# A bound-constrained formulation for complex solution phase minimization

Nicolas Riel [1,2], Boris J.P. Kaus [1,2,3], Albert de Montserrat [4], Evangelos Moulas [1,2,3], Eleanor C.R. Green [5], and Hugo Dominguez [1]

[1]Institute of Geosciences, Johannes Gutenberg-University, Mainz, Germany
[2]Terrestrial Magmatic Systems (TeMaS) research center, Johannes Gutenberg-University, Mainz, Germany
[3]Mainz Institute of Multiscale Modelling (M$^3$ODEL), Johannes Gutenberg-University, Mainz, Germany
[4]Department of Earth Sciences, Institut für Geophysik, ETH Zürich, 8092 Zürich, Switzerland
[5]School of Geography, Earth and Atmospheric Sciences, The University of Melbourne, Victoria 3010, Australia

**Correspondence:** Nicolas Riel (nriel@uni-mainz.de)

**Abstract.**

Prediction of mineral phase assemblages is essential to better understand the dynamics of the solid Earth, such as metamorphic processes, magmatism and the formation of mineral ore deposits. While recently developed thermodynamic databases allow the prediction of stable phase mineral assemblages for an increasing range of pressure, temperature and compositional spaces, the increasing complexity of these databases results in a significant increase of computational cost, hindering our ability to perform realistic models of reactive fluid/magma transport. Presently, prediction of stable phase equilibrium in complex systems is therefore largely limited by how efficiently single phase minimization can be performed, as more than 75 $\%$ of the total computational time is generally dedicated to individual solution phase minimization. This limitation becomes critical for non-ideal solution phase models that involve both a large number of chemical components, and mixing on a large number of sites, resulting in many inequality constraints of the form $0 \leq x_l^M \leq 1$, where $x_l^M$ is the fraction of element $l$ mixing on site $M$.

Here, we present a general reformulation of complex non-ideal solution phases from the thermodynamic database of Holland et al. (2018), which comprises equations of state for multiple mineral solid solutions appearing in magmatic systems, as well as multicomponent silicate melt and aqueous fluid phases. Using a nullspace approach, non-linear inequality constraints governing the site fractions are transformed into equality constraints, and the resulting problem is turned into an bound-constrained optimization problem, subsequently optimized using efficient gradient-based methods. To test our formulation, we apply it to several equations of state for solution phases known for their complexity and compare the results of our approach against classical optimization algorithms supporting inequality constraints.

We find that the the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm yields by far the best performance and stability with respect to the other investigated methods, improving the minimization time of individual solution phase by a factor $\geq$ 10. We estimate that our new approach can improve the computational time of stable phase equilibrium by a factor $\geq 5$, thus potentially allowing to model realistic reactive fluid/magmatic systems by directly integrating phase equilibrium calculations in multiphase thermomechanical codes.

## 1 Introduction

While the last decade has seen significant progress in thermomechanical modeling of complex multiphase systems (e.g., Keller et al., 2013; Taylor-West and Katz, 2015; Keller and Katz, 2016; Keller et al., 2017; Turner et al., 2017; Keller and Suckale, 2019; Rummel et al., 2020; Katz et al., 2022), the coupling with petrological modeling, when addressed at all, remains largely simplified (Riel et al., 2019). There are two key obstacles. First, most phase equilibrium modeling tools (e.g., Perple_X, Theriak_Domino, geoPS, MELTS) (Connolly, 2005; de Capitani and Petrakakis, 2010; Xiang and Connolly, 2021; Ghiorso and Sack, 1995) have been developed with the primary aim of producing phase diagrams and do not offer useful interfaces to integrate with (parallel) geodynamic codes. Second, phase equilibrium modeling is generally achieved by solving a Gibbs energy minimization problem which is computationally challenging. Several numerical strategies have been developed to solve such optimization problems (Ghiorso and Sack, 1995; Connolly, 2005; de Capitani and Petrakakis, 2010; Piro, 2011; Xiang and Connolly, 2021) and some of the most efficient algorithms rely on repeated solution model minimization in order to compute for the most stable mineral assemblage (e.g., de Capitani and Petrakakis, 2010; Xiang and Connolly, 2021; Riel et al., 2022). Although computational performance have been significantly increased over the past few years (Xiang and Connolly, 2021; Riel et al., 2022),single point equilibrium prediction is still costly, with computational times of the order of 10 to 100 ms (e.g., Riel et al., 2022). This limitation effectively precludes direct coupling of phase equilibrium calculations with thermomechanical models, which requires performing from thousands to hundreds of thousands of such calculations every timestep.

In order to account for chemical separation in geodynamic models, several computationally cheaper workarounds have been used. This includes the use of pre-computed set of pseudosections (e.g., Magni et al., 2014; Bouilhol et al., 2015; Rummel et al., 2020), parameterizations (e.g., Jackson et al., 2003, 2018; Hu et al., 2022; Keller et al., 2022) and the ongoing development of neural networks (e.g., Leal et al., 2020; Yuan et al., 2024; Candioti et al., 2024). In (Rummel et al., 2020), the authors generated a database of pre-computed results from phase-equilibria modeling covering the explored/expected compositional, pressure and temperature range of the system. While this approach is powerful, it suffers several limitations. First, to generate a relevant petrological database, the geodynamic model has to be run multiple times in order to characterize the effective pressure, temperature and compositional range of the system. Second, the database is by definition discrete which implies that a compositional tolerance has to be applied when computing the stable phase equilibrium, thus leading to mass conservation issues.

Although the heavy computational requirements of stable phase equilibrium modeling remains a major obstacle for direct coupling, recently developed minimization tools yielded a significant improvement in performance (Xiang and Connolly, 2021; Riel et al., 2022). The recent performance increase mainly results from combining/improving existing minimization methods and making use of gradient-based minimization of individual phases to speed up the computations. Several gradient-based minimization methods are currently employed in the different routines computing phase equilibria. Theriak-Domino (de Capitani and Brown, 1987; de Capitani and Petrakakis, 2010) uses either steepest gradient or Newton-Raphson methods. Minimization

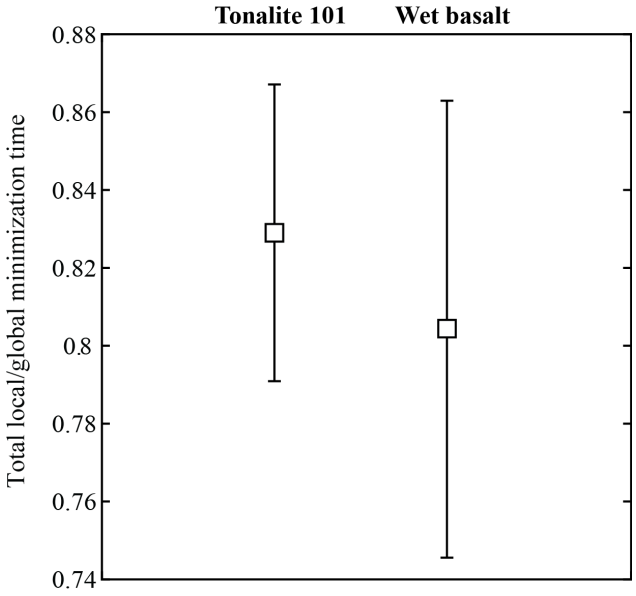


**Figure 1.** Ratio of time spent solving the local problem of the equilibrium composition and order in individual solution models ("total local"), to time spent solving the entire problem of establishing which is the most stable of the possible phase equilibria ("global minimization"), using MAGEMin (Riel et al., 2022) for two representative test cases. Both tonalite and wet basalt bulk-rock compositions are taken from (Holland et al., 2018). In total 619 points were computed from 0 to 12 kbar and from 600 to 900°C and from 800 to 1100°C for the tonalite and the wet basalt case, respectively.

of the solution phase model is achieved using a feasible starting guess and continues until a bound or a site fraction constraint is violated. In our recent phase equilibrium calculation software MAGEMin (Riel et al., 2022), the analytical expressions of the equations of state for solution phases are passed to NLopt software package(Johnson, 2021). Subsequently, the objective function is minimized using the Conservative Convex Separable Approximation with Quadratic penalty (CCSAQ) algorithm (Svanberg, 2002) which solves for inequality-constrained nonlinear programming problems. During the inner iterations, a series of convex sub-problems approximating the objective function and the constraints are generated and solved until the constraints are satisfied (Svanberg, 2002). This procedure is repeated until the solution phase model is minimized. The phase equilibria calculator GeoPS (Xiang and Connolly, 2021) uses the simulated annealing (SA) method. Compared to gradient-based methods, simulated-annealing is a probabilistic technique for approximating the global optimum of a given function (e.g., Pincus, 1970). Here, constraints can be accounted for as penalties on the objective function.

The first release of MAGEMin followed the THERMOCALC software (Powell and Holland, 1988) in treating the physicality of site fractions as a set of inequality constraints expressed as functions of compositional and order variables (Holland et al., 2018) (see e.g., Eq. 11). This approach ensures that all site fractions in the present solution phases satisfy the condition $\geq 0$. Furthermore, since the parameterisation of site fractions enforces that the sum of all site fractions associated with a given site equals unity, this automatically guarantees that all individual site fractions are also $\leq 1$. However, the use of inequality con-

straint in gradient-based methods, results in relatively slow performances and occasional solver failure due to slight violation of inequality constraints. Using the first publicly released version of MAGEMin (Riel et al., 2022), we find that the global minimization time is largely dominated by how fast gradient-based minimization of individual solution phases can be performed, with 75 to 90% of the computation time dedicated to local minimization to find the equilibrium compositions and state of order of solution models (see Figure 1). Therefore, it becomes critically important to improve the minimization time of individual solution phase models to further speed-up the overall phase-equilibrium computational time.

Here, we present a revised implementation of the compositional and order variables (xeos) of (Holland et al., 2018) within MAGEMin that avoids the need to express the site fraction expressions as non-linear inequality constraints. Elimination of these constraints allows using faster bound-constrained optimization methods, thus considerably improving performance and stability of the code. We compare the accuracy and performance of two well-known gradient-based optimization methods: the conjugate gradient (CG) and the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method.

## 2 Methodology

### 2.1 Solution phase formulation

At fixed pressure $P$ and temperature $T$, the total Gibbs energy of solution phase $\lambda$ is given by

$$G_\lambda = \sum_{i=1}^{N_\lambda} \mu_{i(\lambda)} p_{i(\lambda)}. \tag{1}$$

where $G_\lambda$ [J] is the Gibbs energy of the solution phase $\lambda$, $N_\lambda$ the number of end-members of solution phase $\lambda$, $p_{i(\lambda)}$ [mol] is the fraction of end-member $i$ dissolved in solution phase $\lambda$ and $\mu_{i(\lambda)}$ $[\mathrm{J \cdot mol^{-1}}]$ is the molar chemical potential of end-member $i$ in solution phase $\lambda$. An end-member is defined as an independent instance of a solution phase, at a single specified composition, for which the Gibbs energy is fully defined as a function of pressure and temperature only. In a given chemical system, the linear combinations of the end-members span the complete crystallographic site-occupancy space of the solution phase.

The molar chemical potential of a phase is a function of the dissolved end-members within a solution phase (see Ganguly, 2001, for a review)

$$\mu_{i(\lambda)} = g^0_{i(\lambda)} + RT \ln(a^{id}_{i(\lambda)}) + g^{\mathrm{ex}}_{i(\lambda)}, \tag{2}$$

where $R$ $[\mathrm{J \cdot mol^{-1} \cdot K^{-1}}]$ is the ideal gas constant, $T$ [K] is the absolute temperature, $a^{id}_{i(\lambda)}$ [/] is the ideal activity, $g^0_{i(\lambda)}$ $[\mathrm{J \cdot mol^{-1}}]$ the reference molar Gibbs energy of the pure end-member Helgeson (1978); Holland and Powell (1998) and $g^{\mathrm{ex}}_{i(\lambda)}$ $[\mathrm{J \cdot mol^{-1}}]$ is the molar excess energy term Powell and Holland (1993); Holland and Powell (2003). The ideal activity coefficient $a^{id}_{i(\lambda)}$ is generally defined as $a^{id}_{i(\lambda)} = p_{i(\lambda)}$ for molecular mixing, or else for mixing on crystallographic sites as

$$a^{id}_{i(\lambda)} = c_i \prod_s (X^s_{e_{s,i}})^{\nu_s} \tag{3}$$

where $X^s_{e_{s,i}}$ is the site fraction of the element $e_{s,i}$ that appears on site $s$ in end-member $i$ of phase $\lambda$, $\nu_s$ is the number of atoms contained in mixing site $s$ of $\lambda$, and $c_i$ is a normalization constant that ensures that $a^{id}_{i(\lambda)}$ is unity for the pure end-member $i$.

In the asymmetric formalism, $g_i^{ex}$ is given by:

$$g_{i(\lambda)}^{ex} = -\sum_{m=1}^{N_\lambda - 1} \sum_{n>m}^{N_\lambda} (\phi'_m - \phi_m)(\phi'_n - \phi_n) W_{m,n} \left( \frac{2v_i}{v_m + v_n} \right), \tag{4}$$

where $\phi_{m,n}$ is the proportion of end-members $m, n$ weighted by the asymmetry parameters, as $\phi_{m,n} = (p_{m,n} v_{m,n})/(\sum_{k=1}^{N_\lambda} p_k v_k)$, with $v_{m,n,k}$ the van Laar parameters for end-members $m, n, k$. $\phi'_{m,n} = 1$ where $m = n$ and $\phi'_{m,n} = 0$ where $m \neq n$. $W_{m,n}$ $[\mathrm{J \cdot mol^{-1}}]$ is the interaction energy between end-members $m$ and $n$ in the solution.

In Holland et al. (2018), composition (the overall ratios of elements) and order (the distribution of elements over mixing sites) in an xeos are parameterized in terms of an independent set of variables (see example below). Given this formulation, the set of equations 1 to 4 can be directly transformed into the following Gibbs free energy minimization problem as function of the compositional and order variable $x_{cv}$:

$$\min G_\lambda(\mathbf{x}_{cv}) = \sum_{i=1}^{N_\lambda} \mu_{i(\lambda)} p_{i(\lambda)}, \tag{5}$$

subject to the site fraction of the element

$$X^s_{e_{s,i}} \geq 0, \tag{6}$$

and that the compositional and order variables $\mathbf{x}_{cv}$ must be within a lower ($\mathbf{lb}_{cv}$) and upper ($\mathbf{ub}_{cv}$) limit

$$\mathbf{lb}_{\mathrm{cv}} \leq \mathbf{x}_{cv} \leq \mathbf{ub}_{\mathrm{cv}}, \tag{7}$$

where $\mu_{i(\lambda)}$, $p_{i(\lambda)}$, and $X^s_{e_{s,i}}$ are functions of the compositional and order variables $x_{cv}$, and, $\mathrm{lb_{cv}}$ and $\mathrm{ub_{cv}}$ are the lower and upper bounds on the set of compositional and order variables $\mathbf{x}_{cv}$. The first derivative of $f(\mathbf{x}_{cv})$ is given by

$$\frac{\partial f}{\partial \mathbf{x}_{cv}} = \mu_{i(\lambda)} \frac{\partial p_{i(\lambda)}}{\partial \mathbf{x}_{cv}}, \tag{8}$$

and the first derivative of the inequality constraints on the site fraction by

$$\frac{\partial X^s_{e_{s,i}}}{\partial \mathbf{x}_{cv}}. \tag{9}$$

### 2.2 A revised formulation

The solid solutions presented in Holland et al. (2018) are formulated on the basis of exchanging chemical species on a finite number of unique crystallographic sites (Bragg–Williams-type formulation, see Myhill and Connolly, 2021 for more details). A key challenge with this formulation is that minimization has to be performed while keeping site fractions $\geq 0$. Our previous implementation (Riel et al., 2022) imposed these inequalities constraints directly using NLopt (Johnson, 2021), which has a significantly higher numerical cost compared to the bound-constrained minimization algorithms. To simplify the optimization problem and reduce computation time, we use an alternative nullspace formulation (similar to HeFESTo, Stixrude and

Lithgow-Bertelloni, 2011), different from the compositional and order variable approach used in e.g., Holland et al. (2018), that transforms the non-linear inequality constraints on the site fractions into linear equality constraints. This reformulation is illustrated below using olivine as a representative example.

The olivine solid solution model (Holland et al., 2018) contains 2 mixing sites M1 and M2 and represents a phase that can be expressed by the general formula:

$$[\mathrm{Mg}^{2+},\mathrm{Fe}^{2+}]^{\mathrm{M1}}[\mathrm{Mg}^{2+},\mathrm{Fe}^{2+},\mathrm{Ca}^{2+}]^{\mathrm{M2}}\mathrm{SiO}_4 \tag{10}$$

Here $\mathrm{Mg}^{2+}$ and $\mathrm{Fe}^{2+}$ can be exchanged on crystallographic site M1 and $\mathrm{Mg}^{2+}$, $\mathrm{Fe}^{2+}$ and $\mathrm{Ca}^{2+}$ can be exchanged on site crystallographic M2. In Holland et al. (2018), the site fractions are expressed as :

$$X^{\mathrm{M1}}_{\mathrm{Mg}} = 1-x+Q \tag{11}$$
$$X^{\mathrm{M1}}_{\mathrm{Fe}} = x-Q \tag{12}$$
$$X^{\mathrm{M2}}_{\mathrm{Mg}} = (1-x)(1-c)-Q \tag{13}$$
$$X^{\mathrm{M2}}_{\mathrm{Fe}} = x(1-c)+Q \tag{14}$$
$$X^{\mathrm{M2}}_{\mathrm{Ca}} = c \tag{15}$$

where composition has been parameterized using the variables $x = (X^{\mathrm{M1}}_{\mathrm{Fe}} + X^{\mathrm{M2}}_{\mathrm{Fe}})/(X^{\mathrm{M1}}_{\mathrm{Fe}} + X^{\mathrm{M2}}_{\mathrm{Fe}} + X^{\mathrm{M1}}_{\mathrm{Mg}} + X^{\mathrm{M2}}_{\mathrm{Mg}})$ and $c = X^{\mathrm{M2}}_{\mathrm{Ca}}$, and order has been parameterized using the variable $Q = x-(X^{\mathrm{M1}}_{\mathrm{Fe}})/(X^{\mathrm{M1}}_{\mathrm{Fe}} + X^{\mathrm{M1}}_{\mathrm{Mg}})$. Note that we have dropped the ion charges in the notation of the equations.

This parameterisation ensures that the site fractions on each of the individual sites are inherently normalized to 1. Two other types of constraint might be built into the parameterisation in a more complex example: (i) charge balance: if variably-charged ions were mixing, charge balance would be maintained during compositional change, and (ii) equidistribution: the xeos might be simplified by equating two site fractions, typically involving minor elements. The resulting set of composition and order variables is an independent set, that fully and uniquely describes the site occupancies at a given composition and state of order, subject to physical constraints arising from the lattice structure of the mineral. The relationship between the number of composition and order variables and the number of site fractions is then given by:

$$n_{sf} = n_{x_{cv}} + n_{eq_{norm}} + n_{eq_{cb}} + n_{eq_{edist}}, \tag{16}$$

with $n_{sf}$ the number of site fractions, $n_{x_{cv}}$ the number of compositional and order variables, $n_{eq_{norm}}$ the number of constraints arising from the normalization of site fractions on a given site to 1, $n_{eq_{cb}}$ the number of charge balance equations, equal to 0 or 1, and $n_{eq_{edist}}$ the number of equidistribution constraints imposed. Collectively, the normalized charge balance and equidistribution constraints form a set of linear equalities among the site fractions.

In the revised implementation, we retrieve the set of equality constraints $\mathbf{Ax} = \mathbf{b}$ for olivine from the site fractions $\mathbf{X}^s_{e_s} = \{X^{\mathrm{M1}}_{\mathrm{Mg}}, X^{\mathrm{M1}}_{\mathrm{Fe}}, X^{\mathrm{M2}}_{\mathrm{Mg}}, X^{\mathrm{M2}}_{\mathrm{Fe}}, X^{\mathrm{M2}}_{\mathrm{Ca}}\}$, and compositional and order variables $\mathbf{x}_{cv} = \{x, c, Q\}$ as follows. We first take the partial

derivatives of site fractions as functions of the compositional and order variables as:

$$\frac{\partial \mathbf{X}^{s}_{e_{s,i}}}{\partial \mathbf{x}_{cv}} = \begin{bmatrix} -1 & 0 & 1 \\ 1 & 0 & -1 \\ c-1 & x-1 & -1 \\ 1-c & -x & 1 \\ 0 & 1 & 0 \end{bmatrix}, \tag{17}$$

where $\mathbf{X}^{s}_{e_{s,i}}$ is the site fraction of the element $e_{s,i}$ that appears on site $s$, $\mathbf{x}_{cv}$ the set of compositional and order variables $cv$ and $x$, $c$ the compositional and order variables of olivine as defined in Holland et al. (2018). Next, we compute the matrix of site mixing coefficients $\mathbf{A}$ using symbolic expressions as

$$\mathbf{A} = \mathrm{Null}\left(\left(\frac{\partial \mathbf{X}_{e}}{\partial \mathbf{x}_{cv}}\right)^{T}\right) = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}, \tag{18}$$

where Null stands for the null space. We then establish the vector of constraints $\mathbf{b}$ as

$$\mathbf{b} = \{1, 1\}, \tag{19}$$

which ensures that the sum of the site fractions of mixing sites $X^{\mathrm{M1}}$ and $X^{\mathrm{M2}}$ equal unity. Subsequently, given $\mathbf{x} = \mathbf{X}^{s}_{e_s}$ and by linearizing $\mathbf{Ax} = \mathbf{b}$ we can return the set of linear equalities on the olivine site fractions, which comprises two site normalization expressions:

$$1 \times \mathrm{X}^{\mathrm{M1}}_{\mathrm{Mg}} + 1 \times \mathrm{X}^{\mathrm{M1}}_{\mathrm{Fe}} = 1.0, \tag{20}$$

and

$$1 \times \mathrm{X}^{\mathrm{M2}}_{\mathrm{Mg}} + 1 \times \mathrm{X}^{\mathrm{M2}}_{\mathrm{Fe}} + 1 \times \mathrm{X}^{\mathrm{M2}}_{\mathrm{Ca}} = 1.0. \tag{21}$$

Note that, for more complex activity-composition models, and depending on the arbitrary order in which the site fractions are listed, the nullspace operation may not yield expressions that clearly represent the site normalization, charge balance, and equidistribution constraints. However, it will always produce an independent set of linear equalities whose total number is equal to the sum of the number of site normalization constraints, the number of charge balance constraints, and the number of equidistribution constraints. These equalities are mathematically equivalent to the original constraints and can be linearly recombined to recover their straightforward forms.

In the implementation of Riel et al. (2022), MAGEMin solved for the equilibrium composition and state of order of a phase in terms of the variables $x_{cv}$, subject to the constraint that the values of site fractions should be $\geq 0$. To eliminate the site fraction inequalities from the implementation, we now wish to solve directly for the site fractions, while subjecting them to the equality constraints obtained via the nullspace operation. The resulting problem can be expressed as

$$\min f(\mathbf{x}), \tag{22}$$

subject to

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \tag{23}$$

where $x$ represent the $n$ site-fractions and $\mathbf{A}\mathbf{x} = \mathbf{b}$ is the set of $p$ equality constraints (Eq. 18). The linear equality constraints can now be eliminated from the problem, reducing the number of variables solved for back to the number of composition and ordering variables. This is accomplished by parameterizing the feasible set of the constraint equation $\mathbf{A}\mathbf{x} = \mathbf{b}$ using a particular solution $\hat{\mathbf{x}} \in \mathbb{R}^n$ and a matrix $\mathbf{N}\mathbf{z} \in \mathbb{R}^{n \times (n-p)}$ that spans the nullspace of $\mathbf{A}$, such that:

$$\left\{\mathbf{x} \mid \mathbf{A}\mathbf{x} = \mathbf{b}\right\} = \left\{\hat{\mathbf{x}} + \mathbf{N}z \mid z \in \mathbb{R}^{n-p}\right\}. \tag{24}$$

This parameterization can be obtained by performing a full QR decomposition of the constraint matrix $\mathbf{A}$, written as:

$$\mathbf{A} = \mathbf{Q}\begin{bmatrix}\mathbf{R}_1 \\ \mathbf{0}\end{bmatrix} = \begin{bmatrix}\mathbf{Q}_1 \ \mathbf{Q}_2\end{bmatrix}\begin{bmatrix}\mathbf{R}_1 \\ \mathbf{0}\end{bmatrix}, \tag{25}$$

where $\mathbf{Q} \in \mathbb{R}^{m \times m}$ is an orthogonal matrix whose columns $\mathbf{q}_1, \ldots, \mathbf{q}_m$ form an orthonormal basis for $\mathbb{R}^m$. These columns are typically grouped into two blocks: $\mathbf{Q}_1$, which contains the first $p$ columns and spans the image (row space) of $\mathbf{A}$, and $\mathbf{Q}_2$, which contains the remaining $m - p$ columns and spans the nullspace of $\mathbf{A}^T$. With this decomposition, the set of solutions to the linear equality constraints can be written as:

$$\mathbf{x} = \mathbf{Q}_1\mathbf{R}_1^T\mathbf{b} + \mathbf{Q}_2 z, \tag{26}$$

where $\hat{\mathbf{x}} = \mathbf{Q}_1\mathbf{R}_1^T\mathbf{b}$ is a particular solution to $\mathbf{A}\mathbf{x} = \mathbf{b}$, and $\mathbf{Q}_2 z$ is any vector in the nullspace of $\mathbf{A}$.

In other words, the use of the null space of $\mathbf{A}$ ($\mathbf{N}z$) parameterizes the space such that for any step $\Delta z$, $\hat{\mathbf{x}} + \mathbf{N}z\Delta z$ remains in the feasible domain.

Using the elimination method, equation 1 becomes

$$\min f(\mathbf{x}) = G_\lambda(\mathbf{X}^s_{e_{s,i}}) = \sum_{i=1}^{N_\lambda} \mu_{i(\lambda)} p_{i(\lambda)}, \tag{27}$$

and the parameterized first derivative becomes

$$\frac{\partial f}{\partial \mathbf{x}}(z) = \frac{\partial G_\lambda}{\partial \mathbf{X}^s_{e_{s,i}}}(z) = \mathbf{N}_z\left(\left(\frac{\partial G_\lambda}{\partial \mathbf{x}}\right)^T \mathbf{N}_z\right)^T, \tag{28}$$

where $\mathbf{X}^s_{e_{s,i}}$ is the site fraction of the element $e_{s,i}$ that appears on site $s$. Equation 27 is then minimized using the gradient information given by 28 and the methods presented below.

### 2.3 Gradient-type iterative methods

Given a bound-constrained optimization problem:

$$\begin{aligned} &\min f(\mathbf{x}) \\ &\text{subject to } \mathbf{x} > \epsilon \end{aligned} \tag{29}$$

where $f(\mathbf{x})$ is twice continuously differentiable, where $\epsilon$ is a small number, typically $\leq 10^{-8}$. The general gradient-type iterative method to solve this problem is of the form

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k, \tag{30}$$

for iteration $k \geq 0$, where $\mathbf{d}_k$ is the search direction and $\alpha_k$ is the step-length. In this study, we compute the step-length using a Wolfe line search (Wolfe, 1969) such that the inequalities

$$f(\mathbf{x}_k + \gamma_k \alpha_k \mathbf{d}_k) \leq f(\mathbf{x}_k) + \rho \gamma_k \alpha_k \mathbf{g}_k^T \mathbf{d}_k, \tag{31}$$

and

$$\mathbf{g}_{k+1}^T \mathbf{d}_k \geq \sigma \mathbf{g}_k^T \mathbf{d}_k, \tag{32}$$

are satisfied, where $0 < \rho < \sigma < 1$ and $\gamma_x$ is the maximum feasible step-length, computed as

$$\gamma_x = \begin{cases} 1/\frac{\min(\mathbf{x}_{k+1} - \epsilon)}{\mathrm{abs}(\mathbf{d}_k)}, & \text{if any } \mathbf{x}_k + \mathbf{d}_k \leq 0 \\ 1, & \text{otherwise.} \end{cases} \tag{33}$$

The maximum feasible step-length $\gamma_x$ ensures that the values of site-fractions remain $\geq \epsilon$ and form the bounds of the problem. Iterations are then processed until a stopping criterion is satisfied. Because solution phases are not necessarily convex during global Gibbs energy minimization, we set the stopping criteria using the relative change of the objective function. The stopping criteria is met when

$$\mathrm{abs}((f_k - f_{k-1})/f_{k-1}) < \mathrm{tol}, \tag{34}$$

where $\mathrm{tol}$ is a small number typically $\leq 10^{-8}$.

If the descent direction $\mathbf{d}_k$ is simply chosen to be $\mathbf{d}_k =$ -$\mathbf{g}_k$ we obtain the steepest descent algorithm. However, this approach is known to be prone to oscillation (e.g., Nocedal et al., 2002) and slow convergence, and will therefore not be explored. Instead, we test two bound-constrained optimization methods that use the gradient information of the previous iteration(s), namely the conjugate gradient and the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method.

### 2.3.1 Conjugate Gradient method

For the conjugate gradient method, the descent direction is initialized for the first iteration increment k = 0 as

$$\mathbf{d}_k = \text{-}\mathbf{g}_k, \tag{35}$$

and for increment k $\geq$ 1 as

$$\mathbf{d}_k = \text{-}\mathbf{g}_k + \beta_k \mathbf{d}_{k-1}. \tag{36}$$

Here, $g_k$ is the gradient $g(x)$ of function $f(x)$ at point $x_k$, and $\beta_k$ is the conjugate gradient update parameter. Variants of the conjugate gradient method are defined by using different update parameters $\beta_k$ (see for example Hestenes and Stiefel, 1952; Rivaie et al., 2012, 2015). Here, we employ the three-term conjugate gradient method presented by Liu et al. (2018) with the update parameter $\beta_k$ defined in Rivaie et al. (2015)

$$\beta_k = \frac{\mathbf{g}_k^T(\mathbf{g}_k - \mathbf{g}_{k-1} - \mathbf{d}_{k-1})}{\|\mathbf{d}_{k-1}\|^2}, \tag{37}$$

and further extend the descent direction term as

$$\mathbf{d}_k = \text{-}\mathbf{g}_k + \beta_k \mathbf{d}_{k-1} + \theta_k \mathbf{y}_{k-1}, \tag{38}$$

where $\mathbf{y}_{k-1} = \mathbf{g}_k - \mathbf{g}_{k-1}$ and

$$\theta_k = -\frac{\mathbf{g}_k^T \mathbf{d}_{k-1}}{\|\mathbf{d}_{k-1}\|^2}. \tag{39}$$

A useful property of the three-term conjugate gradient method is that the search direction always satisfies the sufficient descent condition without any line search (Liu et al., 2018).

The descent direction is parameterized to satisfy the equality constraints (Eq. 18) such as

$$\mathbf{d}_k^p = \mathbf{N}_z(\mathbf{d}_k^T \mathbf{N}_z)^T. \tag{40}$$

### 2.3.2 BFGS method

The Broyden-Fletcher-Goldfarb-Shanno (BFGS) method is a well-known quasi-Newton method for solving unconstrained and bound-constrained optimization problems (see for instance Fletcher, 1987; Dennis Jr and Schnabel, 1996). The quasi-Newton descent direction is given by

$$\mathbf{d}_k = -\mathbf{B}^{-1}\mathbf{g}_k, \tag{41}$$

where $\mathbf{B}^{-1}$ is the inverse of the Hessian matrix. Here, we approximate $\mathbf{B}^{-1}$ using the Sherman-Morrison formula (Sherman and Morrison, 1950) such as

$$\mathbf{B}_{k+1}^{-1} = \mathbf{B}_k^{-1} + \frac{(\mathbf{s}_k^T \mathbf{y}_k + \mathbf{y}_k^T \mathbf{B}_k^{-1} \mathbf{y}_k)(\mathbf{s}_k \mathbf{s}_k^T)}{(\mathbf{s}_k^T \mathbf{y}_k)^2} - \frac{\mathbf{B}_k^{-1}\mathbf{y}_k \mathbf{s}_k^T + \mathbf{s}_k \mathbf{y}_k^T \mathbf{B}_k^- 1}{\mathbf{s}_k^T \mathbf{y}_k} \tag{42}$$

where $\mathbf{s}_k = \mathbf{x}_{k+1} - \mathbf{x}_k$, $\mathbf{y}_k = \mathbf{g}_{k+1} - \mathbf{g}_k$ and $\mathbf{B}_{k=0}^{-1}$ is initialized with the identity matrix.

Because of the relatively low dimensionality of the solution phase model ($< 20$) we do not consider the limited-memory BFGS method (L-BFGS) and instead update $\mathbf{B}_k^{-1}$ (Eq. 42) during every iteration. Once the problem has converged (i.e., Eq. 34

is satisfied), we reset the Hessian matrix inverse $\mathbf{B}_{k-1}$ to the identity matrix and perform additional iteration(s). This ensures that the problem converges to its local minimum in the event the quality of the approximate Hessian matrix inverse $\mathbf{B}_{k-1}$ is degraded.

As for the conjugate gradient method, the descent direction is parameterized to satisfy the equality constraints (Eq. 18) such as

$$\mathbf{d}_k^p = \mathbf{N}_z(\mathbf{d}_k^T\mathbf{N}_z)^T. \tag{43}$$

## 3 Application

In order to test the bound-constrained solution phase formulation, we selected, from the Holland et al. (2018) set of xeos spaning 11 oxides ($Na_2O$-CaO-$K_2O$-FeO-MgO-$Al_2O_3$-$SiO_2$-$H_2O$-$TiO_2$-O-$Cr_2O_3$), three solution phases with complex features including high-dimensional composition–order spaces and geologically significant solvi: clinoamphibole ($Na_2O$-CaO-$K_2O$-FeO-MgO-$Al_2O_3$-$SiO_2$-$H_2O$-$TiO_2$-O), clinopyroxene ($Na_2O$-CaO-$K_2O$-FeO-MgO-$Al_2O_3$-$SiO_2$-$TiO_2$-O-$Cr_2O_3$) and spinel (FeO-MgO-$Al_2O_3$-$TiO_2$-O-$Cr_2O_3$)(Table 1). We use as starting points the set of feasible points of each discretized solution phase. Discretization of the solution phases is achieved using a compositional variable step of 0.25 which yielded 5498, 4124 and 1521 feasible starting points (or pseudocompounds) for clino-amphibole, clinopyroxene and spinel, respectively. Because gradient-based minimization of solution phase models is achieved with respect to a given Gibbs hyperplane (de Capitani and Brown, 1987; de Capitani and Petrakakis, 2010; Xiang and Connolly, 2021; Riel et al., 2022), we first compute the phase equilibrium at a given pressure, temperature and bulk-rock composition to retrieve the global minimum Gibbs hyperplane using MAGEMin (Table 1). Although any other arbitrary Gibbs hyperplane can be used for this test, we choose a global minimum hyperplane in order to explore a known spinel solvus, All computations were performed on a Linux (x86_64-linux-gnu) operating system, utilizing a 6-core 11th Gen Intel(R) Core(TM) i5-11400H CPU running at 2.70 GHz.

The performance and reliability of the bound-constrained formulations are tested against the inequality constrained formulations using the Sequential Least-Squares Quadratic Programming (SLSQP) (Kraft, 1988, 1994) and Conservative Convex Separable Approximation with Quadratic penalty (CCSAQ) methods (Svanberg, 2002). The minimizations using the SLSQP and CCSAQ methods were computed using the C implementation of NLopt (Jackson et al., 2018) through MAGEMin as described in Riel et al. (2022) and in the scripts provided in supplementary materials. Because the algorithms explored in this study (Julia implementation of CG and BFGS methods) exhibited similar accuracy with residual $\leq 10^{-13}$, the differences in algorithm accuracy are not discussed. Here, a minimization is considered successful when the norm of the distance to the solution is $\leq 10^{-4}$.

**Table 1.** Solution phase models parameters.

| Tested phase | **clino-amphibole** | **clinopyroxene** | **spinel** | **spinel solvus** |
|---|---|---|---|---|
| Pressure [kbar] | 5.0 | 12.0 | 12.0 | 3.26 |
| Temperature [K] | 923.15 | 1323.15 | 1323.15 | 1179.4 |
| Number of points | 4950 | 4121 | 1521 | 1521 |
| Tested methods | CCSAQ, SLSQP, CG, BFGS | CCSAQ, SLSQP, CG, BFGS | CCSAQ, SLSQP, CG, BFGS | BFGS |
| Number of dimensions | 17 | 13 | 10 | 10 |
| Oxides | Reference Gibbs hyperplane [J] | | | |
| $SiO_2$ | -960.9655 | -1011.909631 | -1011.909631 | -1001.730935 |
| $Al_2O_3$ | -1768.2476 | -1829.092564 | -1829.092564 | -1818.611331 |
| $CaO$ | -788.4474 | -819.264126 | -819.264126 | -812.972365 |
| $MgO$ | -678.9683 | -695.467358 | -695.467358 | -689.113013 |
| $FeO$ | -355.2975 | -412.948568 | -412.948568 | -396.911228 |
| $K_2O$ | -914.9708 | -971.890270 | -971.890270 | -966.511310 |
| $Na_2O$ | -839.9561 | -876.544354 | -876.544354 | -882.719670 |
| $TiO_2$ | -1008.3630 | -1073.640927 | -1073.640927 | -1045.994137 |
| $O$ | -263.7269 | -276.590707 | -276.590707 | -249.181839 |
| $Cr_2O_3$ | -1262.6087 | -1380.299631 | -1380.299631 | -1332.815844 |
| $H_2O$ | -368.4674 | - | - | - |

## 4 Discussion

### 4.1 Algorithms performance and reliability

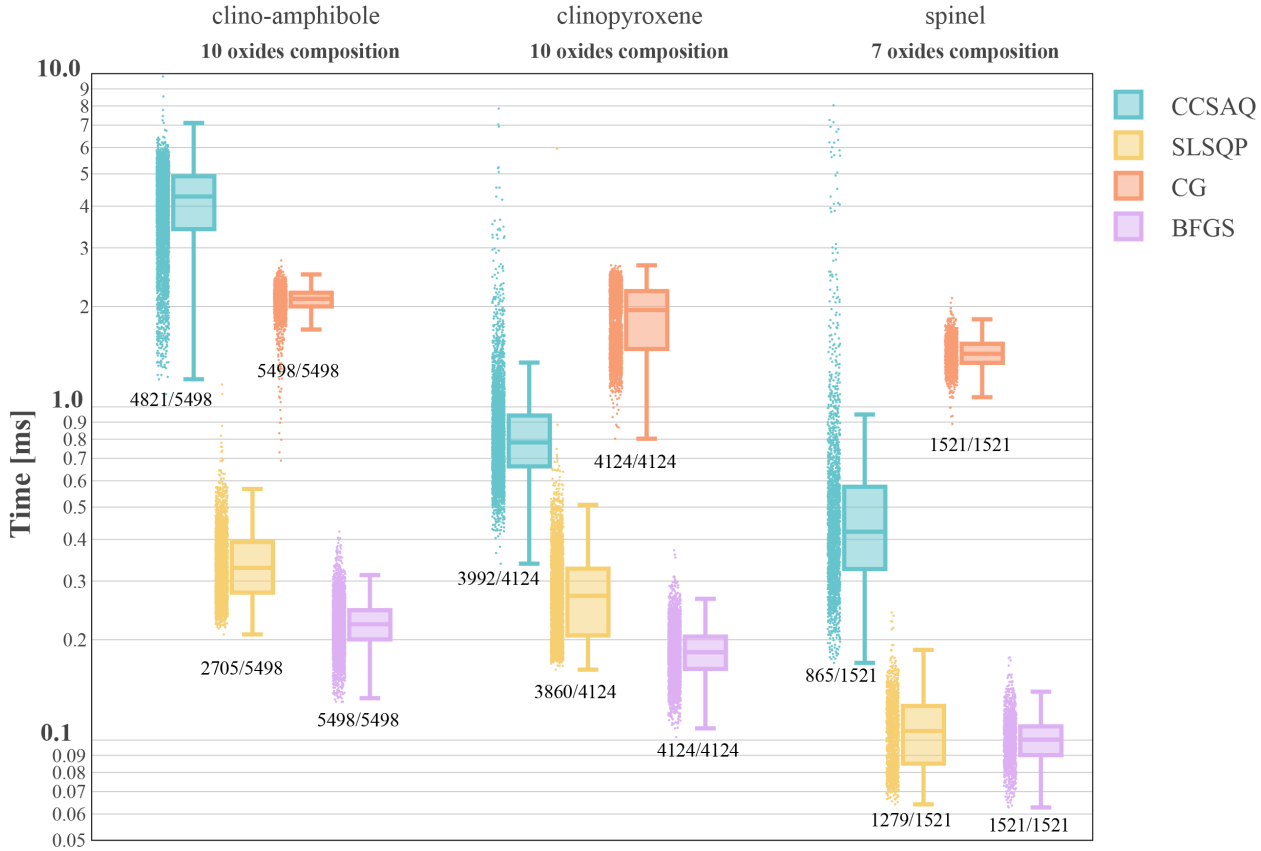


**Figure 2.** Minimization time box plot for tested solution phases and optimization methods. SLSQP, Sequential Least-Squares Quadratic Programming (supporting both inequality and equality constraints); CCSAQ, Conservative Convex Separable Approximation with Quadratic penalty; CG, conjugate gradiend; BFGS, Broyden-Fletcher-Goldfarb-Shanno. 5498, 4124 and 1521 starting points for clino-amphibole, clinopyroxene and spinel, respectively. Starting points were generated by evenly sampling the entire feasible space following the method presented in Riel et al. (2022). The numbers below the boxes show the number of successful minimizations over the total number of tested points.

The box plots depicted in Figure 2 illustrate that the performance of the unconstrained CG method is comparable to that of the inequality-constrained CCSAQ method (implemented via NLopt). While the CG method outperforms CCSAQ for amphibole, with minimization times of $\sim$2100 $\mu$s versus $\sim$4200 $\mu$s, respectively, the efficiency of CCSAQ is larger for problems with lower dimensionality, such as clinopyroxene and spinel. The SLSQP method demonstrates superior efficiency, with average minimization times of $\sim$340 $\mu$s for amphibole, $\sim$270 $\mu$s for clinopyroxene, and $\sim$120 $\mu$s for spinel. However, the BFGS

algorithm outperforms SLSQP, achieving average minimization times of ∼220 $\mu$s for amphibole, 180 $\mu$s for clinopyroxene, and ∼100 $\mu$s for spinel; a performance increase of 20 to 50%. Additionally, the BFGS method's convergence requires between 25 to 90 iterations across different solution phase models, as indicated in Figure 3. Notably, the minimum time per iteration is influenced by the dimensionality of the solution phase model, ranging from ∼4.0 $\mu$s per iteration for clino-amphibole to ∼2.1 $\mu$s for spinel (Figure 3 and table 1).

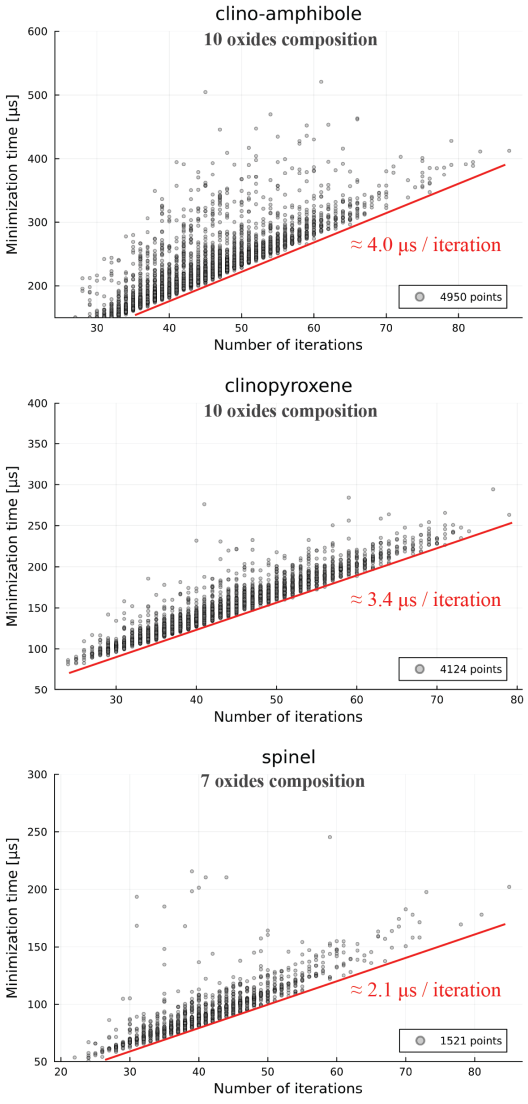


**Figure 3.** Number of iterations versus minimization time for the BFGS method. The red lines show the minimum minimization time per iteration.

Although the average minimization time is a good indicator of the raw performance of the algorithms, reliability of the solvers is of key importance when computing phase equilibria. In this light, we find that the bound-constrained algorithms

(CG and BFGS) are far superior to the inequality constraints ones (CCSAQ and SLSQP). For instance, the bound-constrained methods (BFGS and CG) successfully minimize 100% of the tested starting points (Figure 2) while the inequality constraints methods show a significant amount of unsuccessful minimization reaching up to 50% in some cases e.g., clino-amphibole minimization using SLSQP or spinel using CCSAQ (Fig. 2). The unsuccessful minimizations are related to violated inequality constraints and the inability for the algorithms (SLSQP and CCSAQ) to go back to the feasible domain.

Finally, we tested the BFGS algorithm for an equilibrium between two phases separated by a solvus i.e., an objective function containing more than one local minimum. The parameters of the test are given in table 1 (spinel solvus) and the results are shown in Figure 4. The bound-constrained formulation and the BFGS method perfectly captures the solvus with consistent minimization time similar to those shown in Figure 2

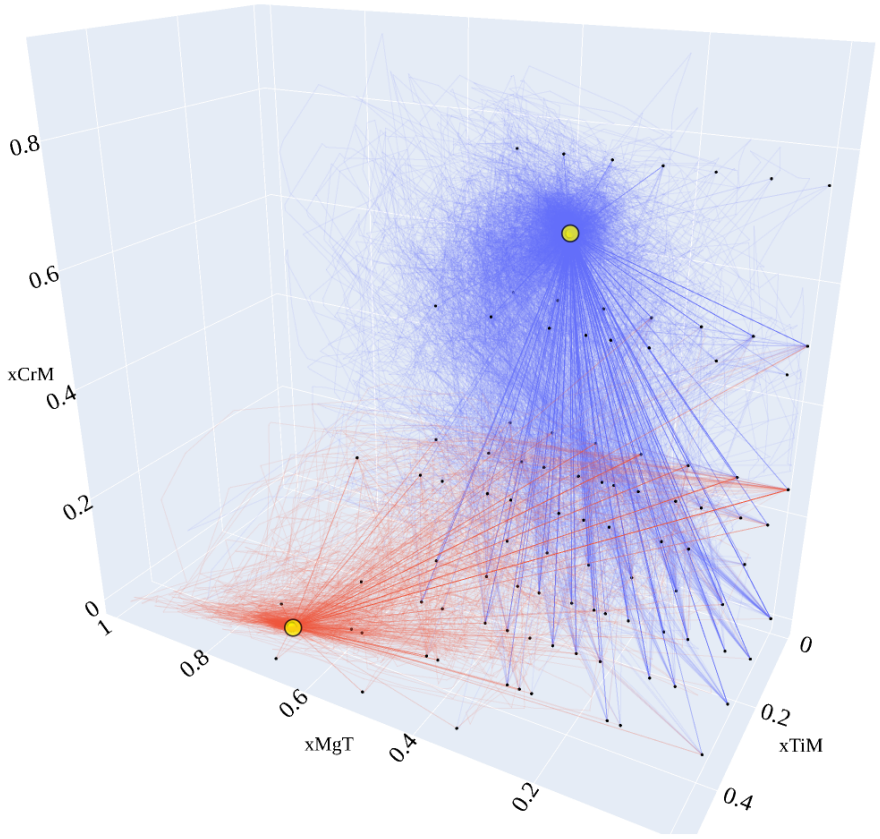


**Figure 4.** Minimization paths for an example of spinel solvus. The spinel solvus has been computed using the Gibbs hyperplane provided in table 1 which was computed using MAGEMin. Black dots, starting point in the $X^M_{Cr}$, $X^T_{Mg}$ and $X^M_{Ti}$ site fraction sub-system; yellow circles with black outline, local minimum of the solvus; red and blue lines, minimization paths from starting point to local minimum. Note that the diagram only displays a 3D sub-system of the full 10D system.

### 4.2 Minimization of perturbed systems

Minimization from discretized starting points allows quantification and comparison of the raw performance and stability of the algorithms (CG, BFGS and CCSAQ, see Figure 2). However, a phase equilibrium calculation employing gradient-based methods generally involves finding a new local minimum under slightly to moderately perturbed conditions between global iterations (de Capitani and Petrakakis, 2010; Riel et al., 2022) i.e., that the distance between the starting point and the ending point during local minimization of individual phases is small. A measure of this distance can be calculated as the variation of the norm of the Gibbs hyperplane: $\|\Delta\Gamma\|_2^2$ , where $\Gamma$ is the chemical potential of the pure components of the system (oxides in our case, Table 1). In order to test the performance of the SLSQP and BFGS methods under small perturbations, we use as starting points the minima obtained by tests 1 to 3 and apply a perturbation to the Gibbs hyperplane (table 1). The perturbation is set by applying a random rotation to the objective function which shifts the local minimum from its current position. We explore the effect of such perturbation by computing 10'000 random rotations per solution phase yielding a range of chemical potential $\|\Delta\Gamma\|_2^2$ varying from 0.0 to ca. 60.0. The results of the minimizations of the rotated systems are presented in Figure 5A-C.

For perturbed conditions, we find that the minimization time of the SLSQP algorithm does not scale with the norm of the perturbation ($\|\Delta\Gamma\|_2^2$) (Figure 5A,B,C). This relationship is independent of the tested solution phase model (Figure 5A-C). Instead, for the BFGS algorithm, the minimization time scales with the norm of the perturbation (Figure 5D-F). For a $\|\Delta\Gamma\|_2^2 < 10.0$ the minimization time is divided by a factor of ca. 2.0 to 3.0 with respect to the mean raw minimization time (Figure 2) resulting in an average time of 70-80 $\mu$s for clino-amphibole, 50-60 $\mu$s for clinopyroxene and 30-40 $\mu$s for spinel.

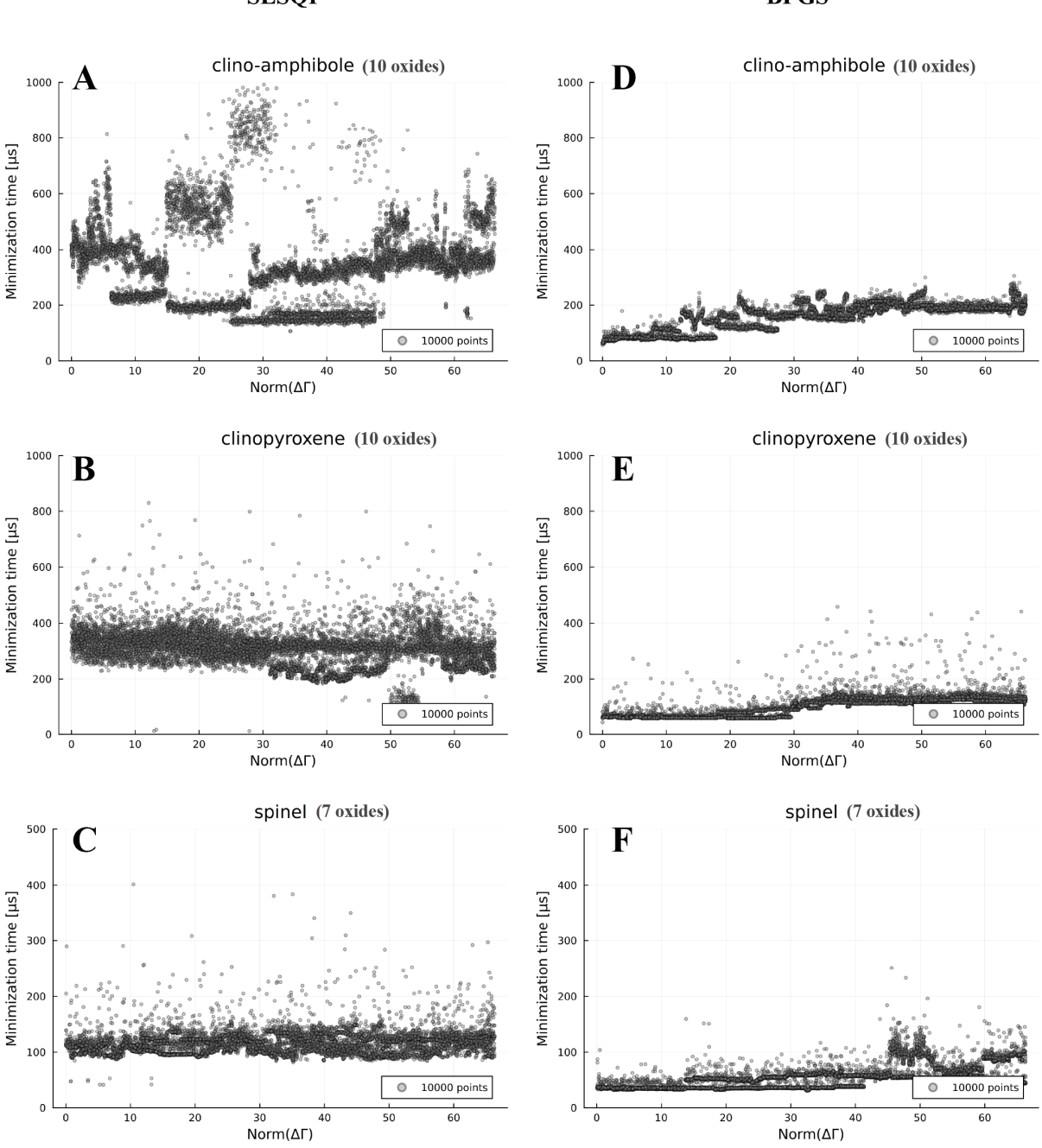


**Figure 5.** Minimization time for perturbed systems. A-C SLSQP algorithm applied to clino-amphibole, clinopyroxene and spinel, respectively. D-F, BFGS algorithm applied to clino-amphibole, clinopyroxene and spinel, respectively. $\|\Delta\Gamma\|_2^2$ is a measure of the distance between the starting guess and the solution. Note that the inequality constrained SLSQP method does not show any correlation between the distance to solution and the minimization. The BFGS method, instead, exhibits a clear trend of decreasing minimization for smaller distances to solution.

We propose a revised xeos implementation using a nullspace approach, that allows using bound-constrained, rather than inequality-constrained, gradient-based optimisation methods. We tested the performance and computational reliability of different algorithms and find that the BFGS method yields the best performance, decreasing the minimization time of individual solution phases by a factor $\geq 10$ compared to CG and CCSAQ methods and by a factor of 1.5 to 2.0 with respect to the SLSQP method (Figure 2). Under slight perturbations, the minimization time is further decreased by a factor of 2.0 reaching down to $\leq 100$ $\mu$s for clino-amphibole, $\leq 80$ $\mu$s for clinopyroxene and $\leq 50$ $\mu$s for spinel.

Regarding computational reliability, bound-constrained optimization methods(such as CG and BFGS), are clearly preferable over methods with inequality constraints (CCSAQ and SLSQP), which exhibit a considerable proportion of unsuccessful minimizations.

Using an bound-constrained nullspace formulation and the BFGS method, can therefore significantly improve the performance of stable phase calculations. Given that in MAGEMin, $\geq 75\%$ of the computational load is dedicated to local minimization of solution phase models, we estimate a general speed-up of stable phase equilibrium prediction by a factor $\geq 5$. Such improvement can potentially opens up the possibility of performing 2D reactive models fluid/magma transport.

*Code availability.* All the Julia scripts and data necessary to reproduce the results and figures of this contribution are provided on Zenodo (10.5281/zenodo.13982544) and Github on https://github.com/ComputationalThermodynamics/SandBox

*Author contributions.* NR developed and implemented the new formulation, made the analysis and wrote the paper. BK, EG, AdM, and VM provided technical support and feedback during development. HD did the additional testing and improvement of the code. All authors actively contributed to the project through regular meetings and provided critical feedback. All authors revised the final version of the paper.

*Competing interests.* BK (co-author) is a member of the editorial board of Geoscientific Model Development.

*Disclaimer.* TEXT

*Acknowledgements.* This study was funded by the European Research Council through the MAGMA project, ERC Consolidator Grant #771143. N.R., B.K. and E.M. would also like to acknowledge the German Research Foundation (DFG) - Project number 521637679 for financial support.

## References


Bouilhol, P., Magni, V., van Hunen, J., and Kaislaniemi, L.: A numerical approach to melting in warm subduction zones, Earth and Planetary Science Letters, 411, 37–44, 2015.

Candioti, L. G., Nathwani, C. L., and Chelle-Michou, C.: Towards fully-coupled thermodynamic-thermomechanical two-phase flow models of transcrustal magmatic systems, in: EGU General Assembly Conference Abstracts, p. 16936, 2024.

Connolly, J. A. D.: Computation of phase equilibria by linear programming: A tool for geodynamic modeling and its application to subduction zone decarbonation, Earth and Planetary Science Letters, 236, 524–541, 2005.

de Capitani, C. and Brown, T. H.: The computation of chemical equilibrium in complex systems containing non-ideal solutions, Geochimica et Cosmochimica Acta, 51, 2639–2652, https://doi.org/https://doi.org/10.1016/0016-7037(87)90145-1, 1987.

de Capitani, C. and Petrakakis, K.: The computation of equilibrium assemblage diagrams with Theriak/Domino software, American Mineralogist, 95, 1006–1016, https://doi.org/doi:10.2138/am.2010.3354, 2010.

Dennis Jr, J. E. and Schnabel, R. B.: Numerical methods for unconstrained optimization and nonlinear equations, SIAM, 1996.

Fletcher, R.: Practical methods of optimization, 1987, John and Sons, Chichester, 1987.

Ganguly, J.: Thermodynamic modelling of solid solutions, EMU Notes in Mineralogy, 3, 37–69, 2001.

Ghiorso, M. S. and Sack, R. O.: Chemical mass transfer in magmatic processes IV. A revised and internally consistent thermodynamic model for the interpolation and extrapolation of liquid-solid equilibria in magmatic systems at elevated temperatures and pressures, Contributions to Mineralogy and Petrology, 119, 197–212, 1995.

Helgeson, H. C.: Summary and critique of the thermodynamic properties of rock-forming minerals, American Journal of Science, 278, 1–229, 1978.

Hestenes, M. R. and Stiefel, E.: Methods of conjugate gradients for solving, Journal of research of the National Bureau of Standards, 49, 409, 1952.

Holland, T. and Powell, R.: Activity–composition relations for phases in petrological calculations: an asymmetric multicomponent formulation, Contributions to Mineralogy and Petrology, 145, 492–501, 2003.

Holland, T. J. B. and Powell, R.: An internally consistent thermodynamic data set for phases of petrological interest, Journal of Metamorphic Geology, 16, 309–343, https://doi.org/https://doi.org/10.1111/j.1525-1314.1998.00140.x, 1998.

Holland, T. J. B., Green, E. C. R., and Powell, R.: Melting of peridotites through to granites: a simple thermodynamic model in the system KNCFMASHTOCr, Journal of Petrology, 59, 881–900, 2018.

Hu, H., Jackson, M. D., and Blundy, J.: Melting, compaction and reactive flow: Controls on melt fraction and composition change in crustal mush reservoirs, Journal of Petrology, 63, egac097, 2022.

Jackson, M., Blundy, J., and Sparks, R.: Chemical differentiation, cold storage and remobilization of magma in the Earth's crust, Nature, 564, 405–409, 2018.

Jackson, M. D., Cheadle, M. J., and Atherton, M. P.: Quantitative modeling of granitic melt generation and segregation in the continental crust, Journal of Geophysical Research: Solid Earth, 108, 2003.

Johnson, S. G.: The NLopt nonlinear-optimization package, https://doi.org/10.7283/633E-1497, 2021.

Katz, R. F., Jones, D. W. R., Rudge, J. F., and Keller, T.: Physics of Melt Extraction from the Mantle: Speed and Style, Annual Review of Earth and Planetary Sciences, 50, 2022.

Keller, T. and Katz, R. F.: The Role of Volatiles in Reactive Melt Transport in the Asthenosphere, Journal of Petrology, 57, 1073–1108, https://doi.org/10.1093/petrology/egw030, 2016.

Keller, T. and Suckale, J.: A continuum model of multi-phase reactive transport in igneous systems, Geophysical Journal International, 219, 185–222, https://doi.org/10.1093/gji/ggz287, 2019.

Keller, T., May, D. A., and Kaus, B. J. P.: Numerical modelling of magma dynamics coupled to tectonic deformation of lithosphere and crust, Geophysical Journal International, 195, 1406–1442, https://doi.org/10.1093/gji/ggt306, 2013.

Keller, T., Katz, R. F., and Hirschmann, M. M.: Volatiles beneath mid-ocean ridges: Deep melting, channelised transport, focusing, and metasomatism, Earth and Planetary Science Letters, 464, 55–68, 2017.

Keller, T., Hanchar, J., Tornos, F., and Suckale, J.: Formation of the El Laco magmatic magnetite deposits by Fe-Si melt immiscibility and bubbly suspension flow along volcano tectonic faults, Authorea Preprints, 2022.

Kraft, D.: A software package for sequential quadratic programming, Forschungsbericht- Deutsche Forschungs- und Versuchsanstalt fur Luft- und Raumfahrt, 1988.

Kraft, D.: Algorithm 733: TOMP–Fortran modules for optimal control calculations, ACM Transactions on Mathematical Software (TOMS), 20, 262–281, 1994.

Leal, A. M., Kyas, S., Kulik, D. A., and Saar, M. O.: Accelerating reactive transport modeling: on-demand machine learning algorithm for chemical equilibrium calculations, Transport in Porous Media, 133, 161–204, 2020.

Liu, J., Feng, Y., and Zou, L.: Some three-term conjugate gradient methods with the inexact line search condition, Calcolo, 55, 1–16, 2018.

Magni, V., Bouilhol, P., and Van Hunen, J.: Deep water recycling through time, Geochemistry, Geophysics, Geosystems, 15, 4203–4216, 2014.

Nocedal, J., Sartenaer, A., and Zhu, C.: On the behavior of the gradient norm in the steepest descent method, Computational Optimization and Applications, 22, 5–35, 2002.

Pincus, M.: A Monte Carlo method for the approximate solution of certain types of constrained optimization problems, Operations research, 18, 1225–1228, 1970.

Piro, M. H. A.: Computation of Thermodynamic Equilibria Pertinent to Nuclear Materials in Multi-Physics Codes, Ph.D. thesis, Royal Military College of Canada (Canada), 2011.

Powell, R. and Holland, T.: An internally consistent dataset with uncertainties and correlations: 3. Applications to geobarometry, worked examples and a computer program, Journal of metamorphic Geology, 6, 173–204, 1988.

Powell, R. and Holland, T.: On the formulation of simple mixing models for complex phases, American Mineralogist, 78, 1174–1180, 1993.

Riel, N., Bouilhol, P., van Hunen, J., Cornet, J., Magni, V., Grigorova, V., and Velic, M.: Interaction between mantle-derived magma and lower arc crust: quantitative reactive melt flow modelling using STyx, Geological Society, London, Special Publications, 478, 65–87, 2019.

Riel, N., Kaus, B. J., Green, E., and Berlie, N.: MAGEMin, an efficient Gibbs energy minimizer: application to igneous systems, Geochemistry, Geophysics, Geosystems, 23, e2022GC010 427, 2022.

Rivaie, M., Mamat, M., June, L. W., and Mohd, I.: A new class of nonlinear conjugate gradient coefficients with global convergence properties, Applied Mathematics and Computation, 218, 11 323–11 332, 2012.

Rivaie, M., Mamat, M., and Abashar, A.: A new class of nonlinear conjugate gradient coefficients with exact and inexact line searches, Applied Mathematics and Computation, 268, 1152–1163, 2015.

Rummel, L., Kaus, B. J. P., Baumann, T. S., White, R. W., and Riel, N.: Insights into the Compositional Evolution of Crustal Magmatic Systems from Coupled Petrological-Geodynamical Models, Journal of Petrology, 61, https://doi.org/10.1093/petrology/egaa029, 2020.

Sherman, J. and Morrison, W. J.: Adjustment of an inverse matrix corresponding to a change in one element of a given matrix, The Annals of Mathematical Statistics, 21, 124–127, 1950.

Stixrude, L. and Lithgow-Bertelloni, C.: Thermodynamics of mantle minerals - II. Phase equilibria, Geophysical Journal International, 184, 1180–1213, https://doi.org/10.1111/j.1365-246X.2010.04890.x, 2011.

Svanberg, K.: A class of globally convergent optimization methods based on conservative convex separable approximations, SIAM Journal on Optimization, pp. 555–573, 2002.

Taylor-West, J. and Katz, R. F.: Melt-preferred orientation, anisotropic permeability and melt-band formation in a deforming, partially molten aggregate, Geophysical Journal International, 203, 1253–1262, 2015.

Turner, A. J., Katz, R. F., Behn, M. D., and Keller, T.: Magmatic focusing to mid-ocean ridges: The role of grain-size variability and non-newtonian viscosity, Geochemistry, Geophysics, Geosystems, 18, 4342–4355, 2017.

Wolfe, P.: Convergence conditions for ascent methods, SIAM review, 11, 226–235, 1969.

Xiang, H. and Connolly, J. A. D.: GeoPS: An interactive visual computing tool for thermodynamic modelling of phase equilibria, Journal of Metamorphic Geology, n/a, https://doi.org/https://doi.org/10.1111/jmg.12626, 2021.

Yuan, Q., Asimow, P. D., Gurnis, M., and Antoshechkina, P. M.: Enabling large-scale geodynamic-geochemical modeling via neural network acceleration, in: AGU Fall Meeting Abstracts, vol. 2024, pp. V41E–3161, 2024.