Revisiting the Theory of Finite Size Scaling in Disordered Systems: ν Can Be Less than $2/d$

Ferenc Pázmándi, Richard T. Scalettar, and Gergely T. Zimányi
Physics Department, University of California, Davis, California 95616
(Received 18 April 1997)

For phase transitions in disordered systems, an exact theorem provides a bound on the finite size correlation length exponent: $\nu_{\mathrm{FS}} \geq 2/d$. It is believed that the intrinsic ν satisfies the same bound. We argue that the standard averaging introduces a noise and a *new* diverging length scale. For $\nu \leq 2/d$ *self-averaging breaks down*, disconnecting ν from ν_{FS}, and the bound applies only for the latter. We illustrate these ideas on two exact examples, with $\nu < 2/d$. We propose a new method of disorder averaging, which is able to capture the intrinsic exponents. [S0031-9007(97)04880-1]

PACS numbers: 75.10.Nr, 05.70.Fh, 72.15.Rn, 75.40.Mg

Using a very general formulation, Ref. [1] presented an exact theorem, which puts constraints on the finite size correlation length exponent ν_{FS} of a large class of disordered systems: $\nu_{\mathrm{FS}} \geq 2/d$, where d is the dimension. This relation is often referred to as the quantum Harris criterion [2]. While many investigations found exponents in accordance with this bound, there is an increasing number of results in contradiction with it. In particular, in a model for charge density waves exact calculations yielded $\nu = 1/2$ below four dimensions [3], and numerical studies on 2D disordered Bose-Hubbard models found $\nu \simeq 0.7$ [4]. Experimentally the Bose glass transition of helium in aerogel [5], and the localization transition in doped semiconductors [6] seem to violate this bound. In this paper we argue that the standard procedure of disorder averaging introduces a noise and an *extrinsic* length scale, resulting in the $\nu_{\mathrm{FS}} \geq 2/d$ bound. For models where the *intrinsic* ν is less than $2/d$ therefore the two exponents are necessarily different. The underlying physical mechanism of this difference is the *breakdown of self-averaging* in systems with $\nu < 2/d$. In agreement with this picture, the equivalence of these exponents has been demonstrated only in specific cases with $\nu \geq 2/d$ [1].

To start our considerations of random systems, we chose the same type of disorder used by Ref. [1]: a binary distribution for, say, a disordered site energy. Typically, physical quantities are calculated by averaging over different disorder realizations. For calculational convenience, the standard method is analogous to the "grand canonical" approach: impurities are put on each site with a given probability p and the averaging is carried out for all possible concentrations of impurities and their configurations. An alternative method, which could be termed the "canonical" approach, keeps the number of impurities fixed, and the average is taken only over the possible configurations of these impurities. For infinite systems the two methods are equivalent. The density fluctuations in the grand canonical method, however, introduce an extra noise. This noise vanishes in the infinite system, but it may alter the results of the finite size scaling. The "canonical averaging" strongly reduces this noise by excluding density fluctuations.

We now argue that the bound obtained in Ref. [1] is only generated by the noise introduced by the "grand canonical averaging." Different choices, such as using canonical averaging, produce different bounds. The theorem of Ref. [1] considers a random system where a phase transition is induced by changing the concentration K of site (or bond) impurities. Let Y be any event depending on disorder realizations in a finite volume, with probability $\mathcal{P}(K)$. This $\mathcal{P}(K)$ is calculated by *averaging* over all disordered configurations, and selecting those compatible with Y. Averaging is performed in the grand canonical way, since fluctuations in the density of impurities are allowed. From these premises the exact statement $|d\mathcal{P}(K)/dK| \leq \mathrm{const} \times \sqrt{N}$ follows, where N is the system size. A closer look at the proof reveals that this result is derived solely from the concentration fluctuations of the impurities, which were *externally introduced in the averaging process* (see the last equation of the proof in Ref. [1]). Thus the bound on $|d\mathcal{P}(K)/dK|$ does not relate to the intrinsic properties of the system under investigation. It only reflects the "resolution" of the grand canonical averaging. In other words, because of the presence of the density fluctuations, the minimal resolvable change in K is $dK \propto 1/\sqrt{N}$. The probability $\mathcal{P}$ can change at most $\mathcal{O}(1)$, immediately explaining the above bound.

On the other hand, if one uses canonical averaging, then the above inequality does not apply. For, in contrast to the previous case, the number of impurities is now well defined. In the present binary example, the resolvable change of K is bounded only by its minimum allowed increment, $1/N$. Hence, $|d\mathcal{P}(K)/dK| \leq N$. Along the lines of Ref. [1], the inequality $\nu_{\mathrm{FS}} \geq 1/d$ now follows. As before, this inequality is characteristic of the canonical averaging only, and does not impose any restriction on the intrinsic exponent ν of the physical system. The physical reason behind this is that both averaging procedures introduce a *new characteristic length scale*, which has the potential to obscure the intrinsic correlation length of the physical system.

While these considerations were presented in the context of finite size scaling, our results are relevant to

 0031-9007/97/79(25)/5130(4)$10.00

disordered mesoscopic experiments where finite N is itself of interest and the choice of grand canonical or canonical averaging is determined by the setup.

It is also important to note that the assumption of a binary disorder plays a crucial role in deriving the above bounds. While the proof was extended to some continuous distributions [1], exceptions exist too. To see this, consider the following simple example, motivated by the quantum phase transition between the so-called Mott-insulator and Bose-glass phases, which takes place in interacting Bose systems with site disorder. At this transition the renormalization group flows are controlled by a fixed point with *zero* hopping strength [7], thus the system reduces to a collection of independent sites with random energies. Let the distribution of the site energy $\epsilon \in [0, K]$ be

$$P(\epsilon) = \frac{\alpha + 1}{K^{\alpha+1}} (K - \epsilon)^{\alpha}, \tag{1}$$

with $\alpha > -1$. We generate N independent ϵ_i ($i = 1, \ldots, N$) from the above distribution. We define the finite-size event Y to occur, when *all* ϵ_i's are smaller than a given value $\mu \in (0, K]$. We fix the value of μ, and drive the transition by changing K. As required by the theorem of [1], the probability $\mathcal{P}$ of Y happening is finite at the critical value of the disorder, $K_c = \mu$. It goes to zero exponentially with the system size N for $K > K_c$. Close to the transition, for $\delta = (K - K_c)/K_c \ll 1$, this probability is

$$\mathcal{P}(N, \delta) \simeq e^{-N\delta^{\alpha+1}}. \tag{2}$$

A characteristic length scale ξ_f can now be defined as a function of δ. It is determined from the system size as $N_f = \xi_f^d$, where $\mathcal{P}(N_f, \delta)/\mathcal{P}(N_f, 0) \sim 1/e$. Defining a critical exponent as $\xi_f \propto \delta^{-\nu_{\rm FS}}$ one arrives at $\nu_{\rm FS} = (\alpha + 1)/d$. For $\alpha < 1$, $\nu_{\rm FS}$ *is less than* $2/d$. While we considered a concrete example, we emphasize that this result can be relevant for *any* transition driven by *local singularities* in the action.

Motivated by the above observations, we now attempt to construct a modified finite size scaling procedure. In contrast to the above described averaging methods, this new *correlated averaging* has the potential to access the inherent exponents of the system. Let us start by observing that any *given* disorder realization in a *finite system* could have been generated from disorder distributions characterized by a *range* of parameters, corresponding to a *range* of the critical control parameter value K_c. This raises the problem of *which* K_c *to use* in a finite size scaling analysis.

The standard procedure answers this question by assuming that one can use a single K_c for all samples generated from the same distribution. However, the above argument suggests that the very same sample may be the realization of distributions with different parameters, leading to an inherent noise in the procedure, similar to the above considered binary examples. The *correlated averaging* procedure eliminates this noise by identifying the critical value of the control parameter K_c^r for *each disorder realization* which it *most likely* corresponds to. In practice this might be difficult, and we return to this question later. For the moment, we only assume that it is possible to identify a K_c^r. We propose that the natural control parameter of the critical behavior is $\Delta = (K - K_c^r)/K_c^r$. The act of averaging should then be performed for the samples with the same Δ.

We propose to adopt the following finite size scaling hypothesis for a generic physical quantity Q,

$$\bar{Q}(L, \Delta) = L^{-y} q(L\Delta^{\nu}), \tag{3}$$

where $q(z)$ is a universal scaling function, and y, ν are the critical exponents for Q, and the inherent correlation length $\xi \propto \Delta^{-\nu}$. Here we assumed that the scaling behavior of Q is characterized by a single length scale. This is a reasonable assumption even for disordered systems where the "typical" and "average" correlations have different exponents [8]: in this case one has to choose a physical quantity Q which is connected to one type of fluctuation only. Note that some aspects of Eq. (3) are already practiced in numerical studies: sizable noise reduction is customarily reached by adjusting the random variables *after* they are generated, for instance, to keep their mean value constant.

Next we assume the validity of Eq. (3) and perform the standard finite size scaling, to demonstrate how that procedure's inherent noise can mask the true critical behavior. Some of the key results of the analysis are: (i) we find that the exponent of the intrinsic correlation length ν might be different from $\nu_{\rm FS}$. Therefore the theorem of Ref. [1] does not provide constraints on the intrinsic exponent ν. (ii) In particular, ν *can be* less than $2/d$. In this case typically $\nu_{\rm FS} = 2/d$.

The standard finite size scaling procedure [9] in disordered systems calls for calculating a physical quantity Q, such as the critical susceptibility, for different values of N and K, the system size and control parameter, each time performing the calculations for a number of disorder realizations. Averaging over the disorder yields $\langle Q(K) \rangle$, and the critical coupling K_c is then identified, for instance, from a crossing pattern [10]. Requiring the collapse of the data when plotted as a function of $L^{1/\nu}\delta$, where $\delta = (K - K_c)/K_c$, determines the exponents.

To make contact between the standard scaling procedure and Eq. (3), a relation between the unique K_c and the fluctuating K_c^r has to be constructed. A simple representation of the inherent noise, or uncertainty, is to assume the validity of the central limit theorem for K_c^r

$$\Delta = \delta + \frac{D}{L^{d/2}} x, \tag{4}$$

where x is a random variable with a distribution width of $\mathcal{O}(1)$. Here D measures the scatter in K_c^r, and δ is the distance from the average critical point K_c.

The standard procedure neglects the fluctuations of K_c^r, which is equivalent to averaging $\bar{Q}$ over the random variable x of Eq. (4),

$$\langle Q \rangle = L^{-y} \left\langle q\left(D^{\nu} L^{1-d\nu/2} \left(x + \frac{\delta L^{d/2}}{D} \right)^{\nu} \right) \right\rangle . \quad (5)$$

Here the x average is denoted by $\langle \cdots \rangle$, corresponding to the standard averaging procedure, as opposed to $\bar{Q}$, the correlated averaging of the new procedure in Eq. (3).

First we analyze the critical point itself, then we shall proceed to extract the critical behavior of the correlation length. At $\delta = 0$ the scaling form for Q is

$$\langle Q \rangle = L^{-y} \langle q(D^{\nu} x^{\nu} L^{1-d\nu/2}) \rangle . \quad (6)$$

For $\nu > 2/d$ the argument of the scaling function approaches zero with increasing system size, and the L dependence of the *averaged* quantity $\langle Q(L) \rangle$ is characterized by the *intrinsic* exponent y. Here we use the customary assumption that the universal scaling function $q(z)$ approaches a finite value as $z \to 0$.

In the $\nu < 2/d$ case, however, the argument of $q(z)$ goes to large values, probing deeply noncritical regions, even though the system is assumed to be *at criticality*. To highlight the consequences of this, we proceed with a generic form for the asymptotic behavior of the scaling function, adopting $q(z) \propto z^{-\beta}$. From Eq. (6) $\langle Q \rangle \propto L^{-\gamma}$, where $\gamma = y + \beta(1 - d\nu/2)$. Clearly the L dependence of $\langle Q \rangle$ is governed by an exponent γ, *different* from the intrinsic y.

Next we develop an understanding of the region in the proximity of the critical point, i.e., the case of finite δ. Let us first focus on $\nu < 2/d$. From Eq. (5) one identifies two scaling regions, governed by *two different* characteristic diverging length scales.

For large system sizes inevitably $D^{\nu} L^{1-d\nu/2} \gg 1$, so the argument of $q(z)$ again extends to large values. Utilizing the previous asymptotic model form,

$$\langle Q \rangle = L^{-\gamma} \hat{q}(\delta L^{d/2}) , \quad (7)$$

from which a length scale can be identified, characterizing the finite size scaling of $\langle Q \rangle$, averaged in the standard way. It diverges with an exponent $\nu_{\mathrm{FS}} = 2/d$ *even though the intrinsic exponent* ν *is less than* $2/d$. This result now demonstrates, in general, what has been observed earlier for the binary example: the standard, or grand canonical averaging introduces a noise, which in turn generates a new length scale and a corresponding new exponent into the analysis.

The other scaling region is reached when $\delta L^{d/2}/D \gg 1$. In this limit

$$\langle Q \rangle = L^{-y} q(\delta^{\nu} L) . \quad (8)$$

As is known, for large values of $\delta^{\nu} L$, the ν exponent is not accessible by finite size scaling [9], hence $\delta^{\nu} L$ should be kept around unity. Therefore the determination of ν requires the study of the region *away from the asymptotics*: large δ and small system sizes. For weak disorder ($D \ll 1$) this window, in fact, might be wide enough for practical purposes. To reiterate, however, studies concentrating on the asymptotic region are bound to see $\nu_{\mathrm{FS}} = 2/d$.

In the case of $\nu > 2/d$ the standard procedure is capable of accessing the intrinsic ν: it can be obtained from $\langle Q \rangle$ by increasing the system size to the extent of $\delta L^{d/2}/D \gg 1$, but keeping $\delta^{\nu} L \propto \mathcal{O}(1)$. This again implies avoiding the "nonscaling" region around $\delta = 0$. For strong disorder and small available system sizes, one can end up again with large arguments of $q(z)$, and consequently in the scaling regime described by ν_{FS} and γ [Eq. (7)]. There are several additional crossover regimes which can be studied based on Eq. (5).

It is far from trivial to identify the infinite system's K_c^r from the finite sample. A solution might be suggested by recalling that for ordered classical magnets, the maximum of the susceptibility of a finite size sample is shifted as $T_c(L) - T_c(\infty) \propto L^{-1/\nu}$. Scaling then can be performed in terms of $T_c(L)$, resulting in the correct exponents. It is natural to expect that the same holds true for K_c^r: $K_c^r(L) - K_c^r(\infty) \propto L^{-1/\nu}$, where $K_c^r(L)$ is extracted from a specific feature of a critical quantity of the *finite-size system*. Using this $K_c^r(L)$ in our new scaling approach should provide the correct exponent ν.

We are thus left with the task of identifying $K_c^r(L)$ of a finite system. For many quantum systems at $T = 0$ a reasonable proposition for $K_c^r(L)$ might be the value of K, where the gap to the first excitation vanishes or has a minimum. For classical systems $K_c^r(L)$ may be identified where a critical susceptibility exhibits a maximum.

To demonstrate the above ideas, consider strongly interacting bosons in a random potential at zero temperature. In Ref. [7] renormalization flows were generated by integrating out the sites with highest excitation energies. For infinite range hopping the renormalization group (RG) equations are *exact*. In particular, at the Mott-insulator to superfluid transition weak disorder is irrelevant and $\nu = 1/d$.

We carried out the finite size scaling analysis of the average local susceptibility at weak but finite disorder for system sizes $N = 64$, 128, and 256. First we used the standard averaging procedure (inset of Fig. 1), and we obtained $\nu_{\mathrm{FS}} \simeq 3/d$ after averaging over 1024 realizations of a uniform disorder distribution of the random potential. The higher-than-expected value of ν_{FS} is related to the singularity of the scaling function. Figure 1 shows the same quantity scaled by using $K_c^r(L)$ extracted from the divergence of the susceptibility for each sample separately. The scaling is convincing, and yields the exact exponent $\nu = 1/d$. The exhibited curves were obtained by averaging over much fewer samples than before, only 16, yet the scaling region extends by more than an order of magnitude further in terms of the scaling variable, $N\Delta$, demonstrating a very effective noise reduction.

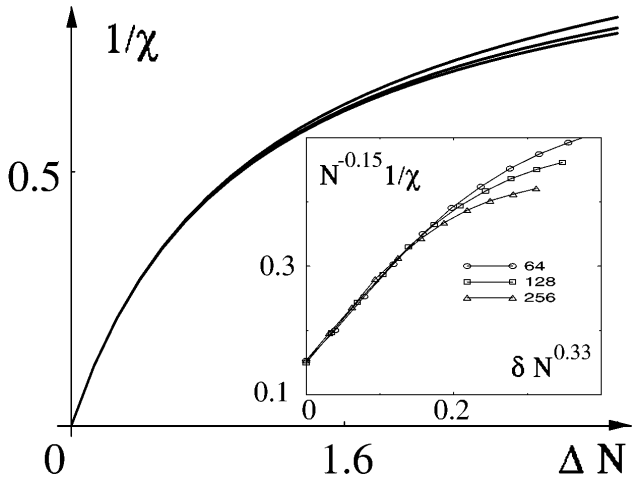

FIG. 1. Scaling plot of the inverse susceptibility using the novel and the standard (inset) averaging procedure for system sizes $N = 64$, 128, and 256.

In some numerical studies, such as in Ref. [4], a $\nu_{\rm FS} < 2/d$ has been reported, using the traditional averaging procedure. We would like to emphasize that this finding can be perfectly accommodated in the present theory. First, our analysis does *not* suggest that $\nu_{\rm FS}$ *must* be greater or equal to $2/d$: this is only the most likely scenario. If, for instance, the fluctuations of K_c^r scale as $L^{-1/\nu}$, then the above analysis yields $\nu_{\rm FS} = \nu$, and thus can be less than $2/d$. Apparently, this is the case in the example of the Mott-insulator to Bose-glass transition in Eq. (2). Second, as emphasized after Eq. (8), if the fluctuations of K_c^r are small and the sample size is not too big, then the intrinsic ν can and will be observed in finite-size scaling. Finally, this theory is *not addressing* the problems associated with distributions with long power-law tails [8] or multicritical fixed points [11].

What is then a possible physical framework to think about the case $\nu < 2/d$? The usual approach imagines dividing the sample to roughly independent blocks of uniform size ξ [12]. For $\nu < 2/d$ the fluctuations of the "local" T_c's of the blocks are bigger than the distance from the true T_c, therefore a self-consistent picture of a sharp transition was believed to be impossible. However, an appropriate modification *can restore* the self-consistency as follows. On the disordered side of such an assumed transition the correlation length must be finite *everywhere* in the sample. Therefore even at $\nu < 2/d$ it should be possible to divide the sample to *finite but unequal* boxes by choosing their sizes to be the same as the *local* correlation length. These variable size boxes will be roughly independent. However, if a uniform partitioning is forced on the system, those boxes extend across the variable partitioning, and thus will exhibit strong correlations. Thus the uniform boxes *cannot be assumed* to be independent realizations of the disordered system. Since this assumption is the foundation of the standard disorder averaging procedure, we conclude that for $\nu < 2/d$ the central phenomenon is the *breakdown of self-averaging*. This manifests itself, for instance, in the broadness of the distribution of some (not all) physical quantities. A similar conclusion was reached for short length scales in [13].

In summary, we reinvestigated the theory of finite size scaling in disordered systems. We found that the standard averaging procedure introduces a new diverging length scale into the problem, therefore the finite size scaling exponent $\nu_{\rm FS}$ may be unrelated to the intrinsic ν. We constructed two explicit examples, where *exact* calculations proved that the intrinsic $\nu < 2/d$. We proposed an alternative averaging, which achieves a remarkable noise reduction and therefore is capable of accessing the intrinsic exponents of the physical problem.

We acknowledge useful discussions with L. Chayes, T. Giamarchi, D. Huse, A. van Otterlo, H. Rieger, S. Sachdev, R. Singh, P. Weichman, and P. Young. This research was supported by NSF DMR-95-28535.

[1] J. T. Chayes, L. Chayes, D. S. Fisher, and T. Spencer, Phys. Rev. Lett. **57**, 2999 (1986); J. Chayes *et al.*, Commun. Math. Phys. **120**, 664 (1989).
[2] A. B. Harris, J. Phys. C **7**, 1671 (1974).
[3] O. Narayan and D. S. Fisher, Phys. Rev. B **46**, 11 520 (1992).
[4] J. Kisker and H. Rieger, cond-mat/9612027.
[5] P. Crowell, F. Van Keuls, and J. Reppy, cond-mat/9612133.
[6] M. Paalanen, T. Rosenbaum, G. Thomas, and R. Bhatt, Phys. Rev. Lett. **48**, 1284 (1982); H. Stupp *et al.*, Phys. Rev. Lett. **71**, 2634 (1993).
[7] F. Pázmándi and G. T. Zimányi, cond-mat/9702169.
[8] D. Fisher, Phys. Rev. B **51**, 6411 (1995).
[9] See the papers in *Finite Size Scaling*, edited by J. L. Cardy (North-Holland, Amsterdam, 1988).
[10] K. Binder, Z. Phys. B **43**, 119 (1981).
[11] R. Singh and M. Fisher, Phys. Rev. Lett. **60**, 548 (1988).
[12] This approach is problematic in the presence of power-law correlations or long range interactions.
[13] A. Aharony and B. Harris, Phys. Rev. Lett. **77**, 3700 (1996).