PNAS RESEARCH ARTICLE | ENGINEERING

Sprain energy consequences for damage localization and fracture mechanics

Houlin Xu[a], Anh Tay Nguyen[b], and Zdeněk P. Bažant[a,b,c,1]

Affiliations are included on p. 8.

Contributed by Zdeněk P. Bažant; received May 28, 2024; accepted August 22, 2024; reviewed by Christian Hellmich and Franz-Josef Ulm

The 2023 smooth Lagrangian Crack-Band Model (slCBM), inspired by the 2020 invention of the gap test, prevented spurious damage localization during fracture growth by introducing the second gradient of the displacement field vector, named the "sprain," as the localization limiter. The key idea was that, in the finite element implementation, the displacement vector and its gradient should be treated as independent fields with the lowest (C_0) continuity, constrained by a second-order Lagrange multiplier tensor. Coupled with a realistic constitutive law for triaxial softening damage, such as microplane model M7, the known limitations of the classical Crack Band Model were eliminated. Here, we show that the slCBM closely reproduces the size effect revealed by the gap test at various crack-parallel stresses. To describe it, we present an approximate corrective formula, although a strong loading-path dependence limits its applicability. Except for the rare case of zero crack-parallel stresses, the fracture predictions of the line crack models (linear elastic fracture mechanics, phase-field, extended finite element method (XFEM), cohesive crack models) can be as much as 100% in error. We argue that the localization limiter concept must be extended by including the resistance to material rotation gradients. We also show that, without this resistance, the existing strain-gradient damage theories may predict a wrong fracture pattern and have, for Mode II and III fractures, a load capacity error as much as 55%. Finally, we argue that the crack-parallel stress effect must occur in all materials, ranging from concrete to atomistically sharp cracks in crystals.

fracture energy | fracture and damage mechanics | second displacement gradient | material rotation gradient | spress–sprain relation

Significance

Commonplace in practice, computational fracture predictions are often erroneous (e.g., for concrete, fiber-composites, geotechnics, earthquake, biomaterials). As recently revealed by the gap test, the problem is that the fracture front is never a point, but exhibits damage over a finite width. This causes the energy required for fracture growth to vary as a function of the previously neglected crack-parallel stresses and gradients of local material rotations. Validated here is a theory in which spurious localization of crack front damage into a point is prevented by means of the second partial derivatives (or a Hessian) of the displacement vector field. The classical strain-gradient theory, which ignores material rotation gradients, is shown to give wrong predictions for shear and shear-compression fractures.

All line crack models, which include the linear elastic fracture mechanics (LEFM), the XFEM, and Barenblatt's cohesive crack model (1–3) [with its reincarnation as Hillerborg's model (4)] cannot reproduce the effect of crack-parallel stresses. The reason why this effect went undetected for a century until the invention of the gap test (5, 6) is that in all of the standardized fracture test specimens, the crack-parallel stresses are negligible ($\bar{\sigma}_{xx} \approx 0$). The recent fads of peridynamics and phase-field model also fail to reproduce the gap test as well as 10 other recently emphasized distinctive fracture tests (7, 8). In addition, all the integral-type nonlocal models and peridynamics (9) have unsurmountable problems with the boundary conditions, which are fatal at the crack faces.

The classical crack band model, CBM, proposed in 1983 (10, 11), has been shown to surpass all of the aforementioned models by far (7). This model uses the finite element (FE) size to represent the width of the crack band, which is set equal to the material characteristic length l_0 best identified from size effect tests. Within each FE, the damage is assumed to be uniform and must be described by a realistic damage law. However, the CBM is not perfect because the variation of the width of crack band front along with the damage distribution across the band cannot be captured and, if a regular mesh is used, a crack band inclined to the mesh lines propagates in a zig-zag way, locally biased by the mesh orientation (even though the overall crack path, dictated by maximization of the energy release, remains essentially correct).

The aforementioned shortcomings are remedied by the new smooth Lagrangian Crack-Band Model, slCBM (12). In slCBM, excessive damage localization is limited by the so-called sprain energy density, Φ, which is defined in ref. 12 as a function of the third-order tensor, η_{ijk}, of the second gradient (or the Hessian) of the vectorial displacement field. The term "sprain" (13) was borrowed from orthopedic medicine where it has long been used to describe "the damage of a ligament over a finite length without a full break" (e.g. ref. 14). The sprain is the second gradient $u_{i,jk}$ of the displacement vector u_i (in Cartesian coordinates x_i, $i, j, k = 1, 2, 3$), multiplied by the material characteristic length

Author contributions: Z.P.B. designed research; H.X., A.T.N., and Z.P.B. performed research; H.X., A.T.N., and Z.P.B. contributed new reagents/analytic tools; H.X., A.T.N., and Z.P.B. analyzed data; H.X. carried out most programing, computations and test data analysis, and helped in formulating the theory and computational approach; A.T.N. helped in programming, computations and conceiving the theory; Z.P.B. conceived the basic theory and helped in the basics of programming approach; and H.X., A.T.N., and Z.P.B. wrote the paper.

Reviewers: C.H., Technische Universität Wien, Vienna; and F.-J.U., Massachusetts Institute of Technology.

The authors declare no competing interest.

[1]To whom correspondence may be addressed. Email: z-bazant@northwestern.edu.

Published September 26, 2024.

l_0, which makes it dimensionless. The partial derivative of the sprain energy density with respect to the sprain tensor was named the "spress" tensor in (12), which is the work-conjugate force variable that corresponds to the sprain tensor. By means of l_0, the sprain and spress are made to possess, conveniently, the same dimensions as the strain and stress although they reflect third-order tensorial quantities. Note that they have nothing to do with continuum homogenization by a uniform strain field.

In the first paper that introduced the sprain energy as a localization limiter (13), the sprain energy attributed to each FE node was differentiated with respect to the nodal displacements. This yielded a set of self-equilibrated curvature-resisting sprain forces, some of which had to be applied on the nodes adjacent to each FE. This required an additional script file that ran through each node. Providing the adjacent node numbers turned out to be a huge computational disruption, which resulted in an increase in the running time of a FE code such as Abaqus by almost two orders of magnitude.

The slCBM formulated in ref. 12 eliminates the use of adjacent nodes by introducing an independent field ζ_{ij} that is an approximation of the actual strain gradient $u_{i,j}$. The ζ_{ij} and $u_{i,j}$ fields are weakly constrained by a Lagrange multiplier tensor λ_{ij}, while the gradient $\zeta_{ij,k}$ gives the curvature (or second gradient, the Hessian) of the displacement vector u_i. This formulation not only obviates the need for applying curvature-resisting sprain forces on the adjacent nodes, but also allows both u_i and ζ_{ij} fields to be assigned the lowest interelement continuity, C_0, which is generally favorable for fracture modeling.

The classical strain-gradient theories (15) may be regarded as a special case in which the gradient $\omega_{ij,k}$ of material rotation $\omega_{i,j}$ is not resisted, which seems to be a physically unrealistic feature (Fig. 1). The error due to omitting $\omega_{ij,k}$ will here be shown to be negligible in symmetric fracture modes but large, with errors exceeding 50%, in shear and shear-compression fracture modes.

Remark 1. CBM and previous experiments: Since the inception of the CBM in 1983, it was clear that the contrast between the success of the CBM predictions for shear and shear-compression cracks in reinforced concrete and the poor predictions of LEFM (2, 16, 17) and of the cohesive (line) crack model (or Hillerborg's fictitious crack model), signaled the importance of triaxial stresses at fracture front. To clarify this argument, Tschegg (18) came up already in 1995 with a pioneering adaptation of the wedge-splitting experiments. However, unlike the gap test which transits from one statically determinate configuration to another and can be evaluated easily without any FEA, Tschegg's test was statically indeterminate and did not have the maximum $\bar{\sigma}_{xx}$ near the notch tip. A proper evaluation of this test called for an optimization of FEA with an assumed triaxial damage law, which was complicated and made the evaluation ambiguous and inconclusive.

Remark 2. Isogeometric analysis (IGA) approach: A recent study (19) reproduced the sprain energy effect by means of IGA, which presents another avenue to ensure the continuity of displacement gradient. However, in the IGA, due to its founding in splines, the displacement gradient continuity is higher than C_0. In contrast to slCBM, the IGA basis precludes a nodal jump in the gradient of displacement gradient–a feature that seems not beneficial in numerical simulation of fracture. This suggests that, for the IGA, one might need to subdivide l_0 into smaller subelements than in ref. 19.

Spress and Sprain. To summarize the theory (12), we decompose the Helmholtz continuum energy density as

$$\bar{\Psi}(\boldsymbol{\epsilon}, \boldsymbol{\xi}) = \Psi(\boldsymbol{\epsilon}) + \Phi(\boldsymbol{\xi}), \qquad [1]$$

where Ψ is the standard strain energy density, Φ is the sprain energy density, $\boldsymbol{\epsilon}$ is the strain tensor and $\boldsymbol{\xi}$ is the sprain tensor

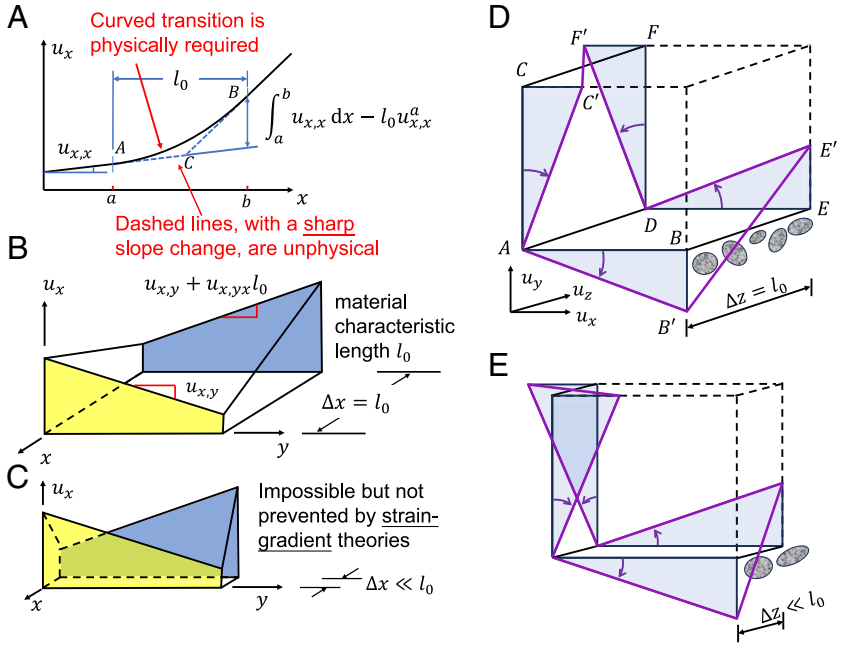

Fig. 1. (*A*) Due to heterogeneity, the dashed line is impossible, and the change of gradient indicated by the curve can happen only over a finite distance dictated by the maximum admissible curvature represented by the characteristic length l_0; (*B* and *C*) if the change of gradient over distance Δx equal to l_0 in figure (*B*) were unresisted (i.e., required no energy), then distance Δx could be made arbitrarily small as in figure (*C*), which would imply sliding on the vertical plane, absurdly with no resistance. This theoretical point suffices to prove that omission of the rotation gradient is wrong. (*D* and *E*) Example of a torsional displacement field when changing the distance between the two rotated planes at different spacing l_0.

(dimensionless); $\boldsymbol{\eta}$ denotes the strain gradient tensor (dimension m^{-1}). Their component definitions in Cartesian coordinates x_i (i = 1,2,3 in 3D, or 1,2 in 2D) are

$$\xi_{ijk} = l_0 u_{i,jk}, \quad \eta_{ijk} = \epsilon_{ij,k}, \quad u_{i,jk} = \eta_{ijk} + \omega_{ij,k}, \qquad [2]$$

The subscripts preceded by a comma denote partial derivatives; u_i = components of displacement vector $\boldsymbol{u}$, $\epsilon_{ij} = (u_{i,j} + u_{j,i})/2$ = small strain tensor (linearized) and $\omega_{ij} = (u_{i,j} - u_{j,i})/2$ = material rotation tensor. The sprain energy, the spress, and the curvature-limiting spress–sprain relation are expressed as

$$\Phi(\boldsymbol{\xi}) = \int s_{ijk} \mathrm{d}\xi_{ijk}, \qquad s_{ijk} = \frac{\partial \Phi(\boldsymbol{\xi})}{\partial \xi_{ijk}}, \qquad [3]$$

$$s_{ijk} = \kappa \big\langle |\boldsymbol{\xi}| - C \big\rangle \xi_{ijk}/|\boldsymbol{\xi}|, \qquad [4]$$

where $\Phi(\boldsymbol{\xi})$ is a function of the dimensionless third-order tensor $\boldsymbol{\xi}$ with components ξ_{ijk}. Here, κ is the sprain stiffness; $|\boldsymbol{\xi}|$ is a suitable norm of $\boldsymbol{\xi}$, for example, $\|\boldsymbol{\xi}\|_2$, and C is the threshold below which the effect of the displacement field curvature on damage is nil. Such a situation prevails in most of the volume of the structure, in which the classical continuum mechanics with no sprain energy applies.

Eq. **4** is not a constitutive relation, because its purpose is a geometric restriction—to set the maximum possible curvature of the vector field (e.g., force the dashed curve in Fig. 1*A* to be replaced by the smooth change of slope over certain minimum distance controlled by the maximum curvature). The spress–sprain relation may be thought to have a character similar to the penalty function used in computer simulations of rigid contact problems, in which there is, of course, no spring at the contact. The spring-like resistance provided by the penalty function is not a material property but a trick, emulated by the spress, which facilitates the FE analysis of contact problems. Different expressions for the penalty function may serve the purpose equally well.

The Question of Spress–Sprain Cross Dependence and Geometric Restriction. The spress–sprain relation is essentially a geometric limitation on the maximum possible curvature (or on the norm or magnitude of the hessian) of the displacement field that is permitted by the size of inhomogeneities in the material microstructure (as shown in Fig. 1*D*). It has nothing to do with the classical homogenization of the heterogeneous microstructure of material over some material representative volume whose size is the material characteristic length l_0. The only material length that must be considered as a material property in the spress–sprain relation is the minimum radius $r_0 = \Delta x/\Delta\theta$ admitted by the maximum size of material inhomogeneities (Fig. 1*D*), such as grains. This radius, or maximum possible curvature $1/r_0$, can be determined only by testing specimens with strongly variable displacement gradient, by inverse analysis of many observations, or by micromechanics. Clearly, r_0 cannot be less than several grain sizes (Fig. 1*E*).

Remark 3. Analogy with molecular interactions: Molecular dynamics is often simplified by considering only pair interactions between two particles characterized by a pair potential, such as the Lennard-Jones potential, whose derivative is an interparticle force. The pair potential completely defines the energy density as a function of the finite strain tensor. But for a more realistic modeling, one must also consider the valence and dihedral angles (20, 21). The valence angle is defined by three particles (or points), such as $\angle ACB$ in Fig. 1*A* or $\angle CAB$ in Fig. 1*D*. A change of this valence angle is resisted by a bending moment, similar to the idea of sprain energy here. The dihedral (or torsional) angle represents the relative (torsional) rotation of two parallel planes. This rotation is defined by four points, such as the top four corners of the gradient line segments on parallel vertical planes in Fig. 1*C*, and is resisted by a torque, same as here. Thus, omitting here the material rotation gradient is similar to omitting the resistance to a change of dihedral angle in molecular dynamics (considering quantum mechanics or DFT would, of course, be physically better justified yet also far more complicated). It may further be noted that the resistance to a change of dihedral angle has no analogy in the strain gradient models, which may be seen as a fundamental deficiency of these models.

Here, the only purpose of the sprain energy and the associated spress–sprain relation is to provide, for convenience, a soft quasi-elastic resistance to an abrupt change of displacement gradient. Although, in the previous study (13), a cross-correlation with volume change ϵ_V was intuitively introduced by a weakening threshold C when the volumetric expansion gets too high, there are no cross-relations introduced in Eq. **4**. In the present calculations, a zero spress is assumed in Eq. **3** when $|\boldsymbol{\xi}| <$ threshold C (although, in the preceding study, a small but nonzero spress below threshold C was introduced to enhance convergence (12)).

Broadening the Concept of Strain Localization Limiter. In 1976 (22), it was shown that FE analysis of a tensioned bar with a strain-softening stress–strain relation predicts, paradoxically, the bar to fail at zero energy dissipation unless the minimum element size is limited to a certain characteristic length l_0 that is a considered as a material property. Without this limitation, the continuum model could not reproduce the transitional size effect, nor the gap test and the variation of energy dissipation rate during crack growth. The strain-based localization limiter for the minimum element size (16, 23) thus became the basis of the crack band model (10) and generally of FE failure analysis of concrete, geomaterials, and fiber composite airframes with a strain-softening stress–strain relation.

The introduction of sprain requires the concept of localization limiter to be broadened to the localization of sprain, which includes localization of material rotation gradient. Since the sprain differs from the strain gradient, $\epsilon_{ij,k}$, only by $\omega_{ij,k}$, one needs to add only a localization limiter on the material rotation gradient. This is accomplished by introducing additional energy density that is a function of $\omega_{ij,k}$ and resists its value becoming unrealistically large.

Lagrange Multiplier Constraint as a Generalized Force Ensuing from Finite Difference Approximation. Consider a one-dimensional (1D) finite element of length h on the interval (x_1, x_2) with nodal displacements u_1 and u_2 at nodes x_1 and x_2, respectively (see figure 2 in ref. 12). The gradient of the axial displacement u within this element is $u' = (u_2 - u_1)/h$. Further, we introduce the nodal approximate independent displacement gradients ζ_1 and ζ_2. Their average over the element is $\bar{\zeta} = (\zeta_1 + \zeta_2)/2$. The advantage of having introduced the independent gradient is that the curvature of u within the element can be expressed as $u'' = (\zeta_2 - \zeta_1)/h$ without any need to consider the adjacent nodes outside the element. The condition for $\bar{\zeta}$ to approximate the actual gradient u' obviously is $u' - \bar{\zeta} \approx 0$. The finite difference solution of the nodal values of u based on all the foregoing equations can be used to enforce on a 1D bar the condition of maximum admissible curvature u''.

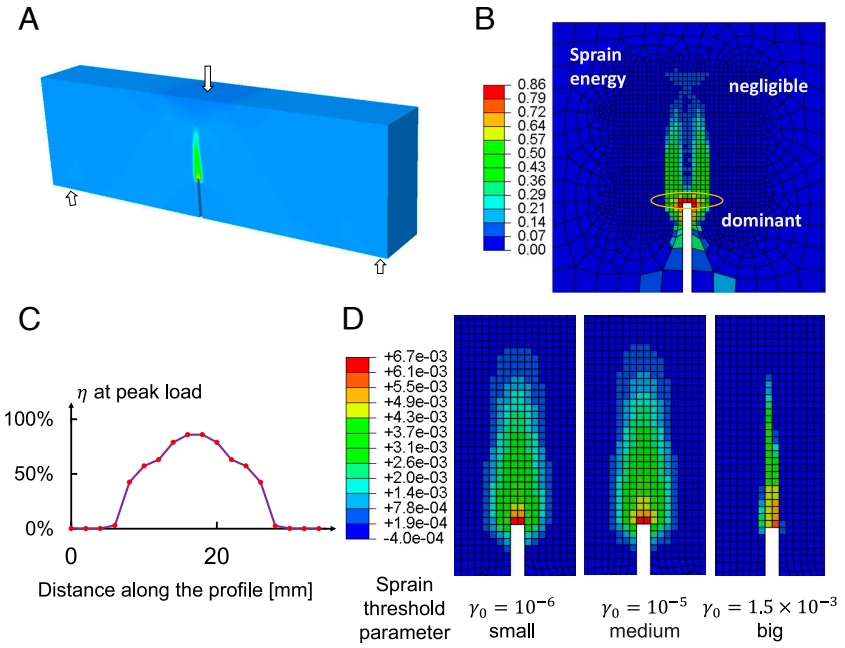

Fig. 2. (*A*) Near-notch damage zone in a symmetric 3PB test specimen using 3D slCBM with the (x,y) plane curvature control: 3-Point Bend test specimen (3PB). (*B*) Map of sprain energy intensity in the near-notch region. (*C*) Profile of sprain energy density ratio η across the damage band front (η = sprain energy density/total energy density); the curvatures at both ends are not restricted because $\eta <$ threshold. (*D*) Near-notch strain profiles for three cases of threshold parameter: 10^{-6}, 10^{-5}, and 1.5×10^{-3} (on the *Right*, the threshold is too large to limit curvature, which makes the situation equivalent to the CBM).

It would, however, be difficult to extend the foregoing finite difference approximation to two or three dimensions (2D or 3D). It is more effective and physically meaningful to recast the foregoing finite difference formulation into an energy-based variational form. To this end, we rewrite the constraint between $\bar{\zeta}$ and u' in the form:

$$\int_{x_1}^{x_2} \delta\lambda(u' - \bar{\zeta})\mathrm{d}x = 0, \qquad [5]$$

This is a variational equation that must be satisfied for any variation $\delta\lambda$ that is physically admissible. This variational equation implies a physical constraint on the gradient difference $\Delta = u' - \bar{\zeta} = 0$, which we already obtained directly from the finite difference approximation in 1D. Eq. **5** represents work and, since Δ is a generalized displacement, λ must be a generalized force, which is independent of Δ and represents a pseudogravity moment density (dimension N·m/m^3, or N/m^2).

Remark 4. Relation to nonlocal models: It is important to note that the present model is not a nonlocal model even though it prevents the softening damage from localizing into a zone narrower than a certain material characteristic length. The nonlocality, as a localization limiter, proposed in 1976 (22), is expressed through the averaging of strain or damage in a heterogeneous solid over a certain finite length, l_0, that is a material constant. The later advent of peridynamics perverted this concept as a supposed interaction between two points (or particles) over a finite distance, l_0, skipping intermediate particles. Such interaction does not exist in solids except on the atomic scale. Thus the peridynamics is a fiction and physically unjustified concept. Labeling it "nonlocal" is incorrect. Its predictions have been shown to conflict with a broad range of experiments (7, 8).

3D Variational Formulation. We now integrate the strain and sprain energy densities over the structure volume, including the applied volume forces $\boldsymbol{f}$ and surface forces $\boldsymbol{p}$, to obtain the expression for the potential energy (i.e., Helmholtz free energy) of the structure. Taking the first variations with respect to u_i, ζ_{ij} and λ_{ij}, we apply the Gauss integral theorem and obtain the following three minimization conditions of equilibrium as constrained by λ_{ij}:

$$\int_V \left\{\sigma_{ij}\cdot\delta\epsilon_{ij} - f_i\cdot\delta u_i - \lambda_{ij}\cdot\delta u_{i,j}\right\}\mathrm{d}V + \int_S p_i u_i \mathrm{d}S$$
$$= \int_V \left\{-\sigma_{ij,j} - f_i + \lambda_{ij,j}\right\}\cdot\delta u_i\,\mathrm{d}V + \int_S (\ldots)\,\mathrm{d}S = 0, \qquad [6]$$

$$\int_V \left\{s_{ijk}\cdot\delta\xi_{ijk} + \lambda_{ij}\cdot\delta\zeta_{ij}\right\}\mathrm{d}V$$
$$= \int_V \left\{-l_0 s_{ijk,k} + \lambda_{ij}\right\}\cdot\delta\zeta_{ij}\,\mathrm{d}V + \int_S (\ldots)\,\mathrm{d}S = 0, \qquad [7]$$

$$\int_V -\left\{u_{i,j} - \zeta_{ij}\right\}\cdot\delta\lambda_{ij}\,\mathrm{d}V = 0, \qquad [8]$$

Here, V and S are the structure volume and surface area, and p_i are the applied surface forces (or tractions). Eqs. **6–8** must hold for any variation, which yields the local equilibrium equations, similar to those for 2D (12):

$$\sigma_{ij,j} + f_i - \lambda_{ij,j} = 0, \qquad l_0 s_{ijk,k} - \lambda_{ij} = 0, \qquad u_{i,j} - \zeta_{ij} = 0, \quad [9]$$

Tackling Various Computational Difficulties. The 3D finite element formulation is based on the three variational Eqs. **6–8**. In 3D, we have 3 degrees of freedom (DoFs) per node for $\boldsymbol{u}$, 9 for $\nabla\boldsymbol{u}$, and 9 for $\boldsymbol{\zeta}$, which is in total 21 DoFs per node. This means 168 DoFs for an 8-node 3D brick element. Obviously, this creates a big computational burden and calls for a supercomputer.

This burden may be alleviated in various ways. One way is to delay the Lagrange multiplier tensor by one load step or, if the step is iterated, by one iteration. Thus, the number of DoF

of a brick element gets reduced to 96, which is still demanding but manageable. For our purpose, it suffices to consider only 3D situations in which the curvature (sprain) threshold C is expected to be exceeded only in plane (x, y), where x is the forward crack direction and y is the crack plane normal. Then, $u_{i,jk}$ needs to be limited only in plane (x, y). With no loss of accuracy, the formulation of a 3D 8-node brick element thus gets simplified as

$$\boldsymbol{u} = [u, v, w]_{1\times 3} \quad \boldsymbol{\zeta} = [\zeta_{ij}]_{2\times 2} \quad \boldsymbol{\lambda} = [\lambda_{ij}]_{2\times 2}, \qquad [10]$$

where $i, j \in \{1, 2\}$. Then, the numbed of DoFs per node is reduced to 11 or 88 per brick element (provided that the threshold C is not exceeded in the (x, z) and (y, z) planes). The 3PB (three-point-bend), the gap test, and the 4PB shear test configurations normally satisfy this condition, using a realistic threshold C.

Here, like in ref. 12, Matlab and Abaqus with a user element were used in FE simulations. It is also effective to use the coupling feature provided by COMSOL multiphysics software. Starting from Eqs. **6**–**8**, the weak variational equations derived from the energy functional provide partial differential equations of solid mechanics coupled with two additional tensorial fields. Using the weak contribution function feature, it was checked that COMSOL produces the same results as Matlab and Abaqus do. For postpeak softening, it provides simulations under crack mouth opening displacement control more seamlessly, with the applied load set as a global unknown variable. Besides, even though no time-dependent evolution equation is specified for $\boldsymbol{\zeta}$ and $\boldsymbol{\lambda}$, COMSOL yields the same displacement field of slCBM regardless of whether the explicit dynamic or static algorithm is used.

Fig. 2*D*, based on a smooth spress–sprain relation, shows that the maximum size h of square or brick elements needed to ensure a smooth strain profile across the damage band is $h \approx l_0/6$, where l_0 is the material characteristic [which is best identified by fitting scaled size effect tests (16)]. The figure also shows differences in the results caused by changing threshold C.

Numerical Simulation of the Gap Test Using the slCBM. Conceived in 2020 and presented in refs. 5 and 6, the gap test has three important consequences for fracture mechanics: 1) The material fracture energy, G_f, varies significantly with the crack-parallel stress, 2) a fracture process zone (FPZ) at the fracture front must have a finite width and 3) the FPZ must be characterized by a damage law that is not scalar but triaxial and tensorial, and must be realistic for each given material, such as microplane model M7 for concrete (24, 25) whose superiority in fitting the distinctive fracture test data was demonstrated in refs. 7 and 8, especially compared to peridynamics and phase field models.

Briefly, in the gap test, the end supports of 3PB specimens are set so as to leave certain gap. The gap is calculated so as to close at the right deflection of the beam (see the schematic picture in Fig. 3). The initial loading is introduced by a pair of elasto-plastic pads, which are installed at the mid-span on both sides of the notch, facing the midspan load applied by the testing machine. The gaps close once the pads reach the desired yield plateau. The bending moments from the reactions at the end supports, then, drive the fracture across the beam. A feature that makes the evaluation unambiguous is that the system remains statically determinate throughout the whole test (to simplify numerical simulations, the constant plastic forces from the yielding pads could thus be replaced by applied dead loads). Another important feature is that stress field near the notch tip is an almost uniform crack-parallel compressive stress.

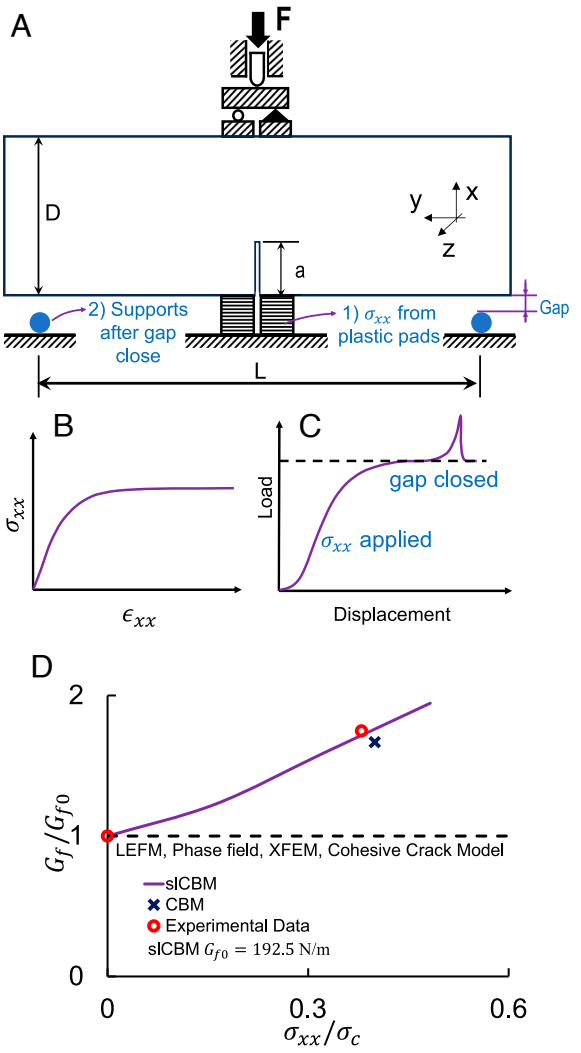

Fig. 3. (*A*) Schematic description for the gap test. (*B*) The stress–strain relationship for the pads demonstrates that plastic pads maintain nearly constant stress level after yielding. (*C*) Applied load vs. displacement in gap test. (*D*) Fracture energy G_f (normalized by G_{f0}) as a function of crack-parallel stress σ_{xx} (normalized by σ_c = 31 MPa) obtained by slCBM and the scalar elasto-plastic damage model compared with two experimental data points, each representing the average of 9 gap tests.

The gap tests are made on geometrically scaled specimens of three sizes, so that the G_f could be evaluated by the size effect method (which is the most accurate method for a heterogeneous material).

The computer results obtained with the slCBM and microplane model M7, shown in Fig. 3, match the experimental findings in which the fracture energy nearly doubles as the crack-parallel compressive stress increases from zero to 0.4 of the compressive strength σ_c. The results of 3D simulations shown here use the curvature control only in plane (x, y). This is sufficient in the case of the gap test because threshold C is never exceeded in planes (x, z) and (y, z). The result in Fig. 3 demonstrates that the slCBM can correctly reproduce the gap test and does so better than the classical CBM as reported in refs. 6 and 8.

It is noteworthy that if, by contrast, the damage constitutive law were changed to the classical scalar damage model, the material fracture energy G_f would remain constant when the applied crack-parallel stress changes (as in phase-field models). Thus the strong effect of crack-parallel compression on G_f

would be entirely missed. This highlights the importance of using a realistic triaxial damage law when large crack-parallel stresses act.

The M7 simulations with the classic CBM have previously been shown to predict a strong effect of transverse $\bar{\sigma}_{zz}$ in 3PB specimens. Ditto for the effect of $\omega_{ij,k}$ observed in Mode III fracture tests such as the torsion test (7). Such effects are likely to be universal.

Distinction from Classical Strain-Gradient, Micropolar, and Other Material Rotation Models. The initial aim of introducing the strain gradients into material constitutive laws was merely to provide homogenization of elastic microstructure with additional local degrees of freedom (28–31). In some studies (32), material rotations implied by the second gradient of displacement were mentioned but no energy density was attributed to them (until 2022). Most studies have used only simplified hypothetical damage laws, which are adequate for illustrative sketches but not for engineering applications.

Beginning in 1984, the strain-gradient continuum has been exploited as a damage localization limiter in quasibrittle fracture modeling (33–35), (36, section 2.9 and equation 13.10.25). An important theory, not intended for fracture, was the strain gradient (nonsoftening) plasticity in which the problem of localization instability does not arise. The objective was a continuum model of elasto-plastic metals accounting for geometrically necessary dislocations, motivated by thin metallic films (37–39).

Fundamentally different were theories such as the micropolar continuum which, too, are irrelevant to damage localization. In these theories, the homogenization dealt with specific material microarchitectures such as flexible frames in which, unlike here, the rotations of beam joints are independent of the rotations of beam members and are unrelated to the continuum displacement field (40–42), (36, sections 2.9 and 2.10).

Strain-Gradient vs. Sprain. Among all materials, the largest variety of fracture tests is available for concrete (see the 11 distinctive tests in ref. 43). In those tests as well as others, the microplane model M7 performed well and much better than the constitutive damage models of classical type, based on subsequent loading surfaces and tensorial invariants. However, M7 has not been implemented in the strain-gradient theory. If the sprain theory using the M7 constitutive law were compared to an existing strain-gradient theory with a different constitutive law, it would be unclear whether differences in their responses should be attributed to the difference in the constitutive law rather than to the replacement of the strain gradient with the sprain.

Here a strain-gradient approach is obtained by deleting the material rotation gradient, $\omega_{ij,k}$, which is accomplished by symmetrizing $u_{i,jk}$ at the outset. This makes the sprain tensor $\xi_{i,jk}$ identical to the strain-gradient tensor η_{ijk}, allowing a meaningful comparison to the sprain model. For clarity, we also ensure that the lack of material rotation gradient $\omega_{ij,k}$ be the only difference in the comparison by keeping in the strain gradient program not only the same constitutive law but also the present Lagrangian multiplier method with the C_0 continuity of strains as independent variables.

Fig. 4 *A–E* shows the change in slCBM when $\omega_{ij,k}$ is suppressed for simulations of 3PB mode I fracture tests, and of the 4PB Mode II shear fracture tests. Note that: 1) The strain-gradient and sprain models are virtually identical if the crack line is the axis of deformation symmetry, which implies Mode I fracture. 2) The slCBM simulations of shear or mixed-mode fracture predict loads up to 55% higher than the strain-gradient model. 3) The slCBM may also predict a different crack pattern. Due to a higher sensitivity to rotation gradient, the slCBM is able to capture the plane crack starting before the peak load, documenting a different crack evolution.

Still bigger differences may logically be suspected for 3D torsion-induced Mode III fracture simulations, in which the $\bar{\sigma}_{xx}$ may cause a major change in the failure mode and fracture shape that cannot be captured by the strain-gradient models at all.

Path Dependence Error in Approximating G_f As Function of $\bar{\sigma}_{xx}$. As shown in figure 9 in ref. 6, the effect of crack-parallel stress σ_{xx} is strongly path-dependent. For the gap test path (first applying $\bar{\sigma}_{xx}$, subsequently growing fracture), the results for concrete can be described by a simple curve-fitting (equation 1 in ref. 6): $G_f/G_{f0} = 1 + a/(1 + b/\xi) - (1 + a + b)\xi^s/(1 + b)$, where $\xi = \bar{\sigma}_{xx}$/compressive strength and G_{f0}, *a, b, s* are constants. However, whether $\bar{\sigma}_{xx}$ is applied first or latter, or simultaneously with fracture loading, makes a huge difference (see figure 9 in ref. 6).

Similar equations could be formulated for other materials and other simple paths. One might want to apply them in FE computations to modify the G_f based on the $\bar{\sigma}_{xx}$ value from the previous load step. That might roughly suffice for anchor pullout, ice sheet pushing on vertical column, shear in a reinforced concrete (RC) beam or slab, longitudinal crack in pressurized fuselage, or in rocket casing. However, in many other situations the G_f, as a function of ξ, is enormously path-dependent, and a FE incremental loading simulation is then required.

Likely Universality of Crack-Parallel Stress Effects. The experiments of Brockmann and Salviato (44) already showed that crack-parallel compression has also a major effect on the fracture energy, G_f, of composites. It causes a major decrease, rather than increase, of the G_f of quasi-isotropic fiber-polymer composites. Based on realistic simulations with the sphero-cylindrical microplane model (45), this effect should be strong in shale (similar to that in concrete but anisotropic) and probably in most rocks, in all composites, printed architected materials, and polymers that exhibit nano-scale crazing.

The simulation by lattice discrete particle model, LDPM, which is the most realistic discrete element type model for concrete, can also capture the crack parallel stress effect at the mesoscale (46), in good agreement with slCBM. The crack-parallel stress effect must even be expected in perfect atomistic fracture of crystals. Why?—Because fracture creates surface energy, which causes that the near-surface atoms must get displaced into higher energy positions and thus they must resist the crack-parallel stresses in a different way (it may be pointed out that the effect of crack-parallel stress on the surface energy was pointed out in a different context already in 1993 (47) but the connection to fracture mechanics has apparently not been noted).

Recent experiments (48) further showed that, in ductile fracture of polycrystalline metals such as aluminum, the long-studied effect of crack-parallel stress σ_{xx} (aka T-stress) cannot be attributed solely to the yielding zone (about 0.1 m in size). Rather, it must be partly attributed to the fact that the fracture in metals is not a line but must have a front that is several crystal sizes wide, which is about 5 to 10 μm. An interesting consequence is an extended size effect law with an intermediate asymptote (equation 40 in ref. 48, spanning the range 1 : 1000). This law allows the G_f of such metals (equal to the J-integral) to be obtained by measuring the maximum loads of relatively small specimens exhibiting large-scale yielding.

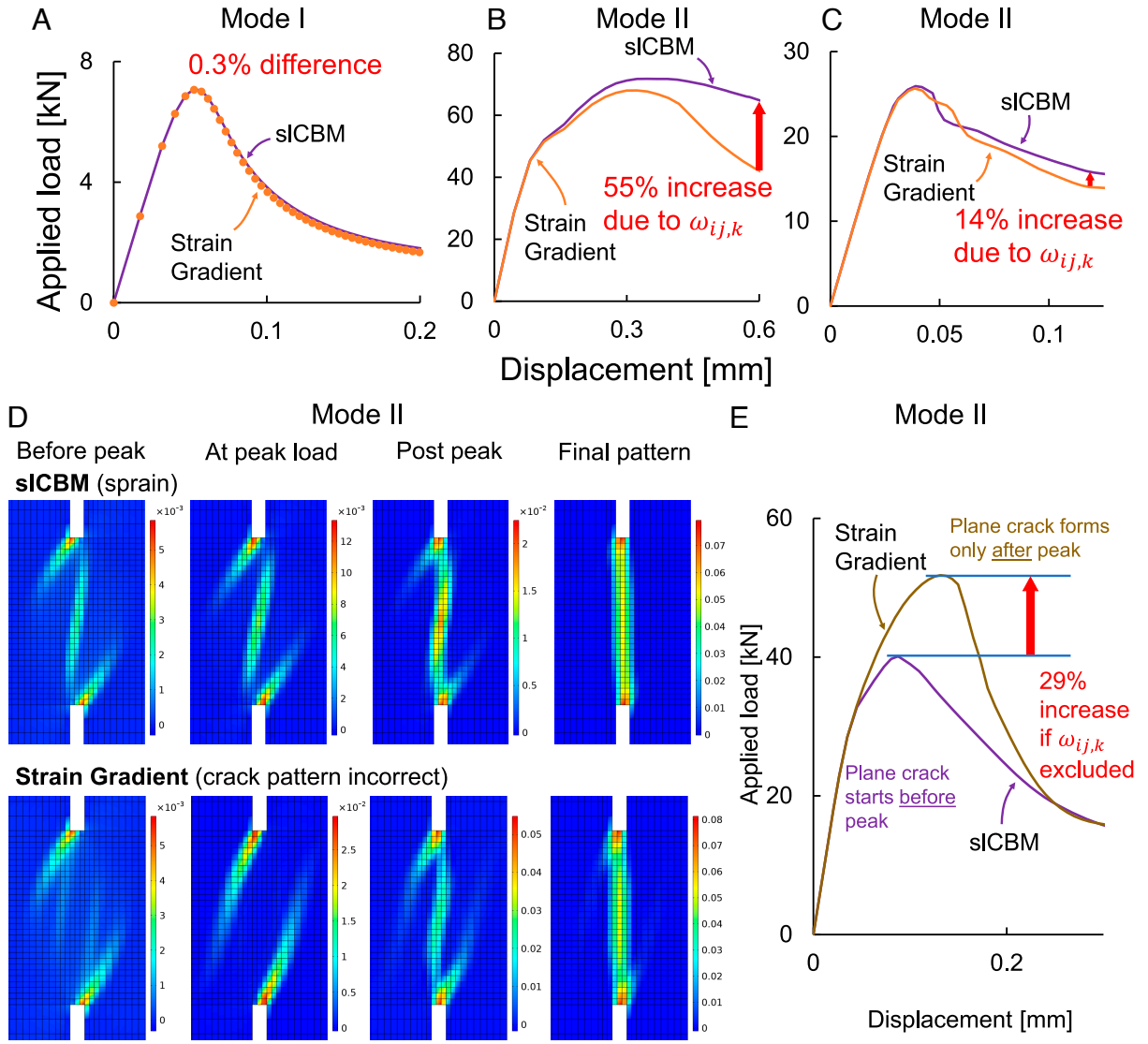

Fig. 4. Comparisons of the sICBM with the Strain Gradient theory: (*A*) In symmetric fracture tests such as 3PB, the responses obtained from both models differ negligibly because the rotation gradients in the notch line vanish and on its sides are not big enough to exceed the threshold. (*B* and *C*) These are Mode II four-point-bend (4PB) Iosipescu shear tests with two different constitutive laws—scalar damage law and Grassl's constitutive model [aka CDPM law featured in software COMSOL (26, 27)]. The material rotation gradients on the notch line are significant and increase in postpeak. (*D* and *E*) The absence of $\omega_{xy,x}$ in strain-gradient model leads to different fracture pattern evolution. After exceeding threshold $C = 5 \cdot 10^{-5}$, κ was kept constant. $\kappa = 1$ MPa in (*A*); $\kappa = 0.11$ MPa in (*B*); $\kappa = 0.105$ MPa in (*C*); and $\kappa = 0.01$ MPa in (*D* and *E*).

Where Crack-Parallel Stresses Matter in Practice. Almost everywhere: • Longitudinal crack in pressurized fuselage • Cracks in the casing of solid-fuel rocket • Shear cracks in aircraft wing, wing box, rudder, stabilizer • Shear cracks in fiber composite wind turbine • Aircraft and automobile crashworthiness, crush cans • Cracks caused by projectile impact • Most thermal cracks • Cracks in inflatable shells • Shear failure of RC beams and slabs (historically the most challenging fracture problem) • Cracks in prestressed concrete • Fracking, esp. with poromechanical fluid-to-solid stress transfer • Sea ice sheet pushing on a fixed structure • Cracks in geology or in seismic events such as Earthquake fault • Fracture in arch dam or arch bridge abutments, in footings • Pullout fracture of anchors in rock or concrete • Particle comminution in projectile impact • Fracture of fatigued plastic-hardening polycrystalline metals • Burst of mine stopes, borehole breakout, failure of tunnels, excavations • Fracture of bone, biomaterials, dental materials • Hydraulic fracturing for deep CO_2 sequestration in peridotite • Fracture in poro-mechanics • Implosion of Oceangate deep submersible near the Titanic wreck; etc.

Conclusions and Main Observations. The fracture mechanics of line cracks, whether LEFM or cohesive, must now be regarded as an illustrative yet crude approximation that is good enough for preliminary estimates. Nevertheless, it will remain indispensable for its instructive value, by describing the essential features of fracture behavior in simple terms. It will remain necessary for getting insight into the effects of various parameters and the scaling. The conclusions are as follows:

1. The smooth crack band with a tensorial Lagrange multiplier constraining the displacement gradient is shown to reproduce correctly the gap tests of various sizes, as well as the size effect tests and Mode II shear tests on a quasibrittle material (concrete).

2. Computer simulations calibrated by experiments confirm that the width of damage zone at the crack band front at peak load is not constant but changes with the crack-parallel stress. Yet the crack front width is controlled by a constant material characteristic length. The width changes play a role in how the crack-parallel stresses affect the fracture energy.
3. The crack-parallel stresses likely affect the fracture energy in all materials—not only concrete and rocks but also fiber composites, ceramics, sea ice, rocks and stiff soils, polymers, polycrystalline metals, and atomistically sharp nanoscale cracks in crystals.
4. The classical strain-gradient theories are a special case of the present theory when the material rotation gradient is omitted. They differ by only less than 1% for symmetric deformation fields in Mode I fractures. The reason is that the rotation gradient does not dominate in the regions on the side of the symmetry line.
5. The strain-gradient theories paradoxically imply that a finite difference in material rotations on two arbitrarily close parallel planes would not be resisted by sprains, in addition to the shear stresses produced by strains. The lack of energy density in the strain-gradient theories attributed to material rotation gradients must lead to insufficient resistance to fracture with a torsional (or Model III) component. This is physically questionable and may incur a significant error in Mode III fractures. Therefore, using these theories in engineering design might be inaccurate when a torsional component is involved [as in torsional Mode III fracture tests (7)].
6. For in-plane (Mode II) shear fractures, the maximum load can be 29% different and the postpeak load 55% higher than the strain gradient prediction, and insufficient resistance to localization of to the rotation gradient may also cause crack pattern evolution difference. The underestimation of shear fracture energy is of similar magnitude. This may particularly matter for simulating impact. Even bigger discrepancies are expected for torsional Mode III fracture. Generally, the errors grow with the sprain stiffness (i.e., with the slope of the spress–sprain curve).
7. Using a realistic triaxial softening damage law (such as microplane model M7 for concrete) is important. The spress–sprain relation is essentially geometric in nature, i.e., a curvature limitation, and is not a material constitutive property. Only the associated material length, of curvature radius, is a constitutive material property, dictated by the size and spacing of material inhomogeneities. Possible coupling with stress–strain constitutive relation may require further study.
8. While crude estimates can be made with a simple formula for the crack-parallel stress effect on the material fracture energy, fracture-sensitive designs will require incremental FE analysis based on the spress and sprain evolution.
9. The classical concept of localization limiter in FE damage analysis needs to be broadened by including the energy density of material rotation gradient as an additional localization limiter.
10. Finally, a benefit for strain-gradient FE programming: If we formulate a strain-gradient FE program as a special case of the FE sprain program in which the displacement gradient tensor is replaced by its symmetric part, the strain-gradient programming benefits from the computational efficiency of treating both the displacements and the strains as independent fields of C_0 continuity constrained by the Lagrangian multiplier.

Data, Materials, and Software Availability. Model inp files data have been deposited in the author's Github repository: github.com/Houlin2018/slCBM (https://github.com/Houlin2018/slCBM) (49). All study data are included in the main text.

ACKNOWLEDGMENTS. This work was partly supported by the Center on Geo-Processes in Mineral Carbon Storage, an Energy Frontier Research Center (EFRC) funded by the US Department of Energy, Office of Science, Basic Energy Sciences at the University of Minnesota under award DE-SC0023429. Further partial support: NSF (Grant No. CMMI-202964) and Army Research Office (ARO) (Grant No. W911NF-19-1-003), both to Northwestern University. Thanks are due to Dr. Yang Zhao, a Postdoctoral Associate at Northwestern, for valuable discussions, and to Karel Matouš, a professor of Aerospace and Mechanical Engineering at the University of Notre Dame, for proposing an initial (curvature-based) variant of the Lagrange multiplier approach while collaborating on the previous paper, whose subsequent (gradient-based) update was essential for success. Thanks are also due to Ellad Tadmor of University of Minnesota for pointing out the work of Mercer et al.

Author affiliations: [a]Department of Civil and Environmental Engineering, Northwestern University, Evanston 60208, IL; [b]Department of Mechanical Engineering, Northwestern University, Evanston 60208, IL; and [c]Department of Materials Science and Engineering, Northwestern University, Evanston 60208, IL

1. G. I. Barenblatt, The mathematical theory of equilibrium cracks in brittle fracture. *Adv. Appl. Mech.* **7**, 55–129 (1962).
2. G. R. Irwin, Analysis of stresses and strains near the end of a crack traversing a plate. *J. Appl. Mech.* **24**, 361–364 (1957).
3. N. Moës, J. Dolbow, T. Belytschko, A finite element method for crack growth without remeshing. *Int. J. Numer. Methods Eng.* **46**, 131–150 (1999).
4. A. Hillerborg, M. Modéer, P. E. Petersson, Analysis of crack formation and crack growth in concrete by means of fracture mechanics and finite elements. *Cem. Concr. Res.* **6**, 773–781 (1976).
5. H. T. Nguyen *et al*., New perspective of fracture mechanics inspired by gap test with crack-parallel compression. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 14015–14020 (2020).
6. H. T. Nguyen, M. Pathirage, G. Cusatis, Z. P. Bažant, Gap test of crack-parallel stress effect on quasibrittle fracture and its consequences. *J. Appl. Mech.* **87**, 071012 (2020).
7. Z. P. Bažant, H. T. Nguyen, A. Abdullah Dönmez, Critical comparison of phase-field, peridynamics, and crack band model M7 in light of gap test and classical fracture tests. *J. Appl. Mech.* **89**, 061008 (2022).
8. Z. P. Bažant, H. T. Nguyen, Proposal of m-index for rating fracture and damage models by their ability to represent a set of distinctive experiments. *J. Eng. Mech.* **149**, 04023047 (2023).
9. Z. P. Bažant, M. Jirásek, Nonlocal integral formulations of plasticity and damage: Survey of progress. *J. Eng. Mech.* **128**, 1119–1149 (2002).
10. Z. P. Bažant, B. H. Oh, Crack band theory for fracture of concrete. *Mat. Constr.* **16**, 155–177 (1983).
11. J. Červenka, Z. P. Bažant, M. Wierer, Equivalent localization element for crack band approach to mesh-sensitivity in microplane model. *Int. J. Numer. Methods Eng.* **62**, 700–726 (2005).
12. A. T. Nguyen, H. Xu, K. Matouš, Z. P. Bažant, Smooth lagrangian crack band model based on spress-sprain relation and lagrange multiplier constraint of displacement gradient. *J. Appl. Mech.* **91**, 031007 (2023).
13. Y. Zhang, Z. P. Bažant, Smooth crack band model-a computational paragon based on unorthodox continuum homogenization. *J. Appl. Mech.* **90**, 041007 (2023).
14. D. T. P. Fong, Y. Hong, L. K. Chan, P. S. H. Yung, K. M. Chan, A systematic review on ankle injury and ankle sprain in sports. *Sports Med.* **37**, 73–94 (2007).
15. E. Kröner, Elasticity theory of materials with long range cohesive forces. *Int. J. Solids Struct.* **3**, 731–742 (1967).
16. Z. P. Bažant, J. L. Le, M. Salviato, *Quasibrittle Fracture Mechanics and Size Effect: A First Course* (Oxford University Press, 2021).
17. A. A. Griffith, Vi., the phenomena of rupture and flow in solids. *Philos. Trans. R. Soc. Lond. Ser. A Contain. Pap. Math. Phys. Character* **221**, 163–198 (1921).
18. E. Tschegg, M. Elser, S. Stanzl-Tschegg, Biaxial fracture tests on concrete-development and experience. *Cem. Concr. Compos.* **17**, 57–75 (1995).
19. H. Nguyen, W. Li, Z. P. Bažant, Y. Bazilevs, Isogeometric smooth crack-band model (isCBM) using spress-sprain relations adapted to microplane theory. *J. Mech. Phys. Solids* **181**, 105470 (2023).
20. M. P. Allen, D. J. Tildesley, *Computer Simulation of Liquids* (Oxford University Press, 2017).
21. Q. Mao *et al*., Classical and reactive molecular dynamics: Principles and applications in combustion and energy systems. *Progr. Energy Combust. Sci.* **97**, 101084 (2023).
22. Z. P. Bažant, Instability, ductility, and size effect in strain-softening concrete. *J. Eng. Mech. Div.* **102**, 331–344 (1976).
23. Z. P. Bažant, J. Planas, *Fracture and Size Effect in Concrete and Other Quasibrittle Materials* (Routledge, 2019).
24. F. C. Caner, Z. P. Bažant, Microplane model M7 for plain concrete. I: Formulation. *J. Eng. Mech.* **139**, 1714–1723 (2013).

25. F. C. Caner, Z. P. Bažant, Microplane model M7 for plain concrete. II: Calibration and verification. *J. Eng. Mech.* **139**, 1724–1735 (2013).
26. P. Grassl, M. Jirásek, Damage-plastic model for concrete failure. *Int. J. Solids Struct.* **43**, 7166–7196 (2006).
27. P. Grassl, D. Xenos, U. Nyström, R. Rempling, K. Gylltoft, CDPM2: A damage-plasticity approach to modelling the failure of concrete. *Int. J. Solids Struct.* **50**, 3805–3816 (2013).
28. E. Cosserat, F. Cosserat, *Théorie des corps déformables (Theory of deformable bodies)* (A. Hermann & Fils, 1909).
29. R. Toupin, Elastic materials with couple-stresses. *Arch. Ration. Mech. Anal.* **11**, 385–414 (1962).
30. R. D. Mindlin, Micro-structure in linear elasticity. *Arch. Ration. Mech. Anal.* **16**, 51–78 (1964).
31. G. Cusatis, X. Zhou, High-order microplane theory for quasi-brittle materials with multiple characteristic lengths. *J. Eng. Mech.* **140**, 04014046 (2014).
32. M. E. Gurtin, A gradient theory of single-crystal viscoplasticity that accounts for geometrically necessary dislocations. *J. Mech. Phys. Solids* **50**, 5–32 (2002).
33. Z. P. Bažant, M. Jirásek, Nonlocal integral formulations of plasticity and damage: Survey of progress. *J. Eng. Mech.* **128**, 1119–1149 (2002).
34. E. C. Aifantis, On the microstructural origin of certain inelastic models. *J. Eng. Mat. Technol.* **106**, 326–330 (1984).
35. Z. P. Bažant, Imbricate continuum and its variational derivation. *J. Eng. Mech.* **110**, 1693–1712 (1984).
36. Z. P. Bažant, L. Cedolin, *Stability of Structures: Elastic, Inelastic, Fracture and Damage Theories* (World Scientific, 2010).
37. H. Gao, Y. Huang, W. Nix, J. Hutchinson, Mechanism-based strain gradient plasticity-I. Theory. *J. Mech. Phys. Solids* **47**, 1239–1263 (1999).
38. Y. Huang, H. Gao, W. Nix, J. Hutchinson, Mechanism-based strain gradient plasticity-II. Analysis. *J. Mech. Phys. Solids* **48**, 99–128 (2000).
39. Z. P. Bažant, Z. Guo, Size effect and asymptotic matching approximations in strain-gradient theories of micro-scale plasticity. *Int. J. Solids Struct.* **39**, 5633–5657 (2002).
40. A. C. Eringen, *Theory of Micropolar Elasticity* (Springer, 1965).
41. A. C. Eringen, D. Edelen, On nonlocal elasticity. *Int. J. Eng. Sci.* **10**, 233–248 (1972).
42. Z. P. Bažant, M. Christensen, Analogy between micropolar continuum and grid frameworks under initial stress. *Int. J. Solids Struct.* **8**, 327–346 (1972).
43. H. T. Nguyen, A. A. Dönmez, Z. P. Bažant, Structural strength scaling law for fracture of plastic-hardening metals and testing of fracture properties. *Extreme Mech. Lett.* **43**, 101141 (2021).
44. J. Brockmann, M. Salviato, The gap test-effects of crack parallel compression on fracture in carbon fiber composites. *Compos. Part A Appl. Sci. Manuf.* **164**, 107252 (2023).
45. C. Li, Z. P. Bažant, H. Xie, S. Rahimi-Aghdam, Anisotropic microplane constitutive model for coupling creep and damage in layered geomaterials such as gas or oil shale. *Int. J. Rock Mech. Min. Sci.* **124**, 104074 (2019).
46. Y. Lyu, M. Pathirage, H. T. Nguyen, Z. P. Bažant, G. Cusatis, Dissipation mechanisms of crack-parallel stress effects on fracture process zone in concrete. *J. Mech. Phys. Solids* **181**, 105439 (2023).
47. J. L. Mercer, M. Y. Chou, Energetics of the Si(111) and Ge(111) surfaces and the effect of strain. *Phys. Rev. B* **48**, 5374–5385 (1993).
48. A. A. Dönmez, H. T. Nguyen, H. Xu, Z. P. Bažant, Crack-parallel stress effect on fracture energy of plastic hardening polycrystalline metal identified from gap test scaling. *J. Mech. Phys. Solids* **173**, 105222 (2023).
49. H. Xu, Smooth Crack Band Model. GitHub. https://github.com/Houlin2018/slCBM. Deposited 6 Sept 2024.