Journal of Statistical Mechanics:
Theory and Experiment

PAPER

Generalized scale behavior and renormalization group for data analysis

To cite this article: Vincent Lahoche *et al J. Stat. Mech.* (2022) 033101

Journal of Statistical Mechanics: Theory and Experiment
An IOP and SISSA journal

PAPER: Quantum statistical physics, condensed matter, integrable systems

Generalized scale behavior and renormalization group for data analysis

Vincent Lahoche[1], **Dine Ousmane Samary**[1,2,*] **and Mohamed Tamaazousti**[1]

[1] Commissariat à l'Énergie Atomique (CEA, LIST), 8 Avenue de la Vauve, 91120 Palaiseau, France

[2] International Chair in Mathematical Physics and Applications (ICMPA-UNESCO Chair), University of Abomey-Calavi, 072B.P.50, Cotonou, Benin

E-mail: vincent.lahoche@cea.fr, dine.ousmanesamary@cipma.uac.bj and mohamed.tamaazousti@cea.fr

Received 1 October 2021
Accepted for publication 4 February 2022
Published 3 March 2022

Online at stacks.iop.org/JSTAT/2022/033101
https://doi.org/10.1088/1742-5468/ac52a6

Abstract. Some recent results showed that the renormalization group (RG) can be considered as a promising framework to address open issues in data analysis. In this work, we focus on one of these aspects, closely related to principal component analysis (PCA) for the case of large dimensional data sets with covariance having a nearly continuous spectrum. In this case, the distinction between 'noise-like' and 'non-noise' modes becomes arbitrary and an open challenge for standard methods. Observing that both RG and PCA search for simplification for systems involving many degrees of freedom, we aim to use the RG argument to clarify the turning point between noise and information modes. The analogy between coarse-graining renormalization and PCA has been investigated in Bradde and Bialek (2017 *J. Stat. Phys.* **167** 462–75), from a perturbative framework, and the implementation with real sets of data by the same authors showed that the procedure may reflect more than a simple formal analogy. In particular, the separation of sampling noise modes may be controlled by a non-Gaussian fixed point, reminiscent of the behaviour of critical systems. In our analysis, we go beyond the perturbative framework using nonperturbative techniques to

*Author to whom any correspondence should be addressed.

 1742-5468/22/033101+30$33.00

investigate non-Gaussian fixed points and propose a deeper formalism allowing us to go beyond power-law assumptions for explicit computations.

Keywords: communication, supply and information networks, renormalisation group

Contents

1. Introduction

The physical description of systems involving a very large number of interacting degrees of freedom remains a difficult task since the discovery of microscopic structures at the beginning of the 20th century. Since the last decade, the big data revolution has provided a new example of such a large dimensional system: the number of degrees of freedom involved in some data sets can easily be of the same order of magnitude as large dimensional physical systems. Note that the difficulty does not come especially from the size of the system, but from the complex relations between degrees of freedoms. Indeed, there is no difficulty in describing a very large number of independent and identical systems, the complexity reduces to the description of a single one of these subsystems ([1–18] and references therein). However, such a reduction breakdown for a system has some complex relations between subsystems; it is, for instance, impossible to reduce the hard complex relations between human cells as the behavior of an isolated cell. More

J. Stat. Mech. (2022) 033101

than an impossibility, such a reduction should be a profound misconception. The same thing occurs for data sets where strong correlations may happen between a vast set of high dimensional data. Principal component analysis (PCA) is based on a systematic dimensional reduction over a finite (and not so big) dimensional space corresponding to the relevant vectors, in the full configuration space, for a given covariance matrix describing correlations between degrees of freedom [19–24]. To be more concrete, a suitably mean-shifted set of data generally takes the form of a big $p \times n$ matrix $X = \{X_{ia}\}$, for $i = 1, \ldots, p$ and $a = 1, \ldots, n$. Generally, p and n are both assumed to be large, in such a way that the ratio p/n remains fixed. The integer p corresponds to the size of the data, whereas n denotes the size of the set. Then, the covariance matrix $\mathcal{C}$ *between data* is an $n \times n$ matrix defined as the average of $X^{\mathrm{T}}X$. Orthodox PCA works well when a relatively small number of discrete eigenvalues are distinguished from the rest of them, allowing to project linearly onto a reduced dimensional subspace. However, it is not generally the case, as pointed out in [24–32], especially for a system with a very large number of degrees of freedom, for which the spectra tend to be continuously distributed. In such a case, the separation between relevant and irrelevant eigenvalues becomes arbitrary. It is this arbitrariness that motivates the renormalization group (RG) analysis.

The RG is a systematic procedure allowing one to describe how a physical system changes at different scales. Originally appearing in field theory as a consequence of the renormalization procedure introduced to solve the problem of 'ultraviolet' (UV) divergences, the RG has been essentially developed in the context of critical phenomena. Rapidly, the RG has become a general framework to describe physical systems having a large number of interacting degrees of freedom, and has been used as a powerful investigation tool in a large variety of physical contexts. In particular, in the context of critical phenomena, RG has been proved to be very appropriate to discuss the question of universality. For some critical systems, and for a sufficiently large scale, the flow becomes dragged toward a finite-dimensional subspace corresponding to the marginal and relevant operator, providing an efficient projection into a finite-dimensional subspace. Such a phenomenon, called the *large river effect*, generally involves a non-Gaussian fixed-point, toward which the mainstream goes from the Gaussian fixed point [33–46]. Such a picture provides a qualitative illustration of how RG work so well to describe the macroscopic properties of large systems; the dimensional reduction provides an efficient description involving a few sets of parameters, so far from the original complexity of the system. In the physical words, macroscopic physics become insensitive to the detailed microscopic structure of the interactions. This is the link that the authors in [19] have explored to meet PCA and RG, in the context where PCA introduces an arbitrariness for the choice of the cut-off between relevant and irrelevant parts of the covariance spectrum. More precisely, the authors argued that RG can be used to distinguish within a large number of degrees of freedom those which are sensitive to a change of scale and those that are insensitive, and they propose a field theoretical model to implement this idea.

The proposed framework was essentially focused on perturbation theory through a partial integration strategy of modes with small variance, essentially inspired from the Wilson–Kadanoff point of view on RG. In this way, the authors essentially focused on dimensional effects, their fixed point being very reminiscent of the well-known Wilson–Fisher (WF) fixed point for critical systems. As a consequence, they pointed out that the presence and the relevance of the fixed point depend on the shape of the eigenvalue distribution $\rho(\lambda)$ for the covariance matrix. The reason for this limitation to the perturbative region is that the authors were essentially interested in distributions not so far from the Gaussian case. However, the analysis that we provide here shows that such an approximation cannot be suitable for realistic data sets, and in particular for small deformations around Marchenko–Pastur (MP) law, for which the scaling dimensions for couplings are positives and larges, justifying the nonperturbative treatment that we propose in this paper. The manuscript is therefore written in a pedagogical style and serves as an introduction to a series of incoming works for an interdisciplinary community.

The outline is the following: in section 2, we give the useful ingredients and definitions allowing us to compute the functional renormalization group (FRG) equation. Particularly, we provide the method used to analyse the scaling behaviour previously given in [19]. In section 3, we study the nonperturbative behaviour of the model through the Wetterich equation. We also give the rigorous way that generating the canonical dimension allows us to derive the scale variation of the coupling and wave function. In section 4, the flow equation is solved in the local potential approximation (LPA). The numerical investigation and the flow diagrams are also given. Section 6 is devoted to the numerical data analysis that helps to understand the relevance of our method compared with [19]. In the same section, we provide our conclusion and remarks.

2. Preliminaries

In this section, we provide some technical aspects useful for the reader in the final steps of the paper. We remain voluntarily close to the vocabulary used in the referenced paper [19], to make contact more easily with the result of the authors.

2.1. Renormalization group for data analysis

Let us consider a physical system involving a large set Φ of N random variables $\Phi = \{\phi_1, \ldots, \phi_N\}$, described by a certain probability distribution:

$$p[\Phi] = \frac{1}{Z}\,\mathrm{e}^{-\mathcal{H}[\Phi]} \tag{1}$$

where, in accordance with physical nomenclature, we call the *Hamiltonian* the functional $\mathcal{H}[\Phi]$ and *partition function* the normalization Z:

$$Z = \int \mathrm{d}\Phi\,\mathrm{e}^{-\mathcal{H}[\Phi]}, \tag{2}$$

where $\mathrm{d}\Phi \equiv \mathrm{d}\phi_1\,\mathrm{d}\phi_2 \ldots \mathrm{d}\phi_N$. It is then useful to define the *generating functional* as follows:

$$Z[J] = \int \mathrm{d}\Phi\,\mathrm{e}^{-\mathcal{H}[\Phi]+\sum_i j_i\phi_i}, \tag{3}$$

J. Stat. Mech. (2022) 033101

which is such a way that the *connected correlation functions* $\langle \phi_i \phi_j \ldots \phi_k \rangle$ are obtained from $\mathcal{W}[J] := \ln Z[J]$ by a simple derivative with respect to the source $J = \{j_i\}$:

$$\langle \phi_i \phi_j \ldots \phi_k \rangle = \frac{\partial \partial \ldots \partial}{\partial j_1 \partial j_2 \ldots \partial j_k} \mathcal{W}[J] \bigg|_{J=0}. \tag{4}$$

Note that $Z[J]$ is usually interpreted as a sourced version of the partition function Z, J being interpreted for instance as an external field, bouncing the system towards a preferred configuration. As recalled in the introduction, the RG is a formalism that allows us to deal with the change of a physical system when the scale changes. In other words, RG supposes a definition of what is microscopic and what is macroscopic. Such a distinction is generally easy for a physical system, like a block of matter; the microscopic scale is identified as the atomic scale and the macroscopic scale to the scale of the magnet itself, involving a large number of microscopic degrees of freedom. For some categories of data, we can eventually define the corresponding notions. For the sets of images, for instance, the microscopic degrees of freedom can be identified with pixels, whereas the macroscopic structures become the ordinary objects 'planes, cats, cars' and so on. However, the situation is not always adequate, and for more abstract sets of data, it may be difficult to identify what is microscopic. On a more abstract level, it is tempting to associate the microscopic level with noisy degrees of freedom, i.e. with the region of small eigenvalues of the covariance matrix spectrum. The RG framework that we discuss in this paper moreover provides a canonical notion of what is microscopic and what is macroscopic.

Indeed from the RG point of view, at the microscopic level, which we usually name as the UV scale in physics, no fluctuations are taken into account, and the distribution $p[\Phi]$ is essentially dominated by *classical configurations* $\Phi_{\mathrm{UV}} \equiv \{\phi_{i,\mathrm{UV}}\}$, corresponding to the minima of the Hamiltonian functional:

$$\frac{\partial \mathcal{H}}{\partial \phi_{i,\mathrm{UV}}} = 0 \quad \forall\, i. \tag{5}$$

In contrast, in the macroscopic level, referred to as an infrared (IR) scale, all the degrees of freedom are integrated out. The effective distribution is then described in terms of the effective Hamiltonian $\Gamma[M]$, defined as the Legendre transform of the generating functional $\mathcal{W}[J]$:

$$\Gamma[M] = \sum_i j_i m_i - \mathcal{W}[J], \tag{6}$$

assuming convexity of $\mathcal{W}[J]$. The effective classical field $M = \{m_i\}$ being defined as:

$$m_i := \frac{\partial \mathcal{W}[J]}{\partial j_i}. \tag{7}$$

The RG transformations provide a path between these two extreme scales, built as a chain of partial integrations of microscopic fluctuations. More precisely, the

Wilson–Kadanoff perspective supposes the existence of slicing in the configuration space of elementary degrees of freedom, such that for each step along the chain, the integration over elementary degrees of freedom into a given slice provides a new effective distribution and a new definition of the UV scale. To be more technical, they assume the existence of finite partitions $\nu_I[\Phi]$ defining a slicing such that $\nu_I \cap \nu_J = \emptyset$.[3] For $I \neq J$ and $\forall\ i$ there exists one and only one I such that $\phi_i \in \nu_I[\Phi]$, and $\Phi = \sum_I \nu_I[\Phi]$. From these partitions, effective distributions $p_1, p_2, \ldots$ may be obtained from integration over each slice:

$$p_I[\Phi_I] = \frac{\mathrm{e}^{-\mathcal{H}_I[\Phi_I]}}{Z_I}, \quad p_{I+1}[\Phi_{I+1}] \propto \int \mathrm{d}\nu_I[\Phi]\mathrm{e}^{-\mathcal{H}_I[\Phi_I]},$$

where $\mathrm{d}\nu_I[\Phi] := \prod_{i|\phi_i\in\nu_I[\Phi]} \mathrm{d}\phi_i$ and $\Phi_I := \sum_{J\geqslant I} \nu_J[\Phi]$. In terms of Hamiltonians, the RG procedure builds a chain of effective Hamiltonian $\mathcal{H}_I$:

$$\mathcal{H} \Rightarrow \mathcal{H}_1 \Rightarrow \mathcal{H}_2 \Rightarrow \cdots. \tag{8}$$

It is along this chain that the relevance of some operators becomes of crude importance. For a sufficient number of steps, only relevant operators survive, and the microscopic irrelevant details are erased. However, even if we dispose of canonical notions for UV and IR, a preferred path to join UV and IR does not necessarily exist, and the results may depend on the path that we consider, i.e. on the arbitrary ordering of the elementary fluctuations in the partial integration chain. As explained before, the UV scales have to correspond with degrees of freedom associated with small eigenvalues of the covariance matrix spectrum, and this intuition is recovered in the field theoretical embedding proposed in [19]. The idea is to interpret the variance matrix $\mathcal{K} := \mathcal{C}^{-1}$ as the kinetic kernel for the N variables ϕ_i for $p[\Phi]$. In this respect, the simpler distribution we can think about is the Gaussian one:

$$\mathcal{H}[\Phi] = \frac{1}{2}\sum_{i,j} \phi_i \mathcal{K}_{ij} \phi_j, \tag{9}$$

the matrix $\mathcal{K}$ being symmetric by construction. The covariance matrix is interpreted as the second-order cumulant of the Gaussian distribution[4].

Finally, the RG can be constructed as for any field theory, starting by integrating out the degrees of freedom with higher variance. In other words, the spectrum of the matrix $\mathcal{K}$ gives a canonical size for the fluctuations and provides a canonical path to join UV and IR scales towards less noisy degrees of freedom. Obviously, for a purely Gaussian distribution, such a procedure is without interest, as the Gaussian distribution is a fixed point of the RG. More interesting is to evaluate the behavior of perturbations to pure Gaussian distributions, and this is the point we discuss in this paper. To be

[3] To be more precise, in a concrete example, the partitions ν_I turn to be distributions over a finite range of wavelength, and may have a non-vanishing but small covering between them.
[4] The 'vacuum' two-point function in the theoretical physics language.

more concrete, and following [19], we denote as $u_i^{(\mu)}$ the set of normalized eigenvectors of $\mathcal{K}$, with eigenvalues λ_μ,

$$\sum_j \mathcal{K}_{ij} u_j^{(\mu)} = \lambda_\mu u_i^{(\mu)}, \quad \sum_i u_i^{(\mu)} u_i^{(\mu')} = \delta_{\mu\mu'}, \tag{10}$$

we straightforwardly deduce that the Gaussian part $\mathcal{H}_{\mathrm{G}}[\Phi]$ reduces to $\mathcal{H}_{\mathrm{G}}[\Phi] = \frac{1}{2}\sum_\mu \lambda_\mu \psi_\mu \psi_\mu$, with $\psi_\mu := \sum_i \phi_i u_i^{(\mu)}$, and each elementary fluctuation has the size:

$$\langle \psi_\mu \psi_\mu \rangle \sim \frac{1}{\lambda_\mu}. \tag{11}$$

Then, in the most general case, for a non-purely Gaussian distribution, the fluctuations can be integrated out following their proper size given by the non-trivial spectrum of the matrix $\mathcal{K}$. We denote as λ_0 the smallest eigenvalue $\lambda_\mu \geqslant \lambda_0$.[5] Anticipating the next section, and in accordance with the field theoretical language, we define the square *momenta* p_μ^2 as $p_\mu^2 := \lambda_\mu - \lambda_0$. The notation reflects the fact that the matrix $\mathcal{K}$ is positively definite by construction; all the eigenvalues have to be positives $\lambda_\mu \geqslant 0\,\forall\mu$ and $p_\mu^2 \geqslant 0$. The isolated eigenvalue λ_0 plays the role of a mass and the Gaussian part of the Hamiltonian takes the form:

$$\mathcal{H}_{\mathrm{G}} = \frac{1}{2}\sum_\mu \psi_\mu (p_\mu^2 + \lambda_0)\psi_\mu. \tag{12}$$

For the reader familiar with quantum and statistical field theory, the previous relation is very reminiscent of the standard kinetic action for scalar field theory, with (square) mass equal to λ_0.

2.2. The model

2.2.1. Maximum entropy estimate. A coarse-graining in information theory reduces the information from each step. This has the consequence that (in the absence of critical lines) the divergence (or mutual entropy) between two distributions has to decrease along with the flow. Thus, RG formalism allows to make contact with the intuition that arbitrariness of the cut-off between noise and information should be connected with a partial integrating process over some microscopic degrees of freedom, in such a way that only a few numbers of parameter survive after some steps. The fact that RG concerns non-Gaussian distributions is clear by constructions. Gaussian distributions are stable for each step of the RG map. In technical terms, the Gaussian fixed point corresponds to a fixed point of the theory, and the distribution remains self-similar at each intermediate scales. Then, the first question is: is the Gaussian distribution stable for a noisy signal? In other words, assuming a little deviation from the Gaussian fixed point, this perturbation must decrease or increase along with the RG flow? As we will see below, the answer is no for the MP distribution. However, it seems clear from the discussion of the previous section that the answer to this question has to depend on

[5] Which is the larger eigenvalue of $\mathcal{C}$.

J. Stat. Mech. (2022) 033101

the shape of the spectral distribution $\rho(\lambda)$. For this reason, we can hopefully detect the presence of information in a signal around a noisy distribution from the universal properties of distributions in the deep IR. These properties can be, for instance, the number of independent correlations functions required to parameterize them, up to the experimental threshold. A physical example of the role played by the shape of the distribution in the asymptotic behavior of the distributions is provided by ferromagnetic metals. Here, the shape of the momenta distribution is $\sim(p^2)^{D/2-1}$. For space dimension $D>4$, the Gaussian fixed point is stable, and Gaussian distributions remain valid to describe the asymptotic behavior of the ferromagnet below the Curie temperature. In contrast, for $D<4$, the Gaussian fixed point becomes unstable, and an interacting theory is required to describe the ferromagnetic transition. Interestingly, such a change of behavior has been stressed in [19], investigating the behavior of the normalized four-point function $\langle\phi_i^4\rangle/\langle\phi_i^2\rangle^2$ by gradually integrating out degrees of freedom. Observed that the behavior of the normalized four-point function is drastically modified when some percentages of the higher eigenvalues of their spectra are suppressed, and this is what we will formalize in this paper.

In [19], the authors especially focused on the following truncation into the theory space:

$$\mathcal{H}[\Phi]:=\frac{1}{2}\sum_{i,j}\phi_i\mathcal{K}_{ij}\phi_j+\frac{g_1}{4!}\sum_i\phi_i^4. \tag{13}$$

In this reference, the authors motivated this choice by the argument that it is a maximum entropy distribution, and thus the least structured possible for our limited knowledge on systems. This is a general principle underlying statistical mechanics [52, 53]. However, this argument imposes a bias, namely the unjustified assumption *a priori* that the four moments are well represented by a local interaction $\sum_i\phi_i^4$. In this paper, we propose a new argument based on the universality of the MP law to justify this model. Indeed, a spectrum like the one in figure 6, close to the MP law, is 'blind' concerning the particular nature of the data. It can be financial data, the activity of neurons in the brain, correlations between genes, or anything else. This is the power of the universality theorems (e.g. MP law), to show that very different systems behave in the same way when they are large (here we assume $N\gg 1$). If we assume that this universality reflects that of an effective field theory for the data, we can hope to grasp the essential aspects of it by choosing a particular model, whose correlation matrices are in the neighborhood of the equivalence class considered (here MP). One can for instance consider a system of discrete spins $s_i\pm 1$ placed on the N nodes of an arbitrary network. The correlations of such a system can be complicated, and in general one has access to only partial information on the true distribution $p_{\text{true}}[\sigma]$ of spin configurations $\sigma=\{s_1,\dots,s_N\}$. Assuming the distribution hse 0 mean and covariance C, the maximum entropy estimate takes the form of standard Gibbs–Boltzmann states, i.e.

$$p[\sigma]\propto\exp\left(\frac{1}{2}\sum_{i,j=1}^N s_iK_{ij}s_j\right),$$

with the constraint: $\sum_{\sigma} p[\sigma]s_i s_j = C_{ij}$. It is now easy, from this expression, to derive an effective field theory for a continuous variable $\phi_i \in \mathbb{R}$, following the well-known procedure allowing to pass from the Ising model to the Ginsburg–Landau theory [51]. Using the properties of the Gaussian integral, we find that the effective distribution for the field ϕ_i follows the exponential law $p[\phi] = \exp(-H[\phi])$, with[6]:

$$H[\phi] = \frac{1}{2}\sum_{i,j=1}^{N} \phi_i \left(C_{ij}^{-1} - \delta_{ij}\right)\phi_j + \frac{1}{12}\sum_{i=1}^{N} \phi_i^4 + \cdots ,$$

which has exactly the structure of the action (13). It should be noted for 2D Ising model that the spectral density changes shape in the vicinity of the critical temperature, behaving like a power law, whereas it agrees with the universal predictions of the MP law for high temperatures [49, 50]. One could thus work with this theory on data about the behavior of discrete spins—which would be quite close to the effective description of neural networks in the brain considered in [54]. However, the universality argument justifies its use for any spectrum, as long as one is close to a universal law (the MP law for our purpose).

2.2.2. RG flow and momentum representation. In [19], the authors showed that, following the choice of the distribution $\rho(\lambda)$ for the eigenvalues of $\mathcal{K}$, the interaction part may increase with the number of RG steps. The origin of this behaviour can be traced from the scaling itself. The intermediate scales between deep UV and deep IR are fixed by the size of the eigenvalues λ_μ. In other words, they provide a scaling, and g_1 may acquire a specific dimension concerning this scaling. This dimension dictates how the coupling constant g_1 is sensitive to the change of scale. In the RG process, this change of scale, in practice, has to correspond to a change of the cut-off Λ corresponding to the upper bound of the spectrum $\lambda_\mu \leqslant \Lambda$. Then, the natural unit along the increasing scales is the one of Λ, and the canonical dimension $d_1(\Lambda)$ of g_1 is defined as its proper dependence under a dilatation of Λ, so that, at first order, we must have for the dimensionless coupling $\tilde{g}_1$:

$$\frac{\mathrm{d}\tilde{g}_1}{\mathrm{d}\ln\Lambda} = -d_1\tilde{g}_1 + \mathcal{O}(\tilde{g}_1^2). \tag{14}$$

The terms of order $\tilde{g}_1^2$ include the effects of the fluctuations, which are progressively integrated out in the RG procedure. To summarize, integrating fluctuations to each step change the fundamental scale $\Lambda \to \Lambda'$ as well as the Hamiltonian, and therefore the couplings constants, whose part of the global modification comes from their proper dimension (provided by the term $-d_1\tilde{g}_1$ in equation (14)). We will return to the exact computation of d_1 in the next section, in which we will present a nonperturbative formalism allowing us to compute non-Gaussian fixed points far from the Gaussian region. A practical limitation of the truncation (13) comes from the choice of the representation. It is easier to do computations in the momentum space; however, translating

[6] Note that zero modes are removed in C, assuming to work in the region of large eigenvalues, where the signal is assumed to be.

the interaction in momentum space without knowledge of the eigenvectors leads to the mysterious quartic coupling:

$$\frac{g_1}{4!}\sum_i \phi_i^4 = \frac{g_1}{4!}\sum_{\{\mu_j\}}\left(\sum_i \prod_{j=1}^{4} u_i^{(\mu_j)}\right)\prod_{j=1}^{4}\psi_{\mu_j}. \tag{15}$$

A naive way to circumvent the difficulty is to note that, after all, the initial interaction is not fundamental. There are no experimental data to justify this interaction, and we may construct an approximation directly in momentum space. The original interaction is reminiscent of the familiar ϕ^4 interaction $\int \phi^4(x)\mathrm{d}x$, restricted to the positive (or negative) region. Heuristically, if we discard the boundary problems, the fields can be decomposed over 'sin' or 'cos' functions instead of ordinary Fourier transform, which are together eigenfunction for the ordinary Laplacian function. The coupling tensor in the bracket in equation (15) then behaves like:

$$\int \cos(k_1x)\cos(k_2x)\cos(k_3x)\cos(k_4x)\mathrm{d}x, \tag{16}$$

which is essentially a sum of deltas of the form $\delta(k_1+\epsilon_2k_2+\epsilon_3k_3+\epsilon_4k_4)$, with $\epsilon_i=\pm 1$. The symmetry of the delta function leads to two distinct couplings, and following this we have one or two positives ϵ_i. By direct inspiration, we choose the following combination of Kronecker deltas:

$$\delta_1\mathcal{H}\sim\sum_{\{\mu_i\}}\left(\tilde{g}_1\delta_{0,p_1^2+p_2^2-p_3^2-p_4^2}+\tilde{g}_2\delta_{0,p_1^2+p_2^2+p_3^2-p_4^2}\right)\prod_{j=1}^{4}\psi_{\mu_j}. \tag{17}$$

We expect that this Hamiltonian must have the same physical content as the original one (13). In particular, and even if it is a little bit caricatural, we expect that ensuring momentum conservation at the vertex level provides a good representation of the original local interaction $\sum_i \phi_i^4$. Unfortunately, the Hamiltonian (17) introduces spurious singular behaviors at the origin for someone loop corrections. From a direct inspection, the problem comes from the positivity of the momenta. A simple way to circumvent this difficulty is the following. Instead of positive eigenvalues, we consider the momentum p as a relative integer $p\in\mathbb{Z}_{\sqrt{\Lambda-\lambda_0}}$, so that p^2 remains distributed following $\rho(\lambda)$, with $\lambda=p^2+\lambda_0$. Moreover, we introduce a new field $\psi(p)$, to distinguish them from the ψ_μ considered above. Then, we chose for $\psi(p)$ the new Hamiltonian:

$$\begin{aligned}\mathcal{H}[\Psi]=&\frac{1}{2}\sum_p\psi(-p)(p^2+\lambda_0)\psi(p)\\&+\frac{\tilde{g}}{4!}\sum_{\{p_i\}}\delta_{0,p_1+p_2+p_3+p_4}\prod_{j=1}^{4}\psi(p_j).\end{aligned} \tag{18}$$

To summarize, from the Hamiltonian [13], we kept essentially three elements:

J. Stat. Mech. (2022) 033101

(a) Without interaction, the correlation functions are essentially given by the eigenvalues of the covariance matrix $\mathcal{C}$.

(b) The square of the momenta p^2 are distributed following the distribution $\rho(\lambda)$.

(c) The interactions are essentially locals in the usual sense.

We can justify (18) in the same way as (13), by exploiting the fact that we suppose to work in the neighborhood of a universal law of a set of matrices (here the MP law). If $\tilde{g}=0$, the model is still Gaussian and describes the two-point correlations as well as (13), and we have no reason to favor one over the other. Both are equivalent in terms of experimental constraints. We are also more interested in the tail of the spectrum, where we look for signal effects[7]. In that regime, the behavior of the spectrum, near the MP class, can be well approximated by a power law. This is moreover what happens to be the case of ordinary field theories in dimension D, whose moment distribution is related to the dimension by $\tilde{\rho}(p^2) \sim (p^2)^{D/2-1}$. If we choose to start from the Gaussian theory defined by (18), then we expect the theory to behave like an ordinary field theory in that regime. However, this raises the dimension issue. Indeed for MP law, the momentum distribution at the tail of the spectrum behaves like $\tilde{\rho}(p^2) \propto (p^2)^{1/2}$ (see equation (45)), which corresponds to a field theory in dimension $D=3$. Obviously, finite-size effects or the presence of a signal will modify the value of this effective dimension, which is not an integer in general. To avoid this difficulty, we set the dimension of the space to 1 in the definition (18), while keeping a non-trivial moment distribution $\rho(p^2)$ (making the underlying 'matter field' unconventional); the last ensuring that the leading effects of the RG are expected to be the same for the effective theory or a more exact and particular model (13). As with any experimentally inspired approach, the numerical analysis in the next section will highlight the relevance of this formalism.

2.2.3. N-scaling. We expect that for our investigations about the stability of the Gaussian distribution, only these three points are really relevant, and we only consider the Hamiltonian (18) for explicit computations using nonperturbative formalism in sections 3 and 4. However, we will continue to use the Hamiltonian (13) as well for some discussion, to make contact with the reference [19]. Note to conclude that, in contrast to the coupling g in (13), the coupling $\tilde{g}$ involved in (18) has to be of order $1/N$ in order to ensure extensivity of the model. This can be checked as follows. In the words of physicists, let us consider a 'one loop correction' to the two-point function (i.e. a correction of order g):

$$\langle \phi_i \phi_j \rangle \sim \mathcal{C}_{ii} \delta_{ij}. \tag{19}$$

Then, setting $i=j$ and summing over i:

$$\sum_i \langle \phi_i \phi_i \rangle \sim \mathrm{Tr}(\mathcal{C}) = \sum_\ell \lambda_\ell^{-1}. \tag{20}$$

[7] This is further motivated in (2021) *Entropy* **23** 1132.

In contrast, let us consider the same kind of correction for the Hamiltonian (18),

$$\langle \psi_\mu \psi_{\mu'} \rangle \sim \delta_{p_\mu, p_{\mu'}} \sum_\ell \lambda_\ell^{-1}, \tag{21}$$

so, setting $p_\mu = p_{\mu'}$ and summing over p_μ, we get:

$$\sum_\mu \langle \psi_\mu \psi_\mu \rangle \sim \left(\sum_\mu 1 \right) \sum_\ell \lambda_\ell^{-1} = N \sum_\ell \lambda_\ell^{-1}. \tag{22}$$

Therefore, to ensure extensivity, the coupling with tilde has to be of order $1/N$, as expected. To make this dependence explicit, we keep the following expression, without tilde:

$$\mathcal{H}[\Psi] = \frac{1}{2} \sum_p \psi(-p)(p^2 + \lambda_0)\psi(p) + \frac{g}{4!N} \sum_{\{p_i\}} \delta_{0, p_1+p_2+p_3+p_4} \prod_{j=1}^4 \psi(p_j). \tag{23}$$

We will return to this scaling at the moment of the derivation of the flow equations in section 4 below.

3. A nonperturbative framework

3.1. The exact RG equation

One expects that the accuracy of the results obtained in [19] may be improved taking into account higher couplings and loop effects, motivating a nonperturbative analysis. The most powerful formalism to keep the nonperturbative effect of the RG is the FRG formalism, essentially based on the Wetterich–Morris equation [36–38]. In this section, we propose constructing a version of this formalism adapted to the PCA investigations. As we recalled, the Wilson–Kadanoff procedure requires splitting into modes, between UV scales (no fluctuations are integrated out) and IR scales (all the fluctuations are integrated out) dictating how the small distance fluctuations are integrated out. In the FRG formalism, this progressive integration of UV modes works with a *momentum-dependent mass* term $\Delta\mathcal{H}_k[\Phi]$ added to the microscopic Hamiltonian $\mathcal{H}[\Phi]$. In momentum representation:

$$\Delta\mathcal{H}_k[\Psi] = \frac{1}{2} \sum_p \psi(-p) r_k(p^2) \psi(p), \tag{24}$$

J. Stat. Mech. (2022) 033101

where k plays the role of a referent momentum scale. The regulator function $r_k(p^2)$ has to satisfy some elementary requirements, in such a way that:

(a) Small distance fluctuations ($p^2 > k^2$) are unaffected by the presence of $\Delta\mathcal{H}_k[\Psi]$ and integrated out.

(b) Long distance fluctuations ($p^2 < k^2$) acquire a large mass and are frozen out.

In this way, the momentum-dependent mass $r_k(p^2)$ must satisfy the elementary requirements:

(a) $r_k(p^2)$ has to have a non-vanishing IR limit, $p^2/k^2 \to 0$,

(b) $r_k(p^2) \to 0$ in the UV limit, for $p^2/k^2 \gg 1$,

(c) $r_k(p^2)$ has to vanish in the limit $k \to 0$, allowing to recover the original partition function when all the degrees of freedom are integrated out,

(d) $r_k(p^2)$ has to be of order Λ for $k^2 \to \Lambda$, Λ referring to the larger eigenvalue of $\mathcal{K}$.

Introducing this mass term into the microscopic Hamiltonian, we replace the global description given by the initial generating functional $\mathcal{Z}[J] := \int \mathrm{d}\Psi\, \mathrm{e}^{-\mathcal{H}[\Psi]+\sum_p j(p)\psi(p)}$ with a one-parameter set of models indexed by k, $\{\mathcal{Z}_k[J]\}$ defined as:

$$\mathcal{Z}_k[J] := \int \mathrm{d}\Psi\, \mathrm{e}^{-\mathcal{H}[\Psi]-\Delta\mathcal{H}_k[\Psi]+\sum_p j(p)\psi(p)}, \tag{25}$$

When the scale k decreases, more and more degrees of freedom are integrated out. The infinitesimal transcription of this goes through a first order differential equation [36–38]:

$$\begin{aligned}\dot{\Gamma}_k &= \frac{1}{2}\sum_\mu \dot{r}_k(p_\mu^2)\left(\Gamma_k^{(2)} + r_k\right)^{-1}_{\mu,-\mu}\\ &= N\int_{p\geqslant 0} \mathrm{d}p\, \tilde{\rho}(p^2)p\dot{r}_k(p^2)\left(\Gamma_k^{(2)} + r_k\right)^{-1}(p,-p),\end{aligned} \tag{26}$$

where, to write the last line, we introduced the momentum density $\tilde{\rho}(p^2)$, related to the eigenvalues density $\rho(\lambda)$ as:

$$\tilde{\rho}(p^2) := \rho(p^2 + \lambda_0) := \frac{1}{N}\sum_\mu \delta(\lambda - \lambda_\mu). \tag{27}$$

Note that the normalization has to be chosen such that the number of degrees of freedom remains equal to N:

$$N\int_{p\geqslant 0} \tilde{\rho}(p^2)p\,\mathrm{d}p = \frac{N}{2}\int \mathrm{d}\lambda\, \rho(\lambda) = \frac{N}{2}. \tag{28}$$

J. Stat. Mech. (2022) 033101

Equation (26) indicates how the *average effective Hamiltonian* Γ_k changes in the windows of scale $[k - \mathrm{d}k, k]$—the dot meaning derivative with respect to the RG parameter $t := \ln k$: $\dot{X} = k\frac{\mathrm{d}}{\mathrm{d}k}X$. We recall that the average effective action is defined as a slightly modified Legendre transform of the free energy $\mathcal{W}_k := \ln\, \mathcal{Z}_k$:

$$\Gamma_k[M] + \Delta\mathcal{H}_k[M] = \sum_p j(p)m(p) - \mathcal{W}_k[J], \tag{29}$$

with $M := \{m(p)\}$. Moreover, note that $\Gamma_k^{(2)}$ in equation (26) denotes the second derivative of the average effective action:

$$\left[\Gamma_k^{(2)}\right]_{\mu\mu'} := \frac{\partial^2\Gamma_k}{\partial m(p_\mu)\partial m(p_{\mu'})}, \tag{30}$$

where $m(p_\mu)$ is defined as: $m(p) = \partial\mathcal{W}_k[J]/\partial j(p)$.

3.2. Scaling dimension

The scaling (or canonical) dimension is defined as the intrinsic dependence of a quantity on the cut-off coming from its dimension. In standard quantum field theory, for instance, the dimensions of the couplings are closely related to renormalizability. In this case, the dimensions are inherited from the referent background space-time where the field is described. The difficulty here comes from the fact that we do not have any background space-time in (25) to fix the dimensions. However, it may be instructive to return to the standard field theory case. Let us for instance consider a free scalar field $\varphi(x)$ defined over $\mathbb{R}^d$; with Hamiltonian:

$$\mathcal{H} = \int \mathrm{d}x\, \varphi(x)(-\Delta)\varphi(x). \tag{31}$$

Here, the role of the matrix $\mathcal{K}$ is played by the Laplacian Δ, whose spectrum fixes the size of the fluctuations and discriminates between IR and UV scales. The eigenvalues will be denoted as p^2, referring to their positivity, and we must have $[p] = -[x]$; $[X]$ denoting the dimension of the quantity X. We fix the definition of the bracket such that $[p] = 1$. The dimension of the fields $\varphi(x)$ can be read directly from the previous expression from the requirement that $\mathcal{H}$ have to be dimensionless, we get:

$$[\varphi] = \frac{d-2}{2}. \tag{32}$$

Now, let us consider the two-point function $\langle\varphi(x)\varphi(x)\rangle$,

$$\langle\varphi(x)\varphi(x)\rangle \sim \frac{1}{p^2}, \tag{33}$$

so that:

$$\int \mathrm{d}x \langle \varphi(x)\varphi(x)\rangle \sim \int^{\Lambda} \frac{\mathrm{d}p}{p^2} \propto \Lambda^{d-2} \tag{34}$$

for some UV cut-off Λ. The dependence on Λ reflects the dimension of the field given by (32), and suggests a way to define dimension without referent background. It must be fixed by the rescaling $\varphi \to \bar{\varphi} = \Lambda^{-[\varphi]}\varphi$ such that $\int \mathrm{d}x\langle\bar{\varphi}(x)\bar{\varphi}(x)\rangle$ becomes essentially independent of Λ for sufficiently large Λ:

$$\frac{\mathrm{d}}{\mathrm{d}\Lambda} \int \mathrm{d}x \langle \bar{\varphi}(x)\bar{\varphi}(x)\rangle \approx 0. \tag{35}$$

By direct inspiration, we fix the rescaling z_Λ of the variables ϕ_i, $\phi_i \to \tilde{\phi}_i := z_\Lambda \phi_i$ such that $\sum_i \langle \tilde{\phi}_i^2\rangle$ becomes Λ-independent:

$$\frac{\mathrm{d}}{\mathrm{d}\Lambda}\sum_i \langle (z_\Lambda \phi_i)^2\rangle = 0 \to \frac{\mathrm{d}\,\ln\, z_\Lambda}{\mathrm{d}\,\ln\,\Lambda} = -\frac{1}{2}\frac{\rho(\Lambda)}{\int \mathrm{d}\lambda \frac{\rho(\lambda)}{\lambda}}. \tag{36}$$

The dimension of the coupling constant must be fixed following the same strategy. With the field dimension $d_1 = -\ln z_\Lambda$ being fixed, the dimension of the quartic coupling g must be fixed from the extensivity argument used in [19]. The authors argue that the average of the Hamiltonian (which is essentially the energy in physics) must be an extensive quantity, and therefore proportional to the number of effective degrees of freedom. This requirement in particular ensures that the entropy is an extensive quantity as well. For some cut-off Λ, the number of effective degrees of freedom N_{eff} is nothing but the number of eigenvalues λ_μ, so then we must have:

$$N_{\text{eff}} := N \int^{\Lambda} \mathrm{d}\lambda\, \rho(\lambda). \tag{37}$$

Therefore, one expects that the rescaling allowing to pass from g_1 to the dimensionless coupling $\bar{g}_1$ must be such that:

$$N g_1 \frac{1}{N}\sum_i \phi_i^4 = N_{\text{eff}}\bar{g}_1 \sum_i \frac{1}{N}(z_\Lambda \phi_i)^4, \tag{38}$$

which fixes the rescaling of the coupling constant as:

$$\frac{\mathrm{d}\,\ln \bar{g}_1}{\mathrm{d}\ln\Lambda} = \rho(\Lambda)\left[\frac{2}{\int \mathrm{d}\lambda \frac{\rho(\lambda)}{\lambda}} - \frac{\Lambda}{\int \mathrm{d}\lambda\, \rho(\lambda)}\right]. \tag{39}$$

J. Stat. Mech. (2022) 033101

The formula can be easily generalized for an interaction involving p fields:

$$\delta\mathcal{H} \propto g_p \sum_{i_1,\dots,i_p} \mathcal{V}_{i_1,\dots,i_p} \prod_{a=1}^{2p} \phi_{i_a} \equiv g_p \mathcal{V}[\phi^{2p}], \tag{40}$$

where the symbol $\mathcal{V}_{i_1,\dots,i_p}$ must be a product of Kronecker delta, identifying indices pairwise. Let us denote as $n(\mathcal{V})$ the number of Kronecker delta. Therefore, equation (38) has to be replaced by:

$$N g_p \frac{1}{N^{p-n(\mathcal{V})}} \mathcal{V}[\phi^{2p}] = N_{\text{eff}} z_\Lambda^{2p} \bar{g}_p \frac{1}{N^{p-n(\mathcal{V})}} \mathcal{V}[\phi^{2p}] \tag{41}$$

leading to:

$$\frac{\mathrm{d}\ \ln \bar{g}_p}{\mathrm{d}\ln\Lambda} = \rho(\Lambda) \left[\frac{p}{\int \mathrm{d}\lambda \frac{\rho(\lambda)}{\lambda}} - \frac{\Lambda}{\int \mathrm{d}\lambda\, \rho(\lambda)} \right]. \tag{42}$$

The canonical dimension that we discussed here corresponds to the one discussed in [19]. However, for FRG applications, we must find the scaling concerning the running scale k, not concerning the fundamental cut-off Λ. Once again, the question is trivial for standard quantum field theory, the dimension being the same as that which we use Λ of k as referent scale. But here, this is not trivial, because the definition of the canonical dimension seems to introduce a Λ dependence for some distributions $\rho(\lambda)$. In contrast, the distribution $\dot{r}_k$ ensures that only windows of momenta around $p_\mu = k$ contribute significantly to the integral in the right-hand side of (26). This is why it more natural to define the canonical scaling for the running scale, and fix the dimensions concerning this parameter, rather than for Λ, which introduces a spurious reference to the microscopic physics. This difficulty may be solved in the same way as we defined the canonical dimension for fields. Returning to the free scalar field $\varphi(x)$, for some regulators $r_k(p^2)$ in the Fourier space, we must have:

$$\int \mathrm{d}x\, \mathrm{d}y \langle \varphi(x) \dot{r}_k(x-y) \varphi(y) \rangle \sim \int^\Lambda \frac{\mathrm{d}p}{p^2} \dot{r}(p^2) \propto k^{d-2+[r_k]}.$$

Therefore, it must be possible to use this relation to define the dimension, as the rescaling of the fields such that the right-hand side scale is $k^{[r_k]}$. Moving on to our field theory for ϕ_i, it is clear from definition of r_k that $[r_k]$ must be equal to 1. Then, one expects that there exists a rescaling z_k of the fields, such that:

$$\frac{\mathrm{d}}{\mathrm{d}k} \sum_{i,j} \langle (z_k\phi_i)(\dot{r}_k r_k^{-1})_{ij}(z_k\phi_j) \rangle \approx 0, \tag{43}$$

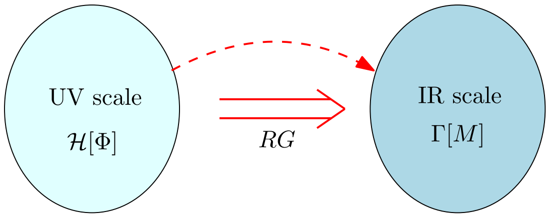

Figure 1. The canonical definition of UV and IR scales from the integration point of view. In the UV scale, no degree of freedom is integrated, whereas they are all integrated out in the IR scale. RG then provides a path through scales, from UV to IR.

for sufficiently large k. The two definitions for z_Λ and z_k seems to be different. However, a moment of reflection shows that they have to coincide at least in the deep UV[8], for very large k and Λ. In fact, it is easy to check that for power-law distributions $\rho(\lambda) \propto \lambda^\alpha$, the right hand sides of equations (38) and (42) become pure numbers depending only on α, which is a general feature of homogeneous distributions $\tilde{\rho}(ap) = a^\beta \tilde{\rho}(p)$. For all cases, it is easy to check that the dimensions are the same using the momentum cut-off Λ and the distribution $\dot{r}_k$, for sufficiently large k. In the next sections, we will compare the flow obtained from spectra with signal and purely noisy signals. The MP distribution usually provides an efficient description of noisy signal, and it corresponds to the asymptotic spectrum of purely i.i.d. random matrices X_{ai}, with arbitrary large p and n, but p/n kept constant [47, 48]. Explicitly, the eigenvalue distribution $\mu(x)$ is the following (figure 1)

$$\mu(x) = \frac{1}{2\pi\sigma^2} \frac{\sqrt{(a_+ - x)(x - a_-)}}{kx}, \tag{44}$$

where:

- $k = p/n$ is the fixed ratio between the size indices of the random matrix X_{ai} with i.i.d. entries, with 0 means and variance $\sigma < \infty$.
- $a_\pm = \sigma^2(1 \pm \sqrt{k})^2$.

Now, let us consider the behavior of the distribution for large x, i.e. for x close to a_+ (the maximal eigenvalue), and expand it in power of $t = x - b$. At leading order, we get straightforwardly in the regime where $1/a_- \gg 1/\lambda$:

$$\tilde{\rho}(p^2) \approx \frac{1+\sqrt{k}}{2\pi\sigma k} \frac{(p^2)^{1/2}}{(p^2+m^2)^2}, \tag{45}$$

[8] Note that the dimensions may be fixed from a purely RG point of view, from the behavior of the RG flow in the vicinity of the Gaussian fixed point. More precisely, requiring that the first leading order perturbative corrections have the same scaling as the corresponding couplings provides a notion of dimension, which reduce to the previous one especially for power-law distributions $\rho(\lambda) \propto \lambda^\alpha$; provided that $\alpha > -1$. For instance, assuming that λ_0 has dimension 1 in Λ, we find that the first leading order quantum correction to mass scales is Λ^α, and therefore requires that the corresponding coupling scale is $\Lambda^{1-\alpha}$ in order to get a global scaling identical to the first term, which is nothing, but that which we get explicitly from equation (42). We will return to this approach in the next section.

https://doi.org/10.1088/1742-5468/ac52a6

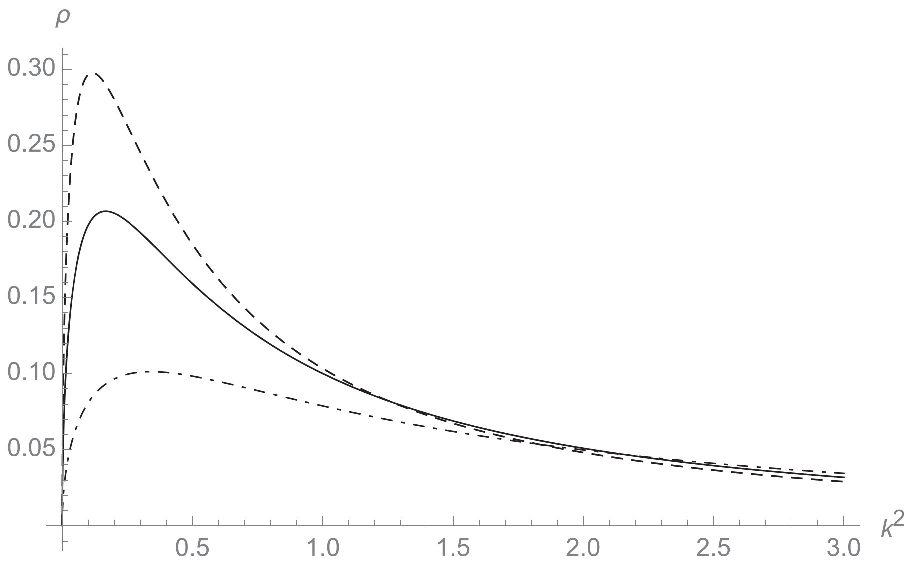

Figure 2. The MP momentum representation of the MP distribution in the deep UV. The dashed curve is for $\sigma^2 = 1.2$, the solid curve for $\sigma^2 = 1$ and the dashed-dotted curve for $\sigma^2 = 0.7$.

where $m^2 := 1/a_+$. The corresponding distribution is plotted in figure 2. Note that the distribution (44) is for the inverse of the kinetic kernel $\mathcal{K} = \mathcal{C}^{-1}$, in contrast with the law (45), as the notations suggest. The relation between μ and ρ could be easily deduced from the fact the number of eigenvalues is the same for $\mathcal{K}$ and $\mathcal{C}$, leading to:

$$\rho(x) = \mu(x^{-1})\frac{1}{x^2}. \tag{46}$$

4. Solving RG using local potential approximation

Solving the exact nonperturbative RG flow equation (26) is a difficult task, even in very special cases. Therefore, extracting some information about this equation requires approximations. The difficulty to solve the exact RG equation (26) may be pointed out as follows. Taking the second derivative of (26) with respect to the classical field m_μ, $\partial^2/\partial m_\mu \partial m_{\mu'}$, we get an equation for $\dot{\Gamma}_k^{(2)}$. Assuming that odd functions $\Gamma_k^{(2n+1)}$ vanish identically, we get that the right-hand side involves $\Gamma_k^{(4)}$ and the effective propagator $G_k := (\Gamma_k^{(2)} + r_k)^{-1}$. Deriving once again two times concerning the classical field, we get an equation for $\dot{\Gamma}_k^{(4)}$, involving G_k, $\Gamma_k^{(4)}$ and $\Gamma_k^{(6)}$, and so on. Taking successive derivatives, we then generate an infinite tower of coupled equations. All the approximation schemes used to solve the RG equations have to aim close this hierarchy. In this paper, we focus on the crude truncation approximation, imposing:

$$\Gamma_k^{(2n)} \approx 0, \tag{47}$$

J. Stat. Mech. (2022) 033101

until a certain n. The restriction to even functions, i.e. $\Gamma_k^{(2n+1)} = 0$ reflect the $\phi \to -\phi$ symmetry of the original microscopic action, and corresponds to expanding the truncated action Γ_k around vanishing classical field $m_\mu = 0$, $\forall\ \mu$. We call the *symmetric phase* the portion of the phase space parametrized like that, and in this introductory paper, we only focus on this approximation. Focusing on this approximation, we consider the following truncation around $n = 3$:

$$\begin{aligned}\Gamma_k[M] = {}& \frac{1}{2}\sum_p m(-p)(p^2 + u_2(k))m(p)\\ &+ \sum_{\{p_i\}} \left(\frac{g(k)}{4!N}\delta_{0,\sum_i p_i}\right)\prod_{j=1}^{4} m(p_j).\end{aligned} \tag{48}$$

Such a truncation defines a parametrization of the *theory space*, i.e. the space of allowed actions, and is called LPA [42–46]; the momentum dependence on the effective vertex being completely discarded. One expects that such an approximation works well for this kind of theory [42]. The dependence on k for the coupling constant $g(k)$ and for effective mass $u_2(k)$ reflects the integration of UV degrees of freedom when k varies on equation (26). Starting calculations still requires two ingredients: the momentum distribution and the regulator function r_k. For the last one, we choose the standard Litim regulator [39, 40], allowing us to do analytic computations:

$$r_k(p^2) = (k^2 - p^2)\theta(k^2 - p^2), \tag{49}$$

the $\theta(x)$ being the Heaviside step function, equal to 1 for $x \geqslant 1$ and to 0 otherwise. Note that physical solutions of the exact RG flow equation (26), in principle, do not depend on r_k. However, the approximation used to solve it generally introduces a dependence on the regulator, which has to be investigated. The regulator (49) has been investigated to be optimal for such dependence for some models [41], which is another practical advantage to making this choice. The remaining ingredient is the momentum distribution. In principle, this distribution has to come from a data set and does not have an analytic form as for the MP law. To keep contact with the reference paper [19], we will firstly consider the example of a power-law distribution $\tilde{\rho}(p^2) \propto p^{2\alpha}$. Note that ordinary field theories provide a non-trivial example of such a power-law behavior, the value of α being trivially connected to the space-dimension through $\alpha = D/2 - 1$. However, for spectra relevant in data analysis, such an approximation cannot be considered better than a caricature or a limit case. For instance, for $p \ll 1$, the MP law (45) is such that $\tilde{\rho}(p^2) \propto (p^2)^{1/2}$, but remains true only in the deep IR sector. However, to investigate the behavior of the RG flow for such a power-law remains an instructive exercise for the unfamiliar reader, to show how the different concepts like canonical dimension work, and to show how this behavior is influenced by the shape of the distribution. We thus briefly discuss this case in the next section, before moving on to more realistic spectra.

J. Stat. Mech. (2022) 033101

4.1. Power law distribution

The flow equations for $\dot{g}$, and $\dot{u}_2$ can be deduced by projection of the exact flow equation (26) along the reduced portion of the full phase space parametrized by the truncation (48). To implement this, let us consider the derivative of (26) with respect to $m(p)$ and $m(p')$. Because $\Gamma_k^{(3)} = 0$, we get:

$$\dot{\Gamma}^{(2)}_{k,\mu_1\mu_2} = -\frac{1}{2}\sum_{\mu} \dot{r}_k(p_\mu^2) G_{k,\mu\mu'} \Gamma^{(4)}_{k,\mu'\mu''\mu_1\mu_2} G_{k,\mu''\mu}. \tag{50}$$

Setting $\mu_1 = \mu_2$, from the truncation (48), it follows that:

$$\dot{\Gamma}^{(2)}_{k,\mu_1\mu_1} = \dot{u}_2. \tag{51}$$

Moreover, the derivative of $r_k(p_\mu)$ can be easily computed: $\dot{r}_k(p^2) = 2k^2\theta(k^2 - p^2)$. Finally, in the symmetric phase, $\Gamma_k^{(2)}$ must be easily computed:

$$\Gamma^{(2)}_{k,\mu_1\mu_2} = \delta_{p_{\mu_1},-p_{\mu_2}} \left(p_{\mu_1}^2 + u_2(k)\right), \tag{52}$$

and (50) reduces to:

$$\dot{u}_2 = -\frac{1}{2}\frac{2k^2}{(k^2+u_2)^2}\sum_{\mu} \theta(k^2 - p_\mu^2)\Gamma^{(4)}_{k,\mu\mu\mu_1\mu_1}\Big|_{p_{\mu_1}=0}. \tag{53}$$

The last sum involves sums over different permutations of external momenta, arising from derivations. However, for large k, all these terms do not provide a significant contribution. Let us consider the contribution arising from the coupling g:

$$\Gamma^{(4)}_{k,\mu_1\mu_2\mu_3\mu_4} = \frac{g}{4!N}\sum_{\pi} \delta_{0,p_{\pi(1)}+p_{\pi(2)}+p_{\pi(3)}+p_{\pi(4)}}, \tag{54}$$

from which we have the sum running over the set of permutation of the four external momenta. Because all the momenta play the same role, all the permutations contribute to the right-hand side of equation (53), with the typical contribution being written as:

$$\frac{1}{N}\sum_{\mu} \theta(k^2 - p_\mu^2)\frac{g}{4!}\delta_{0,p_1+p_2} \sim k^{2\alpha+2}. \tag{55}$$

Therefore, taking into account the scaling of the coupling constant, we get a global k dependence as $k^{2\alpha+2}k^2k^{-4}k^{2-2\alpha} = k^2$, which is nothing but what we expect for the proper scaling of u_2. Note that, as explained in footnote 4 at the end of section 3, the scaling may be deduced directly from these equations. Indeed, assuming that g scales as k^{d_1}, we get that contributions as in (55) scale as $k^{d_1-2+2(\alpha+1)}$. Moreover, the scaling of u_2, which is essentially the lower eigenvalue of the spectrum of the kinetic kernel, must scale as Λ, and therefore as[9] k^2. Then, we must have $d_1 = 2(1-\alpha)$.

[9] All the eigenvalues must be homogeneous to the upper eigenvalue, ensuring that under a global dilatation the shape of the spectrum remains unchanged.

J. Stat. Mech. (2022) 033101

From these observations, and counting the number of relevant contractions in (53), we obtain:

$$\dot{u}_2 = -\frac{g}{1+\alpha}\frac{k^{4+2\alpha}}{(k^2+u_2)^2}. \tag{56}$$

This equation involves couplings having dimensions. Due to the fact that it is suitable to work with dimensionless couplings, we introduce the dimensionless parameters:

$$u_2 =: k^2\bar{u}_2, \quad g =: k^{2(1-\alpha)}\bar{g}, \tag{57}$$

leading to ($\beta_2 := \dot{\bar{u}}_2$):

$$\beta_2 = -2\bar{u}_2 - \frac{\bar{g}}{1+\alpha}\frac{1}{(1+\bar{u}_2)^2}. \tag{58}$$

To find the equation for $\dot{g}$, we proceed in exactly the same way. We take the fourth derivative of the flow equation (26), applying $\partial^4/\partial m_{\mu_1}\partial m_{\mu_2}\partial m_{\mu_3}\partial m_{\mu_4}$, and setting all the external momenta to be equal to:

$$\dot{\Gamma}^{(4)}_{k,\mu_1\mu_1\mu_1\mu_1} = 3\sum_{\mu,\mu',\mu'',\mu''',\mu''''}\dot{r}_k(\mu)G_{\mu\mu'}\Gamma^{(4)}_{k,\mu'\mu''\mu_1\mu_1} \times G_{\mu''\mu'''}\Gamma^{(4)}_{k,\mu'''\mu''''\mu_1\mu_1}G_{\mu''''\mu}.$$

From the truncation, it follows that: $\dot{\Gamma}^{(4)}_{k,0,0,0,0} = \frac{\dot{g}}{N}$. Thus, repeating the same analysis as for the flow of u_2, we get for the dimensionless coupling $\bar{g}$ ($\beta_g := \dot{\bar{g}}$):

$$\beta_g = -2(1-\alpha)\bar{g} + \frac{6}{1+\alpha}\frac{\bar{g}^2}{(1+\bar{u}_2)^3}. \tag{59}$$

The flow equations (58) and (59) exhibit fixed points, which can be easily found by solving the system $\beta_2 = \beta_4 = 0$. In addition to the Gaussian fixed point, with $\bar{g} = \bar{u}_2 = 0$, we get the non-Gaussian fixed point:

$$p = (\bar{u}_2^*, \bar{g}^*) = \left(-1 - \frac{6}{\alpha-7}, 72\frac{\alpha^2-1}{(\alpha-7)^3}\right), \tag{60}$$

which provides an explicit example of how the RG flow depends on α. In particular, we see that the value $\alpha = 1$ corresponds to what we call a critical dimension in standard field theory. For $\alpha < 1$, a fixed point is reminiscent of the standard WF fixed point; with one attractive and one repulsive direction[10]. For $\alpha > 1$, however, the behavior of the RG flow is governed by the Gaussian fixed point, which becomes unstable (see figure 3 for an illustration).

[10] We recall the standard vocabulary in physics: in the vicinity of a fixed point, a direction is relevant (toward the UV scales) if the eigenvalue is positive, irrelevant if it is negative and marginal if it is zero. Moreover, for the Gaussian fixed point, the eigenvalues are nothing but the canonical dimensions.

J. Stat. Mech. (2022) 033101

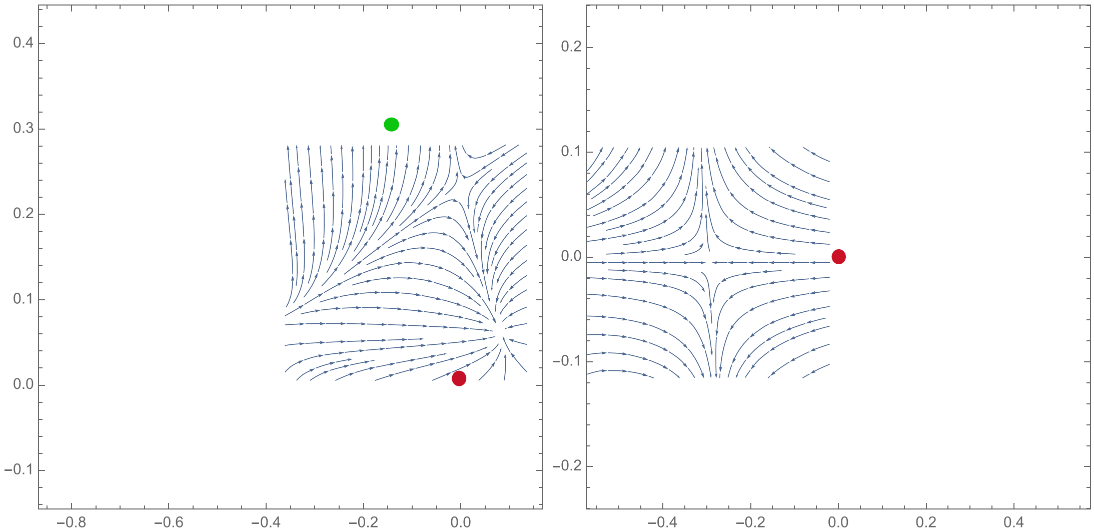

Figure 3. From left to right, numerical RG flow for $\alpha = 0.5$ and $\alpha = 1.5$. In both cases, the red and green points correspond respectively to Gaussian and non-Gaussian fixed points, and the arrows are oriented toward UV scales (from small to big k).

4.2. Flow equations without power law approximation

As a first observation, the flow equations involve loop integrals of type:

$$L := \int_0^k \tilde{\rho}(p^2) p \, \mathrm{d}p, \tag{61}$$

which, for large k, and $\tilde{\rho}(p^2) \propto (p^2)^\alpha$ behaves like $k^{2\alpha+2}$, as computed in the previous section. Therefore:

$$\ln L = (2\alpha + 2)\ln(k) + C, \tag{62}$$

For C being a numerical constant depending on α and $\ln(k) = t$, the scale parameter along the RG flow is abusively called *time*. Therefore, for a power law distribution, $\tau := \ln L$ and t are related by an affine transformation, with $\mathrm{d}\tau$ and $\mathrm{d}t$ being proportional. The two times are, in this respect, essentially physically equivalent. Obviously, such a relation breaks down for arbitrary distributions. However, flow equations simplify using τ rather than t. Let us consider the flow of u_2:

$$\dot{u}_2 = -\frac{2gk^2}{(k^2 + u_2)^2} \int_0^k \tilde{\rho}(p^2) p \, \mathrm{d}p. \tag{63}$$

Computing the derivative:

$$\frac{\mathrm{d}\tau}{\mathrm{d}t} = k^2 \tilde{\rho}(k^2) \frac{1}{\int_0^k \tilde{\rho}(p^2) p \, \mathrm{d}p}, \tag{64}$$

J. Stat. Mech. (2022) 033101

we get:

$$\dot{u}_2 = -\frac{2g}{(1+\bar{u}_2)^2}\tilde{\rho}(k^2)\frac{\mathrm{d}t}{\mathrm{d}\tau}, \tag{65}$$

where we used the fact that $\mathrm{d}t/\mathrm{d}\tau = (\mathrm{d}\tau/\mathrm{d}t)^{-1}$. Multiplying term by term with $\mathrm{d}t/\mathrm{d}\tau$, we then get:

$$\frac{\mathrm{d}\bar{u}_2}{\mathrm{d}\tau} = -2\frac{\mathrm{d}t}{\mathrm{d}\tau}\bar{u}_2 - \frac{2g}{(1+\bar{u}_2)^2}\frac{\tilde{\rho}(k^2)}{k^2}\left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right)^2. \tag{66}$$

Therefore, defining the dimensionless coupling $\bar{g}$ as:

$$g\frac{\tilde{\rho}(k^2)}{k^2}\left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right)^2 =: \bar{g}, \tag{67}$$

we finally obtain:

$$\frac{\mathrm{d}\bar{u}_2}{\mathrm{d}\tau} = -2\frac{\mathrm{d}t}{\mathrm{d}\tau}\bar{u}_2 - \frac{2\bar{g}}{(1+\bar{u}_2)^2}. \tag{68}$$

In the same way, for the coupling, we get:

$$\frac{\mathrm{d}g}{\mathrm{d}\tau} = \frac{12g^2}{(1+\bar{u}_2)^3}\frac{\tilde{\rho}(k^2)}{k^2}\left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right)^2. \tag{69}$$

We have to write the equation for the dimensionless coupling $\bar{g}$, therefore, we need to compute the derivative of the dimensionless coupling given by (67):

$$\bar{g}' = g'\frac{\rho(k^2)}{k^2}\left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right)^2 + 2\bar{g}\left(\frac{t''}{t'} + t'\left(\frac{1}{2}\frac{\mathrm{d}\,\ln\,\tilde{\rho}}{\mathrm{d}t} - 1\right)\right). \tag{70}$$

Therefore:

$$\frac{\mathrm{d}\bar{g}}{\mathrm{d}\tau} = 2\bar{g}\left(\frac{t''}{t'} + t'\left(\frac{1}{2}\frac{\mathrm{d}\,\ln\,\tilde{\rho}}{\mathrm{d}t} - 1\right)\right) + \frac{12\bar{g}^2}{(1+\bar{u}_2)^3}, \tag{71}$$

where:

$$X' := \frac{\mathrm{d}X}{\mathrm{d}\tau}. \tag{72}$$

The role played by the canonical dimension is now played by a more complicated function, defining scale by scale:

$$-\dim(g) := 2\left(\frac{t''}{t'} + t'\left(\frac{1}{2}\frac{\mathrm{d}\,\ln\,\tilde{\rho}}{\mathrm{d}t} - 1\right)\right), \tag{73}$$

which is pictured in figure 4 for the MP distribution. The dimension is positive everywhere, meaning, as expected, that the Gaussian fixed point will be stable placing the cut-off at an arbitrary scale. This scale invariance may be viewed as another

J. Stat. Mech. (2022) 033101

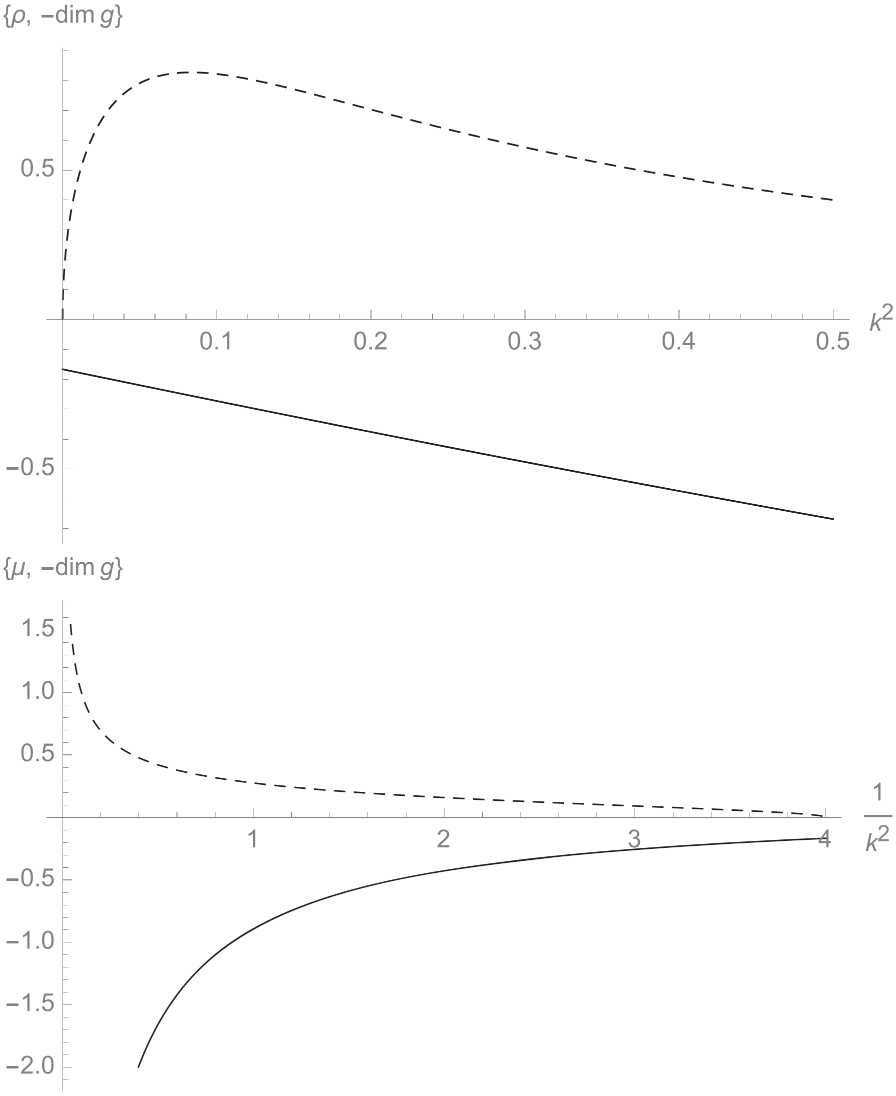

Figure 4. (Top) Canonical dimension (solid line) versus the inverse MP distribution (the dashed line). (Bottom) The canonical dimension (solid line) versus the MP distribution (dashed line).

characterization of noise, and the MP law provides the common analytic representation of noisy signals. The question is now, what happens when the MP law is disturbed by a macroscopic signal? One expects that, for a sufficiently big signal, the canonical dimension will become negative from a certain scale, breaking the scale invariance. As we will see in the next section, this is precisely what happens.

5. Numerical investigation

In this section, as announced, we investigate data sets from an RG point of view. We focus on artificial sets, made of some constant spikes disturbed by a random signal

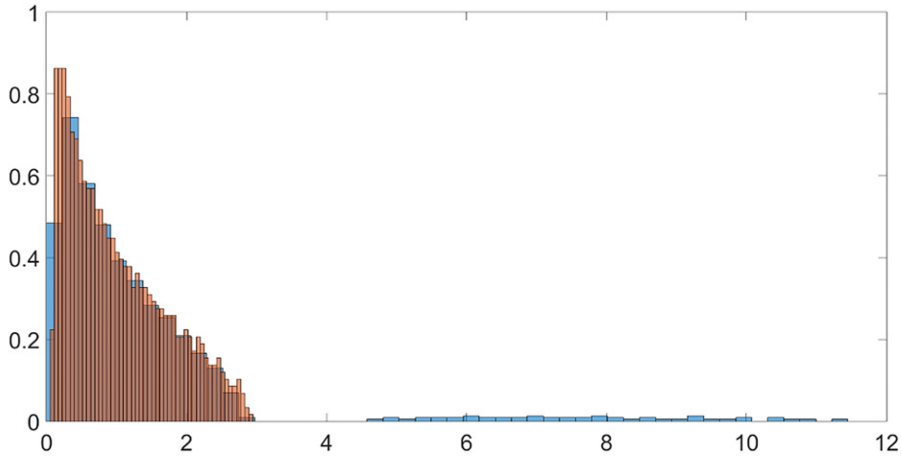

Figure 5. Red histogram: data set with i.i.d. random entries for $p = 1000$, $N = 2000$. Blue histogram: perturbation of the random distribution with a matrix of rank $k = 50$ (defining the size of the signal) and ratio $x = 300$.

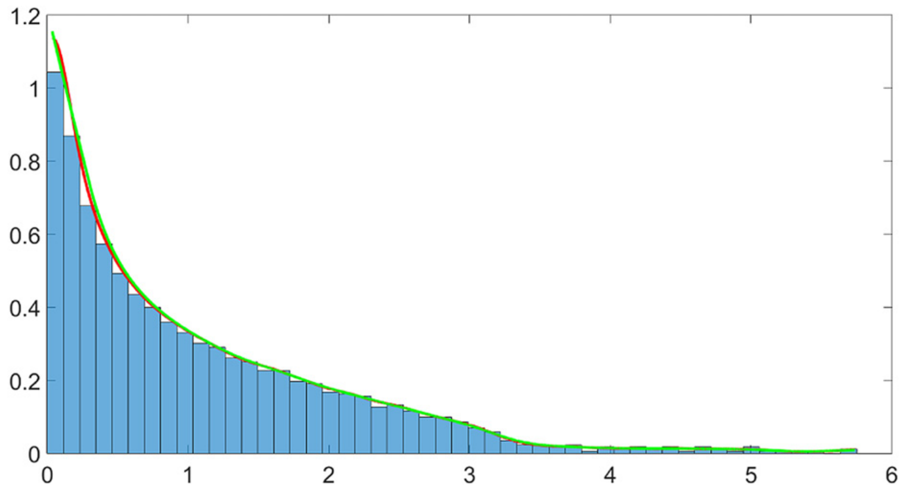

Figure 6. Eigenvalue distribution for $p = 1500$, $N = 2000$; and the eigenvalues of the matrix playing the role of the signal distributed with different weights. For this case, there is no clean separation between noisy and relevant degrees of freedom. The green curve corresponds to a numerical interpolation of the discrete distribution.

materializing with a matrix with random entries, playing the role of the noise. Figure 5 provides a typical spectrum, for purely random entries (the red histogram) and when a non-random signal is added (the blue histogram). In this simple case, the standard PCA could be applied, the largest eigenvalues being far from the bulk, which tend toward the MP distribution for large p and N. The distinction between signal and noise is therefore clear for this example. This is however not the case for the spectrum given in figure 6 below.

J. Stat. Mech. (2022) 033101

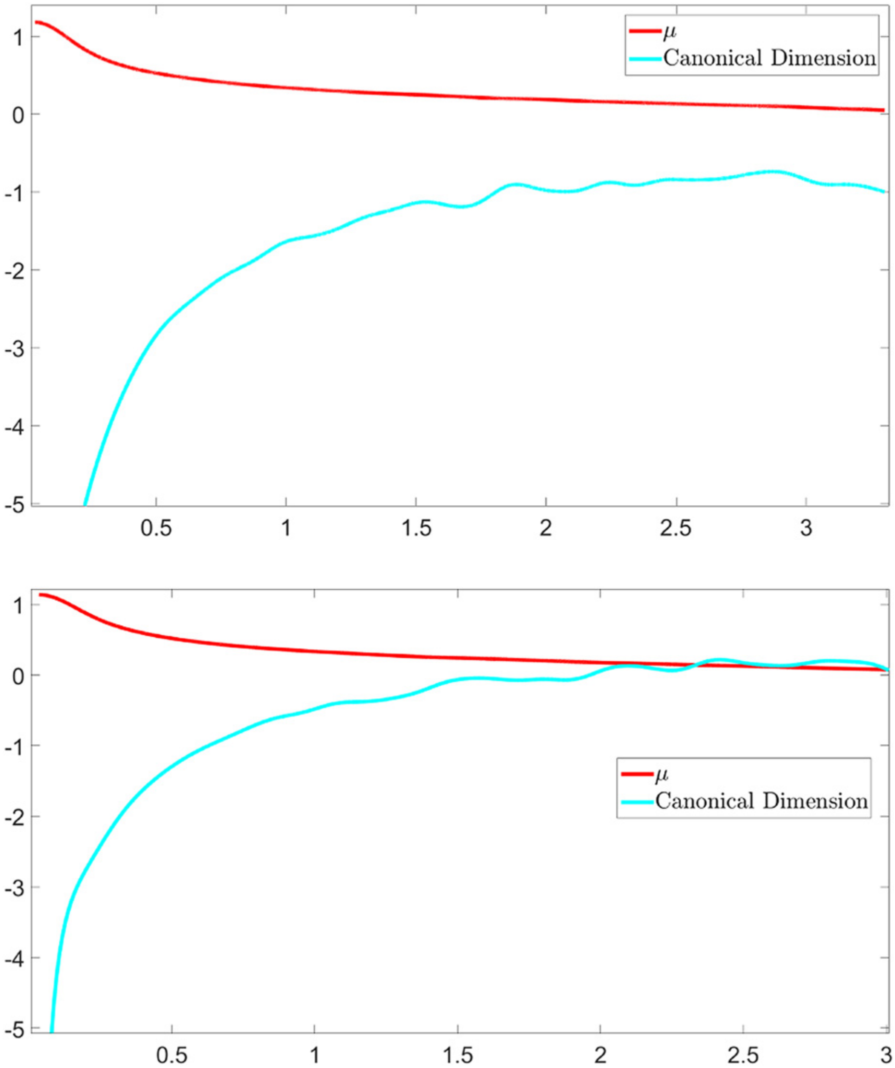

Figure 7. The canonical dimension version of the eigenvalue distribution. Without signal (top) and with signal (bottom).

In this figure, there is no clean separation between what is information and what is noise, with the interesting part of the signal being merged onto the noise. Figure 7 provides the canonical dimension with and without signal. For the case without signal, we recover the same curve as for the MP law, see figure 4, up to some irrelevant irregularities due to the numerical interpolation. Indeed, the computation of the canonical dimension requires a first and second derivative, which are very sensitive to the sharp variations of the interpolation. Adding the signal, we show that the canonical dimension is changed, and increases significantly in the region of large eigenvalues, to become positive from a certain scale. The interpretation of the phenomena is clear from the analysis of section 4.1. For very small eigenvalues, the canonical dimension is essentially unaffected by the signal, and the curve is the same as the one without the signal.

J. Stat. Mech. (2022) 033101

Moving toward the large eigenvalue region, however, the deformation increases and the canonical dimension becomes larger, meaning, from the point of view of the RG that system reaches the critical region. Finally, the canonical dimension becomes positive, and the Gaussian fixed point becomes unstable. This illustrates what we expected in the discussion of section 2. The universal properties of the large scale distributions are affected by the presence of the signal up to a certain range, which we can be 'detected' by the behavior of the RG flow. Note that this observation seems to contradict the assumptions justifying the perturbative treatment in [19]. Indeed, what we show is that a purely noisy signal is not suitably described by a Gaussian distribution. This can be the case however if the strength of the signal makes the scaling dimension for the coupling negative. Interestingly, the canonical dimension becomes positive around 2.2, well before the theoretical end of the MP distribution. At this stage, the signal prevails over the noise. However, the competition between them starts before this point. From the point where the signal goes out, the estimated size for the signal is 52, which has to be compared to the number 65 of eigenvalues for the input signal. Then, in this example, the method allows recovering 80 percent of the original signal.

6. Conclusion and perspectives

In this paper, we propose a revised formalism for the connection between RG and PCA stressed in [19]. The fact that both PCA and RG search for simplifying systems involving a large degree of freedom was an indication that the RG technique could be applied in the PCA context when standard methods fail. This is especially the case for continuous spectra, where no significant 'gap' in eigenvalue exists, and the problem of finding optimal k-means clustering is NP-hard. In that regard, RG is expected to provide an alternative point of view, aiming to extract relevant features by sophisticated coarse-graining. The expected 'separation', in the RG language, plays the role of a cut-off over degrees of freedom and varying it, we expect that noisy and relevant degrees of freedom are distinguished from their influence on the RG behavior.

In this paper, we considered a novel and unconventional field theory, describing some scalar 'matter' filling an abstract space of dimension 1, but having non-trivial momentum distribution fixed by the experimental spectrum of the covariance matrix regarding some large dimensional dataset. This model is expected to be in the same universality class as the effective maximum entropy estimate model considered in [19]. However, the proposed model is more easily adapted for numerical computations that are suited for practical data analysis. In particular, with the model chosen in [19], it is not possible to obtain figures 4 and 7.

This work can be seen as the proposal of this new model, from which we have drawn the first consequences, which are in agreement with the literature and justify *a posteriori* the relevance. For this model, we were in particular able to introduce a nonperturbative framework and discuss arbitrary spectra from a criterion involving an unconventional generalization of the so-called canonical dimension. This scaling dimension is positive for the MP universality class, and for large enough N, which is then a universal statement as well. Aiming to keep control on numerical simulation, we considered artificial data sets, built as a disturbance of a purely noisy signal with a deterministic matrix. Our main result is that disturbance has a non-trivial influence on the power counting, tending to make the interactions irrelevant, qualitatively changing asymptotic states. Finally, an important point to mention is the key role played by the mass. Our formalism introduces a mass, which happens to be the only relevant parameter defining, in the first approximation, the spectrum deformation.

These results added to the ones of [19, 21, 54] show that RG may be fruitfully used as a promising tool in the PCA context for challenging problems.

Surely, some questions remain open and have to be clarified in the future. First of all, the criteria provided by the dimension requires the computation of the first and second derivatives, which are not easy to estimate from discrete spectra. In the analysis of section 5, we used numerical tools to improve the analyticity of the interpolation curve, discarding the most singular points. A finer numerical technique has to be used to improve this point. Secondly, our approximations are limited to the simpler truncation, and deeper investigations of the theory space could provide finer arguments to clarify the cut-off between noisy and relevant degrees of freedom. For instance, one can imagine that higher truncations could reveal finer details over data, breaking scale invariance at higher scales. Another aspect concerns the role of phase transitions, where a very large number of microscopic degrees of freedom behave collectively to generate macroscopic effects. This is a common feature in physical systems involving many degrees of freedom, and our analysis seems to indicate that approaching the separation point between information and noise the behavior of the RG flow is very reminiscent of a system near criticality. One may expect that such an effective description allows improving the criteria, reasoning over the efficient dimensions near the non-Gaussian fixed point rather than the canonical dimensions, only valid in the Gaussian region.

References

[1] Avdoshkin A and Dymarsky A 2019 Euclidean operator growth and quantum chaos (arXiv:1911.09672)

[2] Sünderhauf C, Piroli L, Qi X L, Schuch N and Cirac J I 2019 Quantum chaos in the Brownian SYK model with large finite N: OTOCs and tripartite information *J. High Energy Phys.* JHEP11(2019)038

[3] Franz S, Mézard M, Parisi G and Peliti L 1998 Measuring equilibrium properties in aging systems *Phys. Rev. Lett.* **81** 1758

[4] Bouchaud J P, Cugliandolo L F, Kurchan J and Mezard M 1997 Out of equilibrium dynamics in spin-glasses and other glassy systems (arXiv:cond-mat/9702070)

[5] Bouchaud J-P, Cugliandolo L, Kurchan J and Mézard M 1996 Mode-coupling approximations, glass theory and disordered systems *Physica* A **226** 243

[6] Mezard M and Parisi G 1988 The Euclidean matching problem *J. Physique* **49** 2019–25

[7] Krauth W and Mezard M 1987 Learning algorithms with optimal stability in neural networks *J. Phys. A: Math. Gen.* **20** L745

[8] Carleo G, Cirac I, Cranmer K, Daudet L, Schuld M, Tishby N, Vogt-Maranto L and Zdeborová L 2019 Machine learning and the physical sciences (arXiv:1903.10563)

[9] Loebl M and Zdeborová L 2008 The 3D Dimer and Ising problems revisited *Eur. J. Comb.* **29** 966

[10] Charbonneau P and Yaida S 2017 Nontrivial critical fixed point for replica-symmetry-breaking transitions *Phys. Rev. Lett.* **118** 215701

[11] Yeo J and Moore M A 2012 Renormalization group analysis of the M-p-spin glass model with $p = 3$ and $M = 3$ *Phys. Rev.* B **85** 100405

J. Stat. Mech. (2022) 033101

[12] Castellana M and Parisi G 2010 Renormalization group computation of the critical exponents of hierarchical spin glasses *Phys. Rev.* E **82** 040105

[13] Pezzella U and Coniglio A 1997 Spin glasses and frustrated percolation: a renormalization group approach *Physica* A **237** 353

[14] Dotsenko V S 1987 Renormalisation group theory of spin glasses *J. Phys. C: Solid State Phys.* **20** 5473

[15] Collet P, Eckmann J-P, Glaser V and Martin A 1984 A spin glass with random couplings *J. Stat. Phys.* **36** 89

[16] Aygün E and Erzan A 2011 Spectral renormalization group theory on networks *J. Phys.: Conf. Ser.* **319** 012007

[17] Mehta P and Schwab D 2014 An exact mapping between the variational renormalization group and deep learning (arXiv:1410.3831)

[18] Hattori K, Hattori T and Watanabe H 1987 Gaussian field theories on general networks and the spectral dimensions *Prog. Theor. Phys. Suppl.* **92** 108

[19] Bradde S and Bialek W 2017 PCA meets RG *J. Stat. Phys.* **167** 462–75

[20] Shlens J 2014 A tutorial on principal components analysis (arXiv:1404.1100)

[21] Bradde S, Nourmohammad A, Goyal S and Balasubramanian V 2019 The size of the immune repertoire of bacteria (arXiv:1903.00504)

[22] Bao C 2019 Loop optimization of tensor renormalization: algorithms and applications *Thesis* University of Waterloo

[23] Campeti P, Poletti D and Baccigalupi C 2019 Principal component analysis of the primordial tensor power spectrum *J. Cosmol. Astropart. Phys.* JCAP09(2019)055

[24] Woloshyn R M 2019 Learning phase transitions: comparing PCA and SVM (arXiv:1905.08220)

[25] Benzécri J P 1976 *L'analyse des Données: Leçons sur l'analyse Factorielle et la Reconnaissance des Formes et Travaux du Laboratoire de Statistique de l'Université de Paris VI* vol 1 (Paris: Dunod)

[26] Hotelling H 1933 Analysis of a complex of statistical variables into principal components *J. Educ. Psychol.* **24** 417–41

[27] Abdi H and Williams L J 2010 Principal component analysis *Wiley Interdiscip. Rev. Comput. Stat.* **2** 433–59

[28] Lu H, Plataniotis K N and Venetsanopoulos A N 2011 A survey of multilinear subspace learning for tensor data *Pattern Recognit.* **44** 1540–51

[29] Yue G and Jennifer D 2009 Sparse probabilistic principal component analysi *Journal of Machine Learning Research Workshop and Conf. Proc.*

[30] Foreman S, Giedt J, Meurice Y and Unmuth-Yockey J 2018 Machine learning inspired analysis of the Ising model transition *PoS LATTICE* vol 2018 p 245

[31] Bény C 2018 Inferring relevant features: from QFT to PCA *Int. J. Quantum Inform.* **16** 1840012

[32] Foreman S, Giedt J, Meurice Y and Unmuth-Yockey J 2018 RG-inspired machine learning for lattice field theory *EPJ Web Conf.* **175** 11025

[33] Kadanoff L P *et al* 1967 Static phenomena near critical points: theory and experiment *Rev. Mod. Phys.* **39** 395

[34] Sokal A D, van Enter A C D and Fernandez R 1994 Regularity properties and pathologies of position space renormalization group transformations: scope and limitations of Gibbsian theory *J. Stat. Phys.* **72** 879

[35] Kadanoff L P 2009 More is the same; phase transitions and mean field theories *J. Stat. Phys.* **137** 777

[36] Wilson K G and Fisher M E 1972 Critical exponents in 3.99 dimensions *Phys. Rev. Lett.* **28** 240

[37] Wetterich C 1993 The average action for scalar fields near phase transitions *Z. Phys.* C **57** 451

[38] Wetterich C 1993 Exact evolution equation for the effective potential *Phys. Lett.* B **301** 90

[39] Litim D F 2000 Optimisation of the exact renormalisation group *Phys. Lett.* B **486** 92

[40] Litim D F 2001 Derivative expansion and renormalisation group flows *J. High Energy Phys.* JHEP11(2001)059

[41] Canet L, Delamotte B, Mouhanna D and Vidal J 2003 Optimization of the derivative expansion in the nonperturbative renormalization group *Phys. Rev.* D **67** 065004

[42] Delamotte B 2012 An introduction to the nonperturbative renormalization group *Lect. Notes Phys.* **852** 49

[43] Berges J, Tetradis N and Wetterich C 2002 Non-perturbative renormalization flow in quantum field theory and statistical physics *Phys. Rep.* **363** 223

[44] Gies H and Wetterich C 2002 Renormalization flow of bound states *Phys. Rev.* D **65** 065001

[45] Reuter M and Saueressig F 2011 *J. High Energy Phys.* JHEP12(2011)012

[46] Pawlowski J M, Scherer M M, Schmidt R and Wetzel S J 2017 Physics and the choice of regulators in functional renormalisation group flows *Ann. Phys., NY* **384** 165

[47] Kanzieper E and Singh N 2010 Non-Hermitian Wishart random matrices (I) *J. Math. Phys.* **51** 103510

[48] Lu X and Murayama H 2015 Universal asymptotic eigenvalue distribution of large N random matrices—a direct diagrammatic proof to Marchenko–Pastur law— (arXiv:1410.3503)

[49] Sps V and Seligman T 2014 Time series, correlation matrices and random matrix models *AIP Conf. Proc.* **1575** 196

[50] Sps V, Prosen T, Buca B and Seligman T 2014 Correlation matrices at the phase transition of the Ising model *Europhys. Lett.* **108** 20006
[51] Itzykson C and Drouffe J-M 1991 *Statistical Field Theory* vol 1 (Cambridge: Cambridge University Press)
[52] Jaynes E T 1957 Information theory and statistical mechanics. II *Phys. Rev.* **108** 171
[53] Jaynes E T 1957 Information theory and statistical mechanics. I *Phys. Rev.* **106** 620
[54] Meshulam L, Gauthier J L, Brody C D, Tank D W and Bialek W 2017 Collective behavior of place and non-place neurons in the hippocampal network *Neuron* **96** 1178

https://doi.org/10.1088/1742-5468/ac52a6