Published for SISSA by Springer

Received: *September 4, 2020*
Revised: *November 5, 2020*
Accepted: *November 17, 2020*
Published: *January 7, 2021*

Gentle introduction to rigorous Renormalization Group: a worked fermionic example

Alessandro Giuliani,[a,b] **Vieri Mastropietro**[c] **and Slava Rychkov**[d,e]

[a] *Dipartimento di Matematica e Fisica, Università degli Studi Roma Tre, L.go S. L. Murialdo 1, 00146 Roma, Italy*
[b] *Centro Linceo Interdisciplinare Beniamino Segre, Accademia Nazionale dei Lincei, Via della Lungara 10, 00165 Roma, Italy*
[c] *Dipartimento di Matematica, Università degli Studi di Milano, Via Cesare Saldini 50, 20133 Milano, Italy*
[d] *Institut des Hautes Études Scientifiques, Bures-sur-Yvette, France*
[e] *Laboratoire de Physique de l'Ecole normale supérieure, ENS, Université PSL, CNRS, Sorbonne Université, Université de Paris, F-75005 Paris, France*

E-mail: giuliani@mat.uniroma3.it, vieri.mastropietro@unimi.it

Abstract: Much of our understanding of critical phenomena is based on the notion of Renormalization Group (RG), but the actual determination of its fixed points is usually based on approximations and truncations, and predictions of physical quantities are often of limited accuracy. The RG fixed points can be however given a fully rigorous and non-perturbative characterization, and this is what is presented here in a model of symplectic fermions with a nonlocal ("long-range") kinetic term depending on a parameter ε and a quartic interaction. We identify the Banach space of interactions, which the fixed point belongs to, and we determine it via a convergent approximation scheme. The Banach space is not limited to relevant interactions, but it contains all possible irrelevant terms with short-ranged kernels, decaying like a stretched exponential at large distances. As the model shares a number of features in common with ϕ^4 or Ising models, the result can be used as a benchmark to test the validity of truncations and approximations in RG studies. The analysis is based on results coming from Constructive RG to which we provide a tutorial and self-contained introduction. In addition, we prove that the fixed point is analytic in ε, a somewhat surprising fact relying on the fermionic nature of the problem.

Keywords: Renormalization Group, Nonperturbative Effects

ArXiv ePrint: 2008.04361

https://doi.org/10.1007/JHEP01(2021)026

Contents

---

1 Introduction

Renormalization group (RG) is a pillar of theoretical physics, explaining how long-distance collective behavior emerges from microscopic models. Critical phenomena are thus understood in terms of RG fixed points (Wilson [1–3]) and universality is explained in terms of basins of attractions. While this beautiful picture qualitatively works very well, quantitative applications often lead to practical difficulties. To compute critical exponents, one typically does perturbation theory in a small parameter, like $\epsilon = 4-d$ in the ϵ-expansion [4]. The accuracy of this procedure is limited by the proliferation of Feynman diagrams, and by the slow convergence of Borel-resummed series (while without resummation it normally diverges). As a consequence, predictions of critical exponents using perturbative RG [5, 6] are often less accurate than lattice Monte Carlo simulations or the conformal bootstrap [7].

It should be stressed that Wilson did not consider RG limited to situations with a small coupling. Two strongly coupled RG examples can be found [2]. One is his famous solution of the Kondo problem. The other is less known but no less impressive: an RG calculation for the 2D Ising model in a space of 217 couplings, concluding that "one can do precise calculations using pure renormalization group methods with the only approximations being based on locality."

Other developments in theoretical physics suggest, indirectly, that Wilson ideas are non-perturbatively correct. We can mention here exact results in two-dimensional field theory (see e.g. [8]), obtained by conformal field theory, integrable models, exact S-matrix bootstrap etc., which have provided many examples of exact non-perturbative RG flows, never finding any inconsistency with Wilsonian expectations. In higher dimensions, exact results in supersymmetric theories (see e.g. [9]) as well as the gauge-gravity duality considerations (see e.g. [10, 11]) have always confirmed Wilson ideas.

It is therefore somewhat surprising that the most straightforward interpretation of Wilson's vision, as a non-perturbative machine which would allow non-perturbatively and with an essentially unlimited precision to compute the properties of any RG fixed point of interest, has not so far been achieved. It is fair to say that this was not for the lack of trying, see e.g. [12] for the early attempts.

Two notable, although not fully successful, attempts have been the Functional Renormalization Group (FRG) [13, 14] and Tensor Network Renormalization (TNR) [15–17]. The FRG calculations include couplings with arbitrary powers of fluctuating field, but only up to some finite derivative order. Unfortunately, with more derivatives, FRG results tend to become more and more sensitive to the parameters specifying the regulating function [18, 19]. This has been traced to increasing violations of conformal invariance, except at some special parameter values satisfying a "principle of minimal sensitivity" [20]. It remains to be understood why the convergence to the fixed point does not hold more robustly in FRG. As to the TNR, it works well for simple 2D lattice models such as the 2D Ising model but hasn't been yet as effective in higher dimensions.

So, in spite of these attempts, although it is generally believed that the non-perturbative Wilsonian RG fixed points do exist, at present they often remain Platonic objects, confined to the world of ideas and accessible to us only via approximations of rather limited accuracy. Take e.g. the RG fixed point for the 3D Ising model. If it exists, which Banach space does it belong to? Can we access it via a provably convergent approximation scheme? At the moment these questions are wide open.

Note that Wilson believed in the Ising RG fixed point very concretely: as a fixed-point Hamiltonian invariant under a Kadanoff block-spin transformation. As mentioned above, for 2D Ising, he found an approximate fixed-point Hamiltonian numerically, truncating to a space of 217 lattice spin interactions [2]. But the convergence of his scheme has never been proven, nor has it been implemented in 3D. Incidentally, Wilson did worry about rigorous convergence properties of RG maps; e.g. in [21] a model RG transformation was shown to be convergent for a rescaling parameter larger than 4×10^6.

Of course, as already mentioned, there are nowadays other methods to get precise values of critical exponents, most notably the conformal bootstrap [7]. However, this does not mean that the RG should be abandoned. First, RG is more general than the conformal bootstrap, since many RG fixed points important for physics do not have conformal invariance, such as any fixed point involving time evolution or relaxation, and having a dynamical critical exponent $z \neq 1$. Second, it is quite possible that there exist much better RG implementations, and we just haven't found them yet.

Since theoretical physicists have not been able to find a fully successful implementation of non-perturbative RG in spite of many attempts, can mathematical physics give any hint about what we have been doing wrong? In mathematical physics, the rigorous construction of non-trivial RG fixed points has been achieved in different cases using Constructive RG (CRG). It has been obtained in bosonic scalar field theories[1] [22–26] and interacting fermions [27] with long range interactions, in cases where the system has a scaling dimension differing from marginality by an ε. It has also been achieved in models with marginal interactions of strength λ and asymptotically vanishing beta function, such as 1D interacting fermions [28–31] and 2D spin, vertex and dimer models [32–39]. In all these cases, the non-perturbative existence of a non-trivial RG fixed point, close to the Gaussian or free Fermi one, has been proved for ε or λ sufficiently small, and the critical exponents can be computed at an arbitrary precision in terms of resummed perturbative expansions, with rigorous bounds on the remainder. A feature of the CRG is that the fixed point is found, without any approximation, in a Banach space of interactions where all the irrele-

[1] See section 8.1.8 and appendix K for more details about the rigorously constructed bosonic fixed points.

vant terms are nonlocal, even though fast decaying (e.g. like a stretched exponential): this is in striking contrast with the FRG, where the space of interactions is typically spanned by a sequence of *local* functions of the fluctuation field and its derivatives.[2]

One lesson of all this body of rigorous work is that *weakly coupled* non-perturbative RG is possible, both in the bosonic and fermionic case, although it is easier in the fermionic case because in this case convergent perturbation theory captures full non-perturbative physics. We emphasize that, in general, fermionic perturbation theory is expected to be convergent only in the running rather than the bare coupling, see section 1.3. On the other hand, *strongly coupled* non-perturbative RG has so far been out of reach of mathematical physics research. For this reason we will, as a first step, focus in this paper on the weakly coupled fermionic case, well understood by mathematical physicists, and aim to transfer this knowledge into the theoretical physics realm.

With this in mind, we will present here the rigorous construction of a non-Gaussian fixed point for a fermionic model with weakly relevant quartic interaction. This is possibly the simplest model of this kind where to test field-theoretical RG methods, and a perfect example to provide an introduction to CRG accessible to a wider audience. We will explain how these methods allow one to characterize a non-trivial fixed point without any ad-hoc assumption or any uncontrolled approximation. The above mentioned crucial role played by the space of *mildly nonlocal* interactions, as opposed to expanding all interactions in local functions of the field, will be evident from our presentation.

A complementary goal of our work will be to prove a new result, which is the *analyticity* of the ε-expansion for our non-Gaussian fermionic fixed point. This is in contrast with bosonic ε-expansions, which are, at best, Borel summable. Although we will focus on $\varepsilon > 0$ in much of the paper, eventually we will show analyticity in a complex disk around $\varepsilon = 0$. Although the sign of ε is correlated with the sign of the fixed point quartic interaction, Dyson's argument against analyticity does not apply for our model because fermions are allowed to have quartic interaction of either sign. Moreover, analyticity in ε of the fixed point is not in contradiction with the divergence of perturbation theory in the bare couplings, see section 1.3 for further comments.

Note that in this paper we only construct RG fixed points, and we do not discuss in detail the RG flow between the microscopic model and the constructed fixed points. In any case, the result about analyticity is only valid for the fixed point and does not extend to the full RG flow, whose very structure changes discontinuously with the sign of ε. For positive ε we will have the gaussian model at short distances, perturbed by the relevant quadratic and quartic couplings and flowing at long distances to the nontrivial RG fixed point. For negative ε it will be the other way around: starting from the nontrivial RG fixed point we may flow to the gaussian model at long distances, with the quartic coupling then describing the leading irrelevant interaction at long distances. The former situation is referred to as 'IR fixed point', while the latter as 'UV fixed point'. To avoid any ambiguity we stress that no reverse RG flow is implied: all RG flows are from short to long distances.

[2]For a fair comparison it should be noted that FRG calculations are often performed in terms of the 1PI effective action, not the Wilsonian effective action used here. Also, some FRG schemes do attempt to go beyond the local derivative expansion. See section 8.1.7.

The convergence of our CRG scheme lets us hope that it will be used as a starting point for developing systematic and stable truncation schemes for the fermionic FRG. Although we have been able to prove our theorems only for small ε, it might be that the range of practical applicability of our scheme is order one and includes strongly coupled fixed points — in the future one should try to see if this is the case (see section 8.1.5).

Rigorous non-perturbative constructions of bosonic fixed points present several complications compared to the fermionic case discussed here (see section 8.1.8). Extracting lessons from those constructions for practical RG calculations remains another important open problem for the future.

1.1 The model

The model we consider is schematically described by the following action:

$$\mathrm{aMFT}(\psi) + \nu_0 \int d^d x\, \psi^2 + \lambda_0 \int d^d x\, \psi^4, \tag{1.1}$$

where aMFT is an 'anticommuting Mean Field Theory' of the fermionic[3] field ψ with N components, and the two additional terms are quadratic and quartic interactions preserving $Sp(N)$ global symmetry. Our fermions will be scalars under rotation, rather than spinors. So the model is not reflection-positive, but reflection positivity will play no role in the RG analysis.

Local models of this kind were considered in [40] under the name 'symplectic fermions', and are relevant for the description of polymers and loop-erased random walks [41, 42]. In 3D, interacting symplectic fermion models have been recently considered in the context of dS/CFT correspondence [43]. See also [44, 45] for theoretical studies of related models, and [46–48] for other speculative appearances of symplectic fermions in physics. Our model is a variant of those, with a nonlocal ("long-range") kinetic term of the schematic form $\psi\,(\partial^2)^{\#}\,\psi$ where $(\partial^2)^{\#}$ is a non-integer power of the Laplacian. Similar long-range models in a large N limit were recently considered in [49] in relation to the SYK model.

More precisely, our model is defined as follows: take an even number N of real Grassmann fields $\psi = (\psi_a)_{a=1}^N$ in $\mathbb{R}^d$,[4] with $d = 1, 2, 3$. The reference Gaussian theory (the aMFT mentioned above) is characterized by the following two-point function:

$$\langle \psi_a(x)\psi_b(y)\rangle = \Omega_{ab} P(x-y) \equiv G_{ab}(x,y), \tag{1.2}$$

where $P(x) \propto \frac{1}{|x|^{d/2-\varepsilon}}$ at large distances (see the next section for the explicit expression), with ε small and positive, and Ω_{ab} is the symplectic $N \times N$ matrix:

$$\Omega_{ab} = \begin{bmatrix} & 1 & & & \\ -1 & & & & \\ & & \ddots & & \\ & & & & 1 \\ & & & -1 & \end{bmatrix} \tag{1.3}$$

[3] Fermionic=anticommuting=Grassmann in this paper.

[4] We will not make a distinction between the lower and upper $Sp(N)$ and Euclidean indices whose position is determined only by typographic convenience: $\psi^a \equiv \psi_a$ and $\partial_\mu \psi \equiv \partial^\mu \psi$.

The quadratic and quartic monomials ψ^2 and ψ^4 in (1.1) must be interpreted as $\Omega_{ab}\psi_a\psi_b$ and $(\Omega_{ab}\psi_a\psi_b)^2$. Given the form of $P(x)$, the fields ψ_a are assigned the scaling dimension $[\psi] = d/4 - \varepsilon/2$, so that the quadratic and quartic terms in (1.1) are both relevant, the quartic one being barely so for ε small and positive. The parameter ε plays a role similar to the deviation of spatial dimension d from 4, $\epsilon = 4 - d$ in the Wilson-Fisher ϵ-expansion, which, contrary to ours, is not at present suitable for a rigorous non-perturbative RG analysis because the space of $4 - \epsilon$ dimensions has not been rigorously defined so far.[5] For $d = 1, 2, 3$, there are no other local relevant or marginal terms in addition to those included in (1.1). In perturbative RG, the lowest order RG equations for the fixed point are, letting γ be the scaling parameter:

$$\nu = \gamma^{d/2+\varepsilon}(\nu + I_1\lambda + \cdots), \qquad \lambda = \gamma^{2\varepsilon}(\lambda + I_2\lambda^2 + \cdots), \tag{1.4}$$

where I_1 and I_2 are positive constants given by the one-loop Feynman diagrams. Neglecting the higher-order terms, we get a nontrivial fixed point $\lambda_* = (1 - \gamma^{2\varepsilon})/I_2$, $\nu_* = I_1\lambda_*/(1 - \gamma^{d/2+\varepsilon})$, which is $O(\varepsilon)$, close to the Gaussian one. This is just an approximation, and we want to be sure that the existence of the fixed point is not spoiled by non-perturbative effects caused by higher orders or irrelevant terms. Moreover, we want to define a scheme whose truncations provide arbitrarily good approximations of the actual fixed point, with apriori bounds on the error made.

1.2 Strategy and open questions

Our rigorous construction of the fixed point goes as follows. First, we identity a space which is left invariant by the RG iteration. We cannot restrict to the (finite) space of relevant couplings, since the RG transformation generates the irrelevant interactions, whatever the input action is. Similarly, we cannot restrict to the space of local irrelevant interactions, because it too is not left invariant by the RG map. The right choice turns out to be the span of all possible monomials in ψ and $\partial\psi$ with nonlocal, but sufficiently fast decaying, kernels: this space is left invariant. We stress that this mild nonlocality is unrelated to the long-range character of our reference Gaussian theory; it has to do with the fact that the IR-cutoff propagator is not fully local although short-range; it would be present also for the local kinetic term. Note that we couldn't find an invariant space of nonlocal monomials involving ψ only: in our construction the presence of derivative fields $\partial\psi$ is generated by what we call the *trimming* operation, which consists in extracting from a nonlocal monomial of order 2 or 4 its local part, and in re-expressing the nonlocal remainder in terms of irrelevant monomials of the form $\partial\psi\partial\psi$ or $\psi^3\partial\psi$. Note also that our construction does not exclude the existence of other invariant spaces, with different (non)locality properties of the kernels; in particular, it remains to be seen whether there exists an invariant Banach space consisting of local monomials in ψ and its derivatives (of arbitrary order), but we are not aware of any rigorous result in this sense. The construction of an invariant Banach space of interactions comes, in particular, with a non-perturbative definition of the RG

[5] See however [50] for a non-perturbative analysis in non-integer d using the conformal bootstrap.

JHEP01(2021)026

map: this is achieved via combinatorial cancellations due to the ± signs in the series expansion, ultimately due to the fermionic nature of the fields. In order to take advantage of these cancellations, we need to organize the perturbative expansion in the form of series of determinants, rather than in the more standard form of series of Feynman diagrams. This may be thought of as a smart rearrangement and partial resummation of the perturbative series: while the Feynman diagrams expansion is *not* absolutely convergent, the determinant expansion is. Once the invariant space has been identified and the RG map defined at a non-perturbative level, we prove that the RG map is *contractive* in a suitable neighborhood of the approximate lowest-order fixed point: this implies existence and uniqueness of the actual fixed point in such a neighborhood. (More precisely, the RG map is contractive near the fixed point along all directions but ψ^2, but this complication is easily taken care of.) Remarkably, such fixed point is analytic in ε.

Therefore, the problem of obtaining the correct Banach space to which the fixed point belongs, and of computing the fixed point via a provably convergent approximation scheme, while still open for 3D Ising, is completely solved in our fermionic case, at least when ε is sufficiently small.

Our results have similarities with those of Gawedzki and Kupiainen (GK) [27], with some differences. GK had fermions transforming as spinors and the model (long-range Gross-Neveu) was reflection positive. This is a minor difference and we could have considered their model, the only complication being an extra spinorial index. Their quartic interaction was weakly irrelevant rather than weakly relevant, and so they have obtained an ultraviolet fixed point,[6] while our fixed point for $\varepsilon > 0$ is in the infrared. Our proof establishes estimates on the irrelevant fixed point interactions which are of natural size suggested by perturbation theory. Finally, we establish fixed point analyticity that, as far as we know, has not been previously pointed out. Let us mention that the fixed point we construct can also be obtained by using a different, rigorous, CRG scheme, based on a tree expansion [51], which bypasses the use of the contraction mapping theorem, as well as the apriori definition of an invariant Banach space of irrelevant interactions (see appendix J).

Open questions, to be addressed in future work, include: the computation of critical exponents and their independence from the cutoff, rigorous derivation of conformal invariance and the operator product expansion (OPE) the connection between our mildly nonlocal representation of the fixed point with the local operators used in conformal field theory, the relation with analytic regularization and Wilson-Fisher ϵ-expansion, computer-assisted computation of the radius of convergence, crossover to the local symplectic fermion fixed point for $\varepsilon = \varepsilon_* = O(1)$ in $d = 3$ (Do critical exponents coincide with the Wilson-Fisher ϵ-expansion for the local symplectic fermions in such a limiting case?), etc. See section 8 for a complete list of open problems (9 pages!).

[6] One of the purposes of [27] was to construct rigorously a healthy theory at short distances from a non-renormalizable effective theory at long distances, hence their title. This was made possible by the small parameter (weakly irrelevant interaction). Unfortunately, their paper is often misunderstood as a license to search for the UV theory in terms of IR degrees of freedom even when there is no weak coupling in sight (as e.g. in the asymptotic safety program for gravity).

JHEP01(2021)026

1.3 Convergence, analyticity and non-perturbative nature of the fixed point

Some readers may feel that our result about the analyticity of the ε-expansion contradicts quantum field theory lore, and here we wish to explain why this is not the case.

There are two main reasons for the divergence of perturbation theory in quantum field theory: classical solutions (instantons) and renormalons. Since our theory is fermionic, it does not contain instantons. A related difference of fermions vs bosons is that bosons only make sense for positive quartic while fermions are defined for quartic of any sign, and indeed our fixed point coupling λ_* will be positive or negative depending on the sign of ε.

As for the renormalons and associated divergences (see e.g. reviews [52, 53]), they exist both for fermions and bosons, but only if there is running over a long range of scales. Also in our model, the full RG flow from UV to IR would not be analytic, for reasons similar to renormalons in asymptotically free theories like QCD. However, in this paper we focus exclusively on the fixed point physics, so there is no running, and we are immune to renormalons.

Let us illustrate this point by a short computation, considering for definiteness the weakly relevant quartic case (positive ε). Note that the infrared fixed point can be constructed in two *equivalent* ways. The first, which is the one we use in the rest of this paper, is to construct it as the fixed point of the single step Wilsonian RG transformation. The second, which we briefly discuss here and in appendix J.1, is to construct it dynamically, as the infrared limit of the flow of the running couplings. We will not consider the full flow from the gaussian fixed point, but a "half-flow" which starts at an intermediate scale and flows to the IR fixed point. Even such a "half-flow" is already non-analytic, as we will see.

In our model, the beta function flow equation for $\lambda(t)$ (the running quartic coupling at scale t, where $t \leqslant 0$ is the logarithm of the infrared cutoff scale) at lowest order has the following form:

$$\frac{d\lambda(t)}{dt} = -2\varepsilon\lambda(t) + c_2\lambda^2(t), \tag{1.5}$$

for a suitable positive constant c_2. The solution to (1.5) with initial condition $\lambda(0) = \lambda_0$, which we assume to be positive and smaller than $2\varepsilon/c_2$, is:

$$\lambda(t) = \frac{\lambda_0}{e^{2\varepsilon t} + \frac{c_2\lambda_0}{2\varepsilon}(1 - e^{2\varepsilon t})}. \tag{1.6}$$

The infrared fixed point is $\lambda_* = \lim_{t\to-\infty}\lambda(t) = 2\varepsilon/c_2$, which is obviously analytic in ε. At any finite t, $\lambda(t)$ is analytic in λ_0, but *non-uniformly in* t, as $|t|$ grows. This effect is clearly due to the running and to the nontrivial structure of the RG flow: small positive λ_0 eventually flow to λ_*, while small negative λ_0 flow away to large negative values of the coupling. However, $\lambda(t)$ is Borel-summable in $\lambda_0 > 0$ uniformly in t. The complete flow is more complicated than the toy model (1.5), but it retains the same qualitative features as the above illustration. In our fermionic setting, fixed point observables, such as critical exponents, are expected to be convergent power series in λ_* and, therefore, analytic in ε. Observables (e.g. correlation functions) at intermediate distance scales can be expressed as convergent power series in the *whole sequence* $\{\lambda(t)\}_{t\leqslant 0}$ (see appendix J.1 for further details on this point). Due to the non-analytic dependence of $\lambda(t)$ in the bare coupling λ_0, such observables are expected to be "just" Borel-summable in λ_0.

Finally, let us comment on the setup of massless perturbation theory, i.e. when the gaussian fixed point is perturbed by only the (weakly relevant) quartic coupling, setting mass to zero and working directly in the continuum limit. Such a setup, under the name of "conformal perturbation theory" [54–56], is often considered when perturbing non-gaussian fixed points (see e.g. [57, 58] for recent applications), but it could be used in our problem as well. It is a form of perturbative expansion in the bare coupling. At a small but fixed ε, the first n terms of the resulting perturbative expansion will be finite, where $n \sim 1/\varepsilon$, while subsequent terms have infinite coefficients (because the corresponding integrals diverge at long distances). Thus, the perturbative expansion itself is ill-defined beyond the first few terms in this framework.

Some authors, e.g. ref. [56], argued that this pathology is a possible signal of the appearance of non-analytic terms in the infrared fixed point observables (although, as [56] admits, "their actual presence is unclear at the moment"). It has to be emphasized that we are talking here about the situation when the RG flow leads to a fixed point, and only about the infrared fixed point observables, such as the critical exponents. We are not concerned with the situation when the flow leads to a massive phase, in which case the mass of the particles is indeed generically non-analytic in the bare couplings. While non-analytic terms may affect bosonic relatives of our model (see section 8.1.8), in our fermionic model we rigorously exclude them, see remark 5.1 and appendix H. Our analytic-in-ε fixed point defines the infrared theory in a fully non-perturbative way. Thereby, results based on convergent perturbation theory lead in our case to a fully non-perturbative description. The key point allowing this to happen is that in finite volume Grassmann integrals are finite dimensional. Therefore, in presence of any finite-volume cutoff, there is no room for non-analytic terms. Furthermore, uniformity in the volume of our estimates, in combination with uniqueness theorems for the limit of uniformly convergent analytic functions, imply that the absence of non-analytic terms carries over to the infinite-volume limit, see appendix H for details.

1.4 Summary

The paper is structured as follows. In section 2 we present the model and we state informally our main results. In section 3 we identify an approximate nontrivial fixed point by truncating the RG map at lowest order (explicit perturbative computations are in appendix G). The rest of the paper will be devoted to a non-perturbative proof of its existence: in section 4 we introduce the Banach space of interactions consisting of monomials in the fields with mildly nonlocal interactions, and we equip it with a suitable norm, tailored for our smooth slicing cutoffs (whose properties are in appendix A). In section 5 we show that the assumed form of the interaction is left invariant by the RG map, a fact made apparent rearranging the output via a trimming operation (more details are in appendices B and C). We show also there that the action of the RG map can be expressed as a series which is absolutely convergent in norm; this follows from a number of results described in appendix D, such as determinant bounds for simple fermionic expectations and a suitable representation of connected expectations. Absolute convergence allows to rigorously estimate the action of the RG map, and this allow us in section 6 to prove, see theorem 6.1, the existence of the fixed point, together with its independence of the slicing parameter and its analyticity in ε. This result relies on the crucial Key Lemma 6.1 and its variants, which ensures that the Banach space is invariant and the RG map is contractive. The key lemma is in a sense optimal, as it predicts a dependence on ε of the effective interactions which is exactly the one suggested by perturbation theory; this is obtained by a careful choice of constants done in the proof, presented in section 7 and appendix F. section 8 is devoted to conclusions and open problems. The fact that our convergent analysis fully reconstruct the theory and provides non-perturbative information is proved in appendix H. In appendix I we show that the fixed point can be obtained via a formal series expansion; perturbation theory is similar for boson or fermionic models but for fermions the series converges, a fact offering, see appendix J, a way to construct the RG fixed point alternative to the path via Banach space and contraction method, using instead the direct tree expansion technique. Finally in appendix K a review and comparison with previous results in bosonic theories is presented.

2 Definition of the model and formulation of the problem

Let us now discuss the model more in detail. The propagator (1.2) is defined in terms of $P(x)$, which is chosen in the form

$$P(x) = \int \frac{d^d k}{(2\pi)^d} \hat{P}(k) e^{ikx}, \qquad \hat{P}(k) = \frac{\chi(k)}{|k|^{\frac{d}{2}+\varepsilon}}. \tag{2.1}$$

The function $\chi(k)$ here is a "UV cutoff", a short-distance regulator of the model. We will choose it satisfying the following conditions (see figure 1):

$$\chi \text{ is a radial } C^\infty \text{ function}, \quad 0 \leqslant \chi(k) \leqslant 1, \quad \chi(k) = \begin{cases} 1, & (|k| \leqslant 1/2) \\ 0, & (|k| \geqslant 1), \end{cases} \tag{2.2}$$

In fact we will require something a bit stronger than $\chi \in C^\infty$, namely:

$$\chi \text{ belongs to the Gevrey class } G^s \text{ for some } s > 1. \tag{2.3}$$

This "Gevrey condition" will be defined in section 4.2, see eq. (4.14), and is not used until then. As explained there, it is needed so that the fluctuation propagator $g(x)$ (see section 2.1) decays at infinity as a stretched exponential. There are many cutoff functions satisfying both conditions (2.2) and (2.3); an explicit example is given in appendix A.1.

As a consequence of (2.1) and (2.2), $P(x)$ is uniformly bounded, and its large-x asymptotics is proportional to $1/|x|^{d/2-\varepsilon}$, as stated after (1.2).

We denote by $d\mu_P(\psi)$ the Gaussian Grassmann integration with propagator (1.2), which can be formally written as:

$$\begin{aligned} d\mu_P(\psi) &= D\psi e^{S_2(\psi)}, \\ S_2(\psi) &= \frac{1}{2} \int \frac{d^d k}{(2\pi)^d} \hat{P}(k)^{-1} \Omega_{ab} \psi_a(k) \psi_b(-k). \end{aligned} \tag{2.4}$$

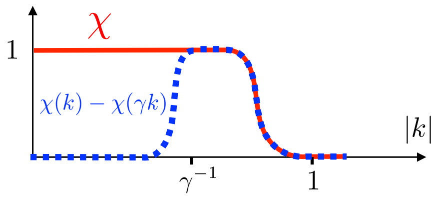


Figure 1. The function $\chi(k)$ (red curve), and the resulting function $\chi(k) - \chi(\gamma k)$, eq. (2.10) (blue dots).

[Since $\hat{P}(k)^{-1}$ is non-analytic in k^2 near $k = 0$, such an action is called "long-range".]

More precisely, $d\mu_P(\psi)$ is characterized by the expectations of an even number $2s$ of fields, via:

$$\begin{aligned}\langle\psi_{a_1}(x_1)\dots\psi_{a_{2s}}(x_{2s})\rangle &\equiv \int d\mu_P(\psi)\psi_{a_1}(x_1)\dots\psi_{a_{2s}}(x_{2s}) \\ &= \sum_\pi (-)^\pi \prod_{i=1}^{s} G_{\pi(a_{2i-1})\pi(a_{2i})}(x_{2i-1}, x_{2i}),\end{aligned} \tag{2.5}$$

where the sum is over all pairings of $2s$ fields and $(-)^\pi$ is the sign of the corresponding permutation.

The full model is defined by an "interacting" Grassmann measure

$$Z^{-1} d\mu_P(\psi) e^{H(\psi)}, \tag{2.6}$$

where the "interaction" $H(\psi)$ is a bosonic function of Grassmann fields, and $Z = \int d\mu_P(\psi)e^{H(\psi)}$ is the partition function. The simplest interaction includes only the local quadratic and quartic terms:[7]

$$H_L(\psi) = \nu \int d^dx \Omega_{ab}\psi_a\psi_b + \lambda \int d^dx (\Omega_{ab}\psi_a\psi_b)^2. \tag{2.7}$$

This interaction has $Sp(N)$ global symmetry rotating the fermion indices, as well as $O(d)$ spatial invariance. We will assume $N \geqslant 4$ so that the quartic interaction does not vanish identically.[8] We will furthermore assume

$$d \in \{1, 2, 3\}, \qquad 0 < \varepsilon < \min(2 - d/2, d/6) = d/6, \tag{2.8}$$

where the second condition guarantees that the two terms in $H_L(\psi)$ are the only $O(d) \times Sp(N)$-invariant interactions which are relevant, see section 5.3.[9]

[7]Here and below we denote the local couplings by ν, λ, rather than by ν_0, λ_0, as in (1.1). The change of notation is meant to highlight the difference between the bare couplings ν_0, λ_0, and the running ones, which will be their meaning from now on. In fact, in the following, we shall construct the interaction H corresponding to the infrared fixed point, whose local quadratic and quartic couplings correspond to the fixed point values ν_*, λ_* computed at lowest order in section 3. The notation ν, λ is used for generic values of the parameter entering the RG equations.

[8]For $N = 2$ the quartic interaction vanishes, while for $N = 4$ it is proportional to $\psi_1\psi_2\psi_3\psi_4$.

[9]The notion of relevance in our setup involving mildly nonlocal kernels will be made precise below in eq. (5.37), and it will agree with the usual rule that the interaction containing l fields and p derivatives is relevant if $l[\psi] + p < d$.

RG transformation of the model will be acting in a more general space of interactions

$$H(\psi) = H_L(\psi) + H_{\rm IRR}(\psi), \tag{2.9}$$

where $H_{\rm IRR}(\psi)$ stands for an infinite number of generally nonlocal (although mildly so) terms corresponding to irrelevant interactions. Like $H_L(\psi)$, interactions in $H_{\rm IRR}(\psi)$ will respect $Sp(N) \times O(d)$ invariance.[10]

Remark 2.1 Eq. (2.6) as written is not immediately meaningful in infinite volume, because partition function is infinite: $Z = \infty$. To give it a rigorous meaning, we should e.g. put the model in finite volume and pass to the limit. To speed up this introductory part of the paper, let us work directly in infinite volume and consider the interacting measure (2.6) in the sense of formal perturbative expansion in $H(\psi)$. In perturbation theory, the normalization factor Z^{-1} in (2.6) means that diagrams with disconnected interaction vertices should be excluded when computing expectations. In the main text we will show that infinite-volume perturbation theory is convergent (this is a general feature of fermionic models at weak coupling). The rigorous definition as a limit from finite volume is postponed to appendix H. Taking this limit will be easy once the infinite volume behavior is understood. See also remark 5.1 below.

2.1 Renormalization map

Let us fix a "rescaling parameter"[11] $\gamma \geqslant 2$. We will define the "renormalization map" which maps $H(\psi)$ to another interaction $H'(\psi)$. It will be a composition of integrating-out and dilatation.

Integrating-out consists in splitting the field ψ as $\psi = \psi_\gamma + \phi$ where ψ_γ is the "low-momentum component" of ψ, and defining the effective interaction $e^{H_{\rm eff}(\psi_\gamma)}$ by eliminating ϕ. Concretely, we split the Grassmann propagator as (see figure 1)

$$P(x) = P_\gamma(x) + g(x), \quad \widehat{P_\gamma}(k) = \frac{\chi(\gamma k)}{|k|^{\frac{d}{2}+\varepsilon}}, \quad \hat{g}(k) = \frac{\chi(k) - \chi(\gamma k)}{|k|^{\frac{d}{2}+\varepsilon}}. \tag{2.10}$$

Note that P_γ is just a rescaled version of P (see eq. (2.14)), while $g(x)$ is called "fluctuation propagator". This decomposition implies factorization of the integration measure $d\mu_P(\psi)$ as

$$d\mu_P(\psi) = d\mu_{P_\gamma}(\psi_\gamma) d\mu_g(\phi), \qquad \psi = \psi_\gamma + \phi, \tag{2.11}$$

where ψ_γ and ϕ are two independent Grassmann fields with propagators P_γ and g. As mentioned, eq. (2.3) will guarantee that $g(x)$ decays at infinity as a stretched exponential.

Correlation functions of ψ_γ with respect to the interacting measure (2.6) can equivalently be computed with respect to the measure

$$d\mu_{P_\gamma}(\psi_\gamma) e^{H_{\rm eff}(\psi_\gamma)} \tag{2.12}$$

[10]In the trimmed representation of section 4.1.1, $H_{\rm IRR}(\psi)$ will consist of $H_{2{\rm R}}, H_{4{\rm R}}, H_{6{\rm SL}}, H_{6{\rm R}}$ and H_ℓ for $\ell \geqslant 8$.

[11]Although at this point any $\gamma > 1$ would do, we assume $\gamma \geqslant 2$ from the start, as some estimates below, specifically eq. (5.43), will require that γ is separated from 1. The fixed point construction will require raising γ even further.

JHEP01(2021)026

(normalization understood) where $e^{H_{\text{eff}}(\psi_\gamma)}$ is defined by "integrating out the fluctuation field" ϕ:[12]

$$e^{H_{\text{eff}}(\psi_\gamma)} = \int d\mu_g(\phi) e^{H(\psi_\gamma+\phi)}. \tag{2.13}$$

Note that the propagator P_γ is related to P via

$$P_\gamma(x) = \gamma^{-2[\psi]} P(x/\gamma) \tag{2.14}$$

with $[\psi] = d/4 - \varepsilon/2$ as above. This motivates to consider the dilatation transformation:

$$\psi_\gamma(x) \mapsto \gamma^{-[\psi]}\psi(x/\gamma), \tag{2.15}$$

which maps the measure (2.12) to the measure $d\mu_P(\psi)e^{H'(\psi)}$ with the same gaussian factor as in (2.6) but with a different interaction:

$$H'(\psi) = H_{\text{eff}}\left[\gamma^{-[\psi]}\psi(\cdot/\gamma)\right]. \tag{2.16}$$

This formula defines the renormalization map $R = R(\varepsilon, \gamma) : H \mapsto H'$ (also called "RG transformation"). Note that R also depends on d, N, χ but this dependence will be left implicit. As a function of γ for a fixed ε, the renormalization map satisfies the semigroup property:

$$R(\varepsilon, \gamma_1)R(\varepsilon, \gamma_2) = R(\varepsilon, \gamma_1\gamma_2). \tag{2.17}$$

Our main goal will be to construct the fixed point of the RG transformation. We would like to remind the reader that although our RG transformation is obtained by integrating out the degrees of freedom with momenta between $\Lambda \sim 1$ and $\Lambda_{\text{IR}} \sim \Lambda/\gamma$, one should not think of Λ_{IR} as some sort of mass which breaks criticality of our fixed point. The correct interpretation is that we have only one RG scale, Λ, while Λ_{IR} entered the game because we find it technically convenient to consider the discrete RG transformation rather than the continuous one, such as Polchinski's equation [59]. At an intuitive level discrete RG transformation can be obtained by integrating the continuous one, and they are expected to have the same fixed points (although to make rigorous sense of the continuous RG may be nontrivial, see remark 5.4 below). In particular, it would be wrong to think that some sort of 'IR cutoff removal' has to be performed with our result to extract the fixed point physics. On the contrary, all of this physics is already contained in the fixed point H_*. E.g., the critical exponents can be obtained by linearizing the RG transformation (the same one which leads to the fixed point), near the fixed point, and computing the eigenvalues.

After this warning, the informal formulation of our main result goes as follows:
Fix $\chi, d \in \{1, 2, 3\}$, and $N \geqslant 4, N \neq 8$. For γ large enough and $\varepsilon > 0$ small enough, there exists a nontrivial interaction $H_*(\varepsilon)$ which is a fixed point of $R(\gamma, \varepsilon)$ for all γ:

$$R(\varepsilon, \gamma)[H_*(\varepsilon)] = H_*(\varepsilon). \tag{2.18}$$

[12]We will drop the ψ_γ-independent term in $H_{\text{eff}}(\psi_\gamma)$, since this constant drops out when normalizing and does not affect the expectations. As we will discuss in appendix H, this constant is finite in finite volume although it becomes infinite in the infinite-volume limit.

Moreover, $H_*(\varepsilon)$ can be extended to an analytic function of ε in a small neighborhood of the origin.

A precise statement is the content of Theorems 6.1 and 6.2, which rely on Key Lemma 6.1 and Abstract Lemma 6.2. The condition $N \neq 8$ comes from requiring a non-vanishing one-loop beta-function. The condition that γ is sufficiently large arises for the following technical reason: the fixed point $H_*(\varepsilon)$ will live in a Banach space, and only for sufficiently large γ will we be able to show that $R(\varepsilon, \gamma)$ is a bounded operator on this Banach space, so that eq. (2.18) makes sense.

The definition of the Banach space requires a suitable representation of the interactions, called *trimmed representation*, and discussed in section 4 below. In order to define it, we will distinguish, quite naturally, the local (relevant) terms from the nonlocal (irrelevant) ones. Moreover, we will rewrite the nonlocal quadratic or quartic interactions in terms of derivative fields, via the trimming operation, defined in section 5.2 below: the usefulness of a representation in terms of derivative fields is to make the irrelevance of these interactions apparent, already at the level of the linearized RG map. An additional feature of the trimmed representation is that it distinguishes a so-called "semilocal" sextic term from the fully nonlocal sextic interaction. This splitting may look strange at first sight. In our setup with a smooth cutoff in momentum space, it is needed to obtain the correct lowest-order approximation to the fixed point, which is in turn important for defining a neighborhood in the Banach space where the RG map (or, better, a suitable rewriting thereof) is contractive.

In order to better motivate it, let us explain more explicitly the structure of the splitting of the sextic term and the intuitive reason behind its definition: we'll do it in the next section, before getting to the formal definition of the trimmed representation.

3 The fixed point equation at lowest order

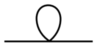
Let us go back to the lowest order fixed point equation (FPE), whose structure was anticipated in eq. (1.4), and let us discuss its derivation more carefully, in view of our choice of a smooth cutoff function. The most naive approximation one can do is to compute the FPE by neglecting all couplings but the relevant ones, ν and λ, and, assuming these couplings to be of order ε, to retain only the dominant contributions to the beta functions for ν and λ, which are of order ε and ε^2, respectively. While very natural, we would like to convince the reader that such a naive approximation leads to a wrong lowest order FPE, whose solution differs from the correct one by $O(\varepsilon)$ rather than $O(\varepsilon^2)$: the important contribution missed by this scheme is the $O(\varepsilon^2)$ contribution to the beta function for the quartic coupling λ, due to the self contraction of the "semilocal" sextic term (the tree graph contribution to the sextic interaction, of order $O(\varepsilon^2)$), see below for details.

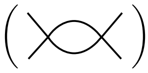
Let us start by describing the most naive approximation (the wrong one). Consider a local interaction, $H(\psi) = H_L(\psi)$, see (2.7), and integrate the fluctuation field via (2.13). After this integrating-out step the local couplings are modified as follows:

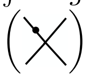
$$\nu \to \nu + \Delta\nu \equiv \nu_{\text{eff}}, \quad \lambda \to \lambda + \Delta\lambda \equiv \lambda_{\text{eff}}, \tag{3.1}$$

JHEP01(2021)026

where the leading contributions to $\Delta\nu$, $\Delta\lambda$ are given by the diagrams

$$\Delta\nu = $$  $$+\cdots$$

$$\Delta\lambda = $$  $$_{\text{loc}} + \ldots \tag{3.2}$$

Here $(\cdot)_{\text{loc}}$ stands for the local part of the nonlocal term generated by the second diagram.[13] Note that the diagram  $= O(\varepsilon^2)$ does not contribute to $\Delta\lambda$ because its local part vanishes. Indeed, the propagator g vanishes in momentum space at $k = 0$, or equivalently $\int d^d x g(x) = 0$. For this reason, ν insertions on external legs never give rise to local terms:  $_{\text{loc}} = 0$. One can easily check by inspection that, starting from $H = H_L$, there are no other contributions of $O(\varepsilon)$ to $\Delta\nu$ and of $O(\varepsilon^2)$ to $\Delta\lambda$, beyond those shown in (3.2).

We now rescale the fields as in (2.16) and find $\nu' = \gamma^{d-2[\psi]}\nu_{\text{eff}}$ and $\lambda' = \gamma^{d-4[\psi]}\lambda_{\text{eff}}$, that is, recalling (3.1),

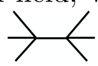
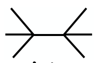
$$\nu' = \gamma^{\frac{d}{2}+\varepsilon}(\nu + I_1\lambda + \ldots), \tag{3.3}$$

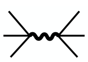
$$\lambda' = \gamma^{2\varepsilon}(\lambda + I_2^{\bowtie}\lambda^2 + \ldots), \tag{3.4}$$

where I_1 and $I_2^{\bowtie}$ are the one-loop diagrams in the two lines of (3.2), respectively. Performing Ω-tensor contractions, one finds $I_2^{\bowtie} \propto N - 8 \neq 0$, since we are assuming $N \neq 8$.[14] It is now extremely tempting to conclude that the fixed point equation for λ is

$$\lambda = \gamma^{2\varepsilon}(\lambda + I_2^{\bowtie}\lambda^2 + \cdots), \tag{3.5}$$

up to terms of $O(\varepsilon^3)$, so that the fixed point is $\lambda_* = (1-\gamma^{2\varepsilon})/I_2^{\bowtie}$ up to an error $O(\varepsilon^2)$; plugging this into the fixed point equation for ν, one would find $\nu_* = I_1\lambda_*/(1-\gamma^{d/2+\varepsilon})$ up to an error $O(\varepsilon^2)$. Even if extremely tempting, **this conclusion is wrong**!

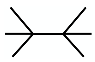
Where is the problem? The point is that neglecting the irrelevant terms, and in particular the sextic one, leads to an error of $O(\varepsilon^2)$ in the FPE for λ; such an error is comparable in size with the term $I_2^{\bowtie}\lambda^2$ that we included above: therefore, dropping blindly the irrelevant terms is not consistent even at the lowest order. To see this, notice that, by starting with a local interaction, $H = H_L$, after having integrated out the fluctuation field, we obtain an effective interaction H_{eff}, whose sextic term contains the tree diagram  $= O(\varepsilon^2)$. Therefore, in order to find an interaction H solving the fixed point equation $H' = H$ at $O(\varepsilon^2)$, we cannot avoid assuming that H contains a sextic irrelevant term with the same structure as . Let us then take $H = H_L + H_{\text{IRR}}$, with H_{IRR} containing the following sextic '$\mathfrak{X}$-term' interaction:

 $$= \Omega_{ab}\Omega_{a'b'}\Omega_{cc'}\int d^d x d^d y (\psi_a\psi_b\psi_c)(x)\mathfrak{X}(x-y)(\psi_{a'}\psi_{b'}\psi_{c'})(y), \tag{3.6}$$

[13] This operation is done in momentum space by evaluating the diagram with all external momenta set to zero. Alternatively, in position space one replaces the kernel of the nonlocal operator by its integral.

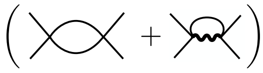
[14] See appendix G for the computations of these coefficients, where we also comment that vanishing of $I_2^{\bowtie}$ is an accident which does not reproduce at higher loops.

with the particular shown contraction of ψ indices. This term might be called 'semilocal': there are two ψ^3 groups interacting via one nonlocal kernel. The gothic $\mathfrak{X}$ is meant to remind about the shape of this diagram. Upon integrating out, the unique new contribution to $\mathfrak{X}$ comes from the tree-level diagram contracting two quartic vertices:



$$\Rightarrow \quad \Delta\mathfrak{X}(x) = -8\lambda^2 g(x). \tag{3.7}$$

At the fixed point, we thus expect $\mathfrak{X} = O(\lambda_*^2) = O(\varepsilon^2)$. On the other hand, $\mathfrak{X}$ gives a direct contribution to $\Delta\lambda$, which therefore has to be included explicitly: the equation for $\Delta\lambda$ thus has to be corrected as follows:

$\Delta\lambda =$  $_{\mathrm{loc}} + \dots$ (3.8)

Assuming that $\mathfrak{X}$ is $O(\varepsilon^2)$ and that all the other irrelevant interactions of order 6 or more in the fields are $O(\varepsilon^3)$ or smaller (while the nonlocal, irrelevant, contributions of order 2 or 4 are $O(\varepsilon^2)$), one can check by inspection that the dots in (3.8) are $O(\varepsilon^3)$. At this point some readers may be thrown out of balance: who has ever seen this second diagram? In fact Wilson and Kogut discuss it, [1], eq. (5.23) and below. They do observe that it is $O(\varepsilon^2)$ and thus would need to be included. Only if one uses a sharp cutoff, then this diagram drops out because its local part then vanishes (momenta along the wavy and curved lines do not overlap). Since we use a smooth cutoff, we have to include both diagrams. The corrected leading approximation to the FPE thus involves ν, λ and the $\mathfrak{X}$-term parametrized by $\mathfrak{X}(x)$; it takes the form

$$\nu = \gamma^{\frac{d}{2}+\varepsilon}[\nu + I_1\lambda + O(\varepsilon^2)], \tag{3.9}$$

$$\lambda = \gamma^{2\varepsilon}[\lambda + I_2^{\bowtie}\lambda^2 + (N-8)\int d^dx\,\mathfrak{X}(x)g(x) + O(\varepsilon^3)], \tag{3.10}$$

$$\mathfrak{X}(x) = \gamma^{2d-6[\psi]}[\mathfrak{X}(x\gamma) - 8\lambda^2 g(x\gamma)], \tag{3.11}$$

the factor $(N-8)$ coming from Ω-tensor contractions. This allows us to compute the fixed point couplings ν_* and λ_* at order ε, while $\mathfrak{X}_*$ will be order ε^2. The approximation is consistent: all the irrelevant terms not explicitly shown contribute to the error terms only. Note that (3.11) allows us to express the fixed point $\mathfrak{X}_*$ in terms of λ as a geometric series:

$$\mathfrak{X}_*(x) = -8\lambda^2 \sum_{n=1}^{\infty} \gamma^{(2d-6[\psi])n} g(x\gamma^n). \tag{3.12}$$

Plugging $\mathfrak{X} = \mathfrak{X}_*$ in the right side of (3.10) gives the FPE for the quartic coupling $\lambda = \gamma^{2\varepsilon}(\lambda + I_2\lambda^2)$ up to an error $O(\varepsilon^3)$, with $I_2 = I_2^{\bowtie} + I_2^{\mathfrak{X}}$ and $I_2^{\mathfrak{X}}$ the constant coming from the term $(N-8)\int d^dx\,\mathfrak{X}(x)g(x)$. We are thus led to the FPE (1.4), which we now expect to be correct at dominant order, contrary to (3.5). Interestingly, in the $\varepsilon \to 0$ limit the sum of the two diagrams $I_2 = I_2^{\bowtie} + I_2^{\mathfrak{X}} \approx \bar{I}_2 \log\gamma$, where $\bar{I}_2$ is independent of the choice of the cutoff function (appendix G). Therefore the fixed point coupling λ_* is universal at order ε. This is similar to the well-known scheme independence of the first two beta-function coefficients.

JHEP01(2021)026

In conclusion, the inclusion of the irrelevant sextic terms is crucial for computing the correct coefficients in the lowest order FPE. One expects that the inclusion of more and more irrelevant terms will produce better and better approximation to the FPE, unless non-perturbative effects come into play. Therefore, in order to compute the exact FPE, we'll assume that H belongs to a space of interactions including all possible irrelevant terms, of arbitrary high order in the fields, as discussed in the next section.

4 The Banach space of interactions

4.1 Representation of interactions by kernels

In order to conveniently represent the interaction, we use the following condensed notation for fields, their first derivatives, and products thereof:

$$\Psi_A = \begin{cases} \psi_a, & A = a, \\ \partial_\mu \psi_a, & A = (a, \mu) \end{cases}, \qquad \Psi(\mathbf{A}, \mathbf{x}) = \prod_{i=1}^{l} \Psi_{A_i}(x_i), \tag{4.1}$$

where $\mathbf{A} = (A_1, \ldots, A_l)$ and $\mathbf{x} = (x_1, \ldots, x_l)$ are finite sequences. $|\mathbf{A}|$ will denote the length of $\mathbf{A}$, and $d(\mathbf{A})$ the number of derivative fields in $\Psi(\mathbf{A}, \mathbf{x})$. See section 5.2.2 for why we allow fields with zero or one (but not more) derivatives.

An interaction $H(\psi)$ is a sum of terms with some kernels $H(\mathbf{A}, \mathbf{x})$:

$$H(\psi) = \sum_{\mathbf{A}} \int d^d\mathbf{x} H(\mathbf{A}, \mathbf{x}) \Psi(\mathbf{A}, \mathbf{x}), \tag{4.2}$$

where $d^d\mathbf{x} = d^d x_1 \ldots d^d x_l$. The kernels satisfy various obvious constraints following from the Grassmann nature of the fields and from the $Sp(N) \times O(d)$ symmetry of the model. E.g., the kernels are assumed antisymmetric.[15] The individual interaction terms being bosonic, "the number of legs" $l = |\mathbf{A}|$ must be even.

The local quadratic and quartic interactions in (2.7) correspond to δ-function kernels:

$$\nu \int d^d x \Omega_{ab} \psi_a \psi_b \leftrightarrow \nu \Omega_{ab} \delta(x_1 - x_2), \tag{4.3}$$

$$\lambda \int d^d x (\Omega_{ab} \psi_a \psi_b)^2 \leftrightarrow \tfrac{1}{3} \lambda q_{abce} \delta(x_1 - x_2) \delta(x_1 - x_3) \delta(x_1 - x_4), \tag{4.4}$$

where $q_{abce} = \Omega_{ab}\Omega_{ce} - \Omega_{ac}\Omega_{be} + \Omega_{ae}\Omega_{bc}$ is totally antisymmetric. We will represent interactions of H_{IRR} by nonlocal kernels rather than expanding them in local interactions.

We divide the kernels into groups ("couplings") H_l according to their number of legs l ($l \geqslant 2$ even):

$$H_l = \{H(\mathbf{A}, \mathbf{x})\}_{|\mathbf{A}|=l}. \tag{4.5}$$

[15] This means $H(\pi\mathbf{A}, \pi\mathbf{x}) = (-)^\pi H(\mathbf{A}, \mathbf{x})$ for any permutation acting simultaneously on $\mathbf{A}$ and $\mathbf{x}$. If not already antisymmetric, the antisymmetrized kernels $H^{\mathcal{A}}(\mathbf{A}, \mathbf{x}) = \frac{1}{l!} \sum_{\pi \in S_l} (-)^\pi H(\pi\mathbf{A}, \pi\mathbf{x})$ define the same interaction. The kernel dependence on $Sp(N)$ indices will be made out of Ω_{ab} tensors. Their dependence on x_i will be an $SO(d)$-invariant tensor built out of $(x_i - x_j)_\mu$ where μ are spatial indices contained in $\mathbf{A}$ (if any), times a function of pairwise distances $|x_i - x_j|$. Finally, $O(d)$ also contains spatial parity $\psi_a(x) \to \psi_a(-x)$, $\Psi(\mathbf{A}, \mathbf{x}) \to (-1)^{d(\mathbf{A})} \Psi(\mathbf{A}, -\mathbf{x})$. Therefore, the kernels will satisfy $H(\mathbf{A}, -\mathbf{x}) = (-1)^{d(\mathbf{A})} H(\mathbf{A}, \mathbf{x})$.

The interaction is thus represented by a coupling sequence $(H_l)_{l\geqslant 2}$. This is a *general* representation. It is useful to introduce a notation also for the restriction of H_l to the kernels with a specified number of derivatives:

$$H_{l,p} = \{H(\mathbf{A},\mathbf{x})\}_{|\mathbf{A}|=l, d(\mathbf{A})=p}. \tag{4.6}$$

We emphasize that $\mathbf{A}$'s are *sequences*: the ordering is important and terms with different orderings appear separately in (4.2). This convention leads to somewhat simpler combinatorics.

4.1.1 Trimmed representation

General representation has too much redundancy in the couplings with a small number of legs. In view of the discussion of section 3, it is convenient to assume that the interaction has a more specific structure. In particular, we want that: H_2 consists of a local term like (4.3) plus an irrelevant term schematically of the form $(\partial\psi)^2$; H_4 consists of a local term like (4.4) plus an irrelevant term schematically of the form $\psi^3\partial\psi$; H_6 consists of a semilocal term like (3.6) plus higher order terms. More precisely, we will assume that the interaction H, to be used as the input for the RG map, is written in the *trimmed* representation, which imposes the following extra requirements on kernels with $l \leqslant 6$:

1. For H_2 we require:

   i. $H_{2,0}$ should be purely local, i.e. be the δ-function kernel reproducing the local quadratic interaction $\nu\int\psi^2$ given in (4.3),

   ii. $H_{2,1} = 0$ (no kernels with one derivative).

   We will denote the nonzero parts of trimmed H_2 as

   $$H_{2\mathrm{L}} = H_{2,0}, \quad H_{2\mathrm{R}} = H_{2,2} \tag{4.7}$$

   With some abuse of notation, we will identify $H_{2\mathrm{L}}$ with the prefactor ν in front of the delta function, and similarly below for $H_{4\mathrm{L}}$ with λ and for $H_{6\mathrm{SL}}$ with $\mathfrak{X}$.

2. For H_4 we require that $H_{4,0}$ should be purely local, i.e. be the δ-function kernel reproducing the local quartic interaction $\lambda\int\psi^4$ given in (4.4). We denote the parts of trimmed H_4 as

   $$H_{4\mathrm{L}} = H_{4,0}, \quad H_{4\mathrm{R}} = \{H_{4,p}\}_{p\geqslant 1}. \tag{4.8}$$

3. For H_6 we demand that it comes split into two pieces:

   $$H_6 = H_{6\mathrm{SL}} + H_{6\mathrm{R}}, \tag{4.9}$$

   where $H_{6\mathrm{SL}}$ contains only a 'semi-local' interaction with no derivatives, of the form[16] (3.6), parametrized by a function $\mathfrak{X}(x)$. Thus we have

   $$\begin{aligned} H_{6,0} &= H_{6\mathrm{SL}} + H_{6\mathrm{R},0}, \\ H_{6,p} &= H_{6\mathrm{R},p} \qquad (p\geqslant 1). \end{aligned} \tag{4.10}$$

[16] Its kernel is the antisymmetrization of $\Omega_{a_1a_2}\Omega_{a_4a_5}\Omega_{a_3a_6}\delta(x_1-x_2)\delta(x_1-x_3)\boldsymbol{\mathfrak{X}}(x_1-x_4)\delta(x_4-x_5)\delta(x_4-x_6)$.

For the moment we do not make any specific requirement on the no-derivative part of $H_{6\mathrm{R}}$.[17]

Mnemonically, L stands for local, SL for semi-local, R for the rest. The trimmed representation thus corresponds to a coupling sequence (H_ℓ) where the index ℓ takes values from the 'trimmed list'

$$\ell \in \mathrm{TL} = \{2\mathrm{L}, 2\mathrm{R}, 4\mathrm{L}, 4\mathrm{R}, 6\mathrm{SL}, 6\mathrm{R}, 8, 10, 12 \dots\}. \tag{4.11}$$

The corresponding number of legs, an integer, will be denoted by $|\ell|$. As for the general representation, we let $H_{\ell,p}$ be the restriction of H_ℓ to the terms with p derivatives. We will always use ℓ for labels from the trimmed list (4.11) and l for integer labels: $l \geqslant 2$ even. When ℓ and l appear in the same equation, they are related by $l = |\ell|$.

Remark 4.1 The reason for requirements 1,2 is as follows. The $H_{2\mathrm{R}}$ and $H_{4\mathrm{R}}$ interactions are irrelevant due to the presence of derivatives, while $H_{2,0}, H_{2,1}, H_{4,0}$ would be relevant by the same criterion, see section 5.3. However, all these "would-be relevant" interactions with arbitrary kernels can be equivalently written as local quadratic and quartic couplings plus an irrelevant $H_{2\mathrm{R}}$ and $H_{4\mathrm{R}}$ (section 5.2). Therefore, requirements 1,2 make manifest the fact that our model has only two relevant couplings: ν and λ.

Requirement 3 originates from the fact that, as discussed in section 3, isolating the semilocal sextic term is important for deriving the correct lowest order FPE; this, in turn, will be crucial for defining the correct neighborhood on which the FPE (or, better, a suitable rewriting thereof) is contractive, see section 6.

Remark 4.2 Even if the input interaction H is in the trimmed representation, in general the interaction H_{eff}, obtained via the integrating-out step (2.13), won't be trimmed. In order to put it in trimmed form, we will need to suitably manipulate the kernels of H_{eff}, via an operation called trimming, discussed in section 5.2 below. This is one of the operations needed for proving that the space of interactions is left invariant by the action of the RG map.

4.2 Norms

The interactions in the trimmed representation form a vector space. In order to promote it to a Banach space, we need to equip it with a norm: for this purpose, we will first specify the norm in the subspace associated with $\ell \in$TL, see (4.11), and then the norm of a trimmed sequence.

4.2.1 The norm of H_ℓ

We will be measuring the size of interaction kernels by means of the weighted L_1 norm

$$\|H(\mathbf{A})\|_w = \int_{x_1=0} d^d\mathbf{x} |H(\mathbf{A}, \mathbf{x})| w(\mathbf{x}), \tag{4.12}$$

[17] Eventually, we shall impose a norm condition which will make the no-derivative part of $H_{6\mathrm{R}}$ smaller by an extra ε factor, compared with $H_{6\mathrm{SL}}$.

where $w(\mathbf{x})$ is a translationally invariant weight function. In view of translation invariance we perform the integral fixing one of the x coordinates to zero. We also let

$$\|H_l\|_w = \max_{|\mathbf{A}|=l} \|H(\mathbf{A})\|_w, \tag{4.13}$$

and similarly $\|H_\ell\|_w$ and $\|H_{\ell,p}\|_w$ are defined as the maximum of weighted norms of all kernels belonging to the corresponding trimmed coupling.[18]

By choosing $w(\mathbf{x})$ growing at infinity appropriately, we will incorporate the information about the decay of the kernels $H(\mathbf{A}, \mathbf{x})$, induced by the decay of the fluctuation propagator $g(x)$. Recall that we are requiring the Gevrey condition (2.3): $\chi \in G^s, s > 1$. Concretely, this means that derivatives of χ of arbitrary order α are uniformly bounded by

$$\max_{k\in\mathbb{R}^d} |\partial^\alpha \chi(k)| \leqslant C^{|\alpha|} |\alpha|^{|\alpha| s} \tag{4.14}$$

for some constant $C = C(\chi) > 0$. The Gevrey condition is stronger than C^∞ but weaker than real analyticity. Importantly for us, Gevrey class contains compactly supported functions. An explicit example of a cutoff functions satisfying both condition (2.2) and (2.3) is given in appendix A.1.

As usual, decay of $g(x)$ is related to the smoothness of its Fourier transform $\hat{g}(k)$, i.e. to the smoothness of $\chi(k)$. It turns out that the Gevrey condition (4.14) implies a stretched exponential decay. Namely, we have the following bound for the fluctuation propagator, as well as its first and second derivatives needed below:

$$|g(x)|, |\partial_\mu g(x)|, |\partial_\mu \partial_\nu g(x)| \leqslant M(x) \equiv C_{\chi 1} e^{-C_{\chi 2}|x/\gamma|^\sigma} \qquad (x \in \mathbb{R}^d), \tag{4.15}$$

where $\sigma = 1/s < 1$. The constants $C_{\chi 1}, C_{\chi 2}$ depend on χ but are independent of γ. See appendix A.2 for a detailed proof, while here we only give two simple remarks. First, the decay scale $x \sim \gamma$ in (4.15) is as expected from the IR momentum cutoff $\sim \gamma^{-1}$. Second, stretched exponential is the best we could hope for: exponential decay ($\sigma = 1$) would require analyticity of χ, incompatible with the compact support.

Kernels $H(\mathbf{A}, \mathbf{x})$ are expected to decay at large separation with the same rate as (4.15). We choose $w(\mathbf{x})$ growing with a similar rate. A convenient choice turns out to be

$$w(\mathbf{x}) = e^{\bar{C}(\mathrm{St}(\mathbf{x})/\gamma)^\sigma}, \tag{4.16}$$

where $\mathrm{St}(\mathbf{x})$ is the Steiner diameter of the set $\mathbf{x}$, defined [60] as the length of the shortest tree τ connecting the points in $\mathbf{x}$ (the tree may contain extra vertices as in figure 2).[19] We will fix $\bar{C} = \frac{1}{2} C_{\chi 2}$ so that $M(x)$ has a finite weighted norm:

$$\|M\|_w = \int d^d x M(x) e^{\bar{C}(|x|/\gamma)^\sigma} = \mathrm{Const}\,.\gamma^d < \infty. \tag{4.17}$$

[18] To avoid any misunderstanding, we stress that $\|H_{6\mathrm{SL}}\|_w$ and $\|H_{6\mathrm{R}}\|_w$ are two independently defined quantities.

[19] To be precise, $\mathrm{St}(\mathbf{x}) = \min_{\mathbf{x}'} \min_{\tau(\mathbf{x}\cup\mathbf{x}')} L(\tau)$, the minimum taken over all possible trees τ with vertices $\mathbf{x} \cup \mathbf{x}'$, with the tree length $L(\tau)$ defined as the sum of the edge lengths. $\mathrm{St}(\mathbf{x})$ coincides with the usual diameter if all points lie on a line (e.g. for sets of 2 points). See appendix E for an explanation of why we use the Steiner diameter.

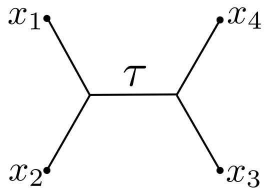


Figure 2. The optimal Steiner tree τ for this configuration of 4 points contains two extra vertices.

Finally, it will be convenient to use definition (4.13) also for the δ-function kernels of the local quadratic and quartic interactions. Since $w = 1$ when all points coincide, we have (see (4.3), (4.4))

$$\|H_{2\mathrm{L}}\|_w = |\nu|, \qquad \|H_{4\mathrm{L}}\|_w = \frac{1}{3}|\lambda|. \tag{4.18}$$

4.2.2 The norm of a trimmed sequence

The norm of an interaction H associated with the trimmed sequence (H_ℓ), $\ell \in$TL, will have the form $\|H\| = \sup_{\ell\in\mathrm{TL}} \|H_\ell\|_w/\delta_\ell$, for a sequence δ_ℓ to be fixed conveniently. In order to decide how to let δ_ℓ scale with ℓ, let us first develop an intuition about the expected size of H_ℓ at the fixed point. If we parametrize H by $(\nu, \lambda, \mathfrak{X}, u)$, with $u = (H_\ell)_{\ell\neq 2\mathrm{L},4\mathrm{L},6\mathrm{SL}} \equiv (u_\ell)_{\ell\in\{2\mathrm{R},4\mathrm{R},6\mathrm{R},8,10,\ldots\}}$, we expect that at the fixed point ν and λ are of order ε, $\mathfrak{X}$ is equal to the kernel $\mathfrak{X}_*$ in eq. (3.12), which is of order ε^2, and u_ℓ is of the order of the corresponding tree graph (i.e., the leading Feynman diagram with vertices all of type λ contributing to the interaction labelled ℓ), namely: of order ε^2 if $\ell = 2\mathrm{R}, 4\mathrm{R}$; of order ε^3 if $\ell = 6\mathrm{R}$; of order $\varepsilon^{l/2-1}$ if $\ell = l \geqslant 8$.

In the following, in order to determine the fixed point, we will fix $\mathfrak{X} = \mathfrak{X}_*$, thought of as a function of λ, see (3.12), and parametrize the fixed point interaction by the remaining coordinates, $y = (\nu, \lambda, u)$. On this subspace, we will use the following norm (depending on the parameters $\delta, A_0, A_0^{\mathrm{R}}, A_1^{\mathrm{R}}, A_2^{\mathrm{R}}, A$):

$$\|y\|_Y = \max\left\{\frac{|\nu|}{A_0\delta}, \frac{|\lambda|}{A_0\delta}, \frac{\|u_{2\mathrm{R}}\|_w}{A_0^{\mathrm{R}}\delta^2}, \frac{\|u_{4\mathrm{R}}\|_w}{A_1^{\mathrm{R}}\delta^2}, \frac{\|u_{6\mathrm{R}}\|_w}{A_2^{\mathrm{R}}\delta^3}, \sup_{l\geqslant 8} \frac{\|u_l\|_w}{A\delta^{l/2-1}}\right\}, \tag{4.19}$$

where, motivated by the intuitive discussion above, the parameter δ will be chosen to be proportional to ε. Eq. (4.19) defines the Banach space of interest.[20] Eventually, the constants $A_0, A_0^{\mathrm{R}}, A_1^{\mathrm{R}}, A_2^{\mathrm{R}}, A$ will be fixed in such a way that the action of the RG map (or better, of a suitable equivalent rewriting thereof, called F in the following, see section 6.2) on the sequence y returns a new sequence y' in the same Banach space Y. Even more, we will show that there is a neighborhood Y_0 in Y on which the fixed point map F is a contraction and, therefore, F admits a unique fixed point in Y_0. All this will be proved in sections 6 and 7 below. As a preparation to these proofs, we need to specify how the RG map explicitly acts on the space of trimmed sequences. This will be discussed in section 5.

[20] It is easy to see in particular that the space is complete with respect to the introduced norm (because weighted L_1 spaces are complete).

In particular, the result of the RG map on the trimmed sequence (H_ℓ) will be expressed in the form of a series, see eq. (5.24) below, which is *absolutely convergent* in the norm of interest, thanks to the bounds discussed in section 5.6 below.

5 The renormalization map

In this section we detail the structure of the RG map, thought of as a map from the vector space of trimmed sequences into itself. We proceed in steps: we first describe the integrating-out map (2.13), assuming the input to be a trimmed sequence (section 5.1). In general, the output of the integrating-out map is not trimmed: therefore, we explain how to make it so, via the trimming operation (section 5.2). Next, we perform the rescaling (2.16) (section 5.3). In section 5.4 we combine these three steps and derive the representation eq. (5.24), which expresses the image of the RG map as a series in the multi-indices $(\ell_i)_{i=1}^n$, $\ell_i \in$TL. Remarkably, this series turns out to be *absolutely convergent* in the relevant norms, i.e., those introduced in section 4.2 above, thanks to the norm bounds discussed in section 5.6.

5.1 Integrating-out map

Let H be an interaction associated with the trimmed sequence (H_ℓ), $\ell \in$TL, and consider the integrating-out map (2.13). For the effective interaction H_{eff} we have a well-known perturbative formula in terms of connected expectations (see appendix D.2):[21]

$$H_{\text{eff}}(\psi) = \sum_{n=1}^{\infty} \frac{1}{n!} \langle \underbrace{H(\psi+\phi); H(\psi+\phi); \ldots; H(\psi+\phi)}_{n \quad \text{times}} \rangle_c. \tag{5.1}$$

We write the interaction in the trimmed representation using the same eq. (4.2) as in the general representation. Kernels $H(\mathbf{A}, \mathbf{x})$ now come from couplings H_ℓ. The $H_{2,1}$ kernels are absent due to trimming requirements. The $H_{6,0}$ kernels are understood as a sums of $H_{6\text{SL}}$ and $H_{6\text{R},0}$ kernels, see (4.10). All the other kernels $H(\mathbf{A}, \mathbf{x})$ are associated with a unique coupling H_ℓ.

Replacing $\psi \to \psi + \phi$ in (4.2), each term gives rise to 'interaction vertices' with 'external' ψ legs and 'internal' ϕ legs. Parametrizing the external legs by a subsequence $\mathbf{B} \subset \mathbf{A}$,[22] and the internal ones by $\overline{\mathbf{B}} = \mathbf{A} \setminus \mathbf{B}$, we write:

$$H(\psi+\phi) = \sum_{\mathbf{A}} \sum_{\mathbf{B} \subset \mathbf{A}} (-)^{\#} \int d^d\mathbf{x} H(\mathbf{A}, \mathbf{x}) \Psi(\mathbf{B}, \mathbf{x}_{\mathbf{B}}) \Phi(\overline{\mathbf{B}}, \mathbf{x}_{\overline{\mathbf{B}}}), \tag{5.2}$$

where $(-)^{\#}$ is the sign, which we won't need to track, produced by reordering the fields to put all ψ's first. The $\mathbf{x}_{\mathbf{B}}$ and $\mathbf{x}_{\overline{\mathbf{B}}}$ are the corresponding restrictions of the coordinate vector $\mathbf{x}$. Substituting (5.2) into (5.1), we obtain the following formula for the kernels of

[21] In eq. (2.13) the argument of H_{eff} was ψ_γ, the low-momentum component of ψ. The momentum-range restriction turns out unimportant for working out (5.1), so we replaced ψ_γ by a generic ψ.

[22] Sequences being ordered sets, a subsequence inherits ordering from the parent sequence.

the effective interaction (see appendix B for more details):

$$H_{\text{eff}}(\mathbf{B}, \mathbf{x_B}) = \mathcal{A} \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{\substack{\mathbf{B}_1, \ldots, \mathbf{B}_n \\ \sum \mathbf{B}_i = \mathbf{B}}} \sum_{\substack{\mathbf{A}_1, \ldots, \mathbf{A}_n \\ \mathbf{A}_i \supset \mathbf{B}_i}} (-)^{\#} \int d^d \mathbf{x}_{\overline{\mathbf{B}}} \mathcal{C}\left(\mathbf{x}_{\overline{\mathbf{B}}}\right) \prod_{i=1}^{n} H(\mathbf{A}_i, \mathbf{x}_{\mathbf{A}_i}), \tag{5.3}$$

where the sum is over all ways to represent $\mathbf{B}$ as a concatenation $\mathbf{B}_1 + \ldots + \mathbf{B}_n$, and then over all ways to extend $\mathbf{B}_i$'s to $\mathbf{A}_i \supset \mathbf{B}_i$. The integration is over points $\mathbf{x}_{\overline{\mathbf{B}}}$, $\overline{\mathbf{B}} = \overline{\mathbf{B}}_1 + \ldots + \overline{\mathbf{B}}_n$, $\overline{\mathbf{B}}_i = \mathbf{A}_i \setminus \mathbf{B}_i$, while the unintegrated parts of the vectors $\mathbf{x}_{\mathbf{A}_i}$ form $\mathbf{x_B} = \mathbf{x}_{\mathbf{B}_1} + \ldots + \mathbf{x}_{\mathbf{B}_n}$, and the integration kernel $\mathcal{C}\left(\mathbf{x}_{\overline{\mathbf{B}}}\right)$ is the connected expectation:

$$\mathcal{C}\left(\mathbf{x}_{\overline{\mathbf{B}}}\right) = \left\langle \Phi(\overline{\mathbf{B}}_1, \mathbf{x}_{\overline{\mathbf{B}}_1}); \ldots ; \Phi(\overline{\mathbf{B}}_n, \mathbf{x}_{\overline{\mathbf{B}}_n}) \right\rangle_c . \tag{5.4}$$

Finally, $\mathcal{A}$ in eq. (5.3) denotes the antisymmetrization operation, see footnote 15. Note that the set of kernels $H_{\text{eff}}(\mathbf{B}, \mathbf{x_B})$ produced by this formula will not in general satisfy the trimming requirements, even if the kernel $H(\mathbf{A}, \mathbf{x})$ did. This will be dealt with in the next section.

We write eq. (5.3) more compactly and abstractly as

$$(H_{\text{eff}})_l = \sum_{(\ell_i)_1^n} \mathcal{S}_l^{\ell_1, \ldots, \ell_n}(H), \qquad H = (H_\ell)_{\ell \in \text{TL}}. \tag{5.5}$$

Each term in (5.5) is numbered by a sequence $(\ell_i)_1^n = (\ell_1, \ldots, \ell_n)$, $n \geqslant 1$, $\ell_i \in \text{TL}$, and by an even $l \geqslant 2$. The map $\mathcal{S}_l^{\ell_1, \ldots, \ell_n}(H)$ in (5.5) is the sum of all terms in (5.3) which have $|\mathbf{B}| = l$ and $H(\mathbf{A}_i, \mathbf{x}_{\mathbf{A}_i}) \in H_{\ell_i}$,while $\mathbf{B}_i$ can be arbitrary, subject to the requirements $\mathbf{B}_i \subset \mathbf{A}_i$, $\sum \mathbf{B}_i = \mathbf{B}$.[23]

Let B_l and B_ℓ be the vector spaces of couplings H_l and trimmed couplings H_ℓ, respectively, and let $B_{\text{trim}} = \bigotimes_{\ell \in \text{TL}} B_\ell$ be the vector space of trimmed coupling sequences $H = (H_\ell)_{\ell \in \text{TL}}$. The map $\mathcal{S}_l^{\ell_1, \ldots, \ell_n}$ then acts from B_{trim} to B_l and is homogeneous of degree n. Of course, this map only depends on the couplings H_{ℓ_i} whose index ℓ_i occurs in the sequence $(\ell_i)_1^n$. If index ℓ_i occurs n_i times, this map has homogeneity degree n_i in H_{ℓ_i}.

It will be also useful to define a closely related map $S_l^{\ell_1, \ldots, \ell_n}$, obtained by replacing $\prod_{i=1}^n H(\mathbf{A}_i, \mathbf{x}_{\mathbf{A}_i}) \to \prod_{i=1}^n h_i(\mathbf{A}_i, \mathbf{x}_{\mathbf{A}_i})$ in (5.3) with independent $h_i \in B_{\ell_i}$. This gives a multilinear map:

$$S_l^{\ell_1, \ldots, \ell_n} : B_{\ell_1} \times \ldots \times B_{\ell_n} \to B_l. \tag{5.6}$$

Note that this map is symmetric, i.e. invariant under the interchanges of indices ℓ_i accompanied by the simultaneous interchange of arguments. By identifying the arguments of $S_l^{\ell_1, \ldots, \ell_n}$, we get back the map $\mathcal{S}_l^{\ell_1, \ldots, \ell_n}$:

$$\mathcal{S}_l^{\ell_1, \ldots, \ell_n}(H) = S_l^{\ell_1, \ldots, \ell_n}(h_1, \ldots, h_n), \qquad h_i = H_{\ell_i} \tag{5.7}$$

It will be very important that the maps $S_l^{\ell_1, \ldots, \ell_n}$ and $\mathcal{S}_l^{\ell_1, \ldots, \ell_n}$ vanish unless $\sum |\ell_i| \geqslant l + 2(n-1)$. Otherwise, the number of fields in the connected expectation (5.4), which is $\sum |\ell_i| - l$, is not enough to get a connected Wick contraction.

[23]As noted above, $H_{6,0}$ kernels are a sum of $H_{6\text{SL}}$ and $H_{6\text{R},0}$. Picking one or the other part of the sum is understood when defining the map $\mathcal{S}_l^{\ell_1, \ldots, \ell_n}$ with $\ell_i = 6\text{SL}$ or $\ell_i = 6\text{R}$.

We will see that the defined maps are continuous with respect to the norms from section 4.2.1, see section 5.6.1 below.

While generally $S_l^{\ell_1,\ldots,\ell_n}$ acts into B_l, in two cases it can be considered to act into B_ℓ:

- For $\ell \geqslant 8$ since then $B_\ell = B_l$.
- For $n = 1$ and $l = |\ell_1|$, because in this case $S_l^{\ell_1}$ is the identity map: $S_l^\ell(H_\ell) = H_\ell$.

Remark 5.1 Eq. (5.3) is going to be the basis for all further considerations. Although derived so far by perturbation theory, in our model this equation will be non-perturbatively true. Let us discuss first why this may be expected on physical grounds. Non-perturbative validity of eq. (5.3) means that in our model the full effective action is captured by perturbation theory, with no extra contributions. Extra "instantonic" contributions are common in models involving bosonic fields, but these are absent in our model since we only have fermions. Perturbation theory may also break down if fermions form a bosonic bound state, but this typically requires a coupling that becomes large when iterating the RG, and in our model all couplings will stay weak. In the main text we will show in particular that the series in the r.h.s. is convergent provided that $H(\psi)$ is sufficiently small, and so $H_{\rm eff}(\psi)$ is well defined.[24] In appendix H we will give a rigorous justification of eq. (5.3), by first deriving this equation in finite volume and then passing to the limit, in line with remark 2.1.

5.2 Trimming

We wish to realize the renormalization map in the space of trimmed couplings. Unfortunately, as mentioned, the kernels $(H_{\rm eff})_l$ provided by eq. (5.5) are not in general trimmed. To correct this, we need an extra "trimming" step, which will find an equivalent trimmed representation of the same interaction. This step corresponds, in different notation, to the rewriting of $H_{\rm eff}$ in the equivalent form $\mathcal{L}H_{\rm eff} + \mathcal{R}H_{\rm eff}$ used in many papers in CRG, see in particular [51] and [71]. In the CRG literature, $\mathcal{L}H_{\rm eff}$ and $\mathcal{R}H_{\rm eff}$ are usually called the "local" and "regularized" parts of the effective interaction, respectively. Before describing this step in detail, let us first discuss what it means for two representations to be *equivalent*.

5.2.1 Equivalent coupling sequences

A coupling sequence (N_l) is called null if the corresponding interaction vanishes as a function of classical Grassmann fields $\psi_a(x)$ (an explicit example of what we mean is discussed

[24] We wish to draw here a parallel with fermionic models of condensed-matter physics, which have convergent perturbation theory at finite temperatures (a notable exception being fermions at finite density in 3d continuous space, unstable with respect to collapse to a point for attractive interaction). One often studied example is the Fermi-Hubbard model, whose perturbative series in the onsite repulsion U has a finite, T-dependent, radius of convergence. A simple extension of the GKL bound for connected expectations discussed in appendix D.4 (see section 3 of [28]; see also [61] and section 6 of [62]) easily implies convergence of the series, but with a far-from-optimal temperature dependence ($U \propto T^{d+1}$, with d the spatial dimension of the lattice). For realistic estimates on the convergence radius, see e.g. [63–68]. The perturbation series for Fermi-Hubbard can be evaluated to high order, and convergence checked, by Diagrammatic Monte Carlo (DiagMC) method [69, 70]. We thank Kris Van Houcke and Felix Werner for discussions about DiagMC. See also footnote 55.

JHEP01(2021)026

below). Two coupling sequences with a null difference are equivalent (represent the same interaction). Interactions are thus identified with coupling sequences modulo this equivalence relation.

The basic mechanism to produce equivalent couplings, referred to as "interpolation", starts with the Newton-Leibniz formula:

$$\psi_a(x) = \psi_a(y) + \int_0^1 ds \partial_s \psi_a(y+s(x-y)) = \psi_a(y) + \int_0^1 ds (x-y)^\mu \partial_\mu \psi_a(y+s(x-y)). \quad (5.8)$$

We can also integrate the r.h.s. in y against some function $f(x,y)$ of unit total integral: $\int dy f(x,y) = 1$. We get a family of "interpolation identities" expressing $\psi^a(x)$ as a weighted linear combination of $\psi^a(y)$ and $\partial_\mu \psi_a(y)$:

$$\psi_a(x) = \int dy \left[f(x,y)\psi_a(y) + f^\mu(x,y)\partial_\mu \psi_a(y) \right], \quad (5.9)$$

where $f^\mu(x,y)$ can be expressed in terms of f.[25]

Now take a single interaction term $\int d\mathbf{x} H(\mathbf{A},\mathbf{x})\Psi(\mathbf{A},\mathbf{x})$ corresponding to some $\mathbf{A}$ with at least one field not differentiated, e.g. the first one: $A_1 = a_1$. Pick a function $f(x,y)$ and replace $\psi_{a_1}(x_1)$ inside this interaction term via the identity (5.9). This generates an equivalent representation of the same interaction of the form

$$\sum_{\mathbf{B}} \int d^d\mathbf{x} \tilde{H}(\mathbf{B},\mathbf{x})\Psi(\mathbf{B},\mathbf{x}), \quad (5.10)$$

where the sum has $d+1$ terms: either $\mathbf{B} = \mathbf{A}$ or it is obtained from $\mathbf{A}$ replacing $A_1 \to (a_1, \mu)$, $\mu = 1, \dots, d$. The corresponding kernels $\tilde{H}(\mathbf{B},\mathbf{x})$ are obtained integrating $H(\mathbf{A},\mathbf{x})$ against f and f^μ. We can also apply this procedure to multiple interaction terms in the original interaction $H(\psi)$, summing up the new kernels $\tilde{H}(\mathbf{B},\mathbf{x})$ to the kernels $H(\mathbf{B},\mathbf{x})$ to which the transformation has not been applied. The resulting total interaction, which we call $\tilde{H}(\psi)$, is equivalent to $H(\psi)$. The difference of coupling sequences (H_l) and $(\tilde{H}_l)$ is null.

We stress that the Newton-Leibniz formula and the interpolation identities will be applied only to those ψ's in the interaction terms which do not carry any derivatives. Then, all the produced terms contain ψ's with at most one derivative. This explains why in (4.1) we allowed fields with zero or one (but not more) derivatives.

5.2.2 Trimming map

We now explain the trimming map, which maps the sequence $(H_{\rm eff})_l$ in (5.5) to an equivalent sequence of trimmed couplings. Of course, the restriction of the sequence $(H_{\rm eff})_l$ to $l \geqslant 8$ is already trimmed (see the end of section 5.1), so we only need to do something for $l \leqslant 6$.

For $l = 6$, we define

$$(H_{\rm eff})_{6\rm SL} = H_{6\rm SL} + S_6^{4\rm L,4L}(H_{4\rm L}, H_{4\rm L}), \quad (5.11)$$

[25]Explicitly $f^\mu(x,y) = \int_0^1 \frac{ds}{(1-s)^d} [(x-z)^\mu f(x,z)]_{z=(y-sx)/(1-s)}$. This is finite if $f(x,y)$ decreases sufficiently fast at large y.

i.e. the two terms in the $(H_{\rm eff})_6$ series which manifestly have the form (3.6). We define $(H_{\rm eff})_{6\rm R}$ as the sum of all the other terms in the $(H_{\rm eff})_6$ series.

For $l = 2, 4$ trimming will involve "localization" and "interpolation". "Localization" extracts the local parts of $(H_{\rm eff})_{2,0}$ and $(H_{\rm eff})_{4,0}$ (see the example below):

$$(H_{\rm eff})_{2\rm L} = T^{2,0}_{2\rm L}(H_{\rm eff})_{2,0}, \qquad (H_{\rm eff})_{4\rm L} = T^{4,0}_{4\rm L}(H_{\rm eff})_{4,0}. \tag{5.12}$$

"Interpolation" rearranges the other components setting to zero the parts of $(H_{\rm eff})_{2\rm R,p}$ and $(H_{\rm eff})_{4\rm R,p}$ in agreement with the trimming requirements, and making sure that the resulting coupling sequence is equivalent as in section 5.2.1. This operation will have the following structure:

$$(H_{\rm eff})_{4\rm R,p} = \begin{cases} 0 & \text{if } p=0 \\ (H_{\rm eff})_{4,1} + T^{4,0}_{4\rm R}(H_{\rm eff})_{4,0} & \text{if } p=1 \\ (H_{\rm eff})_{4,p} & \text{if } p>1, \end{cases} \tag{5.13}$$

$$(H_{\rm eff})_{2\rm R,p} = \begin{cases} 0 & \text{if } p=0,1 \\ (H_{\rm eff})_{2,2} + T^{2,1}_{2\rm R}(H_{\rm eff})_{2,1} + T^{2,0}_{2\rm R}(H_{\rm eff})_{2,0} & \text{if } p=2. \end{cases} \tag{5.14}$$

This can be also written succinctly as

$$(H_{\rm eff})_{4\rm R} = T^4_{4\rm R}(H_{\rm eff})_4, \qquad (H_{\rm eff})_{2\rm R} = T^2_{2\rm R}(H_{\rm eff})_2, \tag{5.15}$$

where (5.13), (5.14) define components of $T^4_{4\rm R}, T^2_{2\rm R}$ in subspaces with a definite number of derivatives.

Consider $l = 4$ as an example. The coupling $(H_{\rm eff})_{4,0}$ corresponds to an interaction the form

$$\Omega_{ab}\Omega_{ce}\int d^d\mathbf{x}F(\mathbf{x})\psi_a(x_1)\psi_b(x_2)\psi_c(x_3)\psi_e(x_4) \tag{5.16}$$

(plus two other terms with $\Omega_{ac}\Omega_{be}$ and $\Omega_{ae}\Omega_{bc}$). Substitute into this an interpolation identity

$$\psi_a(x_1)\psi_b(x_2)\psi_c(x_3)\psi_e(x_4) = (\psi_a\psi_b\psi_c\psi_e)(x_1) + \psi_a(x_1)\int_0^1 dt\partial_t[\psi_b(x_2^t)\psi_c(x_3^t)\psi_e(x_4^t)], \tag{5.17}$$

where $x_i^t = x_1 + t(x_i - x_1)$. The first term gives a local quartic interaction with

$$\lambda = \int_{x_1=0} d^d\mathbf{x}F(\mathbf{x}), \tag{5.18}$$

which defines $T^{4,0}_{4\rm L}$ in (5.12).[26] The second term in (5.17) gives a sum of interactions where one of ψ_b, ψ_c, ψ_e is differentiated: this defines $T^{4,0}_{4\rm R}(H_{\rm eff})_{4,0}$.

The $l = 2$ maps $T^{2,0}_{2\rm L}, T^{2,1}_{2\rm R}, T^{2,0}_{2\rm R}$ are defined analogously. For $T^{2,0}_{2\rm R}$ one needs to apply interpolation twice, to get from a term with no derivatives to a term where both fields carry derivatives. See appendix C for the full construction of these maps, and for the analysis of how they behave with respect to the norms measuring the size of interaction kernels.

[26]This equation can be equivalently written in momentum space as $\lambda = \hat{F}(0,0,0,0)$, i.e. evaluating the kernel with all external momenta set to zero.

Remark 5.2 Note that our trimming map T is just one of infinitely many possible trimming maps, corresponding to different choices of interpolation identities. E.g. instead of (5.17) we could have used

$$\begin{aligned} &\psi_a(x_1)\psi_b(x_2)\psi_c(x_3)\psi_e(x_4) = \\ &\psi_a(x_1)\left[\psi_b(x_1) + \int_0^1 dt\, \frac{d\psi_b}{dt}(x_2^t)\right]\left[\psi_c(x_1) + \int_0^1 ds \frac{d\psi_c}{ds}(x_3^s)]\right]\left[\psi_e(x_1) + \int_0^1 du\, \frac{d\psi_e}{du}(x_4^u)\right]. \end{aligned} \tag{5.19}$$

Such alternative trimming maps $\tilde{T}$ differ from T by a map which is null (gives a null sequence of couplings when applied to any interactions). In our construction we will use T, but any other trimming map satisfying the same norm bounds (see appendix C) would work equally well.

5.3 Dilatation

After having integrated out the fluctuation field and rearranged the result so that it is equivalently rewritten in trimmed form, we rescale the fields, see (2.15)–(2.16). We call this rescaling step *dilatation*, and denote it by D. Note that D preserves the trimmed representation. The action on the kernels is:

$$D : H_{\ell,p}(\mathbf{x}) \mapsto \gamma^{-D_l - p}\gamma^{d(l-1)} H_{\ell,p}(\gamma\mathbf{x}), \tag{5.20}$$

where we recall that $l = |\ell|$, p denotes the number of derivatives in the interaction term, and we denoted

$$D_l = l[\psi] - d = l(d/4 - \varepsilon/2) - d. \tag{5.21}$$

For the special cases $\ell \in \{2\text{L}, 4\text{L}, 6SL\}$ eq. (5.20) becomes:

$$\nu \mapsto \gamma^{-D_2}\nu = \gamma^{\frac{d}{2}+\varepsilon}\nu, \qquad \lambda \mapsto \gamma^{-D_4}\lambda = \gamma^{2\varepsilon}\lambda, \qquad \mathfrak{X}(x) \mapsto \gamma^{-D_6}\gamma^d\mathfrak{X}(\gamma x). \tag{5.22}$$

In terms of the norms of section 4.2.1, the irrelevance condition for $H_{\ell,p}$ will be $D_l + p > 0$, see eq. (5.37) below. As stated in eq. (2.8) we are assuming $d \in \{1,2,3\}$ and $0 < \varepsilon < d/6$. Under these conditions it's easy to check that $D_2, D_4 < 0$,so that ν, λ are relevant, while

$$D_2 + 2, D_4 + 1 > 0, \qquad D_l \geqslant D_6 > 0 \quad (l \geqslant 6). \tag{5.23}$$

so that $H_{2\text{R}}, H_{4\text{R}}, H_{6\text{SL}}, H_{6\text{R}}$ and H_ℓ ($\ell \geqslant 8$) are irrelevant. These interactions comprise H_{IRR} in (2.9).

5.4 Renormalization map in the trimmed representation

The renormalization map R is obtained composing the three operations: integrating-out, then trimming, then dilatation. It is the map defined in section 2.1 but now written in a specific set of coordinates (the trimmed representation). Summarizing sections 5.1, 5.2.2, 5.3, we represent $R = R(\varepsilon, \gamma)$ follows: if $H \in B_{\text{trim}}$ (the vector space of sequences of trimmed couplings), then $R : H \mapsto H' \in B_{\text{trim}}$, with

$$H'_\ell = \sum_{(\ell_i)_1^n} \mathcal{R}_\ell^{\ell_1,\dots,\ell_n}(H), \tag{5.24}$$

where $\mathcal{R}_\ell^{\ell_1,\dots,\ell_n}$ is a homogeneous map of degree n obtained by identifying the arguments in a multilinear map $R_\ell^{\ell_1,\dots,\ell_n}$:

$$\mathcal{R}_\ell^{\ell_1,\dots,\ell_n}(H) = R_\ell^{\ell_1,\dots,\ell_n}(h_1,\dots,h_n), \qquad h_i = H_{\ell_i}. \tag{5.25}$$

This multilinear map can be written explicitly as follows. For $(n;(\ell_1,\cdots,\ell_n)) = (1;\ell)$ we have

$$R_\ell^\ell = D, \tag{5.26}$$

since in this case $S_l^\ell = \mathbb{1}$ and trimming is not needed. In all the other cases $(n;(\ell_1,\cdots,\ell_n)) \neq (1;\ell)$, recalling that $l = |\ell|$, we have

$$\begin{aligned}
R_\ell^{\ell_1,\dots,\ell_n} &= D\begin{cases} S_l^{\ell_1,\dots,\ell_n} & \ell \geqslant 8 \\ T_\ell^l S_l^{\ell_1,\dots,\ell_n} & \ell \in \{2\mathrm{L},2\mathrm{R},4\mathrm{L},4\mathrm{R}\}\end{cases}\\
R_{6\mathrm{SL}}^{\ell_1,\dots\ell_n} &= D\begin{cases} S_6^{4\mathrm{L},4\mathrm{L}} & (\ell_i)_1^n = (4\mathrm{L},4\mathrm{L}) \\ 0 & \text{otherwise}\end{cases}\\
R_{6\mathrm{R}}^{\ell_1,\dots\ell_n} &= D\begin{cases} S_6^{\ell_1,\dots,\ell_n} & (\ell_i)_1^n \neq (6\mathrm{SL}),(4\mathrm{L},4\mathrm{L}) \\ 0 & \text{otherwise}\end{cases}
\end{aligned} \tag{5.27}$$

where $T_\ell^l : B_l \to B_\ell$ is the trimming map whose various components are defined by equations (5.12), (5.13), (5.14), and see eq. (5.11) for $\ell \in \{6\mathrm{SL}, 6\mathrm{R}\}$.

Just as $S_\ell^{\ell_1,\dots,\ell_n}$, the map $R_\ell^{\ell_1,\dots,\ell_n}$ is symmetric (invariant under the interchanges of indices ℓ_i accompanied by the simultaneous interchange of arguments), and it vanishes unless $\sum_i |\ell_i| \geqslant l + 2(n-1)$.

5.5 Fixed point equation

The fixed point equation (FPE) that we will study is

$$(H'_\ell) = (H_\ell). \tag{5.28}$$

with (H'_ℓ) given by (5.24). If we distinguish the components $\ell = 2\mathrm{L}, 4\mathrm{L}, 6\mathrm{SL}$ from the other couplings, denoted $u = (H_\ell)_{\ell\neq 2\mathrm{L},4\mathrm{L},6\mathrm{SL}} \equiv (u_\ell)_{\ell\in\{2\mathrm{R},4\mathrm{R},6\mathrm{R},8,10,\dots\}}$, it reads:

$$\begin{aligned}
\nu &= \gamma^{\frac{d}{2}+\varepsilon}\nu + R_{2\mathrm{L}}^{4\mathrm{L}}(\lambda) + \sum_{(\ell_i)_1^n\neq(2\mathrm{L}),(4\mathrm{L})} R_{2\mathrm{L}}^{\ell_1,\dots,\ell_n}(H_{\ell_1},\dots,H_{\ell_n}),\\
\lambda &= \gamma^{2\varepsilon}\lambda + R_{4\mathrm{L}}^{4\mathrm{L},4\mathrm{L}}(\lambda,\lambda) + R_{4\mathrm{L}}^{6SL}(\mathfrak{X}) + \sum_{(\ell_i)_1^n\neq(4\mathrm{L}),(4\mathrm{L},4\mathrm{L}),(6\mathrm{SL})} R_{4\mathrm{L}}^{\ell_1,\dots,\ell_n}(H_{\ell_1},\dots,H_{\ell_n}),\\
\mathfrak{X}(x) &= R_{6\mathrm{SL}}^{6\mathrm{SL}}(\mathfrak{X}) + R_{6\,\mathrm{SL}}^{4\mathrm{L},4\mathrm{L}}(\lambda,\lambda) = \gamma^{2d-6[\psi]}[\mathfrak{X}(x\gamma) - 8\lambda^2 g(x\gamma)],\\
u_\ell &= \sum_{(\ell_i)_1^n} R_\ell^{\ell_1,\dots,\ell_n}(H_{\ell_1},\dots,H_{\ell_n}), \quad \text{if} \quad \ell \neq 2\mathrm{L},4\mathrm{L},6\mathrm{SL}.
\end{aligned} \tag{5.29}$$

We already observed that, given λ, the FPE for $\mathfrak{X}$ is solved exactly by $\mathfrak{X} = \mathfrak{X}_*$, with $\mathfrak{X}_*$ as in (3.12). Substituting $\mathfrak{X}(x) = \mathfrak{X}_*(x)$ in the remaining equations, the variable $\mathfrak{X}$ is

JHEP01(2021)026

eliminated.[27] Denoting $y = (\nu, \lambda, u)$, we are left with the fixed point equation $y = R(y)$, or in components:

$$\begin{aligned} \nu &= \gamma^{\frac{d}{2}+\varepsilon}(\nu + I_1\lambda) + e_\nu^{(0)}(y), \\ \lambda &= \gamma^{2\varepsilon}(\lambda + I_2\lambda^2) + e_\lambda^{(0)}(y), \\ u &= e_u(y), \end{aligned} \tag{5.30}$$

with $e_\nu^{(0)}, e_\lambda^{(0)}, e_u$ defined via the infinite sums in the right sides of (5.29), and I_1, I_2 defined by

$$R_{2\text{L}}^{4\text{L}}(\lambda) = \gamma^{\frac{d}{2}+\varepsilon} I_1 \lambda, \qquad R_{4\text{L}}^{4\text{L},4\text{L}}(\lambda,\lambda) + R_{4\text{L}}^{6SL}(\mathfrak{X}_*) = \gamma^{2\varepsilon} I_2 \lambda^2. \tag{5.31}$$

These coefficients I_1, I_2 are the same as in section 3. They are given by one-loop Feynman integrals evaluated in appendix G.

Moving the l.h.s. into the r.h.s. and rescaling, we rewrite the system (5.30) as

$$f(y) = 0, \qquad f(y) := \begin{pmatrix} \nu + a\lambda + e_\nu(y) \\ \varepsilon\lambda + b\lambda^2 + e_\lambda(y) \\ u - e_u(y) \end{pmatrix}, \tag{5.32}$$

where

$$\begin{aligned} (a, e_\nu) &= \frac{1}{1-\gamma^{-d/2-\varepsilon}} (I_1, \gamma^{-\frac{d}{2}-\varepsilon} e_\nu^{(0)}), \\ (b, e_\lambda) &= \frac{\varepsilon}{1-\gamma^{-2\varepsilon}} (I_2, \gamma^{-2\varepsilon} e_\lambda^{(0)}). \end{aligned} \tag{5.33}$$

Eq. (5.30) or its equivalent eq. (5.32) are the main equations that we will be solving. Of course, part of the problem is to show that these equations make sense: that is, we need to prove that the infinite sums entering the definitions of $e_\nu^{(0)}, e_\lambda^{(0)}, e_u$ are convergent. We will actually show that, if $y = (\nu, \lambda, u)$ has bounded norm (4.19), say $\|y\|_Y \leqslant 1$, with δ sufficiently small, then the sums defining $e_\nu^{(0)}, e_\lambda^{(0)}, e_u$ are *absolutely* convergent and e_u is contractive. This will be proved in sections 6 and 7 below. In preparation to this, in the next subsection we state the norm bounds satisfied by the multilinear operators $R_\ell^{\ell_1,\ldots,\ell_n}$, which will be central for our proof of convergence.

Remark 5.3 From a more general viewpoint, a fixed point is a sequence of couplings (H_ℓ) such that $(H'_\ell) = R(\varepsilon, \gamma)[(H_\ell)]$ given by (5.24) describes the same *interaction* as (H_ℓ). This will be the case if $(H'_\ell) = (H_\ell)$, as stated in (5.28), or, more generally, if the two sequences differ by a null sequence of couplings (see section 5.2.1). In this sense, the FPE (5.28) discussed above is not the most general we could (and should) consider: the general FPE to be considered reads $(H'_\ell) = (H_\ell) + (N_\ell)$, with (N_ℓ) a null sequence. In this paper, for simplicity, we focus only on the restricted FPE (5.28), and we will show that it has a non-trivial, non-null, solution,[28] which is unique in some neighborhood. The same methods of

[27]Note that although in this paper we take advantage of this possibility, in principle we could have treated $\mathfrak{X}$ on par with all the other irrelevant couplings. The RG map would end up contractive also in the $\mathfrak{X}$ direction, and RG iterations would converge to the same solution $\mathfrak{X} = \mathfrak{X}_*$.

[28]The fixed points we will construct will have nonzero λ and ν, and will therefore be nontrivial. **Lemma.** *Any trimmed coupling sequence with nonzero ν and/or λ is not null.* Proof is left as an exercise.

proof would allow us to show that, for each sufficiently small (N_ℓ), the general FPE has a unique solution, which differs from the one with $N_\ell \equiv 0$ by a null sequence (N'_ℓ). In this sense, we expect that there is a unique interaction (equivalent class of couplings) solving the general FPE. This remains to be shown in full detail, but we prefer not to present this additional proof here, in order not to overwhelm the presentation.

Remark 5.4 Recall that we are considering renormalization maps $R = R(\varepsilon, \gamma)$ with rescaling factor $\gamma \geqslant 2$, in particular γ is separated from 1. Such RG transformations are called "finite" or "discrete". Eq. (5.29) thus sets to zero the "beta-functions" expressing the change of the interaction under a finite RG transformation. In theoretical physics, it is more common to take the limit $\gamma \to 1$ and define an "infinitesimal" or "continuous" RG transformation formally given by the derivative $(d/d\gamma)R_\gamma$ at $\gamma = 1$. At a formal level the continuous RG equation (Polchinski's equation [59]) has fewer terms and looks much simpler than the discrete RG. However, so far it has not been possible to take advantage of this formal simplicity in rigorous constructions of RG fixed points. The problem is to show that solutions to Polchinski's equation have sufficiently good boundedness properties in a Banach space of interactions, and it is not known how to do this without dealing with the finite RG at the intermediate steps of the argument, which brings back the complexity. This problem is open even in fermionic theories.[29]

5.6 Norm bounds

As anticipated in the previous subsection, we now state the norm bounds satisfied by the multilinear operators $R_\ell^{\ell_1,\ldots,\ell_n}$ entering the definitions of $e_\nu^{(0)}, e_\lambda^{(0)}, e_u$. In the case $(n; (\ell_1, \cdots, \ell_n)) = (1; \ell)$, in which R_ℓ^ℓ is defined as in (5.26), we have

$$\|R_\ell^\ell(H_\ell)\|_w \leqslant \begin{cases} \gamma^{-D_2-2}\|H_{2\mathrm{R}}\|_w & \text{if } \ell = 2\mathrm{R}, \\ \gamma^{-D_4-1}\|H_{4\mathrm{R}}\|_w & \text{if } \ell = 4\mathrm{R}, \\ \gamma^{-D_l}\|H_\ell\|_w & \text{if } l = |\ell| \geqslant 6, \end{cases} \tag{5.34}$$

while $|R_{2\mathrm{L}}^{2\mathrm{L}}(\nu)| = \gamma^{-D_2}|\nu|$ and $|R_{4\mathrm{L}}^{4\mathrm{L}}(\lambda)| = \gamma^{-D_4}|\lambda|$. In all the other cases $(n; (\ell_1, \cdots, \ell_n)) \neq (1; \ell)$, in which $R_\ell^{\ell_1,\ldots,\ell_n}$ is defined as in (5.27), we have

$$\|R_\ell^{\ell_1,\ldots,\ell_n}(h_1, \ldots, h_n)\|_w \leqslant \gamma^{-D_l}\rho_l(h_1, \ldots, h_n), \qquad h_i \in B_{\ell_i}, \tag{5.35}$$

$$\rho_l(h_1, \ldots, h_n) := \begin{cases} C_\gamma^{n-1}\prod_{i=1}^n C_0^{|\ell_i|}\|h_i\|_w & \text{if } \quad \sum_i |\ell_i| \geqslant l + 2(n-1) \\ 0 & \text{otherwise} \end{cases}, \tag{5.36}$$

where, as usual, $l = |\ell|$, and, in (5.36), C_γ, C_0 are constants independent of l, n, ℓ_i. In addition, C_0 does not depend on γ, while C_γ does.

[29] See [72] for some global solvability results for Polchinski's equation in bosonic theories with bounded interactions. Ref. [73] attempted to prove local solvability for fermionic theories but their argument has a gap, see [74]. See also an interesting discussion in the conclusions of [75]. Ref. [76] considered continuous RG in a fermionic theory, although that construction was not fully based on continuous RG: they define the effective action via a convergent tree expansion (morally equivalent to using a finite RG), then verify that the continuous RG equations hold when applied to this effective action.

The proof of (5.34) readily follows from the definition of R^ℓ_ℓ, see (5.26), and from the fact that, using the definition of D, see (5.20), and of weighted norm, see section 4.2.1, we have:

$$\|DH_{\ell,p}\|_w = \gamma^{-D_l-p}\|H_{\ell,p}\|_{w(\cdot/\gamma)} \leqslant \gamma^{-D_l-p}\|H_{\ell,p}\|_w. \tag{5.37}$$

Besides proving (5.34), this justifies the rule stated in section 5.3 that the terms with $D_l + p > 0$ are irrelevant. Since $D_l + p = l[\psi] + p$, this rule turns out the same as for the local interactions (see footnote 9).

The proof of (5.35) is more subtle, see the next two subsections, 5.6.1 and 5.6.2.

5.6.1 Bounds for $S_l^{\ell_1,\dots,\ell_n}$

Recall that, if $(n;(\ell_1,\cdots,\ell_n)) \neq (1;\ell)$, then $R_\ell^{\ell_1,\dots,\ell_n}$ is defined in terms of $S_l^{\ell_1,\dots,\ell_n}$ via (5.27). Therefore, in order to prove (5.35), we first need a bound on $S_l^{\ell_1,\dots,\ell_n}$. This is similar to (5.27), with the important difference that there is no scaling factor γ^{-D_l} in the right side:

$$\|S_l^{\ell_1,\dots,\ell_n}(h_1,\dots,h_n)\|_w \leqslant \rho_l(h_1,\dots,h_n), \qquad h_i \in B_{\ell_i}, \tag{5.38}$$

with the same ρ_l as in (5.36) (with, possibly, a different constant C_0). For the full proof of (5.38) see appendix E. Here are the main ideas: from its definition, the map $S_l^{\ell_1,\dots,\ell_n}$ is an integral operator whose kernel is the connected expectation $\mathcal{C}\left(\mathbf{x}_{\overline{\mathbf{B}}}\right)$ (more precisely, it is a sum of $O\left(\text{const}^{\sum l_i}\right)$ integral operators corresponding to different choices of $\mathbf{B}_i$ and $\mathbf{A}_i$). The fermionic connected expectation $\mathcal{C}\left(\mathbf{x}_{\overline{\mathbf{B}}}\right)$ satisfies a crucial bound due to Gawedzki-Kupiainen-Lesniewski (appendix D.4):

$$|\mathcal{C}\left(\mathbf{x}_{\overline{\mathbf{B}}}\right)| = \left|\left\langle \Phi(\overline{\mathbf{B}}_1,\mathbf{x}_{\overline{\mathbf{B}}_1});\dots;\Phi(\overline{\mathbf{B}}_n,\mathbf{x}_{\overline{\mathbf{B}}_n})\right\rangle_c\right| \leqslant C^s \sum_{\mathcal{T}} \prod_{(xx')\in\mathcal{T}} M(x-x'), \tag{5.39}$$

where the sum is over all "anchored trees $\mathcal{T}$ on n groups of points $\mathbf{x}_{\overline{\mathbf{B}}_i}$". These are graphs which become connected trees when each group of points $\mathbf{x}_{\overline{\mathbf{B}}_i}$ is collapsed to a point. There is at least one anchored tree within each connected Wick contraction, and bounding each propagator along the anchored tree by (4.15) we get the product in (5.39). The contribution of remaining $s = \frac{1}{2}(\sum|\overline{\mathbf{B}}_i| - 2(n-1))$ propagators is bounded by C^s. This explains the general structure of (5.39), but the full proof is rather more subtle. The sum of connected graphs defining the connected expectation has to be rewritten as a sum over anchored trees without double counting. For each anchored tree, we then have to sum over the remaining propagator choices, and this whole sum with factorially many terms has to be bounded by C^s. This turns out possible due to fermionic cancelations.

The number of anchored trees is $\leqslant n!4^{\sum|\overline{\mathbf{B}}_i|}$ (appendix D.5), which by the way is much smaller than the total number of connected graphs. This $n!$ cancels with $1/n!$ in (5.3), leaving only exponential factors. When evaluating the weighted norm, the product of M's in (5.39) gives the factor C_γ^{n-1} with $C_\gamma = \|M\|_w = O(\gamma^d)$ by (4.17). This finishes our brief exposition of (5.38); see appendix E for the details.

By (5.38), the multilinear map $S_l^{\ell_1,\dots,\ell_n}$ is continuous. The homogeneous map $\mathcal{S}_l^{\ell_1,\dots,\ell_n}$ related to $S_l^{\ell_1,\dots,\ell_n}$ by identifying some arguments, eq. (5.7), is also continuous.

JHEP01(2021)026

We will also need Frechet derivatives of these maps.[30] Since $S_l^{\ell_1,\dots,\ell_n}$ is a multilinear map, it is Frechet-differentiable and its Frechet derivative in each argument coincides with the map itself. The homogeneous map $\mathcal{S}_l^{\ell_1,\dots,\ell_n}$ is also Frechet-differentiable. The derivative $\nabla\mathcal{S}_l^{\ell_1,\dots,\ell_n}(H)$ is, for a fixed H, a linear operator from $B_{\rm trim}$ to B_l. Using eq. (5.7), the value of this operator on $\delta H \in B_{\rm trim}$ is:

$$[\nabla_H \mathcal{S}_l^{\ell_1,\dots,\ell_n}(H)]\delta H = \sum_{i=1}^{n} S_l^{\ell_1,\dots,\ell_n}(H_{\ell_1},\dots,\delta H_{\ell_i},\dots,H_{\ell_n}). \tag{5.40}$$

Estimating each term in the r.h.s. via (5.38) we get a bound:

$$\|[\nabla_H \mathcal{S}_l^{\ell_1,\dots,\ell_n}(H)]\delta H\|_w \leqslant \sum_{i=1}^{n} \rho_l(H_{\ell_1},\dots,\delta H_{\ell_i},\dots,H_{\ell_n}). \tag{5.41}$$

5.6.2 Bounds for $R_\ell^{\ell_1,\dots,\ell_n}$

From the definition (5.27), we have that $R_\ell^{\ell_1,\dots,\ell_n}$ is related to $S_l^{\ell_1,\dots,\ell_n}$ via the dilatation operator, which we already bounded in (5.37), and via the trimming operator T introduced in section 5.2.2, whose components we still need to bound. The easiest components to bound are the localization maps $T^{2,0}_{\rm 2L}$ and $T^{4,0}_{\rm 4L}$, which map kernels to local kernels and do not increase the norm (see (C.15)):

$$\|(H_{\rm eff})_{\rm 2L}\|_w \leqslant \|(H_{\rm eff})_{2,0}\|_w, \qquad \|(H_{\rm eff})_{\rm 4L}\|_w \leqslant \|(H_{\rm eff})_{4,0}\|_w. \tag{5.42}$$

In turn, the interpolation maps satisfy the bounds (see appendix C):

$$\begin{aligned} \|T^{4,0}_{\rm 4R}(H_{\rm eff})_{4,0}\|_{w(\cdot/\gamma)} &\leqslant C_{\rm R}\gamma\|(H_{\rm eff})_{4,0}\|_w,\\ \|T^{2,1}_{\rm 2R}(H_{\rm eff})_{2,1}\|_{w(\cdot/\gamma)} &\leqslant C_{\rm R}\gamma\|(H_{\rm eff})_{2,1}\|_w,\\ \|T^{2,0}_{\rm 2R}(H_{\rm eff})_{2,0}\|_{w(\cdot/\gamma)} &\leqslant C_{\rm R}\gamma^2\|(H_{\rm eff})_{2,0}\|_w, \end{aligned} \tag{5.43}$$

where $C_{\rm R}$ depends on C_w and σ in (4.16) but not on γ, and we use the fact that $\gamma \geqslant 2$.

Putting together (5.37), (5.38), (5.42) and (5.43), we readily obtain (5.35). In fact, for $\ell \geqslant 8$ and $\ell \in \{6\rm R, 6SL\}$, (5.35) is a consequence of (5.38), and (5.37) with the worst possible $p = 0$. For $\ell \in \{2\rm L, 4L\}$ we also need to use (5.42). Finally, for $\ell \in \{2\rm R, 4R\}$ we additionally have to use (5.43) and rely on the first equality in (5.37). A power of γ that we lose in the r.h.s. of (5.43) is compensated during dilatation, due to the presence of derivatives in the couplings $H_{\rm 2R}$, $H_{\rm 4R}$. Because of the sums in the r.h.s. of (5.13) and (5.14), we get eq. (5.35) with an extra factor of $\gamma^{-2} + \gamma^{-1}C_{\rm R} + C_{\rm R} \leqslant 1 + 2C_{\rm R}$ for $\ell = 2\rm R$ and

[30] Recall that Frechet derivative is a generalization of ordinary derivative to Banach spaces. In general, for a map $f(x)$ from a Banach space Z to another space Z', its Frechet derivative at a point x is defined as a linear operator $\nabla f(x) \in \mathcal{L}(Z, Z')$ having the property that

$$\lim_{\|\delta x\|_Z \to 0} \frac{\|f(x+\delta x) - f(x) - \nabla f(x)\delta x\|_{Z'}}{\|\delta x\|_Z} = 0.$$

In some of our cases of interest, one of the two spaces Z or Z' may be $\mathbb{R}$. When $Z = \mathbb{R}$ we have $\nabla f(x) \in Z'$, and when $Z' = \mathbb{R}$ we have $\nabla f(x) \in \mathcal{L}(Z, \mathbb{R})$, i.e. a linear functional on Z.

$\gamma^{-1} + C_{\rm R} \leqslant 1 + C_{\rm R}$ for $\ell = 4{\rm R}$. We absorb this factor by increasing the constant C_0 in the definition of the function ρ_l in the right side of (5.35).[31]

We will also need Frechet derivatives of the homogeneous map $\mathcal{R}_\ell^{\ell_1,\ldots,\ell_n}$. Since $\mathcal{R}_\ell^\ell = R_\ell^\ell$ is a linear map, its Frechet derivative coincides with it and satisfies the same bound (5.34). In all the other cases $(n;(\ell_1,\cdots,\ell_n)) \neq (1;\ell)$ we have the bound

$$\|[\nabla_H \mathcal{R}_\ell^{\ell_1,\ldots,\ell_n}(H)]\delta H\|_w \leqslant \gamma^{-D_l} \sum_{i=1}^n \rho_l(H_{\ell_1},\ldots,\delta H_{\ell_i},\ldots,H_{\ell_n}), \tag{5.44}$$

which follows from (5.41) just as (5.35) followed from (5.38).

6 Construction of the fixed point

In this and the following section, we finally construct a solution of the FPE $f(y) = 0$, see (5.32), and discuss its uniqueness and regularity properties. The presentation is organized as follows: in section 6.1, we state the main bound on the components of $f(y)$, whose proof (which is one of the main technical contributions of this paper, and uses in a crucial way the bounds stated in section 5.6) is postponed to section 7. Given the bounds of section 6.1, existence and uniqueness at fixed γ of the fixed point follow by a rather general and straightforward argument, discussed in sections 6.2 and 6.3. The independence of the fixed point from γ and its analyticity in ε are simple but remarkable corollaries of our construction, discussed in sections 6.4 and 6.5, respectively.

6.1 Key lemma

In this subsection we formulate, as promised, the estimates for the functions $e_\nu^{(0)}, e_\lambda^{(0)}, e_u$ entering the definition of $f(y)$, see (5.32)–(5.33). We will assume that γ is large enough and that the norm of y, see (4.19), is bounded, say smaller than 1; the constants $A_0, A_0^R, A_1^R, A_2^R, A$ in (4.19) will be fixed in a suitable, γ-dependent, way, and the parameter δ will be chosen sufficiently small (in a γ-dependent way). The smallness of δ is conceptually independent from any stringent requirement on the physical parameter ε: the only needed condition on ε will be that all u_ℓ directions are irrelevant, as guaranteed by eqs. (2.8), (5.23). The conditions that λ is weakly relevant (ε small), and that its one-loop beta-function does not vanish ($N \neq 8$) won't be used here. To emphasize that for the moment the smallness of ε is not used, here we assume that δ is independent of ε. Eventually, the smallness of ε will come back into play in the full contraction argument involving all couplings ν, λ, u_ℓ (sections 6.2 and 6.3): there, δ will be identified with ε up to a constant factor, but here it is logically convenient to keep them separate.

Given γ-dependent constants $A_0 = A_0(\gamma)$, $A_0^{\rm R} = A_0^{\rm R}(\gamma)$, $A_1^{\rm R} = A_1^{\rm R}(\gamma)$, $A_2^{\rm R} = A_2^{\rm R}(\gamma)$, $A = A(\gamma)$, we denote by $\|u\|_{B(\gamma,\delta)}$ the following norm of a vector $u = (H_\ell)_{\ell\neq 2{\rm L},4{\rm L},6{\rm SL}} \equiv (u_\ell)_{\ell\in\{2{\rm R},4{\rm R},6{\rm R},8,10,\ldots\}}$ of irrelevant components (6SL excluded):

$$\|u\|_{B(\gamma,\delta)} = \max\left\{\frac{\|u_{2\rm R}\|_w}{A_0^{\rm R}(\gamma)\delta^2}, \frac{\|u_{4\rm R}\|_w}{A_1^{\rm R}(\gamma)\delta^2}, \frac{\|u_{6\rm R}\|_w}{A_2^{\rm R}(\gamma)\delta^3}, \sup_{\ell\geqslant 8}\frac{\|u_\ell\|_w}{A(\gamma)\delta^{k(\ell)}}\right\}, \tag{6.1}$$

[31] Since $\sum|\ell_i| \geqslant 4$ in any of these cases, it's enough to increase $C_0 \to C_0(1+2C_{\rm R})^{1/4}$.

where $k(\ell) = \frac{|\ell|}{2} - 1$, in terms of which the norm (4.19) of $y = (\nu, \lambda, u)$ can be rewritten

$$\|y\|_{Y(\gamma,\delta)} = \max\left\{\frac{|\nu|}{A_0(\gamma)\delta}, \frac{|\lambda|}{A_0(\gamma)\delta}, \|u\|_{B(\gamma,\delta)}\right\}. \tag{6.2}$$

Note that, compared with (4.19), the symbol Y in (6.2) has an explicit dependence upon γ and δ; the dependence on δ is obvious, the one on γ is meant to emphasize the γ-dependence of the constants A, A_0, etc. and of the weight w (see (4.16)). We are now ready to state the main result of this section.

Lemma 6.1 (Key lemma) *Choose $d \in \{1,2,3\}$, cutoff χ, $N \geqslant 4$, and an ε satisfying (2.8). There exists $\gamma_{\rm key} \geqslant 2$ and*

$$\delta_0(\gamma), A_0(\gamma), \{A_k^{\rm R}(\gamma)\}_{k=0,1,2}, A(\gamma), E_0(\gamma), E_1(\gamma), \tag{6.3}$$

positive continuous functions on $\gamma \geqslant \gamma_{\rm key}$ [whose dependence on γ is omitted in eqs. (6.5)–(6.6) below], with the following property. Take any $\gamma \geqslant \gamma_{\rm key}$, any $0 < \delta \leqslant \delta_0(\gamma)$, and any sequence $y = (\nu, \lambda, u)$ satisfying

$$\|y\|_{Y(\gamma,\delta)} \leqslant 1. \tag{6.4}$$

Then the infinite sums defining the functions $e_\nu^{(0)}, e_\lambda^{(0)}, e_u$ in the right side of (5.30), see (5.29) and following lines, are absolutely convergent, and their sums satisfy:

$$|e_\nu^{(0)}(y)| \leqslant E_0\delta^2, \qquad |e_\lambda^{(0)}(y)| \leqslant E_1\delta^3, \qquad \|e_u(y)\|_{B(\gamma,\delta)} \leqslant \gamma^{-\bar D}, \tag{6.5}$$

where $\bar D = \frac{1}{2}\min\{D_2+2, D_4+1, D_6\}$. In addition,

$$\begin{aligned}
&|\partial_i e_\nu^{(0)}(y)| \leqslant E_0\delta^2/(A_0\delta), && |\partial_i e_\lambda^{(0)}(y)| \leqslant E_1\delta^3/(A_0\delta),\\
&\|\partial_i e_u(y)\|_B \leqslant \gamma^{-\bar D}/(A_0\delta) \quad (i=\nu,\lambda), &&\\
&\|\partial_u e_\nu^{(0)}(y)\|_{\mathcal{L}(B,\mathbb{R})} \leqslant E_0\delta^2, && \|\partial_u e_\lambda^{(0)}(y)\|_{\mathcal{L}(B,\mathbb{R})} \leqslant E_1\delta^3,\\
&\|\partial_u e_u(y)\|_{\mathcal{L}(B,B)} \leqslant \gamma^{-\bar D}, &&
\end{aligned} \tag{6.6}$$

where $B = B(\gamma,\delta)$, $\mathcal{L}(B,\mathbb{R})$ is the space of linear operators from B to $\mathbb{R}$, and similarly for $\mathcal{L}(B,B)$.

We wrote (6.6) in the form which makes apparent that the u-derivatives satisfy the same bounds as the functions themselves, while the bounds for ν, λ-derivatives are worse by $1/(A_0\delta)$ factor. This pattern is natural in view of the assumptions $|\nu| \leqslant A_0\delta$, $|\lambda| \leqslant A_0\delta$ and $\|u\|_B \leqslant 1$ (which are the same as $\|y\|_Y \leqslant 1$). Before presenting the proof of the Key lemma, which is postponed to section 7, we will show that its bounds straightforwardly imply that the FPE $f(y) = 0$ has a unique solution in a suitable neighborhood of the Banach space Y; see the next two subsections, 6.2 and 6.3.

Remark 6.1 The third inequality in (6.5) means that the RG map restricted to the irrelevant directions $\ell = 2{\rm R}, 4{\rm R}, 6{\rm R}, 8, 10, \ldots$ is contractive, as it is natural to expect. Contractivity along the directions with $|\ell| \geqslant 6$ is "easy" to establish: it follows straightforwardly from the bounds (5.34) and (5.35); note, in fact, that $\gamma^{-D_l} \leqslant \gamma^{-\bar{D}} < 1$ for $l = |\ell| \geqslant 6$. See sections 7.1 and 7.2 for the full proof. On the other hand, contractivity along the directions 2R and 4R is more subtle to prove, due to the factor γ^{-D_l}, which is larger than 1 for $l = 2, 4$, in the right side of (5.35). In these cases, we take advantage of the fact that the linearization of the RG map, (5.34), has the good factor γ^{-D_2-2} and γ^{-D_4-1} in the directions $\ell = 2\text{R}$ and 4R, respectively. On the other hand, the nonlinear contributions bounded in (5.35) are small because they are of higher order: loosely speaking, the higher order can be used to compensate the additional bad factor γ^2 or γ, which ultimately originates from the bounds (5.43). More technically, here is where we use the freedom in the choice of the constants A, A_0, A_0^{R} etc., entering the definition the norm (4.19) defining the Banach space: by carefully playing with these γ-dependent constants, we can reabsorb the bad factors γ^2 or γ into their definitions, see sections 7.3, 7.4 and 7.7 for the technical details.

Remark 6.2 In connection with the end of previous remark, we note that the use of a norm involving several constants A, A_0, A_0^{R} etc, rather than a single one, is one original aspect of our proof, and it is the key ingredient allowing us to choose an optimal powers of δ in (4.19) (recall that eventually δ will be chosen proportional to ε, and that the δ-exponents $2, 2, 3, k(\ell) = |\ell|/2-1$ in the right side of (4.19) are dictated by the lowest order contribtions to u_ℓ in perturbation theory and cannot be improved; see the discussion at the beginning of section 4.2.2). If we tried to repeat the proof of Key lemma with a simplified norm with $A = A_0 = A_0^{\text{R}} = \cdots$ we would not succeed in proving the analogues of (6.5) and (6.6). One can however use a simplified norm, and a simplified proof of Key lemma, if one changes the optimal powers of δ to sub-optimal ones, strictly smaller than $2, 2, 3, k(\ell)$. This was the strategy followed in [27] (see the non-optimal powers in their eq. (2.17)). Naively, this strategy leads to an estimate on the fixed-point couplings (analogue of corollary 6.1 below) with sub-optimal powers. However, armed with our analyticity argument from section 6.5, this limitation can be overcome. Namely, once the fixed point existence is proven by working in the sub-optimal Banach space, the argument from section 6.5 still works and shows that it is analytic in a disk around $\varepsilon = 0$. From analyticity, we could then recover the optimal estimates on the fixed-point couplings. Although such a mixed real/complex strategy is possible, here we prefer to keep these two lines of development independent. So, we work with the optimal powers from the start and obtain the optimal estimates with purely real methods (even though it leads to some mild complications in the proof of Key Lemma).

6.2 Abstract analysis

Recall that we are solving $f(y) = 0$ with $y = (\nu, \lambda, u)$ and f given in (5.32). In this subsection we consider u as a vector living in an abstract Banach space B endowed with some norm $\|u\|_B$. This norm will be used to state conditions on the maps e_j guaranteeing the existence and uniqueness of a solution in some neighborhood of y_0, see eq. (6.10). In the next subsection these conditions will be verified with the help of Key Lemma, identifying the norm $\|u\|_B$ with (6.1).

Concerning the rescalings (5.33), note that (the O symbols here and in (6.8) have γ- and ε-independent constants):

$$\frac{\varepsilon}{1-\gamma^{-2\varepsilon}} = (2\log\gamma)^{-1}(1+O(\varepsilon\log\gamma)). \tag{6.7}$$

By the small ε asymptotics of I_1, I_2 from Lemma G.1 in appendix G, we have

$$\begin{aligned} a &= 2(N-2)\left[\int \frac{d^dk}{(2\pi)^d}\frac{\chi(k)}{|k|^{d/2}} + O(\varepsilon\log\gamma)\right], \\ b &= -2(N-8)\left[\frac{S_d}{(2\pi)^d} + O(\varepsilon\log\gamma)\right]. \end{aligned} \tag{6.8}$$

As mentioned in the introduction, we are assuming $N \neq 8$ so that $b \neq 0$. We will also assume $\varepsilon \leqslant c/\log\gamma$ where c is a small γ-independent constant. Under these conditions $a, b, b^{-1} = O(1)$. In particular $b \neq 0$.

Setting e_j ($j = \nu, \lambda, u$) to zero in (5.32), we get an "approximate equation"

$$f_0(y) = 0, \quad f_0(y) = \begin{pmatrix} \nu + a\lambda \\ \varepsilon\lambda + b\lambda^2 \\ u \end{pmatrix}, \tag{6.9}$$

which has a nontrivial solution

$$y_0 = (\nu_0, \lambda_0, u_0) = \left(\frac{a}{b}\varepsilon, -\frac{1}{b}\varepsilon, 0\right). \tag{6.10}$$

Our goal will be to show that the full equation $f(y) = 0$ has a solution of the form $y_0+O(\varepsilon^2)$. Aiming to apply a contraction argument, we rewrite equation $f(y) = 0$ one last time as a fixed point equation for a map $F(y)$. We choose the following rewrite:

$$f(y) = 0 \quad \Longleftrightarrow \quad y = F(y), \qquad F(y) = y - G^{-1}f(y), \tag{6.11}$$

with G an arbitrary invertible linear operator. We would like to choose G so that $F(y)$ is a contraction in a small neighborhood of y_0. Recall that Newton's method for solving nonlinear equations would correspond to $G = \nabla f(y)$. We do not want to deal with the full gradient of the complicated map $f(y)$, and we will instead choose $G = \nabla f_0(y_0)$, cf. (6.9). This "approximated gradient" will be sufficient to make $F(y)$ a contraction. We have

$$G = \begin{pmatrix} 1 & a & 0 \\ 0 & -\varepsilon & 0 \\ 0 & 0 & \mathbb{1} \end{pmatrix}, \qquad G^{-1} = \begin{pmatrix} 1 & a\varepsilon^{-1} & 0 \\ 0 & -\varepsilon^{-1} & 0 \\ 0 & 0 & \mathbb{1} \end{pmatrix}. \tag{6.12}$$

With this choice, the map $F(y)$ takes the form:

$$F(y) \equiv \begin{pmatrix} F^\nu(y) \\ F^\lambda(y) \\ F^u(y) \end{pmatrix} = \begin{pmatrix} -2a\lambda - ab\frac{\lambda^2}{\varepsilon} - e_\nu - a\frac{e_\lambda}{\varepsilon} \\ 2\lambda + b\frac{\lambda^2}{\varepsilon} + \frac{e_\lambda}{\varepsilon} \\ e_u \end{pmatrix}. \tag{6.13}$$

Remark 6.3 The reader may be puzzled: why introduce the new map F rather than use in its place the renormalization map itself, given that eq. (5.30) already has the fixed point form $y = R(y)$? The reason is that contraction argument cannot be applied directly to R, since it is not fully contracting: it is contracting near the fixed point along all directions except ν. Note as well that R is only "barely contracting" in direction λ: its linearization around y_0 has the corresponding eigenvalue equal to $2 - \gamma^{2\varepsilon} = 1 - 2\varepsilon \log \gamma + \cdots$, smaller than 1 but only by $O(\varepsilon)$. This "barely contracting" direction is the reason why we apply the contraction argument in a neighborhood of size ε^2 of y_0 (outside of which even F would not be contracting).

We will aim to apply a contraction argument to $F(y)$ in a neighborhood Y_0 of y_0 defined as

$$Y_0 = \{y : |\nu - \nu_0| \leqslant M_0\varepsilon^2, |\lambda - \lambda_0| \leqslant M_0\varepsilon^2, \|u\|_B \leqslant 1\}, \tag{6.14}$$

whose size depends on ε and on an additional parameter M_0. First of all let us arrange that F maps Y_0 to itself. Writing $\lambda = \lambda_0 + \delta\lambda$, we express $F(y)$ as

$$F(y) = \begin{pmatrix} \nu_0 - \frac{ab}{\varepsilon}(\delta\lambda)^2 - e_\nu - a\frac{e_\lambda}{\varepsilon} \\ \lambda_0 + \frac{b}{\varepsilon}(\delta\lambda)^2 + \frac{e_\lambda}{\varepsilon} \\ e_u \end{pmatrix}. \tag{6.15}$$

We see that $F(Y_0) \subset Y_0$ provided that for any $y \in Y_0$

$$K_1 \max(M_0^2\varepsilon^3, |e_\nu(y)|, \varepsilon^{-1}|e_\lambda(y)|) \leqslant M_0\varepsilon^2, \qquad \|e_u(y)\|_B \leqslant 1, \tag{6.16}$$

where

$$K_1 = K_1(a, b) = \max(1 + |a| + |ab|, 1 + |b|). \tag{6.17}$$

Then (6.16) are satisfied as long as $K_1 M_0 \varepsilon \leqslant 1$ and provided that

$$|e_\nu(y)| \leqslant \frac{M_0}{K_1}\varepsilon^2, \qquad |e_\lambda(y)| \leqslant \frac{M_0}{K_1}\varepsilon^3, \qquad \|e_u(y)\|_B \leqslant 1 \qquad (y \in Y_0). \tag{6.18}$$

We next proceed to arrange that F is a contraction in Y_0. For this we need to specify a Banach space norm on $y = (\nu, \lambda, u)$. We will use the norm

$$\|y\|_Y = \max\{\tilde\varepsilon^{-1}|\nu|, \tilde\varepsilon^{-1}|\lambda|, \|u\|_B\}, \tag{6.19}$$

depending on a parameter $\tilde\varepsilon$. Since in our application ν and λ are $O(\varepsilon)$, while $\|u\|_B$ will be $O(1)$ when identified with (6.1), the natural value for $\tilde\varepsilon$ is order ε so that all terms in (6.19) have the same order. Eventually in section 6.3 we will fix $\tilde\varepsilon = A_0\delta$ so that this norm will coincide with (6.2). However in this section let us keep the ratio $\tilde\varepsilon/\varepsilon$ as a free parameter.

We will next study the gradient ∇F and arrange that its operator norm is less than 1. Here the gradient ∇F is the Frechet derivative which was already discussed in section 6.1.

From (6.13), we compute the gradient ∇F in components:

$$\frac{\partial(F^\nu, F^\lambda, F^u)}{\partial(\nu, \lambda, u)} = \begin{pmatrix} -\partial_\nu e_\nu - a\frac{\partial_\nu e_\lambda}{\varepsilon} & -\frac{2ab}{\varepsilon}(\lambda - \lambda_0) - \partial_\lambda e_\nu - a\frac{\partial_\lambda e_\lambda}{\varepsilon} & -\partial_u e_\nu - a\frac{\partial_u e_\lambda}{\varepsilon} \\ \frac{\partial_\nu e_\lambda}{\varepsilon} & \frac{2b}{\varepsilon}(\lambda - \lambda_0) + \frac{\partial_\lambda e_\lambda}{\varepsilon} & \frac{\partial_u e_\lambda}{\varepsilon} \\ \partial_\nu e_u & \partial_\lambda e_u & \partial_u e_u \end{pmatrix}. \tag{6.20}$$

Various partial derivatives of maps $e_j(\nu,\lambda,u)$ are understood as Frechet derivatives, sometimes with Z or Z' being equal to $\mathbb{R}$. E.g. $\partial_\nu e_u$ is, just like e_u, a B-valued function on Y. On the other hand $\partial_u e_\lambda \in \mathcal{L}(B,\mathbb{R})$, a linear functional on B.

We next proceed to study the norm of $\nabla F(y) \in \mathcal{L}(Y,Y)$ where $y \in Y_0$. Let $\delta y \in Y$, $\|\delta y\|_Y \leqslant 1$ which means

$$|\delta\nu| \leqslant \tilde{\varepsilon}, \quad |\delta\lambda| \leqslant \tilde{\varepsilon}, \quad \|\delta u\|_B \leqslant 1. \tag{6.21}$$

We have

$$\nabla F(y)\delta y = \begin{pmatrix} \partial_\nu F^\nu \delta\nu + \partial_\lambda F^\nu \delta\lambda + \partial_u F^\nu \delta u \\ \partial_\nu F^\lambda \delta\nu + \partial_\lambda F^\lambda \delta\lambda + \partial_u F^\lambda \delta u \\ \partial_\nu F^u \delta\nu + \partial_\lambda F^u \delta\lambda + \partial_u F^u \delta u \end{pmatrix}, \tag{6.22}$$

where all partial derivatives in the r.h.s. are evaluated at y. This implies

$$\begin{aligned} \|\nabla F(y)\|_{\mathcal{L}(Y,Y)} &= \sup_{\|\delta y\|_Y \leqslant 1} \|\nabla F(y)(\delta y)\|_Y \\ &= \sup_{\|\delta y\|_Y \leqslant 1} \max \begin{pmatrix} \tilde{\varepsilon}^{-1}|\partial_\nu F^\nu \delta\nu + \partial_\lambda F^\nu \delta\lambda + \partial_u F^\nu \delta u| \\ \tilde{\varepsilon}^{-1}|\partial_\nu F^\lambda \delta\nu + \partial_\lambda F^\lambda \delta\lambda + \partial_u F^\lambda \delta u| \\ \|\partial_\nu F^u \delta\nu + \partial_\lambda F^u \delta\lambda + \partial_u F^u \delta u\|_B \end{pmatrix} \\ &\leqslant \max \begin{pmatrix} |\partial_\nu F^\nu| + |\partial_\lambda F^\nu| + \tilde{\varepsilon}^{-1}\|\partial_u F^\nu\|_{\mathcal{L}(B,\mathbb{R})} \\ |\partial_\nu F^\lambda| + |\partial_\lambda F^\lambda| + \tilde{\varepsilon}^{-1}\|\partial_u F^\lambda\|_{\mathcal{L}(B,\mathbb{R})} \\ \tilde{\varepsilon}\|\partial_\nu F^u\|_B + \tilde{\varepsilon}\|\partial_\lambda F^u\|_B + \|\partial_u F^u\|_{\mathcal{L}(B,B)} \end{pmatrix} \end{aligned} \tag{6.23}$$

Finally using the explicit form of ∇F components we get that for $y \in Y_0$

$$\begin{aligned} \|\nabla F(y)\|_{\mathcal{L}(Y,Y)} \leqslant K_2 \max\Big\{ &|\partial_\nu e_\nu|, \frac{|\partial_\nu e_\lambda|}{\varepsilon}, M_0\varepsilon, |\partial_\lambda e_\nu|, \frac{|\partial_\lambda e_\lambda|}{\varepsilon}, \frac{\|\partial_u e_\nu\|_{\mathcal{L}(B,\mathbb{R})}}{\tilde{\varepsilon}}, \frac{\|\partial_u e_\lambda\|_{\mathcal{L}(B,\mathbb{R})}}{\varepsilon\tilde{\varepsilon}}, \\ &\tilde{\varepsilon}\|\partial_\nu e_u\|_B, \tilde{\varepsilon}\|\partial_\lambda e_u\|_B, \|\partial_u e_u\|_{\mathcal{L}(B,B)} \Big\}, \end{aligned} \tag{6.24}$$

where we used that $|\lambda - \lambda_0| \leqslant M_0\varepsilon^2$ in Y_0 and defined a constant

$$K_2(a,b) = \max(3 + 3|a| + 2|ab|, 3 + 2|b|). \tag{6.25}$$

We will demand that the following conditions hold uniformly for $y \in Y_0$:

$$\begin{aligned} &|\partial_i e_\nu| \leqslant M_0\varepsilon, && |\partial_i e_\lambda| \leqslant M_0\varepsilon^2, && \|\partial_i e_u\|_B \leqslant \alpha\tilde{\varepsilon}^{-1} \quad (i = \nu, \lambda), \\ &\|\partial_u e_\nu\|_{\mathcal{L}(B,\mathbb{R})} \leqslant M_0\varepsilon\tilde{\varepsilon}, && \|\partial_u e_\lambda\|_{\mathcal{L}(B,\mathbb{R})} \leqslant M_0\varepsilon^2\tilde{\varepsilon}, && \|\partial_u e_u\|_{\mathcal{L}(B,B)} \leqslant \alpha, \end{aligned} \tag{6.26}$$

where α is yet another parameter. Under these conditions eq. (6.24) implies:

$$\|\nabla F(y)\|_{\mathcal{L}(Y,Y)} \leqslant \max(K_2 M_0\varepsilon, K_2\alpha) \qquad (y \in Y_0). \tag{6.27}$$

We restate the conclusions of the above discussion as

Lemma 6.2 (Abstract Lemma) *Suppose that, for a given ε,the constants $M_0, \tilde{\varepsilon}, \alpha$ are such that maps e_j satisfy bounds (6.18) and (6.26) everywhere in Y_0 defined by (6.14). Suppose in addition that (see (6.17), (6.25) for the definition of K_1 and K_2)*

$$K_1 M_0\varepsilon \leqslant 1, \qquad K_2 M_0\varepsilon \leqslant 1/2, \qquad K_2\alpha \leqslant 1/2. \tag{6.28}$$

Then $F(Y_0) \subset Y_0$ and $\|\nabla F(y)\|_{\mathcal{L}(Y,Y)} \leqslant 1/2$ in Y_0, so that F is a contraction in Y_0 and has a unique fixed point there.

6.2.1 Complex version of the Abstract Lemma

By a few minor modifications of the proof of the Abstract Lemma we can get a complex-ε version thereof. This is needed in the proof of fixed point analyticity (section 6.5) and is not used anywhere else. We let $\varepsilon \in \mathbb{C}$, y be an element of the *complex* Banach space $\mathbb{Y}$ with the norm (6.19), and $\mathbb{Y}_0$ (the complex analogue of Y_0, see (6.14)) be defined as:

$$\mathbb{Y}_0 = \{y : |\nu - \nu_0| \leqslant M_0|\varepsilon|^2, |\lambda - \lambda_0| \leqslant M_0|\varepsilon|^2, \|u\|_B \leqslant 1\}. \tag{6.29}$$

Then the following generalization of Lemma 6.2 holds.

Lemma 6.3 (Complex Abstract Lemma) *Suppose that, for a given $\varepsilon \in \mathbb{C}$,the constants $M_0, \tilde{\varepsilon}, \alpha$ are such that maps e_j satisfy bounds the complex analogues of (6.18) and (6.26), i.e.,*

$$|e_\nu(y)| \leqslant \frac{M_0}{K_1}|\varepsilon|^2, \qquad |e_\lambda(y)| \leqslant \frac{M_0}{K_1}|\varepsilon|^3, \qquad \|e_u(y)\|_B \leqslant 1 \qquad (y \in \mathbb{Y}_0), \tag{6.30}$$

and

$$\begin{aligned} &|\partial_i e_\nu| \leqslant M_0|\varepsilon|, && |\partial_i e_\lambda| \leqslant M_0|\varepsilon|^2, && \|\partial_i e_u\|_B \leqslant \alpha\tilde{\varepsilon}^{-1} \quad (i = \nu, \lambda), \\ &\|\partial_u e_\nu\|_{\mathcal{L}(B,\mathbb{R})} \leqslant M_0|\varepsilon|\tilde{\varepsilon}, && \|\partial_u e_\lambda\|_{\mathcal{L}(B,\mathbb{R})} \leqslant M_0|\varepsilon|^2\tilde{\varepsilon}, && \|\partial_u e_u\|_{\mathcal{L}(B,B)} \leqslant \alpha, \end{aligned} \tag{6.31}$$

everywhere in $\mathbb{Y}_0$ defined by (6.29). Suppose in addition that (see (6.17), (6.25) for the definition of K_1 and K_2)

$$K_1 M_0|\varepsilon| \leqslant 1, \qquad K_2 M_0|\varepsilon| \leqslant 1/2, \qquad K_2\alpha \leqslant 1/2. \tag{6.32}$$

Then $F(\mathbb{Y}_0) \subset \mathbb{Y}_0$ and $\|\nabla F(y)\|_{\mathcal{L}(Y,Y)} \leqslant 1/2$ in $\mathbb{Y}_0$, so that F is a contraction in $\mathbb{Y}_0$ and has a unique fixed point there.

Proof. The proof of this lemma is a straightforward repetition of the one of Lemma 6.2, modulo the replacement of ε by $|\varepsilon|$ in a few inequalities. More precisely, a simple critical rereading of the proof shows that, if we leave the definitions of F, see (6.15), and of ∇F, see (6.20) and (6.22), as they are, and we replace Y_0 by $\mathbb{Y}_0$ and ε by $|\varepsilon|$ everywhere in the rest of the proof (in particular in the following places: 1 line after (6.17); in eq. (6.18); in eq. (6.24); in eq. (6.26); and in eq. (6.27)), then we readily obtain the desired claim.

6.3 Fixed point theorem

In this section we will put Key Lemma and Abstract Lemma together and will finally show that the FPE (5.32) has a solution. Namely, we will prove the following result:

Theorem 6.1 *There exists a $\gamma_0 \geqslant 2$ and a positive continuous function $\varepsilon_0(\gamma)$ defined for $\gamma \geqslant \gamma_0$ such that for each $\gamma \geqslant \gamma_0$ and $0 < \varepsilon \leqslant \varepsilon_0(\gamma)$ the fixed point equation (5.32) has a nontrivial solution.*

Proof. We will show, with the help of Key Lemma 6.1, that for $\gamma \geqslant \gamma_0$ and for $0 < \varepsilon \leqslant \varepsilon_0(\gamma)$ conditions of Abstract Lemma 6.2 can be satisfied.

We thus identify the abstract Banach space B in section 6.2 with the space $B(\gamma, \delta)$ in (6.1). We also put

$$\tilde{\varepsilon} = A_0 \delta, \tag{6.33}$$

and identify the space Y from (6.19) with $Y(\gamma, \delta)$ in (6.2). The parameter δ in the Key Lemma will be chosen proportional to ε:

$$\delta = h\varepsilon, \tag{6.34}$$

with h to be fixed momentarily.

Abstract Lemma requires us to examine the neighborhood Y_0 defined in (6.14). By $K_1 M_0 \varepsilon \leqslant 1$, the first of conditions (6.28) (we will make sure to satisfy all of these conditions below), the points of Y_0 will satisfy

$$|\nu|, |\lambda| \leqslant K_3 \varepsilon, \qquad K_3 = K_3(a, b) = \max\left(\left|\frac{a}{b}\right| + \frac{1}{K_1}, \frac{1}{|b|} + \frac{1}{K_1}\right). \tag{6.35}$$

Let us choose

$$h = K_3/A_0. \tag{6.36}$$

By (6.35), we have

$$Y_0 \subset \{y : \|y\|_Y \leqslant 1\}. \tag{6.37}$$

Thus, the basic assumption (6.4) holds in Y_0, and we can use Key Lemma to estimate e_j and their derivatives in Y_0.

We will also fix (see Key Lemma for the definition of $\bar{D}$)

$$\alpha = \gamma^{-\bar{D}}. \tag{6.38}$$

With this identification and (6.33), the bounds on the derivatives of $\partial_i e_u, \partial_u e_u$ requested in (6.26) coincide with the bounds for the same derivatives in (6.6) of the Key Lemma. The request $\|e_u\|_B \leqslant 1$ in Y_0 (eq. (6.18)) is also satisfied by the bound on $\|e_u\|_B$ in (6.5).

Furthermore, we choose γ_0 as

$$\gamma_0 = \max(\gamma_{\text{key}}, (2K_2)^{1/\bar{D}}). \tag{6.39}$$

Then for $\gamma \geqslant \gamma_0$ we have $\gamma \geqslant \gamma_{\text{key}}$ so that we can use Key Lemma, and in addition we satisfy the third condition in (6.28).

Let us now arrange for the conditions in (6.18) and (6.26) concerning e_ν, e_λ, and their derivatives. By eq. (5.33), e_ν, e_λ equal $e_\nu^{(0)}, e_\lambda^{(0)}$ times factors bounded by a γ-dependent constant f_γ. Key Lemma gives estimates for $e_\nu^{(0)}, e_\lambda^{(0)}$ and their derivatives with constants E_0, E_1 in the r.h.s., and e_ν, e_λ will satisfy the same estimates with $E_i \to E_i' = f_\gamma E_i$. Using the proportionality (6.34) between δ and ε, these estimates take the form

$$\begin{aligned}
&|e_\nu| \leqslant E_0' h^2 \varepsilon^2, && |e_\lambda| \leqslant E_1' h^3 \varepsilon^3, \\
&|\partial_i e_\nu| \leqslant (E_0'/A_0) h\varepsilon, && |\partial_i e_\lambda| \leqslant (E_1'/A_0) h^2 \varepsilon^2 \qquad (i = \nu, \lambda), \\
&\|\partial_u e_\nu\|_{\mathcal{L}(B,\mathbb{R})} \leqslant E_0' h^2 \varepsilon^2, && \|\partial_u e_\lambda(\nu, \lambda, u)\|_{\mathcal{L}(B,\mathbb{R})} \leqslant E_1' h^3 \varepsilon^3.
\end{aligned} \tag{6.40}$$

These have the same scaling in ε as the corresponding estimates in (6.18), (6.26) (recall that $\tilde{\varepsilon}/\varepsilon = A_0 h$). So, to satisfy (6.18), (6.26), we simply choose M_0 sufficiently large, namely:

$$M_0 = \max(K_1 E_0' h^2, K_1 E_1' h^3, (E_0'/A_0)h, (E_1'/A_0)h^2). \tag{6.41}$$

We still have to satisfy the first two conditions in (6.28), as well as to make sure that $\delta = h\varepsilon \leqslant \delta_0$. We achieve this by choosing

$$\varepsilon_0(\gamma) = \min\left(\frac{\delta_0}{h}, \frac{1}{K_1 M_0}, \frac{1}{2K_2 M_0}\right). \tag{6.42}$$

For any $0 < \varepsilon \leqslant \varepsilon_0(\gamma)$, conditions of Abstract Lemma are satisfied, and hence a fixed point exist.

Corollary 6.1 *The fixed point whose existence we proved belongs to the neighborhood*

$$\begin{aligned}&|\nu - \nu_0|, |\lambda - \lambda_0| \leqslant M_0 \varepsilon^2,\\ &\|H_{2R}\|_w \leqslant A_0^R h^2 \varepsilon^2, \|H_{4R}\|_w \leqslant A_1^R h^2 \varepsilon^2, \|H_{6R}\|_w \leqslant A_2^R h^3 \varepsilon^3, \|H_l\|_w \leqslant A h^{l/2-1} \varepsilon^{l/2-1},\end{aligned} \tag{6.43}$$

where C_0, A, A_k^R, h are some γ-dependent quantities. Moreover in this neighborhood this is a unique solution of the fixed point equation.

This follows from writing in full the condition $\|u\|_B \leqslant 1$.

6.4 Semigroup property and γ-independence

Theorem 6.1 shows that the renormalization map $R(\varepsilon, \gamma)$ has a fixed point provided that $\gamma \geqslant \gamma_0$ is sufficiently large and $\varepsilon \leqslant \varepsilon_0(\gamma)$ is sufficiently small. We would now like to study how this fixed point depends on various parameters. In this section we will discuss γ-independence, while in the next one we will show that it depends on ε analytically.

The γ-independence at $O(\varepsilon)$ is visible in eq. (6.10), since both a and b become γ-independent as $\varepsilon \to 0$. That it should hold in general can be suspected from the semigroup property (2.17). Indeed, if a certain interaction H_* is a fixed point of $R(\varepsilon, \gamma)$, then by the semigroup property it is also a fixed point of $R(\varepsilon, \gamma^n)$ for any $n \geqslant 2$, as long as $R(\varepsilon, \gamma^n)$ is defined on H_* as a continuous map acting on a neighborhood of a Banach space to which H_* belongs. Combining this simple argument with continuity in γ, one should be able to prove that, in fact, the fixed point is unique and completely independent of γ, at least on a suitable interval of values of γ, such as the one denoted by J below. A full proof of this fact requires a critical re-reading of the proofs of the Key Lemma, of the Abstract Lemma and of the Fixed point theorem, as well as a generalization thereof, providing existence and uniqueness of the general FPE in the space of equivalency classes of couplings modulo null sequences, see remark 5.3 and remark 6.4 below. We won't belabor all the required details here, but we will provide all the elements sufficient for a willing reader to sit down and check the various claims, most of which are just straightforward corollaries of the previous discussion.

Fix an interval $I = [\gamma_{\text{key}}, \bar{\gamma}]$, with γ_{key} the constant of the Key Lemma. From the proof the Key Lemma, see in particular section 7.7, we see that, with no loss of generality, all the functions in (6.3) can be chosen to be decreasing in γ, so that their smallest values in I are those at $\gamma = \bar{\gamma}$, which we denote by $\bar{\delta}_0, \bar{A}_0, \{\bar{A}_k^R\}_{k=0,1,2}, \bar{A}, \bar{E}_0, \bar{E}_1$. It is easy to check that the Key Lemma 6.1 admits the following "uniform" version on I: take $\gamma \in I$, $0 \leqslant \delta \leqslant \bar{\delta}_0$, and any sequence (ν, λ, u) satisfying the inequalities (6.4) with A_0 replaced by $\bar{A}_0$ and $\|u\|_B$ replaced by $\|u\|_{\bar{B}} = \|u\|_{B(\bar{\gamma},\delta)}$; then the conclusions of the lemma, (6.5) and (6.6) hold, with $\delta_0, A_0, \{A_k^R\}_{k=0,1,2}, A, E_0, E_1$ replaced by $\bar{\delta}_0, \bar{A}_0, \{\bar{A}_k^R\}_{k=0,1,2}, \bar{A}, \bar{E}_0, \bar{E}_1$, and B replaced by $\bar{B}$.

Similarly, we can easily obtain a uniform version of the Abstract Lemma and Fixed point theorem for $0 < \varepsilon \leqslant \bar{\varepsilon}_0$, with $\bar{\varepsilon}_0 = \min_{\gamma \in I} \varepsilon_0(\gamma)$, and $\gamma \in I_0 \equiv [\bar{\gamma}_0, \bar{\gamma}]$, with $\bar{\gamma}_0$ defined by the analogue of (6.39) with K_2 replaced by $\bar{K}_2 = \min_{0<\varepsilon\leqslant\bar{\varepsilon}_0, \gamma\in I} K_2$ (note that I_0 is non empty for $\bar{\gamma}$ large enough). We denote by F_γ the function F of section 6.2, in order to emphasize its dependence upon γ. We define $\bar{\nu}_0 = \frac{a_0}{b_0}\varepsilon$ and $\bar{\lambda}_0 = -\frac{1}{b_0}\varepsilon$, with $a_0 = a|_{\varepsilon=0}$ and $b_0 = b|_{\varepsilon=0}$, see (6.8), and let

$$\overline{Y_0} = \{y : |\nu - \bar{\nu}_0| \leqslant \bar{M}_0 \varepsilon^2, |\lambda - \bar{\lambda}_0| \leqslant \bar{M}_0 \varepsilon^2, \|u\|_{\bar{B}} \leqslant 1\}. \tag{6.44}$$

A critical re-reading of the Abstract Lemma and of the Fixed point theorem shows that there exist constants $\bar{h}$ and $\bar{M}_0$ such that, fixing $\delta = \bar{h}\varepsilon$ and using the uniform version of the Key Lemma, then, for any $\gamma \in I_0$, $F_\gamma(\bar{Y}_0) \subset \bar{Y}_0$ and F_γ is continuous for $\gamma \in I_0$. Moreover, letting $\bar{Y}$ be the Banach space with norm

$$\|y\|_{\bar{Y}} = \max\{(\bar{A}_0\delta)^{-1}|\nu|, (\bar{A}_0\delta)^{-1}|\lambda|, \|u\|_{\bar{B}}\}, \tag{6.45}$$

we have $\|\nabla F_\gamma(y)\|_{\mathcal{L}(\bar{Y},\bar{Y})} \leqslant 1/2$ in $\bar{Y}_0$, so that, for any $\gamma \in I_0$, F_γ is a contraction in $\bar{Y}_0$, uniformly in γ, and has a unique fixed point there, denoted $y_*(\gamma)$. Of course, $y_*(\gamma) = \lim_{n\to\infty} F_\gamma^n(\bar{y}_0)$, with $\bar{y}_0 = (\bar{\nu}_0, \bar{\lambda}_0, 0)$. Recalling that F_γ is continuous in γ for $\gamma \in I_0$ and is uniformly contractive there, we find that $y_*(\gamma)$ is continuous in γ for $\gamma \in I_0$, being the uniform limit of a sequence of uniformly continuous functions.

Remark 6.4 The previous discussion, as well as the one of the previous sections, shows that $y_*(\gamma)$ is the unique solution of the *restricted* FPE (5.28), in the sense of remark 5.3. As discussed there, we expect that a generalization of the methods of this paper will allow us to prove the uniqueness of the solution of the general FPE $(H'_\ell) = (H_\ell) + (N_\ell)$ modulo null couplings, provided the null sequence (N_ℓ) is sufficiently small in norm. We will denote by the symbol $\mathfrak{h}_*(\gamma)$ such a (presumed) unique solution in the space of equivalency classes of couplings. Of course, continuity of $y_*(\gamma)$ implies the continuity of $\mathfrak{h}_*(\gamma)$ in the appropriate topology.

By construction, $y_*(\gamma), \gamma \in I_0$, is the unique solution in $\bar{Y}_0$ to the fixed point equation $y = R_\gamma(y)$, where R_γ is the original form of the RG map (before the manipulations (6.11)), given by the right side of (5.30). By its very definition, R_γ satisfies the semigroup property $R_\gamma \circ R_{\gamma'} = R_{\gamma\cdot\gamma'}$+null, so that, if $\mathfrak{h}_*(\gamma) = \mathfrak{h}_*(\gamma')$, then $\mathfrak{h}_*(\gamma) = \mathfrak{h}_*(\gamma') = \mathfrak{h}_*(\gamma \cdot \gamma')$. From this, it follows that $\mathfrak{h}_*(\gamma) \equiv \mathfrak{h}_*(\bar{\gamma})$ for all the values γ in the subset X of I_0 characterized by the following properties: *(i)* $\bar{\gamma} \in X$; *(ii)* if $\gamma \in X$, then $\gamma^{1/n} \in X$, for all natural n such that $\gamma^{1/n} \in I_0$; *(iii)* if $\gamma, \gamma' \in X$, then $\gamma \cdot \gamma' \in X$, as long as $\gamma \cdot \gamma' \in I_0$. Of course, by the continuity of $\mathfrak{h}_*(\gamma)$, the fixed point is constant and equal to $\mathfrak{h}_*(\bar{\gamma})$ on the closure of X, as well. For $\bar{\gamma} \geqslant (\bar{\gamma}_0)^3$, the closure of X contains the sub-interval $J = [\bar{\gamma}^{2/3}, \bar{\gamma}]$.[32] This proves the independence of the fixed point from γ, for any $\gamma \in J$.

Remark 6.5 Another parameter which entered into the renormalization map is the cut-off function χ. The fixed point coupling ν_* depends on χ already at $O(\varepsilon)$, as seen from eq. (6.10), because a depends on χ. That λ_* is χ-independent at $O(\varepsilon)$ is in agreement with the usual lore that the beta-functions for near-marginal couplings and the corresponding fixed-point coupling should not depend on the UV regularization scheme at the first non-trivial order. In higher orders in ε we expect that all couplings will acquire χ dependence. So, in contrast with the γ-independence, the fixed point does depend on χ. In spite of this, we expect on physical grounds that the critical exponents (i.e. eigenvalues of the renormalization map linearized near the fixed point) should be χ-independent. Showing this rigorously is one of the open problems for the future (see section 8).

6.5 Analyticity

In view of the Complex Abstract Lemma 6.3, and of the complex version of the Key Lemma, stated and proved in section 7, see Lemma 7.1, it is easy to show that the fixed point of theorem 6.1 can be extended to an analytic function of ε in a small neighborhood of the origin. More precisely, we get the following:

Theorem 6.2 (Analytic Fixed Point Theorem) *There exists a $\gamma_0 \geqslant 2$ and a positive continuous function $\varepsilon_0(\gamma)$ defined for $\gamma \geqslant \gamma_0$ such that for each $\gamma \geqslant \gamma_0$ and $\varepsilon \in \{z \in \mathbb{C} : |z| \leqslant \varepsilon_0(\gamma)\} \equiv \mathbb{E}_0$ the fixed point equation (5.30) has a solution, analytic in ε, extending the one of theorem 6.1. For any $\varepsilon \in \mathbb{E}_0$, such a solution is the unique solution of the fixed point equation in the complex neighborhood defined by the analogue of (6.43) with $|\varepsilon|$ replacing ε.*

Proof. We let $\delta = h|\varepsilon|$, with h the same as in (6.36). By proceeding as in the proof of theorem 6.1, with $\bar{D}$ defined as in (7.3), we find that F is a contraction on $\mathbb{Y}_0$ for each $\varepsilon \in \mathbb{E}_0 \setminus \{0\}$. Moreover, by Lemma 7.1, F is analytic in ε on the punctured disk $\overline{\mathbb{Y}}_0(\gamma) \equiv \cup_{0<|\varepsilon|\leqslant\varepsilon_0(\gamma)} \mathbb{Y}_0$, and so is y_0. Therefore, $y_n \equiv F^n(y_0)$ is analytic in ε on $\overline{\mathbb{Y}}_0(\gamma)$. Since F is a contraction, y_n converges to a fixed point, call it $y_*(\varepsilon)$,as $n \to \infty$, for any $\varepsilon \in \mathbb{E}_0 \setminus \{0\}$; for any such ε, $y_*(\varepsilon)$ is the unique solution of the fixed point equation in $\mathbb{Y}_0$. By Vitali's theorem on the convergence of sequences of analytic functions, $y_*(\varepsilon)$ is holomorphic in ε on $\overline{\mathbb{Y}}_0(\gamma)$ (monodromy follows from the uniqueness of the solution to the fixed point equation in $\mathbb{Y}_0$). Note that $\lim_{\varepsilon\to 0} \varepsilon y_*(\varepsilon) = 0$; therefore, by Riemann's theorem on removable singularities, $y_*(\varepsilon)$ can be extended to an analytic function of ε on the complex disk of radius $\varepsilon_0(\gamma)$ by letting $y_*(0) = 0$.

[32] For $\bar{\gamma} \geqslant (\bar{\gamma}_0)^3$, we have $\bar{\gamma}^{1/3} \in X$ by (*ii*), and then $\bar{\gamma}^{2/3} \in X$ by (*iii*). So both endpoints of J are in X. Also, if $\gamma_1, \gamma_2 \in J$, then both $\gamma_1^{1/2}, \gamma_2^{1/2} \in X$ by (*ii*) and hence the geometric mean $(\gamma_1\gamma_2)^{1/2} \in X$ by (*iii*). Applying this last statement recursively starting from the endpoints of J, we obtain that X is dense in J.

This result has various consequences. One clear consequence is that since the fixed point is analytic around $\varepsilon = 0$, it has a convergent power series expansion around this point. This is just the perturbative ε-expansion discussed at the level of formal power series in appendix I which is therefore convergent. Another consequence is that the fixed points with real $\varepsilon > 0$ analytically continue to the fixed points with $\varepsilon < 0$. For negative real ε, the quartic interaction is an irrelevant perturbation of the gaussian fixed point (at the linearized level). Thus, the $\varepsilon < 0$ fixed points should be interpreted as UV fixed points: one can RG-flow from them to the gaussian theory, not the other way around. We expect analyticity to be valid also in the long-range Gross-Neveu model of [27] (see the introduction), and in other similar models. See also appendix J for an alternative proof of fixed point analyticity via the tree expansion.

7 Proof of Key lemma

Here we finally prove the Key lemma that, as seen above, is the crucial ingredient for showing the existence and uniqueness of the nontrivial RG fixed point. Rather than proving the Key lemma in the formulation of section 6.1, here we state and prove a generalization of the lemma with complex ε, which is the version used in section 6.5 in the discussion on the analyticity of the fixed point. This does not create any additional complications in the proof.

Let us start by observing that both the fluctuation propagator (2.10) and the rescaling factor $\gamma^{-[\psi]}$ in (2.16) depend analytically on ε. So each individual term $R_\ell^{\ell_1,\dots,\ell_n}$ is analytic in ε, and the sum (5.24) will be analytic when convergent. Let T be a compact subset of the half-plane (see eq. (2.8))

$$T \subset \{\varepsilon \in \mathbb{C} : \operatorname{Re}\varepsilon < d/6\}. \tag{7.1}$$

By Lemma D.3, the constant $C_{\rm GH}$ is uniformly bounded for $\varepsilon \in T$. As a result the multi-linear maps $S_l^{\ell_1,\dots,\ell_n}$ will satisfy estimates (5.38) with uniform (T-dependent) constants for $\varepsilon \in T$. The action of dilatation for complex ε is still given by (5.20), where $D_l = D_l(\varepsilon)$ are complex. We have to replace $D_l \to \operatorname{Re} D_l$ in the norm bounds (5.37) for dilatation, which become

$$\|DH_{\ell,p}\|_w = \gamma^{-\operatorname{Re} D_l - p}\|H_{\ell,p}\|_{w(\cdot/\gamma)} \leqslant \gamma^{-\operatorname{Re} D_l - p}\|H_{\ell,p}\|_w \qquad (\varepsilon \in \mathbb{C}) \tag{7.2}$$

The criterion for irrelevance becomes $\operatorname{Re} D_l - p > 0$. The same replacement has to be done in the right-hand-sides of the estimates for multilinear maps $R_\ell^{\ell_1,\dots,\ell_n}$ in section 5.6, see eqs. (5.34) and (5.35).

The parameter $\bar{D}$ from section 6.1 is redefined for complex $\varepsilon \in T$ as

$$\bar{D} = \bar{D}(T) = \frac{1}{2}\min_{\varepsilon\in T}\{\operatorname{Re} D_2(\varepsilon) + 2, \operatorname{Re} D_4(\varepsilon) + 1, \operatorname{Re} D_6(\varepsilon)\}. \tag{7.3}$$

Note that $\bar{D} > 0$ by assumptions on T. We can now state the generalization of Lemma 6.1 to $\varepsilon \in \mathbb{C}$.

JHEP01(2021)026

Lemma 7.1 (Complex Key Lemma) *Choose $d \in \{1, 2, 3\}$, cutoff χ, $N \geqslant 4$, and a compact set $T \subset \mathbb{C}$ satisfying (7.1). There exists $\gamma_{\rm key} \geqslant 2$ and positive continuous functions (6.3) on $\gamma \geqslant \gamma_{\rm key}$, with the following property. Take any $\gamma \geqslant \gamma_{\rm key}$, any $0 < \delta \leqslant \delta_0(\gamma)$ and any sequence $y = (\nu, \lambda, u)$ satisfying $\|y\|_{Y(\gamma,\delta)} \leqslant 1$, and apply to it the renormalization map $R(\varepsilon, \gamma)$ with any $\varepsilon \in T$. Then the functions $e^{(0)}_\nu, e^{(0)}_\lambda, e_u$ in eq. (5.30) and their derivatives satisfy the bounds (6.5) and (6.6) uniformly in $\varepsilon \in T$. These functions are analytic in ε, being given by convergent series consisting of analytic terms.*

The proof of the Complex Key Lemma is presented in the next subsections, distinguishing various subcases. For instance, in order to prove that $\|e_u\|_{B(\gamma,\delta)} \leqslant \gamma^{-\bar{D}}$, recalling the definition (6.1) of the norm, we will seperately prove that $\|(e_u)_\ell\|_w \leqslant A(\gamma)\delta^{k(\ell)}$ for all $\ell \geqslant 8$, that $\|(e_u)_{6\rm R}\|_w \leqslant A^{\rm R}_2(\gamma)\delta^3$, $\|(e_u)_{4\rm R}\|_w \leqslant A^{\rm R}_1(\gamma)\delta^2$, and $\|(e_u)_{2\rm R}\|_w \leqslant A^{\rm R}_0(\gamma)\delta^2$. For ease of notation, we will drop the dependence on γ from the constants $A(\gamma), A_0(\gamma)$, etc, and simply denote them by A, A_0, etc. Similarly for $Y(\gamma, \delta)$ and $B(\gamma, \delta)$, to be denoted by Y and B, respectively.

7.1 Case $\ell \geqslant 8$

We start from the bound on $\|(e_u(y))_\ell\|_w$ with $\ell \geqslant 8$. From the definitions, see (5.29) and (5.30), we have

$$(e_u(y))_\ell = \sum_{(\ell_i)_1^n} R^{\ell_1,\ldots,\ell_n}_\ell (H_{\ell_1}, \ldots, H_{\ell_n}). \tag{7.4}$$

Using bounds (5.34) and (5.35) on $R^{\ell_1,\ldots,\ell_n}_\ell$ collected we find that

$$\|(e_u(y))_\ell\|_w \leqslant \gamma^{-\operatorname{Re} D_\ell}\|u_\ell\|_w + \gamma^{-\operatorname{Re} D_\ell} \sum_{(\ell_i)_1^n \neq (\ell)} \rho_l[(\ell_i)_1^n], \tag{7.5}$$

where we denoted

$$\rho_l[(\ell_i)_1^n] = \rho_l(H_{\ell_1}, \ldots, H_{\ell_n}), \tag{7.6}$$

and $\rho_l(H_{\ell_1}, \ldots, H_{\ell_n})$ is given in eq. (5.36). Here H_{ℓ_i} should be interpreted as equal to: ν, if $\ell_i = 2\rm L$; λ, if $\ell_i = 4\rm L$; $\mathfrak{X}_*$, if $\ell_i = 6\rm SL$; u_{ℓ_i}, otherwise. Recall that $\rho_l = 0$ unless $\sum_i |\ell_i| \geqslant l + 2(n-1)$.

By using the assumption $\|y\|_{Y(\gamma,\delta)} \leqslant 1$ of Key lemma, writing in full the meaning of this condition (recall the definition of $\|y\|_{Y(\gamma,\delta)}$, eq. (4.19)), we find:

$$\begin{aligned}
\|H_{2\rm L}\|_w + \|H_{2\rm R}\|_w &\leqslant A_0\delta + A^{\rm R}_0\delta^2 =: b_0, \\
\|H_{4\rm L}\|_w + \|H_{4\rm R}\|_w &\leqslant A_0\delta + A^{\rm R}_1\delta^2 =: b_1, \\
\|H_{6\rm SL}\|_w + \|H_{6\rm R}\|_w &\leqslant C_{\gamma 3}A_0^2\delta^2 + A^R_2\delta^3 =: b_2, \\
\|H_\ell\|_w &\leqslant A\delta^{k(\ell)} =: b_{k(\ell)}, \qquad \text{if} \qquad \ell \geqslant 8.
\end{aligned} \tag{7.7}$$

It will be convenient to arrange so that

$$b_k \leqslant A\delta^{\max\{k,1\}}, \quad k \geqslant 0. \tag{7.8}$$

For $k \geqslant 3$ this is true as an equality by the definition of b_k. To have this for $k = 0, 1, 2$ as well, we will assume (we will see later how to satisfy simultaneously all ♠-constraints):

$$(\spadesuit) \quad 2\max(A_0, A_0^R\delta_0, A_1^R\delta_0, C_{\gamma 3}A_0^2 + A_2^R\delta_0) \leqslant A. \tag{7.9}$$

Using these bounds in (7.5) we find that:

$$\|(e_u(y))_\ell\|_w \leqslant \gamma^{-\operatorname{Re} D_\ell}[A\delta^{k(\ell)} + \Delta^{(1)}_{k(\ell)} + \Delta^{(2)}_{k(\ell)}], \tag{7.10}$$

where we defined [here $C = C_2^2$]:

$$\begin{aligned}
\Delta_k^{(1)} &= \sum_{k'=k+1}^{\infty} C^{k'+1} b_{k'}, \\
\Delta_k^{(2)} &= \sum_{(k_i)_{i=1}^n, n\geqslant 2} F_k[(k_i)_1^n], && (7.11)\\
F_k[(k_i)_1^n] &= \begin{cases} C_\gamma^{n-1}\prod_{i=1}^n C^{k_i+1}b_{k_i} & \text{if} \quad \sum_i k_i \geqslant k, \\ 0 & \text{otherwise}. \end{cases} && (7.12)
\end{aligned}$$

We will estimate these sums with the help of the following lemma, imposing assumptions (7.13) which we will arrange in the end by choosing δ_0 and A appropriately. For the proof see appendix F.

Lemma 7.2 *Suppose the nonnegative constants C_γ, C, δ, A satisfy*

$$(\spadesuit) \quad C\delta \leqslant 1/4, \qquad C_\gamma CA\delta \leqslant 1/2, \qquad C_\gamma CA \leqslant 1/2, \tag{7.13}$$

and that $0 \leqslant b_k \leqslant A\delta^{\max\{k,1\}}$ for all $k \geqslant 0$. Then $\Delta_k^{(1)}$ and $\Delta_k^{(2)}$ defined in terms of C, C_γ, b_k by (7.11), (7.11) satisfy

$$\begin{aligned}
\Delta_k^{(1)} &\leqslant A\delta^{k+1}(2C^{k+2}), && (7.14)\\
\Delta_k^{(2)} &\leqslant A\delta^{\max\{k,2\}} \cdot \begin{cases} C_0 = 4C + 8C^2 + 16C^3 & \textit{if } k = 0{,}1 \\ 2(2C)^{k+1} & \textit{if } k \geqslant 2, \end{cases} && (7.15)
\end{aligned}$$

Using (7.14), (7.15) in (7.10), and recalling that we are assuming $\ell \geqslant 8$ (so that $k(\ell) \geqslant 3$), we find

$$\|(e_u(y))_\ell\|_w \leqslant \gamma^{-\operatorname{Re} D_\ell} A\delta^{k(\ell)}[1 + 2C^{k(\ell)+2}\delta + 2(2C)^{k(\ell)+1}], \tag{7.16}$$

It follows that

$$\|(e_u(y))_\ell\|_w \leqslant \gamma^{-\bar{D}} A\delta^{k(\ell)} \qquad (\ell \geqslant 8), \tag{7.17}$$

as long as we impose

$$1 + C^{k(\ell)+1} + 2(2C)^{k(\ell)+1} \leqslant \gamma^{\operatorname{Re} D_\ell - \bar{D}}. \tag{7.18}$$

Given the form of this inequality, it is sufficient to check that it holds for $\ell = 8$, and that the l.h.s. grows slower than r.h.s. as $\ell \to \ell + 2$, which amounts to two requirements:

$$(\spadesuit) \quad 1 + C^4 + 2(2C)^4 \leqslant \gamma^{\operatorname{Re} D_8 - \bar{D}}, \qquad \max(1, C, 2C) \leqslant \gamma^{d/2 - \operatorname{Re}\varepsilon}. \tag{7.19}$$

Next let us estimate derivatives. Consider a vector $\delta y = (\delta\nu, \delta\lambda, \delta u)$ satisfying $\|\delta y\|_Y \leqslant 1$. Consider also a trimmed coupling sequence δH_ℓ which contains the couplings in δy and, in addition, the coupling $\delta H_{6\mathrm{SL}}$ corresponding to the variation of $\mathfrak{X}_*$. We have

$$\nabla_y (e_u(y))_\ell \delta y = \sum_{(\ell_i)_1^n} \sum_{i=1}^n R_\ell^{\ell_1,\dots,\ell_n}(H_{\ell_1},\dots,\delta H_{\ell_i},\dots,H_{\ell_n}), \tag{7.20}$$

and thus

$$\|\nabla_y (e_u(y))_\ell \delta y\|_w \leqslant \gamma^{-\operatorname{Re} D_\ell}\|\delta H_\ell\|_w + \gamma^{-\operatorname{Re} D_\ell} \sum_{(\ell_i)_1^n \neq (\ell)} \sum_{i=1}^n \rho_l(H_{\ell_1},\dots,\delta H_{\ell_i},\dots,H_{\ell_n}) \tag{7.21}$$

Note that $\|\delta H_{6\mathrm{SL}}\|_w \leqslant 2C_{3\gamma} A_0^2 \delta^2$. We will increase $C_{3\gamma}$ by factor 2. Then all couplings δH_ℓ satisfies the same bounds as the bounds on couplings H_ℓ used to estimate $\|(e_u(y))_\ell\|_w$. It follows that the functions ρ_l in the r.h.s. of (7.21) can be estimated in exactly the same way. This gives an estimate of the same form as (7.10), namely

$$\|\nabla_y (e_u(y))_\ell \delta y\|_w \leqslant \gamma^{-\operatorname{Re} D_\ell}[A\delta^{k(\ell)} + \Delta_{k(\ell)}^{(1)} + \tilde{\Delta}_{k(\ell)}^{(2)}], \tag{7.22}$$

where $\tilde{\Delta}_k^{(2)}$ differs from $\Delta_k^{(2)}$ in that $F_k[(k_i)]$ is replaced by

$$\tilde{F}_k[(k_i)_1^n] = nF_k[(k_i)_1^n], \tag{7.23}$$

where the factor n accounts for the sum $\sum_{i=1}^n$ in (7.21). We will increase C_γ in (7.12) by 2 to absorb this factor (note $n \leqslant 2^{n-1}$), so that both $\tilde{F}_k$ and F_k can be considered to satisfy the same bound (7.12).

Then, under the same assumptions that (7.17) was obtained, we will have

$$\|\nabla_y (e_u(y))_\ell \delta y\|_w \leqslant \gamma^{-\bar{D}} A\delta^{k(\ell)} \qquad (\ell \geqslant 8). \tag{7.24}$$

Taking into account the assumed bounds on couplings δy, this inequality is precisely what is asserted in the last line of (6.6) concerning the part of e_u with $\ell \geqslant 8$.

Incidentally, convergence of the series (7.21) also proves that the functions $e_u(y)$ are in fact Frechet differentiable.

The shown method of bounding derivatives is general and will apply to all the other functions that we still have to consider, i.e. $(e_u)_{2\mathrm{R}}, (e_u)_{4R}, (e_u)_{6\mathrm{R}}, e_\nu^{(0)}, e_\lambda^{(0)}$. They are all given by sums of multilinear operators applied to the sequence H_ℓ, and will be estimated using the basic bound (5.36). Whenever we manage to bound such a function by an X, the shown method will naturally bound its u-derivative by the same X, while its ν, λ derivatives by $X/(A_0\delta)$. Note that all bounds (6.6) are of precisely such a form. So we no longer need to discuss derivative bounds, but can focus on estimating the functions themselves.

7.2 Case $\ell = 6\mathrm{R}$

From the definitions (see (5.29) and (5.30) and the third of (5.27)) and the bounds on $R_{6\mathrm{R}}^{\ell_1,\dots,\ell_n}$ we find that

$$\|(e_u(y))_{6\mathrm{R}}\|_w \leqslant \gamma^{-\operatorname{Re} D_6}\|u_{6\mathrm{R}}\|_w + \gamma^{-\operatorname{Re} D_6} \sum_{(\ell_i)_1^n \neq (6\mathrm{SL}),(6\mathrm{R}),(4\mathrm{L},4L)} \rho_6[(\ell_i)_1^n]. \tag{7.25}$$

By repeating a discussion analogous to that of section 7.1, we get the analogue of (7.10), namely

$$\|(e_u(y))_{6\mathrm{R}}\|_w \leqslant \gamma^{-\operatorname{Re} D_6}\left[A_2^{\mathrm{R}}\delta^3 + \Delta_2^{(1)} + \Delta_{2;6\mathrm{R}}^{(2)}\right], \tag{7.26}$$

where $\Delta_{2;6\mathrm{R}}^{(2)}$ is defined analogously to $\Delta_2^{(2)}$, modulo the fact that the contribution from the sequence $(k_i)_{i=1}^n = (1,1)$ is now proportional to $b_1 b_1^{\mathrm{R}}$, with $b_1^{\mathrm{R}} = A_1^{\mathrm{R}}\delta^2$, rather than to b_1^2 (this comes from the constraint $(\ell_i)_{i=1}^n \neq (4\mathrm{L}, 4\mathrm{L})$ in (7.25)):

$$\Delta_{2;6\mathrm{R}}^{(2)} = 2C_\gamma C^4 b_1 b_1^{\mathrm{R}} + \sum_{(k_i)_1^n \neq (1,1)}^{n\geqslant 2} F_2[(k_i)_1^n]. \tag{7.27}$$

It is convenient to define, for any sequence $\varkappa = (k_i)_1^n$,

$$F_{\mathrm{ext}}[\varkappa] = \sum_{\varkappa' \text{ :extends } \varkappa \text{ by } \geqslant 0 \text{ zeros}} F[\varkappa']. \tag{7.28}$$

Using this definition, we split the second term in the r.h.s. of (7.27) into (a) the contributions of sequences $(1,1,0)$, $(2,0)$, their permutations and extensions by zero and (b) sequences with $\sum k_i \geqslant 3$ which form $\Delta_3^{(2)}$. We get

$$\Delta_{2;6\mathrm{R}}^{(2)} = 2C_\gamma C^4 b_1 b_1^{\mathrm{R}} + 2F_{\mathrm{ext}}[(2,0)] + 3F_{\mathrm{ext}}[(1,1,0)] + \Delta_3^{(2)}. \tag{7.29}$$

It is shown in appendix F, see eq. (F.11), that, in the assumptions of Lemma 7.2,

$$F_{\mathrm{ext}}[(k_i)_1^n] \leqslant 4C^{k+1} A\delta^{k+m}, \tag{7.30}$$

where $k = \sum k_i$ and m is the number of zeros in the sequence $(k_i)_1^n$.

Using (7.15) for $\Delta_3^{(2)}$, the basic estimates $b_1^R \leqslant A_1^R\delta^2$, $C_\gamma CA \leqslant 1/2$, and (7.30) we get

$$\Delta_{2;6\mathrm{R}}^{(2)} \leqslant (C^3 A_1^R + [8C^3 + 12C^3 + 2(2C)^4]A)\delta^3, \tag{7.31}$$

so that

$$\|(e_u(y))_{6\mathrm{R}}\|_w \leqslant \gamma^{-\operatorname{Re} D_6}\delta^3\left\{A_2^{\mathrm{R}} + 2C^4 A + C^3 A_1^R + [20C^3 + 2(2C)^4]A\right\}, \tag{7.32}$$

which is smaller than $\gamma^{-\bar{D}} A_2^{\mathrm{R}}\delta^3$, provided that

$$(\spadesuit) \qquad A_2^{\mathrm{R}} + 2C^4 A + C^3 A_1^R + [20C^3 + 2(2C)^4]A \leqslant \gamma^{\operatorname{Re} D_6 - \bar{D}} A_2^{\mathrm{R}}. \tag{7.33}$$

7.3 Case $\ell = 4\mathrm{R}$

From the definitions and the bounds on $R_{4\mathrm{R}}^{\ell_1,\ldots,\ell_n}$ we find that

$$\|(e_u(y))_{4\mathrm{R}}\|_w \leqslant \gamma^{-\operatorname{Re} D_4 - 1}\|u_{4\mathrm{R}}\|_w + \gamma^{-\operatorname{Re} D_4}\sum_{(\ell_i)_1^n \neq (4\mathrm{L}),(4\mathrm{R})} \rho_4[(\ell_i)_1^n], \tag{7.34}$$

(the condition $(\ell_i)_1^n \neq (4\mathrm{L})$ comes from the fact that $R_{4\mathrm{R}}^{4L}$ is identically zero, see the first of (5.27) and the definition of $T_{4\mathrm{R}}^4$ in section 5.2.2; note in particular that, by construction, $T_{4\mathrm{R}}^{4,0}$ annihilates the local quartic kernel associated with $H_{4\mathrm{L}}$) so that

$$\|(e_u(y))_{4\mathrm{R}}\|_w \leqslant \gamma^{-\operatorname{Re} D_4 - 1}\left[A_1^{\mathrm{R}}\delta^2 + \gamma\Delta_1^{(1)} + \gamma\Delta_1^{(2)}\right], \tag{7.35}$$

which gives

$$\|(e_u(y))_{4\mathrm{R}}\|_w \leqslant \gamma^{-\operatorname{Re} D_4 - 1}\delta^2\left[A_1^{\mathrm{R}} + \gamma A(2C^3) + \gamma C_0 A\right]. \tag{7.36}$$

This is smaller than $\gamma^{-\bar{D}} A_1^{\mathrm{R}}\delta^2$, provided that

$$(\spadesuit)\qquad A_1^{\mathrm{R}} + \gamma A(2C^3 + C_0) \leqslant \gamma^{\operatorname{Re} D_4 + 1 - \bar{D}} A_1^{\mathrm{R}}. \tag{7.37}$$

7.4 Case $\ell = 2\mathrm{R}$

From the definitions and the bounds on $R_{2\mathrm{R}}^{\ell_1,\dots,\ell_n}$ we find that

$$\|(e_u(y))_{2\mathrm{R}}\|_w \leqslant \gamma^{-\operatorname{Re} D_2 - 2}\|u_{2\mathrm{R}}\|_w + \gamma^{-\operatorname{Re} D_2} \sum_{(\ell_i)_1^n \neq (2\mathrm{L}),(2R),(4\mathrm{L})} \rho_2[(\ell_i)_1^n], \tag{7.38}$$

(the conditions $(\ell_i)_1^n \neq (2\mathrm{L}),(4\mathrm{L})$ come from the fact that $R_{2R}^{2\mathrm{L}}$ and $R_{2\mathrm{R}}^{4\mathrm{L}}$ are identically zero, see the first of (5.27) and the definition of $T_{2\mathrm{R}}^2$ in section 5.2.2; note in particular that, by construction, $T_{2\mathrm{R}}^{2,0}$ annihilates the local quadratic kernels associated with $H_{2\mathrm{L}}$ and with $S_2^{4\mathrm{L}}(H_{4\mathrm{L}})$) so that

$$\|(e_u(y))_{2\mathrm{R}}\|_w \leqslant \gamma^{-\operatorname{Re} D_2 - 2}\left[A_0^{\mathrm{R}}\delta^2 + \gamma^2\Delta_{0;2\mathrm{R}}^{(1)} + \gamma^2\Delta_0^{(2)}\right], \tag{7.39}$$

where

$$\Delta_{0;2\mathrm{R}}^{(1)} = C^2 b_1^{\mathrm{R}} + \Delta_1^{(1)} \leqslant C^2 A_1^{\mathrm{R}}\delta^2 + A\delta^2(2C_3^3). \tag{7.40}$$

Therefore,

$$\|(e_u(y))_{2\mathrm{R}}\|_w \leqslant \gamma^{-\operatorname{Re} D_2 - 2}\delta^2\left[A_0^{\mathrm{R}} + \gamma^2(C^2 A_1^{\mathrm{R}} + 2C^3 A + C_0 A)\right]. \tag{7.41}$$

This is smaller than $\gamma^{-\bar{D}} A_0^{\mathrm{R}}\delta^2$, provided that

$$(\spadesuit)\qquad A_0^{\mathrm{R}} + \gamma^2(C^2 A_1^{\mathrm{R}} + 2C^3 A + C_0 A) \leqslant \gamma^{\operatorname{Re} D_2 + 2 - \bar{D}} A_0^{\mathrm{R}}. \tag{7.42}$$

7.5 $e_\nu^{(0)}$

From the definitions and the bounds on $R_{2\mathrm{L}}^{\ell_1,\dots,\ell_n}$ we find that

$$|e_\nu^{(0)}(y)| \leqslant \gamma^{-\operatorname{Re} D_2} \sum_{(\ell_i)_1^n \neq (2\mathrm{L}),(2\mathrm{R}),(4\mathrm{L})} \rho_2[(\ell_i)_1^n], \tag{7.43}$$

(the conditions $(\ell_i)_1^n \neq (2\mathrm{L}),(4\mathrm{L})$ come directly from the definition of $e_\nu^{(0)}$, see (5.29) and (5.30), while $(\ell_i)_1^n \neq (2\mathrm{R})$ comes from the fact that $R_{2\mathrm{L}}^{2\mathrm{R}}$ is identically zero, see the first of (5.27) and the definition of $T_{2\mathrm{L}}^2$ in section 5.2.2; note in particular that, by construction, $T_{2\mathrm{L}}^{2,0}$ annihilates the nonlocal quadratic kernel associated with $H_{2\mathrm{R}}$) so that

$$|e_\nu^{(0)}(y)| \leqslant \gamma^{-\operatorname{Re} D_2}[\Delta_{0;2R}^{(1)} + \Delta_0^{(2)}], \tag{7.44}$$

which gives

$$|e_\nu^{(0)}(y)| \leqslant \gamma^{-\operatorname{Re} D_2}\delta^2\left[C^2 A_1^{\mathrm{R}} + 2C^3 A + C_0 A\right]. \tag{7.45}$$

We thus get the first of (6.5), with

$$E_0 = \gamma^{-\operatorname{Re} D_2}\left[C^2 A_1^{\mathrm{R}} + 2C^3 A + C_0 A\right]. \tag{7.46}$$

7.6 $e_\lambda^{(0)}$

From the definitions and the bounds on $R_{4\mathrm{L}}^{\ell_1,\dots,\ell_n}$ we find that

$$|e_\lambda^{(0)}(y)| \leqslant \gamma^{-\operatorname{Re} D_4} \sum_{\substack{(\ell_i)_1^n \neq (4\mathrm{L}),(4\mathrm{R}),(6\mathrm{SL}),\\ (4\mathrm{L},4\mathrm{L}),(4\mathrm{L},2\mathrm{L}),(4\mathrm{R},2\mathrm{L})}} \rho_4[(\ell_i)_1^n], \tag{7.47}$$

(the conditions $(\ell_i)_1^n \neq (4\mathrm{L}),(6\mathrm{SL}),(4\mathrm{L},4\mathrm{L})$ come directly from the definition of $e_\lambda^{(0)}$, see (5.29) and (5.30), while $(\ell_i)_1^n \neq (4\mathrm{R}),(4\mathrm{L},2\mathrm{L}),(4\mathrm{R},2\mathrm{L})$ come from the fact that $R_{4\mathrm{L}}^{4\mathrm{R}}$, $R_{4\mathrm{L}}^{4\mathrm{L},2\mathrm{L}}$ and $R_{4\mathrm{L}}^{4\mathrm{R},2L}$ are identically zero, see the first of (5.27) and the definition of $T_{4\mathrm{L}}^4$ in section 5.2.2; note in particular that, by construction, $T_{4\mathrm{L}}^{4,0}$ annihilates the nonlocal quartic kernels associated with $H_{4\mathrm{R}}$, $S_4^{4\mathrm{L},2\mathrm{L}}(H_{4\mathrm{L}},H_{2\mathrm{L}})$ and $S_4^{4\mathrm{R},2\mathrm{L}}(H_{4\mathrm{R}},H_{2\mathrm{L}})$) so that

$$|e_\lambda^{(0)}(y)| \leqslant \gamma^{-\operatorname{Re} D_4}[\Delta_{1;\lambda}^{(1)} + \Delta_{1;\lambda}^{(2)}], \tag{7.48}$$

where

$$\Delta_{1;\lambda}^{(1)} = C^3 b_2^{\mathrm{R}} + \Delta_2^{(1)} \leqslant C^3 A_2^{\mathrm{R}} \delta^3 + A\delta^3(2C^4), \tag{7.49}$$

$$\Delta_{1;\lambda}^{(2)} = 2C_\gamma C^4 b_1 b_1^{\mathrm{R}} + \sum_{(k_i)_1^n \neq (1,1)}^{n\geqslant 2} F_2[(k_i)_1^n]. \tag{7.50}$$

The sequences with $\sum k_i = 1$ such as $(1,0)$, $(1,0,0)$, etc are excluded from the second term because b_0 insertions then happen on the external legs of a quartic interaction, and they give rise to a vertex with a vanishing local part. We see that $\Delta_{1;\lambda}^{(2)}$ is identical to (7.27) and therefore satisfies the same bound (7.31)

$$\Delta_{1;\lambda}^{(2)} \leqslant (C^3 A_1^R + [8C^3 + 12C^3 + 2(2C)^4]A)\delta^3, \tag{7.51}$$

Therefore we get

$$|e_\lambda^{(0)}(y)| \leqslant \gamma^{-\operatorname{Re} D_4} \delta^3 \left[C^3 A_2^{\mathrm{R}} + 2C^4 A + C^3 A_1^R + [20C^3 + 2(2C)^4]A\right]. \tag{7.52}$$

We thus get the second equation of (6.5), with

$$E_1 = \gamma^{-\operatorname{Re} D_4} \left[C^3 A_2^{\mathrm{R}} + 2C^4 A + C^3 A_1^R + [20C^3 + 2(2C)^4]A\right]. \tag{7.53}$$

7.7 Possibility of all choices

Finally, we need to show that all the ♠-constraints above can be satisfied consistently: eqs. (7.9), (7.13), (7.19), (7.33), (7.37), (7.42). To write them in a more manageable form, let us replace al γ-independent constants in the l.h.s. of the ♠-constraints by their maximum $\bar{C}$ (Recall that C_0 was fixed in terms of C in (7.15)). Also let $\bar{C}_\gamma = \max(C_\gamma, C_{\gamma 3})$. Finally let Z be the minimal of the exponents of γ in the r.h.s. of (7.19), (7.33), (7.37), (7.42) over $\varepsilon \in T$:

$$Z = \min_{\varepsilon \in T} \left\{\operatorname{Re} D_8 - \bar{D}, d/2 - \operatorname{Re}\varepsilon, \operatorname{Re} D_6 - \bar{D}, \operatorname{Re} D_4 + 1 - \bar{D}, \operatorname{Re} D_2 + 2 - \bar{D}\right\}. \tag{7.54}$$

Crucially $Z > 0$ by the assumption on T and the definition of $\bar{D}$. We then get the following list of constraints which, if satisfied, imply the ♠-constraints for any $\varepsilon \in T$:

$$\bar{C} \leqslant \gamma^Z, \tag{7.55}$$

$$\bar{C}\delta_0 \leqslant 1, \quad \bar{C}\bar{C}_\gamma A \leqslant 1, \tag{7.56}$$

$$\max(A_0, A_0^R\delta_0, A_1^R\delta_0, \bar{C}_\gamma A_0^2 + A_2^R\delta_0) \leqslant A/2, \tag{7.57}$$

$$A_2^R + \bar{C}(A_1^R + A) \leqslant \gamma^Z A_2^R, \qquad A_1^R + \bar{C}\gamma A \leqslant \gamma^Z A_1^R, \qquad A_0^R + \bar{C}\gamma^2(A_1^R + A) \leqslant \gamma^Z A_0^R. \tag{7.58}$$

The only remaining varying parameter is γ. We should now choose γ_{key} and $\delta_0, A_0, \{A_k^R\}_{k=0,1,2}, A, E_0, E_1$, which are γ-dependent and positive, so that all these constraints hold for $\gamma \geqslant \gamma_{\text{key}}$.

We can satisfy the first two lines taking γ large, then A and δ_0 small (in this order, because $\bar{C}_\gamma$depends on γ). The remaining constraints are a bit more subtle because A and A_k^R occur both in the l.h.s. and in the r.h.s. To satisfy (7.58) we will require:

$$A_1^R, A \leqslant A_2^R, \qquad \gamma A \leqslant A_1^R, \qquad \gamma^2 A, \gamma^2 A_1^R \leqslant A_0^R, \tag{7.59}$$

$$1 + 2\bar{C} \leqslant \gamma^Z, \qquad 1 + 2\bar{C} \leqslant \gamma^Z, \qquad 1 + \bar{C} \leqslant \gamma^Z. \tag{7.60}$$

The last three constraints on γ are of the same type as (7.55). Joining inequalities in (7.59) to (7.57), the resulting set of constraints reduces to:

$$A_0 \leqslant 0.5A, \qquad \bar{C}_\gamma A_0^2 \leqslant 0.25A, \tag{7.61}$$

$$A_2^R \in [A, 0.25\delta_0^{-1}A], \qquad A_1^R \in [\gamma A, 0.5\delta_0^{-1}A], \qquad A_0^R \in [\gamma^2 A, 0.5\delta_0^{-1}A], \tag{7.62}$$

$$A_1^R \leqslant A_2^R, \qquad \gamma^2 A_1^R \leqslant A_0^R. \tag{7.63}$$

Here's then the final order in which all choices have to be made: γ_{key} is chosen as the minimal $\gamma \geqslant 2$ satisfying (7.55) and (7.60). We then pick any $\gamma \geqslant \gamma_{\text{key}}$ and compute the constant $\bar{C}_\gamma$. We then satisfy (7.56) by choosing:

$$A = (\bar{C}\bar{C}_\gamma)^{-1}. \tag{7.64}$$

We then choose A_0 sufficiently small to satisfy (7.61). Finally, we choose

$$\delta_0 = \min(\bar{C}^{-1}, 1/(2\gamma^3)), \tag{7.65}$$

which satisfies (7.56) and at the same time, thanks to $\delta_0 \leqslant 1/(2\gamma^3)$, allows us to choose A_k^R as follows:

$$A_2^R = 0.25\delta_0^{-1}A, \qquad A_1^R = \gamma A, \qquad A_0^R = 0.5\delta_0^{-1}A. \tag{7.66}$$

Then (7.62) is satisfied, and (7.63) holds as well. Key lemma is proved.

8 Discussion and open problems

In this paper, we discussed what is perhaps the simplest theoretical model to study field-theoretic non-Gaussian fixed points, which is amenable to rigorous analysis: **symplectic**

fermions with a long-range kinetic term and local quartic interaction. Our model is translation and rotation invariant, and the structure of the RG equations is quite similar to models with local kinetic term. This makes our model more realistic than, for example, models with hierarchical interactions (see [77] for an introduction).

Our model depends on 3 physical parameters: the number of dimensions d, the number of fermion species N (assuming $\mathrm{Sp}(N)$ invariance), and a parameter ε in the long-range fermion propagator, which controls the relevance of the quartic fermion interaction. For $0 < \varepsilon \ll 1$ this interaction is weakly relevant, and the beta-function equation for the quartic coupling λ takes the forms $\beta_\lambda = -\varepsilon\lambda + \text{const} \cdot (N-8)\lambda^2 + \cdots$. One thus has the right to expect that, for $N \neq 8$, there exists an RG fixed point with $\lambda = O(\varepsilon)$. Our main result (theorem 6.1) establishes the existence of this fixed point rigorously and non-perturbatively.

Although the path towards this rigorous result was somewhat long, most of the ingredients are rather natural. We introduce an infinite-dimensional Banach space of interactions, whose kernels are essentially local (have to decay very fast at point separation). We work with a smooth momentum space cutoff, so that the UV and IR-cutoff fermion propagator decays very fast in position space, and the almost-locality of the interaction is preserved by an RG step.

An essential feature of our model is that a single RG step leads to a convergent effective action (for weak coupling). Intuitively, this property of fermionic models is due to the Pauli principle, or, equivalently, to fermionic signs leading to cancellations between Feynman diagrams. The formal derivation is somewhat delicate, and we review it pedagogically in appendix D. This is standard in the constructive field theory community, but may appear unexpected to the others. A related detail is that exhibiting these fermionic cancellations requires considering a finite RG step with a rescaling parameter $\gamma > 1$. That's what we do in this paper, as opposed to performing continuous RG à la Polchinski's equation (see remark 5.4).

With these ingredients, we show that the Wilsonian RG map is a well-defined nonlinear operator in the Banach space of interactions, and is a contraction (has derivative whose operator norm is less than 1) along the irrelevant directions. The behavior along the mass direction ν and the quartic λ has to be analyzed separately. These directions are both relevant at the linearized level, with λ becoming irrelevant near the approximate one-loop fixed point. Rigorous bounds on error terms show that these statements remain true at the nonlinear level, at weak coupling. The proofs of these results rely just on some elementary combinatorics, geometric series convergence, and chasing γ^{-D_l} factors suppressing the irrelevant interactions. Given one relevant and infinitely many irrelevant directions, the fixed point equation can then be rigorously solved (for ε sufficiently small) via a variant of Newton's method, appealing to the Banach fixed point theorem.

Although our main interest is in ε real and positive, in which case the fixed point can be thought of as the IR fixed point of an RG flow originating at the gaussian theory, our methods apply as long as ε is small and nonzero. E.g. we can also consider $\varepsilon < 0$. In this case the quartic interaction is irrelevant around the gaussian theory, but relevant around the fixed point whose existence we can prove (which in this case is classified as a UV fixed point). We can also consider complex nonzero ε. Although perhaps lacking clear physical

JHEP01(2021)026

meaning, we can use this as a formal device to show that the fixed point is analytic in ε in a punctured disk around the origin, and thus in the whole disk by Riemann's removable singularity theorem (section 6.5). This is a dramatic conclusion, which implies that our fixed points can be obtained via the convergent perturbative ε-expansion around $\varepsilon = 0$.

8.1 Open questions

We will now list many open questions raised by our work. Some of them are theoretical, while others have potential practical applications to numerical calculations of critical exponents.

8.1.1 Extensions to other nonlocal models

It should be relatively easy to extend our results to many other similar models:

- Models with a symmetry group $G \subset \mathrm{Sp}(N)$, which have several mass terms and quartic couplings consistent with this symmetry. One should be able to find a non-perturbative fixed point in a neighborhood of any isolated solution to the one-loop beta-function equations (as long as all quartic directions have eigenvalues $O(\varepsilon)$, the condition which generalizes non-vanishing one-loop beta-function used in this work).

- Models where different species of fermions have different propagator scaling (different ε). This may include models where some fermions ψ have local kinetic terms (and thus a fixed scaling dimension for a given d), while others ψ' are long-range with tunable dimensions, so that the interaction $\psi^2(\psi')^2$ can be made near-marginal.

- Models with a vanishing one-loop beta function, like our model with $N = 8$. As discussed in appendix G, the two-loop beta-function term λ^3 has a nonzero coefficient [78], giving a perturbative fixed point with $\lambda = O(\sqrt{\varepsilon})$. The non-perturbative existence of such a fixed point and its analyticity properties in ε can be understood almost immediately using the tree expansion method described in appendix J, and a contraction argument should also be possible.

- Our model in $d = 4$. Compared to $d \in \{1, 2, 3\}$ treated here, the local term $(\partial\psi)^2$ would be (weakly) relevant for $\varepsilon > 0$. One thus has to treat it on equal footing with the local ψ^2 and $(\psi^2)^2$ terms. One should be able to construct a non-perturbative fixed point for small ε, helped by the fact that the new coupling is quadratic in ψ. This would be the fermionic analogue of the bosonic problem considered in [22].

- Models where the sextic or higher power $(\psi^2)^p$ is near-marginal, i.e. $[\psi] \approx d/(2p)$, $p \geqslant 3$.

More ambitiously, time may be ripe for a "general theory of fermionic fixed points with scale-invariant kinetic terms (local or long-range) and near-marginal local polynomial interactions". One should be able to prove that any such fixed point showing up in perturbative analysis exists non-perturbatively, rather than writing a new paper for each particular model. The main challenge is to choose an efficient notation, and to cleanly separate the algebraic and analytic aspects of the problem.[33] This future general theory should cover all the above examples, as well as fermionic fields transforming in other rotation representations (e.g. spinors [27]), and even non-rotationally invariant (Lifshitz-type) fixed points having anisotropic scaling.

8.1.2 Further properties of the RG fixed points

Here we proved that the RG fixed points exist, and established a few of there basic properties such as γ-independence (modulo some loose ends), and analyticity in ε. Future work should investigate several other interesting properties, such as:

- Uniqueness of the fixed point as an equivalence class of interactions (i.e. uniqueness of solutions of the general fixed point equation; see remark 5.3);

- Dependence of the fixed point on the UV cutoff function χ (see remark 6.5). In spite of this dependence, the critical exponents are expected to be χ-independent. It is instructive to compare the family of long-range models discussed here with one-parameter families of short-range fixed points, such as the Ashkin-Teller model, 6- and 8-vertex models, and interacting dimer models (see e.g. [35, 39]). In the latter case, the deformation parameter is an exactly marginal coupling, which can renormalize along the RG flow, and so the critical exponents depend on the microscopic details, although if one critical exponent is known, others can be expressed via it (the so called weak universality). In our case, ε is not a coupling but a parameter controlling the nonlocal part of the action, so it does not renormalize. Therefore, the situation is similar to the usual universality, and all critical exponents should be universal functions of ε independent of microscopic details such as the UV cutoff χ.[34] It would be interesting to establish this rigorously. See [81] for a classic intuitive discussion of these issues, in the context of local models.

- Critical exponents. These can be defined, most generally, as eigenvalues of the RG transformation linearized around the fixed point (removing the eigenvalues corresponding to the "redundant operators" [81]).[35] From the densities of the corresponding eigenvectors, one should be able to define the "scaling operators", whose correlation functions with respect to the fixed point interaction have exact scale invariance. One can also study correlation functions of simple operators such as ψ^2. While not exactly scale invariant, they should become so at asymptotically long distances.

[33] One may be inspired by how somewhat similar difficulties have been solved for nonlinear stochastic partial differential equations, another problem which involves renormalization [79, 80].

[34] Note in this respect that the IR scaling dimension of ψ is exactly known and equal to its UV dimension $d/4 - \varepsilon/2$. Therefore the exponent η is trivially known as a function of ε. Even for weak universality, all exponents can be found if one exponent is known, making the conclusion that in our situation all exponents are universal functions of ε less surprising.

[35] Sometimes this is equivalently expressed by introducing perturbing "source terms" and studying their beta-functions.

- Full RG trajectory. By this we mean the theory which interpolates between the gaussian fixed point at short distances and the non-gaussian fixed point at long distances (for $\varepsilon > 0$, while for $\varepsilon < 0$ it is the other way around).

8.1.3 Conformal invariance

The RG fixed points constructed here are expected to be conformally invariant, based on the same intuitive arguments as for the long-range bosonic models [82]. Conformal invariance means the invariance of correlation functions of scaling operators (see above) under the finite-dimensional conformal group $\mathrm{SO}(d+1,1)$. For $d = 1$ these are Möbius transformations, and for $d = 2$ the product of holomorphic and antiholomorphic Möbius transformations.[36] This invariance also implies correspondence between correlation functions in infinite volume as studied here, and correlation functions on a sphere of finite radius (which for $d = 1$ is just a circle with periodic boundary conditions), putting the two manifolds in correspondence via the stereographic projection. Such properties are expected to be generally true based on intuitive physics arguments, and it would be very interesting to see how they emerge rigorously in an explicit model such as ours. In particular, this would provide the first rigorous non-gaussian conformal theory in $d = 3$.[37]

Conserved stress tensor operator plays key role in intuitive discussions of conformal invariance of local theories. Our model being nonlocal (long-range), it does not possess a local stress tensor in d dimensions. One way around this difficulty is to represent the nonlocal kinetic term as arising from a local quadratic action in the $(d + 1)$-dimensional Anti-de-Sitter (AdS) space, of which the d-dimensional space is the boundary, where the quadratic, quartic, and all the irrelevant interactions are localized. This construction is useful for intuitive understanding of conformal invariance (as discussed for bosonic models in [82]), and perhaps also for proving it rigorously.[38]

A key property of local conformally invariant theories is the convergent Operator Product Expansion (OPE).[39] Though nonlocal (long-range), our model also should have this property due to the local AdS representation.[40] It would be very interesting to establish this rigorously. This appears somewhat nontrivial due to the fact that the scaling operators, introduced as densities of linearized RG eigenvectors (see above), will not be exactly local but "mildly nonlocal", with kernels of stretched-exponential decay. It is puzzling why this mild nonlocality does not invalidate the usual intuitive arguments for the OPE, which

[36]Because of the nonlocal kinetic term, there will be no invariance under more general holomorphic transformations, unlike in the case of fixed points of fully local models.

[37]In the last 20 years, starting with Smirnov [83], there was significant progress in showing rigorously conformal invariance of various critical observables of specific 2d lattice models (see [84] for review). Many of these models are exactly solvable in infinite volume, and the main challenge is to show conformal invariance of correlators defined in an arbitrary planar region (see e.g. [85] for the 2d Ising model). A key method used in these works is discrete holomorphicity, which is limited to 2d and to specific models, while RG does not play much of a role. The proof of conformal invariance of our RG fixed points will require very different methods, which should work for any d.

[38]On the contrary it is probably hard to make rigorous sense of the Caffarelli-Silvestre construction from [82], where the higher-dimensional ambient space is flat, but it has non-integer dimension.

[39]See [86] for an introduction for physicists, and [87] for a more mathematical one.

[40]Bosonic cousins of our model have been studied via the numerical conformal bootstrap in [88].

treat scaling operators as living at a point. Somehow, the mild nonlocalities of the scaling operators and of the fixed-point interaction should conspire to produce a fully local OPE. Note that this issue is not specific to our model with a long-range kinetic term, as mild nonlocality of scaling operators would be present also for fully local models such as the 3D Ising model.

8.1.4 Relations with analytic regularization

Analytic dependence of our fixed point on ε implies that the critical exponents should also be analytic in ε. Analyticity of the correlation-length and correction-to-scaling exponents ν and ω follows easily from the tree expansion construction (appendix J), as they can be computed by linearizing the analytic right sides of (J.12) near the fixed point.[41] The exponent η is trivial in our model due to the absence of wavefunction renormalization. Higher exponents may have a subtle analytic structure because of degeneracies of linearized RG eigenvalues at $\varepsilon = 0$. Since our model is non-unitary, some higher critical exponents may become complex even for real ε, forming complex-conjugate pairs.[42]

As already mentioned, it would be interesting to show that the critical exponents are χ-independent. Another problem is to prove rigorously that our critical exponents agree with perturbative techniques by which these exponents are computed in theoretical physics. This is especially interesting given that, as we have shown, perturbation theory converges in the problem at hand.

In theoretical physics, higher-order perturbative computations of critical exponents are usually done by working with a bare action containing only the relevant and nearly-marginal couplings. This uses the fact that, due to the renormalizability of the theory at short distances, one can always find an RG trajectory leading to the fixed point from such a UV theory where all irrelevant couplings are set to zero. Furthermore, theoretical physics calculations are greatly simplified by choosing a "mass-independent regularization scheme", which allows to simply set the mass terms to zero. Examples of such schemes are dimensional regularization and analytic regularization with minimal subtraction, which amount to analytically continue Feynman diagrams in ε, dropping the poles. It is universally believed that any scheme, and in particular a mass independent one, should give the same power series in ε for the critical exponents, but to our knowledge this has never been discussed in full rigor.

8.1.5 Increasing the range of ε

Existence proofs of fixed points in this paper work for $|\varepsilon| \leqslant \varepsilon_0(\gamma)$. We have not attempted to evaluate $\varepsilon_0(\gamma)$ explicitly, although it would be straightforward to do this, following step-by-step our arguments. This may be a good exercise for someone wishing to understand our methods in depth. Both of our methods (contraction and the tree expansion) can be

[41] These series have been computed, via another regulator, for the bosonic $O(N_b)$ long-range models in [78] up to three loops. We thanks Dario Benedetti for sharing a Mathematica notebook. Fermionic series should be obtainable by setting $N_b = -N$. These series are not sufficiently long to test our claim that they are convergent in the fermionic case.

[42] This is similar to how some higher Wilson-Fisher critical exponents become complex in $4 - \epsilon$ dimensions [89].

obtimized to enlarge the range of ε where the fixed point is under control. One simple strategy is to increase the number of terms which are computed explicitly, or estimated more carefully than what is currently done. For ε of order 1, one might have to resort to computer-assisted methods.[43]

An interesting feature which might be revealed by such exploration is the cross-over to the short-range universality class. Namely consider the local symplectic fermion model with the bare action (cf (1.1))

$$\int d^d x(\Omega_{ab}\partial_\mu\psi^a\partial^\mu\psi^b + \nu\psi^2 + \lambda\psi^4). \tag{8.1}$$

Some literature concerning such models was cited in the Introduction. This model is expected to flow to a non-gaussian fixed point for $d = 3$ (although, by the usual arguments, not for $d \leqslant 2$). This fixed point is strongly coupled, and we cannot access it using the techniques of this paper.[44] Physicists study such fixed points by the usual Wilson-Fisher ϵ-expansion working formally in $d = 4 - \epsilon$ and then resumming the series at $\epsilon = 1$. [We will use ϵ to denote $4 - d$ as opposed to the long-range parameter ε.] As mentioned in appendix G, ϵ-expansion for these models is perturbatively equivalent to the ϵ-expansion of bosonic $O(N_b)$ models with $N_b = -N$.

So for $d = 3$ we have a family of long-range fixed points studied here whose critical exponents depend on ε, and the fixed point of (8.1) which we will call "short-range". The scaling dimension of ψ is $[\psi]_{\rm LR}(\varepsilon) = d/4 - \varepsilon/2$ in our models, while it is $[\psi]_{\rm SR} = d/2 - 1 + \eta_{\rm SR}/2$ at the fixed point of (8.1). The short-range $\eta_{\rm SR}$ is given by $\eta_{\rm SR} = \epsilon^2 \frac{2+N_b}{2(8+N_b)^2} + O(\epsilon^3)$, $N_b = -N$, with the series which needs to be Borel-resummed at $\epsilon = 1$.[45]

The subsequent discussion applies for any N for which $[\psi]_{\rm SR} < d/4$, as appears to be the case at least for $N = 4$ (see footnote 45). For such N, we will have $[\psi]_{\rm LR}(\varepsilon) = [\psi]_{\rm SR}$ for $\varepsilon = \varepsilon_* = 2(d/4 - [\psi]_{\rm SR}) > 0$. It can then be conjectured that, for $\varepsilon = \varepsilon_*$, the long-range to short-range crossover will take place. Namely, the long-range fixed point at $\varepsilon = \varepsilon_*$ should become identical to the short-range fixed-point plus a non-interacting gaussian theory of an $\mathrm{Sp}(N)$ symplectic fermion ζ of scaling dimension $d - [\psi]_{\rm LR}(\varepsilon_*)$. This would be analogous to the bosonic long-range models, for which such a crossover has been studied since a long time theoretically (starting in [93, 94], reviewed in [95], section 4.3) and is supported by Monte Carlo simulations [96, 97]. The extra gaussian field ζ is expected by the same arguments as in [58, 98] for the bosonic case.[46] Note that the operator $\psi\zeta$ is marginal for $\varepsilon = \varepsilon_*$ (it

[43] Inspired by Lanford's construction of the Feigenbaum fixed point [90].

[44] We could still prove a result like Key Lemma 6.1, but we would not be able to derive the Fixed Point Theorem 6.1, for lack of a small parameter analogous to ε. Perhaps a computer-assisted method could help.

[45] This series is known up to 7 loops [91, 92], see [6] for the earlier 6 loop results. It is tempting to speculate, by analogy with the long-range case, that the ϵ-expansion series have a finite radius of convergence for negative N_b, while they are known to be only Borel-summable for positive N_b. Numerically, the 6-loop series for $\nu_{\rm SR}, \eta_{\rm SR}, \omega_{\rm SR}$ for $N_b = -4$ seem to be remarkably well behaved. E.g. $\eta_{\rm SR}(N_b = -4) = -0.25\epsilon^3 - 0.25\epsilon^2 - 0.535957\epsilon^4 - 1.25122\epsilon^5 - 3.14893\epsilon^6$. We are grateful to Kay Wiese for communicating this to us.

[46] We will give three reasons: (1) Since the LR (long-range) theory is nonlocal, the theory to which it crosses over cannot be fully local. (2) (Counting of degrees of freedom) The leading spin 2 operator is not conserved in the LR theory. At the crossover its dimension goes to d, which is the stress-tensor dimension of

should be marginally irrelevant for the conjecture to hold). Furthermore, $\varepsilon = \varepsilon_*$ marks the boundary of the region of analyticity of the long-range fixed point, and for $\varepsilon > \varepsilon_*$ the long-range fixed point with real couplings seizes to exist. It would be extremely interesting to provide rigorous evidence for these phenomena.

8.1.6 Extension to non-integer N? To non-integer d?

In quantum field theory, one often likes to continue the number of fields from a positive integer, as it nominally should be, to an arbitrary real or even complex value. For lattice models, such continuation often have geometric meaning, as for the $O(N)$ and Potts models when it can be interpreted respectively in terms of loops and Fortuin-Kastelein clusters. Symmetry meaning of such continuations in terms of Deligne categories was recently discussed in [99]. We wish to discuss how such a continuation can be rigorously performed for the model studied here. First one has to factor out explicitly the products of Ω-tensors out of the couplings, i.e. write

$$H_{2k}(\mathbf{A}, \mathbf{x}) = \Omega_{a_1 a_2}\Omega_{a_3 a_4}\cdots\Omega_{a_{2k-1}a_{2k}}\tilde{H}_{2k}(\widetilde{\mathbf{A}}, \mathbf{x}) \pm \text{permutations},$$

where the kernels $\tilde{H}_{2k}$ are "Ω-free", i.e. no longer depend on the a indices, the sequence $\widetilde{\mathbf{A}}$ containing only μ indices. The RG equations can be rewritten in terms of such Ω-free kernels, with contractions of Ω-tensors giving rise in each order to some factors depending polynomially on N. One should then study such Ω-free fixed point equations. It is tempting to conjecture that one can prove fixed point existence for any N and its analytic dependence on N.[47]

More speculatively, one could also try to perform analytic continuation in the space dimension d. One would have to use rotation invariance to come up with a parametrization of the kernel in terms of scalar functions depending on distances between points, times polynomials in point differences. Expressing the RG equations in terms of scalar functions only, dimension d becomes just a parameter which can in principle be continued to non-integer values. Controlling this continuation will likely require major changes in our arguments (much more so than the continuation in N), since we relied on the existence of the physical position space carrying a positive integration measure in several crucial points. But the stakes are high: if one could prove non-perturbative analyticity in d, it would be the first rigorous result of this kind in almost 50 years since the space of 3.99 dimensions was ushered in by Wilson and Fisher [4].

the SR (short-range) fixed point. However it's still not conserved by continuity, so its divergence represents extra states not present in SR. (3) At the coalescence of LR and SR fixed points there must be a marginal operator, on general grounds and because logarithmic corrections are seen in Monte Carlo simulations. The SR fixed point by itself does not have a marginal operator; it can however be constructed as $\psi\zeta$. More arguments are given in [58, 98] where this picture was proposed and thoroughly tested for consistency.

[47] If this is achieved, the coefficient $N-8$ of the one-loop beta function becomes a new small parameter for N near 8. One could then work for $\varepsilon = 0$ where the quartic is marginal, and construct a fixed point with $\lambda = O(N-8)$ balancing the one-loop term against the two-loop term which has an $O(1)$ coefficient. We are grateful to Dario Benedetti for this comment. This could then be done even in $d = 4$, for a theory of local symplectic fermions (8.1). This would be a rigorous version of the Banks-Zaks fixed points in 4d gauge theories [100–102].

8.1.7 Connections to Functional Renormalization Group

FRG represents the most systematic attempt to implement Wilsonian RG in absence of small parameter; see references in the introduction. We are not aware of any FRG studies of specifically symplectic fermions (local or long-range), although other fermionic models (such as Gross-Neveu, Thirring, or Nambu-Jona-Lasinio), or mixed fermion-boson models with Yukawa interactions have been studied via FRG-like techniques; see e.g. [103–106].

Let us compare our results to FRG calculations. In our theorems, all irrelevant couplings were included, and an RG fixed point was rigorously located in a Banach space of interactions, with a provably convergent way to approach it (for a wide range of cutoff procedures). Any FRG calculation always truncates the space of interactions, so that only a subset of irrelevant couplings is included (even if an infinite one). To our knowledge, there are no rigorous results about the best way to exhaust the space of interactions. What is done instead often looks like a matter of prejudice or of convenience. E.g., for bosonic models with the field φ one typically allows an arbitrary potential $V(\varphi)$ but only a handful of derivatives of φ, because the former is believed (although unproven) to be more important, but also because an arbitrary potential is easy (the so called local potential approximation), while derivatives are hard. This state of affairs is both an invitation to mathematical physicists to weigh in and provide rigorous criteria, and to FRG practitioners to explore different exhaustion schemes.

FRG experts may find instructive that our construction used nonlocal interactions terms parametrized by kernels having fast decay at infinity.[48] In principle, our interactions could be expanded in the basis of local monomials with fields carrying an arbitrarily high number of derivatives (the expansion coefficients would be all finite because of the stretched exponential decay of the kernels). However, we have not found such an expansion necessary. It is an open question if rigorous RG analysis can be carried out with interactions expanded in local monomials, and what would be the appropriate Banach space.

Another difference between our result and FRG is that we work with the full Wilsonian effective action, while most FRG calculations are nowadays performed in terms of the one-particle irreducible (1PI) effective action, which flows according to the Wetterich equation [112], as opposed to Polchinski's equation.[49] Empirically, this seems to give better results. The 1PI effective action may be expected to be a somewhat more local object than the Wilsonian effective action, but it too cannot be fully local. We are not aware of any rigorous fixed point results in terms of the 1PI effective action.

8.1.8 Bosonic fixed points

While this paper deals with fermionic fixed point, most fixed points of interest to physics do contain bosonic fields. A few available rigorous bosonic fixed points are listed in appendix K. Notably, they include the bosonic analogues of the models that we studied here, i.e. long-range bosonic $O(N)$ field theories with weakly relevant quartic bare interactions. Unfortunately, these rigorous constructions remain rather daunting, in spite of serious ped-

[48] See also [107–111] for FRG setups allowing nonlocal momentum dependence in the vertices.

[49] See also a non-partisan review in [113], chapter 5.

agogical work which went into trying to simplify them (e.g. [77]). Further simplification is desirable, however unlikely. A very accessible review can be found in [114].

A major complication in the bosonic case, compared to the fermionic one, is that, in defining the RG map $H \mapsto H'$, the terms involving fluctuation fields φ that are (on a local scale) large compared with their standard deviation must be treated in a distinguished way: rather than dealing with them via resummations of perturbation theory, they are controlled via apriori bounds on the probability of such "large fields" configurations in combination with a "polymer expansion".[50] These additional small/large fields decomposition and polymer expansion add a whole new combinatorial level to the construction, which inevitably leads to several technical complications. To date, essentially all the rigorous works on the construction of bosonic fixed points use a parameterization of the full probability distribution of the form

$$e^H + \mathcal{P}, \tag{8.2}$$

rather than of the more standard Gibbs form e^H: in (8.2), H includes the relevant and marginal parts of the interaction, which are exponentiated, while $\mathcal{P}$ includes the "non-perturbative" large field contributions, which are kept non-exponentiated; this mixed form turns out to be optimal for proving that the RG map preserves the space of interactions. Assuming (8.2), RG fixed point equation becomes $H' = H, \mathcal{P} = \mathcal{P}'$, whose form is quite different from (and quite more involved than) the standard "Exact RG equations", such as Polchinski's or Wetterich's, which typically neglect the contribution from the polymer expansion of the large fields contributions (the $\mathcal{P}$-term).

Acknowledgments

We warmly thank Giuseppe Benfatto for his generous contributions to this work in its preliminary phases. We thank David Brydges and Manfred Salmhofer for communications about the rigorous status of Polchinski's equation, and Marco Serone for discussions about renormalons. We thank Malek Abdesselam, Dario Benedetti, David Brydges, Krzysztof Gawedzki, Antti Kupiainen, Pronob Mitter, Gordon Slade, Kay Wiese for comments on the draft.

SR is grateful to Accademia Nazionale dei Lincei for organizing the conference "Advances in Mathematics and Mathematical Physics" (Rome, September 2017) where the first seeds of this project were planted, and to the organizers of the program "Renormalisation in quantum field theory" (Isaac Newton Institute, Cambridge, Fall 2018) where part of the work was done. AG is grateful to IHES for hospitality in January-February 2020, where part of this work was completed.

[50]Loosely speaking, one proceeds as follows: the volume is paved into boxes of typical length comparable with the correlation length of the fluctuation field; each box is called "good" or "bad" depending on whether the typical size of φ in the box is smaller or larger than a large multiple of the standard deviation, respectively; the probability of a bad box is bounded apriori and proved to be very small in the perturbative parameter: therefore, bad boxes are typically far apart from each other; in other words, they form a rarefied gas, whose partition function can be computed via an analogue of the low fugacity expansion for the pressure of lattice gases (this is the polymer expansion which we referred to in the main text).

JHEP01(2021)026

The work of AG and VM has been supported by: the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (ERC CoG UniCoSM, grant agreement n.724939), and by MIUR, PRIN 2017 project MaQuMA cod. 2017ASFLJR, which are gratefully acknowledged. SR is supported by the Simons Foundation grant 488655 (Simons Collaboration on the Nonperturbative Bootstrap), and by Mitsubishi Heavy Industries as an ENS-MHI Chair holder.

A Gevrey classes and fluctuation propagator bounds

In this appendix we give an explicit example of a compactly supported cutoff function satisfying the Gevrey condition, and prove the stretched exponential bound (4.15) for the fluctuation propagator.

A.1 Cutoff function of Gevrey class

Here we will explain that the set of cutoff functions satisfying conditions (2.2) and (2.3) is not empty. Bump C^∞ functions being standard, we will explain how to satisfy in addition to (2.2) the condition (2.3)$\Longleftrightarrow$(4.14), which we copy here:

$$\sup_{k\in\mathbb{R}^d} |\partial^\alpha \chi(k)| \leqslant C^n n^{ns}, \qquad n = |\alpha| = 0, 1, 2 \ldots \tag{A.1}$$

We will not assume any knowledge about Gevrey classes; see e.g. [115, 116] and [37], appendix C.

Recall the following classic result for analytic function. Let $F(k)$ be a function which allows an analytic continuation from real $k \in \mathbb{R}^d$ to a polydisk D_R, i.e. the region of complex $z \in \mathbb{C}^d$ such that $|z_i - (k_0)_i| \leqslant R$ ($i = 1, \ldots, d$). Then, by the Cauchy integral representation, the derivatives of $F(k)$ at the midpoint of the polydisk are bounded by ($n = |\alpha|$)

$$|\partial^\alpha F(k_0)| \leqslant n! R^{-n} A, \qquad A = \max_{D_R} |F(z)|. \tag{A.2}$$

By (A.2), an analytic $\chi(k)$ would satisfy (A.1) with $s = 1$. However, by (2.2) our $\chi(k)$ is compactly supported, hence cannot be analytic. The best we can hope for is (A.1) with $s > 1$.

Let us first construct a $d = 1$ example of a compactly supported Gevrey-class function. Fix $r > 0$ and consider a C^∞ function (see figure 3)

$$X_0(t) = \begin{cases} 0 & t \leqslant 0 \\ e^{-1/t^r}, & t > 0. \end{cases} \tag{A.3}$$

This function is not compactly supported, but this will be corrected below. For now let us check that it is Gevrey class, namely that it satisfies the $d = 1$ analogue of (A.1):

$$|\partial^n X_0(t)| \leqslant C^n n^{ns}, \qquad n = 0, 1, 2 \ldots \tag{A.4}$$

with $s = 1 + 1/r$. (C stands for various positive constants which can change from one line to the next.) Indeed, consider the function

$$X_0(t + z) = e^{-1/(t+z)^r}. \tag{A.5}$$

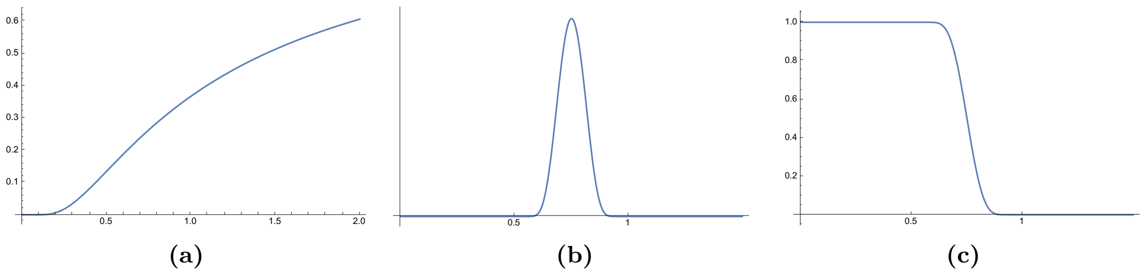


Figure 3. (a) Function $X_0(t)$ for $r = 1$; (b,c) The corresponding functions $X_1(t)$ and $X(t)$.

In the disk of complex $|z| < \kappa t$, where $\kappa > 0$ is sufficiently small, this function is analytic and bounded in absolute value by e^{-C/t^r}.[51] By the $d = 1$ case of the Cauchy estimate (A.2), we have:

$$|\partial^n X_0(t)| \leqslant n!(\kappa t)^{-n} e^{-C/t^r}, \tag{A.6}$$

from where (A.4) follows via elementary maximization of the r.h.s. over t.

From $X_0(t)$ which vanishes at $t \leqslant 0$, we pass to a function of compact support $[1/2, 1]$:

$$X_1(t) = X_0(t - 1/2)X_0(1 - t). \tag{A.7}$$

By the Leibniz rule, it's easy to verify that $X_1(t)$ also satisfies (A.4). Finally, we put

$$X(t) = \int_{|t|}^{\infty} X_1(t')dt' \tag{A.8}$$

which is constant for $|t| \leqslant 1/2$, vanishes for $|t| \geqslant 1$, and still satisfies (A.4). We rescale it so that $X(0) = 1$. See figure 3.

The function $X(t)$ is an explicit example of a cutoff function satisfying conditions (2.2) and (2.3) in $d = 1$. Infinitely many examples of this sort can be given multiplying $X_0(t)$ in (A.3) by an analytic function and repeating the construction.

The function $\chi(k)$ in d dimensions will be given in terms of the 1d function by

$$\chi(k) = X(|k|). \tag{A.9}$$

While eq. (A.1) can be verified using the chain rule, we will instead give a more robust argument via analytic continuation and the Cauchy estimate. The function $f(k) = |k|$ is real analytic for $|k| \in [1/2, 1]$ where derivatives of $X(t)$ are nonzero. Generally, a composition $X(f(k))$ of a Gevrey class function $X(t)$ and a real analytic function $f(k)$ remains in the Gevrey class. For the proof, let $f(z)$ be analytic continuation into a polydisk $|z_i - k_i| \leqslant R$ (we can choose $R = 1/4$ for $f(k) = |k|$, $1/2 \leqslant |k| \leqslant 1$). Writing $X(f(z)) = X(f(k) + [f(z) - f(k)])$ and Taylor-expanding we have

$$X(f(z)) = \sum_{i=0}^{\infty} \frac{1}{i!} \partial^i X(f(k))[f(z) - f(k)]^i. \tag{A.10}$$

[51] κ here depends on r. We can choose it so that $\mathrm{Re}[(1+\zeta)^{-r}] > 1/2$ for $|\zeta| < \kappa$.

Suppose we want to compute $\partial^\alpha[X(f(k))]$, $|\alpha| = n$. We can compute this derivative by differentiating the Taylor series (A.10) truncated to $i \leqslant n$, since all terms with $i > n$ are anyway higher order:

$$\partial_k^\alpha[X(f(k))] = \partial_z^\alpha \Phi(z)|_{z=k}, \qquad \Phi(z) = \sum_{i=0}^{n} \frac{1}{i!}\partial^i X(f(k))[f(z) - f(k)]^i. \tag{A.11}$$

The function $\Phi(z)$ is analytic. It can be bounded in the polydisk by

$$|\Phi(z)| \leqslant C^n n^{n(s-1)}, \tag{A.12}$$

using (A.4) for $X(t)$, and that $f(z)$ is bounded in the polydisk. From here using (A.2) we get (A.1).

A.2 Fluctuation propagator bounds

A.2.1 k-space

Recall that the Fourier transform of $g(x)$ is given by eq. (2.10) which we copy here:

$$\hat{g}(k) = [\chi(k) - \chi(\gamma k)]/|k|^{\frac{d}{2}+\varepsilon}. \tag{A.13}$$

In this subsection we will show, using (A.1), that, for any $k \in \mathbb{R}^d$ and any $n = |\alpha| \geqslant 0$,

$$|\partial^\alpha \hat{g}(k)| \leqslant C \frac{(C\gamma)^n n^{ns}}{|k|^{d/2+\varepsilon}}. \tag{A.14}$$

(C will denote γ- and n-independent constants which may change from one equation to the next.)

We first estimate the derivatives of $1/|k|^{d/2+\varepsilon}$. Consider the analytic continuation of $F(k) = 1/|k|^{d/2+\varepsilon}$ into the polydisk centered at $k \neq 0$ of radius $R = \frac{1}{2}\max_i |k_i|$. The maximum of $|F(z)|$ in this polydisk is bounded by $CF(k)$. We conclude by (A.2) that[52]

$$\left|\partial^\alpha \frac{1}{|k|^{d/2+\varepsilon}}\right| \leqslant n! R^{-n} C F(k) \leqslant \frac{C^n n!}{|k|^{d/2+\varepsilon+n}}. \tag{A.15}$$

Finally since $|k| \geqslant 1/(2\gamma)$ on $\mathrm{supp}[\chi(k) - \chi(\gamma k)]$ we conclude

$$\left|\partial^\alpha \frac{1}{|k|^{d/2+\varepsilon}}\right| \leqslant \frac{(C\gamma)^n n!}{|k|^{d/2+\varepsilon}} \qquad \text{on}\, \mathrm{supp}[\chi(k) - \chi(\gamma k)]. \tag{A.16}$$

Now let us prove (A.14). By the Leibniz rule we have

$$\partial^\alpha \hat{g}(k) = \sum_{\beta \leqslant \alpha} Q_\beta \partial^\beta \frac{1}{|k|^{d/2+\varepsilon}} \times \partial^{\alpha-\beta}[\chi(k) - \chi(\gamma k)], \tag{A.17}$$

[52]For $d = 1$ bound (A.16) would be easy to get by repeated differentiation deriving an explicit formula for the l.h.s. The analytic continuation argument is more robust to show that the same estimate is true for any d.

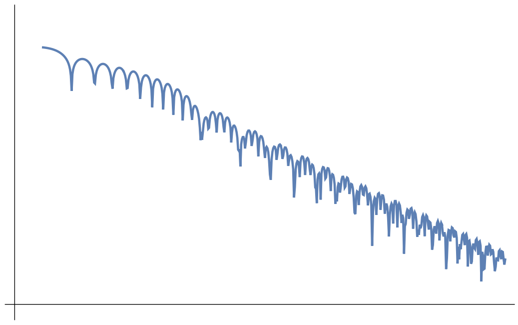


Figure 4. The Fourier transform of the function $X(t)$ from figure 3(c), plotted in log scale against $|x|^{1/s}$ where $s = 1 + 1/r = 2$. The expected $\exp(-C|x|^{1/s})$ decay is visible.

where $Q_\beta = \prod_{i=1}^{d} \binom{\alpha_i}{\beta_i}$. We estimate the ∂^β derivative on the support of $\chi(k) - \chi(\gamma k)$ via (A.16), while the $\partial^{\alpha-\beta}$ factor by (A.1) with $C \to C\gamma$. Combining these two estimates via ($n = |\alpha|$)

$$\begin{aligned} (C\gamma)^{|\beta|}(C\gamma)^{|\alpha-\beta|} &= (C\gamma)^n, \\ |\beta|^{|\beta|}|\alpha-\beta|^{s|\alpha-\beta|} \leqslant n^{s|\beta|}n^{s|\alpha-\beta|} &= n^{sn}, \end{aligned} \tag{A.18}$$

and using that $\sum_\beta Q_\beta = 2^n$, we get (A.14).

A.2.2 x-space

Finally we show (4.15). Consider first the bound for $g(x)$. We use the standard trick that the Fourier transform of $(-ix)^\alpha g(x)$ is $\partial^\alpha \hat{g}(k)$, hence

$$\sup_x |x^\alpha g(x)| \leqslant (2\pi)^{-d} \|\partial^\alpha \hat{g}\|_{L^1}. \tag{A.19}$$

Bound (A.14) then implies (note that $\hat{g}(k)$ has compact support and that $\int_{|k|\leqslant 1} \frac{d^d k}{|k|^{d/2+\varepsilon}} < \infty$)

$$\sup_x |x^\alpha g(x)| \leqslant C(C\gamma)^n n^{ns}, \tag{A.20}$$

which can be rewritten as

$$|g(x)| \leqslant C u^{-n} n^{ns} \equiv C e^{ns \log \frac{n}{u^{1/s}}}, \qquad u = C|x|/\gamma. \tag{A.21}$$

From here (4.15) for $g(x)$ follows by choosing n optimally as $n = \lfloor u^{1/s}/e \rfloor$.

The Fourier transforms of the first and second derivatives of $g(x)$ are $g_1(k) = k_\mu \hat{g}(k)$ and $g_2(k) = k_\mu k_\nu \hat{g}(k)$. Using (A.14) for $\hat{g}(k)$, it's easy to see that $g_1(k), g_2(k)$ satisfy the same type of bounds. Thus the bounds for the first and second derivatives of $g(x)$ follow by the same argument. In fact derivative of any order will have the same kind of decay, only the constants will degrade.

As an illustration, we plot in figure 4 the numerically computed Fourier transform of the function $X(t)$ from figure 3(c). The plot shows the expected $\exp(-C|x|^{1/2})$ decay.

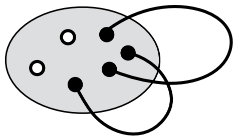


Figure 5. This figure represents a term in $H_{\rm eff}$ corresponding to $n=1$ in (B.1): just one vertex with $|\mathbf{A}|=6$ (gray oval). Empty circles are the external fields and filled circles are the internal ones. The generated $H_{\rm eff}$ term has $l=|\mathbf{B}|=2$. The 4 internal fields are contracted (just one possible contraction is shown).

B Details about $H_{\rm eff}$

In this appendix we give more details about the derivation of eq. (5.3). Plugging (5.2) into (5.1), we represent $H_{\rm eff}(\psi)$ as

$$H_{\rm eff}(\psi)=\sum_{n=1}^{\infty}\frac{1}{n!}\sum_{\mathbf{A}_1,\ldots,\mathbf{A}_n}\sum_{\substack{\mathbf{B}_1,\ldots,\mathbf{B}_n\\ \mathbf{B}_i\subset\mathbf{A}_i}}(-)^{\#} \tag{B.1}$$

$$\times\int d^d\mathbf{x}\Psi(\mathbf{B}_1,\mathbf{x}_{\mathbf{B}_1})\ldots\Psi(\mathbf{B}_n,\mathbf{x}_{\mathbf{B}_n})\prod_{i=1}^{n}H(\mathbf{A}_i,\mathbf{x}_{\mathbf{A}_i})\left\langle\Phi(\overline{\mathbf{B}}_1,\mathbf{x}_{\overline{\mathbf{B}}_1});\ldots;\Phi(\overline{\mathbf{B}}_n,\mathbf{x}_{\overline{\mathbf{B}}_n})\right\rangle_c.$$

Let us explain this in words. We sum over even-length sequences $\mathbf{A}_i$ indexing terms in H. We introduce a coordinate sequence $\mathbf{x}$ of length $|\mathbf{A}_1|+\ldots+|\mathbf{A}_n|$ to be integrated over. We further sum over subsequences $\mathbf{B}_i\subset\mathbf{A}_i$ selecting which fields inside the $H(\mathbf{A}_i,\mathbf{x}_{\mathbf{A}_i})$ term are external. The fields from the complements $\overline{\mathbf{B}}_i=\mathbf{A}_i\setminus\mathbf{B}_i$ are internal, to be contracted in the connected expectation. We use $\mathbf{x}_{\mathbf{A}_i}$, $\mathbf{x}_{\mathbf{B}_i}$, $\mathbf{x}_{\overline{\mathbf{B}}_i}$ to denote the part of the vector $\mathbf{x}$ for the corresponding subsequence. The $(-)^{\#}$ is the sign, which we don't need to track, of the permutation reordering sequence $\mathbf{A}_1+\ldots+\mathbf{A}_n$ to $\mathbf{B}_1+\ldots+\mathbf{B}_n+\overline{\mathbf{B}}_1+\ldots+\overline{\mathbf{B}}_n$.

There are several distinguished groups of terms in (B.1):

- Terms with $n=1$ and $\mathbf{B}_1=\mathbf{A}_1$. They involve no contractions and their sum gives back H.
- Terms with $n=1$ and $\mathbf{B}_1\neq\mathbf{A}_1$. These involve a single $H(\mathbf{A},\mathbf{x})$ vertex with several fields identified as internal and contracted among themselves, while the rest remaining external (see figure 5).
- Terms with $n\geqslant 2$, which therefore correspond to contractions of several vertices. Since (B.1) involves connected expectation, we have to sum over contractions such that the graph becomes connected when every interaction vertex is shrunk to a point (see figure 6).
- Terms for which all $\mathbf{B}_k$ are empty, meaning that all fields are contracted. These sum up to a ψ-independent constant (infinite when working in infinite volume as we are). As mentioned in section 2.1, footnote 12, this constant will be dropped.

Finally, eq. (5.3) follows by rewriting (B.1) in the form (4.2).

JHEP01(2021)026

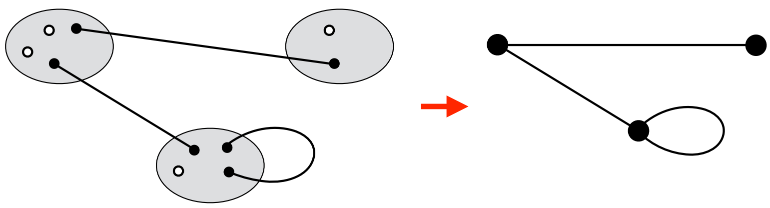


Figure 6. *Left:* a contraction of three vertices, $n = 3$ in (B.1). The generated $H_{\rm eff}$ term has $l = 4$ (the total number of empty circles). *Right:* the graph of contractions obtained when every gray oval is shrunk to a point (denoted by a fat dot). This graph is connected, as it should be because we are considering connected expectations.

C Trimming details

This appendix deals with the trimming map introduced in section 5.2.2, and with how it behaves with respect to the weighted norms. This map takes the interaction $H_{\rm eff}$ in a general representation and returns an equivalent trimmed representation. Consider the parts of $\mathcal{H} = H_{\rm eff}$ which need to be set to zero: $\mathcal{H}_{4,0}$, $\mathcal{H}_{2,0}$, $\mathcal{H}_{2,1}$. Recall that $\mathcal{H}_{l,p}$ corresponds to l-leg interactions with p derivatives:

$$\mathcal{H}_{l,p} \leftrightarrow \sum_{|\mathbf{A}|=l, d(\mathbf{A})=p} \int d^d\mathbf{x} \mathcal{H}(\mathbf{A},\mathbf{x})\Psi(\mathbf{A},\mathbf{x}). \tag{C.1}$$

The dependence on $Sp(N)$ indices should be given by possible invariant tensors:

$$\mathcal{H}_{4,0} \leftrightarrow \mathbf{A} = (a,b,c,e), \quad \mathcal{H}(\mathbf{A},\mathbf{x}) = \Omega_{ab}\Omega_{ce}F_1(\mathbf{x}) - \Omega_{ac}\Omega_{be}F_2(\mathbf{x}) + \Omega_{ae}\Omega_{bc}F_3(\mathbf{x}) \tag{C.2}$$
$$\mathcal{H}_{2,0} \leftrightarrow \mathbf{A} = (a,b), \quad \mathcal{H}(\mathbf{A},\mathbf{x}) = \Omega_{ab}G(\mathbf{x}) \tag{C.3}$$
$$\mathcal{H}_{2,1} \leftrightarrow \mathbf{A} = ((a,\mu),b), \quad \mathcal{H}(\mathbf{A},\mathbf{x}) = \Omega_{ab}K_1^\mu(\mathbf{x}), \tag{C.4}$$
$$\mathbf{A} = (a,(b,\mu)), \quad \mathcal{H}(\mathbf{A},\mathbf{x}) = \Omega_{ab}K_2^\mu(\mathbf{x}). \tag{C.5}$$

Recall that kernels $\mathcal{H}(\mathbf{A},\mathbf{x})$ are antisymmetric, $\mathcal{H}(\pi\mathbf{A},\pi\mathbf{x}) = (-)^\pi\mathcal{H}(\mathbf{A},\mathbf{x})$. This implies various symmetry relations for the functions F, G, K. E.g. G is symmetric, F_1, F_2, F_3 are all related by permutations of their arguments, and finally $K_1^\mu(x_1,x_2) = K_2^\mu(x_2,x_1)$.

By spatial parity (see footnote 15) we have $K_1^\mu(-x_1,-x_2) = -K_1^\mu(x_1,x_2)$. Combined with translational and rotational invariance this implies that $K_1^\mu(x_1,x_2) = (x_1-x_2)^\mu\bar{K}(|x_1-x_2|)$.

The map $T_{\mathbf{2R}}^{\mathbf{2,1}}$ takes $\mathcal{H}_{2,1}$ and returns an equivalent interaction of $H_{2,2}$ type. Consider the part of $\mathcal{H}_{2,1}$ with $\mathbf{A} = ((a,\mu),b)$. Using the interpolation identity

$$\psi_b(x_2) = \psi_b(x_1) + \int_0^1 dt\partial_t[\psi_b(x_1 + t(x_2 - x_1))], \tag{C.6}$$

this term is mapped onto the sum of two terms. The first one is

$$\Omega_{ab}\int d^dx_1 \left[\int d^dx_2 K_1^\mu(x_1,x_2)\right](\partial_\mu\psi_a\psi_b)(x_1), \tag{C.7}$$

which vanishes because as mentioned above $K_1^\mu(x_1,x_2)$ is odd in x_1-x_2. So no local terms are generated in the case at hand. The second term is

$$\Omega_{ab}\int d^dx_1d^dx_2K_1^\mu(x_1,x_2)\partial_\mu\psi_a(x_1)(x_2-x_1)_\nu\int_0^1 dt\partial_\nu\psi_b(x_1+t(x_2-x_1)). \tag{C.8}$$

Changing integration variables from d^dx_2 to d^dy with $y=x_1+t(x_2-x_1)$ and doing the integral over t we have an identity

$$\int d^dx_2K_1^\mu(x_1,x_2)(x_2-x_1)_\nu\int_0^1 dt\partial_\nu\psi_b(x_1+t(x_2-x_1))=\int d^dyL^{\mu\nu}(x_1,y)\partial_\nu\psi_b(y), \tag{C.9}$$

with the help of which we rewrite (C.8) as

$$\Omega_{ab}\int d^dx_1d^dyL^{\mu\nu}(x_1,y)\partial_\mu\psi_a(x_1)\partial_\nu\psi_b(y). \tag{C.10}$$

which is an interaction of type $H_{2,2}$ as promised, and we associate it with $T^{2,1}_{2\mathrm{R}}(\mathcal{H}_{2,1})$. The action on the part of $\mathcal{H}_{2,1}$ with $\mathbf{A}=(a,(\mu,b))$ is analogous and we have to add it to the previous result.

The maps $T^{2,0}_{2\mathrm{L}}$ and $T^{2,0}_{2\mathrm{R}}$ take $\mathcal{H}_{2,0}$ and return an equivalent interaction which is a sum of $H_{2\mathrm{L}}$ and $H_{2,2}$ type interactions. Using the interpolation identity (C.6), an $\mathcal{H}_{2,0}$ interaction is mapped to a sum of two terms. The first term is the local quadratic interaction and we associate it with $T^{2,0}_{2\mathrm{L}}$:

$$T^{2,0}_{2\mathrm{L}}(\mathcal{H}_{2,0})=\nu\Omega_{ab}\int d^dx_1(\psi_a\psi_b)(x_1),\qquad \nu=\int d^dx_2G(x_1,x_2). \tag{C.11}$$

(By translational invariance the integral $\int d^dx_2G(x_1,x_2)$ is x_1-independent). The second term is

$$\Omega_{ab}\int d^d\mathbf{x}\,G(x_1,x_2)\psi_a(x_1)(x_2-x_1)_\nu\int_0^1 dt\partial_\nu\psi_b(x_1+t(x_2-x_1)), \tag{C.12}$$

and similarly to (C.10) we can write it after a change of variable and t-integration as

$$\Omega_{ab}\int d^dx_1d^dy\,G^\nu(x_1,y)\psi_a(x_1)\partial_\nu\psi_b(y). \tag{C.13}$$

This is of type $H_{2,1}$ which we already considered. Acting on it with the map $T^{2,1}_{2\mathrm{R}}$ we will get an equivalent interaction of type $H_{2,2}$. This final result is $T^{2,0}_{2\mathrm{R}}(\mathcal{H}_{2,0})$.

The maps $T^{4,0}_{4\mathrm{L}}$ and $T^{4,0}_{4\mathrm{R}}$ are constructed with the help of the interpolation identity (5.17), which maps $\mathcal{H}_{4,0}$ to an equivalent sum of two interactions, the first of which defines $T^{4,0}_{4\mathrm{L}}$ as it is a local quartic interaction with the coupling

$$\lambda=\int_{x_1=0}d^d\mathbf{x}[F_1(\mathbf{x})+F_2(\mathbf{x})+F_3(\mathbf{x})]=3\int_{x_1=0}d^d\mathbf{x}F_1(\mathbf{x}), \tag{C.14}$$

while the second term is an interaction of $H_{4,1}$ type which is associated with $T^{4,0}_{4\mathrm{R}}$.

We now consider **weighted norm estimates** for the introduced maps. Since $w\geqslant 1$, the localization maps have simply norm one (factor 3 in (C.14) cancels with 1/3 in (4.18)):

$$\|T^{2,0}_{2\mathrm{L}}(\mathcal{H}_{2,0})\|_w\leqslant\|\mathcal{H}_{2,0}\|_w,\qquad \|T^{4,0}_{4\mathrm{L}}(\mathcal{H}_{4,0})\|_w\leqslant\|\mathcal{H}_{4,0}\|_w. \tag{C.15}$$

On the other hand, due to the factors like $(x_2 - x_1)$ in (C.8), the interpolation maps will not preserve the norm $\|\cdot\|_w$. Let us aim instead for an inequality of the type $\|T(\mathcal{H})\|_{w'} \leqslant$ Const $.\|\mathcal{H}\|_{w''}$ where w' is a slightly weaker weight than w'' (i.e. growing slower than w'' at infinity). Eventually we will choose $w' = w(\cdot/\gamma)$, $w'' = w$.

For $T^{2,1}_{2R}$, we need to estimate the w'-norm of $L^{\mu\nu}$ in (C.10) in terms of the w''-norm of K_1^μ. The relation between $L^{\mu\nu}$ and K_1^μ is encoded by the identity (C.9) which by translational invariance and renaming $\partial_\nu\psi^b$ by f can be written equivalently as

$$\int d^d y L^{\mu\nu}(0,y) f(y) = \int d^d x K_1^\mu(0,x) x^\nu \int_0^1 dt f(tx), \tag{C.16}$$

where $f(y)$ is an arbitrary function. The actual expression for $L^{\mu\nu}$ in terms of K_1^μ can be written by e.g. choosing $f(y) = \delta(y - y_0)$ but we don't need it. We write the norm of $L^{\mu\nu}$ as

$$\begin{aligned}\|L^{\mu\nu}\|_{w'} &= \int d^d y |L^{\mu\nu}(0,y)| w'(0,y) = \int d^d y L^{\mu\nu}(0,y)\Sigma(y) w'(0,y) \\ &= \int d^d x K_1^\mu(0,x) x^\nu \int_0^1 dt [\Sigma(y) w'(0,y)]_{y=tx} \\ &\leqslant \int d^d x |K_1^\mu(0,x)||x| w'(0,x), \\ &\leqslant C_1 \|K_1^\mu\|_{w''},\end{aligned} \tag{C.17}$$

where in the first line we defined $\Sigma(y) = \mathrm{sign} L^{\mu\nu}(0,y)$, in the second line we used identity (C.16) with $f(y) = \Sigma(y) w'(0,y)$, in the third line we used that $|\Sigma| = 1$ and assumed that the weight w' is monotonically increasing. Finally, in the last line we assumed the inequality:

$$|x| w'(0,x) \leqslant C_1 w''(0,x) \qquad (x \in \mathbb{R}^d). \tag{C.18}$$

Multiplying the bound (C.17) by 2 to account for the contribution of K_2^μ, we conclude

$$\|T^{2,1}_{2\mathrm{R}}(\mathcal{H}_{2,1})\|_{w'} \leqslant 2C_1 \|\mathcal{H}_{2,1}\|_{w''} \tag{C.19}$$

For $T^{2,0}_{2R}$, very similar considerations will apply. Recall that we have to apply the interpolation identity twice, and each time we will pay a factor of $|x|$ in the weight function. So we get a bound

$$\|T^{2,0}_{2\mathrm{R}}(\mathcal{H}_{2,0})\|_{w'} \leqslant C_2 \|\mathcal{H}_{2,0}\|_{w''} \tag{C.20}$$

under the condition

$$|x|^2 w'(0,x) \leqslant C_2 w''(0,x) \qquad (x \in \mathbb{R}^d). \tag{C.21}$$

For $T^{4,0}_{4R}$, the interpolation identity (5.17) will give rise to an extra factor $|x_2| + |x_3| + |x_4|$ in the weight function. We will therefore obtain:

$$\|T^{4,0}_{4\mathrm{R}}(\mathcal{H}_{4,0})\|_{w'} \leqslant C_3 \|\mathcal{H}_{4,0}\|_{w''}, \tag{C.22}$$

assuming

$$(|x_2| + |x_3| + |x_4|) w'(0,x_2,x_3,x_4) \leqslant C_3 w''(0,x_2,x_3,x_4) \qquad (x_2,x_3,x_4 \in \mathbb{R}^d). \tag{C.23}$$

Finally, we specialize to the case of interest for us: $w'' = w$, $w' = w(\cdot/\gamma)$ where w is our weight (4.16). We leave it as an elementary exercise to show that, for $\gamma \geqslant 2$, inequalities (C.18), (C.21), (C.23) hold with $C_1, C_3 = O(\gamma)$, $C_2 = O(\gamma^2)$ and the constants in the O bounds depend only on C_w and σ in (4.16). [Here $\gamma \geqslant 2$ is useful as C_i would blow up in the limit $\gamma \to 1$, $w' \to w''$.] Bounds (5.43) follow.

D Determinant bounds for fermionic expectations

This appendix discusses the determinant bounds (Gram-Hadamard and Gawedzki-Kupiainen-Lesniewski) for the simple and connected[53] fermionic expectations. These are standard in mathematical physics, but will be unfamiliar to most theoretical physicists, usually concerned with computing, not bounding. These bounds are closely related to the Pauli principle and, at a formal level, to the fermionic expectation being a determinant (hence the name). We will also review the Brydges-Battle-Federbush (BBF) formula, a clever integral representation for connected expectations, useful to derive bounds (and perhaps for other things).

Classic sources (citing previous literature) are [117] for the bounds and [118] for the BBF formula. Other presentations are in [119–121].

D.1 Simple expectations

We are interested in the expectations of the fluctuation field $\phi_a(x)$ which is a gaussian Grassmann field with the propagator $\langle \phi_a(x)\phi_b(y)\rangle = \Omega_{ab} g(x-y)$. We use the notation $\Phi(\mathbf{A}, \mathbf{x})$ for field products as in (4.1). The simple fermionic expectations are given by, see (2.5),

$$\langle \Phi(\mathbf{A}, \mathbf{x})\rangle \equiv \langle \Phi_{A_1}(x_1)\dots\Phi_{A_{2s}}(x_r)\rangle = \sum (-)^p \times \text{Wick contractions}, \tag{D.1}$$

where a Wick contraction is a product of s propagators like

$$\langle \Phi_{A_1}(x_1)\Phi_{A_2}(x_2)\rangle \dots \langle \Phi_{A_{2s-1}}(x_{r-1})\Phi_{A_{2s}}(x_r)\rangle, \tag{D.2}$$

or any other pairing where fields are ordered as $p(1)\dots p(r)$ and $(-)^p = \pm 1$ in (D.1) is the sign of the corresponding permutation p.

Eq. (D.1) contains factorially many terms, but there are cancellations because of the signs. To see this, we rewrite (D.1) as a determinant [122]. As a model, take gaussian Grassmann fields ξ and $\bar\xi$ with propagator

$$\langle \xi(x)\bar\xi(y)\rangle = g(x-y). \tag{D.3}$$

Then,

$$\langle \xi(x_1)\dots\xi(x_s)\bar\xi(y_1)\dots\bar\xi(y_s)\rangle = \pm \det M, \quad M_{ij} = g(x_i - y_j). \tag{D.4}$$

In the general case (D.1), fields ϕ_a carry indices $a = 1\dots N$, and propagator is $\Omega_{ab} g(x-y)$. Renaming odd-a fields as ξ_α, even-a as $\bar\xi_\alpha$, the ξ_α-$\bar\xi_\alpha$ pairs ($\alpha = 1\dots N/2$) are decoupled,

[53] Connected expectations are referred to as ‘truncated’ in mathematical physics.

with propagator $\propto \delta_{\alpha\alpha'}$. The number of ξ and $\bar{\xi}$ fields in the non-vanishing expectation must be the same, let x_i and y_j be their coordinates. Then the expectation (D.1) is, up to a sign, the determinant of the $s \times s$ matrix:

$$\det \mathcal{M}, \quad \mathcal{M}_{ij} = \delta_{\alpha_i \alpha_j} \Gamma_{ij}(x_i - y_j). \tag{D.5}$$

Here Γ_{ij} is either $g(x-y)$ or its derivative if some fields carry derivatives. We will estimate it with the help of

Lemma D.1 (Gram-Hadamard inequality) *For (f_i), (h_i) $(i = 1 \dots s)$ two lists of vectors in a Hilbert space, let M_{ij} be the $s \times s$ matrix of their inner products: $M_{ij} = (f_i, h_j)$. Then $D(f,h) = \det M$ satisfies an upper bound:*

$$|D(f,h)| \leqslant \prod_{i=1}^{s} \|f_i\| \|h_i\|. \tag{D.6}$$

Proof. For h_i orthonormal, this holds interpreting the determinant as the volume of parallelepiped formed by f_i (this case is known as Hadamard's inequality). By rescaling, the inequality remains true for h_i orthogonal of arbitrary length. We will next reduce the general case to this special case.

Out of general h_i, we build $\tilde{h}_i$ by Gram-Schmidt: $\tilde{h}_1 = h_1$, $\tilde{h}_2 = h_2 - \alpha h_1 \perp \tilde{h}_1$ (projection of h_2 on the subspace orthogonal to $\tilde{h}_1$), $\tilde{h}_3$ projection of h_3 on the subspace orthogonal to $\tilde{h}_1, \tilde{h}_2$ etc. By properties of determinants:

$$D(f,h) = D(f,\tilde{h}). \tag{D.7}$$

From the special case, the r.h.s. is bounded by $\prod \|f_i\| \|\tilde{h}_i\|$, and $\|\tilde{h}_i\| \leqslant \|h_i\|$ since it's a projection. Q.E.D.

To apply this result, we have to write the matrix elements (D.5) as products of vectors in a Hilbert space. Without indices and derivatives we have

$$g(x_i - y_j) = (f_i, h_j), \tag{D.8}$$

introducing two families of L^2 functions in momentum space (a trick due to [123]):[54]

$$f_i(k) = e^{-ikx_i} |\hat{g}(k)|^{1/2} \hat{g}(k)/|\hat{g}(k)|, \quad h_j(k) = e^{-iky_j} |\hat{g}(k)|^{1/2}, \tag{D.9}$$

When some fields carry derivatives, we just include a factor ik_μ into the corresponding function. Finally with indices, we view f_i and h_j as vector functions, multiplying (D.9) by the unit vectors in the directions α_i, α_j, whose inner product reproduces the Kronecker $\delta_{\alpha_i \alpha_j}$. We just proved

Lemma D.2 (Gram-Hadamard bound) *We have the bound*

$$\begin{aligned} &|\langle \Phi_{A_1}(x_1) \dots \Phi_{A_{2s}}(x_{2s}) \rangle| \leqslant (C_{\rm GH})^s \quad \textit{with} \\ &C_{\rm GH} = \max \left(\int \frac{d^d k}{(2\pi)^d} |\hat{g}(k)|, \int \frac{d^d k}{(2\pi)^d} (k_1)^2 |\hat{g}(k)| \right). \end{aligned} \tag{D.10}$$

[54] $\hat{g}(k)$ will be non-negative if $\chi(k)$ is non-negative and monotonic, but we state this part of the argument for a general complex $\hat{g}(k)$.

This bound is related to the Pauli principle, as can be seen from the following alternative proof. We can represent $\Phi_A(x)$ as operators acting on a Hilbert space (fermionic Fock space). Fermionic occupation numbers being either zero or one, operators $\Phi_A(x)$ turn out to have a finite norm $\|\Phi_A\|$, and expectation then grow at most as a power $\|\Phi_A\|^{2s}$. This should be contrasted with the bosonic case, when the operator norm would have been infinite (even for a simple harmonic oscillator).

Lemma D.3 *Let $\hat{g}(k)$ be as in (2.10), with $\chi(k)$ satisfying (2.2). Then $C_{\rm GH}$ is uniformly bounded over $\gamma \geqslant 2$ and $\varepsilon \in T$ where $T \subset \mathbb{C}$ is any compact subset of the complex half-plane (7.1).*

Proof. We have $C_{\rm GH} \leqslant \int_{|k|\leqslant 1} d^d k (2\pi |k|)^{-(d/2+\mathrm{Re}\,\varepsilon)}$, uniformly bounded since $\max_T \mathrm{Re}\,\varepsilon < d/6$ by (7.1).

D.2 Connected expectations

Dividing the points $\mathbf{x}$ into n groups $\mathbf{x}_1, \dots, \mathbf{x}_n$, connected (also called 'truncated') expectations are given by

$$\langle \Phi(\mathbf{A}_1, \mathbf{x}_1); \dots ; \Phi(\mathbf{A}_n, \mathbf{x}_n)\rangle_c = \sum (-)^p \times \text{connected Wick contractions.} \tag{D.11}$$

Connected Wick contractions form a subset of terms from (D.1), those for which the graph of propagators becomes connected when each group of points $\mathbf{x}_i$ is collapsed into one point. The signs $(-)^p$ are the same as in (D.1).

Because $g(x)$ decays at infinity, connected expectations are small when any two groups $\mathbf{x}_i$ and $\mathbf{x}_j$ get far apart. We need a bound incorporating both this fact and the cancellations due to signs. This will be done via a clever generalization of the determinant representation to connected expectations.

To begin with, via (D.1) and (D.11), the simple and connected expectations satisfy two relations ($\Phi_i \equiv \Phi(\mathbf{A}_i, \mathbf{x}_i)$). First, they coincide for a single group of points:

$$\langle \Phi_i \rangle_c = \langle \Phi_i \rangle. \tag{D.12}$$

Second, simple expectation can be computed by partitioning n group of fields into subsets, taking products of connected expectations within each subset, and summing over all ways of partitioning:[55]

$$\langle \prod_{i=1}^{n} \Phi_k \rangle = \sum_{\Pi \in \text{partitions of } \{1 \dots n\}} (-)^\pi \prod_{Y\in\Pi} \langle \Phi_{Y_1}; \Phi_{Y_2}; \dots \rangle_c. \tag{D.13}$$

The $(-)^\pi$ is the parity of the permutation bringing fields in the r.h.s. into the original order in the l.h.s. (it's not the same permutation as in (D.11).

[55]This formula is more rapid that (D.11) to compute connected expectations: one recursively expresses them via the usual expectations, which in turn are evaluated by the determinant formula (D.5) [124]. This observation speeds up the Diagrammatic Monte Carlo algorithm from footnote 24 [124] achieving polynomial complexity [125]. One wonders if the BBF formula (D.47) below could give an alternative practical way to evaluate connected expectations.

JHEP01(2021)026

Reading (D.13) from right to left, one can recursively computes connected expectations from simple ones. E.g. for $n = 2, 3$ we have

$$\langle \Phi_1 \Phi_2 \rangle = \langle \Phi_1 ; \Phi_2 \rangle_c + \langle \Phi_1 \rangle \langle \Phi_2 \rangle, \tag{D.14}$$

$$\begin{aligned} \langle \Phi_1 \Phi_2 \Phi_3 \rangle &= \langle \Phi_1 ; \Phi_2 ; \Phi_3 \rangle_c + \langle \Phi_1 ; \Phi_2 \rangle_c \langle \Phi_3 \rangle + \langle \Phi_1 \rangle \langle \Phi_2 ; \Phi_3 \rangle_c \\ &\quad + (-)^{N_2 N_3} \langle \Phi_1 ; \Phi_3 \rangle_c \langle \Phi_2 \rangle + \langle \Phi_1 \rangle \langle \Phi_2 \rangle \langle \Phi_3 \rangle. \end{aligned} \tag{D.15}$$

In the r.h.s. we replaced $\langle \Phi_k \rangle_c = \langle \Phi_k \rangle$ by (D.12). N_k is the number of fields in Φ_k. From (D.14) we find $\langle \Phi_1 ; \Phi_2 \rangle_c = \langle \Phi_1 \Phi_2 \rangle - \langle \Phi_1 \rangle \langle \Phi_2 \rangle$; substituting this into (D.15) we find $\langle \Phi_1 ; \Phi_2 ; \Phi_3 \rangle_c$; etc. So (D.13) provides an alternative definition of connected expectations, a useful starting point for what follows.

Let $m = \sum_{k=1}^{n} |\mathbf{x}_k|$ be the total number of points. In section D.1 we wrote the simple expectation as a determinant of the matrix $\mathcal{M}$ defined in (D.5). Introducing auxiliary Grassmann variables η_i and $\bar{\eta}_j$ ($m/2$ of each type), we write it then as a Grassmann integral

$$\langle \prod_{k=1}^{n} \Phi(\mathbf{A}_k, \mathbf{x}_k) \rangle = \int \prod d\eta_i d\bar{\eta}_j e^{V}, \tag{D.16}$$

where we defined the potential function

$$V = \sum_{i,j} \mathcal{M}_{ij} \eta_i \bar{\eta}_j. \tag{D.17}$$

Here i, j index the fields classified in section D.1 as $\xi_{\alpha_i}(x_i)$ and $\bar{\xi}_{\alpha_j}(y_j)$. Depending in which group $\mathbf{x}_k$ their positions x_i and y_j belong, we subdivide V as

$$V = \frac{1}{2} \sum_{k,l=1}^{n} V_{kl}, \tag{D.18}$$

$$V_{kl} = \sum_{i,j: x_i \in \mathbf{x}_k, y_j \in \mathbf{x}_l} \mathcal{M}_{ij} \eta_i \bar{\eta}_j + (k \leftrightarrow l). \tag{D.19}$$

Define $V(X)$ and $\psi(X)$ on any finite subset $X \subset \{1 \dots n\}$ by

$$V(X) = \frac{1}{2} \sum_{k,l \in X} V_{kl}, \quad \psi(X) = e^{V(X)}. \tag{D.20}$$

We can think of $V(X)$ as the total potential energy for a group of points with pairwise interactions. Define connected part $\psi_c(X)$ recursively by the following equations:

$$\psi_c(X) = \psi(X) \quad \text{if } |X| = 1, \tag{D.21}$$

$$\psi(X) = \sum_{\Pi \in \text{partitions of } X} \prod_{Y \in \Pi} \psi_c(Y). \tag{D.22}$$

Crucially, the form these equations is such that integrating them in η_i, $\bar{\eta}_j$, we land precisely on (D.12), (D.13) (including the $(-)^{\pi}$ sign), provided we identify

$$\langle \Phi_{Y_1} ; \Phi_{Y_2} ; \dots \rangle_c = \int \prod d\eta_i d\bar{\eta}_j \psi_c(Y), \tag{D.23}$$

where the integral is over the subset of Grassmann variables belonging to Y (which means $x_i \in \mathbf{x}_k, y_j \in \mathbf{x}_l$, where $k, l \in Y$). Computing connected expectations is thus reduced to finding $\psi_c(Y)$ in terms of V. We will consider this problem in general, for an arbitrary symmetric V_{kl}. That our V_{kl} is given by eq. (D.19) will become important again only in section D.4.

There is a standard formula for ψ_c:

$$\psi_c(X) = \sum_{G \in \text{connected graphs on } X,\ |G|=|X|} \prod_{kl \in G} (e^{V_{kl}} - 1) \prod_{k \in X} e^{\frac{1}{2} V_{kk}}. \tag{D.24}$$

But this is not very useful for our purposes: plugging it into (D.23) just gives back eq. (D.11) (perhaps not surprisingly as both (D.11) and (D.24) involve connected graphs; we leave the proof as an exercise.) The number of terms in (D.24) is asymptotically $2^{\binom{n}{2}}$, since for large n almost all graphs on n points are connected (e.g. [126], example II.15).

D.3 Brydges-Battle-Federbush (BBF) formula

We will now derive a remarkable formula for $\psi_c(X)$ with much fewer terms. First two simplifying remarks: 1) The diagonal interactions V_{kk} enter into $\psi_c(X)$ as a trivial common factor $\exp(\frac{1}{2} \sum_{k \in X} V_{kk})$. So we will set them to zero and reinstate in the final result. We consider a complete graph on n points with pairs kl as the graph edges e, and we write

$$V(X) = \sum_{e} V(e). \tag{D.25}$$

2) It is enough to aim for the equation

$$e^{V(X)} = \sum_{Y \ni 1} e^{V(X \setminus Y)} \psi_c(Y), \tag{D.26}$$

from which (D.22) follows by iterating. Every term in the r.h.s. has Y and $X \setminus Y$ "decoupled" in the sense that no edges linking them are involved. Let us describe a general decoupling mechanism.

Let $F(X)$ be any sum of pairwise interactions $F(e)$, and $Z \subset X$ a subset we wish to decouple. We say that an edge e "exits Z" (written $e \dashv Z$) if it is of the form $e = kl$ where $k \in Z$, $l \in X \setminus Z$. In other words, an edge exits Z if it connects Z to $X \setminus Z$. Introduce a variable $s \in [0, 1]$ and a new pairwise interaction with exiting edges rescaled by s, others left intact:

$$F\begin{bmatrix} s \\ Z \end{bmatrix}(e) = \begin{cases} sF(e), & \text{if } e \dashv Z, \\ F(e) & \text{otherwise.} \end{cases} \tag{D.27}$$

We also define $F\begin{bmatrix} s \\ Z \end{bmatrix}(X)$ summing this new interaction:

$$F\begin{bmatrix} s \\ Z \end{bmatrix}(X) = \sum_{e \subset X} F\begin{bmatrix} s \\ Z \end{bmatrix}(e), \tag{D.28}$$

Then $F\begin{bmatrix}1\\Z\end{bmatrix}(X) = F(X)$ and $F\begin{bmatrix}0\\Z\end{bmatrix}(X) = F(X\setminus Z) + F(Z)$ is decoupled. Therefore we have

$$e^{F(X)} = e^{F\begin{bmatrix}0\\Z\end{bmatrix}(X)} + \int_0^1 ds\partial_s \exp F\begin{bmatrix}s\\Z\end{bmatrix}(X) = e^{F(X\setminus Z)}e^{F(Z)} + \sum_{e\dashv Z}\int_0^1 dsF(e)\exp F\begin{bmatrix}s\\Z\end{bmatrix}(X). \tag{D.29}$$

The first term is decoupled. In the second one we have a sum over exiting edges. Below, each of these summands will be decoupled further with respect to the sets $Z \sqcup e$ obtained by joining to Z the outside vertex of the edge e. ($\sqcup$ denotes "vertex union").

Think of $\begin{bmatrix}s\\Z\end{bmatrix}$ as an operation, which can be applied repeatedly. E.g. $F\begin{bmatrix}s_1\\Z_1\end{bmatrix}\begin{bmatrix}s_2\\Z_2\end{bmatrix}$ means first rescale by s_1 all edges $e_1 \dashv Z_1$, then by s_2 all edges $e_2 \dashv Z_2$ by s_2. In general these operations don't commute.

To derive (D.26), we first apply this general mechanism with $Z = \{1\}$ consisting of one point, and $F(X) = V(X)$. We obtain

$$e^{V(X)} = e^{V(X\setminus\{1\})} + \sum_{e_1\dashv\{1\}}\int_0^1 ds_1 V(e_1)\exp V\begin{bmatrix}s_1\\\{1\}\end{bmatrix}(X). \tag{D.30}$$

The first term gives the first term in the r.h.s. of (D.26), with $Y = \{1\}$ and $\psi_c(Y) = 1$. For each term in the sum over e_1, we define the set $Z_{e_1} = \{1\} \sqcup e_1$ and decouple with respect to Z_{e_1}, i.e. apply (D.29) for $F = V\begin{bmatrix}s_1\\\{1\}\end{bmatrix}$ and the rescaled interaction $F\begin{bmatrix}s_2\\Z_{e_1}\end{bmatrix} = V\begin{bmatrix}s_1\\\{1\}\end{bmatrix}\begin{bmatrix}s_2\\Z_{e_1}\end{bmatrix}$:

$$e^{V\begin{bmatrix}s_1\\\{1\}\end{bmatrix}(X)} = e^{V(X\setminus Z_{e_1})}e^{V\begin{bmatrix}s_1\\\{1\}\end{bmatrix}(Z_{e_1})} + \sum_{e_2\dashv Z_{e_1}}\int_0^1 ds_2 V\begin{bmatrix}s_1\\\{1\}\end{bmatrix}(e_2)\exp V\begin{bmatrix}s_1\\\{1\}\end{bmatrix}\begin{bmatrix}s_2\\Z_{e_1}\end{bmatrix}(X). \tag{D.31}$$

In the first term we used $F(X\setminus Z_{e_1}) = V(X\setminus Z_{e_1})$, since $F = V$ outside Z_{e_1}. Plugging (D.31) into (D.30) we get:

$$\begin{aligned} e^{V(X)} &= e^{V(X\setminus\{1\})} \\ &\quad + \sum_{e_1\dashv\{1\}} e^{V(X\setminus Z_{e_1})}\int_0^1 ds_1 V(e_1)\exp V\begin{bmatrix}s_1\\\{1\}\end{bmatrix}(Z_{e_1}) \\ &\quad + \sum_{e_1\dashv\{1\}}\sum_{e_2\dashv Z_{e_1}}\int_0^1 ds_1 ds_2 V(e_1)V\begin{bmatrix}s_1\\\{1\}\end{bmatrix}(e_2)\exp V\begin{bmatrix}s_1\\\{1\}\end{bmatrix}\begin{bmatrix}s_2\\Z_{e_1}\end{bmatrix}(X). \end{aligned} \tag{D.32}$$

To compare with (D.26), we rewrite the second line as

$$\sum_{e_1\dashv\{1\}} e^{V(X\setminus Z_{e_1})}\psi_c(Z_{e_1}), \qquad \psi_c(Z_{e_1}) = \int_0^1 ds_1 V(e_1)\exp V\begin{bmatrix}s_1\\\{1\}\end{bmatrix}(Z_{e_1}). \tag{D.33}$$

This equation defines ψ_c for sets of two points.

We then continue iterating. For each term in the third line of (D.32), we define $Z_{e_1e_2} = Z_{e_1} \sqcup e_2$ and decouple with respect to this set. This allows us to rewrite the third line as:

$$\sum_{e_1 \dashv \{1\}} \sum_{e_2 \dashv Z_{e_1}} e^{V(X \backslash Z_{e_1e_2})} \int_0^1 ds_1 ds_2 V(e_1) V\begin{bmatrix} s_1 \\ \{1\} \end{bmatrix}(e_2) \exp V \begin{bmatrix} s_1 \\ \{1\} \end{bmatrix}\begin{bmatrix} s_2 \\ Z_{e_1} \end{bmatrix}(Z_{e_1e_2}), \quad \text{(D.34)}$$

plus a triple integral term which we don't write. To compare this with (D.26), we write it as

$$\sum_{Z \ni \{1\}, |Z|=3} e^{V(X\backslash Z)} \psi_c(Z). \quad \text{(D.35)}$$

We should define $\psi_c(Z)$ summing over all possible orders of adding edges so that the final set $Z_{e_1e_2} = Z$, which means the two added edges should be picked from Z. We thus have

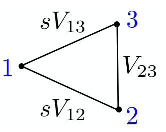
$$\psi_c(Z) = \sum_{e_i \subset Z, e_1 \dashv \{1\}, e_2 \dashv Z_{e_1}} \int_0^1 ds_1 ds_2 V(e_1) V\begin{bmatrix} s_1 \\ \{1\} \end{bmatrix}(e_2) \exp V \begin{bmatrix} s_1 \\ \{1\} \end{bmatrix}\begin{bmatrix} s_2 \\ Z_{e_1} \end{bmatrix}(Z). \quad \text{(D.36)}$$

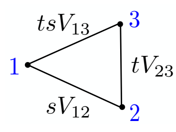
Continuing to iterate, we obtain the following general formula for $|Z| = n$, $Z \ni \{1\}$, which involves $n-1$ added edges and integrations:

$$\psi_c(Z) = \sum_{e_i \subset Z, e_1 \dashv \{1\}, e_2 \dashv Z_{e_1}, e_3 \dashv Z_{e_1e_2}, \ldots} \quad \text{(D.37)}$$
$$\times \int_0^1 \prod_{k=1}^{n-1} ds_k \left\{ V(e_1) V\begin{bmatrix} s_1 \\ \{1\} \end{bmatrix}(e_2) V\begin{bmatrix} s_1 \\ \{1\} \end{bmatrix}\begin{bmatrix} s_2 \\ Z_{e_1} \end{bmatrix}(e_3) \cdots \right\} \exp V \begin{bmatrix} s_1 \\ \{1\} \end{bmatrix}\begin{bmatrix} s_2 \\ Z_{e_1} \end{bmatrix} \cdots \begin{bmatrix} s_{n-1} \\ Z_{e_1e_2\ldots e_{n-2}} \end{bmatrix}(Z).$$

This is valid under the simplifying assumption $V_{kk} = 0$ made above, while for general V we must multiply by $\exp(\frac{1}{2}\sum_{k \in Z} V_{kk})$. Another remark: definition (D.21), (D.22) shows that $\psi_c(Z)$, like $\psi(Z)$, must be symmetric with respect to permutations. Eq. (D.37) is not manifestly symmetric as it selects $\{1\} \in Z$, although of course it produces symmetric results, see an example below. This lack of manifest symmetry will not be a problem in our applications.

Example D.1 It's instructive to check this formula for three points: $Z = \{1, 2, 3\}$. We thus have three edges $e = 12, 13, 23$. The rescaled interactions are ($s_1 = s, s_2 = t$):

$V\begin{bmatrix} s \\ \{1\} \end{bmatrix}$: 

$V\begin{bmatrix} s \\ \{1\} \end{bmatrix}\begin{bmatrix} t \\ \{12\} \end{bmatrix}$: 

(D.38)

Thus the term in (D.36) corresponding to $e_1 = 12$ reads:

$$\int_0^1 ds dt V_{12}(sV_{13} + V_{23}) e^{sV_{12} + tsV_{13} + tV_{23}}. \quad \text{(D.39)}$$

For $e_1 = 13$ we get the same integral with $2 \leftrightarrow 3$. Doing the two integrals and summing, we obtain the same expression as by expanding out (D.24):

$$\psi_c(123) = 2 - e^{V_{12}} - e^{V_{13}} - e^{V_{23}} + e^{V_{12}+V_{13}+V_{23}}. \quad \text{(D.40)}$$

Continuing with the general case, it is not hard to write out all parts of (D.37) explicitly. Define a function $g : \{1, \dots, n\} \to \{1, \dots, n\}$ which tells us the order of adding vertices when adding edges e_l: $g(1) = 1$ and $g(l)$ is the endpoint of the edge e_{l-1}. So $e_{l-1} = g(k)g(l)$ with some $k < l$. In the exponential of (D.37) the contribution V_e of this edge appears rescaled by

$$r_e = s_k s_{k+1} \cdots s_{l-1}, \tag{D.41}$$

while in the prefactor $\{\dots\}$ $V_{e_{l-1}}$ appears rescaled by

$$s_k s_{k+1} \cdots s_{l-2} \tag{D.42}$$

(or 1 if $k = l - 1$). This expression is consistent with applying $\partial_{s_{l-1}}$ to (D.41), as it must be by (D.29).

It follows from this discussion that every summand in (D.37) can be written as

$$\prod_{k=1}^{n-1} V_{e_k} \times \int_0^1 \prod_{k=1}^{n-1} ds_i f(\mathbf{s}) \exp \sum_e r_e V_e, \tag{D.43}$$

where r_e are given by (D.41) and $f(\mathbf{s})$ is some product of s's. Its dependence on the choice of added edges can be made completely explicit, but for us it will suffice to know that $f \geqslant 0$. Furthermore, we would like to view the integral in (D.43) as performed over r_e's, not over s's, writing it as

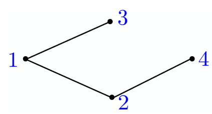
$$\prod_{k=1}^{n-1} V_{e_k} \times \int d\mu(\mathbf{r}) \exp \sum_e r_e V_e, \tag{D.44}$$

where $d\mu(\mathbf{r})$ is some nonnegative measure, push-forward of the measure $\prod_{k=1}^{n-1} ds_k f(\mathbf{s})$ to r_e's. This singular, delta-function-like, measure is concentrated on r_e's which can be represented as (D.41).

For the final repackaging, consider the graph which is the union of all added edges:

$$T = e_1 \cup e_2 \cup \ldots \cup e_{n-1}. \tag{D.45}$$

By construction, this is a tree with n vertices $\{1 \dots n\}$. Note that the same tree T may appear from different sequences of edges, all satisfying the constraints $e_1 \dashv \{1\}, e_2 \dashv Z_{e_1}$, etc in (D.37). E.g. the following tree may arise from $\{e_1, e_2, e_3\} = \{13, 12, 24\}$ or $\{12, 13, 14\}$ or $\{12, 24, 13\}$:



(D.46)

Terms corresponding to the same tree T have the same prefactor $\prod_{k=1}^{n-1} V_{e_k}$ in (D.44), although the measures $d\mu(\mathbf{r})$ will be different. Let us group all such terms into one term per tree T, summing their measures into some total measure. Reintroducing as well the trivial $\exp(\frac{1}{2} \sum V_{kk})$ factor, we obtain the remarkable BBF formula [127, 128]; our presentation followed [118].

Lemma D.4 (BBF formula) *Let T run through all trees with n vertices $\{1\dots n\}$. There exist non-negative measures $d\mu_T(\mathbf{r})$ so that, for any pairwise interaction V:*

$$\psi_c(1\dots n) = \sum_T \prod_{e\in T} V_e \int d\mu_T(\mathbf{r}) \exp(\sum_e r_e V_e + \tfrac{1}{2}\sum_{k=1}^{n} V_{kk}). \tag{D.47}$$

In addition, these measures have the following two properties:

- *For each $\mathbf{r}$ in the support of $d\mu_T(\mathbf{r})$, there exists a bijection $g : \{1,\dots,n\} \to \{1,\dots,n\}$ and a set of $n-1$ numbers $s_k \in [0,1]$ such that for all $k < l$*

$$r_{g(k)g(l)} = s_k s_{k+1} \cdots s_{l-1}. \tag{D.48}$$

- *The measures $d\mu_T(\mathbf{r})$ are probability measures, i.e. they have total weight 1: $\int d\mu_T(\mathbf{r}) = 1$.*

The first property is clear, since $d\mu_T(\mathbf{r})$ were obtained as sums of such measures. The second property can be checked by a trick: apply the general formula to a particular conveniently chosen potential V. Pick a tree T and consider V such that $V_e = \varepsilon$ for $e \in T$ and $V = 0$ otherwise. That $\int d\mu_T(\mathbf{r}) = 1$ follows by comparing, for small ε, the equations

$$\psi_c(1\dots n) = \begin{cases} \varepsilon^n \int d\mu_T(\mathbf{r}) + \text{higher order} & \text{from (D.47)}, \\ \varepsilon^n + \text{higher order} & \text{from (D.24)}. \end{cases} \tag{D.49}$$

Remark D.1 The measure $d\mu_T$ may be thought as having several components, corresponding to different bijections g, whose number equals the number of ways to grow the tree T by adding edges. Each component can be pulled back to an integral over s_k's, and the total weight of all components is 1. This can also be verified by an explicit computation (Lemma A.4 in [119] or Lemma 2.3 in [121]).

The number of terms in the BBF formula is much smaller than in the standard formula (D.24): it grows as the number of trees on n points, which is n^{n-2} (Cayley).

We we will use the BBF formula to prove bounds on the connected fermionic expectations. One wonders if this formula can also be useful to *evaluate* connected expectations, e.g. performing the integral numerically, rather than just prove bounds. We are not aware of such applications (see also footnote 55).

D.4 Gawedzki-Kupiainen-Lesniewski (GKL) bound

We will now present a bound on fermionic connected expectations due to Gawedzki and Kupiainen [129]. Its physical origin, like for the Gram-Hadamard bound (D.10), is the Pauli principle. Our exposition follows Lesniewski [117], who gave an elegant proof based on the BBF formula.[56]

[56] Original proofs [123, 129] were based on an improved Gram-Hadamard inequality for the simple expectations, transferred to connected expectations via cluster expansion techniques. See also [75] for an alternative approach.

We start with eq. (D.23) copied here for $Y = \{1, \ldots, n\}$:

$$\langle \Phi(\mathbf{A}_1, \mathbf{x}_1); \ldots; \Phi(\mathbf{A}_n, \mathbf{x}_n) \rangle_c = \int \prod d\eta_i d\bar{\eta}_j \psi_c(1 \ldots n). \tag{D.50}$$

Recall that we introduced the potential function (D.17) (see (D.5) for matrix $\mathcal{M}$) subdivided in symmetric pairwise interactions V_{kl}, $1 \leqslant k, l \leqslant n$, see (D.18), (D.19). The BBF formula (D.47) gives a general expression for $\psi_c(1 \ldots n)$ in terms of V_{kl}.

What happens with every term in the BBF formula when we plug in expressions (D.19) for V_{kl} and do the Grassmann integral? First of all let us look at the prefactor $\prod_{e \in T} V_e$. By (D.19), every $V_e = V_{kl}$ is a sum of terms $\mathcal{M}_{ij} \eta_i \bar{\eta}_j$ where $x_i \in \mathbf{x}_k$, $y_j \in \mathbf{x}_l$ or vice versa and $\mathcal{M}_{ij}$ is a propagator, see (D.5). We choose in each V_e, $e \in T$, one of such possible propagator terms. The graph with all chosen propagators as edges is called an "anchored tree $\mathcal{T}$ on the n groups of points $\mathbf{x}_i$". When each group of points $\mathbf{x}_i$ is contracted to one points, the anchored tree $\mathcal{T}$ becomes a tree (the tree T in the case at hand). We will say that "$\mathcal{T}$ comes from T".

Let us look at a term corresponding to a fixed anchored tree $\mathcal{T}$ coming from T. The variables η and $\bar{\eta}$ of the vertices along the anchored tree are saturated by the prefactor. Doing the Grassmann integral over these "tree-saturated" variables gives a product of $n-1$ propagators along $\mathcal{T}$:

$$\prod_{\text{along } \mathcal{T}} \Gamma_{ij}(x_i - y_j). \tag{D.51}$$

We are left with the integral of the exponent over the remaining $m - 2(n-1)$ variables.[57] (By the rules of Grassmann integration, the tree-saturated variables can now be set to zero in the exponent.) The potential function in the exponent of (D.47) will have the form

$$\sum_{ij} \mathcal{N}_{ij} \eta_i \bar{\eta}_j, \qquad \mathcal{N}_{ij} = r_{k(i)k(j)} \mathcal{M}_{ij}, \tag{D.52}$$

summing over the remaining variables, so $\mathcal{N} = \mathcal{N}(\mathbf{r})$ is an $s \times s$ matrix with $s = \frac{1}{2}(m - 2(n-1))$. If variables i, j belong to two different groups of points $\mathbf{x}_{k(i)}$, $\mathbf{x}_{k(j)}$, then $r_{k(i)k(j)} = r_e$ in (D.52), where $e = k(i)k(j)$ is the edge of the tree T, progenitor of the anchored tree, and r_e is the rescaling factor in the BBF formula. Terms from the same group, coming from $\frac{1}{2} \sum V_{kk}$ in (D.47), should not be rescaled: so we set $r_{kk} = 1$. The Grassmann integral over the remaining variables is then the determinant of the so defined matrix $\mathcal{N}$. To summarize, the connected expectation (D.50) can be represented as

$$\sum_{\mathcal{T}} \prod_{\text{along } \mathcal{T}} \Gamma_{ij}(x_i - y_j) \int d\mu_T(\mathbf{r}) \det \mathcal{N}. \tag{D.53}$$

Aiming to bound $\det \mathcal{N}$ by the Gram-Hadamard inequality (D.6), we wish to represent $\mathcal{N}_{ij}$ as a product of vectors in a Hilbert space. For $\mathcal{M}_{ij}$ such a representation was given in section D.1. To deal with the extra factor $r_{k(i)k(j)}$ we will use

[57] Recall that $m = \sum_{k=1}^{n} |\mathbf{x}_k|$ is the total number of points, and we have one Grassmann variable per point.

Lemma D.5 *Let* $s_k \in [0,1], k = 1 \dots n-1$. *There exist* n *unit vectors* $u_k = u_k(\mathbf{s}) \in \mathbb{R}^n$ *such that* $(u_k, u_l) = s_k \dots s_{l-1}$ *for all* $1 \leqslant k < l \leqslant n$.

Proof. Let v_k be the standard orthonormal basis in $\mathbb{R}^n$. We put $u_1 = v_1$. Take u_2 the unit-length linear combination of u_1 and v_2 which has projection $s_1 u_1$ on $V_1 = \mathrm{span}(v_1)$, explicitly $u_2 = s_1 u_1 + (1 - s_1^2)^{1/2} v_2$. Take u_3 the unit-length linear combination of u_2 and v_3 which has projection $s_2 u_2$ on $V_2 = \mathrm{span}(v_1, v_2)$, explicitly $u_3 = s_2 u_2 + (1 - s_2^2)^{1/2} v_3$. Continuing in this fashion, we end up with a sequence of unit vectors $u_k \in V_k$ whose orthogonal projections on the previous V_{k-1} are

$$P_{V_{k-1}}(u_k) = s_{k-1} u_{k-1}. \tag{D.54}$$

Computing (u_k, u_l), $k < l$, via orthogonal projections $u_l \to V_{l-1} \to V_{l-2} \dots$ gives precisely $s_k \cdots s_{l-1}$. Q.E.D.

Now for any component of $d\mu_T$ measure, i.e. one particular bijection g in the BBF formula, we satisfy (D.48) via

$$r_{kl} = (u_{g^{-1}(k)}, u_{g^{-1}(l)}), \tag{D.55}$$

which is also symmetric in k, l and $r_{kk} = 1$. Finally the rescaling factor in the $\mathcal{N}$ matrix:

$$r_{k(i)k(j)} = (u_{g^{-1}(k(i))}, u_{g^{-1}(k(j))}). \tag{D.56}$$

In section D.1 we showed that $\mathcal{M}_{ij} = (f_i, h_j)$ where f, h are vectors in a Hilbert space. Considering the tensor product of those vectors with $u_{g^{-1}(k(i))} \in \mathbb{R}^n$, we obtain an inner product representation for $\mathcal{N}_{ij}$ elements. Since u's have unit length, by the same argument which led to (D.10) we obtain a bound with the same constant C_{GH}:

$$|\det \mathcal{N}| \leqslant (C_{\mathrm{GH}})^s, \quad s = \frac{1}{2}(m - 2(n-1)). \tag{D.57}$$

This is true for any r_e lying in the support of the measure $d\mu_T$. We can also integrate this bound since $d\mu_T$ has weight 1. We conclude that the connected expectation is bounded by

$$(C_{\mathrm{GH}})^s \sum_T \sum_{\mathcal{T} \text{ comes from } T} |\text{eq. (D.51)}| = (C_{\mathrm{GH}})^s \sum_{\mathcal{T}} |\text{eq. (D.51)}|\,. \tag{D.58}$$

where we used that every anchored tree comes from one and only one tree. We finally get:

Lemma D.6 (GKL bound) *Fermionic connected expectations are bounded by*

$$|\langle \Phi(\mathbf{A}_1, \mathbf{x}_1); \dots; \Phi(\mathbf{A}_n, \mathbf{x}_n)\rangle_c| \leqslant (C_{\mathrm{GH}})^s \sum_{\mathcal{T}} \prod_{\text{along } \mathcal{T}} |\Gamma_{ij}(x_i - y_j)|, \tag{D.59}$$

where $s = \frac{1}{2}\sum_{i=1}^n |\mathbf{x}_i| - (n-1)$, C_{GH} *is from* (D.10), *the sum is over all anchored trees* $\mathcal{T}$ *on* n *groups of points* $\mathbf{x}_k$, *and the product of propagators is along* $\mathcal{T}$.

The Γ_{ij} here are either propagators, or their first derivatives with respect to x_i and/or y_j. Since we are assuming (4.15), we can replace $|\Gamma_{ij}(x_i - y_j)|$ by $M(x_i - y_j)$ in the r.h.s. of the bound. See also figure 7 for an illustration. For $n = 1$ the GKL bound reduces to the Gram-Hadamard bound (D.10).

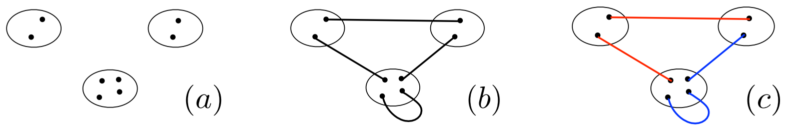


Figure 7. This illustrates the $n = 3$ case of the connected expectation. (a) Three groups of points. (b) A particular connected Wick contraction. (c) Red: an anchored tree consisting of $n-1$ propagators. Blue: remaining s propagators.

D.5 Bound on the number of anchored trees

We will prove that the number of anchored trees on n groups of points $\mathbf{x}_i$ is bounded by

$$N_{\mathcal{T}} \leqslant n! 4^{\sum_{i=1}^n |\mathbf{x}_i|}. \tag{D.60}$$

Recall that an anchored tree is a graph which becomes a tree when each group $\mathbf{x}_i$ is collapsed to a point. The graphs are labeled (i.e. the vertices are distinguishable).

By Cayley's formula, the number of labeled trees T on n points is n^{n-2}. Its proof via Prüfer sequences [130] simultaneously gives a finer result: the number of labeled trees specifying degrees d_k of each vertex k is

$$\frac{(n-1)!}{(d_1-1)! \cdots (d_n-1)!}. \tag{D.61}$$

To get an anchored tree $\mathcal{T}$ out of T, we choose a propagator for each edge $e = kl$. There are at most $m_k m_l$ choices per edge, with $m_k = |\mathbf{x}_k|$ the number of points in the group k, so at most $\prod_{k=1}^n (m_k)^{d_k}$ choices in total. Multiplying by (D.61) and summing over all possible degrees, gives an upper bound on the number of anchored trees:

$$N_{\mathcal{T}} \leqslant (n-1)! \sum_{(d_k)_{k=1}^n} \prod_{k=1}^n \frac{(m_k)^{d_k}}{(d_k-1)!}. \tag{D.62}$$

Summing over each d_k independently from 1 to ∞ (which is an overestimate since e.g. degrees are limited from above by $n-1$), results in a further bound

$$N_{\mathcal{T}} \leqslant (n-1)! \prod_{k=1}^n F(m_k), \qquad F(x) = \sum_{p=1}^{\infty} \frac{x^p}{(p-1)!} = x e^x. \tag{D.63}$$

We have $F(x) \leqslant B^x$ with $B = e^{1+1/e} \approx 3.93$, and so (D.60) follows. For a proof not relying on (D.61) see [121], Lemma 2.4.

E Proof of $S_l^{\ell_1,\ldots,\ell_n}$ norm bound

Our goal here is to prove the bound (5.38). Let $h_i \in H_{\ell_i}$ and denote $\widetilde{H}_l = S_l^{\ell_1,\ldots,\ell_n}(h_1,\ldots,h_n) \in B_l$. Unpacking the definition in section 5.1, the kernels of $\widetilde{H}_l$ are

$$\widetilde{H}_l(\mathbf{B}, \mathbf{x_B}) = \mathcal{A} \frac{1}{n!} \sum_{\substack{\mathbf{B}_1,\ldots,\mathbf{B}_n \\ \sum \mathbf{B}_i = \mathbf{B}}} \sum_{\substack{\mathbf{A}_1,\ldots,\mathbf{A}_n \\ \mathbf{A}_i \supset \mathbf{B}_i}} (-)^{\#} K_{(\mathbf{B}_i,\mathbf{A}_i)_1^n}(\mathbf{x_B}), \tag{E.1}$$

$$K_{(\mathbf{B}_i,\mathbf{A}_i)_1^n}(\mathbf{x_B}) = \int d^d \mathbf{x}_{\overline{\mathbf{B}}} \mathcal{C}(\mathbf{x}_{\overline{\mathbf{B}}}) \prod_{i=1}^n h_i(\mathbf{A}_i, \mathbf{x}_{\mathbf{A}_i}).$$

Let us count the terms in the sum (E.1) corresponding to a fixed $\mathbf{B}$ and fixed lengths $|\mathbf{A}_i| = l_i$. It is easy to see that these terms are in one-to-one correspondence with sequences $\mathbf{R}$ of total length $l_1 + l_2 + \cdots + l_n$ extending the sequence $\mathbf{B}$. [Every such sequence $\mathbf{R}$ can be cut into sequences $\mathbf{A}_1, \ldots, \mathbf{A}_n$, uniquely since the lengths $|\mathbf{A}_i|$ are kept fixed, and then $\mathbf{B}_i$ can be extracted, uniquely, as the part of $\mathbf{B}$ falling into $\mathbf{A}_i$.] It follows that the number of terms in (E.1) is bounded by:

$$\text{Number of terms in (E.1)} \;=\; \binom{\sum l_i}{l} (Nd+N)^{\sum l_i - l} \leqslant 2^{\sum l_i} \times (Nd+N)^{\sum l_i}, \qquad \text{(E.2)}$$

where in the first equality $Nd+N$ is the maximal number of choices for every element of the sequence $\mathbf{R}\backslash\mathbf{B}$, see (4.1), assuming that they are assigned independently. [This counting does not take into account that some of these terms would vanish by constraints imposed by the rotation and $\mathrm{Sp}(N)$ invariances.]

We next pick some $\mathbf{B}_i, \mathbf{A}_i$ and consider the corresponding term $K(\mathbf{x_B}) = K_{(\mathbf{B}_i,\mathbf{A}_i)_1^n}(\mathbf{x_B})$ in (E.1). Its integration kernel $\mathcal{C}(\mathbf{x}_{\overline{\mathbf{B}}})$, eq. (5.4), is bounded by the GKL bound (D.59) with $\mathbf{x}_i \equiv \mathbf{x}_{\overline{\mathbf{B}}_i}$:

$$|\mathcal{C}(\mathbf{x}_{\overline{\mathbf{B}}})| = \left|\left\langle \Phi(\overline{\mathbf{B}}_1, \mathbf{x}_{\overline{\mathbf{B}}_1}); \ldots; \Phi(\overline{\mathbf{B}}_n, \mathbf{x}_{\overline{\mathbf{B}}_n})\right\rangle_c\right| \leqslant (C_{\mathrm{GH}})^s \sum_{\mathcal{T}} \prod_{(xx')\in\mathcal{T}} M(x-x'), \qquad \text{(E.3)}$$

where $s = \frac{1}{2}\sum\left|\overline{\mathbf{B}}_i\right| - (n-1) \leqslant \frac{1}{2}\sum l_i$. We wish to bound the norm of $K(\mathbf{x_B})$:

$$\|K\|_w = \int_{x_1=0} d^d\mathbf{x_B} |K(\mathbf{x_B})|\, w(\mathbf{x_B}) \qquad \text{(E.4)}$$

with the weight function (4.16). Let $\mathcal{T}$ be any anchored tree in (E.3). If τ_i are trees connecting points of $\mathbf{x}_{\mathbf{A}_i}$, then $\mathcal{T} \cup \tau_1 \cup \ldots \cup \tau_n$ connects points in $\mathbf{x_B}$. Therefore, we have a bound:

$$\mathrm{St}(\mathbf{x_B}) \leqslant \sum_{i=1}^n \mathrm{St}(\mathbf{x}_{\mathbf{A}_i}) + \prod_{(xx')\in\mathcal{T}} |x-x'|. \qquad \text{(E.5)}$$

Raising this to the power σ and using the elementary inequality

$$(\sum p_i)^\sigma \leqslant \sum p_i^\sigma \qquad (p_i \geqslant 0, 0<\sigma\leqslant 1), \qquad \text{(E.6)}$$

we conclude

$$w(\mathbf{x_B}) \leqslant \prod_{i=1}^n w(\mathbf{x}_{\mathbf{A}_i}) \prod_{(xx')\in\mathcal{T}} w(\{x,x'\}). \qquad \text{(E.7)}$$

Using in (E.4) this bound, the definition of $K(\mathbf{x_B})$, and the bound (E.3), we get

$$\begin{aligned}\|K\|_w &\leqslant (C_{\mathrm{GH}})^s \sum_{\mathcal{T}} \int_{x_1=0} d^d\mathbf{x}_{\mathbf{B}\cup\overline{\mathbf{B}}} \prod_{(xx')\in\mathcal{T}} M(x-x')w(\{x,x'\}) \prod_{i=1}^n h_i(\mathbf{A}_i, \mathbf{x}_{\mathbf{A}_i}) w(\mathbf{x}_{\mathbf{A}_i}) \\ &= (C_{\mathrm{GH}})^s N_{\mathcal{T}} \|M\|_w^{n-1} \prod_{i=1}^n \|h_i\|_w. \end{aligned} \qquad \text{(E.8)}$$

The latter equality is shown via an "amputating tree leaves" argument. Given an anchored tree $\mathcal{T}$, we can find a leaf: a group of points $\mathbf{x}_{\overline{\mathbf{B}}_k} \subset \mathbf{x}_{A_k}$ connected to the rest by just one edge of $\mathcal{T}$, call it (zz') where $z \in \mathbf{x}_{A_k}$. Amputating the leaf consists of two steps. First, keep z fixed and integrate over all the other leaf points, which gives a factor $\|h_k\|_w$. Second, integrate over z (keeping z' fixed), which gives a factor $\|M\|_w$. Then find the next leaf and continue the amputation. We get the same result for each $\mathcal{T}$. By appendix D.5, the number of anchored trees

$$N_{\mathcal{T}} \leqslant n!4^{\sum |\mathbf{x}_{\overline{\mathbf{B}}_i}|} = n!4^{\sum l_i - l} \leqslant n!4^{\sum l_i}.$$

To summarize, the number of terms in (E.1) is bounded by (E.2), the norm of each individual term by (E.8), s by $\frac{1}{2}\sum l_i$ and $N_{\mathcal{T}}$ by $n!4^{\sum l_i}$. The latter $n!$ cancels with $\frac{1}{n!}$ in (E.1). A further useful fact is that the antisymmetrization operator $\mathcal{A}$ (footnote 15) does not increase the norm (it averages over all permutations with signs, and the norm is defined as the maximum over all permutations). Taking all of these into account, we get a bound

$$\|\tilde{H}_l\|_w \leqslant C_\gamma^{n-1} \prod_{i=1}^{n} C_0^{l_i} \|h_i\|_w \tag{E.9}$$

with $C_\gamma = \|M\|_w$, $C_0 = 8(Nd + N)\sqrt{C_{\mathrm{GH}}}$. This is the bound (5.38) in the case $\sum l_i \geqslant l + 2(n-1)$. In the opposite case $S_l^{\ell_1,\ldots,\ell_n}$ vanishes (section 5.1), so there is nothing to prove.

Remark E.1 We can see how some of the above steps justify choices made in the main text. The Steiner diameter is tailor-made for (E.5). Stretched exponential weight is handy because of (E.6). The L_1 norm (as opposed to any other L_p) works great when recursively amputating tree leaves.

F Estimates of $\Delta_k^{(1)}, \Delta_k^{(2)}$

In this appendix we prove Lemma 7.2. To estimate $\Delta_k^{(1)}$, we use $b_k \leqslant A\delta^k$for $k \geqslant 1$, sum the geometric progression, and use $C\delta \leqslant 1/2$, which gives what we need:

$$\Delta_k^{(1)} \leqslant C^{k+2} A \delta^{k+1} \frac{1}{1 - C\delta} \leqslant 2C^{k+2} A\delta^{k+1} \qquad (k \geqslant 0). \tag{F.1}$$

The estimate of $\Delta_k^{(2)}$ is more subtle and will require several steps. First we introduce the tool of **extending by zeros**. For any sequence $\varkappa = (k_i)_1^n$, we say that a sequence $\varkappa'$ "extends $\varkappa$ by m zeros" if it is obtained from $\varkappa$ inserting m zeros in arbitrary places. This definition is also valid if $\varkappa$ itself already contains some zeros (this will be useful). For $m = 0$ we have $\varkappa' = \varkappa$. E.g. (0,2,0) extends (2) by 2 zeros, and (0,1,0,1) extends $(0, 1, 1)$ by 1 zero.

We define $F_{\mathrm{ext}}[\varkappa]$ as the sum of $F[\varkappa']$ over all $\varkappa'$ extending $\varkappa$ by an arbitrary number of zeros $m \geqslant 0$:

$$F_{\mathrm{ext}}[\varkappa] = \sum_{m=0}^{\infty} \sum_{\varkappa' \text{ :extends } \varkappa \text{ by } m \text{ zeros}} F[\varkappa']. \tag{F.2}$$

For a fixed m, the number of sequences $\varkappa'$ is $\leqslant \binom{m+n}{m}$ where n is the length of $\varkappa$. [$= \binom{m+n}{m}$ if the original sequence $\varkappa$ does not have zeros.] Since they all have $F[\varkappa'] = F[\varkappa](C_\gamma C b_0)^m$, we obtain

$$F_{\rm ext}[\varkappa] \leqslant F[\varkappa] \sum_{m=0}^{\infty} \binom{m+n}{m} (C_\gamma C b_0)^m = \frac{1}{(1-C_\gamma C b_0)^{n+1}} F[\varkappa].$$

Recall $b_0 \leqslant A\delta$. Thus using $C_\gamma C A\delta \leqslant 1/2$ we have

$$F_{\rm ext}[\varkappa] \leqslant 2^{n(\varkappa)+1} F[\varkappa]. \tag{F.3}$$

Next we will **sum over sequences with a fixed $\sum k_i$**. Namely we define

$$\Phi_k = \sum_{n=2}^{\infty} \sum_{(k_i)_1^n, \sum k_i = k} F[(k_i)_1^n] \quad (k \geqslant 0) \tag{F.4}$$

We first estimate Φ_k's and then convert into the estimate for $\Delta_k^{(2)} = \sum_{k'=k}^{\infty} \Phi_{k'}$.

Consider first $k \geqslant 2$ (see below for the simpler $k = 0, 1$). We can obtain all sequences in Φ_k extending by zeros sequences which satisfy $\sum k_i = k$, $k_i \geqslant 1$ (and whose length is therefore at most k). Thus

$$\Phi_k \leqslant \sum_{n=1}^{k} \sum_{(k_i)_1^n, \sum k_i = k, k_i \geqslant 1} F_{\rm ext}[(k_i)_1^n] \leqslant \sum_{n=1}^{k} 2^{n+1} \sum_{(k_i)_1^n, \sum k_i = k, k_i \geqslant 1} F[(k_i)_1^n], \tag{F.5}$$

where we used (F.3). Note that although $n \geqslant 2$ in (F.4), we need $n \geqslant 1$ in (F.5) to include the one-term sequence (k) whose two-term extensions $(k, 0)$ and $(0, k)$ appear in (F.4).[58]

Using $b_{k_i} \leqslant A\delta^{k_i}$ and $\sum k_i = k$ we have, for sequences with all $k_i \geqslant 1$,

$$F[(k_i)_1^n] \leqslant (C_\gamma)^{n-1} \prod_{i=1}^{n} C^{k_i+1} A\delta^{k_i} = (C_\gamma)^{n-1} C^{k+n} A^n \delta^k = (C_\gamma C A)^{n-1} C^{k+1} A\delta^k. \tag{F.6}$$

Finally, by elementary combinatorics the number of n-term sequences $(k_i)_1^n$ satisfying $\sum k_i = k, k_i \geqslant 1$ is $\binom{k-1}{n-1}$. Plugging all this information into (F.5), we have

$$\Phi_k \leqslant 4C^{k+1} A\delta^k \sum_{n=1}^{k} \binom{k-1}{n-1} (2C_\gamma C A)^{n-1} = 4C^{k+1} A\delta^k (1 + 2C_\gamma C A)^{k-1}, \tag{F.7}$$

which implies

$$\Phi_k \leqslant 2C(2C\delta)^k A \qquad (k \geqslant 2). \tag{F.8}$$

once we use

$$C_\gamma C A \leqslant 1/2. \tag{F.9}$$

For future use, let us derive a bound on $F[(k_i)]$, $F_{\rm ext}[(k_i)]$ under the condition (F.9), this time allowing for $k_i \geqslant 0$. Using $b_k \leqslant A\delta^{\max(k,1)}$ we then have the same bound (F.6) but

[58] Their contribution is suppressed by b_0 compared to (k), but taking this suppression into account does not lead to a better estimate because of other terms present in (F.5).

with an extra factor δ^m where m is the number of zeros in the sequence. Using (F.9), (F.3) we have

$$F[(k_i)_{i=1}^n] \leqslant 2^{1-n} C^{k+1} A\delta^{k+m} \qquad \left(\sum k_i = k, k_i \geqslant 0, m = \#\{k_i = 0\}\right), \tag{F.10}$$

$$F_{\text{ext}}[(k_i)_1^n] \leqslant 4C^{k+1} A\delta^{k+m}. \tag{F.11}$$

Finally let us bound Φ_k for $k = 0, 1$. For $k = 0$, eq. (F.4) involves the sequence $(0, 0)$ and its extensions by zeros. We then have by (F.11):

$$\Phi_0 = F_{\text{ext}}[(0, 0)] \leqslant 4CA\delta^2. \tag{F.12}$$

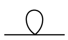
For $k = 1$, eq. (F.4) involves the sequences (1,0), (0,1) and their extensions by zeros. By (F.11):

$$\Phi_1 = 2F_{\text{ext}}[(1, 0)] \leqslant 8C^2 A\delta^2. \tag{F.13}$$

Finally we estimate $\Delta_k^{(2)} = \sum_{k'=k}^{\infty} \Phi_{k'}$ summing (F.8) in geometric progression, which is possible since $2C\delta \leqslant 1/2$, and adding (F.12), (F.13) when needed. We thus obtain:

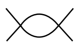
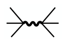
$$\begin{aligned} \Delta_k^{(2)} &\leqslant C(2C\delta)^k A \qquad (k \geqslant 2), \\ \Delta_1^{(2)} &= \Phi_1 + \Delta_2^{(2)} \leqslant (8C^2 + 16C^3)A\delta^2, \\ \Delta_0^{(2)} &= \Phi_0 + \Phi_1 + \Delta_2^{(2)} \leqslant (4C + 8C^2 + 16C^3)A\delta^2. \end{aligned} \tag{F.14}$$

G One-loop coefficients I_1 and I_2

Here we evaluate the coefficients I_1, I_2 in the fixed-point equations from section 5.5, and prove their needed properties. Thinking in momentum space, computing the contributions to the effective ν and λ shown in (5.31) involves setting to zero the external momenta in the corresponding Feynman diagrams (see footnote 26). To compute I_1 we consider the one-loop diagram  , where the vertex is the local quartic coupling and the field propagating in the loop is the fluctuating component φ. This gives

$$I_1 = 2(N-2) \int \frac{d^d k}{(2\pi)^d} \frac{\rho(k)}{|k|^{d/2+\varepsilon}}, \tag{G.1}$$

where we denoted $\rho(k) = \chi(k) - \chi(\gamma k)$. To work out the combinatorial prefactor $2(N-2)$, rewrite the local quartic as $Q(\psi) = \frac{1}{3} q_{abcd}\psi_a\psi_b\psi_c\psi_d$ (integration over x understood) with $q_{abcd} = \Omega_{ab}\Omega_{cd} - \Omega_{ac}\Omega_{bd} + \Omega_{ad}\Omega_{bc}$ totally antisymmetric. Then $Q(\psi + \phi)$ contains the quadratic in ϕ term $2q_{abcd}\psi_a\psi_b\phi_c\phi_d$, from where we get the term $2q_{abcd}\Omega_{cd}\psi_a\psi_b = 2(N - 2)\Omega_{ab}\psi_a\psi_b$in the effective action (times the k-integral). Identities $\Omega_{ac}\Omega_{bc} = \delta_{ab}$, $\Omega_{ab}\Omega_{ab} = N$ are helpful.

Similarly, to compute I_2 we consider the diagrams  (with two local quartic vertices and two ϕ-propagators) and  (where we start with the diagram  describing the $H_{6\text{SL}}$ interaction, with the wavy line denoting $\hat{\mathfrak{X}}_*(k)$, and contract two out of six external vertices

JHEP01(2021)026

with the φ-propagator). Their contributions to I_2 are given by

$$
\begin{aligned}
I_2^{\bowtie} &= -4(N-8)\int \frac{d^dk}{(2\pi)^d}\frac{\rho(k)^2}{|k|^{d+2\varepsilon}},\\
I_2^{\mathfrak{X}} &= -8(N-8)\int \frac{d^dk}{(2\pi)^d}\frac{\rho(k)R(k)}{|k|^{d+2\varepsilon}},
\end{aligned}
\tag{G.2}
$$

where we wrote the Fourier transform of $\mathfrak{X}_*(x)$ as

$$
\widehat{\mathfrak{X}}_*(k) = -\,8\lambda^2\frac{R(k)}{|k|^{\frac{d}{2}+\varepsilon}}, \qquad R(k) = \sum_{n=1}^{\infty}\gamma^{4\varepsilon n}\rho(k/\gamma^n). \tag{G.3}
$$

To compute the prefactor in $I_2^{\bowtie}$, we use $Q(\psi+\phi) = 2q_{abcd}\psi_a\psi_b\phi_c\phi_d$ and the identity

$$
q_{abcd}q_{a'b'c'd'}\Omega_{cc'}\Omega_{dd'} = (N-4)\Omega_{ab}\Omega_{a'b'} + 2\Omega_{aa'}\Omega_{bb'} + 2\Omega_{ab'}\Omega_{a'b}. \tag{G.4}
$$

The prefactor in $I_2^{\mathfrak{X}}$ can be computed similarly.[59] Thus

$$
I_2 = I_2^{\bowtie} + I_2^{\mathfrak{X}} = -4(N-8)\int \frac{d^dk}{(2\pi)^d}\frac{\rho(k)^2 + 2\rho(k)R(k)}{|k|^{d+2\varepsilon}}. \tag{G.5}
$$

Now that we computed I_1 and I_2, let us discuss their properties. $I_1 \propto N-2$ is explained by the fact that for $N=2$ the quartic interaction vanishes (see the Introduction). That $I_2 \neq 0$ for $N=2$ is not a contradiction, since for a vanishing quartic interaction the change in λ is anyway unphysical.

Note that perturbative beta-functions of symplectic fermion models can be obtained from the beta-functions of bosonic $O(N_b)$ models by setting formally $N_b = -N$. This is valid for local models [44, 48] and extends to nonlocal (long-range) models considered here. The $(N-8)$ factor in I_2 is thus related to the well-known (N_b+8) factor in the bosonic $O(N_b)$ model one-loop beta-function [131]. The long-range bosonic $O(N_b)$ beta function is known at three loops [78].[60] The two-loop term in [78], (3.16) is proportional to $5N_b+22$ and does not vanish for $N_b \to -8$.

Therefore, vanishing of I_2 for $N=8$ is an accident unrelated to any symmetry, which does not repeat in higher orders. With vanishing λ^2 and nonzero λ^3 term in λ', there will be a perturbative fixed point with $\lambda = O(\sqrt{\varepsilon})$ at $N=8$, and it should be possible to justify its existence non-perturbatively (see section 8.1.1). In this paper we stick to the generic case $N \neq 8$.

Because the integrals in (G.1) and (G.5) are cutoff at both UV and IR momenta, I_1 and I_2 depend analytically on ε. At small ε they behave as follows:

[59]The following argument explains why it is the double of $I_2^{\bowtie}$. Let $Q_a = (\delta/\delta\psi_a)Q$, $Q_{ab} = (\delta^2/\delta\psi_a\delta\psi_b)Q$ be the first and second functional derivatives of $Q(\psi)$ at $\psi=0$. The $\bowtie$ diagram is $-\frac{1}{4}Q_{ab}Q_{cd}\Omega_{ac}\Omega_{bd}$ times the loop integral. The $\mathfrak{X}$-term can be written as $-\frac{1}{2}\int d^dxd^dy\, Q_a(x)Q_c(y)\Omega_{ac}\mathfrak{X}(x-y)$. It is then clear that the $I_2^{\mathfrak{X}}$ diagram gives $-\frac{1}{2}Q_{ab}Q_{cd}\Omega_{ac}\Omega_{bd}$ times the corresponding loop integral.

[60]From the two-loop level it differs from the local bosonic $O(N_b)$ beta function (e.g. [132],(11.98)). Even the dependence on N_b is different, due to the absence of wavefunction renormalization diagrams in the nonlocal model.

JHEP01(2021)026

Lemma G.1 *We have, for $|\varepsilon| \leqslant 1/\log\gamma$,*

$$I_1 = 2(N-2)\left[(1-\gamma^{-d/2})\int \frac{d^dk}{(2\pi)^d}\frac{\chi(k)}{|k|^{d/2}} + O(\varepsilon\log\gamma)\right], \tag{G.6}$$
$$I_2 = -4(N-8)\left[\frac{S_d}{(2\pi)^d}\log\gamma + O(\varepsilon(\log\gamma)^2)\right], \tag{G.7}$$

where S_d is the area of the unit sphere in $\mathbb{R}^d$ and the constants in O are γ- and ε-independent.

Note that I_2 is χ-independent as $\varepsilon \to 0$, even though $I_2^{\bowtie}$ and $I_2^{\mathfrak{X}}$ separately depend on χ. This is not accidental, see remark 6.5.

Proof. The $|k|^{-d/2-\varepsilon}$ factor in the integrand in (G.1) can be estimated as follows:

$$\tfrac{1}{|k|^{d/2+\varepsilon}} = \tfrac{1}{|k|^{d/2}} + \tfrac{|k|^{-\varepsilon}-1}{|k|^{d/2}} = \tfrac{1}{|k|^{d/2}} + \tfrac{O(\varepsilon\log\gamma)}{|k|^{d/2}} \qquad (1/(2\gamma) \leqslant |k| \leqslant 1). \tag{G.8}$$

The integral in (G.1) is thus given, modulo $O(\varepsilon\log\gamma)$ error, by

$$\int \frac{d^dk}{(2\pi)^d}\frac{\rho(k)}{|k|^{d/2}} = (1-\gamma^{-d/2})\int \frac{d^dk}{(2\pi)^d}\frac{\chi(k)}{|k|^{d/2}}, \tag{G.9}$$

thus proving (G.6). Let us next prove (G.7). Recall that $\gamma \geqslant 2$. We have

$$\begin{aligned} &\operatorname{supp}\rho \subset \{1/(2\gamma) \leqslant |k| \leqslant 1\},\\ &\rho \equiv 1 \quad \text{on} \quad \{1/\gamma \leqslant |k| \leqslant 1/2\}. \end{aligned} \tag{G.10}$$

From here it follows that $R(k) = \gamma^{4\varepsilon}\rho(k/\gamma) = \gamma^{4\varepsilon}(1-\chi(k))$ on the support of ρ. We then rewrite the numerator of the integrand in (G.5) as:

$$\begin{aligned} \rho(k)^2 + 2\rho(k)R(k) &= \rho(k)^2 + 2\rho(k)(1-\chi(k))[1+(\gamma^{4\varepsilon}-1)]\\ &= \rho(k)^2 + 2\rho(k)(1-\chi(k)) + O(\varepsilon\log\gamma),\\ &= F(k) - F(\gamma k) + O(\varepsilon\log\gamma), \qquad F(k) = 2\chi(k) - \chi(k)^2, \end{aligned} \tag{G.11}$$

and the $|k|^{-d-2\varepsilon}$ factor similarly to (G.8). Collecting the error terms, the integral in (G.5) is

$$\int \frac{d^dk}{(2\pi)^d}\frac{F(k)-F(\gamma k)}{|k|^d} + O(\varepsilon(\log\gamma)^2). \tag{G.12}$$

Like $\chi(k)$, the function $F(k) \equiv 1$ for $|k| \leqslant 1/2$ and vanishes for $|k| \geqslant 1$. The integral in (G.12) can now be computed:

$$\int \frac{d^dk}{(2\pi)^d}\frac{F(k)-F(\gamma k)}{|k|^d} = \frac{S_d}{(2\pi)^d}\log\gamma, \tag{G.13}$$

e.g. by separating the integration region into the region close to the origin plus the rest, and using the properties of $F(k)$ given above. In particular, the answer is χ-independent. Q.E.D.

JHEP01(2021)026

H Finite volume and non-perturbative validity of H_{eff}

In this appendix we will provide details mentioned in Remarks 2.1 and 5.1, regarding a definition of our model in finite volume, and a rigorous derivation of eq. (5.3). Our plan is as follows. First we consider general aspects of gaussian and interacting Grassmann fields in finite volume and with a UV cutoff, and explain that because the effective number of Grassmann variables is finite, all path integrals are manifestly well defined. Then we apply this to the effective action in finite volume, and show that perturbation theory, if convergent, gives the correct answer. Finally we pass to the infinite volume limit and show that it agrees with eq. (5.3).

General aspects. Working in the finite volume $\mathcal{V} = [-V/2, V/2]^d$ with periodic boundary conditions, we Fourier-expand the fields as

$$\psi_a(x) = \frac{1}{V^d} \sum_{k \in K_V} \psi_{a,k} e^{ikx}, \tag{H.1}$$

where $K_V = (2\pi/V)\mathbb{Z}^d \cap \operatorname{supp} \chi$ is the finite set of Fourier momenta which belong to $\operatorname{supp} \chi$. We truncate away all other momenta because they have zero propagator. The finite volume gaussian measure $d\mu_{P,V}$ is a finite-dimensional measure over Grassmann variables $\psi_{a,k}$:

$$\begin{aligned} d\mu_{P,V}(\psi) &= \mathrm{Pf}^{-1} \prod_{\psi_{a,k}, k \in K_V} d\psi_{a,k} e^{S_{2,V}(\psi)}, \\ S_{2,V} &= \frac{1}{2V^d} \sum_{k \in K_V} \hat{P}(k)^{-1} \Omega_{ab} \psi_{a,k} \psi_{b,-k}, \end{aligned} \tag{H.2}$$

where the normalization factor $\mathrm{Pf} > 0$ is the Pfaffian of $S_{2,V}$ which is an antisymmetric quadratic form in $\psi_{a,k}$'s. This is a meaningful finite volume version of the formal eq. (2.4) in infinite volume. The propagator $\langle \psi_a(x) \psi_b(y) \rangle = \Omega_{ab} P_V(x-y)$ where P_V is the periodic version of (2.1):

$$P_V(x) = \frac{1}{V^d} \sum_{k \in (2\pi/V)\mathbb{Z}^d} \hat{P}(k) e^{ikx}, \quad \hat{P}(k) \text{ as in (2.1)}. \tag{H.3}$$

For fixed x and $V \to \infty$, we have $P_V(x) \to P(x)$, just as for the higher-order expectations. In this sense we can say that $d\mu_{P,V} \to d\mu_P$, the infinite volume measure defined in section 2.

The interacting Grassmann measure is then defined as

$$Z_V^{-1} d\mu_{P,V}(\psi) e^{sH_V(\psi)}, \tag{H.4}$$

where H_V is given by the finite volume analogue of (4.2):

$$H_V(\psi) = \sum_{\mathbf{A}} \int_{\mathcal{V}^l} d^d\mathbf{x} H_V(\mathbf{A}, \mathbf{x}) \Psi(\mathbf{A}, \mathbf{x}), \tag{H.5}$$

and $Z_V = \int d\mu_{P,V}(\psi) e^{sH_V(\psi)}$ is the partition function. Factor s multiplying $H_V(\psi)$ in (H.4) is for further convenience, eventually we will set $s = 1$. Using Fourier expansion (H.1), we express $H_V(\psi)$ as a series in $\psi_{a,k}$'s with finite coefficients (assuming that the kernels $H_V(\mathbf{A},\mathbf{x})$ are in L_1). By eq. (H.2) and the usual rules of Grassmann integration, Z_V and $d\mu_{P,V}(\psi)e^{sH_V(\psi)}$ are well defined and are polynomials in s, because there are only finitely many $\psi_{a,k}$'s and only finitely many terms from the Taylor expansion of $e^{sH_V(\psi)}$ will contribute. In particular Z_V is finite. The measure (H.4) will therefore be well defined as long as $Z_V \neq 0$.

Effective action in finite volume. Consider next eq. (2.13) in finite volume. Defining $d\mu_{g,V}(\phi)$ as $d\mu_{P,V}(\psi)$, we consider

$$I(s,\psi) = \int d\mu_{g,V}(\phi)e^{sH_V(\psi+\phi)}. \tag{H.6}$$

By the arguments as above, we have $I(s,\psi) = e^{sH_V(\psi)}p(s,\psi)$ where $p(s,\psi)$ is a polynomial in s.

We would like to find $H^V_{\text{eff}}(s,\psi)$ so that

$$e^{H^V_{\text{eff}}(s,\psi)} = I(s,\psi). \tag{H.7}$$

Let us define $H^V_{\text{eff}}(s,\psi)$ by the perturbative expansion (see eqs. (5.1), (5.3)):

$$H^V_{\text{eff}}(s,\psi) = \sum_{\mathbf{B}} \int_{\mathcal{V}^{|\mathbf{B}|}} d^d\mathbf{x} H^V_{\text{eff}}(s,\mathbf{B},\mathbf{x_B})\, \Psi(\mathbf{B},\mathbf{x_B}), \tag{H.8}$$

where $H^V_{\text{eff}}(s,\mathbf{B},\mathbf{x_B})$ are given by the finite volume analogue of eq. (5.3) replacing $H \to sH_V$:

$$H^V_{\text{eff}}(s,\mathbf{B},\mathbf{x_B}) = \mathcal{A}\sum_{n=1}^{\infty}\frac{s^n}{n!}\sum_{\sum \mathbf{B}_i = \mathbf{B},\mathbf{A}_i\supset\mathbf{B}_i}(-)^{\#}\int_{\mathcal{V}^{|\overline{\mathbf{B}}|}} d^d\mathbf{x}_{\overline{\mathbf{B}}}\mathcal{C}_V\left(\mathbf{x}_{\overline{\mathbf{B}}}\right)\prod_{i=1}^{n} H_V(\mathbf{A}_i,\mathbf{x}_{\mathbf{A}_i}), \tag{H.9}$$

where $\mathcal{C}_V$ is as in (5.4) only with finite-volume propagators. From the arguments like in Key lemma, we will be able to show that this series converges and defines $H^V_{\text{eff}}(s,\mathbf{B},\mathbf{x_B})$ as analytic L_1-valued functions in the disk $|s| < 2$ (Lemma H.1 below, Part (b)). Since, by perturbation theory, $e^{H^V_{\text{eff}}(s,\psi)}$ and $I(s,\psi)$ have the same Taylor series in s, we conclude eq. (H.7) is satisfied in the disk $|s| < 2$ where they are both analytic, in particular at $s = 1$. This proves that (H.9) gives the correct effective action in finite volume.

Effective action in infinite volume. For an infinite volume interaction $H(\psi)$ given by (4.2), we consider the corresponding finite-volume interaction (H.5) with kernels given by periodization (we are assuming translational invariance):

$$H_V(\mathbf{A},(0,x_2,\dots,x_l)) = \sum_{r_i\in\mathbb{Z}^d, i=2\dots l} H(\mathbf{A},(0,x_2+r_2V,\dots,x_l+r_lV)). \tag{H.10}$$

To prove that (5.3) is the correct effective action in infinite volume, we will show that it can be obtained as a $V\to\infty$ limit of the kernels of H^V_{eff}, in the precise sense of Part (c) of the following lemma. (Part (b) was used above to justify the effective action in finite volume.)

JHEP01(2021)026

Lemma H.1 *There exists $A > 0$ and $\delta > 0$ such that, for any infinite volume interaction satisfying*

$$\|H_l\|_w \leqslant A\delta^{\min(1,l/2-1)} \qquad (l \geqslant 2), \tag{H.11}$$

and defining the finite volume interactions by (H.10) for any $V \geqslant 1$, we have

(a) the kernels of H_{eff} and of H^V_{eff} given by eqs. (5.3) and by (H.9) with $s = 1$ are well defined (the series is convergent in L_1);

(b) the kernels of $H^V_{\text{eff}}(s)$ defined by (H.9) are well defined and analytic L_1-valued functions in the disk $|s| < 2$;

(c) for any $\mathbf{B}$ we have $H^V_{\text{eff}}(\mathbf{B}, \mathbf{x}) \to H_{\text{eff}}(\mathbf{B}, \mathbf{x})$ as $V \to \infty$ in the sense of L_1 norm on any fixed bounded subset of $(\mathbb{R}^d)^l$.

Proof. Claim (a) for H^V_{eff} is a consequence of (b), which we prove as follows: consider the L_1 norm (with, as usual, one of the points fixed to the origin) of the n-th term of the series for H^V_{eff},

$$\sum_{\sum \mathbf{B}_i = \mathbf{B}, \mathbf{A}_i \supset \mathbf{B}_i} \frac{|s|^n}{n!} \int_{\mathcal{V}^{|\mathbf{A}|}, x_1=0} d^d\mathbf{x}_\mathbf{A} \, |\mathcal{C}_V(\mathbf{x}_{\overline{\mathbf{B}}})| \prod_{i=1}^n |H_V(\mathbf{A}_i, \mathbf{x}_{\mathbf{A}_i})| \tag{H.12}$$

where $\mathbf{A} = \mathbf{A}_1 + \cdots + \mathbf{A}_n$. Recall that $\mathcal{C}_V(\mathbf{x}_{\overline{\mathbf{B}}})$ is as in (5.4) with finite volume propagator g_V replacing g. Here:

$$g_V(x) = \frac{1}{V^d} \sum_{k \in (2\pi/V)\mathbb{Z}^d} \hat{g}(k) e^{ikx} = \sum_{r \in \mathbb{Z}^d} g(x + rV). \tag{H.13}$$

From the Fourier representation of g_V (first equality in (H.13)), we see that g_V can be written in Gram form, as in (D.8), with f_i and h_i as in (D.9), with the only difference that the finite volume scalar product between f_i and h_j should be interpreted as $(f_i, h_j) = \frac{1}{V^d} \sum_{k \in (2\pi/V)\mathbb{Z}^d} \overline{\hat{f}_i(k)} \hat{h}_j(k)$. Therefore, the Gram-Hadamard bound (D.10) holds, with C_{GH} replaced by $C_{\text{GH},V}$, which is defined by the same expression as C_{GH}, modulo the replacement of $\int \frac{d^dk}{(2\pi)^d}$ by the corresponding Riemann sum. Moreover, from the real space representation of g_V (second equality in (H.13)) and (4.15), we see that g_V satisfies a bound analogous to (4.15) itself, with $|x/\gamma|$ replaced by $\|x\|/\gamma$ and $\|x\| = \min_{r \in \mathbb{Z}^d} |x + rV|$ the norm on the torus, and with the constant $C_{\chi 1}$ replaced by a larger one, but still independent of V and γ. We denote by $M_V(x)$ the analogue of the right side of (4.15) with these two replacements. From these considerations, we see that $\mathcal{C}_V(\mathbf{x}_{\overline{\mathbf{B}}})$ is bounded as in (E.3),

$$|\mathcal{C}_V(\mathbf{x}_{\overline{\mathbf{B}}})| \leqslant (C_{\text{GH},V})^{\frac{1}{2}\sum_i l_i} \sum_{\mathcal{T}} \prod_{(xx') \in \mathcal{T}} M_V(x - x'). \tag{H.14}$$

Thanks to these considerations, proceeding as in appendix E, we get the analogue of (E.8), namely

$$\begin{aligned} &\int_{\mathcal{V}^{|\mathbf{A}|}, x_1=0} d^d\mathbf{x}_\mathbf{A} \, |\mathcal{C}_V(\mathbf{x}_{\overline{\mathbf{B}}})| \prod_{i=1}^n |H_V(\mathbf{A}_i, \mathbf{x}_{\mathbf{A}_i})| \\ &\leqslant \quad (C_{\text{GH},V})^{\frac{1}{2}\sum_i l_i} \sum_{\mathcal{T}} \int_{\mathcal{V}^{|\mathbf{A}|}, x_1=0} d^d\mathbf{x}_\mathbf{A} \prod_{(xx') \in \mathcal{T}} M_V(x - x') \prod_{i=1}^n |H_V(\mathbf{A}_i, \mathbf{x}_{\mathbf{A}_i})| , \end{aligned} \tag{H.15}$$

which, by computing the integrals, can be further bounded as

$$(\text{H.15}) \leqslant (C_{\text{GH},V})^{\frac{1}{2}\sum_i l_i} N_{\mathcal{T}} \|M_V\|_1^{n-1} \prod_{i=1}^{n} \|H_{V,l_i}\|_1. \tag{H.16}$$

Now, by using (H.10), we see that $\|H_{V,l_i}\|_1 \leqslant \|H_{l_i}\|_w$, which is bounded by $A\delta^{\min(1,l/2-1)}$, thanks to (H.11). Recalling also (E.2) and the fact that $\mathcal{N}_{\mathcal{T}} \leqslant n!4^{\sum l_i}$, we find

$$(\text{H.12}) \leqslant |s|^n A^n \|M_V\|_1^{n-1} \sum_{(l_i)_1^n} \prod_{i=1}^{n} (C_V)^{\sum l_i} \delta^{\min(1,l_i/2-1)} \tag{H.17}$$

where $C_V = 16N(d+1)C_{\text{GH},V}^{1/2}$. Note that both C_V and $\|M_V\|_1$ are uniformly bounded in V. Positive even integers l_i satisfy $\sum_i l_i \geqslant l+2(n-1)$, but here it will be enough to extend the sum to arbitrary $l_i \geqslant 2$. Therefore, with a suitable V-independent constant C, we get a bound

$$(\text{H.17}) \leqslant \left(\frac{C|s|\delta}{1-C\delta}\right)^n, \tag{H.18}$$

from which summability in n follows, for all $|s| < 2$, if δ is sufficiently small. Of course, item (a) for H_{eff} in infinite volume follows from the same argument.[61]

Remark H.1 According to the discussion after eq. (H.5), we also need to make sure that the finite volume partition function is nonzero. The constant, ψ-independent term in the effective action can be estimated by the same argument as above, and it is given by the torus volume times a convergent series, in particular it is finite. Hence the partition function, which is its exponential, is nonzero. The effective action is thus well defined. Once we know that the ψ-independent term is finite, we may drop it as we did throughout.

Let us now prove (c) (cf. [65], appendix D). We fix a bounded subset of $(\mathbb{R}^d)^l$ that, without loss of generality, we assume to be centered in the origin, and we call it $\mathcal{V}_0$. We want to prove that the sum over n of

$$\sum_{\substack{\sum \mathbf{B}_i = \mathbf{B}, \\ \mathbf{A}_i \supset \mathbf{B}_i}} \frac{1}{n!} \int_{\mathcal{V}_0^{|\mathbf{B}|}} d^d\mathbf{x}_{\mathbf{B}} \left| \int_{\mathcal{V}^{|\overline{\mathbf{B}}|}} d^d\mathbf{x}_{\overline{\mathbf{B}}} \mathcal{C}_V(\mathbf{x}_{\overline{\mathbf{B}}}) \prod_{i=1}^{n} H_V(\mathbf{A}_i, \mathbf{x}_{\mathbf{A}_i}) - \int_{\mathbb{R}^{|\overline{\mathbf{B}}|d}} d^d\mathbf{x}_{\overline{\mathbf{B}}} \mathcal{C}(\mathbf{x}_{\overline{\mathbf{B}}}) \prod_{i=1}^{n} H(\mathbf{A}_i, \mathbf{x}_{\mathbf{A}_i}) \right| \tag{H.19}$$

goes to zero as $V \to \infty$. We will in fact prove that the sum of (H.19) over n goes to zero exponentially fast in V as $V \to \infty$. We rewrite the integral over $\mathcal{V}^{|\overline{\mathbf{B}}|}$ by multiplying the integrand by $1 = 1(\text{St}_V(\mathbf{x}_{\mathbf{A}}) \leqslant V/4) + 1(\text{St}_V(\mathbf{x}_{\mathbf{A}}) > V/4)$ [here, if $\mathbf{x} = (x_1, \ldots, x_l)$, the finite volume Steiner diameter $\text{St}_V(\mathbf{x})$ is the length of the shortest tree on the torus (possibly with extra vertices) which connects all the points in $\mathbf{x}$. Note that $\text{St}_V(\mathbf{x}) \leqslant \min_{\mathbf{r}\in\mathbb{Z}^{dl}} \text{St}(\mathbf{x}+\mathbf{r}V)$],

[61]This case is also a consequence of Key Lemma but we preferred to give an independent argument to demonstrate how much simpler it is to show the convergence than, as in Key Lemma, to get an optimal bound on the sum.

and similarly for the integral over $\mathbb{R}^{|\overline{\mathbf{B}}|d}$, with the finite volume Steiner diameter replaced by the standard, infinite volume, one. In view of this manipulation, we bound (H.19) from above by $|\mathcal{V}_0|(R_{1,n}+R_{2,n}+R_{3,n})$, where, letting x_1 being the first coordinate in the list $\mathbf{x_B}$:

$$\begin{aligned}
R_{1,n} &= \sum_{\substack{\sum \mathbf{B}_i = \mathbf{B} \\ \mathbf{A}_i \supset \mathbf{B}_i}} \frac{1}{n!} \int_{\mathcal{V}^{|\mathbf{A}|}, x_1=0} d^d\mathbf{x_A} \, |\mathcal{C}_V(\mathbf{x}_{\overline{\mathbf{B}}})| \prod_{i=1}^n |H_V(\mathbf{A}_i, \mathbf{x}_{\mathbf{A}_i})| \, (\mathrm{St}_V(\mathbf{x_A}) > V/4), \\
R_{2,n} &= \sum_{\substack{\sum \mathbf{B}_i = \mathbf{B} \\ \mathbf{A}_i \supset \mathbf{B}_i}} \frac{1}{n!} \int_{\mathbb{R}^{|\mathbf{A}|d}, x_1=0} d^d\mathbf{x_A} \, |\mathcal{C}_V(\mathbf{x}_{\overline{\mathbf{B}}})| \prod_{i=1}^n |H_V(\mathbf{A}_i, \mathbf{x}_{\mathbf{A}_i})| \, 1(\mathrm{St}(\mathbf{x_A}) > V/4), \qquad \text{(H.20)} \\
R_{3,n} &= \sum_{\substack{\sum \mathbf{B}_i = \mathbf{B} \\ \mathbf{A}_i \supset \mathbf{B}_i}} \frac{1}{n!} \int_{\mathbb{R}^{|\mathbf{A}|d}, x_1=0} d^d\mathbf{x_A} \, \Bigg| \mathcal{C}_V(\mathbf{x}_{\overline{\mathbf{B}}}) \prod_{i=1}^n H_V(\mathbf{A}_i, \mathbf{x}_{\mathbf{A}_i}) \\
&\qquad\qquad - \mathcal{C}(\mathbf{x}_{\overline{\mathbf{B}}}) \prod_{i=1}^n H(\mathbf{A}_i, \mathbf{x}_{\mathbf{A}_i}) \Bigg| \, 1(\mathrm{St}(\mathbf{x_A}) \leqslant V/4),
\end{aligned}$$

where, in the definition of $R_{3,n}$, we used the fact that $1(\mathrm{St}_V(\mathbf{x_A}) \leqslant V/4)$ is the same as $1(\mathrm{St}(\mathbf{x_A}) \leqslant V/4)$, provided we identify the points of the torus $\mathcal{V}$ closer than $V/4$ to the origin with the corresponding points in $\mathbb{R}^d$.

In order to bound $R_{1,n}$ we proceed as we did above for (H.12), with only a few differences: consider the analogue of (H.15) that, compared with that equation, has the additional constraint $1(\mathrm{St}_V(\mathbf{x_A}) > V/4)$ under the integral sign. In the second line, we multiply and divide each factor $M_V(x-x')$ by $w_V(x,x')$ and each factor $H_V(\mathbf{A}_i, \mathbf{x}_{\mathbf{A}_i})$ by $w_V(\mathbf{x}_{\mathbf{A}_i})$, where $w_V(\mathbf{x})$ is the finite volume analogue of $w(\mathbf{x})$, namely $w_V(\mathbf{x}) = w(\mathbf{x}) = e^{C_w(\mathrm{St}_V(\mathbf{x})/\gamma)^\sigma}$. We collect together all the factors $1/w_V(x,x')$ and $1/w_V(\mathbf{x}_{\mathbf{A}_i})$ and note that, on the support of $1(\mathrm{St}_V(\mathbf{x_A}) > V/4)$,

$$\left(\prod_{(xx')\in\mathcal{T}} \frac{1}{w_V(x,x')} \right) \left(\prod_{i=1}^n \frac{1}{w_V(\mathbf{x}_{\mathbf{A}_i})} \right) \leqslant e^{-C_w(\mathrm{St}_V(\mathbf{x_A})/\gamma)^\sigma} \leqslant e^{-C_w(V/(4\gamma))^\sigma}. \qquad \text{(H.21)}$$

Therefore, we can bound the analogue of (H.15) by the analogue of the right side of (H.16), that is

$$(C_{\mathrm{GH},V})^{\frac{1}{2}\sum_i l_i} N_{\mathcal{T}} e^{-C_w(V/(4\gamma))^\sigma} \|M_V\|_{w_V}^{n-1} \prod_{i=1}^n \|H_{V,l_i}\|_{w_V}. \qquad \text{(H.22)}$$

Note also that, thanks to (H.10) and (H.11), $\|H_{V,l_i}\|_{w_V} \leqslant \|H_{l_i}\|_w \leqslant A\delta^{\min(1,l_1/2-1)}$. Putting things together, we get the analogue of (H.18):

$$R_{1,n} \leqslant e^{-C_w(V/(4\gamma))^\sigma} \left(\frac{C\delta}{1-C\delta} \right)^n, \qquad \text{(H.23)}$$

for a suitable V-independent constant C. Clearly, for δ small enough, the sum over n of $R_{1,n}$ converges and goes to zero exponentially as $V \to \infty$. Analogous discussion and bounds are valid for $R_{2,n}$.[62]

[62]This discussion also makes clear that finite-volume convergence statements in Parts (a) and (b) can be easily generalized to weighted L_1 norm with weight w_V.

Let us now consider $R_{3,n}$. We rewrite the difference

$$\mathcal{C}_V\left(\mathbf{x}_{\overline{\mathbf{B}}}\right) \prod_{i=1}^{n} H_V(\mathbf{A}_i, \mathbf{x}_{\mathbf{A}_i}) - \mathcal{C}\left(\mathbf{x}_{\overline{\mathbf{B}}}\right) \prod_{i=1}^{n} H(\mathbf{A}_i, \mathbf{x}_{\mathbf{A}_i}). \tag{H.24}$$

in telescopic form as the sum of $n+1$ terms, in each of which either a difference $\mathcal{C}_V\left(\mathbf{x}_{\overline{\mathbf{B}}}\right) - \mathcal{C}\left(\mathbf{x}_{\overline{\mathbf{B}}}\right)$ or $H_V(\mathbf{A}_i, \mathbf{x}_{\mathbf{A}_i}) - H(\mathbf{A}_i, \mathbf{x}_{\mathbf{A}_i})$ appears. The terms with $H_V - H$ can be bounded via an analogue of (H.16), with the important difference that one of the factors $\|H_{V,l_i}\|_1$ is replaced by (denoting $\mathbf{x}_{\mathbf{A}_i} = (x_1, \ldots, x_{l_i})$ and $\mathbf{r} = (r_1, \ldots, r_{l_i})$)

$$\begin{aligned} &\int_{\mathbb{R}^{|\mathbf{A}_i|d}, x_1=0} d^d\mathbf{x}_{\mathbf{A}_i} \left|H_V(\mathbf{A}_i, \mathbf{x}_{\mathbf{A}_i}) - H(\mathbf{A}_i, \mathbf{x}_{\mathbf{A}_i})\right| 1(\mathrm{St}(\mathbf{x}_{\mathbf{A}_i}) \leqslant V/4) \qquad \text{(H.25)} \\ \leqslant &\sum_{\substack{\mathbf{r} \in \mathbb{Z}^{dl_i} \\ r_1 = 0, \mathbf{r} \neq \mathbf{0}}} \int_{\mathbb{R}^{dl_i}, x_1=0} d^d\mathbf{x}_{\mathbf{A}_i} \left|H(\mathbf{A}, \mathbf{x}_{\mathbf{A}_i} + \mathbf{r}V)\right| \frac{w(\mathbf{x}_{\mathbf{A}_i} + \mathbf{r}V)}{w(\mathbf{x}_{\mathbf{A}_i} + \mathbf{r}V)} 1(\mathrm{St}(\mathbf{x}_{\mathbf{A}_i}) \leqslant V/4), \end{aligned}$$

where in passing from the first to the second line we used the definition (H.10) and we multiplied and divided by $w(\mathbf{x}_{\mathbf{A}_i} + \mathbf{r}V)$. Now, note that, on the support of $1(\mathrm{St}(\mathbf{x}_{\mathbf{A}_i}) \leqslant V/4)$, $|\mathbf{x}_{\mathbf{A}_i} + \mathbf{r}V| > V/2$ for any $\mathbf{r} \neq \mathbf{0}$. Therefore, the second line of (H.25) can be bounded from above by $\|H_{l_i}\|_w / w(V/2)$, where $1/w(V/2) = e^{-C_w (V/(2\gamma))^\sigma}$ represents the desired exponentially small gain as $V \to \infty$.

Consider now the contribution to $R_{3,n}$ associated with the difference $\mathcal{C}_V\left(\mathbf{x}_{\overline{\mathbf{B}}}\right) - \mathcal{C}\left(\mathbf{x}_{\overline{\mathbf{B}}}\right)$. Recall that both $\mathcal{C}(\mathbf{x}_{\overline{\mathbf{B}}})$ and $\mathcal{C}_V(\mathbf{x}_{\overline{\mathbf{B}}})$ can be written in terms of the BBF formula that, see (D.53), can be written as

$$\mathcal{C}(\mathbf{x}_{\overline{\mathbf{B}}}) = \sum_{\mathcal{T}} \prod_{(xx') \in \mathcal{T}} g(x - x') \int d\mu_{\mathcal{T}}(\mathbf{r}) \det \mathcal{N}, \tag{H.26}$$

where $\mathcal{N} = \mathcal{N}(\mathbf{r})$ is a Gram matrix, i.e., with elements represented as a suitable scalar product. Of course, $\mathcal{C}_V(\mathbf{x}_{\overline{\mathbf{B}}})$ admits a representation analogous to (H.26), with g replaced by g_V and $\mathcal{N}$ replaced by $\mathcal{N}_V$. Using (H.26) and the analogous representation for $\mathcal{C}_V(\mathbf{x}_{\overline{\mathbf{B}}})$, we write the difference $\mathcal{C}_V\left(\mathbf{x}_{\overline{\mathbf{B}}}\right) - \mathcal{C}\left(\mathbf{x}_{\overline{\mathbf{B}}}\right)$ in telescopic form, as the sum of terms in each of which either a difference $g_V(x - x') - g(x - x')$ or $\det \mathcal{N}_V - \det \mathcal{N}$ appears. In the former terms, recalling (H.13), we write $g_V(x) - g(x) = -\sum_{r \neq 0} g(x + rV)$ and, proceeding as we did for the bound of the terms with $H_V - H$, we see that they are exponentially small in V, and their sum over n too.

We are left with the term involving the difference $\det \mathcal{N}_V - \det \mathcal{N}$, which we rewrite once again in telescopic form as

$$\det \mathcal{N}_V - \det \mathcal{N} = \sum_{i,j=1}^{s} (\det \mathcal{N}_V^{(i,j)} - \det \mathcal{N}_V^{(i,j)'}), \tag{H.27}$$

where s is the linear size of the matrices, $\mathcal{N}_V^{(i,j)}$ is the matrix whose elements with label smaller or equal to (resp. larger than) (i, j) in the lexicographic order are equal to the elements of $\mathcal{N}_V$ (resp. $\mathcal{N}$), and $(i, j)'$ is the label immediately preceding (i, j) in the lexicographic order (if $(i,j) = (1,1)$, we interpret $\mathcal{N}_V^{(1,1)'} \equiv \mathcal{N}$). Since $\mathcal{N}_V^{(i,j)}$ and $\mathcal{N}_V^{(i,j)'}$ differ in only one element, expanding in minors along row i we have

$$\det \mathcal{N}_V^{(i,j)} - \det \mathcal{N}_V^{(i,j)'} = (-1)^{i+j} \left((\mathcal{N}_V)_{i,j} - \mathcal{N}_{i,j} \right) \det \widehat{\mathcal{N}}_V^{(i,j)}, \tag{H.28}$$

where $\widehat{\mathcal{N}}_V^{(i,j)}$ denotes the matrix $\mathcal{N}_V^{(i,j)}$ with both the i-th row and the j-th column removed. Recall that both $\mathcal{N}_V$ and $\mathcal{N}$ are Gram matrices; in particular, they can be written as $(\mathcal{N}_V)_{k,l} = (f_{V,k}, h_{V,l})$ and $\mathcal{N}_{k,l} = (f_k, h_l)$ for appropriate vectors f_V, h_V, f, h in two apriori different Hilbert spaces $\mathcal{H}_V$ and $\mathcal{H}$. Remarkably, also $\widehat{\mathcal{N}}_V^{(i,j)}$ is in Gram form, that is, for any $k \in \{1, \ldots, s\} \setminus \{i\}$ and any $l \in \{1, \ldots, s\} \setminus \{j\}$, we can write $(\widehat{\mathcal{N}}_V^{(i,j)})_{k,l} = (F_k, H_l)$, where $(\cdot, \cdot)$ denotes the scalar product in $\mathcal{H}_V \oplus \mathcal{H}$, and F_k, H_l are the following vectors in $\mathcal{H}_V \oplus \mathcal{H}$:

$$F_k = \begin{cases} (f_{V,k}, 0) & \text{if } k < i \\ (0, f_k) & \text{if } k > i \end{cases} \quad \text{and} \quad H_l = (h_{V,l}, h_l). \tag{H.29}$$

Therefore, $\det \widehat{\mathcal{N}}_V^{(i,j)}$ can be bounded qualitatively in the same way as $\det \mathcal{N}_V$ or $\det \mathcal{N}$, so that, using (H.28) into (H.27), and recalling that $(\mathcal{N}_V)_{i,j} - \mathcal{N}_{i,j}$ is proportional to $g_V - g$, we find that the term in $R_{3,n}$ involving the difference $\det \mathcal{N}_V - \det \mathcal{N}$ is bounded qualitatively as all the other terms, that is, they are exponentially small in V, and their sum over n too. This concludes the proof of Lemma H.1. $\square$

Consider e.g. $H(\psi)$ corresponding to the fixed point whose existence we proved. By corollary 6.1, this interaction satisfies bounds (6.43) which for sufficiently small ε are stronger than (H.11). Therefore, the effective action is indeed given by (5.3) as we assumed all along.

I Fixed point in a formal power series expansion

In this appendix we will show that eq. (5.32) $f(y) = 0$ can be solved in a formal power series expansion in ε. This is rather easy, compared to the proof of the existence of an actual solution given in section 6. We introduce a positive grading function on the couplings $y_i \in \{\nu, \lambda, u_{2\mathrm{R}}, u_{4\mathrm{R}}, u_{6\mathrm{R}}, (u_\ell)_{\ell \geqslant 8}\}$:

$$\mathrm{gr}(\nu) = \mathrm{gr}(\lambda) = 1, \quad \mathrm{gr}(u_{2\mathrm{R}}) = \mathrm{gr}(u_{4\mathrm{R}}) = 2, \quad \mathrm{gr}(u_{6\mathrm{R}}) = 3, \quad \mathrm{gr}(u_\ell) = k(l) = \frac{l}{2} - 1 (l \geqslant 8). \tag{I.1}$$

We also define grading of a product as a sum of gradings. It is then easy to check that each function e_{y_i} is a sum of terms whose grading is $\geqslant \mathrm{gr}(y_i)$, with strict inequality for ν, λ (table 1). This motivates the following

Theorem I.1 *Equation $f(y) = 0$ has a unique solution where couplings are formal power series in ε starting from:*

$$\nu = \frac{a}{b}\varepsilon + O(\varepsilon^2), \quad \lambda = -\frac{1}{b}\varepsilon + O(\varepsilon^2), \quad y_i = O(\varepsilon^{\mathrm{gr}(y_i)}) \quad (y_i \in \{u_{2\mathrm{R}}, u_{4\mathrm{R}}, u_{6\mathrm{R}}, (u_\ell)_{\ell \geqslant 8}\}). \tag{I.2}$$

Function	Grading	Notable present terms	Notable absent terms
e_ν	$\geqslant 2$	$\lambda^2, \nu\lambda, u_{4\mathrm{R}}$	$\nu, u_{2\mathrm{R}}, \lambda$
e_λ	$\geqslant 3$	$\nu\lambda^2, \lambda^3, \nu\mathfrak{X}_\lambda, \lambda\mathfrak{X}_\lambda, \lambda u_{4R}, u_{6\mathrm{R}}$	$\lambda, u_{4\mathrm{R}}, \lambda^2, \mathfrak{X}_\lambda, \nu\lambda, \nu u_{4\mathrm{R}}$
$e_{2\mathrm{R}}$	$\geqslant 2$	$u_{2\mathrm{R}}, u_{4\mathrm{R}}, \lambda^2$	ν, λ
$e_{4\mathrm{R}}$	$\geqslant 2$	$u_{4\mathrm{R}}, \lambda^2$	λ
$e_{6\mathrm{R}}$	$\geqslant 3$	$u_{6\mathrm{R}}, \lambda^3, \nu\lambda^2, \nu\mathfrak{X}_\lambda, u_{4\mathrm{R}}\lambda, u_8$	$\mathfrak{X}_\lambda, \lambda^2$
e_ℓ, $\ell \geqslant 8$	$\geqslant k(l)$	$u_\ell, \lambda^{k(l)}, \sum_{l_1+\ldots+l_n \geqslant l+2(n-1)} H_{\ell_1}\cdots H_{\ell_n}$	

Table 1. Grading of terms in functions e_{y_i}. We only show the variables on which the terms depend. E.g. $u_{2\mathrm{R}}$ and λ^2 in $e_{2\mathrm{R}}$ stand for $R^{2\mathrm{R}}_{2\mathrm{R}}(u_{2\mathrm{R}}) = Du_{2\mathrm{R}}$ and $R^{4\mathrm{L},4\mathrm{L}}_{2\mathrm{R}}(\lambda,\lambda)$, respectively.

Proof. Parameter ε enters (5.32) through the explicit term $\varepsilon\lambda$. In addition, all the other coefficients such as a, b and the multilinear kernels from the r.h.s. of (5.29) also depend on ε. This dependence originates from the fluctuation propagator $g(x)$ defined in (2.10), and it is nonsingular as $\varepsilon \to 0$ in our setup involving the UV and IR cutoffs. Below we will keep track only of the explicit dependence on ε from the $\varepsilon\lambda$ term, which we denote ϵ. All other coefficients will be treated as constants. We will give an algorithm to expand the solution as a formal power series in ϵ. To produce a power series in ε, one would have to set $\epsilon \to \varepsilon$ and additionally expand all coefficients in ε.

We start by rescaling the couplings $y_i \to \epsilon^{\mathrm{gr}(y_i)} y_i$. We will abuse notation denoting the rescaled couplings by the same letters. We have to show that the rescaled couplings have unique power series expansions starting at $O(1)$. The equations $f(y) = 0$ in terms of the rescaled couplings can be written as (the explanations and the definition of $e_{y_i,k}$ are given after (I.6))

$$\text{couplings of grading 1}: \begin{cases} -\nu - a\lambda = \sum_{k\geqslant 1} \epsilon^k e_{\nu,k} \\ -\lambda - b\lambda^2 = \sum_{k\geqslant 1} \epsilon^k e_{\lambda,k+1} \end{cases} \tag{I.3}$$

$$\text{couplings of grading 2}: \begin{cases} (1-D)u_{2\mathrm{R}} - R^{4\mathrm{R}}_{2\mathrm{R}} u_{4\mathrm{R}} = \sum_{k\geqslant 0} \epsilon^k e_{2\mathrm{R},k} \\ (1-D)u_{4\mathrm{R}} = \sum_{k\geqslant 0} \epsilon^k e_{4\mathrm{R},k} \end{cases} \tag{I.4}$$

$$\text{couplings of grading 3}: \begin{cases} (1-D)u_{6\mathrm{R}} - R^{8}_{6\mathrm{R}} u_8 = \sum_{k\geqslant 0} \epsilon^k e_{6\mathrm{R},k} \\ (1-D)u_8 = \sum_{k\geqslant 0} \epsilon^k e_{8,k} \end{cases} \tag{I.5}$$

$$\text{couplings of grading } \geqslant 4: (1-D)u_\ell = \sum_{k\geqslant 0} \epsilon^k e_{\ell,k} \quad (\ell \geqslant 10), \tag{I.6}$$

For each coupling, $e_{y_i,k}$ denotes the part of e_{y_i} which contains the terms of grading exactly $\mathrm{gr}(y_i) + k$. In addition we separated the linear terms Du_i as well as $R^{4\mathrm{R}}_{2\mathrm{R}} u_{4\mathrm{R}}$ and $R^8_{6\mathrm{R}} u_8$ from $e_{u_i,0}$ in (I.4)–(I.6). With this definition the remaining $e_{u_i,0}$ are at least quadratic in its arguments. Note that the r.h.s. of (I.3) are $O(\epsilon)$, while the other equations have r.h.s. $O(1)$. Note also the shift $k \to k+1$ in $e_{\lambda,k+1}$ in (I.3).

Step 1. The $O(1)$ parts of ν and λ known, $\nu = \frac{a}{b} + O(\epsilon)$, $\lambda = -\frac{1}{b} + O(\epsilon)$, let us solve for the $O(1)$ parts of the other couplings. Firstly, note that the $O(1)$ parts of the r.h.s. of (I.4)–(I.6), $e_{u_i,0}$, having grading exactly $\mathrm{gr}(u_i)$ and being at least quadratic, are computable in terms of the $O(1)$ parts of the couplings with smaller grading. Secondly, the linear operators in the l.h.s. of (I.4)–(I.6) are invertible. Indeed, the operator $(1-D)$ is invertible on each irrelevant coupling subspace as is clear from definition (5.20).[63] eqs. (I.4) and (I.5) involve a matrix-triangular operator with $(1-D)$ on the diagonal, hence also invertible. By these two observations, $O(1)$ parts of all couplings are uniquely determined starting from ν and λ and going recursively up in grading.

Step 2a. Now suppose we computed expansions of all couplings up to and including $O(\epsilon^N)$ (call it "inductive hypothesis 1"), and we want to solve for the ϵ^{N+1} terms. For any quantity $\alpha = \sum \epsilon^n \alpha_n$ we denote by $[\alpha]_n = \alpha_n$ the ϵ^n coefficient. We start with (I.3) and take its ϵ^{N+1} part:

$$-[\nu]_{N+1} - a[\lambda]_{N+1} = \sum_{k=1}^{N+1} [e_{\nu,k}]_{N+1-k}, \tag{I.7}$$

$$(-1-2b[\lambda]_0)[\lambda]_{N+1} = b\sum_{k=1}^{N} [\lambda]_k[\lambda]_{N+1-k} + \sum_{k=1}^{N+1} [e_{\lambda,k+1}]_{N+1-k}. \tag{I.8}$$

All the terms in the r.h.s. are computable by inductive hypothesis 1. Since $-1-2b[\lambda]_0 = 1$ we can compute first $[\lambda]_{N+1}$ and then $[\nu]_{N+1}$.

Step 2b. The remaining couplings are treated recursively going up in grading as before. Suppose all couplings of grading lower than u_i are already known up to and including $O(\epsilon^{N+1})$ (call it "inductive hypothesis 2"). Consider the equation for u_i (if there are two couplings having the same grading we should study their equations together as in Step 1) and take its ϵ^{N+1} part. In the l.h.s. we have an invertible linear operator, same as in Step 1, acting on $[u_i]_{N+1}$, while in the r.h.s. we have

$$\sum_{k=0}^{N+1} [e_{u_i,k}]_{N+1-k}. \tag{I.9}$$

For $k \geqslant 1$ this is computable by inductive hypothesis 1, and for $k=0$ by inductive hypothesis 2, since $e_{u_i,0}$ has grading $\mathrm{gr}(u_i)$ and is at least quadratic, which means it involves only lower-grading couplings. Therefore, we can compute $[u_i]_{N+1}$ and continue the induction.[64] This finishes the proof of the theorem.

Remark I.1 At the level of formal series expansions bosonic and fermionic fixed point are quite analogous. Consider e.g. the bosonic model (K.1). In perturbation theory, we could parametrize its fixed point by an interaction written in terms of kernels, like in (4.2). We could derive a perturbative renormalization map acting on the sequence of kernels

[63] We have $(1-D)^{-1} = 1 + D + D^2 + \dots$ and the series converges in L_1 if $\gamma^{-D_l - p} < 1$ which is the condition for irrelevance.

[64] Steps 2a, 2b can be unified, at the price of rendering the argument less explicit, by moving the nonlinear functions $e_{u_i,0}$ to the l.h.s. and noting that the Jacobian of the resulting nonlinear infinite matrix function of y in the l.h.s. is invertible at the point $(\nu, \lambda, u_i) = \left(\frac{a}{b}, -\frac{1}{b}, u_i^{(0)}\right)$ where $u_i^{(0)}$ are the $O(1)$ values of u_i computed in Step 1.

in the trimmed representation, similarly to section 5. We could then find, exactly as in theorem I.1, a fixed point in a formal power series in ε.

This analogy breaks down beyond perturbation theory. As stated in remark 5.1 and proved in appendix H, perturbative expansion captures the full fermionic effective action at small coupling. This does not hold for bosons due to large field effects ("instantons") at arbitrarily small couplings. Furthermore, rigorous non-perturbative studies of bosonic models parametrize irrelevant interactions not by kernels as in (4.2), but by a more complicated "polymer expansion" (see [114] and section 8.1.8).

Another difference is that for bosons, the formal power series solution in ε is expected to be only asymptotic, like the ϵ-expansion series for the Wilson-Fisher fixed point in $d = 4 - \epsilon$ dimensions [133], necessitating Borel resummations for the critical exponents [5]. The same considerations should apply to the long-range bosonic model (K.1). On the other hand, for fermions we have established in section 6.5 that the fixed point depends analytically on ε. This implies that the formal power series solution will be convergent for small ε. For a direct proof of convergence of the fermionic power series expansion via tree expansion, see the next appendix.

J Fixed point via the tree expansion

In the main sections of this paper we provided an explicit rigorous construction of a non-trivial RG fixed point, by finding the appropriate Banach space, which the RG map acts on, and by proving its contractivity in an appropriate neighborhood of this space. In the case of fermionic theories, as in the case at hand, the fixed point can also be found by a different strategy, which bypasses the construction of the Banach space and the contractivity argument, and is based on an expansion in tree diagrams, sometimes referred to as "Gallavotti-Nicolò" trees [134], see also [51, 71] and [119] for a review in the context of interacting fermionic theories. While the two constructions build on the same general foundations from sections 1–5 (trimmed representation of the interaction, weighted norm $\|\cdot\|_w$ for measuring the size of couplings, and the norm bounds on D and $R_\ell^{\ell_1,\dots,\ell_n}$), the tree expansion is closer in spirit to a direct combinatorial proof convergence of the formal ε-expansion of appendix I. Let us briefly describe here the construction of the fixed point via trees.

The starting point is the fixed point equation for the irrelevant couplings u_ℓ, with $\ell = 2\mathrm{R}, 4\mathrm{R}, 6\mathrm{R}, 8, 10, \dots$, which we rewrite, extracting the term $(n; (\ell_1, \cdots, \ell_n)) = (1; \ell)$ explicitly:

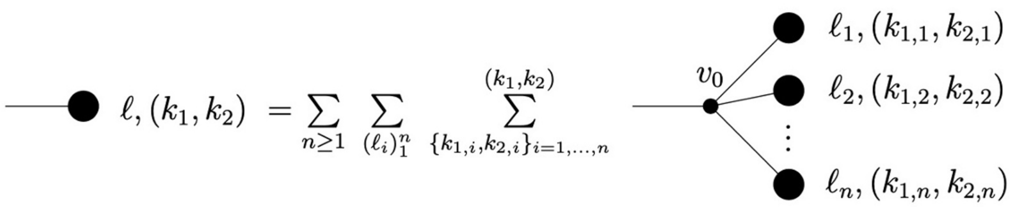
$$(1 - D)u_\ell = \sum_{n\geqslant 1} \sum_{(\ell_i)_1^n}^{*} R_\ell^{\ell_1,\dots,\ell_n}(H_{\ell_1}, \dots, H_{\ell_n}), \tag{J.1}$$

where the $*$ on the sum indicates the constraint that, if $n = 1$, then $\ell_1 \neq \ell$. We look for a solution in power series in ν and λ, $u_\ell = \sum_{k_1,k_2\geqslant 0} u_\ell^{(k_1,k_2)} \nu^{k_1} \lambda^{k_2}$, with $u_\ell^{(k_1,k_2)} = 0$ unless $k_1 + k_2 \geqslant |\ell|/2 - 1$ and, moreover, $k_1 + k_2 \geqslant 2$ for $\ell = 2\mathrm{R}, 4\mathrm{R}$ and $k_1 + k_2 \geqslant 3$ for $\ell = 6\mathrm{R}$.

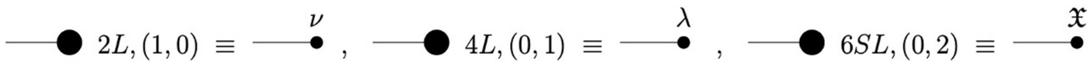
By plugging this ansatz in (J.1), we get[65]

$$u_\ell^{(k_1,k_2)} = \sum_{n\geqslant 1} \sum_{(\ell_i)_1^n}^{*} \sum_{\{k_{1,i},k_{2,i}\}_{i=1,2}}^{(k_1,k_2)} (1-D)^{-1} R_\ell^{\ell_1,\dots,\ell_n}\left(H_{\ell_1}^{(k_{1,1},k_{2,1})},\dots,H_{\ell_n}^{(k_{1,n},k_{2,n})}\right), \tag{J.2}$$

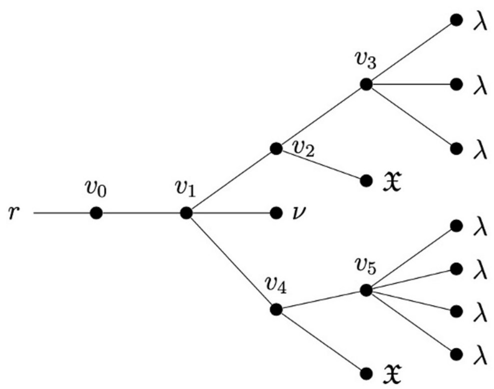
where the third sum runs over integers $k_{1,i}, k_{2,i} \geqslant 0$, with $i = 1,\dots,n$, such that $k_{1,1} + \cdots + k_{1,n} = k_1$ and $k_{2,1} + \cdots + k_{2,n} = k_2$. Moreover, the arguments $H_{\ell_i}^{(k_{1,i},k_{2,i})}$ should be interpreted as being equal to $u_{\ell_i}^{(k_{1,i},k_{2,i})}$ if $\ell_i \neq$ 2L, 4L, 6SL, to ν if $\ell_i = 2L$ (in which case we set $(k_{1,i}, k_{2,i}) = (1,0)$), to λ if $\ell_i =$ 4L (in which case we set $(k_{1,i}, k_{2,i}) = (0,1)$), and to $\mathfrak{X}_*$ if $\ell_i =$ 6SL (in which case we set $(k_{1,i}, k_{2,i}) = (0,2)$). Eq. (J.2) can be graphically represented as follows:



where the vertex labelled v_0 in the right side represents the action of $(1-D)^{-1}R_\ell^{\ell_1,\dots,\ell_n}$; we shall say that the n lines (or "branches") labelled $\ell_1,\dots,\ell_n$ "enter the vertex v_0"; similarly, we'll say that the branch to the left of v_0 "exits" from v_0: it carries the label ℓ and its left endpoint is called *root*. In the special cases $(\ell_i, (k_{1,i}, k_{2,i})) = (2\text{L}, (1,0)), (4\text{L}, (0,1)), (6\text{SL}, (0,2))$, the big dots in the right side will be reinterpreted as small dots with labels ν, λ, $\mathfrak{X}_*$, respectively:



Iterating the graphical equation above until the endpoints are all small dots with labels ν, λ, or $\mathfrak{X}_*$, we obtain an expansion in tree diagrams of the following form:



[65]Note that since u_ℓ is irrelevant, $(1-D)$ is an invertable operator, see footnote 63.

In this example, the branch labels are left implicit; each vertex v_r, with $r = 1, \ldots, 5$, is associated with the action of an operator $(1-D)^{-1} R_\ell^{\ell_1,\ldots,\ell_n}$, with ℓ the label of the line exiting from v_r, n the number of lines entering it, and $\ell_1, \ldots, \ell_n$ their labels. If m_ν, m_λ and $m_{\mathfrak{X}}$ are the numbers of endpoints of type ν, λ and $\mathfrak{X}_*$, respectively (in the example, $m_\nu = 2$, $m_\lambda = 6$ and $m_{\mathfrak{X}} = 2$), and ℓ_0 is the label of the line exiting from v_0, then the tree contributes to $u_{\ell_0}^{(m_\nu, m_\lambda + 2m_{\mathfrak{X}})}$.

Let us now use the norm bounds on D and $R_\ell^{\ell_1,\ldots,\ell_n}$, see sections 5.3 and 5.6, to bound the value of any such tree in the $\|\cdot\|_w$ norm, assuming $|\nu|, |\lambda| \leqslant \delta$. The norm bounds on D imply that the norm of $(1-D)^{-1}$ is bounded by some $d_\gamma > 1$ (uniformly in ℓ). Using this and eq. (5.35), the norm of the value of a tree is bounded by the product of the norms of its endpoints, which is bounded by

$$\delta^{m_\nu + m_\lambda} (B_\gamma \delta^2)^{m_{\mathfrak{X}}} \leqslant (B_\gamma \delta)^k, \qquad k = m_\nu + m_\lambda + 2m_{\mathfrak{X}}, \tag{J.3}$$

times the product over vertices v_r (i.e. all vertices v which are not endpoints):

$$\prod_{v \text{ not e.p.}} d_\gamma C_\gamma^{n_v - 1} \gamma^{-D_{l_v}} C_0^{\sum_{i=1}^{n_v} |\ell_i(v)|}, \tag{J.4}$$

where $l_v = |\ell_v|$, with ℓ_v the label of the line exiting from v, and $\ell_i(v)$ the label of the i-th line entering v. Note that $\ell_i(v) = \ell_{v'}$ for some v', which may be an endpoint. Therefore, (J.4) can be equivalently rewritten as

$$C_0^{-|\ell_{v_0}|} \times \prod_{v \text{ not e.p.}} d_\gamma C_\gamma^{n_v - 1} \gamma^{-D_{l_v}} C_0^{l_v} \times \prod_{v \text{ e.p.}} C_0^{l_v}. \tag{J.5}$$

Replacing the first factor by 1, and including the factor (J.3) associated with the endpoints, the value of a single tree is bounded by

$$(B_\gamma C_0^4 \delta)^k \prod_{v \text{ not e.p.}} d_\gamma C_\gamma^{n_v - 1} \gamma^{-D_{l_v}} C_0^{l_v}, \tag{J.6}$$

where we also used that $\sum_{v \text{ e.p.}} l_v = 2m_\nu + 4m_\lambda + 6m_{\mathfrak{X}} \leqslant 4k$.

As the next step we have to sum over all possible values of ℓ_i's which label tree branches. We take γ sufficiently large so that $\gamma^{-(d/4-\varepsilon/2)} C_0 < 1$, implying (recall $D_l = l(d/4 - \varepsilon/2) - d$)

$$\sum_{l=2}^{\infty} \gamma^{-D_l} (C_0)^l < \infty. \tag{J.7}$$

Then the sum $\sum_{\{\ell_v\}}$of (J.6) is bounded by

$$(B_\gamma C_0^4 \delta)^k \prod_{v \text{ not e.p.}} \tilde{d}_\gamma C_\gamma^{n_v - 1}. \tag{J.8}$$

Furthermore, it is easy to see that

$$\sum_{v \text{ not e.p.}} (n_v - 1) = m - 1, \tag{J.9}$$

JHEP01(2021)026

where $m_\cdot = m_\nu + m_\lambda + m_\mathfrak{x} \leqslant k$ is the total number of the endpoints, and

$$\sum_{v \text{ not e.p.}} 1 \leqslant \frac{1}{2} \sum_{v \text{ e.p.}} l_v, \tag{J.10}$$

as implied by $|\ell| \leqslant \sum_i |\ell_i| - 2$ in each vertex v (note the constraint (5.36) for $n \geqslant 2$, and that for $n = 1$ the sum in (J.1) starts from $l_1 = l + 2$). Using the last two relations, we can estimate (J.8) by

$$(B_\gamma C_0^4 C_\gamma \tilde{d}_\gamma^2 \delta)^k. \tag{J.11}$$

Finally we have to sum this expression over all trees subject to the constraint on the endpoints $m_\nu + m_\lambda + 2m_\mathfrak{x} = k$. The number of such trees is smaller than 4^k (see, e.g., Lemma A.1 of [119]). We conclude that the contribution to u_ℓ of order k is bounded in the $\|\cdot\|_w$ norm by $(A_\gamma \delta)^k$, for some γ-dependent constant A_γ. This implies that the tree expansion for u_ℓ is convergent for $\delta \leqslant \delta_0(\gamma)$ small enough. This concludes the construction of the irrelevant couplings u_ℓ in terms of the fixed point relevant couplings ν and λ, which implies automatically that u_ℓ are analytic in ν and λ in the neighborhood $|\nu|, |\lambda| \leqslant \delta_0(\gamma)$ of the origin in $\mathbb{C}^2$. The same argument also establishes analyticity of u_ℓ in ε as long as it belongs to the complex half-plane $\mathrm{Re}\,\varepsilon < d/6$ where couplings u_ℓ are all irrelevant.

We are left with the beta function equations for the relevant couplings. Via the same strategy, we find that they are given by the first two equations of (5.30), with $e_\nu^{(0)}$ and $e_\lambda^{(0)}$ expressed in terms of two tree expansions, convergent for δ sufficiently small. In conclusion, at the fixed point, ν and λ satisfy

$$\nu = \gamma^{\frac{d}{2}+\varepsilon}(\nu + I_1\lambda) + e_\nu^{(0)}(\nu, \lambda, \varepsilon), \qquad \lambda = \gamma^{2\varepsilon}(\lambda + I_2\lambda^2) + e_\lambda^{(0)}(\nu, \lambda, \varepsilon), \tag{J.12}$$

with $e_\nu^{(0)}$ and $e_\lambda^{(0)}$ two analytic functions of $\nu, \lambda, \varepsilon$ for $|\nu|, |\lambda| \leqslant \delta_0(\gamma)$, $\mathrm{Re}\,\varepsilon < d/6$. Moreover, we have $e_\nu^{(0)}$ and $e_\lambda^{(0)}$ of order δ^2 and δ^3, respectively where $\delta = \max(|\nu|, |\lambda|)$ and for ε in any compact subset of $\mathrm{Re}\,\varepsilon < d/6$. I_1 and I_2 also depend analytically on ε. By the analytic implicit function theorem, these equations have a unique solution $\nu_*(\varepsilon)$, $\lambda_*(\varepsilon)$ which is defined in the disk $|\varepsilon| \leqslant \varepsilon_0(\gamma)$ and is analytically close to the lowest order approximated solution $\lambda_0 = (1 - \gamma^{2\varepsilon})/I_2$, $\nu_0 = I_1\lambda_0/\left(1 - \gamma^{\frac{d}{2}+\varepsilon}\right)$.

Remark J.1 The condition of γ large is not truly required for the tree expansion to converge; in fact it converges for any $\gamma > 1$ (for λ, ν sufficiently small). In the proof above, large γ was needed due to the pessimistic way in which we bounded some combinatorial factors in the previous sections. Note that the origin of the factor $(C_0)^l$ in (J.7) has to be traced back to the factors $C_0^{|\ell_i|}$ in (5.36). A critical rereading of the proof leading to those factors shows that the factor $C_0^{\sum_{i=1}^{n_v} |\ell_i(v)|}$ in (J.4) can be replaced by $C_\mathrm{R}\binom{\sum_{i=1}^{n_v} |\ell_i(v)|}{l_v} C_0^{\sum_{i=1}^{n_v} |\ell_i(v)| - l_v}$.[66] The product of the factors $C_0^{\sum_{i=1}^{n_v} |\ell_i(v)| - l_v}$ over the vertices v that are not endpoints gives

[66]See in particular the first equality in (E.2) and the second one in (E). The factor $C_\mathrm{R} > 1$ is the constant in (5.43); it appears only if $l_v = 2, 4$. In the main text it was absorbed into C_0 (footnote 31), but now we keep it explicit, since it blows up as $\gamma \to 1$, due to the blowup of constants $C_{1,2,3}$ from the end of appendix C.

$C_0^{\sum_{v\, e.p.} l_v - l_{v_0}}$, with v_0 the vertex attached to the root; this does not need any condition on γ. The product of $C_{\rm R}$'s is bounded by $C_{\rm R}^{2k}$ because of (J.10). Finally, the sum over the branch labels, given the type of endpoints, reduces to $\sum_{\{\ell_v\}} \prod_{v \text{ not e.p.}} \gamma^{-D_{l_v}\left(\frac{\sum_{i=1}^{n_v} |\ell_i(v)|}{l_v}\right)}$, which can be bounded as explained in appendix A.6.1 of [119], leading to a factor smaller than $\left(\frac{1}{1-\gamma^{-\alpha}}\right)^{4k}$, for some $\alpha > 0$ and any $\gamma > 1$.

Remark J.2 In Wilsonian RG, the single RG step contains all the information about the fixed point, so that we should feel free to forget about what happened in the RG past. The fact that the tree expansion represents the fixed point kernels by tree diagrams with several levels may superficially seem to go against this standard idea. This is not the case: the tree expansion as presented in this appendix is just a way to solve the fixed point equation for a single RG step.[67] In general, trees with several levels are exponentially suppressed, as compared to "short" ones: this is the so called *short-memory property* (see e.g. the remark after (7.26) in [28]).

The tree expansion can be easily adapted to the construction of the full Wilsonian RG flow of the effective couplings from the ultraviolet to the infrared, see section J.1 below for a brief discussion of this fact, and to the computation of correlation functions and critical exponents. It has been used to construct them in several 1D fermionic theories [28–31] and 2D statistical mechanics models at criticality [32–39], whose nontrivial fixed points are all in the Luttinger liquid universality class. In these cases, the construction of the fixed point requires a proof that the beta function for the quartic coupling is asympotically vanishing in the infrared limit; historically, the proof of vanishing beta functions in these models has first been proved via a comparison with the Luttinger model exact solution [28], and later via a combined use of Ward Identities and Schwinger-Dyson equations [30]. We are not aware of a construction of nontrivial fixed points in the Luttinger liquid universality class via methods different from the tree expansion; it would be an interesting exercise to reproduce their construction via a contraction argument in a suitable Banach space of interactions, as done in this paper for long-range symplectic fermions with quartic interaction.

The fermionic nature of models such as the one studied in this paper makes the approach based on the tree expansion an extremely efficient tool for constructing the RG fixed point, arguably simpler than the one based on the contraction argument in a Banach space. However, we do not expect that the tree expansion is as a general scheme as the other, which is, to date, the only available technique for constructing nontrivial bosonic fixed points, see appendix K, and it looks the most promising (at least conceptually) for approaching the non-perturbative problem of constructing very non-Gaussian fixed points in the vicinity of approximate fixed points (possibly computed via the truncation of some other alternative scheme, such as numerical FRG). This explains the reason why in the main sections of this paper we decided to follow the scheme based on the contraction argument in a Banach space: it provides a benchmarch for other approaches, like the Functional

[67] To make an analogy with something already seen, consider the construction of the fixed point sextic semilocal term in section 3: $\mathfrak{X}_*$ is the solution to the single step equation (3.11), and its explicit expression in terms of λ involves a sum over many "levels", i.e., the integers n in (3.12).

Renormalization Group (whose conceptual similarity allows for a direct comparison of results), and it displays general features which do not depend on the specific, fermionic, nature of the problem.

J.1 On the flow of the effective couplings

A tree expansion analogous to the one described above for the construction of the fixed point can be used to compute the whole sequence of effective potentials associated with the Wilsonian RG flow from the ultraviolet to the infrared scales. Such a generalized tree expansion was described in several previous reviews on the subject, see in particular [119]. Suppose that we interested in constructing the model formally defined by the interacting Grassmann measure $Z^{-1}d\mu_P(\psi)e^{H^{(0)}(\psi)}$ at all distances (rather than being interested "just" in the construction of its infrared fixed point, as done in this paper), where $d\mu_P(\psi)$ is the Grassmann Gaussian integration with the propagator $P(x)$ in (2.1) (with fixed ultraviolet cutoff but without any infrared one) and $H^{(0)}$ is a local interaction like the one in (2.7), with bare couplings ν_0, λ_0. The partition function $\int d\mu_P(\psi)e^{H^{(0)}(\psi)}$ and the closely related generating function of correlations can be computed iteratively, by first integrating out momenta in the annulus of radii γ^{-1} and 1, then in the one of radii γ^{-2} and γ^{-1}, and so on. In formulae, this means rewriting the propagator $P(x)$ as $\sum_{h\leqslant 0} g^{(h)}(x)$, with $g^{(h)}(x) = \gamma^{h(d/2-\varepsilon)}g(\gamma^h x)$ and $g(x)$ the same as in (2.10); correspondingly, the fluctuation field ψ is decomposed as $\psi = \sum_{h\leqslant 0}\psi^{(h)}$ and the Grassmann Gaussian integration $d\mu_P(\psi)$ as $\prod_{h\leqslant 0} d\mu_{g^{(h)}}(\psi^{(h)})$ (cf. with (2.11)), in terms of which we define the sequence (cf. with (2.13)):

$$e^{H^{(h-1)}_{\rm eff}(\psi)} = \int d\mu_{g^{(h)}}(\psi^{(h)})e^{H^{(h)}_{\rm eff}(\psi+\psi^{(h)})}, \quad h \leqslant 0,$$

with $H^{(0)}_{\rm eff} \equiv H^{(0)}$. After appropriate rescaling, we obtain the effective potentials (cf. with (2.16))

$$H^{(h)}(\psi) = H^{(h)}_{\rm eff}(\gamma^{h[\psi]}\psi(\cdot\, \gamma^h)),$$

which satisfy the RG equation

$$R[H^{(h)}] = H^{(h-1)},$$

with $R = R(\varepsilon, \gamma)$ the same renormalization map introduced after (2.16). A mild generalization of the construction of this paper allows us to prove that, for λ_0 *positive* and sufficiently small (we are taking here ε positive and small, as well), there exists a choice of ν_0 such that the whole sequence of effective potentials $\{H^{(h)}\}_{h\leqslant 0}$ is well defined, they all belong to the same Banach space (the same we used to construct the fixed point) and $\lim_{h\to-\infty} H^{(h)} = H_*$, where H_* is the fixed point constructed above. Correspondingly, the local part of $H^{(h)}$ is parametrized by two running coupling constants ν_h, λ_h, which interpolate between the bare values ν_0, λ_0 and the fixed point values $\nu = \lim_{h\to-\infty}\nu_h$, $\lambda = \lim_{h\to-\infty}\lambda_h$. In particular, the sequence $\{\nu_h, \lambda_h\}_{h\leqslant 0}$ is small, uniformly in the scale index h.

Remarkably, the effective potentials $H^{(h)}$ can be expressed in terms of a convergent tree expansion, analogous to the one described above, with the important difference that now the endpoints carry a scale label and are, therefore, associated with couplings ν_k or

λ_k, computed at scales $h < k \leqslant 0$. From the knowledge of the effective potentials, one can reconstruct all the observables one is interested in, including the correlation functions computed at arbitrary finite distance (before any scaling limit is taken): these will be expressed as convergent expansions *in the whole sequence of running couplings* $\{\nu_h, \lambda_h\}_{h\leqslant 0}$. See, e.g., chapters 12, 13, 14 in [119]. Note that the existence of such a convergent expansion does not imply convergence of the naive perturbation theory in the bare couplings: schematically, the relation between λ_h and λ_0 has the same features as the one between $\lambda(t)$ and λ_0 in (1.6); in particular, λ_h is analytic in λ_0 non-uniformly in h, while it is Borel summable in λ_0 uniformly in h. Therefore, pre-scaling-limit observables are expected to be, at best, Borel summable in λ_0. On the contrary, observables at the fixed point, such as critical exponents, are expressed as convergent expansions in the fixed point couplings ν_*, λ_* only, and, therefore, recalling that ν_* and λ_* are analytic in ε, they can be proved to be analytic functions of ε, as well.

K Rigorously constructed bosonic fixed points

In this appendix we will mention some existing rigorous constructions of non-gaussian bosonic fixed points. Earlier works not directly focusing on such fixed points, but instrumental for acquiring rigorous RG control in bosonic theories, include [135–139].

In 1998, Brydges, Dimock and Hurd [22] gave the first construction of a fixed point in a bosonic scalar field theory with a long-range interactions. In analogy to (1.1), their bare action can be written schematically as

$$\mathrm{MFT}(\varphi) + \nu \int d^d x\, \varphi^2 + \lambda \int d^d x\, \varphi^4, \tag{K.1}$$

i.e. a gaussian scale-invariant Mean Field Theory of a bosonic field φ in $\mathbb{R}^d$ of dimension $[\varphi] = d/4 - \varepsilon/2$ with a quadratic and quartic interactions. They considered the model in $d = 4$, which necessitated adding to (K.1) one more relevant local interaction $\int d^d x (\partial\varphi)^2$. Contrary to what the title of [22] may suggest, it does not provide a rigorous definition of the Wilson-Fisher fixed point in $d = 4 - \varepsilon$. The two models differ already in their perturbative critical exponents. E.g. the scaling dimension of φ gets corrections at $O(\varepsilon^2)$ in the Wilson-Fisher model, while such corrections are absent in the model of [22] at any order in ε.

In 2000, Mitter and Scoppola [140] studied a different model perturbing MFT by a δ-function interaction:

$$\mathrm{MFT}(\varphi) + g \int dx\, \delta^{(N)}(\varphi(x)), \tag{K.2}$$

where φ is an N-component field in $\mathbb{R}^{d=1}$. The δ-function penalizes configurations when $\varphi(x)$ passes through zero, which physically describes repulsion of a polymer from an impurity. The scaling dimension of this interaction is $-N[\varphi]$. They constructed a fixed point of this model in the case when $[\varphi]$ is negative and small while N is large so that $-N[\varphi] = 1 - \varepsilon$ is close to marginality.

In 2003, Brydges, Mitter, and Scoppola [23] constructed a fixed point of exactly the model (K.1). Following [140], they used fluctuation covariance of finite support in x-space, simplifying the proof compared to the construction in [22].[68] Although nominally $d = 3$ in [23], the proof should apply also for $d = 1, 2$ [25, 141]. See also the nice review in [114]. Further work in this direction was done by Abdesselam [24] who constructed a full renormalization group trajectory from MFT at short distances to the fixed point of [23] at long distances. More recently, Slade [26] considered an analogous fixed point for an n-component field φ. He also considered the case $n = 0$, corresponding to the self-avoiding random walk. This formal limit is analyzed rigorously by considering a theory of two scalar bosons and two scalar fermions whose global symmetry is OSp(2|2). Unlike in our model, there is no quartic interactions for fermions in [26] because there are only two of them. Such a "supersymmetric" model was also studied earlier by Mitter and Scoppola [142].

Physically, model (K.1) should describe the critical point of the long-range Ising model [93] (see also [143]), and much is known or conjectured about it. The critical point is expected to have conformal invariance [82]. At $\varepsilon = \varepsilon_c(d)$ the critical point should cross over to the local Wilson-Fisher fixed point plus a decoupled Gaussian sector [58]. See also [78] for higher-loop perturbative computations of critical exponents.

References

[1] K.G. Wilson and J.B. Kogut, *The Renormalization group and the ϵ-expansion*, Phys. Rept. **12** (1974) 75 [INSPIRE].

[2] K.G. Wilson, *The Renormalization Group: Critical Phenomena and the Kondo Problem*, Rev. Mod. Phys. **47** (1975) 773 [INSPIRE].

[3] K.G. Wilson, *The renormalization group and critical phenomena*, Rev. Mod. Phys. **55** (1983) 583 [INSPIRE].

[4] K.G. Wilson and M.E. Fisher, *Critical exponents in 3.99 dimensions*, Phys. Rev. Lett. **28** (1972) 240 [INSPIRE].

[5] R. Guida and J. Zinn-Justin, *Critical exponents of the N vector model*, J. Phys. A **31** (1998) 8103 [cond-mat/9803240] [INSPIRE].

[6] M.V. Kompaniets and E. Panzer, *Minimally subtracted six loop renormalization of $O(n)$-symmetric ϕ^4 theory and critical exponents*, Phys. Rev. D **96** (2017) 036016 [arXiv:1705.06483] [INSPIRE].

[7] D. Poland, S. Rychkov and A. Vichi, *The Conformal Bootstrap: Theory, Numerical Techniques, and Applications*, Rev. Mod. Phys. **91** (2019) 015002 [arXiv:1805.04405] [INSPIRE].

[8] G. Mussardo, *Statistical field theory: an introduction to exactly solved models in statistical physics*, Oxford University Press (2010).

[68] In this paper we used fluctuation covariance of finite support in Fourier space, not in x-space, see (2.10). Unlike in [23], our fluctuation covariance has zero integral in x-space, which was useful at times although not essential.

[9] J. Terning, *Modern supersymmetry: Dynamics and duality*, Oxford University Press (2006).

[10] J. de Boer, E.P. Verlinde and H.L. Verlinde, *On the holographic renormalization group*, *JHEP* **08** (2000) 003 [hep-th/9912012] [INSPIRE].

[11] M. Bianchi, D.Z. Freedman and K. Skenderis, *How to go with an RG flow*, *JHEP* **08** (2001) 041 [hep-th/0105276] [INSPIRE].

[12] T. Burkhardt and J. van Leeuwen, *Real-space Renormalization*, Springer (1982).

[13] J. Berges, N. Tetradis and C. Wetterich, *Nonperturbative renormalization flow in quantum field theory and statistical physics*, *Phys. Rept.* **363** (2002) 223 [hep-ph/0005122] [INSPIRE].

[14] B. Delamotte, *An Introduction to the nonperturbative renormalization group*, *Lect. Notes Phys.* **852** (2012) 49 [cond-mat/0702365] [INSPIRE].

[15] M. Levin and C.P. Nave, *Tensor renormalization group approach to 2D classical lattice models*, *Phys. Rev. Lett.* **99** (2007) 120601 [cond-mat/0611687] [INSPIRE].

[16] G. Evenbly and G. Vidal, *Tensor Network Renormalization*, *Phys. Rev. Lett.* **115** (2015) 180405 [arXiv:1412.0732].

[17] M. Hauru, C. Delcamp and S. Mizera, *Renormalization of tensor networks using graph independent local truncations*, *Phys. Rev. B* **97** (2018) 045111 [arXiv:1709.07460] [INSPIRE].

[18] I. Balog, H. Chaté, B. Delamotte, M. Marohnic and N. Wschebor, *Convergence of Nonperturbative Approximations to the Renormalization Group*, *Phys. Rev. Lett.* **123** (2019) 240604 [arXiv:1907.01829] [INSPIRE].

[19] G. De Polsi, I. Balog, M. Tissier and N. Wschebor, *Precision calculation of critical exponents in the $O(N)$ universality classes with the nonperturbative renormalization group*, *Phys. Rev. E* **101** (2020) 042113 [arXiv:2001.07525] [INSPIRE].

[20] I. Balog, G. De Polsi, M. Tissier and N. Wschebor, *Conformal invariance in the nonperturbative renormalization group: a rationale for choosing the regulator*, *Phys. Rev. E* **101** (2020) 062146 [arXiv:2004.02521] [INSPIRE].

[21] K.G. Wilson, *A model of coupling constant renormalization*, *Phys. Rev. D* **2** (1970) 1438 [INSPIRE].

[22] D. Brydges, J. Dimock and T.R. Hurd, *A nonGaussian fixed point for ϕ^4 in $4-\epsilon$ dimensions*, *Commun. Math. Phys.* **198** (1998) 111 [INSPIRE].

[23] D.C. Brydges, P.K. Mitter and B. Scoppola, *Critical $\Phi^4_{3,\epsilon}$*, *Commun. Math. Phys.* **240** (2003) 281 [hep-th/0206040] [INSPIRE].

[24] A. Abdesselam, *A Complete Renormalization Group Trajectory Between Two Fixed Points*, *Commun. Math. Phys.* **276** (2007) 727 [math-ph/0610018] [INSPIRE].

[25] P.K. Mitter, *Long range ferromagnets: Renormalization group analysis*, talk presented at LPTHE, Université Pierre et Marie Curie, Paris, 24 October 2013, https://hal.archives-ouvertes.fr/cel-01239463.

[26] G. Slade, *Critical Exponents for Long-Range $O(n)$ Models Below the Upper Critical Dimension*, *Commun. Math. Phys.* **358** (2018) 343 [arXiv:1611.06169] [INSPIRE].

[27] K. Gawędzki and A. Kupiainen, *Renormalization of a Nonrenormalizable Quantum Field Theory*, *Nucl. Phys. B* **262** (1985) 33 [INSPIRE].

[28] G. Benfatto, G. Gallavotti, A. Procacci and B. Scoppola, *β-function and Schwinger functions for a many fermions system in one-dimension. Anomaly of the Fermi surface*, *Commun. Math. Phys.* **160** (1994) 93 [INSPIRE].

[29] G. Benfatto and V. Mastropietro, *Renormalization Group, hidden symmetries and approximate Ward Identities in the XYZ model*, Rev. Math. Phys. **13** (2001) 1323.

[30] G. Benfatto and V. Mastropietro, *Ward identities and chiral anomaly in the Luttinger liquid*, Commun. Math. Phys. **258** (2005) 609 [cond-mat/0409049] [INSPIRE].

[31] G. Benfatto, P. Falco and V. Mastropietro, *Functional Integral Construction of the Thirring model: Axioms verification and massless limit*, Commun. Math. Phys. **273** (2007) 67 [hep-th/0606177] [INSPIRE].

[32] V. Mastropietro, *Ising Models with Four Spin Interaction at Criticality*, Comm. Math. Phys. **244** (2004) 595.

[33] A. Giuliani and V. Mastropietro, *Anomalous Universality in the Anisotropic Ashkin-Teller Model*, Comm. Math. Phys. **256** (2005) 681.

[34] A. Giuliani and V. Mastropietro, *Anomalous Critical Exponents in the Anisotropic Ashkin-Teller Model*, Phys. Rev. Lett. **93** (2004) 190603.

[35] G. Benfatto, P. Falco and V. Mastropietro, *Extended Scaling Relations for Planar Lattice Models*, Comm. Math. Phys. **292** (2009) 569.

[36] G. Benfatto, P. Falco and V. Mastropietro, *Universal Relations for Non Solvable Statistical Models*, Phys. Rev. Lett. **104** (2010) 075701 [arXiv:0909.2707] [INSPIRE].

[37] A. Giuliani, V. Mastropietro and F.L. Toninelli, *Height fluctuations in interacting dimers*, Ann. Inst. H. Poincaré Probab. Statist. **53** (2017) 98.

[38] A. Giuliani, V. Mastropietro and F.L. Toninelli, *Haldane relation for interacting dimers*, J. Stat. Mech. (2017) 034002.

[39] A. Giuliani, V. Mastropietro and F.L. Toninelli, *Non-integrable dimers: Universal fluctuations of tilted height profiles*, Comm. Math. Phys. **377** (2020) 1883.

[40] H.G. Kausch, *Curiosities at $c = -2$*, hep-th/9510149 [INSPIRE].

[41] H. Saleur, *Polymers and percolation in two-dimensions and twisted $N = 2$ supersymmetry*, Nucl. Phys. B **382** (1992) 486 [hep-th/9111007] [INSPIRE].

[42] K.J. Wiese and A.A. Fedorenko, *Field Theories for Loop-Erased Random Walks*, Nucl. Phys. B **946** (2019) 114696 [arXiv:1802.08830] [INSPIRE].

[43] D. Anninos, T. Hartman and A. Strominger, *Higher Spin Realization of the dS/CFT Correspondence*, Class. Quant. Grav. **34** (2017) 015009 [arXiv:1108.5735] [INSPIRE].

[44] L. Fei, S. Giombi, I.R. Klebanov and G. Tarnopolsky, *Critical $Sp(N)$ models in $6 - \epsilon$ dimensions and higher spin dS/CFT*, JHEP **09** (2015) 076 [arXiv:1502.07271] [INSPIRE].

[45] A. Stergiou, *Symplectic critical models in $6 + \epsilon$ dimensions*, Phys. Lett. B **751** (2015) 184 [arXiv:1508.03639] [INSPIRE].

[46] A. LeClair, *Quantum critical spin liquids, the 3D Ising model, and conformal field theory in 2+1 dimensions*, cond-mat/0610639 [INSPIRE].

[47] A. LeClair, *3D Ising and other models from symplectic fermions*, cond-mat/0610817 [INSPIRE].

[48] A. LeClair and M. Neubert, *Semi-Lorentz invariance, unitarity, and critical exponents of symplectic fermion models*, JHEP **10** (2007) 027 [arXiv:0705.4657] [INSPIRE].

[49] D.J. Gross and V. Rosenhaus, *A line of CFTs: from generalized free fields to SYK*, JHEP **07** (2017) 086 [arXiv:1706.07015] [INSPIRE].

[50] S. El-Showk, M. Paulos, D. Poland, S. Rychkov, D. Simmons-Duffin and A. Vichi, *Conformal Field Theories in Fractional Dimensions*, *Phys. Rev. Lett.* **112** (2014) 141601 [arXiv:1309.5089] [INSPIRE].

[51] G. Gallavotti, *Renormalization theory and ultraviolet stability for scalar fields via renormalization group methods*, *Rev. Mod. Phys.* **57** (1985) 471 [INSPIRE].

[52] G. Altarelli, *Introduction to renormalons*, in *5th Hellenic School and Workshops on Elementary Particle Physics*, (1996) pp. 221–236.

[53] M. Beneke, *Renormalons*, *Phys. Rept.* **317** (1999) 1 [hep-ph/9807443] [INSPIRE].

[54] A.B. Zamolodchikov, *Renormalization Group and Perturbation Theory Near Fixed Points in Two-Dimensional Field Theory*, *Sov. J. Nucl. Phys.* **46** (1987) 1090 [INSPIRE].

[55] A.W.W. Ludwig and J.L. Cardy, *Perturbative Evaluation of the Conformal Anomaly at New Critical Points with Applications to Random Systems*, *Nucl. Phys. B* **285** (1987) 687 [INSPIRE].

[56] A. Cappelli and J.I. Latorre, *Perturbation Theory of Higher Spin Conserved Currents Off Criticality*, *Nucl. Phys. B* **340** (1990) 659 [INSPIRE].

[57] Z. Komargodski and D. Simmons-Duffin, *The Random-Bond Ising Model in 2.01 and 3 Dimensions*, *J. Phys. A* **50** (2017) 154001 [arXiv:1603.04444] [INSPIRE].

[58] C. Behan, L. Rastelli, S. Rychkov and B. Zan, *A scaling theory for the long-range to short-range crossover and an infrared duality*, *J. Phys. A* **50** (2017) 354002 [arXiv:1703.05325] [INSPIRE].

[59] J. Polchinski, *Renormalization and Effective Lagrangians*, *Nucl. Phys. B* **231** (1984) 269 [INSPIRE].

[60] Wikipedia, *Steiner tree problem.* https://en.wikipedia.org/wiki/Steiner_tree_problem.

[61] W.A. de S. Pedra and M. Salmhofer, *Determinant bounds and the matsubara uv problem of many-fermion systems*, *Comm. Math. Phys.* **282** (2008) 797.

[62] A. Giuliani, *Order, disorder and phase transitions in quantum many body systems*, *Le Scienze* **150** (2016) 3 [arXiv:1711.06991] [INSPIRE].

[63] G. Benfatto, A. Giuliani and V. Mastropietro, *Fermi Liquid Behavior in the 2D Hubbard Model at Low Temperatures*, *Annales Henri Poincaré* **7** (2006) 809.

[64] A. Giuliani and V. Mastropietro, *Rigorous construction of ground state correlations in graphene: Renormalization of the velocities and Ward identities*, *Phys. Rev. B* **79** (2009) 201403.

[65] A. Giuliani and V. Mastropietro, *The Two-Dimensional Hubbard Model on the Honeycomb Lattice*, *Comm. Math. Phys.* **293** (2010) 301.

[66] V. Mastropietro, *Interacting Weyl semimetals on a lattice*, *J. Phys. A* **47** 465003.

[67] V. Mastropietro, *Weyl semimetallic phase in an interacting lattice system*, *J. Stat. Phys.* **157** (2014) 830.

[68] A. Giuliani, V. Mastropietro and M. Porta, *Anomaly non-renormalization in interacting Weyl semimetals*, arXiv:1907.00682 [INSPIRE].

[69] K. Van Houcke, E. Kozik, N. Prokof'ev and B. Svistunov, *Diagrammatic Monte Carlo*, *Physics Procedia* **6** (2010) 95 [arXiv:0802.2923]

[70] N. Prokof'ev, *Diagrammatic Monte Carlo*, in proceedings of *Autumn School on Correlated Electrons*, 16–20 September 2019, Forschungszentrum Jülich, Germany, E. Pavarini, E. Koch and S. Zhang eds., *Many-Body Methods for Real Materials. Modeling and Simulation*, Vol. 9, http://www.cond-mat.de/events/correl19.

[71] G. Benfatto and G. Gallavotti, *Renormalization Group*, Princeton University Press (1995).

[72] D.C. Brydges and T. Kennedy, *Mayer expansions and the Hamilton-Jacobi equation*, J. Stat. Phys. **48** (1987) 19.

[73] D.C. Brydges and J.D. Wright, *Mayer expansions and the Hamilton-Jacobi equation. II. Fermions, dimensional reduction formulas*, J. Stat. Phys. **51** (1988) 435.

[74] J.D. Wright and D. Brydges, *Erratum: Mayer expansions and the hamiltonian–jacobi equation. ii. fermions, dimensional reduction formulas*, J. Stat. Phys. **97** (1999) 1027.

[75] M. Salmhofer and C. Wieczerkowski, *Positivity and Convergence in Fermionic Quantum Field Theory*, J. Stat. Phys. **99** (2000) 557.

[76] M. Disertori and V. Rivasseau, *Continuous constructive fermionic renormalization*, Annales Henri Poincaré **1** (2000) 1 [hep-th/9802145] [INSPIRE].

[77] R. Bauerschmidt, D.C. Brydges and G. Slade, *Introduction to a renormalisation group method*, vol. 2242. Springer, 2019, 10.1007/978-981-32-9593-3 [arXiv:1907.05474] [INSPIRE].

[78] D. Benedetti, R. Gurau, S. Harribey and K. Suzuki, *Long-range multi-scalar models at three loops*, J. Phys. A **53** (2020) 445008 [arXiv:2007.04603] [INSPIRE].

[79] M. Hairer, *A theory of regularity structures*, Inventiones mathematicae **198** (2014) 269.

[80] Y. Bruned, M. Hairer and L. Zambotti, *Algebraic renormalisation of regularity structures*, Inventiones mathematicae **215** (2019) 1039.

[81] F. Wegner, *The Critical State, General Aspects*, in C. Domb and M. Green eds., *Phase Transitions and Critical Phenomena*, Vol. 6, Academic Press (1976), pp. 8–126.

[82] M.F. Paulos, S. Rychkov, B.C. van Rees and B. Zan, *Conformal Invariance in the Long-Range Ising Model*, Nucl. Phys. B **902** (2016) 246 [arXiv:1509.00008] [INSPIRE].

[83] S. Smirnov, *Critical percolation in the plane: conformal invariance, Cardy's formula, scaling limits*, C. R. Mathematique **333** (2001) 239.

[84] H. Duminil-Copin and S. Smirnov, *Conformal invariance of lattice models*, in D. Ellwood, C. Newman, V. Sidoravicius and W. Werner eds., *Probability and Statistical Physics in Two and More Dimensions*, American Mathematical Society (2012) arXiv:1109.1549.

[85] D. Chelkak, C. Hongler and K. Izyurov, *Conformal invariance of spin correlations in the planar ising model*, Ann. Math. **181** (2012) 1087 [arXiv:1202.2838].

[86] S. Rychkov, *EPFL Lectures on Conformal Field Theory in $D \geqslant 3$ Dimensions*, SpringerBriefs in Physics (2016), 10.1007/978-3-319-43626-5 [arXiv:1601.05000] [INSPIRE].

[87] S. Rychkov, *3D Ising Model: a view from the Conformal Bootstrap Island*, Comptes Rendus Physique **21** (2020) 185 [arXiv:2007.14315] [INSPIRE].

[88] C. Behan, *Bootstrapping the long-range Ising model in three dimensions*, J. Phys. A **52** (2019) 075401 [arXiv:1810.07199] [INSPIRE].

[89] M. Hogervorst, S. Rychkov and B.C. van Rees, *Unitarity violation at the Wilson-Fisher fixed point in 4-ϵ dimensions*, Phys. Rev. D **93** (2016) 125025 [arXiv:1512.00013] [INSPIRE].

[90] O.E. Lanford, *A computer-assisted proof of the Feigenbaum conjectures*, Bull. Amer. Math. Soc. (N.S.) **6** (1982) 427 [https://projecteuclid.org:443/euclid.bams/1183548786].

[91] O. Schnetz, *Numbers and Functions in Quantum Field Theory*, Phys. Rev. D **97** (2018) 085018 [arXiv:1606.08598] [INSPIRE].

[92] O. Schnetz, *HyperlogProcedures*, Maple package, https://www.math.fau.de/person/oliver-schnetz/.

[93] M.E. Fisher, S.-k. Ma and B.G. Nickel, *Critical Exponents for Long-Range Interactions*, *Phys. Rev. Lett.* **29** (1972) 917 [INSPIRE].

[94] J. Sak, *Recursion relations and fixed points for ferromagnets with long-range interactions*, *Phys. Rev. B* **8** (1973) 281.

[95] J.L. Cardy, *Scaling and renormalization in statistical physics*, Cambridge University Press, Cambridge, U.K. (1996).

[96] E. Luijten and H.W.J. Blöte, *Boundary between long-range and short-range critical behavior in systems with algebraic interactions*, *Phys. Rev. Lett.* **89** (2002) 025703 [cond-mat/0112472].

[97] M.C. Angelini, G. Parisi and F. Ricci-Tersenghi, *Relations between short-range and long-range ising models*, *Phys. Rev. E* **89** (2014) 062120 [arXiv:1401.6805].

[98] C. Behan, L. Rastelli, S. Rychkov and B. Zan, *Long-range critical exponents near the short-range crossover*, *Phys. Rev. Lett.* **118** (2017) 241601 [arXiv:1703.03430] [INSPIRE].

[99] D.J. Binder and S. Rychkov, *Deligne Categories in Lattice Models and Quantum Field Theory, or Making Sense of $O(N)$ Symmetry with Non-integer N*, *JHEP* **04** (2020) 117 [arXiv:1911.07895] [INSPIRE].

[100] W.E. Caswell, *Asymptotic Behavior of Nonabelian Gauge Theories to Two Loop Order*, *Phys. Rev. Lett.* **33** (1974) 244 [INSPIRE].

[101] A.A. Belavin and A.A. Migdal, *Calculation of anomalous dimensions in non-abelian gauge field theories*, *Pisma Zh. Eksp. Teor. Fiz.* **19** (1974) 317.

[102] T. Banks and A. Zaks, *On the Phase Structure of Vector-Like Gauge Theories with Massless Fermions*, *Nucl. Phys. B* **196** (1982) 189 [INSPIRE].

[103] A. Jakovác, A. Patkós and P. Pósfay, *Non-Gaussian fixed points in fermionic field theories without auxiliary Bose-fields*, *Eur. Phys. J. C* **75** (2015) 2 [arXiv:1406.3195] [INSPIRE].

[104] F. Gehring, H. Gies and L. Janssen, *Fixed-point structure of low-dimensional relativistic fermion field theories: Universality classes and emergent symmetry*, *Phys. Rev. D* **92** (2015) 085046 [arXiv:1506.07570] [INSPIRE].

[105] H. Gies, T. Hellwig, A. Wipf and O. Zanusso, *A functional perspective on emergent supersymmetry*, *JHEP* **12** (2017) 132 [arXiv:1705.08312] [INSPIRE].

[106] L. Dabelow, H. Gies and B. Knorr, *Momentum dependence of quantum critical Dirac systems*, *Phys. Rev. D* **99** (2019) 125019 [arXiv:1903.07388] [INSPIRE].

[107] J.-P. Blaizot, R. Mendez-Galain and N. Wschebor, *Non perturbative renormalisation group and momentum dependence of n-point functions (I)*, *Phys. Rev. E* **74** (2006) 051116 [hep-th/0512317] [INSPIRE].

[108] J.-P. Blaizot, R. Mendez-Galain and N. Wschebor, *Non perturbative renormalization group and momentum dependence of n-point functions. II.*, *Phys. Rev. E* **74** (2006) 051117 [hep-th/0603163] [INSPIRE].

[109] F. Benitez, J.-P. Blaizot, H. Chate, B. Delamotte, R. Mendez-Galain and N. Wschebor, *Non-perturbative renormalization group preserving full-momentum dependence: implementation and quantitative evaluation*, *Phys. Rev. E* **85** (2012) 026707 [arXiv:1110.2665] [INSPIRE].

[110] N. Hasselmann, *Effective average action based approach to correlation functions at finite momenta*, *Phys. Rev. E* **86** (2012) 041118 [arXiv:1206.6121] [INSPIRE].

[111] F. Rose and N. Dupuis, *Nonperturbative renormalization-group approach preserving the momentum dependence of correlation functions*, *Phys. Rev. B* **97** (2018) 174514 [arXiv:1801.03118] [INSPIRE].

[112] C. Wetterich, *Exact evolution equation for the effective potential*, Phys. Lett. B **301** (1993) 90 [arXiv:1710.05815] [INSPIRE].

[113] R. Gurau, V. Rivasseau and A. Sfondrini, *Renormalization: an advanced overview*, arXiv:1401.5003 [INSPIRE].

[114] P.K. Mitter, *The Exact renormalization group*, math-ph/0505008 [INSPIRE].

[115] L. Rodino, *Linear Partial Differential Operators in Gevrey Spaces*, World Scientific, Singapore (1993).

[116] S.-Y. Chung and J. Chung, *There Exist No Gaps between Gevrey Differentiable and Nowhere Gevrey Differentiable*, Proc. Amer. Math. Soc. **133** (2005) 859.

[117] A. Lesniewski, *Effective Action for the Yukawa$_2$ Quantum Field Theory*, Commun. Math. Phys. **108** (1987) 437 [INSPIRE].

[118] D.C. Brydges, *A Short Course on Cluster Expansions*, in *Summer School in Theoretical Physics, Session XLIII: Critical Phenomena, Random Systems, Gauge Theories*, Les Houches, France, 1 August–7 September 1984, pp. 129–183. http://www.math.ubc.ca/ db5d/Seminars/les_houches_84.pdf.

[119] G. Gentile and V. Mastropietro, *Renormalization group for one-dimensional fermions: A review on mathematical results*, Phys. Rep. **352** (2001) 273.

[120] A. Giuliani, *The ground state construction of the two-dimensional Hubbard model on the honeycomb lattice*, in J. Frohlich, M. Salmhofer, V. Mastropietro, W. De Roeck, and L.F. Cugliandolo eds., *Quantum Theory from Small to Large Scales: Lecture Notes of the Les Houches Summer School*, Volume 95, August 2010, Oxford University Press (2012).

[121] V. Mastropietro, *Non-Perturbative Renormalization*, World Scientific (2008), http://dx.doi.org/10.1142/6748.

[122] E.R. Caianiello, *Number of Feynman graphs and convergence*, Nuovo Cim. **3** (1956) 223.

[123] J. Feldman, J. Magnen, V. Rivasseau and R. Seneor, *A Renormalizable Field Theory: The Massive {Gross-Neveu} Model in Two-dimensions*, Commun. Math. Phys. **103** (1986) 67 [INSPIRE].

[124] R. Rossi, *Determinant Diagrammatic Monte Carlo Algorithm in the Thermodynamic Limit*, Phys. Rev. Lett. **119** (2017) 045701.

[125] R. Rossi, N. Prokof'ev, B. Svistunov, K. Van Houcke and F. Werner, *Polynomial complexity despite the fermionic sign*, Europhys. Lett. **118** (2017) 10004 [arXiv:1703.10141].

[126] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambirdge University Press (2009).

[127] I. Battle, Guy A. and P. Federbush, *A Phase Cell Cluster Expansion for Euclidean Field Theories. Part 1*, Annals Phys. **142** (1982) 95 [INSPIRE].

[128] D. Brydges and P. Federbush, *A New Form of the Mayer Expansion in Classical Statistical Mechanics*, J. Math. Phys. **19** (1978) 2064 [INSPIRE].

[129] K. Gawędzki and A. Kupiainen, *Gross-Neveu Model through convergent perturbation expansions*, Commun. Math. Phys. **102** (1985) 1 [INSPIRE].

[130] Wikipedia, *Prüfer sequence*, https://en.wikipedia.org/wiki/Pruefer_sequence.

[131] K.G. Wilson, *Feynman graph expansion for critical exponents*, Phys. Rev. Lett. **28** (1972) 548 [INSPIRE].

[132] J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena*, 4th ed., Oxford University Press (2002).

[133] E. Brézin, J.C. Le Guillou and J. Zinn-Justin, *Perturbation Theory at Large Order. 1. The ϕ^{2N} Interaction*, *Phys. Rev. D* **15** (1977) 1544 [INSPIRE].

[134] G. Gallavotti and F. Nicolo, *Renormalization theory in four-dimensional scalar fields. I*, *Commun. Math. Phys.* **100** (1985) 545 [INSPIRE].

[135] G. Benfatto, M. Cassandro, G. Gallavotti, F. Nicolo, E. Olivieri, E. Presutti et al., *Ultraviolet stability in euclidean scalar field theories*, *Commun. Math. Phys.* **71** (1980) 95 [INSPIRE].

[136] G. Gallavotti, *On the ultraviolet stability in statistical mechanics and field theory*, *Ann. Mat. pura ed applicata* **120** (1979) 1.

[137] K. Gawędzki and A. Kupiainen, *A rigorous block spin approach to massless lattice theories*, *Commun. Math. Phys.* **77** (1980) 31 [INSPIRE].

[138] K. Gawedzki and A. Kupiainen, *Massless lattice ϕ^4_4 theory: Rigorous control of a renormalizable asymptotically free model*, *Comm. Math. Phys.* **99** (1985) 197, https://projecteuclid.org/euclid.cmp/1103942678.

[139] D. Brydges and H.T. Yau, *Grad Phi perturbations of massless Gaussian fields*, *Commun. Math. Phys.* **129** (1990) 351 [INSPIRE].

[140] P.K. Mitter and B. Scoppola, *Renormalization group approach to interacting polymerized manifolds*, *Commun. Math. Phys.* **209** (2000) 207 [hep-th/9812243] [INSPIRE].

[141] P.K. Mitter, personal communication.

[142] P.K. Mitter and B. Scoppola, *The Global renormalization group trajectory in a critical supersymmetric field theory on the lattice $\mathbb{Z}^3$*, *J. Statist. Phys.* **133** (2008) 921 [arXiv:0709.3941] [INSPIRE].

[143] M. Suzuki, Y. Yamazaki and G. Igarashi, *Wilson-type expansions of critical exponents for long-range interactions*, *Phys. Lett. A* **42** (1972) 313.