Phase Field Model of Fracture: An Accessible Introduction

Conor Rowan

Spring 2024

1 Introduction

The phase field model is a popular and relatively recent method for modeling damage and fracture in structures. Though it is heavily researched, there are not many "textbook" presentations of the model. This write-up is meant to be a more accessible way to learn about the model than reading papers. Papers tend to assume a certain level of familiarity with the basics already, and in my opinion do not act as effective pedagogical resources. The phase field model has a very nice interpretation once one becomes familiar with it, though the process of getting to that point may be arduous. Hopefully this helps!

2 Phase Field Model of Fracture

One might trace the origins of the phase field model of fracture back to the work of Griffith's in the early 20th century. The problem that Griffith analyzed is a classical result in linear elastic fracture mechanics, and gets at the essential components of the phase field model of fracture. Griffith analyzed brittle fracture of a semi-infinite plate with an existing crack of length a, as shown in Figure 1. He argued that as the crack grows, a triangular region of material unloads and becomes stress free. This corresponds to a "liberation" of strain energy from the structure. However, the crack forms by breaking atomic bonds holding the material together, and this requires energy. Thus crack formation also acts to increase in the energy of the system. Assuming that the external loads do no work as the crack grows, and that the stress state outside of the triangular region is unaffected by crack growth, we can write the change in the structure's energy as a result of crack growth as

$$\Delta U = 2\gamma\Delta a - \frac{1}{2}\sigma\epsilon\Delta A = 2\gamma\Delta a - \frac{\sigma^2}{2E}\Big((a+\Delta a)^2\beta - a^2\beta\Big)$$

where ΔA is the increase in the area shielded by the crack. The two triangular regions of unloaded material change size with crack growth, and this corresponds to a loss of stored strain energy. The first term is the work required to form two

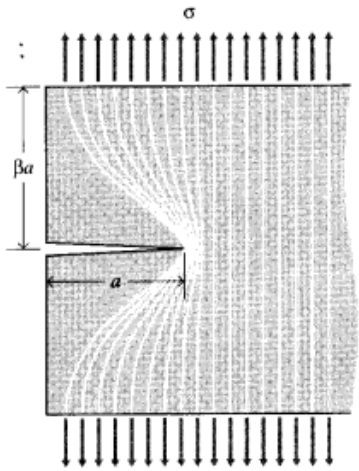

Figure 1: Two triangular regions with base a and height βa are shielded from carrying any load by the crack. The height of this region has no obvious a priori value, and is controlled by the parameter β. This parameter can be estimated by appealing to other physical considerations, but for our purposes we treat it as a given.

new surfaces (one on either side of the crack), where we assume that the energy associated with crack formation is directly proportional to the crack length. The parameter γ is material specific, controlling the energy required to break atomic bonds, and thus the resistance to crack formation. The second term is the released strain energy as a result of advancing the crack from length a to $a + \Delta a$, where the strain energy of the plate is computed using a uniaxial stress assumption. We compute this strain energy by multiplying the energy density with the area. Expanding the above expression and neglecting the quadratic dependence on the increment in crack growth, we obtain

$$\Delta U = \Big(2\gamma - \frac{\beta\sigma^2}{E}a\Big)\Delta a$$

Griffith's argued that the crack is stable when increasing its length increases the energy of the system, i.e. $\Delta U > 0$. Alternatively, unstable crack growth occurs when increasing the crack leads to a loss of energy. The transition point is when $\Delta U = 0$, which gives a condition for the crack being at a critical point

$$2\gamma = \frac{\beta\sigma^2}{E}a$$

This equation can be used to determine the critical stress for a given crack length, or critical crack length for a given stress. Basically, Griffith's analysis states that cracks grow when it is energetically favorable to do so. The phase

field model of fracture takes this insight and translates it into a framework amenable to computation. Inspired by continuum damage models, we introduce a damage (or phase field) variable $\phi \in [0, 1]$ which continuously interpolates between a state of undamaged material ($\phi = 0$) and fully fractured material ($\phi = 1$). This is a definition. Analogous to Griffith's approach, we can model the liberation of strain energy as a result of cracking by computing the stored strain energy of the structure as

$$\mathcal{E}_u = \int_\Omega (\phi - 1)^2 \Psi d\Omega$$

where $\Psi = \frac{1}{2}\sigma_{ij}\epsilon_{ij}$ is the usual elastic strain energy density and Ω defines the volume of the structure. Note that the function $g(\phi) = (\phi-1)^2$ satisfies $g(1) = 0$ and $g(0) = 1$. This means that undamaged material experiences no reduction in strain energy, and fully fractured material stores no energy. At intermediate values, there is a continuous reduction in the storage of strain energy at each point in the structure. This is analogous to the Griffith's model. There are other choices for what we call the "degradation function" $g(\phi)$, but we will stick with the form given above. In other words, this quadratic form of energy loss, though it acts as a constitutive relation of sorts, is more-or-less arbitrary. It can be shown to have certain desirable properties, however. For example, it can be shown that $g'(1) = 0$ ensures that damage does not grow beyond $\phi = 1$. Thus, the normalization condition on the phase field field variable can be enforced automatically. Perhaps you can see why this is after reading through these notes. To reiterate: this expression for the total strain energy of the structure states that damage acts to degrade the stored strain energy at each point.

Griffith's wrote *the change in energy* of the elastic system as a competition between energy released due to crack formation and energy required to form cracks, using this to find a condition for crack stability. In the phase field model, we will write the total energy of the system in terms of an unknown displacement $u(x)$ and crack pattern $\phi(x)$. We have already written the total stored strain energy of the potentially damaged structure. *The insight of this model is that the displacement and phase field should be such that the total energy is at a minimum.* Whereas Griffith's model assumed a pre-existing crack, simple geometry, a given direction of crack propagation, and did not model stable crack growth, the phase field model will apply to arbitrarily complex geometries, and it will model the formation and stable growth of cracks. Like the Griffith's model, we expect that when the loads are small, it is "energetically favorable" to store energy in the form of strain. This is because there are serious "overhead" energy requirements to open up any cracks, controlled by a material parameter like γ. But at some point, it will be energetically "worth it" to open cracks in order to release strain energy. *The phase field model is essentially a way of quantifying that energy can be stored either through strain or through cracks, and the state of the system we actually observe is the one which minimizes the total energy.* Having seen the mechanism of strain energy storage in the presence of damage,

and having motivated the intentions of this model, we can now specify the way in which energy is "stored" in cracks. To be clear, the interpretation of "stored" energy here is maybe unclear, because we cannot recovery the energy that goes into forming cracks (unlike elastic stored energy). So it might be better to call this dissipated energy. I will leave this up to you to make sense of, but the point of the word "stored" energy is that crack formation is an energy sink of some sort.

The energy associated with crack formation is computed using the damage field $\phi(x)$. How exactly this is done turns out to be where the real artistry of the model comes into play. The first thing we do is define a new constitutive parameter $G \geq 2\gamma$ which measures the material's resistance to fracture accounting not just for the work required to break atomic bonds, but also localized plastic flow in the vicinity of the crack tip. For a crack of area Γ, the energy associated with fracture is then $G\Gamma$. The point here is that only accounting for broken atomic bonds underestimates the energy dissipated in the opening of a crack. We need a volumetric representation of the presence of cracks, so that we can monitor the formation of damage at every point in the structure. Analogous to a strain energy density, we want a "crack density" d such that the total stored energy associated with crack formation is

$$\mathcal{E}_\phi = \int G d(\phi) d\Omega$$

This expression requires that cracks cannot be represented by bands of damage with zero width, otherwise we are integrating a field of measure zero. When the damage localizes to zero-width bands, the structure could be fully fractured with no energy associated with crack formation. Thus, cracks need to be smeared out over some finite width in order for the crack density approach to work. Why not simply use $\gamma(\phi) = \frac{1}{2}G\phi^2$, meaning that the total energy associated with damage is just the sum of all the local stored energies? The quadratic dependence on the field variable is typical of an expression for energy, though its meaning is not obvious here. Perhaps we could think of there being a crack opening force $G\phi$, such that the work it does is $\int G\phi d\phi = \frac{1}{2}G\phi^2$. This is a reasonable thought, but nothing prevents this from localizing to zero measure bands which cause the structure to fail to carry loads without any associated energy "cost" of crack growth. Plus, the physical interpretation of the energy's quadratic dependence on the phase field is a bit hand-wavey. We need to select a crack density functional which gives rise to localized, but finite-width cracks. It turns out that the following crack density functional accomplish this

$$d(\phi) = \frac{1}{2\ell}\left(\phi^2 + \ell^2 \frac{\partial \phi}{\partial x_i}\frac{\partial \phi}{\partial x_i}\right)$$

It is as if we added a penalty on the gradients of the damage ϕ which incur a cost for localization, where gradients are large. Both the formation of the crack and its "sharpness" incur an energetic cost. This causes the crack to spread out

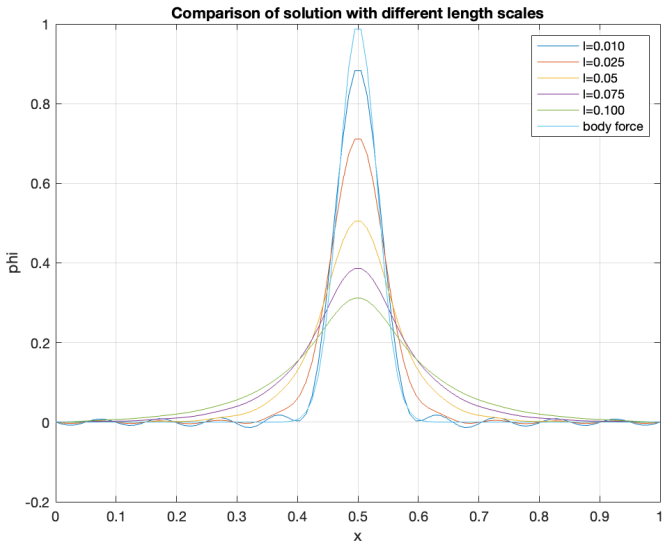

Figure 2: The length scale dictates how localized the crack is. When the length scale is large, gradients of the damage are penalized more heavily leading to more diffuse crack patterns. The damage pattern is a smeared out version of the body force as a result of the gradients in the energy functional. This solution is obtained using spectral basis functions with zero boundaries.

over a finite width. The parameter ℓ is called the "length scale" and it controls how much gradients of the damage are penalized. We can think of this as scaling an almost diffusion-like term. Given this diffusion behavior, we do not expect the crack to be more localized than the force that drives its formation. In other words, the length scale parameter will control how much whatever force drives crack growth is smeared out in the formation of the crack. See the next section for a more in-depth exploration of this crack density function. To get a feel for how this functional works in practice, we can compute its minimum, as we will do as part of solution to the phase field model. But first, we need to add a volumetric driving force term. For now, we can think of the elastic deformation of the body driving the formation of cracks in some unknown way. The energy stored in crack formation is

$$\Pi_\phi = \int \frac{G}{2\ell}\left(\phi^2 + \ell^2 \frac{\partial \phi}{\partial x_i}\frac{\partial \phi}{\partial x_i}\right) - F\phi dx$$

where F is the a generalized force driving cracak growth. For our purposes, we treat the volumetric "crack-driving force" F as known. We are ignoring the part of the phase field model which involves the displacement in order to explore the behavior of the crack density functional. We want to see the kind of crack

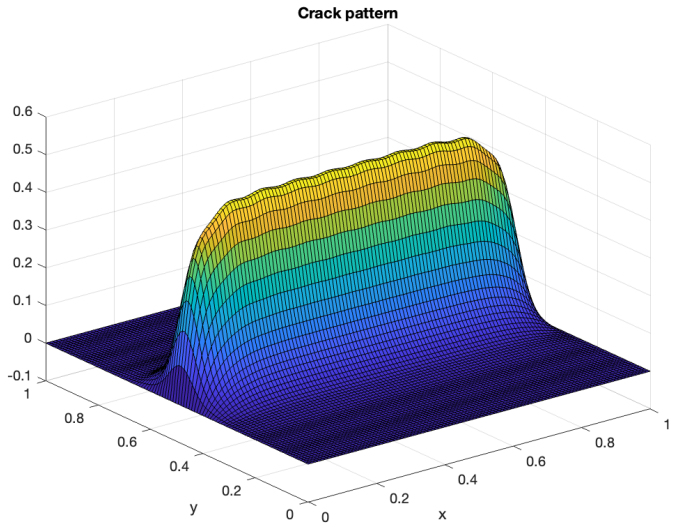

Figure 3: Two dimensional crack pattern obtained from minimizing the crack energy functional. The body force is a narrow Gaussian bump along $y = 1/2$. This solution is obtained using spectral basis functions with zero boundaries.

patterns the chosen density functional gives. The minimum of this functional is given by

$$\delta\Pi_\phi = \int \frac{G}{\ell}\left(\phi\delta\phi + \ell^2 \frac{\partial\phi}{\partial x_i}\frac{\partial\delta\phi}{\partial x_i}\right) - F\delta\phi dx$$

With standard methods, we discretize the displacement in terms of a set of basis functions f_j, meaning that $\phi = \sum_j \phi_j f_j$. The variation $\delta\phi$ is discretized with same set of basis functions. Noting that the variation is arbitrary, we obtain a linear system for the phase field coefficients:

$$\left(\int \frac{G}{\ell}\left(f_i f_j + \ell^2 \frac{\partial f_i}{\partial x_k}\frac{\partial f_j}{\partial x_k}\right)dx\right)\phi_j = \int F f_i dx$$

See Figures 2 and 3 for some example solutions of the crack problem. This crack density functional leads to localized crack patterns which are zero outside regions where the driving force is large. The size of the length scale dictates how localized the solution is. We see from these results that this choice of crack density will encourage the formation of narrow bands of damage as we expect from actual fracture phenomena.

Having justified this choice of crack density, we can now write the total stored energy of the system as

$$\mathcal{E}_{u+\phi} = \int_\Omega (\phi - 1)^2 \Psi + \frac{G}{2\ell}\left(\phi^2 + \ell^2 \frac{\partial \phi}{\partial x_i}\frac{\partial \phi}{\partial x_i}\right) d\Omega$$

This is quite similar to the Griffith's model. The first term in the energy gauges how energy is released as a result of crack growth unloading material. The second term quantifies how energy is absorbed in the formation of cracks. Now, we know that the total potential energy functional whose minimum solves the elastic problem is the difference between the stored energy of the system and the work of external forces. This is the case for the phase field problem as well. Thus, we can finally write the energy functional which governs the phase field problem as

$$\Pi = \int_\Omega (\phi - 1)^2 \Psi + \frac{G}{2\ell}\left(\phi^2 + \ell^2 \frac{\partial \phi}{\partial x_i}\frac{\partial \phi}{\partial x_i}\right) - b_i u_i d\Omega - \int_{\partial\Omega} t_i u_i dS$$

where the external work of body forces and tractions have been subtracted from the stored energy. This is called the isotropic phase field model, because there is no distinction between what forms of strain energy drive crack growth. In one dimension, this means that compression and tension contribute equally to the formation of cracks. In higher dimensions, this means that volumetric strains drive fracture as much as strains causing distortion. We know this is not physical–tension should more readily cause fracture than compression, and we expect that a material will not yield from hydrostatic stress states. Thus, we can introduce the anisotropic phase field model, which accounts for the influence of different types of strain energy on crack formation. Conceptually, the anisotropic phase field model is a simple modification:

$$\Pi = \int_\Omega (\phi - 1)^2 \Psi^+ + \Psi^- + \frac{G}{2\ell}\left(\phi^2 + \ell^2 \frac{\partial \phi}{\partial x_i}\frac{\partial \phi}{\partial x_i}\right) - b_i u_i d\Omega - \int_{\partial\Omega} t_i u_i dS$$

The strain energy has been decomposed into a "positive" part, which is released as damage forms, and a "negative" part, which is not affected by the formation of damage. This decomposition must respect $\Psi = \Psi^+ + \Psi^-$. When one writes out the strong form of the governing equations associated with this energy functional, it becomes clear that only the positive energy Ψ^+ drives crack growth. One interpretation of this is that even when the material is fully fractured, it can store strain energy in the form of Ψ^-. For example, a fractured bar can still carry compressive loads. There are various ways to decompose the energy. In one dimension, it is clear that Ψ^+ is tensile strain energy, and Ψ^- is compressive strain energy. In higher dimensions, there is more freedom in choosing the form of this decomposition. Different decompositions correspond to different constitutive assumptions about what causes fracture. To give a feel for how this might work, the 1D energy decomposition would be

$$\Psi^+ = \frac{1}{2}E\left\langle \frac{\partial u}{\partial x}\right\rangle_+^2, \quad \Psi^- = \frac{1}{2}E\left\langle \frac{\partial u}{\partial x}\right\rangle_-^2$$

where the we make use of the notation

$$\langle x \rangle_+ = \begin{cases} x, & x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

$$\langle x \rangle_- = \begin{cases} 0, & x \geq 0 \\ -x, & \text{otherwise} \end{cases}$$

Two and three dimensional phase field energy decompositions make use of similar operations.

The final aspect of the phase field model we need to consider is history dependence. We imagine the formation of cracks as a process which progresses with the quasi-static application of external forces. Fracture is an inherently path-dependent phenomenon. For example, the application of a cyclic load may not bring the structure back to its initial state if cracks form. But, there is nothing at this point that prevents cracks from closing ("healing") when loads are removed. Enforcing the irreversibility of cracks is the final ingredient of the phase field model. We thus need to track the state of the structure over the entire load path in order to understand the damage that results. Even when the load is not cyclic, we imagine loading as a process which progressively forms damage. Thus, if a structure is loaded statically by external tractions $\underline{t}(\underline{x})$, the phase field model treats this loading as incremental with external tractions

$$(t_1)\underline{t}(\underline{x}) \rightarrow (t_2)\underline{t}(\underline{x}) \rightarrow \cdots \rightarrow (1)\underline{t}(\underline{x}), \quad t_1 < t_2 < \cdots < 1$$

The displacement and phase field solution under a static load are modeled as the end result of a process of monotonically increasing loads. The energy at load step k for an anisotropic phase field model is thus given by

$$\Pi^k = \int_\Omega (\phi^k - 1)^2 \Psi^+ \left(\frac{\partial \underline{u}^k}{\partial \underline{x}} \right) + \Psi^- \left(\frac{\partial \underline{u}^k}{\partial \underline{x}} \right) + \frac{G}{2\ell} \left(\phi^{2,k} + \ell^2 \frac{\partial \phi^k}{\partial x_i} \frac{\partial \phi^k}{\partial x_i} \right) d\Omega - \int_\Omega b_i^k u_i^k d\Omega - \int_{\partial\Omega} t_i^k u_i^k dS$$

The minimum of Π^k is the displacement and phase field solution at load step k. The crack irreversibility constraint can be written simply as

$$\phi^k \geq \phi^{k-1} \quad \forall x \in \Omega$$

During the ramping up of the external forces, nowhere in the domain can the crack heal. In some cases of monotonic load stepping this constraint may be met without explicit enforcement. This is certainly the case in one spatial dimension. But for a more complex structure, a crack may form and unload some previously damaged material surrounding it. This unloading could cause the unphysical reversal of damage without explicit enforcement of the constraint. See Figure 4 for further elaboration.

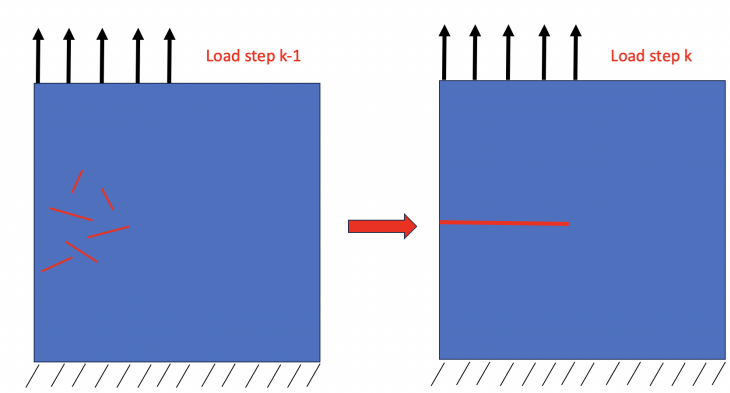

Figure 4: An example situation in which cracks could heal under monotonic loading. A plate is loaded and begins to form diffuse damage at load step $k-1$, then localizes into a sharp crack at load step k. Because the sharp crack unloads the surrounding material, the original diffuse damage could heal unless the irreversibility constraint is explicitly enforced.

3 Crack Density Function

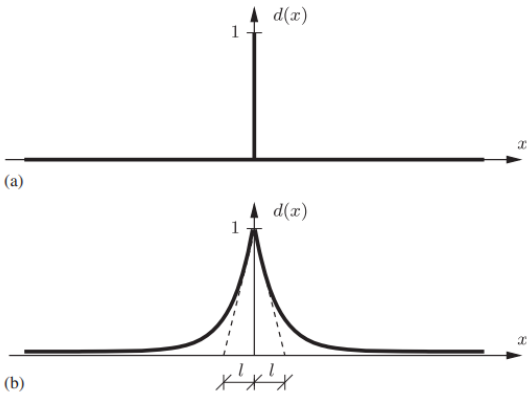

Figure 5: Sharp crack geometry, and the diffuse representation of the crack with length scale ℓ. This figure is borrowed from this paper.

We have claimed that the energy associated with the opening of cracks in a homogeneous material (constant fracture energy G) is

$$\mathcal{E}_\phi = \frac{G}{2\ell} \int \phi^2 + \ell^2 \frac{\partial \phi}{\partial x_i} \frac{\partial \phi}{\partial x_i} d\Omega$$

The material parameter G measures the energy required to open a unit surface area of crack, meaning that it has units J/m^2. It can be seen that this implies the phase field is unit-less. The total energy expended in opening cracks is then $G\Gamma$ where Γ is the total area of cracks formed. This implies that for a given phase field ϕ, the crack area is

$$\Gamma = \frac{1}{2\ell}\int \phi^2 + \ell^2 \frac{\partial \phi}{\partial x_i}\frac{\partial \phi}{\partial x_i} d\Omega$$

It is not at all clear why this should be the case. It makes sense that the total crack area could be computed by integration the damage over the domain, but why the "density" of crack area takes this form is not obvious. We can begin to familiarize ourselves with this construction of the crack area by considering a simple example. A sharp crack in an infinite 1D bar with cross-section A will be spread out over a finite width, as shown in Figure 5. The sharp crack is

$$\phi(x) = \begin{cases} 1 & x = 0 \\ 0 & \text{otherwise} \end{cases}$$

whereas the "smeared" crack is represented by

$$\phi(x) = e^{-|x|/\ell}$$

The length scale ℓ controls how narrow the crack is. Smaller length scales lead to sharper crack geometries. This crack shape is not based on any physics, it is simply a reasonable seeming approach to spread the crack over a finite width. Now, observe that this function satisfies the following differential equation:

$$\phi(x) - \ell^2 \frac{\partial^2 \phi}{\partial x^2} = 0, \quad \phi(0) = 1, \quad \frac{\partial \phi}{\partial x}(-\infty) = \frac{\partial \phi}{\partial x}(\infty) = 0$$

This is a bit of strange differential equation given that it has no forcing term and is driven by a prescribed value, which would need to be enforced as a constraint. The calculus of variations can be used to verify that this differential equation arises from a variational principle stated as

$$I(\phi) = \frac{A}{2}\int_{-\infty}^{\infty} \phi^2 + \ell^2 \left(\frac{\partial \phi}{\partial x}\right)^2 dx$$

The cross-sectional area appears because we integrate over the volume of the bar, but make use of fact that the phase field only depends on the axial coordinate x. The differential equation expresses the condition for a minimum of this energy-like functional. The smeared approximation of the sharp crack we have given can be conceptualized as the minimum to this energy-like functional given the constraint that $\phi(0) = 1$ and the two zero Neumann boundary conditions at $\pm\infty$. Seeing how the minima of this energy functional interpolates between the two boundary conditions and a prescribed value of the damage gives a sense of its behavior. In this case, it seems to prefer sharp gradients which quickly drive the solution down to zero. The length scale ℓ controls how localized the solution

is. Note that governing equation $\phi = \ell^2 \phi''$ states that the solution equals its curvature at each point. This means that if the solution is nonzero, it must be changing rapidly in a convex manner. This helps us understand its tendency to localize, and particularly to form a cusp at $x = 0$.

If we take the diffuse crack $\phi = e^{-|x|/\ell}$ and plug it into the functional $I(\phi)$, one can easily show that

$$I(e^{-|x|/\ell}) = A\ell$$

Remember that we have used this diffuse crack approximation to model complete fracture of the bar. This means that, by construction, we expect the area exposed from cracking to be the cross-sectional area A, thus the total energy expended in fracture is GA. We do not want the crack area to depend on how we smooth out the sharp crack geometry with the length scale ℓ. Thus, we can see that the value of the minimum of the new functional $\Gamma(\phi) = I(\phi)/\ell$ is simply the cross-sectional area A. From this particular example, we are led to the conclusion we hoped for, namely that the crack area is given by

$$\Gamma = \frac{1}{2\ell} \int_\Omega \phi^2 + \ell^2 \left(\frac{\partial \phi}{\partial x} \right)^2 d\Omega$$

where $\int (\cdot) d\Omega = A \int (\cdot) dx$. Now the fracture area for the fully formed crack is independent of the length scale, as desired. This is a bit of an unsatisfying argument, though. To summarize: a particular form of diffuse crack was assumed which seemed to arise as the minimum of an energy functional. This energy functional was modified slightly such that its value when evaluated at the crack was the cross-sectional area, which in this particular case was also the crack area. Is this grounds to claim that in all cases, the functional Γ takes in the phase field and approximates the crack area? It does not seem that the 1D example can be pushed any further to help answer this question. Nor does it seem like an analogous analytical example is available in 2D. Thus, we can turn to a computational example. Consider the following problem, depicted schematically in Figure 6. We have a square plate with a notch of length a. We imagine this notch as fully fractured material and that the notch is at $x_2 = L/2$, thus we have $\phi(s, L/2) = 1$ for $0 < s < a$. We then use the energy functional Γ to interpolate the damage prescribed at the notch into the entire domain of the plate. In 2D, the crack area functional is

$$\Gamma = \frac{1}{2\ell} \int \phi^2 + \ell^2 \frac{\partial \phi}{\partial x_i} \frac{\partial \phi}{\partial x_i} dA$$

Note that the zero Neumann boundary conditions are enforced weakly by this energy functional. We need to find a phase field which minimizes this functional subject to the constraint that $\phi = 1$ along the notch. Call this solution $\tilde{\phi}$. We want to verify with a computational example that $\Gamma(\tilde{\phi}) \approx a$. In other words, we know that the crack length should be approximately the length of the notch

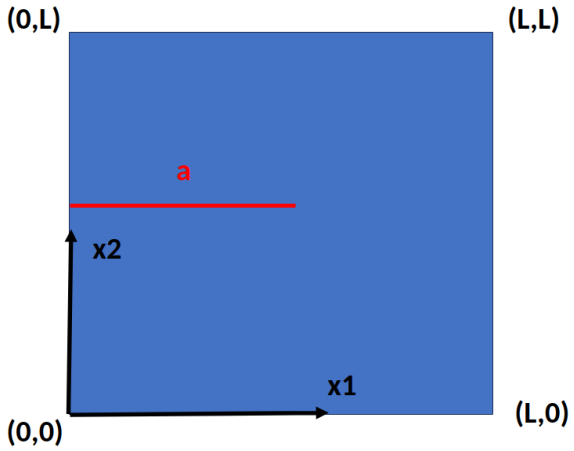

Figure 6: A square plate with side length L and a notch of length a, along which $\phi = 1$. The boundary conditions are $\frac{\partial \phi}{\partial x_i} n_i = 0$ where n is the local outward facing normal along the four edges.

(because we prescribed it), and we want to verify that when Γ is evaluated at its minimum, we return an approximation of the crack length. We can discretize the phase field ϕ with a neural network, and define an objective

$$\Pi = \frac{1}{2\ell} \int \phi^2 + \ell^2 \frac{\partial \phi}{\partial x_i} \frac{\partial \phi}{\partial x_i} dA + \frac{\lambda}{2} \int_0^a \Big(\phi(s, L/2) - 1 \Big)^2 ds$$

where λ is a penalty parameter used to enforce the constraint the crack is fully formed along the notch. We can compute the gradient of this with respect to the neural network parameters defining the phase field, and plug this in to an unconstrained optimization algorithm to determine the minimizer $\tilde{\phi}$. Once we have the phase field solution for a given notch length a from solving the optimization problem, we can compute $\Gamma(\tilde{\phi})$ and compare it to the notch length. See Figures 7-9 for results. Like the 1D example, these results suggest that the crack area functional leads to localized solutions, and its value corresponds to the total crack area for a given phase field.

4 Strong Form of the Phase Field Model

For completeness, we will derive the strong form of the governing equations for the total potential energy of the phase field model. Even if these equations are not used in a numerical implementation, they do provide some additional insight into the mechanics of the model. The total potential for the anisotropic phase field model is

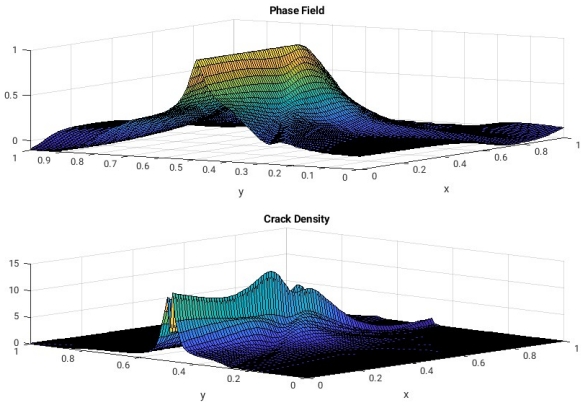

Figure 7: Surface plot of phase field and crack density in a plate with side length $L = 1$, a notch of length $a = 1/2$, and length scale of $\ell = 0.1$. The phase field interpolates the damage in the notched plate by minimizing the functional Γ. We can see that outside the crack, the phase field quickly drops to zero. The crack length is computed by evaluating Γ at the solution. For the plate with side length $L = 1$, we compute the crack length as $\Gamma = 0.7$. It makes sense that the crack length from the phase field is larger than the notch given it is smeared out over a larger region as a result of the interpolation.

$$\Pi = \int_\Omega (\phi-1)^2 \Psi^+\left(\frac{\partial \underline{u}}{\partial \underline{x}}\right) + \Psi^-\left(\frac{\partial \underline{u}}{\partial \underline{x}}\right) + \frac{G}{2\ell}\left(\phi^2 + \ell^2 \frac{\partial \phi}{\partial x_i}\frac{\partial \phi}{\partial x_i}\right) d\Omega \\ - \int_\Omega b_i u_i d\Omega - \int_{\partial\Omega} t_i u_i dS$$

We ignore load stepping here for simplicity, though it is simple to factor in if one wishes to make use of the strong form. We also assume going forward that the dependence of the strain energies on the displacement gradient is implicit. The strong form of the governing equations is derived by computing the condition for a minimum of the total potential with the calculus of variations. This reads

$$\delta\Pi = 0 = \int 2(\phi-1)\Psi^+\delta\phi + (\phi-1)^2 \frac{\partial \Psi^+}{\partial\left(\frac{\partial u_i}{\partial x_j}\right)}\frac{\partial \delta u_i}{\partial x_j} + \frac{\partial \Psi^-}{\partial\left(\frac{\partial u_i}{\partial x_j}\right)}\frac{\partial \delta u_i}{\partial x_j} \\ + \frac{G}{\ell}\phi\delta\phi + G\ell\frac{\partial\phi}{\partial x_i}\frac{\partial\delta\phi}{\partial x_i} - b_i\delta u_i d\Omega - \int_{\partial\Omega} t_i \delta u_i dS$$

Note that the two variations δu_i and $\delta\phi$ are independent. The sets of terms they multiply must be zero independently. Grouping terms and integrating by

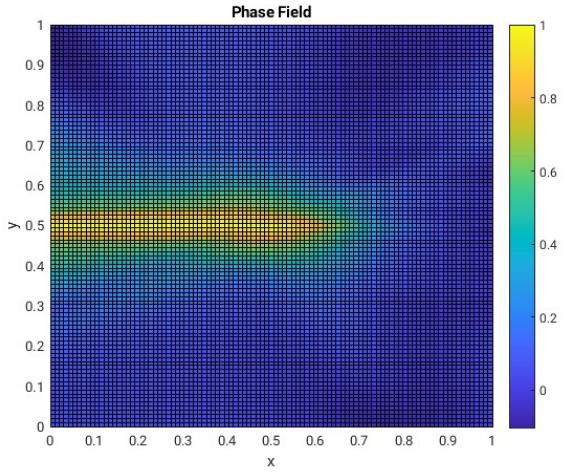

Figure 8: Heat map of the phase field computed by minimizing Γ for a notch of length $1/2$ and length scale $\ell = 0.1$. The crack area functional behaves in 2D as we expect–its minimum corresponds to sharp crack geometries, and its value corresponds to the total area of cracks formed.

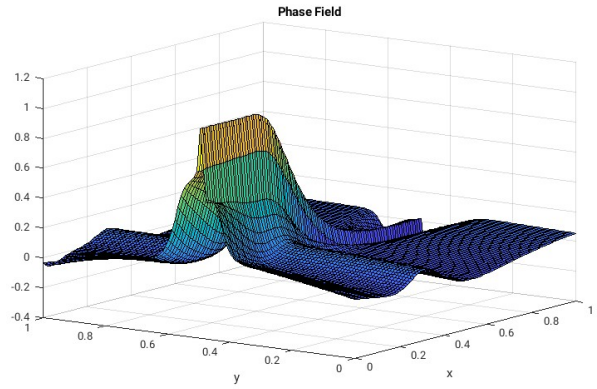

Figure 9: Phase field computed by minimizing the crack area functional for a notch of length $a = 1/4$ and length scale of $\ell = 0.05$. Evaluating Γ at this solution gives a crack area of 0.4, which again is slightly more than the notch length. This seems reasonable in light of the fact that the continuous interpolation spreads the crack over a larger region than the notch.

parts the spatial gradients off the test functions, we obtain two separate integral equations

$$\int \left(\frac{\partial}{\partial x_j} \left((\phi - 1)^2 \frac{\partial \Psi^+}{\partial \left(\frac{\partial u_i}{\partial x_j} \right)} + \frac{\partial \Psi^-}{\partial \left(\frac{\partial u_i}{\partial x_j} \right)} \right) + b_i \right) \delta u_i d\Omega = 0$$

$$\int \left(2(\phi - 1)\Psi^+ + \frac{G}{\ell}\phi - G\ell \frac{\partial^2 \phi}{\partial x_i \partial x_i} \right) \delta \phi d\Omega = 0$$

As is standard in the calculus of variations, we argue that because the test functions (or variations) are arbitrary, what they multiply under the integral must be zero pointwise. This leads to the strong form of the governing equations for the anisotropic phase field problem:

$$\frac{\partial}{\partial x_j} \left((\phi - 1)^2 \frac{\partial \Psi^+}{\partial \left(\frac{\partial u_i}{\partial x_j} \right)} + \frac{\partial \Psi^-}{\partial \left(\frac{\partial u_i}{\partial x_j} \right)} \right) + b_i = 0$$

$$\frac{G}{\ell} \left(\phi - \ell^2 \frac{\partial^2 \phi}{\partial x_i \partial x_i} \right) = 2(1 - \phi)\Psi^+$$

Remembering that derivatives of the strain energy density with respect to the displacement gradient (strain) give stresses, the first equation, which is the balance of linear momentum, says that the stress degrades with the damage, but only for the part of the strain energy associated with fracture. In other words, larger strains generate smaller stresses as the damage increases. Intuitively, this makes sense. The second equation shows that that the damage is driven by the strain energy associated with fracture, and the driving force decreases with increasing damage. It seems that the Laplacian of the damage acts as a sort of diffusive regularization to prevent the formation of zero width bands of damage.

5 A Particular Numerical Implementation

A fully-connected neural network is used to discretize the solution of the phase field fracture model. We will have separate networks for the displacement u and phase field ϕ. In Figure 10, the neural network discretization of the displacement is shown. The spatial coordinates are taken in at the input layer, and are passed through a series of affine transformations and nonlinear activation steps, per the architecture of a traditional fully-connected network. The neural network parameters build the functional relationship between the spatial coordinate and the solution. Note that unlike a traditional discretization strategy, the solution is a nonlinear function of the parameters. Because the neural network is defined globally over the computational domain, it is a kind of spectral discretization. Remember that the phase field is constrained to $\phi \in [0, 1]$. This constraint can be enforced quite simply with the neural network approach. We can simply pass the output of the neural network discretizing the phase field through a function whose range is in this interval. One such example of this is

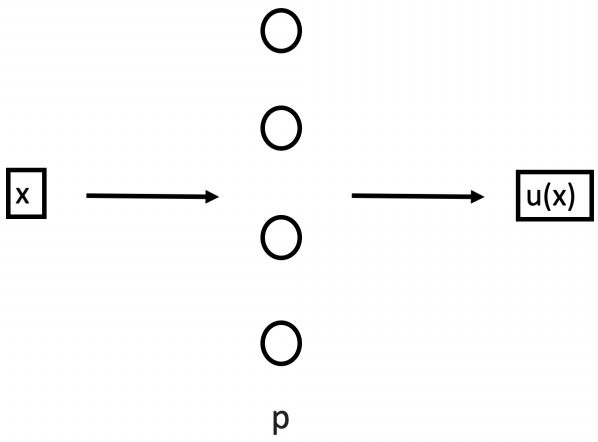

Figure 10: The spatial coordinate(s) x are taken as an input to a multi-layer perceptron neural network which discretizes the PDE solution $u(x)$. The "degrees of freedom" of the discretization are the weights and biases of the neural network, which are collected into a vector $\underline{p}_u$. A separate neural network of the same form, but with parameters $\underline{p}_\phi$ is used to discretize the phase field. The parameters of both networks define the solution the fracture problem, and are collected into a single vector $\underline{\theta} = [\underline{p}_u, \underline{p}_\phi]^T$.

$$\phi = \frac{1}{2}\Big(1 + \tanh\Big(\tilde{\phi}(x, y; \underline{p})\Big)\Big)$$

where $\tilde{\phi}$ is a fully-connected neural network for the phase field which does not necessarily respect the normalization condition. The function ϕ on the other hand respects the normalization condition by construction. This is how the phase field will be discretized. No such operation is necessary for the displacement.

Using neural network discretizations of the solution, the Deep Ritz Method makes use of the variational form of a partial differential equation to compute the parameters by minimizing the total potential energy. When a PDE is variational, there is an associated "energy" functional whose minimum corresponds to a solution of the strong form of the governing equation. Not all PDE's are variational, though many problems of engineering interest such as heat conduction, linear elasticity, and nonlinear elasticity, have associated variational principles. As we have seen, this is also the case for the phase field model of fracture. A generic energy functional Π depends on the solution to the PDE through the parameters of the neural network. This can be written as

$$\Pi\Big(u(x;\underline{p})\Big) = \int_\Omega f\Big(u(x;\underline{p}), \frac{\partial u}{\partial x}(x;\underline{p})\Big) d\Omega$$

where u is the solution, $\partial u/\partial x$ is the spatial gradient, Ω is the domain over which the solution is defined, f is an "energy density," and $\underline{p}$ are the neural network parameters. Different energy densities will correspond to different physical models. Because the solution is discretized, a minimum of the energy can be computed by taking its gradient with respect to the neural network parameters. This reads

$$\frac{\partial \Pi}{\partial p_j} = \int_\Omega \frac{\partial f}{\partial u}\frac{\partial u}{\partial p_j} + \frac{\partial f}{\partial(\frac{\partial u}{\partial x})}\frac{\partial^2 u}{\partial x \partial p_j} d\Omega = 0$$

This is a system of nonlinear equations, even when the underlying physical model is linear. This is because the $\partial u/\partial p_j$ is the gradient of a neural network with respect to its parameters, which is a nonlinear function. This contrasts with traditional spectral or finite element discretizations, which lead to linear systems for linear PDE's. When treated as a system of equations, this corresponds to the weak form of the governing equations, where $\partial u/\partial p_j$ are the test functions. Due to its nonlinearity, the neural network approximates the solution on a manifold, and this nonlinear test function is a local tangent. This is Galerkin orthogonality for a nonlinear solution approximation: the PDE residual is orthogonal to the local tangent of the approximation space. Instead of explicitly solving the nonlinear system of equations, we can use the fact the system expresses a condition for a minimum and use the gradient as a search direction in an optimization framework. This avoids computing higher derivatives of the system, which would be required to form the Jacobian matrix in a Newton-Raphson method, for example. Thus, we compute the parameters such that the energy takes on a minimum, where the gradient is used as a search direction at each step in the optimization.

Remember that we need to gradually ramp up the loads in order to compute a physically meaningful solution to the fracture problem. With the Deep Ritz method, we can simply find the minimum of the total potential at load step k using the previous converged solution as an initial guess. The energy for the anisotropic phase field model at load step k is

$$\Pi^k = \int_\Omega (\phi^k - 1)^2 \Psi^+\left(\frac{\partial \underline{u}^k}{\partial \underline{x}}\right) + \Psi^-\left(\frac{\partial \underline{u}^k}{\partial \underline{x}}\right) + \frac{G}{2\ell}\left(\phi^{2,k} + \ell^2 \frac{\partial \phi^k}{\partial x_i}\frac{\partial \phi^k}{\partial x_i}\right) - b_i^k u_i^k d\Omega - \int_{\partial\Omega} t_i^k u_i^k dS \quad (1)$$

For the Deep Ritz method, any optimization algorithm will require the gradient of this expression. Remember that the displacement and phase field and parameterized by neural networks. Collecting the parameters for both field variables into a single vector $\underline{\theta}$, we can write the gradient of the energy as

$$\frac{\partial \Pi^k}{\partial \theta_j} = \int_\Omega 2(\phi^k - 1)\frac{\partial \phi^k}{\partial \theta_j}\Psi^+ + (\phi^k - 1)^2 \frac{\partial \Psi^+}{\partial \epsilon_a^k}\frac{\partial \epsilon_a^k}{\partial \theta_j} + \frac{\partial \Psi^-}{\partial \epsilon_a^k}\frac{\partial \epsilon_a^k}{\partial \theta_j} + \frac{G}{\ell}\phi^k \frac{\partial \phi^k}{\partial \theta_j} + G\ell \frac{\partial \phi^k}{\partial x_i}\frac{\partial^2 \phi^k}{\partial x_i \partial \theta_j} - b_i^k \frac{\partial u_i^k}{\partial \theta_j} d\Omega - \int_{\partial\Omega} t_i^k \frac{\partial u_i^k}{\partial \theta_j} dS$$

This is a bit of a mess. There is no way around it. We have assumed that the strain, which is a function of the displacement at the current load step, is written in vector form as is common in finite element implementations. The spatial and parameter gradients should be computed with symbolic or automatic differentiation. Loading steeping is accomplished by driving the problem with different body forces and/or tractions at each step. Using the solution from the past load step to initialize the current load step will expedite convergence. The final aspect of implementation that needs to be discussed is enforcing crack irreversibility. There are two main approaches to this. The first is very simple: just add irreversibility as a constraint to the energy minimization. For optimization algorithms which handle constraints, we supplement the objective with an inequality constraint of the form

$$g(\underline{\theta}^k) = \phi(x_i; \underline{\theta}^{k-1}) - \phi(x_i; \underline{\theta}^k) \leq 0$$

where the set of spatial points x_i are often the integration points. This constraint says that the updated phase field at the current load step must be greater than or equal to the phase field at the past load step at every integration point. This is the most obvious method, but it may be desirable in some cases to avoid a constrained optimization problem. The alternative approach makes use of a history variable:

$$\mathcal{H}^k(x) = \max(\Psi^{+,k-1}(x), \Psi^{+,k-2}(x), \dots)$$

The history variable is the maximum value of the positive strain energy over the whole load stepping history at each point in the structure. Like the irreversibility constraint, it is evaluated at integration points in practice. The history variable replaces Ψ^+ as driving force of crack growth in the phase field governing equation, whose strong form we derived in the previous section. This ensures that the "force" causing crack growth never decreases, and thus the damage does not reverse. Unfortunately, this method is less amenable to an energy formulation of the problem, as there is not an associated variational principle with this new governing equation for the damage. In other words, the history variable approach is no well-suited for use with the Deep Ritz method, though it is quite convenient for any solver making use of the weak form of the governing equations. This method has been briefly outlined only for the sake of completeness. In summary, a Deep Ritz implementation of the phase field model of fracture requires: 1) a neural network discretization of the displacement and phase field (which may be done with one or two networks), 2) a method to compute parameter gradients of the energy (automatic or symbolic

differentiation), 3) a method to compute integrals over the computational domain to build the energy and its gradient at given parameter values, and 4) an optimization algorithm that can handle the irreversibility constraint.