Dehydration-induced instabilities at intermediate depths in subduction zones

Nicolas Brantut, Ioannis Stefanou, Jean Sulem

▶ **To cite this version:**

Nicolas Brantut, Ioannis Stefanou, Jean Sulem. Dehydration-induced instabilities at intermediate depths in subduction zones. Journal of Geophysical Research : Solid Earth, 2017, 10.1002/2017JB014357 . hal-01568176

HAL Id: hal-01568176
https://enpc.hal.science/hal-01568176

Submitted on 24 Jul 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Dehydration-induced instabilities at intermediate depths in subduction zones

Nicolas Brantut[1], Ioannis Stefanou[2] and Jean Sulem[2]

N. Brantut, Department of Earth Science, University College London, Gower Street, London WC1E 6BT, UK. (n.brantut@ucl.ac.uk)

[1]Rock and Ice Physics Laboratory and Seismological Laboratory, Department of Earth Sciences, University College London, London, UK.

[2]Université Paris-Est, Laboratoire Navier, CNRS UMR 8205, École des Ponts – ParisTech, France.

Abstract. We formulate a model for coupled deformation and dehydration of antigorite, based on a porosity-dependent yield criterion and including shear enhanced compaction. A pore pressure and compaction instability can develop when the net volume change associated with the reaction is negative, i.e., at intermediate-depth in subduction zones. The instability criterion is derived in terms of the dependence of the yield criterion on porosity: if that dependence is strong, instabilities are more likely to occur. We also find that the instability is associated with strain localisation, over characteristic length scales determined by the hydraulic diffusivity, the elasto-plastic parameters of the rock, and the reaction rate. Typical lower bounds for the localisation length are of the order of 10 to 100 m for antigorite dehydration and deformation at 3 GPa. The fluid pressure and deformation instability is expected to induce stress build-up in the surrounding rocks forming the subducted slab, which provides a mechanism for the nucleation and propagation of intermediate-depth earthquakes.

1. Introduction

During prograde metamorphism in subduction zones, hydrous phases such as serpentines progressively dehydrate, forming free fluid phases at depth. Such dehydration reactions are systematically associated with a net decrease in solid volume (the reactions forming solid products denser than reactants), and with a variable change in fluid volume, the sign of which being controlled by the pressure and temperature conditions at which the reaction occurs. From initially nonporous metamorphic rocks (such as antigorite serpentinite), dehydration reactions therefore produce, at least transiently, a porous rock saturated with fluids, the rheology of which is markedly different from the original rock [e.g. *Rutter et al.*, 2009]. The occurrence of metamorphic dehydration reactions has therefore a great impact on the stress/strain state in subduction zones.

One key specific impact of dehydration reactions is their potential to trigger unstable faulting and earthquakes, a phenomenon generally termed "dehydration embrittlement". This phenomenon corresponds to the transition from ductile to brittle deformation due to a dehydration-induced increase in pore fluid pressure. It is often thought that dehydration-embrittlement is one of the main causes of, or is at least linked to, intermediate-depth earthquakes in subduction zones [e.g., *Hacker et al.*, 2003b]. Dehydration embrittlement has been observed experimentally [e.g., *Raleigh and Paterson*, 1965; *Murrell and Ismail*, 1976] and is well explained theoretically when the reaction produces an excess fluid volume, i.e., typically at relatively low pressure conditions (e.g., less than around 2.5 GPa in antigorite). Under those conditions, the excess fluid volume generated by the reaction tends to increase the pore fluid pressure, reducing the effective stress, and therefore

bringing back the material into the brittle field. At higher pressure, where the total volume change of the dehydration reaction is negative, laboratory experiments indicate that dehydration embrittlement and earthquake instability might still occur [e.g. *Jung et al.*, 2004], but the exact mechanism remains unclear. One possibility is that the reaction products are plastically weak and facilitate brittle deformation in their surroundings [e.g. *Rutter et al.*, 2009; *Brantut and Sulem*, 2012]. Another potential mechanism is that the porosity generated by the reaction is rapidly compacted, hence producing a pore pressure rise which could bring back the material into the brittle regime.

Coupling between mechanical compaction and dehydration reactions has been investigated in detail in the case of viscous rock rheology: in his seminal study, *Connolly* [1997] (followed by *Connolly and Podladchikov* [1998] and *Connolly and Podladchikov* [2004], summarised recently in *Connolly* [2004]) used a viscous compaction rheology coupled to devolatilisation reactions together with a power-law relation between porosity and permeability, and determined that compaction would drive pore fluid pressure up to near-lithostatic values, while producing intermittent upward motions of fluid (so-called porosity-waves). Despite the great success of this model for the prediction of fluid extraction from the lower crust (or along subduction zones, see *Skarbek and Rempel* [2016]), one key assumption is that the compaction behaviour is essentially driven by a viscous creep process, and does not include the instantaneous response of the material.

Serpentinites, and more specifically the high pressure form, antigorite, are known to behave in a semi-brittle manner even at high pressure and temperature [e.g. *Chernak and Hirth*, 2010; *Proctor and Hirth*, 2016], i.e., antigorite deformation systematically involves a significant degree of microcracking, and its behaviour at high pressure is similar

to cataclastic flow. Furthermore, laboratory experiments [*Rutter et al.*, 2009] show that dehydrated and partially dehydrated serpentinite behave essentially like classical porous rocks, and that concepts of porous rock mechanics can be used to describe their mechanical behaviour. Therefore, a purely viscous constitutive law may not capture all the features and potential instabilities associated with coupled compaction and dehydration in antigorite.

Here, we model coupled dehydration and deformation in antigorite using as a first approximation a time-independent inelastic flow law which includes strain hardening, strain-dependent dilatancy/compaction and a porosity-dependent yield envelope. Our approach is based on the concepts typically used to model the behaviour of porous rocks [e.g *Rudnicki and Rice*, 1975; *Issen and Rudnicki*, 2000; *Wong and Baud*, 2012; *Stefanou and Sulem*, 2014], and dehydration has here an indirect effect by contributing to the overall change in porosity and fluid pressure. In this framework, two types of instabilities can arise: a rate-independent bifurcation related to the constitutive behaviour of the rock, and a reaction-driven, rate-dependent instability due to the growth to small pore pressure perturbations. We specifically focus on the behaviour of serpentinite under conditions such that the dehydration reaction produces a negative total volume change, with the aim of determining whether pore pressure instabilities can occur.

2. Model and governing equations

2.1. Fluid mass balance and pore pressure change

During dehydration, antigorite becomes a porous aggregate, with a porosity n filled with water at a pressure denoted p_{f}. We consider that the porosity of the aggregate is connected (at least at the scale of interest, here of the order of 10 to 100 metres), and that

the fluid flows through the rock according to Darcy's law with a permeability k. The fluid pressure is modified by two independent contributions: the elasto-plastic compaction of the rock (bulk volumetric strain ϵ, taken negative in compression), and the generation of fluids from the dehydration reaction. These assumptions lead to the following governing equation for pore pressure (see full derivation in Appendix A1):

$$\frac{\partial p_\mathrm{f}}{\partial t} = \frac{k}{\eta c_\mathrm{b}} \nabla^2 p_\mathrm{f} - \frac{1}{c_\mathrm{b}} \frac{\partial \epsilon}{\partial t} + \frac{m_\mathrm{d}^0 (1 + \rho_\mathrm{f} \Delta_\mathrm{r} V_\mathrm{s})}{\rho_\mathrm{f} c_\mathrm{b}} \frac{\partial \xi}{\partial t}, \tag{1}$$

where m_d^0 is the total fluid mass that can be released by the reaction per unit rock volume, ρ_f is the fluid density, $\Delta_\mathrm{r} V_\mathrm{s}$ is the solid volume change of the reaction, $\partial \xi / \partial t$ is the reaction rate, and

$$c_\mathrm{b} = (1 - n) c_\mathrm{s} + n c_\mathrm{f} \tag{2}$$

is an effective compressibility combining the compressibility of the solid c_s, the compressibility of the fluid c_f, and the porosity n. Note here that c_b is *not* the usual storage capacity, because we did not split the volumetric strain rate into an elastic and plastic one.

2.2. Rheology

It is well established experimentally that antigorite aggregates undergo a brittle to ductile transition at confining pressures of the order of 300 to 400 MPa [*Escartín et al.*, 1997], and that this transition depends weakly on temperature (within antigorite's stability field). Near the dehydration temperature of antigorite, the ductile behaviour remains dominated by cataclastic flow even at mantle pressures [*Chernak and Hirth*, 2010; *Proctor and Hirth*, 2016], which is most likely due to the strong [001] cleavage plane and the insufficient number of independent slip systems in antigorite single crystals. Across the stability boundary

of antigorite, a very significant volume change occurs: the solid volume tends to decrease by up to around 25%, generating a porosity occupied by pressurised water. Therefore, at pressure and temperature conditions near its stability boundary, antigorite is expected to behave very much like a ductile (cataclastic) porous rock. Such a behaviour has been well documented in lizardite by *Rutter et al.* [2009], and we assume here that the same behaviour applies to antigorite.

In order to describe the elasto-plastic behaviour of antigorite, we introduce the yield function $f(\sigma, \zeta)$, where σ is the stress tensor (here we follow the sign convention of continuum mechanics and take compressive stresses as negative) and ζ is an internal variable on which the yield cap depends. The yield function limits the elastic domain in the stress space ($f < 0$, see Figure 1). It is assumed that inelastic strain increments are generated when the stress state lies on the yield surface ($f = 0$) and if loading is taking place. ζ can be identified to either the porosity of the material, or more directly to the reaction progress in the case of a pure chemical control over the material's strength [*Stefanou and Sulem*, 2014; *Sulem and Stefanou*, 2016]. While either option could be deemed acceptable in the light of the available experimental data from *Rutter et al.* [2009], we will develop our model assuming that the primary control on the rock's strength is given by its porosity (a robust observation in porous rocks, see *Wong and Baud* [2012]). We therefore equate incremental changes in the internal variable ζ to irreversible (inelastic) porosity changes. We also assume, in accordance with experimental observations, that the material undergoes strain hardening. The incremental constitutive behaviour resulting from our

assumptions is written as follows (see full derivation in Appendix A2):

$$dp' = \frac{GK\left[(1+(h-f'\beta)/G)d\epsilon - \beta d\gamma\right]}{h-f'\beta+G+\beta\mu K} + \frac{f'\beta K m_{\rm d}^0 \Delta_{\rm r} V_{\rm s}}{h-f'\beta+G+\beta\mu K} d\xi, \tag{3}$$

$$d\tau = \frac{GK\left[-\mu d\epsilon + \big((h-f'\beta)/K+\beta\mu\big)d\gamma\right]}{h-f'\beta+G+\beta\mu K} + \frac{f' G m_{\rm d}^0 \Delta_{\rm r} V_{\rm s}}{h-f'\beta+G+\beta\mu K} d\xi, \tag{4}$$

where p' is the Terzaghi effective mean stress ($p' = p + p_{\rm f}$, where p is the mean stress), τ is the shear stress (taken equal to the square root of the second invariant of the deviatoric stress tensor), K and G are the bulk and shear elastic moduli of the rock, respectively, γ is the shear strain (taken equal to the square root of the second invariant of the deviatoric strain tensor), h is the strain hardening coefficient, β is the dilatancy factor and $f' = \partial f/\partial\zeta$ is the dependency of the yield cap on porosity. We observe in equations 3 and 4 that the effect of the variation of f with porosity (terms in f') on the mechanical behaviour is entirely captured by the modified hardening modulus $h - f'\beta$. The chemical coupling appears through the product $f'\Delta_{\rm r}V_{\rm s}$: the reaction has only an indirect effect, which is to modify the porosity. The solid volume change is always negative (porosity creation), and f' is positive (the yield surface shrinks with increasing porosity), so that the overall effect of the reaction is to weaken the material.

2.3. Reaction rate

A very general formulation of mineral reaction rates is given by [*Lasaga and Rye*, 1993]

$$\frac{\partial \xi}{\partial t} = \kappa A_{\rm rlm} s |\Delta G|^{n_{\rm r}}, \tag{5}$$

where κ is the temperature-dependent kinetic constant (typically following an Arrhenius law), $A_{\rm rlm}$ is the specific surface area of the rate-limiting mineral, ΔG is the Gibbs energy change of the reaction, s is the opposite of the sign of ΔG, and $n_{\rm r}$ is the order of the reaction. Under isothermal conditions, and for small departures from equilibrium, we can

expand ΔG in terms of pore pressure; only retaining the leading order term, we have [*Wang and Wong*, 2003]

$$\Delta G \approx c'(p_{\rm f} - p_{\rm eq}), \tag{6}$$

where $c' = \partial \Delta G / \partial p_{\rm f}$ and $p_{\rm eq}$ is the pore pressure at equilibrium. Following *Wang and Wong* [2003], we can rewrite 5 as

$$\frac{\partial \xi}{\partial t} = s r_0 \frac{A_{\rm rlm}}{A^0_{\rm rlm}} |1 - p_{\rm f}/p_{\rm eq}|^{n_{\rm r}}, \tag{7}$$

where r_0 is a reference reaction rate, and $A^0_{\rm rlm}$ is the surface area of the rate limiting mineral at the reference rate. The change in surface area is not very well constrained by experimental data. Hence, for simplicity, in the following we will consider the ratio $A_{\rm rlm}/A^0_{\rm rlm} \approx 1$. This simplification is valid only when the reaction progress is small, i.e., when depletion of the reaction is negligible ($\xi \ll 1$).

2.4. Geometry and stress equilibrium

We consider a simple system made of a uniform horizontal layer of antigorite, sufficiently extended so that lateral strains can be neglected (i.e., the system is invariant in the plane of the layer). In this geometry, the vertical stress is given by $p - 2\tau/\sqrt{3}$, and stress equilibrium requires that:

$$\frac{\partial}{\partial y}\left(p - \frac{2}{\sqrt{3}}\tau\right) = 0. \tag{8}$$

where y denotes the vertical coordinate. The boundary condition is a constant applied vertical stress, which implies that

$$\frac{\partial p}{\partial t} - \frac{2}{\sqrt{3}}\frac{\partial \tau}{\partial t} = 0. \tag{9}$$

In such a geometry, only the vertical strain component is nonzero. Because no lateral deformation is allowed, the vertical strain is equal to the volumetric strain. Therefore,

the shear and volumetric strains are related to one another by:

$$\gamma + \frac{2}{\sqrt{3}}\epsilon = 0. \tag{10}$$

Using relations (9) and (10) in the incremental constitutive formulation (Equations 3, 4), and combining with the fluid mass conservation (Equation 1) and expression (7) for the reaction rate, we arrive at a single, nonlinear diffusion equation that governs the fluid pressure (see details in Appendix A3):

$$\frac{\partial p_\mathrm{f}}{\partial t} = \frac{Mk/\eta}{1+c_\mathrm{b}M}\frac{\partial^2 p_\mathrm{f}}{\partial y^2} + \frac{Mm_\mathrm{d}^0(1/\rho_\mathrm{f}+\Delta_\mathrm{r}V_\mathrm{s}) - f'X}{1+c_\mathrm{b}M} sr_0 \left|1-\frac{p_\mathrm{f}}{p_\mathrm{eq}}\right|^{n_\mathrm{r}}, \tag{11}$$

where

$$M = \frac{GK}{h - f'\beta + G + \beta\mu K}\left[\left(1+\frac{2}{\sqrt{3}}\mu\right)\left(1+\frac{2}{\sqrt{3}}\beta\right) + (h-f'\beta)\left(\frac{1}{G}+\frac{4}{3K}\right)\right], \tag{12}$$

and

$$X = -m_\mathrm{d}^0\Delta_\mathrm{r}V_\mathrm{s}\frac{\beta K - 2G/\sqrt{3}}{h - f'\beta + G + \beta\mu K}. \tag{13}$$

3. Parameters

3.1. Yield function

The model described above contains a number of parameters that ought to be constrained from experimental data. The yield cap can be constrained from the extensive dataset of *Rutter et al.* [2009] on intact and dehydrated blocks of lizardite. The data used are shown in Figure 2. For simplicity, we use a modified Cam-clay yield surface, given by

$$f(p',\tau) = \tau - C\sqrt{(b+p')(p^*-p')} = 0, \tag{14}$$

where C is the critical state line ratio, b is the tensile strength, and p^* is the compaction yield pressure (following the notation of *Wong and Baud* [2012]). Specific values of C, b and p^* for intact, partially and fully dehydrated serpentinites are reported in Figure 2.

In accordance with observations on porous sandstones, the critical compaction pressure p^* decreases with increasing porosity. *Zhang et al.* [1990] proposed a grain crushing model in which p^* scales with the porosity n as $p^* \propto n^{-3/2}$. For the serpentinite samples dehydrated at 35 MPa effective pressure and deformed at room temperature (squares in Figure 2), we find a reasonable fit with $p^* = 130$ MPa, while the inferred porosity of the sample was around 19%. The partially dehydrated samples (diamonds in Figure 2) had a porosity of around 4%, and we determine a p^* of 380 MPa. The relationship between p^* and n for these two sample types does not seem compatible with the scaling proposed by *Zhang et al.* [1990]. Here we will use an empirical scaling $p^* \propto 1/n$ [*Rutter and Glover*, 2012], with a constant of proportionality equal to 19 MPa. The corresponding yield caps are reported as dotted lines in Figure 2.

Obviously, the yield cap described by Equation 14 is not appropriate as the porosity approaches zero, since in that case p^* diverges. However, the focus of this work is the description of rocks that are already undergoing dehydration, i.e., in which the porosity is never exactly zero. Furthermore, in absence of a more complete dataset on dehydrated and partially dehydrated serpentinite, the Cam-clay yield surface is one of the simplest yield criterion which is closed at high pressure (i.e., the material can fail by pure hydrostatic compaction). Hence, our choice for the yield function should be viewed as a first order approximation which incorporates the essential qualitative elements of the behaviour of dehydrating serpentinite: a yield cap that is closed at high pressure and that shrinks with increasing porosity.

3.2. Mechanical and hydraulic parameters

The elastic properties of serpentinite can be obtained from Voigt-Reuss-Hill averages of the single crystal properties, and are given by *Bezacier et al.* [2010]: $K = 67.9$ GPa, $G = 38.5$ GPa. The average Poisson's ratio is hence $\nu = 0.26$. Note that we are modelling dehydrating serpentinite, and hence the average elastic properties of the rock should be made dependent upon the evolving rock mineralogical composition and porosity. However, we focus here on the initiation of the dehydration reaction, and hence expect that the pure antigorite end-member is a good approximation to the overall properties of the rock at the beginning of dehydration.

The friction coefficient μ is given by the local slope of the yield envelope (equation A24):

$$\mu = \frac{\partial f}{\partial p'} = C\frac{2p' + b - p^*}{2\sqrt{(p' + b)(p^* - p')}}. \tag{15}$$

In the framework of associated plasticity, we could assume that the dilatancy factor is merely equal to the friction coefficient. However, it is well known that rocks do not follow associated flow rules, and hence we shall leave the dilatancy factor β as a free parameter, and explore how it influences the stability of compaction in our model. Likewise, we will leave the hardening coefficient as a free parameter, in order to encompass the widest possible range of behaviours.

The permeability of the rock is expected to vary as a function of porosity, and hence be impacted by the compaction of the rock. However, these second order controls on permeability should only influence the behaviour of the material at large times, and not the initiation of the instability. For instance, using power-law permeability-porosity relationships, *Connolly* [1997] has shown that a dehydrating rock can generate travelling porosity pulses, a well known feature of nonlinear parabolic equations. Since we want

to focus on the initiation of the instability, we shall assume a constant value for the permeability, keeping in mind that this assumption should be relaxed when modelling the long-term evolution of the system. The permeability of dehydrating serpentinite was measured by *Tenthorey and Cox* [2003], who report $k = 10^{-22}$ m^2 for the intact material and $k = 10^{-20}$ to 10^{-18} m^2 during dehydration at 600°C and 700°C, respectively. Here we choose $k = 10^{-20}$ m^2 as a representative value for serpentinites which are dehydrating not far from equilibrium.

The viscosity of the fluid η can be precisely determined from interpolation of experimental measurements. We will use here the formulation of the International Association for the Properties of Water and Steam, 2008 (http://iapws.org/relguide/visc.pdf), and compute the appropriate viscosity at the target pressure and temperature conditions. The resulting viscosity of water ranges from $\eta = 6.9 \times 10^{-5}$ to 4.9×10^{-7} Pa s at $p_\mathrm{f} = 2$ GPa and 600°C and $p_\mathrm{f} = 5$ GPa and 700°C, respectively.

The effective compressibility of the rock, c_b, does not play any role in the stability of the system; it is only required for the computation of the full numerical solutions, and acts as a scaling factor for the pore pressure rate. Here we compute c_b from Equation (2), assuming that $c_\mathrm{s} = 1/K$.

3.3. Chemical parameters

The parameters associated with the chemical reaction can be obtained from the thermodynamic properties of the mineral (and fluid) involved. At elevated pressure and

temperature, there are three distinct dehydration reactions involving antigorite:

$$\text{antigorite} \rightarrow 4 \text{ talc} + 18 \text{ forsterite} + 27 \text{ water}, \tag{16}$$

$$\rightarrow 14 \text{ forsterite} + 10 \text{ enstatite} + 31 \text{ water} \tag{17}$$

$$\rightarrow (14/5) \text{ phase A} + (71/5) \text{ enstatite} + (113/5) \text{ water}. \tag{18}$$

The properties of each mineral can be extracted from thermodynamic tables [e.g. *Holland and Powell*, 1998; *Hacker et al.*, 2003a], and are reported in Table 1. The thermodynamic properties of water are determined as a function of pressure and temperature from the IAPWS formulation 95.

The computation of the solid volume change $\Delta_r V_s$ requires the knowledge of the molar volumes of each solid phase at the pressure and temperature conditions of each reaction. The molar volumes are computed following the approach explained in *Hacker et al.* [2003a], which is recalled in Appendix B for completeness.

The phase boundaries and net volume change of each reaction (16, 17, 18) are shown as a function of pressure and temperature in Figure 3. Average volume changes along each phase boundary are presented in Table 2.

In addition, the knowledge of the molar volume of antigorite allows to compute precisely the total potential mass of water releasable by each reaction, m_d^0. Because antigorite is not very compressible, the main factor influencing m_d^0 is the stoichiometry of the reactions. The computed averages of m_d^0 for each reaction are shown in Table 2.

The reaction kinetics of antigorite as a function of pressure is not well constrained by existing experimental data, which typically focus on the effect of temperature. However, the formulation 5 is general, and hence kinetic parameters obtained from experiments in which ΔG is imposed from a temperature over- or under-step should also be valid in the

case when ΔG changes due to pressure fluctuations. Here we extract kinetic parameters from the study of *Eggler and Ehmann* [2010], in which the dehydration kinetics of antigorite was determined at 2 GPa (i.e., for reaction 17) as a function of temperature. At a temperature T, the rate of antigorite dehydration is given as

$$\text{rate} = \kappa' A_{\text{rlm}} s \left| \frac{\Delta G}{RT} \right|^{n_\text{r}} \quad \text{in mol}_{\text{atg}}/\text{cm}^3_{\text{rock}}/\text{s}, \tag{19}$$

where R is the gas constant. In terms of reaction progress ξ, we rewrite equation 19 as

$$\frac{\partial \xi}{\partial t} = s r_0 \frac{A_{\text{rlm}}}{A^0_{\text{rlm}}} |1 - p_\text{f}/p_\text{eq}|^{n_\text{r}}, \tag{20}$$

where

$$r_0 = V_\text{m} A^0_{\text{rlm}} \kappa' \left| \frac{p_\text{eq} c'}{RT} \right|^{n_\text{r}}. \tag{21}$$

The molar volume of antigorite V_m can be taken as the average along the phase boundary between antigorite and forsterite, enstatite and water, using equation B6. This yields $V_\text{m} \approx 1.73 \times 10^{-3}$ m^3/mol$_{\text{atg}}$. The initial specific surface area of antigorite, A^0_{rlm}, depends on the grain size and shape; for square prisms with width W and length L, $A^0_{\text{rlm}} = (2W^2 + 4LW)/LW^2$. Using $W = 50$ μm and $L = 10$ μm, we find $A^0_{\text{rlm}} = 2.8 \times 10^5$ m^2/m^3. The rate constant κ' given by *Eggler and Ehmann* [2010] is 9.2×10^{-11} mol$_{\text{atg}}$/m^2/s. The coefficient c', as defined in equation 6, is the net volume change of the reaction per unit mole of antigorite:

$$c' = \nu_\text{f} M_\text{f} (1/\rho_\text{f} + \Delta_\text{r} V_\text{s}). \tag{22}$$

Using the parameter value reported in Table 2 for the reaction of antigorite into forsterite and enstatite, we obtain $c' \approx 8.1 \times 10^{-5}$ m^3/mol$_{\text{atg}}$. The equilibrium pressure p_eq and the appropriate temperature T can be found from the phase boundary (see Figure 3). As a representative value, we choose $p_\text{eq} = 3$ GPa and $T \approx 640$°C. Finally, *Eggler and Ehmann*

[2010] report that the exponent n_r is equal to one, i.e., the kinetic is linear. Combining all the above parameters into our lumped kinetic parameter r_0, we find a representative value of

$$r_0 \approx 1.47 \times 10^{-6}/s. \tag{23}$$

4. Stability analysis

The system governed by Equation (11) involves two coupled phenomena: mechanical deformation (compaction), through the parameter M, and metamorphic effects with fluid production and porosity creation through the source term in (11). In this Section, we detail the possible sources of mechanical and chemical instabilities, and derive the key stability conditions in terms of the model parameters.

4.1. Mechanical instability

If $M < 0$, we immediately observe that Equation 11 is a diffusion equation with a negative diffusivity. This corresponds to an unstable system as non uniformities become more localised rather than more diffuse with increasing time. In fact, the condition $M < 0$ is strictly equivalent to the compaction localisation condition derived by *Issen and Rudnicki* [2000, their equation 19] for axisymmetric compression:

$$\frac{h - f'\beta}{G} < \frac{h_{\text{crit}}}{G} = -\frac{3K}{3K + 4G}\left(1 + \frac{2}{\sqrt{3}}\mu\right)\left(1 + \frac{2}{\sqrt{3}}\beta\right), \tag{24}$$

with the additional term $-f'\beta$ occurring here due to the explicit dependence of the yield cap on porosity. For a yield cap shrinking with increasing porosity, as expected and observed in dehydrating serpentinites (Figure 2), f' is positive, and β is negative (due to shear-enhanced compaction at high pressure). Hence, the critical hardening modulus is reduced when the yield cap is assumed to be directly dependent on porosity; this merely

reflects that the yield cap expansion due to shear hardening is in fact offset by a further yield cap expansion (resp. shrinkage) due to shear-induced compaction (resp. dilatancy).

Note, in passing, that we have derived here a slightly more general case for the compaction instability including pore fluid pressurisation effects; the stability condition, without chemical effects, is in fact $M/(1+c_\mathrm{b}M) < 0$. This has no practical consequence on the criterion because c_b is usually very small.

4.2. Chemical instability and overpressure development

The case of interest here is when $M > 0$ (Equation (12)), i.e., a mechanically stable case. Equation 11 is a nonlinear diffusion equation. We examine here the case when the total volume change of the reaction is negative, for which the reaction kinetics increases when p_f increases beyond p_eq. Before delving into the full analysis including the effect of pore pressure diffusion, it is instructive to first analyse the undrained case; the governing equation for pore pressure reduces to

$$\frac{\partial p_\mathrm{f}}{\partial t} = \frac{M m_\mathrm{d}^0 (1/\rho_\mathrm{f} + \Delta_\mathrm{r} V_\mathrm{s}) - f' X}{1 + c_\mathrm{b} M} r_0 \left(\frac{p_\mathrm{f}}{p_\mathrm{eq}} - 1 \right)^{n_\mathrm{r}}. \tag{25}$$

Assuming constant parameters, this equation has an analytical solution, which is

$$p_\mathrm{f}(t) = p_\mathrm{eq} + (p_\mathrm{f}^0 - p_\mathrm{eq}) \exp\left(\frac{M m_\mathrm{d}^0 (1/\rho_\mathrm{f} + \Delta_\mathrm{r} V_\mathrm{s}) - f' X}{p_\mathrm{eq}(1 + c_\mathrm{b} M)} r_0 t \right) \tag{26}$$

if $n_\mathrm{r} = 1$, and

$$p_\mathrm{f}(t) = p_\mathrm{eq} + \left(p_\mathrm{eq}^{-n_\mathrm{r}} (1 - n_\mathrm{r}) r_0 \frac{M m_\mathrm{d}^0 (1/\rho_\mathrm{f} + \Delta_\mathrm{r} V_\mathrm{s}) - f' X}{1 + c_\mathrm{b} M} t + \left(p_\mathrm{f}^0 - p_\mathrm{eq} \right)^{1 - n_\mathrm{r}} \right)^{1/(1-n_\mathrm{r})} \tag{27}$$

if $n_\mathrm{r} > 1$, where p_f^0 is the initial pore pressure in the system (a small perturbation above the equilibrium pressure).

In both cases (linear and nonlinear kinetics), the evolution of pore pressure is an unbounded growth if

$$Mm_{\mathrm{d}}^{0}(1/\rho_{\mathrm{f}} + \Delta_{\mathrm{r}}V_{\mathrm{s}}) - f'X > 0. \tag{28}$$

For linear kinetics, the growth is exponential, while for nonlinear kinetics the growth corresponds to a finite time blow-up. In practice, this distinction is unimportant since pore pressure diffusion, as well as other nonlinearities not accounted for in our simplified system (such as the depletion of the reactant or the change in mechanical properties with the evolving deformation and mineralogy of the rock), are expected to strongly change the evolution of pore pressure at large times. Despite these subtleties, the solution of the undrained problem yields a key condition (Inequality 28) to observe a potential pore pressure runaway. This condition for instability can be expressed as a pair of conditions in terms of the dilatancy factor β and the dependence of the yield cap on porosity f' (see Appendix C1 for details):

$$\beta < \beta_{\mathrm{crit}} = -(\sqrt{3}/2)\left(\frac{3}{2}\frac{1-\nu}{1-2\nu}\frac{1}{\rho_{\mathrm{f}}\Delta_{\mathrm{r}}V_{\mathrm{s}}} + 1\right)^{-1}, \tag{29}$$

$$f' > f'_{\mathrm{crit}} = \sqrt{3}\frac{1-\nu}{1-2\nu}\frac{h - h_{\mathrm{crit}}}{1-\beta/\beta_{\mathrm{crit}}}\left(1 + \frac{1}{\rho_{\mathrm{f}}\Delta_{\mathrm{r}}V_{\mathrm{s}}}\right). \tag{30}$$

Using the numerical values detailed in the previous section, we remark that β_{crit} is always positive. Hence, for the cases of interest where $\beta < 0$ (shear enhanced compaction), we always have $\beta < \beta_{\mathrm{crit}}$, and thus the condition for instability is simply $f' > f'_{\mathrm{crit}}$.

The stability boundaries ($h = h_{\mathrm{crit}} + f'\beta$ and $f' = f'_{\mathrm{crit}}$) are shown as function of f'/G and β in Figure 4, where we have assumed $\beta = \mu$. For $h > 0$, the mechanical compaction bifurcation arises only for negative values of f', while the reaction-driven compaction and pore pressure instability occurs only for positive values of f', leaving an area of

stability between these two boundaries. For $h = 0$ (which seems relatively plausible for antigorite, see *Escartín et al.* [1997]), the stability boundaries do not overlap; the area of stability expands with decreasing (negative) values of β. Because the friction coefficient μ only appears in the expression for $h_{\rm crit}$, changing μ independently from β (which is a reasonable choice for most rocks) has a moderate impact on the stability diagram for the pore pressure instability (condition (30)), but introduces large changes for the bifurcation criterion (condition (24)).

Using our assumed form for the yield function f, we note that f' diverges as porosity decreases towards zero: in practice, f' is potentially very large at the initiation of the dehydration reaction (when porosity is very low). If the applied stress is such that $\beta < 0$, as in the case at high effective pressure, then the reaction-driven instability is very probable at the onset of dehydration. This implies that the pore fluid pressure will quickly rise to decrease the effective stress; the pressure (solid and fluid) will hence equilibrate very suddenly, generating a transient pore pressure pulse.

4.3. Strain localisation and pore pressure build-up

The considerations above are restrained to the undrained system. The diffusion of pore fluids will tend to stabilise the pore pressure runaway (if it occurs) by draining the rock over a certain length scale. There is no general analytical solution available for equations of the form (11); here we restrict our analysis to a study of the stability of the system to small departures from equilibrium (which corresponds to $p_{\rm f} = p_{\rm eq}$).

As stated in Section 3.3, the experimental data of *Eggler and Ehmann* [2010] show that the near-equilibrium kinetics is linear, $n_{\rm r} = 1$. In that case, the stability analysis detailed in Appendix C2 shows that pore pressure runaways are possible if $f' > f'_{\rm crit}$ (same as

condition 30) and if the spatial wavelength of the perturbation λ is such that

$$\lambda > \lambda_{\rm crit} = 2\pi\sqrt{\frac{p_{\rm eq}(k/\eta)[1 - f'\beta/(h - h_{\rm crit})]}{r_0 m_{\rm d}^0(1/\rho_{\rm f} + \Delta_{\rm r}V_{\rm s})(1 - f'/f'_{\rm crit})}}. \tag{31}$$

We observe that the critical wavelength depends on a number of parameters, some of them well constrained (equilibrium pressure $p_{\rm eq}$, permeability k, fluid viscosity η, and the net volume change of the reaction), and others much more poorly known (essentially, all the parameters associated with the mechanical behaviour, including f' and β). However, as discussed above, the value of f' at the onset of the reaction is expected to be very large since small increments in porosity have large effects on the yield cap when the rock is initially non porous. We can use this fact to our advantage by noticing that the critical wavelength $\lambda_{\rm crit}$ tends to a constant, nonzero value for $f' \gg f'_{\rm crit}$:

$$\lambda_{\rm crit} \sim 2\pi\sqrt{\frac{p_{\rm eq}(k/\eta)}{r_0 m_{\rm d}^0 \Delta_{\rm r} V_{\rm s}}\sqrt{3}\frac{\nu - 1}{2\nu - 1}\frac{1}{\beta^{-1} - \beta_{\rm crit}^{-1}}}. \tag{32}$$

Equation (32) provides a simple lower bound for the critical wavelength, which is, quite remarkably, independent from the hardening modulus h.

The value of $\lambda_{\rm crit}$ is shown in Figure 5 as a function of f' and β for the parameters relevant to the dehydration of antigorite into enstatite and forsterite. As expected, the wavelength tends to the constant given by (32) at large values of f', and we confirm that this limit value has only a mild dependence (square root) on the dilatancy factor β.

Based on the parameter values outlined above for antigorite, we estimate typical values for $\lambda_{\rm crit}$ of the order of 10 to 100 m. In our model, this length scale corresponds to the characteristic width over which pore pressure builds up, and compaction (negative volumetric strain) localises.

4.4. Numerical tests

We performed a series of numerical computations in order to explore further the effects of potential nonlinearities associated with the variations of mechanical, hydraulic and chemical parameters during deformation and reaction. We included a power-law dependence of the permeability on porosity, $k \propto \zeta^3$ (where we recall that ζ is the porosity of the rock minus its poroelastic variations), and accounted for depletion of antigorite by using a simple first-order approximation for the reactant surface area, $A_{\rm rlm}/A^0_{\rm rlm} \approx (1-\xi)$, in Equation (7). The numerical method is described fully in Appendix D, and the Matlab® implementation and source code is available online at `http://github.com/nbrantut/Compaction_Dehydration`. The solution for pore fluid pressure, strain, reaction progress and porosity is computed within a layer of width L, with periodic boundary conditions, and an initial sinusoidal infinitesimal pore pressure perturbation is added to the homogeneous initial conditions. We chose a representative example by using an initial pore pressure of 3 GPa, an initial total mean stress of 3.61 GPa (i.e., an initial shear stress of 0.87 GPa), and an initial porosity of 3 %. Using the initial, reference parameters, we find that the critical wavelength for instability is $\lambda_{\rm crit} \approx 0.08L$, so that we expect some compaction localisation (at least transiently).

The volumetric strain profile within the layer is shown as a function of time in Figure 6(a). The initial ($\epsilon(y,0)=0$) and final (at $t \times r_0 = 100$) profiles are highlighted in black, and intermediate stages are shown light grey. Over time, a net volumetric strain localisation develops around the centre of the layer, and remains there permanently at large times. The time evolution of the compaction localisation instability is better observed in Figure 6(b), which shows the volumetric strain in the centre of the layer ($y=0$) and

on the edge ($y = \pm L/2$). The peak compaction at the centre develops quite rapidly, initially accelerating, and then develops over a timescale of the order of $t \times r_0 \approx 10$, and then stabilises. At the edges of the layer, the compaction develops more slowly and a strong strain gradient develops in the initial phase of the instability; once the compaction stabilises in the centre (at around $t \times r_0 \approx 20$), the strain becomes more homogeneous as the edges also compact further until the whole process eventually stabilises, leaving only a slight strain heterogeneity near the centre.

The evolution of all other key variables as a function of time is shown in Figure 7. The onset of the compaction instability is marked by a rapid acceleration of the reaction progress, porosity and pore pressure changes localised in the centre of the layer (solid lines, $t \times r_0$ between 10 and 15). As the reaction progresses further the differences between the reaction progress, porosity, strain and pore pressure at the centre and at the edges progressively decrease. As the reaction approaches completion, the porosity stabilises at around 27% while the pore pressure remains very high, corresponding to a mean effective stress of around 0.07 GPa. At this point, no further compaction is possible because the periodic boundary conditions effectively ensure that the system is undrained.

The full stress path of each material element in the layer is drawn in Figure 8 in the effective mean stress, shear stress (p', τ) space. As the pore pressure and porosity increase in the layer, the yield surface shrinks and the stress state evolves to maintain mechanical equilibrium. The stress paths of different elements in the layer are not exactly the same, since heterogeneities in pore pressure and stress develop between the centre (solid line) and the edges (dashed line). As the reaction approaches completion, the effective stress becomes very small ($p'/\sigma_{\mathrm{n}} \approx 0.019$, i.e., $p' \approx 0.07$ GPa) and the material cannot compact

further because of the residual (small) nonzero strength, illustrated by the final yield surface.

Overall, the numerical simulation confirms that a compaction instability is possible, but that it is only transient as the system tends again to a homogeneous state while the reaction approaches completion. The typical timescale for instability is of the same order of magnitude as $1/r_0$, i.e., around 7 to 8 days. In our simulations, we impose periodic boundary conditions so that the layer is effectively undrained: hence, the porosity and pore pressure remain high after the reaction is completed. We performed complementary simulations using drained boundaries as another end-member case scenario, and observed that there is a long-term compaction occurring after the initial instability, and the pore pressure returns progressively towards its equilibrium value over timescales determined by the drained length across the layer.

5. Discussion

5.1. Model assumptions and limitations

The key assumption of the model presented here is that ductile deformation of antigorite is essentially time-independent. This approximation is justified by experimental observations of antigorite deformation at elevated pressures and temperature, showing that cataclastic mechanisms typically dominate [*Chernak and Hirth*, 2010; *Amiguet et al.*, 2014; *Auzende et al.*, 2015] and that fully plastic flow is unlikely to occur due to the large crystal anisotropy and lack of available slip systems in antigorite. Furthermore, partially dehydrated serpentinite has also been shown to deform very similarly to porous sandstones [*Arkwright et al.*, 2008; *Rutter et al.*, 2009], which motivates the use of a closed yield envelope.

However, there are also clear experimental indications that ductile flow of antigorite depends on strain rate [e.g. *Hilairet et al.*, 2007; *Amiguet et al.*, 2012], and time-dependency may not be negligible if deformation occurs over very long timescales. We can test whether time-dependent plastic flow contributes significantly to deformation by estimating the strain rates developing during the instability in our model. As mentioned in the previous Section, the characteristic timescale over which the instability develops is determined by the reaction kinetics, $1/r_0$. The typical strain achieved during the instability is of the order of a few percents, so that the strain rate is of the order of $r_0/100$, which is around 10^{-8} s^{-1}. Using that strain rate, the typical shear flow stress extrapolated from the plastic flow laws given in *Amiguet et al.* [2012, their figure 6] is of the order of 0.1 GPa. Therefore, our model based on time-independent ductile flow is broadly consistent with the rheology of serpentinites as long as relatively low shear stresses are considered (around 0.1 GPa).

However, under near-isostatic conditions, the shear stress is not expected to be large enough to produce significant viscous flow within the timescale of instability. The contribution of viscous flow to isostatic compaction typically scales with the inverse of the porosity [e.g. *Wilkinson and Ashby*, 1975; *McKenzie*, 1984], so that the driving effective mean stress should be of the order of 1 GPa to achieve strain rates of around 10^{-8} s^{-1} for a porosity of around 10%. It is therefore likely that viscous compaction in the absence of shear stresses slows down the development of the pore pressure instability.

A number of coupled deformation-dehydration physical models have been developed based on time-dependent rheology of rocks [e.g. *Connolly*, 1997; *Connolly and Podladchikov*, 1998; *Connolly*, 2004; *Skarbek and Rempel*, 2016]. One key parameter exhibited by these models is the compaction length scale, which is related to the hydraulic diffu-

sivity, compressibility and bulk viscosity of the rock. In this context, the timescale for compaction is entirely determined by the rheology of the rock. In our approach using time-independent ductile deformation, a characteristic length for pore pressure development and compaction (λ_c, Equation (31)) also arises, associated with the hydraulic diffusivity, rheology and reaction rate. This length λ_c is clearly analogue to the one obtained from viscous compaction, and corresponds to the limiting case where the timescale for compaction is determined by the reaction rate. For typical crustal metamorphic reactions, *Connolly* [1997] determines that the minimum compaction length is of the order of 10 to 100 m. Quite interestingly, we find that the minimum compaction length in our model is also of the order of 10 to 100 m for antigorite dehydration at intermediate depth (see Figure 5). This similarity in compaction length scale between the two types of model essentially arises from the similarity in timescales between viscous flow and reaction rate.

Here we made the assumption that the material deforms under isothermal conditions. Although a full analysis including thermal effects is beyond the scope of the present work, we discuss here qualitatively how changes in temperature can arise and modify the behaviour of the material. Firstly, the dehydration reaction of antigorite is endothermic, which constitutes a significant heat sink. Secondly, the irreversible work done by inelastic deformation corresponds to energy dissipation and is a heat source. At 4 GPa and 625°C, the enthalpy change of the reaction is 442 kJ per mole of antigorite (computed from the database of *Holland and Powell* [1998]), i.e., around 2.5×10^2 MJ m^{-3}. For a reaction rate of the order of 10^{-8} s^{-1}, the rate of heat absorbed by the reaction is of the order of 2.5 J m^{-3} s^{-1}. The rate of work of the volumetric strain is given by $\dot{\epsilon} p'$, and for an effective stress p' of around 100 MPa and a characteristic strain rate of the order of 10^{-8} s^{-1}, the

rate of heat generated by inelastic deformation is of the order of 1 J m^{-3} s^{-1}. Although the exact quantities are only approximate, the rate of heat absorbed by the reaction and released by deformation are of the same order of magnitude, and might compensate each other. If significant changes in temperature occur, for instance a cooling due to the endothermic character of the reaction overcoming the heat generated by deformation, the reaction kinetics will also be modified according to the corresponding change in ΔG (Equation (5)). In that case, heat flow across the deformed/reacted zone might be the rate limiting process. Assuming a heat diffusivity of 10^{-6} m^2 s^{-1}, the characteristic diffusion time across a layer of 100 m in thickness is 10^{10} s, and it drops to 10^8 s for a 10 m width layer. Therefore, we expect heat flow to be limiting only when the deformation (and reaction) rate becomes faster than 10^{-10} s^{-1} (respectively, 10^{-8} s^{-1}) in a thick (respectively, thin) layer. In our simulations, such strain rates are achieved transiently during the instability, so that thermal effects might affect the behaviour of the system only after the instability has initiated.

Despite the limitations outlined above, the model formulation is quite general and could be applicable to most devolatilisation reactions in subduction zones but also at shallower depths in the crust. This is especially relevant for relatively cold parts of the crust (or subduction zones), where the viscous creep rates of rocks are slow compared to reaction kinetics, so that the hypothesis of time-independent rheology would be justified. Not all reactions are expected to generate a pore pressure instability, and the surprising result here is that instabilities only arise when the total volume change of the reaction is negative (i.e., at high pressure for antigorite). Under shallow crustal conditions, most devolatilisation reactions are expected to produce a positive volume change, so instaiblities should be

investigated on a case-by-case basis. In any case, the governing equation for pore pressure (Equation (11)) remains valid and can be used to make predictions for fluid flow in active prograde metamorphic settings, such as deep sedimentary basins (where gypsum and clay mineral dehydrate), around rising plutons and magma chambers, and of course along subduction zones.

5.2. Effective stress in subduction zones and implications for intermediate-depth earthquakes

One of the key outcome of our model is that antigorite dehydration at intermediate depth results in a very rapid build-up of pore pressure, reducing the effective mean stress towards near-zero values while maintaining a significant open porosity (at least as long as the fluids are trapped inside the dehydrating layer). This pore pressure build-up occurs despite the net negative volume change of the reaction, and is primarily driven by the collapse of porosity. Hence, we expect the effective mean stress to remain near zero throughout the regions where dehydration proceeds, independently from the net volume change associated with the reaction. While this is also expected in models involving viscous rock deformation, we find here that the path towards low effective stress states is unstable, but only in the case when the net volume change is negative $1/\rho_{\mathrm{f}} + \Delta_{\mathrm{r}} V_{\mathrm{s}} < 0$. This instability arises because the compaction tends to increase pore pressure, and takes the system further away from equilibrium. If the net volume change from the reaction is positive, any compaction and pore collapse would increase pore pressure and bring the system back to equilibrium, unless unrealistic amounts of shear-induced dilatancy occur.

One interesting outcome of our model is that we show that the dehydration and compaction process produces significant shear deformation in the rock, and not just pure

volumetric compaction. In our approach, we used uniaxial strain boundary conditions in order to simulate a simple, tractable problem. In nature, the stress state and boundary conditions are necessarily more complex. The existence of shear stresses, even very small (0.01 to 0.1 GPa), is expected to produce significant shear strains during the pore pressure build-up. The dehydrating body of antigorite then acts as a very deformable layer or inclusion, amplifying the stresses around it. In addition, pore fluid diffusion outside the dehydrating body also contributes to decreasing the effective stress in the surrounding rocks.

The combination of relatively rapid stress amplification and effective pressure reduction provides a reasonable mechanism for the inception of brittle deformation in the surrounding ultramafic rocks (peridotites and metagabbros). Indeed, at the temperatures and pressures relevant to antigorite dehydration and at strain rates as high as 10^{-8} s^{-1} during the pore pressure instability, peridotites have a high strength and are unlikely to accomodate deformation in a purely viscous manner, thus raising elastic stresses and favoring brittle deformation. Such a mechanism is essentially a kind of dehydration embrittlement, but in the surrounding rocks and not in the dehydrating serpentinite itself. While this process had been suggested in earlier works by *Kirby* [1987] or *Rutter et al.* [2009], our model provides first-order quantitative constrains on its likelyhood by establishing a closed-form stability criterion that depends on the rock rheology and reaction kinetics (Equation (28)).

As a final note of caution regarding our interpretations in terms of fluid pressure at intermediate depths, we recall here that the concept of effective stress makes sense only if a uniform pore pressure can be defined for a representative volume element of the rock.

This is most likely the case when porosity is larger than several percents, above the percolation threshold [*Guéguen and Palciauskas*, 1994]. However, over long timescales, the progressive drainage of the pore fluid outside of the dehydrating zones tends to allow compaction to reduce porosity; this reduction in porosity occurs concomitantly with surface diffusion and dissolution-precipitation processes that heal and seal the pore space, leaving a disconnected pore network and free fluids present only as fluid inclusions [e.g. *Smith and Evans*, 1984]. In the long term, pore pressure is not a well defined concept and neither is effective stress. There are currently insufficient constrains on healing mechanisms in silicates to draw definitive quantitative conclusions on the persistence and connectivity of the pore space at intermediate depths in subduction zones, but we expect that the short timescale of the instability (typically of the order of 10 days in our simulations) and the large porosity generated by the reaction (around 20%) ensure a reasonable pore network connectivity and validates the use of the concept of effective stress.

6. Conclusions

We developed a model to simulate coupled deformation and dehydration of antigorite at intermediate depths in subduction zones. Our model shows that dehydration can lead to unstable pore pressure rise and deformation when the net volume change of the reaction is negative, due to a positive feedback between pore fluid pressure, compaction and dehydration rate. The stability criterion (Equation (30)) is controlled by the dependence of the yield envelope on porosity (parameter $f' = \partial f / \partial \zeta$); using estimates for f' consistent with well established mechanics of porous rocks, we find that antigorite dehydration leads to unstable deformation under typical intermediate-depth conditions in subduction zones.

Furthermore, we also show that the instability is associated with localised deformation and fluid pressure over a characteristic length scale controlled by hydraulic diffusivity, rheological parameters and reaction rate. A lower bound for this characteristic length is of the order of 10 to 100 m, commensurate with the viscous compaction length scale obtained in models using time-dependent rheologies.

Our model predicts that the typical strain rates during the instability are of the order of 10^{-8} s^{-1}. At such rates and at the relatively cool dehydration temperature of antigorite (around 600°C), the surrounding, chemically stable peridotites and metagabbroic rocks forming the subducted slab have a high strength, and are expected to build up elastic stresses. The elevated pore pressures associated with the dehydration reaction can therefore transiently bring the surrounding rocks back into the brittle field, thus allowing the nucleation and propagation of earthquakes. This mechanism is clearly a type of dehydration embrittlement, but is crucially based on a rapid stress transfer between weak dehydrating rocks and strong surrounding ultramafic rocks, in a manner conceptually similar to the early model of *Kirby* [1987] for transformation-induced instabilities.

Appendix A: Derivation of governing equations

A1. Fluid pressure

For a reacting porous medium, the continuity equations for the solid skeleton and the fluid are [*Coussy*, 2004]

$$\frac{\partial\big(\rho_\mathrm{s}(1-n)\big)}{\partial t} + \mathrm{div}\big(\rho_\mathrm{s}(1-n)\mathbf{v}_\mathrm{s}\big) = -r, \tag{A1}$$

$$\frac{\partial(n\rho_\mathrm{f})}{\partial t} + \mathrm{div}(n\rho_\mathrm{f}\mathbf{v}_\mathrm{f}) = +r, \tag{A2}$$

where ρ_s is the density of the solid, ρ_f is the density of the fluid, n is the Eulerian porosity, $\mathbf{v}_s$ is the velocity of the solid and $\mathbf{v}_f$ is the velocity of the fluid. In the above equations, r denotes the rate at which fluid mass is generated from solid mass. This will be our definition for the reaction rate. Neglecting gradients in ρ_s and ρ_f, the combination of equations A1 and A2 leads to

$$\frac{1-n}{\rho_s}\frac{\partial \rho_s}{\partial t} + \frac{n}{\rho_f}\frac{\partial \rho_f}{\partial t} + \mathrm{div}(\mathbf{v}_s) + \mathrm{div}\big(n(\mathbf{v}_f - \mathbf{v}_s)\big) = r(1/\rho_s - 1/\rho_f). \tag{A3}$$

We can relate the divergence of the relative fluid velocity with respect to the solid to the gradient in fluid pressure by using Darcy's law [*Coussy*, 2004]:

$$n(\mathbf{v}_f - \mathbf{v}_s) = -(k/\eta)\mathrm{grad}(p_f), \tag{A4}$$

where p_f is the fluid pressure, k is the permeability of the material, and η is the viscosity of the fluid. The divergence of the velocity of the solid is the bulk volumetric (Eulerian) strain rate:

$$\mathrm{div}(\mathbf{v}_s) = \frac{\partial \epsilon}{\partial t}, \tag{A5}$$

where ϵ is the bulk volumetric strain. The combination of relations A4 and A5 with equation A3 yields:

$$\frac{1-n}{\rho_s}\frac{\partial \rho_s}{\partial t} + \frac{n}{\rho_f}\frac{\partial \rho_f}{\partial t} + \frac{\partial \epsilon}{\partial t} - \frac{k}{\eta}\nabla^2 p_f = r(1/\rho_s - 1/\rho_f). \tag{A6}$$

The variation of the fluid density can expressed as

$$\frac{1}{\rho_f}\frac{\partial \rho_f}{\partial t} = c_f \frac{\partial p_f}{\partial t}, \tag{A7}$$

where c_f is the compressibility of the fluid. The variation of the solid density is decomposed into two contributions:

$$\frac{1}{\rho_s}\frac{\partial \rho_s}{\partial t} = c_s \frac{\partial p_s}{\partial t} + \frac{1}{\rho_s}\frac{\partial \rho_s}{\partial m} r, \tag{A8}$$

where c_s is the compressibility of the constituents of the solid skeleton, and m denotes the mass of fluid released by the chemical reaction. The last term in A8 corresponds to the evolution of the average density of the solid skeleton as the reaction proceeds (this term would be zero if the solid was transforming entirely into a fluid, without generating solid products). Combining relations A7 and A8 into equation A6 yields:

$$\big((1-n)c_\mathrm{s}+nc_\mathrm{f}\big)\frac{\partial p_\mathrm{f}}{\partial t}=\frac{k}{\eta}\nabla^2 p_\mathrm{f}-\frac{\partial \epsilon}{\partial t}+r\left(\frac{1}{\rho_\mathrm{f}}-\frac{1}{\rho_\mathrm{s}}-\frac{1-n}{\rho_\mathrm{s}}\frac{\partial \rho_\mathrm{s}}{\partial m}\right). \tag{A9}$$

We now need to express the evolution of the average density of the solid as a function of the reaction progress. We are interested in the following type of chemical reaction:

$$\mathrm{mineral}\,0 \longrightarrow \sum_i \nu_i \mathrm{mineral}\,i+\nu_\mathrm{f}\mathrm{fluid}, \tag{A10}$$

where $\nu_{i,\mathrm{f}}$ are stoichiometric coefficients. Denoting ξ the reaction progress and m_d^0 the total mass of fluid that can be released by the reaction (per unit volume of rock), the solid volume is expressed as

$$V_\mathrm{s}=\frac{m_\mathrm{d}^0}{\nu_\mathrm{f}M_\mathrm{f}}\left(M_0(1-\xi)/\rho_0+\sum_i(\nu_i M_i/\rho_i)\xi\right), \tag{A11}$$

where M_i is the molar mass of constituent i. The average solid mass is denoted m_s. The density of the solid is $\rho_\mathrm{s}=m_\mathrm{s}/V_\mathrm{s}$; hence we have:

$$\frac{1}{\rho_\mathrm{s}}\frac{\partial \rho_\mathrm{s}}{\partial m}=\frac{1}{m_\mathrm{s}}\frac{\partial m_\mathrm{s}}{\partial m}-\frac{1}{V_\mathrm{s}}\frac{\partial V_\mathrm{s}}{\partial m}. \tag{A12}$$

The conservation of mass imposes that $\partial m_\mathrm{s}/\partial m=-1$, so that

$$\frac{1}{\rho_\mathrm{s}}\frac{\partial \rho_\mathrm{s}}{\partial m}=\frac{1}{V_\mathrm{s}}\left(-\frac{1}{\rho_\mathrm{s}}-\frac{\partial V_\mathrm{s}}{\partial m}\right). \tag{A13}$$

The last term in parenthesis of the previous equation corresponds to the solid volume change of the reaction, which we denote $\Delta_\mathrm{r}V_\mathrm{s}$. Keeping in mind that $m=m_\mathrm{d}^0\xi$, we can

use relation A11 to express $\Delta_{\rm r}V_{\rm s}$ as

$$\Delta_{\rm r}V_{\rm s} = -\frac{M_0}{\rho_0 \nu_{\rm f} M_{\rm f}} + \frac{\sum_i \nu_i M_i/\rho_i}{\nu_{\rm f} M_{\rm f}}. \tag{A14}$$

Finally, we have to keep in mind that $V_{\rm s} = 1 - n$ by definition, so that equation A13 becomes

$$\frac{1-n}{\rho_{\rm s}} = -\frac{1}{\rho_{\rm s}} - \Delta_{\rm r}V_{\rm s}. \tag{A15}$$

The combination of the relation above with the mass balance equation A9, eventually leads to the governing equation (1) for pore fluid pressure.

A2. Constitutive behaviour

We assume that the dehydrating rock is elasto-plastic. We introduce the yield function $f(\sigma, \zeta)$, where σ is the stress tensor and ζ is an internal variable on which the yield cap depends. ζ can be identified to either the finite porosity of the material, or more directly to the reaction progress in the case of a pure chemical control over the material's strength. These two options will be discussed later on. We also assume, in accordance with experimental observations, that the material undergoes strain hardening. For the sake of simplicity, f and g were assumed linear in terms of τ. In that case, the consistency condition for plastic loading is therefore

$$\left(\frac{\partial f}{\partial \sigma}\right)^{\rm T} d\sigma + \frac{\partial f}{\partial \zeta} d\zeta - h d\lambda = 0, \tag{A16}$$

where h is the hardening modulus and $d\lambda$ is a positive infinitesimal scalar (so-called plastic increment). If we now introduce the plastic potential $g(\sigma)$, the elasto-plastic stress increment is then given by

$$d\sigma = \mathbf{M}^{\rm el} d\epsilon - d\lambda \mathbf{M}^{\rm el} \frac{\partial g}{\partial \sigma}, \tag{A17}$$

where $\mathbf{M}^{\mathrm{el}}$ is the elastic tensor and ϵ is the total strain tensor. The combination of equations A16 and A17 allows the determination of the plastic increment $d\lambda$ and yields the full incremental constitutive relation

$$d\sigma = \mathbf{M}^{\mathrm{ep}} d\epsilon + \mathbf{\Psi} d\zeta, \tag{A18}$$

where

$$\mathbf{M}^{\mathrm{ep}} = \mathbf{M}^{\mathrm{el}} - \frac{\mathbf{M}^{\mathrm{el}} \dfrac{\partial g}{\partial \sigma} \left(\dfrac{\partial f}{\partial \sigma} \right)^{\mathrm{T}} \mathbf{M}^{\mathrm{el}}}{h + \left(\dfrac{\partial f}{\partial \sigma} \right)^{\mathrm{T}} \mathbf{M}^{\mathrm{el}} \dfrac{\partial g}{\partial \sigma}} \tag{A19}$$

and

$$\mathbf{\Psi} = - \frac{\mathbf{M}^{\mathrm{el}} \dfrac{\partial f}{\partial \zeta} \dfrac{\partial g}{\partial \sigma}}{h + \left(\dfrac{\partial f}{\partial \sigma} \right)^{\mathrm{T}} \mathbf{M}^{\mathrm{el}} \dfrac{\partial g}{\partial \sigma}}. \tag{A20}$$

We further assume that the material is isotropic. The scalars p' and τ are used here and they represent, respectively, the Terzaghi effective mean stress (i.e. the difference between the total mean stress and the pore pressure, $p' = \mathrm{tr}(\sigma)/3 + p_{\mathrm{f}}$) and the shearing stress intensity. The shearing stress intensity τ is defined as the square root of the second invariant of the deviatoric part, s, of the stress tensor: $\tau = \sqrt{\frac{1}{2} s_{ij} s_{ij}}$. The Einstein summation convention is adopted and the indices i, j take values $1, 2, 3$. In the (p', τ) space, the stress vector σ is defined as:

$$\sigma = \begin{pmatrix} p' \\ \tau \end{pmatrix}. \tag{A21}$$

Likewise, the strain vector can be described by the volumetric strain ϵ and shear strain γ:

$$\epsilon = \begin{pmatrix} \epsilon \\ \gamma \end{pmatrix}. \tag{A22}$$

The elastic tensor is written then:

$$\mathbf{M}^{\mathrm{el}} = \begin{pmatrix} K & 0 \\ 0 & G \end{pmatrix}, \tag{A23}$$

where K is the bulk modulus of the porous material, and G its shear modulus. For a general plastic behaviour, the derivatives of the yield function and plastic potential are expressed as follows:

$$\frac{\partial f}{\partial \sigma} = \begin{pmatrix} \mu \\ 1 \end{pmatrix}, \quad \frac{\partial g}{\partial \sigma} = \begin{pmatrix} \beta \\ 1 \end{pmatrix}, \tag{A24}$$

where μ is the friction coefficient, and β the dilatancy factor. For the sake of simplicity, f and g were assumed linear in terms of τ. The full expression for the incremental stress-strain relation in the (p', τ) space becomes:

$$dp' = \frac{GK\left[(1 + h/G)d\epsilon - \beta d\gamma\right]}{h + G + \beta\mu K} - \frac{\beta K}{h + G + \beta\mu K}\frac{\partial f}{\partial \zeta}d\zeta, \tag{A25}$$

$$d\tau = \frac{GK\left[-\mu d\epsilon + (h/K + \beta\mu)d\gamma\right]}{h + G + \beta\mu K} - \frac{G}{h + G + \beta\mu K}\frac{\partial f}{\partial \zeta}d\zeta. \tag{A26}$$

In expressions A25 and A26, the factor $\partial f/\partial\zeta$ corresponds to the dependency of the yield cap on the internal variable (or material parameter) ζ. As summarised by *Wong and Baud* [2012], yield caps for porous rocks can be scaled by the critical pressure for hydrostatic pore collapse, usually denoted P^*. It has also been observed [e.g. *Zhang et al.*, 1990] that P^* is scaled by the product of the grain size and the porosity of the rock, to the power 3/2. This dependency of f on P^*, and of P^* on porosity implies that the yield cap can be considered as a function of the porosity of the rock. Hence, a natural choice for the parameter ζ is the total nominal porosity (i.e., the current finite porosity resulting from deformation and reaction, minus any variations produced by elastic deformations). In such a framework, we can write

$$d\zeta = -m_{\mathrm{d}}^0 \Delta_{\mathrm{r}} V_{\mathrm{s}} d\xi + d\epsilon - dp'/K. \tag{A27}$$

Using this expression for ζ into Equations (A25) and (A26), we finally arrive at the incremental constitutive formulation of Equations (3) and (4) in the main text.

A3. Uniaxial compaction

652 The combination of relations 10 and 3 yields

$$\frac{\partial p}{\partial t} + \frac{\partial p_\mathrm{f}}{\partial t} = \frac{GK\left[(1+(h-f'\beta)/G) + 2\beta/\sqrt{3}\right]}{h - f'\beta + G + \beta\mu K}\frac{\partial \epsilon}{\partial t} - \frac{f'\beta K}{h - f'\beta + G + \beta\mu K}\frac{\partial \xi}{\partial t}. \quad \text{(A28)}$$

653 Assuming a constant σ_n over time, equation 9 yields

$$\frac{\partial p}{\partial t} = \frac{2}{\sqrt{3}}\frac{\partial \tau}{\partial t}, \quad \text{(A29)}$$

654 which we combine to equation A28, making use of relation 4, to obtain

$$\frac{\partial p_\mathrm{f}}{\partial t} = M\frac{\partial \epsilon}{\partial t} - f'X\frac{\partial \xi}{\partial t}, \quad \text{(A30)}$$

655 where

$$M = \frac{GK}{h - f'\beta + G + \beta\mu K}\left[\left(1+\frac{2}{\sqrt{3}}\mu\right)\left(1+\frac{2}{\sqrt{3}}\beta\right) + (h-f'\beta)\left(\frac{1}{G}+\frac{4}{3K}\right)\right], \quad \text{(A31)}$$

656 and

$$X = -m_\mathrm{d}^0\Delta_\mathrm{r}V_\mathrm{s}\frac{\beta K - 2G/\sqrt{3}}{h - f'\beta + G + \beta\mu K}. \quad \text{(A32)}$$

657 Now we can use equation 1 to express the volumetric strain rate:

$$\frac{\partial \epsilon}{\partial t} = \frac{k}{\eta}\frac{\partial^2 p_\mathrm{f}}{\partial y^2} + \left(m_\mathrm{d}^0(1/\rho_\mathrm{f} + \Delta_\mathrm{r}V_\mathrm{s})\right)\frac{\partial \xi}{\partial t} - c_\mathrm{b}\frac{\partial p_\mathrm{f}}{\partial t}. \quad \text{(A33)}$$

658 We finally use the expression of the volumetric strain rate given by A33 into equation A30

659 to obtain

$$\frac{\partial p_\mathrm{f}}{\partial t} = \frac{Mk/\eta}{1+c_\mathrm{b}M}\frac{\partial^2 p_\mathrm{f}}{\partial y^2} + \frac{Mm_\mathrm{d}^0(1/\rho_\mathrm{f}+\Delta_\mathrm{r}V_\mathrm{s}) - f'X}{1+c_\mathrm{b}M}\frac{\partial \xi}{\partial t}. \quad \text{(A34)}$$

660 Using the chemical kinetics established in Equation 7, we finally arrive at Equation (11)

661 of the main text.

Appendix B: Volume change for antigorite dehydration

662 The density of a phase as a function of temperature is given by

$$\rho(T) = \rho^\circ e^{-\Phi}, \tag{B1}$$

663 where ρ° is the density under standard conditions (at $T = T_0 = 25°\text{C}$), and

$$\Phi = \ln(V_\text{m}(T)/V_\text{m}^\circ) = \alpha^\circ \left(T - T_0 - 20(\sqrt{T} - \sqrt{T_0})\right). \tag{B2}$$

664 The density as a function of pressure is given by

$$\rho(P) = \rho^\circ(1 + 2\ell)^{3/2}, \tag{B3}$$

665 where ℓ is the linear strain calculated from the bulk modulus K and its derivative with
666 pressure $K' = \partial K/\partial P$:

$$P/K = 3\ell(1 + 2\ell)^{5/2}\left[1 - 2(3 - 3K'/4)\ell + \frac{\ell^2}{6}\left(4(3 - 3K'/4)(4 - 3K') + 5(3K' - 5)\right)\right]. \tag{B4}$$

667 The total change in density as a function of pressure and temperature is finally obtained
668 from

$$\rho(P, T) = [\rho(P)/\rho^\circ]\rho(T), \tag{B5}$$

669 which implies that the molar volume is

$$V_\text{m}(P, T) = M/\rho(P, T) = V_\text{m}^\circ(1 + 2\ell)^{-3/2}e^{\Phi}. \tag{B6}$$

Appendix C: Stability analysis

C1. Derivation of criterion

670 The inequality (28) can be rewritten as

$$f'\left[\Delta_\text{r}V_\text{s}(\beta K - 2G/\sqrt{3}) - \beta(K + 4G/3)(1/\rho_\text{f} + \Delta_\text{r}V_\text{s})\right] > -(K + 4G/3)(h - h_\text{crit})(1/\rho_\text{f} + \Delta_\text{r}V_\text{s}). \tag{C1}$$

Assuming that the material is nominally stable, i.e, $h > h_{\rm crit}$, and considering that the reaction has a total negative volume change ($1/\rho_{\rm f} + \Delta_{\rm r}V_{\rm s} < 0$), the rhs of Inequality (C1) is a positive quantity. If the term in brackets on the lhs is negative, f' would have to be also negative in order to satisfy the inequality. This is a contradiction since the material is porosity-softening and $f' > 0$. So a first requirement for the instability to be possible is that the bracketed term is positive, which implies that

$$\beta \left[\frac{1}{\rho_{\rm f}} + \frac{4G}{3K+4G}\Delta_{\rm r}V_{\rm s} \right] < \frac{2}{\sqrt{3}} \frac{-3G}{3K+4G}\Delta_{\rm r}V_{\rm s}. \tag{C2}$$

Considering that the material is compactant ($\beta < 0$) and that the solid volume change of the reaction is negative ($\Delta_{\rm r}V_{\rm s} < 0$), Inequality (C2) is satisfied when either

$$\frac{1}{\rho_{\rm f}} + \frac{4G}{3K+4G}\Delta_{\rm r}V_{\rm s} > 0, \tag{C3}$$

or

$$\frac{1}{\rho_{\rm f}} + \frac{4G}{3K+4G}\Delta_{\rm r}V_{\rm s} < 0 \text{ and } \beta > \beta_{\rm crit}, \tag{C4}$$

where $\beta_{\rm crit}$ is defined in Equation (29) of the main text. In that case, the instability criterion in terms of f' is given by (see Equation (C1))

$$f' > \frac{-(K+4G/3)(h-h_{\rm crit})(1/\rho_{\rm f} + \Delta_{\rm r}V_{\rm s})}{\Delta_{\rm r}V_{\rm s}(\beta K - 2G/\sqrt{3}) - \beta(K+4G/3)(1/\rho_{\rm f} + \Delta_{\rm r}V_{\rm s})}, \tag{C5}$$

which is exactly the same as Equation (30) after some rearrangements.

C2. Linear analysis

Denoting $p'_{\rm f}$ the small perturbation of $p_{\rm f}$ above $p_{\rm eq}$, we rewrite the governing equation for pore pressure (11) as follows:

$$\frac{\partial p'_{\rm f}}{\partial t} = \frac{Mk/\eta}{1+c_{\rm b}M}\frac{\partial^2 p'_{\rm f}}{\partial y^2} + \frac{Mm^0_{\rm d}(1/\rho_{\rm f} + \Delta_{\rm r}V_{\rm s}) - f'X}{1+c_{\rm b}M} r_0 \left(\frac{p'_{\rm f}}{p_{\rm eq}} \right)^{n_{\rm r}}. \tag{C6}$$

The qualitative behaviour of the system depends on the value of n_r.

For $n_\mathrm{r} = 1$ (i.e., the dehydration reaction kinetics is approximated to be first-order), the equation is linear and we can perform a linear stability analysis to explore how the system evolves. We assume no flux conditions at the boundaries of a dehydrating layer of thickness W (which simulates a serpentinite layer embedded in an impermeable host rock). Then we look for solutions of (C6) (with $n_\mathrm{r} = 1$) in the form

$$p'_\mathrm{f} = A\cos(2\pi y/\lambda)\exp(St), \tag{C7}$$

where A is the amplitude of the perturbation, S is its growth factor, and λ is its wavelength. Our suggested solution must be consistent with the prescribed boundary conditions, hence we require that $\lambda = W/k$, $(k = 1, 2, \ldots)$. The perturbation is unstable $(S > 0)$ if (1) $f' > f'_\mathrm{crit}$ and (2) the wavelength is greater than a critical wavelength λ_crit:

$$\lambda > \lambda_\mathrm{crit} = 2\pi\sqrt{\frac{p_\mathrm{eq}k/\eta}{r_0[m^0_\mathrm{d}(1/\rho_\mathrm{f} + \Delta_\mathrm{r}V_\mathrm{s}) - f'X/M]}}. \tag{C8}$$

After some algebra, the critical wavelength can be rewritten as in (31) in the main text.

For $n_\mathrm{r} > 1$, the situation is mathematically more complicated. Indeed, one can immediately observe in equation (C6) that the reaction term does *not* appear in a first order stability analysis (it is elevated to a power greater than 1). However, one could hope to make useful analytical predictions without resorting to a full numerical treatment. Let us assume that the perturbation has a characteristic amplitude A and a characteristic length scale L, i.e., $p'_\mathrm{f}(y,t) = A(t)g(y/L(t))$ where g is a non dimensional function of the order of 1. The governing equation for p'_f is then

$$\frac{\partial p'_\mathrm{f}}{\partial t} = \frac{A}{L^2}\frac{Mk/\eta}{1 + c_\mathrm{b}M}g''(y/L) + \frac{Mm^0_\mathrm{d}(1/\rho_\mathrm{f} + \Delta_\mathrm{r}V_\mathrm{s}) - f'X}{1 + c_\mathrm{b}M}r_0A^{n_\mathrm{r}}g^{n_\mathrm{r}}(y/L). \tag{C9}$$

The reaction term will be dominant, i.e., the system will be *unstable*, if

$$L \gg \sqrt{\frac{1}{A^{n_\mathrm{r}-1}} \frac{p_\mathrm{eq} k/\eta}{r_0[m_\mathrm{d}^0(1/\rho_\mathrm{f} + \Delta_\mathrm{r} V_\mathrm{s}) - f' X/M]}} = \frac{A^{(1-n_\mathrm{r})/2}}{2\pi} \lambda_\mathrm{crit}. \tag{C10}$$

The condition given in (C10) is very similar to the one obtained in (C8) for the linear case, but here we see that the amplitude of the perturbation A appears in the definition of the critical length scale. Hence, according to (C10), the system is unstable only if the wavelength *and the amplitude* of the perturbations are large enough to overcome diffusion. In any case, we note that the reaction will only dominate the system at early times; one can show that diffusion cannot be neglected everywhere when the system evolves with time. Here we are only interested in the behaviour at early times, because we have assumed that the parameters of the equations are constant. For further evolution of the system, the full nonlinearities should be included and it is not worth going too far mathematically with our simplified system.

Appendix D: Numerical methods

The numerical solution of the fully coupled, nonlinear system is obtained by discretising the governing equation for pore pressure (11) in space using a centered finite difference stencil, and then solving for reaction progress, pore pressure, volumetric strain, total mean stress, porosity and shear stress as a coupled system of ordinary differential equations (ODEs). In practice, we normalise the governing equations by using the magnitude of the imposed total normal stress σ_n as the stress scale, the reaction rate $1/r_0$ as the time scale, and the thickness of the antigorite layer (denoted L). We use a centered finite difference approximation of the second-order spatial derivatives of pore pressure, with a grid defined by points $y_i = i\Delta y$, and implement periodic boundary conditions at the edges $y = 0$ and

$y/L = 1$. The full system of ODEs is then:

$$\frac{\partial \xi_i}{\partial t} = s(1-\xi_i)|p_{\mathrm{f}i}/p_{\mathrm{eq}} - 1|^{n_\mathrm{r}}, \tag{D1}$$

$$\frac{\partial p_{\mathrm{f}i}}{\partial t} = \frac{M}{1+c_\mathrm{b}M}\frac{1}{\eta\Delta y^2}\left(k_{i+1/2}\left(p_{\mathrm{f}i+1}-p_{\mathrm{f}i}\right) - k_{i-1/2}\left(p_{\mathrm{f}i}-p_{\mathrm{f}i-1}\right)\right) + \frac{Mm_\mathrm{d}^0(1/\rho_\mathrm{f}+\Delta_\mathrm{r}V_\mathrm{s}) - f'X}{1+c_\mathrm{b}M}\frac{\partial \xi_i}{\partial t}, \tag{D2}$$

$$\frac{\partial \epsilon_i}{\partial t} = \frac{1}{M}\left(\frac{\partial p_{\mathrm{f}i}}{\partial t} + f'X\frac{\partial \xi_i}{\partial t}\right), \tag{D3}$$

$$\frac{\partial p_i}{\partial t} = \frac{1}{\sqrt{3}/2+\beta-f'/K}\left((f'/K-\beta)\frac{\partial p_{\mathrm{f}i}}{\partial t} - f'\frac{\partial \epsilon_i}{\partial t} + f'm_\mathrm{d}^0\Delta_\mathrm{r}V_\mathrm{s}\frac{\partial \xi_i}{\partial t}\right), \tag{D4}$$

$$\frac{\partial \zeta_i}{\partial t} = -m_\mathrm{d}^0\Delta_\mathrm{r}V_\mathrm{s}\frac{\partial \xi_i}{\partial t} + \frac{\partial \epsilon_i}{\partial t} - \frac{1}{K}\left(\frac{\partial p_i}{\partial t} + \frac{\partial p_{\mathrm{f}i}}{\partial t}\right), \tag{D5}$$

$$\frac{\partial \tau_i}{\partial t} = -\beta\left(\frac{\partial p_i}{\partial t} + \frac{\partial p_{\mathrm{f}i}}{\partial t}\right) - f'\frac{\partial \xi_i}{\partial t}, \tag{D6}$$

where subscripts i indicate variables at point x_i, and

$$k_{i\pm1/2} = (k_i + k_{i\pm1})/2. \tag{D7}$$

In addition, a consistency check is performed by computing the total normal stress:

$$\frac{\partial \sigma_{\mathrm{n}i}}{\partial t} = \frac{\partial p_i}{\partial t} - \frac{2}{\sqrt{3}}\frac{\partial \tau_i}{\partial t}, \tag{D8}$$

and verifying a posteriori that it remains constant throughout space and time. All the parameters that are stress dependent, namely, β, μ and f', are updated at every time and space step to account for the nonlinearities. The ODEs are solved by using Matlab's `ode15s` solver. The full Matlab code is available online at the following url: `www.github.com/nbrantut/Compaction_Dehydration.git`.

Acknowledgments. This work was supported by the UK Natural Environment Research Council through grants NE/K009656/1 and NE/M016471/1 to NB. We are thankful to Bruce Hobbs and John Rudnicki for their valuable review comments.

References

Amiguet, E., B. Reynard, R. Caracas, B. Van de Moortèle, N. Hilairet, and Y. Wang, Creep of phyllosilicates at the onset of plate tectonics, *Earth Planet. Sci. Lett.*, *345-348*, 142–150, 2012.

Amiguet, E., B. Van de Moortèle, P. Cordier, N. Hilairet, and B. Reynard, Deformation mechanisms and rheology of serpentines in experiments and in nature, *J. Geophys. Res.*, *119*, doi:10.1002/2013JB010791, 2014.

Arkwright, J. C., E. H. Rutter, K. H. Brodie, and S. Llana-Fúnez, Role of porosity and dehydration reactions on the deformation of hot-pressed serpentinite aggregates, *J. Geol. Soc. London*, *165*, 639–649, 2008.

Auzende, A.-L., J. Escartin, N. P. Walte, S. Guillot, G. Hirth, and D. J. Frost, Deformation mechanics of antigorite serpentinite at subduction zone conditions from experimentally and naturally deformed rocks, *Earth Planet. Sci. Lett.*, *411*, 229–240, 2015.

Bezacier, L., B. Reynard, J. D. Bass, C. Sanchez-Valle, and B. Van de Moortèle, Elasticity of antigorite, seismic detection of serpentinites, and anisotropy in subduction zones, *Earth Planet. Sci. Lett.*, *289*, 198–208, 2010.

Brantut, N., and J. Sulem, Strain localisation and slip instability in a strain-rate hardening, chemically weakening material, *J. Appl. Mech.*, *73*(3), 031004, doi: 10.1115/1.4005880, 2012.

Chernak, L. J., and G. Hirth, Deformation of antigorite serpentinite at high temperature and pressure, *Earth Planet. Sci. Lett.*, *29*, 23–33, 2010.

Connolly, J. A. D., Devolatilization-generated fluid pressure and deformation-propagated fluid flow during prograde regional metamorphism, *J. Geophys. Res.*, *102*(B8), 18,149–

18,173, 1997.

Connolly, J. A. D., The mechanics of metamorphic fluid expulsion, *Elements*, *6*, 165–172, 2004.

Connolly, J. A. D., and Y. Y. Podladchikov, Compaction-driven fluid flow in viscoelastic rock, *Geodin. Acta*, *11*(2-3), 55–84, 1998.

Connolly, J. A. D., and Y. Y. Podladchikov, Fluid flow in compressive tectonic settings: Implications for midcrustal seismic reflectors and downward fluid migration, *J. Geophys. Res.*, *109*, B04201, doi:10.1029/2003JB002822, 2004.

Coussy, O., *Poromechanics*, John Wiley & Sons, Ltd, Chichester, England, 2004.

Eggler, D. H., and A. N. Ehmann, Rate of antigorite dehydration at 2 GPa applied to subduction zones, *Am. Min.*, *95*, 761–769, 2010.

Escartín, J., G. Hirth, and B. Evans, Nondilatant brittle deformation of serpentinites: Implications for Mohr-Coulomb theory and the strength of faults, *J. Geophys. Res.*, *102*(B2), 2897–2913, 1997.

Guéguen, Y., and V. Palciauskas, *Introduction to the Physics of Rocks*, Princeton University Press, Princeton, New Jersey (USA), 1994.

Hacker, B. R., G. A. Abers, and S. M. Peacock, Subduction factory 1. Theoretical mineralogy, densities, seismic wave speeds, and H_2O content, *J. Geophys. Res.*, *108*(B1), 2029, doi:10.1029/2001JB001127, 2003a.

Hacker, B. R., S. M. Peacock, G. A. Abers, and S. D. Holloway, Subduction factory 2. Are intermediate-depth earthquakes in subducting slabs linked to metamorphic dehydration reactions?, *J. Geophys. Res.*, *108*(B1), 2030, doi:10.1029/2001JB001129, 2003b.

Hilairet, N., B. Reynard, Y. Wang, I. Daniel, S. Merkel, N. Nishiyama, and S. Petitgirard, High-pressure creep of serpentine, interseismic deformation, and initiation of subduction, *Science*, *318*(5858), 1910–1913, 2007.

Holland, T. J. B., and R. Powell, An internally consistent thermodynamic data set for phases of petrological interest, *J. Metamorphic Geol.*, *16*, 309–343, 1998.

Issen, K. A., and J. W. Rudnicki, Conditions for compaction bands in porous rock, *J. Geophys. Res.*, *105*(B9), 21,529–21,536, 2000.

Jung, H. H., H. W. Green, and L. F. Dobrzhinetskaya, Intermediate depth earthquake faulting by dehydration embrittlement with negative volume change, *Nature*, *428*, 545–549, 2004.

Kirby, S. H., Localized polymorphic phase transformations in high-pressure faults and applications to the physical mechanism of deep earthquakes, *J. Geophys. Res.*, *92*(B13), 13,789–13,800, 1987.

Lasaga, A. C., and D. M. Rye, Fluid flow and chemical reaction kinetics in metamorphic systems, *Am. J. Sci.*, *293*, 361–404, 1993.

McKenzie, D., The generation and compaction of partially molten rock, *J. Petrol.*, *25*(3), 713–765, 1984.

Murrell, S. A. F., and I. A. H. Ismail, The effect of decomposition of hydrous minerals on the mechanical properties of rocks, *Tectonophysics*, *31*, 207–258, 1976.

Proctor, B., and G. Hirth, "Ductile to brittle" transition in thermally stable antigorite gouge at mantle pressures, *J. Geophys. Res.*, *121*, doi:10.002/2015JB012710, 2016.

Raleigh, C. B., and M. S. Paterson, Experimental deformation of serpentinite and its tectonic implications, *J. Geophys. Res.*, *70*(16), 3965–3985, 1965.

789 Rudnicki, J. W., and J. R. Rice, Conditions for the localization of deformation in pressure-
790 sensitive dilatant materials, *J. Mech. Phys. Solids*, *23*, 371–394, 1975.
791 Rutter, E. H., and C. Glover, The deformation of porous sandstones; are Byerlee friction
792 and the critical state line equivalent?, *J. Struct. Geol.*, *44*, 129–140, 2012.
793 Rutter, E. H., S. Llana-Fùnez, and K. H. Brodie, Dehydration and deformation of intact
794 cylinders of serpentinite, *J. Struct. Geol.*, *31*, 29–43, doi:10.1016/j.jsg.2008.09.008, 2009.
795 Skarbek, R. M., and A. W. Rempel, Dehydration-induced porosity waves and episodic
796 tremor and slip, *Geochem. Geophys. Geosyst.*, *17*, doi:10.002/2015GC006155, 2016.
797 Smith, D. L., and B. Evans, Diffusional crack healing in quartz, *J. Geophys. Res.*, *89*(B6),
798 4125–4135, 1984.
799 Stefanou, I., and J. Sulem, Chemically induced compaction bands: Triggering conditions
800 and band thickness, *J. Geophys. Res.*, *119*(2), 880–899, 2014.
801 Sulem, J., and I. Stefanou, Thermal and chemical effects in shear and compaction bands,
802 *Geomech. Energy Environ.*, *6*, 4–21, 2016.
803 Tenthorey, E., and S. F. Cox, Reaction-enhanced permeability during serpentinite dehy-
804 dration, *Geology*, *31*(10), 921–924, 2003.
805 Wang, W. H., and T.-F. Wong, Effects of reaction kinetics and fluid drainage on the
806 development of pore pressure excess in a dehydrating system, *Tectonophysics*, *370*,
807 227–239, doi:10.1016/S0040-1951(03)00188-4, 2003.
808 Wilkinson, D. S., and M. F. Ashby, Pressure sintering by power law creep, *Acta Metall.*,
809 *23*, 1277–1285, 1975.
810 Wong, T.-F., and P. Baud, The brittle-ductile transition in porous rock: a review, *J.*
811 *Struct. Geol.*, *44*, 25–53, 2012.

812 Zhang, J., T.-F. Wong, and D. M. Davis, Micromechanics of pressure-induced grain crush-
813 ing in porous rocks, *J. Geophys. Res.*, *95*(B1), 341–351, 1990.

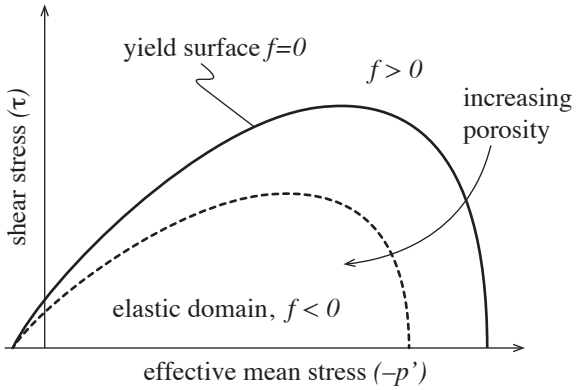

Figure 1. Schematic of the yield surface in the stress space (p', τ). The yield surface is capped at elevated pressures, which corresponds to the possibility of yield under purely isotropic stress conditions. With increasing porosity, the yield surface tends to shrink (reducing the stress range for a purely elastic behaviour), as represented by the dashed line.

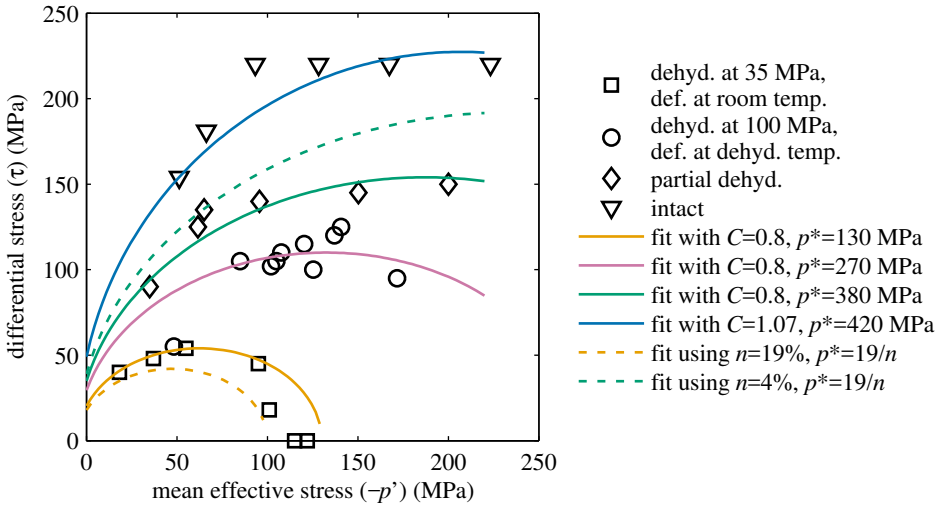

Figure 2. Yield surfaces of intact, partially and fully dehydrated serpentinites from *Rutter et al.* [2009], and fits with a modified Cam-clay model (equation 14). In all fits, we choose a tensile strength $b = 5$ MPa.

Table 1. Thermodynamic properties of phases involved in the dehydration reactions of antigorite. Data from *Holland and Powell* [1998].

	Molar weight	Molar volume	Thermal expansion	Bulk modulus	
	M	V_m°	α°	K°	$\partial K/\partial P$
Phase	g/mol	cm^3/mol	$\times 10^{-5}$ K^{-1}	GPa	
Atg, antigorite	4536	1754.7	4.7	67.9	2.77
Talc	379.7	136.4	3.7	41.6	6.5
Fo, forsterite	140.7	43.7	6.1	127	5.37
Ens, enstatite	200.8	62.6	5.1	106	8.5
PhA, phase A	456.3	154.4	8.3	97.4	6.0

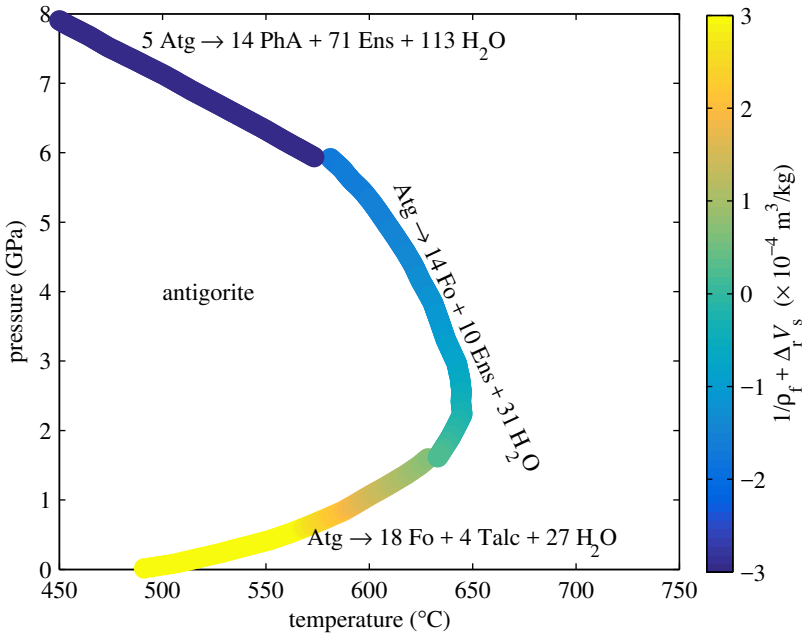

Figure 3. Net volume change $1/\rho_f + \Delta_r V_s$, in m^3 per unit mass of water, as a function of pressure and temperature, for each reaction. Atg: antigorite; Fo: forsterite; Ens: enstatite; PhA: phase A.

Table 2. Average fluid mass and volume change associated with the dehydration reactions of antigorite.

	Mass of releasable fluid	Volume change		
	m_d^0	$1/\rho_f + \Delta_r V_s$		
Reaction	(kg/m^3)	($\times 10^{-4}$ m^3/kg)		
		average	min.	max.
Atg → 4 Talc + 18 Fo + 27 H_2O	276	6.18	0.73	212.55
Atg → 14 Fo + 10 Ens + 31 H_2O	323	−1.02	−1.71	0.25
5 Atg → 14 PhA + 71 Ens + 113 H_2O	240	−3.09	−3.22	−2.93

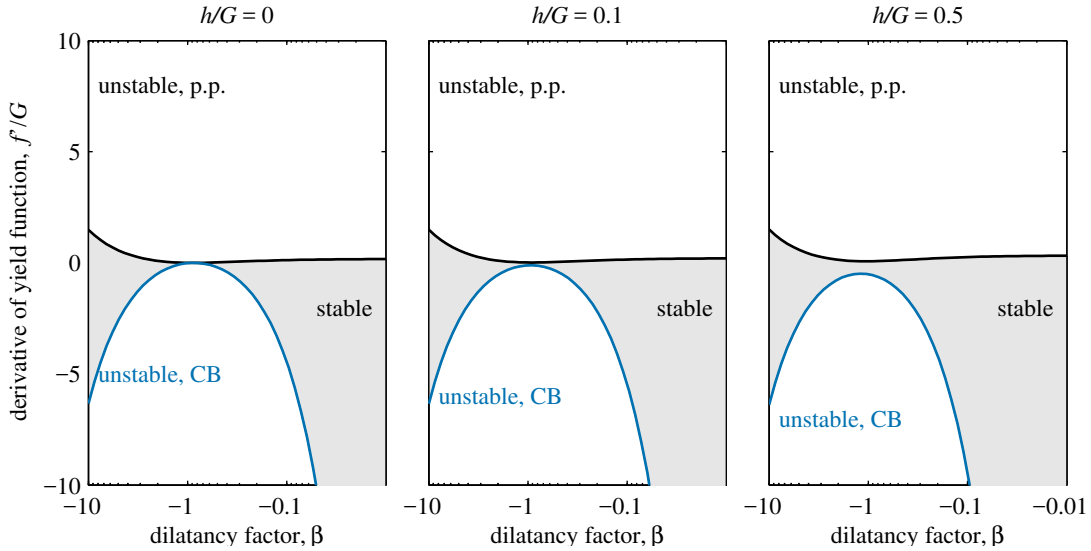

Figure 4. Stability boundaries for a dehydrating material with a porosity-dependent yield function. The dilatancy factor β has been assumed equal to the friction coefficient μ (i.e., we used the so-called associated plasticity assumption) in order to limit the number of free parameters. Black curves correspond to stability boundary for the pore pressure runaway (condition 30), and blue curves correspond to stability boundary for conventional (mechanical) compaction instability (condition 24).

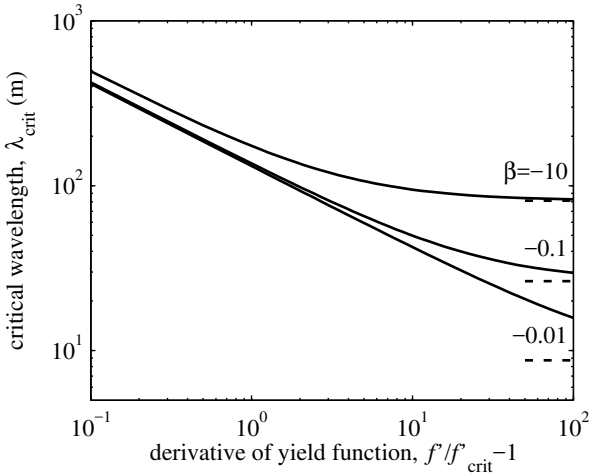

Figure 5. Critical unstable wavelength $\lambda_{\rm crit}$. Values are reported in metres. The parameter values used here are those relevant to the dehydration of antigorite into enstatite and forsterite. For simplicity we assumed $h = 0$ and associated plasticity ($\beta = \mu$). The wavelength becomes infinite along the stability boundary, i.e., for f' approaching $f'_{\rm crit}$, and approaches a constant (dashed lines) given by Equation (32) for $f' \gg f'_{\rm crit}$.

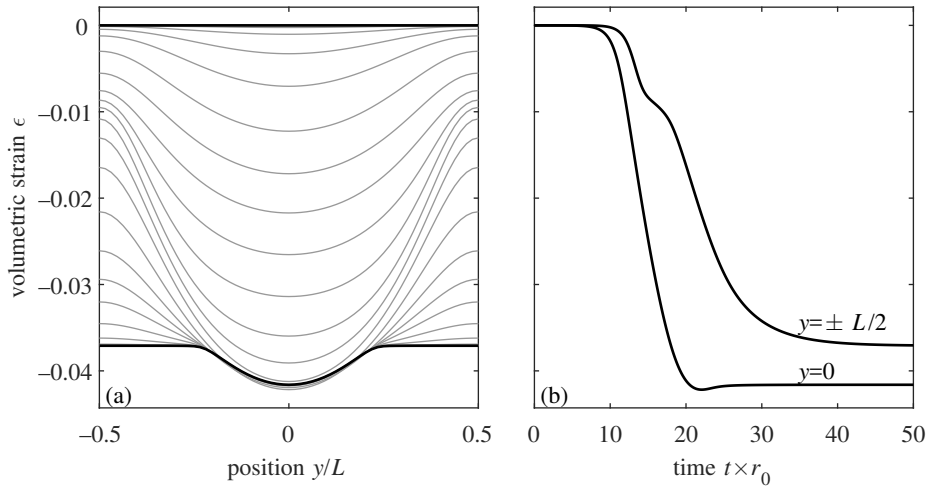

Figure 6. Nonlinear evolution of the compaction instability, modelled numerically. (a) Volumetric strain profiles as a function of time; initial and final profiles are shown in thick black lines, and intermediate profiles are in light grey. (b) Time evolution of volumetric strain at the centre ($y = 0$) and at the edges ($y = \pm L/2$) of the modelled layer.

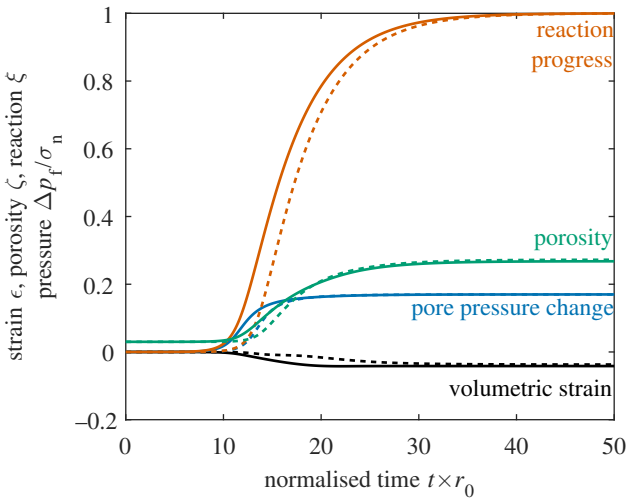

Figure 7. Volumetric strain ϵ (black), reaction progress ξ (orange), porosity ζ (green) and normalised pore pressure change $p_\mathrm{f} - p_\mathrm{f0}/\sigma_\mathrm{n}$ (blue) as a function of normalised time $t \times r_0$. Solid lines show the evolution at the centre of the layer ($y = 0$), and dotted lines correspond to the evolution at the edge of the layer ($y = \pm L/2$).

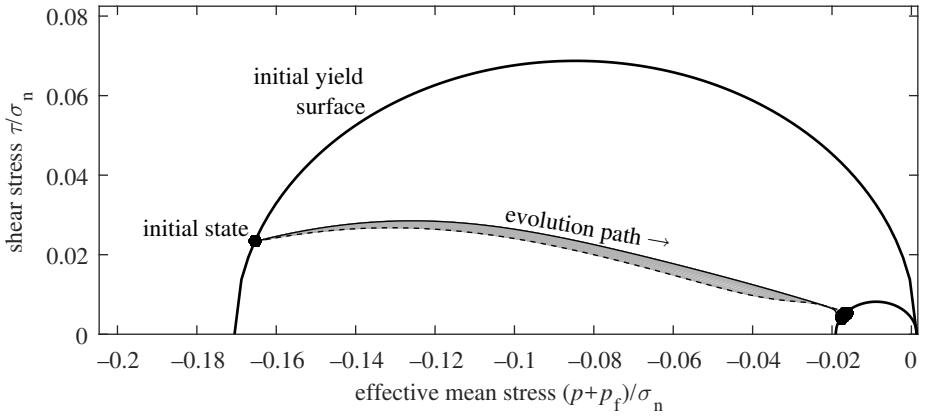

Figure 8. Stress path of material elements within the antigorite layer. The initial and final stress states are given by the filled black circles. The trajectory of the element in the centre of the layer is given by the solid black line, and that of the element at the edge is given by the dashed black line. The trajectories of all other elements are given as the thin grey lines. In this simulation, the total normal stress is $\sigma_n = 3.71$ GPa, so that the initial effective mean stress is 0.61 GPa and the initial shear stress is 0.09 GPa.