Is the cause of size effect on structural strength fractal or energetic–statistical?

Zdeněk P. Bažant a,*, Arash Yavari b

a Departments of Civil Engineering and Materials Science, Northwestern University, 2145 Sheridan Road, CEE-Tech A135, Evanston, IL 60208, USA
b Graduate Aeronautical Laboratories and Department of Mechanical Engineering, California Institute of Technology, Pasadena, CA 91125, USA

Received 23 October 2003; received in revised form 10 March 2004; accepted 11 March 2004

Abstract

The size effect on structural strength is an important phenomenon with a very old history. Unfortunately, despite abundant experimental evidence, this phenomenon is still not taken into account in most specifications of the design codes for concrete structures, as well as the design practices for polymer composites, rock masses and timber. The main reason appears to be a controversy between two different theories of size effect, namely the theory based on energetic–statistical scaling and the theory based on ideas from fractal geometry. This paper aims to critically analyze these two theories, examine their hypotheses and point out the limitations, in order to help code-writing committees choose a rational basis for their work. The paper begins by reviewing the theory of energetic size effect and the efforts to explain the size effect by fractal geometry. The advantages and disadvantages in modeling the structural size effect by fractals are pointed out. Certain flaws in the fractal theory of size effect are illuminated and it is shown that various aspects of this theory lack a sound physical or mathematical basis, or both. The paper ends by recommending how engineering designers and code writers should take the size effect into account.

1. Introduction

1.1. Motivation

Quasibrittle failures of concrete structures, as well as rock masses, sea ice bodies, fiber–polymer composites, ceramics and timber, exhibit a large statistical scatter. In static testing of very large structures, proper similitude is hard to achieve because of the self-weight. A statistically significant set of test results on a sufficiently large number of identical large structures cannot be obtained because of prohibitive costs. The largest concrete structures, sea ice bodies or rock formations are way beyond the range of failure testing.
Consequently, the design of large structures must rely on extrapolation from test results on much smaller laboratory specimens (Fig. 1(top)). Having a good theory is, therefore, imperative.

Currently there exist two fundamentally different theories of size effect in quasibrittle structures: (1) the energetic–statistical theory, and (2) the fractal theory. The conflict between these two theories is a serious impediment to progress in structural design codes and practice. Despite a presidential inaugural address with the motto *ars sine scientia nihil est* (‘art without science is nothing’) [95], society committees will not adopt a scientific approach to size effect if they get from the literature the impression that there is no generally accepted scientific theory of this phenomenon.

For example, a few recent scaled tests of shear failure of large reinforced concrete beams [1,74], together with a few older ones [93,120], convinced most members of a code-making committee that a significant nonstatistical size effect indeed exists. But many members, aware of an apparent irreconcilable disagreement among theoreticians, feel reluctant to incorporate a scientifically based size effect formula into the building code specifications. They prefer, therefore, a purely empirical approach in which the existing small-size laboratory data are fitted by some simple intuitive formula approved by the vote of a committee, and this formula is then simply extended to large sizes.

Such an empirical approach, however, is doomed to yield an incorrect formula, for two reasons: (1) very high scatter (Fig. 1(bottom)), and (2) vast predominance of test data in the small-size range (Fig. 1(top)) which makes extrapolation to the large sizes of main concern highly uncertain [42]. Because of scatter, the existing test data can be fitted almost equally well by different formulas giving very different extrapolations to large sizes. How a purely empirical evaluation of test data can mislead is explained by Fig. 2, illustrating the kind of approach that has recently been taken by a certain society committee and has been used to claim that a code formula for the size effect on shear of reinforced concrete beams should have the form of a power law of exponent −1/3 (this would, in fact, overestimate the strength of extremely large structures by almost 100%). In Fig. 2(top) it is assumed that hypothetical test data of a limited size range, obtained for four different concretes, perfectly agree with the curves of the theoretical size effect law, with different material parameters for each concrete. Paying no attention to the theory and considering the data for the four concretes as one statistical set, this committee fits the available data by a straight line in the biloga-
rithmic plot, for which a slope of $-1/3$ is obtained (as shown in Fig. 2(top right)). However, such empirical extrapolation of the available test data is unsafe (causing an error of about 80% at very large structure sizes; Fig. 2(bottom)). Obviously, if the size range of available data in Fig. 2(top) happened to be different, a power law of any exponent between 0 and $-1/2$ could be obtained by such a fallacious approach to data analysis. To obtain a realistic extrapolation, a sound theory, justified by geometrically scaled tests of sufficient size range, performed on one and the same concrete, must be invoked for interpreting the data.

The foregoing example makes it clear that, until the theoretical disagreement is clarified, until it is demonstrated what is and what is not a sound theory, rational design code specifications guarding against the size effect will hardly be introduced by code-making committees into the concrete design codes of various countries. The purpose of this article is to provide such clarification, and thus to facilitate innovative designs and improvements of design safety and economy.

The problem of codes is particularly acute for structural engineering, and not only because large structure tests are unavailable. In that branch of engineering, many thousands of structures are designed annually. Each of them is different, and most must be designed quickly. This makes simple formulas, capturing the main trends, indispensable, despite the difficulty in developing such formulas. By contrast, for example, aircraft engineering does not need such formulas. Only a few large aircraft are designed per decade, which means that design codes are not needed, and sophisticated analyses and large-scale tests are affordable.

1.2. Size effect and its background

In the classical theories of elasticity, plasticity and continuum damage mechanics, the failure criterion is expressed in terms of stresses and strains, and no characteristic material length $\ell$ is present. This causes the nominal strength of geometrically similar structures, defined as $\sigma_N = P/bD$, to be independent of structure.
size $D$, i.e., there is no size effect ($P = \text{maximum load}, b = \text{structure thickness}$). The dependence of $\sigma_N$ on $D$, called the size effect, is generally caused by the existence of some sort of characteristic length, $\ell$.

Interest in the size effect is very old, in fact older than the mechanics of materials itself [75]. The primeval scaling idea was Galileo’s [87] invention of the concepts of stress and strength, the size-independence of which correctly determines the scaling of failure loads for many materials. The first meaningful theory of size effect, understood as the deviation from Galileo’s idea of constant material strength, was the statistical weakest-link theory. Although the basic concept of that theory was qualitatively proposed already in the 17th century by Mariotte [102], almost three centuries had elapsed until the correct and complete mathematical formulation of extreme value statistics (or weakest-link model) was developed by Fisher and Tippett [84], who were the first to derive what later came to be known as the Weibull distribution. Based on experiments on fatigue fracture of metals and heuristic arguments, Weibull [122] introduced this probability distribution into the theory of fatigue failure of metals and ceramics and obtained a power law for the statistical size effect.

For about half a century afterwards, Weibull’s theory reigned supreme, all the experimentally observed size effects, in all the materials, being attributed to Weibull statistical theory. Serious discrepancies, however, transpired from various experiments, first conducted on concrete [121] and then on other materials, all of which are now termed quasi-brittle (rocks, sea ice, fiber composites, toughened ceramics, dry snow slabs, wood, paper, etc.). These are materials that lack plasticity and are characterized by gradual softening in a fracture process zone (FPZ) that is not negligible compared to structure size $D$. These discrepancies led to the development of a new deterministic energetic theory [8,9,12,21–23,29,31–34,40,103], in which the size effect is explained by stress redistribution and the associated energy release due to development of large cracks or a large FPZ prior to failure. Amalgamation with the Weibull theory then led to a general energetic–probabilistic theory [19,20,35–38,44], in which the failure probability at a point of a continuum depends on a weighted average strain in a certain characteristic neighborhood of the point, rather than on the stress at that point. As far as the statistical mean of size effect in quasi-brittle structures is concerned, the probabilistic generalization of the energetic theory was shown to be necessary only for failures occurring right at fracture initiation in large enough structures in which the FPZ is negligible compared to $D$ [19,37,38]. For the variance and probability distribution of size effect, the probabilistic generalization of the energetic theory is, of course, always necessary.

2. Overview of energetic–statistical theory of size effect

2.1. Classical Weibull statistical theory

The classical statistical theory of size effect applies to structures that fail as soon as a macroscopic fracture initiates in one small material element of the structure. This theory is based on the weakest link model, the three-dimensional continuous generalization of which yields the following distribution of failure probability (Fig. 3(left)):

$$P_1(\sigma_N) = 1 - \exp \left[ -\int_V c(\sigma(x), \sigma_N) dV(x) \right]$$  \hspace{1cm} (1)

Here $\sigma = \text{stress tensor field just before failure}$, $x = \text{position vector}$, $V = \text{spatial domain occupied by the structure}$, and $c(\sigma) = \text{function giving the spatial concentration of failure probability of the material}$ ($= V^{-1}_r \times \text{failure probability of material representative volume } V_r$) (see [86,119]); $c(\sigma) \approx \sum_i P_i(\sigma_i) / V_0$ where $\sigma_i = \text{principal stresses } (i = 1, 2, 3)$ and $P_1(\sigma) = \text{failure probability (cumulative) of the smallest possible test specimen of volume } V_0$ (or representative volume of the material) subjected to uniaxial tensile stress $\sigma$ [122–125].
Here $m$, $s_0$, $r_u = \sigma_u$ are material constants; $m = \text{Weibull modulus}$, a material parameter which characterizes the coefficient of variation of strength scatter and varies widely (between 5 and 50); $s_0 = \text{scale parameter}$; $r_u = \text{strength threshold}$ (which may usually be taken as 0), and $V_0 = \text{reference volume}$ understood as the volume of small test specimens on which $c(\sigma)$ was measured (for recent generalization for plastic crack-tip blunting, see e.g., [46,96,114]).

Distribution (1) with (2) generally implies the mean and the coefficient of variation of the nominal strength to be

$$\bar{\sigma}_N \propto \Gamma(1 + m^{-1})(D_0/D)^{n_d/m}, \quad \omega = [\Gamma(1 + 2m^{-1})/\Gamma^2(1 + m^{-1}) - 1]^{1/2}$$

in which $n_d = \text{number of dimensions in which the structure is scaled}$ ($n_d = 1, 2, 3$). Because, according to Weibull theory, $\omega$ is independent of $D$, the measurement of $\omega$ is normally used as the easiest way to identify $m$ from tests (but unfortunately it is usually forgotten to check whether the same $\omega$ is obtained from size effect tests and scatter test specimens of various sizes, which is a condition of validity of Weibull theory; e.g. [127,136]).

As far as quasibrittle structures are concerned, application of the classical Weibull theory faces a number of objections:

1. The fact that the Weibull size effect is a power law
   $$\sigma_N \propto D^{-n_d/m}$$
   implies the absence of any characteristic length, which cannot be correct if the material exhibits an FPZ of a finite, non-negligible, size [2,11].
2. The energy release due to stress redistribution caused by macroscopic FPZ or stable crack growth before $P_{\text{max}}$ gives rise to a deterministic size effect which is ignored.
3. Every structure is mathematically equivalent to a uniaxially stressed bar (or chain, Fig. 3). 
4. The tests of concrete, ice, fiber composites, rocks and ceramics show a much stronger size effect than explicable by Weibull theory.
5. The $m$-values inferred from scatter are not independent of structure size $D$ and do not match the $m$-values determined from size effect tests [15].
6. Spatial correlations of material failure probabilities caused by nonlocal behavior are not taken into account.
7. The predicted difference in size effect between two- and three-dimensional similarity ($n_d = 2$ or 3) in flexure tests is excessive.
Consequently, in the case of concrete structures, the mean size effect of the classical Weibull theory appears applicable only to the failure of extremely thick plain (unreinforced) structures, e.g., to the flexural fracture of arch dams.

2.2. Deterministic energetic theory of quasibrittle size effect

In quasibrittle materials, the failure is characterized by both a critical energy per unit area (or the fracture energy, \( G_1 \), of dimension Pa m) and a critical stress (or the tensile strength, \( f'_0 \), of dimension Pa). From dimensional analysis, it is clear that a material with such properties must possess a characteristic length, \( \ell_0 = \frac{EG_1}{f'_0^2} \) (\( E = \) Young’s elastic modulus); \( \ell_0 \) approximately characterizes the length of FPZ, as proposed by Irwin [94] (and for concrete promulgated by Hillerborg et al. [90, 91]). Linear elastic fracture mechanics applies asymptotically when \( \ell_0/D \rightarrow 0 \), and in that limit case \( f'_0 \) becomes irrelevant. The cohesive (or fictitious) crack model (originated by Barenblatt [3, 4]) is needed when \( \ell_0/D \) is neither negligible nor very large. In the asymptotic limit \( D/\ell_0 \rightarrow 0 \), \( G_1 \) becomes irrelevant and the cohesive crack model leads to an elastic body with a perfectly plastic crack of yield strength \( f'_0 \).

The general approximate size effect laws for quasibrittle failures described by the cohesive crack model or nonlocal damage models can be most generally derived by asymptotic matching—a procedure in which the large-size asymptotic behavior described by LEFM is expanded into an asymptotic series in terms of powers of \( \ell_0/D \), the small-size asymptotic behavior described by plasticity is expanded into an asymptotic series in terms of powers of \( D/\ell_0 \), and both series are then approximately matched to obtain a size effect law approximately applicable over the entire size range [21, 28]. For structures with pre-existing notches or large pre-existing fatigue-weakened cracks, this leads to the size effect law [9]

\[
\sigma_N = B f'_0 \left(1 + \frac{D}{D_0}\right)^{1/2}, \quad B f'_0 = \sqrt{\frac{EG_1}{g'(x_0)c_i}} \quad D_0 = c_i \frac{g'(x_0)}{g(x_0)}
\]

(5)

Here \( B \) is a dimensionless geometry-dependent parameter, which depends on the geometry of the structure and of the crack, and can be calculated from the cohesive crack model [41]; \( P \) is the applied load; \( K_1 \) is stress intensity factor; \( x_0 = a_0/D \); \( a_0 \) is initial crack length; \( D_0 \) is geometry-dependent parameter called the transitional size; and \( c_i = \) half-length of FPZ proportional to \( \ell_0 \); the primes denote derivatives; and \( g(x_0) = K_1^2(x_0)b^2D/P^2 \) is dimensionless LEFM energy release function for the given geometry \( (g(x_0) = k^2(x_0)) \) where \( k(x_0) \) is dimensionless stress intensity factor [40]. The first formula in (5) has also been derived in several other ways—analytically, from the cohesive crack model [21, Eq. (9.40), 115]; by a combination of dimensional analysis with asymptotic matching [43]; and by asymptotic expansions of equivalent LEFM [16]; and was also verified by simulations with nonlocal finite elements, discrete elements (lattice or particle models), and crack band model (see also [110]).

For structures failing at crack initiation, right after the FPZ has formed, the asymptotic matching leads to a different size effect law [10, 18, 21, 24, 31, 32]:

\[
\sigma_N = \sigma_x \left(1 + \frac{rD_b}{\eta D_b + D}\right)^{1/r}, \quad \sigma_x = \sqrt{\frac{EG_1}{c_i g'(0)}}, \quad D_b = \left\langle \frac{-g''(0)c_i}{4g'(0)} \right\rangle
\]

(6)

where \( r, \eta = \) empirical positive constants (of the order of 1); \( \langle \ldots \rangle \) denotes the positive part of the argument; \( B = \) dimensionless geometry-dependent parameter; and \( D_b = \) constant that can be regarded as the thickness of the boundary layer of cracking in beam flexure.

There exists also a third case of size effect. This is the case of stable growth of a large crack in which the cohesive stresses have not been reduced to zero by previous fatigue; see [21]. But the plots of the corresponding formula and of (5) are rather similar and hardly distinguishable in comparison with experiments.
Let us now give a simple intuitive argument leading to equation (5). Consider the rectangular panel in Fig. 3(right), which is initially under a uniform stress equal to $\sigma_N$. Introduction of a crack of length $a$ with an FPZ of a certain length and width $h$ may be approximately imagined to relieve the stress, and thus release the strain energy, from the triangular zones on the flanks of the crack band shown in Fig. 3(right). The slope $k$ of the effective boundary of the stress relief zone need not be known; what is important is that, for geometrically similar panels, $k$ is independent of the size $D$. It is also assumed that the situations at failure are approximately geometrically similar, i.e., $a/D \approx$ constant (which is true in many, but not all situations, and must be justified experimentally or computationally). The stress reduction in the shaded triangular zones, of areas $ka^2/2$ (Fig. 3(right)), causes the strain energy release $U_a = 2(ka^2/2)\sigma_N^2/2E$ (for the case $b = 1$). The stress drop within the crack band of width $h$ causes further strain energy release $U_b = ha\sigma_N^2/E$. The total energy dissipated by fracture is $W = aG_f$, where $G_f$ is the fracture energy, a material property representing the energy dissipated per unit area of the fracture surface. Energy balance during static failure requires that $\partial(U_a + U_b)/\partial a = dW/da$. Setting $a = D(a/D)$ (where $a/D$ is approximately a constant if the failures for different structure sizes are geometrically similar), the solution of the last equation for $\sigma_N$ yields the approximate size effect law in (5) (see [9] and Fig. 4(top left)).

Likewise, one can give a simple intuitive argument [31] leading to equation (6). For the sake of illustration, let us consider the modulus of rupture test, i.e., the flexural failure of a simply supported beam of span $L$ with a rectangular cross-section of depth $D$ and width $b$, subjected to a concentrated load $P$ at midspan (Fig. 5(top left)). Due to material heterogeneity, causing the quasibrittle behavior, the maximum
Load is not decided by the stress \( \sigma_1 = 3PL/2bD^2 \) at the tensile face, but by the average stress value within a boundary layer of microcracking having thickness \( c_f \) that is a material property (equal roughly to one to three maximum inhomogeneity sizes). This value may be approximated by the stress \( \bar{\sigma} \) roughly at distance \( c_f/2 \) from the tensile face (which is at the middle of FPZ). Because \( \bar{\sigma} = \sigma_1 - \sigma'_1 c_f/2 \) where \( \sigma'_1 = \) stress gradient = \( 2\sigma_1/D \), and also because \( \bar{\sigma} = \sigma_0 = \) intrinsic tensile strength of the material, the failure condition \( \bar{\sigma} = \sigma_0 \) yields \( P/bD = \sigma_N = \sigma_0/(1 - D_b/D) \) where \( D_b = (3L/2D)c_f \), which is a constant because for geometrically similar beams \( L/D = \) constant. This expression, however, is unacceptable for \( D/D_b \to 0 \). But since the foregoing derivation is valid only for small enough \( c_f/D_1 \), one may replace this formula with (6), which is asymptotically equivalent for \( D/D_b \to 0 \) (Fig. 5(top middle)). This formula happens to also satisfy the asymptotic conditions of plastic failure for \( D/D_b \to 0 \), and thus is acceptable for the whole range of \( D \) (\( r \) is any positive constant). The values \( r = 1 \) or 2 have been used for concrete, while \( r \approx 1.44 \) gives the optimum fit of the concrete test data from the literature [18].

Derivations of (5) and (6), applicable to arbitrary structure geometry, have been given in terms of asymptotic analysis based on Rice’s path-independent \( J \)-integral [18,21], and also on the basis of equivalent LEFM [15].
2.3. Combined statistical–energetic size effect in quasibrittle materials

The Weibull statistical theory and the energetic theory (which arose from the fracture propagation concept of Griffith) have been combined in the nonlocal Weibull theory [35,36,44], which has both aforementioned theories as its asymptotic limits. The deterministic energetic size effect is obtained for not too large structure sizes, and the Weibull statistical size effect is obtained as the asymptotic limit for very large structures \((D \to \infty)\), provided that the failure (or instability) occurs at macrocrack initiation and that the structure geometry is positive. \(^1\)

The nonlocal concept was proposed, within the context of elasticity, in the 1960s [77,78,92], and has later been extended to hardening plasticity [79]. In the 1980s, it was adapted to strain-softening continuum damage mechanics and strain-softening plasticity [25,109], with three motivations:

1. to serve as a computational ‘trick’ (localization limiter) eliminating spurious mesh sensitivity and incorrect convergence of finite element simulations of damage;
2. to reflect the physical causes of nonlocality, which are three:
   a. material heterogeneity,
   b. energy release due to microcrack formation, and
   c. microcrack interactions; and
3. to simulate the experimentally observed size effects that are stronger than those explicable by Weibull theory. Because of material heterogeneity, the macroscopic continuum stress at a material point must depend mainly on the average deformation of a representative volume of the material surrounding that point, rather than on the local stress or strain at that point.

In the deterministic nonlocal theory for strain-softening damage or plasticity, the spatial averaging must be applied only to the inelastic part \(\varepsilon''\) of the total strain \(\varepsilon\) (or some of its parameters), rather than to the total strain itself [109]. Accordingly, the cumulative failure probability \(P_1(\sigma)\), considered in the classical Weibull theory as a function of the local stress tensor \(\sigma\) at material point \(x\), must be replaced by a function of a nonlocal variable [24,35,36,39,44]. The nonlocal stress is not acceptable as a nonlocal variable because it decreases when the average strain increases. A suitable nonlocal variable is the nonlocal strain or, more precisely, the nonlocal inelastic part of strain. The material failure probability is thus defined in the nonlocal Weibull theory as

\[
P_1 = \left( \frac{\sigma}{s_0} \right)^m
\]

where

\[
\bar{\sigma}(x) = E : [\varepsilon(x) - \bar{\varepsilon}''(x)], \quad \bar{\varepsilon}''(x) = \frac{1}{\bar{z}(x)} \int_V \alpha(s - x) \varepsilon''(s) dV(s)
\]

in which \(\varepsilon''\) = inelastic part of strain tensor; \(E\) = initial elastic moduli tensor; \(\bar{z}(x)\) = normalizing factor of \(\alpha(s - x)\); and \(\alpha(s - x)\) = a bell-shaped nonlocal weight function whose effective spread is characterized by

\(^1\) A ‘positive’ geometry is a geometry of structure (including geometry of crack and loading) for which the stress intensity factor increases as the crack extends at constant load (negative geometry— the opposite). Structures of positive geometry fail, under controlled load (or dead load), as soon as the process zone forms. Nearly all notched fracture specimens are of positive geometry. The exceptions are a large panel with a small crack loaded by concentrated loads on the crack faces, the reverse-tapered double-cantilever specimen, and the Charpy notched specimen; so is the flexural failure of unreinforced beams. Many reinforced concrete structures have an initially negative geometry, and so do many structures with a compression zone in front of fracture tip, e.g., a gravity dam with a dipping crack produced by over flow.
characteristic (material) length $\ell_0$, which for example governs the minimum possible spacing of parallel cracks and is in general different from (and much smaller than) $\ell_0$ characterizing the FPZ length.

The nonlocality makes the Weibull integral over a body with crack tip singularity converge for any value of Weibull modulus $m$, and it also introduces into the Weibull theory spatial statistical correlation. Numerical calculations of bodies with large cracks or notches showed that the randomness of material strength is almost irrelevant for the size effect on the mean $\sigma_N$ [44]. Therefore, the energetic mean size effect law (5) for the case of large fatigued cracks or large notches remains almost unaffected by material randomness. Intuitively, the reasons are that: (1) a significant contribution to Weibull integral comes only from the FPZ, the size of which remains constant if the structure size is increased, and (2) mechanics almost dictates the crack path, so that the FPZ cannot sample locations of different local strength.

One case in which the statistical size effect on the mean strength $\sigma_N$ is important is the failure at crack initiation in structures of positive geometry that are much larger than the inhomogeneity size. This is the case of bending of very thick plain concrete beams or plates (exemplified by the flexural failure of an arch dam, typically about 10 m thick) [35,36]. In this case, Eq. (6) needs a correction for large $D$ [22]. Based on asymptotic matching arguments, this equation needs to be generalized as [19] (Fig. 5(top right and bottom))

$$\sigma_N = \sigma_0 \left[ \left( \frac{D_b}{\eta D_b + \frac{r D_b}{\eta D_b + D}} \right)^{n_D/m} + \frac{r D_b}{\eta D_b + D} \right]^{1/r}$$

where $n_D/m < 1$; $\eta, r$ = empirical constants. This formula, which gives the statistical mean of size effect, asymptotically approaches the Weibull size effect as $D \to \infty$. It was shown to fit quite well the bulk of the existing test data on the modulus of rupture of concrete and to closely agree with numerical predictions of the nonlocal Weibull theory over the size range 1:1000 [36,39]. Parameter $\eta$ can be taken as zero for the fitting of the existing test data, however, a nonzero value is needed for the purposes of asymptotic matching, in order to match the asymptotic behavior of the cohesive crack model (or crack band model) for $D \to 0$ (normally approached only for sizes less than the size of material inhomogeneities). Aside from the

Fig. 6. Universal (mean, energetic–statistical) size effect law describing the transition between failures at crack initiation and failures for large notches or cracks formed in a stable manner before failure [43] ($x = a_0/D = \text{relative\ notch\ or\ crack\ length}; \ell_i, \ell_0; r$ = constants, $g_0 = g(x_0); g'_0 = g'(x_0)$) (the upper rim of the surface shown, $r = 0$, may be compared to MFSL).
two aforementioned asymptotic limits, the formula also satisfies, as a third asymptotic condition, the requirement that the deterministic formula (6) must be recovered for \( m \to \infty \). The statistical distribution of the size effect has also been deduced; it represents a transition from the Gaussian distribution for \( D \to 0 \) to the Weibull distribution for \( D \to \infty \) [24,39].

A more difficult fracture scaling problem arises in the transitional situation in which there is a notch or initial crack not much larger than the material inhomogeneities, e.g., the aggregate size or the thickness of a layer in a laminate. Information on this transition is computationally best generated by random particle-lattice models for concrete microstructure (e.g., [26]), but can also be obtained by nonlocal Weibull model [39]. Asymptotic matching approach to this problem has led to a universal mean size effect law presented in [43] and pictured in Fig. 6 (for the deterministic case, see [32]). The formula of this law, shown in Fig. 6, satisfies six asymptotic conditions: (1) for a large \( a_0 \) (initial crack or notch length) and \( D \to \infty \)—convergence to LEFM; (2) for large \( a_0 \) and \( D \to 0 \)—to a plastic limit (with no size effect) required by cohesive crack model, crack band model or nonlocal models; (3) for \( a_0 = 0 \) (failure at crack initiation, and \( D \to \infty \)—to pure Weibull scaling; (4) for \( a_0 = 0 \) and \( D \to 0 \)—again to the aforementioned plastic limit; (5) for \( m \to \infty \)—to deterministic theory; and finally (6) for \( l_0 \ll D \ll \eta l_0 \) (i.e., between the nonlocal characteristic length \( l_0 \) and FPZ length \( \eta l_0 \), as explained in [24])—convergence to an intermediate asymptotic behavior (in the sense rigorously mathematically discussed in [6]).

3. Fractal theories of size effect on nominal strength of structures

3.1. Fractal characterization of crack randomness, disorder and roughness

The fact that the surface roughness of cracks in many materials as well as the distributions of microcracks are physical fractals, i.e., can be described over a certain limited range by fractal concepts, is not in doubt (see [57,99,108,111,116,117,129] and references therein). To date, many attempts have been made to relate the fractal dimension (or roughness) of a crack to material properties such as the toughness, and some experimental studies have been conducted to find some universality in the fractal dimension (or roughness exponent, see [55]). However, it seems that there is no simple universal relationship between fractal dimension and fracture toughness or fractal energy, and no relationship between fractal dimension and material characteristic length (length and width of the fracture process zone). Looking at the literature on the fractal properties of fracture, one can see that most studies are experimental.

An important point to note is that simply knowing the fractal dimension of a fracture surface or a distribution of microcracks does not help in understanding the mechanics of failure. There exist only a few theoretical studies on the mechanical consequences of fractality of cracks (see [12,13,15,21,27,49–53,130–135]). Unfortunately, there have been conceptual mistakes in some recent papers. One can also find some unnecessarily sophisticated and unjustified generalizations of classical results that contribute nothing but complexity and are practically unusable.

Several researchers have tried to use fractals in the modeling of structural size effect. Among these, one can mention Weiss [126], Borodich [52,54], Gol’dshtein and Mosolov [89] and Bažant [15]. Gol’dshtein and Mosolov [89] used fractals and multifractals for the purpose of size effect, although with limited results. Self-affine crack roughness [47,73] was considered by Weiss [126] and Morel and coworkers [104–106] to model the size effect on different scales using the transition property of self-affine curves. Borodich [52] used self-similar fractals in the modeling of multiple fracture. What particularly calls for a critical examination in this paper, which is focused on the size effect law needed for design, are the formulations of Carpinteri and co-workers [58,70]. It will be seen, however, that these formulations, while thought-provoking, lack a solid basis and have various fundamental flaws.
3.2. The so-called ‘multifractal’ scaling law (MFSL)

A possible role of fractality in the size effects was discussed in 1990 in relation to sea ice [45]. The partly fractal nature of crack surfaces and of the distribution of microcracks in concrete was considered in 1994 as the physical origin of the size effects observed on quasibrittle structures, and the so-called ‘multifractal’ scaling law (MFSL), which can be used for concrete structures failing at fracture initiation from a smooth surface, was proposed [58]. This law reads

\[ \sigma_N = \sqrt{A_1 + (A_2/D)} \]  \hspace{1cm} (10)

where \( A_1, A_2 = \) constants; see also [59–62,64–67]. In an attempt for mathematical derivation (critically discussed in Appendix A), it was argued [58] that (i) the fractal nature of crack surface or microcrack distribution, or both, requires the scaling \( \sigma_N \propto \sqrt{D} \), while (ii) the disappearance of fractality at very large-scales implies the vanishing of size effect, i.e. \( \sigma_N = \) constant. In view of these two assumed opposite asymptotic properties, the MFSL formula (10) was then proposed as a way to achieve a smooth transition.

What Carpinteri calls ‘multifractals’ is different from the concept of multifractals in mathematics and should properly be called ‘scale-dependent fractals’ (see [72]) (a multifractal is a measure, supported on some geometrical set; see e.g. Feder [82] for details).

The MFSL does not capture, in contrast to formula (9), the transition to Weibull size effect for very large sizes. But that the MFSL allows a good fit of the test data on the modulus of rupture of concrete (i.e., the flexural strength of unreinforced beams) on the laboratory scale is undeniable. However, a formula identical to MFSL has been derived from fracture mechanics [12,18,21,31] based on a reasonable hypothesis already discussed—namely that there exists a boundary layer of cracking that has a fixed thickness and causes stress redistribution with energy release before the maximum load is reached. In fact, the MFSL, Eq. (10), is found to be identical to the special case of formula (6), as well as its probabilistic generalization (9), if one sets [18]

\[ r = 2, \quad \eta = 0, \quad m \to \infty \]  \hspace{1cm} (11)

\[ A_1 = EG_t/c_1g'(0), \quad A_2 = -EG_tg''(0)/2c_1[g'(0)]^2 \]  \hspace{1cm} (12)

The last two equations relate the MFSL parameters to \( g(x) \), the dimensionless energy release function of LEFM, and thus they automatically provide the geometry dependence of \( A_1 \) and \( A_2 \), if it is accepted that the explanation of (10) is energetic rather than ‘multifractal’.

3.3. Comparisons and problems of MFSL

Two kinds of comparisons of MFSL need to be discussed: one meaningless and one meaningful. First the meaningless. The proponents of MFSL, since 1995 until the present (see e.g. [63], or Fig. 2 in [70], or Figs. 10–13 in [71]), have never compared MFSL in their works to (6). Rather, they have habitually compared it to (5). But this equation, which is the size effect law for structures with large notches or with large cracks at maximum load, has never been claimed to apply for structures failing at the initiation of a macroscopic crack, the sole case for which the MFSL can be applied. Comparisons of these two formulas, intended for different situations, only sow confusion and make no sense at all.

Now the meaningful comparison, which is a comparison of (10) with (6). This comparison is crucial and must be clarified before the size effect can be introduced into practical design.

Although the fact that many features of disorder in the microstructure of concrete and other quasibrittle materials can be characterized in terms of fractal concepts is not questioned, a number of objections must be raised against the proposition that the size effects on the nominal strength of structures, and the MFSL (Eq. (10)) in particular, follow from the hypothesis of fractal microstructure. They are as follows:
1. The stress redistribution and energy release phenomena associated with large cracks or notches, or large FPZ, or both, must cause a size effect. They are by now well proven from many angles, experimentally as well as theoretically. Therefore, if fractality of the microstructure should be regarded as a physical source of size effect on the structural level, it cannot be introduced as a replacement of the stress redistribution and energy phenomena. It could come only as some additional feature, a refinement. For example, it is certainly possible that microstructure fractality could help to explain the statistical scatter about the mean energetic size effect or some properties of the fracture energy, the strength limit and the softening stress-separation law of a cohesive crack.

2. In contrast to the energetic concept, Eq. (12), the fractal concept of structural size effect does not provide information on the dependence of the size effect parameters on the geometry of the structure. It cannot provide it unless the fractal boundary value problem (which is not even well-defined) is solved. This fact alone suffices to render the MFSL almost useless for practical applications.

3. The ‘MFSL’ was based on a series of hypotheses but does not follow from these hypotheses by a valid mathematical procedure (see Appendix A). In particular, the argument based on dimensional analysis [118] is mathematically inconsistent.

4. For very small sizes \( D \), the MFSL (Eq. (10)) gives the scaling law \( \sigma_N \propto D^{-1/2} \). The asymptotic small-size power scaling law of exponent \(-1/2\) (which corresponds to the value \( r = 2 \) in (6)) agrees reasonably well with the data for concrete. In view of the universality of this exponent in the random walk and percolation models, this is no surprise. However, the value \(-1/2\) is an unproven conjecture which does not follow from the fractal hypothesis (Appendix A). Moreover, it has not been explained why the scaling law exponent should be independent of the fractal dimension \( D_f \) that characterizes the cracking morphology. Should not the exponent be different for very different fractal dimensions \( D_f \), for example when \( D_f \) deviates from the Euclidean dimension by 40% or by 2%? In the extreme, one may consider the difference of the fractal dimension from the Euclidean dimension to be almost vanishing, for instance 0.001%, and in that case one must expect the scaling exponent to be very close to 0 because a sudden jump from \(-1/2\) to 0 as \( D_f \rightarrow 0 \) would obviously be irrational. This argument demonstrates that the exponent of the power scaling law cannot be independent of the fractal dimension if the fractal hypothesis is adopted (see [133] for similar considerations and recovery of classical results as a limit case of their fractal counterpart).

5. The lack of fractality on a large-scale is assumed to imply the absence of size effect. In comparison with the small-size value \(-1/2\), this implies that the size effect exponent would depend on the fractal dimension. However, as already pointed out, this conflicts with the fact that, for very small sizes, the exponent is implicitly considered as independent of that dimension.

6. If the hypothesis of fractal origin of the structural size effect were justified, and if the argument for fractal-based scaling were valid, then an \( n \)-fold increase of beam width would have to cause the same size effect on \( \sigma_N \) as an \( n \)-fold increase of beam depth. But this is not the case. The beam width has almost no effect on the modulus of rupture. This empirical conflict alone suffices to reject the hypothesis of fractal origin of structural size effect.

7. In microscopic observations, the fractality of fracture surfaces and microcrack distributions is observed for up to only about 1.5 orders of magnitude of refinement. Such a range of fractality is much narrower than the range important for size effect laws. Besides, within such a narrow range, a rough crack regarded as a fractal can be described about equally well by statistical models for surface roughness. Normally, the fractality needs to be experimentally observed through about six orders of magnitude for the fractal scaling to be considered a very good model.

8. The mathematical arguments advanced in support of the fractal concept invoked [68,69] the renormalization group transformation [5,6]. This transformation relates the power scaling law for one size range to a different power scaling law for an adjacent size range, giving the intersection of these two power laws (which appear as straight lines of different slopes in a bi-logarithmic plot). For quasibrittle materials, however, the salient question is not only the intersection but also the gradual transition between these
two power laws in the size coordinate. This transition occupies several orders of magnitude (about 3) of structure size $D$ and represents the size range of practical interest. The fractal approach as formulated so far [68,69] says nothing about this transition, while the energetic approach based on stress redistribution does.

The fractal arguments have also been extended to provide an alternative explanation for the dependence of fracture energy of quasibrittle material on structure size, and for the apparent variation of fracture energy (evaluated according to LEFM) during the initial crack growth, called the $R$-curve (resistance curve) [68]. However, both phenomena have already been explained by nonfractal concepts—by the initial growth of FPZ and by nonlocal continuum damage models simulating this growth [40]). It is not denied that fractal concepts could play some role. However, this role would have to be a refinement of the energetic (or stress redistribution) mechanism, rather than its replacement (the same as in point 1 above).

Questions also arise with regard to the recent attempts at general fractal-based continuum mechanics in which the stress, strain, energy density and mass density have fractal dimensions [68,69]. The difficulty of developing such a radical theory is apparent from the fact that the concept of fractal stress has so far been enunciated only in the uniaxial setting, i.e., merely as a force per unit measure of a lacunar cross-section rather than a multidimensional tensor. The same uniaxial limitation also applies to the concept of fractal strain. In more than one dimension, it is necessary to define the surface on which a force (the fractal stress) is supposed to act. However, a fractal surface is nowhere differentiable, and so the normal to such a surface is not defined. Thus the generalizations made in [68,69] appears to be unjustified and useless, achieving nothing but artificial complexity (see Section 4.4).

It may be noted that a similar controversy about the applicability of fractal models has developed in turbulence [21], a field that is still far from fully understood despite a century of research. Dimotakis and Catrakis [72,76] recently explain that the problem of turbulence is too complex to be modeled by a fractal of constant fractal dimension and show that a fractal with scale-dependent variable fractal dimension would have to be used.

3.4. Crack characterization by self-affine roughness and simulation of $R$-curve

While a fractal curve is obtained by self-similar disturbances relative to the local direction of the curve (i.e., by repeatedly scaling down each multisegment section and substituting it for a single segment on the lower-scale), a self-affine curve is obtained by imposing lateral disturbances transverse to the direction of the global fracture direction, which are self-similar only in the transverse direction (Fig. 7(top left)).

Morel and coworkers [104–106] studied the consequences of self-affine crack roughness with a focus on mildly heterogeneous materials in which the source of crack roughness appears to be the interaction of the crack front with microcracks. To simulate their experimental observations of crack surfaces in wood (pine and spruce), they described the fluctuations of the asperity height $\Delta h$ along the crack curve as a transition between two power laws.

Specifically, Morel et al. considered a generalization of the Family-Vicsek law [81]—an ‘anomalous’ scaling law according to which $\Delta h(l,y) \approx A_l^{\zeta \text{loc}} \zeta(y)^{z_{\text{loc}}}$ for $l \ll \zeta$ (large-scale of observation) and $\Delta h(l,y) \approx A^{\zeta(y)}$ for $l \gg \zeta$ (small-scale of observation) where $l =$ length of observation window (or ruler length), $\zeta(y) = B y^{j/z}$, $z_{\text{loc}} =$ local roughness exponent, $\zeta =$ global roughness exponent, $y =$ distance from notch tip, and $A, B, z =$ constants ($z$ called the dynamic exponent).

In [105], Morel et al. present an energetic analysis of the size effect on the nominal strength $\sigma_N$, analogous to that in [15]. The result is similar to that in Fig. 7(bottom left) from [15], except that they assumed not one power law but a transition between two power laws for crack roughness, which causes that the size effect curve found (Fig. 3 in [105]) is the same as a transition in Fig. 7 from the nonfractal curve for small sizes (which eliminates the rising portion of the curve) to the fractal curve for large sizes, with an asymptotic
slope significantly less than $-1/2$. The transition from one size effect curve to another is caused by the fact that the ‘anomalous’ scaling law possesses a lower bound on the scale of crack roughness.

As already mentioned, such a milder asymptotic curve would disagree with the experimental evidence for concrete, sea ice, some rocks and fiber composites. In [105], a good agreement is nevertheless found with tests of geometrically similar notched wood specimens. However, the size range of these tests has been limited and it seems that a size effect curve as in Fig. 4(top right), with a long transition to the terminal slope of $-1/2$, could also match these test results (the asymptotic slope $-1/2$ being approached only beyond the size range of these tests). The reason that an asymptotic slope milder than $-1/2$ is obtained in [105] is that the assumed large-scale power law for crack roughness is considered to have no upper bound. Whether this is true deserves further study, especially since generally the ranges of fractal regimes are not unbounded.

Although the $R$-curves (e.g., [40]) are not the focus of this study, it may be mentioned that the growth of roughness of a self-affine crack leads to $R$-curve behavior. This is explained by noting that if the nominal (projected) crack length is increased by $\Delta a$, its measure increases in proportion to $(\Delta a)^{1/H-1}$, and thus the critical energy release rate increases as

$$G_R^f \sim 2\gamma_f (\Delta a)^{1/H-1}$$  \hspace{1cm} (13)

where $\gamma_f$ is a specific surface energy per unit of a fractal measure. This means that $G_R^f$ starts from $2\gamma_f$ and increases with the projected length. The self-affine roughness model of Morel et al. has led to a good agreement with $R$-curve measurements on wood [104–106]. However, such $R$-curves can be simulated more simply by the cohesive crack model, crack band model and nonlocal models, which better describe the fact
that the fracture process zone is finite, while the self-affine and fractal models for the $R$-curve presume the fracture behavior to be describable point-wise. It is also strange that the power law exponents for the $R$-curves for pine and spruce are found in these studies to be rather different (0.42 and 0.73), which suggests lack of universality.

4. Some fundamental problems of structural size effect based on material fractality

4.1. The concept of fractal crack

Before proceeding further, it will be helpful to review some basic ideas of fractality. When Mandelbrot introduced fractals [98,100], he concentrated on similarities between mathematical and physical (or empirical, natural) fractals. However, the modern papers on the subject study these two kinds of fractals separately [97] or at least they emphasize the difference between them [50]. The former fractals are observed in nature on a bounded region of size scale, while the latter kind gives mathematical models of these observations in the ideal limit case in which the size range tends to infinity. For example, if the length of a rough curve is measured by a ruler of a certain length $d_0$, representing the resolution (Fig. 7(left)), the measured length will evidently depend on $d_0$. If the curve is a physical fractal over a certain limited range, then this is described by the equation

$$a_d = \delta_0 (a/\delta_0)^{d_f}$$

(14)

where $a$ is the projected (or smooth, Euclidean) crack length, and exponent $d_f > 1$. This exponent is usually called the fractal dimension (or the physical fractal dimension). The relation (14) can be modeled by mathematical fractals. In this case, the resolution $\delta_0 \to 0$ while $a_d \to \infty$. Fractal dimensions of mathematical fractals can be defined in various ways. However, all of them are based on covering a fractal set by objects such as cubes, squares or line segments, the size of which is at least $\delta_0$, considering the sums of $\delta_0$ over the covered domain for various values of exponent $s$, and calculating a fractal measure of the set. Generally, the measure is either 0 or $\infty$. The value $s = d_f$ at which the measure jumps from 0 to $\infty$ is called the mathematical (in particular, Hausdorff or Hausdorff–Besikovitch) fractal dimension [80,98]. Note that only for $s = d_f$ the fractal measure of a set may have a nonzero bounded value [80] (it was suggested to call such sets as $d_f$-measurable [51]).

For a mathematical fractal model of a crack, Borodich noted [48,49,107] that the total energy dissipation $\mathcal{W}$ would be infinite if a finite amount of energy $G_f$ were assumed to be dissipated per unit crack length because the crack length $a_d \to \infty$ (he even called this as the ‘fractal cracking paradox’ [49,50]). To overcome this conceptual difficulty, he refined Barenblatt’s idea [5,7] to refer physical quantities to units of fractal measure $m_{df}$ and introduced a new concept of specific energy-absorbing capacity of a fractal crack $\beta(d_f)$, which may also be called the fractal fracture energy, having a dimension that is not J/m$^2$ but $1/m^{d_f+1}$. Note that originally the Barenblatt–Borodich idea was applied only to scalar objects such as the mass [5] or energy [48]. For example, if $G_{df}$ is the fracture energy related to $m_{df}(a)$ then $G_{df}m_{df}(a)$ is the fracture energy of a fractal crack of nominal length $a$. Using this notation one can calculate the fracture energy of a fractal crack of nominal length $ka$ ($k = constant$). Because $m_{df}(ka) = k^{d_f}m_{df}(a)$, the energy is found to increase $k^{d_f}$ times. Note, however, that the idea cannot be applied to vectorial quantities, introducing terms such as the force per unit of a fractal measure, because, mathematically, no normal to a fractal surface exists.

For a physical fractal crack, one can introduce the energy $G_f$ with the standard dimension J/m$^2$ and attribute it to the lower limit of validity of (14) [48]. For both mathematical and physical fractal cracks, one may then write

$$\mathcal{W}/b = G_f a^{d_f}$$

(15)
in which \( b = \) width of body; \( \dot{\gamma} \) = total energy dissipation; and \( G_0 \) represents either the fractal fracture energy \( G_{d_1} \) or the fracture energy \( G_{t1} \) for the smallest \( \delta_0 \) in (14). Based on this fractal concept of fracture energy, quasibrittle fracture propagation of a fractal crack has been analyzed [15] under the assumption that standard continuum mechanics and energy balance conditions are applicable on the scale of the structure as well as the FPZ. Some recent studies [68,69] developed the Barenblatt–Borodich idea to refer various physical quantities to a unit of the fractal measure \( m_{d_1} \) of the object. In these studies, however, this idea was applied to vectorial quantities such as the force, stress and strain, which is not acceptable.

4.2. Could crack surface fractality or self-affine roughness be the cause of size effect?

Let us now summarize the analysis in [15], which showed that the answer to the above question is negative. A crack was considered as a fractal curve in a plane (Fig. 7(left)). Eqs. (14) and (15) were taken as the point of departure. To take into account finite (projected) length \( 2c_t \) of the FPZ, the approximation by equivalent LEFM was used. In this approximation, the tip of the equivalent LEFM (sharp) crack is assumed to lie in the middle of the FPZ. For the case of specimens with notches or structures failing only after the development of large traction-free (fatigued) cracks, the matching of the large-size and small-size asymptotic expansions was shown to yield, instead of Eq. (5), the following approximate fractal size effect law:

\[
\sigma_N = \sigma_0^{D(D_t-1)/2} \left( 1 + \frac{D}{D_0} \right)^{-1/2} \quad \text{(16)}
\]

For the limit case \( D \gg D_0 \), corresponding to the fractal generalization of LEFM, this yields the large-size asymptotic behavior

\[
\sigma_N \propto D^{(D_t/2) - 1} \quad \text{where} \quad -1/2 \leq (D_t/2) - 1 \leq 2
\]

because \( 1 \leq D_t \leq 2 \). Obviously, the condition that the limit case of LEFM, \( \sigma_N \propto D^{-1/2} \), must be recovered for \( D_t \to 1 \) is satisfied.

For failure at crack initiation, the asymptotic analysis furnished, instead of Eq. (6), the result:

\[
\sigma_N = \sigma_0^{D(D_t-1)/2} \left( 1 + \frac{D_0}{D} \right) \quad \text{(18)}
\]

The large-size asymptotic behavior is

\[
\sigma_N \propto D^{(D_t-1)/2} \quad \text{where} \quad 0 \leq (D_t - 1)/2 \leq 1/2
\]

The plots of (16) and (18) are shown in Fig. 7(bottom) in comparison with the nonfractal laws (5) and (6) representing the limit case for \( D_t \to 1 \).

By judging the consequences for size effect, one may decide whether or not the hypothesis of fractal fracture energy is realistic [15,21]. Fig. 7(bottom) reveals that the fractal case disagrees with the available experimental evidence. For the case of structures failing only after large crack growth, the rising portion of the plot has never been seen in experiments, and there are many data showing that the asymptotic size effect is equal or very close to a power law of exponent \(-1/2\) ([21,40] e.g.), rather than an exponent of much smaller magnitude predicted from the fractal hypothesis. This is clear by looking at Fig. 7(bottom). For the case of structures failing at crack initiation, the kind of plots seen in this figure, with a rising size effect curve for large sizes, is also never observed. Thus it is inevitable to conclude that the hypothesis of a size effect caused by crack fractality, with a fractal fracture energy, is contradicted by test data and thus untenable.
The existence of limited fractal characteristics of fracture surfaces in various materials is of course not disputed, but crack surface fractality cannot cause a size effect in the mean. As for describing the statistical scatter of size effect about the mean, fractals might nevertheless be helpful.

What is the physical reason for the unrealistic consequences of the fractal hypothesis? Doubtless it is the fact that the crack front is not sharp but is surrounded by a large fracture process zone involving microcracks and frictional slips (in Fig. 7(right)). The development of a wide zone of extensive microcracking has been evidenced by locating sources of acoustic emissions, as well as by thermographic and holographic measurements. Besides, the length difference between the partially fractal crack curve and its smooth projection cannot account for the huge difference in energy dissipation because fractality of cracks surface is limited to only about 1.5 orders of magnitude, whereas the fracture energy $G_f$ of quasibrittle materials such as concrete is known to be several orders of magnitude larger than the surface energy $\gamma$ of the solid from which the structure is made. Consequently, far more energy is dissipated in the volume of the FPZ than on the surface of the final crack, which is created by coalescence of some of the microcracks in the FPZ. An additional reason for this discrepancy is that, in concrete, frictional slips in the fracture process zone account for more than 50% of energy dissipation, as corroborated by the fact that only a part of crack opening is recovered upon unloading.[14]

As an alternative approach, mode I fracture in a solid with fractal crack has recently been studied by Yavari et al.[133] in terms of a fractal stress intensity factor $K_f$. This study, which has also been extended to self-affine cracks, has been limited to a generalization of LEFM, i.e., quasibrittle materials have not been considered. The asymptotic near-tip field of stress tensor $\sigma_{ij}$ is in this study assumed to be written in a separated form similar to the expression in LEFM, but with different exponents:

$$\sigma_{ij}(r, \theta) = K_f^i r^{-\alpha} f_{ij}(\theta)$$  \hspace{1cm} (20)

where (see [128,135])

$$\alpha = \begin{cases} 1 - D_f/2, & 1 \leq D_f \leq 2 \quad [K_f^i] = FL^{-1-D_f/2} \quad \text{self-similar cracks} \\ 1 - 1/2H, & 1/2 \leq H < 1 \\ 0, & 0 < H \leq 1/2 \quad [K_f^i] = FL^{-1-1/2H} \quad \text{self-affine cracks} \end{cases}$$  \hspace{1cm} (21)

Here $F, L =$ quantities of force and length dimension; $H$ is the roughness exponent of self-affine crack, and the physical dimension of $K_f^i$ depends on $H$ or the fractal dimension $D_f$. The stress tensor, of course, is well defined only farther away from the crack, which is a measure-zero set in the plane. It is considered that the fractal fracture toughness $K_{fc}^i$ (critical fractal stress intensity factor) is a function of fractal dimension $D_f$ or roughness exponent $H$ of the crack. Similar to classical LEFM, it is assumed that a fractal crack propagates if $K_f^i = K_{fc}^i$. Yavari [135] and Yavari et al. [133] further restrict consideration to a mode I crack of length $2a$ so small that there exists a remote stress field $\sigma$ not disturbed by the crack (a situation typical of steel or fine-grained ceramics, but not relevant to quasibrittle materials). Based on dimensional arguments and in analogy to LEFM, they show that

$$K_f^i = \chi(D_f) \sigma \sqrt{\pi a^{2-D_f}}$$  \hspace{1cm} (23)

where $a$ is the projected length of the fractal crack, $\sigma$ is the remote stress, and $\chi$ is a coefficient depending on $D_f$. Considering an infinite body in which the only dimension is the crack length, they set $D = a$ and $\sigma = \sigma_N$. Inversion of (23) then yields the scaling law:

$$\sigma = \sigma_N = \frac{K_f^i}{\chi(D_f) \sqrt{\pi D_f^{2-D_f}}} \propto D_f^{D_f/2-1}$$  \hspace{1cm} (24)
This scaling relation is identical to the special limiting case of Bažant’s [15] analysis of quasibrittle fractal fracture, Eq. (17). This analysis rests on two, quite plausible, hypotheses—that the first law of thermodynamics (energy balance) must apply globally, and that the solutions to the fractal and nonfractal boundary value problems are the same except near the fractal crack surface. In view of our inability to solve boundary value problems with fractal boundaries, this equivalence is important to retroactively justify the hypotheses that the fractal stress intensity factor can be defined as in (20) and that its relationship to a remote uniform stress field \( \sigma \) can be written in the form of (23).

4.3. Could lacunar fractality be the cause of size effect?

In view of the difficulties with crack surface fractality, as just described, it was proposed to deal with another type of fractality—the lacunar (or rarefying) fractality of the solid caused by an array of microcracks (Fig. 7(top right)) or microvoids. First, we need to summarize the critical analysis of this hypothesis given in [15].

From distance, one can see in the material only large cracks. But, looking closer, one can discern that each crack is discontinuous and consists of shorter mini-cracks, with mini-gaps between them. Looking still closer, one can discern that each mini-crack is also discontinuous and again consists of shorter microcracks with microgaps between them, etc. Refinement ad infinitum produces a fractal set, with a fractal dimension \( D_f \) that is less than the Euclidean dimension of the space (Fig. 7 top right). It seems that the microcrack systems in concrete might be described by this type of fractality, but only for about one order of magnitude of crack size, which is hardly enough to justify a fractal treatment.

More recently, it has been argued that lacunar fractality could be the cause of size effect on the structural level [61,62,66,68,69]. For a small-scale, the fractal dimension \( d_f \) of the arrays of microcracks is considered to be distinctly less than 1, and for a large-scale equal to 1. It is supposed that, for the failure of a small structure, the small-scale matters, and for the failure of a large structure, the large-scale matters. Hence, as it was argued, there should be a transition from a power scaling law corresponding to small-scale fractality to another power scaling law corresponding to large-scale fractality, the latter having size exponent 0 for the nominal strength, i.e., no size effect. Thus, as it was proposed, the size effect in a plot of \( \log \sigma_N \) versus \( \log D \) would be given by a transitional curve between an inclined asymptote and a horizontal asymptote. The inclined asymptote was considered to have the slope \(-1/2\) (Fig. 7(bottom)), which leads to the MFSL (Eq. (10)).

However, even though for some of the existing test data on the modulus of rupture the slope \(-1/2\) is not unreasonable, the mathematical argument that was used to arrive at this slope is incorrect (Appendix A). Besides, there exist, for specimens of sizes as small as possible for the given aggregate size, other test data that indicate that the initial slope can be much less than \(-1/2\). As already pointed out, the formula of MFSL is a special case for \( r = 2 \) of formula (6) derived strictly from fracture mechanics, without any recourse to fractals; however, the optimal fit of the available test data is obtained for \( r = 1.44 \) [18,36].

On closer scrutiny, the hypothesis of lacunar fractality appears to lead to the classical Weibull statistical theory. If the failure is assumed to be controlled by lacunar fractality of the solid, rather than large cracks, it obviously implies that the failure occurs at crack initiation. In that case, the mathematical formulation must be akin to Weibull theory [15]. Labeling the aforementioned small and large-scales of observations of the lacunar material by superscripts \( A \) and \( B \), the Weibull distributions of the strength of a small material element at small and large-scales may be written as

\[
\varphi(\sigma(x); D_i^t) = \left( \frac{\sigma_N S(\zeta) c^t_i}{\sigma^A_0 - \sigma^A_0} \right)^m
\]  

(25)
in which the stress in the small material element of random strength has been written as \( \sigma = \sigma_N S(\xi) \); function \( S \) must be the same for all sizes of geometrically similar structures; and \( \xi = x/D = \text{dimensionless coordinate for the nonfractal (nonlacunar) case on the global structural level.} \)

For the fractal (lacunar) case, this is generalized as \( \sigma = \sigma_N S(\xi)^{C_1/D_f} \) because the stress of the material element, in the case of lacunar structure of the solid, must be considered to have a nonstandard, fractal dimension. Obviously, the Weibull constants \( ^\theta r_0 \) and \( ^\theta r_u \) must now be considered to have fractal dimensions as well, but Weibull modulus \( m \) must not. An equation of the type of Eqs. (25) or (26) was written by Carpinteri, et al., however, further analysis consisted of geometric and intuitive arguments. We will now sketch a recently published mechanical analysis [17].

In Weibull theory of failure at the initiation of macroscopic fracture, every structure is equivalent to a long bar of variable cross-section [44] (Fig. 9). The lacunarity argument means that a small structure is considered subdivided into small material elements, in which a low fractal exponent \( D_f < 1 \) is what matters (Fig. 9(a)), while a large structure is subdivided into proportionately larger material elements with \( D_f = 1 \) (Fig. 9(c)). However, a direct comparison of these small and large material elements is not objective, since structures made of the same material must be compared. The large material elements of the large structure (Fig. 9(c)) must be divisible into the small elements considered for the small structure (Fig. 9(a)), which may be identified with the representative volume of the material for which the material properties are defined. If the large elements were not divisible into the small ones, it would imply that the material of the small structure is not the same. So we must consider that the large material elements can be subdivided into the small material elements, as shown in Fig. 9(b). Accordingly, we may now calculate the failure probability of the large structure on the basis of the refined subdivision into the small elements, as shown in Fig. 9(b). We note that the failure probability \( P_f \) of the large structure subdivided into large elements \( \Delta V_{B_j} \) \((j = 1, 2, \ldots N)\), and the failure probability \( P^B_{ij} \) of the large element \( \Delta V_{B_j} \) of the large structure subdivided into small elements \( \Delta V_{A_{ij}} \), must satisfy the following relations based on the weakest link model underlying Weibull theory (based on the fact that the failure of one element implies failure of the whole structure):

\[
-\ln(1 - P_f) = \sum_j \varphi(\sigma_N S_j^\beta; D_{ij}^\beta) \Delta V_{B_j}/V_r
\]

(27)

\[
-\ln(1 - P_{ij}^B) = \sum_j \varphi(\sigma_N S_{ij}^\beta; D_{ij}^\beta) \Delta V_{A_{ij}}/V_r
\]

(28)
Now, since we may subdivide each element $B$ of the large structure into the small elements $A$ if the material is the same, we have the recursive condition:

$$- \ln(1 - P_i) = - \sum_j \ln(1 - P_{ij}^B) = \sum_j \sum_i \varphi(\sigma_{NJ}^B; D_l^B) \Delta V_{Ai}/V_r$$

(29)

Now, upon equating this to (27), we see that, in order to meet the requirement of the objective existence of the same material, the Weibull characteristics on scales $A$ and $B$ must be different, and precisely such that

$$\varphi(\sigma_{NJ}^B; D_l^B) = (\Delta V_{Ai}^B)^{-1} \sum_i \varphi(\sigma_{NJ}^A; D_l^A) \Delta V_{Ai}^A$$

(30)

From Eqs. (29) and (30) it follows that, for structures made of the same material failing at crack initiation (i.e., following the weakest link model), the consideration of different scales cannot yield different scaling laws. The same power law size effect must ensue from the hypothesis of lacunar fractality of the material, regardless of the scale considered. So the lacunarity argument leads to Weibull theory and offers nothing new in terms of the size effect on structural strength, although it might offer something new for the understanding of failure mechanism on the microlevel.

In summary, the scaling law of the nominal strength of a structure failing at the initiation of fracture in a lacunar fractal solid must be identical to the scaling of the classical Weibull theory. The only difference is that the values of Weibull parameters may depend on the lacunar fractality. This aspect could be quantified if the values of these parameters could be predicted by some sort of fractal micromechanics.

Before closing this subject, it should be noted that the concept of lacunar fractality should be used more carefully. According to Mandelbrot [101], who discussed the idea of lacunarity in his treatise [98], the concept of fractal lacunarity appears dubious from the mathematical viewpoint. Noting that the lacunarity
of fractals comes into play only when fractals have the same fractal dimension but different distributions of holes (lacunes), Mandelbrot argued that a measure of lacunarity could be the multiplicative pre-factors of fractal scaling laws, and that the fractal dimension itself cannot describe lacunarity. As explained by Mandelbrot [101], a single fractal dimension cannot properly describe a fractal set with holes. Describing lacunarity requires more than just a number. There are many examples of fractals with the same fractal dimensions but with very different distributions of holes. This fact should be taken into account by any lacunar fractal model because media with different distribution of holes (or defects, microcracks) have, in general, different mechanical behaviors.

4.4. Problems of continuum mechanics for bodies with fractal surfaces of discontinuity

To make the fractal approach to size effect more fundamental, it has recently been attempted to develop ‘continuum mechanics of fractal media’ (CMFM). However, it has been unclear what is CMFM. In continuum mechanics, each representative volume element is treated as a mathematical point. This implies the assumption that the size of the representative volume element is very small compared to the smallest dimensions of the structure. The structures can have a very complicated microstructure, but these details are not seen in the homogenized continuum. Consider three cases in which one might be tempted to introduce CMFM.

4.4.1. Continuum with a fractal microstructure

An example may be a fractal distribution of twins. Its description, however, does not require any generalization of continuum mechanics. Although the macroproperties might be affected by fractal characteristics, the homogenized continuum is simply a continuum and the standard continuum mechanics is applicable.

4.4.2. Continuum with a fractal distribution of microcracks, voids or other defects

Some researchers have in this case been tempted to use the so-called lacunar fractals for such a medium [61,62,66,68,69]. A lacunar fractal in $\mathbb{R}^n$ is a spongy-shaped subset with fractal dimension strictly less than $n$. Such a set may be thought of as a topological ball with infinitely many holes in it. A natural concept consistent with the axioms of the classical continuum mechanics is to consider the distributed holes or defects as a special microstructure. Its homogenization leads to the usual continuum. Some authors [69] nevertheless proposed to define a ‘fractal stress’ as a density of force per unit of a fractal measure. However, to accept such a fractal concept would require overcoming two problems. First, as already explained, it would be necessary to define clearly what is meant by a fractal continuum. Second, the definition of a fractal density of force would have to be generalized to more than one dimension. We will come back to this at the end of this section.

To be more precise, consider now a closed fractal curve in $\mathbb{R}^2$ (Fig. 8(right)). According to Jordan’s theorem, this curve partitions the plane into two subsets. If the interior region, denoted as $\Omega$, is bounded, then it must have a finite area, i.e.,

$$ \text{vol}(\Omega) < \infty $$

This means that it is meaningful to speak about a fractal surface in the framework of continuum mechanics. Having a body $\mathcal{B}$, we could consider a sub-body $\mathcal{P}$ with a fractal boundary $\partial \mathcal{P}$. This sub-body (Fig. 8(right)) has no unit normal vector on $\partial \mathcal{P}$ but has a finite volume and hence a finite mass. If a fractal set $\Omega \subset \mathbb{R}^3$ has fractal dimension $D_l < 3$, then (by definition of fractal dimension) $\text{vol}(\Omega) = 0$. Every (time-independent) problem in continuum mechanics is studied in $\mathbb{R}^3$, i.e., the continuum body $\mathcal{B}$ is embedded in $\mathbb{R}^3$. In any continuum mechanics problem, a finite body must certainly be considered to have a finite and nonzero mass (because the notion of a ‘body’ with zero mass would be physically meaningless). It is argued
that, in a continuum body with defects, any cross-section as a subset of \( R^2 \) is fractal, with some fractal dimension \( d_f < 2 \) associated with microcracks and other defects. But this means that the continuum \( \mathcal{B} \) has dimension \( D_f < 3 \) and hence has volume zero. It follows that the body has a zero mass unless the mass density is assumed to be fractal as well. Obviously, the mass density would change with any evolution of \( D_f \) (as assumed in the proposed derivation of MFSL, Appendix A). However, this would be absurd. On the scale of a macroscopic homogenized continuum, there are no fractal holes. Knowing the distribution of microcracks and cavities, the microstructure can in principle be homogenized as a continuum with effective properties depending on the distribution of defects. This again demonstrates that CMFM for case 2 is a meaningless artifice.

4.4.3. Continuum with fractal surfaces of discontinuity (fractal cracks)

There is experimental evidence that fracture surfaces are fractals within some limited range of scales. This is true even for materials with relatively mild inhomogeneity, for which the source of crack fractality is believed to be the interaction of a dynamically propagating crack front with material inhomogeneities (see [56,112,113]). The first question that comes to mind is whether it is acceptable to have a fractal crack, a fractal interface, a fractal surface of discontinuity, etc., in a continuum. In the writers’ opinion, the answer to this question should be yes. Existence of a fractal surface of discontinuity (say, in the deformation gradient) does not contradict any accepted fundamental concept. However, the classical Hadamard jump condition does not apply here, because a unit normal is not defined on a fractal surface. All that needs to be known is that the displacement field is continuous everywhere along the fractal surface.

Fractal roughness of cracks, however, creates enormous mathematical problems. For many years, fracture mechanicians have had great success working with smooth cracks. However, assuming a non-smooth crack makes the elasticity problem virtually intractable.

It may be helpful to classify physical quantities as primary and secondary. The primary quantities are those quantities that can be measured directly. Examples are the force, mass, energy, etc. No matter what the geometry of a continuum is, the force has always the same meaning and the same physical dimension because it can be measured directly. Similarly, the mass is a primary quantity and, in any continuum mechanics formulation, must have the same physical dimension. The secondary quantities are those defined through the primary quantities. For example, the surface traction is a secondary quantity defined in terms of force and area. Secondary quantities can be geometry dependent and could have nonstandard physical dimensions. For example, the stress intensity factor and energy release rate are secondary quantities and can have nonstandard dimensions for a fractal crack.

The first step in formulating a continuum theory of bodies with fractal surfaces of discontinuity is to define traction or a similar quantity on a fractal surface. Suppose that a fractal surface \( \mathcal{S} \) separates a body \( \mathcal{B} \) into two sub-bodies \( \mathcal{B}_1 \) and \( \mathcal{B}_2 \). These sub-bodies exert forces on each other across \( \mathcal{S} \). The force is a primary quantity and always exists no matter how rough and irregular \( \mathcal{S} \) is. Note that, for points in \( \mathcal{B} \setminus \mathcal{S} \), the usual stress tensor is defined and we just need to worry about points that lie on \( \mathcal{S} \) (a measure-zero set).

One may be tempted to define a fractal traction as

\[
\mathbf{t}_f(x, t) = \lim_{\Delta m_{D_f} \to 0} \frac{\Delta \mathbf{F}}{\Delta m_{D_f}}
\]

where \( \Delta m_{D_f} \) is a fractal measure. This and similar definitions, however, are mathematically inconsistent because they ignore the fact that a single fractal dimension is not enough to describe a fractal surface. The above definition does not recover the classical definition in the limit \( D_f \to 2 \) (see [135]). This shows that further geometrical information about \( \mathcal{S} \) is necessary, to be able to define a reasonable ‘fractal traction’. All that we can say is that traction has the dimension of force per unit of a fractal measure. Similarly, the driving force (or configurational force) on a fractal interface must have the dimension of energy per unit of
a fractal measure. Thus it now becomes clear that ‘continuum mechanics of fractal media’ in the sense of [69] is not meaningful.

In [69], some of the concepts of classical continuum mechanics are nevertheless generalized to ‘fractal media’. This generalization, however, tends to be misleading because it goes too far in making several unjustified assumptions which lead to some incorrect conclusions.

A ‘fractal stress’ in [69] is defined as force per unit of a fractal measure in the case of a solid that is a fractal with fractal dimension less than three (a ‘spongy’ solid). While this ‘fractal stress’ is defined only for a uniaxial tension specimen, the same concept is used in a three-dimensional solid without any justification. However, the ‘fractal stress tensor’ is not even defined in [69]. Defining such a density of force might be acceptable in one dimension, but not in three dimensions (see [133,135] for more details). Eq. (5) of [69] defines a fractal traction and it is then implicitly assumed that it exists at any ‘singular point’. One could consider this traction to be some measure supported on the fractal solid (i.e., a measure being zero everywhere else), but again it is not made clear what is meant by a continuum.

In Section 2.3 of [69], a ‘fractal strain’ is defined, but again only in one dimension. The fact that the strain, no matter how defined, must be some measure of the local deformation of the solid is ignored. Later in [69], the three-dimensional ‘fractal strain’ is defined in a peculiar way which will be discussed shortly. A very special deformation for defining the one-dimensional ‘fractal strain’ is considered—namely the deformation of a bar with a fractal distribution of localized deformations.

In view of postulating the ‘fractal stress’ and ‘fractal strain’, Section 4 of [69] then introduces a constitutive equation with fractional derivatives, without offering any sound physical or mathematical grounds. Section 5 all of a sudden defines a ‘fractal strain’ field as a fractional gradient of the displacement field. It is not clear at all why the strain is defined in this strange way and why such a quantity is supposed to describe the local deformation of a solid.

Section 5.2 mentions a stress vector, however, without defining it for a three-dimensional solid. Thus the ‘equilibrium equations’ in Eq. (22) are meaningless, and it is not even clear why these should be regarded as equilibrium equations. To determine the equilibrium equations for any system, the natural approach would be to begin with the integral form of the balance of linear momentum and of angular momentum, and then try to adapt it to fractality. Instead, [69] merely looks at the classical equilibrium equations and, without considering the physical meaning of a fractional divergence, replaces the divergence operator by a fractional divergence. Last but not least, the fractal ‘principle of virtual work’, as used in [69], is simply a replacement of the classical terms of the principle of virtual work by some physically meaningless quantities that have fractional dimensions.

In short, it is not justified to generalize the classical continuum mechanics to ‘fractal media’ by simply replacing the classical differential operators by some fractional operators and stress and strain tensors with some ill-defined ‘fractal’ stress and strain tensors.

5. Conclusions

1. The so-called ‘multifractal scaling law’ (MFSL) is identical to a special case of an energetic–statistical scaling law for failures at crack initiation, which has been derived from fracture mechanics taking into account the finiteness of the thickness of the boundary layer of cracking.

2. Comparing and contrasting the MFSL to Bažant’s original size effect law [9] (as seen in the papers by the proponents of MFSL, e.g. [63,70]), makes no sense and is misleading. The former is applicable only for failures occurring at macrocrack initiation (from a zone of microcracking), while the latter, derived from fracture mechanics, can be applied only for structures with a notch or a large stress-free (fatigued) crack formed prior to the maximum load. These two types of failure necessarily follow different size effect laws.
3. The existing mathematical derivation of MFSL from fractal concepts includes problematic steps which invalidate it. The MFSL does not mathematically follow from the fractal hypotheses made by its proponents (Appendix A).

4. The proposition that a structural size effect is caused by lacunar fractality of the material, particularly in the fracture process zone, leads to Weibull statistical theory of size effect. This theory is applicable only to unnotched structures that have a positive geometry (i.e., fail at macrocrack initiation) and are far larger than the fracture process zone as well as the material inhomogeneities (a situation that does not arise in the quasibrittle range of structural response).

5. The recent proposal [69] for introducing fractal stresses, fractal strains, fractal thermodynamic potentials, etc., is not mathematically consistent and practically usable, and it brings about unnecessary complexity (as explained in Section 4.4).

6. The fractal theories proposed thus far cannot predict the dependence of size effect law parameters on the structure geometry. On the other hand, the energetic and statistical–energetic theories are able to predict this dependence.

7. The fractal concepts could, at most, serve only as a refinement, but not a replacement, of the size effect caused by energy release due to stress redistribution engendered by a large crack or a large fracture process zone. The energetic and statistical sources of size effect are undeniable.

8. The theory of size effect is now ripe for implementation in design codes and engineering practice. The arguments summarized in this paper show that the correct approach is the energetic–statistical theory of size effect.

Acknowledgements

Partial financial support under Grant ONR-N00014-02-1-0622 from the Office of Naval Research to Northwestern University (monitored by Dr. Yapa D.S. Rajapakse) to Northwestern University, and under Grant 0740-357-A466 from the Department of Transportation through Infrastructure Technology Institute of Northwestern University, is gratefully acknowledged. Thanks are also due to G.I. Barenblatt, E. Bouchaud, S. Moreland G. Ravichandran for helpful comments on a draft of this paper. Furthermore, the first author is indebted to F.M. Borodich, and the second author to P.E. Dimotakis and K. Bhattacharya, for valuable discussions.

Appendix A. Problematic steps in mathematical derivation of MFSL

For a mathematical derivation of MFSL, its proponents have generally been citing reference [58], even in the latest works. MFSL was formulated as a smooth transition between the small-size asymptotic size effect, which was taken as \( \sigma_N \propto D^{-1/2} \), and the large-size asymptotic behavior considered as free of size effect, i.e. \( \sigma_N = \text{constant} \). The problem is with the claim that the fractal concept mathematically leads to the small-size asymptotic scaling law \( \sigma_N \propto D^{-1/2} \), in other words, to the slope \(-1/2\) of the left-side asymptote of MFSL in a plot of \( \log \sigma_N \) versus \( \log D \). It has apparently passed unnoticed that the derivation of this scaling law has many serious problems, which will now be summarized.

1. Eqs. (1)–(8) of [58] are used to claim that the nominal (ultimate) strength of a body containing fractal defects of fractal dimension \( 2 - d_r \) must obey the size effect law \( \sigma_u = \sigma_u^* b^{-d_r} \) where \( \sigma_u^* = \text{constant} \), \( b = D/D_1 = \text{dimensionless size ratio of geometrically similar bodies} \) \( (D_1 = \text{reference size}) \), and \( d_r = \text{dimensional decrement from the Euclidean dimension of space} \). The proposed derivation [58] rests on the simplifying assumption that the stress field is a uniformly distributed uniaxial tensile stress.
Further it rests on the implied hypothesis (obviously questionable) that any stress redistribution and energy release that may occur prior to the maximum load can be ignored (in other words, the energetic size effect is ruled out a priori). The exponent \(-1/2\) attributed to the small-size asymptotic scaling law is supposed to be solely a consequence of a peculiar situation called the ‘extreme disorder’. The microstructure is considered to be replete with small planar defects having orientations characterized by spherical coordinates \((\phi, \theta)\) of their normals. The probability density distribution, \(p(a)\), of defects of all sizes \(a\) is characterized in [58] by Eq. (21a) or (34), which may be, for clarity, rewritten as

\[
[p(k^b a_{\text{max}}) \Delta a] n \Delta \Omega = 1, \quad n = \rho(kb)^3, \quad \Delta \Omega = \frac{\Delta \phi \Delta \theta}{2\pi}
\]

where \(a\) = defect size, \(a_{\text{max}}\) = maximum defect size in the body, \(\Delta \Omega\) = element of the normalized surface area \(\Omega\) of a sphere whose points characterize the orientation of defect normal, \(\rho\) = ratio of the number \(n\) of defects in a certain volume to that volume = spatial density of defects (with the dimension of length\(^{-3}\)); \(k\) denotes the dimensionless size ratio, apparently the same quantity that was in an earlier part of [58] denoted differently (namely as \(b\)); and \(b\) = exponent which reduces to 1 for the nonfractal case (Eq. (21a) of [58]). For this equation to be dimensionally correct, it must further be assumed that \(b\) doubtless was intended, from this point on, to represent not the dimensionless size but the actual characteristic size of structure (with the dimension of length). Furthermore, two corrections are required in this equation: (1) the expression given for \(\Delta \Omega\) must be replaced with an element of solid angle, \(\Delta \Omega = \sin \theta \Delta \phi \Delta \theta / 4\pi\); and (2) integrations over \(\phi\) and \(\theta\) on the left-hand side of (A.1) need to be carried out because defects \(a_{\text{max}}\) can have any of all possible orientations, not just one arbitrary orientation \((\phi, \theta)\). But even after these corrections, one must note that the defects of maximum size cannot have the same probability distribution of \(a\) as the ensemble of all defects (as considered in Eqs. (22)–(32) of [58]) but could have only one of the three possible extreme value distributions (Fréchet [83], Gumbel [84,85] or Weibull), of which only the Weibull distribution would be realistic here because a non-negative threshold on \(a\) must exist.

2. Aside from the relation \(\sigma_a = u_b b^{-d_e}\), the subsequent argument tacitly implies further three hypotheses which are questionable. The first implied hypothesis is self-similarity of the distributions of defect sizes in geometrically similar bodies of different sizes (otherwise it would be illegitimate to use Eqs. (21b)–(30) and (35) of [58] to obtain an inverse power distribution for \(p(a)\). As a second implied hypothesis, the maximum size of defects is simply assumed to scale up with the body size \(b\), and to do so as a power function (Eq. (37) of [58]), which is written as \(a_{\text{max}}(k) = k^b a_{\text{max}}(1)\) where \(b = 3/(N + 1)\) and is then used to infer that \(\sigma_a \propto a_{\text{max}}^{-1}\) and \(\sigma_a \propto b^{-d}\). Here \(N\) = constant such that \(1 \leq N < \infty\), and \(k\) (instead of \(b\)) stands for the dimensionless size ratio. As a third implied hypothesis, the maximum defect size \(a_{\text{max}}\) is treated as nonrandom when the scaling is considered, although in reality it should more properly be considered as randomly distributed according to Weibull distribution. These three implied hypotheses seem to be rather arbitrary and unreasonably restrictive, apparently made for convenience of the argument (they seem particularly dubious when the body size exceeds the size of a fully developed FPZ).

3. There is a logical gap in passing from Eqs. (29) to Eq. (31) in [58]. The reason for the sudden appearance of exponent \(N\), allegedly ‘measuring in some way the degree of disorder’, is unclear. So is the sudden incorporation of exponent \(\zeta\) to account for some unspecified ‘secondary effects’ (presumably including the stress redistribution with energy release?).

4. Why the condition called the ‘maximum disorder’ (for which \(N\) is supposed to tend to 2, as argued below Eqs. (39) and (40) in [58]), rather than just some degree of disorder, should occur for a vanishing structure size? This alleged property is unproven, even under the aforementioned assumptions.

In consequence of the aforementioned problems with the derivation, the property that the left-size asymptote of the MFSL in a bi-logarithmic plot should have the slope \(-1/2\) must be considered as unproven by the fractal argument in [58]. It does not logically ensue from fractal concepts. If only the fractal
viewpoint is considered, this property is merely an empirical assumption, which happens to yield an acceptable fit of the test data on the modulus of rupture. On the other hand, from the viewpoint of fracture mechanics, this property, representing a special case of (6), is of course reasonable.

References


Letter to the Editor

Comments on “Is the cause of size effect on structural strength fractal or energetic-statistical?” by Bažant & Yavari [Engng Fract Mech 2005;72:1–31]

Alberto Carpinteri *, Bernardino Chiaia, Pietro Cornetti, Simone Puzzi

Department of Structural Engineering and Geotechnics, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy

Received 1 February 2007; accepted 7 February 2007
Available online 20 February 2007

Abstract

Aim of this letter to the Editor is at replying to the criticisms raised by Bažant and Yavari [Bažant ZP, Yavari A. Is the cause of size effect on structural strength fractal or energetic – statistical? Engng Fract Mech 2005;72:1–31] against the fractal approach to the size-scale effects on the mechanical properties of materials and the concept of the Multi-Fractal Scaling Law presented by Carpinteri [Carpinteri A. Scaling laws and renormalization groups for strength and toughness of disordered materials. Int J Solids Struct 1994;31:291–302]. These criticisms will be analysed thoroughly, showing how they also contain some mistakes and misunderstandings. The presented elucidations should redirect the discussion to a more correct scientific debate.

© 2007 Elsevier Ltd. All rights reserved.

Keywords: Size-scale effects; Fractals; Strength; Fracture energy

1. Introduction

In the paper by Bažant and Yavari [1] several criticisms are raised against the concept of the Multi-Fractal Scaling Law presented by Carpinteri [2]; the Authors question its validity and even argue that it lacks sound physical and mathematical basis. Quoting from their paper (p. 13, item 3): “The ‘MFSL’ was based on a series of hypotheses but does not follow from these hypotheses by a valid mathematical procedure”.

The paper by Bažant and Yavari, as well as two previous papers of the first author [3,4], seriously question the scientific work by Carpinteri and co-workers. Research on the “fractal conjecture” by the group from the Politecnico di Torino, carried out since 1992, has resulted in more than 60 peer-reviewed papers published in the most prestigious journals in the fields of fracture mechanics, mechanics of materials and structural engineering. During these years, Carpinteri and co-workers have constantly presented their results in the most...
important conferences throughout the world, accepting the debate and receiving support from several colleagues.

It is worth noting that experimental tests, as evidenced from the first publications early in 1992 [5,6], have often supported the soundness of the fractal approach and of the Multi-Fractal Scaling Law (MFSL). Indeed, Carpinteri and co-workers have always defined clearly the limits of this law, i.e., that when a strong, energy-driven fracture process is activated, as in the presence of deep notches in the structure, an opposite curvature in the bi-log strength vs. size diagram should be considered. This is exactly the case when Bažant’s original Size Effect Law (SEL) applies. On the contrary, when the role of microstructural disorder and of self-similar features (i.e., fractality) dominates the damage and fracturing processes, the MFSL permits interpolation of the experimental data better than SEL.

On the other hand, Bažant appears to appreciate this, as his strategy in arguing against the “fractal argument” clearly changed during the years. At the beginning, Bažant simply tried to oppose his 1984 formula (SEL [7]) to the MFSL by Carpinteri. His opposition exploited the same tactics he followed to demonstrate that Weibull-type size effect is not applicable to concrete structures.

At the FraMCoS-2 Conference of Zurich in 1995, a report [8] was distributed to the participants showing how the experimental tests in the literature strongly support the MFSL conjecture when large notches are absent. The reaction of Bažant was a fierce opposition to the existence of a finite asymptotic strength for large structural sizes, which was in contradiction with the SEL. In the following years, however, Bažant realised that laboratory tests, and more generally, structural mechanics, could not be forced to fit SEL. Therefore, he introduced the so-called “Universal SEL” [9,10], where an asymptote can be reached for both small and large sizes, and, more importantly, the same MFSL upwards concavity may be obtained under certain values of the parameters.

After the introduction of the Universal Size Effect Law, the scientific community tried to redirect the discussion into a widely accepted framework, through the RILEM Technical Committee “Quasi-brittle fracture scaling and size effect”, chaired by Bažant himself, with the presence of prominent scientists in the field. The work of the Committee was finalised in the publication of its final report [11], wherein the different theories of size effects (including Weibull’s) were described and compared.

In this document, when describing the Multifractal Theory, it is reported that Bažant demonstrated that the MFSL for strength can be obtained as a special case of the Universal SEL. The original SEL (1984) is never quoted in the report, and a non-zero finite term is included in the formula which takes into account a finite strength at large sizes. Although some skepticism about the fractal argument was outlined in the same report by Bažant, Gettu, Jirasek, Planas and Xi, the other members of the Committee did not take a position and, in other papers [12,13], expressed their independent point of view, also showing results in favour of the MFSL.

The criticism by Bažant and co-workers rests on very weak bases. As a matter of fact, the mathematical foundations of the multifractal theory may be considered in the framework of Renormalisation Group Theory. The consistency of the 1/2 hypothesis for the exponents of the MFSLs, originally based on simple statistical arguments, has been also proven by relating the fractal exponents for tensile strength, critical strain and fracture energy [14,15]. Eventually, the so-called fractal mechanics was introduced in more recent years [16–19], based on Fractional Calculus.

Other mistakes by Bažant should be noted. For instance, after more than 10 years, Bažant still seems to confuse lacunar fractals (where stress is defined) with invasive fractals (where energy is dissipated). This situation makes serious and correct scientific debate more difficult.

Finally, we recall that important code-of-practice formulae taking into account size effects (see e.g. the FIB formulas [20] suggested for shear strength in reinforced concrete) agree with the MFSL conjecture.

2. Slope of the MFSL asymptote at the smaller scales

Bažant and Yavari [1] aim at showing that the scaling law for strength at the smaller scales, i.e. \( \sigma_N \propto D^{-1/2} \), cannot be based on the statistical treatment presented in [2]. They insist on this point throughout the whole paper; quoting again from [1]:

\[D = \text{size of the structure}.\]
(i) (p. 13, item 4): “the value $-1/2$ is an unproven conjecture which does not follow from the fractal hypothesis”;

(ii) (p. 26, item 1): “The exponent $-1/2$ attributed to the small-size asymptotic scaling law is supposed to be solely a consequence of a peculiar situation called ‘the extreme disorder’”;

(iii) (p. 26, last paragraph): “the property that the left-size asymptote of the MFSL in a bi-logarithmic plot should have the slope $-1/2$ must be considered as unproven by the fractal argument. [...] If only the fractal viewpoint is considered, this property is merely an empirical assumption”.

In this section, we will review and reject these criticisms against the fractal interpretation of the size effects, by clarifying some aspects that have been misunderstood and confused by Bažant and Yavari, and showing how their discussion contains flaws and mistakes.

This slope, not only follows from the statistical treatment presented in [2], but is also explained in the framework of the Fractal Cohesive Model [14], that has been confirmed by experiments, in our opinion, very convincingly [14,15]. In this framework, indicating with $d_\sigma$, $d_e$ and $d_G$ the non-integer exponents for tensile strength, critical strain and fracture energy, respectively, it has been shown that the following equation can be written

$$d_\sigma + d_e + d_G = 1$$

(1)

At the smaller scales, the collapse is governed by the canonical critical strain $e_c$ and continuum damage mechanics holds. In this case, the damage is diffused (with uniform strain in the bulk) and one obtains $d_e = 0$. Thus, the previous relation becomes $d_\sigma + d_G = 1$. On the other hand, the maximum value for $d_G$ is $1/2$, since it implies a fractal dimension of the dissipation domain $D_G = 2.5$. This would correspond to a Brownian crack surface due to kinematic reasons of crack opening and closing. As a consequence, $d_\sigma = 1/2$ is the limit value at the smaller scales.

In any case, the criticisms of the statistical treatment presented by Carpinteri [2] are unjustified. On p. 26 of their paper, Bazan and Yavari [1] affirm that “defects of maximum size $a_{\text{max}}$ cannot have the same probability distribution of $a$ as the ensemble of all defects, but could have only one of the three possible extreme value distributions (Fréchet, Weibull or Gumbel), of which only the Weibull distribution would be realistic here because a non-negative threshold on $a$ exists”. In other words, they state that $a_{\text{max}}$ cannot have a power-law distribution and, consequently, the assumption $a_{\text{max}}/b = \text{const.}$ (from which the $-1/2$ (LEFM) slope of the size-scale effects follows) should be unjustified.

A first remark concerns the second part of the statement, which is definitely wrong: the limit distribution for a heavy tailed distribution, such as the Pareto, or the Cauchy, is not the Weibull, as erroneously stated by Bažant and Yavari [1], but the Fréchet (this result was already used by Freudenthal [21] almost 40 years ago). Moreover, the existence of “a non-negative threshold on $a$” (presumably upper) is, in any case, merely speculative and unproven.

A second remark concerns the fact, rigorously provable in the framework of Extreme Value Theory, that $a_{\text{max}}/b = \text{const.}$ on average [22]. In other words, although it is true that “defects of maximum size $a_{\text{max}}$ cannot have the same probability distribution of $a$ as the ensemble of all defects”, it can be shown that, starting from a power-law distribution of flaw sizes, the maximum defect size is proportional, on average, to the structural scale.

In addition, this result has been confirmed by Monte Carlo numerical simulations [22]. Therefore, the hypothesis $a_{\text{max}}/b = \text{const.}$ is well justified both theoretically (in two different alternative ways) and numerically.

In conclusion, the fractal-statistical treatment in [2] should be considered valid. The exponent $-1/2$ attributed to the small-size asymptotic scaling law is definitely not an assumption, or “solely a consequence of a peculiar situation called ‘the extreme disorder’”. Rather, it is a consequence of LEFM when a fractal distribution (with power-law tail) describes the flaw size distribution inside the material.

3. Further considerations

(1) The paper by Bažant and Yavari is ill-posed also in its title. In fact, fractals can be deterministic or random, i.e. statistically self-similar [2]. The former are only mathematical models, whereas the latter are the shapes usually met in real systems. Thus, fractals and statistics are not in contrast. Furthermore, an ener-
getic-statistical approach, based on a truncated power-law distribution of the defect size inside a material coupled with LEFM considerations, has been recently proposed in the literature by Carpinteri and co-workers [23–27]. Interestingly, the size effects on tensile strength and fracture energy provided by this model substantially coincide with the ones provided by the multifractal approach. Again, fractal geometry, statistical distributions and energy balances do not contradict each other.

(2) P. 13, item 4: Bazant and Yavari have not caught the relation between the mono-fractal and the multifractal laws. The former should be seen as an approximation of the latter, valid only in a certain size-scale range.

(3) P. 13, item 7: The statement that “the fractality needs to be experimentally observed through about six orders of magnitude for the fractal scaling to be considered a very good model” is without any justification. It is understood that the proportion between smallest and largest aggregates should reflect such a ratio, whereas the fractality – or better the renormalisation group – prevailingly springs from the ratios that the considered property presents at the different observation scales.

(4) P. 18, third paragraph: The results obtained by Yavari in several papers dealing with fractal cracks are summarized, including the results about the size effect. Thus, it seems that some researchers can argue about fractals and size effect (Yavari and co-workers) and some others cannot (Carpinteri and co-workers). Even worse, the “alternative approach” by Yavari is not new, at least for self-similar cracks: Eq. (20) coincides with Eq. (15) in Carpinteri and Chiaia [28].

(5) P. 19, Section 4.3, first and second paragraphs: Bazant and Yavari confuse the ligament fractality (our conjecture) with the crack fractality (a conjecture not taking anywhere).

(6) P. 22, last paragraph, and p. 23, first paragraph: Bazant and Yavari state that a null volume implies a null mass. This is not true since, for example, the archetype of fractal solids, the Menger sponge, has a fractal mass density, so that the mass is finite even if the volume is zero.

(7) P. 23, Section 4.4.3, last paragraph: A definition of fractal stress vector is given although, as stated in the same paper, that definition is not rigorous because the vector normal to an invasive fractal surface is not defined. Nevertheless, that “fractal stress” is used throughout [29]. On the other hand, the fractal stress vector defined in [17] has formally the same expression but is defined on a lacunar fractal set, which presents a unique normal.

(8) P. 24, second and third paragraphs: The fractal generalisation of the stress tensor is immediate, if we consider lacunar fractal infinitesimal areas even in the case of the shearing stresses. Obviously, the same generalisation holds also for the strain tensor, included the shearing strains.

(9) P. 24, fourth paragraph: Bazant and Yavari state that in Carpinteri et al. [17] we used fractional derivatives in the definition of the constitutive equations. This is an incorrect statement, since we simply wrote Eq. (12) to show an example of application of fractional derivatives in mechanics. Furthermore, in Carpinteri et al. [17], there is no need of constitutive equation because that paper deals with the Principle of Virtual Work and, as should be well known, its validity is independent of the material behaviour.

(10) P. 26, second paragraph: Bazant and Yavari erroneously affirm that the exponent β “reduces to 1 for the non-fractal case”. By carefully reading p. 299 in [2] it is clear that β has nothing to do with fractals. This exponent simply characterizes the power-law behaviour of the maximum defect size $a_{\text{max}}$.

(11) P. 26, item 3: Bazant and Yavari state: “...exponent $N$, allegedly ‘measuring in some way the degree of disorder’, is unclear”. On the contrary, the exponent $N$ in Eq. (29) in [2] is an evident measure of disorder: as $N$ increases, the tail of the distribution becomes lighter, and in the limit for $N \rightarrow 1$ the tail disappears, being $P(a_0) = 1$. The maximum defect size is in this case deterministic and equal to $a_0$. On the contrary, as $N$ decreases, the tail becomes larger and the probability of finding defects of size exceeding $a_0$ increases, indicating a more disordered microstructure.

References


Reply

Response to A. Carpinteri, B. Chiaia, P. Cornetti and S. Puzzi’s Comments on “Is the cause of size effect on structural strength fractal or energetic-statistical?”

Zdeněk P. Bažant a,*, Arash Yavari b

a McCormick Institute Professor and W.P. Murphy Professor of Civil Engineering and Materials Science, Northwestern University, 2145 Sheridan Road, CEEIA135, Evanston, IL 60208, United States
b Assistant Professor of Civil Engineering, School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA 30332, United States

Received 1 February 2007; accepted 7 February 2007
Available online 6 July 2007

Abstract

Carpinteri et al.’s discussion is very welcome for it gives an opportunity to clarify long-running disagreements on the problem of size effect, important to several engineering fields. However, the discussion misinterprets many points of Bažant and Yavari’s paper and attempts to raise new issues. This response presents recent experimental results contradicting applicability of Carpinteri’s “multifractal scaling law” (MFSL), and refutes the discussers’ arguments on their proposed concepts of “fractal mechanics”, on the statistical size effect, on the validity of mathematical derivation of MFSL and its asymptotic slope, and on various other aspects of scaling of quasibrittle failure.

© 2007 Elsevier Ltd. All rights reserved.

1. Introduction

A factual discussion of the differences between the energetic or energetic-statistical theory and the fractal theories has been overdue for a long time. Therefore, we sincerely welcome the discussion of our paper by the group of researchers from the Politecnico di Torino. However, since the discussion misinterprets many points from our paper and also attempts to raise many new issues, a detailed response is called for.

2. Response to discussers’ “introduction” and to their comparisons with energetic and energetic-statistical theories of size effect

Carpinteri, Chiaia, Cornetti and Puzzi (henceforth called the discussers) begin by pointing out that the hypothesis of fractal origin of the experimentally observed size effect on structural strength, and the

DOI of original article: 10.1016/j.engfracmech.2007.02.006
Corresponding author. Tel.: +1 847 491 4025; fax: +1 847 491 4011.
E-mail addresses: z-bazant@northwestern.edu (Z.P. Bažant), arash.yavari@ce.gatech.edu (A. Yavari).

0013-7944/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.engfracmech.2007.02.026
“multifractal scaling law” (MFSL) in particular, have been promulgated by the senior discusser (Carpinteri) in “more than 60 peer-reviewed papers published in the most prestigious journals” and “in the most important conferences throughout the world”. This is true. We leave it up to the reader to decide whether it invalidates our critique.

2.1. Experiments showing inapplicability of MSFL to reinforced concrete structures

The discussers write that “the experimental tests, as evidenced from the first publications in 1992, have often supported the soundness of the fractal approach and of the MFSL”. This statement would be accurate only if one considered only older experiments on not too large unreinforced (plain) concrete structures, and only those failing at macro-crack initiation (which requires the structure geometry to be positive [1]). The senior discusser has long claimed that the MFSL applied to all quasibrittle failures of concrete structures except notched test specimens. This claim is not true, and we are compelled to make it clear by giving a synopsis of the main experimental evidence, as available today.

The discussers ignore the difference between the size effects of two types: type 1, which occurs in unnotched structures (of positive geometry) [1,2] at fracture initiation from a boundary layer of cracking in structures, and type 2, which occurs in unnotched or notched structures that reach maximum load only after large stable crack growth. Their fractal geometric arguments do not distinguish between these two cases, and so they assume that the MFSL applies to both. Yet a very different size effect law is required for each [1–3].

The discussers accept that the size effect in notched structures is energetic and different from MFSL. In reality, though, a notch acts similarly to a crack, and is almost perfectly equivalent to a pre-existing (initial) crack when the crack-bridging cohesive stresses in this crack have been reduced to zero by prior load cycling and fatigue, as typical in practice. So, if they accept one case to be energetic, why don’t they accept the other? Both cases lead to size effect laws with exactly the same asymptotes [2,3], and even if the crack-bridging stresses have not been reduced to zero by prior load cycles and fatigue, there is only a small and negligible difference in the sharpness of the transition between the asymptotes (see [4,2,3] for the difference between the type 2 and type 3 size effect laws). Another reason why both cases are essentially equivalent is that, if the crack propagation (in a non-homogeneously stressed structure) is stable, the crack path is almost completely dictated by the laws of mechanics (maximization of entropy production or energy release rate), and does not vary perceptibly because of material randomness, nor fractality of fracturing or damage. So, it is known in advance where approximately the crack will go, which is almost the same situation as for notches.

These facts become clearer by considering the widely studied case of diagonal shear fracture of longitudinally reinforced concrete beams without stirrups. As confirmed by many experiments as well as finite element simulations, the diagonal cracks have essentially similar paths in small and large beams and about the same relative length at maximum load, extending through about 80% of the cross section depth (see the bold curves ending by a circle in Fig. 1). There is no chance that the crack would propagate with a much steeper or much milder average slope for one size than for another, start at very dissimilar positions, or have a very different relative length at the moment of failure (faint curves ending by a cross in Fig. 1). Thus the fracture process zones at maximum loads of small and large geometrically similar beams are found to lie at the same homologous locations (in Fig. 1), as if a precut notch had a tip at that location. The senior discusser has since 1994 insisted repeatedly, in numerous papers, that the MFSL is applicable to all unnotched structures, including the aforementioned case of beam shear. Initially, the experimental evidence and its statistical analysis was not unambiguous in this regard. Broader-range test data for normal concrete, of a size range sufficiently broader than the width of the scatter band, were at that time unavailable, or geometrical scaling of test specimens was not adhered to. This situation, and the excessive scatter of old test data, were the reasons why, in the senior discusser’s widely circulated 1995 report (the discussers’ Ref. [8]), the MFSL seemed to fit various old data equally well as, and in some cases better than, the energetic size effect law (SEL) of type 2 (Eq. (5) in Ref. [5]).

For the shear failure of reinforced concrete beams (with non-reduced aggregate sizes), the situation in 1995 is documented in Fig. 2, which shows all the classical test data on beam shear reported between 1962 and 1994 [6–10]. Note that the closeness of fits of these data by the SEL (energetic size effect law of type 2) and by the MFSL is not very different. Although, in our opinion, the fits by SEL in Fig. 2 are somewhat better, the only clear conclusion that can be drawn exclusively from these data or those compiled in the senior discusser’s 1995
report is that a size effect exists, but not which formula is correct. So, by 1995, one had to rely either on fracture mechanics arguments, or on nonlocal finite element simulation. Although experimental evidence was by 1995 available from reduced-scale model tests of geometrically scaled concrete beams [11] (Fig. 3, left), most practicing engineers (as well as the senior discusser) were unwilling to accept it because the maximum aggregate size \( (d_a = 4.8 \text{ mm}) \) was deemed to be too small.

Now, 12 years later, two test series carried out at the University of Toronto (summarized in [12,13]) make the experimental evidence clearer; see Fig. 3 (middle and right). These large-size tests were approximately geometrically scaled and had a broader size range than those in Fig. 2. They were conducted on beams with normal aggregate size, and reached the beam depth of \( d = 1.89 \text{ m} \).
Fig. 3 shows the optimum fits of the two Toronto test series by the SEL (solid curves) and by the MFSL (dashed curves). Also shown is the optimum fit of the older reduced-scale model tests [11]. Now note that the SEL fits the data points well, while the MFSL does not. In particular, note that the final asymptote of slope $-1/2$, corresponding to the asymptotic power law $d^{-1/2}$, is supported by the data. The final slope of $-1/2$ further implies that the discussers’ intuitive argument about the loss of disorder at a large enough scale is unrealistic (it may be noted that ACI Committee 446, chaired by W. Gerstle, voted unanimously in 1993 that any size effect formula to be considered for revising code provisions for beam shear must terminate with the large-size asymptotic slope of $-1/2$).

From the ACI database [14] (shown in Fig. 1 of our paper [5] and in [12]), a clear statistical trend can be extracted using refined statistical analysis. This database collects 398 beam test results obtained in hundreds of laboratories throughout the world (and is an extension of the 1984 Northwestern University database with 296 test results reported in [15]). A simple statistical regression of all the points in this database, attempted previously by several authors, cannot give a meaningful trend of the size effect on the nominal (or average) beam shear strength $v_c$, because other important influencing parameters, such as the steel ratio $\rho_w$, the relative shear span $a/d$ and the maximum aggregate size $d_a$, vary in that database arbitrarily and highly non-uniformly with respect to beam size $d$.

Such primitive statistical techniques showed that power laws of various exponents, the MFSL, and the energetic size effect law, all fit the database almost equally well (or equally badly) [16]. Therefore, special statistical techniques are needed to extract any meaningful and unbiased information by purely statistical means. This objective necessitates suppressing the bias implied by lack of statistical design of parameter sampling. A suitable statistical technique has recently been introduced in [17] (and summarized in [18]). It does provide a clear size effect trend; see Fig. 4 (top, and bottom left). Here the range of beam depths $d$ in the database is subdivided into five equal-ratio size intervals (Fig. 4). They range from 3 to 6 in., from 6 to 12 in., from 12 to 24 in., from 24 to 48 in., and from 48 to 96 in. (1 in. = 25.4 mm). The borders between the size intervals are chosen to form a geometric (rather than arithmetic) progression because what matters for size effect is the ratio of sizes, not their difference (note that, e.g., from $d = 4$ to $4 + 20$ in., the size effect is strong, from 400 to 400 + 20 in. negligible). The averages and the distributions of the values of $\rho_w$, $a/d$ and $d_a$ in these intervals of the ACI database are very different. Because, as generally agreed, the effect of the required concrete strength $f'_c$ is adequately captured by assuming the shear strength of cross section, $v_c$, to be proportional to $\sqrt{f'_c}$, the ratio $v = \frac{v_c}{\sqrt{f'_c}}$, where $v_c$ and $f'_c$ are both given in psi (1 psi = 6895 Pa) may be considered to depend only on $\rho_w$, $a/d$ and $d_a$.

To filter out the effect of influencing parameters other than $d$, one must, within each interval of $d$, gradually (step by step) restrict the relevant values or influencing parameters $\rho_w$, $a/d$ and $d_a$ by adjusting the upper and lower limits until the averages of these relevant values within each interval of $d$ would become, for each interval of $d$, about the same (within a given tolerance). To ensure statistically unbiased treatment, no data point...
within the upper and lower limits of each interval of $d$ for each influencing parameter may be left out, and no data point outside those limits may be included. A computer optimization algorithm has been written for extracting such relevant data from the ACI database. The algorithm has been run to extract three subsets of the ACI database, corresponding to three different average values, $q_w/C_{25} = 1.5\%, 2.5\%$ and $0.9\%$.

The centroids of the extracted relevant data points within each of the aforementioned 5 intervals of $d$ are shown as the bold diamond points in Fig. 4 (top left and right, and bottom left) in the plot of $\frac{v_c}{\sqrt{f_c}}$ versus $\log(d)$. For $q_w = 1.5\%$ and $2.5\%$, the last of the five intervals of $d$ (which is the case of very large beams) had to be left empty because the ACI database does not contain within that interval sufficient data points giving these average values of $q_w$. The subsets of extracted relevant data points are shown as the faint empty circles. For $q_w = 1.5\%$, the algorithm delivered a subset of 128 relevant data points for which the precise values of the averages within each interval of $d$ were $q_w = 1.51\%, 1.50\%, 1.50\%, 1.50\%$, with the corresponding averages $a/d = 3.45, 3.33, 3.33, 3.23$, and $d_a = 0.66, 0.67, 0.66, 0.65$ in. For $q_w = 2.5\%$, the algorithm delivered a subset of 157 relevant data points for which the precise values of the averages within each interval of $d$ were $q_w = 2.55\%, 2.51\%, 2.48\%, 2.44\%$, with the corresponding averages $a/d = 3.33, 3.33, 3.34, 3.33$, and $d_a = 0.67, 0.67, 0.66, 0.67$ in. For $q_w = 0.9\%$, the algorithm delivered a subset of only 24 relevant data points for which the precise values of the averages within each interval of $d$ were $q_w = 0.91\%, 0.94\%, 0.94\%, 0.91\%, 0.74\%$, with the corresponding averages $a/d = 2.94, 2.94, 2.94, 2.94, 2.86$, and $d_a = 0.39, 0.39, 0.39, 0.39, 0.39$ in.

Under the assumption that the statistical weight of each size interval centroid in Fig. 4 is the same, the foregoing procedure [17] is now used to obtain the optimum least-square fit of these 4 or 5 centroids with the SEL (type 2 energetic size effect law, Eq. (5) in [5]), which is written here as $v_c/\sqrt{f_c} = C(1 + d/d_0)^{-1/2}$ where $C$, $d_0 = \text{free constants to be found by the fitting algorithm}$ (note that Eq. 5 of our paper [5] has a misprint: the exponent should be $-1/2$, not $1/2$).

Fig. 4. Top left: data extracted from the ACI beam shear database obtained by a statistically unbiased algorithm that restricts the ranges of the steel ratio, shear span and maximum aggregate size so that their mean values be about the same in each interval shown; and fit of the interval centroids of these data (diamond points) by the size effect law for failure after large stable crack growth. Top right and bottom left: the same but for higher and lower steel ratios. Bottom right: optimum fit (in transformed coordinates) of the entire ACI database on the shear strength of longitudinally reinforced three-point loaded concrete beams by a formula based on size effect law of type 2 [12] (the dashed curves represent the optimum fit by the MFSL, and the circle size is proportional to weight in regression).
The fits of the diamond points (solid curves in Fig. 4 top left and right, and bottom left) are seen to be quite satisfactory; they have very small coefficients of variation of errors, \( \alpha = 2.5\%, 1.7\% \) and \( 5.1\% \), respectively (standard deviation of errors divided by data mean). The trends of the diamond points are seen to be quite systematic. Note all the three trends show a negative curvature, agreeing with the SEL and contradicting the MFSL. The terminal trends of the centroids agree with the asymptote of slope \(-1/2\), characterizing the SEL \([15,16,12,19]\), and give no hint of an approach to the asymptote of slope 0, characterizing the MFSL. Finally note that the trends of the centroids disagree also with the Weibull statistical theory, which would require a straight line of slope cca \(-1/12\).

Another way to obtain from the database a meaningful result is to conduct a multivariate nonlinear regression with a rational size effect formula in which the strong influences of \( q_w \), \( d/a \) and \( d_a \) on the parameters are taken realistically into account. These influences have been incorporated into the coefficients of SEL (see the formula in Fig. 4 bottom right). This led to the optimum fit of the entire ACI database of 398 points shown in Fig. 4 (right) \([12]\), plotted in transformed coordinates of the formula. Despite scatter, the trend of the energetic size effect is clearly confirmed. The MFSL cannot be plotted against the entire ACI database because the influences of \( \rho_w \), \( d/a \) and \( d_a \) on its coefficient are not known and are impossible to determine by the discussers’ fractal mechanics’. Nevertheless, the positive curvature of the MFSL disagrees with the trend in Fig. 4 (right).

Since the MFSL cannot apply to reinforced concrete beams failing after large crack growth, it was misleading, in numerous papers of the senior discusser, to present the MFSL and the SEL as two competing models. The former applies only to type 1 size effect (at small sizes), and the latter only to type 2 size effect. Models that apply to different situations cannot be in competition.

The discussers are silent about two other inconsistencies. When applying the MFSL to type 2 failures, they ignore the rate of energy release caused by a large crack growing stably before the maximum load. Yet the fact that this energy release causes a size effect is undeniable and easily understood (see the discussion about Fig. 3 (right) in [5], and point 1 on page 564 of [20]). Even if the fractal source of size effect for type 2 failures were accepted, the size effect of the energy release would have to be superposed on it because it is inevitable. Likewise, in applying the MFSL to type 1 failures, the discussers ignore both the size effect due to stress redistribution engendered by a finite fracture process zone (or boundary layer of cracking), and the Weibull statistical size effect, which is undeniable in the case of failures at crack initiation in a material of random strength.

The discussers are further silent about the fact that the dependence of MFSL coefficients on structural geometry is not predicted by their fractal mechanics. By contrast, the energetic size effect law (of type 1 as well as 2) for failures at crack initiation does predict this dependence, through the limit values of the energy release rate derivatives (Eq. (6) in [5]).

Referring to the RILEM report [20], the discussers are not right in criticizing that “a non-zero term is included in the formula” of the SEL. As a matter of fact, the original SEL [19] is cited in [20] as Ref. [53] and is presented in Eq. (11). It is true that “a non-zero term” \( \sigma_R \) (the need for which was pointed out in 1987 [21]) is added, but this is done only for the sake of brevity, and a few lines below it is stated that the residual strength \( \sigma_R \) is usually zero, with two exceptions – unbroken fibers crossing the crack, or transition at large sizes to a residual frictional plastic mechanism as in compression kink bands or the Brazilian test [22]. We can see nothing that could be criticized.

2.2. Discussers’ comments on statistical size effect and Weibull theory

The discussers write that Z.P. Bažant (ZPB) “exploited the same tactics he followed to demonstrate that Weibull-type size effect is not applicable to concrete structures”. This is a distortion of our position, which must be explained. We certainly do not deny that the Weibull-type statistical size effect on the mean structural strength exists in concrete structures, though not without important limitations. It exists only in those structures that fail (under load control) at the moment of macro-crack initiation and are sufficiently large compared to the aggregate size. Very large unreinforced concrete structures (such as large arch dams or retaining walls failing by flexure) qualify, but reinforced concrete structures failing after large crack growth generally do not. Neither do small plain concrete beams. What needs to be understood is that (aside from other reasons [23,2]) there are four strong reasons for these limitations:
(1) Large cracks grow stably prior to the maximum load, traversing typically 80% of the cross section in the case of beam shear. Since, as already pointed out, the crack path is dictated by mechanics, the fracture process zone at maximum load lies at one precise location, and so a random local strength at other locations in the beam does not matter. In such structures, strength randomness affects only the shape of the probability density function (pdf) of structural strength but not the mean, and causes no appreciable size effect on the mean strength [23]. Here the size effect, seen in experiments, is energetic (non-statistical), caused by the energy release due to stress redistribution engendered by large crack growth.

(2) The contribution to Weibull probability integral is proportional to $r = \left( \frac{\sigma(x)}{\sigma_{\text{max}}} \right)^m$ where $\sigma(x) =$ first principal stress at point of coordinate $x$, $\sigma_{\text{max}} =$ maximum of principal stresses within the structure, and $m =$ Weibull modulus $\approx 24$ for concrete [24]. Because of stress concentration near the tip of a large crack, the stresses decay rapidly and at points where $\sigma(x) \leq 0.9\sigma_{\text{max}}$, one has $r \leq 0.08$, and where $\sigma(x) \leq 0.5\sigma_{\text{max}}$, one has $r \leq 6 \times 10^{-8}$. So, virtually the only non-negligible contribution to the Weibull probability integral for structural strength comes from the fracture process zone, whose size is essentially independent of the structure size, and thus can cause no size effect [23]. Besides, the Weibull probability integral could diverge if the singular elastic stress field of a crack were substituted. Hence, there is no appreciable statistical size effect if a large crack forms before the maximum load.

(3) The third reason, which precludes the use of classical Weibull statistics to both unreinforced and reinforced concrete structures of normal sizes (as well as to most structural parts made of laminates or coarse-grained ceramics), is that the equivalent number [25,26] of representative volume elements (RVE) contained in a normal concrete structure is not large enough (because of insufficient ratio of cross section size to aggregate size, as well as stress non-uniformity).

(4) To obtain acceptable (in fact, barely acceptable) data fits, the classical Weibull theory is often considered to have a non-zero threshold. But recently it was shown [25,26] that the threshold in Weibull theory must always be zero, or else the Maxwell–Boltzmann distribution of interatomic bond strength would be contradicted.

While the discussers object to certain ‘tactics’ against the Weibull statistical theory, they themselves actually never use that theory. They argue that some sort of strength randomness is in some way communicated through fractals. But the way it is supposed to be communicated is to us mathematically incomprehensible.

Since the discussers emphasize the statistical aspect of the material, it is strange that they accept a horizontal large-size asymptote of size effect, as exhibited by MFSL. According to the Weibull statistical theory (and also the chain-of-RVEs model [25,26]), the slope of that asymptote would have to be $-n_d/m$, which is about $-1/12$ for two-dimensional scaling of failure of unreinforced concrete structure ($m =$ Weibull modulus $\approx 24$ for concrete, and $n_d =$ number of dimensions of fracture scaling).

2.3. Why do the discussers compare MFSL to the universal SEL?

In [5], the universal size effect law (U-SEL) was mentioned only to illustrate a unified description of size effect clarifying the transition between the energetic size effects of type 1 (failures at crack initiation) and type 2 (failures after large crack growth). It has not been mentioned in regard to MFSL because it is irrelevant to the comparison of MFSL with the energetic size effect. Yet the discussers write: “The reaction of Bažant was a fierce opposition . . . Therefore he introduced the so-called Universal SEL”. The word “therefore” is misplaced (because the U-SEL was shown on the screen of an opening lecture at that same conference at which the discussers’ Ref. [8] was distributed).

More importantly, none of us has ever mentioned the U-SEL in relation to the MFSL, nor to the fractal theories of the senior discusser. The repeated mentions of this law by the discussers in regard to MFSL only cloud the issue and evade the real problem.

In all the discussers’ statements, the words “Universal SEL” should be replaced by the words “the energetic size effect law for failures at crack initiation”, or simply the words “type 1 size effect law”. Only then the discussers’ commentary would make sense (although it would still be unjustified). The MFSL should be compared only to the type 1 size effect law because it is usable (as an empirical formula for a limited size
range) only for small enough structures failing at crack initiation, and not when small or large macro-cracks
grow stably before the maximum load.

2.4. Discussers’ other points

The discussers appeal to the well-known renormalization group theory as a foundation of their ‘multifractal’ theory. That is unwarranted. The renormalization group theory does not sanction the ‘multifractal theory’. That theory merely describes the transition from one power law scaling to another power law scaling, which is a characteristic not only of the MFSL but also of the energetic and energetic-statistical size effect laws (thus, a foundation in the renormalization group theory could just as well be invoked for the latter). What is crucial is the gradual transition between these two power-law scalings, which spreads over several orders of magnitude depending on strain localization instabilities. On that, the renormalization group theory says nothing.

For the sake of accuracy, various marginal comments of the discussers need to be corrected. E.g., the discussers write “Although some scepticism . . . was outlined . . . by Bazˇant, Gettu, Jirásek, Planas and Xi, the other members of the Committee did not take a position and, in other papers, expressed their independent point of view . . . in favour of the MFSL” (actually, not ‘members’, but only one member, van Mier, who expressed that point of view, in a rather non-specific way).

3. Response to “slope of the MFSL asymptote”

3.1. The concept of “fractal mechanics”

The discussers’ concept of the so-called “fractal mechanics” is interesting but not well defined. As explained in [5], it involves some simplistic extensions of linear elasticity. The arguments presented in [27], and also those offered by the discussers, conflict with some universal principles of mechanics. Note the following points, to wit.

In an axially loaded bar, the strain is a normalized displacement, i.e., the ratio of the total relative displacement to the original length of the bar. But this can have nothing to do with the cross section of the bar, as implied in [27,28]. Thus the one-dimensional definition of strain proposed in [28, Eq. (3)] is not rigorously justified. In the prototypal one-dimensional problem of an axially loaded bar, one can define a “fractal normal stress”, similar to what was defined for a cohesive theory in [29], but such a definition is limited to forces in a fixed direction. If it is desired to introduce a normal stress definition as the density of force on a fractal cross section, what should be modified is the stress–strain relation, and not the strain. In that sense, the constitutive equations become scale dependent, but not the strain.

The recent works of the senior discusser’s team go even further and, without any sensible justification, define the strain as the fractional derivative of the displacement field. What they present as a “formal derivation” [28, see the paragraph after Eq. (18)] looks to us as a collection of various undefined quantities with undefined connections.

In this regard, note that linear elasticity can be rigorously derived from nonlinear elasticity by linearization about a given deformation mapping [30]. The linearized strain represents the linearization of deformation gradient and thus happens to depend on the first derivatives of the displacement field. The deformation gradient (which is the derivative map of the deformation mapping) maps an infinitesimal line element in the reference configuration to its deformed form in the current configuration.

Now it is unclear what a fractional derivative would mean in this fundamental context. Defining a quantity as the fractional derivative of the displacement field is meaningless. The possibility of some new measures of strain is not excluded, but any such measure would have to be based on sound geometric arguments.

Another surprising aspect in the so-called “fractal mechanics” [27] is the differential equation of equilibrium, which is expressed in terms of some fractional divergence of an undefined stress tensor. In classical continuum mechanics, the local balance of linear momentum is obtained by localization of the global balance of linear momentum away from discontinuities. This yields the Cauchy theorem and the local balance of linear momentum (or equilibrium equations), i.e. \( \text{div}\sigma + \rho b = \rho a \) (where \( \sigma = \text{stress tensor}, b = \text{body force}, a = \))
acceleration, $\rho = $ mass density). However, the discussers’ definition of a “fractal stress tensor” is unclear, and so is their replacement of the divergence operator with a fractional operator. It is also unclear how the stress could be defined on a lacunar fractal if the energy is supposed to dissipate on an invasive fractal.

3.2. Problems with MFSL asymptotic slope $-1/2$

Trying to defend the small-size asymptotic slope of MFSL, the discussers say nothing about one simple but strong objection to its value $-1/2$. Since the fractal dimension $\delta$ of the cracking morphology is not constant but varies from one material to another, how can the exponent $n = -1/2$ of the small-size power-law asymptote of MFSL be treated as independent $\delta$?

Consider that $\delta$ varies and approaches the Euclidean dimension $\delta_{Eu}$. In the discussers’ view and in the derivation of MFSL ([31] or discussers’ Ref. [2]), exponent $n$ remains equal to $-1/2$ even when, for example, $\delta = \delta_{Eu} + 10^{-9}$ (where $\delta_{Eu}$ = Euclidean dimension). Then $n$ is supposed to jump discontinuously to 0 as $\delta = \delta_{Eu}$.

In a credible theory, $n$ would have to approach 0 continuously, i.e., $\lim_{\delta \to \delta_{Eu}} n = 0$ should hold. The discontinuity of $n$ as a function of $\delta$ is physically unacceptable. Thus the arguments in [31] that led to asserting that $d_a = 1/2$ are hazy and, to us, mathematically incomprehensible.

In regard to their Eq. (1), the discussers claim that “at the smaller scales . . . continuum mechanics holds”, and “damage is diffused and one obtains $d_a = 0$”. This cannot be true. At the smaller scale, the discreteness of the microstructure of the material, for instance concrete, cannot be ignored, and thus the continuum mechanics approximation of concrete as such cannot hold.

Furthermore, the discussers’ claim that “the maximum value for $d_a$ is 1/2” looks to us as nothing but a questionable conjecture. In fact, the variation of the critical energy release rate $G$ in the sense of an $R$-curve is not a physical fact but merely an artifice allowing approximations by linear elastic fracture mechanics. According to the cohesive crack model, which is a physically more realistic approach, the critical $G$, i.e., the fracture energy, is a constant, yet the $R$-curve (representing the variation of the critical $J$-integral near the crack tip) can be predicted by this model (and so can the variation of this integral when the crack front is close enough to the boundary to interact).

The discussers’ response to our statement in [5] that the “defects of maximum size cannot have the same probability distribution of $a$ as the ensemble of all defects (as considered in Eqs. (22)–(32) of [31]) but could have only one of the three possible extreme value distributions (Fréchet, Gumbel or Weibull)” sidetracks the issue that we raised. The Fréchet distribution would, of course, be acceptable (provided that the defect size distribution had a Pareto tail), but that is not what we criticized.

What we criticized as incorrect is that the derivation of the MFSL in [31] considered, for the maxima, the same distribution as for the whole ensemble of defects, and particularly not the Fréchet distribution (note that a consistent statistical theory of crack propagation using the Fréchet extreme value distribution was presented in Section 12.6 of [1]).

Stochastic simulations with the nonlocal Weibull statistical theory or with the probabilistic nonlocal damage mechanics [24,32, e.g.] demonstrate that the small-size asymptotic power law of size effect of type 1 can have an exponent different from $-1/2$. Since no analytical solutions of boundary value problems with the fractal theory have been presented by the discussers, it would be welcome to see at least some numerical solution of the fractal field equations of some boundary value problem, and its comparisons with experiments. We expect that such solutions would show that the small-size asymptotic exponent of size effect is not restricted to $-1/2$, even under the hypotheses of their ‘fractal theory’.

A fundamental model for the statistical aspect of the size effect at crack initiation (type 1) in a heterogeneous material is the chain-of-RVEs model [25,26,33]. This is a weakest-link model that, in contrast to Weibull theory, has a finite, rather than infinite, number of links, each of which is imagined to correspond to one RVE. Based on the Maxwell–Boltzmann statistics of atomic energies, each RVE of a heterogeneous brittle material must have a Gaussian strength distribution onto which a power-law tail with a zero threshold is necessarily grafted at the probability of the order of 0.001. For large sizes, this model asymptotically reduces to the classical Weibull theory, while for small sizes it yields the correct (experimentally and computationally confirmed) deviation from the Weibull power-law size effect (whose exponent is $-n_o/m$). This is the same as predicted by
the energetic analysis of size effect (the reason is that the RVE is considered to have a finite size and the equivalent number of RVEs is finite). The predicted small-size asymptotic slope of size effect (plotted in a logarithmic scale) is, in general, different from $-1/2$ [25,26].

The MFSL, with its initial asymptotic slope fixed as $-1/2$, and the chain-of-RVEs model (whose mean coincides with the energetic-statistical size effect law and also with the mean of nonlocal Weibull theory), are mutually exclusive. They cannot both be correct. Eventually, the science and engineering community will have to choose.

The discussers try to justify the small-size asymptotic slope of $-1/2$ by claiming that “the maximum defect size is proportional, on the average, to the structural scale”. It is not clear whether the discussers consider the defects to be the initial microscopic defects, i.e., microcracks (smaller than about the maximum inhomogeneity sizes), or the macrocracks produced by load. There are two points to consider in this regard.

First consider the initial microcracks, or flaws. The distribution of their sizes is supposed to be an objective material property, not alterable by human will once the concrete is cast. So how can it depend on the structural scale, e.g., the depth of a reinforced concrete beam, which is a subjective property chosen by the engineer at will?

On the other hand, if the maximum ‘defect’ is considered to be the macro-crack produced by loading, the size of which at maximum load does depend on the beam depth, then there is a different problem. Wouldn’t macrocrack formation cause stress redistribution with energy release? Wouldn’t that energy release inevitably cause an energetic, rather than fractal, size effect? Wouldn’t that size effect have to be taken into account? Isn’t it true that the energy release increases with structure size roughly quadratically, while the energy dissipated by the crack increases approximately linearly? [2, e.g.].

Alas, the argumentation relative to the maximum defect size is not mathematically comprehensible to us.

3.3. Criticisms of MFSL derivation whose refutation was not attempted

Aside from the discussers’ points rebutted later in Section 4, the interested reader should note that the authors made no attempt to refute the following problematic and mathematically invalid steps in the ‘derivation’ of MFSL, as identified in [5]: seven of the nine points mentioned in item 1 on page 26 of [5], all the three points in item 2, and all of item 4.

So it must be concluded that the MFSL does not logically follow from the hypothesis of fracture or damage fractality. It is merely an empirical formula which is good enough only for the type 1 size effect (at crack initiation), and only for sizes not so large that Weibull statistical size effect would intervene. In that range, the MFSL is equivalent to a special case of the energetic and energetic-statistical size effects. For very small sizes, though, the MFSL has a questionable asymptote, and for very large sizes it fails to capture the inclined asymptote of the power-law size effect of Weibull statistical theory, which must occur for failures at crack initiation.

Although MFSL has the same asymptotes as the earlier CEB-FIP formula [34] \( A + \sqrt{B/D} \), it certainly cannot be claimed that it provides a theoretical foundation for that formula, which was introduced purely empirically.

4. Responses to points made in “further considerations”

(1) We find this point to be mere playing with words. Careful reading of [5] makes it clear what we meant by ‘fractals’ and by ‘energetic-statistical’.
(2) Mathematically, a multi-fractal is not what the discussers call ‘multi-fractal’, as explained in [5].
(3) The requirement for “six orders of magnitude” of scales [35] (see also item 8 on page 564 of [20]) is necessary to ensure that the scaling property be unambiguously fractal, i.e., that it cannot be described equally well by some non-fractal theory (e.g., autocorrelated statistical roughness). The fractal scaling property must be verified over a sufficiently broad range of scales, since not everything with apparently fractal scaling in a narrow range of scales is properly modelled as a fractal. Of course, the two limits of this interval depend on the microstructure scale and on the macroscopic characteristic length of the problem.
(4) Our paper [5] certainly did not exclude the possibility of modelling certain aspects of fracture mechanics by fractals. What [5] warns about is the danger of introducing some ill-defined fractal-motivated concepts into mechanics. Furthermore, the discussers pointed out the similarity of Eq. 15 in an earlier Carpinteri’s paper to Eq. (10) in [29], which describes the asymptotic stress distribution around a fractal crack (they actually write “Eq. (20)” but the number “20” was most likely a misprint). This comment, however, is surprising since Eq. (10) was not presented in [29] as original. In all the papers by Yavari and co-workers on this subject [29,36–40], proper reference has been made to those who studied this problem for the first time, i.e. Mosolov [41], and Gol’dstein and Mosolov [42,43]. However, none of these original sources has been cited in the long series of papers by the senior discusser. In this context, we wish to emphasize that, aside from Mosolov and Gol’dstein’s pioneering contributions, we see, despite our critiques, valuable contributions in the early and recent works on fractal aspects of fracture by many other researchers (see, e.g. [44–53] and other works cited in [5]).

(5) There is no such confusion in [5]. Careful reading reveals that crack fractality and the so-called ligament fractality were properly distinguished. Besides, the related sentence “after more than ten years, Bazant still seems to confuse lacunar fractals with invasive fractals” (in the penultimate paragraph of the discussers’ Section 2) is baseless.

(6) In principle, one can work with an entity having a zero volume and a finite fractal measure, and associate with it a fractal mass density. Our argument in [5] does not exclude this possibility. The point is the following. Would it be meaningful to work with a body of zero volume within the setting of classical continuum mechanics? The answer is no, if one wants to use the established continuum concepts. For example, if a body has the fractal dimension of 2.55 (locally, in the neighborhood of a point) it makes no sense to think of the “stress” and “strain”, or even mass density, as a field in the classical sense, since these quantities cannot be defined everywhere in the ambient space.

(7) This comment is surprising because, in subsection 4.4.3 of [5], we give no definition of stress. We explicitly say that “One may be tempted to define a fractal traction...” but we do not use the equation following this statement, and we explain in detail why one must be very careful in defining the stress on fractal surfaces, and why the concepts of the so-called “fractal mechanics” are unjustified and should be avoided. Besides, [29] discussed similar problems and again made no use of any “fractal stress tensor”. In that work, it was explained that, in one-dimensional problems, one can define fractal densities of force although their naive extension to three-dimensional problems is impossible.

(8) Any definition of stress and strain should be based on sound mathematical arguments, but those used by the senior discusser are not of such a kind. For a given system, one should start with a proper description of kinematics and define appropriate strain measures. This is not done in the so-called “fractal mechanics” in [27]. The existence of Cauchy stress is a consequence of balance of linear momentum and explicitly depends on smoothness of the surfaces on which traction acts. In any well-defined “fractal mechanics”, one needs to address this issue and define a proper measure of stress. Again, this is not done in [27].

(9) Eq. (12) in [27] is a constitutive equation. This equation, whatever one may want to call it, uses a meaningless quantity, and this was the point made in [5]. Fractal derivatives have been applied in physics – for example, in describing anomalous diffusion, in characterizing viscoelasticity, and also in describing dissipation in the framework of Lagrangian mechanics. We certainly did not imply in [5] that fractional derivatives could not be useful. Our point simply is that one cannot simply take a fractional derivative of a displacement field and call it a “fractal strain”.

(10) This statement of ours is attacked outside its context. Exponent $\beta$, of course, does not characterize fractality, but if $\beta$ would not reduce to 1 then standard continuum mechanics could not be the limit case of Carpinteri et al.’s “fractal mechanics”, which is a fundamental requirement.

(11) The point is not whether exponent $N$ is used in [31] as some sort of a measure of disorder, but whether it is realistic to use it for that purpose. The author provides no justification, and the reason for its introduction in the MFSL derivation in [31] is indeed unclear.

In conclusion, it must be reasserted that “the ‘MFSL’ has been based on a series of hypotheses but does not follow from these hypotheses by a valid mathematical procedure”.

4.1. Closing comments

The tone of the discussion and the personal emotive comments it contains suggest that our attempt at a factual criticism has offended the feelings of the discussers. If that is so, we are truly sorry. It was definitely not our intention to cause that. What is needed is an unemotional scientific debate of the salient differences between competing viewpoints.

Although the discussion and our response are focussed on concrete structures, further debate of size effects will be important for all the fields of engineering affected by quasibrittle (or cohesive) fracture. This for example includes the design of large load-bearing composite parts for aircraft and ship, predictions of the load capacity of sea ice and the forces it exerts on obstacles, estimates of the danger of landslides, snow avalanches, and rock burst in mines, safety of nuclear containments and waste storage, and reliability of micro-electronic components and nano-devices.

Introducing the size effect into design codes for concrete structures, which affect how many thousands of concrete structures are built, is a necessity. One reason why progress has been delayed for 15 years is the unresolved scientific conflict between the energetic-statistical and fractal theories of size effect. Reaching a consensus on the theory is essential since it is costly to test a very large structure to destruction and outright prohibitive to conduct a statistically significant number of such tests. Thus it is not surprising that, in the case of shear of reinforced concrete beams [12,16,18,54], 86% of all available test data pertain to beam depths less than 0.5 m, and 99% to depths less than 1.1 m, while, by contrast, the world-record box girder of the Babeldaoob-Koror Bridge in Palau, whose fatal collapse in 1996 was marked by a large inclined shear-compression fracture emanating from the support, was 14.2 m deep. In extrapolations of the laboratory test data to such a structure size, the SEL and the MFSL curves in Figs. 2–4 differ by factors between 2.4 and 3.7, the MFSL being on the unconservative side. Clearly, the present debate has grave engineering consequences [55].

That a serious problem exists is also clear from past experience with catastrophic collapses of large concrete structures involving fracture. According to the statistics reported in [56,57], the frequency of collapses of very large structures has been more than 1 in a thousand (per lifetime), while for normal size concrete structures it has been about one in a million, which is what is generally required [58,56] to ensure that concrete structures (or aircrafts, ships, nuclear plants, etc.) would not add significantly to other hazards that people inevitably face. One in a thousand is intolerable, and a way to move forward in the ongoing polemic must be found.

Acknowledgements

Financial support from the US National Science Foundation through Grant CMS-0556323 to Northwestern University, and from the US Department of Transportation through Northwestern University Infrastructure Technology Institute Grant 0740-357-A210, is gratefully acknowledged.

References