nu\rho} \chi^\rho] \\
\quad \times \left[ (g^{\mu\beta} - h^{\mu\beta}) \nabla_\beta \chi^\nu + g^{\mu\beta} \delta \Gamma^\nu_{\beta\rho} \chi^\rho + (g^{\nu\beta} - h^{\nu\beta}) \nabla_\beta \chi^\mu + g^{\nu\beta} \delta \Gamma^\mu_{\beta\rho} \chi^\rho \right] \\
\quad + \bar{\mu} \left[ (\nabla_\mu \chi^\mu)^2 + 2 \delta \Gamma^\mu_{\mu\rho} \chi^\nu (\nabla_\rho \chi^\nu) \right] + \mathcal{O}(h^2) . \tag{B.2}
\]

To simplify the above expression further, we may use the following identities,

\[
g_{\nu\alpha} \delta \Gamma^\alpha_{\mu\rho} \chi^\rho + g_{\mu\alpha} \delta \Gamma^\alpha_{\nu\rho} \chi^\rho \\
= \frac{1}{4} g_{\nu\alpha} g^{\alpha\beta} \chi^\rho [-\nabla_\beta h_{\mu\rho} + \nabla_\mu h_{\beta\rho} + \nabla_\rho h_{\mu\beta}] + \frac{1}{2} g_{\mu\alpha} g^{\alpha\beta} \chi^\rho [-\nabla_\beta h_{\nu\rho} + \nabla_\nu h_{\beta\rho} + \nabla_\rho h_{\nu\beta}] \\
= \frac{1}{2} \chi^\rho [-\nabla_\nu h_{\mu\rho} + \nabla_\mu h_{\nu\rho} + \nabla_\rho h_{\mu\nu}] + \frac{1}{2} \chi^\rho [-\nabla_\mu h_{\nu\rho} + \nabla_\nu h_{\mu\rho} + \nabla_\rho h_{\mu\nu}] = \chi^\rho \nabla_\rho h_{\mu\nu} , \tag{B.3}
\]
as well as,
\[ g^\mu_\beta \delta \Gamma^\nu_\sigma \chi^\sigma + g^{\nu_\beta} \delta \Gamma^\mu_\sigma \chi^\sigma \]
\[ = \frac{1}{2} g^{\mu_\sigma} g^{\nu_\alpha} \chi^\sigma \left[ -\nabla_\alpha h_{\beta\sigma} + \nabla_{\beta\sigma} h_{\alpha\sigma} + \nabla_\sigma h_{\alpha\beta} \right] + \frac{1}{2} g^{\nu_\beta} g^{\mu_\alpha} \chi^\sigma \left[ -\nabla_\alpha h_{\beta\sigma} + \nabla_{\beta\sigma} h_{\alpha\sigma} + \nabla_\sigma h_{\alpha\beta} \right] \]
\[ = \frac{1}{2} g^{\mu_\beta} g^{\nu_\alpha} \chi^\sigma \left[ -\nabla_\alpha h_{\beta\sigma} + \nabla_{\beta\sigma} h_{\alpha\sigma} + \nabla_\sigma h_{\alpha\beta} \right] + \frac{1}{2} g^{\mu_\beta} g^{\nu_\alpha} \chi^\sigma \left[ -\nabla_\beta h_{\alpha\sigma} + \nabla_{\alpha\sigma} h_{\beta\alpha} + \nabla_\sigma h_{\beta\alpha} \right] \]
\[ = \chi^\sigma \nabla_\sigma h^{\mu_\nu} , \quad \text{(B.4)} \]
and finally,
\[ \delta \Gamma^\mu_\nu \chi^\nu = \frac{1}{2} \chi^\mu g^{\mu_\alpha} \left[ -\nabla_\alpha h_{\mu\nu} + \nabla_\nu h_{\alpha\mu} \right] = \frac{1}{2} \chi^\alpha \nabla_\alpha h . \quad \text{(B.5)} \]

Using these identities we have derived Eq. (3.14) in the main text. We decompose the Lagrangian into terms quadratic in the derivative of the perturbation \( \delta \xi^\alpha \), terms linear in the derivative of the perturbation \( \delta \xi^\alpha \) and terms independent of the derivative of the perturbation \( \delta \xi^\alpha \). The quadratic terms from Eq. (3.15) are

**Quadratic Terms**
\[ \frac{\mu}{2} \left( \nabla_\mu \delta \xi^\alpha \right)^2 - \frac{1}{2} \left( g_{\nu\alpha} + h_{\nu\alpha} \right) \nabla_\mu \delta \xi^\alpha + \frac{1}{2} \left[ \nabla_\mu \delta \xi^\alpha - h^{\mu_\beta} \nabla_\beta \delta \xi^\nu \right] \times \left[ \nabla_\nu \delta \xi^\mu - h^{\nu_\beta} \nabla_\beta \delta \xi^\mu \right] \]
\[ = \frac{1}{2} \left[ \mu \delta^{\alpha_\beta} \beta_\delta - g_{\alpha_\beta} g^{\nu_\mu} - \delta^{\nu_\beta} g^{\nu_\alpha} \right] \left( \nabla_\mu \delta \xi^\alpha \nabla_\nu \delta \xi^\beta \right) + \frac{\mu}{2} \delta^{\mu_\beta} \beta_\delta \left( \nabla_\mu \delta \xi^\alpha \nabla_\nu \delta \xi^\beta \right) \]
\[ = \frac{1}{2} \left[ \mu + \delta \nu \right] \delta^{\nu_\mu} \beta_\delta - g_{\alpha_\beta} g^{\nu_\mu} - \delta^{\nu_\beta} g^{\nu_\alpha} - \left( \delta^{\nu_\alpha} + h_{\nu\alpha} \right) \left( \delta^{\mu_\beta} - h^{\mu_\beta} \right) \left( \nabla_\mu \delta \xi^\alpha \nabla_\nu \delta \xi^\beta \right) . \quad \text{(B.6)} \]
which has been used to derive Eq. (3.16) in the main text. Similarly the linear terms become,

**Linear Terms**
\[ \frac{\mu}{2} \left[ \left( \xi^\alpha + \delta \xi^\alpha \right) \nabla_\alpha h \left( \nabla_\mu \delta \xi^\mu \right) \right] \]
\[ - \frac{1}{2} \left[ g_{\nu\alpha} + h_{\nu\alpha} \right] \nabla_\mu \delta \xi^\alpha + \frac{1}{2} \left[ \nabla_\mu \delta \xi^\alpha - h^{\mu_\beta} \nabla_\beta \delta \xi^\nu \right] \times \left[ \nabla_\nu \delta \xi^\mu - h^{\nu_\beta} \nabla_\beta \delta \xi^\nu \right] \]
\[ = \frac{1}{2} \left[ \left( \xi^\alpha + \delta \xi^\alpha \right) \nabla_\alpha h \left( \nabla_\mu \delta \xi^\mu \right) \right] \]
\[ - \frac{1}{2} \left[ g_{\nu\alpha} + h_{\nu\alpha} \right] \nabla_\mu \delta \xi^\alpha - \frac{1}{2} \left[ \nabla_\mu \delta \xi^\alpha + h_{\nu\mu} \nabla_\nu \delta \xi^\mu \right] \times \left[ \nabla_\nu \delta \xi^\mu - h^{\nu_\beta} \nabla_\beta \delta \xi^\nu \right] \]
\[ = \frac{1}{2} \left[ \left( \xi^\alpha + \delta \xi^\alpha \right) \nabla_\alpha h \left( \nabla_\mu \delta \xi^\mu \right) \right] \]
\[ - \frac{1}{2} \nabla_\mu \delta \xi^\nu \left[ h_{\nu\mu} \nabla_\nu \delta \xi^\alpha + h_{\nu\mu} \nabla_\nu \delta \xi^\alpha + \left( \xi^\mu + \delta \xi^\mu \right) \nabla_\rho h_{\nu\mu} \right] + \mathcal{O}(h^2) . \quad \text{(B.7)} \]

and the term independent of the derivatives of \( \delta \xi^\alpha \) reads,

**No Derivative Terms**
\[ - \frac{1}{2} \left[ h_{\nu\mu} \nabla_\nu \xi^\alpha + h_{\nu\mu} \nabla_\nu \xi^\alpha \left( \xi^\nu + \delta \xi^\nu \right) \nabla_\rho h_{\nu\mu} \right] \times \left[ - h^{\mu_\beta} \nabla_\beta \delta \xi^\nu + \frac{1}{2} \left( \xi^\sigma + \delta \xi^\sigma \right) \nabla_\sigma h^{\mu_\nu} \right] \]
\[ = \mathcal{O}(h^2) . \quad \text{(B.8)} \]

These results have been quoted in the main text. The contribution to the action arising out of the quadratic term presented in Eq. (3.16) becomes,

\[ \sqrt{-g} \text{ Quadratic Term} \]
\[ = \frac{1}{2} \sqrt{-g} \left( 1 + \frac{1}{2} h \right) \]
where the last relation defines the terms $Q_1, Q_2, Q_3$ and $Q_4$ respectively. The variation of each these terms for a variation of the perturbed Killing vector takes the following form,

$$
\Delta Q_1 = \frac{1}{2} \sqrt{-g} \left( \mu \delta_\alpha^\mu \delta_\beta^\nu - g_{\alpha\beta} g^{\mu\nu} - \delta_\alpha^\mu \delta_\beta^\nu \right) \left[ \nabla_\mu \Delta (\delta^\alpha_\xi) \nabla_\nu \delta^\beta_\xi + \nabla_\mu \delta^\alpha_\xi \nabla_\nu \Delta (\delta^\beta_\xi) \right]
$$

$$
= \nabla_\mu \Delta (\delta^\alpha_\xi) \nabla_\nu \delta^\beta_\xi + \nabla_\mu \delta^\alpha_\xi \nabla_\nu \Delta (\delta^\beta_\xi)
$$

$$
= \text{Tot. Derv.} - \sqrt{-g} \Delta (\delta^\alpha_\xi) \left( \mu \delta_\alpha^\mu \delta_\beta^\nu - g_{\alpha\beta} g^{\mu\nu} - \delta_\alpha^\mu \delta_\beta^\nu \right) \nabla_\mu \nabla_\nu \delta^\beta_\xi
$$

$$
= \text{Tot. Derv.} - \sqrt{-g} \Delta (\delta^\alpha_\xi) \left( \mu g_{\alpha\beta} \nabla^\beta (\nabla_\nu \delta^\alpha_\xi) - g_{\alpha\beta} \nabla^\beta \nabla_\nu \delta^\alpha_\xi - \nabla_\beta \nabla_\alpha \delta^\beta_\xi \right)
$$

$$
= \text{Tot. Derv.} - \sqrt{-g} \Delta (\delta^\alpha_\xi) \left( \mu g_{\alpha\beta} \nabla^\beta (\nabla_\nu \delta^\alpha_\xi) - g_{\alpha\beta} \nabla^\beta \nabla_\nu \delta^\alpha_\xi - \nabla_\beta \nabla_\alpha \delta^\beta_\xi \right)
$$

$$
= \text{Tot. Derv.} + \frac{1}{2} \sqrt{-g} \Delta (\delta^\alpha_\xi) \left( (1 - \mu) g_{\alpha\beta} \nabla^\beta (\nabla_\nu \delta^\alpha_\xi) + g_{\alpha\beta} \nabla^\beta \delta^\beta_\xi + R_{\alpha\beta} \delta^\beta_\xi \right),
$$

and,

$$
\Delta Q_2 = \frac{1}{2} \sqrt{-g} h \left( \mu \delta_\alpha^\mu \delta_\beta^\nu - g_{\alpha\beta} g^{\mu\nu} - \delta_\alpha^\mu \delta_\beta^\nu \right) \left[ \nabla_\mu \Delta (\delta^\alpha_\xi) \nabla_\nu \delta^\beta_\xi + \nabla_\mu \delta^\alpha_\xi \nabla_\nu \Delta (\delta^\beta_\xi) \right]
$$

$$
= \nabla_\mu \Delta (\delta^\alpha_\xi) \nabla_\nu \delta^\beta_\xi + \nabla_\mu \delta^\alpha_\xi \nabla_\nu \Delta (\delta^\beta_\xi)
$$

$$
= \text{Tot. Derv.} - \frac{1}{2} \sqrt{-g} \Delta (\delta^\alpha_\xi) h \left( \mu \delta_\alpha^\mu \delta_\beta^\nu - g_{\alpha\beta} g^{\mu\nu} - \delta_\alpha^\mu \delta_\beta^\nu \right) \nabla_\mu \nabla_\nu \delta^\beta_\xi
$$

$$
= \text{Tot. Derv.} - \frac{1}{2} \sqrt{-g} \Delta (\delta^\alpha_\xi) h \left( (1 - \mu) g_{\alpha\beta} \nabla^\beta (\nabla_\nu \delta^\alpha_\xi) + g_{\alpha\beta} \nabla^\beta \delta^\beta_\xi + R_{\alpha\beta} \delta^\beta_\xi \right)
$$

$$
= \text{Tot. Derv.} + \frac{1}{2} \sqrt{-g} \Delta (\delta^\alpha_\xi) \left( \mu \delta_\alpha^\mu \delta_\beta^\nu - g_{\alpha\beta} g^{\mu\nu} - \delta_\alpha^\mu \delta_\beta^\nu \right) \nabla_\nu \delta^\beta_\xi,
$$

and,

$$
\Delta Q_3 = \frac{1}{2} \sqrt{-g} \left( 1 + \frac{1}{2} h \right) \delta_\alpha^\mu \delta_\beta^\nu \left[ \nabla_\mu \Delta (\delta^\alpha_\xi) \nabla_\nu \delta^\beta_\xi + \nabla_\mu \delta^\alpha_\xi \nabla_\nu \Delta (\delta^\beta_\xi) \right]
$$

$$
= \nabla_\mu \Delta (\delta^\alpha_\xi) \nabla_\nu \delta^\beta_\xi + \nabla_\mu \delta^\alpha_\xi \nabla_\nu \Delta (\delta^\beta_\xi)
$$

$$
= \text{Tot. Derv.} - \sqrt{-g} \Delta (\delta^\alpha_\xi) \left( 1 + \frac{1}{2} h \right) \delta_\alpha^\mu \delta_\beta^\nu \nabla_\mu \nabla_\nu \delta^\beta_\xi - \sqrt{-g} \Delta (\delta^\alpha_\xi) \nabla_\mu \left( 1 + \frac{1}{2} h \right) \delta_\alpha^\mu \delta_\beta^\nu \nabla_\nu \delta^\beta_\xi
$$

$$
= \text{Tot. Derv.} - \sqrt{-g} \Delta (\delta^\alpha_\xi) \left( 1 + \frac{1}{2} h \right) \delta_\alpha^\mu \nabla_\alpha \nabla_\beta \delta^\beta_\xi - \frac{1}{2} \sqrt{-g} \Delta (\delta^\alpha_\xi) \nabla_\alpha h \delta_\mu \nabla_\beta \delta^\beta_\xi,
$$

and finally,

$$
\Delta Q_4 = \frac{1}{2} \sqrt{-g} \left( g_{\alpha\beta} h^{\mu\nu} - h_{\alpha\beta} g^{\mu\nu} + \delta^\mu_\alpha h^\nu_\beta - \delta^\nu_\alpha h^\mu_\beta \right) \left[ \nabla_\mu \Delta (\delta^\alpha_\xi) \nabla_\nu \delta^\beta_\xi + \nabla_\nu \delta^\alpha_\xi \nabla_\mu \Delta (\delta^\beta_\xi) \right]
$$

$$
= \sqrt{-g} \left( g_{\alpha\beta} h^{\mu\nu} - h_{\alpha\beta} g^{\mu\nu} \right) \nabla_\mu \Delta (\delta^\alpha_\xi) \nabla_\nu \delta^\beta_\xi
$$

$$
= \text{Tot. Derv.} - \sqrt{-g} \Delta (\delta^\alpha_\xi) \nabla_\mu \left[ (g_{\alpha\beta} h^{\mu\nu} - h_{\alpha\beta} g^{\mu\nu}) \nabla_\nu \delta^\beta_\xi \right]
$$
\[ \text{Tot. Derv. } - \sqrt{-g} \Delta (\delta \xi^\alpha) \left[ (g_{\alpha\beta} h^{\mu\nu} - h_{\alpha\beta} g^{\mu\nu}) \nabla_\mu \nabla_\nu \delta \xi^\beta + \nabla_\nu \delta \xi^\beta \right] (g_{\alpha\beta} \nabla_\mu h^{\mu\nu} - g^{\mu\nu} \nabla_\mu h_{\alpha\beta}) \]

\[ \text{Tot. Derv. } + \sqrt{-g} \Delta (\delta \xi^\alpha) \left[ -g_{\alpha\beta} h^{\mu\nu} \nabla_\mu \nabla_\nu \delta \xi^\beta + h_{\alpha\beta} \Box \delta \xi^\beta - g_{\alpha\beta} \left( \nabla_\nu \delta \xi^\beta \right) \nabla_\mu h^{\mu\nu} + \left( \nabla^\mu \delta \xi^\beta \right) \nabla_\mu h_{\alpha\beta} \right]. \]

(B.13)

The linear term is already linear in the perturbation \( h_{\alpha\beta} \). Thus contribution to the action can be further simplified, yielding,

\[ \sqrt{-g} \text{ Linear Terms} = \sqrt{-g} \left[ \frac{\mu^2}{2} \left( (\xi^\alpha + \delta \xi^\alpha) \nabla_\alpha \nabla_\mu (\nabla_\mu \delta \xi^\mu) \right) + \frac{1}{2} \sqrt{-g} \left[ g_{\alpha\nu} \nabla_\mu \delta \xi^\alpha + g_{\mu\sigma} \nabla_\nu \delta \xi^\sigma \right] h_{\mu\beta} \nabla_\beta \xi^\nu - \frac{1}{4} \sqrt{-g} \left[ g_{\sigma \nu} \nabla_\mu \delta \xi^\sigma + g_{\mu \nu} \nabla_\sigma \delta \xi^\nu \right] \left( \xi^\alpha + \delta \xi^\alpha \right) \nabla_\sigma h^{\mu\nu} - \frac{1}{2} \sqrt{-g} \left( \nabla^{\mu} \delta \xi^\nu \left( \xi^\sigma + \delta \xi^\sigma \right) \nabla_\sigma h^{\mu\nu} - \frac{1}{2} \sqrt{-g} \left( \nabla^{\mu} \delta \xi^\nu \left( \xi^\sigma + \delta \xi^\sigma \right) \nabla_\sigma h^{\mu\nu} \right) \nabla_\mu h_{\nu\mu} \right] \equiv L_1 + L_2 + L_3, \]

(B.14)

where the last line defines the three terms \( L_1, L_2 \) and \( L_3 \) respectively. Further simplification yields,

\[ L_1 = \sqrt{-g} \left[ \frac{\mu^2}{2} \left( (\xi^\alpha + \delta \xi^\alpha) \nabla_\alpha \nabla_\mu (\nabla_\mu \delta \xi^\mu) \right) \right], \]

(B.15)

\[ L_2 = \frac{1}{2} \sqrt{-g} \left[ g_{\alpha\nu} \nabla_\mu \delta \xi^\alpha + g_{\mu\sigma} \nabla_\nu \delta \xi^\sigma \right] h_{\mu\beta} \nabla_\beta \xi^\nu - \frac{1}{2} \sqrt{-g} \left[ g_{\sigma \nu} \nabla_\mu \delta \xi^\sigma + g_{\mu \nu} \nabla_\sigma \delta \xi^\nu \right] \left( \xi^\alpha + \delta \xi^\alpha \right) \nabla_\sigma h^{\mu\nu} \]

(B.16)

\[ L_3 = -\frac{1}{4} \sqrt{-g} \left[ g_{\alpha\nu} \nabla_\mu \delta \xi^\alpha + g_{\mu\sigma} \nabla_\nu \delta \xi^\sigma \right] \left( \xi^\sigma + \delta \xi^\sigma \right) \nabla_\sigma h^{\mu\nu} = -\frac{1}{2} \sqrt{-g} \left( \nabla^{\mu} \delta \xi^\nu \left( \xi^\sigma + \delta \xi^\sigma \right) \nabla_\sigma h^{\mu\nu} \right) \nabla_\mu h_{\nu\mu} \]

(B.17)

Under the variation, each of these terms yield,

\[ \Delta L_1 = \sqrt{-g} \left[ \frac{\mu^2}{2} \left( \Delta (\delta \xi^\alpha) \nabla_\alpha \nabla_\mu (\nabla_\mu \delta \xi^\mu) \right) + \sqrt{-g} \left[ \frac{\mu^2}{2} \left( (\xi^\alpha + \delta \xi^\alpha) \nabla_\alpha \nabla_\mu \Delta (\delta \xi^\mu) \right) \right] \]

(B.18)

\[ \Delta L_2 = -\sqrt{-g} \left[ \frac{\mu^2}{2} \left( \nabla_\mu \Delta (\delta \xi^\alpha) \right) h_{\mu\beta} \nabla_\beta \xi^\mu + h^{\mu\beta} \nabla_\alpha \xi^\beta \right] \]

(B.19)

\[ \Delta L_3 = -\sqrt{-g} \left[ g_{\alpha\nu} \nabla_\mu \Delta (\delta \xi^\alpha) \left( \xi^\sigma + \delta \xi^\sigma \right) \nabla_\sigma h^{\mu\nu} - \sqrt{-g} \left[ g_{\sigma \nu} \left( \nabla_\mu \delta \xi^\sigma \right) \Delta (\delta \xi^\alpha) \nabla_\sigma h^{\mu\nu} \right] \right] \]

(B.20)

These expressions for the quadratic and linear derivative terms of \( \delta \xi^\alpha \) have been used to express the Lagrangian as in Eq. (3.18) in the main text. We have now derived all the necessary equations in order to determine the variation of the Lagrangian for the perturbed Killing vector \( \delta \xi^\alpha \), presented in Eq. (3.18). To start with, we would like to point out one identity, which will become helpful for our later purposes. This reads,

\[ \nabla_\mu \nabla_\sigma h^{\mu\nu} = \nabla_\sigma \nabla_\mu h^{\mu\nu} + \left[ \nabla_\mu, \nabla_\sigma \right] h^{\mu\nu} = \frac{1}{2} \nabla_\sigma \nabla^\nu h + R_{\rho \mu \alpha} h^{\rho\nu} + R_{\rho \mu \sigma} h^{\rho\nu} = \frac{1}{2} \nabla_\sigma \nabla^\nu h + R_{\rho \mu \sigma} h^{\rho\nu} - R_{\rho \sigma \mu} h^{\rho\nu} \]

(B.21)
At the first step we combine two terms from the quadratic part of the action and two terms from the linear part of the action. This leads to the following expression,

\[ \Delta \left( Q_1 + Q_2 + L_1 + L_3 \right) \]

\[ = \sqrt{-g} \Delta (\delta \xi^\alpha) \left[ (1 - \mu) \, g_{\alpha \beta} \nabla^\beta (\nabla_\nu \delta \xi^\nu) + g_{\alpha \beta} \nabla^\beta - g_{\alpha \beta} h^{\mu \nu} \nabla_\mu \nabla_\nu \delta \xi^\beta \right. \]

\[ + h_{\alpha \beta} \nabla^\beta - g_{\alpha \beta} (\nabla_\nu \delta \xi^\beta) \nabla_\mu h^{\mu \nu} + (\nabla^\beta \delta \xi^\beta) \nabla_\mu h^{\mu \alpha} + \frac{\mu}{2} \nabla_\alpha (\nabla_\mu \delta \xi^\mu) - \nabla_\alpha \{ (\xi^\rho + \delta \xi^\rho) \nabla_\rho \}
\]

\[ - (\nabla^\beta \delta \xi^\beta) \nabla_\alpha h_{\mu \beta} + (\nabla^\beta \delta \xi^\beta) \nabla_\sigma h_{\mu \alpha} + \nabla^\beta \delta \xi^\beta \nabla_\sigma h_{\mu \alpha} \]

\[ + (\xi^\sigma + \delta \xi^\sigma) \nabla_\alpha \left( \frac{1}{2} \nabla_\sigma \nabla^\nu h + R_{\rho \sigma \mu} h^{\rho \nu} - R_{\rho \sigma \mu} h^{\rho \nu} \right) \]

\[ = \sqrt{-g} \Delta (\delta \xi^\alpha) \left[ (1 - \mu) \, g_{\alpha \beta} \nabla^\beta (\nabla_\nu \delta \xi^\nu) + g_{\alpha \beta} \nabla^\beta - g_{\alpha \beta} h^{\mu \nu} \nabla_\mu \nabla_\nu \delta \xi^\beta \right. \]

\[ + (\nabla^\beta \delta \xi^\beta) \nabla_\mu h_{\alpha \beta} + (\nabla^\beta \delta \xi^\beta) \nabla_\mu h_{\alpha \beta} - (\nabla^\beta \delta \xi^\beta) \nabla_\sigma h_{\alpha \beta} + \frac{\mu}{2} \nabla_\alpha \{ (\xi^\rho + \delta \xi^\rho) \nabla_\rho \}
\]

\[ + h_{\alpha \beta} \nabla^\beta - g_{\alpha \beta} (\nabla_\nu \delta \xi^\beta) \nabla_\mu h^{\mu \nu} + (\nabla^\beta \delta \xi^\beta) \nabla_\sigma h_{\mu \alpha} + \frac{\mu}{2} \nabla_\alpha (\nabla_\mu \delta \xi^\mu) \]

\[ + (\xi^\sigma + \delta \xi^\sigma) \nabla_\alpha \left( \frac{1}{2} \nabla_\sigma \nabla^\nu h + R_{\rho \sigma \mu} h^{\rho \nu} - R_{\rho \sigma \mu} h^{\rho \nu} \right) \]

As we can see, the terms appearing in the first line are recognizable if we keep in mind the AKE for the vector \( \xi^\mu + \delta \xi^\mu \).

Upon adding the variation of the term \( Q_2 \) from the quadratic expression, we obtain,

\[ \Delta \left( Q_1 + Q_2 + Q_4 + L_1 + L_3 \right) \]

\[ = \sqrt{-g} \Delta (\delta \xi^\alpha) \left[ (1 + \frac{1}{2} \xi) \left( g_{\alpha \beta} + h_{\alpha \beta} \right) \left\{ (1 - \mu) \nabla^\beta (\nabla_\nu \delta \xi^\nu) + \nabla^\beta + R_{\rho \sigma \mu} h^{\rho \nu} \right\} \right. \]

\[ + 2 \left( \nabla^\beta \delta \xi^\beta \right) \nabla_\mu h_{\alpha \beta} - \frac{1}{2} \nabla_\alpha h_{\mu \beta} \nabla_\sigma h_{\mu \alpha} - \frac{1}{2} \nabla_\alpha \{ (\xi^\rho + \delta \xi^\rho) \nabla_\rho \}
\]

\[ - h_{\alpha \beta} \nabla^\beta - g_{\alpha \beta} (\nabla_\nu \delta \xi^\beta) \nabla_\mu h^{\mu \nu} - (\nabla^\beta \delta \xi^\beta) \nabla_\rho h^{\rho \nu}
\]

\[ + h_{\alpha \beta} \nabla^\beta - g_{\alpha \beta} (\nabla_\nu \delta \xi^\beta) \nabla_\mu h^{\mu \nu} + (\nabla^\beta \delta \xi^\beta) \nabla_\sigma h_{\mu \alpha} + \frac{\mu}{2} \nabla_\alpha (\nabla_\mu \delta \xi^\mu) \]

\[ + (\xi^\sigma + \delta \xi^\sigma) \nabla_\alpha \left( \frac{1}{2} \nabla_\sigma \nabla^\nu h + R_{\rho \sigma \mu} h^{\rho \nu} - R_{\rho \sigma \mu} h^{\rho \nu} \right) \] .

The above variation can be simplified further, if we use the following identities:

\[ - g_{\alpha \beta} (\nabla_\nu \delta \xi^\beta) \nabla_\mu h^{\mu \nu} + \frac{1}{2} (\nabla_\mu h) \, g_{\alpha \beta} g^{\mu \nu} \nabla_\nu \delta \xi^\beta = - \frac{1}{2} g_{\alpha \beta} (\nabla_\nu \delta \xi^\beta) \nabla_\mu h^{\mu \nu} + \frac{1}{2} (\nabla_\mu h) \, g_{\alpha \beta} g^{\mu \nu} \nabla_\nu \delta \xi^\beta = 0 \] ,

\[ \frac{\mu}{2} \nabla_\alpha h (\nabla_\mu \delta \xi^\mu) - \frac{1}{2} (\nabla_\mu h) \, \mu \, \delta_{\alpha \beta} \delta_{\nu \sigma} \nabla_\nu \delta \xi^\beta = \frac{\mu}{2} \nabla_\alpha h (\nabla_\mu \delta \xi^\mu) \] .

Using these identities, the variation of some of the quadratic and linear order terms in the action for the perturbed Killing vector becomes,

\[ \Delta \left( Q_1 + Q_2 + Q_4 + L_1 + L_3 \right) \]

\[ = \sqrt{-g} \Delta (\delta \xi^\alpha) \left[ \left( 1 + \frac{1}{2} \xi \right) \left( g_{\alpha \beta} + h_{\alpha \beta} \right) \left\{ (1 - \mu) \nabla^\beta (\nabla_\nu \delta \xi^\nu) + \nabla^\beta + R_{\rho \sigma \mu} h^{\rho \nu} \right\} \right. \]

\[ - g_{\alpha \beta} h^{\mu \nu} \nabla_\mu \nabla_\nu \delta \xi^\beta + 2 \left( \nabla^\beta \delta \xi^\beta \right) \left( \nabla_\rho h_{\alpha \beta} - \frac{1}{2} \nabla_\alpha h_{\mu \beta} \right) - \frac{1}{2} \nabla_\alpha \{ (\xi^\rho + \delta \xi^\rho) \nabla_\rho \} \]
\begin{align}
- h_{\alpha\beta} (1 - \mu) \nabla^\beta (\nabla_\mu \delta \xi^\nu - h_{\alpha\beta} R^\beta_\rho \delta \xi^\rho + (\nabla^\mu \xi^\sigma) \nabla_\sigma h_{\mu\alpha} + \frac{\delta \mu}{2} \nabla h (\nabla_\mu \delta \xi^\mu) + \frac{1}{2} (\nabla_\beta h) \nabla_\alpha \delta \xi^\beta \\
+ (\xi^\sigma + \delta \xi^\sigma) g_{\nu\alpha} \left( \frac{1}{2} \nabla_\sigma \nabla^\nu h + R_{\rho\sigma h^{\rho\nu} - R^{\nu}_{\alpha\rho \mu} h^{\mu\rho}} \right). \tag{B.26}
\end{align}

Again the appearance of the terms in the first line are recognizable from the form of the full AKE. Proceeding further, we incorporate all the contributions from the quadratic order terms, yielding,

\begin{align}
\Delta \left( Q_1 + Q_2 + Q_3 + Q_4 + L_1 + L_3 \right) \\
= \sqrt{-g} \Delta (\delta \xi^\alpha) \left[ \left( 1 + \frac{1}{2} h \right) (g_{\alpha\beta} + h_{\alpha\beta}) \left( 1 - \mu \right) \nabla^\beta (\nabla_\nu \delta \xi^\nu) + \Box \delta \xi^\beta + R^\beta_\rho \delta \xi^\rho - \mu \nabla^\beta \nabla_\sigma \delta \xi^\sigma \right] \\
- g_{\alpha\beta} h^{\mu\nu} \nabla_\mu \nabla_\nu \delta \xi^\beta + 2 \left( \nabla^{(\mu} \delta \xi^{\beta)} \right) \left( \nabla_\mu h_{\alpha\beta} - \frac{1}{2} \nabla h_{\mu\beta} \right) - \frac{\mu + \delta \mu}{2} \nabla \left\{ (\xi^\rho + \delta \xi^\rho) \nabla_\rho h \right\} \\
- h_{\alpha\beta} (1 - \mu) \nabla^\beta (\nabla_\nu \delta \xi^\nu) - h_{\alpha\beta} R^\beta_\rho \delta \xi^\rho + (\nabla^\mu \xi^\sigma) \nabla_\sigma h_{\mu\alpha} + \frac{\delta \mu}{2} \nabla h (\nabla_\mu \delta \xi^\mu) + \frac{1}{2} (\nabla_\beta h) \nabla_\alpha \delta \xi^\beta \\
+ \mu h_{\alpha\beta} \nabla^\beta \nabla_\sigma \delta \xi^\sigma - \frac{1}{2} \nabla h \delta \mu \nabla_\beta \delta \xi^\beta + (\xi^\sigma + \delta \xi^\sigma) g_{\nu\alpha} \left( \frac{1}{2} \nabla_\sigma \nabla^\nu h + R_{\rho\sigma h^{\rho\nu} - R^{\nu}_{\alpha\rho \mu} h^{\mu\rho}} \right) \right] \\
= \sqrt{-g} \Delta (\delta \xi^\alpha) \left[ \left( 1 + \frac{1}{2} h \right) (g_{\alpha\beta} + h_{\alpha\beta}) \left( 1 - \mu \right) \nabla^\beta (\nabla_\nu \delta \xi^\nu) + \Box \delta \xi^\beta + R^\beta_\rho \delta \xi^\rho - \mu \nabla^\beta \nabla_\sigma \delta \xi^\sigma \right] \\
- g_{\alpha\beta} h^{\mu\nu} \nabla_\mu \nabla_\nu \delta \xi^\beta + 2 \left( \nabla^{(\mu} \delta \xi^{\beta)} \right) \left( \nabla_\mu h_{\alpha\beta} - \frac{1}{2} \nabla h_{\mu\beta} \right) - \frac{\mu + \delta \mu}{2} \nabla \left\{ (\xi^\rho + \delta \xi^\rho) \nabla_\rho h \right\} \\
- h_{\alpha\beta} (1 - \mu) \nabla^\beta (\nabla_\nu \delta \xi^\nu) - h_{\alpha\beta} R^\beta_\rho \delta \xi^\rho + (\nabla^\mu \xi^\sigma) \nabla_\sigma h_{\mu\alpha} + \mu h_{\alpha\beta} \nabla^\beta \nabla_\sigma \delta \xi^\sigma + \frac{1}{2} (\nabla_\beta h) \nabla_\alpha \delta \xi^\beta \\
+ (\xi^\sigma + \delta \xi^\sigma) g_{\nu\alpha} \left( \frac{1}{2} \nabla_\sigma \nabla^\nu h + R_{\rho\sigma h^{\rho\nu} - R^{\nu}_{\alpha\rho \mu} h^{\mu\rho}} \right) \right]. \tag{B.27}
\end{align}

The above analysis provides many of the simplifications we are after. It is now time to consider the variation of the full action building from bits and pieces of the above expressions. The full expression for the variation takes the form

\begin{align}
\Delta \left( Q_1 + Q_2 + Q_3 + Q_4 + L_1 + L_2 + L_3 \right) \\
= \sqrt{-g} \Delta (\delta \xi^\alpha) \left[ \left( 1 + \frac{1}{2} h \right) (g_{\alpha\beta} + h_{\alpha\beta}) \left( 1 - \mu \right) \nabla^\beta (\nabla_\nu \delta \xi^\nu) + \Box \delta \xi^\beta + R^\beta_\rho \delta \xi^\rho - \mu \nabla^\beta \nabla_\sigma \delta \xi^\sigma \right] \\
- g_{\alpha\beta} h^{\mu\nu} \nabla_\mu \nabla_\nu \delta \xi^\beta + 2 \left( \nabla^{(\mu} \delta \xi^{\beta)} \right) \left( \nabla_\mu h_{\alpha\beta} - \frac{1}{2} \nabla h_{\mu\beta} \right) - \frac{\mu + \delta \mu}{2} \nabla \left\{ (\xi^\rho + \delta \xi^\rho) \nabla_\rho h \right\} \\
- h_{\alpha\beta} (1 - \mu) \nabla^\beta (\nabla_\nu \delta \xi^\nu) - h_{\alpha\beta} R^\beta_\rho \delta \xi^\rho + (\nabla^\mu \xi^\sigma) \nabla_\sigma h_{\mu\alpha} + \mu h_{\alpha\beta} \nabla^\beta \nabla_\sigma \delta \xi^\sigma + \frac{1}{2} (\nabla_\beta h) \nabla_\alpha \delta \xi^\beta \\
+ \mu h_{\alpha\beta} \nabla^\beta \nabla_\sigma \delta \xi^\sigma - \frac{1}{2} \nabla h \delta \mu \nabla_\beta \delta \xi^\beta + (\xi^\sigma + \delta \xi^\sigma) g_{\nu\alpha} \left( \frac{1}{2} \nabla_\sigma \nabla^\nu h + R_{\rho\sigma h^{\rho\nu} - R^{\nu}_{\alpha\rho \mu} h^{\mu\rho}} \right) \right]. \tag{B.27}
\end{align}
\[- h_{\alpha \beta} R^\beta \partial_\alpha \xi^\rho + h^{\mu \beta} R_{\rho \mu \alpha \beta} \xi^\rho + \nabla_\mu h^{\mu \beta} \nabla_\alpha \xi_\beta + (\xi^\rho + \delta \xi^\rho) \left( \frac{1}{2} \nabla_\rho \nabla_\alpha h + R_{\rho \sigma} h^\rho_{\alpha} - R_{\alpha \rho \mu} h^{\mu \rho} \right) \]  
\[= \sqrt{-g} \Delta (\delta \xi^\alpha) \left[ (1 + \frac{1}{2} \mu) (g_{\alpha \beta} + h_{\alpha \beta}) \left\{ (1 - \mu) \nabla^{\beta} (\nabla_\nu \delta \xi^\nu) + \Box \delta \xi^\beta + R^\beta_{\mu \delta} \delta \xi^\rho - \mu \nabla^{\beta} \nabla_\sigma \delta \xi^\sigma \right\} \right. \]
\[- g_{\alpha \beta} h^{\mu \nu} \nabla_\mu \nabla_\nu \delta \xi^\beta + 2 \left( \nabla^{(\mu} \delta \xi^{\beta)} \right) \left( \nabla_\mu h_{\alpha \beta} - \frac{1}{2} \nabla_\alpha h_{\mu \beta} \right) - \frac{\mu + \delta \mu}{2} \nabla_\alpha \left\{ (\xi^\rho + \delta \xi^\rho) \nabla_\rho h \right\} - R_{\alpha \rho \mu} h^{\mu \rho} \delta \xi^\sigma \]
\[- h_{\alpha \beta} (1 - \mu) \nabla^{\beta} (\nabla_\nu \delta \xi^\nu) + \delta \mu h_{\alpha \beta} \nabla^{\beta} \nabla_\sigma \delta \xi^\sigma \]  
\[= \sqrt{-g} \Delta (\delta \xi^\alpha) \left[ (1 + \frac{1}{2} \mu) (g_{\alpha \beta} + h_{\alpha \beta}) \left\{ (1 - \mu) \nabla^{\beta} (\nabla_\nu \delta \xi^\nu) + \Box \delta \xi^\beta + R^\beta_{\mu \delta} \delta \xi^\rho - \mu \nabla^{\beta} \nabla_\sigma \delta \xi^\sigma \right\} \right. \]
\[- g_{\alpha \beta} h^{\mu \nu} \nabla_\mu \nabla_\nu \delta \xi^\beta + 2 \left( \nabla^{(\mu} \delta \xi^{\beta)} \right) \left( \nabla_\mu h_{\alpha \beta} - \frac{1}{2} \nabla_\alpha h_{\mu \beta} \right) + \frac{1 - \mu - \delta \mu}{2} \nabla_\alpha \left\{ (\xi^\rho + \delta \xi^\rho) \nabla_\rho h \right\} - R_{\alpha \rho \mu} h^{\mu \rho} \delta \xi^\sigma \]
\[- h_{\alpha \beta} (1 - \mu) \nabla^{\beta} (\nabla_\nu \delta \xi^\nu) + \delta \mu h_{\alpha \beta} \nabla^{\beta} \nabla_\sigma \delta \xi^\sigma \]  

This completes the expression for the variation of the action (the integrand, to be precise) and the above expression after neglecting terms $\mathcal{O}(h^2)$ has been presented in Eq. (3.19) in the main text.

### Appendix C: Perturbed almost Killing equation

In this appendix, we will provide an explicit derivation of the perturbed AKE. The perturbation involves — (a) the perturbation of the background spacetime $g_{\mu \nu}$ and (b) the perturbation of the Killing vector $\xi^\alpha$. It is useful to start with the following expressions,

\[\delta \Gamma^{\alpha \mu \nu} = \frac{1}{2} g^{\alpha \beta} \left( - \nabla_\beta h_{\mu \nu} + \nabla_\mu h_{\beta \nu} + \nabla_\nu h_{\beta \mu} \right),\]  
\[\delta R_{\mu \nu} = \frac{1}{2} \left( - \Box h_{\mu \nu} - \nabla_\alpha \nabla_\nu h + \nabla_\alpha \nabla_\mu h^\alpha_{\nu} + \nabla_\mu \nabla_\alpha h^\alpha_{\nu} \right).\]  

We also have the following result,

\[
\nabla_\alpha \nabla_\mu h^\alpha_{\nu} = [\nabla_\alpha, \nabla_\mu] h^\alpha_{\nu} + \nabla_\mu \nabla_\alpha h^\alpha_{\nu} = R_{\beta \mu \beta \nu} - R_{\alpha \mu \beta \nu} h^{\alpha \beta} + \nabla_\mu \nabla_\alpha h^\alpha_{\nu}\]  

The perturbation of the Ricci tensor becomes,

\[\delta R_{\mu \nu} = \frac{1}{2} \left( - \Box h_{\mu \nu} - \nabla_\alpha \nabla_\nu h + R_{\beta \mu \beta \nu} - R_{\alpha \mu \beta \nu} h^{\alpha \beta} + \nabla_\mu \nabla_\alpha h^\alpha_{\nu} + R_{\beta \nu \beta \mu} h^\alpha_{\beta} + \nabla_\nu \nabla_\alpha h^\alpha_{\mu} \right),\]  

which yields Eq. (3.22) upon simplification using the following identity, $R_{\alpha \mu \beta \nu} = R_{\beta \mu \alpha \nu}$. For our future results, we will also need the following identity,

\[
\nabla_\alpha \nabla_\beta V^\mu = \partial_\alpha (\nabla_\beta V^\mu) + \Gamma^{\mu}_{\alpha \rho} \nabla_\beta V^\rho - \Gamma^{\rho}_{\alpha \beta} \nabla_\rho V^\mu \]
\[= \partial_\alpha (\nabla_\beta V^\mu + \Gamma^{\mu}_{\beta \rho} V^\rho) + \Gamma^{\mu}_{\alpha \rho} \left( \partial_\beta V^\rho + \Gamma^{\rho}_{\beta \sigma} V^\sigma \right) - \Gamma^{\rho}_{\alpha \beta} (\partial_\rho V^\mu + \Gamma^{\mu}_{\rho \sigma} V^\sigma) \]
\[= \partial_\alpha (\nabla_\beta V^\mu + \delta \Gamma^{\mu}_{\beta \rho} V^\rho) + \Gamma^{\mu}_{\alpha \rho} \left( \nabla_\beta V^\rho + \delta \Gamma^{\rho}_{\beta \sigma} V^\sigma \right) - \delta \Gamma^{\rho}_{\alpha \beta} (\nabla_\rho V^\mu + \delta \Gamma^{\mu}_{\rho \sigma} V^\sigma) \]
\[= \partial_\alpha (\nabla_\beta V^\mu + \delta \Gamma^{\mu}_{\beta \rho} V^\rho) + \Gamma^{\mu}_{\alpha \rho} \left( \nabla_\beta V^\rho + \delta \Gamma^{\rho}_{\beta \sigma} V^\sigma \right) + \delta \Gamma^{\mu}_{\alpha \beta} \nabla_\beta V^\rho \]
\[- \delta \Gamma^{\rho}_{\alpha \beta} (\nabla_\rho V^\mu + \delta \Gamma^{\mu}_{\rho \sigma} V^\sigma) - \delta \Gamma^{\rho}_{\alpha \beta} \nabla_\rho V^\mu + \Omega(h^2) \]
\[= \nabla_\alpha (\nabla_\beta V^\mu) + \nabla_\alpha \left( \delta \Gamma^{\rho}_{\beta \rho} V^\rho \right) + \delta \Gamma^{\rho}_{\alpha \beta} \nabla_\beta V^\rho - \delta \Gamma^{\rho}_{\alpha \beta} \nabla_\rho V^\mu \]

where in the last line, terms $\mathcal{O}(h^2)$ have been neglected. Thus by appropriate contraction of the above identity we have Eq. (3.23). Further, from Eq. (C.1) it follows that,

\[\delta \Gamma^{\beta}_{\beta \rho} = \frac{1}{2} g^{\beta \sigma} \nabla_\rho h_{\beta \sigma} = \frac{1}{2} \nabla_\rho h.\]  

(C.6)
Thus Eq. (3.23) can also be expressed as,
\[
\nabla_\alpha \nabla_\beta V^\beta = \nabla_\alpha \left( \nabla_\beta V^\beta \right) + \frac{1}{2} \nabla_\alpha \left( V^\rho \nabla_\rho h \right) .
\]  
(C.7)

Similarly from Eq. (C.5) it also follows that,
\[
\Box V^\mu = g^{\alpha\beta} \nabla_\alpha \nabla_\beta V^\mu = \left( g^{\alpha\beta} - h^{\alpha\beta} \right) \nabla_\alpha \nabla_\beta V^\mu
\]
\[
= g^{\alpha\beta} \left[ \nabla_\alpha \left( \nabla_\beta V^\mu \right) + \nabla_\beta \left( \delta^{\mu}_{\rho \nu} V^\rho \right) + \delta^{\mu}_{\rho \beta} \nabla_\beta V^\rho - \delta^{\beta}_{\rho \nu} \nabla_\rho V^\nu \right] - h^{\alpha\beta} \nabla_\alpha \left( \nabla_\beta V^\mu \right)
\]
\[
= \left( g^{\alpha\beta} - h^{\alpha\beta} \right) \nabla_\alpha \nabla_\beta V^\mu + 2 \delta^{\mu}_{\rho \beta} \nabla_\beta V^\rho - g^{\alpha\beta} \delta^{\rho}_{\nu \beta} \nabla_\rho V^\nu + V^\rho g^{\alpha \beta} \nabla_\alpha \left( \delta^\mu_{\beta \rho} \right)
\]
\[
= \left( g^{\alpha\beta} - h^{\alpha\beta} \right) \nabla_\alpha \nabla_\beta V^\mu + g^{\mu \sigma} \left( - \nabla_\sigma h_{\alpha \rho} + \alpha h_{\alpha \rho \sigma} + \rho h_{\alpha \sigma \rho} \right) \nabla^\sigma V^\rho
\]
\[
- \frac{1}{2} g^{\alpha \beta} \left( g^{\rho \sigma} (- \nabla_\sigma h_{\alpha \beta} + \alpha h_{\sigma \beta \alpha}) \right) \nabla^\rho V^\mu + \frac{1}{2} V^\rho g^{\alpha \beta} \nabla_\alpha \left\{ g^{\mu \nu} (- \nabla_\nu h_{\beta \rho} + \beta h_{\nu \beta \rho}) \right\}
\]
\[
= \left( g^{\alpha \beta} - h^{\alpha \beta} \right) \nabla_\alpha \nabla_\beta V^\mu + \left( - \nabla_\sigma h_{\alpha \rho} + \alpha h_{\alpha \rho \sigma} + \rho h_{\alpha \sigma \rho} \right) \nabla^\sigma V^\rho
\]
\[
- \frac{1}{2} \left( - \nabla^\rho h + 2 \nabla_\alpha h^{\alpha \rho} \right) \nabla_\rho V^\mu + \frac{1}{2} V^\rho \left( \Box h_\mu^\rho + \nabla^\alpha \nabla_\rho h_\mu^\rho \right) .
\]  
(C.8)

We now have the following identity for the double derivative of the perturbation \( h_{\alpha \beta} \),
\[
\nabla^\beta \nabla_\rho h_\mu^\beta \nabla_\beta - \nabla^\beta \nabla^\mu h_{\beta \rho} = \left[ \nabla_\beta, \nabla_\rho \right] h_{\beta \rho} + \nabla_\rho \nabla_\beta h_{\beta \mu} - \left[ \nabla_\beta, \nabla_\mu \right] h_{\beta \rho} - \nabla_\beta \nabla^\beta h_{\beta \rho}
\]
\[
= \nabla_\beta \nabla_\rho h_{\beta \mu} - \nabla^\mu h_{\beta \rho} + R^\beta_{\sigma \beta \rho} h_{\sigma \mu} + R^\alpha_{\rho \beta \sigma} h_{\beta \mu} - R^\beta_{\sigma \mu \rho} h_{\sigma \beta} - R^\beta_{\beta \mu \rho} h_{\sigma \sigma} + \nabla_\rho \nabla_\beta h_{\beta \mu} - \nabla_\rho \nabla^\beta h_{\beta \rho} + R_{\sigma \rho \beta} h_{\sigma \mu} - R_{\sigma \rho \mu} h_{\beta \sigma} ,
\]  
(C.9)

where in the last line we have used the result, \( R^\beta_{\sigma \beta \rho} h_{\sigma \mu} = R^\beta_{\beta \rho \mu} h_{\sigma \sigma} \), such that both the Riemann dependent terms cancel each other. Substituting this result in Eq. (C.8), we obtain,
\[
\Box V^\mu = \left( g^{\alpha \beta} - h^{\alpha \beta} \right) \nabla_\alpha \nabla_\beta V^\mu + \left( - \nabla^\mu h_{\alpha \rho} + \nabla_\rho h_{\alpha \mu} + \nabla^\mu h_\alpha^\mu \right) \nabla^\alpha V^\rho
\]
\[
- \frac{1}{2} \left( - \nabla^\rho h + 2 \nabla_\alpha h^{\alpha \rho} \right) \nabla_\rho V^\mu + \frac{1}{2} V^\rho \left( \Box h_\mu^\rho + \nabla^\alpha \nabla_\rho h_\mu^\rho \right) + \nabla^\rho \nabla_\beta h_{\beta \rho} - \nabla^\rho \nabla_\beta h_{\beta \rho} + R_{\sigma \rho \beta} h_{\sigma \mu} - R_{\sigma \rho \mu} h_{\beta \sigma} ,
\]  
(C.10)

the further simplification of which results in Eq. (3.24).

**Appendix D: Hamiltonian for the almost Killing equation**

In this appendix, we will construct the Hamiltonian for the perturbed Killing vector which will use in the stability analysis. The starting point is the action for the perturbed Killing vector δμν, which takes the following form,
\[
\mathcal{A}[\delta \xi] = \int d^4x \sqrt{-g} \left\{ \frac{1}{2} \left( 1 + \frac{1}{h^2} \right) \left\{ \bar{\mu} \delta^\mu_{\alpha \beta} - (h_{\alpha \beta} + \delta_{\alpha \beta}) (g^{\mu \nu} - h^{\mu \nu}) - (\delta^\mu_{\nu} + h^\nu_{\mu}) \left( \delta^\mu_{\beta} - h^\mu_{\beta} \right) \right\} \left( \nabla_\mu \delta \xi^\alpha \nabla_\nu \delta \xi^\beta \right)
\]
\[
+ \frac{\bar{\mu}}{2} \left\{ (\xi^\alpha + \delta \xi^\alpha) \nabla_\alpha h (\nabla_\nu \delta \xi^\mu) \right\} - (\nabla_\mu \delta \xi^\alpha) \left\{ h_{\alpha \beta} \nabla^\mu \delta \xi^\beta + h^\mu_{\beta} \nabla_\alpha \delta \xi^\beta \right\} - g_{\alpha \sigma} \nabla_\mu \delta \xi^\alpha (\xi^\sigma + \delta \xi^\sigma) \nabla_\sigma \delta \xi^\mu \right\} .
\]  
(D.1)

The above structure of the action, though instructive, depends on terms quadratic in the perturbation. Since we are working in the linear regime, we can neglect all the terms \( O(h^2) \) and hence the above action can be expressed as,
\[
\mathcal{A}[\delta \xi] = \int d^4x \sqrt{-g} \left\{ \frac{1}{2} \left( 1 + \frac{1}{h^2} \right) \left\{ \bar{\mu} \delta^\mu_{\alpha \beta} - (h_{\alpha \beta} + \delta_{\alpha \beta}) (g^{\mu \nu} - h^{\mu \nu}) - (\delta^\mu_{\nu} + h^\nu_{\mu}) \left( \delta^\mu_{\beta} - h^\mu_{\beta} \right) \right\} \left( \nabla_\mu \delta \xi^\alpha \nabla_\nu \delta \xi^\beta \right)
\]
\[
+ \frac{\bar{\mu}}{2} \left\{ (\xi^\alpha + \delta \xi^\alpha) \nabla_\alpha h (\nabla_\nu \delta \xi^\mu) \right\} - (\nabla_\mu \delta \xi^\alpha) \left\{ h_{\alpha \beta} \nabla^\mu \delta \xi^\beta + h^\mu_{\beta} \nabla_\alpha \delta \xi^\beta \right\} - g_{\alpha \sigma} \nabla_\mu \delta \xi^\alpha (\xi^\sigma + \delta \xi^\sigma) \nabla_\sigma \delta \xi^\mu \right\} .
\]  
(D.2)

Note that the last two terms in the first line of the above expression will not contribute. These terms are antisymmetric under the following exchange \( (\mu \leftrightarrow \nu) \) and \( (\alpha \leftrightarrow \beta) \) simultaneously, while \( \nabla_\mu \delta \xi^\alpha \nabla_\nu \delta \xi^\beta \) is symmetric under the said exchange. The structure of the action then simplifies to Eq. (4.1). The total derivative terms that we have thrown away in the derivation of the field equation for \( \delta \xi^\mu \) in the previous section yields the momentum \( P^\mu \) defined in Eq. (4.3) in the main text.
In order to determine all the time derivative parts of the Hamiltonian quadratic in the perturbed Killing vector, we need both the contributions from the Lagrangian as well as the momentum. The contribution from the Lagrangian becomes,

\[
L_2 = \frac{1}{2} \left( 1 + \frac{1}{h} \right) \left\{ \bar{\mu} \delta^0_\alpha \delta^0_\beta - g_{\alpha \beta} g^{00} - h_{\alpha \beta} g^{00} + g_{\alpha \beta} h^{00} - \delta^0_\alpha \delta^0_\beta \right\} \left( \delta \xi^\alpha \delta \xi^\beta \right)
\]

\[
= \frac{1}{2} \left( 1 + \frac{1}{h} \right) \left\{ \bar{\mu} - 1 \right\} \left( \delta \xi^0 \delta \xi^0 \right) + \frac{1}{2} \left( 1 + \frac{1}{h} \right) \left\{ -g_{\alpha \beta} g^{00} - h_{\alpha \beta} g^{00} + g_{\alpha \beta} h^{00} \right\} \left( \delta \xi^\alpha \delta \xi^\beta \right)
\]

\[
= \frac{1}{2} \left( 1 + \frac{1}{h} \right) \left\{ \bar{\mu} - 1 \right\} \left( \delta \xi^0 \delta \xi^0 \right) + \frac{1}{2} \left( 1 + \frac{1}{h} \right) \left\{ g_{ij} + h_{ij} + g_{ij} h^{00} \right\} \left( \delta \xi^i \delta \xi^j \right)
\]

\[
+ \frac{1}{2} \left( 1 + \frac{1}{h} \right) \left\{ h^{00} - h^{00} \right\} \left( \delta \xi^0 \delta \xi^0 \right) + \left( 1 + \frac{1}{h} \right) \delta \xi^i \delta \xi^i . \quad (D.3)
\]

Similarly, the contribution from the momentum yields,

\[
P^0_\alpha \delta \xi^\alpha = \left( 1 + \frac{1}{h} \right) \left( \bar{\mu} \delta^0_\alpha \delta^0_\beta - g_{\alpha \beta} g^{00} - \delta^0_\alpha \delta^0_\beta + g_{\alpha \beta} h^{00} - h_{\alpha \beta} g^{00} \right) \delta \xi^\alpha \delta \xi^\beta
\]

\[
= \left( 1 + \frac{1}{h} \right) \left\{ \bar{\mu} - 1 \right\} \left( \delta \xi^0 \delta \xi^0 \right) + \left( 1 + \frac{1}{h} \right) \left\{ -g_{\alpha \beta} g^{00} - h_{\alpha \beta} g^{00} + g_{\alpha \beta} h^{00} \right\} \left( \delta \xi^\alpha \delta \xi^\beta \right)
\]

\[
= \left( 1 + \frac{1}{h} \right) \left\{ \bar{\mu} - 2 \right\} \left( \delta \xi^0 \delta \xi^0 \right) + \left( 1 + \frac{1}{h} \right) \left\{ g_{ij} + h_{ij} + g_{ij} h^{00} \right\} \left( \delta \xi^i \delta \xi^j \right) + 2 \left( 1 + \frac{1}{h} \right) \delta \xi^i \delta \xi^i . \quad (D.4)
\]

Thus the part of the Hamiltonian which is quadratic in the time derivative of the perturbed Killing vector becomes the one given by Eq. (4.7) in the main text. Similarly, double space derivatives of the perturbed Killing vector within the Lagrangian yields,

\[
L_2 = \frac{1}{2} \left( 1 + \frac{1}{h} \right) \left\{ \bar{\mu} \delta^i_\alpha \delta^j_\beta - g_{\alpha \beta} g^{ij} - h_{\alpha \beta} g^{ij} + g_{\alpha \beta} h^{ij} - \delta^i_\alpha \delta^j_\beta \right\} \left( \partial_i \delta \xi^\alpha \partial_j \delta \xi^\beta \right)
\]

\[
= \frac{1}{2} \left( 1 + \frac{1}{h} \right) \left\{ \bar{\mu} - 1 \right\} \left( \partial_i \delta \xi^\alpha \partial_j \delta \xi^\beta \right) + \frac{1}{2} \left( 1 + \frac{1}{h} \right) \left\{ -g_{00} g^{ij} - h_{00} g^{ij} + h_{00} h^{ij} \right\} \left( \partial_i \delta \xi^0 \partial_j \delta \xi^0 \right)
\]

\[
+ \frac{1}{2} \left( 1 + \frac{1}{h} \right) \left\{ -g_{ab} g^{ij} - h_{ab} g^{ij} \right\} \left( \partial_i \delta \xi^a \partial_j \delta \xi^b \right) - \left( 1 + \frac{1}{h} \right) \delta \xi^i \delta \xi^j \left( \partial_i \delta \xi^0 \partial_j \delta \xi^0 \right)
\]

\[
= \frac{1}{2} \left( 1 + \frac{1}{h} \right) \left\{ \bar{\mu} - 1 \right\} \left( \partial_i \delta \xi^\alpha \partial_j \delta \xi^\beta \right) + \frac{1}{2} \left( 1 + \frac{1}{h} \right) \left\{ g^{ij} - h_{00} g^{ij} - h^{ij} \right\} \left( \partial_i \delta \xi^0 \partial_j \delta \xi^0 \right)
\]

\[
+ \frac{1}{2} \left( 1 + \frac{1}{h} \right) \left\{ -g_{ab} g^{ij} - h_{ab} g^{ij} \right\} \left( \partial_i \delta \xi^a \partial_j \delta \xi^b \right) - \left( 1 + \frac{1}{h} \right) \delta \xi^i \delta \xi^j \left( \partial_i \delta \xi^0 \partial_j \delta \xi^0 \right) . \quad (D.5)
\]

The momentum part does not contribute, since it depends on single time derivative of the perturbed Killing vector and hence we obtain the contribution of double space derivative parts to the Hamiltonian in Eq. (4.8).

**Appendix E: Hyperbolicity of the AKE**

In this appendix, we will present the detailed derivation regarding the hyperbolicity of the perturbed AKE. Collecting all the terms involving double derivatives of \( \delta \xi^\alpha \) and expressing the metric as, \( g^{\alpha \beta} = g^{\alpha \beta} - n^n n^\beta + s^a s^\beta \), where \( n^n \) is the time-like unit vector and \( s^a \) is the space-like unit vector, we obtain from Eq. (3.29),

\[
\left( q^{\alpha \beta} - n^n n^\beta + s^a s^\beta \right) \partial_a \partial_b \delta \xi^\mu + \left( 1 - \mu - \delta \mu \right) \left( q^{\mu \alpha} - n^\mu n^\alpha + s^\mu s^\alpha \right) \delta^\nu_\mu \partial_\mu \partial_\nu \delta \xi^\alpha
\]

\[
- \left( h^{\mu \nu} \delta^\rho_\mu \partial_\rho \partial_\mu \delta \xi^\nu - (1 - \mu - \delta \mu) \delta^\mu_\rho \partial_\rho \partial_\mu \delta \xi^\nu \right) = 0 . \quad (E.1)
\]

Defining \( n^n \partial_\alpha = \partial_n, s^a \partial_\alpha = \partial_a \) and \( q^{\alpha \beta} = q^{AB} \delta^\alpha_A \delta^\beta_B \), we can rewrite various terms appearing in the above equation as

\[
\text{Term 1} = q^{AB} \partial_A \partial_B \delta \xi^\mu - \partial^2 \delta \xi^\mu + \partial^2 \delta \xi^\mu , \quad (E.2)
\]
Term 2 = \((1 - \mu - \delta \mu) (q^{\alpha\alpha} - n^\alpha n^\alpha + s^\alpha s^\alpha) (q^\mu - n^\mu n^\mu + s^\mu s^\mu) \partial_\alpha \partial_\beta \xi^\alpha \),
\[\text{Term 3} = -h^{\mu\nu} (q^\nu - n^\nu n^\nu + s^\nu s^\nu) (q^\mu - n^\mu n^\mu + s^\mu s^\mu) \partial_\alpha \partial_\beta \xi^\mu \]
\[\text{Term 4} = \frac{2}{2} hl^{\beta} (q_\beta^\rho - n_\rho n_\rho + s_\rho s_\rho) \partial_\alpha (\partial_\beta \xi^\beta - \partial_\alpha \xi^\alpha + \partial_\alpha \delta^\beta) + (1 - \mu - \delta \mu) h^{\alpha\mu} \partial_\rho (\partial_\alpha \xi^\alpha - \partial_\rho \delta^\alpha + \partial_\rho \delta^\alpha) \]
\[\text{Collecting all these terms, we obtain the perturbed Killing equation along } n^\mu \text{ as,}
\[\left( q^{AB} - h^{AB} \right) \partial_A \partial_B \delta^\xi^B = \left( 2 - \mu - \delta \mu \right) \left( 1 + h^{nm} \right) \partial^2 \delta^\xi^B + \left( 1 - h^{ss} \right) \partial^2 \delta^\xi^B \]
\[\text{Along identical lines, the perturbed Killing equation along } s^\mu \text{ becomes,}
\[\left( q^{AB} - h^{AB} \right) \partial_A \partial_B \delta^\xi^B = \left( 1 + h^{nm} \right) \partial^2 \delta^\xi^B + \left( 2 - \mu - \delta \mu \right) \left( 1 - h^{ss} \right) \partial^2 \delta^\xi^B \]
\[\text{Finally, the perturbed equation along the angular direction becomes,}
\[\left( q^{AB} - h^{AB} \right) \partial_A \partial_B \delta^\xi^P = \left( 1 + h^{nm} \right) \partial^2 \delta^\xi^P + \left( 1 - h^{ss} \right) \partial^2 \delta^\xi^P \]
\[\text{From the above set of equations we can retrieve the equations of [12] by setting } h_{\mu\nu} = 0 = \delta \mu. \text{ It remains to be seen whether these equations are hyperbolic or not. For that purpose, let us rewrite these equations in an appropriate manner. Let us start with Eq. (E.5) and write it as,}
\[\left( 2 - \mu - \delta \mu \right) \left( 1 + h^{nm} \right) \partial^2 \delta^\xi^B = \left( 1 - h^{ss} \right) \partial^2 \delta^\xi^B + \left( 3 - \mu - \delta \mu \right) h^{ns} \partial_\alpha \partial_\beta \partial^\alpha \delta^\beta - \left( 1 - \mu - \delta \mu \right) h^{ss} \partial_\alpha \partial_\beta \partial^\alpha \delta^\beta \]
\[\text{which, after neglecting the transverse derivatives, yields Eq. (4.10) in the main text. This equation can also be re-expressed as,}
\[\partial^2 \delta^\xi^B = \frac{1}{2 - \mu - \delta \mu} \left[ \frac{2}{2} hl^{\beta} \partial_\alpha (\partial_\beta \xi^\beta - \partial_\alpha \delta^\beta + \partial_\alpha \delta^\beta) \right] + \left( 3 - \mu - \delta \mu \right) h^{ns} \partial_\alpha \partial_\beta \partial^\alpha \delta^\beta \]
\[\text{Performing a similar analysis, we obtain the following from Eq. (E.6),}
\[\left( 1 + h^{nm} \right) \partial^2 \delta^\xi^P = \left( 2 - \mu - \delta \mu \right) \left( 1 - h^{ss} \right) \partial^2 \delta^\xi^P + \left( 3 - \mu - \delta \mu \right) h^{ns} \partial_\alpha \partial_\beta \partial^\alpha \delta^\beta \]
\[\left( 1 + h^{nm} \right) \partial^2 \delta^\xi^P = \left( 2 - \mu - \delta \mu \right) \left( 1 - h^{ss} \right) \partial^2 \delta^\xi^P + \left( 3 - \mu - \delta \mu \right) h^{ns} \partial_\alpha \partial_\beta \partial^\alpha \delta^\beta \]
\[-(1 - \mu - \delta\mu) h^{sn} \left( \frac{1}{(2 - \mu - \delta\mu)(1 + h^{nn})} \partial_s^2 \delta s^n + \frac{1 - \mu - \delta\mu}{(2 - \mu - \delta\mu)} \partial_s \partial_s \delta s^n \right) \]

which, ignoring transverse coordinates, has been presented as Eq. (4.11) in the main text. Further simplification of this equation yields:

\[
\partial_s^2 \delta s^n = \frac{(3 - \mu - \delta\mu)}{(1 + h^{nn})} \partial_s \partial_s \delta s^n - \frac{(1 - \mu - \delta\mu)^2 h^{sn}}{(2 - \mu - \delta\mu)(1 + h^{nn})} \partial_s \partial_s \delta s^n - \frac{1 - \mu - \delta\mu}{(1 + h^{nn})} \partial_s \partial_s \delta s^n
\]

Finally, from Eq. (E.7) we obtain,

\[
(1 + h^{nn}) \partial_s^2 \delta s^P = (1 - h^{ss}) \partial_s^2 \delta s^P + 2 h^{ns} \partial_s \partial_s \delta s^P + (1 - \mu - \delta\mu) h^{Ps} \partial_s \partial_s \delta s^n
\]

which in some form, again neglecting transverse coordinates, has been presented as Eq. (4.12) in the main text. This expression can be further rewritten as,

\[
\partial_s^2 \delta s^P = \frac{(1 - h^{ss})}{(1 + h^{nn})} \partial_s^2 \delta s^P - \frac{(1 - \mu - \delta\mu) h^{Ps}}{(2 - \mu - \delta\mu)(1 + h^{nn})} \partial_s \partial_s \delta s^n - \frac{(1 - \mu - \delta\mu) h^{Ps}}{(1 + h^{nn})} \partial_s \partial_s \delta s^n
\]

Hence we can write,

\[
\partial_n \begin{bmatrix} \partial_s \delta s^n \\ \partial_s \delta s^n \\ \partial_s \delta s^A \\ \partial_n \delta s^n \\ \partial_n \delta s^n \\ \partial_n \delta s^A \end{bmatrix} = P^s \partial_s \begin{bmatrix} \partial_s \delta s^n \\ \partial_s \delta s^n \\ \partial_s \delta s^A \\ \partial_n \delta s^n \\ \partial_n \delta s^n \\ \partial_n \delta s^A \end{bmatrix}
\]

where \( P_s \) is the principal symbol defined in the main text. Note that as the perturbations vanish, we obtain \( A = 1/(2 - \mu) \), \( B = 0 = C \), \( D = (1 - \mu)/(2 - \mu) \), \( E = 0 \), \( F = 2 - \mu \), \( G = -(1 - \mu) \), \( H = 0 = I = J \), \( K = 1 \) and \( L = 0 = M = N \). Thus we obtain,

\[
\partial_n \begin{bmatrix} \partial_s \delta s^n \\ \partial_s \delta s^A \\ \partial_n \delta s^n \\ \partial_n \delta s^A \\ \partial_n \delta s^n \\ \partial_n \delta s^A \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ \frac{1 - \mu}{2 - \mu} & 0 & 0 & 0 & 1 & 0 \\ 0 & 2 - \mu & 0 & -(1 - \mu) & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \partial_s \begin{bmatrix} \partial_s \delta s^n \\ \partial_s \delta s^n \\ \partial_s \delta s^A \\ \partial_n \delta s^n \\ \partial_n \delta s^n \\ \partial_n \delta s^A \end{bmatrix}
\]

which coincides with earlier results.
Appendix F: Perturbation of Noether current and Noether charge

In this appendix we will discuss the perturbation of the Noether current and the Noether charge due to the gravitational perturbation and the perturbation of the AKE. To start with, one may consider the following expression for the Noether current associated with the GKV as the diffeomorphism vector field,

\[ J^\mu = \bar{\nabla}_\nu (\bar{\nabla}_\nu \chi^\nu - \bar{\nabla}_\mu \chi^\mu) = \bar{\nabla}_\nu (\bar{\nabla}_\nu |h\rangle + \bar{\nabla}_\mu \bar{\nabla}_\nu \chi^\nu) - \Box \chi^\mu \]

which has been used to derive Eq. (6.1) in the main text.

Further, using the result, we obtain,

\[ \bar{R}_\sigma = g^\mu\nu \bar{R}_\sigma = (g^\mu\nu - h^\mu\nu) (R_{\alpha\sigma} + \delta R_{\alpha\sigma}) = R^\nu\sigma - h^\nu\alpha R_{\alpha\sigma} + g^\nu\alpha \delta R_{\alpha\sigma} + O(h^2). \]

We also have the following result for the mixed component of the Ricci tensor in the perturbed background given by \( \bar{g}_{\mu\nu} \),

\[ \bar{R}_\sigma = g^\mu\nu \bar{R}_\sigma = (g^\mu\nu - h^\mu\nu) (R_{\alpha\sigma} + \delta R_{\alpha\sigma}) = R^\nu\sigma - h^\nu\alpha R_{\alpha\sigma} + g^\nu\alpha \delta R_{\alpha\sigma} + O(h^2). \]

Using the above results for the perturbation of the Ricci tensor and the Ricci scalar, the perturbation of the Einstein tensor becomes,

\[ \delta G_{\alpha\beta} = \delta R_{\alpha\beta} - \frac{1}{2} \delta \Box \delta R = \frac{1}{2} \left( -\Box h_{\alpha\beta} - \nabla^\alpha \nabla^\beta h + \nabla^\alpha \nabla_\mu h^\mu_{\beta} + \nabla_\mu \nabla^\alpha h^\mu_{\beta} + \nabla^\alpha \nabla_\mu \nabla^\mu h_{\alpha\beta} - \nabla^\alpha \nabla_\nu \nabla^\nu h^\alpha_{\beta} - 2R_{\rho\mu\nu\lambda} h^\alpha_{\beta} h^\rho_{\mu} \right) \]

Along identical lines, the perturbation of the Ricci scalar can be expressed as,

\[ \delta R = \delta^\nu \delta R^\nu = \frac{1}{2} \delta^\nu \left( -\Box h^\nu - \nabla^\alpha \nabla_\nu h^\nu + \nabla^\alpha \nabla_\mu h^\mu_{\nu} + \nabla_\nu \nabla^\alpha h^\alpha_{\nu} + \nabla^\alpha \nabla_\mu \nabla^\mu h^\nu - \nabla^\alpha \nabla_\nu \nabla^\nu h^\alpha_{\nu} - 2R^\nu_{\mu\sigma\nu} h^\mu_{\nu} \right) \]

Using the above results for the perturbation of the Ricci tensor and the Ricci scalar, the perturbation of the Einstein tensor becomes,

\[ \delta G_{\alpha\beta} = \delta R_{\alpha\beta} - \frac{1}{2} \delta \Box \delta R = \frac{1}{2} \left( -\Box h_{\alpha\beta} - \nabla^\alpha \nabla^\beta h + \nabla^\alpha \nabla_\mu h^\mu_{\beta} + \nabla_\mu \nabla^\alpha h^\mu_{\beta} + \nabla^\alpha \nabla_\mu \nabla^\mu h_{\alpha\beta} - \nabla^\alpha \nabla_\nu \nabla^\nu h^\alpha_{\beta} - 2R_{\rho\mu\nu\lambda} h^\alpha_{\beta} h^\rho_{\mu} \right) \]

Before proceeding further, we may consider the following two cases:

- In the Lorenz gauge, the above perturbation expressions for the Ricci and the Einstein tensor becomes,

\[ \delta G_{\alpha\beta} = \frac{1}{2} \left( -\Box h_{\alpha\beta} + \nabla^\alpha \nabla_\mu h^\mu_{\beta} + \nabla_\mu \nabla^\alpha h^\mu_{\beta} + \nabla^\alpha \nabla_\mu \nabla^\mu h_{\alpha\beta} - \nabla^\alpha \nabla_\nu \nabla^\nu h^\alpha_{\beta} - 2R_{\rho\mu\nu\lambda} h^\alpha_{\beta} h^\rho_{\mu} \right) \]

\[ \delta R_{\alpha\beta} = \frac{1}{2} \left( -\Box h_{\alpha\beta} + \nabla^\alpha \nabla_\mu h^\mu_{\beta} + \nabla_\mu \nabla^\alpha h^\mu_{\beta} + \nabla^\alpha \nabla_\mu \nabla^\mu h_{\alpha\beta} - \nabla^\alpha \nabla_\nu \nabla^\nu h^\alpha_{\beta} - 2R_{\rho\mu\nu\lambda} h^\alpha_{\beta} h^\rho_{\mu} \right) \]
• For a vacuum background (and $\delta R = 0$), with the transverse-traceless gauge condition, we obtain,

$$\delta G^\alpha_\beta = \frac{1}{2} (\Delta h^\alpha_\beta - 2 R^\alpha_\beta \rho h^{\mu \rho}) = \delta R^\alpha_\beta . \quad (F.9)$$

Interestingly, using the following identity,

$$\nabla^\nu (\nabla_\sigma \chi^\sigma) = \tilde{g}^\nu_\alpha \nabla_\alpha (\nabla_\sigma \chi^\sigma)$$

$$= \tilde{g}^\nu_\alpha \left[ \nabla_\alpha (\nabla_\sigma \chi^\sigma) + \frac{1}{2} \chi^\sigma \nabla_\alpha \nabla_\sigma h + \frac{1}{2} \nabla_\rho h (\nabla_\alpha \chi^\rho) \right], \quad (F.10)$$

the Noether current for the perturbed spacetime becomes,

$$\tilde{J}^\nu = 2 (R^\nu_\sigma + \delta R^\nu_\sigma) \chi^\sigma + (2 - \bar{\mu}) \tilde{g}^\nu_\alpha \left[ \nabla_\alpha (\nabla_\sigma \chi^\sigma) + \frac{1}{2} \chi^\sigma \nabla_\alpha \nabla_\sigma h + \frac{1}{2} \nabla_\rho h (\nabla_\alpha \chi^\rho) \right] + \tilde{g}^\nu_\alpha \nabla_\alpha (2 - \bar{\mu})$$

$$\nabla^\nu \left[ \nabla_\sigma \chi^\sigma + \frac{1}{2} \chi^\sigma \nabla_\sigma h + \frac{1}{2} \nabla_\rho h (\nabla_\sigma \chi^\rho) \right]$$

$$= (-\Box h^\nu_\sigma - \nabla^\nu \nabla_\sigma h + \nabla^\nu \nabla_\mu h^\mu_\sigma + \nabla_\sigma \nabla_\mu h^{\mu \nu} + R^\nu_\mu h^\mu_\sigma - R^\sigma_\mu h^{\mu \nu} - 2 R^\nu_\mu \rho h^{\mu \rho}) \chi^\sigma$$

$$+ 2 R^\nu_\sigma \chi^\sigma + (2 - \bar{\mu}) (g^{\nu_\alpha} - h^{\nu_\alpha}) \left[ \nabla_\alpha (\nabla_\sigma \chi^\sigma) + \frac{1}{2} \chi^\sigma \nabla_\alpha \nabla_\sigma h + \frac{1}{2} \nabla_\rho h (\nabla_\alpha \chi^\rho) \right]$$

$$+ \tilde{g}^\nu_\alpha \left( \nabla_\alpha \chi^\sigma + \frac{1}{2} \chi^\sigma \nabla_\alpha h \right) \nabla_\alpha (2 - \bar{\mu}) \quad (F.11)$$

This equation can be simplified further using the Lorenz gauge condition, which reads,

$$\nabla_\alpha \left( h^\alpha_\rho _\beta - \frac{1}{2} \delta^\alpha_\beta h \right) = 0 . \quad (F.12)$$

Hence the second, third and fourth terms in the first line of the above expression for Noether current will cancel and the Noether current can be written in the following form

$$\tilde{J}^\nu = 2 R^\nu_\sigma \chi^\sigma + (+ - \Box h^\nu_\sigma + R^{\nu_\mu} h_{\mu} \sigma - R_{\sigma \mu} h^{\mu \nu} - 2 R^{\nu_\mu \rho} h^{\mu \rho} \delta \xi^\sigma)$$

$$+ (2 - \bar{\mu}) (g^{\nu_\alpha} - h^{\nu_\alpha}) \left( \nabla_\sigma \xi^\sigma + \nabla_\sigma \delta \xi^\sigma + \frac{1}{2} (\xi^\sigma + \delta \xi^\sigma) \nabla_\sigma h \right)$$

$$+ (g^{\nu_\alpha} - h^{\nu_\alpha}) \left( \nabla_\sigma \xi^\sigma + \nabla_\sigma \delta \xi^\sigma + \frac{1}{2} (\xi^\sigma + \delta \xi^\sigma) \nabla_\sigma h \right) \nabla_\alpha (2 - \bar{\mu}) \quad (F.13)$$

If we assume the background spacetime to admit a Killing vector $\xi^\sigma$, then it follows that $\nabla_\sigma \xi^\sigma = 0$. Hence we obtain,

$$\tilde{J}^\nu = J^\nu + 2 R^\nu_\sigma \delta \xi^\sigma + (+ - \Box h^\nu_\sigma + R^{\nu_\mu} h_{\mu} \sigma - R_{\sigma \mu} h^{\mu \nu} - 2 R^{\nu_\mu \rho} h^{\mu \rho} \delta \xi^\sigma)$$

$$+ (2 - \bar{\mu}) (g^{\nu_\alpha} - h^{\nu_\alpha}) \nabla_\alpha \left( \nabla_\sigma \delta \xi^\sigma + \frac{1}{2} (\xi^\sigma + \delta \xi^\sigma) \nabla_\sigma h \right)$$

$$+ (g^{\nu_\alpha} - h^{\nu_\alpha}) \left( \nabla_\sigma \delta \xi^\sigma + \frac{1}{2} (\xi^\sigma + \delta \xi^\sigma) \nabla_\sigma h \right) \nabla_\alpha (2 - \bar{\mu}) \quad (F.14)$$

The perturbation $\delta J^\nu = \tilde{J}^\nu - J^\nu$ then takes the following form:

$$\delta J^\nu = 2 R^\nu_\sigma \delta \xi^\sigma + 2 \delta R^\nu_\sigma (\xi^\sigma + \delta \xi^\sigma)$$

$$+ (2 - \bar{\mu}) (g^{\nu_\alpha} - h^{\nu_\alpha}) \nabla_\alpha \left( \nabla_\sigma \delta \xi^\sigma + \frac{1}{2} (\xi^\sigma + \delta \xi^\sigma) \nabla_\sigma h \right)$$

$$- \left[ g^{\nu_\alpha} \left( \nabla_\sigma \delta \xi^\sigma + \frac{1}{2} (\xi^\sigma + \delta \xi^\sigma) \nabla_\sigma h \right) - h^{\nu_\alpha} \left( \nabla_\sigma \delta \xi^\sigma \right) \right] \nabla_\alpha (2 - \bar{\mu}) \quad (F.15)$$
as well as
\[
\delta J^\nu = 2R^\nu\rho\delta\xi^\rho + \left[2\delta G^\nu + \delta\xi^\nu \left(-\frac{1}{2}\Box h - R_{\mu\rho}h^{\mu\rho}\right)\right] (\xi^\sigma + \delta\xi^\sigma) \\
+ (2 - \bar{\mu}) (g^{\nu\alpha} - h^{\nu\alpha}) \nabla_\alpha (\nabla_\sigma \delta\xi^\sigma) + \frac{1}{2} (2 - \bar{\mu}) g^{\nu\alpha} \nabla_\alpha \left[\left(\xi^\sigma + \delta\xi^\sigma\right) \nabla_\sigma h\right] \\
- \left[g^{\nu\alpha} \left(\nabla_\sigma \delta\xi^\sigma + \frac{1}{2} \left(\xi^\sigma + \delta\xi^\sigma\right) \nabla_\sigma h\right) - h^{\nu\alpha} \left(\nabla_\sigma \delta\xi^\sigma\right)\right] (\nabla_\alpha \delta\mu) ,
\]
(Eq. 6.4)

which have been used in Eq. (6.4) in the main text.

The above expression for the change in Noether current simplifies considerably in the following two cases:

- For vacuum spacetime in transverse traceless gauge, the Noether current becomes:
  \[
  \delta J^\nu = \left(-\Box h^\nu - 2 R_{\mu\sigma\rho}h^{\mu\sigma\rho}\right) (\xi^\sigma + \delta\xi^\sigma) \\
  + (2 - \bar{\mu}) (g^{\nu\alpha} - h^{\nu\alpha}) \nabla_\alpha (\nabla_\sigma \delta\xi^\sigma) - \left[g^{\nu\alpha} \left(\nabla_\sigma \delta\xi^\sigma\right) - h^{\nu\alpha} \left(\nabla_\sigma \delta\xi^\sigma\right)\right] (\nabla_\alpha \delta\mu) .
  \]
  (F.17)

- On the other hand, if we keep \( \delta\mu \) to be constant, i.e., spacetime coordinate independent, it follows from Eq. (F.23) that the change in Noether current amounts to,
  \[
  \delta J^\nu = 2R^\nu\rho \delta\xi^\rho + \left(-\Box h^\nu + R^\nu_{\mu\sigma\rho}h^{\mu\rho}\right) (\xi^\sigma + \delta\xi^\sigma) \\
  + (2 - \bar{\mu}) (g^{\nu\alpha} - h^{\nu\alpha}) \nabla_\alpha (\nabla_\sigma \delta\xi^\sigma) + \frac{1}{2} (2 - \bar{\mu}) g^{\nu\alpha} \nabla_\alpha \left[\left(\xi^\sigma + \delta\xi^\sigma\right) \nabla_\sigma h\right] .
  \]
  (F.18)

The change in the Noether current can also be expressed in terms of matter fields flowing into the spacetime and for this purpose it is instructive to write down the perturbed Einstein’s equations, which yield,
\[
\delta R^\alpha\beta = 8\pi \delta\tilde{T}^\alpha\beta ; \quad \tilde{T}^\alpha\beta = T^\alpha\beta - \frac{1}{2} \delta T^\alpha\beta .
\]

(F.19)

Using Eq. (F.4), we obtain,
\[
\frac{1}{2} \left(-\Box h^\nu_{\beta} - \nabla^\alpha \nabla_\beta h + \nabla^\alpha \nabla_\mu h^\mu_{\beta} + \nabla^\beta \nabla_\mu h^\mu_{\alpha} + R^{\alpha\mu}h_{\mu\beta} - R_{\beta\mu}h^{\mu\alpha} - 2R^{\alpha}_{\mu\beta\rho}h^{\mu\rho}\right) = 8\pi \delta\tilde{T}^\alpha\beta .
\]

(F.20)

While using both Eq. (F.4) and Eq. (F.5) we obtain, from Einstein’s equation
\[
\frac{1}{2} \left(-\Box h^\nu_{\beta} - \nabla^\alpha \nabla_\beta h + \nabla^\alpha \nabla_\mu h^\mu_{\beta} + \nabla^\beta \nabla_\mu h^\mu_{\alpha} + R^{\alpha\mu}h_{\mu\beta} - R_{\beta\mu}h^{\mu\alpha} - 2R^{\alpha}_{\mu\beta\rho}h^{\mu\rho}\right) = 8\pi \delta T^\alpha\beta .
\]

(F.21)

One can impose Lorenz gauge condition, which will eliminate the second, third and fourth term in the above expression. Further spacetime being vacuum demands the Ricci tensor to vanish and hence we obtain,
\[
-\Box h^\nu_{\beta} - 2R^{\alpha}_{\mu\beta\rho}h^{\mu\rho} = 16\pi \delta\tilde{T}^\alpha\beta ,
\]

(F.22)
as the perturbed Einstein’s equations in vacuum spacetime. Thus from Eq. (F.23), i.e., using perturbed Einstein’s equations, we obtain the perturbed Noether current in the vacuum background spacetime to read,
\[
\delta J^\nu = \left(-\Box h^\nu - 2 R_{\mu\sigma\rho}h^{\mu\sigma\rho}\right) (\xi^\sigma + \delta\xi^\sigma) \\
+ (2 - \bar{\mu}) \left\{\left(g^{\nu\alpha} - h^{\nu\alpha}\right) \nabla_\alpha (\nabla_\sigma \delta\xi^\sigma) + \frac{1}{2} g^{\nu\alpha} \nabla_\alpha \left[\left(\xi^\sigma + \delta\xi^\sigma\right) \nabla_\sigma h\right]\right\} \\
- \left[g^{\nu\alpha} \left(\nabla_\sigma \delta\xi^\sigma + \frac{1}{2} \left(\xi^\sigma + \delta\xi^\sigma\right) \nabla_\sigma h\right) - h^{\nu\alpha} \left(\nabla_\sigma \delta\xi^\sigma\right)\right] (\nabla_\alpha \delta\mu) \\
= 16\pi \delta\tilde{T}^\nu\sigma (\xi^\sigma + \delta\xi^\sigma) \\
+ (2 - \bar{\mu}) \left\{\left(g^{\nu\alpha} - h^{\nu\alpha}\right) \nabla_\alpha (\nabla_\sigma \delta\xi^\sigma) + \frac{1}{2} g^{\nu\alpha} \nabla_\alpha \left[\left(\xi^\sigma + \delta\xi^\sigma\right) \nabla_\sigma h\right]\right\} \\
- \left[g^{\nu\alpha} \left(\nabla_\sigma \delta\xi^\sigma + \frac{1}{2} \left(\xi^\sigma + \delta\xi^\sigma\right) \nabla_\sigma h\right) - h^{\nu\alpha} \left(\nabla_\sigma \delta\xi^\sigma\right)\right] (\nabla_\alpha \delta\mu) ,
\]

(F.23)
where the perturbed Einstein’s equations have been used. Using the form of the perturbed AKE as presented in Eq. (3.29), we further obtain,

\[
\delta J'' = 16\pi \delta T''_{\sigma} (\xi^\sigma + \delta \xi^\sigma) \\
+ (2 - \bar{\mu}) \left\{ (g'^{\alpha\nu} - h^{\alpha\nu}) \nabla_\alpha (\nabla_\sigma \delta \xi^\sigma) + \frac{1}{2} g'^{\alpha\nu} \nabla_\alpha \left[ (\xi^\sigma + \delta \xi^\sigma) \nabla_\sigma h \right] \right\} \\
- (g^{\alpha\beta} - h^{\alpha\beta}) \left[ \nabla_\alpha \nabla_\beta \delta \xi^\mu + R^\nu_{\mu\alpha\beta} \delta \xi^\rho \right] - 2 \left( \nabla_\alpha h^\mu_{\rho} - \frac{1}{2} \nabla^\mu h_{\alpha\rho} \right) \nabla^{(\alpha} \delta \xi^{\rho)} \\
- (1 - \mu - \delta \mu) \left[ (g'^{\alpha\nu} - h^{\alpha\nu}) \nabla_\alpha (\nabla_\sigma \delta \xi^\sigma) + \frac{1}{2} g'^{\alpha\nu} \nabla_\alpha \left[ (\xi^\sigma + \delta \xi^\sigma) \nabla_\sigma h \right] \right] \\
= 16\pi \delta T''_{\sigma} (\xi^\sigma + \delta \xi^\sigma) \\
+ \left\{ (g'^{\alpha\nu} - h^{\alpha\nu}) \nabla_\alpha (\nabla_\sigma \delta \xi^\sigma) + \frac{1}{2} g'^{\alpha\nu} \nabla_\alpha \left[ (\xi^\sigma + \delta \xi^\sigma) \nabla_\sigma h \right] \right\} \\
- (g^{\alpha\beta} - h^{\alpha\beta}) \left[ \nabla_\alpha \nabla_\beta \delta \xi^\mu + R^\nu_{\mu\alpha\beta} \delta \xi^\rho \right] - 2 \left( \nabla_\alpha h^\mu_{\rho} - \frac{1}{2} \nabla^\mu h_{\alpha\rho} \right) \nabla^{(\alpha} \delta \xi^{\rho)} . \tag{F.24} \]

Note that in the vacuum background spacetime with transverse traceless gauge condition, the above equation can also be expressed as,

\[
\delta J'' = 16\pi \delta T''_{\sigma} (\xi^\sigma + \delta \xi^\sigma) \\
+ \left\{ (g'^{\alpha\nu} - h^{\alpha\nu}) \nabla_\alpha (\nabla_\sigma \delta \xi^\sigma) + \frac{1}{2} g'^{\alpha\nu} \nabla_\alpha \left[ (\xi^\sigma + \delta \xi^\sigma) \nabla_\sigma h \right] \right\} - \Box \delta \xi^\sigma . \tag{F.25} \]

Using the decomposition of \(\xi^\mu = \xi'^{\mu}_1 + \xi'^{\mu}_2\), and keeping terms linear in the perturbation, we immediately obtain

\[
\delta J'' = 16\pi \delta T''_{\sigma} \xi^\sigma \\
+ \left\{ \nabla^\nu (\nabla_\sigma \delta \xi^\sigma) + \frac{1}{2} \nabla^\nu (\xi^\sigma \nabla_\sigma h) \right\} - \Box \delta \xi^\sigma . \tag{F.26} \]

Another way of expressing the same is,

\[
\delta J'' = 16\pi \delta T''_{\sigma} \xi^\sigma \\
+ (2 - \mu) \left\{ \nabla^\nu (\nabla_\sigma \delta \xi^\sigma) + \frac{1}{2} \nabla^\nu (\xi^\sigma \nabla_\sigma h) \right\} . \tag{F.27} \]

Finally, the result that surface gravity does not change at first order in vacuum spacetime with dilute radiation can be derived along the following lines. The surface gravity associated with the GKV field can be derived and expanded in terms of the metric perturbation as,

\[
\kappa^2 = -\frac{1}{2} \left( \tilde{\nabla}_\mu \chi_\nu \right) \left( \tilde{\nabla}_\nu \chi_\mu \right) = -\frac{1}{2} g^{\mu\nu} g_{\alpha\beta} \left( \tilde{\nabla}_\mu \chi_\alpha \right) \left( \tilde{\nabla}_\nu \chi_\beta \right) \\
= -\frac{1}{2} \left( g^{\mu\nu} - h^{\mu\nu} \right) \left( g_{\alpha\beta} + h_{\alpha\beta} \right) \left( \nabla_\mu \chi_\alpha + \delta \Gamma^\alpha_{\mu\rho} \chi^\rho + \nabla_\rho \chi_\alpha + \delta \Gamma^\rho_{\nu\sigma} \chi_\alpha \right) \\
= -\frac{1}{2} \left( g^{\mu\nu} g_{\alpha\beta} - g_{\alpha\beta} h^{\mu\nu} + g^{\mu\nu} h_{\alpha\beta} \right) \left( \nabla_\mu \chi_\alpha \nabla_\nu \chi_\beta + \delta \Gamma^\alpha_{\mu\rho} \chi^\rho \nabla_\nu \chi_\beta + \nabla_\nu \chi_\alpha \delta \Gamma^\beta_{\nu\sigma} \chi_\sigma \right) \\
= -\frac{1}{2} \left[ g^{\mu\nu} g_{\alpha\beta} \left( \nabla_\mu \chi_\alpha \nabla_\nu \chi_\beta + \delta \Gamma^\alpha_{\mu\rho} \chi^\rho \nabla_\nu \chi_\beta + \nabla_\nu \chi_\alpha \delta \Gamma^\beta_{\nu\sigma} \chi_\sigma \right) - \left( g_{\alpha\beta} h^{\mu\nu} - g^{\mu\nu} h_{\alpha\beta} \right) \nabla_\mu \chi_\alpha \nabla_\nu \chi_\beta \right] \\
= -\frac{1}{2} g^{\mu\nu} g_{\alpha\beta} \left[ \nabla_\mu (\xi'^\alpha_1 + \delta \xi'^\alpha_1) \nabla_\nu (\xi'^\beta_1 + \delta \xi'^\beta_1) \right] - \frac{1}{2} g^{\mu\nu} g_{\alpha\beta} \left( \delta \Gamma^\alpha_{\mu\rho} \chi^\rho \delta \xi'^\beta_1 + \nabla_\nu \delta \xi'^\alpha_1 \nabla_\mu \chi_\beta \right) \\
+ \frac{1}{2} \left( g_{\alpha\beta} h^{\mu\nu} - g^{\mu\nu} h_{\alpha\beta} \right) \nabla_\mu \xi'^\alpha_1 \nabla_\nu \chi_\beta \\
= -\frac{1}{2} g^{\mu\nu} g_{\alpha\beta} \left( \nabla_\mu \chi_\alpha \nabla_\nu \chi_\beta \right) - \frac{1}{2} g^{\mu\nu} g_{\alpha\beta} \left[ \nabla_\mu \chi_\alpha \nabla_\nu \delta \xi'^\beta_1 + \nabla_\mu \delta \xi'^\alpha_1 \nabla_\nu \chi_\beta \right]
\[-\frac{1}{2}(\nabla_{\beta}h_{\mu\rho} + \nabla_{\mu}h_{\beta\rho} + \nabla_{\rho}h_{\beta\mu})\xi^\rho\nabla^\mu\xi^\beta + \frac{1}{2}(g_{\alpha\beta}h^{\mu\nu} - g^{\mu\nu}h_{\alpha\beta})\nabla_{\mu}\xi^\alpha\nabla_{\nu}\xi^\beta = \kappa^2 - \nabla_{\mu}\xi^\alpha\nabla^\mu\delta\xi^\alpha - (\nabla_{\mu}h_{\beta\rho})\xi^\rho\nabla^\mu\xi^\beta\] (F.28)

From the results of Section A for first order perturbation, under the assumptions of vacuum background with dilute radiation, it follows that \(\delta\xi^\alpha\) and \(\xi^\alpha\) satisfies identical equations. Thus, \(\delta\xi^\alpha\) is proportional to the background Killing vector field \(\xi^\alpha\) and thus the surface gravity can be equated to \(\kappa\) by a mere rescaling of the parameter. In addition, the term, \((\nabla_{\mu}h_{\beta\rho})\xi^\rho\nabla^\mu\xi^\beta\) will also vanish, since it is always convenient to choose \(\xi^\alpha = (\partial/\partial\lambda)^\alpha\) and choose a gauge such that \(h_{\mu\lambda} = 0\), for all \(\mu\). Thus the surface gravity to first order will not change in a vacuum spacetime with \(\mu = 2\) or with dilute radiation field.

[1] R. A. Matzner, J. Math. Phys. (N.Y.) 9, 1657 (1968); C. Beetle, (2008), arXiv:0808.1745; C. Beetle and S. Wilder, Classical Quantum Gravity 31, 075009 (2014).
[2] J. Brunekreef and M. Reitz, (2020), to be published.
[3] D. Garfinkle and C. Gundlach, Classical Quantum Gravity 16, 4111 (1999).
[4] A. I. Harte, Classical Quantum Gravity 25, 205008 (2008).
[5] C. H. Taubes, J. Math. Phys. (N.Y.) 19, 1515 (1978).
[6] C. Bona, J. Carot, and C. Palenzuela-Luque, Phys. Rev. D 72, 124010 (2005).
[7] A. Komar, Phys. Rev. 113, 934 (1959).
[8] A. Komar, Phys. Rev. 127, 1411 (1962).
[9] M. Ruiz, C. Palenzuela, and C. Bona, Phys. Rev. D 89, 025011 (2014).
[10] J. C. Feng, Phys. Rev. D 98, 104005 (2018).
[11] J.-J. Peng, Communications in Theoretical Physics 72, 065402 (2020); J.-J. Peng and C.-L. Zou, (2019), arXiv:1910.00339 [gr-qc].
[12] J. C. Feng, E. Gasperin, and J. L. Williams, Phys. Rev. D 100, 124034 (2019).
[13] J. Katz, Classical Quantum Gravity 2, 423 (1985); D. Bak, D. Cangemi, and J. Jackiw, Phys. Rev. D 49, 5173 (1994); J. Katz, J. Bičák, and D. Lynden-Bell, Phys. Rev. D 55, 5957 (1997); Y. N. Obukhov and G. F. Rubilar, Phys. Rev. D 73, 124017 (2006); Phys. Rev. D 74, 064002 (2006).
[14] N. Deruelle, J. Katz, and S. Ogushi, Classical Quantum Gravity 21, 1971 (2004); Y. N. Obukhov, G. F. Rubilar, and J. G. Pereira, Phys. Rev. D 74, 104007 (2006); Y. N. Obukhov and D. Puetzfeld, Phys. Rev. D 87, 081502 (2013); Phys. Rev. D 90, 024004 (2014); Y. N. Obukhov, F. Portales-Oliva, D. Puetzfeld, and G. F. Rubilar, Phys. Rev. D 92, 104010 (2015); J. Schmidt and J. Bičák, J. Math. Phys. (N.Y.) 59, 042501 (2018).
[15] V. Iyer and R. M. Wald, Phys. Rev. D 50, 846 (1994), arXiv:gr-qc/9403028.
[16] T. Jacobson and A. Mohd, Phys. Rev. D 92, 124010 (2015), arXiv:1507.01054 [gr-qc].
[17] K. Prabhu, Class. Quant. Grav. 34, 035011 (2017), arXiv:1511.00388 [gr-qc].
[18] S. Chakraborty and R. Dey, Phys. Lett. B 786, 432 (2018), arXiv:1806.05840 [gr-qc].
[19] P. Aneesh, S. Chakraborty, S. J. Hoque, and A. Virmani, Class. Quant. Grav. 37, 205014 (2020), arXiv:2004.10215 [hep-th].
[20] T. Padmanabhan, Gen. Rel. Grav. 46, 1673 (2014), arXiv:1312.3253 [gr-qc].
[21] S. Chakraborty and T. Padmanabhan, Phys. Rev. D 92, 104011 (2015), arXiv:1508.04060 [gr-qc].
[22] S. Chakraborty, JHEP 08, 029 (2015), arXiv:1505.07272 [gr-qc].
[23] https://github.com/justincfeng/mathematica-files/blob/a7f627be270e2de69e9e82ed7b6a1e3e640d8b/2020/AKEPert.nb.
[24] J. C. Feng and R. A. Matzner, Gen. Relativ. Gravit. 50, 99 (2018); E. Sudarshan and N. Mukunda, Classical Dynamics: A Modern Perspective (Krieger, Melbourne, FL, 1983); R. A. Matzner and L. C. Shepley, Classical Mechanics (Prentice-Hall, Englewood Cliffs, NJ, 1991); P. Weiss, Proc. R. Soc. Lond. A 156, 192 (1936).
[25] C. Gundlach and J. M. Martín-García, Classical Quantum Gravity 23, S387 (2006).
[26] D. Hilditch, Int. J. Mod. Phys. A28, 1340015 (2013).
[27] D. Lynden-Bell and J. Bičák, Phys. Rev. D 96, 104053 (2017).
[28] S. Chakraborty and T. Padmanabhan, Phys. Rev. D 101, 064023 (2020), arXiv:1909.00096 [gr-qc].
[29] J. M. Martín-García, “xAct: tensor computer algebra.” (2020), http://www.xact.es/.