Spontaneous generation of geometry in four dimensions

Jorge Alfaro

Pontificia Universidad Católica de Chile, Avenida Vicuña Mackenna 4860, Santiago, Chile

Domènec Espriu and Daniel Puigdomènech

*Departament d'Estructura i Constituents de la Matèria, Institut de Ciències del Cosmos (ICCUB),
Universitat de Barcelona, Martí i Franquès, 1, 08028 Barcelona, Spain*

We present the extension to four dimensions of a Euclidean two-dimensional model that exhibits spontaneous generation of a metric. In this model gravitons emerge as Goldstone bosons of a global $SO(4) \times GL(4)$ symmetry broken down to $SO(4)$. The microscopic theory can be formulated without having to appeal to any particular space-time metric and only assumes the preexistence of a manifold endowed with an affine connection. In this sense the microscopic theory is quasitopological. There are indications that suggest that the model may be renormalizable, a fact that seems to be due to the impossibility of constructing counterterms without a preexisting metric. The vierbein appears as a condensate of the fundamental fermions defining a vacuum that, if the background affine connection is set to zero, is the Euclidean continuation of de Sitter space-time. If perturbatively small affine connections are introduced on this background, fluctuations of the metric (i.e., gravitons) appear and are described by an effective theory at long distances whose more relevant operators correspond to the Einstein-Hilbert action with a cosmological constant. This effective action is derived in the large N limit, N being the number of fermion species in the fundamental theory. The counterterms required by the microscopic theory are directly related to the cosmological constant and Newton constant and their couplings could eventually be adjusted to the physical values of M_p and Λ. The relevance of higher-dimensional operators is also briefly discussed.

I. INTRODUCTION

It has been pointed out several times in the literature (see, e.g., [1–4]) that gravitons could perhaps be considered as Goldstone bosons of some broken symmetry. The nonrenormalizability of gravity would therefore be analogous to the bad ultraviolet properties of effective models of hadrons, for instance. This is exactly the point of view that we adopt in this paper. This work is the extension to four dimensions of a previous analysis in two dimensions [5].

In our view, concrete implementations of this idea have been lacking so far (see however [6,7]). By concrete proposal we mean a field theory that does not contain the graviton field as an elementary degree of freedom. It should not even contain the tensor $\eta_{\mu\nu}$ either, as this already implies the use of some background metric and thus the notion of geometry; gravitons being fluctuations around this flat background. Instead, one would like to see all the metric degrees of freedom emerging dynamically, like pions appear dynamically after chiral symmetry breaking in QCD. We would also prefer the underlying theory to be in some sense "simpler" than gravity, renormalizable if possible. It was shown in [5] that a theory with all these characteristics can be found in two dimensions.

The purpose of the present work is to extend the $D = 2$ model to the far more interesting case of $D = 4$. In spite of complications brought up by the higher dimensionality,

most of the nice features of the $D = 2$ case seem to persist in $D = 4$. The amount of divergent terms remains under control, at least on shell, and the use of the equations of motion ensures that the final effective theory is precisely that of Einstein-Hilbert plus a cosmological term, complemented by higher-dimensional terms (in the sense of terms containing higher derivatives of the induced metric) normalized by a computable dimensionful constant. We cannot claim at this point that the model here introduced is renormalizable, but we will provide some indications in this direction. Verification of this point beyond the calculations presented in this work remains a very interesting issue in our view.

The topic of "emergent gravity" has recently been a popular one [8] but with the exceptions of the earlier works of [6,7] we have been unable to find any proposal meeting our rather restricted criteria of not assuming any preexisting metric structure whatsoever. On the other hand the proposals in [6,7] appear to be intractable and quantitative results are hard or impossible to get. Of course, there is nothing wrong in assuming a preexisting metric (such as Minkowski or Euclidean), but in a scenario where gravitons emerge dynamically it appears natural that the theory should select itself the vacuum where it chooses to live. Therefore we shall not assume any predefined metric in space-time and assume only the minimal structure of endowing space-time with an affine connection. Once a

metric is spontaneously generated, the equations of motion relate the induced metric to the connection.

Traditionally a major stumbling block in the program that will be developed below has been the so-called Weinberg-Witten theorem [9] (see also [10]). The apparent pathology of theories intending to generate dynamically gauge bosons (including gravitons in this category) lies in the fact that the energy-momentum tensor has to be identically zero if massless particles with spin ≥ 1 appear and one insists on the energy-momentum tensor being Lorentz covariant. However, our results, while not constituting a mathematical proof, indicate that one can indeed get in $D = 4$ an effective low-energy theory with massless composite gravitons, so it is legitimate to ask why the Weinberg-Witten theorem would not apply. We note something peculiar to our proposal; namely, the energy-momentum tensor (derived in [5] for $D = 2$ and which is exactly the same in $D = 4$) does not have tangent (Lorentz) indices. In fact Lorentz indices are of an internal nature in the present approach as we will see below. The connection between Lorentz and world indices appears only after a vierbein is dynamically generated. But then one is exactly in the same situation as general relativity where the applicability of [9] is excluded.

The dynamical generation of geometry, combined with the usual renormalization group arguments have rather interesting consequences. Geometry and distance are induced rather than fundamental concepts. At sufficiently short scales, when the effective action does not make sense anymore, the physical degrees of freedom are fermionic. Below that scale there is not even the notion of distance: in a sense that is the shortest scale that can exist. This precludes the existence of an ultraviolet fixed point advocated by some [11] but also indicates that at short distances gravity is non-Wilsonian as suggested by others [12] in an holographic context.

This paper is organized as follows: in Sec. II we review the similarities and differences between the $D = 2$ and $D = 4$ theories and discuss possible counterterms. In Sec. III we study the equations of motion and revise the calculational set-up. Section IV is devoted to the explicit one-loop calculation of the one, two and three-point functions of the model. In Sec. V we summarize the divergent content of the effective theory and make use of the equations of motion to prove the on shell renormalizability of the model at leading order. In Sec. VI the final effective action is written down and the connection between the free constants of the theory and the physical universal constants of gravity is discussed. Finally in Sec. VII we summarize our results and discuss possible extensions.

II. TWO-DIMENSIONAL MODEL REVIEW AND FOUR-DIMENSIONAL EXTENSION

It was shown in [5] that a consistent and apparently renormalizable $D = 2$ model reproducing gravity at long

distances could be built. The same model can be considered in $D = 4$. The free Lagrangian density is

$$
\begin{aligned}
\mathcal{L}_0 = {} & i\bar{\psi}_a \gamma^a (\partial_\mu + iw_\mu^{bc}\sigma_{bc})\chi^\mu \\
& + i\bar{\chi}^\mu \gamma^a (\partial_\mu + iw_\mu^{bc}\sigma_{bc})\psi_a,
\end{aligned}
\tag{1}
$$

where ψ_a and χ^μ are two species of fermions transforming, respectively, under Lorentz $(a, b \ldots$ are the tangent indices, which can be considered internal ones for our purposes) and Diffeomorphisms $(\mu, \nu \ldots$ are world indices labelling the manifold coordinates x^μ, globally defined on the manifold, with tangent vectors taken to be orthonormal with respect to the tangent space $SO(D)$ metric). A spin connection is added to the derivative to preserve the Lorentz $\times$ Diff symmetry[1] under local coordinate transformations. It is important to notice that there is no need to have a metric defined in the manifold as long as χ^μ transforms as a spinorial density because then $\Gamma_{\nu\rho}^\mu$ does not enter the covariant derivative, only w_μ^{ab}. If we keep this spin connection fixed, i.e., we do not consider it to be a dynamical field for the time being, there is no invariance under general coordinate transformations, but only under the global group $G = SO(D) \times GL(D)$. Notice once more that the spin connection is the only geometrical quantity introduced.

The interaction term in the model, in Euclidean conventions, is provided by

$$
\mathcal{L}_I = iB_\mu^a(\bar{\psi}_a \chi^\mu + \bar{\chi}^\mu \psi_a) + c \det(B_\mu^a),
\tag{2}
$$

which obviously does not require any metric to be formulated either, assuming that B_μ^a transforms as a vector under Lorentz $\times$ Diff. We will assume that we have N species of the previous fermions but we will not add an additional index to avoid complicating the notation. We emphasize that Lorentz symmetry acts as an internal symmetry at this point.

The object of the interaction (2) is to trigger the spontaneous breaking of the global symmetry via fermion condensation. Upon use of the equations of motion for the auxiliary field B_μ^a

$$
\bar{\psi}_a \chi^\mu + \bar{\chi}^\mu \psi_a = -ic \frac{1}{(D-1)!} \epsilon_{aa_2\ldots a_D} \epsilon^{\mu\mu_2\cdots\mu_D} B_{\mu_2}^{a_2}\ldots B_{\mu_D}^{a_D}
\tag{3}
$$

and thus

$$
\langle \bar{\psi}_a \chi^\mu + \bar{\chi}^\mu \psi_a \rangle \neq 0 \Rightarrow \langle B_\mu^a \rangle \neq 0.
\tag{4}
$$

---

[1] We actually use Euclidean conventions but still refer to $SO(D)$ as Lorentz symmetry. Note that (1) is not the usual Dirac coupling of fermions to a connection (that requires a metric). The field χ^μ has a spin $1/2$ and $3/2$ components in general, although this statement makes little sense unless a metric in the manifold is defined.

If a nonzero value for the fermion condensate appears then the field B_μ^a necessarily acquires a nonzero expectation value B_μ^a (the reciprocal is not necessarily true, but it will be true too in our case). Such condensation was seen to happen in $D = 2$ (in the large N limit) and it will be present in $D = 4$ even for finite values of N as we will see in a moment thus defining a nontrivial vacuum. Small perturbations above this vacuum will yield the effective theory of the quantum excitations of the theory. We will use a perturbative approach corresponding to a weak field expansion around the solution for $w_\mu^{ab} = 0$; the value of the connection appears implicitly on the left-hand side of (3). We shall first consider the case $w_\mu^{ab} = 0$.

If $w_\mu^{ab} = 0$ the vacuum of the theory is expected to be translational invariant, i.e., we should obtain a constant value for B_μ^a, possibly zero. This constant value is obtained from the gap equation derived from the effective potential

$$V_{\text{eff}} = c \det(B_\mu^a) - 2N \int \frac{d^D k}{(2\pi)^D} \text{tr}(\log(\gamma^a k_\mu + iB_\mu^a)). \quad (5)$$

Note that the $2N$ preceding the integral comes from the $2N$ species of fermions present. As for the flat measure used for the integration, this corresponds to the functional trace of the differential operator.[2] As such the trace is independent of the particular basis that is used to compute it. Any other basis, if used correctly, would yield the same result.

Deriving (5) with regard to B_μ^a we obtain

$$c \frac{D}{D!} \epsilon_{aa_2\ldots a_D} \epsilon^{\mu\mu_2\ldots\mu_D} B_{\mu_2}^{a_2} \ldots B_{\mu_D}^{a_D}$$
$$- 2Ni \, \text{tr} \int \frac{d^D k}{(2\pi)^D} (\gamma^a k_\mu + iB_\mu^a)^{-1}|_a^\mu = 0. \quad (6)$$

This equation has a general nontrivial solution corresponding to $B_\mu^a = M\delta_\mu^a$ (or any $SO(D) \times GL(D)$ global transformations of this). This is analogous to the more familiar phenomenon of chiral symmetry breaking in strong interactions and any value of B_μ^a in the $SO(D) \times GL(D)$ orbit is equivalent. For simplicity we will take $B_\mu^a = M\delta_\mu^a$ and then the gap equation reduces to an equation for M. In $D = 4$

$$cM^3 - 2N \int \frac{d^D k}{(2\pi)^D} \frac{M}{k^2 + M^2} = 0$$
$$cM^3 + N\frac{M^3}{8\pi^2}\left(\frac{2}{\epsilon} - \log\frac{M^2}{4\pi\mu^2} - \gamma + 1\right) = 0, \quad (7)$$

whose formal solution is

$$M^2 = \mu^2 e^{8\pi^2 c(\mu)/N}, \quad (8)$$

where $\mu\frac{dc}{d\mu} = -\frac{N}{4\pi^2}$, making M a renormalization-group invariant. In the previous, we introduced the usual mass scale μ to preserve the correct dimensionality of the D-dimensional integral as dimensional regularization is used. For the solution to actually exist we have to require $c > 0$ if $M > \mu$. If $\mu > M$ the solution exists only if $c < 0$. Therefore $c > 0$ will be the case we are interested in on physical grounds.

Note that B_μ^a has the right structure to be identified as the vierbein, and as it was shown in [5], it consistently reemerges in the $D = 2$ effective theory to form the determinant of the spontaneously generated metric.

The free fermion propagator of the theory in the broken phase can then be easily found after replacing B_μ^a by its vacuum expectation value. With a $D = 4$ matrix notation

$$\Delta^{-1}(k)^i{}_j = \frac{-i}{M}\left(\delta^i{}_j - \frac{\gamma^i(\slashed{k} - iM)k_j}{k^2 + M^2}\right). \quad (9)$$

A particularity of $D = 2$ was that the most general form for B_μ^a (in Euclidean conventions) is a conformal factor times a scale M times a δ_μ^a. This means that perturbations around the minimum of the potential can only have one physical degree of freedom, the conformal parameter. The other degrees of freedom in B_μ^a can be removed by suitable coordinate transformations and are thus unphysical (recall that the microscopic theory is fully generally covariant—even without a metric).

The main difference of the $D = 4$ case with respect to the $D = 2$ case is that the maximum number of possible physical degrees of freedom for a perturbation around the value $B_\mu^a = M\delta_\mu^a$ grows up to six instead of one, making the calculation much more complex. Clearly, considering a uniparametric family of perturbations is far too simple in $D = 4$ and does not yield enough information to find the long-distance effective action unambiguously. To bypass this difficulty, but still keeping the calculation manageable, we have chosen to restrict our considerations to diagonal perturbations, where

$$B_j^i(x) = M\delta_j^i e^{-(\sigma_i(x)/2)} \quad \text{(no sum over} i). \quad (10)$$

This form contains 4 degrees of freedom (rather than six) but is rich enough for our purposes.[3] The validity of our conclusions rely on the assumption that the effective action should be covariant (exactly as the microscopic theory is). This was actually checked in the $D = 2$ case using heat kernel techniques. Here we have performed partial checks but we have to assume that covariance holds to draw our conclusions.

---

[2] Note that "plane waves" are eigenmodes of the differential operator $\gamma^a\nabla_\mu$ if the connection w_μ^{ab} is set to zero. The connection itself is treated perturbatively in the subsequent.

[3] Note that these perturbations do not correspond to pure gauge degrees of freedom as they lead to nonzero values for the curvature, which is gauge-invariant; i.e., they necessarily involve physical degrees of freedom.

Note that once a dynamical value for e^a_μ is generated we can write terms such as $M\bar{\psi}_a e^a_\mu \chi^\mu$, where e^a_μ would correspond to the vierbein. A large number of Goldstone bosons are produced. The original symmetry group $G = SO(4) \times GL(4)$ has $\frac{4(4-1)}{2} + 4^2$ generators. After the breaking $G \rightarrow H$, with $H = SO(4)$, there are 16 broken generators, as expected. Since the metric must be symmetric, at most 10 Goldstone bosons can enter the perturbation. Four of those can be removed by a gauge choice, leaving the before mentioned six. Finally within each gauge choice a residual gauge freedom will in general allow for the removal of four more. In this respect our counting is analogous to the one in general relativity. The final number of physical degrees of freedom will be two.

Another difference with respect to the $D = 2$ case is that the integrals involved in the perturbative calculation have potentially a much worse ultraviolet behavior in $D = 4$. We postpone to Secs. III and IV the explicit calculations that indicate that the nice characteristics found in the $D = 2$ model, in particular, renormalizability, seem to persist in the $D = 4$ case. However the ultimate reason for the apparent renormalizability lies in the very limited number of counterterms that can be written without a metric (and the usual assumption that the ultraviolet behavior is unaltered by the phenomenon of spontaneous symmetry breaking).

A. Possible counterterms

Before tackling the perturbative derivation of the effective action it is important to list the possible invariants one can write in this theory without making use of a metric. In $D = 2$ we had, as already explained, two invariants that could be constructed without having to appeal to a metric, namely $\mathcal{L}_0$ and $\mathcal{L}_I$. The latter, upon use of the parametrization $B^a_\mu = M\delta^a_\mu e^{-(\sigma/2)}$, reduces to

$$\frac{1}{2!} \int B^a_\mu B^b_\nu \epsilon^{\mu\nu} \epsilon_{ab} d^2x = M^2 \int \sqrt{g} d^2x, \qquad (11)$$

i.e., is the cosmological term. In addition, there is the curvature term which in $D = 2$ in terms of the connection is simply $\int d^2x dw$, purely topological; therefore we do not expect it to appear in the perturbative calculation. Then, apart from the free kinetic term for the fermions, there is only one invariant term that can be written down without a metric. Or what is tantamount, only one possible counterterm remains to absorb any divergence appearing in the perturbative calculation after integrating out the fermions. This fact enforces the renormalizability of the $D = 2$ model in the large N limit in spite of the bad ultraviolet behavior of the integrals. This argument was supported by the explicit calculations presented in [5].

Now in $D = 4$ we can write, in addition to $\mathcal{L}_0$ and $\mathcal{L}_I$, one more counterterm

$$\mathcal{S}_R = \frac{1}{2} \int R_{[\mu\nu]ab} B^a_\rho B^b_\sigma \epsilon^{\mu\nu\rho\sigma} d^4x, \qquad (12)$$

where $R_{[\mu\nu]ab} = [\nabla_{\mu ac}, \nabla_{\nu cb}]$. After integrating the fermion fields only

$$\mathcal{S}_D = \frac{1}{4!} \int B^a_\mu B^b_\nu B^c_\rho B^d_\sigma \epsilon_{abcd} \epsilon^{\mu\nu\rho\sigma} d^4x, \qquad (13)$$

which was already present in $\mathcal{L}_I$, and $\mathcal{S}_R$ can appear as genuine divergences if general covariance is preserved. We will denote by $\mathcal{L}_D$ and $\mathcal{L}_R$ the respective Lagrangian densities.

Another counterterm one could write without making use of a metric, namely, the Gauss-Bonnet topological invariant in $D = 4$, which is of $\mathcal{O}(p^4)$ in the usual momentum counting.

We did not include the term $\mathcal{S}_R$ in our action to start with because it does not contain the fermionic fields. It does not modify the equation of motion (3) for the auxiliary field B^a_μ or the gap Eq. (7) either, if the connection w_μ is set to zero as we did in the previous section (recall that we use a weak field expansion and $w_\mu = 0$ is used to determine the vacuum). However, we see that $\mathcal{S}_R$ is an allowed counterterm in $D = 4$ and therefore it needs to be included in the initial action. In fact, any divergence in the theory must be reabsorbable in the two terms $\mathcal{S}_D$ and $\mathcal{S}_R$, as they are the only local counterterms one can write before the symmetry breaking, i.e., before the generation of the metric.

When the auxiliary field B^a_μ is identified with the vierbein, the parametrization (10) and the equations of motion are used the two counterterms reduce to

$$M^4 \int \sqrt{g} d^4x, \qquad M^2 \int \sqrt{g} R d^4x, \qquad (14)$$

respectively; i.e., the familiar cosmological and Einstein terms. This will be explained in more detail in the next section.

III. EQUATIONS OF MOTION

Let us write explicitly what the $D = 4$ counterterms look like once we replace B^a_μ by its vacuum expectation value plus perturbations around it. To keep the notation simple, let us consider the case in (10) when $\sigma_i(x) = \sigma(x)$ (conformally flat perturbation). After substituting the solution $B^a_\mu = M e^a_\mu = M e^{-\sigma/2} \delta^a_\mu$ we have

$$\mathcal{L}_D = \frac{1}{4!} B^a_\mu B^b_\nu B^c_\rho B^d_\sigma \epsilon_{abcd} \epsilon^{\mu\nu\rho\sigma} = M^4 e^{-2\sigma} \qquad (15)$$

and

$$\frac{1}{2} R_{[\mu\nu]ab} B^a_\rho B^b_\sigma \epsilon^{\mu\nu\rho\sigma}$$
$$= \frac{1}{2}(\partial_\mu w_\nu{}^{\mu\nu} - \partial_\nu w_\mu{}^{\mu\nu} + w_\mu{}^{\mu c} w_\nu{}^{c\nu} - w_\nu{}^{\mu c} w_\mu{}^{c\nu}) e^{-\sigma} M^2.$$
$$(16)$$

Note that because in the vacuum solution $e_\mu^a = \delta_\mu^a$ we can use indiscriminately Greek and Latin indices; they are lowered and raised with a metric proportional to the identity.

Let us now work out the equations of motion for the full Lagrangian $\mathcal{L}_0 + \mathcal{L}_I + \mathcal{L}_R$. In Sec. II we already discussed the equations of motion for the field B_μ^a when $w_\mu^{ab} = 0$. In addition we have

$$\frac{\delta \mathcal{L}}{\delta w_\mu{}^{ab}} = \partial_\rho \left(\frac{\delta(\mathcal{L}_0 + \mathcal{L}_I + \mathcal{L}_R)}{\delta \partial_\rho w_\mu{}^{ab}} \right) - \frac{\delta(\mathcal{L}_0 + \mathcal{L}_I + \mathcal{L}_R)}{\delta w_\mu{}^{ab}}$$

$$= \frac{1}{2} (-\partial_a \sigma \delta_b^\mu + \partial_b \sigma \delta_a^\mu - \delta_a^\mu w_{\nu b}{}^\nu$$

$$- \delta_b^\mu w_\nu{}^\nu{}_a + w_{ab}{}^\mu + w_b{}^\mu{}_a) e^{-\sigma}$$

$$- \frac{1}{M^2} (\bar{\psi}_c \gamma^c \sigma_{ab} \chi^\mu + \bar{\chi}^\mu \gamma^c \sigma_{ab} \psi_c) = 0. \quad (17)$$

To solve (17) we will only consider the lowest-order term in the $1/M^2$ expansion, following the usual counting rules in effective Lagrangians based on a momentum expansion. The solution for the connection in a conformally flat metric is then

$$w_\mu{}^{ab} = \frac{1}{2} (\partial^a \sigma \delta_\mu^b - \partial^b \sigma \delta_\mu^a). \quad (18)$$

This is the relation one obtains from the usual condition between the spin connection and the vierbein in general relativity (19), characteristic of the Palatini formalism [13]

$$w_\mu^{ab} = e_\nu^a \partial_\mu E^{\nu b} + e_\nu^a E^{\rho b} \Gamma_{\mu\rho}^\nu \quad (19)$$

particularized to a conformally flat metric given by $e_\mu^a = \delta_\mu^a e^{-(\sigma/2)}$ ($E^{\rho b}$ is the inverse vierbein). Making use of (18) in (16), that is on shell, we are now allowed to identify the curvature in terms of the scalar field σ

$$\mathcal{L}_R|_{(\text{on shell})} = M^2 \sqrt{g} R = \frac{3}{2} \left(\Box \sigma - \frac{1}{2} \partial_\mu \sigma \partial^\mu \sigma \right) e^{-\sigma} M^2. \quad (20)$$

Note that in the particular case of a vierbein corresponding to a conformally flat metric one can integrate by parts either of the terms in (20) to obtain the other

$$\sqrt{g} R = \frac{3}{2} M^2 \left(\Box \sigma - \frac{1}{2} \partial_\mu \sigma \partial^\mu \sigma \right) e^{-\sigma}$$

$$= \frac{3}{4} M^2 (\Box \sigma) \left(1 - \sigma + \frac{\sigma^2}{2} - \frac{\sigma^3}{6} + \cdots \right) \quad (21)$$

This term plus a constant times (15) are the only divergences that should appear in the final effective theory upon integration of the fermionic fields for this particular type of perturbations above the vacuum (i.e., those interpretable as a conformally flat metric).

As previously mentioned we shall consider a more general type of perturbations; namely, we will use the diagonal parametrization of the perturbations around the vacuum solution given by (10). This is not the most general one in $D = 4$, but it is enough for our purposes. After the identification of B_μ^a with the vierbein, this corresponds to a metric

$$g_{\mu\nu} = \begin{pmatrix} e^{-\sigma_1(x)} & 0 & 0 & 0 \\ 0 & e^{-\sigma_2(x)} & 0 & 0 \\ 0 & 0 & e^{-\sigma_3(x)} & 0 \\ 0 & 0 & 0 & e^{-\sigma_4(x)} \end{pmatrix}. \quad (22)$$

This parametrization provides enough generality to the calculation. We can now derive the equivalent expression to (18) for the general diagonal perturbation using (19) to obtain[4]

$$w_\mu{}^{ab} = \frac{1}{2} \left[\begin{pmatrix} e^{(\sigma_1/2)} & 0 & 0 & 0 \\ 0 & e^{(\sigma_2/2)} & 0 & 0 \\ 0 & 0 & e^{(\sigma_3/2)} & 0 \\ 0 & 0 & 0 & e^{(\sigma_4/2)} \end{pmatrix}^{a\rho} \begin{pmatrix} \partial_\rho \sigma_1 e^{-(\sigma_1/2)} & 0 & 0 & 0 \\ 0 & \partial_\rho \sigma_2 e^{-(\sigma_2/2)} & 0 & 0 \\ 0 & 0 & \partial_\rho \sigma_3 e^{-(\sigma_3/2)} & 0 \\ 0 & 0 & 0 & \partial_\rho \sigma_4 e^{-(\sigma_4/2)} \end{pmatrix}^b_\mu \right.$$

$$\left. - \begin{pmatrix} e^{(\sigma_1/2)} & 0 & 0 & 0 \\ 0 & e^{(\sigma_2/2)} & 0 & 0 \\ 0 & 0 & e^{(\sigma_3/2)} & 0 \\ 0 & 0 & 0 & e^{(\sigma_4/2)} \end{pmatrix}^{b\rho} \begin{pmatrix} \partial_\rho \sigma_1 e^{-(\sigma_1/2)} & 0 & 0 & 0 \\ 0 & \partial_\rho \sigma_2 e^{-(\sigma_2/2)} & 0 & 0 \\ 0 & 0 & \partial_\rho \sigma_3 e^{-(\sigma_3/2)} & 0 \\ 0 & 0 & 0 & \partial_\rho \sigma_4 e^{-(\sigma_4/2)} \end{pmatrix}^a_\mu \right]. \quad (23)$$

Making use of the equations of motion one can compute the corresponding $\mathcal{L}_R$ for the general case and expand it in the σ fields. The result up to two sigma fields reads

---

[4]Again we emphasize that although it may seem strange to see latin indices in the derivatives this should not confuse the reader. After the symmetry breaking a vierbein is generated relating world indices with tangent space ones through δ_μ^a. In expression (23) we have compiled the entries for w_μ^{ab} in a bimatrix form, but they *should not* be multiplied; only the index ρ is summed up.

$$\mathcal{L}_R|_{(\text{on shell})} = M^2\sqrt{g}R$$
$$= M^2\Big[\partial_3^2\sigma_4 + \partial_2^2\sigma_4 + \partial_1^2\sigma_4 + \partial_4^2\sigma_3 + \partial_2^2\sigma_3 + \partial_1^2\sigma_3 + \partial_4^2\sigma_2 + \partial_3^2\sigma_2 + \partial_1^2\sigma_2 + \partial_4^2\sigma_1 + \partial_3^2\sigma_1 + \partial_2^2\sigma_1$$
$$- \tfrac{1}{2}(\partial_3\sigma_1\partial_3\sigma_2 + \partial_4\sigma_1\partial_4\sigma_2 + \partial_2\sigma_1\partial_2\sigma_3 + \partial_4\sigma_1\partial_4\sigma_3 + \partial_2\sigma_1\partial_2\sigma_4 + \partial_3\sigma_1\partial_3\sigma_4 + \partial_1\sigma_2\partial_1\sigma_3 + \partial_4\sigma_2\partial_4\sigma_3$$
$$+ \partial_1\sigma_2\partial_1\sigma_4 + \partial_3\sigma_2\partial_3\sigma_4 + \partial_1\sigma_3\partial_1\sigma_4 + \partial_2\sigma_3\partial_2\sigma_4) + \mathcal{O}(\sigma^3)\Big]. \tag{24}$$

More details on the calculation of (24) can be found in the Appendix. Ignoring for a moment the Gauss-Bonnet invariant, the divergent terms from the perturbative calculation for the general perturbation should match on shell either with (24) or with

$$\mathcal{L}_D|_{(\text{on shell})} = M^4\sqrt{g} = M^4 e^{-(\sum_i \sigma_i/2)}. \tag{25}$$

An extension to the most general perturbation with the full 6 degrees of freedom should be possible but would require much more effort, which we consider unnecessary at this point as the above parametrization provides enough redundancy. Since the coefficients for the terms in the effective action are universal there should be no loss of generality in the present approach. This of course assumes that general covariance is kept all along the derivation of the effective action and by the regulator, as it should be the case in dimensional regularization.

So far we have explained how the $D = 2$ model can be consistently extended to $D = 4$ preserving the key features. We study small perturbations around a constant vacuum expectation value for the field B_μ^a (which does not need to be small itself) corresponding to the solution of the gap equation for $w_\mu^{ab} = 0$. In such a theory one can write a limited number of counterterms without making use of a metric. These counterterms are consistent with the usual terms of GR once used the equations of motion. With all these ingredients we are ready to move to the actual perturbative derivation of the effective action.

IV. ONE-LOOP STRUCTURE FOR A GENERAL DIAGONAL PERTURBATION

The effective action that describes perturbations above the trivial vacuum

$$w_\mu^{ab} = 0, \qquad B_\mu^a = M\delta_\mu^a, \tag{26}$$

will be given by a polynomial expansion in powers of $w_\mu(x)$, $\sigma_i(x)$ and their derivatives obtained after integration of the fundamental degrees of freedom. In this section we will derive this effective action diagrammatically.

We shall use the diagonal perturbation (10) with four degrees of freedom for the vierbein perturbations. For

simplicity, we will calculate only the one-point and two-point functions for this rather general case and then particularize to the conformal case [$\sigma_i(x) = \sigma(x)$] to compute some three-point functions.
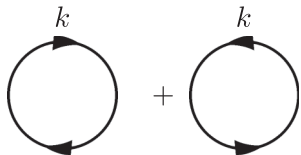
Since perturbation theory in this model has some peculiar features (note, in particular, the behavior of the fermion propagator) in what follows we shall provide enough details so that the diagrammatic calculation can be reproduced.

Starting from the Lagrangian density $\mathcal{L}_0 + \mathcal{L}_I$ described in Sec. II (note that $\mathcal{L}_R$ plays no role whatsoever in the integration of the N species of fermions), and using a parametrization of B_μ^a given by (10), the interaction vertices are

$$\sigma_i \dashrightarrow \qquad i\frac{1}{2}M\delta^i_{\ \mu}$$

$$\qquad -i\frac{1}{8}M\delta^i_{\ \mu} \tag{27}$$

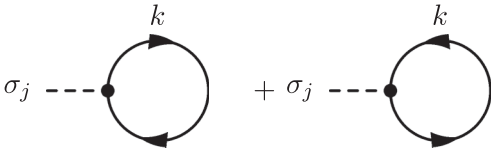$$\sigma_i \dashrightarrow \qquad i\frac{1}{48}M\delta^i_{\ \mu}$$

A. One- and two-point functions for the fields σ_i

With the rules described above and using the propagator (9) we can calculate the first one-loop diagrams for $D = 4 - \epsilon$. We will not include the factor N in the diagrammatic results presented below. The vacuum bubble diagram is
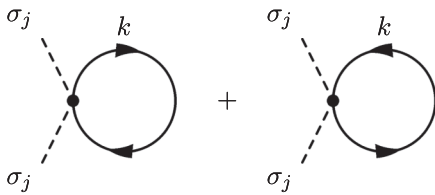
$$= -2\text{Tr}\left[\int \frac{d^D k}{(2\pi)^D}\Delta^{-1}(k)^\mu{}_a(-i)\delta^a{}_\mu\right] = -\frac{M^3}{2\pi^2}\left(\frac{2}{\epsilon} - \log\left(\frac{M^2}{4\pi\mu^2}\right) - \gamma + 1\right).$$

(28)

We also compute the one-point function for the different vertices

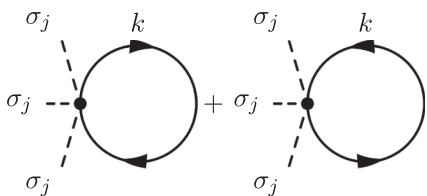$$= -\sum_{j=1}^{4}\sigma_j\frac{2}{2}\text{Tr}\left[\int \frac{d^D k}{(2\pi)^D}\Delta^{-1}(k)^\mu{}_j(i)M\delta^j{}_\mu\right] = \sum_{j=1}^{4}\frac{\sigma_j M^4}{16\pi^2}\left(\frac{2}{\epsilon} - \log\left(\frac{M^2}{4\pi\mu^2}\right) - \gamma + 1\right),$$

(29)

$$= -\sum_{j=1}^{4}\sigma_j^2\frac{2!\,2}{2!\,2\cdot 2}\text{Tr}\left[\int \frac{d^D k}{(2\pi)^D}\Delta^{-1}(k)^\mu{}_j(-i)M\delta^j{}_\mu\right] = -\sum_{j=1}^{4}\sigma_j^2\frac{M^4}{32\pi^2}\left(\frac{2}{\epsilon} - \log\left(\frac{M^2}{4\pi\mu^2}\right) - \gamma + 1\right).$$

(30)

Let us, for this particular diagram, clarify what the origin of the numerical factor is. In the numerator, the 2! comes from the combinatorial possible connections of the external fields σ_i. The other 2 is due to the two species of fermions and it is present in all diagrams. In the denominator, $2!2\cdot 2$ comes from the vertex. Since it is a one-point function there are no additional factors, however for n-point functions the corresponding $n!$ will be present in the denominator.

$$= -\sum_{j=1}^{4}\frac{3!\,2}{3!\,2\cdot 2\cdot 2}\text{Tr}\left[\int \frac{d^D k}{(2\pi)^D}\Delta^{-1}(k)^\mu{}_j(i)M\delta^j{}_\mu\right] = \sum_{j=1}^{4}\frac{\sigma_j^3 M^4}{64\pi^2}\left(\frac{2}{\epsilon} - \log\left(\frac{M^2}{4\pi\mu^2}\right) - \gamma + 1\right).$$

(31)

Next diagram is the two-point function

$$= -\sum_{j=1}^{4}\sum_{l=1}^{4}\sigma_j\sigma_l\frac{2!\,2}{2!\cdot 2\cdot 2}\mathrm{Tr}\left[\int\frac{d^Dk}{(2\pi)^D}iM\delta^l{}_\mu\Delta^{-1}(k)^\mu{}_j iM\delta^j{}_\nu\Delta^{-1}(k+p)^\nu{}_l\right]$$

$$= \sum_{j=1}^{4}\sum_{\substack{l=1\\ l\neq j}}^{4}\left[-\frac{\sigma_j\sigma_l M^4}{16\pi^2}\left(\frac{2}{\epsilon}-\log\left(\frac{M^2}{4\pi\mu^2}\right)-\gamma+\frac{2}{3}\right)-\frac{\sigma_j\sigma_l p^2 M^2}{48\pi^2}\left(\frac{2}{\epsilon}-\log\left(\frac{M^2}{4\pi\mu^2}\right)-\gamma-\frac{1}{3}\right)\right.$$

$$\left.+\frac{\sigma_j\sigma_l(p_j^2+p_l^2)M^2}{48\pi^2}\left(\frac{2}{\epsilon}-\log\left(\frac{M^2}{4\pi\mu^2}\right)-\gamma-\frac{1}{2}\right)\right]$$

$$+\sum_{j=1}^{4}\left[+\frac{\sigma_j^2 M^4}{32\pi^2}+\frac{\sigma_j^2 M^2 p^2}{32\pi^2}-\frac{\sigma_j^2 M^2 p_j^2}{32\pi^2}\right]+\mathcal{O}(p^4).$$

(32)

The numerical factor in this case is composed by 2! 2 in the numerator from the possible contractions of external fields times the two species of fermions. And 2! in the denominator from the diagram being a two-point function and finally the $1/2$ from each vertex. We will not elaborate on the combinatorial factors anymore but we write all factors explicitly, even if the notation may be a bit cumbersome, in order to facilitate the check of our results. By $\mathcal{O}(p^4)$ we mean finite higher order in p^2 contributions.

B. Diagrams with w_μ^{ab}

Now we turn to the diagrams that contain a field w_μ^{ab}. The corresponding vertex is

$$i\gamma^a w_\mu^{bc}\sigma_{bc}=\frac{\gamma^a}{4}[\gamma_b,\gamma_c]. \quad (33)$$

The one- and two-point functions yield

$$= -2\mathrm{Tr}\left[\int\frac{d^Dk}{(2\pi)^D}i\gamma^2\sigma_{bc}\Delta^{-1}(k)_a^\mu\right]=0 \quad (34)$$

then

$$= -\sum_{j=1}^{4}\sigma_j\frac{2}{2!\cdot 2}\mathrm{Tr}\left[\int\frac{d^Dk}{(2\pi)^D}i\gamma^a\sigma_{bc}\Delta^{-1}(k)^\mu{}_d iM\delta^d{}_\nu\Delta^{-1}(k+p)^\nu{}_a\right]=0. \quad (35)$$

This suggests that diagrams containing only one field w_μ^{ab} are zero. For two w_μ^{ab} fields we have

$$
= -\frac{2! \cdot 2}{2!} \operatorname{Tr} \left[\int \frac{d^D k}{(2\pi)^D} i\gamma^a \sigma_{bc} \Delta^{-1}(k)^\mu{}_d i\gamma^d \sigma_{ef} \Delta^{-1}(k+p)^\nu{}_a \right]
$$

$$
= \left(\frac{2}{\epsilon} - \log\left(\frac{M^2}{4\pi\mu^2}\right) - \gamma + \frac{1}{4} \right) \left[\frac{M^2}{4\pi^2} (\delta_{be}\delta_c^\nu \delta_f^\mu - \delta_{be}\delta_{cf}\delta^{\mu\nu} + \delta_{bf}\delta_{ce}\delta^{\mu\nu} - \delta_{bf}\delta_c^\nu \delta_e^\mu \right.
$$

$$
\left. + \delta_b^\nu \delta_{cf}\delta_e^\mu - \delta_b^\nu \delta_{ce}\delta_f^\mu) \right] + \frac{M^2}{16\pi^2} \left(\delta_{bf}\delta_e^\nu \delta_c^\mu - \delta_{be}\delta_c^\mu \delta_f^\nu + \delta_b^\mu \delta_{ce}\delta_f^\nu - \delta_b^\mu \delta_e^\nu \delta_{cf} \right)
$$

$$
+ \frac{1}{\epsilon} F^{\mu\nu}{}_{bcef}(p^2) + \mathcal{O}(p^2)
$$

$$
= \left(\frac{2}{\epsilon} - \log\left(\frac{M^2}{4\pi\mu^2}\right) - \gamma + \frac{1}{4} \right) \left[\frac{M^2}{4\pi^2} D^{\mu\nu}{}_{bcef} \right] + \frac{M^2}{16\pi^2} E^{\mu\nu}{}_{bcef} + \frac{1}{\epsilon} F^{\mu\nu}{}_{bcef}(p^2) + \mathcal{O}(p^2) \tag{36}
$$

where $F^{\mu\nu}{}_{bcef}$ is a complicated structure composed of external momenta and Kronecker deltas of order $\mathcal{O}(p^2)$. This divergence is of higher order, in the $1/M^2$ expansion, than the one of $D^{\mu\nu}{}_{bcef}$. Now, taking into account that $w_\mu^{bc} = -w_\mu^{cb}$, we can show that

$$
\begin{aligned}
D^{\mu\nu}{}_{bcef} w_\mu^{bc} w_\nu^{ef} &= 4 w_\mu^{\nu b} w_\nu^{\mu b} - 2 w_\mu^{be} w_\mu^{be} = 0; \\
E^{\mu\nu}{}_{bcef} w_\mu^{bc} w_\nu^{ef} &= 4 w_\mu^{\mu b} w_\nu^{b\nu}
\end{aligned} \tag{37}
$$

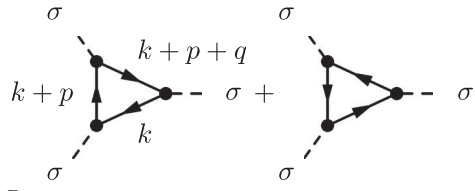More details can be found in the Appendix where we will show how the combination of the divergences proportional to $F^{\mu\nu}{}_{bcef}$ appearing in this diagram combine to reproduce exactly the $\mathcal{O}(p^4)$ terms of the Gauss-Bonnet invariant.

C. Three-point functions

In order to keep the calculations simple, we particularize to the case $B_\mu^a = M e^{-(\sigma(x)/2)} \delta_\mu^a$. The previous results, (28)–(37) are all valid taking $\sigma_i = \sigma$, $i = 1, 2, 3, 4$. With this simplification we can easily further compute more diagrams. For the field σ we have
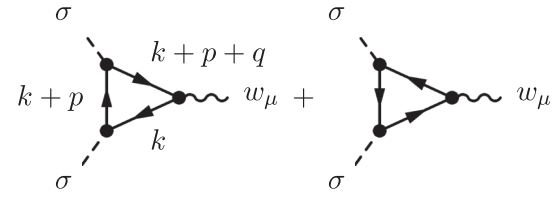

$$
= -\frac{2\,2\,2}{2!\,2!\cdot 2\cdot 2\cdot 2} \operatorname{Tr} \left[\int \frac{d^D k}{(2\pi)^D} (-i) M \delta_\mu^a \Delta^{-1}(k)^\mu{}_b i M \delta_\nu^b \Delta^{-1}(k+p+q)^\nu{}_a \right]
$$

$$
- \frac{2\,2\,2}{2!\,2!\cdot 2\cdot 2\cdot 2} \operatorname{Tr} \left[\int \frac{d^D k}{(2\pi)^D} (-i) M \delta_\mu^a \Delta^{-1}(k+p)^\mu{}_b i M \delta_\nu^b \Delta^{-1}(k+p+q)^\nu{}_a \right]
$$

$$
- \frac{2\,2\,2}{2!\,2!\cdot 2\cdot 2\cdot 2} \operatorname{Tr} \left[\int \frac{d^D k}{(2\pi)^D} (-i) M \delta_\mu^a \Delta^{-1}(k+q)^\mu{}_b i M \delta_\nu^b \Delta^{-1}(k+p+q)^\nu{}_a \right]
$$

$$
= \frac{9M^4}{16\pi^2} \left(\frac{2}{\epsilon} - \log\left(\frac{M^2}{4\pi\mu^2}\right) - \gamma + \frac{1}{3} \right)
$$

$$
+ \frac{M^2(p^2 + (p+q)^2 + q^2)}{32\pi^2} \left(\frac{2}{\epsilon} - \log\left(\frac{M^2}{4\pi\mu^2}\right) - \gamma - \frac{2}{3} \right) + \mathcal{O}(p^4) \tag{38}
$$

And also

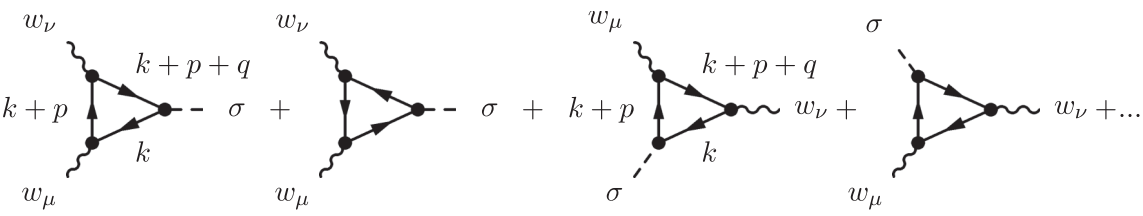$$= -\frac{3! \, 2}{3! \, 2 \cdot 2 \cdot 2} \text{Tr} \left[\int \frac{d^D k}{(2\pi)^D} iM\Delta^{-1}(k)^{\mu}{}_b iM\delta^b{}_{\nu}\Delta^{-1}(k+p+q)^{\nu}{}_d iM\delta^d{}_{\rho}\Delta^{-1}(k+p)^{\rho}{}_a \right]$$

$$- \frac{3! \, 2}{3! \, 2 \cdot 2 \cdot 2} \text{Tr} \left[\int \frac{d^D k}{(2\pi)^D} iM\Delta^{-1}(k)^{\mu}{}_b iM\delta^b{}_{\nu}\Delta^{-1}(k+p+q)^{\nu}{}_d iM\delta^d{}_{\rho}\Delta^{-1}(k+q)^{\rho}{}_a \right]$$

$$= \frac{3M^4}{8\pi^2} \left(\frac{2}{\epsilon} - \log \left(\frac{M^2}{4\pi\mu^2} \right) - \gamma - \frac{2}{3} \right) - M^2 \left(\frac{p^2}{32\pi^2} + \frac{(p+q)^2}{32\pi^2} + \frac{q^2}{32\pi^2} \right) + \mathcal{O}(p^4) \qquad (39)$$

On the other hand, for the field w_{μ}^{ab} we can compute

$$= -\frac{2! \, 2}{3! \, 2 \cdot 2} \text{Tr} \left[\int \frac{d^D k}{(2\pi)^D} i\gamma^a \sigma_{bc}\Delta^{-1}(k)^{\mu}{}_d iM\delta^d{}_{\nu}\Delta^{-1}(k+p+q)^{\nu}{}_e iM\delta^e{}_{\rho}\Delta^{-1}(k+p)^{\rho}{}_a \right]$$

$$- \frac{2! \, 2}{3! \, 2 \cdot 2} \text{Tr} \left[\int \frac{d^D k}{(2\pi)^D} i\gamma^a \sigma_{bc}\Delta^{-1}(k)^{\mu}{}_d iM\delta^d{}_{\nu}\Delta^{-1}(k+p+q)^{\nu}{}_e iM\delta^e{}_{\rho}\Delta^{-1}(k+q)^{\rho}{}_a \right] = 0. \qquad (40)$$

And finally

$$= -\frac{3 \, 2! \, 2}{3! \cdot 2} \text{Tr} \left[\int \frac{d^D k}{(2\pi)^D} i\gamma^a \sigma_{bc}\Delta^{-1}(k)^{\mu}{}_d \gamma^d \sigma_{ef}\Delta^{-1}(k+p+q)^{\nu}{}_g iM\delta^g{}_{\rho}\Delta^{-1}(k+p)^{\rho}{}_a \right]$$

$$- \frac{3 \, 2! \, 2}{3! \cdot 2} \text{Tr} \left[\int \frac{d^D k}{(2\pi)^D} i\gamma^a \sigma_{bc}\Delta^{-1}(k)^{\mu}{}_d i\gamma^d \sigma_{ef}\Delta^{-1}(k+p+q)^{\nu}{}_g iM\delta^g{}_{\rho}\Delta^{-1}(k+q)^{\rho}{}_a \right]$$

$$= \left(\frac{2}{\epsilon} - \log \left(\frac{M^2}{4\pi\mu^2} \right) - \gamma - \frac{3}{4} \right) \left[-\frac{M^2}{4\pi^2} D^{\mu\nu}{}_{bcef} \right] - \frac{M^2}{16\pi^2} E^{\mu\nu}{}_{bcef} + \mathcal{O}(p^2). \qquad (41)$$

With $D^{\mu\nu}{}_{bcef}$ and $E^{\mu\nu}{}_{bcef}$ being the same as in (36).

V. SUMMARY OF DIVERGENCES

In the previous section we obtained the results of the one-, two- and three-point functions for a general diagonal perturbation, sometimes particularizing to a conformally flat metric to ease the notation. Let us now summarize the results.

The divergent part of diagrams (28)–(31) together with the M^4 piece of diagram (32) add up in the effective action to[5]

$$\frac{M^4 e^{-\left(\sum_i \sigma_i/2\right)}}{8\pi^2}\left(\frac{2}{\epsilon} - \log\left(\frac{M^2}{\mu^2}\right)\right). \quad (42)$$

Note that this term's dimensionality matches (15). Furthermore, it can be proved that the divergent terms in (32) proportional to $M^2 p^2$ are precisely those corresponding to (24) in momentum space, thus allowing us to recover the first orders of $\mathcal{L}_R$ for the general diagonal perturbation, which on shell correspond to $\sqrt{g}R$.

Diagram (36) has one divergent term proportional to $F^{\mu\nu}{}_{bcef}$ which is of higher order, i.e., $\mathcal{O}(\sqrt{g}R^2)$. Before addressing this apparent new divergence let us particularize to the case of a conformally flat perturbation above the vacuum. Taking $\sigma_i = \sigma$, (28)–(31) plus (32) add up in the effective action to

$$\frac{M^4 e^{-2\sigma}}{8\pi^2}\left(\frac{2}{\epsilon} - \log\left(\frac{M^2 e^{-\sigma}}{4\pi\mu^2}\right) - \gamma + \frac{3}{2}\right)$$
$$= \frac{M^4 e^{-(2\sigma + (\sigma/2)\epsilon)}}{8\pi^2}\left(\frac{2}{\epsilon} - \log\left(\frac{M^2}{4\pi\mu^2}\right) - \gamma + \frac{3}{2}\right). \quad (43)$$

An important remark is in order at this point. Note the peculiar form of $e^{-(2\sigma + (\sigma/2)\epsilon)}$: this factor corresponds to the determinant of a conformally flat metric in $D = 4 - \epsilon$ dimensions and it is a remnant of the fact that we used dimensional regularization to calculate the momentum integrals. Of course $\lim_{\epsilon\to 0}\sqrt{g_D} = \sqrt{g}$ (where g_D is the determinant of the D-dimensional metric), but this is telling us that in order to regularize our integrals it is not enough to add a mass scale to match the dimensionality; an ϵ power of the determinant of the metric is also needed to ensure diffeomorphism invariance. That is $\mu^2 \to \mu^2 e^{-\sigma}$. Then (43) would read

$$\frac{M^4 e^{-2\sigma}}{8\pi^2}\left(\frac{2}{\epsilon} - \log\left(\frac{M^2}{4\pi\mu^2}\right) - \gamma + \frac{3}{2}\right). \quad (44)$$

Continuing with the divergences, diagrams (38) and (39) contain terms of order M^4 that are the subsequent orders of the expansion of (43) in terms of σ. As for the terms of order p^2, one has to express them in position space. The result of diagram (32) for instance is

$$-\frac{\sigma p^2 \sigma M^2}{16\pi^2}\left(\frac{2}{\epsilon} - \log\left(\frac{M^2}{4\pi}\right) - \gamma - \frac{2}{3}\right) \quad (45)$$

that in position space reads

$$\frac{\sigma\square\sigma M^2}{16\pi^2}\left(\frac{2}{\epsilon} - \log\left(\frac{M^2}{4\pi}\right) - \gamma - \frac{2}{3}\right). \quad (46)$$

The next diagrams we consider are (38) plus (39)

$$\frac{M^2\sigma^2(p^2 + (p+q)^2 + q^2)\sigma}{32\pi^2}\left(\frac{2}{\epsilon} - \log\left(\frac{M^2}{4\pi}\right) - \gamma - \frac{5}{3}\right), \quad (47)$$

or in position space

$$-\frac{3M^2\sigma^2\square\sigma}{32\pi^2}\left(\frac{2}{\epsilon} - \log\left(\frac{M^2}{4\pi}\right) - \gamma - \frac{5}{3}\right). \quad (48)$$

Now it is clear that the full calculation at order $M^2 p^2$ resums to the following term in the effective action

$$\frac{M^2\square\sigma e^{-\sigma}}{32\pi^2}\left(\frac{2}{\epsilon} - \log\left(\frac{M^2}{4\pi\mu^2}\right) - \gamma + \frac{1}{3}\right). \quad (49)$$

Note that this term has the same structure that $\sqrt{g}R$ for a conformally flat metric. This divergence can be absorbed by redefining $\mathcal{L}_R$ and using the equations of motion. This is already telling us that the theory is renormalizable only on shell,[6] namely when the spin connection w_μ^{ab} corresponds to the Levi-Civita one. In our approach this identification is forced by the use of the equations of motion.

For a general diagonal perturbation one has to consider the momentum-dependent $O(p^2)$ divergent pieces in (32) and similar diagrams with more external scalar legs. As a check we can see that the momentum-dependent terms with two σ_i fields faithfully reproduce the $O(\sigma^2)$ piece in the curvature term (24) thus confirming the general covariance of the effective action. Details are relegated to the Appendix.

Let us now retake the issue of the apparent new divergences emerging from (36). To see if they really contribute to the final effective action we have to express them in terms of the σ fields using the available equations of motion. Then in principle, they must either vanish or correspond to a valid counterterm. We argued in Sec. II that there is a third possible counterterm in $D = 4$, the Gauss-Bonnet term, which is a total derivative and should not contribute to the dynamics. In the Appendix it is shown how the lower-order divergence vanishes and how the higher-order term indeed corresponds to a piece of the Gauss-Bonnet term.

---

[5]Note that factors $1/n!$, where n is the number of identical external legs, and a sign flip, are needed to reconstruct the term in the effective action from the diagrammatic calculation.

[6]Please note that this is quite unrelated to the well known fact that pure gravity at one loop is finite on shell. The latter result corresponds to performing a one-loop calculation with gravitons. Here instead we integrate the microscopic degrees of freedom that supposedly generate the gravitons after spontaneous symmetry breaking and generation of the metric degrees of freedom.

VI. EFFECTIVE ACTION AND PHYSICAL CONSTANTS

We are now ready to write the effective action we obtain on shell; that is once the spin connection is set to the value obtained after use of the equations of motion and the gap equation is used. We shall present details only for a vierbein corresponding to a conformally flat metric but as previously discussed we have a good check of its validity for the divergent parts of a general diagonal perturbation above the vacuum.

We recall our conventions. We have used Euclidean conventions so that the (emerging) metric has signature $(+, +, +, +)$. The effective action at long distances is defined by the functional integral

$$\int [dg] \exp(-S[g]), \tag{50}$$

where

$$g_{\mu\nu} = \eta_{ab} e^a_\mu e^b_\nu = \frac{1}{M^2} \eta_{ab} B^a_\mu B^b_\nu, \tag{51}$$

according to our discussion in Sec. II.

The effective action obtained after the diagrammatic calculation of the previous sections is

$$S_{\text{eff}} = \int d^4x \Big(c'M^4 e^{-2\sigma} - N \frac{M^4}{8\pi^2} e^{-2\sigma} \Big(\log\Big(\frac{M^2}{\mu^2}\Big) - \frac{3}{2} \Big)$$
$$+ A'M^2 \Box \sigma e^{-\sigma}$$
$$- N \frac{M^2}{32\pi^2} \Box \sigma e^{-\sigma} \Big(\log\Big(\frac{M^2}{\mu^2}\Big) - \frac{28}{3} \Big) \Big) + \dots, \tag{52}$$

where $c' = c + \frac{N}{8\pi^2}(\frac{2}{\epsilon} + \log 4\pi - \gamma)$ and $A' = A + \frac{N}{8\pi^2} \times (\frac{2}{\epsilon} + \log 4\pi - \gamma)$ are renormalized coupling constants that have absorbed the divergences. The $\overline{\text{MS}}$ subtraction scheme is assumed. Note that the finite part of the term proportional to M^2 has received a contribution from the diagrams containing only w^{ab}_μ fields, see the Appendix. Making use of the gap Eq. (6) we can write the previous expression as

$$S_{\text{eff}} = \int d^4x \Big(N \frac{M^4}{16\pi^2} e^{-2\sigma} + A' \Box \sigma e^{-\sigma} M^2$$
$$- N \frac{M^2}{32\pi^2} \Box \sigma e^{-\sigma} \Big(\log\Big(\frac{M^2}{\mu^2}\Big) - \frac{28}{3} \Big) \Big) + \cdots. \tag{53}$$

The resulting effective theory thus describes a geometry with a cosmological term. Sometimes it is stated in the literature, see the first reference of [4], that if gravity is an emergent phenomenon and gravitons are Goldstone bosons all interactions should be of a derivative nature and the cosmological constant problem would be in a sense solved. This is not so, as we see a cosmological terms is generated necessarily (both in $D = 2$ and $D = 4$), at least in the present approach.

The previous result is not exact of course. The effective action is in fact an infinite series containing higher-order derivatives, starting with terms of $O(\sqrt{g}R^2)$ and so on, which are represented by the dots in the previous expression. In fact, as we have seen, a counterterm proportional to Gauss-Bonnet [of order $\mathcal{O}(\sqrt{g}R^2)$] is required; finite terms will appear too. The effective action should also contain a nonlocal finite piece corresponding to the conformal anomaly (of dimension four in $D = 4$ [14]). The conformal anomaly was indeed reproduced in the previous work in $D = 2$ [5]. Note that any dimension four term that is generated will be accompanied by a factor of N. The dimension six terms will be of $O(N/M^2)$ and so forth. It would be natural to redefine the constant A' to include this factor of N in order to keep the counting of powers of N homogeneous.

Appealing to covariance arguments we can now express (52) in terms of invariants

$$S_{\text{eff}} = \int d^4x \Big[\frac{N}{16\pi^2} M^4 \sqrt{g} + \Big(A' - \frac{N}{48\pi^2} \Big(\log\Big(\frac{M^2}{\mu^2}\Big) - \frac{28}{3} \Big) \Big)$$
$$\times M^2 \sqrt{g} R + \dots \Big]. \tag{54}$$

Next we recall that the classical Einstein action corresponding to the Euclidean conventions is [15]

$$S = -\frac{M_P^2}{32\pi} \int d^4x \sqrt{g}(R - 2\Lambda). \tag{55}$$

Now identifying

$$\frac{N}{16\pi^2} M^4 = 2\Lambda \frac{M_P^2}{32\pi}$$
$$M^2 \Big(A' - \frac{N}{48\pi^2} \Big(\log\Big(\frac{M^2}{\mu^2}\Big) - \frac{28}{3} \Big) \Big) = -\frac{M_P^2}{32\pi}, \tag{56}$$

we indeed obtain

$$S_{\text{eff}} = -\frac{M_p^2}{32\pi} \int d^4x \sqrt{g}(R - 2\Lambda) + \mathcal{O}(p^4). \tag{57}$$

As we see from the previous discussion, the integration of the fermions (assumed to be the fundamental degrees of freedom in the theory) yields a positive cosmological constant. As for the value of M_P^2, the Planck mass squared, the sign is not really automatically defined. More on this latter.

A. Fine-tuning and running of the constants

To ensure that the action is renormalization-group-invariant, thus observable, the following beta function for each free constant in the theory must be obeyed

$$\mu \frac{dc'}{d\mu} = -\frac{N}{4\pi^2} \qquad \mu \frac{dA'}{d\mu} = -\frac{N}{24\pi^2}. \tag{58}$$

This running has nothing to do with the one generated by graviton exchange and it is thus unrelated to the presence or absence of asymptotic safety that some authors advocate for gravity. At scales $\mu \gg M$ the relevant degrees of freedom are not gravitons, but the $2N$ fermions appearing in the microscopic Lagrangian. On the other hand, at the moment that fermions become the relevant degrees of freedom, geometry loses its meaning. There is then no "shorter" distance than M^{-1}, or at the very least this regime cannot be probed. Note that to realize our physical assumption of having the fermions as fundamental degrees of freedom we should have $c > 0$ as discussed in Sec. II.

These equations do not reflect the complete running of the dimensionless couplings associated with $\mathcal{L}_D$ and $\mathcal{L}_R$, i.e., the constants associated with the cosmological and Einstein-Hilbert terms, but only the one obtained at leading order in N. That is, the "graviton" loops are not included here; they are suppressed by one power of N if N is large. To see this last statement we recall that the usual power counting rules show that the exchange of the vierbein degrees of freedom would be accompanied by a factor of M_P^{-2}, suppressed by $1/N$. Leaving these corrections aside, we note that the two free couplings of the theory have a running that is opposed in sign to the one found in $D = 2$.

It is probably useful to appeal to the QCD analogy. At long distances strong interactions are well-described by the pion chiral Lagrangian, parametrized by f_π or the $O(p^4)$ coefficients, generically named low-energy constants (LEC). The LEC are a complicated function of α_s, the coupling constant of QCD. The microscopic theory proposed in this paper is the analogous of QCD, while the resulting effective theory (57) is the counterpart of the chiral effective Lagrangian. Then M_P and Λ are the LEC of the present theory. The running of α_s does not have an immediate translation on the LEC while in the present model, because of its simplicity, the consequences of the running in the microscopic particle reflects directly in Λ and M_P. But in addition these constants have an additional running (analogous to doing pion loops in the chiral Lagrangian). The counting of powers of N disentangles both types of running.

At some scale, $q \sim M$ the effective theory stops making sense. At that moment the relevant degrees of freedom change and, as a result, the metric disappears. Exactly in the same way as for large momentum transfers we do not see pions but quarks. Of course, if there is no metric there is no geometry and, in particular, the notion of distance disappears altogether at length scales below M^{-1}. From this point of view, gravity is non-Wilsonian.

Let us now try to make contact with the value that the LEC take in gravity. Clearly, there is enough freedom in the theory (by adjusting A' and M) to reproduce any values of Λ and M_P. But we also want higher-order terms to be small for the effective theory to make sense in a reasonable range of momenta. We may even get rid of all of the high

order [$\mathcal{O}(p^4)$ and beyond] if we take $M \to \infty$ and at the same time we take $N \to 0$ in a prescribed form. Then, in the actual limit, which corresponds to a "quenched" approximation, we exactly reproduce Einstein-Hilbert Lagrangian, with a cosmological constant, and nothing else. Of course in this limit, the presumed fundamental degrees of freedom disappear completely and we have all the way up to $\mu = \infty$ Einstein's theory —with all its ultraviolet problems; there are no fundamental degrees of freedom providing form factors to cut off the offending divergent integrals.[7]

Of course the $N \to 0$ limit is just the opposite one to the one we have used. All our diagrammatic results are exact in the $N \to \infty$ limit and presumably get large corrections as N approaches zero, but the general features of the model should survive.

Note that M is a fixed quantity in the model and if M_P^2 increases, Λ decreases. Taking the actual observed or estimated values of these two parameters we get the value $NM^4 \sim 10^{18}$ m^{-4}, which is a very low scale. One may think that this may already represent unacceptably large corrections from higher-order operators. However, this is not necessarily so because the bounds on R^2 terms are very weak. For instance, the bound $k < 10^{74}$ has been quoted for a generic coefficient [16] k of the $\mathcal{O}(p^4)$ terms. Thus, a relatively low scale for M cannot be really excluded observationally by studying gravitational effects alone and one should be aware of this. However, our own intuition tells us that M should be much larger than the value quoted above as the notion of metric certainly makes sense at much shorter distances. We can increase the value of M as much as we want by decreasing the value of N, as previously indicated. We shall not elaborate further on this as it seems too premature an speculation.

Finally we note that the sign of Newton's constant is not determined *a priori* in this theory due to the subtraction required from the counterterm in $\mathcal{L}_\mathcal{R}$. This ties nicely with some of the early discussions on induced gravity [17].

VII. SUMMARY

Let us summarize our findings and comment on possible implications of our work. We have proposed a model where $D = 4$ gravity emerges from a theory without any predefined metric. The minimal input is provided by assuming a differential manifold structure endowed with an affine connection. Nothing more. The Lagrangian can be defined without having to appeal to a particular metric or vierbein.

---

[7]Note that resolving the vertices' sigularities is not enough to mitigate the divergences of gravity as a loop of, e.g., Dirac fermions generates itself new divergences of $\mathcal{O}(p^4)$. It is the combination of this with the absence of a metric tensor in the unbroken phase that might help, as in the mechanism proposed here.

Gravity and distance are induced rather than fundamental concepts in this proposal. At sufficiently short scales, when the effective action does not make sense anymore, the physical degrees of freedom are fermionic. At such short scales there is not even the notion of a smaller scale.

The relative technical simplicity of this proposal constitutes its main virtue when compared with previous proposals [6,7], where even semiquantitative discussions appear impossible. Here one is able to derive in full detail the effective action.

A very important aspect of the model is the apparent improvement of the ultraviolet behavior. After integration of the fundamental degrees of freedom all the divergences that appear to the order we have computed in external fields can seemingly be absorbed in the redefinition of the cosmological constant and the Planck mass (as seen from the effective theory point of view, even though the respective counterterms do not have this meaning in the underlying theory). With the running, dictated by the corresponding beta functions, both quantities are renormalization-group-invariant. In addition the Gauss-Bonnet invariant is renormalized too. This mitigation of the short-distance divergences happens in spite of the bad ultraviolet behavior of the propagator and the ultimate reason, we think, is that these are the *only* counterterms that can be written without having to assume an underlying metric that does not exist before spontaneous symmetry breaking takes place.

At long distances the fluctuations around the broken vacuum are the relevant degrees of freedom and are described by an effective theory whose lowest-dimensional operators are just those of ordinary $D = 4$ gravity. They of course exhibit the usual divergences of quantum gravity but this now poses in principle no problem as we know that at very high energies this is not the right theory. For $q \sim M$ one starts seeing the fundamental degrees of freedom. Gravitons are the Goldstone bosons of a broken global symmetry. We have seen how the barrier of the Weinberg-Witten no-go theorem could be overcome.

In a sense the fundamental fermions resolve the point-like 3-graviton, 4-graviton, etc., interactions into extended form factors and this is the reason for the mitigation of the terrible ultraviolet behavior of quantum gravity. However this is only part of the story, because this could be equally achieved by using Dirac fermions coupled to gravity (or any other field for that matter). This would in fact be just a reproduction of the old program of induced gravity [17] and therefore not that interesting. The really novel point in this proposal is that the microscopic fermion action does not contain any metric tensor at all. Then not only is the metric and its fluctuations —the gravitons— spontaneously generated, but the possible counterterms are severely limited in number.

We stop short of making any strong claims about the renormalizability of the model. We can just say that, from our calculations and our experience with the model, renormalizability is a plausible hypothesis (our results actually amount to an heuristic proof in the large N limit). Likewise we do not insist in that the one presented is the *sole* possibility to carry out the present program, although it looks fairly unique. Clearly a number of issues need further study before the present proposal can be taken seriously but we think that the results presented here are of sufficient interest to make them known to researchers in the field.

A number of extensions and applications come to our mind. Perhaps the most intriguing one from a physical point of view would be to investigate in this framework singular solutions in GR such as black holes. A more in-depth study of the renormalizability issue is certainly required too as there are issues related to the renormalization group to be addressed in the present setting.

Note added in proof. This work is a continuation of [5]. We list here a few errors found in [5] in order to facilitate the comparison. The equations cited in what follows refer to [5]. The first term on the left-hand side of Eq. (34) should be divided by $n!$. The second term on the left-hand side of Eq. (35) has to be multiplied by 2. In the result of Eq. (56), the term proportional to M^2 has to be multiplied by 2. Equation (60) has to be multiplied by a global factor of 2, since we refer to the total contribution to the effective action. In Eq. (63), the three terms not proportional to M^2 need to be divided by 2. Finally there is a typo in the second paragraph after Eq. (63); the expressions are valid as long as the characteristic momenta are $k < M$.

APPENDIX

In this Appendix we include the explicit calculation of the different terms appearing in (36) showing how they correspond on shell to different terms in the action. We also include, for completeness, how the result of (32) in the diagonal parametrization of the metric used in the text yields precisely (24).

1. $D^{\mu\nu}{}_{bcef}$, $E^{\mu\nu}{}_{bcef}$ and $F^{\mu\nu}{}_{bcef}$

We saw in Sec. (4.2) that diagram (36) contains three different terms, two of them being divergent, let us show how they either cancel or can be accommodated in the available counterterms. Let us write them down together with the w^{ab}_μ fields.

$$D^{\mu\nu}{}_{bcef} w_\mu^{bc} w_\nu^{ef} = 4w_\mu^{\nu b} w_\nu^{\mu b} - 2w_\mu^{be} w_\mu^{be}$$

$$= 2(w_1^{12})^2 - 2(w_1^{21})^2 + 2(w_1^{13})^2 - 2(w_1^{31})^2$$

$$+ 2(w_1^{14})^2 - 2(w_1^{41})^2 + \cdots$$

$$+ 2(w_4^{34})^2 - 2(w_4^{43})^2 = 0. \tag{A1}$$

So this divergence cancels regardless of the parametrization we choose.

Let us write the second one now considering a conformally flat parametrization of the metric

$$g_{\mu\nu} = e^{-\sigma(x)} \delta_{\mu\nu}. \tag{A2}$$

Then we have

$$E^{\mu\nu}{}_{bcef} w_\mu^{bc} w_\nu^{ef} = 4w_\mu^{\mu b} w_\nu^{b\nu}. \tag{A3}$$

Making use of Eq. (18), that is,

$$w_\mu{}^{ab} = \tfrac{1}{2}(\partial^a \sigma \delta_\mu^b - \partial^b \sigma \delta_\mu^a), \tag{A4}$$

we obtain

$$E^{\mu\nu}{}_{bcef} w_\mu^{bc} w_\nu^{ef} = -9[(\partial_1\sigma)^2 + (\partial_2\sigma)^2 + (\partial_3\sigma)^2 + (\partial_4\sigma)^2]$$

$$= -9\partial_\mu\sigma\partial_\mu\sigma. \tag{A5}$$

Recall this term appeared both in (36) and in (41). Summing their contributions in the way we did to reconstruct the effective action for the σ field diagrams we obtain

$$\frac{9M^2}{32\pi^2}\partial_\mu\sigma\partial_\mu\sigma(1-\sigma+\cdots) \rightarrow \frac{9M^2}{32\pi^2}\partial_\mu\sigma\partial_\mu\sigma e^{-\sigma} \rightarrow \frac{9M^2}{32\pi^2}e^{-\sigma}\Box\sigma. \tag{A6}$$

This is just a finite contribution to $\sqrt{g}R$ and it is reflected in (52).

The Gauss-Bonnet term corresponding to such a metric perturbation reads

$$\mathcal{L}_{GB} = \sqrt{g}(R^2 + 4R_{\mu\nu}R^{\mu\nu} - R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma})$$

$$= -4\partial_3\partial_4\sigma\partial_3\partial_4\sigma + 4\partial_4^2\sigma\partial_3^2\sigma - 4\partial_2\partial_4\sigma\partial_2\partial_4\sigma + 4\partial_4^2\sigma\partial_2^2\sigma - 4\partial_2\partial_3\sigma\partial_2\partial_3\sigma + 4\partial_3^2\sigma\partial_2^2\sigma - 4\partial_1\partial_4\sigma\partial_1\partial_4\sigma$$

$$+ 4\partial_4^2\sigma\partial_1^2\sigma - 4\partial_1\partial_2\sigma\partial_1\partial_2\sigma + 4\partial_2^2\sigma\partial_1^2\sigma - 4\partial_1\partial_3\sigma\partial_1\partial_3\sigma + 4\partial_3^2\sigma\partial_1^2\sigma - 3\partial_4\sigma\partial_4\sigma\partial_4^2\sigma - \partial_4^2\sigma\partial_3\sigma\partial_3\sigma$$

$$- 4\partial_4\sigma\partial_3\sigma\partial_3\partial_4\sigma - \partial_4\sigma\partial_4\sigma\partial_3^2\sigma - 3\partial_3\sigma\partial_3\sigma\partial_3^2\sigma - \partial_4^2\sigma\partial_2\sigma\partial_2\sigma - \partial_3^2\sigma\partial_2\sigma\partial_2\sigma - 4\partial_4\sigma\partial_2\sigma\partial_2\partial_4\sigma$$

$$- 4\partial_3\sigma\partial_2\sigma\partial_2\partial_3\sigma - \partial_4\sigma\partial_4\sigma\partial_2^2\sigma - \partial_3\sigma\partial_3\sigma\partial_2^2\sigma - 3\partial_2\sigma\partial_2\sigma\partial_2^2\sigma - \partial_4^2\sigma\partial_1\sigma\partial_1\sigma - \partial_3^2\sigma\partial_1\sigma\partial_1\sigma$$

$$- \partial_2^2\sigma\partial_1\sigma\partial_1\sigma - 4\partial_4\sigma\partial_1\sigma\partial_1\partial_4\sigma - 4\partial_3\sigma\partial_1\sigma\partial_1\partial_3\sigma - 4\partial_2\sigma\partial_1\sigma\partial_1\partial_2\sigma - \partial_4\sigma\partial_4\sigma\partial_1^2\sigma - \partial_3\sigma\partial_3\sigma\partial_1^2\sigma$$

$$- \partial_2\sigma\partial_2\sigma\partial_1^2\sigma - 3\partial_1\sigma\partial_1\sigma\partial_1^2\sigma. \tag{A7}$$

The last term we have to explore is the piece $\frac{1}{\epsilon}F^{\mu\nu}{}_{bcef}$. Let us explicitly write this term together with the w_μ^{bc} fields

$$\frac{1}{\epsilon}F^{\mu\nu}{}_{bcef}w_\mu^{bc}w_\nu^{ef} = \frac{1}{\epsilon}\left[-\frac{\delta^{\mu\nu}\delta_{be}p_c p_f}{12\pi^2} + \frac{\delta^{\mu\nu}\delta_{bf}p_e p_c}{12\pi^2} + \frac{\delta^{\mu\nu}\delta_{ce}p_b p_f}{12\pi^2} - \frac{\delta^{\mu\nu}\delta_{cf}p_b p_e}{12\pi^2} - \frac{\delta_b^\mu \delta_e^\nu \delta_{cf}p^2}{12\pi^2} + \frac{\delta_b^\mu \delta_e^\nu p_c p_f}{12\pi^2} \right.$$

$$+ \frac{\delta_b^\mu \delta_f^\nu \delta_{ce}p^2}{12\pi^2} - \frac{\delta_b^\mu \delta_f^\nu p_e p_c}{12\pi^2} - \frac{\delta_e^\mu \delta_{bf}p^\nu p_c}{12\pi^2} + \frac{\delta_e^\mu \delta_b^\nu p_c p_f}{12\pi^2} - \frac{\delta_e^\mu \delta_c^\nu p_b p_f}{12\pi^2} + \frac{\delta_e^\mu \delta_{cf}p^\nu p_b}{12\pi^2} + \frac{\delta_c^\mu \delta_e^\nu \delta_{bf}p^2}{12\pi^2}$$

$$- \frac{\delta_c^\mu \delta_e^\nu p_b p_f}{12\pi^2} - \frac{\delta_c^\mu \delta_f^\nu \delta_{be}p^2}{12\pi^2} + \frac{\delta_c^\mu \delta_f^\nu p_b p_e}{12\pi^2} + \frac{\delta_f^\mu \delta_{be}p^\nu p_c}{12\pi^2} - \frac{\delta_f^\mu \delta_b^\nu p_e p_c}{12\pi^2} + \frac{\delta_f^\mu \delta_c^\nu p_b p_e}{12\pi^2} - \frac{\delta_f^\mu \delta_{ce}p^\nu p_b}{12\pi^2}$$

$$\left. + \frac{\delta_{be}\delta_c^\nu p^\mu p_f}{12\pi^2} - \frac{\delta_{be}\delta_{cf}p^\mu p^\nu}{12\pi^2} - \frac{\delta_{bf}\delta_c^\nu p^\mu p_e}{12\pi^2} + \frac{\delta_{bf}\delta_{ce}p^\mu p^\nu}{12\pi^2} - \frac{\delta_b^\nu \delta_{ce}p^\mu p_f}{12\pi^2} + \frac{\delta_b^\nu \delta_{cf}p^\mu p_e}{12\pi^2} \right]w_\mu^{bc}w_\nu^{ef}$$

$$= \frac{1}{3\pi^2}\frac{1}{\epsilon}\left[\partial_c w_\mu^{bc}\partial_e w_\mu^{be} + w_\mu^{\mu c}\Box w_\nu^{\nu c} - \partial_c w_\mu^{\mu c}\partial_e w_\nu^{\nu e} - \partial_c w_\mu^{bc}\partial_e w_b^{\mu e} + \partial_c w_\mu^{bc}\partial^\nu w_\nu^{\mu b} \right.$$

$$\left. + \partial^\mu w_\mu^{b\nu}\partial_e w_\nu^{eb} + \frac{1}{2}\partial^\mu w_\mu^{bc}\partial^\nu w_\nu^{bc} \right] \tag{A8}$$

$$\frac{1}{\epsilon}F^{\mu\nu}{}_{bcef}w^{bc}_\mu w^{ef}_\nu = \frac{6}{\pi^2}\frac{1}{\epsilon}[-\partial^2_4\sigma\partial^2_1\sigma + \partial_1\partial_4\sigma\partial_1\partial_4\sigma - \partial^2_3\sigma\partial^2_4\sigma + \partial_3\partial_4\sigma\partial_3\partial_4\sigma - \partial^2_2\sigma\partial^2_4\sigma + \partial_2\partial_4\sigma\partial_2\partial_4\sigma$$
$$- \partial^2_2\sigma\partial^2_3\sigma + \partial_2\partial_3\sigma\partial_2\partial_3\sigma - \partial^2_3\sigma\partial^2_1\sigma + \partial_1\partial_3\sigma\partial_1\partial_3\sigma - \partial^2_2\sigma\partial^2_1\sigma + \partial_1\partial_2\sigma\partial_1\partial_2\sigma]. \tag{A9}$$

Now it is easy to see that (A9) corresponds to the second and third lines in (A7). This divergent contribution is part of the Gauss-Bonnet term, which, although being a total derivative, is a valid counterterm. The rest of (A7) contains three sigma fields and is not present in (36) as should be the case. These remaining terms would be generated in the triangular diagram with three external w^{ab}_μ fields and would come with a divergent coefficient (we have not computed such diagram). Note also that both (A9) and the first two lines of (A7) can be integrated by parts to make them vanish. This happens because in the conformally flat metric perturbations terms from the diagrams with two and three external w fields can not be related to each other integrating by parts. Therefore, they must vanish independently as the Gauss-Bonnet term is a total derivative after all.

This is a particularity of the conformally flat parametrization of the metric perturbation and would not hold for a generally diagonal parametrization. In that case terms

generated in the two-point function can be transformed into the terms appearing in the three-point function by integration by parts and one would require a full calculation to show there is a match with an independent calculation of the Gauss-Bonnet term.

2. $\sqrt{g}R$ for the general diagonal parametrization of the metric

Let us consider a generally diagonal metric with 4 degrees of freedom

$$g_{\mu\nu} = \begin{pmatrix} e^{-\sigma_1(x)} & 0 & 0 & 0 \\ 0 & e^{-\sigma_2(x)} & 0 & 0 \\ 0 & 0 & e^{-\sigma_3(x)} & 0 \\ 0 & 0 & 0 & e^{-\sigma_4(x)} \end{pmatrix}. \tag{A10}$$

We saw that the corresponding expression for the curvature is Eq. (24)

$$\mathcal{L}_R|_{\text{(on shell)}} = M^2\sqrt{g}R$$
$$= M^2\Big[\partial^2_3\sigma_4 + \partial^2_2\sigma_4 + \partial^2_1\sigma_4 + \partial^2_4\sigma_3 + \partial^2_2\sigma_3 + \partial^2_1\sigma_3 + \partial^2_4\sigma_2 + \partial^2_3\sigma_2 + \partial^2_1\sigma_2 + \partial^2_4\sigma_1 + \partial^2_3\sigma_1$$
$$+ \partial^2_2\sigma_1 - \tfrac{1}{2}(\partial_3\sigma_1\partial_3\sigma_2 + \partial_4\sigma_1\partial_4\sigma_2 + \partial_2\sigma_1\partial_2\sigma_3 + \partial_4\sigma_1\partial_4\sigma_3 + \partial_2\sigma_1\partial_2\sigma_4 + \partial_3\sigma_1\partial_3\sigma_4$$
$$+ \partial_1\sigma_2\partial_1\sigma_3 + \partial_4\sigma_2\partial_4\sigma_3 + \partial_1\sigma_2\partial_1\sigma_4 + \partial_3\sigma_2\partial_3\sigma_4 + \partial_1\sigma_3\partial_1\sigma_4 + \partial_2\sigma_3\partial_2\sigma_4) + \mathcal{O}(\sigma^3)\Big]. \tag{A11}$$

We consider now the divergent part proportional to M^2 of the result of (32)

$$\frac{2}{\epsilon}\sum_{j=1}^4\sum_{\substack{l=1 \\ l\neq j}}^4\Big[-\frac{\sigma_j\sigma_l p^2 M^2}{48\pi^2} + \frac{\sigma_j\sigma_l(p_j^2 + p_l^2)M^2}{48\pi^2}\Big]$$
$$= \frac{2}{\epsilon}\frac{M^2}{48\pi^2}[-p^2(\sigma_1\sigma_2 + \sigma_1\sigma_3 + \sigma_1\sigma_4 + \sigma_2\sigma_3 + \sigma_2\sigma_4 + \sigma_3\sigma_4)$$
$$+ \sigma_1\sigma_2(p_1^2 + p_2^2) + \sigma_1\sigma_3(p_1^2 + p_3^2) + \sigma_1\sigma_4(p_1^2 + p_4^2) + \sigma_2\sigma_3(p_2^2 + p_3^2)$$
$$+ \sigma_2\sigma_4(p_2^2 + p_4^2) + \sigma_3\sigma_4(p_3^2 + p_4^2)]$$
$$= \frac{2}{\epsilon}\frac{M^2}{48\pi^2}[-\sigma_1\sigma_2(p_3^2 + p_4^2) - \sigma_1\sigma_3(p_2^2 + p_4^2) - \sigma_1\sigma_4(p_2^2 + p_3^2)$$
$$- \sigma_2\sigma_3(p_1^2 + p_4^2) - \sigma_2\sigma_4(p_1^2 + p_3^2) - \sigma_3\sigma_4(p_1^2 + p_2^2)]. \tag{A12}$$

This last expression, when expressed in position space, corresponds exactly to (A11) except for a numerical factor and minus the second derivatives of the fields which are total derivatives and do not appear in the perturbative calculation.

[1] C. J. Isham, A. Salam, and J. A. Strathdee, Ann. Phys. (N.Y.) **62**, 98 (1971).

[2] A. B. Borisov and V. I. Ogievetsky, Theor. Math. Phys. **21**, 1179 (1974); E. A. Ivanov and V. I. Ogievetsky, Lett. Math. Phys. **1**, 309 (1976).

[3] H. C. Ohanian, Phys. Rev. **184**, 1305 (1969).

[4] P. Kraus and E. T. Tomboulis, Phys. Rev. D **66**, 045015 (2002); V. A. Kostelecky and R. Potting, Phys. Rev. D **79**, 065018 (2009); G. E. Volovik, Physica B (Amsterdam) **162**, 222 (1990).

[5] J. Alfaro, D. Espriu, and D. Puigdomènech, Phys. Rev. D **82**, 045018 (2010).

[6] D. Amati and J. Russo, Phys. Lett. B **248**, 44 (1990); J. Russo, Phys. Lett. B **254**, 61 (1991).

[7] A. Hebecker and C. Wetterich, Phys. Lett. B **574**, 269 (2003); C. Wetterich, Phys. Rev. D **70**, 105004 (2004).

[8] In addition to those listed in [4]: S. C. Zhang and J. Hu, Science **294**, 823 (2001); C. Barcelo and M. Visser, Int. J. Mod. Phys. D **10**, 799 (2001); H. Grosse, H. Steinacker, and M. Wohlgenannt, J. High Energy Phys. 04 (2008) 023; B. McElrath, arXiv:0812.2696; F. Girelli, S. Liberati, and L. Sindoni, Phys. Rev. D **79**, 044019 (2009); T. Padmanabhan, Int. J. Mod. Phys. D **17**, 591 (2008); P. Horava, J. High Energy Phys. 03 (**2009**) 020; M. M. Anber, U. Aydemir, and J. F. Donoghue, Phys. Rev. D **81**, 084059 (2010); C. Xu and P. Horava, Phys. Rev. D **81**, 104033 (2010); E. W. Mielke, Phys. Rev. D **83**, 044004 (2011); L. Sindoni, SIGMA **8**, 027 (2012).

[9] S. Weinberg and E. Witten, Phys. Lett. B **96**, 59 (1980).

[10] A. Jenkins, Int. J. Mod. Phys. D **18**, 2249 (2009).

[11] S. Weinberg, in *General Relativity: An Einstein Centenary Survey*, edited by S. W. Hawking and W. Israel (Cambridge University Press, Cambridge, England, 1979), p. 790.

[12] See, e.g., I. Heemskerk and J. Polchinski, J. High Energy Phys. 06 (2011) 031; G. Dvali and C. Gomez, arXiv:1005.3497.

[13] A. Palatini, Rend. Circ. Mat. Palermo **43**, 203 (1919).

[14] See, e.g., M. Asorey, E. V. Gorbar, and I. L. Shapiro, Classical Quantum Gravity **21**, 163 (2004).

[15] S. M. Christensen and M. J. Duff, Nucl. Phys. **B170**, 480 (1980); Phys. Lett. B **76**, 571 (1978).

[16] K. S. Stelle, Gen. Relativ. Gravit. **9**, 353 (1978).

[17] Ya. B. Zeldovich, JETP Lett. **6**, 316 (1967); A. Sakharov, Sov. Phys. Dokl. **12**, 1040 (1968); O. Klein, Phys. Scr. **9**, 69 (1974); S. Adler, Rev. Mod. Phys. **54**, 729 (1982).