



Cite this: *Soft Matter*, 2016,
12, 8069

Fracture toughness of hydrogels: measurement and interpretation

Rong Long^a and Chung-Yuen Hui^{*b}

The fracture mechanics of hydrogels, especially those with significantly enhanced toughness, has attracted extensive research interests. In this article we discuss the experimental measurement and theoretical interpretation of the fracture toughness for soft hydrogels. We first review the definition of fracture toughness for elastic materials, and the commonly used experimental configurations to measure it. In reality most gels are inelastic. For gels that are rate insensitive, we discuss how to interpret the fracture toughness associated with two distinct scenarios: crack initiation and steady-state crack propagation. A formulation to estimate energy dissipation during steady-state crack propagation is developed, and connections to previous models in the literature are made. For gels with rate-dependent behaviors, we review the physical mechanisms responsible for the rate-dependence, and outline the difficulties to rigorously define the fracture toughness for both crack initiation and propagation. We conclude by discussing a few fundamental questions on the fracture of tough gels that are yet to be answered.

Received 24th July 2016,
Accepted 11th September 2016

DOI: 10.1039/c6sm01694d

www.rsc.org/softmatter

1. Introduction

Hydrogels are polymer networks swollen by water molecules. The permeability to water or solute molecules and biocompatibility of

hydrogels have made them useful for food products,¹ personal care,² drug delivery vehicles,^{3–6} and tissue engineering scaffolds.^{6–8} Recently many new applications emerge where hydrogels are required to bear or even generate mechanical forces.⁹ Examples include artificial cartilage,¹⁰ autonomous sealants for microfluidic channels^{11,12} or oil wells,¹³ soft actuators,^{14,15} and stretchable ionic conductors.^{16,17} For these applications, the mechanical properties of hydrogels are critical parameters to enable high-fidelity simulation and design. Significant efforts have been devoted to modeling the multi-faceted mechanical properties of hydrogels, including

^a Department of Mechanical Engineering, University of Colorado Boulder, Boulder, CO 80309, USA

^b Field of Theoretical and Applied Mechanics, Sibley School of Mechanical and Aerospace Engineering, Cornell University, Ithaca, NY 14853, USA.
E-mail: ch45@cornell.edu



Rong Long

Rong Long is an Assistant Professor in the Mechanical Engineering Department at the University of Colorado Boulder. Prior to that he was an Assistant Professor at the University of Alberta in 2013–2014 and a Research Associate at the University of Colorado in 2012. He received his PhD degree in Theoretical and Applied Mechanics from Cornell University in 2011, and BS degree in the same field from University of Science and Technology of China in 2006.

Dr Long's current research focuses on the nonlinear mechanics of soft materials, including constitutive modeling of active soft materials, fracture mechanics, contact mechanics, adhesion and biomechanics.



Chung-Yuen Hui

Chung-Yuen Hui received an undergraduate degree in physics and mathematics from the University of Wisconsin-Madison. He has a master degree in Applied Mathematics and a PhD in Applied Mechanics, both from Harvard University. He joined the Cornell faculty in 1981 and is currently the Joseph-Ford professor of Mechanical and Aerospace Engineering. His current research is on the mechanical behavior of soft materials. He was the recipient

of several teaching awards, including the Tau Beta Pi Teaching award in 2005. He received the 3M award for excellence in adhesion science in 2011 and has published over 260 papers in international journals.

nonlinear elasticity,¹⁸ viscoelasticity,^{19–22} poroelasticity,^{23–26} sensitivity to pH values²⁷ and temperature,²⁸ damage^{29,30} and fracture.^{31–39} Among these properties, the ability of hydrogels to sustain mechanical stress and to resist fracture is crucial, *e.g.* for setting the operation limit and service life of hydrogel-based devices. Typically macroscopic fracture originates from the growth of small defects (*e.g.* cracks) due to material damage induced by the amplified local stress field surrounding the defects. A metric widely used to characterize a material's ability to resist defect growth is the fracture toughness. In linear elastic fracture mechanics (LEFM),^{40,41} toughness is usually defined in reference to the initiation and growth of a pre-existing crack under prescribed loading; it can refer to the critical stress intensity factor upon crack growth (unit: $\text{Pa m}^{1/2}$) or the critical energy required per unit area of crack growth (unit: J m^{-2}). The latter definition, also known as fracture energy Γ_c , is more appropriate for soft gels, since the concept of stress intensity factor may not apply for cracks undergoing large deformations.⁴²

Intrinsically hydrogels are brittle due to the reduced areal density of polymer chains caused by their large water content. Typical values of Γ_c are estimated to be on the order of 10 J m^{-2} (*e.g.* alginate gel),^{43,44} which is several orders of magnitude less than that of natural rubber ($\sim 10^4 \text{ J m}^{-2}$). This disadvantage is overcome by the developments of hydrogels with enhanced toughness where Γ_c can be as large as 10^3 to 10^4 J m^{-2} while the gel maintains $\sim 90 \text{ wt\%}$ of water.^{44,45} A representative example of the tough hydrogels is the double network (DN) gel pioneered by Gong *et al.*⁴⁵ The key mechanism behind the dramatically enhanced fracture toughness is the energy dissipation due to the breakage of the sacrificial network near the crack tip while the other network maintains the macroscopic integrity of the gel. Many other physical mechanisms have been explored to enhance the fracture toughness, including the introduction of the crystalline phase,⁴⁶ hybrid networks with irreversible and reversible crosslinks,^{47–49} composite gels with micro-⁵⁰ or nano-particles,^{51–53} slider ring crosslinks,⁵⁴ and protein unfolding.⁵⁵ Examples of tough hydrogels and their fracture resistance are illustrated in Fig. 1. As reviewed by Zhao,⁴³ all toughening mechanisms share a common theme, *i.e.* introducing energy dissipation mechanisms into the polymer network while maintaining macroscopic integrity. These mechanisms are usually triggered by a certain threshold of stress. When coupled to the amplified stress field near the crack tip, these energy dissipation mechanisms lead to the formation of a dissipation zone surrounding the crack tip. As a result, the energetic driving force supplied by the remote loading cannot be fully delivered to the crack tip. Effectively the crack tip is shielded from the remote loading and the apparent fracture toughness is enhanced.

Driven by the need to further enhance the fracture resistance of hydrogels, the proper measurement and interpretation of fracture toughness become a frequent question of interest. The tensile strength, *i.e.* maximum stress upon fracture in a uni-axial tensile test, may serve as an indicator of fracture resistance. However, it does not solely characterize fracture behavior since it is also highly dependent on the stiffness of the material. For example, silica glass is known to be brittle⁴¹ with a $\Gamma_c \sim 10 \text{ J m}^{-2}$,

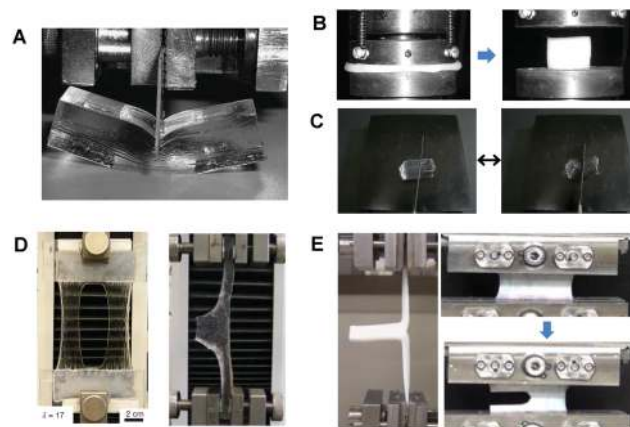


Fig. 1 (A) Double network gel consisting of a stiff brittle network and a soft extensible network. Reproduced from ref. 57 with permission from The Royal Society of Chemistry. (B) Micro-sphere composite gel. Reproduced from ref. 50 with permission from John Wiley & Sons Inc. (C) Hydrogel with slider-ring cross-linkers. Reproduced from ref. 54 with permission. (D) Hybrid alginate–polyacrylamide gel. Fracture resistance is demonstrated by stretching cracked specimens. Reproduced from ref. 44 by permission from Macmillan Publisher Ltd. (E) Polyampholytes physical hydrogels. Fracture toughness is measured using tearing test and pure shear fracture test. Reproduced from ref. 49 by permission from Macmillan Publisher Ltd.

much smaller than that of the DN gel ($\Gamma_c \sim 10^3 \text{ J m}^{-2}$);⁵⁷ but the tensile strength of glass ($\sim 100 \text{ MPa}$)⁵⁶ is much larger than the DN gel ($1\text{--}10 \text{ MPa}$)⁵⁷ because of the large Young's modulus of glass ($\sim 70 \text{ GPa}$) in comparison to that of the DN gel ($0.1\text{--}1 \text{ MPa}$). Similarly, high fracture strain in a tensile test does not necessarily imply high fracture toughness either. The maximum stress or strain criterion does not work well even if one restricts its usage to the same material. For example, larger specimens are likely to have more and larger flaws, and as a result, the maximum stress can be much less than that of a smaller sample of the same shape.⁵⁶ Likewise, specimens with different shapes (*e.g.* a dog-bone tensile specimen *versus* a rectangular thin-sheet specimen) and subjected to different loading conditions can fail very differently.

The objective of this review is to elucidate the concept of fracture toughness for hydrogels, especially those with energy dissipation mechanisms and highly resistant to fracture. Because hydrogel fracture is only studied until recently, many interpretations and modeling of experiments are still primarily based on linear elastic fracture mechanics (LEFM) in which material damage was assumed to be confined in a very small region surrounding the crack tip. In principle LEFM does not apply for tough hydrogels where the enhanced toughness is due to a large dissipation zone surrounding the crack tip. This situation is parallel to the fracture of elastic-plastic or high temperature materials which has stimulated the development of nonlinear fracture mechanics for metals.^{58–61} For tough hydrogels, a number of models have been proposed to link the crack-tip dissipation to the fracture toughness,^{62–64} and an excellent review of the toughening principle and various physical dissipation mechanisms are recently provided in Zhao.⁴³ However, as noted by Creton and Ciccotti,⁶⁵ a systematic

theoretical framework that rigorously defines the fracture toughness for soft dissipative materials is yet to be developed. This is the focus of our review.

The plan of this paper is as follows. We first briefly review the experimental methods to measure fracture toughness (Section 2) in predominantly elastic solids. Most of the results in this section are valid for large deformation, as long as the material is hyperelastic. This is followed by a discussion on the physical interpretation of toughness for tough gels with rate insensitive damage mechanisms (Section 3). The characterization of fracture for rate-dependent hydrogels is discussed in Section 4. In Section 5 we will discuss a list of future challenges related to the toughness of hydrogels.

2. Experimental measurement of fracture toughness

2.1 Background: energy release rate and J -integral

The essential goal of fracture mechanics is to define a criterion for crack growth and to link this criterion to the physical mechanisms of fracture. Typically such a criterion involves the comparison between two distinct quantities: one is a material property which quantifies the resistance to crack growth, *e.g.* the fracture toughness Γ_c , and the other describes the driving force for crack growth as governed by external loading and specimen geometry. For a crack to grow, the latter must reach or exceed the former. There are two equivalent approaches in LEFM to define the driving force for crack growth: the energetic approach and the field approach. In the energetic approach, the driving force is the energy release rate G , defined as the change in the sum of the elastic strain energy stored in the crack specimen and the potential energy of the loading system per unit area of crack growth. On the other hand, the field approach defines the driving force for crack growth according to the amplitude of the singular stress field surrounding the crack tip. In LEFM, the crack tip stress field follows universal structures⁴¹ and exhibits a square root singularity. As a result, the amplitudes of the singular terms in the stress field, known as the stress intensity factors, are used in the crack growth criterion. Alternatively, these amplitudes are uniquely related to a path independent integral surrounding the crack tip, known as the J -integral. For example, in LEFM J is uniquely related to the stress intensity factors.⁶⁷ For cracks subjected to the plane stress or strain conditions with finite deformation, the J -integral is defined as follows:⁶⁶

$$J = \int_C \left(WN_1 - \frac{\partial u_\alpha}{\partial X_1} S_{\alpha\beta} N_\beta \right) d\xi \quad \alpha, \beta = 1, 2, \quad (2.1)$$

where W is the strain energy density, u_α is the in-plane displacement field, $S_{\alpha\beta}$ is the first Piola–Kirchhoff stress, C is a smooth contour in the reference configuration which encloses the crack tip, starting from the lower crack face and ending on the upper crack face as shown in Fig. 2, ξ is the arc length and N_β are the Cartesian components of the unit outward normal vector of C .

The energetic approach to fracture is attractive because it is difficult to determine the state of stress and deformation fields

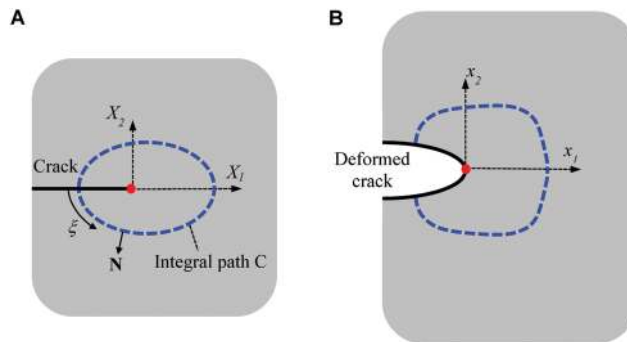


Fig. 2 Schematic of the path C for the J -integral. (A) Undeformed configuration of the crack where the path C is defined. The path C is illustrated as a dashed line starting at the lower crack face and ending on the upper crack face, and encloses the crack tip. (B) The deformed configuration of the crack to illustrate the distortion of path C after deformation.

near the tip of cracks even for purely elastic solids. For soft elastic material subjected to large deformation, the behavior of these tensor fields is highly dependent on the constitutive model.⁴² Fortunately, for purely elastic solids, the J -integral is independent of paths such as C (see Fig. 2) and it is equal to the energy release rate G . We emphasize that this conclusion is based on an implicit assumption that the crack grows in a self-similar manner. In other words, the crack extends straightly along its length, and as the crack tip moves by an infinitesimal amount, the crack tip stress and deformation fields translate with respect to a coordinate system that is centered at the crack tip. Of course, no material is perfectly elastic all the way to the crack tip; the stress/strain singularities must ultimately be removed by material damage in combination with inelastic behavior. For brittle solids, these nonlinearities are localized in a region surrounding the crack tip that is much smaller than the typical specimen dimensions (*e.g.* crack length), a condition called small scale yielding (SSY).⁴¹ As long as SSY holds, there exists a region close to the crack tip, but away from the non-linear zone, where the singular field still dominates, and this justifies $J \approx G$. As a result, the energetic and field approaches are equivalent in LEFM. However, for soft solids, especially for tough hydrogels, the SSY assumption is usually not satisfied since the size of the damage zone can be very large. In addition, cracks may not grow in a self-similar way. For example, secondary cracks may initiate on the blunted crack surface for soft gels, leading to rough crack surfaces or tortuous crack paths.^{36,68} In such cases, one may not always compute an energy release rate G , and even if it exists, will not be equal to J . Further, J may not be path independent. These issues will be discussed in detail in Section 2.3.

2.2 Common specimens for fracture testing

In this section we summarize four commonly used experimental configurations in the literature of hydrogel fracture. They are: (1) the pure shear (PS) test,⁶⁹ (2) the simple extension (SE) test,⁶⁹ (3) the single edge crack (SEC) test,⁷⁰ and (4) the tearing test.⁷¹ These specimens allow easy calculation of the energy release rate G during crack propagation, which is then taken as a measurement of fracture toughness Γ_c . The PS, SE and

tearing configurations have the advantage that the stress and strain fields for a steadily growing crack are invariant to crack length. In the following, we assume that the materials are elastic and the crack grows in a self-similar manner, and hence $J = G$.

2.2.1 Pure shear test. The PS test was first proposed in Rivlin and Thomas⁶⁹ to test fracture of rubber samples and has been recently adopted to characterize gel fracture.^{32,33,36,44,49,72,73}

The undeformed sample is a long thin strip of width L_0 , height $2H_0$ and thickness b_0 with $L_0 \gg 2H_0$ and b_0 (see Fig. 3A). A long crack of length c ($c \gg 2H_0$) lies in the middle between the top and bottom boundaries of the strip which are clamped to the loading device. Typically, a uniform vertical displacement $\pm\Delta$ (or displacement rate $\pm\dot{\Delta}$) is imposed on the top and bottom of the strip. The condition that L_0 and $c \gg 2H_0$ enables the translational invariance of the stress and strain fields, *i.e.* they remain unchanged in a moving coordinate system centered at the crack tip as the crack propagates. In addition, if $L_0 \gg c$, the material far ahead of the crack tip is under spatially uniform deformation with principle stretches $\lambda_1 = 1$, $\lambda_2 = \lambda_s$, $\lambda_3 = 1/\lambda_s$, where the subscripts 1, 2, and 3 denote the direction parallel to the undeformed crack, perpendicular to the undeformed crack and along the thickness, respectively. As derived in Rivlin and Thomas,⁶⁹ the energy release rate is independent of crack length and is given by

$$G = 2\bar{W}(\lambda_s)H_0, \quad \lambda_s = 1 + \frac{\Delta}{H_0}. \quad (2.2)$$

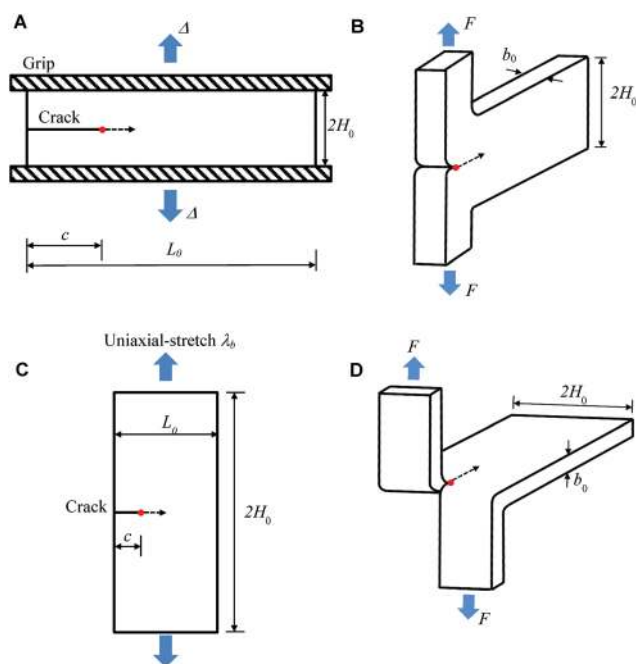


Fig. 3 Fracture test configurations. (A) Pure shear. The thickness of the specimen is b_0 (not shown in the figure), and the grips assumed to be rigid. (B) Simple extension. (C) Single edge crack. (D) Tearing. For (A and C) the undeformed geometry is shown to illustrate specimen dimensions, while for (B and D) the deformed geometry is shown to illustrate the loading mode. The crack tip is illustrated as a red point and the dashed arrow indicates the projected crack path.

where $\bar{W}(\lambda_s)$ is the strain energy density in material points far ahead of the crack tip. For an isotropic incompressible elastomer, the strain energy density W depends only on two scalar invariants I_1 and I_2 of the deformation gradient tensor, *i.e.* that is, $W(I_1, I_2)$. For example, for a Mooney–Rivlin solid,

$$W(I_1, I_2) = c_1(I_1 - 3) + c_2(I_2 - 3), \quad (2.3)$$

where c_1 and c_2 are material parameters. For the PS configuration, $\bar{W}(\lambda_s)$ can be calculated using $W(I_1, I_2)$ while setting

$$I_1 = I_2 = \lambda_s^2 + \lambda_s^{-2} + 1. \quad (2.4)$$

In practice, $\bar{W}(\lambda_s)$ is typically measured by subjecting an uncracked sample to tension under the pure shear constraint ($\lambda_1 = 1$) and calculating the area under the measured stress-strain curve.

2.2.2 Simple extension test. This configuration was also proposed by Rivlin and Thomas⁶⁹ to determine the tearing energy of rubber. Recently it has been applied to study fracture of DN gels.^{74–76} Fig. 3B shows the specimen in the deformed state. Similar to the PS test, the undeformed sample geometry here is defined by the length L_0 , height $2H_0$ and thickness b_0 . However, unlike the PS test, here the two arms of the specimen at the cracked end are clamped and peeled apart as shown in Fig. 3B. The energy release rate G was determined by Rivlin and Thomas,⁶⁹ which can be written as

$$G = \frac{2\lambda_a F}{b_0} - 2\tilde{W}(\lambda_a)H_0, \quad (2.5)$$

where F is the force applied to the two arms, λ_a and \tilde{W} are the stretch ratio and elastic strain energy density of the arms, respectively. The function $\tilde{W}(\lambda_a)$ can be calculated by recognizing that the arms are under uni-axial tension where

$$I_1 = \lambda_a^2 + 2\lambda_a^{-1}, \quad I_2 = 2\lambda_a + \lambda_a^{-2}. \quad (2.6)$$

Note that in Tanaka *et al.*⁷⁴ and other works,⁷¹ the elastic deformation of the two arms is neglected, *i.e.* ($\lambda_a = 1$, $\tilde{W} = 0$), and so eqn (2.5) is simplified to $G = 2F/b_0$. This approximation should not be applied if the arms are severely stretched with a large λ_a .

2.2.3 Single edge crack test. This test configuration was used by Greensmith⁷⁰ to determine the fracture energy of several vulcanized natural rubbers, and was recently used to study gel fracture.^{53,77,78} Using a compliance method, Greensmith⁷⁰ found that the energy release rate for short cracks with length $c \ll L_0$ is approximately given by

$$G = \frac{6}{\sqrt{\lambda}} \tilde{W}(\lambda_b)c. \quad (2.7)$$

where $\tilde{W}(\lambda_b)$ is the strain energy density of an uncracked sample subjected to a uniaxial stretch λ_b , and L_0 is the width of the sample which is assumed to be much less than the sample height $2H_0$ (see Fig. 3C). Eqn (2.7) has two limitations:

- It is valid only for small crack lengths and for small to moderate strains, and has not been verified for large strains. Furthermore, it can be shown that eqn (2.7) does not reduce to the LEFM result in the limit of $\lambda_b = 1$. In this limit, it

underestimates the energy release rate by about 25% in comparison to the LEFM result.⁷⁹

- The tensile behavior of the rubbers used by Greensmith⁷⁰ was shown to obey the Mooney–Rivlin model (see eqn (2.3)). It is not clear how well this expression works for elastomers with different strain hardening behaviors, especially at large deformation.

2.2.4 Tearing test. The tearing test, also known as the trousers test, was used to characterize the fracture of rubber⁷¹ and elastomers,⁸⁰ and also applied to tough gels.^{49,81} Unlike the three configurations above where the crack is primarily deformed in the opening mode, or referred to as Mode I in LEFM, in tearing tests the crack is deformed by out-of-plane shear loading. As shown in Fig. 3D, the two arms of a pre-cracked specimen are oppositely displaced to impose the tearing load. The out-of-plane loading results in complex three-dimensional stress and deformation fields in the testing specimen. However, this can be circumvented by taking the energetic approach to derive the energy release rate G . It turns out that for the tearing test G can also be calculated using eqn (2.5) provided that F is the tearing force, b_0 is the specimen thickness (see Fig. 3D) and λ_a is the stretch ratio in the two arms. If the elastic deformation of the two arms is neglected, G can be approximated as $G = 2F/b_0$ which is often used in the literature.^{71,80,81}

2.3 Stationary versus propagating cracks

In the four experimental configurations above, the cracks are all assumed to be propagating. There have been experiments in the literature that start with stationary cracks, *e.g.* using the PS geometry, where one can increase the applied load until it reaches a point when the crack starts to grow. We refer to such tests as crack initiation. For crack initiation, the energy release rate G is not well defined since the definition of G inherently involves crack extension. Indeed, how the crack extends after it starts to grow cannot be determined from the crack initiation data. A more appropriate way to characterize crack initiation is to use the field approach, *i.e.* the criterion for the onset of crack growth is based on the J -integral:

$$J|_{\text{initiation}} = \Gamma_c^{\text{init}}, \quad (2.8)$$

where Γ_c^{init} is a material property describing the fracture toughness at crack initiation. The key idea is that for a stationary crack in an elastic solid, the J -integral uniquely characterizes the crack tip stress and strain fields. For example, in a linear elastic solid with Young's modulus E , the stress intensity factor K_I for a plane stress Mode I crack uniquely characterizes the crack tip fields and is related to the J -integral by

$$J = K_I^2/E, \quad (2.9)$$

and hence the criteria governing crack initiation can also be expressed in terms of J . The situation is more complicated for large deformation, since the behavior of the crack tip stress field depends on the strain energy density function W . Nevertheless, as long as the material is elastic, the dominant stress field near the tip of a stationary crack can be shown to be uniquely related to J .⁴²

For predominantly elastic solids, the crack propagation tests (*e.g.* using the four configurations summarized in Section 2.2)

would also give a critical energy release rate Γ_c . A natural question then arises: is Γ_c^{init} equal to the Γ_c measured for propagating cracks? In other words, do the crack initiation tests give the same toughness measurement as the crack propagation tests? Theoretically the answer should depend on whether the crack grows in a self-similar manner after being initialized. If so, the concept of energy release rate G can be applied to the crack initiation test as well, and $G = J$. However, if the crack deviates from its projected path after initiation, or if the size of the dissipation zone (even if it is small) increases as the crack growth, then Γ_c^{init} and Γ_c are not necessarily the same. For the latter case, assuming SSY, the fracture toughness Γ_c is a function of crack length increment $\Delta c = c - c_0$, where c and c_0 are the current and initial crack length, respectively, that is

$$\Gamma_c(\Delta c) = G(c) = J(c), \quad (2.10)$$

with $\Gamma_c(\Delta c \rightarrow 0^+) = \Gamma_c^{\text{init}}$.

Finally, it should be noted that G and J are purely mechanical quantities, they are completely determined by the continuum fields and the sample geometry, whereas $\Gamma_c(\Delta c)$ is a material property that cannot be determined by continuum analysis without additional information about the fracture process. It is common to conflate the two, but here we try to make this distinction as clear as possible.

3. Fracture toughness for hydrogels with rate-independent dissipation mechanisms

In this section, we consider rate-independent tough hydrogels with damage mechanisms that are modulated by mechanical deformation. A representative example of such gels is the DN gels where both polymer networks are chemically cross-linked. For these gels, the sacrificial bonds cannot heal and their mechanical responses are rate insensitive. Wang and Hong²⁹ called such materials pseudo-elastic since the loading and unloading behaviors can be described by two separate hyper-elastic models.

The main difficulty of applying the elastic fracture mechanics results in Section 2 to interpret fracture tests of tough gels is that these gels are not elastic and the SSY condition is usually not satisfied. Indeed, it is the extensive inelastic behavior that enables toughening. As a result, it is necessary to clarify the usage of energy release rate G . Although one may still calculate the strain energy for such inelastic gels, which is dependent on loading histories, it would be too simplistic to interpret G in the same way as elastic materials, *i.e.* the change in the stored strain energy and potential energy of loading system per unit area of crack growth. For inelastic gels, only a small part of the work done to a material point is stored as strain energy. The rest is lost due to inelastic mechanisms (*e.g.* damage or plastic deformation). Therefore, if G is still defined using the stored strain energy, it would underestimate the energetic driving force required for crack growth, since it does not account for the additional energy dissipation accompanying crack growth.

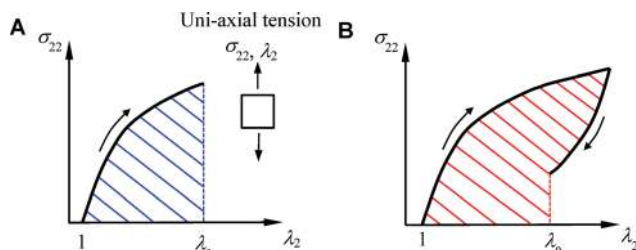


Fig. 4 An example of the loading history dependence of inelastic materials with loading–unloading hysteresis. Consider uni-axial tension along the X_2 direction with nominal stress σ_{22} and the stretch ratio λ_2 . (A) Monotonic loading to λ_0 . (B) Loading and unloading back to the same stretch ratio λ_0 . The stress work required for the two cases is different and is given by the shaded areas.

We argue that for inelastic gels G should be defined as the total mechanical work required per unit area of crack growth. This work consists of two parts: (i) the external work by the loading system; (ii) the stress work by the material elements. For the stress work, if a material element unloads during crack growth, it does positive work to the neighboring material elements by releasing the stored strain energy. If a material element is further loaded during crack growth, it does negative work to the neighboring material elements. For elastic materials, the stress work is equal to the change in strain energy of the material element. This is not true for inelastic materials with loading–unloading hysteresis, where the stress work depends on the particular loading path that a material element went through during crack growth, as illustrated in Fig. 4 using uni-axial tension as an example. Therefore, eqn (2.2), (2.5) and (2.7) to calculate G based on the assumptions of elasticity and SSY are not automatically applicable to tough gels. The detailed loading history of each material element near the crack tip has to be examined. This is not an easy task given the highly amplified and multi-axial deformation field near the crack tip. In the following we shall call G the effective energy release rate since a part of it is dissipated by hysteresis and the rest is released to perform the local fracture process. Next we review two scenarios for crack growth in tough gels that have been used in the literature, and examine the interpretation of fracture toughness in a rigorous manner.

3.1 Crack initiation for rate independent gels

One way for measuring the toughness Γ_c of tough gels is to start with a stationary crack in a PS sample (see Fig. 3A). A critical effective energy release rate G_c can be calculated using eqn (2.2) by measuring the applied stretch ratio λ_c at the onset of crack growth. The function $\overline{W}(\lambda_c)$ should now be interpreted as the stress work done per unit volume, instead of strain energy density, to stretch an uncracked sample to λ_c under the pure shear constraint. This critical effective energy release rate G_c is then taken as the fracture toughness Γ_c .

However, a closer examination reveals that eqn (2.2) is not applicable to crack initiation. Eqn (2.2) is based on the assumption of translational invariance, which means that as the crack grows, the stress and deformation fields in the sample remain

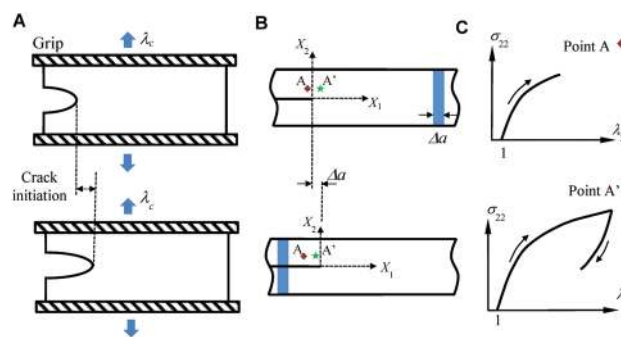


Fig. 5 Crack initiation in inelastic gels. (A) The deformed crack in the PS configuration at the critical stretch ratio λ_c before (top) and after (bottom) crack initiation. (B) The corresponding undeformed configurations before (top) and after (bottom) initiation. After crack growth, the coordinates of point A' with respect to the translating coordinate system X_1 – X_2 is the same as those of point A before crack growth. (C) Stress–strain states at point A before crack growth (top) and at A' after crack growth (bottom). Note that both points are subjected to complex multi-axial stress–strain states. Here we schematically illustrate the dominant stress and strain components, *i.e.* nominal normal stress σ_{22} and stretch ratio λ_2 .

unchanged with respect to a translating coordinate system that is always centered at the crack tip. If this assumption holds, the work done to grow the crack length by an increment Δa is equivalent to moving a strip with width Δa far ahead of the crack tip (stretched to λ_c) to far behind of the crack tip (completely relaxed; see Fig. 5B), and hence eqn (2.2) is derived. The translation invariance breaks down for crack initiation in tough gels since the amount of crack growth is exactly zero at initiation. Moreover, translation invariance is violated even if the crack grows straightly ahead. This point is illustrated in Fig. 5. Before crack initiation, a material point A behind the crack tip (X_{1A} , X_{2A}) is monotonically loaded as the applied stretch λ increases. After the crack grows by Δa , the material point A' occupies the location (X_{1A} , X_{2A}) in the translating coordinate system X_1 – X_2 , as shown in Fig. 5B. However, the material point A' is first loaded before crack initiation and then unloaded once the crack grows. Because of the loading–unloading hysteresis, the point A' after crack growth does not possess the same stress and deformation state as that at point A before crack growth (see Fig. 5C). In principle, an extra term of the stress work accounting for the change in the crack-tip stress and deformation fields due to crack initiation needs to be included in G , but this term is difficult to compute in practice.

Now that the G_c calculated using eqn (2.2) is not the critical energy release rate, what is the theoretical nature of this quantity? Here we propose a justification using the J -integral based on the field approach. Before the onset of crack growth, the crack is stationary. If the specimen boundary is subjected to monotonic loading, then it is a reasonable assumption that this condition is valid for every material point; further, because of rate insensitivity, the material behavior is indistinguishable from a hyperelastic solid. This means that the near tip fields are uniquely characterized by the path independent J -integral. For elastic solids, the J -integral at the onset of crack growth is equal

to the G_c described above, which can be taken as a measurement of fracture toughness

$$J_{\text{initiation}} = G_c = \Gamma_c(\Delta c = 0). \quad (3.1)$$

Eqn (3.1) also applies to the inelastic gels as long as unloading does not occur anywhere in the sample before crack initiation. Therefore, G_c measured using crack initiation is not the energy release rate, but rather the critical value for the J -integral. The two approaches of fracture, energetic and field, are equivalent only for elastic solids, but not for the inelastic gels. Similar arguments can be extended to the other testing configurations reviewed in Section 2.

Energy dissipation does not fit into the above interpretation of crack initiation toughness, since this criterion does not involve any material unloading and hysteresis. However, even if the crack does not grow, there is undeniably material damage and experiments have shown that damage enhances crack initiation toughness.^{44,49} To reconcile the field approach for crack initiation with this physical reality, we propose that material damage can help reduce stress concentration at the stationary crack tip and thus delay fracture. To illustrate this point, we consider a class of hyperelastic material models, *i.e.* the generalized neo-Hookean model, with the following strain energy density function:

$$W = \frac{\mu}{2b} \left\{ \left[1 + \frac{b}{n}(I_1 - 3) \right]^n - 1 \right\} \quad (3.2)$$

where μ is the small strain shear modulus, b and n are material parameters which control strain hardening, and I_1 is the trace of the Cauchy-Green tensor. Specifically we compare two cases: (i) $n = 1$ where eqn (3.2) reduces to that of the neo-Hookean model; (ii) $n = 0.55$ and $b = 1$. As shown in Fig. 6A, Case II exhibits a uni-axial stress-strain curve that is flatter than Case I, which we intend to simulate the softened loading curve due to material damage. The crack tip stress field based on eqn (3.2) for a Mode-I plane stress crack has been solved in the literature^{82,83} and is rather complicated. To make our point, we consider the distribution of the nominal stress component σ_{22} directly ahead of the crack tip. We assume that this opening stress component controls the initiation of Mode-I cracks.

Using the results summarized in Long & Hui,⁴² we have

$$\text{Case I } (n = 1): \frac{\sigma_{22}}{\mu} = \left(\frac{J}{\mu\pi r} \right)^{0.5}; \quad (3.3)$$

$$\text{Case II } (n = 0.55, b = 1): \frac{\sigma_{22}}{\mu} = 0.83 \left(\frac{J}{\mu\pi r} \right)^{0.091},$$

where r is the distance to the crack tip in the undeformed configuration. Note that in practice, the singularity in σ_{22} as r approaches zero can be regularized by local failure mechanisms (*e.g.* such as breaking of the soft network) in the very vicinity of the crack tip. Eqn (3.3) reveals that to achieve the same stress level ahead of the crack tip at a certain r , a significantly larger J is needed for Case II which simulates the case with damage (see Fig. 6B). On the other hand, the vertical stretch λ_2 directly ahead of the crack tip is

$$\text{Case I } (n = 1): \lambda_2 = \left(\frac{J}{\mu\pi r} \right)^{0.5}; \quad (3.4)$$

$$\text{Case II } (n = 0.55, b = 1): \lambda_2 = 2.2 \left(\frac{J}{\mu\pi r} \right)^{0.91},$$

which means in Case II the material must accommodate a much larger deformation λ_2 at the crack tip as shown in Fig. 6C. This is consistent with the principle proposed by Zhao⁴³ that high extensibility of the damaged network is required to achieve enhanced toughness.

Note that our argument here is based on the implicit assumption that crack initiation is controlled by the nominal stress σ_{22} . Other criteria may be proposed, which may result in a different conclusion. For example, the true stress τ_{22} directly ahead of the crack tip follows the distribution below

$$\text{Case I } (n = 1): \frac{\tau_{22}}{\mu} = \frac{J}{\mu\pi r}; \quad (3.5)$$

$$\text{Case II } (n = 0.55, b = 1): \frac{\tau_{22}}{\mu} = 1.82 \frac{J}{\mu\pi r},$$

where Case II shows slightly larger τ_{22} than that of Case I. A crack growth criterion based on the true stress τ_{22} would suggest a less difference in crack initiation between these two materials. How to impose a physically based criterion for crack initiation in terms of the crack tip fields is a question yet to be

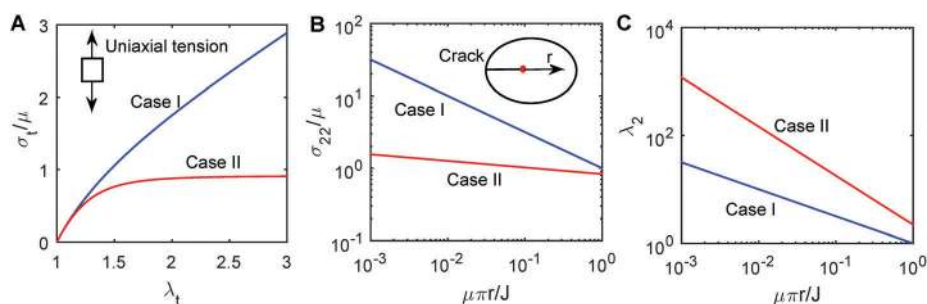


Fig. 6 (A) Uni-axial tensile response for Case I and II: normalized nominal tensile stress σ_t/μ versus stretch ratio λ_t . (B) The nominal stress σ_{22}/μ directly ahead of the crack tip where r is the distance from the crack tip in the undeformed configuration. (C) Stretch ratio λ_2 directly ahead of the crack tip.

answered, which will probably involve extensive experimental and multi-scale modeling efforts.

So far we have assumed monotonic loading for every material point before the onset of crack growth. A possible exception is when the gel exhibits necking behavior, as shown by the experiments of Na *et al.*⁸⁴ and Nakajima *et al.*⁸⁵ Necking is also captured by the model of Wang and Hong.²⁹ For these gels, unloading can occur locally before crack growth even if the external loading is monotonic. In this case, the use of eqn (3.1) may not be justified.

3.2 Steady state crack propagation for rate independent gels

When a crack starts to propagate in a PS specimen, it is reasonable to assume that it will eventually reach a steady state. Modeling the transition from initiation to steady state crack propagation is a difficult problem and has not been systematically studied. Hence we only consider the steady state crack propagation. For this case, the effective energy release rate G needed to drive steady-state crack propagation is well defined and can be easily computed. Therefore, it is convenient to use the energetic approach to define an effective fracture toughness Γ_c . It has been well recognized that most of the energy release rate is dissipated, and only a small fraction of it is used to drive the fracture process. In tough gels, the amount of dissipation is strongly tied to the local fracture process, and here we review this coupling to gain insight into the interplay between toughening and local failure mechanisms. The basic idea is simple: the actual amount of energy released to drive crack propagation is determined by subtracting the energy dissipation from the energy supplied by the loading device. This principle has been utilized in the literature to derive models for the toughness of DN gels.^{62–64} We attempt to state it in a rigorous manner. We will use the PS configuration as an example, but the same argument can be extended to other testing configurations where steady state crack propagation is possible, such as the simple extension (SE) test.

The steady state condition implies that the stress and strain fields everywhere in the sample are invariant with respect to an observer moving together with the crack tip. In contrast to the crack initiation test, here the translational invariance of stress and strain fields allows us to apply eqn (2.2). However, since the material is no longer elastic, the function $\overline{W}(\lambda_s)$ in eqn (2.2) needs to be interpreted as the stress work per unit undeformed volume to stretch the material to λ_s , where $\lambda_s = 1 + \Delta/H_0$ is the applied stretch for steady state crack propagation. In the PS geometry, this stress work is performed solely by σ_{22} , *i.e.* the nominal stress component in the X_2 direction (see Fig. 7B). Therefore, we obtain

$$G = 2H_0 \overline{W}(\lambda_s) = 2H_0 \int_1^{\lambda_s} \sigma_{22} d\lambda_2, \quad (3.6)$$

For tough gels, most of G is consumed by energy dissipation. To compute the actual energy available to drive the crack, consider a material element in the slab located at $X_2 = Y$ (B in Fig. 7C). The steady state condition implies that the deformation history of B can be determined by fixing the crack at the

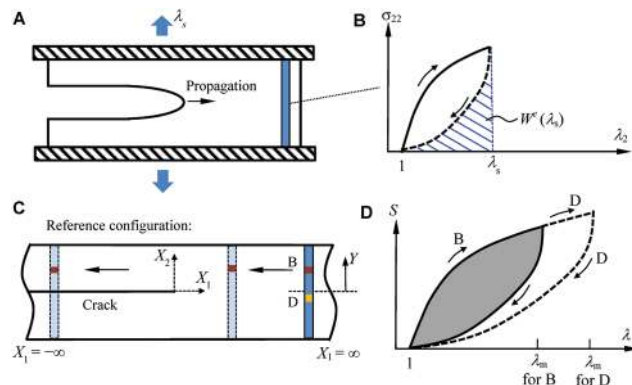


Fig. 7 (A) Deformed configuration for steady-state crack propagation in the PS geometry. (B) The stress–strain relation of the slab far ahead of the crack tip. The stress work $\overline{W}(\lambda_s)$ is the area under the loading curve (solid line), while the stored strain energy density $W^e(\lambda_s)$ is shown as the shaded area. (C) Energy dissipation can be computed for imaginarily moving a slab from $X_1 = \infty$ to $-\infty$. (D) The loading histories of points B and D as they are translated from $X_1 = \infty$ to $-\infty$. The shaded area is the energy dissipation density at point B. Point D is closer to $Y = 0$, and it can reach to higher maximum stretch λ_m than point B, which means larger dissipation density at point D.

origin and translating this element from $X_1 = \infty$ all the way to $X_1 = -\infty$. Since the gel is rate independent, the speed of translation has no effect on the mechanics. As this element moves closer to the crack tip, it undergoes additional loading or unloading. This additional load is multi-axial, *e.g.* the element is subjected to shear and can rotate and stretch in different directions. When a material element starts to unload depends on its vertical position Y . For example, the material elements located near $Y = H_0$ or $-H_0$ start to unload earlier than those near $Y = 0$. Indeed, the strain energy released by unloading the former material elements (near $Y = H_0$ or $-H_0$) is used to further load those near $Y = 0$, since the applied displacement Δ remains fixed. Nevertheless, all the material elements at any Y should be completely relaxed when they reach $X_1 = -\infty$, provided that the material does not retain any permanent or plastic deformation after unloading. We emphasize that if plastic deformation exists, the material elements are under residual stresses when they are translated to far behind the crack tip. This scenario is not hypothetical; Yu *et al.*⁷⁵ observed a wake of damaged material trailing behind the crack in their tearing experiments.

We emphasize that initially when the slab ($-H_0 \leq Y \leq H_0$) is located at $X_1 = \infty$, only part of the stress work density $\overline{W}(\lambda_s)$ is stored in the material as elastic strain energy density, this is denoted by $W^e(\lambda_s)$ and illustrated by the shaded area in Fig. 7B. As the slab moves from $X_1 = \infty$ to $X_1 = -\infty$, deformation in the slab is redistributed by the process described above, which further dissipates energy. Therefore, the energy release rate available to the crack tip is in general less than $2H_0 W^e(\lambda_s)$.

Assisted by the above picture, we can compute the energy dissipation for a material element located at Y by following its loading history, which consists of two parts. First, a material element starts at an initially relaxed state (zero stress and strain) and then stretched to the deformation state at $X_1 = \infty$.

For the PS configuration, this deformation is characterized by $\lambda_1 = 1$, $\lambda_2 = \lambda_s$, and $\lambda_3 = 1/\lambda_s$. Second, the material element undergoes loading and unloading as it translates from $X_1 = \infty$ to $X_1 = -\infty$. We denote the energy dissipation per unit volume by $W_{-\infty}(Y)$ which can be written as

$$W_{-\infty}(Y) = \int_{\gamma(Y)} \sigma_{ij} dF_{ij}, \quad (3.7)$$

where F_{ij} is the deformation gradient tensor and σ_{ij} is the nominal stress. Here $\gamma(Y)$ represents the aforementioned deformation history of the material element at Y . The loading path $\gamma(Y)$ depends on Y . Summing the energy dissipation along Y , we obtain the following expression for the energy dissipated per unit crack extension, G_D :

$$G_D = 2 \int_0^{H_0} W_{-\infty}(Y) dY. \quad (3.8)$$

The energy flow to the crack tip per unit crack extension is denoted as the local energy release rate G_{tip} , which is defined by

$$G_{\text{tip}} = G - G_D. \quad (3.9)$$

For tough gels, during crack propagation $G \gg G_{\text{tip}}$ which implies $G_D \gg G_{\text{tip}}$. If we make the additional assumption that the local failure process is also rate independent, then G_{tip} for steady state crack growth must be a constant independent of crack speed. Thus, the crack growth condition is $G_{\text{tip}} = \Gamma_{\text{intrinsic}}$. Here it is important to note that the mechanics of crack growth is controlled by the intrinsic toughness $\Gamma_{\text{intrinsic}}$, which is much smaller than the effective fracture toughness Γ_c . In experiments, it is Γ_c and not $\Gamma_{\text{intrinsic}}$ that can be measured. Yet a small increase in $\Gamma_{\text{intrinsic}}$ can substantially increase the effective toughness, as shown in the next section.

The above formulation is exact as long as the steady state condition is satisfied (whether or not a plastic wake exists). However, to evaluate G_D , one must specify a multi-axial material model which may not be readily available, impose a local failure condition, and then solve the stress and deformation fields in the crack sample using the full set of continuum equations – a non-trivial problem that requires numerical solutions. Nevertheless, the basic idea can be illustrated by approximating the complex multi-axial deformation as the slab is translated from $X_1 = \infty$ to $-\infty$ by a simple constrained tension state, *i.e.* with $\lambda_1 = 1$, $\lambda_2 = \lambda$, $\lambda_3 = 1/\lambda$. The corresponding nominal tensile stress σ_{22} is denoted as S . The loading history $\gamma(Y)$ corresponding to this approximation can be described as follows. The material element is first stretched from $\lambda = 1$ to λ_s . As it is translated from $X_1 = \infty$ to $-\infty$, λ further increases to some maximum value λ_m and is then unloaded to 1 (here we assume for simplicity that there is no plastic wake). Since the maximum stretch λ_m is a function of the vertical coordinate of the material element Y , eqn (3.9) becomes

$$G_{\text{tip}} = G - G_D = 2H_0 \int_1^{\lambda_s} S^+ d\lambda - 2 \int_0^{H_0} U(\lambda_m(Y)) dY, \quad (3.10)$$

where $U(\lambda_m) = \int_0^{\lambda_m} S^+ d\lambda + \int_{\lambda_m}^0 S^- d\lambda$ is the area of the hysteresis loop in the stress–strain diagram under constrained tension

(see Fig. 7D). The “+” and “−” signs in the superscript of S denote the loading and unloading branches, respectively. The function $\lambda_m(Y)$ in eqn (3.10) depends on the sample geometry, loading configuration, bulk material behavior as well as the intrinsic toughness $\Gamma_{\text{intrinsic}}$, and is difficult to determine. To determine $\lambda_m(Y)$, a full-field solution of the stretch ratio $\lambda(X_1, X_2)$ in the translating coordinate system (X_1, X_2) is needed. A possible approximation to simplify eqn (3.10) is to replace $\lambda_m(Y)$ by its average value λ_m^a in the slab $-H_0 \leq Y \leq H_0$, and treat λ_m^a as a fitting parameter to match experimental data. Alternatively, eqn (3.10) can be simplified based on experimental observations (*e.g.* for the DN gels as discussed in Section 3.3). Finally, it is important to note that G_D in general depends on the local failure process as well as the height H_0 of the PS specimen. This implies that dissipation can be influenced by the specimen size. Additional assumptions are needed to render G_D independent of the specimen size, *e.g.* damage and hysteresis only occur in a subset of the slab $-H_0 \leq Y \leq H_0$. This will be illustrated in the next section.

3.3 Models of fracture toughness in DN gels

Two theoretical models connecting the toughening mechanism of DN gels to fracture toughness were proposed independently by Brown⁶² and Tanaka⁶³ in 2007. These two models are very similar. They both obtained expressions for the effective toughness for steady state crack growth. The main difference is how they estimate the size of the damage zone, and this leads to very different scaling for the toughness. Here we review their analyses in some detail.

Both Brown and Tanaka considered steady state propagation in the PS configuration. For simplicity, the tensile behavior of an unnotched sample is idealized as that shown in Fig. 8A. Strictly speaking, in this tensile behavior the normal strain in the X_1 direction should be constrained to be compatible with pure shear geometry. The DN gel is assumed to exhibit a yielding behavior. When the stress is below a critical yield stress σ_a , the gel is linear elastic with a Young's modulus of E_1 that is mostly governed by the stiff network. When the stress

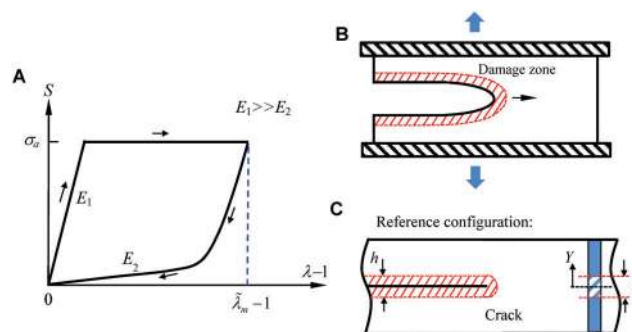


Fig. 8 (A) Idealized tensile stress–strain curve for the DN gels. (B) Damage zone surrounding the crack in the deformed configuration. (C) Illustration of the damage zone in the reference configuration. The thickness of the damage is h . This means, according to our formulation, in the vertical slab ahead of the crack tip only material points with $-h/2 \leq Y \leq h/2$ will be damaged.

reaches σ_a , the first stiff network is progressively damaged until a maximum strain of $\tilde{\lambda}_m - 1$ where the first network is completely broken. After that the gel again behaves as a linear elastic solid but with a much smaller modulus E_2 that is governed by the second soft network. The critical stress σ_a was identified to be approximately the necking stress of DN gels in a tensile test of an unnotched specimen. Brown⁶² estimated σ_a to be ~ 0.3 MPa for DN gels which is close to the experimentally measured value (~ 0.2 MPa⁸⁴ or 0.7 MPa⁵⁷). The maximum strain $\tilde{\lambda}_m - 1$ depends on the relative volume ratio of the first network, and is taken as a material parameter. In Brown⁶² and Tanaka,⁶³ this maximum strain is estimated to be ~ 10 based on the experimental observation.

With this simplified bulk material behavior for the DN gels, both Brown⁶² and Tanaka⁶³ envisioned that during crack propagation a damage zone surrounding the crack will be formed where the first stiff network is completely broken. Within this damage zone, every material point has experienced the maximum strain $\tilde{\lambda}_m - 1$ and hence behaves as a linear solid with Young's modulus E_2 . While outside the damage zone, the material is assumed to be undamaged with modulus E_1 . The damage zone was assumed to be in the shape of a strip surrounding the crack surface with a width of h in the reference configuration. According to Fig. 8C, the material elements inside $-h/2 \leq Y \leq h/2$ are subjected to damage. This definition is consistent with that in Brown.⁶² In Tanaka,⁶³ the definition of h is not as clear. It appears to denote half of the damage zone width in the deformed state in Fig. 1 and 2 of Tanaka,⁶³ but then h seems to represent the full damage zone width when the dissipation energy is estimated (see eqn (1) of Tanaka⁶³).

Based on the assumptions above, the following expression for the effective energy release rate was derived.

$$G \approx \sigma_a(\tilde{\lambda}_m - 1)h. \quad (3.11)$$

To place this result into our theoretical framework described in Section 3.2, specifically eqn (3.10), we first note that for DN gels, $G \gg G_{\text{tip}}$ and thus

$$G \approx G_D = 2 \int_0^{H_0} U(\lambda_m(Y)) dY. \quad (3.12)$$

Outside the damage zone (*i.e.* $h/2 < Y \leq H_0$ and $-H_0 \leq Y < -h/2$), there is no bulk material dissipation which implies $U(\lambda_m(Y)) = 0$. Within the damage zone ($-h/2 \leq Y \leq h/2$), $\lambda_m(Y)$ is uniform and is equal to $\tilde{\lambda}_m$. In addition, we neglect $\int_{\tilde{\lambda}_m}^0 S^- d\lambda$ in comparison to $\int_0^{\tilde{\lambda}_m} S^+ d\lambda$ when computing $U(\lambda_m(Y))$ in light of Fig. 8A. Therefore, we have

$$G = 2 \int_0^{h/2} U(\tilde{\lambda}_m) dY \approx h \int_0^{\tilde{\lambda}_m} S^+ d\lambda \approx \sigma_a(\tilde{\lambda}_m - 1)h. \quad (3.13)$$

The next step is to estimate the damage zone size, h . Since the damage zone is assumed to be a long strip with modulus E_2 , it is modeled as a pure shear specimen with a traction free crack. Failure occurs when the crack grows in this strip, that is, when the local energy release rate G_{tip} is equal to $\Gamma_{\text{intrinsic}}$, which is taken to be the fracture toughness of the second

network (soft). It is the calculation of this energy release rate G_{tip} that the two models depart from each other. Specifically, Brown⁶² imposed a displacement condition on the boundary of the damage zone ($-h/2 \leq Y \leq h/2$) whereas Tanaka⁶³ imposed a traction condition of σ_a on this boundary. The local energy release rate in Brown's model was calculated to be

$$G_{\text{tip}}^{\text{Brown}} = E_2(\tilde{\lambda}_m - 1)^2 h/2. \quad (3.14)$$

The macroscopic crack grows when the second network is fractured, at which h reaches the maximum value h_{max} governed by $G_{\text{tip}}^{\text{Brown}} = \Gamma_{\text{intrinsic}}$. Using eqn (3.14), h_{max} is

$$h_{\text{max}} = \frac{2\Gamma_{\text{intrinsic}}}{E_2(\tilde{\lambda}_m - 1)^2}. \quad (3.15)$$

Note that in Brown's model, the traction acting on the boundary of the damaged strip (*i.e.* at $Y = \pm h/2$) is $E_2(\tilde{\lambda}_m - 1)$ and not σ_a . At first glance, this seems to contradict the fact that σ_a is needed to maintain the damaged material in equilibrium with the material outside the damage zone. However, if $E_2(\tilde{\lambda}_m - 1) = \sigma_a$, then eqn (3.13) and (3.15) imply that

$$G = 2\Gamma_{\text{intrinsic}}, \quad (3.16)$$

and the toughening due to damage of the first network is lost. In fact, toughening is due to $E_2(\tilde{\lambda}_m - 1) \ll \sigma_a$. This means that the strip must suddenly unload from σ_a to $E_2(\tilde{\lambda}_m - 1)$ once λ reaches $\tilde{\lambda}_m$. In Brown's model,⁶² this sudden unloading provides the energy to drive fracture, and he envisions that the material unloads immediately as the crack grows, as illustrated by the red dotted line in Fig. 9A. The area of the orange triangle is the energy release rate given by eqn (3.14). In contrast, in Tanaka's model, σ_a is prescribed on the strip; the energy available for fracture is due to elastic unloading with slope E_u , *i.e.*,

$$G_{\text{tip}}^{\text{Tanaka}} = \frac{\sigma_a^2 h}{2E_u}, \quad (3.17)$$

where $G_{\text{tip}}^{\text{Tanaka}}$ is the area of triangle highlighted in blue (see Fig. 9B). Similarly, by setting $G_{\text{tip}}^{\text{Tanaka}} = \Gamma_{\text{intrinsic}}$, the thickness of the damage zone is obtained as follows

$$h_{\text{max}} \approx \frac{2E_u \Gamma_{\text{intrinsic}}}{\sigma_a^2}. \quad (3.18)$$

Substituting eqn (3.15) and (3.18) into eqn (3.13), we obtain the ratio of the effective fracture toughness of the composite network, Γ_c , to the fracture toughness of the second network, $\Gamma_{\text{intrinsic}}$:

$$\frac{\Gamma_c}{\Gamma_{\text{intrinsic}}} = \begin{cases} \frac{2\sigma_a}{E_2(\tilde{\lambda}_m - 1)} & \text{Brown}^{62} \\ \frac{2E_u(\tilde{\lambda}_m - 1)}{\sigma_a} & \text{Tanaka}^{63} \end{cases}. \quad (3.19)$$

The amplification ratios in eqn (3.19) were estimated to be ~ 40 in Brown⁶² and ~ 10 in Tanaka,⁶³ which can be considered to be consistent on the order of magnitude. Note that the ratios

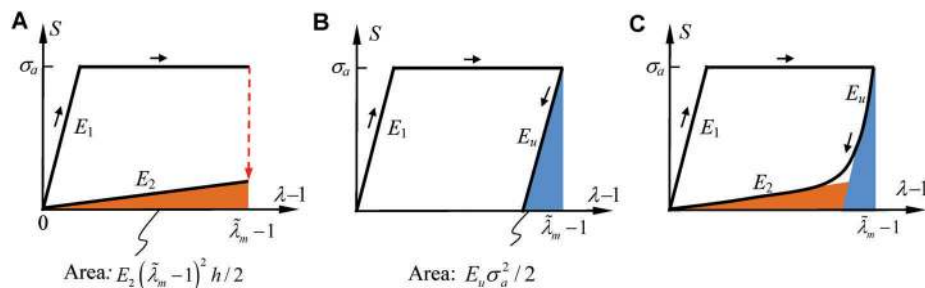


Fig. 9 (A) The method of Brown⁶² to estimate G_{tip} . The red arrow indicates a sudden unloading from σ_a to $E_2(\tilde{\lambda}_m - 1)$. (B) The method of Tanaka⁶³ to estimate G_{tip} . (C) The two models of Brown and Tanaka can be reconciled by calculating the area underneath the entire unloading curve, which is approximately the sum of the areas of the two triangles adopted by Brown⁶² and Tanaka.⁶³

given in eqn (3.19) are large. Therefore, a small increase in $\Gamma_{\text{intrinsic}}$ can lead to a large increase in effective toughness Γ_c . Although both predictions are in reasonable agreement with experimental data, the scaling is clearly different. For example, the amplification ratio in Brown's model increases with σ_a whereas the opposite is true for Tanaka's model.

One way of reconciling this difference is to note that the actual unloading curve is better represented by the one shown in Fig. 9C. Therefore, a more consistent estimate of G_{tip} can be calculated by summing the areas of the two triangles adopted by Brown⁶² and Tanaka⁶³ (see Fig. 9C) and then multiplying by h , that is,

$$G_{\text{tip}} = \left[\frac{\sigma_a^2}{2E_u} + \frac{E_2(\tilde{\lambda}_m - 1)^2}{2} \right] h, \quad (3.20)$$

where E_u is the unloading modulus (which could be different from E_1 due to hardening, see Tanaka⁶³). Enforcing the fracture condition, $G_{\text{tip}} = \Gamma_{\text{intrinsic}}$, we have

$$h_{\text{max}} = \frac{2E_u\Gamma_{\text{intrinsic}}}{\sigma_a^2 + E_uE_2(\tilde{\lambda}_m - 1)^2}. \quad (3.21)$$

Therefore, eqn (3.19) becomes:

$$\frac{\Gamma_c}{\Gamma_{\text{intrinsic}}} = \frac{2E_u\sigma_a(\tilde{\lambda}_m - 1)}{\sigma_a^2 + E_uE_2(\tilde{\lambda}_m - 1)^2}. \quad (3.22)$$

The result of Brown⁶² is recovered when $E_u \rightarrow \infty$, while the result of Tanaka⁶³ is recovered when $E_2 \rightarrow 0$.

These two models captured the qualitative toughening mechanism of DN gels. However, a more detailed analysis is needed for accurate quantitative prediction of the fracture toughness. This was achieved in a recent work reported by Zhang *et al.*⁶⁴ They considered steady-state crack propagation under PS configuration and derived the following scaling relation between the effective and local energy release rates:

$$G = \frac{G_{\text{tip}}}{1 - \alpha R_{\text{max}}}, \quad (3.23)$$

where α is an unknown positive coefficient and was determined from finite element simulation. The factor $1/(1 - \alpha R_{\text{max}}) > 1$

represents the amplification ratio of G_{tip} , where R_{max} is the maximum value of the hysteresis ratio R_m defined as

$$R_m = \frac{\int_0^{\lambda_m} S^+ d\lambda + \int_{\lambda_m}^0 S^- d\lambda}{\int_0^{\lambda_m} S^+ d\lambda}, \quad (3.24)$$

i.e. the ratio of dissipated energy *versus* the total stress work done on a material element in a loading–unloading cycle. The hysteresis ratio R_m ($0 \leq R_m < 1$) is a function of the peak stretch λ_m that a material element has experienced. R_{max} in eqn (3.23) is the value of R_m associated with the maximum stretch λ_{max} experienced by all material elements in the slab ($-H_0 \leq Y \leq H_0$).

To make a connection between eqn (3.23) and the framework outlined in Section 3.2, we first write eqn (3.10) in the following form using the definition of R_m :

$$G_{\text{tip}} = G - G_D = 2H_0 \int_1^{\lambda_s} S^+ d\lambda - 2 \int_0^{H_0} \left(R_m \int_1^{\lambda_m} S^+ d\lambda \right) dY. \quad (3.25)$$

Although λ_m is a function of Y and $\lambda_{\text{max}} = \max(\lambda_m)|_{-H_0 \leq Y \leq H_0}$, scaling-wise we can write eqn (3.25) as

$$G_{\text{tip}} = G - G_D \approx 2H_0 \left(\int_1^{\lambda_s} S^+ d\lambda - \eta R_{\text{max}} \int_1^{\lambda_{\text{max}}} S^+ d\lambda \right), \quad (3.26)$$

where $0 < \eta < 1$ is a parameter accounting for the variation of $\lambda_m(Y)$ where λ_{max} is the maximum stretch experienced by all points in the slab ($-H_0 \leq Y \leq H_0$). This maximum stretch should occur at material elements close to $Y = 0$ due to stress concentration at the crack tip. To drive the additional loading of the material elements around $Y = 0$ as they are translated from $X_1 = \infty$ to $X_1 = -\infty$, stress work is needed which comes from strain energy released by unloading the material elements closer to $Y = \pm H_0$. In other words, the level of maximum stretch λ_{max} is set by the external load λ_s . If we can assume λ_{max} and λ_s are related such that $\int_1^{\lambda_{\text{max}}} S^+ d\lambda$ is proportional to $\int_1^{\lambda_s} S^+ d\lambda$, *i.e.*,

$$\int_1^{\lambda_{\text{max}}} S^+ d\lambda = \beta \int_1^{\lambda_s} S^+ d\lambda, \quad (3.27)$$

then eqn (3.26) can be reduced to:

$$G_{\text{tip}} = (1 - \eta\beta R_{\text{max}})2H_0 \int_1^{\lambda_s} S_{22} d\lambda = (1 - \eta\beta R_{\text{max}})\Gamma, \quad (3.28)$$

which reproduces the result in Zhang *et al.*⁶⁴ (see eqn (3.23)) if η and β are combined into a single parameter α . Our derivation of eqn (3.28) relies on eqn (3.27) which makes a strong assumption on how deformation is distributed in the PS specimen. To verify its validity, a detailed solution of the deformation fields during steady-state crack propagation is needed, which can be achieved by finite element simulations. This was carried out by Zhang *et al.*⁶⁴ using the neo-Hookean material with a modified Ogden–Roxburgh model⁸⁶ to capture damage. It would be useful to test the validity of eqn (3.27) for other damage models, *e.g.* the pseudo-elastic model by Wang and Hong.²⁹ Nevertheless, the scaling law given by eqn (3.23) was shown to successfully predict the critical load for steady-state propagation in a hybrid polyacrylamide–alginate tough hydrogel⁴⁴ under the PS configuration.⁶⁴

We emphasize that the scaling analysis in Zhang *et al.*⁶⁴ is for steady state crack propagation. It may not be applied to crack initiation where the energetic approach is not applicable for inelastic materials. Indeed, Zhang *et al.*⁶⁴ made an excellent point on the difference between crack initiation and propagation. Using a finite element model, they showed that for inelastic materials with strong loading–unloading hysteresis, the critical stretch to achieve steady-state crack propagation is much larger than that needed for crack initiation, whereas for elastic materials the critical stretch for crack propagation and initiation is approximately the same. This is consistent with the points we made in Section 3.1.

It should be noted that the results of Brown⁶² and Tanaka⁶³ (eqn (3.19)) are stated in terms of the fracture toughness Γ_c and $\Gamma_{\text{intrinsic}}$. The amount of energy dissipation is mainly determined by when the second network fractures, *i.e.*, by $\Gamma_{\text{intrinsic}}$. In contrast, the scaling relation of Zhang *et al.*⁶⁴ (eqn (3.23)) is for the energy release rates G and G_{tip} . In this case, the energy dissipation G_D is determined by G , *i.e.*, how large the external loading is to drive the loading–unloading process and hence dissipation. These two approaches can be reconciled if one applies the fracture criterion: $G_{\text{tip}} = \Gamma_{\text{intrinsic}}$.

Finally, it is worth mentioning that Zhang *et al.*⁶⁴ proposed that $\Gamma_{\text{intrinsic}}$ can be measured by preloading the gel before a crack is introduced. In this way, the gel is already damaged before the fracture test, and thus the dissipation term G_D can be reduced or even vanish. They measured the fracture toughness for an interpenetrating polyacrylamide–alginate gel using the PS geometry but with different degrees of pre-stretch. It was found that the toughness Γ_c converged to a value of 400 J m^{-2} for sufficiently large pre-stretch. This was taken as the value of $\Gamma_{\text{intrinsic}}$. However, this value is much larger than the values assumed by Brown⁶² and Tanaka⁶³ ($\sim 10 \text{ J m}^{-2}$). Note that the value of $\Gamma_{\text{intrinsic}} \sim 10 \text{ J m}^{-2}$ is consistent with the prediction of the Lake–Thomas theory for simple gels.⁴³ Correspondingly, the amplification ratio found in the experiment of Zhang *et al.*⁶⁴ (~ 2.7) is much less than that predicted by Brown⁶² (~ 40) and Tanaka⁶³ (~ 10).

4. Rate dependent fracture of gels

4.1 Physical mechanisms of rate dependence

In practice, many gels exhibit rate dependent fracture behaviors.^{36,77,87–92} Because of rate dependence, the onset of

crack growth in a crack initiation test can be sensitive to the applied strain rate. In a crack propagation test, the effective fracture toughness Γ_c depends on the crack propagation speed. The physical mechanisms responsible for such a rate dependence are complex and multi-faceted due to the rich mechanical behaviors of gels. Here we summarize three main mechanisms: (i) rate dependent crack-tip fracture process; (ii) poroelasticity; and (iii) viscoelasticity.

Rate dependent crack tip fracture process. A representative example for the rate dependent crack-tip fracture process is the gelatin gel studied by Baumberger *et al.*^{32,33,68,90} Gelatin gels consist of physically associated networks swollen by solvent molecules (*e.g.* water) and exhibit much lower toughness (on the order of 10 J m^{-2}) than that of DN gels. The fracture experiments of Baumberger *et al.*^{32,33} suggested that the fracture process is localized in a small zone surrounding the crack tip, in which chains creep due to viscoplastic disentanglement until they are pulled out of the gel matrix. Outside this region the gel is elastic. This chain pull-out process is resisted by the viscous drag of solvents. For example, Baumberger *et al.*^{32,33,90} observed that their gelatin gels are toughened by increasing solvent viscosity. In addition, the toughness can decrease significantly if a small drop of the solvent is introduced to wet the moving crack tip,³² indicating the effect of viscous flow of the solvent on the crack tip fracture process. Wetting the crack tip can also lead to crack branching instabilities.⁶⁸

Recently Lefranc and Bouchaud⁹¹ show that the toughness of agar gels ($\sim 1 \text{ J m}^{-2}$) increases with crack speed, despite that rheological tests suggest that the agar gels are mostly elastic with rate-insensitive storage and loss modulus in a wide frequency range (0.1–100 Hz). This rate-dependence of toughness was also attributed to viscous chain pull-out or stress-accelerated chain dynamics, both localized at the crack tip region. Since these physical gels are quite brittle, the fracture process is highly localized at the crack tip and the SSY condition is satisfied. Therefore, the full power of elastic fracture mechanics can be used to understand the fracture of these brittle gels.

Poroelastic effects. Poroelasticity refers to the coupling between solvent diffusion in the polymer network and mechanical stresses.^{23–26} Solvent diffusion can disturb the crack tip stress field and local fracture mechanisms, thus altering the fracture condition. The rate of solvent diffusion is often characterized by an effective diffusion coefficient D , which depends on the permeability of the network, solvent viscosity and gel modulus. This allows one to estimate a length scale R_p to characterize the size of a region around the crack tip within which the poroelastic effect can be important. For the crack initiation test with loading time t_e , this length scale is estimated by $R_p \sim \sqrt{Dt_e}$,³⁸ whereas for steady state crack propagation with speed V , R_p is estimated as D/V .^{93,94} If R_p is small compared to the characteristic geometrical dimensions then the poroelastic effect may be lumped into the intrinsic toughness by appealing to the SSY condition. However, if R_p is large (*e.g.* due to slow loading or slow crack propagation), poroelasticity needs to be explicitly considered in the fracture formulation. Recently, Bouklas *et al.*³⁹ developed a theoretical framework that accounts

for poroelasticity in gel fracture for a special case where the crack growth is assumed to be much slower than the solvent diffusion. Based on this assumption, Bouklas *et al.*³⁹ found that the energy release rate for a slowly propagating crack is given by the J -integral in eqn (2.1) modified by an additional term to account for the dissipation due to solvent diffusion.

Fracture in poroelastic solids is sensitive to the boundary condition for solvent transport at the crack surfaces, *i.e.* whether the crack surfaces are immersed in a reservoir of solvents. Bouklas *et al.*³⁹ showed that for a stationary crack under a fixed displacement loading, the intensity of the crack tip field, as reflected by their modified J -integral, increases with time if the crack surfaces are exposed to air and thus have zero solvent flux. This is consistent with a previous model by Wang and Hong⁹⁵ to explain the delayed fracture in gels.⁹⁶ If the crack surfaces are immersed in solvents, the crack tip field evolves differently and its intensity may even decrease with time, a feature also found in other works.^{38,97}

Viscoelastic effects. Viscoelasticity is an important time dependent mechanism that can substantially enhance the toughness of gels.^{98–100} Here we focus on viscoelasticity due to reversible physical bonds.^{19–22,101} These reversible bonds have been introduced into gels as sacrificial bonds to enhance toughness and yet allow the gels to heal after damage.^{44,49} Since physical bonds can reform after breakage, upon unloading and resting the gels can partially or fully recover to their original state depending on the resting period, reminiscent of a classic viscoelastic solid. More importantly, the viscoelastic behavior can be tuned using different types of physical crosslinks. For example, Grindy *et al.*¹⁰² have recently demonstrated how to tune the relaxation spectrum of a hydrogel by using multiple kinetically distinct metal-ligand crosslinks. The hypothesis that macroscopic viscoelastic behavior is controlled by bond breaking and healing kinetics has been tested by comparing the mechanical behavior of a poly(vinylalcohol) (PVA) dual crosslinked gel with a theory capturing the bond kinetics.^{21,22} This theory can accurately capture experimental data of uni-axial tension^{21,22} and torsion tests¹⁰³ subjected to complex loading histories. This dual-crosslinked gel has a simple chemical structure, which is ideal as a model system to probe the connection between microscopic bond kinetics and macroscopic viscoelasticity. However, the fracture toughness of this gel is much lower than gels with more complex molecular structures,^{49,104} and quantitative models for the rate-dependent mechanical behavior of these tougher gels are yet to be developed.

For rate-dependent gels, especially those toughened by a large amount of viscoelasticity, it is difficult to design and interpret fracture tests, since the material can undergo very large deformation in addition to exhibiting hysteretic behaviors that depend on the loading history and the local failure process. Although there have been experiments studying crack growth in viscoelastic gels with reversible physical bonds,^{49,77} researchers are still struggling to quantify the fracture behavior of these gels. Here we outline some of these challenges.

4.2 Crack initiation for viscoelastic gels

Crack initiation tests involving viscoelastic gels are difficult to interpret. First, the energy release rate is not well defined. Since

the amount of loading and unloading hysteresis depends on the loading history, it is difficult to tally the work and energy dissipation at every material point, similar to the discussion in Section 3.1 for materials with rate-independent damage. Second, although one can still measure an “initiation toughness” Γ_c^{init} following the procedures described in Section 3.1, the J -integral based interpretation of the initiation toughness Γ_c^{init} (see Section 3.1) is no longer valid for viscoelastic materials even if one assumes monotonic loading throughout the fracture specimen. Indeed, for rate dependent materials, the J -integral is not path independent. This is because at a given global loading rate, *e.g.* $\dot{\lambda}_s$ in the PS configuration, the local strain rates at individual material points around the crack tip are amplified non-uniformly. Such non-uniformity in local strain rates caused by stress concentration may depend on the specimen geometry, *e.g.* the size and shape of the specimen. Thus, the material points in a fracture specimen follow different stress-strain relations due to viscoelasticity. Therefore, it is in general not possible to find a nonlinear elastic solid to reproduce the same stress and strain fields in a viscoelastic fracture specimen. Without the J -integral based interpretation, Γ_c^{init} does not necessarily reflect the intensity of crack tip fields in viscoelastic materials, and should not be used as a criterion for crack initiation. For example, such Γ_c^{init} is likely to be dependent on the loading history and specimen geometry.

Even in special cases where one can apply the J -integral interpretation for viscoelastic materials, the critical load at the onset of crack growth, and hence Γ_c^{init} , will still be sensitive to the loading history. For example, take the simplest case of a linear viscoelastic solid, and consider two different loading histories. In history 1, the sample is subjected to very fast loading rate so the material has no time to relax. In this case the material behaves like a linear elastic solid with instantaneous Young's modulus E_0 . The J integral for this case is well-defined (since material is linearly elastic)

$$J = K_I^2/E_0 \quad (\text{fast loading, plane stress}). \quad (4.1)$$

On the other hand, if the sample is loaded very slowly so complete relaxation occurs throughout the sample (history 2), then the material behaves like a linear elastic solid with the long time Young's modulus E_∞ . For this case the J integral is also well-defined and is given by

$$J = K_I^2/E_\infty \quad (\text{slow loading, plane stress}). \quad (4.2)$$

According to eqn (4.1) and (4.2), a crack initiation criterion based on $J = \Gamma_c^{\text{init}}$ will lead to very different stress intensity factors at initiation, for the fast loading (history 1), this is:

$$K_{\text{Ic}}^{\text{fast}} = \sqrt{E_0 \Gamma_c^{\text{init}}}, \quad (4.3)$$

whereas for the slow loading case,

$$K_{\text{Ic}}^{\text{slow}} = \sqrt{E_\infty \Gamma_c^{\text{init}}}. \quad (4.4)$$

Since typically $E_0/E_\infty \gg 1$, the crack initiation criterion $J = \Gamma_c^{\text{init}}$ leads to the physically unreasonable result that the crack tip stress level at initiation is very different for fast and slow loading.

Note that we have assumed a simple loading history above, *i.e.* monotonic loading with a constant strain rate $\dot{\lambda}_s$. One can impose different loading histories such as cyclic loading, stress relaxation under a fixed displacement, or creep under a fixed force. The “toughness” or “ Γ_c^{init} ” measured at the onset of crack growth is expected to vary depending on the loading history.

In summary, the concept of fracture toughness, either based on the energetic or the J integral approach, is not well defined for crack initiation tests involving viscoelastic materials, since the onset of crack growth depends on the loading history, the local failure process, as well as specimen geometry. More studies are needed to establish a theoretically sound criterion for crack initiation in viscoelastic materials.

4.3 Crack propagation for viscoelastic gels

The inapplicability of the J integral approach to characterize crack initiation foreshadows the difficulty in quantifying crack propagation. The concerns on the loading history dependence discussed in Section 4.2 still remain. A way to bypass this difficulty is to restrict discussion to steady state crack propagation. In this case, locally around the crack tip, the loading histories and rates of material points are governed by the crack speed. As long as steady state crack growth exists, the effective energy release rate can be determined using the PS or SE specimens reviewed in Section 2. The problem is that steady state crack growth is difficult to achieve in practice since it requires special specimen geometries and boundary conditions. For example, in a PS specimen containing a rate independent gel, a propagating crack will eventually reach the steady state if the applied displacement Δ is held at some appropriate fixed value. However, in a viscoelastic gel, a fixed displacement Δ will cause the material far ahead of the crack tip to relax. Therefore, to drive crack growth at a fixed speed, it is necessary to move the grips apart at some unknown rate. This means that the strain far away from the crack tip increases with time – violating translation invariance. A similar problem is encountered in the SE specimen, where the arms can creep under a constant force F during crack propagation. One possible way to achieve steady state crack propagation is to have a sufficiently slow crack speed V so that the material far ahead of the crack tip has achieved the fully relaxed state, but this may be difficult to achieve in practice.

The above discussion invariably raises the question about the existence of steady state crack growth. Here we introduce the concept of quasi-steady state crack growth. A crack is undergoing quasi-steady state growth if the change in crack speed V and applied loading Δ is small while the crack traverses the damage zone ahead of it whose size is denoted by R_d . For a PS specimen, this condition is:

$$\max[|\dot{V}|/V, \dot{\Delta}/\Delta] \ll R_d/V. \quad (4.5)$$

While eqn (3.9) is still valid for quasi-steady state crack growth, the amount of energy flow to the crack tip G_{tip} , and hence the intrinsic energy $\Gamma_{\text{intrinsic}}$ is in general a function of crack speed V , that is, $\Gamma_{\text{intrinsic}}(V)$. However, the energy dissipated per unit crack extension, G_D , is now not only a function of

crack speed, but also depend on the history of loading as well as the specimen geometry. Indeed, for the same specimen, different loading histories can lead to different crack growth rates and different amounts of dissipation. As an example, Mayumi *et al.*⁷⁷ recently studied crack propagation in a viscoelastic gel with reversible ionic bonds. Using the SEC configuration with a fixed geometry (see Section 2.2.3), the crack achieved a constant speed V for a range of applied strain rates, but the computed effective energy release rate based on eqn (2.7) varied drastically. This is inconsistent with the expectation that $\Gamma_{\text{intrinsic}}$ should be a constant for the same crack speed V . Besides the fact that eqn (2.7) is approximate, it also includes both G_D and $G_{\text{tip}}(V)$. Using an approximation procedure, Mayumi *et al.*⁷⁷ subtracted the dissipation from the effective energy release rate given by eqn (2.7), and found that the local energy release rate $G_{\text{tip}}(V)$ is approximately constant for the same crack speed V . We emphasize that the estimate for dissipation in Mayumi *et al.*⁷⁷ is at best approximate. A more accurate calculation would require the solution of the full deformation fields in the specimen during crack propagation.

Finally, for gels where rate dependent material behavior is confined to a small region surrounding the crack tip (SSY), it is possible to lump the viscoelastic dissipation into an effective critical energy release rate which depends only on the crack speed V . The contribution of viscoelastic dissipation can be calculated following the method outlined in Section 3.2. Of course, in this case the hysteresis loop of material points surrounding the crack tip depends on the crack speed V and can be strongly coupled to the local failure mechanism. Persson and Brener¹⁰⁵ developed a model to estimate the energy dissipation for crack propagation in viscoelastic solids, but their formulation was based on the assumption of linear viscoelasticity and small deformation. Analytical or numerical models based on finite strain and nonlinear viscoelastic models, *e.g.* for gels with reversible bonds,^{21,22} are not yet available.

5. Future challenges

We have discussed a number of concepts and theoretical formulations for the characterization of fracture in tough hydrogels. Many fundamental questions and challenges still remain to be answered. In this section we list some of these challenges.

5.1 Rate independent gels

Ideally the toughness Γ_c^{init} for crack initiation that is defined based on the J -integral approach should be independent of specimen geometry, provided that the crack is loaded in the same mode (*e.g.* plane stress and Mode-I loading). Sun *et al.*⁴⁴ performed extensive tests to verify that the initiation toughness Γ_c^{init} for their alginate–polyacrylamide hybrid gel is insensitive to the sample geometry. Specifically they tested PS specimens with different heights $2H_0$, SE specimens and SEC specimens, and obtained approximately consistent Γ_c^{init} . Whether this is true for other gels is yet to be tested with more experimental data.

A more fundamental question is how to reconcile the field and energy approaches for crack initiation. Is Γ_c^{init} based on the critical J -integral intrinsically related to the energy delivered to the crack tip and the energy dissipation as the crack starts to grow?

For steady state crack propagation, the amount of dissipation can depend on the specimen geometry even if the intrinsic toughness $\Gamma_{\text{intrinsic}}$ is a material constant. Therefore, $\Gamma_{\text{intrinsic}}$ appears to be a more important parameter for the prediction of crack propagation condition compared to Γ_c . How can one directly measure $\Gamma_{\text{intrinsic}}$ in experiments? As mentioned in Section 3.3, Zhang *et al.*⁶⁴ developed an interesting method, *i.e.* to impose pre-stretch to the material before the fracture test so that the sacrificial network is already damaged. It would be useful to apply this method to a variety of tough gels. However, it should be noted that this method yielded a quite large $\Gamma_{\text{intrinsic}}$ ($\sim 400 \text{ J m}^{-2}$)⁶⁴ for an interpenetrating polyacrylamide–alginate gel. This value is higher than the fracture toughness of the pure polyacrylamide gel ($\sim 100\text{--}250 \text{ J m}^{-2}$) or the pure alginate gel ($\sim 10\text{--}35 \text{ J m}^{-2}$) according to Sun *et al.*⁴⁴ It is also much higher than the theoretical prediction based on Lake–Thomas theory ($\sim 10 \text{ J m}^{-2}$).⁴³ A possible explanation for this discrepancy is that the damaged network can still play a significant role in resisting the local fracture process. Further development of this and other techniques to experimentally probe the link between the network structure and the fracture process at the crack tip will be extremely useful to understand the factors controlling $\Gamma_{\text{intrinsic}}$.

Theoretical and numerical modeling can shed light on some of these issues. Numerical methods such as nonlinear finite element methods can be used to study transition from crack initiation to crack propagation in different specimens, which may provide insights into possible physical relation between the Γ_c values measured for crack propagation and the J integral based Γ_c^{init} for crack initiation.

5.2 Rate dependent gels

For rate dependent gels, *e.g.* viscoelastic gels with a mixture of chemical and reversible physical bonds,^{21,22} rigorous theory and appropriate experimental configurations to define fracture resistance for crack initiation and propagation are yet to be developed. The concept of toughness, either based on the energy release rate or the J -integral, in theory cannot be applied to characterize fracture of viscoelastic gels. As yet there is no replacement for these concepts, a different line of attack is to use the field approach instead of the energy approach. In this approach, a unique relation is sought between the crack propagation speed and some appropriate crack tip parameters that reflects the intensity of the crack-tip stress and strain fields. Similar approaches have been used to characterize creep crack growth in high temperature materials.¹⁰⁶ However, this approach requires a multi-axial time dependent constitutive model for the gel, which is not readily available in the literature except for some special cases.^{21,22,103}

5.3 Fatigue of tough gels

Fatigue refers to the failure of materials under cyclic loading. So far very little has been known about the fatigue properties of

tough gels. It is not clear what physical mechanisms determine the fatigue resistance of tough gels and how energy dissipation plays a role here, especially for cyclic loadings with low amplitudes. Intuitively, one would expect the early versions of DN gels to have low fatigue resistance, since they can be permanently damaged when irreversible breakage of covalent bonds in the stiff network occurs. But this hypothesis remains to be verified by experimental data. On the other hand, gels with reversible bonds have demonstrated the capability to self-heal after damage.^{21,49} It would be very interesting to study how such bond reversibility affects the fatigue resistance. However, fatigue test data for such self-healing gels are currently lacking. The theoretical framework to define an appropriate fatigue crack propagation criterion for gels is yet to be developed. Extensive experimental tests and theoretical analysis will be needed for this still primitive area. A particular important component is the measurement and modeling of breaking and reforming kinetics of the reversible bonds. Recently a model for a poly(vinylalcohol) (PVA) gel with reversible bonds has been shown to well capture experimental data.^{21,22} This model should be extended for gels with more complex molecular structures and kinetic mechanisms, *e.g.* stress-dependent bond breaking kinetics. At this time, there is a lack of experimental data on how physical bonds break and reform in such gel networks, especially near the crack tip where the high stresses there can affect these kinetic processes.

5.4 Structure–property relation for gel fracture

Ultimately the goal is to understand how the molecular structure of tough gels, *e.g.* chain lengths and the relative ratio of different networks, is related to the energy dissipation around the crack tip. Efforts to link the molecular structure and toughness have been made for the DN gels.^{107–109} It is useful to extend such efforts to a general class of tough gels with similar physical structures, which will provide important insights for further improvement of the toughness. For example, how is the intrinsic toughness $\Gamma_{\text{intrinsic}}$ related to the network structure? How does the coupling between the stiff and soft networks affect the formation of a crack tip dissipation zone? Modeling and simulations can play an important role in this area of study.

6. Summary

This article focuses on the emerging area of fracture mechanics of tough hydrogels. We start by reviewing key concepts in elastic fracture mechanics and draw attention to their limitations. For gels that are rate insensitive, we argue that the fracture toughness at crack initiation Γ_c^{init} is characterized by the J integral. However, characterizing crack propagation is more complicated since it involves unloading. Hence we limit the discussion to steady state crack propagation. For this case, the energy release rate can be separated into a dissipative part and a part that is responsible for energy flow to the crack tip. We note that fracture toughness for crack growth in tough gels

can be specimen dependent, and give examples to illustrate that the intrinsic fracture toughness controls the amount of dissipation, even though it is usually much smaller. We also make connection to previous models explaining the toughening mechanism of DN gels.

The toughness concept is difficult to apply for viscoelastic gels, especially those with reversible physical bonds. For this case it is not possible to define a unique fracture toughness for crack initiation since the stress and strain states in a fracture specimen depend on the entire loading history and specimen geometry. For crack propagation, we suggest the idea that the crack propagation speed should be correlated with the local energy release rate G_{tip} , which can be very difficult to compute.

Finally, we emphasize that the concept of rate-independence is relative. Gels are inherently rate dependent (e.g. due to poroelasticity). Rate independence in this work refers to the length and time scale where the mechanical behavior and fracture resistance of gels are insensitive to rate-dependent mechanisms.

Acknowledgements

The authors are thankful to Costantino Creton for fruitful discussions. R. Long acknowledges the support from the Michelin Chair program for a visiting professorship at ESPCI Paris, France. C. Y. Hui acknowledges support from National Science Foundation, Award no. 1537087.

References

- D. J. McClements, E. A. Decker, Y. Park and J. Weiss, *Crit. Rev. Food Sci. Nutr.*, 2009, **49**, 577–606.
- M. J. Zohuriaan-Mehr, H. Omidian, S. Doroudiani and K. Kabiri, *J. Mater. Sci.*, 2010, **45**, 5711–5735.
- B. Jeong, Y. H. Bae, D. S. Lee and S. W. Kim, *Nature*, 1997, **388**, 860–862.
- N. A. Peppas, P. Bures, W. Leobandung and H. Ichikawa, *Eur. J. Pharm. Biopharm.*, 2000, **50**, 27–46.
- Y. Qiu and K. Park, *Adv. Drug Delivery Rev.*, 2001, **53**, 321–339.
- N. A. Peppas, J. Z. Hilt, A. Khademhosseini and R. Langer, *Adv. Mater.*, 2006, **18**, 1345–1360.
- J. L. Drury and D. J. Mooney, *Biomaterials*, 2003, **24**, 4337–4351.
- M. S. Shoichet, *Macromolecules*, 2010, **43**, 581–591.
- P. Calvert, *Adv. Mater.*, 2009, **21**, 743–756.
- K. Yasuda, J. P. Gong, Y. Katsuyama, A. Nakayama, Y. Tanabe and E. Kondo, *Biomaterials*, 2005, **26**, 4468–4475.
- D. J. Beebe, J. S. Moore, J. M. Bauer, Q. Yu, R. H. Liu, C. Devadoss and B. H. Jo, *Nature*, 2000, **404**, 588–590.
- L. Dong and H. Jiang, *Soft Matter*, 2007, **3**, 1223–1230.
- S. Cai, Y. Lou, P. Ganguly, A. Robisson and Z. Suo, *J. Appl. Phys.*, 2010, **107**, 103535.
- L. Ionov, *Mater. Today*, 2014, **17**, 494–503.
- B. P. Lee and S. Konst, *Adv. Mater.*, 2014, **26**, 3415–3419.
- C. Keplinger, J. Y. Sun, C. C. Foo, P. Rothenmund, G. M. Whitesides and Z. Suo, *Science*, 2013, **341**, 984–987.
- B. Chen, J. J. Lu, C. H. Yang, J. H. Yang, J. Zhou, Y. M. Chen and Z. Suo, *ACS Appl. Mater. Interfaces*, 2014, **6**, 7840–7845.
- C. Storm, J. J. Pastore, F. C. MacKintosh, T. C. Lubensky and P. A. Janmey, *Nature*, 2005, **435**, 191–194.
- X. Zhao, N. Huebsch, D. J. Mooney and Z. Suo, *J. Appl. Phys.*, 2010, **107**, 063509.
- T. Narita, K. Mayumi, G. Ducouret and P. Hebraud, *Macromolecules*, 2013, **46**, 4174–4183.
- R. Long, K. Mayumi, C. Creton, T. Narita and C. Y. Hui, *Macromolecules*, 2014, **47**, 7243–7250.
- J. Guo, R. Long, K. Mayumi and C. Y. Hui, *Macromolecules*, 2016, **49**, 3497–3507.
- W. Hong, X. Zhao, J. Zhou and Z. Suo, *J. Mech. Phys. Solids*, 2008, **56**, 1779–1793.
- S. A. Chester and L. Anand, *J. Mech. Phys. Solids*, 2010, **58**, 1879–1906.
- Y. Hu, X. Zhao, J. J. Vlassak and Z. Suo, *Appl. Phys. Lett.*, 2010, **96**, 121904.
- Z. L. Kalcioğlu, R. Mahmoodian, Y. Hu, Z. Suo and K. J. van Vliet, *Soft Matter*, 2012, **8**, 3393–3398.
- R. Macrombe, S. Cai, W. Hong, X. Zhao, Y. Lapusta and Z. Suo, *Soft Matter*, 2010, **6**, 784–793.
- S. Cai and Z. Suo, *J. Mech. Phys. Solids*, 2011, **59**, 2259–2278.
- X. Wang and W. Hong, *Soft Matter*, 2011, **7**, 8578–8581.
- X. Zhao, *J. Mech. Phys. Solids*, 2012, **60**, 319–332.
- G. W. Scherer, *J. Non-Cryst. Solids*, 1992, **144**, 8171–8178.
- T. Baumberger, C. Caroli and D. Martina, *Nat. Mater.*, 2006, **5**, 552–555.
- T. Baumberger, C. Caroli and D. Martina, *Eur. Phys. J. E*, 2006, **21**, 81–89.
- E. Bouchbinder, A. Livne and J. Fineberg, *Phys. Rev. Lett.*, 2008, **101**, 264302.
- E. Bouchbinder, A. Livne and J. Fineberg, *J. Mech. Phys. Solids*, 2009, **57**, 1568–1577.
- M. E. Seitz, D. Martina, T. Baumberger, V. R. Krishnan, C. Y. Hui and K. R. Shull, *Soft Matter*, 2009, **5**, 447–456.
- A. Livne, E. Bouchbinder, I. Svetlizky and J. Fineberg, *Science*, 2010, **327**, 1359–1363.
- C. Y. Hui, R. Long and J. Ning, *J. Appl. Mech.*, 2013, **80**, 021014.
- N. Bouklas, C. M. Landis and R. Huang, *J. Appl. Mech.*, 2015, **82**, 081007.
- T. L. Anderson, *Fracture Mechanics: Fundamental and Applications*, CRC Press, Boca Raton, 3rd edn, 2005.
- A. T. Zehnder, *Fracture Mechanics*, Springer, New York, 2012.
- R. Long and C. Y. Hui, *Extreme Mechanics Letters*, 2015, **4**, 131–155.
- X. Zhao, *Soft Matter*, 2014, **10**, 672–687.
- J. Y. Sun, X. Zhao, W. R. K. Illeperuma, O. Chaudhuri, K. H. Oh, D. J. Mooney, J. J. Vlassak and Z. Suo, *Nature*, 2012, **489**, 133–136.
- J. P. Gong, Y. Katsuyama, T. Kurokawa and Y. Osada, *Adv. Mater.*, 2003, **15**, 1155–1158.

- 46 N. A. Peppas and E. W. Merrill, *J. Biomed. Mater. Res.*, 2004, **11**, 423–434.
- 47 H. J. Kong, E. Wong and D. J. Mooney, *Macromolecules*, 2003, **36**, 4582–4588.
- 48 K. J. Henderson, T. C. Zhou, K. J. Otim and K. R. Shull, *Macromolecules*, 2010, **43**, 6193–6201.
- 49 T. L. Sun, T. Kurokawa, S. Kuroda, A. B. Ihsan, T. Akasaki, K. Sato, M. A. Haque, T. Nakajima and J. P. Gong, *Nat. Mater.*, 2013, **12**, 932–937.
- 50 T. Huang, H. Xu, K. Jiao, L. Zhu, H. R. Brown and H. Wang, *Adv. Mater.*, 2007, **19**, 1622–1626.
- 51 K. Haraguchi and T. Takehisa, *Adv. Mater.*, 2002, **14**, 1120–1124.
- 52 Q. Wang, J. L. Mynar, M. Yoshida, E. Lee, M. Lee, K. Okuro, K. Kinbara and T. Aida, *Nature*, 2010, **436**, 339–343.
- 53 W. C. Lin, W. Fan, A. Marcellan, D. Hourdet and C. Creton, *Macromolecules*, 2010, **43**, 2554–2563.
- 54 A. B. Imran, K. Esaki, H. Gotoh, T. Seki, K. Ito, Y. Sakai and Y. Takeoka, *Nat. Commun.*, 2014, **5**, 5124.
- 55 A. E. X. Brown, R. I. Litvinov, D. E. Discher, P. K. Purohit and J. W. Weisel, *Science*, 2009, **325**, 741–744.
- 56 A. A. Griffith, *Philos. Trans. R. Soc., A*, 1921, **221**, 582–593.
- 57 J. P. Gong, *Soft Matter*, 2010, **6**, 2583–2590.
- 58 J. R. Rice and G. F. Rosengren, *J. Mech. Phys. Solids*, 1968, **16**, 1–12.
- 59 J. W. Hutchinson, *J. Mech. Phys. Solids*, 1968, **16**, 13–31.
- 60 J. W. Hutchinson, *J. Appl. Mech.*, 1983, **50**, 1042–1051.
- 61 R. O. Ritchie, *Journal of Engineering Materials and Technology*, 1983, **105**, 1–7.
- 62 H. R. Brown, *Macromolecules*, 2007, **40**, 3815–3818.
- 63 Y. Tanaka, *European Physics Letters*, 2007, **78**, 56005.
- 64 T. Zhang, S. Lin, H. Yuk and X. Zhao, *Extreme Mechanics Letters*, 2015, **4**, 1–8.
- 65 C. Creton and M. Ciccotti, *Reports of Progress in Physics*, 2016, **79**, 40616.
- 66 J. K. Knowles and E. Steinberg, *Arch. Ration. Mech. Anal.*, 1972, **44**, 83–87.
- 67 J. R. Rice, *J. Appl. Mech.*, 1968, **35**, 370–386.
- 68 T. Baumberger and O. Ronsin, *Eur. Phys. J. E*, 2010, **31**, 51–58.
- 69 R. S. Rivlin and A. G. Thomas, *J. Polym. Sci.*, 1953, **10**, 291–318.
- 70 H. W. Greensmith, *J. Appl. Polym. Sci.*, 1963, **7**, 993–1002.
- 71 A. N. Gent, *Langmuir*, 1996, **12**, 4492–4496.
- 72 G. Miquelard-Garnier, D. Hourdet and C. Creton, *Polymer*, 2009, **50**, 481–490.
- 73 M. A. Haque, T. Kurokawa, G. Kamita and J. P. Gong, *Macromolecules*, 2011, **44**, 8916–8924.
- 74 Y. Tanaka, R. Kuwabara, Y. H. Na, T. Kurokawa, J. P. Gong and Y. Osada, *J. Phys. Chem. B*, 2005, **109**, 11559–11562.
- 75 Q. M. Yu, Y. Tanaka, H. Furukawa, T. Kurokawa and J. P. Gong, *Macromolecules*, 2009, **42**, 3852–3855.
- 76 M. A. Haque, T. Kurokawa and J. P. Gong, *Polymer*, 2012, **53**, 1805–1822.
- 77 K. Mayumi, J. Guo, T. Narita, C. Y. Hui and C. Creton, *Extreme Mechanics Letters*, 2016, **6**, 52–59.
- 78 H. Guo, N. Sanson, D. Hourdet and A. Marcellan, *Adv. Mater.*, 2016, **28**, 5857–5864.
- 79 H. Tada, P. C. Paris and G. R. Irwin, *Stress analysis of cracks handbook*, ASME Press, New York, 3rd edn, 2000.
- 80 M. W. Keller, S. R. White and N. R. Sottos, *Adv. Funct. Mater.*, 2007, **17**, 2399–2404.
- 81 D. R. King, T. L. Sun, Y. Huang, T. Kurokawa, T. Nonoyama, A. J. Crosby and J. P. Gong, *Mater. Horiz.*, 2015, **2**, 584–591.
- 82 P. H. Geubelle and W. G. Knauss, *J. Thermoelasticity*, 1994, **35**, 61–98.
- 83 R. Long, V. R. Krishnan and C. Y. Hui, *J. Mech. Phys. Solids*, 2011, **59**, 672–695.
- 84 Y. H. Na, Y. Tanaka, Y. Kawauchi, H. Furukawa, T. Sumiyoshi, J. P. Gong and Y. Osada, *Macromolecules*, 2006, **39**, 4641–4645.
- 85 T. Nakajima, H. Furukawa, J. P. Gong, E. K. Lin and W. Wu, *Macromol. Symp.*, 2010, **291–292**, 122–126.
- 86 R. W. Ogden and D. G. Roxburgh, *Proc. R. Soc. A*, 1999, **455**, 2861–2877.
- 87 T. van Vliet and P. Walstra, *Faraday Discuss.*, 1995, **101**, 359–370.
- 88 Y. Tanaka, K. Fukao and Y. Miyamoto, *Eur. Phys. J. E*, 2000, **3**, 395–401.
- 89 C. Gamonpilas, M. N. Charalambides and J. G. Williams, *J. Mater. Sci.*, 2009, **44**, 4976–4986.
- 90 T. Baumberger and O. Ronsin, *J. Chem. Phys.*, 2009, **130**, 061102.
- 91 M. Lefranc and E. Bouchaud, *Extreme Mechanics Letters*, 2014, **1**, 97–103.
- 92 H. Xin, H. R. Brown, S. Naficy and G. M. Spinks, *Polymer*, 2015, **65**, 253–261.
- 93 A. Ruina, *Influence of coupled deformation-diffusion effects on the retardation of hydraulic fracture*, MSc thesis, Brown University, Providence, RI, 1978.
- 94 R. Long, M. Lefranc, E. Bouchaud and C. Y. Hui, *Extreme Mechanics Letters*, 2016, DOI: 10.1016/j.eml.2016.05.005.
- 95 X. Wang and W. Hong, *Soft Matter*, 2012, **8**, 8171–8178.
- 96 D. Bonn, H. Kellay, M. Prochnow, K. Ben-Djemaa and J. Meunier, *Science*, 1998, **280**, 265–267.
- 97 G. Noselli, A. Lucantonio, R. M. McMeeking and A. DeSimone, *J. Mech. Phys. Solids*, 2016, **94**, 33–46.
- 98 G. M. Kavanagh and S. B. Ross-Murphy, *Prog. Polym. Sci.*, 1998, **23**, 533–562.
- 99 Y. Hu and Z. Suo, *Acta Mech. Solida Sin.*, 2012, **25**, 441–458.
- 100 Q. M. Wang, A. C. Mohan, M. L. Oyen and X. Zhao, *Acta Mech. Sin.*, 2014, **30**, 20–27.
- 101 K. Mayumi, A. Marcellan, G. Ducouret, C. Creton and T. Narita, *ACS Macro Lett.*, 2013, **2**, 1065–1068.
- 102 S. C. Grindy, R. Learsch, D. Mozhdehi, J. Cheng, D. G. Barrett, Z. Guan, P. B. Messersmith and N. Holten-Andersen, *Nat. Mater.*, 2015, **14**, 1210–1216.
- 103 R. Long, K. Mayumi, C. Creton, T. Narita and C. Y. Hui, *J. Rheol.*, 2015, **59**, 643–665.
- 104 T. L. Sun, F. Luo, T. Kurokawa, S. N. Karobi, T. Nakajima and J. P. Gong, *Soft Matter*, 2015, **11**, 9355–9366.

- 105 B. N. J. Persson and E. A. Brener, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2005, **71**, 036123.
- 106 H. Riedel, *Fracture at High Temperatures*, Springer-Verlag, Berlin, Germany, 1987, ISBN: 978-3-642-82963-5.
- 107 V. R. Tirumala, T. Tominaga, S. Lee, P. D. Butler, E. K. Lin, J. P. Gong and W. L. Wu, *J. Phys. Chem.*, 2008, **112**, 8024–8031.
- 108 T. Nakajima, H. Furukawa, Y. Tanaka, T. Kurokawa, Y. Osada and J. P. Gong, *Macromolecules*, 2009, **42**, 2184–2189.
- 109 T. Nakajima, H. Furukawa, Y. Tanaka, T. Kurokawa and J. P. Gong, *J. Polym. Sci., Part B: Polym. Phys.*, 2011, **49**, 1246–1254.