Micro-macro mechanics of damage and healing in rocks

Chloé Arson

Abstract. This paper presents the state of the art of the theory of rock damage and healing mechanics, with a particular emphasis on the strategies available to relate the micro-scale of crystals, cracks and pores to the scale of a Representative Elementary Volume (REV). We focus on mechanical degradation and recovery of stiffness and strength. Damage and healing models formulated in the author's group are used as examples to illustrate and compare the reviewed micro-macro approaches, which include fabric enrichment, micromechanical formulations and homogenization schemes. This manuscript was written for doctoral students or researchers relatively new to the field of damage mechanics of geomaterials. Equations are provided to explain how to formulate a thermodynamically consistent model from scratch. Reviewing damage and healing modeling strategies led to the following conclusions: (i) The framework of hyperplasticity, which does not require any postulate on the existence or uniqueness of yield functions and which automatically ensures thermodynamic consistency, was never applied to Continuum Damage Mechanics (CDM). There may be an avenue to improve state-of-the-art damage and healing models in a similar framework of “hyper-damage mechanics”. (ii) In damage softening models, the mesh dependence of the width of the damage localization zone is currently alleviated by non-local regularization. Perhaps the next step is to couple micro-macro damage and healing models at the REV scale to discrete fracture mechanics at a larger scale to understand how damage and healing localization occurs. (iii) There may be an opportunity to use fabric-enriched models to capture the effect of microstructure organization on both mechanical properties and permeability. (iv) Coupling chemomechanical damage and healing processes across the scales would be useful to model the competition between damage and healing whenever both can occur at the same temperature and pressure conditions. (v) Many challenges still exist to implement healing models in the Finite Element Method, especially in regards to the mapping of net damage.

Keywords. mechanics, healing processes, rock mechanics, fabric, micro-mechanics, homogenization, fracture mechanics, thermodynamics
1. Introduction

In rock mechanics, damage refers to the degradation of macroscopic properties, usually stiffness and strength, due to the propagation of cracks. Rocks are endowed with so-called intrinsic cracks that are modeled as initial damage or initial fabric, and have anisotropic stiffness and strength properties, which depend on the orientation of the bedding. Damage is a phenomenological variable that is used to quantify the loss of elastic deformation energy. Its evolution depends on postulates on how the energy dissipates at the macroscopic scale of a Representative Elementary Volume (REV) and/or at the microscopic scale of the cracks. Healing, in this context, refers to the recovery of mechanical properties due to crack rebonding. In this paper, we will focus on self-healing, which happens in rock without the external input of chemicals to seal the cracks.

Continuum Damage Mechanics (CDM) models were initially formulated to predict the degradation of stiffness and strength properties of metals. With the increasing energy demand and the important environmental issues that arose from waste management, rocks became an interesting object of study in civil engineering, fracture mechanics and material science - not only in geology. Compared to fracture mechanics, CDM is seen as a computationally efficient framework to predict the changes of stress and deformation at the bulk scale, which is well suited for engineering design. The first attempts to extend CDM to rock mechanics date back to the 1990's, with the works of French researchers like Dragon and Chaboche, who grounded their formulations on micromechanical analyses presented by Kachanov in the 1980's. Two damage mechanics schools emerged: the first puts emphasis on the thermodynamic consistency of phenomenological laws expressed at the REV scale and the second focuses on fracture micromechanics and homogenization-based upscaling methods. In the late 1990's, models were improved to account for anisotropic unilateral effects induced by crack closure. In the early 2000's, researchers coupled CDM models to the Biot's theory to account for hydro-mechanical couplings, and later, to plasticity models. In the 2010's, the first theories of damage and healing mechanics were born, with the pioneering works of Ju and Voyiadjis in particular. However, the current mathematical frameworks do not yet allow accounting for both crack closure and crack rebonding, and require using tensor variables that are unrelated to microstructure. As a result, evolution laws are expressed at the bulk scale and are disconnected from the physical processes that drive crack healing and the subsequent changes of stiffness and permeability.

Here, we explain how to couple Continuum Damage and Healing Mechanics to microstructure evolution, in order to understand why rocks break and why some rocks recover from mechanical damage, and why some do not. We take two rocks as example materials: granite and halite. Granite is considered a good candidate for hosting nuclear waste disposals due to its high strength and stiffness properties. It is also a rock extensively studied by geomorphologists to understand how soil forms. One thesis is that tectonic stresses drive fracture propagation in the granite bedrock, hence the formation of saprolite, which then degrades into soil. Another thesis is that saprolite production is controlled by chemical weathering of Fe-bearing minerals that expand as they weather, which creates stresses that are sufficient to fracture rock. Salt rock is an attractive host for deep waste disposals and gas storage, due to favorable creep characteristics and low gas permeability. Under favorable temperature and stress conditions, cracks in salt rebond due to diffusive mass transfer and pressure solution. Although polycrystalline halite is known for deformation by isochoric dislocation and diffusion processes, cracking is also an important grain-scale process at lower mean stress. The relationship between these micro-mechanisms and macroscopic strain evolution is still not well understood, especially in transient states. How and why do complex macroscopic phenomena (e.g., chemical weathering, damage, healing, creep) emerge from elementary microscopic processes?

In this paper, we examine these fundamental scientific questions through an extensive literature review that presents the basic principles of damage and healing mechanics, compares state-of-the-art modeling approaches and summarizes the contributions made by the author over the past ten years. This paper was written for doctoral students or researchers relatively new to the field of damage mechanics of geomaterials. Equations are provided to explain how to formulate a thermodynamically consistent model from scratch. In Section 2, we present the theory of CDM, its extension to healing processes and the strategies available to address softening behavior and residual strains. In Section 3, the damage/healing variable is defined as a convolution of moments of probability of microstructure descriptors. Evolution laws are postulated at the macroscopic scale and allow predicting microstructure changes. In Section 4, crack propagation is modeled based on fracture mechanics and the REV energy potentials are obtained by superposition, through an integration of elementary potentials over all the possible crack orientation angles. In Section 5, the damage-healed stiffness tensor is obtained by solving a matrix-inclusion problem. For each section, we provide a detailed literature review and we present examples taken from research done in the author's group.

2. Phenomenological models of damage and healing

In phenomenological CDM, the relationship between the damage variable $D$ and the stiffness tensor of the REV stems from mathematical requirements on the stiffness tensor. The damage evolution laws are expressed at the REV scale, on the basis of constitutive postulates. Most CDM models are constructed on the basis of hyper-elasticity, in which the damaged elastic stress/strain relationship derives from an energy density function. The REV free energy is expressed as a polynomial function of $D$ and of either stress $\sigma$ or deformation $\epsilon$, so as to ensure stiffness positiveness, symmetry
and constitutive requirements. The damage threshold function and the damage evolution law depend on constitutive assumptions and they can take any form, as long as the model is thermodynamically consistent. The damage variable is usually a scalar, a vector (containing components of tensile and compressive damage for instance) or a second-order tensor. Higher-order damage tensors were seldom used, to account for crack interaction [Kachanov, 1992, Lubarda and Krajcinovic, 1994], to represent the effects of sliding cracks [Halm and Dragon, 1998] or to model damaged materials with little symmetries [Cowin, 1985, Zysset and Curnier, 1995].

2.1. Free energy

Consider a second-order damage tensor $D = kF$, in which $k = Tr(D)$ and $F$ is a trace-less second-order fabric tensor. The most general form of the REV Helmholtz free energy density $\psi$ that ensures that the undamaged stress/strain relationship is linear is given by:

\[
\psi = \psi(e, k, F) = \frac{c_1}{2} Tr^2(e) + \frac{c_2}{2} Tr(e \cdot e) + \frac{c_3}{2} Tr^2(e \cdot F) + \frac{c_4}{2} Tr(e \cdot e + \frac{c_5}{2} Tr^2(e \cdot F) + \frac{c_6}{2} Tr((e \cdot F)^2)
\]

where the $c_i$ are functions of $k$ and of the two invariants of $F$, the expressions of which are given in [Zysset and Curnier, 1995]. Often, $D$ is defined as the second-order crack density tensor, the components of which are given by:

\[
D_{ij} = \sum_{k=1}^{N} \rho^{(k)} n_i^{(k)} n_j^{(k)}
\]

with $\rho^{(k)}$ the volume fraction of the $k^{th}$ family of cracks and $n^{(k)}$ the direction orthogonal to the plane of the $k^{th}$ crack family. If cracks do not interact, the expression of the free energy does not contain any damage term of order two or higher [Halm and Dragon, 1998].

The expression of the damaged stiffness tensor $\tilde{C}$ is derived from Equation 1, as follows:

\[
\tilde{C} = \tilde{\sigma}^2 \psi(e, k, F)
\]

In some models, the expression of the damaged (or damaged-healed) elasticity tensor is postulated and the expression of the free energy is calculated $a posteriori$ to check the thermodynamic consistency of the model [Arson et al., 2012, Chaboche, 1992, 1993, Voyiadis and Kattan, 2017]. If the damage variable $D$ is a second-order tensor, a fourth-order damage operator $M(D)$ is introduced to define a damage effective stress $\tilde{\sigma}$ [Lemaitre and Desmorat, 2005]:

\[
\tilde{\sigma} = M(D) : \sigma
\]

$\tilde{\sigma}$ can be understood as the stress that acts on the effective surface area of the cross section of a damaged material. The simplest operator that satisfies the symmetry requirements for the stress tensor is the operator proposed by Cordebois and Sidoroff [1982]:

\[
\tilde{\sigma} = \frac{1}{2}(\delta - D)^{-1/2} \cdot \sigma \cdot (\delta - D)^{-1/2}
\]

in which $\delta$ is the second-order identity tensor. The stiffness tensor is then obtained by applying the Principle of Equivalent Elastic Strain (PEES) or the Principle of Equivalent Elastic Energy (PEEE) [Hansen and Schreyer, 1994]. The PEES states that the elastic strain $\varepsilon^E$ induced by the actual stress $\sigma$ applied to the damaged material (of stiffness $\tilde{C}$) is equal to the effective elastic strain $\varepsilon^E$ induced by the effective stress $\tilde{\sigma}$ applied to the intact (non-damaged) material (of stiffness $C_0$):

\[
\varepsilon^E = \tilde{C}^{-1} : \sigma = C_0^{-1} : \tilde{\sigma} = \varepsilon^E
\]

Substituting Equation 4 into Equation 6, one gets:

\[
\tilde{C} = M^{-1}(D) : C_0
\]

The PEEE states that the elastic strain energy stored by the damaged material under the actual stress is equal to that stored by the intact material under the effective stress:

\[
\frac{1}{2} \sigma : \tilde{C}^{-1} : \sigma = \frac{1}{2} \tilde{\sigma} : C_0^{-1} : \tilde{\sigma},
\]

from which one gets:

\[
\tilde{C} = M^{-1}(D) : C_0 : M^{-T}(D)
\]

For some damage operators, the damaged stiffness tensor obtained by applying the PEEE is not symmetric. That is why the PEEE is usually preferred.

2.2. Damage evolution and thermodynamic consistency

The damage evolution law is constrained by the requirement of dissipation positiveness. This requirement depends on the damage work-conjugate variable, called the “damage driving force” or the “energy release rate”. The expression of the damage driving force is obtained through thermodynamic conjugation relationships, by differentiating the expression of the free energy density by damage. The subsection below explains the main steps of the derivation and recalls the thermodynamic consistency requirements.

2.2.1. Thermodynamic principles

The first law of thermodynamics expresses a principle of energy conservation. Let us ignore pore fluids in this review section, and let us consider a solid $\Omega$. The energy of $\Omega$ is conserved if, and only if, any variation of the sum of the internal energy and kinetic energy of $\Omega$ during an increment of time remains equal to the sum of the work input and heat exchange during that time increment:

\[
K + \dot{E} = W + Q
\]

in which $K$, $E$, $W$ and $Q$ are respectively the rates (material time derivatives) of kinetic energy, internal energy, mechanical work and heat exchange. Now introducing the mass density $\rho$, the local velocity $v$ and the internal energy density $e$, Equation 10 becomes:

\[
\frac{D}{Dt} \int_{\Omega} \rho \left( \frac{1}{2} \rho \dot{v} \cdot \dot{v} + e \right) d\Omega = \int_{\Gamma} (\rho \dot{\mathbf{b}} \cdot \mathbf{v} + r) d\Gamma + \int_{\Omega} (\mathbf{t} \cdot \mathbf{v} - q) d\Omega
\]

Open Geomechanics, 2020, article no. 1
Chloé Arson, Micro-macro mechanics of damage and healing in rocks

3
In which \( \mathbf{b} \) is the body force applied to \( \Omega \) per unit mass, \( \mathbf{t} \) is the traction force applied to \( \Omega \) per unit area, \( \mathbf{q} \) is the outward heat flux through the boundary \( \Gamma \) of the system \( \Omega \) and \( r \) is the rate of heat generation per unit mass. We note \( \mathbf{n} \) the unit vector normal to \( \Gamma \). Using the Gauss divergence theorem, the mass conservation equation and the momentum balance equation, we obtain the following form of the first law of thermodynamics:

\[
\rho \dot{\mathbf{e}} = \mathbf{a} : \mathbf{n} - \nabla \cdot \mathbf{q} + r
\]

in which the small strain assumption \( \dot{\mathbf{e}} = \nabla \cdot \dot{\mathbf{u}} \) is used with \( \mathbf{a} \) denoting the Cauchy stress tensor.

The second law of thermodynamics states that energy is dissipated during irreversible processes: the rate of entropy increase of \( \Omega \) is never less than the rate of entropy increase due to the heat source \( r \) and the heat flux \( \mathbf{q} \):

\[
\frac{\mathcal{D}S}{\mathcal{D}t} \geq - \int_{\Omega} r \, d\Omega - \int_{\Gamma} \mathbf{q} \cdot \mathbf{n} \, d\Gamma
\]

in which \( T \) is the absolute temperature and \( S \) is the entropy of \( \Omega \). Using the Gauss divergence theorem, Equation 13 becomes:

\[
\rho \dot{s} + \nabla \left( \mathbf{q} \right) - r \geq 0
\]

in which \( s \) is the entropy per unit mass.

Combining Equations 12 and 14, and introducing the definition of the Helmholtz free energy:

\[
\psi = e - Ts,
\]

we obtain the so-called Clausius-Duhem inequality, as follows:

\[
\mathbf{a} : \dot{\mathbf{e}} - \rho \left( \dot{\psi} + s \dot{T} \right) - \mathbf{q} \cdot \nabla T \geq 0
\]

2.2.2. Thermodynamic conjugation relationships

Helmholtz free energy depends on state variables (elastic strain and temperature for solids and dry porous media, plus pore pressures for fluid-saturated media) and on internal variables (such as damage, healing and hardening variables). We note \( \chi \) the vector of internal variables. The rate of Helmholtz free energy density is:

\[
\dot{\psi} (\mathbf{E}, T, \chi) = \frac{\partial \psi}{\partial \mathbf{E}} : \dot{\mathbf{E}} + \frac{\partial \psi}{\partial T} : \dot{T} + \frac{\partial \psi}{\partial \chi} : \dot{\chi}
\]

Combining equations 16 and 17, we get:

\[
\left( \mathbf{a} - \rho \frac{\partial \psi}{\partial \mathbf{E}} \right) : \dot{\mathbf{e}} + \sigma : \dot{\mathbf{e}}^{irr} - \rho \left( s + \frac{\partial \psi}{\partial T} \right) \dot{T} - \rho \frac{T}{\partial \chi} : \dot{\chi} - \mathbf{q} \cdot \nabla T \geq 0
\]

in which \( \dot{\mathbf{e}}^{irr} \) is the irreversible strain tensor; \( \dot{\mathbf{e}} = e - \dot{\mathbf{E}} \).

For isothermal reversible processes, the Clausius-Duhem inequality is an equality, from which we get:

\[
\sigma = \rho \frac{\partial \psi}{\partial \mathbf{E}}
\]

Similarly, for non-isothermal irreversible processes, the inequality is an equality from which we get:

\[
s = - \frac{\partial \psi}{\partial T}
\]

As a result, the dissipation inequality reduces to:

\[
\phi := \mathbf{a} : \dot{\mathbf{e}}^{irr} - \rho \frac{\partial \psi}{\partial \chi} : \dot{\chi} - \mathbf{q} \cdot \nabla T \geq 0
\]

in which \( \phi \) is the density of energy dissipation.

Closing the formulation of a phenomenological CDM model requires not only the expression of the free energy density, but also, the introduction of evolution laws for each of the internal variables. These evolution laws ought to represent the constitutive behavior of the system under study (\( \Omega \)) and must satisfy the positiveness of the dissipation functional \( \phi \) - a condition known as thermodynamic consistency.

2.2.3. Thermodynamic consistency

A sufficient condition to ensure thermodynamic consistency in Equation 21 is to express the evolution laws by differentiating a non-negative convex dissipation potential \( \phi \) that depends on the tensor of generalized force \( \left( \mathbf{a}, -\rho \frac{\partial \psi}{\partial \chi}, -\frac{\nabla T}{T} \right) \):

\[
\dot{\mathbf{e}}^{irr} = \frac{\partial \phi}{\partial \mathbf{a}}, \quad \dot{\chi} = -\frac{\partial \phi}{\partial \rho \frac{\partial \psi}{\partial \chi}}, \quad \dot{\mathbf{q}} = -\frac{\partial \phi}{\partial \frac{\nabla T}{T}}
\]

In rate-dependent CDM, evolution laws are often chosen empirically. It is possible however to ensure thermodynamic consistency by expressing \( \phi \) as the sum of a force potential \( z \) (the true stress potential) and a flow potential \( w \) (conjugate to \( z \) relative to the rate of internal variable). For more details, the reader is referred to [Houlsby and Puzrin, 2007] (Chapter 11).

In rate-independent CDM, a damage potential \( g_d \) is defined as a function of the generalized stress vector and of the internal variables, such that \( g_d = 0 \) [Houlsby and Puzrin, 2007]. Ensuring the maximum entropy production between two equilibrium states then boils down to minimizing the following Lagrangian functional [Hansen and Schreyer, 1994]:

\[
\mathcal{L} = -\phi + \dot{\lambda}_d g_d
\]

in which \( \dot{\lambda}_d \) is the damage multiplier and \( g_d \) is a non-negative convex function. A sufficient condition to minimize \( \mathcal{L} \) is to satisfy the following conditions:

\[
\frac{\partial \mathcal{L}}{\partial \mathbf{a}} = 0, \quad \frac{\partial \mathcal{L}}{\partial \rho \frac{\partial \psi}{\partial \chi}} = 0, \quad \frac{\partial \mathcal{L}}{\partial \frac{\nabla T}{T}} = 0,
\]

which provides

\[
\dot{\mathbf{e}}^{irr} = \dot{\lambda}_d \frac{\partial g_d}{\partial \mathbf{a}}, \quad \dot{\chi} = -\dot{\lambda}_d \frac{\partial g_d}{\partial \rho \frac{\partial \psi}{\partial \chi}}, \quad \dot{\mathbf{q}} = -\dot{\lambda}_d \frac{\partial g_d}{\partial \frac{\nabla T}{T}}
\]

In classical CDM, it is postulated that the damage function \( f_d \) (equivalent to the yield function in plasticity) exists and that it is unique. \( f_d \) is typically a function of the generalized force \( \left( \mathbf{a}, -\rho \frac{\partial \psi}{\partial \chi}, -\frac{\nabla T}{T} \right) \) and of the internal variables \( \left( \dot{\mathbf{e}}^{irr}, \dot{\chi}, \dot{\mathbf{q}} \right) \). The evolution laws are said to be associate if \( f_d = g_d \), i.e., if the flow of internal variables is orthogonal to the damage surface \( f_d \). In this case, the damage multiplier can be found by using the consistency conditions:
\[ f_d = \dot{f}_d = 0, \text{ from which we typically get:} \]
\[ \frac{\partial f_d}{\partial \sigma} : a + \frac{\partial f_d}{\partial \rho \frac{\partial \psi}{\partial \chi}} : \frac{D}{D t} \left( \rho \frac{\partial \psi}{\partial \chi} \right) + \frac{\partial f_d}{\partial \chi} : \frac{\partial D}{\partial T} \left( \nabla T \right) = 0 \]  
(26)

It was shown by Desmorat [2006] that, for symmetric second-order damage tensors and for particular forms of the Helmholtz free energy that are continuously differentiable, dissipation positiveness is ensured if, and only if, the rate of damage is non-negative, i.e. the rates of damage eigenvalues are non-negative.

In hyper-plasticity [Collins and Houlsby, 1997, Houlsby and Puzrin, 2007], it is assumed that the generalized stress obtained by differentiating the free energy by the internal variable is equal to the general dissipative stress obtained by differentiating the dissipation by the rate of internal variable. The existence and uniqueness of the yield function follows from that assumption. The flow rules are associative in the sense of generalized stress but they can be non-associative in the sense of true stress. The formulation requires only two postulates: the expression of a free energy potential and the expression of the dissipation or of the yield function in generalized stress. By construction, the yield function in generalized stress is locally convex, which automatically ensures the thermodynamic consistency of the model. The principles of hyper-plasticity can be applied to CDM to ensure thermodynamic consistency with a minimum of postulates. Let us define the damage generalized stress as:
\[ \chi^d := - \frac{\partial \psi}{\partial D} \]  
(27)
In the absence of irreversible deformation and under isothermal conditions, the dissipation is expressed as:
\[ \phi = \chi^d : D \]  
(28)
Most models are formulated with a homogeneous dissipation function of order one in the damage rate, which ensures thermodynamic consistency [Collins and Houlsby, 1997]. Assuming that \( \phi \) is homogeneous or order one in \( D \), we have:
\[ \phi = \frac{\partial \phi}{\partial D} : D, \quad \chi^d := \frac{\partial \phi}{\partial D} \]  
(29)
in which \( \chi^d \) is the damage generalized dissipation stress. Equations 28 and 29 yield:
\[ \left( \chi^d - \chi^d \right) : D = 0 \]  
(30)
which indicates that \( \chi^d - \chi^d \) should be normal to the direction of damage flow. Applying the principles of hyperplasticity to CDM, we assume:
\[ \chi^d = \chi^d \]  
(31)
The Legendre transform \( \xi \) of \( \phi \) is a degenerate form that is identically equal to zero:
\[ \phi + \xi = \chi^d : D, \quad \xi = 0 \]  
(32)
\( \xi = 0 \) indicates the “yield” (or damage) threshold. The damage function is defined only to within an arbitrary multiplicative constant \( \lambda_d \), which plays the role of a damage multiplier: \( \lambda = \lambda_d f_d \). The flow rule, associate in the sense of generalized stress, is obtained from Equation 32:
\[ D = \lambda_d \frac{\partial f_d}{\partial \chi^d} \]  
(33)
The damage multiplier is then obtained from the consistency condition: \( f_d = \dot{f}_d = 0 \). Some constraints on strains or damage can be included in the model. For instance, if a constraint is imposed on the damage rate: \( c(D) = 0 \), a Lagrange multiplier \( \lambda' \) is introduced to define the modified dissipation potential:
\[ \phi' = \phi + \lambda' c(D) \]  
(34)
and the damage generalized dissipation stress is defined as:
\[ \chi^d = \frac{\partial \phi'}{\partial D} = \frac{\partial \phi}{\partial D} + \lambda' \frac{\partial c}{\partial D} \]  
(35)
in which \( \lambda' \) is obtained from the condition \( \left[ \chi^d - \chi^d \right] = 0 \) and Equation 27, with a known free energy potential. Despite its attractiveness, the framework of hyper-plasticity was never applied to CDM. To the author’s best knowledge, the only paper that distinguishes between damage generalized stress and damage generalized dissipative stress is that of Darabi and collaborators [2012], but the proposed model is developed for viscous processes, for which there is no need for a damage “threshold” function.

2.3. Anisotropic damage criteria
Quasi-brittle materials such as concrete, rock, and ceramic composites, exhibit a complex mechanical behavior at the macro-scale, including stress-induced damage and stiffness anisotropy, non-linear stress/strain relationships and volumetric dilation, unilateral effects due to crack closure, and a transition from brittle to ductile behavior at increasing confining stress [Chiarelli et al., 2003, Krajcinovic et al., 1991]. All of these effects can be explained by the nucleation and propagation of micro-cracks at the grain boundaries and/or from pore spaces. CDM models are phenomenological in nature, which implies that damage tensors essentially measure the damaged mechanical effects rather than the microstructure evolution itself [Swoboda and Yang, 1999b]. As a result, damage evolution laws are arbitrarily crafted to match a macroscopic behavior (usually represented by stress/strain curves), and do not represent any clear physical mechanism. The behavior of brittle geomaterials depends on the sign of the applied stress/strain. Therefore, the damage driving force (i.e. the energy release rate that is work-conjugate to damage) has to be split into positive and negative components, which are introduced in two different damage criteria (one for tension, one for compression) [Comi et al., 2007, Frémond and Nedjar, 1996, Lubarda et al., 1994]. Although the singularities of the damage surfaces raise convergence issues in numerical simulation, CDM models were implemented in Finite Element Methods for practical engineering purposes [Jin et al., 2017, Xu and Arson, 2015] and were successfully used...
to predict damage-induced anisotropy and confinement-induced strengthening in rock subject to compression [Shao and Rudnicki, 2000, Shao et al., 2005]. However, multiple non-linear damage phenomena require more constitutive parameters that are often not related to any microstructure or mechanical property, which raises calibration challenges [Halm and Dragon, 1998]. Typically, damage evolution depends on distinct yield criteria and damage potentials [Comi and Perego, 2001, Zhu and Arson, 2014]. These so-called bi-dissipative models are based on complex mathematical formulations (challenging to implement in FEM) and depend on a large number of parameters (challenging to calibrate).

Stress and damage are not the only drivers of mechanical anisotropy. The mechanical behavior of rocks is strongly dependent on the orientation of the bedding, layering and crack patterns [Niandou et al., 1997, Sone and Zoback, 2013]. Rock strength depends on the confining pressure and the loading orientation with respect to microstructure [Cho et al., 2012, Heng et al., 2015, Nasser et al., 2003, Ye et al., 2016]. In CDM, intrinsic anisotropy (i.e. anisotropy of mechanical behavior in the absence of damage) is accounted for in the expression of the free energy and/or of the damage criterion. For instance, Hill [1948] extended the von Mises yield criterion to orthotropic ductile materials, by using six quadratic stress terms. To further account for the strength difference in tension and compression, Hoffman [1967] added three linear terms of stress into Hill’s failure criterion. Tsai and Wu [1971] then expressed failure criteria in terms of all the linear and quadratic stress terms. Other approaches were proposed to introduce projection tensors in the yield criteria or in the expression of the free energy [Boehler and Sawczuk, 1977, Cazacu et al., 1998, Rouabhi et al., 2007]. Cazacu and Cristescu [1999] extended the Mises-Schleicher yield criterion initially expressed for isotropic materials to transversely isotropic materials, by using the fourth order characteristic tensor. The microstructure tensor can also be constructed with eigenvectors representing the axes of material symmetry to capture the orientation dependence of strain hardening, softening, damage and plasticity [Chen et al., 2010, Pietruszczak et al., 2002, Pietruszczak and Mroz, 2000, 2001]. Thermodynamic models were also proposed, in which the free energy was expressed in terms of microstructure tensor and strain invariants [Halm et al., 2002, Nedjar, 2016].

Below, we present the damage criteria proposed by Jin and Arson [2018a] to model crack initiation and propagation in shale. Phenomenological damage evolution laws are constructed so as to distinguish the mechanical response of the material in tension and compression, along the direction perpendicular to the bedding plane and within the bedding plane. Inspired from the stress invariants used in Hill’s yield criterion [Hill, 1948] and in Hashin’s failure criterion for unidirectional fiber composites [Hashin, 1980], we introduce the following strain measures, which are strain invariants if axis 1 is normal to the bedding planes:

\[
\begin{align*}
I_1 &= \varepsilon_{11} \\
I_2 &= \varepsilon_{22} + \varepsilon_{33} \\
I_3 &= \frac{1}{4} (\varepsilon_{22} - \varepsilon_{33})^2 + \varepsilon_{23} \\
I_4 &= \varepsilon_{12}^2 + \varepsilon_{13}^2 \\
I_5 &= 2\varepsilon_{12}\varepsilon_{13}\varepsilon_{23} - \varepsilon_{22}\varepsilon_{33} - \varepsilon_{33}\varepsilon_{22}
\end{align*}
\]

where \( I_3 \) is the square of the maximum transverse shear strain while \( I_4 \) is the square of the maximum axial shear strain. Following the form of Hill’s and Hashin’s models, we choose a quadratic damage criterion. Based on the invariants defined above, the most general form of a transversely isotropic quadratic failure criterion is

\[
A_1 I_1^2 + A_2 I_2^2 + A_3 I_3 + A_4 I_4 + B_{12} I_1 I_2 = 1
\]

in which \( A_1, A_2, A_3, A_4 \) and \( B_{12} \) are material parameters. Noting \( x_1, x_2 \) and \( x_3 \) the directions of space, and assuming that the direction normal to the bedding plane is \( x_1 \), shale can fail in sliding mode, in which failure is controlled by the tensile and shear strength of the bedding planes, and in non-sliding mode, in which failure is controlled by the strength of the matrix material. Failure in sliding mode is controlled by strain components \( \varepsilon_{11}, \varepsilon_{12} \) and \( \varepsilon_{13} \). In non-sliding mode, the normal direction of the failure surface is contained in the bedding plane. Due to material isotropy in the bedding plane \((x_2 - x_3)\), failure in non-sliding mode is controlled by all strain components except \( \varepsilon_{11} \). Consequently, the two failure criteria are expressed in the following form:

\[
\begin{align*}
A_1 I_1^2 + A_4 I_4 &= 1 \\
A_3 I_3 + A_4 I_4 &= 1
\end{align*}
\]

for the sliding mode and the non-sliding mode, respectively. We define the equivalent strain measures as \( \varepsilon_{\text{eq}}^{\text{t}}/\kappa_1^0 = \sqrt{A_1 I_1^2 + A_4 I_4} \) and \( \varepsilon_{\text{eq}}^{\text{c}}/\kappa_3^0 = \sqrt{A_3 I_3 + A_4 I_4} \). Noting \( \kappa_0^0 \) (respectively \( \kappa_0^0 \)) and \( \kappa_1^0 \) (respectively \( \kappa_1^0 \)) the initial tensile and compressive strain thresholds for the sliding mode (respectively for the non-sliding mode), we have \( \kappa_1^0 \leq \varepsilon_{\text{eq}}^{\text{t}} \leq \kappa_0^0 \) and \( \kappa_1^0 \leq \varepsilon_{\text{eq}}^{\text{c}} \leq \kappa_0^0 \) in tension, \( \kappa_1^0 \leq \varepsilon_{\text{eq}}^{\text{c}} \leq \kappa_0^0 \) in compression. Two loading surfaces are used to distinguish micro-crack propagation in the axial and transverse directions. For tensile damage, we consider the two following damage criteria:

\[
\begin{align*}
g_{11}(\varepsilon, \kappa_1) &= \varepsilon_{\text{eq}}^{\text{t}} - \kappa_1 \\
g_{21}(\varepsilon, \kappa_2) &= \varepsilon_{\text{eq}}^{\text{c}} - \kappa_2
\end{align*}
\]

Similarly, we construct the two compressive loading surfaces in axial and transverse directions as:

\[
\begin{align*}
g_{12}(\varepsilon, \kappa_1) &= \varepsilon_{\text{eq}}^{\text{c}} + \eta((\sigma_2 + \sigma_3)/2) - \kappa_1 \\
g_{22}(\varepsilon, \kappa_2) &= \varepsilon_{\text{eq}}^{\text{t}} + \eta(\sigma_1) - \kappa_2
\end{align*}
\]

where \( \eta \) controls the influence of the confining stress on compressive damage. Note that the Mc Auley brackets are introduced to account for compressive confining stress only. \( \kappa_1 \) and \( \kappa_2 \) are hardening variables. Damage evolution laws
are functions of $\kappa_1$ and $\kappa_2$. Although calibrated for shale, the model can be used to solve a wide range of engineering problems involving the mechanical integrity of structural members, borehole stability, or delamination of composites. Anisotropy is accounted for at the microstructure scale and at the phenomenological scale of the REV. Damage constitutive laws are direction-specific, which makes it possible to represent several concurrent damage mechanisms in the macroscopic response, and to interpret the failure mechanisms that control the damage process zone.

2.4. Unilateral effects due to crack closure and friction

Consider for instance that the damage tensor has three eigenvalues $D_1$, $D_2$ and $D_3$ and three eigenvectors $n_1$, $n_2$ and $n_3$. Suppose that $D_1$ occurred because of a tensile stress in direction $n_1$. Imagine that this stress is relaxed, and then, that a compression is applied in direction $n_1$. Physically, it is expected that the compressive loading will cause the cracks normal to $n_1$ to close. As a result, the stiffness and the compressive strength of the REV should be recovered in direction $n_1$, but not in the other directions. Unilateral effects refer to the recovery of stiffness and strength in the directions normal to closed cracks. Chaboche [1993] proposed to account for unilateral effects by expressing the damaged stiffness tensor defined in Equation 3 as follows:

$$\tilde{C} := \tilde{C} + \eta \sum_{k=1}^{3} H(-Tr(P_k \cdot \epsilon)) P_k : (C_0 - \tilde{C}) : P_k \quad (41)$$

in which compression is counted negative. $C_0$ is the stiffness tensor of the REV in the absence of damage:

$$C_0 = \frac{d^2 \psi_0}{d\epsilon^2}, \quad \psi_0 = \frac{c_1}{2} Tr^2 (\epsilon) + \frac{c_2}{2} Tr(\epsilon \cdot \epsilon) \quad (42)$$

$H$ is the Heaviside jump function: $H(x) = 0$ if $x \leq 0$ and $H(x) = 1$ otherwise. $P_k = n_k \otimes n_k \otimes n_k$ is the projection operator in the principal base of the damage tensor, and $0 \leq \eta \leq 1$ is a constitutive parameter introduced to account for partial recovery (in case tensile damage is still effective upon crack closure).

The rotation of the principal base of damage induced by crack closure makes classical CDM frameworks inconsistent [Chaboche, 1992]. Most models proposed for mechanical crack closure are either isotropic, or restricted to mode I fracture propagation, e.g., [Ortiz, 1985]. An exception is the model proposed by Halm and Dragon [1998] for cohesive materials, which accounts for the recovery of the shear modulus due to friction. The density of Helmholtz free energy is written as:

$$\psi = \frac{c_1}{2} Tr^2 (\epsilon) + \frac{c_2}{2} Tr(\epsilon \cdot \epsilon) + \tilde{c}_1 Tr(\epsilon \cdot \epsilon) + \tilde{c}_2 Tr(\epsilon \cdot D) + \frac{3}{2} \sum_{k=1} H(-n_k \cdot n_k) [-\tilde{c}_1 D_2 \epsilon \cdot P_k : \epsilon - \tilde{c}_4 Tr(\epsilon \cdot D)] + \frac{3}{2} \sum_{k=1} H(-n_k \cdot n_k) [2\tilde{c}_4 Tr(\gamma \cdot D) - \tilde{c}_4 Tr(\gamma \cdot \epsilon \cdot D)]$$

in which $\tilde{c}_4 = c_4 / Tr(D)$, $\tilde{c}_2 = c_2 / Tr(D)$ and $\gamma$ is the second-order sliding variable, defined as $\gamma_{ij} = sym(\epsilon_{ik} n_k n_j)$ (where i,j,k denote the number of the damage eigenvector).

Later, a discrete approach was proposed to represent the effects of open and closed cracks with a minimum set of independent damage tensors of various orders [Bargellini et al., 2006]. Although elegant, this type of formulation was not used to solve boundary value problems. For the same level of complexity (and for a similar number of constitutive parameters to calibrate), micro-mechanical damage models are usually preferred because micro-mechanical variables have a physical meaning that can be interpreted within the framework of fracture mechanics.

2.5. Healing mechanics

Porosity and permeability are good indicators of clogging induced by precipitation. Nevertheless, crack filling does not necessarily imply crack rebonding. Only loading and unloading cycles can provide a measure of potential stiffness recovery by crack rebonding [Fuenkajorn and Phueakphum, 2011]. Numerous studies were conducted at the microscopic scale on halite because this natural geomaterial can heal even in the absence of impurities [Senseny et al., 1992, Spiers et al., 1990]. Phenomenological healing models proposed for salt rock assume that both crack propagation and rebonding are time-dependent, which makes it possible to represent cracking and healing effects by a viscoplastic dilatant deformation [Chan et al., 1998, Hou, 2003]. Microscopic dislocation processes in salt rock are known to originate isochoric viscoplastic deformation during transient creep. At the transition to secondary creep, dislocation multiplication is counteracted and balanced by recovery phenomena such as cross-slip, diffusion-controlled climb and recrystallization processes [Senseny et al., 1992]. At low pressure, dislocation pile-ups and strain incompatibilities between grains induce stress concentrations, which result in micro-cracking, brine inflow, salt rock dissolution, and potential cavern collapse followed by subsidence. Under favorable temperature and pressure conditions, fluid-assisted Diffusive Mass Transfer (DMT) [Smith and Evans, 1984], solution precipitation creep [Carter and Heard, 1970, Raj, 1982] and pressure-solution [De Boer, 1977, Rutter, 1976, Spiers et al., 1990] can accelerate crack healing, increase rock stiffness, decrease rock permeability and therefore improve geostorage safety. Figure 1(D) shows the principle of pressure-solution: the difference of chemical potential between the solid-solid interface and the solid-fluid interface drives dissolution at the grain contacts ($\Delta \mu_s$), transportation via the grain boundary ($\Delta \mu_d$) and precipitation at the free pore walls ($\Delta \mu_p$). Creep models and deformation maps are based on the postulate that a one-to-one relationship exists between microscopic mechanisms and deformation rates (see Fig. 1(A-C) for example). The underlying assumption is that the rate of macroscopic deformation equals the rate of the fastest micro-mechanism, which does not apply to heterogeneous rock microstructures, in which independent mechanisms can occur simultaneously. Similar strategies were used to model carbonates [Baud et al., 2009, Dautrait]
Some authors combined rate-dependent or rate-independent damage and healing variables [Miao et al., 2011], but most damage and healing models are formulated with a “net damage” variable defined as the difference between the CDM damage variable and a healing variable [Arson et al., 2012, Ju and Yuan, 2012]. For instance, the author’s group proposed a phenomenological model of damage and healing for halite [Zhu and Arson, 2015], in which the stiffness tensor depends on the difference between a second-order damage tensor that grows with positive differential stress, and a volumetric healing tensor that evolves over time according to a diffusion equation. The model accounts for unilateral effects and thus distinguishes crack closure and crack rebonding. The concept of healing by diffusive mass transfer (DMT) was initially introduced to model atomic interactions in cracked glass [Wiederhorn and Townsend, 1970], and refers to crack rebonding by the migration of ions in the lattice forming the solid mass. In the case of halite, solid DMT occurs from the center of salt crystals (grains) to crystal edges, which results in inter-granular crack rebonding. The diffusion coefficient depends on temperature, which allows capturing the acceleration of healing at higher temperature. Model predictions depend on temperature, which allows capturing the acceleration of healing by diffusive mass transfer (DMT) was initially introduced to model atomic interactions in cracked glass [Wiederhorn and Townsend, 1970], and refers to crack re-bonding by the migration of ions in the lattice forming the solid mass. In the case of halite, solid DMT occurs from the center of salt crystals (grains) to crystal edges, which results in inter-granular crack rebonding. The diffusion coefficient depends on temperature, which allows capturing the acceleration of healing at higher temperature. Model predictions highlight the increased efficiency of healing with time and temperature. For instance, a similar level of healing was achieved after a period of hold of 8.5 hours at 593 K and after a period of hold of 3 hours at 693 K. All macroscopic internal variables are expressed as functions of moments of probability of microstructure descriptors; it is thus possible to calibrate the model from microscopic and macroscopic experimental data. More details on CDM fabric enrichment will be provided in Section 3.

A thermodynamically consistent theory of damage and healing mechanics was first proposed by Barbero and collaborators [2005]. The mechanical dissipation is expressed as:

\[ \phi = \sigma : \dot{e}^{D} - \rho \frac{\partial \psi}{\partial D} : \dot{D} - \rho \frac{\partial \psi}{\partial H} : \dot{H} \]  

(44)

where \( \dot{e}^{D} \), \( D \) and \( H \) are the plastic deformation, damage and healing tensors, respectively. Here, the healing tensor is a diagonal second-order tensor, defined in the same way as the damage tensor in Equation 2. Damage and healing components are projected onto three principal directions. Decoupled plasticity, damage and healing potentials are introduced, respectively: \( g_{p} = 0 \), \( g_{d} = 0 \) and \( g_{h} = 0 \), such that it is possible to minimize the dissipation by minimizing the following Lagrangian functional:

\[ \mathcal{L} = -\phi + \dot{\lambda}_{p} g_{p} + \dot{\lambda}_{d} g_{d} + \dot{\lambda}_{h} g_{h} \]  

(45)

A sufficient condition to minimize \( \mathcal{L} \) is to ensure that the partial derivatives of \( \mathcal{L} \) relative to the components of the generalized force vector are zero, which, for decoupled potentials, provides the following evolution laws of the internal variables:

\[ \dot{e}^{D} = \dot{\lambda}_{p} \frac{\partial g_{p}}{\partial \sigma} \quad \dot{D} = -\dot{\lambda}_{d} \frac{\partial g_{d}}{\partial \rho} \quad \dot{H} = -\dot{\lambda}_{h} \frac{\partial g_{h}}{\partial \rho} \]  

(46)

Flow rules are non-associate, i.e. the authors postulate the existence and uniqueness of a plastic yield function \( f_{p} \), a damage function \( f_{d} \) and a healing function \( f_{h} \) that all satisfy the consistency conditions:

\[ f_{i} = 0 \Rightarrow \dot{f}_{i} = 0, \quad i = p, d, h \]  

(47)

Equation 47 allows calculating the dissipation multipliers \( \dot{\lambda}_{p}, \dot{\lambda}_{d}, \dot{\lambda}_{h} \). The free energy is decomposed into a damaged-healed elastic deformation energy and damage, healing and plasticity internal energy functionals, as follows:

\[ \psi = \frac{1}{2} (\epsilon - \epsilon^{D}) \cdot \mathcal{C}(D, H) : (\epsilon - \epsilon^{D}) + \psi_{d} + \psi_{h} + \psi_{p} \]  

(48)

in which \( \psi_{d} \), \( \psi_{h} \) and \( \psi_{p} \) depend on damage, healing and plasticity hardening variables, respectively. A new mapping is introduced to extend the concept of damage effective stress \( \bar{\sigma} \) to damage and healing configurations:

\[ \bar{\sigma} = (detX_{DH})^{-1} : (X_{DH})^{1/2} : \sigma : (X_{DH})^{1/2} := M^{-1} : \sigma \]  

(49)

in which \( X_{DH} \) is an operator that translates the evolution of material surface orthogonal to each direction of space with damage and healing. The damaged-healed stiffness tensor \( \mathcal{C}(D, H) \) is then obtained by invoking the PEEG.

Voyadjis’s group (Voyadjis et al., 2011a,b) accounted for couplings between rate-independent damage, healing and plasticity. Helmholtz free energy is assumed to be a function of plastic deformation, damage and healing, as well as plasticity, damage and healing kinetic hardening variables (respectively: \( \alpha, D^{K}, H^{K} \)) and plasticity, damage and healing isotropic hardening variables (respectively: \( p, D^{I}, H^{I} \)). Additionally, the free energy is decomposed as:

\[ \psi = \frac{1}{2} (\epsilon - \epsilon^{D}) \cdot \mathcal{C}(D, H) : (\epsilon - \epsilon^{D}) + \psi_{*}(\alpha, D^{K}, H^{K}, p, D^{I}, H^{I}) \]  

(50)

Note that Equation 48 is a particular form of Equation 50. The dissipation is written:

\[ \phi = \alpha : \dot{e}^{D} - \rho \frac{\partial \psi_{*}(\alpha, D^{K}, H^{K}, p, D^{I}, H^{I})}{\partial \alpha} \]  

\[ -\rho \frac{\partial \psi_{*}(\alpha, D^{K}, H^{K}, p, D^{I}, H^{I})}{\partial \rho} \frac{\partial \psi_{*}(\alpha, D^{K}, H^{K}, p, D^{I}, H^{I})}{\partial \rho} : \dot{D}^{I} \]  

\[ -\rho \frac{\partial \psi_{*}(\alpha, D^{K}, H^{K}, p, D^{I}, H^{I})}{\partial \rho} \frac{\partial \psi_{*}(\alpha, D^{K}, H^{K}, p, D^{I}, H^{I})}{\partial \rho} : \dot{H}^{I} \]  

(51)

Authors introduce a plasticity yield function \( f_{p} \), a damage “yield” function \( f_{d} \) and a healing “yield” function \( f_{h} \) and assume that flow rules are associate, so that evolution laws can be obtained by minimizing the following Lagrangian functional:

\[ \mathcal{L} = -\phi + \dot{\lambda}_{p} f_{p} + \dot{\lambda}_{d} f_{d} + \dot{\lambda}_{h} f_{h} \]  

(52)

which, if the yield functions are coupled, provides the following evolution laws for the internal variables:

\[ \dot{e}^{D} = \dot{\lambda}_{p} \frac{\partial f_{p}}{\partial \alpha} + \dot{\lambda}_{d} \frac{\partial f_{d}}{\partial \rho} + \dot{\lambda}_{h} \frac{\partial f_{h}}{\partial \rho} \]  

\[ \dot{D} = -\dot{\lambda}_{p} \frac{\partial f_{p}}{\partial \rho} - \dot{\lambda}_{d} \frac{\partial f_{d}}{\partial \rho} - \dot{\lambda}_{h} \frac{\partial f_{h}}{\partial \rho} \]  

\[ \dot{H} = -\dot{\lambda}_{p} \frac{\partial f_{p}}{\partial \rho} - \dot{\lambda}_{d} \frac{\partial f_{d}}{\partial \rho} - \dot{\lambda}_{h} \frac{\partial f_{h}}{\partial \rho} \]  

(53)
in which the Lagrangian multipliers are obtained by using the consistency conditions:  \( \dot{f}_p = \dot{f}_d = \dot{f}_h = 0 \). The evolution laws of the hardening variables are obtained in a similar way, with some simplifications if plastic (respectively, damage, healing) functions only depend on plastic (respectively, damage, healing) hardening variables. The damaged-healed stiffness tensor is derived from the PEEE, in which the effective stress is related to the actual stress by an operator that depends on both damage and healing:

\[
M^{-1} = \left[ M_D^{-1} + (I - M_D^{-1}) : M_H^{-1} \right]^{-1}
\]

\[
M_D = [(I - D) \otimes (I - D)]^{-1/2}
\]

\[
M_H = [H \otimes H]^{-1/2}
\]

in which \( I \) is the fourth-order identity tensor.

The theory of damage and healing mechanics was then adapted for decoupled rate-dependent damage and healing processes [Darabi et al., 2012]. The dissipation is written:

\[
\phi = -\rho \frac{\partial \psi}{\partial \dot{D}} : \dot{D} - \rho \frac{\partial \psi}{\partial \dot{H}} : \dot{H}
\]

and it is assumed that the damage and healing driving forces can be expressed as:

\[
-\rho \frac{\partial \psi}{\partial \dot{D}} = \dot{\lambda}_d \frac{\partial \phi}{\partial \dot{D}}, \quad -\rho \frac{\partial \psi}{\partial \dot{H}} = \dot{\lambda}_h \frac{\partial \phi}{\partial \dot{H}}
\]

The expressions of the free energy \( \psi \) and of the dissipation \( \phi \) are postulated, which yields the damage and healing evolution laws. Since both are viscous phenomena, no damage or healing threshold function is needed.

2.6. Irreversible deformation

Crack opening is sometimes assumed to generate irreversible deformation due to geometric incompatibilities at the faces of the cracks that close. Figure 2 shows a typical stress/strain curve for a uniaxial loading with such residual deformation. The total deformation \( \epsilon \) is the sum of the irreversible damage deformation \( \epsilon^{id} \), which remains after unloading, the reversible damage deformation \( \epsilon^{rd} \), which occurs due to the degradation of stiffness after damage growth, and the purely elastic strain \( \epsilon^e \), which would exist in the absence of damage. The total elastic (reversible) deformation is noted \( \epsilon^E \). The residual stress \( \sigma_R \) represents the amount of
stress that needs to be applied beyond unloading to return to a state of zero deformation. Classical CDM allows modeling the degradation of stiffness, i.e., \( \dot{e}^\text{id} \). The theory of plasticity allows capturing irreversible deformation such as \( \dot{e}^\text{id} \). CDM and plasticity are thus different, but complementary. It is natural to couple CDM with the theory of plasticity to capture both the degradation of stiffness and the occurrence of irreversible deformation due to crack propagation. Damage mechanics and plasticity are both theories of continuum mechanics that allow predicting the evolution of state and internal variables from energy postulates. Typically, the expressions of the REV free energy and of the dissipation potential are assumed, based on phenomenological observations and thermodynamic consistency requirements.

Several strategies have been proposed in CDM to calculate the increment of \( \dot{e}^\text{id} \) and close the formulation without introducing a plastic potential. Implications on the expression of the free energy depend on whether the non-damaged model is classical elastic, hypo-elastic or hyper-elastic. Halm and Dragon [1998] proposed the first CDM formulations with irreversible damage deformation, within a classical elasticity framework. Swoboda and Yang [1999a, 1999b] later proposed a similar macroscopic CDM model, in which the macroscopic damage evolution is related to micro-crack propagation laws governed by fracture mechanics. Stress is thermodynamically conjugated to the total deformation. Using Legendre transform [Collins and Houlsby, 1997], the REV Gibbs free energy \( \mathcal{G} \) is related to the Helmholtz free energy as follows:

\[
\psi (\dot{\epsilon}, D) + \mathcal{G} (\sigma, D) = \sigma : \dot{\epsilon} \tag{57}
\]

In classical elasticity:

\[
\dot{\epsilon} = \dot{\epsilon}^E + \dot{\epsilon}^\text{id}, \quad \dot{\epsilon}^E = \tilde{C}^{-1} : \sigma \tag{58}
\]

in which \( \tilde{C} \) is the damaged stiffness tensor of the REV. Assuming that the non-damaged elastic behavior is linear (a common assumption in CDM), and assuming that \( \dot{e}^\text{id} \) does not depend on the current state of stress (but only on the loading history of the material), Equations 57 and 58 yield:

\[
\dot{\epsilon} = \dot{\epsilon}^E + \dot{\epsilon}^\text{id} \tag{59}
\]

\[
\dot{\epsilon}^E = \tilde{C}^{-1} : \sigma + \left( \frac{\partial \tilde{C}^{-1}}{\partial \sigma} : \sigma \right) : D \tag{60}
\]

\[
\dot{\epsilon}^\text{id} = \alpha D \tag{61}
\]

Note that Equation 61 relies on the assumption that variables \( \dot{\epsilon} \) and \( D \) are independent. However, while the total elastic deformation depends only on the current state of stress (and accumulated damage), the irreversible damage deformation does depend on the variations of damage. Hence, it is not possible to consider that \( \dot{\epsilon} \) and \( D \) are thermodynamically independent. This contraction was previously highlighted by the author [Arson, 2014].

Now, if the non-damaged model is hypo-elastic:

\[
\dot{\epsilon} = \dot{\epsilon}^E + \dot{\epsilon}^\text{id} \tag{62}
\]

\[
\dot{\epsilon}^E = \tilde{C}^{-1} : \sigma \tag{63}
\]

\[
\dot{\epsilon}^\text{id} = \alpha D \tag{64}
\]

The inequality of dissipation is the same as Equation 61 and the constraint on the expression of Gibbs free energy is the following:

\[
\frac{\partial \mathcal{G}}{\partial \sigma^2} = \tilde{C}^{-1}, \quad \frac{\partial^2 \mathcal{G}}{\partial D \partial \sigma} = \alpha \tag{65}
\]

It can be seen that the requirements expressed in 63 cannot be met, because \( \tilde{C} \) is a function of damage.

If the non-damaged model is hyper-elastic, stress is thermodynamically conjugated to the total elastic deformation,
To summarize:

- In classical elasticity, the introduction of $\varepsilon_{id}$ in the formulation poses a problem to express the inequality of dissipation. The dependence of $\varepsilon_{id}$ to damage is usually ignored, and that is a contradiction.
- In hypo-elasticity, it is impossible to account for irreversible damage deformation without introducing a plastic potential that translates the dependence of irreversible deformation to the stress path and not only to the evolution of damage.
- In hyper-elasticity, it is possible to close the formulation of a CDM model with irreversible damage deformation without resorting to any extra plasticity potential, but the expression of the inequality of dissipation is not classic. While for certain classes of CDM models, thermodynamic consistency is ensured as long as the damage rate is positive [Desmorat, 2006], it is not the case in hyper-elastic CDM models with irreversible damage deformation.

Accounting for irreversible damage deformation with only one internal variable is likely to lead to inconsistencies. Additionally, this kind of formulation often fails short when the material response is very different in tension and compression, like in rocks [Xu and Arson, 2014].

A sound alternative is to calculate irreversible deformation by means of a plastic potential. Note that the damage and plastic potentials can be coupled, but the formulation resorts to at least two internal variables, which avoids the inconsistencies mentioned above. A general framework for coupling damage and plasticity models was given in [Al-Rub and Voyiadjis, 2003, Hansen and Schreyer, 1994, Luccioni et al., 1996]. The basic principles are the same as in models that couple damage and healing. The dissipation is written in the same way as in Equation 51, without the healing terms:

$$\psi(\varepsilon^E, D) + \mathcal{G}(\sigma, D) = \sigma : \varepsilon^E$$

The reduced inequality of dissipation now writes:

$$\sigma : \varepsilon^d - \psi \geq 0, \quad \sigma : \varepsilon_{id} - \frac{\partial \psi}{\partial D} : D \geq 0, \quad \left(\alpha \sigma - \frac{\partial \psi}{\partial D}\right) : D \geq 0$$

(66)

To summarize:

- $\psi(\varepsilon^E, D) + \mathcal{G}(\sigma, D) = \sigma : \varepsilon^E$ as proposed in [Al-Rub and Voyiadjis, 2003, Jin and Arson, 2017, 2018b]:

$$\phi = \sigma : \varepsilon^p - \rho \frac{\partial \phi}{\partial \varepsilon^p} : \varepsilon^p + \alpha - \rho \frac{\partial \psi}{\partial D} : D$$

(67)

Assuming the existence of a plastic potential $g_p = 0$ and of a damage potential $g_d = 0$, the maximization of dissipation is ensured if the associated Lagrangian functional $\mathcal{L}$ reaches a minimum. In this case, we have:

$$\mathcal{L} = -\phi + \lambda_p g_p + \lambda_d g_d$$

(68)

A sufficient condition is to impose that the partial derivatives of $\mathcal{L}$ in reference to the components of the generalized force vector are zero, which yields the following evolution equations:

$$\varepsilon^d = \lambda_p \frac{\partial g_p}{\partial \sigma} + \lambda_d \frac{\partial g_d}{\partial \sigma}, \quad \frac{D}{\partial \varepsilon^p} : D^\varepsilon = -\lambda_p \frac{\partial g_p}{\partial \varepsilon^p} - \lambda_d \frac{\partial g_d}{\partial \varepsilon^p}$$

(69)

Note that in [Al-Rub and Voyiadjis, 2003], it is assumed that:

$$\dot{\varepsilon}^d = \lambda_p \frac{\partial g_p}{\partial \sigma}, \quad \dot{\varepsilon}_{id} = \lambda_d \frac{\partial g_d}{\partial \sigma}$$

(70)

which is equivalent to decoupling the irreversible deformation due to plasticity from that due to damage. As explained above, caution is needed in defining the free energy and the dissipation potentials in order to ensure the non-conditional positiveness of the dissipation functional.

Numerous damage-plasticity models were proposed for rocks starting in the early 2000’s [Bennett and Borja, 2018, Chen et al., 2012, Chiarelli et al., 2003, Conil et al., 2004, Dufour et al., 2012, Le Pense et al., 2016, Parisio and Laloui, 2017, Parisio et al., 2015, Salari et al., 2004, Shen et al., 2001, Zhang et al., 2016, Zhou et al., 2008]. Note that the list is far from exhaustive. Many other types of phenomenological damage-plasticity models exist for concrete, but reviewing them here would be beyond the scope of this paper.

### 2.7. Non-local CDM

Finite Element models of damage or strain softening exhibit mesh dependency. The finer the mesh, the narrower the damage zone. Since the energy dissipated per unit volume is finite, the vanishing damage zone volume causes the structure to fail with zero energy dissipation, which is physically unrealistic [Bazant and Pijaudier-Cabot, 1988, Pijaudier-Cabot and Bažant, 1987]. From a mathematical standpoint, partial differential equations lose their hyperbolic (respectively elliptic) character in dynamic problems (respectively in quasi-statistic problems), thus making the initial-boundary problem ill-posed [De Vree et al., 1995, Peerlings et al., 2012, Le Pense et al., 2016, Parisio and Laloui, 2017]. To avoid numerical inconsistencies, damage softening models are formulated with non-local variables, which allows simulating finite localization zones, the size of which is independent of the mesh size. Internal lengths are used to scale the influence of a variable defined at $x$ on a point located at $x+dx$. The main strategies are [Lasry and Belytschko, 1988]: (1) spatial averaging of state variables (integral nonlocal formulation); (2) introduction of spatial gradients of state variables (differential nonlocal formulation); (3) microstructure enrichment; and (4) introduction of rate dependency. Rate dependency occurs naturally if diffusion equations are used for deformation and dissipation evolution laws: the diffusion coefficient is inversely proportional to the square of an internal length parameter (e.g., percolation distance for DMT). Although
convenient, the use of a diffusive creep law for brittle-elastic solids can lead to thermodynamic inconsistencies [Turcotte and Shcherbakov, 2006]. Hence we restrict our review to integral and differential non-local formulations and to microstructure enrichment.

### 2.7.1. Integral non-local formulation

In integral formulations [Bažant and Jirásek, 2002, Bazant and Ožbolt, 1990], space averages are weighted by attenuation functions. If \( \eta(x) \) is a local field variable in a solid body occupying a domain \( V \), the corresponding non-local field variable \( \bar{\eta}(x) \) is defined as:

\[
\bar{\eta}(x) = \int_V \alpha'(x, \xi) \eta(\xi) d\xi
\]

(71)

where \( \alpha'(x, \xi) \) is a weight function. The local and integral values of a uniform field at a point are equal, hence the weight function satisfies a normalization condition that can be expressed as:

\[
\int_V \alpha'(x, \xi) d\xi = 1
\]

(72)

The normalized weight function is defined as:

\[
\bar{\alpha}'(x, \xi) = \frac{\alpha(\|x - \xi\|)}{\int_V \alpha(\|x - \xi\|) d\xi}
\]

(73)

where the weight function \( \alpha(x, \xi) \) is typically a Gaussian (normal) distribution function or a Bell-shaped function and always depends on an internal characteristic length \( l_{c} \), which is of the same order of magnitude as the maximum size of the material’s inhomogeneities. The length \( l_{c} \) can be determined experimentally by comparing the responses of specimens in which the damage remains distributed with the response of fractured specimens, in which damage localizes [Bažant and Pijaudier-Cabot, 1989]. It can also be determined by comparing the simulation results for various values of \( l_{c} \) with the experimental response [Geers et al., 1999].

Integral non-local regularization was performed on the averaged energy release rate \( \omega_{Y_{\text{max}}(\varepsilon)} \) [Pijaudier-Cabot and Bažant, 1987], the damage variable \( \omega_{Y_{\text{max}}(\varepsilon)} \) [Bazant and Pijaudier-Cabot, 1988], the equivalent strain \( \omega_{Y_{\text{max}}(\varepsilon)} \) [Bažant and Lin, 1988], the specific fracture strain \( \gamma = \Omega/(1 - \Omega) \) [Pijaudier-Cabot and Bažant, 1987], the inelastic stress \( \omega(\varepsilon) \) [Jirásek, 1998], the inelastic stress rate \( \dot{\omega}(\varepsilon) \) [Jirásek, 1998] and the inelastic stress calculated from the non-local strain \( \omega(\varepsilon) \) [Bažant et al., 1996]. Jirásek [1998] demonstrated that only averaging the equivalent strain, the energy release rate or the specific fracturing strain can correctly reproduce large post-peak deformation or complete fracture. Other integral non-local models lead to spurious residual stresses and to a dilatation of the softening zone.

Jin and Arson used an integral non-local formulation to model geomaterials that exhibit tensile softening [Jin and Arson, 2018c]. The free energy is expressed in a similar way as in Equation 1, with different coefficients for open and closed cracks. Damage evolution laws in tension and compression are postulated so as to obtain damage patterns that conform to experimental observations:

\[
\dot{D} = \lambda_1 \dot{D}_t = \lambda_1 \begin{bmatrix} \langle \varepsilon_t \rangle^2 / \ell_t^2 & 0 & 0 \\ 0 & \langle \varepsilon_c \rangle^2 / \ell_c^2 & 0 \\ 0 & 0 & \langle \varepsilon_c \rangle^2 / \ell_c^2 \end{bmatrix},
\]

(74)

\[
\dot{D} = \lambda_c \dot{D}_c = \lambda_c \begin{bmatrix} \langle \varepsilon_t \rangle^2 / \ell_t^2 & 0 & 0 \\ 0 & \langle \varepsilon_c \rangle^2 / \ell_c^2 & 0 \\ 0 & 0 & \langle \varepsilon_c \rangle^2 / \ell_c^2 \end{bmatrix}.
\]

(75)

in which the Lagrange multipliers \( \lambda_1 \) and \( \lambda_c \) are determined from consistency conditions applied to the damage criteria:

\[
f_t = \dot{\varepsilon}_t - (\kappa_t + \alpha_t \text{Tr}D)
\]

\[
f_c = \dot{\varepsilon}_c + \eta \text{Tr} - (\kappa_c + \alpha_c \text{Tr}D)
\]

(76)

Equivalent strains for tension \( (\dot{\varepsilon}_t) \) and compression \( (\dot{\varepsilon}_c) \) are defined as:

\[
\dot{\varepsilon}_t = \sum_{i=1}^{3} \langle \varepsilon_i \rangle^2, \quad \text{if Tr} > 0
\]

\[
\dot{\varepsilon}_c = \sum_{i=1}^{3} \langle \varepsilon_i \rangle^2, \quad \text{if Tr} \leq 0
\]

(77)

in which \( \varepsilon_i \) are the principal strain components and \( \varepsilon_t \) are the principal deviatoric strain components calculated as \( \varepsilon_i = \varepsilon_t - \text{Tr}\varepsilon/3 \).

The model was calibrated against triaxial compression tests performed on shale [Amendt et al., 2013]. As an illustration, we present the predictions made with the calibrated model at the material point for two stress paths commonly studied in experimental rock mechanics (note that these simulations are not replicates of specific physical experiments). Figure 3 shows the stress-strain curve and the evolution of damage during a uniaxial tension/compression test, indicating elastic tensile loading (A-B), post-damage softening (B-C), elastic unloading (C-D), elastic compressive loading with unilateral effects (D-E), elastic tensile reloading (E-F-G), and again, post-damage softening (G-H). Figure 4 shows the stress-strain curve and the evolution of damage components obtained for triaxial compression test under various confining pressures. For all confining pressures considered, triaxial compression results in lateral damage. Results show that the calibrated model captures general trends observed experimentally for quasi-brittle materials [Germanovich et al., 1994, Sahouryeh et al., 2002, Savalli and Engelnder, 2005, Yuan and Harrison, 2006], such as the dependence of damage development on the confining pressure, compressive hardening and the difference of up to one order of magnitude between the tensile and compressive yield stresses.

In order to account for the non-local nature of damage, the equivalent strains that control damage evolution (Equation 77) are replaced by their weighted average defined on an influence domain \( V \), as follows:

\[
\dot{\varepsilon}_i^{\text{nl}}(x) = \int_V \alpha'(x, \xi) \varepsilon_i(\xi) dV(\xi), \quad (i = t, c)
\]

(78)

where \( x \) is the position vector of the material point considered, and \( \xi \) is the position vector of points in the influence domain of \( x \), \( \alpha'(x, \xi) \) is the normalized Bell-shaped function.
Figure 3. Simulation of a uniaxial tension-unloading-compression-tension loading sequence for a single element. Modified and reprinted from [Jin and Arson, 2018c] with permission from Elsevier.

Figure 4. Simulation of triaxial compression tests under various confining pressures for a single element. Modified and reprinted from [Jin and Arson, 2018c] with permission from Elsevier.

Figure 5 shows the results of FEM simulations of three-point bending tests, with the local and the non-local formulations of the damage model and for two mesh refinements. The internal length parameter in these simulations was $l_c = 0.01$ m. The element size was 0.065 m for the coarse mesh and about 0.002 m for the fine mesh, so that the ratio characteristic length to element size was 1.5 for the fine mesh and 5 for the fine mesh. Results show that non-local enhancement avoids mesh dependency during crack development: the width of the process zone is the same for both mesh refinements (marked with a rectangle in Figure 5). However, the shape of the damage process zone is still mesh dependent because the absence of micro-crack interaction in the proposed model makes it impossible to capture the total tensile stress relaxation that happens after the peak tensile stress has been reached. It is expected that this issue can be solved by coupling continuum damage mechanics and discrete fracture mechanics.

### 2.7.2. Differential non-local formulation

In differential formulations [Bažant and Jirásek, 2002, de Borst et al., 1999], local field variables are developed in Taylor series [Askes et al., 2000, Askes and Sluys, 2002, De Vree et al., 1995, Peerlings et al., 1996b]:

$$\eta(x) = \eta(x) + c_1 \Delta^2 \eta(x) + c_2 \Delta^4 \eta(x) + c_3 \Delta^6 \eta(x) + \ldots$$  \hspace{1cm} (79)

$\Delta^2$ denotes the Laplacian operator and the coefficients $c_i$ depend on the weight function $\alpha$ and on the averaging volume $V$. Note that the odd-order derivatives of $\eta$ cancel out because the non-local weight function is symmetric. Since $\eta(x)$ explicitly depends upon higher order derivatives, $C^1$ shape functions are required in Finite Element Analysis (FEA). To circumvent this limitation, Peerlings et al. [1998, 1996a] calculated the second derivative of Equation 79, multiplied it by $c_4$ and injected the result back into Equation 79, as follows:

$$\eta(x) - c_1 \Delta^2 \eta(x) = \eta(x) + (c_2 - c_1^2) \Delta^4 \eta(x) + (c_3 - c_1 c_2) \Delta^6 \eta(x) + \ldots$$ \hspace{1cm} (80)

Ignoring the terms or order higher than 4, it is sufficient to employ $C^0$-continuous interpolation functions for displacements in FEA.

The coefficients $c_i$ can be expressed explicitly in terms of $l_c$ if the weight function is a Gaussian distribution. Because the non-local variable $\eta(x)$ is explicitly related to the local variable $\eta(x)$, Equation 79 is referred to as the explicit differential-type non-local model, and can rewritten as:

$$\eta(x) = \eta(x) + \frac{1}{2} l_c^2 \Delta^2 \eta(x) + \frac{1}{8} l_c^4 \Delta^4 \eta(x) + \frac{1}{48} l_c^6 \Delta^6 \eta(x) + \ldots$$ \hspace{1cm} (81)

Similarly, Equation 80 is referred to as the implicit differential-type non-local model because $\eta(x)$ is implicitly related to the local variable $\eta(x)$. The complete development in series with the Gaussian weight function is expressed as:

$$\eta(x) - \frac{1}{2} l_c^2 \Delta^2 \eta(x) + \frac{1}{8} l_c^4 \Delta^4 \eta(x) - \frac{1}{48} l_c^6 \Delta^6 \eta(x) + \ldots = \eta(x)$$ \hspace{1cm} (82)

Note that both series are equivalent when no truncation of higher order is made. However, at same truncation order, continuity requirements on shape functions are different for explicit and implicit series. A detailed account on this topic is provided in [Jin, 2018]. The implicit second-order scheme, expressed as:

$$\eta^2(x) - \zeta \Delta^2 \eta(x) = \eta(x),$$ \hspace{1cm} (83)

is wildly invoked to simulate softening in problems of dynamics [Askes et al., 2000, De Borst et al., 1995], brittle damage [Askes and Sluys, 2002, De Borst et al., 1995, De Vree et al., 1995] and plasticity [de Borst et al., 1999, De Borst et al., 1995]. The parameter $\zeta$ with the dimension of a length squared is related to the internal length. As indicated by Geers et al. [1998], applying Equation 83 with a constant $\zeta$ provides inconsistent predictions in mode I crack propagation, because the damage zone becomes wider and wider in a direction perpendicular to the crack, where the material should unload. It is possible to overcome this limitation by accounting for the transient behaviour of the gradient parameter during damage evolution.
2.7.3. Microstructure enrichment

Enriched continuum models have additional degrees of freedom for microscopic translations and rotations. Toupin [1964], Mindlin [1964, 1968] and Germain [1973] developed the theory of microstructure enriched elasticity. Vernerey and collaborators [2007] extended the formulation to microstructure enriched elastoplasticity for hierarchical materials. The microscopic displacement, noted \( u'_i(x'_i) \), is defined in reference to the macroscopic displacement, noted \( u_i(x_i) \). Expanding \( u'_i(x'_i) \) in Taylor's series and truncating the higher order terms, one gets the microscopic displacement field of a so-called micromorphic continuum of degree one:

\[
u'_i = u_i + \chi_{ij} x'_j\tag{84}
\]

The kinematic description of a micromorphic continuum of degree one only depends on the macroscopic displacement field \( u_i \) and on the gradient of micro-displacement \( \partial u'_i = y_{ij} \). By definition, \( y_{ij} \) is a second order tensor, the symmetric part of which, \( \frac{1}{2}(y_{ij} + y_{ji}) \), is called the micro-strain rate tensor, and the anti-symmetric part of which, \( \frac{1}{2}(y_{ij} - y_{ji}) \), is named the micro-rotation rate tensor. Noting the macroscopic strain \( \epsilon_{ij} = \frac{1}{2}(\partial_i u_j + \partial_j u_i) \), the relative deformation \( \eta_{ij} = \partial_i u_j - y_{ij} \) and the micro-deformation gradient \( \kappa_{ijk} = \partial_i y_{jk} = y_{ij,k} \), the volumetric Helmholtz free energy \( \psi \) can be expressed in terms of the deformation tensors \( \epsilon_{ij}, \eta_{ij}, \kappa_{ijk} \). The stress conjugates are:

\[
s_{ij} = \frac{\partial \psi}{\partial \eta_{ij}}, \quad \nu_{ijk} = \frac{\partial \psi}{\partial \kappa_{ijk}} \tag{85}
\]

The state of stress in a micromorphic material of degree one is thoroughly defined by the Cauchy symmetric stress \( \sigma_{ij} \), the micro-structure relative stress tensor \( s_{ij} \), and the third order double stress tensor \( \nu_{ijk} \). From a physical point of view, the Cauchy stress represents the macroscopic average of forces per unit area, the micro-stress can be interpreted as a spatial average of forces arising from the nonlocal behavior of the microstructure, and the double stress (or coupled stress) represents the spatial average of the microscopic moments per unit area. The micro-relative-stress is the stress necessary to balance the couple stress at micro-scale and can be thought of as an additional stress that represents the nonlocal interactions within the microstructure. Note that microstructure-enriched models, like the gradient-enriched formulations, involve non-physical variables (e.g., third-order tensors), which raises a number of issues for numerical implementation [Zhao et al., 2005].

Micro-polar media, governed by Cosserat models [Cosserat and Cosserat, 1909], are a particular case of

---

**Figure 5.** Horizontal damage component (i.e. vertical crack density) obtained by simulating a three-point bending test without and with non-local enhancement, for various mesh densities. Only one half of the beam is shown. Modified and reprinted from [Jin and Arson, 2018c] with permission from Elsevier.
micromorphic material of degree one, in which only microrotations are accounted for (i.e. microscopic translations are ignored) [Chambon et al., 2001, Matsushima et al., 2000, Vernerey et al., 2007]. Second gradient models are also models of micromorphic media of degree one, in which it is assumed that the gradient of micro deformation is equal to the macro deformation. The micropolar and second gradient theories were successfully applied to rocks [Sulem and Vardoulakis, 2014, Tamagnini et al., 2001] and granular materials [Chambon et al., 2001]. Chambon and collaborators [Chambon et al., 2001, 2004, Tamagnini et al., 2001] formulated the strain gradient theory relative to a reference configuration. They assumed that the deformation gradient followed a multiplicative decomposition and that the second gradient of the motion could be decomposed into an elastic and a plastic part. In their model, the free energy expression is used so as to formulate the evolution laws (yield functions, hardening laws) for both the plastic strain and the hyper strain tensor. As an alternative to the plasticity theory for modeling the nonlinear irreversible behavior, CDM models coupled with the strain gradient theory provide another way to circumvent scale effects encountered in conventional damage models and to simulate strain softening. For example, isotropic and anisotropic damage models were proposed with second strain gradient enhancement [Zhao et al., 2005, Zhou et al., 2002], in which the conventional strain and the strain gradient are split into an undamaged elastic part and a damaged irreversible part, and the increment of work-conjugated Cauchy stress and higher order stress are updated thanks to the damage yield function and the damage potential. Both the isotropic and anisotropic models were used to predict shear bands and to investigate size effects.

3. Fabric enrichment

Fabric-enriched models are essentially phenomenological models in which the damage tensor is defined as a convolution of moments of probability of microstructure descriptors. The free energy and the dissipation potential are postulated at REV scale and constitutive equations are derived following the method explained in Section 2. In this section, we review the fabric tensors most used in geomechanics, and we apply fabric enrichment to the modeling of damage in salt rock, within Zysset’s theoretical framework (Equation 1).

3.1. Fabric tensors in geomechanics

In structural geology, fabric is typically described with the distributions of crack length, aperture and orientation [Chester et al., 2004, Long et al., 1982, Wilson et al., 2003]. In material sciences, poly-dispersed media are often characterized by the pore size distribution, the lineal-path function, the chord length probability density function (pdf), the void nearest-surface distribution function and the void exclusion probability [Blum and Eisenlohr, 2009, Lu and Torquato, 1992a,b, Philleo, 1983, Reid, 1955, Torquato and Lu, 1993, Torquato et al., 1990], see Table 1 for the definitions (at the end of the paper). In particulate mechanics, the fabric tensor was used as an internal variable in the Anisotropic Critical State Theory (ACST) initially proposed by Dafalias's group [Dafalias and Manzari, 2004, Dafalias et al., 2004, Fu and Dafalias, 2011, Gao et al., 2014, Li and Dafalias, 2011]. In the ACST, any deviatoric second-order tensor can be used as a measure of fabric. The main inconvenience is that the governing equations of the model depend on the fabric tensor in the critical state, which is not known a priori.

Some examples of microstructure descriptors commonly used to define fabric tensors are given in Table 1 (at the end of the paper), in which rock is viewed as an assembly of different types of minerals and of cracks of different geometries. All the different minerals and cracks are defined as “phases”. Figure 6 shows two fabric tensors proposed by Li and Li for granular media [Li and Li, 2009]. Both tensors require a Voronoi-Delaunay tessellation of the granular medium. The contact-based fabric tensor depends on the contact vector from void centroid to grain contacts, and characterizes the arrangement of so-called “solid cells”, which each contains a grain and its assigned void space. The assigned void space is calculated by dividing each void element into void tetrahedra whose vertices are the three vertices of a Delaunay boundary surface and the center of the void element, and by joining these tetrahedra to the solid elements that share the same Delaunay boundary surfaces. The void-based fabric tensor depends on the contact vector from grain centroid to grain contacts, and characterizes the arrangement of so-called “void cells”, which each contains a pore (or void) and its assigned solid space. The latter is calculated by dividing each grain into solid tetrahedra whose vertices are the three vertices of a Delaunay boundary surface and the center of the grain, and by joining these tetrahedra to the void elements that share the same Delaunay boundary surfaces. The contact-based fabric tensor and the void-based fabric tensor can be used to characterize damage in rocks made of polycrystals of cemented aggregates. Cracks affect the geometry of the grains (crystals or aggregates) and of the pores, which makes the contact-based and void-based fabric tensors particularly efficient to track the evolution of intra- and inter-granular crack sizes, shapes, orientations and connectivity.

State-of-the-art CDM models do not capture the full range of possible rock fabrics and textures encountered in nature [Vernon, 2018]. For instance, connections between cracks imply enhanced hydraulic crack interaction, but not necessarily mechanical crack interaction, especially if cracks are randomly oriented and if the distribution of cracks is dense [Schubnel et al., 2006]. Two different damage variables are necessary to model permeability enhancement and stiffness degradation [Maleki and Pouya, 2010] and the REV should be adapted to the properties homogenized [Lacy et al., 1999]. In what follows, we provide an example of fabric-enriched CDM model, in which the damage variable is defined as a convolution of moments of pdfs of microstructure descriptors.
Salt is a polycrystalline material made of bonded crystals (called grains in the following). Salt rocks' stiffness depends both on the mechanical properties of individual grains and on rock microstructure, which changes with damage. Here, we summarize the micro-macro mechanical model proposed by the author's group in [Shen et al., 2017] to predict the evolution of stiffness, deformation and microstructure in halite. A series of oedometer tests and triaxial tests on reagent-grade granular salt in dry conditions and under a temperature of 150°C were performed in Chester’s group at Texas A&M University. The diameter of salt particles ranged between 0.300mm and 0.355mm. Samples were 6.426cm high and 1.905cm in diameter. The tests were conducted at a constant rate of 0.034mm/s, which corresponds to slightly more than 3% per minute. The initial porosity was 40% in oedometer tests. Triaxial tests were conducted on consolidated samples with a 6% initial porosity.

During the oedometer tests, salt exhibited higher stiffness and lower porosity for higher axial stress. In the cyclic triaxial tests, salt samples yielded when differential stress reached 40MPa. 2D-microscopic images, parallel to the axial direction, were obtained at several stages of the loading path, both in the oedometer and in the cyclic triaxial tests. Image analyses were performed to capture the evolution of microstructure. For microstructure analyses, we represented grains by equivalent ellipses of same area as the actual grain shapes observed in the micrographs. After examining the evolution of a number of descriptors, we decided to interpret the evolution of damage in terms of the actual grain shapes observed in the micrographs. After

The pdfs of the four microstructure descriptors determined during the oedometer tests are presented in Figure 6. Fabric descriptors defined by Li and Li [2009]. $f(n)$: norm of contact vector from void centroid to grain contacts; $v(n)$ norm of contact vector from grain centroid to grain contacts; $H_0$, $G_0$: normalization coefficients. Figure adapted and reprinted from [Li and Li, 2009] with permission from ASCE.

**3.2. Example: Fabric enriched damage model for salt rock**

The distribution of branch orientations is more uniform than that of grain orientations. Branches have a slight preferential orientation parallel to the direction of the minimum principal stress. Grain orientations are initially relatively uniform, but salt grains take a preferential orientation under compression: the major axis of grains rotates to align with the direction of minimal compression principal stress. With the compaction of the sample, the local solid volume fraction increases, and the distribution of void sizes becomes more uniform. Due to the dislocations and indentations that occur under high contact forces, grain solidity decreases during consolidation.

A fabric tensor $\bar{F}$ can be defined for each descriptor, by integrating the corresponding pdf. Since only 2D images are available, fabric tensors are $2 \times 2$ and take the following expression:

$$\bar{F}_{ij} = \int_{\Omega} n_i n_j E(\Gamma) d\Gamma, \quad (i, j = 1, 2) \quad (86)$$

where $\bar{F}$ is a symmetric second-rank tensor; $\Gamma$ is the whole solid angle, and equals to $2\pi$ for 2D images; $E(\Gamma)$ is a pdf. The components of the fabric tensors are calculated as follows:

$$\bar{F}_{11} = \frac{1}{N} \sum_{k=1}^{N} (p_k)^2 (\sin \theta_k)^2 \quad (87)$$

$$\bar{F}_{12} = \frac{1}{N} \sum_{k=1}^{N} (p_k)^2 \sin \theta_k \cos \theta_k \quad (88)$$

$$\bar{F}_{22} = \frac{1}{N} \sum_{k=1}^{N} (p_k)^2 (\cos \theta_k)^2 \quad (89)$$

$N$ is the total number of measures (or angles) considered in the image. For the local solid volume fraction fabric tensor, $p_k$ is the local solid volume fraction of each polygon and $\theta_k$ is the solid angle, and equals to $2\pi$ for 2D images; $E(\Gamma)$ is a pdf.
is the angle between the horizontal and the line connecting a polygon's center with the image's center. For the grain solidity fabric tensor, \( p_k \) is the grain solidity of each grain and \( \theta_k \) is the angle between the horizontal and the line connecting a grain's center with the image's center. For the tensors of grain orientation and branch orientation, \( \theta_k \) is the orientation angle of the vector \( n \) in reference to a set axis in the image. We note \( G_i \) and \( B^i \) the fabric tensors of grain orientation and branch orientation for image \( i \), respectively. We note \( L_i \) and \( S^i \) the traceless components of the fabric tensors of local solid volume fraction and grain solidity for image \( i \), respectively. After calculating the fabric components according to Equations 90, the fabric tensors of local solid volume fraction and grain solidity are normalized and their traceless components are calculated. We note that in each fabric tensor, the \( F_{ij} \) component is negligible in front of the diagonal components. We thus consider that all fabric tensors are diagonal and orthogonal, and we define a normalized fabric tensor \( H^i \) to characterize the total anisotropy induced by grain orientation, branch orientation, distribution of local solid volume fraction, and distribution of grain solidity, as follows:

\[
H^i = \gamma^i G_i B^i L_i S^i
\]

(91)

where \( \gamma \) is a normalizing coefficient used to make \( Tr(H) \) equal to 1. The second-rank fabric tensor \( H \) can be written as \( kH + K \). \( k \) is a scalar, and \( K \) is a traceless second-rank tensor. Calculating \( H^i \) for each image \( i \) allows tracking the evolution of fabric with deformation.

Based on the expression of the free energy of an elastic medium enriched with microstructure given by Zysset [Zysset and Curnier, 1995], the stiffness tensor of salt rock can be expressed as a function of \( k \), \( K \) and material properties, as expressed in Equation 1. The coefficients \( a_i \) are functions of \( k \) and Lamé like constants \( \mu_c \) and \( \lambda_c \). Considering that the salt rock Young's modulus is an exponential function of porosity [Turner and Cowin, 1987], we propose to relate the salt rock Young's modulus is an exponential function of porosity \( \phi \) to the average local solid fraction \( \alpha_l \) and to average grain solidity \( \alpha_s \) as follows:

\[
\lambda_c = \lambda_o (ma_l + na_s)
\]

(92)

\[
\mu_c = \mu_o (ma_l + na_s)
\]

(93)

We calibrated the parameters \( m \) and \( n \) against the experimental oedometer modulus (i.e., the ratio of axial stress by axial strain). After calibration, the model was used to calculate the fabric tensor \( H^i \) for each image and to deduce the damaged stiffness tensor, according to Equation 1. Figure 8 shows how the oedometer modulus calculated numerically (red dots) fits the damaged oedometer modulus measured experimentally (solid line).

For cyclic triaxial tests, image analysis was performed on samples with 2.4% axial strain and 4.3% axial strain. The pdfs are presented in Figure 9. Branches orient in the direction of maximum principal compression stress, forming columns of supporting grains. Grains exhibit a slight preferential orientation parallel to the direction of the minimum principal stress. During the triaxial loading path, cracks open, the size of the voids increases, the coordination number of grains decreases, and a lower local solid volume fraction is observed. During triaxial loading, grain solidity decreases and plastic deformation accumulates in the grains.

The global fabric tensor constructed based on the moments of pdfs in Equation 86 can only capture the total loss of contact between grains, which appears as intergranular cracks in the microstructure images. In some cases, the connection is in fact weakened rather than broken. In other words, there is a need to represent joint damage between grains (where damage could be a scalar with a value between 0 and 1). Joint damage cannot be observed by classical scattered electron microscopy, but the effect of joint damage can be measured via mechanical tests. To capture both the cracks that are observable in microscopy images and those that are not, a micro-mechanical approach is preferred.

4. Micro-mechanics enrichment

By contrast with phenomenological models, micro-mechanics-enriched models are formulated with micro-scale...
damage variables (typically, crack displacement jumps and crack densities). The REV free energy is obtained by applying the principle of superposition: it is the sum of the elastic strain energy stored in the non-damaged matrix and of the strain energy stored in the crack displacement jumps. Damage/crack threshold and evolution functions are expressed at the microscopic scale, and the REV dissipation is obtained by summation (or integration).

4.1. Micro-plane theory

In the micro-plane theory [Bažant and Oh, 1985, Bažant and Prat, 1988], constitutive laws are formulated on a plane and then integrated over all the possible plane orientations of space to calculate the REV constitutive law. It is assumed that either the stresses on the weak planes within the material are the components of the macroscopic stress tensor (static constraint), or that the strain on the weak planes are the components of the macroscopic strain tensor (kinematic constraint). The constitutive relations for each plane are formulated independently and the equilibrium between the REV stress \( \sigma \) and the micro-plane stresses is satisfied in a weak sense, by applying the principle of virtual work. Over a REV of volume \( \Omega \), the principle of virtual work is expressed as [Carol and Bazant, 1997, Carol et al., 2001]:

\[
\frac{4\pi}{3} \epsilon : \delta \epsilon = 2 \int_{\Omega} [\sigma_{N}\delta \epsilon_{N} + \sigma_{T} : \delta \epsilon_{T}] d\Omega \quad (94)
\]

in which \( \epsilon \) is the macroscopic strain, \( \sigma_{N} \) and \( \sigma_{T} \) are the normal and shear stresses that apply on the micro-plane, \( \epsilon_{N} \) and \( \epsilon_{T} \) are the micro-plane normal and shear strains. For a kinematic constraint:

\[
\delta \epsilon_{N} = N : \delta \epsilon, \quad \delta \epsilon_{T} = T : \delta \epsilon \quad (95)
\]

\[
N_{ij} = n_i n_j, \quad T_{kij} = \frac{1}{2} [n_i \delta_{jk} + n_j \delta_{ik} - 2n_i n_j n_k]
\]

where \( n \) is the normal to the micro-plane. Since the individual components of the virtual strain tensor are independent, the combination of Equations 94 and 95 provides:

\[
\sigma = \frac{3}{2\pi} \int_{\Omega} \sigma_{N} N d\Omega + \frac{3}{2\pi} \int_{\Omega} \sigma_{T} : T d\Omega \quad (96)
\]

The static constraint was extensively used in metal plasticity and applied with great success in the so-called “slip theory”. The kinematic approach was applied by Bazant’s research group to describe geomaterials with strain softening, especially concrete. In some models, an additional assumption is made to state that the volumetric, deviatoric and shear responses on each microplane are mutually independent [Bazant and Ožbolt, 1990, Bažant and Prat, 1988]. In other micro-plane models, the microplane inelastic behavior is characterized by stress-strain boundaries [Bažant et al., 2000, 1996, Caner and Bažant, 2012], which can be treated as yield limits. Within the boundaries, the response is incrementally elastic, i.e. the elastic moduli are incrementally degraded due to damage. When the accumulated stress and strain get outside of the boundaries at some incremental step, they immediately drop back to the boundaries and then follow the elastic boundary. It is possible to catch hysteresis effects during cyclic loading. In another category of micro-plane models, a kinematically constrained microplane system is coupled to a statically constrained microplane system to simulate cohesive tensile cracking [Bažant and Caner, 2005]. Microplane models have been extensively widely in engineering applications, however, one should bear in mind that the weak planes do not correspond to any geometric feature of the microstructure: there is no physical principle that dictates the evolution laws of the weak planes.

4.2. Micro-mechanical damage models

In micro-mechanical models, the expression of the REV free energy stems from micro-mechanical equations (often, fracture mechanics equations) used to represent micro-crack initiation, propagation, opening, closure and frictional sliding. The displacement jump across crack faces [Budiansky and O’Connell, 1976] is used as a basis to upscale the effective properties of the damaged REV [Kachanov, 1992, 1993] and to express the corresponding energy potentials [Kachanov, 1982a,b, Pensee and Kondo, 2003, Pensée et al., 2002]. The evolution law is based on fracture mechanics and can represent Mode I splitting [Gambartotta and Lagomarsino, 1993, Krajcínovic et al., 1991], Mode II friction sliding [Gambartotta and Lagomarsino, 1993] or mixed mode wing crack development [Kachanov, 1982b, Nemat-Nasser and Obata, 1988]. In order to account for crack interactions, one can explicitly express the stress field as the sum of the stress that results from external loading and of the stress that results from crack interaction [Paliwal and Ramesh, 2008].

The expression of the free energy that was established on the basis of damaged stiffness symmetry and positiveness requirements can be obtained from micro-mechanics. Here, we show that the free energy of the REV takes the form of Equation 1 for a dilute distribution of penny-shaped cracks, embedded in an isotropic linear elastic matrix of compliance tensor \( S^{-m} \). Each microscopic crack is characterized by its normal direction \( \vec{n} \) and its radius \( a \). For a dilute distribution of non-interacting cracks, the stress field that applies to micro-crack faces, \( \sigma^{m}(x) \), is the direct projection of the macro stress \( \sigma(x) \) on crack faces. Consequently, for each crack, the local stress that applies at the crack faces is self-equilibrating and the matrix stress is equal to the macro stress. Considering a penny shaped crack of radius \( a \) subjected to a uniformly distributed normal stress \( p \) (respectively shear stress \( \tau \)) at its faces and embedded in an infinite elastic medium with Young’s modulus \( E_0 \) and Poisson’s ratio \( \nu_0 \), the average normal (respectively shear) displacement jump, also known as Crack Opening Displacement (COD), is expressed as:

\[
\langle u_{n} \rangle = \frac{16}{3} \left( 1 - \frac{\nu_0}{2} \right) \frac{1}{\pi E_0} \frac{pa}{a} \quad (97)
\]

\[
\langle u_{t} \rangle = \frac{32}{3} \left( 2 - \frac{\nu_0}{2} \right) \frac{1}{\pi E_0} \frac{\tau a}{a}
\]

If the family contains \( N \) cracks of same orientation \( \vec{n}_i \), the volume fraction of the normal and shear displacement
jumps can be calculated as follows:

\[ \beta_i = \frac{N}{\Omega_i} \left( |u_{i0}| \right) \pi \alpha_i^2 = \rho_i c_i \sigma \cdot (\n_i \otimes \n_i) = \rho_i c_i \sigma \cdot (\n_i \otimes \n_i) \]

\[ \gamma_i = \frac{N}{\Omega_i} \left( |u_{i0}| \right) \pi \alpha_i^2 = \rho_i c_i (\sigma^d \cdot \n_i - (\n_i \cdot \sigma^d \cdot \n_i)) \]

\[ = \rho_i c_i (\sigma \cdot \n_i - (\n_i \cdot \sigma \cdot \n_i)) \n_i \]

\[ c_0 = \frac{16}{3} - v_0^2 \]

\[ c_1 = \frac{32}{3} \frac{1 - v_0^2}{E_0} \]

(98)

where \( \overline{T} = \sigma \cdot \n_i - (\n_i \cdot \sigma \cdot \n_i) \n_i \). \( \rho_i = \frac{N \alpha_i^2}{\Omega_i} \) is the crack density parameter along the direction \( \n_i \) with \( \Omega_i \) the volume of the REV. Details are available in [Budiansky and O’Connell, 1976] and in [Kachanov, 1992]. Note that the value of \( \rho_i \) can exceed one.

With the soil mechanics convention, the unilateral contact condition at crack faces can be expressed as

\[ [u_n] \geq 0, \quad \sigma_{nn} = -\n \cdot \sigma \cdot \n \leq 0, \quad [u_n] \sigma_{nn} = 0 \]

(99)

The average strain due to the displacement jumps of the micro-cracks of the family with normal \( \n_i \) is calculated as:

\[ e^d = \frac{N}{\Omega_i} \int_{\partial \omega_i} [u_n]|(\n_i \otimes \n_i)|dS 

+ \frac{N}{2\pi} \int_{\partial \omega_i} ([\n_i] \otimes \n_i + [\n_i] \otimes [\n_i])|dS 

\]

\[ = \beta_i \n_i \otimes \n_i + \frac{1}{2} (\gamma_i \n_i + \n_i \n_i) \]

(100)

According to the principle of superposition, the Helmholtz free energy \( W^* \) of the REV containing \( N \) cracks of orientation \( \n_i \) is the sum of the elastic deformation energy of the matrix and the energy stored in the micro cracks displacement jumps. The Gibbs energy (free enthalpy) is obtained by Legendre transformation, as follows:

\[ G^* = \sigma : e^E - W^* \]

(101)

in which \( e^E = e^E + e^d \) is the REV elastic strain. As a result, \( G^* \) is expressed as:

\[ G^* = \frac{1}{2} \sigma : S_{nn} \cdot \sigma \]

\[ + \frac{1}{2} c_0 \rho_i H(\n_i \cdot \sigma \cdot \n_i) \sigma \cdot N_i \cdot \sigma \cdot T_i : T_i \]

(102)

where \( H(\cdot) \) is the Heaviside jump function and \( \overline{n_i} \cdot \sigma \cdot \overline{n_i} = \sigma_{nn}^i \) is the normal stress at the crack face. The fourth order normal (respectively, tangent) operator \( N_i \) (respectively, \( T_i \)) is defined by:

\[ N_i = N_{ijkl}^i = n_i^a j^a n_j^a n_k^a \]

\[ T_i = T_{ijkl}^i = \frac{1}{4} n_i^a n_j^a \delta_{jk} n_k^a + n_j^a n_k^a \delta_{jk} n_i^a + \delta_{ik} n_j^a n_k^a - n_i^a n_j^a n_k^a \]

(103)

in which \( n^a \) is the unit normal vector of each direction \( a \).

For more than one crack orientation, we can calculate the total Gibbs energy of the REV by integrating \( G^* \) for a distribution of crack orientations \( \rho(n) \), over the unit sphere \( S^2 = \left\{ \overline{n} : |\overline{n}| = 1 \right\} \), as follows:

\[ G = \frac{1}{2} \sigma : S_{nn} \cdot \sigma \]

\[ + \frac{1}{8\pi} \int_{S^2} \rho(\overline{n}) (c_i H(\sigma_{nn}^i) \sigma : N_i + c_i \sigma : T_i : T_i) \sigma dS \]

(104)

At the scale of the REV, the second order crack density tensor \( \rho \) is defined in such a way that: \( \rho(\overline{n}) = \overline{n} \cdot \rho \cdot \overline{n} \). The second order damage tensor is defined as follows:

\[ D = \frac{1}{4\pi} \int_{S^2} \rho(\overline{n}) (\overline{n} \otimes \overline{n}) dS \]

\[ = \frac{1}{4\pi} \int_{0}^{2\pi} \int_{0}^{\pi} \rho(\overline{n}) (\overline{n} \otimes \overline{n}) \sin \theta d\phi d\theta \]

(105)

It can be shown mathematically (see [Lubarda et al., 1994, Yang et al., 1999] for details) that the crack density function \( \rho(\overline{n}) \) is related to the damage tensor as follows:

\[ \rho(\overline{n}) = \frac{3}{2} (\overline{n} \otimes D \cdot \overline{n} - Tr D) \]

(106)

After introducing the relation 106 in the expression of Gibbs energy (Equation 104), we obtain the macroscopic free enthalpy as a function of the second order damage tensor \( D \), as follows:

\[ G(\sigma, D) = \frac{1}{2} \sigma : S_{nn} \cdot \sigma + a_1 Tr D (\sigma : D)^2 + a_2 Tr (\sigma \cdot \sigma \cdot D) \]

\[ + a_3 Tr \sigma \cdot D (\sigma \cdot D + \sigma \cdot D) + a_4 Tr D Tr (\sigma \cdot \sigma) \]

(107)

The expression of the free enthalpy obtained from micromechanical principles in Equation 107 is similar to that assumed in a number of purely phenomenological models, e.g. [Hayakawa and Murakami, 1997, Kachanov, 1980]. We also note that Equation 107 is the Legendre transform of the Helmholtz free energy expressed in Equation 1, in which the higher order terms in damage have been truncated.

### 4.3. Example: The Discrete Equivalent Wing Crack Damage (DEWCD) model

Micro-mechanical approaches are particularly suitable to model wing cracks that propagate at the tips of sliding cracks under compression. The sliding wing crack model was initially presented in the pioneering work of Brace and Bombolakis [1963]. Since then, numerous studies were devoted to the mechanisms of crack propagation in brittle solids under compression, for instance: [Ashby and Hallam, 1986, Dyskin and Salganik, 1987, Horii and Nemat-Nasser, 1986, Lehner and Kachanov, 1986, Nemat-Nasser and Obata, 1988, Scholtes and Donzé, 2012]. As an example, we present a summary of the Discrete Equivalent Wing Crack Damage (DEWCD) model proposed by the author’s group to predict non-frictional crack propagation under compression [Jin and Arson, 2017]. If the unilateral contact condition is satisfied, cracks propagate due to normal tensile stresses, according to the following mode I propagation criterion:

\[ f_{\rho}(\sigma, a_i) = a_i^2 - \frac{K_f}{\sqrt{a_i}} \]

(108)
In order to account for hardening, the material toughness \( \theta \) is determined by projecting the damage tensor defined in Chloé Arson, Micro-macro mechanics of damage and healing in rocks, Open Geomechanics, 2020, article no. 1.

Shear stress on the main crack exerts tension on the wing cracks. 

Wing cracks propagate normal to the minimum principal stress (following experimental evidence).

**Figure 10.** Wing crack propagation model under compression. Modified and reprinted from [Jin and Arson, 2017] with permission from Elsevier.

Assuming that the tensile driving force \( F \) is uniformly distributed along the faces of the fictitious planar crack (Figure 10), we define the wing micro crack propagation criterion as follows:

\[
\frac{F}{\sigma^2} = \frac{\cos(\theta)\max_i a_i}{\min_i a_i} - \frac{K_c}{\sqrt{\sigma_{min}}} 
\]

(109)

where \( a_i \) is the radius of the fictitious wing crack, which can be determined by projecting the damage tensor defined in Equation 2 as follows:

\[
a_{i} = \left( \frac{\Omega_i}{N_i} \right)^{1/3} \left[ n_i \cdot D \cdot \bar{n}_i \right]^{1/3}
\]

(110)

in which \( N_i \) is the number of cracks in the \( i \)th crack family, \( \Omega_i \) is the volume of the REV and \( \bar{n}_i \) is the unit vector normal to the family of wing cracks of orientation \( \theta \), and is expressed as:

\[
\bar{n}_i = \bar{n}_i \cos(\theta_i) + \bar{n}_i \sin(\theta_i)
\]

(111)

In order to account for hardening, the material toughness is expressed as a hyperbolic function of the crack radius, as follows:

\[
K_c = \frac{a^{3/2}}{K_0 + a} \quad (112)
\]

where \( a = a_{\theta i} \) for a crack in compression (leading to the propagation of wing cracks), and \( a = a_i \) for a crack in tension. We introduce discrete damage potentials (expressed in terms of \( Y_i \)) in a homogeneous function of degree one, as follows:

\[
g_d(n_i) = Y_i - C_0 \quad (113)
\]

in which \( Y_i \) is the variable that is work-conjugate to the density of cracks oriented with a normal \( \bar{n}_i \). \( Y_i \) is a driving force. \( C_0 \) is the damage threshold: if \( Y_i \) exceeds \( C_0 \), then cracks normal to \( \bar{n}_i \) propagate. Following a non-associate flow rule, the inelastic strain and damage increments can be computed from the damage potential as

\[
\epsilon^{in} = \sum_{i=1}^{M} \lambda_i \frac{\partial g_d(n_i)}{\partial \sigma}, \quad \rho = \lambda_i \frac{\partial g_d}{\partial Y_i} = \dot{\lambda}_i \quad (114)
\]

where \( \dot{\lambda}_i \) is Lagrange multiplier for each family of crack with normal \( \bar{n}_i \). The incremental damage density is calculated from the increments of crack radius and from the consistency rule:

\[
\Delta \rho_i = \frac{N_i}{|\Omega_i|} \Delta(a_i^3) \quad (115)
\]

Once the increment of crack radius is obtained for each crack orientation (\( \Delta(a_i), \Delta(a_{\theta i}) \)), the increment of damage tensor \( \Delta D \) can be updated as

\[
\Delta D = \sum_{i=1}^{M} \frac{N_i}{|\Omega_i|} \Delta(a_i^3) \bar{n}_i \otimes \bar{n}_i + \sum_{i=1}^{M} \frac{N_i}{|\Omega_i|} \Delta(a_{\theta i}^3) \bar{n}_{\theta i} \otimes \bar{n}_{\theta i} \quad (116)
\]

in which \( M \) is the total number of crack families (or orientations) considered. Ultimately, the increment of crack density is obtained by projecting the increment of damage tensor in each of the \( 2 \times 21 \) directions considered in the quadrature:

\[
\Delta \rho_i = \bar{n}_i \cdot \Delta D \cdot \bar{n}_i \quad (117)
\]

The DEWCD model was calibrated against triaxial compression tests performed on shale [Jin and Arson, 2017]. Figure 11 presents the stress-strain paths and damage evolution predicted with the calibrated DEWCD model for a uniaxial tension test followed by unloading and uniaxial compression. The model predicts that the specimen yields at \( \sigma_y = 12 \) MPa in uniaxial tension. The introduction of a damage potential together with non-associate flow rules in the DEWCD model allows capturing the occurrence of residual inelastic strains after unloading. Simulation results also highlight unilateral effects induced by crack closure in compression.

Figure 12 shows the results of the simulation of a uniaxial compression cycles of increasing amplitude. According to Kachanov’s calculations [1992], a damage of 0.3 corresponds to the initiation of crack interactions. Considering that above this damage threshold, the REV has reached failure, the present simulations indicate that the uniaxial tensile strength is 30 MPa (Figure 11), and the uniaxial compression strength is 180 MPa (Figure 12). In addition, the yield
stress predicted by the DEWCD model in uniaxial compression is 50 MPa. We conclude that the DEWCD model predicts values of yield stress and strength that are in the range of values expected for a rock material like shale. It is also worth noting that the DEWCD model predicts realistic damage evolution for uniaxial compression, as the axial damage (crack planes perpendicular to the loading axis) does not propagate, and the lateral damage components grow exponentially after the yield stress is reached. In the simulations done with the DEWCD model, damage during the second loading cycle initiates at a lower stress value than the maximum stress value reached during the first cycle, which indicates that the DEWCD model can capture hysteric effects.

4.4. Example: Combining micro-mechanics with phenomenological damage functions

The DEWCD model proposed in Subsection 4.2 provides excellent predictions at the Gauss Point, with 5 material parameters that have a physical meaning. However, the projection of the crack density due to wing crack growth to the damage tensor, and the subsequent projection of the damage tensor onto the 2 × 21 discrete orientations, necessary to calculate the damaged stiffness tensor, make it challenging to implement the DEWCD model in a finite element code. Here, we present a model in which the expression of the REV free energy is obtained from micro-mechanics principles but in which micro-crack propagation is governed by phenomenological CDM for ease of numerical implementation. Here, we present a model in which the expression of the REV free energy is obtained from micro-mechanics principles but in which micro-crack propagation is governed by phenomenological CDM for ease of numerical implementation. Details are published in [Jin and Arson, 2018b].

We define the damage driving force in direction $i$ as the energy release rate necessary to propagate a unit crack density in that direction. The energy release rate must exceed the crack resistance $R(\rho_i)$ to allow the crack boundary to grow. The yield criterion adopted in the proposed model is inspired from the Drucker-Prager model, and is expressed as follows:

$$f_i(\rho_i, Y_i) = Y_i - a_i \text{Tr}\sigma - R(\rho_i)$$  \hspace{1cm} (118)

From a mechanical point of view, the expression of the resistance curve $R(\rho_i)$ controls the hardening or softening behavior after the initial yield surface is reached. We consider that $R(\rho_i)$ is a linear function of the crack density $\rho_i$ [Pensee and Kondo, 2003] and we emphasize that our model is only applicable for dilute distributions of micro-cracks, i.e. before crack coalescence and before the peak of strength. In addition, we distinguish the increase of open crack density in Modes I & II (when the unilateral condition is satisfied for the $i^{th}$ microplane direction) and the increase of closed crack density in Mode II (when the unilateral condition is not satisfied for the $i^{th}$ microplane direction), as follows:

$$f_i(\rho_i, Y_i) = Y_i - a_i \text{Tr}\sigma - k(1 + \eta \rho_i)$$  \hspace{1cm} (119)

where $k = k_c, \eta = \eta_c$ if cracks of the $i^{th}$ family are closed, and $k = k_o, \eta = \eta_o$ if cracks of the $i^{th}$ family are open. Each crack yield criterion $f_i$ is associated with one particular crack family. The macroscopic yield surface is the boundary of the elastic domain intersected by all the activated crack yield surfaces.

For each active microplane direction, the closed crack criterion is activated if the macroscopic stress projected on the crack plane is a compression, and the open crack criterion is activated if the macroscopic stress projected on the crack plane is a tension. Note that in Equation 119, the crack yield criterion $f_i$ can be rewritten in the form of a function of stress and crack density only, because the energy release rate is a function of stress. As a result, the increment of crack density of an activated crack family ($f_i > 0$) can be readily calculated by means of the consistency condition under controlled stress conditions:

$$\dot{f}_i(\rho_i, Y_i) = \frac{\partial f_i}{\partial \sigma} : \dot{\sigma} + \frac{\partial f_i}{\partial \rho_i} \dot{\rho}_i = 0$$  \hspace{1cm} (120)

Figure 13 provides a representation of the yield surfaces of the closed and open cracks. The consistency conditions expressed for all activated crack families form a system of
equations that can be solved simultaneously. By contrast, only one consistency condition is used in purely phenomenological CDM models, which limits the number of crack propagation modes considered.

5. Homogenization schemes

Most homogenization schemes used in CDM are based on Eshelby’s theory [Eshelby, 1957]. The REV is viewed as a homogeneous matrix that contains sets of inclusions. An inclusion set is characterized by micro-structure properties (such as size, shape, orientation) and/or material properties (such as mechanical behavior or chemical potential). Postulates on free energy and dissipation are made at the microscopic scale [Pensee and Kondo, 2003, Pensée et al., 2002], for each phase (i.e. for the matrix and for each set of inclusions). The REV constitutive equations are derived from a matrix-inclusion law [Dormieux et al., 2006b]. Homogenization schemes were compared in several studies [Ju, 1991, Krajcinovic and Sumarac, 1989]. Here, we recall the basic equations of Eshelby’s theory, we review homogenization-based CDM and we present homogenization-based models of time-dependent damage and healing formulated by the author’s group.

5.1. A review of homogenization-based damage models for rocks

Eshelby’s theory consists in finding the stress and strain fields in each phase by solving a matrix-inclusion interaction problem. For a domain Ω subjected to a uniform strain field \( \sigma_\infty \) in the far field, the inclusion-matrix interaction problem is stated as:

\[
\begin{align*}
\forall x \in \Omega, \quad & e(x) = E_\infty + e^d(x) \\
\forall x \in \Omega_0, \quad & e(x) = e^0(x), \quad \sigma^0(x) = C_0 : e^0(x) \\
\forall x \in \Omega_i, \quad & e(x) = e^i(x), \quad \sigma^i(x) = C^i : e^i(x) \\
\forall x \in \Omega_i, \quad & \sigma^i(x) = C_0 \left( e^i(x) - e^*_i(x) \right)
\end{align*}
\]

in which \( \Omega_i \) is the domain of the i\(^{th}\) inclusion, \( \Omega_0 \) is the domain of the matrix. \( e^d \) is the disturbance strain field (for a homogeneous medium, the strain field is uniform and equal to the far field strain, hence the disturbance strain field is zero). \( \sigma^0, \sigma^i \) and \( C_0 \) are respectively the stress field, the strain field and the stiffness tensor of the matrix. \( C^i, \sigma^i, e^i, \) and \( e^*_i \) are respectively the stiffness tensor, the stress field, the strain field and the so-called eigenstrain of the i\(^{th}\) inclusion.

The inclusion eigenstrain is defined as the strain field that is necessary to apply to the inclusion in order to consider that this inclusion has the same stiffness tensor as the matrix.

Eshelby’s theory states that if the matrix is homogeneous, linearly elastic and infinitely extended, and if the inclusions are ellipsoidal, then the eigenstrain and disturbance strain fields are uniform in each inclusion, and:

\[
\forall x \in \Omega_i, \quad e^d(x) = S^i : e^*_i, \quad S^i = P^i : C_0 \tag{122}
\]

in which \( S^i \) is the so-called Eshelby tensor and \( P^i \) is the so-called P-tensor, which depends on the Green’s function. From Equations 121 and 122, one gets the other form of the matrix-inclusion interaction law:

\[
e^i = \left( \delta + S^i : C_0^{-1} : C^i - C_0 \right)^{-1} : E_\infty \tag{123}
\]

in which \( \delta \) is the second-order identity tensor.

The following concentration tensors (\( A \)) are defined:

\[
e^i = A_{\infty}^i : E_\infty, \quad A_0 : E_\infty = E, \quad e^i = A^i : E \tag{124}
\]

in which \( E = \langle \epsilon \rangle \) is the macroscopic strain field, defined as the space average of the local strain field. From Equations 123 and 124, for N inclusions, one gets:

\[
A_{\infty}^i = \left( \delta + P^i : \left( C^i - C_0 \right) \right)^{-1} \tag{125}
\]

\[
A_0 = \sum_{i=0}^{N} \phi_i \left( \delta + P^i : \left( C^i - C_0 \right) \right)^{-1}
\]

in which \( \phi_0 \) is the volume fraction of the matrix and \( \phi_i \) is the volume fraction of the i\(^{th}\) inclusion set. The stiffness tensor

---

**Figure 12.** Stress-strain behavior and damage evolution predicted with the DEWCD model, for a stress path that comprises two cycles of uniaxial compression loading - elastic unloading. Modified and reprinted from [Jin and Arson, 2017] with permission from Elsevier.
The stiffness of each phase is usually given in the problem, and needs to be calculated iteratively. The self-consistent homogenization scheme [Budiansky and O’Connell, 1976, Ju and Lee, 1991, Lee and Ju, 1991], there is no concept of matrix, i.e. all phases have comparable volume fractions. To solve the inclusion-matrix interaction problem, it is assumed that the matrix stiffness tensor is equal to the REV stiffness tensor: \( C_0 = C_{\text{hom}} \). The expression of the REV stiffness tensor in Equation 127 is thus implicit and needs to be calculated iteratively. The self-consistent approach was used by Sumarac and Krajcinovic [1987] for investigating damage by micro-cracking and by many other authors to study the nonlinear behavior and load induced anisotropy due to crack closure and frictional contact [Horii and Nemat-Nasser, 1983, Zhu et al., 2009].

After the pioneering work of Eshelby [Eshelby, 1957], Kröner [1961] extended the self-consistent method to the elastic-plastic behavior of polycrystalline materials, assuming that the presence of inclusions only induced a perturbation of the elastic strain of the matrix (so-called elastic inclusion-matrix interaction). Hill [1965a, 1965b] replaced the model of inclusion subjected to a free strain placed in the matrix by a cavity embedded in an infinite matrix, and subjected at its wall to surface traction and displacements. The local stress in the inclusions (\( \sigma \)) and the far-field stress in the matrix (\( \overline{\sigma} \)) are coupled to the microscopic strain (\( \varepsilon \)) and the macroscopic strain (\( \overline{\varepsilon} \)) by the following relationship:

\[
\sigma - \overline{\sigma} = -L^*: (\varepsilon - \overline{\varepsilon})
\]  

in which \( L^* \) is Hill’s tensor. For spherical inclusions embedded in an isotropic matrix, \( L^*_{ijkl} \) is expressed as:

\[
L^*_{ijkl} = \frac{\mu^*}{4 - 5v} \left[ (3 - 5v)\delta_{ij}\delta_{kl} + \frac{7 - 5v}{2} (\delta_{ik}\delta_{jl} + \delta_{ij}\delta_{jk}) \right]
\]  

in which \( v \) and \( \mu^* \) are the Poisson’s ratio and the shear modulus of the matrix. The link between Hill’s tensor and Eshelby’s tensor was established by Hill himself [1965a, 1965b] and can be stated as follows:

\[
L^* : S = C_0 : (\delta - S)
\]

Furthermore, Hill formulated the interaction model with an incremental method instead of the original secant method, which facilitated further developments to account for the non-linear behavior of the matrix [Cleary et al., 1980, Ju, 1991, Nemat-Nasser, 2004, Nemat-Nasser and Hori, 2013, Willis, 1977, Zeng et al., 2015], plastic accommodation
Chloé Arson, Micro-macro mechanisms of damage and healing in rocks

5.2. Example: Self-consistent micro-macro approach for modeling viscous damage in halite

In quasi-brittle polycrystalline materials like crystalline rocks, damage by cracking or by cleavage on some crystallographic weakness planes dominates plastic and viscous deformation [Feng and Yu, 2010, Ju, 1996, Wang et al., 2015, Zeng et al., 2014, Zhu et al., 2016, Zhu and Shao, 2015]. Here, we summarize the self-consistent homogenization scheme proposed in [Pouya et al., 2016] to predict the viscous damage and fatigue behavior of halite polycrystals from mono-crystal slip mechanisms. We propose a homogenization scheme based on Hill’s incremental interaction model [Hill, 1965a], in which we account for the heterogeneity of the elastic stiffness tensor that results from different damage mechanisms occurring at the grain scale. In order to focus on the effects of grain breakage on macroscopic viscoplastic strains, we disregard the viscous accommodation of the matrix in the inclusion-matrix interaction model (evidenced in [Mercier and Molinari, 2009, Rougier et al., 1994] for instance).

Halite is a Face-Centered Cubic (FCC) crystal. Due to electronic interaction forces between ions, the planes along which sliding requires the minimum energy input are the six {1 01} planes. Based on experimental correlations [Pouya et al., 2016a, Dormieux et al., 2006a, Levasseur et al., 2009, Nguyen et al., 2016] and benchmarked in [Levasseur et al., 2013, Zhu et al., 2008], the major principal local stress of a grain exceeds 2 MPa, the number of unbroken grains is noted \( N_g \), the number of broken grains is noted \( N_b \), and the total number of grains in the REV is noted \( N \). The dilute scheme estimate for the elastic matrix yields the following effective bulk modulus \( (\tilde{\kappa}) \) and shear modulus \( (\tilde{\mu}) \):

\[
\tilde{\kappa} = \frac{N_g}{N} \kappa = \frac{N - N_b}{N} \kappa = (1 - D) \kappa
\]

\[
\tilde{\mu} = \frac{N_g}{N} \mu = \frac{N - N_b}{N} \mu = (1 - D) \mu
\]

in which the damage variable is defined as \( D = N_b/N = 1 - N_g/N \). The micro-macro stress computational method is explained in detail in [Pouya et al., 2016]. For each time increment, the damage criterion is first checked in each grain; grains in which the stress exceeds the tensile strength are deleted. The sliding mechanism is then activated in the non-broken grain to update the stresses with the viscoplastic strain development. The model was calibrated against creep tests.

Here, we show the impact of grain breakage on the creep deformation regime of the salt polycrystal and on the response of the salt polycrystal to cyclic loading. We start by simulating the three following long-term creep tests: (i) “no damage” case: grain damage accounted for under low creep stress (5 MPa); (ii) “non-damage model” case: grain damage not accounted for under high creep stress (7 MPa); (iii) “damage model” case: grain damage accounted for under high creep stress (7 MPa). Results are presented in Figure 14, which shows the projections of the principal micro-stresses on the radial and vertical axes. Compression is plotted in the bottom left quadrant, and tension is plotted in the top right quadrant. In the “no damage” case, the macroscopic strain rate reaches a steady state in the secondary creep phase, which indicates an absence of accelerated creep. Microscopic tensile stresses remain below 2 MPa, which implies that none of the mono crystals breaks, even though grain breakage is taken into account in the model. In the “non-damage model” case, the macroscopic strain rate remains almost constant, and microscopic stresses exceed the tensile strength limit, which indicates that grains would have broken if damage had been accounted for. In the “damage model” case, the microstress goes to zero in some of the grains, and the macroscopic strain rate increases abruptly at the end of the creep test. These results indicate that the model can capture the accelerated creep regime (tertiary creep) caused by grain breakage. The simulations reproduce the creep rates observed in the experiments reported in Fuenkajorn and Phueakphum [2010] as well as the occurrence of tertiary creep in the form of accelerated creep.

We now study the response of the polycrystal to a sinusoidal cyclic loading of frequency 0.03 Hz (= 2592 day⁻¹)
for several maximum loading stresses and loading amplitudes. For small maximum stress and small loading amplitude, a limit cycle is reached. The number of broken grains \( N_b \) remains low, and Young’s modulus decreases to a finite limit value. For larger maximum stress or larger loading amplitude, failure occurs after a limited number of cycles. Figure 15 shows the results obtained for a cyclic loading with a maximum loading stress of 20 MPa and an amplitude of 10 MPa. Some grains experience zero stress, which indicates that they are broken. Grain breakage first occurs at the 6th cycle, and the whole polycrystal fails after the 8th cycle. The increasing number of broken grains is visible in the progressive reduction of the Young’s modulus in the macroscopic stress-strain curve. These results are in agreement with the experimental results reported in [Fuenkajorn and Phueakphum, 2010]. Damage initiates after a low number of cycles, and the behavior becomes extremely brittle thereafter: total failure usually occurs less than five cycles after the initiation of damage.

One of the main limitations of the model presented here is the assumption of elastic matrix-inclusion interaction. In order to account for damage accommodation, the authors later derived the interaction law in Equation 128 as follows [Pouya et al., 2019]:

\[
\delta \sigma - \delta \bar{\sigma} = -L^* (D) : (\delta \varepsilon - \delta \bar{\varepsilon}) - \left( \frac{\partial L^* (D)}{\partial D} : \delta D \right) : (\varepsilon - \bar{\varepsilon}) \tag{134}
\]

The last term translates damage accommodation, i.e. the irreversible alteration of the inclusion-matrix interaction stiffness tensor due to damage \( D \). Another limitation of the previous model is the assumption that the REV matrix is isotropic. One original aspect of the inclusion-matrix model presented in [Pouya et al., 2019] is that both the grains and the matrix exhibit ellipsoidal anisotropy and that an explicit expression of the Hill’s tensor is derived to account for matrix ellipsoidal anisotropy.

5.3. Example: A self-consistent model of healing for salt rock

We summarize a self-consistent approach proposed by the author’s group in [Shen and Arson, 2019a] to predict the mechanical healing of halite subject to pressure solution.
Inclusions are represented by hollow spheres in which a spherical pore lies at the intersection between three orthogonal grain contact planes. The inclusion thus contains a pore and eight grain fractions, see Figure 16. Noting \( r_g \) the radius of the hollow spherical inclusion, \( W \) the thickness of the shell around the pore, and assuming that \( W \) is uniform around the pore, the spherical void’s diameter is \( 2r_g - 2W \).

The plane that contains the axes \( x \) and \( y \) (respectively \( y \) and \( z \), \( x \) and \( z \)) is noted \( XY \) (respectively \( YZ \), \( XZ \)). The REV is the volume that contains a representative distribution of inclusion types, which can be defined in terms of inclusion size, pore size and inclusion orientation (or grain contact plane orientation).

At the inclusion scale, under normal stress, the solid mineral is dissolved at contact planes, the solute diffuses along the contact planes towards the pore inside the inclusion, and precipitates at the pore wall. Let us consider one pair of contact planes. Due to mass conservation principles, the mass of mineral that diffuses radially from the contact plane towards the pore inside the inclusion, through the annulus of area \( \Delta \omega \), is equal to the mass of mineral that is dissolved at the contact plane. Moreover, the radial diffusion flux \( J(w) \) is related to the chemical potential \( \mu \) according to Fick’s first law. The energy dissipation per unit volume, noted \( \Delta \omega_e \), is equal to \(-J(w)\partial \mu / \partial w \) [Lehner, 1995].

Under low stress conditions, we assume that the solute concentration \( C(w) \) along the grain boundary is equal to that in the pore, \( C_o \) [Plumakers and Spiers, 2015, Rutter, 1983, Spiers and Schütjens, 1999]. Integrating the radial energy dissipation increment over the distance that goes from a point at the periphery of the inclusion to the pore wall, we obtain the expression of the total dissipation \( \Delta \omega \omega_e \) on a contact plane. Assuming that the inclusion work input is entirely dissipated by pressure solution, the total dissipation energy is equal to the work done by the compressive stress on the contact plane. The dissolution velocity normal to the contact plane \( XY \) is expressed as:

\[
V_c^{XY} = \frac{D' S c_0 \Omega^2 \sigma_{n}^{XY} (2r_gW - W^2)}{R^* T \int_0^W ( \frac{2(r_gW - w)^2}{r_gW - w} ) dw}
\]  

(135)

where \( D' \) is the grain boundary diffusion coefficient; \( S \) is the fluid film thickness laying between these two planes; \( \sigma_{n}^{XY} \) is the normal effective stress on plane \( XY \). The rate of mineral volume dissolved, \( V_{c}^{XY} \), depends on the dissolution velocity on plane \( XY \) and on the area of the contact surface: \( V_{c}^{XY} = (2r_gW - W^2) \pi V_c^{XY} \). The dissolution velocity and the rate of volume dissolved on planes \( YZ \) and \( XZ \) can be obtained in the same way. The chemical viscous strain rate \( \epsilon_{c} \) is calculated from the dissolution velocity on planes \( YZ \) as follows:

\[
\epsilon_{c} = \frac{V_{c}^{YZ}}{r_g}
\]  

(136)

\( \epsilon_{c} \) and \( \epsilon_{*} \) can also be determined in a similar way. The mineral only dissolves on the contact planes. In addition, we assume that the mineral precipitates on the pore wall uniformly. Thus, the change of thickness \( \delta W_l \) at time step \( t_l \) can be calculated from the total volume of mineral dissolved and from the pore’s surface area \( A_s \) at time step \( t_{l-1} \). Diffusion and precipitation processes are illustrated for plane \( XY \) in Figure 16. We finally obtain:

\[
\delta W_l = \frac{\tilde{V}_l^{XY} + \tilde{V}_l^{YZ} + \tilde{V}_l^{XZ}}{A_s} \delta t
\]  

(137)

where \( \delta t \) is the time step.

Note that the stress and strain fields in each inclusion are assumed to be uniform, i.e. we assume that pressure-solution produces uniform changes of stress and strain in the hollow spherical inclusion. Since we present simulations done with uniform distributions of inclusion orientations, the matrix is isotropic. We choose an inclusion constitutive model that reflects a linear dependence of the shear and bulk moduli of the matrix \( (\mu^* \text{ and } k^*) \) to the porosity of the REV \( (\phi^*) \), and we assume that the Poisson’s ratio of the
matrix $\nu$ is a material constant [Pouya et al., 2016]. The decrease of REV porosity induces an increase of the matrix shear modulus $\mu^*$, which results in a change of Hill's tensor $L^*$. The local stresses and strains in the inclusions can be obtained as follows:

$$\delta\sigma_{\alpha} = \delta\sigma_{\alpha}^i + L^* \delta\epsilon_{\alpha} + C_{\alpha}^{-1} : \delta\sigma_{\alpha} - \delta\epsilon_{\alpha}$$

(138)

where the superscript $i$ indicates the time step and the subscript $a$ refers to a phase (inclusion type). By definition, the REV stress increment $\delta\sigma^i$ (respectively the REV strain increment $\delta\epsilon^i$) is the average of the local stress increments $\delta\sigma_{\alpha}^i$ (respectively the average of the local strain increments $\delta\epsilon_{\alpha}^i$). Using the expression of the dissolution velocity on each contact plane, the local chemical strain increment is obtained by calculating the change of inclusion radius induced by pressure solution under local stress $\sigma_{\alpha}^i$ and local stress change $\delta\sigma_{\alpha}^i$.

The volumetric strain curves obtained by Spiers for wet salt aggregates [Spiers and Brzesowsky, 1993] are used to calculate the evolution of brine-saturated halite porosity during uniaxial creep tests. Following the description of the experiments, the inclusion size was chosen uniform, equal to 0.1375 mm. Based on the value of initial porosity, the initial mean pore radius was found to be 0.103 mm. We assumed a lognormal distribution for the void radius, with a coefficient of variance (COV) of 0.0001 mm$^2$. Experimental results obtained under axial stresses of 1.1MPa and 3.1MPa were used for the calibration of the elastic and chemical properties of the inclusions. Inclusions of different orientations or different initial void radius develop different microscopic principal stresses $\sigma_p$. Figure 17 shows the principal stresses in the inclusions for the test conducted under an axial stress of 1.1 MPa, after 100,000 seconds (27h40) and after 250,000 seconds (69h25). Like Figures 14.c and 14.d, Figure 17 shows the projections of the principal micro-stresses on the radial and vertical axes. Compression is plotted in the bottom left quadrant, and tension is plotted in the top right quadrant. The dots with a lighter color represent inclusions with smaller initial voids. The major (tensile) principal stresses are almost perpendicular to the loading axis, while the minor (compressive) principal stresses tend to align with the loading direction. The magnitudes of the minor and major principal stresses increase with time. The minor principal stress is nearly zero in some inclusions, while it reaches a value close to -1.5 MPa in some other inclusions. Inclusions with larger initial voids tend to have smaller principal stresses.

We simulate uniaxial creep tests on REVs that contain 200 uniformly oriented inclusions, and we study the sensitivity of the micro-macro model of chemo-mechanical healing to the inclusion size ($r_g$, set equal to the grain size), to the initial porosity and to the coefficient of variance (COV) of the void radius distribution. We vary one parameter at a time, keeping the other two constant: $r_g$ is equal to 0.05mm, 0.15mm or 0.25mm; the initial porosity of the REV is set to 10%, 20% or 40%; the void radius COV is equal to 0.01, 0.05 or 0.1. The results of the parametric study are shown in Figure 18. The rate of halite densification is known to be highly dependent on salt grain size. There are less inter-granular contact planes in the REV if salt crystals are larger. When salt polycrystals are saturated with brine, the size of salt grains not only affects the contact area between grains, but also controls the length of the diffusion path. Results show that a smaller grain size results in a faster decrease of porosity. It is interesting to note that porosity stabilizes to a non-zero value. This is because over time, the macroscopic stress is sustained by less and less inclusions: the inclusions under sustained by less and less inclusions: the inclusions under low stress undergo pressure solution at a negligible rate, and therefore, the full healing time cannot be reached during the

**Figure 16.** The inclusion model in the homogenization scheme of healing. (a) Schematic representation of the inclusion. (b) Mass transfer governed by pressure solution, in contact plane XY. Modified and reprinted from [Shen and Arson, 2019a] with permission from Elsevier.
Correspondingly, the distribution of the void sizes departs more and more from the initial uniform size distribution. Under same COV, void size heterogeneity increases with the initial porosity. In heterogeneous samples, smaller voids heal fast, and larger voids do not heal, which leads to a non-uniform healing rate over time. For smaller initial porosity, we observe a uniform healing rate. The rate of porosity change decreases when the void radius COV increases. This can be explained by the fact that the stiffness of inclusions increases when the void size decreases. As a result, inclusions with small voids undergo higher microscopic compressive stress, and heal faster. As healing proceeds in the inclusions that contain smaller pores, the void radius COV increases, which enhances the difference of healing rate between inclusions that contain small pores and inclusions that contain larger pores.

As a result, it was observed that in samples containing inclusions with different initial void sizes, inclusions with larger voids had a negligible healing rate, and were slowing down the overall healing rate of the REV. In samples with uniformly distributed void sizes, the healing rate was faster, because all inclusions contributed to the healing of the REV. In specimens with smaller grain sizes, principal stresses were more widely distributed in magnitude and the healing rate was higher. For uniform void size distributions, the healing rate increased with initial porosity, but the final porosity change did not depend on the initial porosity of the sample.

5.4. Example: A Mori-Tanaka model of chemical weathering for granite

Biotite is an abundant mineral in many granitic rocks. Upon weathering, biotite expands. The dilation of biotite minerals is thought to be the cause of crack propagation in granite. Here, we summarize the Mori-Tanaka homogenization scheme proposed by the author’s group in [Shen et al., 2019] to understand how weathering at the biotite inclusion scale can trigger damage at the REV scale. Following previous studies [Buss et al., 2008, Goodfellow et al., 2016], we focus on the mechanism that induces the largest expansion: the replacement of interlayer of potassium cations by hydrated magnesium cations, which transforms biotite to vermiculite and smectite.

We use the weathered biotite production rate to calculate the rate of change of the volume of weathered biotite \( V_w \) (m\(^3\)) in the bedrock REV. The biotite production rate is assumed to be zero at the onset of biotite weathering. Following empirical observations of biotite weathering [White and Brantley, 2003], the weathering rate \( R \) is a decreasing power-law function of time:

\[
\frac{dV_w}{dt} = v_m \frac{dQ}{dt}, \quad \frac{dQ}{dt} = RS, \quad R = 3.001 \times 10^{-5} t^{-0.603} \tag{139}
\]

Here \( v_m \) is the molar volume of weathered biotite: \( v_m = 2.10 \times 10^{-4} \) m\(^3\) mol\(^{-1}\). \( Q \) (mol) is the number of moles of weathered biotite in the REV, \( R \) (mol m\(^{-2}\) s\(^{-1}\)) is the weathering rate of biotite per unit mineral surface area, and \( S \) (m\(^2\)) is the total biotite surface area within the REV.

Biotite crystals are represented by oblate spheroidal inclusions. We assume that biotites remain oblate spheroids throughout the weathering process (with variable size and aspect ratio, though). According to experimental observations made by Banfield [1988], the conversion of biotite to vermiculite increases the width of biotite layers by 40%. Under these assumptions, the chemical strain rate in the direction of the smallest axis of a biotite oblate spheroid (called thickness in the following), noted \( \dot{\epsilon}_c \), is equal to \( 0.4\dot{V}_w/V_b \), where \( V_b \) is the volume of a biotite crystal. In the absence of data, we consider at first that the elastic behavior of weathered biotite is linear isotropic. Voigt approximation is used to obtain the shear modulus and bulk modulus of the composite made of biotite and vermiculite, at each stage of the weathering process.
A Mori-Tanaka homogenization scheme is established to predict the evolution of the REV stiffness tensor. Each inclusion is assumed to be embedded in an infinitely extended solid space. Using the stiffness of this infinite medium ($C_0$) as the matrix stiffness, the homogenized stiffness $C_{\text{hom}}$ of the REV is obtained. The relation between the macroscopic stress $\bar{\sigma}$ and the macroscopic strain $\bar{\epsilon}$ is as follows:

$$\bar{\sigma} = C_{\text{hom}} : \bar{\epsilon} - \sum_{i=1}^{n} \phi_i C_i : \eta_i : A_i \quad (140)$$

where $\phi_i$ is the phase fraction (in this case, matrix phase vs. biotite phase, with different phases for different types of biotite inclusions, defined by their size, shape and orientation); $C_i$ is the stiffness of phase $i$; $\eta_i$ is the chemical eigenstrain in each phase (i.e. the strain field that would exist in each phase in absence of the other phases around them). In this study, $\eta_i$ is induced by the chemical weathering strain $\epsilon_c$. $A_i$ is the concentration tensor of each phase, which can be calculated based on Eshelby’s theory. The total strain field in each phase $\epsilon_i$ is given by:

$$\epsilon_i = A_i : \epsilon - \sum_{j=1}^{n} B_{ij} \eta_j \quad (141)$$

where $B_{ij}$ is the influence tensor [Pichler and Hellmich, 2010]. At each time step, the expansion of biotite crystals triggers a stress redistribution in each phase of the REV. With the increase of stress and strain in the matrix, damage accumulates. We propose a phenomenological damage model to account for this form of energy dissipation in the REV. The damage variable is noted $D$. Based on the mathematical requirements of the elastic and damaged elastic stiffness tensor, the Helmholtz free energy is expressed as [Halm and Dragon, 1998]:

$$\Psi = \frac{1}{2} \lambda (tr \epsilon)^2 + \mu tr (\epsilon \cdot \epsilon) + a tr (\epsilon) tr (D \epsilon) + \beta tr (D \epsilon^2) \quad (142)$$

where $\alpha$ and $\beta$ are damaged material parameters.

We conduct sensitivity analyses to understand in which conditions biotite weathering can trigger damage in granite. First, we study the effects of boundary conditions. We consider three biotite inclusion families. In all biotite families, the thickness of biotite oblate spheroids is $1 \text{ mm}$. Each biotite family is characterized by a different initial radius: $10 \text{ mm}$, $6 \text{ mm}$ or $3 \text{ mm}$. The volume fraction of each biotite family (or phase) is $5\%$. All biotite inclusions have the same orientation, with their smaller dimension (‘thickness’ $c$) parallel to the loading axis. Figure 19 presents the results obtained when simulating the effect of biotite weathering under uniaxial, isotropic and oedometer stress conditions, with a stress magnitude of 10 MPa. Results confirm that the net deformation of biotite inclusions is less than the chemical eigenstrain, due to the mechanical reaction applied by the matrix. The sharp evolution of biotite thickness and radius translates into an abrupt evolution of damage and volumetric strain at the REV level. The lateral constraint applied in the oedometer conditions delays the accumulation of damage; the expansion of biotite in the thickness direction is slower than in the other tests. In the case of free expansion, isotropic loading and uniaxial loading, the radius of biotite spheroidal inclusions decreases due to the Poisson’s effects in the matrix. In oedometer tests on the contrary, the radius of biotite inclusions increases.

Second, we study the effects of depth and of biotite inclusion characteristics on weathering. Figure 20 shows the predicted time evolution of damage in the matrix in various scenarios. A more extensive set of parametric studies is presented in [Shen et al., 2019].

Simulation results show that the physical characteristics of biotite crystals - most importantly aspect ratio and volumetric fraction - have a profound effect on the evolution of bedrock damage during biotite weathering. These characteristics exert particularly strong influences on the timing of the onset of damage, which occurs earlier under higher biotite abundances and smaller biotite aspect ratios. Biotite orientation, by contrast, exerts a relatively weak influence on damage. Our simulations further show that damage development is strongly influenced by the boundary conditions, with damage initiating earlier under laterally confined boundaries than under unconfined boundaries. Additionally, results show that depth has no influence on damage development. To sum up, simulations suggest that relatively minor differences in biotite populations can drive significant differences in the progression of rock weakening.

Figure 18. Analysis of the sensitivity of the healing rate to (a) grain size (GR), (b) initial porosity (IP), which indicates the influence of the initial mean pore radius, and (c) pore radius coefficient of variance (COV). Modified and reprinted from [Shen and Arson, 2019a] with permission from Elsevier.
Figure 19. Effects of the loading conditions on the weathering deformation of aligned biotite inclusions (with thickness direction parallel to the loading direction). From left to right, top to bottom: inclusion thickness; inclusion radius; matrix damage; REV volumetric strain.

6. Conclusion

Like in the theory of plasticity, closing the formulation of a CDM model requires postulates on the REV free energy, the dissipation potential and on the “yield” functions. If the non-damaged material is linear elastic, then the Helmholtz free energy (respectively Gibbs free energy) is a polynomial function, quadratic in deformation (respectively, stress) and quadratic in damage. The most general expression of the free energy potential depends on nine constitutive coefficients. To reduce the number of parameters to calibrate, it is common to derive the damaged elasticity tensor from a principle of equivalence between the current damaged state and a reference non-damaged state subjected to an effective stress. The effective stress is related to the current stress by a damage operator. The damage potential is most often chosen to be a non-negative convex function, which ensures the thermodynamic consistency of the CDM model. The damage (or healing) “yield” functions are equal to the damage (or healing) potentials if flow rules are associate. In the literature, unilateral effects of crack closure were mainly accounted for by introducing a Heaviside function in the expression of the free energy, to deactivate stiffness damage in compression directions. But damage rotation challenges thermodynamic consistency and numerical implementation. Healing is often modeled as net damage, i.e., as the difference between a damage tensor and a healing variable. The main challenge with that approach is the mapping of stresses and other variables on the principal planes of the damage and healing variables, which are not necessary equal. Otherwise, accounting for healing in CDM is similar to coupling plasticity and damage: a healing potential $g_h$ is introduced in addition to the damage potential $g_d$, and evolution laws of damage and healing are obtained by solving a problem of minimization under constraint ($g_d = g_h = 0$). Representing damage degradation and irreversible deformation with only one damage variable and only one dissipation potential is only possible if the non-damaged model is hyper-elastic – otherwise, the formulation is thermodynamically inconsistent. A recurrent issue in CDM is the simulation of the post-peak softening behavior. Mesh dependency is currently alleviated by non-local regularization.

Fabric-enriched CDM models are essentially phenomenological models in which the damage tensor is defined as a convolution of moments of probability of microstructure descriptors. The free energy and the dissipation potential are postulated at REV scale. Fabric enrichment is
attractive because fabric is measurable by image analysis and provides a way to predict microstructure evolution through a macroscopic damage variable. However, the coupling between micro- and macro-mechanisms is only one way, since damage evolution allows predicting microstructure changes, but not the opposite.

In micro-mechanics enriched CDM, a two-way coupling exists between the evolution of microstructure descriptors, which is dictated by postulated evolution laws, and the evolution of damage, which influences the state of stress (and hence the evolution of microstructure). The polynomial expression of the free energy that is postulated in phenomenological models to satisfy symmetry and positiveness requirements can be derived from micromechanical principles. Micromechanical evolution laws require mechanical testing for validation, and not only SEM images. The microplane theory falls in the category of micro-mechanics enriched CDM. Constitutive laws are formulated on a plane and then integrated over all the possible plane orientations of space to calculate the REV constitutive law. It is assumed that either the weak plane stresses or the weak plane strains are components of the macroscopic stress (or the macroscopic strain). The equilibrium between the REV stress and the micro-plane stresses is satisfied in a weak sense, by applying the principle of virtual work. The projection of the macroscopic damage tensor onto multiple discrete crack orientations (usually 42 or 74 for numerical integration purposes), necessary to calculate the damaged stiffness tensor, makes the implementation of micromechanical damage models in a finite element code challenging. It is common to formulate the REV free energy based on micro-mechanics principles but to assume that micro-crack propagation is governed by phenomenological laws, for ease of numerical implementation.

Figure 20. Sensitivity of weathering to: (a) depth under oedometric conditions (imposed vertical stress, fixed horizontal displacements) and proportional stress conditions (imposed vertical and horizontal stresses); (b) biotite volume fraction (or abundance); (c) biotite aspect ratio; and (d) biotite orientation (only one family of biotite inclusions, orientation angle $\theta$ in reference to the horizontal axis). In the reference case, the biotite volume fraction is 15%, the biotite aspect ratio is 3, the biotite orientation is 30° and the boundary conditions are proportional loading with a vertical stress of 10 MPa and a horizontal stress equal to half of the vertical stress. Modified and reprinted from [Shen et al., 2019] with permission from Wiley.
In homogenization-based models, the REV is viewed as a homogeneous matrix that contains several sets of inclusions. Each set of inclusions is endowed with a specific stiffness and a specific constitutive behavior. For a strain controlled problem (respectively, stress controlled problem), the strain (respectively, stress) in all phases is equal to the boundary strain (respectively, stress) field plus a perturbation strain (respectively, stress) field. In the Eshelby theory, the stress (respectively, strain) field of the inclusions is expressed as a function of the stiffness (respectively, compliance) tensor of the matrix and as a function of a so-called eigenstrain (respectively, eigenstress) field. For ellipsoidal inclusions embedded in an infinite isotropic matrix, the perturbation strain (respectively, stress) field is related to the eigenstrain (respectively eigenstress) field through the Eshelby tensor. Additionally, the eigenstrain (respectively, eigenstress) is uniform in each inclusion. In Hill's theory, the micro-macro relationship is derived by solving a problem in which inclusions are viewed as cavities subjected to displacements or tractions at their boundaries. Hill's tensor relates the difference between the micro and macro stresses to the difference between the micro and macro strains. A simple relationship exists between the Hill's tensor and the Eshelby tensor. Different homogenization schemes are based on different microstructure assumptions. In the dilute scheme, the volume fraction of inclusions is negligible and inclusions are assumed to not interact. In the Mori-Tanaka scheme, the average of the perturbation field is assumed to be zero in the matrix, so that the average matrix field (strain or stress) is equal to the field applied in the far field (strain or stress). In the self-consistent method, there is no concept of matrix, and each inclusion is viewed as a domain embedded in a uniform medium that has the mechanical properties of the REV (yet unknown). The author's group proposed a variety of homogenization-based models of damage or healing, with account for the chemical eigenstrains of the inclusions. However, coupling different damage and healing processes within the same model is still a work in progress.

Some thoughts on how to expand the framework of rock damage and healing mechanics follow.

(i) The framework of hyperplasticity, which does not require any postulate on the existence or uniqueness of yield functions and which automatically ensures thermodynamic consistency, was never applied to CDM. There may be an avenue to improve state-of-the-art damage and healing models in a similar framework of “hyper-damage mechanics”.

(ii) Mesh dependency in softening models is currently alleviated by non-local regularization. However, healing can only happen where damage localizes. A linkage can be established between preferential directions of crack propagation and preferential directions of crack re bonding via a postulated representation of microstructure, for instance via a homogenization scheme. Perhaps the next step is to couple micro-macro damage and healing models at the REV scale to discrete fracture mechanics at a larger scale to understand how damage and healing localization occurs.

(iii) It is not common to encounter a CDM model in which two damage variables are defined to describe the evolution of two different microstructure traits. There may be an opportunity to define damage variables by their effects (e.g. by microstructure descriptors) rather than their cause (e.g., compression, tension, shear), especially for fabric-enriched models that aim to capture the effect of microstructure organization on both mechanical properties and permeability.

(iv) To date, no theoretical framework has been proposed to couple chemo-mechanical damage and healing processes across the scales. Such a framework would be particularly useful to model the competition between damage and healing whenever both can occur at the same temperature and pressure conditions (for instance, dislocation glide vs. pressure solution in salt).

(v) To the author's best knowledge, healing models were never implemented in Finite Element codes to solve engineering boundary value problems before the author's group [Shen and Arson, 2019b]. Many challenges still have to be overcome, especially in regards to the mapping of net damage.

Conflicts of Interest

The author declares no conflicts of interest.

Acknowledgements

This research was funded by the U.S. National Science Foundation, under grants CMMI 1362004/1361996: “Collaborative research: Linking Salt Rock Deformation Regimes to Microstructure Organization”, CMMI 1552368: “CAREER: Multiphysics Damage and Healing of Rocks for Performance Enhancement of Geo-Storage Systems - A Bottom-Up Research and Education Approach” and EAR 1755321: “Collaborative research: Microscopic fracturing and macroscopic weakening: A novel model for bedrock fracturing by biotite weathering”. This work would not have been possible without the assistance and support of Dr. Hao Xu, Dr. Cheng Zhu, Dr. Wencheng Jin, Mr. Xianda Shen, Ms. Tingting Xu, Dr. Soleen Le Pense, all current or former students of the author. The author is also thankful to the following collaborators: Dr. Fred Chester, Dr. Judi Chester, Dr. Amade Pouya, Dr. Ken Ferrier and Dr. Niki West.

References


*Manuscript received 7th July 2019, revised 30th September 2019, accepted 20th November 2019.*
Table 1. A few grain and void microstructure descriptors commonly used to define fabric tensors. Note: Grains are represented by virtual ellipses with same second moment.

<table>
<thead>
<tr>
<th>Name</th>
<th>Sketch</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grain size distribution</td>
<td><img src="image" alt="Grain area or volume" /></td>
<td>Pdf of a grain size (typically, largest segment that can be fit in the virtual ellipse).</td>
</tr>
<tr>
<td>Grain orientation distribution</td>
<td><img src="image" alt="Grain orientation" /></td>
<td>Pdf of the orientation of the major axis of the virtual ellipses.</td>
</tr>
<tr>
<td>Grain aspect ratio distribution</td>
<td><img src="image" alt="Grain aspect ratio" /></td>
<td>Pdf of the ratio between the smaller and larger axes of a virtual ellipse.</td>
</tr>
<tr>
<td>Grain roundness distribution</td>
<td><img src="image" alt="Grain roundness" /></td>
<td>Pdf of the ratio of the volume (or area) of a grain over the volume of a sphere (or area of a circle) whose diameter equals the length of the virtual ellipse's major axis.</td>
</tr>
<tr>
<td>Grain solidity distribution</td>
<td><img src="image" alt="Grain solidity" /></td>
<td>Pdf of the ratio of a grain area over the area of the grain's circumscribed circle.</td>
</tr>
<tr>
<td>Branch length distribution</td>
<td><img src="image" alt="Branch length" /></td>
<td>Pdf of the lengths of the segments that link the centroids of two grains in contact (defined as a branch).</td>
</tr>
<tr>
<td>Branch orientation distribution</td>
<td><img src="image" alt="Branch orientation" /></td>
<td>Pdf of the orientation of a branch.</td>
</tr>
<tr>
<td>Coordination number distribution</td>
<td><img src="image" alt="Coordination number" /></td>
<td>Pdf of the number of grains in contact with a given grain.</td>
</tr>
<tr>
<td>Local solid volume fraction distribution</td>
<td><img src="image" alt="Local solid volume fraction" /></td>
<td>Pdf of the solid volume fraction over a polygon with edges matching grains centroids.</td>
</tr>
<tr>
<td>Pore size distribution</td>
<td><img src="image" alt="Pore area or volume" /></td>
<td>Pdf of a pore size (typically, largest segment that can be fit in a virtual ellipse)</td>
</tr>
<tr>
<td>Lineal-path function $L(z)$</td>
<td><img src="image" alt="Lineal-path function" /></td>
<td>Probability of finding a line segment of length $z$ wholly in one of the phases (e.g. grain or pore), when randomly thrown into the sample.</td>
</tr>
<tr>
<td>Chord length distribution</td>
<td><img src="image" alt="Chord" /></td>
<td>Pdf of line segments contained fully within a phase (e.g., grain or pore) and extending between two points on the object's surface (say, pore surface or mineral surface).</td>
</tr>
<tr>
<td>Grain/void nearest-surface distribution $h_j(r)$</td>
<td><img src="image" alt="Grain/void nearest-surface distribution" /></td>
<td>Probability that at an arbitrary point, the nearest surface of a particle or pore of phase $j$ lies at a distance between $r$ and $r+dr$.</td>
</tr>
<tr>
<td>Void exclusion probability $E_V(r)$</td>
<td><img src="image" alt="Void exclusion probability" /></td>
<td>Probability of finding a spherical region of radius $r$ that is empty of solid particle center.</td>
</tr>
</tbody>
</table>